@'

MF
ANALIZĂ FUNCŢIONALĂ, PROBABILITĂŢI şi STATISTICĂ MATEMATICĂ

Paul Flondor, Mircea Olteanu

Cuprins

Introducere

Partea de Analiză funcţională tratează următoarele subiecte: spaţii metrice, spaţii normate, spaţii Hilbert, măsură şi integrală, serii Fourier, operatori liniari şi continui cât şi aplicaţii ale acestor noţiuni în teoria matematică a sistemelor. Textul se adresează studenţilor din universităţile tehnice, deci prezentarea noţiunilor şi a rezultatelor s-a făcut tinându-se cont de programa de matematică din aceste tip de universităţi. Cu unele excepţii, cursul conţine noţiunile, enunţurile şi demonstraţiile esenţiale; totuşi, ţinând cont de scopul principal al lucrării, în unele capitole s-a evitat intarea în detaliile tehnice ale unor construcţii sau demonstraţii, compensându-se această omisiune cu exemple semnificative. Sursele bibliografice au ca scop reamintirea rezultatelor de algebră liniară şi de analiză matematică utilizate, cât şi completări şi aprofundări ale noţiunilor de analiză funcţională.

Partea de probabilităţi şi Statistică Matematică (MF.08-MF14) acoperă materia unui curs de 14 ore dedicat acestor importante subiecte pentru aplicaţiile inginereşti. Prezentarea evită o formalizare strictă tip definiţie-teoremă -demonstraţie concentrându-se pe precizarea noţiunilor şi rezultatelor fundamentale şi pe ilustrarea acestora cu exemple semnificative. Teoria matematică modernă a probabilităţilor necesită un aparat tehnic dezvoltat pe teoria măsurii şi a integralei abstracte. Ne bazăm, pentru acest subiect, pe capitolul MF.05 dar reluăm, pe scurt, în contextul câmpurilor de probabilitate rezultatele importante. Bibliografia ajută pe studentul interesat să aprofundeze noţiunile, să parcurgă demonstrţii, să găsească exerciţii variate si, evident, să se pregătească pentru examen. Fiecare tema importantă are, în bibliografie, un text accesibil pe Internet. In capitolul MF15 sunt rezovate şi propuse o serie de exerciţii şi probleme menite să ajute atât la aprofundarea noţiunilor cât şi la autoevaluarea studentului în pregătirea examenului.

Capitolul 1

MF.01. Spaţii metrice

Cuvinte cheie :

distanţă, spaţiu metric, şir convergent, contracţie, punct fix, metoda aproximatiilor succesive, funcţie continuă.

În acest capitol presupunem cunoscute noţiunile de topologie pe spaţiul $\mathbf{R^n}$ studiate la cursul de Calcul diferenţial şi integral (MC), ca de exemplu: şir convergent, şir Cauchy, funcţie contină,mulţime deschisă, mulţime închisă, mulţime mărginită, mulţime compactă, etc. Sursele bibliografice recomandate sunt: [B01], [C01], [D01], [F02], [R01], [S02].

1.1 MF.01.1. Noţiuni generale

1. Definiţii

Fie X o mulţime nevidă. O **distanţă** (sau metrică) pe X este orice aplicaţie

$$d : X \times X \mapsto [0\ ,\ \infty)$$

cu proprietăţile:
(i) $d(x, y) = 0$ dacă şi numai dacă $x = y$ (pozitiv definită).
(ii) $d(x, y) = d(y, x), \forall x, y \in X$ (simetrie).
(iii) $d(x, z) \leq d(x, y) + d(y, z), \forall x, y \in X$ (inegalitatea triunghiului).

Prin **spaţiu metric** se înţelege o mulţime X pe care s-a definit o distanţă d ca mai sus; notaţia uzuală pentru un spaţiu metric este (X, d) sau, dacă distanţa d se subînţelege, se notează simplu X.
Dacă $Y \subseteq X$, atunci (Y, D) se numespaţiu metric indus (de distanţa d restricţionată la Y).

2. Exemple

a. Mulţimea numerelor reale, **R**, este spaţiu metric cu distanţa uzuală

$$d(x,y) = |x-y|, \forall x,y \in \mathbf{R}.$$

b. Pe mulţimea numerelor complexe, **C**, distanţa uzuală este

$$d(z,w) = |z-w|, \forall z,w \in \mathbf{C}.$$

c. Fie $\mathbf{R^n} = \{x = (x_1, x_2, ..., x_n) \,|\, x_j \in R\}$. Distanţa uzuală (euclidiană):

$$d_2(x,y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}, \quad \text{unde } x = (x_1, x_2, ..., x_n),\ y = (y_1, y_2, ..., y_n).$$

Tot pe $\mathbf{R^n}$ se mai pot defini şi distanţele:

$$d_1(x,y) = \sum_{k=1}^{n} |x_k - y_k|,$$

$$d_\infty(x,y) = \max_{j=1,2,...,n} |x_j - y_j|.$$

d. Fie $\mathbf{C^n} = \{x = (x_1, x_2, ..., x_n) \,|\, x_j \in \mathbf{C}\}$. Distanţa uzuală (euclidiană):

$$d_2(x,y) = \sqrt{\sum_{k=1}^{n} |x_k - y_k|^2},$$

cu notaţii evidente.

e. Fie $\ell^2(\mathbf{N}) = \{x : N \mapsto \mathbf{R} \mid \sum_n |x(n)|^2 < \infty\}$. Generalizarea distanţei euclidiene la cazul infinit dimensional (pe $\ell^2(\mathbf{N})$) este

$$d(x,y) = \sqrt{\sum_n |x(n) - y(n)|^2}.$$

f. Fie $A \neq \emptyset$ şi fie $\mathcal{M}(A) = \{f : A \mapsto R \mid f \text{ mărginită}\}$. Distanţa "supremum" pe spaţiul funcţiilor mărginite este

$$d_\infty(f,g) = \sup_{x \in A} |f(x) - g(x)|.$$

g. Fie $a, b \in \mathbf{R}$ şi fie $\mathcal{C}[a, b] = \{f : [a, b] \mapsto R \mid f \text{ continuă}\}$. Distanţa uzuală pe spaţiul funcţiilor continue este

$$d_\infty(f, g) = \sup_{x \in [a,b]} |f(x) - g(x)|.$$

h. Tot pe spaţiul $\mathcal{C}[a, b]$, o altă distanţă este

$$d_1(f, g) = \int_a^b |f(t) - g(t)|\, dt.$$

i. Fie $\mathcal{B} = \{0, 1\}$ codul (alfabetul) binar. Pe produsul cartezian

$$\mathcal{B}^n = \{(x_1, x_2, ...x_n) \mid x_j \in \mathcal{B}, \forall\, j = 1, 2, ..., n\}$$

definim distanţa Hamming:

$$d((x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n)) = \sum_{j=1}^{n} |x_j - y_j|.$$

Distanţa Hamming măsoară de fapt numărul de necoincidenţe dintre elementele $(x_1, x_2, ..., x_n)$ şi $(y_1, y_2, ..., y_n)$.

j. Distanţa Hausdorff; fie $K(\mathbf{R^n})$ familia mulţimilor compacte (închise şi mărginite) din $\mathbf{R^n}$. Pentru orice $x \in \mathbf{R^n}$ şi $A \in K(\mathbf{R^n})$ definim:

$$\operatorname{dist}(x, A) = \inf_{a \in A} \| x - a \|,$$

unde, $\| \cdot \|$ este norma euclidiană pe $\mathbf{R^n}$. Distanţa Hausdorff pe mulţimea $K(\mathbf{R^n})$ este:

$$d_H : K(\mathbf{R^n}) \times K(\mathbf{R^n}) \mapsto [0, \infty),$$

$$d_H(A, B) = \max\{\sup_{a \in A} \operatorname{dist}(a, B)\ ,\ \sup_{b \in B} \operatorname{dist}(b, A)\}.$$

k. Spaţiu metric discret; orice mulţime nevidă A se poate organiza ca spaţiu metric (discret) cu distanţa "discretă" $\rho(x, y) = \begin{cases} 1 & \text{dacă} \quad x \neq y \\ 0 & \text{dacă} \quad x = y \end{cases}$

3. Definiţii

Dacă (X, d) este un spaţiu metric, atunci prin **şir** de elemente din X se înţelege orice aplicaţie

$$x : \mathbf{N} \mapsto X.$$

Se notează $x(n) = x_n, \forall n \in \mathbf{N}$ (termenul de rang n al şirului).
Un şir $x_n \in X$ se numeşte **şir convergent** dacă există $a \in X$ astfel încât

$$\forall \epsilon > 0, \exists n_\epsilon \in N \text{ cu proprietatea } d(x_n, a) < \epsilon, \forall n \geq n_\epsilon.$$

În acest caz a se numeşte limita şirului şi se notează $a = \lim\limits_{n\to\infty} x_n$ sau $x_n \to a$. O formulare echivalentă este: şirul $x_n \in X$ este convergent dacă există $a \in X$ astfel încât $\lim_{n\to\infty} d(x_n, a) = 0$. Se demonstrează uşor că limita unui şir convergent este unică. Un şir care nu este convergent se numeşte divergent.

Un şir $x_n \in X$ se numeşte **şir Cauchy** (sau fundamental) dacă

$$\forall \epsilon > 0, \exists n_\epsilon \in N \text{ astfel încât } d(x_n, x_m) < \epsilon, \forall n, m \geq n_\epsilon.$$

Un spaţiu metric se numeşte **spaţiu metric complet** dacă orice şir Cauchy este convergent.

Fie d_1 şi d_2 două distanţe pe o aceeaşi mulţime X. Metricele d_1 şi d_2 se numesc echivalente dacă există $\alpha > 0$ şi $\beta > 0$ astfel încât

$$d_1(x, y) \leq \alpha d_2(x, y) \leq \beta d_1(x, y), \forall x, y \in X.$$

In acest caz, (dacă d_1 şi d_2 sunt echivalente), se demonstrează că un şir $x_n \in X$ este convergent (respectiv Cauchy) în spaţiul metric (X, d_1) dacă şi numai dacă este convergent (respectiv Cauchy) în spaţiul metric (X, d_2). De exemplu, pe $\mathbf{R^n}$ metricele d_1, d_2 şi d_∞ sunt echivalente.

Dintre spaţiile metrice complete uzuale, amintim (a se vedea exemplul 2): $(\mathbf{R^n}, d_2)$ (dar şi cu d_1 şi d_∞), analog $\mathbf{C^n}$, $(\ell^2(\mathbf{N}), d_2)$, $(\mathcal{M}(A), d_\infty)$ $(\mathcal{C}[a, b]\,,\, d_\infty)$, $(K(\mathbf{R^n})\,,\, d_H)$, spaţiul metric discret.

4. Exemple

i. Aşa cum se ştie din cursul de analiză matematică, în spaţiul metric $(\mathbf{R^n}, d_2)$, convergenţa însemnă convergenţă pe componente, i.e. (cu notaţii evidente),

$$\lim_{p\to\infty} (x_1^{(p)}, x_2^{(p)}, ..., x_n^{(p)}) = (a_1, a_2, ..., a_n)$$

dacă şi numai dacă

$$\lim_{p\to\infty} x_1^{(p)} = a_1, \lim_{p\to\infty} x_2^{(p)} = a_2, ..., \lim_{p\to\infty} x_n^{(p)} = a_n.$$

ii. Convergenţa în spaţiul metric al funcţiilor continue cu distanţa d_∞ este convergenţa uniforma. Fie, de exemplu, în spaţiul metric $(\mathcal{C}[0, 1], d_\infty)$, şirurile $f_n(t) = t^n - t^{n+1}$ şi $g_n(t) = t^n - t^{2n}$. Amândouă tind punctual la funcţia nulă (notată 0). Şirul f_n este convergent:

$$d_\infty(f_n, 0) = \sup_{t\in[0,1]} |f_n(t)| = f_n(\frac{n}{n+1}) \to 0, \text{ când } n \to \infty,$$

în timp ce şirul g_n este divergent:

$$d_\infty(g_n, 0) = \sup_{t\in[0,1]} g_n(t) = g_n(\frac{1}{\sqrt[n]{2}}) = \frac{1}{4}.$$

1.2 MF.01.2. Principiul contracţiei

Fie (X, d) un spaţiu metric şi fie $f : X \mapsto X$. Aplicaţia f se numeşte **contracţie** (pe X) dacă există $k \in [0, 1)$ astfel încât:

$$d(f(x), f(y)) \leq k \cdot d(x, y), \forall x, y \in X.$$

Numărul k se numeşte factor de contracţie.

5. Teorema de punct fix a lui Banach (Principiul contracţiei)
Fie (X, d) un spaţiu metric complet şi fie $f : X \mapsto X$ o contracţie de factor k. Atunci există un unic punct $\xi \in X$ astfel încât $f(\xi) = \xi$.
Punctul ξ de mai sus se numeşte **punct fix** al aplicaţiei f.
Demonstraţie

Construcţia lui ξ (numită **metoda aproximaţiilor succesive**) se face astfel: fie $x_0 \in X$, arbitrar fixat şi fie şirul (numit şirul aproximaţiilor succesive) definit prin recurenţă $x_{n+1} = f(x_n)$. Vom demonstra că şirul x_n este convergent şi limita sa este punctul fix căutat. În plus, are loc formula (evaluarea erorii):

$$d(x_n, \xi) \leq \frac{k^n}{1-k} \cdot d(x_0, x_1), \forall n \in \mathbf{N}.$$

Vom arăta mai întâi că şirul x_n este şir Cauchy; avem:

$$d(x_n, x_{n+1}) = d(f(x_{n-1}, f(x_n)) \leq kd(x_{n-1}, x_n) \leq$$

$$\leq k^2 d(x_{n-2}, x_{n-1}) \leq ... \leq k^n d(x_0, x_1), \forall n \in \mathbf{N}.$$

De aici, pentru orice $p, n \in \mathbf{N}$, rezultă (din inegalitatea triunghiului):

$$d(x_n, x_{n+p}) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + ... + d(x_{n+p-1}, x_{n+p}) \leq$$

$$\leq (k^n + k^{n+1} + ... + k^{n+p-1}) d(x_0, x_1) = \frac{k^n(1-k^p)}{1-k} d(x_0, x_1) \leq \frac{k^n}{1-k} \cdot d(x_0, x_1).$$

Deoarece $k^n \to 0$, rezultă că şirul x_n este şir Cauchy, deci convergent (spaţiul metric X a fost presupus complet). Rezultă că există $\xi \in X$ astfel încât $\xi = \lim_{n\to\infty} x_n$. Egalitatea $f(\xi) = \xi$ rezultă din:

$$0 \leq d(x_n, f(\xi)) = d(f(x_{n-1}), f(\xi)) \leq d(x_{n-1}, \xi) \to 0, \text{ când } n \to \infty$$

şi din unicitatea limitei. Unicitatea punctului fix: presupunem prin absurd că există $\eta \in X$, $\eta \neq \xi$ astfel încât $f(\eta) = \eta$; atunci:

$$d(\xi, \eta) = d(f(\xi), f(\eta)) \leq kd(\xi, \eta),$$

şi deci simplificând prin $d(\xi, \eta) \neq 0$, rezultă $k = 1$, contradicţie.

6. Exemple

i. Să se aproximeze cu o eroare mai mică decât 10^{-3} soluţia reală a ecuaţiei $x^3 + 4x - 1 = 0$.

Ecuaţia are o singură soluţie reală, $\xi \in (0, 1)$. Vom aplica metoda aproximaţiilor succesive. Fie $X = [0, 1]$ şi $f : X \mapsto X$, $f(x) = \frac{1}{x^2+4}$. Ecuaţia este echivalentă cu $f(x) = x$, iar spaţiul metric X este complet (cu metrica uzuală indusă din $\mathbf{R}$); demonstrăm acum că f este contracţie pe X. Derivata este $f'(x) = \frac{-2x}{(x^2+4)^2}$, iar $\sup_{x \in X} |f'(x)| = -f'(1) = \frac{2}{25} < 1$, deci f este contracţie cu factorul factorul de contracţie $k = \frac{2}{25}$. Şirul aproximaţiilor succesive este

$$x_0 = 0, x_{n+1} = f(x_n) = \frac{1}{x_n^2 + 4};$$

evaluarea erorii:

$$|x_n - \xi| < \frac{k^n}{1-k}|x_0 - x_1| = \frac{1}{3} \cdot \left(\frac{2}{25}\right)^n,$$

deci $\xi \approx x_3 = f\left(\frac{16}{65}\right) = 0,2355072$.

Aceeaşi ecuaţie se poate rezolva aproximativ şi folosind contracţia $g(x) = \frac{1}{4}(1 - x^3), x \in [0, 1]$. În acest caz factorul de contracţie este $k = \sup\limits_{x \in (0,1)} |g'(x)| = \dfrac{3}{4}$. Metoda aproximaţiilor succesive converge mult mai încet în acest caz, $\xi \approx x_6$.

ii. Să se demonstreze că pentru orice $a \in (-1, 1)$ şi pentru orice funcţie continuă $h : [0, 1] \mapsto \mathbf{R}$ există o singură funcţie continuă $u : [0, 1] \mapsto \mathbf{R}$ astfel încât $u(x) = h(x) + a \sin u(x)$, $\forall x \in [0, 1]$.

Studiem ecuaţia în spaţiul metric $(\mathcal{C}([0, 1]), d_\infty))$. Fie aplicaţia

$$F : \mathcal{C}([0, 1]) \mapsto \mathcal{C}([0, 1]),\ (F(f))(x) = h(x) + a \sin f(x), \forall x \in [0, 1].$$

Este suficient să demonstrăm că F este contracţie (funcţia u din concluzie este punctul fix al contracţiei). Fie $f, g \in \mathcal{C}([0, 1])$; atunci:

$$d_\infty(F(f), F(g)) = \sup_{x \in [0,1]} |a(\sin f(x) - \sin g(x))| \leq |a| \cdot d_\infty(f, g),$$

deci F este contracţie cu factorul $|a|$.

7. Metoda lui Newton
Fie $a, b \in R$ şi fie $f : [a, b] \mapsto R$ o funcţie de două ori derivabilă astfel încât $f'(x) \neq 0, \forall x \in [a, b]$. Presupunem că ecuaţia $f(x) = 0$ are o soluţie $\xi \in [a, b]$. Metoda lui Newton aproximează ξ cu şirul aproximaţiilor succesive generat de contracţia $g(x) = x - \dfrac{f(x)}{f'(x)}$. Să se găsească o condiţie suficientă pentru a se putea aplica metoda aproximaţiilor succesive şi să se interpreteze geometric metoda.
Condiţia $|g'(x)| \leq k < 1$ (pentru ca g să fie contracţie) este echivalentă cu $|f(x)f''(x)| < k(f'(x))^2, \forall x \in [a, b]$. Dacă această condiţie este îndeplinită, atunci şirul aproximaţiilor succesive este $x_n = x_{n-1} - \dfrac{f(x_{n-1})}{f'(x_{n-1})}$. Geometric, x_n este abscisa punctului de intersecţie al tangentei la graficul funcţiei f în punctul $(x_{n-1}, f(x_{n-1}))$ cu axa Ox. Într-adevăr, ecuaţia tangentei este $y - f(x_{n-1}) = f'(x_{n-1})(x - x_{n-1})$ şi pentru $y = 0$ se obţine x_n. Condiţia de convergenţă, $|g'(x)| \leq k < 1$ este satisfăcută dacă punctul iniţial, x_0, este suficient de apropiat de ξ; aceasta rezultă din faptul că $g'(\xi) = 0$, şi deci există o vecinătate V a punctului ξ astfel încât $|g'(x)| < k < 1, \forall x \in V$; (de obicei se ia $k = \frac{1}{2}$)

8. Fie $a > 0$ şi $p \in N$. Să se aplice metoda lui Newton pentru a construi un şir al aproximaţiilor succesive pentru $\sqrt[p]{a}$.
Fie $f(x) = x^p - a$ şi $g(x) = x - \dfrac{f(x)}{f'(x)} = \dfrac{1}{p}\left((p-1)x + ax^{1-p}\right)$. Aplicaţia g este contracţie pe un interval care conţine $\sqrt[p]{a}$:

$$|g'(x)| = \frac{p-1}{p}\;\left|\,1 - ax^{-p}\,\right| < \frac{p-1}{p} < 1\,, \forall x > \sqrt[p]{\frac{a}{2}}.$$

In concluzie, pentru a construi un şir al aproximaţiilor succesive, trebuie ca primul termen, x_0, să fie ales astfel încât $x_0 > \sqrt[p]{\frac{a}{2}}$.

9. Aplicaţie la sisteme liniare
Fie $A = (a_{ij})$ o matrice pătratică de ordinul $n \in N$ cu elemente $a_{ij} \in \mathbf{R}$, $\forall i, j \in \{1, 2, .., n\}$ şi fie $b_i \in R$, $\forall i \in \{1, 2, ..., n\}$. Pentru ce valori ale parametrului $\lambda \in R$ se poate aplica metoda aproximaţiilor succesive sistemului liniar: $x_i = \lambda \displaystyle\sum_{k=1}^{n} a_{ik} x_k + b_i$, $\forall\, 1 \leq i \leq n$?
Fie d oricare din distanţele echivalente pe $\mathbf{R^n}$ (de exemplu d_1, d_2, d_∞) şi fie

$$T : \mathbf{R^n} \mapsto \mathbf{R^n}, T(x_1, x_2, .., x_n) = \lambda \cdot \left(\sum_{k=1}^{n} a_{1k} x_k + b_1, ..., \sum_{k=1}^{n} a_{nk} x_k + b_n\right).$$

Se impune condiţia ca T să fie contracţie şi se obţin valorile cerute pentru λ. Vom face în continuare calculul pentru distanţa euclidiană, $d = d_2$ (propunem cititorului să studieze şi celelalte două cazuri). Pentru orice vectori

$x = (x_1, x_2, ..., x_n)$ şi $y = (y_1, y_2, ..., y_n)$ din $\mathbf{R^n}$, avem (aplicăm inegalitatea lui Schwarz):

$$d_2(T(x), T(y)) = |\lambda| \sqrt{\sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik}(x_k - y_k) \right)^2} \leq$$

$$\leq |\lambda| \sqrt{\sum_{i=1}^{n} \left(\left(\sum_{k=1}^{n} a_{ik}^2 \right) \left(\sum_{k=1}^{n} (x_k - y_k)^2 \right) \right)} = |\lambda| \sqrt{\sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik}^2} \cdot d_2(x, y).$$

Deci T este contracţie dacă şi numai dacă $|\lambda| \leq \left(\sqrt{\sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik}^2} \right)^{-1}$.

10. Ecuaţii integrale de tip Fredholm

Fie $a, b \in \mathbf{R}$, $a < b$ şi fie $K : [a, b] \times [a, b] \mapsto \mathbf{R}$ o funcţie continuă (numită nucleu). Pentru orice funcţie continuă $f : [a, b] \mapsto \mathbf{R}$ şi pentru orice $\lambda \in \mathbf{R}$, fie ecuaţia (cu necunoscuta ϕ):

$$\phi(x) = \lambda \int_a^b K(x, y)\phi(y)dy + f(x), \ \forall x \in [a, b].$$

Să se afle pentru ce valori ale lui $\lambda \in R$ se poate aplica metoda aproximaţiilor succesive ecuaţiei considerate.
Studiem ecuaţia în spaţiul metric complet $(\mathcal{C}([a, b]), d_\infty)$. Fie aplicaţia

$$U : \mathcal{C}([a, b]) \mapsto \mathcal{C}([a, b]), (U\phi)(x) = \lambda \int_a^b K(x, y)\phi(y)dy + f(x).$$

Ecuaţia Fredholm se scrie $U(\phi) = \phi$. Pentru orice $\phi, \psi \in \mathcal{C}([a, b])$, calculăm:

$$d_\infty(U\phi, U\psi) = \sup_{x \in [a,b]} |(U\phi)(x) - (U\psi)(x)| =$$

$$= \sup_{x \in [a,b]} |\lambda \int_a^b K(x, y)(\phi(y) - \psi(y))dy| \leq$$

$$\leq |\lambda| \sup_{x \in [a,b]} \int_a^b |K(x, y)| \cdot |\phi(y) - \psi(y)|dy \leq$$

$$\leq |\lambda| \parallel K \parallel_\infty \int_a^b |\phi(y) - \psi(y)|dy,$$

unde, am notat $\parallel K \parallel_\infty = \sup_{(x,y) \in [a,b]^2} |K(x, y)|$. Rezultă deci

$$d_\infty(U\phi, U\psi) \leq |\lambda| \parallel K \parallel_\infty (b - a)d_\infty(\phi, \psi), \ \forall \phi, \psi \in \mathcal{C}([a, b]).$$

Aplicaţia U este contracţie dacă şi numai dacă $|\lambda| < \dfrac{1}{(b-a) \parallel K \parallel_\infty}$. În această ipoteză, construim un şir al aproximaţiilor succesive: fie $\phi_0 \in \mathcal{C}([a,b])$ şi $\phi_{n+1} = U\phi_n$. Calculăm primii termeni ai şirului:

$$\phi_1(x) = (U\phi_0)(x) = \lambda \int_a^b K(x,y)\phi_0(y)dy + f(x), \forall x \in [a,b].$$

$$\phi_2(x) = (U\phi_1)(x) = \lambda \int_a^b K(x,y) \left(\lambda \int_a^b K(y,t)\phi_0(t)dt + f(y) \right) dy + f(x) =$$

$$= f(x) + \lambda \int_a^b K(x,y)f(y)dy + \lambda^2 \int_a^b \int_a^b K(x,y)K(y,t)\phi_0(t)dtdy =$$

$$= f(x) + \lambda \int_a^b K(x,y)f(y)dy + \lambda^2 \int_a^b \left(\phi_0(t) \int_a^b K(x,y)K(y,t)dy \right) dt.$$

Definim nucleele iterate:

$$K_1(x,y) = K(x,y),$$

$$K_2(x,y) = \int_a^b K(x,t)K_1(t,y)dt,$$

$$K_3(x,y) = \int_a^b K(x,t)K_2(t,y)dt,$$

....

$$K_n(x,y) = \int_a^b K(x,t)K_{n-1}(t,y)dt.$$

Cu aceste notaţii, soluţia ecuaţiei Fredholm este:

$$\xi(x) = f(x) + \sum_{n \geq 1} \lambda^n \int_a^b K_n(x,y)f(y)dy, \ \forall x \in [a,b].$$

Fie $R(x,y,\lambda) = \sum\limits_{n \geq 0} \lambda^n K_{n+1}(x,y)$ rezolventa ecuaţiei; atunci:

$$\xi(x) = f(x) + \lambda \int_a^b R(x,y,\lambda)f(y)dy.$$

1.3 MF.01.3. Noţiuni de topologie

În acest paragraf vom studia câteva tipuri remarcabile de mulţimi care se pot defini într-un spaţiu metric.

11. Definiţii
Fie (X, d) un spaţiu metric şi fie $a \in X$, fixat; pentru orice $r > 0$ definim:

$$B(a,r) = \{x \in X \; ; \; d(x,a) < r\} \text{ (bila deschisă cu centrul în } a \text{ şi rază } r)$$

$$B(a,r) = \{x \in X \; ; \; d(x,a) \leq r\} \text{ (bila închisă cu centrul în } a \text{ şi rază } r)$$

$$S(a,r) = \{x \in X \; ; \; d(x,a) = r\} \text{ (sfera cu centrul în } a \text{ şi rază } r)$$

Vecinătate a punctului a este orice mulţime care conţine o bilă centrată în a. Mulţimea vecinătăţilor lui a se va nota $\mathcal{V}_a$.

O submulţime $D \subseteq X$ se numeşte deschisă dacă $\forall a \in D, \exists\, r > 0$ astfel încât $B(a,r) \subseteq D$; rezultă că o mulţime deschisă este vecinătate pentru orice punct al ei.
O submulţime $F \subseteq X$ se numeşte închisă dacă $X \setminus D$ (complementara) este mulţime deschisă.

12. Proprietăţile mulţimilor deschise, închise
Fie (X, d) un spaţiu metric. Următoarele proprietăţi rezultă direct din definiţii.
i. $\emptyset$ şi X sunt mulţimi deschise (deci şi închise).
ii. Reuniune arbitrară de mulţimi deschise este deschisă.
iii. Intersecţie finită de mulţimi deschise este deschisă.
iv. Reuniune finită de mulţimi închise este închisă.
v. Intersecţie arbitrară de mulţimi închise este închisă.
vi. O mulţime $F \subseteq X$ este închisă dacă şi numai dacă pentru orice şir convergent $x_n \in F$ rezultă $\lim\limits_{n\to\infty} x_n \in F$.
vii. Dacă (X, d) este complet, atunci pentru orice submulţime închisă $Y \subseteq X$, spaţiul metric indus (Y, d) este complet.

13. Definiţii
O submulţime $M \subseteq X$ se numeşte mărginită dacă există $a \in X$ şi $r > 0$ astfel încât $M \subseteq B(a,r)$.

O mulţime $K \subseteq X$ se numeşte compactă dacă pentru orice familie de mulţimi deschise $(D_i)_{i\in J}$ cu proprietatea $\bigcup\limits_{i\in J} D_i \supseteq K, \exists I \subseteq J$, I finită astfel încât $\bigcup\limits_{i\in I} D_i \supseteq K$. Dacă spaţiul metric X este compact, se poate demonstra ca el este şi complet.

14. Observaţie
În cursul de analiză matematică s-a demonstrat că în $\mathbf{R^n}$ o mulţime este compactă dacă şi numai dacă este închisă şi mărginită.

14. Teorema Borel-Lebesgue
Fie (X, d) un spaţiu metric. Se poate demonstra următoarea caracterizare cu şiruri a mulţimilor compacte:
$K \subseteq X$ este mulţime compactă dacă şi numai dacă orice şir de elemente din K conţine un subşir convergent.

15 Definiţii
Fie (X, d) un spaţiu metric şi fie $A \subseteq X$.
Un punct $a \in X$ se numeşte punct de acumulare al mulţimii A dacă pentru orice $r > 0$, avem $B(a, r) \cap A \setminus \{a\} \neq \emptyset$. Caracterizarea cu şiruri a punctelor de acumulare este:

Punctul $a \in X$ este punct de acumulare al mulţimii A dacă şi numai dacă există un şir $x_n \in A$ astfel încât $x_n \neq a \, \forall n \in N$ şi $\lim\limits_{n\to\infty} x_n = a$.

Un punct $a \in X$ se numeşte punct interior lui A dacă există $r > 0$ astfel încât $A \cap B(a, r) \neq \emptyset$. Mulţimea punctelor interioare lui A se numeşte interiorul lui A şi se notează A^o. Interiorul lui A este cea mai mare submulţime deschisă inclusă în A (este egală şi cu reuniunea tuturor submulţimilor deschise incluse în A).

Un punct $a \in X$ se numeşte punct aderent lui A dacă pentru orice $r > 0$, avem $B(a, r) \cap A \neq \emptyset$. Mulţimea punctelor aderente ale lui A se numeşte aderenţa (sau închiderea) mulţimii A şi se notează $\overline{A}$. Închiderea lui A este cea mai mică mulţime închisă care o conţine pe A (este egală şi cu intersecţia tuturor mulţimilor închise care o conţin pe A).

Un punct al mulţimii A se numeşte punct izolat al lui A dacă nu este punct de acumulare al lui A. Mulţimea A se numeşte perfectă dacă este închisă şi nu conţine puncte izolate.
Mulţimile finite nu au puncte de acumulare. În plus, se poate demonstra că orice mulţime perfectă este nenumărabilă.

Spaţiul metric X se numeşte conex dacă nu există două submulţimi simultan deschise (sau închise) D_1 şi D_2 cu proprietăţile:

$$D_1 \neq \emptyset,\ D_2 \neq \emptyset,\ X = D_1 \cup D_2 \text{ şi } D_1 \cap D_2 = \emptyset.$$

O submulţime $A \subseteq X$ se numeşte conexă dacă spaţiul metric (indus) (A, d) este conex. Rezultă următoarea caracterizare a submulţimilor conexe: o submulţime $A \subseteq X$ este conexă dacă şi numai dacă nu există două submulţimi deschise D_1 şi D_2 astfel încât $D_1 \cap D_2 = \emptyset$, $D_1 \cap A \neq \emptyset$, $D_2 \cap A \neq \emptyset$ şi $A \subseteq D_1 \cup D_2$.
În mulţimea numerelor reale o mulţime este conexă dacă şi numai dacă este

interval.

1.4 MF.01.4. Funcţii continue

16. Definiţie
Fie (X, d) şi (Y, d') două spaţii metrice şi fie $a \in X$. O aplicaţie $f : X \mapsto Y$ se numeşte **continuă** în a dacă pentru orice vecinătate $V \in \mathcal{V}_{f(a)}$, există o vecinătate $U \in \mathcal{V}_a$ astfel încât $f(U) \subseteq V$.
Funcţia f se numeşte continuă (sau continuă pe X) dacă este continuă în orice $a \in X$.
Ca şi în cazul real, există caracterizări echivalente pentru continuitate:

17. Teoremă
Cu notaţiile de mai sus, următoarele afirmaţii sunt echivalente:
(a) Funcţia f este continuă în a.
(b) $\forall \varepsilon > 0$, $\exists \delta_\varepsilon > 0$ astfel încât dacă $d(x, a) < \delta_\varepsilon$, atunci $d'(f(x), f(a)) < \varepsilon$.
(c) Pentru orice şir x_n de elemente din X cu proprietatea $x_n \to a$, rezultă $f(x_n) \to f(a)$.
Demonstraţie (a) $\Rightarrow$ (b) Fie $\varepsilon > 0$ şi fie $V = B(f(a), \varepsilon) \in \mathcal{V}_{f(a)}$. Deoarece f este continuă în a, există $U \in \mathcal{V}_a$ astfel încât $f(U) \subseteq V$; alegem $\delta_\varepsilon > 0$ cu proprietatea $B(a, \delta_\varepsilon) \subseteq U$, ceea ce încheie demonstraţia.
(b) $\Rightarrow$ (c) Fie $x_n \to a$ şi fie $\varepsilon > 0$ fixat. Conform lui (b), există $\delta_\varepsilon > 0$ astfel încât dacă $d(x, a) < \delta_\varepsilon$, atunci $d'(f(x), f(a)) < \varepsilon$. Deoarece $x_n \to a$, de la un anumit rang $N_\varepsilon \in \mathbf{N}$ avem $d(x_n, a) < \delta_\varepsilon, \forall n \geq N_\varepsilon$. Rezultă $d'(f(x_n, f(a)) < \varepsilon$, $\forall n \geq N_\varepsilon$, deci $f(x_n) \to f(a)$.
(c) $\Rightarrow$ (a) Presupunem prin absurd că f nu este continuă în a, deci există $V \in \mathcal{V}_{f(a)}$ astfel încât pentru orice $U \in \mathcal{V}_a$ să avem: $f(U) \not\subseteq V$. Fie, pentru orice $n \in \mathbf{N}$, $U = B(a, \frac{1}{n})$ şi fie $x_n \in U$ astfel încâ $f(x_n) \notin V$; am construit deci un şir x_n care tinde la a, dar $f(x_n) \not\to f(a)$, contradicţie.

18. Observaţie
Din (c) de mai sus, rezultă imediat: compunerea a dou

Folosind caracterizările echivalente de mai sus, se demonstrează următoarele proprietăţi ale funcţiilor continue:

18. Teoremă
Fie (X, d) şi (Y, d') două spaţii metrice şi fie $f : X \mapsto Y$; atunci următoarele afirmaţii sunt echivalente:
(a) f este continuă.
(b) $f(\overline{A}) \subseteq \overline{f(A)}$, pentru orice $A \subseteq X$.
(c) $f^{-1}(D)$ este mulţime deschisă (în X) pentru orice mulţime deschisă $D \subseteq Y$.

(d) $f^{-1}(F)$ este mulţime închisă (în X) pentru orice mulţime închisă $F \subseteq Y$.

19. Exemple
(a) Funcţiile constante sunt continue.
(b) Proiecţiile canonice $p_j : \mathbf{R^n} \mapsto \mathbf{R}$, $p_j(x_1, x_2, ..., x_n) = x_j$ sunt continue (pentru orice $j = 1, 2, ..., n$). De aici rezultă că o funcţie $f : \mathbf{R^n} \mapsto \mathbf{R^m}$ este continuă dacă şi numai dacă este continuă "pe componente".

Din analiza reală, ştie că imaginea printr-o funcţie continuă a unui interval compact din $\mathbf{R}$ este de asemenea interval compact. În continuare vom studia generalizări ale acestui rezultat pentru funcţii continue pe spaţii metrice.

20. Teoremă
Fie (X, d) şi (Y, d') două spaţii metrice şi fie $f : X \mapsto Y$ o funcţie continuă. Dacă $A \subseteq X$ este o mulţime compactă, atunci $f(A)$ este compactă.
Demonstraţie Fie $(D_j)_{j\in J}$ o acoperire cu mulţimi deschise a lui $f(A)$, deci $f(A) \subseteq \cup_{j\in J} D_j$. Rezultă $A \subseteq f^{-1}(\cup_{j\in J} D_j) = \cup_{j\in J} f^{-1}(D_j)$, iar $f^{-1}(D_j)$ sunt mulţimi deschise pentru orice $j \in J$ (conform Teoremei 18). Rezultă că $(f^{-1}(D_j))_{j\in J}$ este o acoperire cu mulţimi deschise a mulţimii compacte A, deci există o submulţime finită $I \subseteq J$ astfel încât $A \subseteq \cup_{j\in I} f^{-1}(D_j)$. Aplicând f, obţinem:

$$f(A) \subseteq f(\cup_{j\in I} f^{-1}(D_j)) = \cup_{j\in I} f(f^{-1}(D_j)) \subseteq \cup_{j\in I} D_j,$$

şi deci $f(A)$ este mulţime compactă.

O consecinţă importantă se referă la funcţiile continue cu valori reale:
21. Corolar
Fie (X, d) un spaţiu metric compact şi fie $f : X \mapsto \mathbf{R}$ o funcţie continuă. Atunci funcţia f este mărginită (i.e. mulţimea $f(X)$ este mărginită) şi îşi atinge marginile, i.e. există $\xi, \eta \in X$ astfel încât $\sup_{x\in X} f(x) = f(\xi)$ şi $\inf_{x\in X} f(x) = f(\eta)$.
Demonstraţie Conform teoremei 20, mulţimea $f(X)$ este compactă (în $\mathbf{R}$) şi deci este închisă şi mărginită; rezultă că f este mărginită şi în plus, deoarece $f(X)$ este ňchisă, rezultă că $\sup_{x\in X} f(x)$ şi $\inf_{x\in X} f(x)$ sunt elemente din $f(X)$.

22. Teoremă
Fie (X, d) şi (Y, d') două spaţii metrice şi fie $f : X \mapsto Y$ o funcţie continuă. Dacă X este mulţime conexă, atunci imaginea $f(X)$ este mulţime conexă.
Demonstraţie Presupunem prin absurd că $f(X)$ nu este conexă; atunci există mulţimile deschise şi disjuncte U şi V astfel încât $f(X) \cap U$ şi $f(X) \cap V$ sunt nevide şi $f(X) \subseteq U \cap V$. Fie $S = f^{-1}(U)$ şi $T = f^{-1}(V)$; atunci S

şi T sunt nevide, deschise, disjuncte şi $X = S \cup T$, deci X nu este conex, contradicţie.

Pentru funcţii continue cu valori reale, obţinem:

23. Corolar
Fie (X, d) un spaţiu metric conex şi fie $f : X \mapsto \mathbf{R}$ o funcţie continuă astfel încât există $a, b \in X$ cu proprietatea $f(a)f(b) < 0$; atunci există $\xi \in X$ astfel încât $f(\xi) = 0$.
Demonstraţie Mulţimea $f(X)$ fiind conexă în $\mathbf{R}$, este un interval care conţine numerele de semne contrare $f(a)$ şi $f(b)$, deci $0 \in f(X)$.

Capitolul 2

MF.02. Spaţii normate

Cuvinte cheie:

normă, spaţiu normat, operator liniar, funcţională liniară, spaţiu dual.

În acest capitol vom prezenta noţiuni şi rezultate generale din teoria spaţiilor normate. Vom presupune cunoscute conceptele uzuale din teoria spaţiilor vectoriale şi a spaţiilor metrice. Sursele bibliografice recomandate sunt: [C02], [D01], [H02], [O01], [S02].

2.1 MF.02.1. Noţiuni generale

1. Definiţie
Fie X un spaţiu vectorial complex.
O aplicaţie $\| \|: X \longrightarrow [0\,,\,\infty)$ cu proprietăţile:
(a) $\| x+y \| \leq \| x \| + \| y \|$
(b) $\| \alpha x \| = |\alpha| \; \| x \|$
(c) $\| x \| = 0 \Longleftrightarrow x = 0$,
pentru orice $x\,,\, y \in X$ şi $\alpha \in \mathbf{C}$, se numeşte **normă**. O aplicaţie care verifică doar condiţiile **(a)** şi **(b)** se numeşte **seminormă**.
Perechea $(X\,,\| \|)$ se numeşte **spaţiu normat**. Orice spaţiu normat este şi spaţiu metric, distanţa dintre x şi y fiind, prin definiţie, $d(x,y) = \| x-y \|$. Dacă în plus orice şir Cauchy este convergent, atunci $(X\,,\,\| \|)$ se numeşte spaţiu Banach (sau spaţiu normat complet). Se poate demonstra uşor că operaţiile algebrice sunt continue: dacă $\lim\limits_{n\to\infty} x_n = x$ şi $\lim\limits_{n\to\infty} y_n = y$, atunci $\lim\limits_{n\to\infty}(x_n+y_n) = x+y$ şi analog pentru înmulţirea cu scalari. Dacă X este un spaţiu vectorial şi $\| \|_1\,,\| \|_2$ sunt două norme pe X, atunci ele se numesc echivalente dacă există c şi k două constante pozitive astfel încât $\| x \|_1 \leq c \| x \|_2 \leq k \| x \|_1$; în acest caz $\lim\limits_{n\to\infty} x_n = x$ în $\| \|_1 \Longleftrightarrow \lim\limits_{n\to\infty} x_n = x$

în $\| \|_2$, deci structurile topologice coincid.
Fie X şi Y două spaţii vectoriale; o aplicaţie $T : X \to Y$ se numeşte **aplicaţie liniară** (sau **operator liniar**) dacă:

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty, \, \forall x, y \in X, \, \forall \alpha, \beta \in \mathbf{C}.$$

Spaţiile vectoriale X şi Y se numesc izomorfe dacă există $T : X \to Y$ o aplicaţie liniară şi bijectivă. În acest caz, T se numeşte izomorfism de spaţii vectoriale. Este simplu de demonstrat că inversul unui izomorfism de spaţii vectoriale este de asemenea operator liniar.
Două spaţii normate $(X, \| \|_1)$ şi $(Y, \| \|_2)$ se numesc spaţii normate izomorfe dacă există între ele un izomorfism F de spaţii vectoriale cu proprietatea $\| F(x) \|_2 = \| x \|_1$; în acest caz, F se numeşte izomorfism de spaţii normate. Dacă aplicaţia liniară F verifică egalitatea $\| F(x) \|_2 = \| x \|_1$, (fără a fi neapărat bijectivă), atunci F se numeşte izometrie liniară.
O noţiune importantă care se poate defini într-un spaţiu normat este cea de serie convergentă.

2. Definiţie
Fie $(X, \| \|)$ un spaţiu normat şi fie $(x_n)_{n \in N}$ un şir de elemente din X. Spunem că seria $\sum\limits_{n \in N} x_n$ este convergentă la $x \in X$ (numit în acest caz suma seriei) dacă şirul sumelor parţiale, $s_n = \sum\limits_{k=1}^{n} x_k$ converge la x. Seria $\sum\limits_{n \in N} x_n$ se numeşte absolut convergentă dacă seria (de numere reale şi pozitive) $\sum\limits_{n \in N} \| x_n \|$ este convergentă. Cu o demonstraţie asemănătoare celei de la serii de numere reale se poate arăta că într-un spaţiu Banach orice serie absolut convergentă este convergentă, reci proca fiind, în general, falsă. Este interesant faptul că această proprietate caracterizează spaţiile normate complete.

3. Propoziţie
Un spaţiu normat $(X, \| \|)$ este complet dacă şi numai dacă pentru orice şir $(x_n)_n \subset X$ cu proprietatea că seria $\sum\limits_{n \in N} \| x_n \|$ este convergentă, rezultă că seria $\sum\limits_{n \in N} x_n$ este convergentă.
Demonstraţie. Să presupunem că $(X, \| \|)$ este un spaţiu normat în care este verificată ipoteza din enunţ şi fie $(x_n)_n \subset X$ un şir Cauchy. Fie $k \in \mathbf{N}$ şi fie $n_k \in \mathbf{N}$ astfel încât pentru orice $i, j \geq n_k$ să avem $\| x_i - x_j \| < 2^{-k}$. Dacă definim $y_1 = x_1$ şi $y_k = x_{n_k} - x_{n_{k-1}}$, pentru orice $k \in N$ şi $k \geq 2$, atunci seria $\sum\limits_{k \in N} \| y_k \|$ este convergentă. Din ipoteză rezultă că seria $\sum\limits_{k \in N} y_k$ este convergentă şi deci şirul $(x_{n_k})_{k \in N}$ converge la un element $x \in X$. Este uşor de arătat că şirul $(x_n)_n$ converge la x. Cealaltă implicaţie o lăsăm ca

exerciţiu.

Încheiem acest paragraf cu o listă de spaţii Banach ce vor fi citate frecvent în continuare.

4. Exemple
(i) Fie $n \in \mathbf{N}$ fixat şi fie

$$\mathbf{C^n} = \{x = (x_1, x_2, .., x_n)\,;\, x_j \in \mathbf{C},\, \forall j = 1, 2, .., n\}$$

cu structura uzuală de spaţiu vectorial. Cu **norma euclidiană**:

$$\| x \|_2 = \sqrt{\sum_{j=1}^{n} |x_j|^2},$$

$\mathbf{C^n}$ este spaţiu Banach; propunem cititorului ca exerciţiu afirmaţia că următoarele două norme sunt echivalente cu norma
euclidiană:

$$\| x \|_1 = \sum_{j=1}^{n} |x_j|,$$

$$\| x \|_\infty = \max\{|x_j|\,, 1 \leq j \leq n\}.$$

Cu aceleaşi norme, şi $\mathbf{R^n}$ este spaţiu Banach.

(ii) Spaţiile $\ell^p(\mathbf{Z})$ şi $\ell^p(\mathbf{N})$
Fie $p \in R\,,\, p \geq 1$, fixat şi fie

$$\ell^p(Z) = \{x : Z \to C\,;\, \sum_{n \in Z} |x(n)|^p \text{ este convergentă}\}.$$

Facem precizarea că dacă $(a_n)_{n \in Z}$ este un şir de numere complexe indexat după $\mathbf{Z}$, atunci seria $\sum\limits_{n \in Z} a_n$ este convergentă dacă seriile $\sum\limits_{n=-\infty}^{0} a_n$ şi $\sum\limits_{n=1}^{\infty} a_n$ sunt amândouă convergente.
Este evident că pentru orice $\alpha \in \mathbf{C}$ şi $x \in \ell^p(\mathbf{Z})$, şirul $(\alpha x)(n) = \alpha x(n)$ este în $\ell^p(\mathbf{Z})$. Fie acum $x, y \in \ell^p(\mathbf{Z})$ şi fie $(x + y)(n) = x(n) + y(n)$. Notând

$$\| x \|_p = \left(\sum_{n \in Z} |x(n)|^p \right)^{\frac{1}{p}},$$

atunci, din inegalitatea lui Minkovski, (a se vedea cap. 5), rezultă inegalitatea $\| x + y \|_p \leq \| x \|_p + \| y \|_p$; celelalte proprietăţi din definiţia spaţiului normat sunt evidente. Demonstrăm în continuare completitudinea. Dacă $(x^k)_{k \in N}$ este un şir Cauchy în $\ell^p(\mathbf{Z})$ şi dacă $\epsilon > 0$, atunci există $k_\epsilon \in \mathbf{N}$ astfel încât $\| x^k - x^j \|_p < \epsilon$, $\forall k, j \geq k_\epsilon$. De aici rezultă:

$$\sum_{n \in Z} |x^k(n) - x^j(n)|^p < \epsilon^p , \forall k, j \geq k_\epsilon, \tag{2.1}$$

şi deci pentru orice $n \in \mathbf{N}$, avem $|x^k(n) - x^j(n)| < \epsilon, \forall k, j \geq k_\epsilon$, ceea ce arată că şirul $(x^k(n))_{k \in N}$ este şir Cauchy (de numere complexe) pentru orice $n \in \mathbf{N}$, deci este convergent; fie $x(n) = \lim_{k \to \infty} x^k(n)$. Din relaţia (1.1), pentru $j \to \infty$, rezultă $\sum_{n \in Z} |x^k(n) - x(n)|^p \leq \epsilon^p, \forall k \geq k_\epsilon$, adică $x^k - x \in \ell^p(\mathbf{Z})$, $\forall k \geq k_\epsilon$. De aici rezultă că $x \in \ell^p(Z)$ şi $\| x^k - x \|_p \leq \epsilon, \forall k \geq k_\epsilon$, ceea ce încheie demonstraţia.
Analog se definesc spaţiile $\ell^p(\mathbf{N}) = \{x : \mathbf{N} \to \mathbf{C}\, ;\, \sum_{n=0}^{\infty} |x(n)|^p < \infty\}$.

(iii) Spaţiul $\ell^\infty(\mathbf{Z})$
Fie $\ell^\infty(\mathbf{Z}) = \{x : Z \to C\, ;\, x$ şir mărginit $\}$. Cu operaţiile uzuale, definite în exemplul anterior, $\ell^\infty(\mathbf{Z})$ este spaţiu vectorial. Este uşor de arătat că aplicaţia $\| x \|_\infty = \sup_{n \in Z} |x(n)|$ este normă. Pentru a demonstra completitudinea, să considerăm $(x^k)_{k \in N}$ un şir Cauchy în $\ell^\infty(Z)$ şi $\epsilon > 0$. Atunci există $k_\epsilon \in N$ asfel încât $\| x^k - x^j \|_\infty < \epsilon$, $\forall k, j \geq k_\epsilon$, şi deci şirul $(x^k(n))_{k \in N}$ este şir Cauchy (de numere complexe) pentru orice $n \in Z$. Există deci $x(n) = \lim_{k \to \infty} x^k(n)$. Şirul x astfel construit este în $\ell^\infty(\mathbf{Z})$ şi este limita (în $\ell^\infty(\mathbf{Z})$) a lui $(x^k)_{k \in N}$:

$$\| x^k - x \|_\infty = \sup_{n \in Z} \lim_{j \to \infty} |x^k(n) - x^j(n)| \leq \epsilon.$$

Analog se defineşte spaţiul $\ell^\infty(\mathbf{N})$.

(iv) Spaţiul funcţiilor continue $\mathcal{C}(\mathcal{D})$
Fie $\mathcal{D}$ un spaţiu metric compact şi fie

$$\mathcal{C}(\mathcal{D}) = \{f : \mathcal{D} \to \mathbf{C}\, ;\, f \text{ continuă}\}.$$

Cu operaţiile uzuale, $\mathcal{C}(\mathcal{D})$ este spaţiu vectorial. Structura de spaţiu Banach este definită de norma supremum:

$$\| f \|_\infty = \sup_{t \in \mathcal{D}} |f(t)|.$$

Dacă $\mathcal{D} = [a, b] \subseteq \mathbf{R}$, atunci obţinem spaţiul $\mathcal{C}[a, b]$ (cu topologia convergenţei uniforme).

Alte exemple importante de spaţii normate sunt spaţiile de funcţii integrabile prezentate în capitolul 5.

2.2 MF.02.2. Operatori liniari

5. Definiţie
Fie $(X, \| \|)$ şi $(Y, \| \|)$ două spaţii normate (complexe). Reamintim că un **operator** $T : X \to Y$ se numeşte **liniar** dacă

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)\,,\ \forall \alpha, \beta \in C,\ \forall x, y \in X.$$

Operatorul T se numeşte continuu în punctul $x_o \in X$ dacă pentru orice $\epsilon > 0$ există $\delta > 0$ astfel încât pentru orice $x \in X$ cu proprietatea $\| x - x_o \| < \delta$, să avem $\| Tx - Tx_o \| < \epsilon$, sau, într-o formulare echivalentă (cu şiruri), dacă $\forall (x_n)_n \subset X$ cu proprietatea $\lim\limits_{n\to\infty} x_n = x_o$, rezultă $\lim\limits_{n\to\infty} Tx_n = Tx_o$. Operatorul T se numeşte continuu dacă este continuu în orice punct. Vom nota mulţimea operatorilor liniari şi continui de la X în Y cu $\mathcal{L}(X, Y)$. Dacă $X = Y$, vom nota această mulţime cu $\mathcal{L}(X)$. Aşa cum vom arăta în capitolul 3, orice operator liniar pe C^n este continuu. Pe spaţii normate infinit dimensionale, există operatori liniari care nu sunt continui.

6. Observaţie
Mulţimea $\mathcal{L}(X, Y)$ se poate organiza ca spaţiu vectorial cu operaţiile uzuale: $(T + S)(x) = Tx + Sx$ şi $(\alpha T)x = \alpha Tx$, pentru orice operatori $T, S \in \mathcal{L}(X)$ şi pentru orice $x \in X$, $\alpha \in C$.
Vom nota cu O operatorul nul şi cu $-T$ opusul lui T.

7. Propoziţie
Fie $(X, \| \|)$ şi $(Y, \| \|)$ două spaţii normate şi fie $T : X \to Y$ un operator liniar. Următoarele afirmaţii sunt echivalente:
(a) T este continuu în 0.
(b) T este continuu.
(c) Există $M > 0$ astfel încât $\| Tx \| \leq M \ \| x \|$, $\forall x \in X$.
(d) Pentru orice submulţime mărginită $A \subseteq X$, submulţimea $T(A) \subseteq Y$ este de asemenea mărginită.
Demonstraţie Implicaţia **(a)** $\Rightarrow$ **(b)** o propunem ca exerciţiu.
(b) $\Rightarrow$ **(c)**. Din continuitatea lui T în 0, (şi $T(0) = 0$), rezultă că pentru orice $\epsilon > 0$, există $\delta > 0$ astfel încât $\| Ty \| \leq \epsilon$, $\forall y \in X$ cu proprietatea $\| y \| \leq \delta$. Fie $x \in X$; atunci, scriind inegalitatea de mai sus pentru $\epsilon = 1$ şi $y = \delta \| x \|^{-1} x$, obţinem $\| Tx \| \leq \frac{1}{\delta} \| x \|$.
Implicaţia **(c)** $\Rightarrow$ **(d)** este şi ea evidentă.
(d) $\Rightarrow$ **(a)**. Fie $\epsilon > 0$ şi fie $\delta > 0$ astfel încât $\| \frac{\delta}{\epsilon} x \| \leq 1$. Din ipoteza **(d)** rezultă $\| T(\frac{\delta}{\epsilon} x) \| \leq \delta$; în concluzie, dacă $\| x \| \leq \frac{\epsilon}{\delta}$, rezultă că $\| Tx \| \leq \epsilon$, ceea ce arată că T este continuu în origine.

8. Definiţie

Pentru orice $T \in \mathcal{L}(X,Y)$, definim:

$$\| T \| = \inf\{M > 0\,;\, \| Tx \| \leq M \| x \|\,,\, \forall x \in X\};$$

$$\| T \|_1 = \sup\{\| Tx \|\,;\, x \in X \text{ şi } \| x \| = 1\};$$

$$\| T \|_2 = \sup\{\| Tx \|\,;\, x \in X \text{ şi } \| x \| \leq 1\}.$$

Să observăm că buna definiţie a lui $\| \, \|$ este asigurată de punctul (c) din propoziţia anterioară. Facem de asemenea precizarea că notaţiile $\| T \|_1$ şi $\| T \|_2$ vor fi folosite numai în cursul demonstraţiei lemei următoare.

9. Lemă
(a) Pentru orice $T \in \mathcal{L}(X,Y)$, avem $\| T \| = \| T \|_1 = \| T \|_2$.
(b) Aplicaţia $T \to \| T \|$ este o normă pe spaţiul vectorial $\mathcal{L}(X,Y)$.
Demonstraţie (a) Demonstrăm mai întâi inegalitatea:

$$\| Tx \| \leq \| T \| \| x \|\,,\, \forall x \in X. \tag{2.2}$$

Fie $\mathcal{D} = \{M > 0\,;\, \| Tx \| \leq M \| x \|\,,\, \forall x \in X\}$. Deoarece $\| T \| = \inf \mathcal{D}$, rezultă că există un şir $M_n \in \mathcal{D}$ astfel încât $\| T \| = \lim\limits_{n\to\infty} M_n$ şi $\| T(x) \| \leq M_n \| x \|$. Din aceste două relaţii, rezultă, pentru $n \to \infty$, inegalitatea (3.1). Folosind acum (3.1), obţinem:

$$\| T \|_2 = \sup\{\| Tx \|\,;\, x \in X \text{ şi } \| x \| \leq 1\} \leq$$

$$\leq \sup\{\| T \| \| x \|\,;\, x \in X \text{ şi } \| x \| \leq 1\} = \| T \|,$$

şi deci am demonstrat :

$$\| T \|_2 \leq \| T \| \,. \tag{2.3}$$

Pentru orice $x \in X$, $x \neq 0$, vectorul $u = \| x \|^{-1} x$ are normă 1 şi deci:

$$\frac{1}{\| x \|} \| Tx \| = \| Tu \| \leq \sup\{\| Ty \|\,;\, y \in X \text{ şi } \| y \| = 1\} = \| T \|_1,$$

şi deci am demonstrat inegalitatea:

$$\| Tx \| \leq \| T \|_1 \| x \|, \tag{2.4}$$

pentru orice $x \in X$; (pentru $x = 0$, (3.3) este evidentă). Din (3.3) rezultă că $\| T \|_1 \in \mathcal{D}$ şi deci din definiţia lui $\| T \|$, rezultă:

$$\| T \|_1 \geq \| T \| \,. \tag{2.5}$$

Cum inegalitatea $\| T \|_1 \leq \| T \|_2$ este evidentă, din (3.2) şi (3.4) rezultă egalitatea de la punctul **(a)**.

(b) Dacă $\| T \| = 0$, atunci $Tu = 0\,, \forall u \in X$ cu proprietatea $\| u \| = 1$; fie $x \in X$, oarecare. Atunci $u = \| x \|^{-1} x$ are normă 1 şi deci Tu=0, adică $Tx = 0$ şi deci $T = O$. Celelalte 2 proprietăţi ale normei rezultă din proprietăţile corespunzătoare ale normelor din X şi Y.

9. Definiţie
Din Lema anterioară rezultă că $\mathcal{L}(X, Y)$ poate fi organizat ca spaţiu normat cu norma din definiţia 8. Atunci când nu se va specifica în mod explicit contrariul, pe spaţiul $\mathcal{L}(X, Y)$ se va subînţelege topologia definită de această normă.

10. Exemplu

Fie, pe spaţiul Banach al funcţiilor continue $\mathcal{C}[a, b]$, operatorul de înmulţire cu variabila independentă:

$$T : \mathcal{C}[a, b] \mapsto \mathcal{C}[a, b],\ (Tf)(t) = tf(t),\ \forall t \in [a, b].$$

Se demonstrează că T este operator liniar şi continuu, având norma $\| T \| = b$.

11. Teoremă
Fie $(X, \| \, \|)$ şi $(Y, \| \, \|)$ spaţii normate. Dacă Y este complet, atunci $\mathcal{L}(X, Y)$ este spaţiu Banach.
Demonstraţie Fie T_n un şir Cauchy în $\mathcal{L}(X, Y)$, deci pentru orice $\epsilon > 0$, există $n_\epsilon \in N$ astfel încât

$$\| T_n - T_m \| < \epsilon\,, \forall n, m \geq n_\epsilon. \tag{2.6}$$

Din inegalitatea (3.1), aplicată operatorului $T_n - T_m$ şi din (3.5), rezultă:

$$\| T_n x - T_m x \| < \epsilon \| x \|\,, \forall n, m \geq n_\epsilon\,, \forall x \in X. \tag{2.7}$$

Din inegalitatea (3.6) rezultă că pentru orice $x \in X$ şirul $(T_n x)_{n \in N}$ este şir Cauchy în spaţiul Y, care, conform ipotezei, este complet. Fie deci $Tx = \lim\limits_{n \to \infty} T_n x$, pentru orice $x \in X$. Liniaritatea operatorului T astfel definit este imediată. Demonstrăm acum continuitatea lui T. Din inegalitatea (3.6), pentru $m \to \infty$ rezultă:

$$\| T_n x - Tx \| \leq \epsilon \| x \|,\ \forall x \in X,\ \forall n \geq n_\epsilon. \tag{2.8}$$

Din propoziţia 3 (**b**$\Leftrightarrow$**c**) şi din inegalitatea (3.7) rezultă că operatorul $T_n - T$ este continuu şi deci $T = T_n - (T_n - T) \in \mathcal{L}(X, Y)$. Tot din (3.7) şi din definiţia normei în $\mathcal{L}(X, Y)$ rezultă că $\| T_n - T \| \leq \epsilon$, $\forall n \in N$, ceea ce arată că şirul T_n este convergent la T.

12. Definiţie
Fie X un spaţiu normat. Orice aplicaţie $f : X \to \mathbf{C}$ se numeşte funcţională (pe spaţiul X). Mulţimea funcţionalelor liniare şi continue se notează cu X' şi se numeşte **dualul** lui X. Deoarece spaţiul $\mathbf{C}$ este complet, din teorema precedentă (pentru $Y = \mathbf{C}$) rezultă că X' este spaţiu Banach. Evident, aceeaşi construcţie i se poate aplica şi spaţiului X'; se obţine spaţiul Banach X'', (numit al doilea dual al lui X, sau bidualul). Dacă $x \in X$, atunci aplicaţia $\phi_x : X' \to \mathbf{C}$, $\phi_x(f) = f(x)$, este în X'' şi $\| \phi_x \| = \| x \|$, ([4],p.120). In felul acesta, orice spaţiu Banach X este izomorf (în mod canonic) cu un subspaţiu din X''. Dacă în plus aplicaţia $X \ni x \to \phi_x \in X''$, este un izomorfism de spaţii Banach, atunci X se numeşte spaţiu reflexiv (în general această aplicaţie nu este surjectivă).

Nu există o teoremă de reprezentare a funcţionalelor liniare şi continue pe un spaţiu Banach arbitrar; dăm în continuare caracterizările dualelor unor spaţii Banach uzuale.

13. Exemple
(i) Fie $\alpha \in \ell^\infty(\mathbf{Z})$; atunci aplicaţia

$$f_\alpha : \ell^1(\mathbf{Z}) \to \mathbf{C},\ f_\alpha(x) = \sum_{n \in Z} \alpha(n) x(n)$$

este o functională liniară şi continuă pe spaţiul $\ell^1(\mathbf{Z})$ cu proprietatea

$$\| f_\alpha \| = \| \alpha \|_\infty .$$

Reciproc, pentru orice funcţională liniară şi continuă f pe spaţiul $\ell^1(\mathbf{Z})$, există $\alpha \in \ell^\infty(\mathbf{Z})$ astfel încât $f = f_\alpha$.
Demonstraţie Fie $\alpha \in \ell^\infty(\mathbf{Z})$ şi fie f_α ca în enunţ; este evident că f_α este liniară. Continuitatea sa rezultă din inegalitatea:

$$|f_\alpha(x)| = |\sum_{n \in Z} \alpha(n) x(n)| \leq \left(\sum_{n \in Z} |x(n)| \right) \sup_{n \in Z} |\alpha(n)|,\ \forall x \in \ell^1(Z).$$

Tot de aici rezultă şi inegalitatea $\| f_\alpha \| \leq \| \alpha \|_\infty$. Demonstrăm acum inegalitatea inversă. Pentru aceasta, fie $\sigma_k \in \ell^1(\mathbf{Z})$ definit prin:

$$\sigma_k(n) = \begin{cases} 1 & \text{dacă} \quad n = k \\ 0 & \text{dacă} \quad n \neq k \end{cases}$$

Evident, $\| \sigma_k \|_1 \leq 1$, şi deci

$$\| f_\alpha \| = \sup\{ |f_\alpha(y)| \, ; \, \| y \|_1 \leq 1 \} \geq |f_\alpha(\sigma_k)| = |\alpha(k)|.$$

Luând acum supremumul după k, rezultă $\| f_\alpha \| \geq \| \alpha \|_\infty$.
Demonstrăm acum afirmaţia reciprocă; fie f o funcţională liniară şi continuă

pe spaţiul $\ell^1(\mathbf{Z})$ şi fie σ_k definit mai sus.
Pentru orice $x \in \ell^1(\mathbf{Z})$, fie seria (de elemente din $\ell^1(\mathbf{Z})$), $\sum\limits_{k \in Z} x(k)\sigma_k$. Este uşor de arătat că această serie converge în spaţiul $\ell^1(\mathbf{Z})$ la şirul x, deci

$$x = \sum_{k \in Z} x(k)\sigma_k.$$

Rezultă deci (folosind liniaritatea şi continuitatea lui f):

$$f(x) = \sum_{k \in Z} x(k)f(\sigma_k),\ \forall x \in \ell^1(Z).$$

Definim şirul $\alpha(k) = f(\sigma_k)$, $\forall k \in Z$; rezultă că $\| \alpha \|_\infty \leq \| f \|$ (deci α este mărginit) şi $f = f_\alpha$.
În mod analog se poate identifica şi dualul spaţiului $\ell^1(\mathbf{N})$ cu $\ell^\infty(\mathbf{N})$.

Dăm în continuare, fără demonstraţie, caracterizarea dualului spaţiului $\ell^p(\mathbf{Z})$.

(ii) Dualul spaţiului $\ell^p(\mathbf{Z})$
Fie $p > 1$ şi fie $q > 1$ astfel încât $\frac{1}{p} + \frac{1}{q} = 1$. Dacă $\alpha \in \ell^q(Z)$, atunci aplicaţia

$$f_\alpha : \ell^p(\mathbf{Z}) \to \mathbf{C},\ f_\alpha(x) = \sum_{n \in Z} \alpha(n)x(n),$$

este o funcţională liniară şi continuă cu proprietatea $\| f \| = \| \alpha \|_q$.
Reciproc, pentru orice funcţională liniară şi continuă f pe spaţiul $\ell^p(\mathbf{Z})$, există $\alpha \in \ell^q(\mathbf{Z})$ astfel încât $f = f_\alpha$.

14. Observaţie
Convergenţa în spaţiul normat $\mathcal{L}(X,Y)$ se numeşte **convergenţă uniformă**. Pe mulţimea $\mathcal{L}(X,Y)$ se mai pot defini şi alte tipuri de convergenţă; spunem că şirul de operatori $T_n \in \mathcal{L}(X,Y)$ converge **punctual** (sau **tare-operatorial**) la $T \in \mathcal{L}(X,Y)$ dacă

$$\lim_{n\to\infty} T_n x = Tx\ ,\ \forall x \in X;$$

spunem că şirul $T_n \in \mathcal{L}(X,Y)$ converge **slab-operatorial** la operatorul $T \in \mathcal{L}(X,Y)$ dacă

$$\lim_{n\to\infty} f(T_n x) = f(Tx)\ ,\ \forall x \in X\ ,\ \forall f \in X'.$$

Lăsăm ca exerciţiu cititorului afirmaţiile: convergenţa uniformă o implică pe cea punctuală, iar aceasta pe cea slabă, reciprocele fiind, în general, false.

15. Definiţie
Fie X, Y, Z trei spaţii normate şi fie $T \in \mathcal{L}(X, Y)$ şi $S \in \mathcal{L}(Y, Z)$. Operatorul $ST \in \mathcal{L}(X, Z)$, definit prin $(ST)(x) = S(Tx)$ se numeşte **produsul** operatorilor S şi T. Să mai observăm că din inegalităţile :

$$\| STx \| \leq \| S \| \| Tx \| \leq \| S \| \| T \| \| x \|, \forall x \in X,$$

rezultă $\| ST \| \leq \| S \| \| T \|$. Să considerăm acum un operator $T \in \mathcal{L}(X)$. El se numeşte **inversabil** dacă există un alt operator $T^{-1} \in \mathcal{L}(X)$ astfel încât $TT^{-1} = T^{-1}T = I$, unde, I este operatorul identic, adică $Ix = x$, $\forall x \in X$. Este simplu de observat că dacă T este bijectiv, atunci T^{-1} este liniar; în general însă, operatorul T^{-1} nu este continuu; dacă $X = C^n$ este un spaţiu finit dimensional, atunci T^{-1} este continuu în virtutea faptului că pe spaţii finit dimensionale orice aplicaţie liniară este şi continuă, (un rezultat care va fi demonstrat în capitolul următor).

Prezentăm în continuare două rezultate referitoare la inversabilitatea operatorilor liniari şi continui.

16. Propoziţie
Fie X, Y două spaţii normate şi fie $T \in \mathcal{L}(X, Y)$. Dacă T este bijectiv, atunci următoarele afirmaţii sunt echivalente:
(a) T^{-1} este operator continuu.
(b) Există $m > 0$ astfel încât $\| Tx \| \geq m \| x \|$, $\forall x \in X$. In acest caz, are loc inegalitatea $\| T^{-1} \| \leq m^{-1}$.
Un operator arbitrar (nu neapărat bijectiv) care satisface condiţia **(b)** se numeşte **mărginit inferior**. Este evident că un operator mărginit inferior este injectiv.
Demonstraţie (a)$\Rightarrow$ (b) Dacă T^{-1} este operator continuu, atunci, conform propoziţiei 3, există $M > 0$ astfel încât $\| T^{-1}y \| \leq M \| y \|$,
$\forall y \in Y$. Notând $x = T^{-1}y \in X$, rezultă $\| x \| \leq M \| Tx \|$, $\forall x \in X$ şi deci luând $m = M^{-1}$ relaţia **(b)** este verificată.
(b)$\Rightarrow$ (a) Fie $x \in X$ şi fie $y = Tx$; din ipoteză avem:

$$\| T^{-1}y \| = \| x \| \leq \frac{1}{m} \| Tx \| = \frac{1}{m} \| y \|,$$

deci conform propoziţiei 3 operatorul T^{-1} este continuu şi $\| T^{-1} \| \leq m^{-1}$.

Un rezultat fundamental este: (pentru demonstraţie se poate consulta [O01]):

17. Teorema lui Banach
Fie $T \in \mathcal{L}(H)$ un operator bijectiv; atunci inversul T^{-1} este continuu, deci $T^{-1} \in \mathcal{L}(H)$.

2.3 MF.02.3. Teorema Hahn-Banach

18. Definiţie
Fie X un spaţiu vectorial real sau complex. O funcţională $p : X \to \mathbf{R}$ se numeşte **subliniară** dacă $p(x + y) \leq p(x) + p(y)$ şi $p(\alpha x) = \alpha p(x)$, pentru orice $x, y \in X$ şi $\alpha \geq 0$. Din definiţie rezultă imediat proprietăţile $p(0) = 0$ şi $-p(-x) \leq p(x)$, $\forall x \in X$.
În demonstraţia teoremei principale din acest paragraf (teorema Hahn-Banach) vom folosi lema lui Zorn, pe care o reamintim în continuare.

19. Lema lui Zorn
Fie $(\mathcal{A}, \leq)$ o mulţime (parţial) ordonată. O submulţime $\mathcal{B} \subseteq \mathcal{A}$ se numeşte **total ordonată** dacă $\forall a, b \in \mathcal{B}$, atunci $a \leq b$ sau $b \leq a$. Se numeşte **majorant** al mulţimii $\mathcal{B}$ orice element $c \in \mathcal{A}$ astfel încât
$a \leq c$, $\forall a \in \mathcal{B}$. Spunem că $m \in \mathcal{A}$ este un **element maximal** al lui $\mathcal{A}$ dacă pentru orice $x \in \mathcal{A}$ cu proprietatea $m \leq x$, rezultă $x = m$. Mulţimea $\mathcal{A}$ se numeşte **inductiv ordonată** dacă orice submulţime total ordonată a lui $\mathcal{A}$ admite majoranţi.

Lema lui Zorn afirmă că orice mulţime nevidă inductiv ordonată admite un element maximal.

20. Teorema Hahn-Banach
Fie X un spaţiu vectorial real şi fie p o funcţională subliniară pe X. Fie Y un subspaţiu în X şi fie $g : Y \to R$, o funcţională liniară cu proprietatea $g(x) \leq p(x)$, $\forall x \in Y$. Atunci există $f : X \to \mathbf{R}$, o funcţională liniară cu proprietăţile: $f(x) = g(x)$, $\forall x \in Y$ şi $f(x) \leq p(x)$, $\forall x \in X$.
Se spune că f prelungeşte pe g la întreg spaţiul cu păstrarea inegalităţii $f(x) \leq p(x)$, $\forall x \in X$.
Demonstraţie În cele ce urmează, dacă h este o funcţională liniară, vom nota cu $D(h)$ subspatiul din X pe care este ea definită. Notaţia $g \preceq h$ va însemna că h este o funcţională liniară cu proprietăţile $D(h) \supseteq Y$, $h(y) = g(y)$, $\forall y \in Y$ şi $h(x) \leq p(x)$, $\forall x \in D(h)$. Fie:

$$\mathcal{A} = \{h : D(h) \to \mathbf{R} \, ; \, g \preceq h\}.$$

Mulţimea $\mathcal{A}$ este nevidă deoarece îl conţine pe g. Se demonstrează fără dificultate că $\mathcal{A}$ este inductiv ordonată; fie f elementul maximal dat de Lema lui Zorn. Demonstraţia se încheie dacă $D(f) = X$. Presupunem prin absurd că există $a \in X - D(f)$. Construim funcţionala
$h : D(h) = D(f) + a\mathbf{R} \to \mathbf{R}$, $h(x + at) = f(x) + \alpha t$, unde, α este o constantă reală neprecizată încă. Vom demonstra că putem alege α astfel încât $h \in \mathcal{A}$, ceea ce ar constitui o contradicţie cu maximalitatea lui f (incluziunea $D(h) \supset D(f)$ este, în mod evident, strictă). Relaţia pe care trebuie să

o satisfacă α pentru ca $h \in \mathcal{A}$ este:

$$f(x) + \alpha t \leq p(x + \alpha t)\,,\ \forall x \in D(f),\ \forall t \in \mathbf{R},$$

sau, echivalent

$$f(x) - p(x - a) \leq p(x + a) - f(x), \tag{2.9}$$

pentru orice $x \in D(f)$. Din ipoteză, pentru orice $x, y \in D(f)$ avem:

$$f(x) + f(y) \leq p(x + y) \leq p(x + a) + p(y - a),$$

deci pentru orice $x, y \in D(f)$,avem:

$$f(y) - p(y - a) \leq p(x + a) - f(x),$$

ceea ce încheie demonstraţia.

21. Corolar
Fie X un spaţiu vectorial real şi p o funcţională subliniară pe X. Atunci, pentru orice $x_o \in X$ există o funcţională liniară f pe X astfel încât $f(x_o) = p(x_o)$ şi $f(x) \leq p(x)\,, \forall x \in X$.
Pentru demonstraţie, se aplică teorema Hahn-Banach pentru
$Y = \{\alpha x_o\,;\, \alpha \in \mathbf{R}\}$ şi $g(\alpha x_o) = \alpha p(x_o)$.

O problemă importantă de geometrie în a cărei rezolvare teorema Hahn-Banach este un instrument esenţial este separarea submulţimilor nevide, convexe şi disjuncte dintr-un spaţiu normat.

22. Definiţie
Fie $(X, \|\ \|)$ un spaţiu normat (real) şi fie $f : X \to R$ o funcţională liniară neidentic nulă. Pentru orice $\alpha \in \mathbf{R}$, mulţimea:

$$Y(f, \alpha) = \{x \in X\,;\, f(x) = \alpha\}$$

se numeşte **hiperplanul** de ecuaţie $f = \alpha$.

Propunem ca exerciţiu afirmaţia: hiperplanul $Y(f, \alpha)$ este submulţime închisă în X dacă şi numai dacă funcţionala f este continuă.

Fie A şi B două submulţimi nevide în X. Spunem că hiperplanul $Y(f, \alpha)$ **separă nestrict** A de B dacă:

$$f(x) \leq \alpha,\ \forall x \in A \text{ şi } f(x) \geq \alpha,\ \forall x \in B.$$

Spunem că hiperplanul $Y(f, \alpha)$ **separă strict** A de B dacă există $\epsilon > 0$ astfel încât:

$$f(x) \leq \alpha - \epsilon,\ \forall x \in A \text{ şi } f(x) \geq \alpha + \epsilon,\ \forall x \in B.$$

Din punct de vedere geometric, separarea înseamnă că A şi B se găsesc "de o parte şi de alta a lui $Y(f, \alpha)$".

23. Teoremă
Fie X un spaţiu normat (real) şi fie $A \subset X$ şi $B \subset X$ două submulţimi nevide, convexe şi disjuncte.
(a) Dacă A este submulţime deschisă, atunci există o funcţională liniară şi continuă f pe X şi $\alpha \in R$ astfel încât hiperplanul (închis) $Y(f, \alpha)$ separă nestrict A de B.
(b) Dacă A este submulţime închisă şi B este compactă, atunci există o funcţională liniară şi continuă f pe X şi $\alpha \in R$ astfel încât hiperplanul (închis) $Y(f, \alpha)$ separă strict A de B.
Demonstraţie (a) Vom face demonstraţia în trei etape.

Etapa I. Fie $K \subset X$ o submulţime convexă astfel încât $0 \in K$. Atunci aplicaţia:

$$p : X \to R,\ p(x) = \inf\{a > 0\,;\ a^{-1}x \in K\},$$

este o funcţională subliniară pe X astfel încât există $M > 0$ cu proprietatea $0 \le p(x) \le M \parallel x \parallel$, $\forall x \in X$. În plus, are loc egalitatea:

$$K = \{x \in X\,;\ p(x) < 1\}.$$

Pentru orice $\rho > 0$, vom nota cu $B(0, \rho)$ bila deschisă de centru 0 şi rază ρ din X. Fie $r > 0$ astfel încât $B(0, r) \subset K$; este evident că $p(x) \le r^{-1} \parallel x \parallel$, $\forall x \in X$, deci putem lua $M = r^{-1}$. Faptul că $p(tx) = tp(x)$, $\forall x \in X$, $\forall t > 0$, este evident. Demonstrăm acum prin dublă incluziune egalitatea $K = \{x \in X\,;\ p(x) < 1\}$. Dacă $x \in K$, atunci, (deoarece K este mulţime deschisă), există $\epsilon > 0$ (suficient de mic) astfel încât $(1 + \epsilon)x \in K$, deci $p(x) \le (1 + \epsilon)^{-1} < 1$. Invers, dacă $p(x) < 1$, atunci există $t \in (0, 1)$ astfel încât $t^{-1}x \in K$, deci $x = t(t^{-1}x) + (1 - t)0 \in K$.
Demonstrăm acum proprietatea $p(x + y) \le p(x) + p(y)$, $\forall x, y \in X$.
Fie $x, y \in X$ şi $\epsilon > 0$. Din egalitatea $K = \{x \in X\,;\ p(x) < 1\}$, rezultă:

$$\frac{x}{p(x) + \epsilon} \in K \text{ şi } \frac{y}{p(y) + \epsilon} \in K,$$

şi deci, deoarece K este convexă, rezultă:

$$\frac{tx}{p(x) + \epsilon} + \frac{(1 - t)y}{p(y) + \epsilon} \in K,\ \forall t \in [0, 1].$$

În particular, pentru $t = (p(x) + p(y) + 2\epsilon)^{-1}\,(p(x) + \epsilon)$, obţinem:

$$\frac{x + y}{p(x) + p(y) + 2\epsilon} \in K,$$

şi deci $p(x+y) < p(x)+p(y)+2\epsilon$, ceea ce încheie demonstraţia primei etape.

Etapa a II-a Fie $K \subset X$ o submulţime nevidă, convexă şi deschisă şi fie $x_o \in X$ astfel încât $x_o \notin K$. Atunci există o funcţională liniară şi continuă $f : X \to R$, astfel încât $f(x) < f(x_o)$, $\forall x \in K$. În particular, hiperplanul (închis) $Y\,(f, f(x_o))$ separă nestrict $\{x_o\}$ de K.
Putem presupune, fără a restrânge generalitatea că $0 \in K$ (făcând eventual o translaţie). Fie p funcţionala subliniară introdusă în etapa I, adică $p(x) = \inf\{a > 0\,;\, a^{-1}x \in K\}$. Considerăm subspaţiul vectorial generat de x_o: $\mathcal{V} = \{tx_o\,;\, t \in R\}$. Fie $g : \mathcal{V} \to R$, $g(tx_o) = t$. Este evident că g este o funcţională liniară şi că ea verifică inegalitatea
$g(x) \leq p(x)$, $\forall x \in \mathcal{V}$. Conform teoremei Hahn-Banach, există o funcţională liniară $f : X \to R$ astfel încât

$$f(x) = g(x),\ \forall x \in \mathcal{V} \text{ şi } f(x) \leq p(x),\ \forall x \in X.$$

Din etapa I, avem că $f(x) \leq p(x) \leq M \parallel x \parallel$, $\forall x \in X$, deci f este continuă. De asemenea, $f(x_o) = 1$ şi deci, folosind egalitatea (demonstrată în etapa I) $K = \{x \in X\,;\, p(x) < 1\}$, rezultă $f(x) < 1$, $\forall x \in K$.

Etapa a III-a Demonstrăm acum enunţul teoremei. Fie A şi B ca în enunţ şi fie $K = \{x - y\,;\, x \in A,\, y \in B\}$. Este simplu de arătat că mulţimea K este convexă şi deschisă şi $0 \notin K$. Conform celor demonstrate în etapa a II-a, (pentru $x_o = 0$), există o funcţională liniară şi continuă $f : X \to R$, astfel încât $f(u) < 0$, $\forall u \in K$, adică:

$$f(x) < f(y),\ \forall x \in A,\ \forall y \in B.$$

De aici rezultă că există $\alpha \in R$ astfel încât

$$\sup_{x \in A} f(x) \leq \alpha \leq \inf_{y \in B} f(y),$$

ceea ce arată că hiperplanul (închis) $Y(f, \alpha)$ separă nestrict A de B.
(b) Fie A şi B ca în enunţ şi fie $\epsilon > 0$, suficient de mic, astfel încât mulţimile:

$$A_\epsilon = A + B(0, \epsilon) \text{ şi } B_\epsilon = B + B(0, \epsilon)$$

să fie disjuncte şi deschise. Cum A_ϵ şi B_ϵ sunt şi nevide şi convexe, din **(a)** rezultă că există un hiperplan $Y(f, \alpha)$ care separă nestrict A_ϵ de B_ϵ, adică:

$$f(x + \epsilon u) \leq \alpha \leq f(y + \epsilon u),\ \forall x \in A,\ \forall y \in B,\ \forall u \in B(0, 1).$$

Deoarece $f(z) \leq \parallel f \parallel$, $\forall z \in B(0, 1)$, rezultă:

$$f(x) + \epsilon \parallel f \parallel \leq \alpha \leq f(y) - \epsilon \parallel f \parallel,\ \forall x \in A,\ \forall y \in B,$$

ceea ce arată că hiperplanul $Y(f, \alpha)$ separă strict A de B întrucât f nu este identic nulă.

Revenim acum la problema prelungirii funcţionalelor liniare şi studiem cazul spaţiilor vectoriale complexe.

24. Observaţie
Dacă X este un spaţiu vectorial complex şi p este o seminormă pe X, atunci p este şi funcţională subliniară; în plus, următoarele relaţii se verifică imediat:

$$p(-x) = p(x) \text{ şi } |p(x) - p(y)| \leq p(x - y).$$

Demonstrăm în continuare teorema Hahn-Banach pentru spaţii vectoriale complexe.

25. Teoremă
Fie X un spaţiu vectorial complex, p o seminormă pe X şi $Y \subseteq X$ un subspaţiu vectorial. Atunci, pentru orice funcţională liniară $g : Y \to C$ astfel încât $|g(x)| \leq p(x)\,, \forall x \in Y$, există o funcţională liniară
$f : X \to C$ cu proprietăţile:

$$f(x) = g(x)\,, \forall x \in Y \text{ şi } |f(x)| \leq p(x)\,, \forall x \in X.$$

Demonstraţie Fie $g : Y \to \mathbf{C}$ ca în enunţ şi fie $g_1\,,\, g_2 : Y \to R$, funcţionale liniare reale astfel încât $g(x) = g_1(x) + ig_2(x)$. Atunci, explicitând egalitatea $g(ix) = ig(x)$, obţinem $g_1(ix) = -g_2(x)$, deci $g(x) = g_1(x) - ig_1(x)$. În plus, din ipoteză rezultă $|g_1(x)| \leq p(x)\,,$
$\forall x \in Y$. Aplicând teorema Hahn-Banach (cazul real) funcţionalei reale g_1 rezultă că există $f_1 : X \to R$ funcţională liniară reală astfel încât $f_1(x) = g_1(x)\,, \forall x \in Y$ şi $f_1(x) \leq p(x)\,, \forall x \in X$.
Definim $f : X \to \mathbf{C}\,,\, f(x) = f_1(x) - if_1(ix)$; se demonstrează prin calcul direct că f satisface concluzia teoremei.
Prezentăm în continuare câteva consecinţe (pe spaţii normate) ale teoremei Hahn-Banach.

26. Corolar
Fie $(X, \| \|)$ un spaţiu normat complex şi fie $Y \subseteq X$ un subspaţiu. Atunci, pentru orice funcţională liniară şi continuă $g : Y \to \mathbf{C}$ există $f \in X'$ astfel încât

$$f(x) = g(x)\,, \forall x \in Y \text{şi } \| f \| = \sup\{|g(x)|\,;\, x \in Y\,,\, \| x \| \leq 1\}.$$

Demonstraţie Fie $m = \sup\{|g(x)|\,;\, x \in Y\,,\, \| x \| \leq 1\}$ şi fie
$p : X \to \mathbf{R}\,,\, p(x) = m \| x \|$. Aplicând teorema Hahn-Banach funcţionalei g şi seminormei p, demonstraţia se încheie.

27. Corolar
Fie $(X, \| \|)$ un spaţiu normat. Atunci, pentru orice $x_o \in X$ există $f_o \in X'$

astfel încât:

$$\| f_o \| = \| x_o \| \text{ şi } f_o(x_o) = \| x_o \|^2 .$$

În particular, dacă $f(x_o) = 0$, $\forall f \in X'$, atunci $x_o = 0$.
Demonstraţie Se aplică corolarul precedent pentru
$Y = \{\alpha x_o\,;\, \alpha \in \mathbf{C}\}$ şi $g(\alpha x_o) = \alpha \| x_o \|^2$.

28. Corolar
Fie $(X, \| \, \|)$ un spaţiu normat. Atunci, pentru orice $x \in X$, are loc egalitatea:

$$\| x \| = \sup\{ |f(x)|\,;\, f \in X' \text{ şi } \| f \| = 1\}.$$

În plus, există $f_x \in X'$ astfel încât $\| x \| = |f_x(x)|$.
Demonstraţie Fie $x \in X$ şi fie $f \in X'$ astfel încât $\| f \| = 1$.
Din inegalitatea $|f(x)| \leq \| f \| \| x \| = \| x \|$, rezultă

$$\| x \| \geq \sup\{ |f(x)|\,;\, f \in X' \text{ şi } \| f \| = 1\}.$$

Din corolarul 26, rezultă existenţa unei funcţionale $h_x \in X'$ astfel încât $\| h_x \| = \| x \|$ şi $h_x(x) = \| x \|^2$; fie $f_x = \| x \|^{-1} h_x$. Atunci $\| f_x \| = 1$ şi $f_x(x) = \| x \|$, ceea ce încheie demonstraţia.

29. Observaţie
Din corolarul 27 rezultă că dacă X, Y sunt spaţii normate şi
$T \in \mathcal{L}(X, Y)$, atunci :

$$\| T \| = \sup\{ |g(Tx)|\,;\, x \in X\,,\, \| x \| \leq 1\,,\, g \in Y'\,,\, \| g \| \leq 1\}.$$

În particular, dacă $X = Y = (H, <\,,\,>)$ este un spaţiu Hilbert, atunci, din teorema lui Riesz, rezultă egalitatea:

$$\| T \| = \sup\{| < Tx, y > |\,;\, x, y \in H,\ \| x \| \leq 1,\ \| y \| \leq 1\}.$$

30. Corolar
Fie X un spaţiu normat şi fie $Y \subset X$ un subspatiu care nu este dens în X, adică $\overline{Y} \neq X$. Atunci există o funcţională neidentic nulă $f \in X'$ astfel încât $f(x) = 0\,,\, \forall x \in Y$.
Demonstraţie Vom face demonstraţia pentru cazul real. Considerăm ca funcţională subliniară distanţa la subspaţiul Y : $p(x) = \mathrm{dist}(x, Y) =$
$= \inf\{ \| x - y \|\,;\, y \in Y\}$. Evident, $p(x) = 0 \Leftrightarrow x \in \overline{Y}$. Fie $x_o \in X - \overline{Y}$; conform corolarului 19, există $f \in X'$ astfel încât $f(x_o) = p(x_o)$ şi $f(x) \leq p(x)\,,\, \forall x \in X$, ceea ce încheie demonstraţia.

Capitolul 3

MF.03. Operatori pe $\mathbf{C}^n$

Cuvinte cheie

operator liniar, spectru, subspaţiu invariant, valoare proprie, vector propriu, diagonalizare, operator adjunct, operator normal, calcul funţional.

În acest capitol vom studia operatorii liniari şi continui pe spaţiul $\mathbf{C^n}$. O parte din rezultatele acestui capitol vor fi generalizate la cazul operatorilor liniari şi continui definitţi pe spaţii Hilbert infinit dimensionale în capitolul 6. Sursele bibliografice recomandate: [B01], [F02], [O01], [S02].

3.1 MF.03.1. Noţiuni de algebră liniară

În acest paragraf vom reaminti unele rezultate de algebră liniară finit dimensională.

1. Definiţii

Fie $\mathbf{C^n} = \{\mathbf{x} = (\mathbf{x_1}, \mathbf{x_2}, .., \mathbf{x_n})\,;\, \mathbf{x_j} \in \mathbf{C}\}$ complex de dimensiune n, cu produsul scalar şi norma uzuale:

$$< x, y >= \sum_{j=1}^{n} x_j \overline{y_j}\ ,\ \| x \|= \sqrt{\sum_{j=1}^{n} |x_j|^2}.$$

Prin (**operator liniar**) este orice aplicaţie $T : C^n \to C^n$ cu proprietatea

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y),\ \forall \alpha, \beta \in C,\ \forall x, y \in C^n.$$

Vom nota cu $\mathcal{L}(C^n)$ mulţimea operatorilor liniari pe C^n.
Cu operaţiile uzuale de adunare şi înmulţire cu scalari:

$$(T + S)(x) = Tx + Sx,\ (\alpha T)x = \alpha Tx,\ \forall \alpha \in C, \forall x \in C^n,\ \forall T, S \in \mathcal{L}(C^n),$$

mulţimea $\mathcal{L}(C^n)$ este spaţiu vectorial; vom nota cu O operatorul nul şi cu I aplicaţia identică.
Produsul (compunerea) a doi operatori $T, S \in \mathcal{L}(C^n)$ este, prin definiţie

$$(TS)x = T(Sx),\ \forall x \in C^n.$$

Evident, operatorul TS este şi el liniar. Proprietăţile produsului sunt binecunoscute: asociativ, distributiv faţă de adunare şi admite ca element neutru operatorul identic I; el nu este comutativ. Dacă un operator $T \in \mathcal{L}(C^n)$ este injectiv şi surjectiv (deci bijectiv), atunci există şi este unic un operator, T^{-1}, de asemenea liniar, (numit inversul) lui T) astfel încât $TT^{-1} = T^{-1}T = I$. Operatorul T se numeşte în acest caz inversabil. Pentru orice operatori inversabili $T, S \in \mathcal{L}(C^n)$ şi $0 \neq \alpha \in \mathbf{C}$, se verifică afirmaţiile:
(a) $\left(T^{-1}\right)^{-1} = T$.
(b) $(\alpha T)^{-1} = \alpha^{-1}T^{-1}$.
(c) Produsul TS este inversabil şi $(TS)^{-1} = S^{-1}T^{-1}$.
Oricărui operator liniar T, i se pot asocia două subspaţii vectoriale remarcabile: nucleul, notat $\mathrm{Ker}(T)$, şi imaginea, notată $\mathrm{Im}(T)$, definite prin:

$$\mathrm{Ker}(T) = \{x \in \mathbf{C^n}\,;\, \mathbf{Tx} = \mathbf{0}\} \text{ şi } \mathrm{Im}(\mathbf{T}) = \{\mathbf{Tx}\,;\, \mathbf{x} \in \mathbf{C^n}\}.$$

Evident, operatorul T este injectiv dacă şi numai dacă $\mathrm{Ker}(T) = \{0\}$ şi este surjectiv dacă şi numai dacă $\mathrm{Im}(T) = C^n$.
O noţiune ce va fi frecvent utilizată în continuare este aceea de **subspaţiu invariant** pentru un operator. Un subspaţiu $\mathcal{X} \subseteq C^n$ se numeşte invariant pentru operatorul T dacă:

$$\forall x \in \mathcal{X} \Rightarrow Tx \in \mathcal{X}.$$

Nucleul şi imaginea unui operator sunt subspaţii invariante pentru acel operator.

2. Observaţie

Un rezultat important în legătură cu inversabilitatea operatorilor din $\mathcal{L}(\mathbf{C^n})$ (şi care nu este adevărat pentru aplicaţii liniare de la $\mathbf{C^n}$ în $\mathbf{C^m}$, $m \neq m$ şi nici pentru aplicaţii liniare pe un spaţiu vectorial infinit dimensional) este următorul:

Pentru orice operator $T \in \mathcal{L}(\mathbf{C^n})$, următoarele afirmaţii sunt echivalente:
(a) T este inversabil.
(b) T este injectiv.
(c) T este surjectiv.
Demonstraţia poate fi găsită în [B01].

3. Definiţie
Fie $\mathcal{B} = \{u_1, u_2, .., u_n\}$ o bază (nu neapărat ortonormală) în $\mathbf{C^n}$ şi fie $T \in \mathcal{L}(C^n)$ un operator fixat. Fie, pentru orice $\imath, \jmath \in \{1, 2, .., n\}$,

$$a_{\imath\jmath} = < Tu_\jmath, u_\imath > .$$

Matricea $M_T^{\mathcal{B}} = (a_{\imath\jmath})_{\imath,\jmath}$ se numeşte matricea operatorului T în baza $\mathcal{B}$.
Se verifică prin calcul direct egalitatea:

$$Tu_\imath = \sum_{\jmath=1}^{n} a_{\imath\jmath} u_\jmath, \ \forall \imath \in \{1, 2, .., n\}.$$

Mai general, dacă $x = (x_1, x_2, .., x_n) \in \mathbf{C^n}$ este un vector arbitrar, atunci

$$(Tx)_\imath = \sum_{\jmath=1}^{n} a_{\imath\jmath} x_\jmath, \ \forall \imath \in \{1, 2, .., n\}.$$

Matriceal, relaţia de mai sus se scrie $Tx = M_T^{\mathcal{B}} x$, vectorul x fiind aici vector coloană.
În cazul în care $\mathcal{B}$ este baza canonică, vom nota cu M_T matricea lui T în această bază.
Utilitatea asocierii $T \to M_T^{\mathcal{B}}$ este dată de următoarea teoremă ([B01]):

4. Teoremă
Fie $\mathcal{M}_n$ mulţimea matricelor pătratice de ordinul n cu elemente complexe şi fie $\mathcal{B}$ o bază fixată în $\mathbf{C^n}$.
(a) Aplicaţia $\mathcal{L}(\mathbf{C^n}) \ni T \to M_T^{\mathcal{B}} \in \mathcal{M}_n$ este un izomorfism de spaţii vectoriale şi:

$$M_{T+S}^{\mathcal{B}} = M_T^{\mathcal{B}} + M_S^{\mathcal{B}}, \ M_{\alpha T}^{\mathcal{B}} = \alpha M_T^{\mathcal{B}},$$

pentru orice $\alpha \in \mathbf{C}$ şi $T, S \in \mathcal{L}(\mathbf{C^n})$.
(b) Aplicaţia $\mathcal{L}(\mathbf{C^n}) \ni T \to M_T^{\mathcal{B}} \in \mathcal{M}_n$ este un izomorfism de inele şi

$$M_{TS}^{\mathcal{B}} = M_T^{\mathcal{B}} M_S^{\mathcal{B}}, \ \forall T, S \in \mathcal{L}(\mathbf{C^n}).$$

În particular, operatorul T este inversabil dacă şi numai dacă matricea sa (în orice bază, deoarece $\mathcal{B}$ a fost aleasă arbitrar) este nesingulară:

$$\det M_T^{\mathcal{B}} \neq 0.$$

5. Definiţie
Fie $\mathcal{B} = \{u_1, u_2, .., u_n\}$ şi $\mathcal{V} = \{v_1, v_2, .., v_n\}$ două baze fixate în $\mathbf{C^n}$. Pentru orice vector $v_\jmath \in \mathcal{V}$ există (şi sunt unici) $p_{\imath\jmath} \in \mathbf{C}$ astfel încât:

$$v_\imath = \sum_{\jmath=1}^{n} p_{\imath\jmath} u_\jmath.$$

Matricea $\mathcal{P} = (p_{\imath\jmath})_{\imath\jmath}$ se numeşte **matricea de trecere** de la baza $\mathcal{B}$ la baza $\mathcal{V}$. Dacă $S : C^n \to \mathbf{C^n}$ este aplicaţia liniară definită (pe bază) prin

$$Su_\jmath = v_\jmath, \ \forall \jmath \in \{1, 2, .., n\},$$

atunci $\mathcal{P}$ este matricea lui S în baza $\mathcal{B}$:

$$\mathcal{P} = M_S^{\mathcal{B}}.$$

Este uşor de demonstrat că orice matrice de trecere este nesingulară, şi, reciproc, orice matrice nesingulară este matricea de trecere între două baze (bine alese).

6. Observaţie
Fie acum un operator liniar $T \in \mathcal{L}(\mathbf{C^n})$; atunci legătura dintre matricele lui T în bazele $\mathcal{B}$ şi $\mathcal{V}$ şi matricea de trecere $\mathcal{P}$ între aceste două baze este ([B01]):

$$M_T^{\mathcal{V}} = \mathcal{P}^{-1} M_T^{\mathcal{B}} \mathcal{P}.$$

Din faptul că matricea unui operator depinde de alegerea bazei, decurge in mod natural problema găsirii unei baze în care matricea operatorului să aibă o formă cât mai simplă, şi anume formă diagonală. Această problemă (numită "**diagonalizarea** operatorilor liniari" pe $\mathbf{C^n}$) constituie subiectul central al acestui capitol. Menţionăm că analogul infinit dimensional al diagonalizării (deci pentru operatori definiţi pe spaţii Hilbert infinit dimensionale) este unul din scopurile principale ale capitolului 6.

7. Definiţie
Fie $T \in \mathcal{L}(C^n)$.
(a) T se numeşte operator diagonal dacă matricea sa în baza canonică (a lui C^n) este matrice diagonală.
(b) Spunem că T este diagonalizabil în sens algebric dacă există o bază a lui $\mathbf{C^n}$ în care matricea lui T să fie matrice diagonală.
(c) Spunem că T este diagonalizabil în sens geometric dacă există o bază ortonormală a lui $\mathbf{C^n}$ în care matricea lui T să fie matrice
diagonală.
Evident, avem implicaţiile:

$$\textbf{(a)} \Rightarrow \textbf{(c)} \Rightarrow \textbf{(b)}.$$

În acest paragraf vom reaminti (fără demonstraţii) principalele rezultate cu privire la operatorii diagonalizabili în sens algebric, iar în ultimul paragraf al acestui capitol vom studia (şi caracteriza) operatorii diagonalizabili în sens geometric.
Pentru diagonalizarea în sens algebric recomandăm [B01],p.75-90, unde sunt prezentate demonstraţiile complete ale rezultatelor ce urmează.

Instrumentele esenţiale pentru studiul diagonalizării sunt polinomul caracteristic, **vectorii proprii** şi **valorile proprii**.

8. Teorema Hamilton-Cayley
Fie $A \in \mathcal{M}_n$ şi fie I_n matricea unitate de ordinul n. Polinomul caracteristic al matricei A, este, prin definiţie,

$$P_A(z) = \det\left(zI_n - A\right).$$

Evident, P_A este un polinom de gradul n cu coeficienţi complecşi (şi de variabilă complexă z).
Dacă A şi B sunt două matrice pentru care există o matrice nesingulară $\mathcal{P}$ astfel încât $B = \mathcal{P}^{-1}A\mathcal{P}$, atunci se demonstrează că polinoamele lor caracteristice sunt egale: $P_A = P_B$. În particular, această proprietate se poate aplica în cazul în care matricele A şi B sunt matricele (în două baze diferite) ale aceluiaşi operator $T \in \mathcal{L}(\mathbf{C^n})$. Rezultă deci că putem defini polinomul caracteristic al operatorului $T \in \mathcal{L}(\mathbf{C^n})$ prin egalitatea:

$$P_T(z) = \det\left(zI_n - M_T^{\mathcal{B}}\right),$$

baza $\mathcal{B}$ fiind arbitrară.

Fie acum un polinom arbitrar, $f(z) = \sum\limits_{k=0}^{m} a_k z^k$; prin definiţie, polinomul de matricea A definit de f, este matricea:

$$f(A) = a_m A^m + a_{m-1}A^{m-1} + ... + a_1 A + a_o I_n.$$

Teorema Hamilton-Cayley ([B01],p.76) afirmă că

$$P_A(A) = O_n,$$

unde, O_n este matricea nulă de ordinul n.

9. Valori proprii şi vectori proprii
Fie $T \in \mathcal{L}(\mathbf{C^n})$. **Spectrul** operatorului T este, prin definiţie, mulţimea ([1],p.79):

$$\sigma(T) = \{\lambda \in \mathbf{C}\ ;\ \text{operatorul } \lambda I - T \text{ nu este inversabil}\}.$$

Mulţimea **valorilor proprii** ale operatorului T (sau **spectrul punctual**) este, prin definiţie:

$$\sigma_p(T) = \{\lambda \in \mathbf{C}\ ;\ \text{operatorul } T \text{ nu este injectiv}\}.$$

Incluziunea $\sigma_p(T) \subseteq \sigma(T)$ este evidentă. Dar, deoarece pe spaţii finit dimensionale un operator liniar este inversabil dacă şi numai dacă este injectiv,

rezultă că avem egalitatea $\sigma(T) = \sigma_p(T)$. În concluzie, un număr λ este în spectru dacă şi numai dacă λ este valoare proprie. Din această cauză, spectrul unui operator pe $\mathbf{C^n}$ se mai numeşte şi mulţimea valorilor proprii. Vom vedea că pe spaţii infinit dimensionale această proprietate nu mai este adevărată, spectrul punctual fiind, în general, o submulţime strictă a spectrului; există chiar exemple de operatori care nu au valori proprii, dar al căror spectru este nevid (a se vedea, de exemplu, operatorii de translaţie din cap.6).
Este acum evident că spectrul operatorului T este format din rădăcinile polinomului caracteristic asociat lui T:

$$\sigma(T) = \{\lambda \in \mathbf{C}\ ;\ P_T(\lambda) = 0\}.$$

În particular, rezultă că spectrul unui operator pe $\mathbf{C^n}$ este o mulţime nevidă şi finită.
De asemenea, $\lambda \in \sigma(T)$ dacă şi numai dacă există $x \in \mathbf{C^n}$, $x \neq 0$, astfel încât $Tx = \lambda x$. Un astfel de vector x se numeşte **vector propriu** asociat valorii proprii λ. Mulţimea vectorilor proprii asociaţi unei valori proprii fixate, λ, (la care adăugăm şi vectorul nul), este, în mod evident egală cu subspaţiul $\mathrm{Ker}(\lambda I - T)$.
Subspaţiile de vectori proprii au proprietăţile:
(a) Sunt subspaţii invariante pentru operatorul T, deci:

$$\forall x \in \mathrm{Ker}(\lambda I - T) \Rightarrow Tx \in \mathrm{Ker}(\lambda I - T).$$

(b) Dacă λ şi μ sunt două valori proprii distincte ale lui T, atunci:

$$\mathrm{Ker}(\lambda I - T) \cap \mathrm{Ker}(\mu I - T) = \{0\}.$$

Fie $\lambda \in \sigma(T)$. Multiplicitatea lui λ ca rădăcină a polinomului caracteristic P_T se numeşte dimensiunea (multiplicitatea) algebrică a lui λ şi o vom nota $n(\lambda)$. Evident, suma dimensiunilor algebrice ale tuturor valorilor proprii este egală cu n. Dimensiunea (multiplicitatea) geometrică a valorii proprii λ (notată $r(\lambda)$) este, prin definiţie, egală cu dimensiunea subspaţiului $\mathrm{Ker}(\lambda I - T)$. În general, are loc inegalitatea:

$$r(\lambda) \leq n(\lambda),\ \forall \lambda \in \sigma(T).$$

Încheiem acest paragraf recapitulativ cu rezultatul principal în legătură cu diagonalizarea în sens algebric a operatorilor liniari pe $\mathbf{C^n}$.

10. Teoremă (Criteriul de diagonalizare algebrică)
Fie $T \in \mathcal{L}(\mathbf{C^n})$; următoarele afirmaţii sunt echivalente:
(a) T este diagonalizabil în sens algebric.
(b) Există o bază $\mathcal{B} = \{u_1, u_2, .., u_n\}$ a lui $\mathbf{C^n}$ formată din vectori proprii ai operatorului T.

(c) $r(\lambda) = n(\lambda),\ \forall \lambda \in \sigma(T)$.
(d) $\bigoplus\limits_{\lambda \in \sigma(T)} \ker(\lambda I - T) = \mathbf{C^n}$.

În ipoteza că T este diagonalizabil în sens algebric, matricea sa în baza $\mathcal{B}$ are pe diagonală valorile proprii ale lui T, iar matricea de trecere $\mathcal{P}$ de la baza canonică la baza $\mathcal{B}$ are drept coloane vectorii proprii din baza $\mathcal{B}$. În concluzie, dacă $\sigma(T) = \{\lambda_1, \lambda_2, .., \lambda_n\}$, (fiecare valoare proprie fiind repetată de un număr egal cu dimensiunea sa algebrică), atunci:

$$M_T^{\mathcal{B}} = \mathrm{diag}(\lambda_1, \lambda_2, .., \lambda_n) = \mathcal{P}^{-1} M_T \mathcal{P}, [B01].$$

11. Observaţie

În capitolul 2, s-a definit norma unui operator liniar şi continuu între două spaţii normate şi s-a studiat spaţiul Banach $\mathcal{L}(\mathbf{C^n})$.

Un rezultat remarcabil, adevărat numai în cazul finit dimensional, este următorul.

12. Teoremă

Orice operator liniar $T \in \mathcal{L}(\mathbf{C^n})$ este continuu.

Demonstraţie Fie $T \in \mathcal{L}(\mathbf{C^n})$ şi fie $\mathcal{B} = \{e_1, e_2, .., e_n\}$ baza canonică a lui $\mathbf{C^n}$. Fie $K = \max\{\| Te_1 \|, \| Te_2 \|, .., \| Te_n \|\}$; fie $x = \sum\limits_{\jmath=1}^{n} x_\jmath e_\jmath$ un vector din $\mathbf{C^n}$. În mod evident avem:

$$|x_\jmath| \leq \| x \|, \forall 1 \leq \jmath \leq n.$$

Folosind inegalitatea triunghiului şi inegalităţile anterioare, rezultă:

$$\| Tx \| = \| T \left(\sum_{\jmath=1}^{n} x_\jmath e_\jmath \right) \| = \| \sum_{\jmath=1}^{n} x_\jmath T e_\jmath \| \leq$$

$$\leq \sum_{\jmath=1}^{n} |x_\jmath| \, \| Te_\jmath \| \leq nK \| x \| .$$

Din propoziţia precedentă, ((**b**)$\Leftrightarrow$ (**c**)), rezultă că T este aplicaţie continuă.

3.2 MF.03.2. Diagonalizarea operatorilor normali

În acest paragraf vom caracteriza operatorii care sunt diagonalizabili în sens geometric (cf. definiţiei 7).

13. Lemă (adjunctul unui operator)

Pentru orice $T \in \mathcal{L}(\mathbf{C^n})$, există un unic operator, $T^\star \in \mathcal{L}(\mathbf{C^n})$ astfel încât:

$$< Tx, y > = < x, T^\star y >, \forall x, y \in \mathbf{C^n}.$$

Operatorul $T^\star$ se numeşte **adjunctul** lui T. Demonstraţia se bazează în mod esenţial pe teorema lui Riesz de reprezentare a funcţionalelor liniare şi continue pe un spaţiu Hilbert (capitolul 6). În capitolul 6, paragraful 1 vom face demonstraţia pentru un spaţiu Hilbert arbitrar.
Alte proprietăţi remarcabile ale adjunctului sunt (a se vedea capitolul 6 pentru demonstraţii în cazul general al spaţiilor Hilbert):
(i) $(\alpha T + \beta S)^\star = \overline{\alpha} T^\star + \overline{\beta} S^\star$, $\forall \alpha, \beta \in C$, $T, S \in \mathcal{L}(\mathbf{C^n})$.
(ii) $(T^\star)^\star = T$, $\forall T \in \mathcal{L}(\mathbf{C^n})$.
(iii) $(TS)^\star = S^\star T^\star$, $\forall T, S \in \mathcal{L}(\mathbf{C^n})$.
(iv)$\| T^\star T \| = \| T \|^2$, $\forall T \in \mathcal{L}(\mathbf{C^n})$.
(v) $\| T^\star \| = \| T \|$, $\forall T \in \mathcal{L}(\mathbf{C^n})$.
(vi) Dacă $T \in \mathcal{L}(\mathbf{C^n})$ este inversabil, atunci $\left(T^{-1}\right)^\star = (T^\star)^{-1}$.

14. Propoziţie
Fie $T \in \mathcal{L}(\mathbf{C^n})$ şi fie $\mathcal{B} = \{u_1, u_2, .., u_n\}$ o bază ortonormală în $\mathbf{C^n}$. Dacă $M_T^{\mathcal{B}} = (a_{\imath\jmath})_{\imath\jmath}$ este matricea lui T în baza $\mathcal{B}$, atunci $M_{T^\star}^{\mathcal{B}} = (\overline{a_{\jmath,\imath}})_{\imath\jmath}$.
În particular, dacă elementele $a_{\imath\jmath}$ sunt reale, atunci matricea lui $T^\star$ este transpusa matricei lui T.
Demonstraţie Fie $M_{T^\star}^{\mathcal{B}} = (b_{\imath\jmath})_{\imath\jmath}$; conform definiţiei 3, avem:

$$b_{\imath\jmath} = < T^\star u_\jmath, u_\imath > = < u_\jmath, T u_\imath > =$$

$$= \overline{< T u_\imath, u_\jmath >} = \overline{a_{\jmath\imath}}, \ \forall \imath, \jmath \in \{1, 2, .., n\},$$

ceea ce încheie demonstraţia.
Cu ajutorul noţiunii de adjunct, putem defini câteva clase remarcabile de operatori:

15. Definiţie
Fie $T \in \mathcal{L}(\mathbf{C^n})$.
(a) T se numeşte autoadjunct dacă $T = T^\star$.
(b) T se numeşte unitar dacă $TT^\star = T^\star T = I$.
Este uşor de observat (deoarece $\mathbf{C^n}$ are dimensiune finită), că în definiţia dată mai sus este suficientă doar condiţia $T^\star T = I$ (de exemplu), cealaltă fiind o consecinţă. Pe spaţii Hilbert infinit dimensionale amândouă condiţiile sunt necesare.
(c) T se numeşte pozitiv (şi vom nota $T \geq O$) dacă

$$< Tx, x > \geq 0, \ \forall x \in C^n.$$

Aşa cum vom vedea, orice operator pozitiv este autoadjunct. În cazul unui spaţiului real $\mathbf{R^n}$ definiţia de mai sus nu mai implică $T = T^\star$; de aceea, în cazul $\mathbf{R^n}$ în definiţia operatorului pozitiv se cere şi condiţia de a fi autoadjunct.
(d) T se numeşte **normal** dacă $TT^\star = T^\star T$.

Există o analogie între operatori liniari şi numere complexe în care operatorii autoadjuncţi corespund numerelor reale, operatorii unitari corespund numerelor complexe de modul 1, iar operatorii pozitivi corespund numerelor pozitive. De exemplu, orice operator $T \in \mathcal{L}(\mathbf{C^n})$ se poate scrie în mod unic sub forma $T = A + iB$, cu A şi B operatori autoadjuncţi, această descompunere fiind analogul descompunerii (Carteziene) a unui număr complex $z = a + ib$, cu $a, b \in \mathbf{R}$; într-adevăr, dacă $A = \frac{1}{2}(T + T^\star)$ şi $B = \frac{1}{2i}(T - T^\star)$, atunci A şi B sunt autoadjuncţi şi $A + iB = T$.

$$T \in \mathcal{L}(C^n) \leftrightarrow z \in C.$$
$$A = A^\star \leftrightarrow z = \overline{z} \in R.$$
$$U^\star U = I \leftrightarrow z\overline{z} = |z|^2 = 1.$$

Menţionăm că, aşa cum vom vedea în continuare, există şi alte rezultate care întăresc această analogie, inclusiv în cazul infinit dimensional (a se vedea capitolul 6).

Revenim acum la problema diagonalizării în sens geometric, care constituie subiectul central al acestui paragraf; rezultatul fundamental (pe care îl vom demonstra în teorema 26) este:

T este diagonalizabil în sens geometric $\Leftrightarrow T$ este normal.

Primul rezultat se referă la operatorii unitari.

16. Teoremă
Fie $U \in \mathcal{L}(\mathbf{C^n})$; următoarele afirmaţii sunt echivalente:
(a) U este operator unitar.
(b) U este inversabil şi U^{-1} este unitar.
(c) $< Ux, Uy > = < x, y >$, $\forall x, y \in \mathbf{C^n}$.
(d) U transformă orice bază ortonormală în bază ortonormală.
Demonstraţie (a)⇒ (b) Dacă U este unitar, atunci, din definiţie, U este inversabil şi $U^{-1} = U^\star$; operatorul U^{-1} este unitar deoarece:

$$\left(U^{-1}\right)^\star U^{-1} = (U^\star)^{-1} U^{-1} = (UU^\star)^{-1} = I.$$

(b)⇒ (c) Pentru orice $x, y \in \mathbf{C^n}$, avem:

$$< x, y > = < U^{-1}Ux, y > = < Ux, \left(U^{-1}\right)^\star y > = < Ux, Uy > .$$

(c)⇒ (d) Fie $\mathcal{B} = \{u_1, u_2, .., u_n\}$ o bază ortonomală, deci

$$< u_\imath, u_\jmath > = \delta_i^\jmath, \text{ (simbolul lui Kronecker)}.$$

Mulţimea $\{Uu_1, Uu_2, .., Uu_n\}$ este bază ortonormală deoarece:

$$< Uu_\imath, Uu_\jmath > = < u_\imath, u_\jmath > = \delta_i^\jmath, \ \forall \imath, \jmath \in \{1, 2, .., n\}.$$

(d)$\Rightarrow$ **(c)** Fie $\{e_1, e_2, .., e_n\}$ baza canonică şi fie

$$x = \sum_{i=1}^{n} x_i e_i \text{ şi } y = \sum_{i=1}^{n} y_i e_i.$$

Deoarece, conform ipotezei, $\{Ue_1, Ue_2, .., Ue_n\}$ este tot bază ortonormală, avem:

$$< Ux, Uy >=< \sum_{i=1}^{n} x_i U e_i, \sum_{j=1}^{n} y_j U e_j >=$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j < Ue_i, Ue_j >= \sum_{i=1}^{n} x_i y_i =< x, y > .$$

(c)$\Rightarrow$ **(a)** Propunem mai întâi ca exerciţiu următoarea afirmaţie: doi operatori $T, S \in \mathcal{L}(\mathbf{C^n})$ sunt egali dacă şi numai dacă

$$< Tx, y >=< Sx, y >, \ \forall x, y \in \mathbf{C^n}.$$

Pentru orice $x, y \in \mathbf{C^n}$, avem:

$$< U^\star Ux, y >=< Ux, Uy >=< x, y >,$$

şi deci $U^\star U = I$.

17. Observaţie

Punctul **(d)** din teorema de mai sus arată că matricele operatorilor unitari sunt exact matricele de trecere între două baze ortonormale. Mai mult, matricea (într-o bază ortonormală) a unui operator unitar are drept coloane vectori ortonormali; o astfel de matrice se numeşte matrice ortogonală.
Din egalitatea (c) din teorema 16 rezultă $\| Ux \|=\| x \|, \forall x \in \mathbf{C^n}$, deci operatorii unitari sunt izometrii liniare; pe spaţiul $\mathbf{C^n}$ se poate demonstra şi reciproca: orice izometrie liniară este operator unitar. Într-adevăr, dacă

$$\| Ux \|=\| x \|, \forall x \in \mathbf{C^n},$$

atunci $< Ux, Ux >=< x, x >, \forall x \in \mathbf{C^n}$ şi deci

$$< U^\star Ux, x >=< x, x >, \forall x \in \mathbf{C^n}.$$

Demonstraţia se încheie dacă folosim următorul rezultat adevărat numai pe spaţii Hilbert complexe (demonstraţia, care este elementară, se găseşte în capitolul 6, propoziţia 9):
Fie $T \in \mathcal{L}(C^n)$ un operator arbitrar; dacă $< Tx, x >= 0, \forall x \in C^n$, atunci $T = O$.
Vom vedea (în capitolul 6) că pe spaţii Hilbert infinit dimensionale
există izometrii liniare neinversabile.

Utilitatea operatorilor unitari pentru problema diagonalizării în sens geometric este conţinută în următoarea teoremă.

18. Teoremă
(a) Un operator $D \in \mathcal{L}(\mathbf{C^n})$ este operator diagonal dacă şi numai dacă vectorii bazei canonice sunt vectori proprii pentru D.
(b) Un operator $T \in \mathcal{L}(\mathbf{C^n})$ este operator diagonalizabil în sens geometric dacă şi numai dacă există o bază ortonormală a lui $\mathbf{C^n}$ formată din vectori proprii ai lui T, sau, echivalent, există un operator unitar U astfel încât operatorul $D = U^{-1}TU$ să fie operator diagonal.
Demonstraţie (a) Dacă $D \in \mathcal{L}(\mathbf{C^n})$ este un operator diagonal, atunci, prin definiţie, matricea sa în baza canonică este:

$$M_D = \begin{pmatrix} \lambda_1 & 0 & . & . & . & 0 \\ 0 & \lambda_2 & . & . & . & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & . & . & . & \lambda_n \end{pmatrix},$$

unde, $\lambda_1, \lambda_2, .., \lambda_n$ sunt valorile proprii ale lui D. Dacă $\{e_1, e_2, .., e_n\}$ este baza canonică, atunci evident $De_\imath = \lambda_\imath e_\imath$ şi deci $e_\imath$ este vector propriu asociat valorii proprii $\lambda_\imath$.
Reciproc, dacă $De_\imath = \lambda_\imath e_\imath$, atunci elementele matricei lui D (în baza canonică), sunt:

$$a_{\imath\jmath} =< De_\jmath, e_\imath >= \begin{cases} \lambda_\imath & \text{dacă} \quad \imath = \jmath \\ 0 & \text{dacă} \quad \imath \neq \jmath. \end{cases}$$

(b) Dacă T este un operator diagonalizabil în sens geometric, atunci, din definiţie, există o bază ortonormală $\mathcal{B} = \{u_1, u_2, .., u_n\}$ astfel încât

$$M_T^{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & . & . & . & 0 \\ 0 & \lambda_2 & . & . & . & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & . & . & . & \lambda_n \end{pmatrix}.$$

Fie D operatorul (diagonal) a cărui matrice în baza canonică este $M_T^{\mathcal{B}}$. Dacă U este operatorul definit de relaţiile

$$Ue_\imath = u_\imath, \forall \imath \in \{1, 2, .., n\},$$

atunci U este operator unitar (conform teoremei 16(d)) şi $D = U^{-1}TU$. Conform celor demonstrate la punctul **(a)**, rezultă că

$$De_\imath = \lambda_\imath e_\imath, \forall \imath \in \{1, 2, .., n\}.$$

Rezultă deci $TUe_\imath = \lambda_\imath Ue_\imath$, adică

$$Tu_\imath = \lambda_\imath u_\imath, \forall \imath \in \{1, 2, .., n\},$$

deci $u_1, u_2, .., u_n$ sunt vectori proprii ai lui T.
Reciproc, fie $\mathcal{B} = \{u_1, u_2, .., u_n\}$ o bază ortonormală formată din vectori proprii ai operatorului $T \in \mathcal{L}(\mathbf{C^n})$, deci

$$Tu_\imath = \lambda_\imath u_\imath,\ \forall \imath \in \{1, 2, .., n\}.$$

Fie U operatorul definit prin $Ue_\imath = u_\imath$, $\forall \imath$. Atunci U este operator unitar (cf. teoremei 16(d)); rezultă că operatorul $U^{-1}TU$ este operator diagonal deoarece vectorii bazei canonice sunt vectori proprii:

$$U^{-1}TUe_\imath = U^{-1}Tu_\imath = U^{-1}(\lambda_\imath u_\imath) = \lambda_\imath U^{-1}u_\imath = \lambda_\imath e_\imath,\ \forall \imath.$$

Rezultă deci că matricea lui T în baza $\mathcal{B}$ este diagonală, deci T este operator diagonalizabil în sens geometric.

Trecem acum la studiul operatorilor normali; înainte de a demonstra teorema de diagonalizare (în sens geometric) pentru această clasă de operatori, vom prezenta mai întâi câteva proprietăţi uzuale ale acestora.
Reamintim că un operator $T \in \mathcal{L}(\mathbf{C^n})$ se numeşte normal dacă el comută cu adjunctul său: $TT^\star = T^\star T$.

19. Propoziţie
Dacă $T \in \mathcal{L}(\mathbf{C^n})$ este operator normal, atunci

$$\| Tx \| = \| T^\star x \|,\ \forall x \in \mathbf{C^n}.$$

Menţionăm că reciproca acestei afirmaţii este şi ea adevărată.
Demonstraţie Pentru orice $x \in \mathbf{C^n}$, avem:

$$\| Tx \|^2 = < Tx, Tx > = < x, T^\star Tx > =$$

$$= < x, TT^\star x > = < T^\star x, T^\star x > = \| T^\star x \|^2 .$$

20. Consecinţă
Dacă $T \in \mathcal{L}(\mathbf{C^n})$ este operator normal, atunci Ker(T) =Ker$(T^\star)$.
Folosind propoziţia precedentă, demonstraţia este evidentă:

$$\| Tx \| = 0 \Leftrightarrow \| T^\star x \| = 0.$$

21. Propoziţie
Fie $T \in \mathcal{L}(\mathbf{C^n})$ un operator normal. Atunci, pentru orice $\lambda \in \mathbf{C}$ şi $x \in \mathbf{C^n}$, avem:

$$Tx = \lambda x \Leftrightarrow T^\star x = \overline{\lambda} x.$$

Deci, dacă λ este valoare proprie pentru T, iar x este un vector propriu (al lui T) corespunzător valorii proprii λ, atunci, $\overline{\lambda}$ este valoare proprie pentru

$T^\star$, iar x este vector propriu (al lui $T^\star$) corespunzător valorii proprii $\overline{\lambda}$.
Demonstraţie Fie $\lambda \in \sigma(T)$; atunci, pentru orice $x \in \mathbf{C^n}$, avem:

$$Tx = \lambda x \Leftrightarrow x \in \mathrm{Ker}(\lambda I - T),$$

şi deci este suficient să demonstrăm egalitatea:

$$\mathrm{Ker}(\lambda I - T) = \mathrm{Ker}(\overline{\lambda} I - T^\star).$$

Se verifică direct că operatorul $\lambda I - T$ este normal, iar adjunctul său este $\overline{\lambda} I - T^\star$; aplicând acum consecinţa 20 operatorului (normal) $\lambda I - T$, demonstraţia se încheie.

22. Propoziţie
Fie $T \in \mathcal{L}(\mathbf{C^n})$ un operator normal şi fie $\lambda \neq \mu$ două valori proprii distincte ale sale. Dacă $Tx = \lambda x$ şi $Ty = \mu y$, atunci $x \perp y$.
Demonstraţie Fie T, λ, μ, x, y ca în enunţ; atunci , conform propoziţiei precedente $T^\star y = \overline{\mu} y$ şi deci:

$$\lambda < x, y > = < \lambda x, y > = < Tx, y > = < x, T^\star y > = < x, \overline{\mu} y > = \mu < x, y > .$$

Deoarece $\lambda - \mu \neq 0$, rezultă $< x, y > = 0$, adică $x \perp y$.

23. Observaţie
Din algebra liniară se ştie că pentru un operator liniar arbitrar, vectorii proprii corespunzători unor valori proprii distincte sunt liniari independenţi; propoziţia anterioară afirmă că pentru operatorii normali, aceştia sunt perpendiculari. O formulare echivalentă este: pentru orice $\lambda, \mu \in \sigma(T), \lambda \neq \mu$, avem $\mathrm{Ker}(\lambda I - T) \perp \mathrm{Ker}(\mu I - T)$.

Pentru a putea demonstra teorema de diagonalizare pentru operatorii normali, mai sunt necesare două rezultate cu caracter general.

24. Lemă
Fie $A, B \in \mathcal{L}(\mathbf{C^n})$. Dacă $AB = BA$, atunci A şi B au (cel puţin) un vector propriu comun.
Demonstraţie Fie $\lambda \in \sigma(A)$ şi fie $x \neq 0$ un vector propriu corespunzător: $Ax = \lambda x$.
Din egalitatea $AB = BA$ rezultă prin inducţie $AB^k = B^k A, \forall k \in N$. Aplicând egalităţii $Ax = \lambda x$ operatorul B, obţinem: $BAx = \lambda Bx$, adică $ABx = \lambda Bx$; în concluzie, vectorul Bx este şi el vector propriu pentru operatorul A (corespunzător tot valorii proprii λ). Analog, aplicând în continuare $B^2, B^3, \ldots$, rezultă

$$AB^k x = \lambda B^k x, \ \forall k \in N,$$

şi deci toţi vectorii $B^k x$, $k \in N$, sunt vectori proprii ai operatorului A, corespunzători valorii proprii λ. Deoarece dimensiunea lui $\mathbf{C^n}$ este finită, rezultă că numai un număr finit dintre aceştia sunt liniari independenţi; fie

$$\{x, Bx, B^2x, .., B^{p-1}x\}$$

primii p vectori liniari independenţi şi fie $\mathcal{X}$ subspaţiul liniar generat de ei. Proprietăţile subspaţiului $\mathcal{X}$ sunt:
(i) $\dim(\mathcal{X}) = p$.
(ii) $\forall y \in \mathcal{X}$ este vector propriu pentru operatorul A.
(iii) $\mathcal{X}$ este subspaţiu invariant pentru operatorul B, adică $B(\mathcal{X}) \subseteq \mathcal{X}$.
Proprietatea **(i)** este evidentă; pentru a demonstra **(ii)** este suficient să observăm că, în general, orice combinaţie liniară de vectori proprii (corespunzători toţi aceleeaşi valori proprii) este în continuare vector propriu.
Demonstrăm acum **(iii)**; pentru aceasta, este suficient să demonstrăm că pentru orice vector (din bază) $B^q x \in \mathcal{X}$ rezultă $B(B^q x) \in \mathcal{X}$. Dar $B^{q+1}x \in \{x, Bx, B^2x, ...\}$ şi deci conform alegerii lui p rezultă că $B^{q+1}x$ este o combinaţie liniară a vectorilor $\{x, Bx, .., B^{p-1}x\}$, adică $B^{q+1}x \in \mathcal{X}$.
Fie $B|_{\mathcal{X}} : \mathcal{X} \to \mathcal{X}$ restricţia operatorului B la subspaţiul $\mathcal{X}$; operatorul $B|_{\mathcal{X}}$ are cel puţin o valoare proprie (deoarece $\dim(\mathcal{X}) \geq 1$) şi deci există cel puţin un vector propriu $y \in \mathcal{X}$ al operatorului B. Deoarece toţi vectorii din $\mathcal{X}$ sunt vectori proprii pentru A, rezultă că y este un vector propriu comun operatorilor A şi B.

25. Lemă
Fie $T \in \mathcal{L}(\mathbf{C^n})$ şi fie $\mathcal{X}$ un subspaţiu invariant pentru T. Atunci subspaţiul ortogonal, $\mathcal{X}^\perp$, este invariant pentru operatorul $T^\star$.
Demonstraţie Pentru orice $y \in \mathcal{X}^\perp$ şi $x \in \mathcal{X}$, deoarece $Tx \in \mathcal{X}$, avem:

$$< T^\star y, x >=< y, Tx >= 0.$$

Rezultă deci că $T^\star y \perp x$, $\forall x \in \mathcal{X}$, adică $T^\star y \in \mathcal{X}^\perp$.
Evident, are loc şi implicaţia reciprocă: dacă $\mathcal{X}^\perp$ este invariant la $T^\star$, atunci $\mathcal{X}$ este invariant la T.

Demonstrăm în continuare principalul rezultat al acestui paragraf.

26. Teorema de diagonalizare pentru operatori normali
Fie $T \in \mathcal{L}(\mathbf{C^n})$; următoarele afirmaţii sunt echivalente:
(i) T este operator normal.
(ii) T este operator diagonalizabil în sens geometric.
Demonstraţie Vom începe cu implicaţia mai uşoară: **(ii)**$\Rightarrow$**(i)**. Dacă T este operator diagonalizabil în sens geometric, atunci, conform teoremei 18 (b), există un operator unitar $U \in \mathcal{L}(\mathbf{C^n})$ astfel încât operatorul $D = U^\star T U$ să fie operator diagonal. Deoarece $DD^\star = D^\star D$ (egalitate evidentă), rezultă:

$$TT^\star = U^\star DU\,(U^\star DU)^\star = U^\star DUU^\star D^\star U = U^\star DD^\star U =$$

$$= U^\star D^\star DU = U^\star D^\star UU^\star DU = (U^\star DU)^\star U^\star DU = T^\star T,$$

şi deci T este operator normal.
Demonstrăm acum implicaţia **(i)**$\Rightarrow$**(ii)**. Fie $T \in \mathcal{L}(\mathbf{C^n})$ un operator normal, deci $TT^\star = T^\star T$. Pentru a demonstra că T este diagonalizabil în sens geometric este suficient, conform teoremei 18(b), să construim o bază ortonormală a lui $\mathbf{C^n}$ formată din vectori proprii ai operatorului T. Deoarece operatorii T şi $T^\star$ comută, din lema 24 rezultă că există $u_1 \in \mathbf{C^n}$ un vector propriu comun pentru T şi $T^\star$. Fie $\mathcal{X}_1$ subspaţiul liniar generat de u_1 şi fie $\mathcal{X}_1^\perp$ ortogonalul său. Proprietăţile subspaţiilor $\mathcal{X}_1$ şi $\mathcal{X}_1^\perp$ sunt:
(i) $\mathcal{X}_1 \bigoplus \mathcal{X}_1^\perp = \mathbf{C^n}$.
(ii) $\dim \mathcal{X}_1 = 1$ şi $\dim \mathcal{X}_1^\perp = n-1$.
(iii) $\mathcal{X}_1$ şi $\mathcal{X}_1^\perp$ sunt invariante la T şi $T^\star$.
Primele două proprietăţi sunt evidente. Subspaţiul $\mathcal{X}_1$ este invariant şi la T şi la $T^\star$ deoarece u_1 este vector propriu atât pentru T cât şi pentru $T^\star$. Conform lemei 25, rezultă că subspaţiul $\mathcal{X}_1^\perp$ este şi el invariant pentru operatorii T şi $T^\star$. Considerăm restricţiile operatorilor T şi $T^\star$ la subspaţiul $\mathcal{X}_1^\perp$:

$$T|_{\mathcal{X}_1^\perp} : \mathcal{X}_1^\perp \to \mathcal{X}_1^\perp,$$

$$T^\star|_{\mathcal{X}_1^\perp} : \mathcal{X}_1^\perp \to \mathcal{X}_1^\perp.$$

Aplicând acum lema 24 operatorilor $T|_{\mathcal{X}_1^\perp}$ şi $T^\star|_{\mathcal{X}_1^\perp}$, (care comută între ei), rezultă că există $u_2 \in \mathcal{X}_1^\perp$ care este vector propriu comun operatorilor T şi $T^\star$. Fie $\mathcal{X}_2$ subspaţiul liniar generat de vectorii u_1 şi u_2 şi fie $\mathcal{X}_2^\perp$ ortogonalul său. Deoarece vectorii u_1 şi u_2 sunt perpendiculari
(din construcţie), rezultă că dimensiunea spaţiului $\mathcal{X}_2$ este 2; cu un raţionament analog celui de mai sus, se demonstrează că subspaţiile $\mathcal{X}_2$ şi $\mathcal{X}_2^\perp$ sunt invariante la T şi $T^\star$. Repetând acum construcţia anterioară (considerăm restricţiile operatorilor T şi $T^\star$ la subspaţiile $\mathcal{X}_2$ şi $\mathcal{X}_2^\perp$, etc), obţinem o mulţime $\{u_1, u_2, .., u_n\}$ cu proprietăţile:
(i) $u_\imath \perp u_\jmath$, $\forall \imath \neq \jmath$.
(ii) $u_\imath$ este vector propriu pentru operatorii T şi $T^\star$, $\forall \imath \in \{1, 2, .., n\}$.
Considerând acum

$$v_\imath = \| u_\imath \|^{-1} u_\imath, \forall \imath \in \{1, 2, .., n\},$$

rezultă că mulţimea $\{v_1, v_2, .., v_n\}$ este o bază ortonormală a lui $\mathbf{C^n}$ formată din vectori proprii ai operatorului T (şi ai lui $T^\star$), ceea ce încheie demonstraţia.

Înainte de a enunţa o primă consecinţă importantă a teoremei de mai sus, introducem o nouă clasă de operatori liniari.

27. Definiţie
Fie $\mathcal{X}$ un subspaţiu în $\mathbf{C^n}$. Atunci, conform teoremei proiecţiei pe un

subspaţiu închis (a se vedea şi teorema proiecţiei, capitolul 4), orice $x \in \mathbf{C^n}$ admite o descompunere unică $x = y + z$ cu $y \in \mathcal{X}$ şi $z \in \mathcal{X}^{\perp}$. Considerăm operatorul (liniar)

$$P_{\mathcal{X}} : \mathbf{C^n} \to \mathbf{C^n},\ P_{\mathcal{X}}x = y.$$

Operatorul $P_{\mathcal{X}}$ se numeşte **proiecţia** pe subspaţiul $\mathcal{X}$. Este evident că $P_{\mathcal{X}}^2 = P_{\mathcal{X}}$; se demonstrează de asemenea fără dificultate că $P_{\mathcal{X}}$ este autoadjunct. Un studiu aprofundat al operatorilor de proiecţie (pe un spaţiu Hilbert arbitrar) va fi prezentat în capitolul 6, paragraful 2. În cele ce urmează vom folosi următoarele proprietăţi (demonstraţiile sunt imediate).
Dacă $\mathcal{X} \perp \mathcal{Y}$, atunci:
(i) $P_{\mathcal{X}}P_{\mathcal{Y}} = P_{\mathcal{Y}}P_{\mathcal{X}} = O$.
(ii) Operatorul sumă $P_{\mathcal{X}} + P_{\mathcal{Y}}$ este de asemenea proiecţie, subspaţiul de proiecţie corespunzător fiind suma (directă) a subspaţiilor $\mathcal{X}$ şi $\mathcal{Y}$.

3.3 MF.03.3. Calcul funcţional

28. Formula de descompunere spectrală pentru operatori normali
Fie $T \in \mathcal{L}(\mathbf{C^n})$ un operator normal şi fie $\lambda_1, \lambda_2, .., \lambda_m$ valorile sale proprii (distincte). Fie, pentru orice $\imath \in \{1, 2, .., m\}$, $P_\imath$ operatorul de proiecţie pe subspaţiul vectorilor proprii asociaţi valorii proprii $\lambda_\imath$. Atunci:
(i) $P_\imath P_\jmath = O$, $\forall \imath \neq \jmath$.
(ii) $P_1 + P_2 + .. + P_m = I$.
(iii) $T = \lambda_1 P_1 + \lambda_2 P_2 + .. + \lambda_m P_m$.
Formula **(iii)** se numeşte descompunerea spectrală a lui T; în plus, această descompunere este unică.
Demonstraţie Prima relaţie este adevărată deoarece, conform propoziţiei 22, pentru un operator normal vectorii proprii corespunzători unor valori proprii distincte sunt ortogonali. Din teorema 26, rezultă că există o bază a lui $\mathbf{C^n}$ formată din vectori proprii ai operatorului T şi deci suma (directă) a tuturor subspaţiilor de vectori proprii este $\mathbf{C^n}$; acest fapt justifică egalitatea **(ii)**. Pentru a demonstra formula de descompunere spectrală, fie, (ca în teorema 26(a)), $T = UDU^{-1}$, unde U este operator unitar, iar D este operator diagonal (matricea sa în baza canonică are pe diagonala principală valorile proprii ale lui T). Este evident că $D = \lambda_1 E_1 + \lambda_2 E_2 + .. + \lambda_m E_m$, unde, $E_\imath$ este proiecţia pe subspaţiul vectorilor proprii ai lui D asociaţi valorii proprii $\lambda_\imath$; reamintim că, în baza teoremei 18(a), vectorii din baza canonică sunt vectori proprii pentru D. Demonstraţia se încheie observând că

$$P_\imath = UE_\imath U^{-1}.$$

Lăsăm demonstraţia unicităţii ca exerciţiu.
Din demonstraţie rezultă de asemenea şi formula de descompunere spectrală

a adjunctului:

$$T^\star = \sum_{\imath=1}^{m} \overline{\lambda_\imath} P_\imath.$$

29. Observaţie
Deoarece operatorii autoadjuncţi şi operatorii unitari sunt în mod evident operatori normali, din teorema 26 rezultă că aceşti operatori sunt diagonalizabili în sens geometric; propunem cititorului să găsească o demonstraţie directă pentru teorema de diagonalizare a operatorilor autoadjuncţi.

În finalul acestui capitol vom da câteva aplicaţii remarcabile ale teoremelor 26 şi 28; pentru completări, recomandăm [9].

30. Observaţie
Se demonstrează fără dificultate următoarele implicaţii:
(i) Dacă T este operator autoadjunct, atunci valorile sale proprii sunt numere reale.
(ii) Dacă T este operator pozitiv, atunci valorile sale proprii sunt numere pozitive.
(iii) Dacă T este proiector, atunci $\sigma(T) \subseteq \{0, 1\}$.
(iv) Dacă T este operator unitar, atunci valorile sale proprii sunt numere complexe de modul 1.
Este remarcabil faptul că pentru operatorii normali sunt adevărate şi reciprocele acestor afirmaţii.

31. Teoremă
Fie $T \in \mathcal{L}(\mathbf{C^n})$ un operator normal; atunci:
(i) T este operator autoadjunct dacă şi numai dacă valorile sale proprii sunt numere reale.
(ii) T este operator pozitiv dacă şi numai dacă valorile sale proprii sunt numere pozitive.
(iii) T este proiector dacă şi numai dacă $\sigma(T) \subseteq \{0, 1\}$.
(iv) T este operator unitar dacă şi numai dacă valorile sale proprii sunt numere complexe de modul 1.
Menţionăm că există operatori (dar nu normali) care au toate valorile proprii reale, dar nu sunt autoadjuncţi, etc.
Demonstraţie Vom demonstra numai implicaţiile "$\Leftarrow$".
Fie, conform teoremei 28,

$$T = \sum_{\imath=1}^{m} \lambda_\imath P_\imath \text{ şi } T^\star = \sum_{\imath=1}^{m} \overline{\lambda_\imath} P_\imath,$$

descompunerile spectrale ale operatorilor (normali) T şi $T^\star$.
(i) Este clar că dacă $\lambda_\imath$ sunt numere reale, atunci T este autoadjunct.

(ii) Dacă $\lambda_\imath \geq 0$, atunci pentru orice $x \in \mathbf{C^n}$, avem:

$$< Tx, x >= \sum_{\imath=1}^{m} \lambda_\imath < P_\imath x, x >= \sum_{\imath=1}^{m} \lambda_\imath < P_\imath^2 x, x >=$$

$$= \sum_{\imath=1}^{m} \lambda_\imath < P_\imath x, P_\imath x >= \sum_{\imath=1}^{m} \lambda_\imath \parallel P_\imath x \parallel^2 \geq 0,$$

deci T este operator pozitiv.
(iii) Dacă $\sigma(T) \subseteq \{0, 1\}$, atunci operatorul T este suma unor proiecţii pe subspaţii ortogonale, deci este el însuşi o proiecţie.
(iv) Dacă $|\lambda_\imath| = 1$, $\forall \imath$, atunci:

$$TT^\star = \sum_{\imath=1}^{m} \sum_{\jmath=1}^{m} \lambda_\imath \lambda_\jmath P_\imath P_\jmath = \sum_{\imath=1}^{m} |\lambda_\imath|^2 P_\imath^2 = \sum_{\imath=1}^{m} P_\imath = I,$$

ceea ce arată că T este operator unitar.

32. Definiţie (calcul funcţional polinomial)
Fie $T \in \mathcal{L}(\mathbf{C^n})$ şi fie $p(z) = \sum\limits_{k=0}^{m} a_k z^k$ un polinom cu coeficienţi complecşi. O definiţie naturală pentru "valoarea lui p în T " este
$p(T) = \sum\limits_{k=0}^{m} a_k T^k$; în această formulă $T^0 = I$. Este uşor de demonstrat că pentru orice două polinoame p, q şi $\alpha, \beta \in \mathbf{C}$, avem:
(i) $(\alpha p + \beta q)(T) = \alpha p(T) + \beta q(T)$.
(ii) $(pq)(T) = p(T)q(T)$.
Aplicaţia $p \to p(T)$ se numeşte **calculul funcţional** (polinomial) al operatorului T. Extinderea acestei aplicaţii la alte clase de funcţii este o problemă importantă.
Demonstrăm mai întâi legătura dintre calculul funcţional şi teorema de descompunere spectrală pentru operatori normali.

33.Propoziţie
Fie $T \in \mathcal{L}(\mathbf{C^n})$ un operator normal şi fie $T = \sum\limits_{\imath=1}^{m} \lambda_\imath P_\imath$ descompunerea sa spectrală. Atunci, pentru orice polinom p, avem:

$$p(T) = \sum_{\imath=1}^{m} p(\lambda_\imath) P_\imath.$$

Demonstraţie Este suficient să demonstrăm că pentru orice $k \in \mathbf{N}$, avem:

$$T^k = \sum_{\imath=1}^{m} \lambda_\imath^k P_\imath.$$

Deoarece $P_\imath P_\jmath = O$ dacă $\imath \neq \jmath$ şi $P_\imath^2 = P_\imath$, rezultă:

$$T^2 = \left(\sum_{\imath=1}^{m} \lambda_\imath P_\imath\right)\left(\sum_{\jmath=1}^{m} \lambda_\jmath P_\jmath\right) = \sum_{\imath=1}^{m}\sum_{\jmath=1}^{m} \lambda_\imath \lambda_\jmath P_\imath P_\jmath = \sum_{\imath=1}^{m} \lambda_\imath^2 P_\imath.$$

Egalitatea pentru k oarecare rezultă prin inducţie.

34. Definiţie (calcul funcţional)
Fie $T \in \mathcal{L}(\mathbf{C^n})$ un operator normal având descompunerea spectrală $T = \sum_{\imath=1}^{m} \lambda_\imath P_\imath$ şi fie f o funcţie de variabilă complexă al cărei domeniu de definiţie include spectrul operatorului T. În acest caz definim:

$$f(T) = \sum_{\imath=1}^{m} f(\lambda_\imath) P_\imath.$$

Din propoziţia 33 rezultă că pentru funcţii polinomiale această definiţie coincide cu definiţia 32. Se verifică simplu următoarele proprietăţi:
(i) $(\alpha f + \beta g)(T) = \alpha f(T) + \beta g(T)$, (liniaritate),
(ii) $(fg)(T) = f(T)g(T)$, (multiplicativitate) $\forall \alpha, \beta \in \mathbf{C}$ şi pentru orice funcţii f, g definite pe spectrul lui T.
Să observăm că din această definiţie rezultă că operatorul $f(T)$ este şi el normal, iar $\sum_{\imath=1}^{m} f(\lambda_\imath) P_\imath$ este descompunerea sa spectrală.
De exemplu, dacă $f(z) = \overline{z}$, atunci $f(T) = T^\star$; Fie $g(z) = \frac{1}{z}$. Dacă operatorul T este şi inversabil (deci 0 nu este în spectrul său), atunci are sens $g(T)$; obţinem $g(T) = T^{-1}$.

O proprietate importantă a calculului funcţional este teorema de transformare a spectrului.

35. Teoremă (de transformare a spectrului)
Fie $T \in \mathcal{L}(\mathbf{C^n})$ un operator normal şi fie f o funcţie definită pe spectrul lui T; atunci:

$$\sigma\left(f(T)\right) = \{f(\lambda)\,;\, \lambda \in \sigma(T)\}.$$

Vom nota în continuare mulţimea din membrul drept al acestei egalităţi cu $f\left(\sigma(T)\right).$

Demonstraţie Fie $T = \sum_{\imath=1}^{m} \lambda_\imath P_\imath$ descompunerea spectrală a operatorului T; atunci, din descompunerea spectrală $f(T) = \sum_{\imath=1}^{m} f(\lambda_\imath) P_\imath$, rezultă că valorile proprii ale lui $f(T)$ sunt $f(\lambda_1), f(\lambda_2), .., f(\lambda_m)$, ceea ce încheie demonstraţia.

O altă aplicaţie remarcabilă a calculului funcţional este existenţa rădăcinii pătrate pozitive pentru operatori pozitivi.

36. Teoremă (rădăcina pătrată)
Fie $T \in \mathcal{L}(\mathbf{C^n})$ un operator pozitiv. Atunci există un unic operator pozitiv $S \in \mathcal{L}(\mathbf{C^n})$ astfel încât $T = S^2$; operatorul S se numeşte rădăcina pătrată pozitivă a lui T şi se notează cu $\sqrt{T}$.
Demonstraţie Deoarece T este operator pozitiv, avem incluziunea: $\sigma(T) \subseteq [0, \infty)$. Rezultă deci că funcţia radical $f(t) = \sqrt{t}$ este definită pe spectrul operatorului T; fie $S = f(T) = \sqrt{T}$. Fie $\mathrm{id}(t) = t$ funcţia identică. Deoarece $f^2 = \mathrm{id}$, din multiplicativitatea calculului funcţional rezultă că $S^2 = \mathrm{id}(T) = T$.
Din teorema de transformare a spectrului rezultă că

$$\sigma(\sqrt{T}) = \{\sqrt{\lambda}\,;\ \lambda \in \sigma(T)\} \subset [0, \infty),$$

şi deci, conform teoremei 31(ii) operatorul $\sqrt{T}$ este pozitiv.
Unicitatea lui S rezultă din unicitatea formulei de descompunere spectrală; dacă $T = \sum\limits_{\imath=1}^{m} \lambda_\imath P_\imath$ este descompunerea spectrală a lui T, atunci

$$S = \sum_{\imath=1}^{m} \sqrt{\lambda_\imath} P_\imath = \sqrt{T}$$

este descompunerea spectrală a lui $\sqrt{T}$.

37. Consecinţă
Fie $A, B \in \mathcal{L}(C^n)$ doi operatori pozitivi; dacă $AB = BA$, atunci produsul AB este de asemenea pozitiv.
Pentru demonstraţie trebuie observat că dacă A şi B comută, atunci $\sqrt{A}$ şi $\sqrt{B}$ comută şi ei.

O aplicaţie a existenţei rădăcinii pătrate este existenţa unei descompuneri analoage descompunerii polare de la numere complexe. Se ştie că orice număr complex nenul z se poate scrie (în mod unic) sub forma $z = ru$, unde $r = |z| > 0$ şi $|u| = 1$.

38. Teoremă (descompunerea polară)
Pentru orice $T \in \mathcal{L}(\mathbf{C^n})$ există un unic operator pozitiv $P \in \mathcal{L}(\mathbf{C^n})$ şi un operator unitar (nu neapărat unic) $U \in \mathcal{L}(\mathbf{C^n})$ astfel încât $T = UP$. Dacă în plus operatorul T este inversabil, atunci U este unic determinat.
Demonstraţie Vom face mai întâi demonstraţia în ipoteza că T este inversabil, apoi vom trata cazul general. Fie $P = \sqrt{T^\star T}$ şi fie
$V = PT^{-1}$; atunci, dacă notăm $U = V^{-1}$, obţinem $T = UP$, unde P este un operator pozitiv. Mai avem de arătat că U este unitar. Pentru aceasta arătăm că V este unitar; deoarece $V^\star = (T^\star)^{-1}P$, rezultă:

$$V^\star V = (T^\star)^{-1} PPT^{-1} = (T^\star)^{-1} T^\star T T^{-1} = I,$$

şi deci V este unitar. Pentru a demonstra unicitatea lui P, să presupunem că $UP = T = U_oP_o$ este o altă descompunere polară a lui T. Din egalitatea $UP = U_oP_o$, prin trecere la adjuncţi rezultă $PU^\star = P_oU_o^\star$ şi deci:

$$P^2 = PU^\star UP = P_oU_o^\star U_oP_o = P_o^2.$$

Deoarece rădăcina pătrată pozitivă este unică, rezultă că $P = P_o$. Pentru a demonstra unicitatea lui U să observăm că dacă T este inversabil atunci şi $P = U^{-1}T$ este inversabil şi deci din egalitatea $UP = U_oP$ obţinem $U = U_o$. Considerăm acum cazul general; operatorul P se construieşte la fel: $P = \sqrt{T^\star T}$. Construim acum U; pentru aceasta, să observăm că pentru orice $x \in \mathbf{C^n}$, avem:

$$\| Px \|^2 = < Px, Px > = < P^2x, x > = < T^\star Tx, x > = \| Tx \|^2 .$$

Definim operatorul U mai întâi pe subspaţiul $\mathrm{Im}(P)$ prin $U(Px) = Tx$, $\forall x \in \mathbf{C^n}$. Definiţia este corectă, în sensul că dacă
$Px_1 = Px_2$, atunci $Tx_1 = Tx_2$; pentru aceasta folosim egalitatea demonstrată mai sus:

$$0 = \| P(x_1 - x_2) \| = \| T(x_1 - x_2) \| .$$

Tot din egalitatea $\| Px \| = \| Tx \|$, rezultă că $U : \mathrm{Im}(P) \to \mathrm{Im}(T)$ este o izometrie:

$$\| U(Px) \| = \| Tx \| = \| Px \|, \forall x \in C^n.$$

De aici rezultă că subspaţiile $\mathrm{Im}(P)$ şi $\mathrm{Im}(T)$ au aceeaşi dimensiune (fiind izomorfe) şi deci şi ortogonalele lor au dimensiuni egale; fie

$$W : (\mathrm{Im}(P))^\perp \to (\mathrm{Im}(T))^\perp$$

o izometrie liniară arbitrară (există, deoarece cele două subspaţii sunt izomorfe). Prelungim U pe întregul $\mathbf{C^n}$, punând $U = W$ pe $(\mathrm{Im}(P))^\perp$. Rezultă deci că U este o izometrie pe $\mathbf{C^n}$, adică

$$\| Ux \| = \| x \|, \forall x \in \mathbf{C^n}.$$

Conform observaţiei 17 rezultă că U este operator unitar; egalitatea $UP = T$ este de asemenea verificată şi deci demonstraţia este completă.

Variante infinit dimensionale ale rezultatelor de mai sus vor fi studiate în capitolul 6.

Capitolul 4

MF.04. Spaţii Hilbert

Cuvinte cheie

produs scalar, spaţiu Hilbert, ortogonalitate, funcţională liniară, proiecţie, bază ortonormală, serie Fourier.

Pricipalul concept geometric ce nu poate fi definit satisfăcător într-un spaţiu Banach este perpendicularitatea. Noţiunea de produs scalar este instrumentul care permite construirea unei teorii geometrice apropiate de cea euclidiană în cadrul abstract al spaţiilor vectoriale. În acest paragraf prezentăm noţiunile şi rezultatele de bază din teoria spaţiilor Hilbert. Deoarece unele dintre acestea vor fi date fără demonstraţii, recomandăm următoarele surse bibliografice pentru completarea informaţiei: [B01], [C02], [D02], [F01], [H01], [O01], [R02],[S02].

4.1 MF.04.1. Geometria spaţiilor Hilbert

1. Definiţie

Fie X un spaţiu vectorial complex; se numeşte **produs scalar pe** X orice aplicaţie $<,>$: $X \times X \to \mathbf{C}$ care, pentru orice $x, y, z \in X$ şi α, $\beta \in \mathbf{C}$, verifică proprietăţile:

(a) $< \alpha x + \beta y\, , z >= \alpha < x, z > + \beta < y, z >$

(b) $< x, y >= \overline{< y, x >}$

(c) $< x, x > \geq 0$

(d) $< x, x >= 0 \Leftrightarrow x = 0$.

Perechea $(X, <,>)$ se numeşte spaţiu cu produs scalar sau spaţiu prehilbertian.

2. Observaţie

Fie $(H, <,>)$ un spaţiu cu produs scalar. Atunci, pentru orice

vectori $x, y \in H$, avem:
(a) relaţia de polarizare:

$$< x, y >=$$

$$= \frac{1}{4}(< x+y, x+y > - < x-y, x-y > +$$

$$+i < x+iy, x+iy > -i < x-iy, x-iy >).$$

(b) inegalitatea lui Schwarz:

$$| < x, y > | \leq \sqrt{< x, x >< y, y >}.$$

(c) Aplicaţia $\| \|: H \to [0, \infty)$, $\| x \| = \sqrt{< x, x >}$ este o normă pe H; o vom numi **norma definită de produsul scalar**.
(d) Produsul scalar este aplicaţie continuă: dacă $\lim\limits_{n\to\infty} x_n = x$ şi $\lim\limits_{n\to\infty} y_n = y$, atunci $\lim\limits_{n\to\infty} < x_n, y_n > = < x, y >$.
Demonstraţiile se pot găsi în: [B01], [D02].

3. Definiţie
Fie $(H, < , >)$ un spaţiu prehilbertian şi fie $\| \|$ norma indusă de produsul scalar. Dacă $(H, \| \|)$ este complet atunci $(H, < , >)$ se numeşte **spaţiu Hilbert**. Doi vectori $x, y \in H$ se numesc **ortogonali** (sau **perpendiculari**) dacă $< x, y > = 0$. Vectorul x se numeşte ortogonal pe submulţimea nevidă $M \subseteq H$ (şi vom nota $x \perp M$) dacă x este ortogonal pe toţi vectorii din M. Ortogonalul mulţimii M este, prin definiţie, $M^\perp = \{y \in H\,;\, y \perp x, \forall x \in M\}$. Este simplu de arătat că $M^\perp$ este subspaţiu vectorial închis în H (se foloseşte continuitatea produsului scalar). Propunem de asemenea ca exerciţiu egalitatea (aici bara înseamnă închiderea mulţimii respective):
$\left(K^\perp\right)^\perp = \overline{K}$, pentru orice subspaţiu $K \subseteq H$. Dacă $M \neq \{0\}$, atunci $M^\perp \neq H$. Submulţimea nevidă $M \subseteq H$ se numeşte ortogonală dacă $x \perp y$, $\forall x, y \in M$ şi ortonormală dacă, în plus, $\| x \| = 1$, $\forall x \in M$.
Două spaţii Hilbert $(H_1, < , >_1)$ şi $(H_2, < , >_2)$ se numesc izomorfe dacă există un izomorfism U de spaţii vectoriale de la H_1 la H_2 astfel încât $< Ux, Uy >_2 = < x, y >_1$. Aplicaţia U se numeşte în acest caz izomorfism de spaţii Hilbert sau operator unitar.
Următoarele două proprietăţi sunt generalizări ale unor rezultate din geometria elementară.

4. Propoziţie
(a) Fie $(H, < , >)$ un spaţiu prehilbertian. Atunci:

$$\| x+y \|^2 + \| x-y \|^2 = 2(\| x \|^2 + \| y \|^2), \forall x, y \in H.$$

(b) Reciproc, dacă $(X, \| \|)$ este un spaţiu normat astfel încât este verificată

egalitatea de la punctul **(a)**, atunci există un produs scalar $< , >$ pe X astfel încât $\| x \|^2 = < x, x >$; (legea paralelogramului).
(c) Pentru orice mulţime finită ortogonală de vectori $\{x_1, x_2, .., x_n\}$ din spaţiul prehilbertian H, are loc egalitatea (teorema lui Pitagora):

$$\| \sum_{j=1}^{n} x_j \|^2 = \sum_{j=1}^{n} \| x_j \|^2 .$$

Pentru demonstraţii se pot consulta [B01], [D02], [O01].

Următorul rezultat este fundamental în studiul spaţiilor Hilbert. Reamintim că o submulţime M a unui spaţiu vectorial se numeşte convexă dacă $\alpha x + (1 - \alpha) y \in M$, $\forall x, y \in M$, $\forall \alpha \in [0, 1]$.

5. Teoremă
Fie H un spaţiu Hilbert şi fie $M \subseteq H$ o mulţime nevidă, închisă şi convexă. Atunci există şi este unic un vector $x_M \in M$ astfel încât

$$\| x_M \| = \inf\{\| x \| ; x \in M\}.$$

Demonstraţie Să notăm cu $\delta = \inf\{\| x \| ; x \in M\}$ şi fie $(x_n)_n$ un şir de elemente din M astfel încât $\lim\limits_{n\to\infty} \| x_n \| = \delta$. Deoarece M este mulţime convexă, rezultă că pentru orice $n, m \in N$ avem $\frac{1}{2}x_n + \frac{1}{2}x_m \in M$ şi deci:

$$\| \frac{x_n + x_m}{2} \|^2 \geq \delta^2 , \forall n, m \in N.$$

Aplicând legea paralelogramului vectorilor $\frac{1}{2}x_n$ şi $\frac{1}{2}x_m$ şi folosind inegalitatea de mai sus, rezultă:

$$\| \frac{x_n - x_m}{2} \|^2 = 2 \| \frac{x_n}{2} \|^2 + 2 \| \frac{x_m}{2} \|^2 - \| \frac{x_n + x_m}{2} \|^2 \leq$$
$$\leq \frac{1}{2}(\| x_n \|^2 + \| x_m \|^2) - \delta^2$$

şi deci :

$$\| x_n - x_m \|^2 \leq 2 (\| x_n \|^2 + \| x_m \|^2) - 4\delta^2.$$

De aici rezultă că $\limsup\limits_{n,m\to\infty} \| x_n - x_m \|^2 = 0$, ceea ce arată că $(x_n)_n$ este şir Cauchy; fie $x_M = \lim\limits_{n\to\infty} x_n$. Deoarece M este mulţime închisă, rezultă că $x_M \in M$, iar din continuitatea normei avem $\| x_M \| = \delta$. Pentru a demonstra unicitatea, presupunem prin absurd că există x_M şi y_M în M, diferiţi, astfel încât $\| x_M \| = \| y_M \| = \delta$. Aplicând legea paralelogramului vectorilor x_M şi y_M şi repetând raţionamentul anterior, rezultă $\| x_M - y_M \| = 0$, ceea ce constituie o contradicţie.

6. Consecinţă (teorema proiecţiei pe un subspaţiu închis)
Fie H un spaţiu Hilbert şi fie $K \subseteq H$ un subspaţiu închis. Atunci, pentru orice $x \in H$, există şi este unic $y \in K$ astfel încât $x - y \in K^{\perp}$ şi $\| x - y \| \leq \| x - z \|$, pentru orice $z \in K$. Elementul y (care depinde în mod evident de x şi K) se numeşte proiecţia lui x pe K. Pentru demonstraţie se aplică teorema precedentă.
Rezultatul următor este o generalizare în spaţii Hilbert a descompunerii unui vector după două direcţii perpendiculare din geometria euclidiană.

7. Teoremă (descompunerea ortogonală)
Fie H un spaţiu Hilbert şi fie K un subspaţiu închis al său. Atunci, pentru orice vector $x \in H$ există $y \in K$ şi $z \in K^{\perp}$ astfel încât $x = y + z$; în plus, această descompunere este unică. Vom nota această descompunere ortogonală $H = K \bigoplus K^{\perp}$.
Demonstraţie Fie $x \in H$, arbitrar fixat. Aplicând teorema 5 mulţimii nevide, convexe şi închise $M = \{x - u \,,\, u \in K\}$, rezultă că există un vector unic $z \in M$ cu proprietatea $\| z \| = \inf\{\| v \| \,;\, v \in M\}$. Fie $u \in K$ astfel încât $\| u \| = 1$; atunci $z - < z, u > u \in M$ şi deci:

$$\| z \|^2 \leq \| z - < z, u > u \|^2 = < z - < z, u > u, z - < z, u > u > =$$

$$= \| z \|^2 - \overline{< z, u >} < z, u > - < z, u > \overline{< z, u >} + | < z, u > |^2 =$$

$$= \| z \|^2 - | < z, u > |^2,$$

ceea ce este posibil numai dacă $< z, u > = 0$. Am demonstrat deci că $z \in K^{\perp}$. Din definiţia lui M rezultă că există $y \in K$ astfel încât $x = y + z$. Pentru a demonstra unicitatea, presupunem prin absurd că există $y_1, y_2 \in K$ şi $z_1, z_2 \in K^{\perp}$ astfel încât $y_1 \neq y_2$, $z_1 \neq z_2$ şi $x = y_1 + z_1 = y_2 + z_2$. De aici rezultă că $y_1 - y_2 = z_2 - z_1$; dar $y_1 - y_2 \in K$ şi $z_2 - z_1 \in K^{\perp}$, şi deci $y_1 - y_2 = z_2 - z_1 \in K \cap K^{\perp} = \{0\}$, contradicţie care încheie demonstraţia.

8. Definiţie
Fie X un spaţiu normat. Se numeşte **funcţională liniară** pe X orice aplicaţie liniară $f : X \to \mathbf{C}$. Aşa cum am văzut în capitolul 2, funcţionalele liniare şi continue au un rol deosebit în studiul spaţiilor Banach. Pe spaţii Hilbert este adevărat următorul rezultat remarcabil (teorema lui Riesz de reprezentare a funcţionalelor liniare şi continue pe un spaţiu Hilbert).

9. Teorema lui Riesz
Fie H un spaţiu Hilbert.
(a) Pentru orice $y \in H$, fixat, aplicaţia $f_y : H \to \mathbf{C}$, $f_y(x) = < x, y >$ este funcţională liniară şi continuă.

(b) Reciproc, dacă f este o funcţională liniară şi continuă pe H, atunci există şi este unic un vector $y_f \in H$ astfel încât $f(x) =< x, y_f >$, $\forall x \in H$; în plus, are loc egalitatea:

$$\sup\{|f(x)|\,;\, x \in H \text{ şi } \| x \|= 1\} =\| y_f \| .$$

Demonstraţie Punctul **(a)** este evident (pentru continuitate se foloseşte inegalitatea lui Schwarz).
(b) Fie $\mathrm{Ker}(f) = \{x \in H \mid f(x) = 0\}$ nucleul aplicaţiei f. Dacă $\mathrm{Ker}(f) = H$, atunci f este identic nulă şi deci putem lua $y_f = 0$. Dacă $\mathrm{Ker}(f) \neq H$, atunci există $z \in \mathrm{Ker}(f)^{\perp}$ cu proprietatea $\| z \|= 1$ şi $f(z) \neq 0$. Pentru orice $x \in H$, vectorul $x - (\frac{f(x)}{f(z)})z$ este în $\mathrm{Ker}(f)$, şi deci:

$$f(x) = f(x) < z, z >=< \frac{f(x)}{f(z)}z, \overline{f(z)}z >=$$

$$=< x - \frac{f(x)}{f(z)}z, \overline{f(z)}z > + < \frac{f(x)}{f(z)}z, \overline{f(z)}z >=< x, \overline{f(z)}z >,$$

şi deci putem alege $y_f = \overline{f(z)}\ z$. Unicitatea lui y_f este imediată. Din inegalitatea lui Schwarz, rezultă

$$\sup\{|f(x)|;\, x \in H \text{ şi } \| x \|= 1\} = \sup\{\, | < x, y_f > |;\, x \in H \text{şi } \| x \|= 1\} \leq$$

$$\leq \sup\{\| x \|\| y_f \|\,;\, x \in H \text{ şi } \| x \|= 1\} =\| y_f \| .$$

Pentru a demonstra cealaltă inegalitate, să observăm că:

$$|f\left(\frac{y_f}{\| y_f \|}\right)| =\| y_f \| .$$

În particular, rezultă că supremumul este atins în punctul $\frac{y_f}{\|y_f\|}$.

4.2 MF.04.2. Serii Fourier

10. Definiţie
Fie H un spaţiu Hilbert. Se numeşte **bază ortonormală** în H orice submulţime $\mathcal{B} = \{\varepsilon_\imath\}_{\imath \in \mathcal{J}}$ cu proprietăţile:
(i) $< \varepsilon_\imath, \varepsilon_\jmath >= \delta^\imath_\jmath$, $\forall \imath, \jmath \in \mathcal{B}$; (am notat cu $\delta^\imath_\jmath$ simbolul lui Kronecker).
(ii) Subspaţiul vectorial generat de $\mathcal{B}$ este dens în H.
Se demonstrează că în orice spaţiu Hilbert există cel puţin o bază ortonormală, ([B01], [D02]); de asemenea, orice două baze ortonormale ale aceluiaşi spaţiu Hilbert H au acelaşi număr de elemente, numit **dimensiunea** lui H. Spaţiile Hilbert care admit baze ortonormale cel mult numărabile (card $(\mathcal{B}) \leq \aleph_\circ$), se numesc **separabile**. Cum în această lucrare vom considera numai spaţii Hilbert separabile, de aici înainte, prin spaţiu Hilbert vom

înţelege un spaţiu Hilbert separabil.
Fie $\mathcal{B} = \{\varepsilon_n\}_{n\in N}$ o bază ortonormală fixată şi fie $x \in H$ un vector arbitrar fixat; **coeficienţii Fourier** (în raport cu baza $\mathcal{B}$), ai lui x sunt, prin definiţie, numerele $\hat{x}(n) =< x, \varepsilon_n >$. Vom nota cu $\hat{x} : N \to \mathbf{C}$ **şirul coeficienţilor Fourier**. Seria $\sum\limits_{n\in N} \hat{x}(n)\varepsilon_n$ se numeşte **seria Fourier** asociată lui x (în baza $\mathcal{B}$).

11. Teoremă
În ipotezele şi notaţiile de mai sus, seria Fourier converge la x şi are loc (egalitatea lui Parseval):

$$\| x \|^2 = \sum_{n\in N} |\hat{x}(n)|^2.$$

Demonstraţie Fie u_n şirul sumelor parţiale asociat seriei Fourier; pentru orice $k \in \{1, 2, .., n\}$, avem:

$$< u_n, \varepsilon_k >= \sum_{j=1}^{n} \hat{x}(j) < \varepsilon_j, \varepsilon_k >= \hat{x}(k) =< x, \varepsilon_k >,$$

ceea ce arată că $x - u_n \perp \varepsilon_k$, $\forall k \leq n$. Fie, pentru orice $n \in N$ subspaţiul H_n generat de $\{\varepsilon_1, \varepsilon_2, .., \varepsilon_n\}$; H_n este subspaţiu închis (deoarece este finit dimensional) şi $x - u_n \in H_n^{\perp}$. Fie acum un vector arbitrar $v \in H_n$; conform teoremei lui Pitagora, avem:

$$\| x - v \|^2 = \| x - u_n \|^2 + \| v - u_n \|^2 \geq \| x - u_n \|^2 .$$

De aici rezultă că u_n este proiecţia vectorului x pe subspaţiul H_n, conform consecinţei 6. Fie $\epsilon > 0$; deoarece subspaţiul liniar generat de $\mathcal{B}$ este dens în H, există $n(\epsilon) \in N$ şi un vector $z \in H_{n(\epsilon)}$ astfel încât $\| z - x \| < \epsilon$. Fie $n \geq n(\epsilon)$; aplicând din nou teorema proiecţiei, rezultă (deoarece $z \in H_n$):

$$\| x - u_n \| \leq \| x - z \| < \epsilon,$$

ceea ce arată că $\lim\limits_{n\to\infty} u_n = x$.
Pentru a demonstra egalitatea lui Parseval, să observăm că pentru orice $n \in \mathbf{N}$, avem:

$$\| u_n \|^2 =< \sum_{j=1}^{n} \hat{x}(j)\varepsilon_j, \sum_{k=1}^{n} \hat{x}(k)\varepsilon_k >=$$

$$= \sum_{j,k=1}^{n} \hat{x}(j)\overline{\hat{x}(k)} < \varepsilon_j, \varepsilon_k >= \sum_{j=1}^{n} |\hat{x}(j)|^2.$$

Pentru $n \to \infty$, se obţine egalitatea lui Parseval.

12. Observaţie
Seria Fourier a vectorului x se mai numeşte şi dezvoltarea (în baza $\mathcal{B}$) a lui x. Se poate demonstra că această dezvoltare este unică, deci dacă seria $\sum\limits_{n\in\mathbf{N}} \alpha_n \varepsilon_n$ converge la x, atunci $\alpha_n = \hat{x}(n)$. De asemenea, un calcul direct arată că pentru orice $x, y \in H$ are loc egalitatea: $< x, y >= \sum\limits_{n\in N} \hat{x}(n)\overline{\hat{y}(n)}$.
Încheiem acest capitol cu exemple de spaţii Hilbert şi noţiuni despre transformarea Fourier.

4.3 MF.04.3. Exemple

13. Spaţiul Banach $(\mathbf{C^n}, \| \|_2)$ este spaţiu Hilbert, produsul scalar fiind $< x, y >= \sum\limits_{j=1}^{n} x_j \overline{y_j}$. Baza canonică a spaţiului vectorial $\mathbf{C^n}$ este bază ortonormală. Nu este dificil de demonstrat că orice spatiu Hilbert complex (real) de dimensiune n este izomorf cu $\mathbf{C^n}$, (respectiv $\mathbf{R^n}$).

14. Spaţiul $\ell^2(\mathbf{Z})$
Folosind legea paralelogramului, se demonstrează că dintre spaţiile Banach $\ell^p(\mathbf{Z})$ numai $\ell^2(\mathbf{Z})$ este spaţiu Hilbert, produsul scalar fiind

$$< x, y >= \sum_{n\in N} x(n)\overline{y(n)}.$$

Fie, pentru orice $n \in \mathbf{Z}$, şirul $\sigma_n : \mathbf{Z} \to \mathbf{C}$, definit prin:

$$\sigma_n(m) = \begin{cases} 1 & \text{dacă} \quad m = n \\ 0 & \text{dacă} \quad m \neq n \end{cases}$$

Atunci mulţimea $(\sigma_n)_{n\in Z}$ este bază ortonormală (numită baza canonică) în $\ell^2(\mathbf{Z})$. Dacă $x \in \ell^2(\mathbf{Z})$, atunci şirul coeficienţilor săi Fourier este $\hat{x}(n) = x(n)\,, \forall n \in \mathbf{Z}$, iar seria sa Fourier este $\sum\limits_{n\in Z} x(n)\sigma_n$. Un subspaţiu închis inclus în $\ell^2(\mathbf{Z})$ este

$$h^2(\mathbf{Z}) = \{x \in \ell^2(\mathbf{Z})\,;\, x(n) = 0\,, \forall n < 0\}.$$

Evident că $\ell^2(N)$ se poate identifica cu acest subspaţiu, prelungind şirurile cu 0 pentru $n < 0$. Se poate arăta că orice spaţiu Hilbert (separabil) H este izomorf cu un spaţiu de tip ℓ^2. Într-adevăr, dacă $\mathcal{B} = \{\varepsilon_n\}_{n\in N}$ este o bază ortonormală a lui H, atunci aplicaţia

$$H \ni x \to \hat{x} \in \ell^2(\mathbf{N})$$

este un izomorfism de spaţii Hilbert: [B01], [C02], [D02].

15. Un rezultat similar cu cel din exemplul anterior este adevărat şi pentru spaţiile $L^p(\Omega, \mu)$, (a se vedea Capitolul 5): $L^p(\Omega, \mu)$ este spaţiu Hilbert dacă şi numai dacă $p = 2$, produsul scalar fiind definit prin relaţia $< f, g >= \int_\Omega f\, \overline{g}\, d\mu$.

Vom studia în continuare spaţiul Hilbert al funcţiilor de pătrat integrabil pe cercul unitate.

16. Definiţie
Fie $\mathcal{S}^1$ cercul unitate (considerat cu măsura Lebesgue) şi fie $L^2(\mathcal{S}^1)$ spaţiul Hilbert al funcţiilor de pătrat integrabil cu produsul scalar:

$$< f, g >= \frac{1}{2\pi} \int_0^{2\pi} f(e^{it})\overline{g(e^{it})}dt$$

şi norma:

$$\| f \|_2 = \frac{1}{\sqrt{2\pi}} \sqrt{\int_0^{2\pi} |f(e^{it})|^2 dt}.$$

Vom defini în continuare o bază ortonormală remarcabilă în $L^2(\mathcal{S}^1)$. Fie, pentru orice $n \in Z$, funcţia:

$$\omega_n(e^{it}) = e^{int}$$

şi fie mulţimea (numărabilă) $\{\omega_n\}_{n \in Z}$.
Are loc următorul rezultat fundamental:

17. Teoremă
Mulţimea $\{\omega_n\}_{n \in Z}$ este bază ortonormală în spaţiul $L^2(\mathcal{S}^1)$.
Pentru demonstraţie: [B01], [F02] [D02]. În continuare vom subînţelege că pe spaţiul $L^2(\mathcal{S}^1)$ a fost fixată baza ortonormală $\{\omega_n\}_{n \in Z}$.
Şirul coeficienţilor Fourier asociaţi unei funcţii $f \in L^2(\mathcal{S}^1)$ este:

$$\widehat{f} : Z \to \mathbf{C},\ \widehat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) e^{-int} dt,$$

iar seria (de funcţii) Fourier corespunzătoare este:

$$\sum_{n \in Z} \widehat{f}(n) \omega_n.$$

Sumele parţiale ale acestei serii se numesc **polinoame trigonometrice**:

$$S_n(e^{it}) = \sum_{k=-n}^{n} \widehat{f}(k) e^{ikt}.$$

Seria de funcţii $\sum_{n \in Z} \widehat{f}(n)\omega_n$ converge în spaţiul $L^2(\mathcal{S}^1)$ la funcţia f, adică:

$$\lim_{n \to \infty} \| f - S_n \|_2 = 0.$$

Teorema 11 nu dă informaţii despre alte tipuri de convergenţă specifice spaţiilor de funcţii (convergenţă punctuală sau convergenţă uniformă, de exemplu) care se pot pune în legătură cu seria Fourier. Există în această direcţie câteva teoreme clasice: Fejer, Dini, Dirichlet.

18. Definiţie (transformarea Fourier pe spaţiul $L^2(\mathcal{S}^1)$)
Fie $f \in L^2(\mathcal{S}^1)$; din identitatea lui Parseval rezultă faptul că şirul $\widehat{f}$ aparţine spaţiului $\ell^2(\mathbf{Z})$ şi $\| f \|_2 = \| \widehat{f} \|_2$. Rezultă deci că aplicaţia:

$$\mathcal{F} : L^2(\mathcal{S}^1) \to \ell^2(\mathbf{Z}) \, , \, \mathcal{F}(f) = \widehat{f}$$

este izometrie liniară; $\mathcal{F}$ se numeşte **transformarea Fourier** (între spaţiile $L^2(\mathcal{S}^1)$ şi $\ell^2(\mathbf{Z})$), iar $\widehat{f}$ se numeşte **transformata Fourier** (sau **Fourier-Plancherel**) a funcţiei f.
Din teorema 11 şi din completitudinea spaţiului $L^2(\mathcal{S}^1)$, rezultă că aplicaţia $\mathcal{F}$ este şi surjectivă: pentru orice $x \in \ell^2(\mathbf{Z})$, seria $\sum_{n \in Z} x(n)\, \omega_n$ converge în spaţiul $L^2(\mathcal{S}^1)$, deci defineşte o funcţie f (de fapt o clasă de echivalenţă de funcţii egale a.p.t.) care are în mod evident proprietatea $\mathcal{F}(f) = x$, ([B01], [F02], [D01]). În concluzie, transformarea Fourier $\mathcal{F}$ este un izomorfism de spaţii Hilbert având ca inversă aplicaţia

$$\mathcal{F}^{-1} : \ell^2(Z) \to L^2(\mathcal{S}^1), \, \mathcal{F}^{-1}(x) = \sum_{n \in Z} x(n)\, \omega_n.$$

Menţionăm că în egalitatea de mai sus $\sum_{n \in Z} x(n)\omega_n$ semnifică suma seriei în sensul normei $\| \, \|_2$.

Restricţia aplicaţiei $\mathcal{F}^{-1}$ la subspaţiul $\ell^1(\mathbf{Z}) \subset \ell^2(\mathbf{Z})$ admite o formulă punctuală explicită.

19. Teoremă
Dacă $\alpha \in \ell^1(\mathbf{Z})$, atunci:

$$\left(\mathcal{F}^{-1}\alpha\right)(e^{it}) = \sum_{n \in Z} \alpha(n) e^{int}, \, \forall e^{it} \in \mathcal{S}^1.$$

Demonstraţie Deoarece $\alpha \in \ell^1(\mathbf{Z})$, seria $\sum_{n \in Z} \alpha(n)e^{int}$ converge absolut şi uniform pe $\mathcal{S}^1$:

$$\sum_{n \in Z} |\alpha(n)e^{int}| \leq \sum_{n \in Z} |\alpha(n)| = \| \alpha \|_1, \, \forall e^{it} \in \mathcal{S}^1.$$

Fie f suma seriei de mai sus. Atunci f este o funcţie continuă şi mărginită; în particular, rezultă că $f \in L^2(\mathcal{S}^1)$, deci îi putem calcula coeficienţii Fourier:

$$\widehat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})e^{-int}dt = \frac{1}{2\pi}\int_0^{2\pi}\left(\sum_{m\in Z}\alpha(m)e^{imt}\right)e^{-int}dt =$$

$$= \frac{1}{2\pi}\sum_{m\in Z}\alpha(m)\int_0^{2\pi} e^{it(m-n)}dt = \alpha(n),\ \forall n \in \mathbf{Z}.$$

Comutarea seriei cu integrala este justificată deoarece amândouă sunt absolut convergente şi ca urmare se poate aplica teorema lui Fubini. În concluzie, $\mathcal{F}f = \alpha$, deci $\mathcal{F}^{-1}\alpha = f$, (egalitatea este adevărată peste tot, deoarece f este funcţie continuă), ceea ce încheie demonstraţia.

20. Observaţie

Conform celor de mai sus, restricţia lui $\mathcal{F}^{-1}$ la subspaţiul $\ell^1(\mathbf{Z})$ ia valori în mulţimea funcţiilor continue (definite pe cerc). Fie

$$A(\mathcal{S}^1) = \{\mathcal{F}^{-1}x\,;\, x \in \ell^1(\mathbf{Z})\}.$$

Se poate demonstra că $A(\mathcal{S}^1)$ este o submulţime densă în $\mathcal{C}(\mathcal{S}^1)$ (în norma $\|\ \|_\infty$); cum $\mathcal{C}(\mathcal{S}^1)$ este la rândul ei densă în $L^2(\mathcal{S}^1)$, (în norma $\|\ \|_2$), rezultă că $A(\mathcal{S}^1)$ este o submulţime densă în $L^2(\mathcal{S}^1)$, adică:

$$\forall f \in L^2(\mathcal{S}^1),\ \forall \epsilon > 0,\ \exists x \in \ell^1(Z) \text{ astfel încât } \| f - \mathcal{F}^{-1}x \|_2 < \epsilon.$$

Capitolul 5

MF.05. Măsură şi integrală

Cuvinte cheie

măsură, integrală, funcţie măsurabilă, funcţie integrabilă, serie trigonometrică.
O parte din rezultatele din acest capitol vor fi prezentate fără demonstraţii; recomandăm ca surse bibliografice: [H02], [F02], [R01], [S02].

5.1 MF.05.1. Spaţii cu măsură

1. Definiţie
Fie X o mulţime nevidă şi fie $\mathcal{P}(X)$ mulţimea părţilor lui X. O submulţime $\mathcal{A} \subseteq \mathcal{P}(X)$ se numeşte $\sigma-$ algebră pe X dacă verifică următoarele proprietăţi:
i. $X \in \mathcal{A}$.
ii. dacă $A \in \mathcal{A}$ atunci $X \setminus A \in \mathcal{A}$.
iii. dacă $A_n \in \mathcal{A}$, $\forall n \in N$ atunci $\cup_{n \in N} A_n \in \mathcal{A}$.
În acest caz $(X, \mathcal{A})$ se numeşte spaţiu măsurabil iar elementele σ-algebrei $\mathcal{A}$ se numesc mulţimi măsurabile.

2. Propoziţie
Dacă $\mathcal{A}$ este o σ-algebră pe X, atunci:
i. $\emptyset \in \mathcal{A}$.
ii. dacă $A, B \in \mathcal{A}$ atunci $A \bigcup B \in \mathcal{A}$, $A \cap B \in \mathcal{A}$, $A \setminus B \in \mathcal{A}$.
iii. dacă $A_n \in \mathcal{A}$, $\forall n \in N$ atunci $\bigcap\limits_{n \in N} A_n \in \mathcal{A}$.
iv. dacă $(\mathcal{A})_{i \in J}$ sunt σ-algebre pe X atunci intersecţia $\bigcap\limits_{i \in J} \mathcal{A}_i$ este σ-algebră pe X.

3. Definiţii
Dacă $\mathcal{C} \subseteq \mathcal{P}(X)$ atunci σ-algebra generată de $\mathcal{C}$ se notează $\mathcal{A}_{\mathcal{C}}$ şi este definită prin

$$\mathcal{A}_{\mathcal{C}} = \bigcap \{\mathcal{B} \mid \mathcal{B} \text{ este } \sigma - \text{algebră pe } X, \text{ şi } \mathcal{B} \supseteq \mathcal{C}\}.$$

Dacă (Y, d) este un spaţiu metric, atunci σ-algebra mulţimilor Boreliene pe Y este σ-algebra generată de familia mulţimilor deschise din Y.
În cazul particular $Y = \mathbf{R}$, σ-algebra mulţimilor Boreliene (pe $\mathbf{R}$) coincide cu σ-algebra generată de oricare din următoarele tipuri de intervale:
$\mathcal{C}_1 = \{(-\infty, b) \mid b \in \mathbf{R}\}$
$\mathcal{C}_2 = \{(a, \infty) \mid a \in \mathbf{R}\}$
$\mathcal{C}_3 = \{(a, b) \mid a, b \in \mathbf{R}\}$,
deoarece orice mulţime deschisă din $\mathbf{R}$ este reuniune cel mult numărabilă de intervale deschise.

4. Definiţie
Fie $(X, \mathcal{A})$ un spaţiu cu măsură şi (Y, d) un spaţiu metric; o aplicaţie $f : X \mapsto Y$ se numeşte măsurabilă dacă $f^{-1}(B) \in \mathcal{A}$, $\forall B$ mulţime Boreliană din Y.
Dacă $A \subseteq X$, atunci funcţia caracteristică $\chi_A : X \mapsto R$ este măsurabilă dacă şi numai dacă $A \in \mathcal{A}$.
O aplicaţie $s : X \mapsto R$ se numeşte simplă (sau etajată) dacă mulţimea $s(X)$ este finită, sau, echivalent, dacă există $A_1, A_2, ..., A_n \in \mathcal{P}(X)$ şi $\alpha_1, \alpha_2, ..., \alpha_n \in R$ astfel încât $s = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$. Evident, s este măsurabilă dacă şi numai dacă mulţimile $A_1, A_2, ..., A_n$ sunt măsurabile.

Are loc următorul rezultat de aproximare:
5. Propoziţie
Orice funcţie măsurabilă şi pozitivă este limita punctuală a unui şir crescător de funcţii simple măsurabile şi pozitive.

6. Definiţie
Fie $(X, \mathcal{A})$ un spaţiu măsurabil; o aplicaţie

$$\mu : \mathcal{A} \mapsto [0, \infty]$$

se numeşte măsură (pe X) dacă:
i. $\mu(\emptyset) = 0$
ii. pentru orice $(A_n)_{n \in N} \subseteq \mathcal{A}$ astfel încât $A_n \cap A_m = \emptyset$, $\forall n \neq m$ rezultă

$$\mu\left(\bigcup_{n \in N} A_n\right) = \sum_{n \in N} \mu\left(A_n\right).$$

$(X, \mathcal{A}, \mu)$ se numeşte spaţiu cu măsură. Proprietatea **ii** de mai sus se numeşte numărabil-aditivitate.

7. Propoziţie
Dacă $(X, \mathcal{A}, \mu)$ este un spaţiu cu măsură, atunci:
i. $\forall\, A, B \in \mathcal{A},\ A \subseteq B\ \Rightarrow\ \mu(A) \leq \mu(B)$. (monotonie).
ii. pentru orice $(A_n)_{n\in N} \subseteq \mathcal{A}$ rezultă

$$\mu\left(\bigcup_{n\in N} A_n\right) \leq \sum_{n\in N} \mu\left(A_n\right) \quad \text{(numărabil-subaditivitate).}$$

8. Exemple de măsuri
Se demonstrează că următoarele aplicaţii sunt măsuri:
a. Fie $X \neq \emptyset$, fie $\mathcal{P}(X)$ mulţimea părţilor lui X şi fie $a \in X$, un element arbitrar fixat. Măsura Dirac concentrată în punctul a este, prin definiţie:

$$\delta_a : \mathcal{P}(X) \mapsto [0, \infty),\ \delta_a(A) = \begin{cases} 1 & \text{dacă} \quad a \in A \\ 0 & \text{dacă} \quad a \notin A \end{cases}$$

b. Măsura de numărare pe X este, prin definiţie:

$$\mu_c : \mathcal{P}(X) \mapsto [0, \infty],\ \mu_c(A) = \begin{cases} \text{card}A & \text{dacă} \quad A \text{ este finită} \\ \infty & \text{dacă} \quad A \text{ este infinită} \end{cases}$$

c. Presupunem în plus că mulţimea X este finită. Măsura de probabilitate pe X este, prin definiţie:

$$P : \mathcal{P}(X) \mapsto [0, 1],\ P(A) = \frac{\text{card}(A)}{\text{card}(X)}.$$

9. Definiţie
O mulţime măsurabilă $A \in \mathcal{A}$ se numeşte de măsură nulă dacă $\mu(A) = 0$.
Două funcţii măsurabile se numesc egale aproape peste tot (se notează $f = g\,(a.p.t.)$) dacă mulţimea $\{x \in X \mid f(x) \neq g(x)\}$ este de măsură nulă.
Relaţia de egalitate aproape peste tot este relaţie de echivalenţă pe mulţimea funcţiilor măsurabile.

5.2 MF.05.2. Funcţii integrabile

10. Definiţii
Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură şi $s : X \mapsto [0, \infty]$, $s = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$ o funcţie

simplă, pozitivă, măsurabilă. Integrala lui s în raport cu măsura μ este, prin definiţie:

$$\int_X s d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i).$$

Dacă $f : X \mapsto [0, \infty]$ este o funcţie măsurabilă pozitivă atunci integrala lui f în raport cu măsura μ este:

$$\int_X f d\mu = \sup\{\int_X s d\mu \mid s \text{ funcţie simpla măsurabilă, } 0 \leq s \leq f\}.$$

Dacă $f : X \mapsto [0, \infty]$ este o funcţie măsurabilă pozitivă şi dacă $A \in \mathcal{A}$, atunci integrala lui f pe mulţimea A în raport cu măsura μ este:

$$\int_A f d\mu = \int_X f \cdot \chi_A d\mu.$$

Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură; o funcţie măsurabilă $f : X \mapsto \mathbf{C}$ se numeşte integrabilă dacă $\displaystyle\int_X |f| d\mu < \infty$.
Mulţimea funcţiilor integrabile este spaţiu vectorial cu operaţiile uzuale de adunare şi inmulţire cu scalari.

Fie $f : X \mapsto \mathbf{C}$ o funcţie integrabilă; atunci $f = u + iv$, unde u şi v sunt funcţii măsurabile reale; descompunând $u = u^+ - u^-$, $v = v^+ - v^-$ (aici u^+, v^+ şi u^-, v^- sunt părţile pozitive şi respectiv negative ale lui u şi v), atunci integrala lul f în raport cu măsura μ este

$$\int_X f d\mu = \int_X u^+ d\mu - \int_X u^- d\mu + i \left(\int_X v^+ - \int_X v^- d\mu \right).$$

Integrala astfel definită are proprietăţile:

i. $\displaystyle\int_X (\alpha f + \beta g) \, d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu$, $\forall f, g$ integrabile şi $\forall \, \alpha, \beta \in \mathbf{C}$.

ii. $\displaystyle\left| \int_X f d\mu \right| \leq \int_X |f| d\mu$, $\forall \, f$ integrabilă.

11. Exemple de integrale

Fie δ_a măsura Dirac şi fie $f : X \mapsto \mathbf{R}$. Atunci:

$$\int_X f d\delta_a = f(a).$$

Dacă μ_c este măsura de numărare pe $\mathbf{N}$ şi $f : \mathbf{N} \mapsto \mathbf{R}$, atunci:

$$\int_N f d\mu_c = \sum_{n \in N} f(n),$$

în ipoteza că seria din membrul drept este convergentă sau are suma $\pm\infty$.
Fie P măsura de probabilitate şi fie $f : X \mapsto \mathbf{R}$. Atunci:

$$\int_X f dP = \frac{\sum_{x \in X} f(x)}{\text{card}(X)}.$$

Dăm în continuare câteva rezultate fundamentale din teoria funcţiilor integrabile.

12. Teorema de convergenţă monotonă a lui Lebesgue
Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură şi fie $f_n : X \mapsto [0, \infty]$ un şir crescător de funcţii măsurabile: $f_n \leq f_{n+1}$, $\forall n \in N$. Dacă f este limita punctuală a şirului f_n, atunci

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

În particular, dacă $f_n : X \mapsto [0, \infty]$, atunci:

$$\int_X \sum_{n\in N} f_n d\mu = \sum_{n\in N} \int_X f_n d\mu.$$

Teorema de convergenţă dominată a lui Lebesgue
Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură şi fie $f_n : X \mapsto \mathbf{C}$ un şir de funcţii măsurabile cu proprietăţile:
i. f_n converge punctual la funcţia f.
ii. există g o funcţie integrabilă astfel încât $|f_n| \leq g$.
Atunci f este funcţie integrabilă şi $\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu$.

13. Spaţii de funcţii p-integrabile
Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură şi fie $1 \leq p < \infty$. Considerăm mulţimea:

$$\mathcal{L}^p(X, \mu) = \{f : X \mapsto \mathbf{C} \mid f \text{ măsurabilă şi } \int_X |f|^p d\mu < \infty\}.$$

Evident, pentru $p = 1$ se obţine mulţimea funcţiilor integrabile.
$\mathcal{L}^p(X, \mu)$ este spaţiu vectorial, iar aplicaţia

$$\| f \|_p = \left(\int_X |f|^p d\mu \right)^{\frac{1}{p}}$$

este o seminormă pe $\mathcal{L}^p(X, \mu)$. Din relaţia $\| f \|_p = 0$ rezultă $f = 0$ (a.p.t.); fie $L^p(X, \mu)$ mulţimea claselor de echivalenţă în raport cu relaţia de egalitate a.p.t. Atunci $(L^p(X, \mu), \| \ \|_p)$ este spaţiu normat (spaţiul funcţiilor p-integrabile). În plus, se demonstrează că $(L^p(X, \mu), \| \ \|_p)$ este spaţiu Banach.

14. Măsura Lebesgue
Fie $\mathbf{R^k}$ spaţiul euclidian k-dimensional şi fie $-\infty \leq a_i \leq b_i \leq \infty$, $\forall\, i = 1, 2..., k$. Un paralelipiped în $\mathbf{R^k}$ este orice mulţime de forma:

$$P = \{(x_1, x_2, ..., x_k) \mid a_i \leq x_i \leq b_i, \, \forall i = 1, 2, ..., k\}.$$

Inegalităţile nestricte pot fi înlocuite şi de inegalităţi stricte. Prin definiţie, mulţimea vidă şi $\mathbf{R^k}$ sunt paralelipipede.

Măsura (Lebesgue) a unui paralelipiped este definită prin:

$$\mu\,(P) = \Pi_{i=1}^{n}(b_i - a_i).$$

În cazurile particulare $k = 1, 2, 3$ se obţin noţiunile uzuale de lungime, arie, volum.

O submulţime $E \subseteq \mathbf{R^k}$ se numeşte elementară dacă există $P_1, P_2, ..., P_n$ paralelipipede astfel încât $E = \bigcup\limits_{i=1}^{n} P_i$.

Notăm cu $\mathcal{E}$ familia mulţimilor elementare din $\mathbf{R^k}$.

Orice mulţime elementară se poate scrie ca reuniune de paralelipipede disjuncte două câte două. Dacă $E = \bigcup\limits_{i=1}^{n} P_i$ este o astfel de descompunere, atunci măsura Lebesgue a lui E este: $\mu(E) = \sum\limits_{i=1}^{n} \mu(P_i)$. Se poate arăta că $\mu(E)$ nu depinde de descompunerea considerată.

15. Proprietăţi

Proprietăţile aplicaţiei μ pe familia mulţimilor elementare sunt:

i. dacă $A, B \in \mathcal{E}$ atunci $A \cup B, A \cap B, A \setminus B$ sunt mulţimi elementare.

ii. dacă $A, B \in \mathcal{E}$ astfel încât $A \cap B = \emptyset$ atunci $\mu(A \cup B) = \mu(A) + \mu(B)$.

iii. pentru orice $A \in \mathcal{E}$ şi $\varepsilon > 0$ există $F, G \in \mathcal{E}$, F închisă şi G deschisă astfel încât:

$$F \subseteq A \subseteq G$$

$$\mu(G) - \varepsilon < \mu(A) < \mu(F) + \varepsilon.$$

Aplicaţia μ se prelungeşte la toate părţile lui R^k; fie $A \subseteq R^k$ şi fie

$$\mu^\star(A) = \inf\{\sum_{n\in N} \mu(A_n) \mid A \subseteq \bigcup_{n\in N} A_n,\ A_n \in \mathcal{E},\ A_n \text{ deschisă } \forall n \in N\}.$$

Aplicaţia $\mu^\star$ se numeşte măsură exterioară; principalele proprietăţi sunt:

i. $\mu^\star(A) \geq 0$, $\forall A \subseteq R^k$.

ii. dacă $A_1 \subseteq A_2$ atunci $\mu^\star(A_1) \leq \mu^\star(A_2)$.

iii. dacă $E \in \mathcal{E}$ atunci $\mu^\star(E) = \mu(E)$.

iv. $\mu^\star\left(\bigcup\limits_{n\in N} A_n\right) \leq \sum\limits_{n\in N} \mu^\star(A_n)$, $\forall A_n \subseteq R^k$.

Se demonstrează că există o σ-algebră de părţi ale lui R^k, notată $\mathcal{M}$ astfel încât restricţia $\mu^\star : \mathcal{M} \mapsto [0, \infty]$ este măsură. Măsura astfel obţinută (notată μ) se numeşte măsura Lebesgue (în R^k), iar elementele lui $\mathcal{M}$ se numesc mulţimi măsurabile Lebesgue.

Principalele proprietăţi ale spaţiului cu măsură $\left(\mathbf{R^k}, \mathcal{M}, \mu\right)$ sunt:

i. $\mathcal{M}$ conţine mulţimile Boreliene.

ii. dacă $A \in \mathcal{M}$ atunci $\mu(A) = \inf\{\mu(D) \mid D$ deschisă şi $D \supseteq A\}$.
iii. dacă $A \in \mathcal{M}$ atunci $\mu(A) = \sup\{\mu(K) \mid K$ compactă şi $K \subseteq A\}$.
iv orice mulţime compactă are măsură Lebesgue finită.
v. dacă $A \in \mathcal{M}$, $\mu(A) = 0$ şi $B \subseteq A$ atunci $B \in \mathcal{M}$ şi $\mu(B) = 0$.
vi. dacă $A \in \mathcal{M}$ atunci pentru orice $x \in R^k$ mulţimea (translatată) $A + x = \{a + x \mid a \in A\}$ este măsurabilă Lebesgue şi $\mu(A + x) = \mu(A)$.

16. Integrala Lebesgue

Dacă f este o funcţie integrabilă în raport cu măsura Lebesgue (în $\mathbf{R^k}$), atunci integrala corespunzătoare (pe o mulţime A) se notează

$$\int_A f(x_1, x_2, ..., x_k) dx_1 dx_2 ... dx_k.$$

În cazurile particulare (uzuale) $k = 1, 2, 3$ se folosesc notaţiile:

$$\int_A f(x)dx, \ \int\int_A f(x, y)dxdy, \ \int\int\int_A f(x, y, z)dxdydz.$$

Legătura cu integrabilitatea în sens Riemann

i. Dacă $f : [a, b] \mapsto \mathbf{R}$ este o funcţie integrabilă Riemann (pe intervalul compact $[a, b]$), atunci f este şi integrabilă în raport cu măsura Lebesgue şi cele două integrale sunt egale.
ii. Dacă $f : [a, b] \mapsto \mathbf{R}$ este o funcţie mărginită atunci ea este integrabilă Riemann dacă şi numai dacă mulţimea punctelor sale de discontinuitate are măsura Lebesgue nulă (se spune că f este continuă a.p.t.).
iii. Există funcţii care sunt integrabile Lebesgue dar nu sunt integrabile Riemann; de exemplu, funcţia lui Dirichlet (pe intervalul $[0, 1]$) nu este integrabilă Riemann dar este integrabilă Lebesgue (integrala sa este 0, pentru că funcţia este nulă a.p.t.).
iv. Dacă $\int_a^b f(x)dx$ este o integrală Riemann improprie absolut convergentă atunci f este integrabilă Lebesgue şi integralele sunt egale.
Există însă integrale Riemann improprii convergente $\int_a^b f(x)dx$ (dar nu absolut convergente) pentru care funcţia f nu este integrabilă Lebesgue; de exemplu $f(x) = \frac{\sin x}{x}$ pe intervalul $(0, \infty)$.

Teorema lui Fubini

În continuare notăm $(x, y) \in R^{k+p}$, măsura Lebesgue în R^k cu dx, măsura Lebesgue în R^p cu dy şi măsura Lebesgue în R^{k+p} cu $dxdy$.
Fie $f : R^{k+p} \mapsto R$ o funcţie integrabilă Lebesgue; atunci:

$$\int_{R^k} \left(\int_{R^p} f(x, y)dy \right) = \int_{R^{k+p}} f(x, y)dxdy = \int_{R^p} \left(\int_{R^k} f(x, y)dx \right) dy.$$

Următoarele cazuri particulare ale rezultatului de mai sus sunt frecvent utilizate în aplicaţii.

17. Exemple
i. Fie $\varphi,\ \phi : [a, b] \mapsto R$ două funcţii continue astfel încât $\varphi \leq \phi$ şi fie mulţimea

$$K = \{(x, y) \in R^2 \mid x \in [a, b],\ \varphi(x) \leq y \leq \phi(x)\}.$$

Dacă $f : K \mapsto R$ este o funcţie continuă, atunci f este integrabilă Lebesgue pe K şi:

$$\int\int_K f(x, y)dxdy = \int_a^b \left(\int_{\varphi(x)}^{\phi(x)} f(x, y)dy \right) dx.$$

În particular, aria mulţimii K este:

$$\mu(K) = \int\int_K dxdy = \int_a^b (\phi(x) - \varphi(x))\, dx.$$

ii. Fie $D \subseteq R^2$ o mulţime compactă, fie $\varphi,\ \phi : D \mapsto R$ două funcţii continue astfel încât $\varphi \leq \phi$ şi fie

$$\Omega = \{(x, y, z) \in R^3 \mid (x, y) \in D,\ \varphi(x, y) \leq z \leq \phi(x, y)\}.$$

Dacă $f : \Omega \mapsto R$ este o funcţie continuă, atunci f este integrabilă Lebesgue pe Ω şi:

$$\int\int\int_\Omega f(x, y, z)dxdydz = \int\int_D \left(\int_{\varphi(x,y)}^{\phi(x,y)} f(x, y, z)dz \right) dxdy.$$

În particular, volumul lui Ω este:

$$\mu(\Omega) = \int\int\int_\Omega dxdydz = \int\int_D (\phi(x, y) - \varphi(x, y))\, dxdy.$$

Formula schimbării de variabile
Fie $A \subseteq R^n$ o mulţime deschisă şi fie $\Lambda : A \mapsto \Lambda(A) \subseteq R^n$ un difeomorfism. Pentru orice funcţie continuă $f : \Lambda(A) \mapsto R$, avem:

$$\int_{\Lambda(A)} f(x)dx = \int_A (f \circ \Lambda)(y)|J_\Lambda(y)|dy,$$

unde J_Λ este iacobianul difeomorfismului Λ.

18. Exemple de spaţii de funcţii integrabile
i. Dacă pe mulţimea numerelor naturale, N, se consideră măsura de numărare, atunci, în acest caz, spaţiul funcţiilor p-integrabile este spaţiul şirurilor p-absolut sumabile:

$$\ell^p(N) = \{x : N \mapsto C \ \mid\ \sum_{n \in N} |x(n)|^p < \infty\}.$$

Analog se defineşte spaţiul $\ell^p(Z)$ al şirurilor bilaterale p-absolut sumabile; evident, norma în aceste cazuri este:

$$\| x \|_p = \left(\sum_{n \in N} |x(n)|^p \right)^{\frac{1}{p}}, \ \forall x \in \ell^p(N).$$

ii. Dacă pe mulţimea numerelor reale, R, se consideră măsura Lebesgue, atunci notăm:

$$L^p(R) = \{ f : R \mapsto C \ | \ f \text{ măsurabilă şi } \int_{-\infty}^{\infty} |f(x)| \, dx < \infty \}.$$

Norma în acest caz este

$$\| f \|_p = \left(\int_{-\infty}^{\infty} |f(x)| \, dx \right)^{\frac{1}{p}}, \ \forall f \in L^p(R).$$

Analog, se pot considera intervale $(a, b) \subset R$ şi se obţin spaţiile $L^p(a, b)$ corespunzătoare.

iii. Un caz particular remarcabil este spaţiul funcţiilor periodice (de perioadă 2π definite pe R) de pătrat integrabil, definit după cum urmează; dacă $\mathcal{P}$ este mulţimea funcţiilor periodice $f : R \mapsto C$, f de perioadă 2π, $f(x) = f(x + 2\pi), \forall x \in R$, atunci:

$$L^2[0, 2\pi] = \{ f \in \mathcal{P} \ | \ f \text{ măsurabilă şi } \int_0^{2\pi} |f(t)| dt < \infty \},$$

integrala fiinds în raport cu măsura Lebesgue. Evident, în locul intervalului $[0, 2\pi]$ se poate lua orice interval de lungime 2π; o altă alegere uzuală este intervalul $[-\pi, \pi]$. Norma pe spaţiul $L^2[0, 2\pi]$ este:

$$\| f \|_2 = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 \, dt}, \ \forall f \in L^2[0, 2\pi].$$

Analog se defineşte spaţiul funcţiilor periodice (de perioadă 2π) integrabile:

$$L^1[0, 2\pi] = \{ f \in \mathcal{P} \ | \ f \text{ măsurabilă şi } \int_0^{2\pi} |f(t)| \, dt < \infty \}.$$

Norma pe spaţiul $L^1[0, 2\pi]$ este:

$$\| f \|_1 = \frac{1}{2\pi} \int_0^{2\pi} |f(t)| \, dt, \ \forall f \in L^1[0, 2\pi].$$

Spaţiul funcţiilor esenţial mărginite

Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură şi fie $f : X \mapsto [0, \infty]$ o funcţie măsurabilă; considerăm mulţimea

$$M = \{ t \in R \mid \mu \left(f^{-1}(t, \infty] \right) = 0 \}.$$

Prin definiţie, supremumul esenţial al lui f este:

$$\text{esssup} f = \infty \text{ dacă } M = \emptyset \text{ şi } \text{esssup} f = \inf M \text{ dacă } M \neq \emptyset.$$

Notăm $\| f \|_\infty = \text{esssup}|f|$; funcţia f se numeşte esenţial mărginită dacă $\| f \|_\infty < \infty$. Mulţimea funcţiilor esenţial mărginite se notează $L^\infty(X, \mu)$ şi este spaţiu Banach cu norma $\| \ \|_\infty$.
Şi în acest caz putem particulariza spaţiul cu măsură $(X, \mathcal{A}, \mu)$, obţinându-se (ca mai sus) spaţiile $(\ell^\infty(N), \| \ \|_\infty)$, $(L^\infty(R), \| \ \|_\infty)$, $(L^\infty[0, 2\pi], \| \ \|_\infty)$.

5.3 MF.05.3. Funcţii convexe, inegalităţi

19. Definiţie
O funcţie $\phi : (a, b) \mapsto R$ se numeşte convexă dacă:

$$\phi((1-\lambda)x + \lambda y) \leq (1-\lambda)\phi(x) + \lambda\phi(y), \forall x, y \in (a, b), \ \forall \lambda \in [0, 1],$$

sau, echivalent:

$$\frac{\phi(t) - \phi(s)}{t - s} \leq \frac{\phi(u) - \phi(t)}{u - t}, \ \forall a < s < t < u < b.$$

Se demonstrează fără dificultate că o funcţie convexă pe un interval deschis este continuă. Un exemplu remarcabil de funcţie convexă pe R este funcţia exponenţială, $\phi(x) = e^x$.

20. Inegalitatea lui Jensen Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură astfel încât $\mu(X) = 1$ şi fie $a, b \in \overline{R}, \ a < b$. Atunci, pentru orice funcţie integrabilă $f : X \mapsto (a, b)$ şi pentru orice funcţie convexă $\phi : (a, b) \mapsto R$, are loc inegalitatea:

$$\phi\left(\int_X f d\mu\right) \leq \int_X (\phi \circ f) d\mu.$$

Soluţie
Fie $t = \displaystyle\int_X f d\mu \in R$ şi fie $\alpha = \displaystyle\sup_{s \in [a,t]} \frac{\phi(t) - \phi(s)}{t - s}$. Propunem ca exerciţiu inegalitatea:

$$\phi(s) \geq \phi(t) + \alpha(s - t), \forall s \in (a, b).$$

În particular, pentru $s = f(x)$, obţinem:

$$\phi(f(x)) \geq \phi(t) + \alpha(f(x) - t), \forall x \in X.$$

Integrând ultima inegalitate pe X în raport cu măsura μ, obţinem:

$$\int_X (\phi \circ f) d\mu \geq \int_X \phi(t) d\mu + \alpha\left(\int_X f d\mu - t\right),$$

adică:

$$\int_X (\phi \circ f) d\mu \geq \phi \left(\int_X f d\mu) \right).$$

21. Inegalitatea lui Holder

Fie $p > 0$ şi $q > 0$ astfel încât $\dfrac{1}{p} + \dfrac{1}{q} = 1$; p şi q se numesc în acest caz conjugate.

a. Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură. Atunci, pentru orice funcţii măsurabile $f, g : X \mapsto [0, \infty)$, are loc inegalitatea lui Holder:

$$\int_X fg \, d\mu \leq \left(\int_X f^p \, d\mu \right)^{\frac{1}{p}} \left(\int_X g^q \, d\mu \right)^{\frac{1}{q}}.$$

b. Fie $a_1, a_2, ..., a_n$ şi $b_1, b_2, ..., b_n$ numere reale pozitive; atunci:

$$a_1 b_1 + a_2 b_2 + ... + a_n b_n \leq (a_1^p + a_2^p + ... + a_n^p)^{\frac{1}{p}} \cdot (b_1^q + b_2^q + ... + b_n^q)^{\frac{1}{q}}.$$

Soluţie

a. Fie $A = \left(\int_X f^p \, d\mu \right)^{\frac{1}{p}}$ şi $B = \left(\int_X g^q \, d\mu \right)^{\frac{1}{q}}$. Dacă $A = \infty$ sau $B = \infty$, atunci inegalitatea este evidentă. Dacă $A = 0$ sau $B = 0$, atunci, conform exerciţiului 3, $f = 0 (a.p.t.)$ sau $g = 0 \, (a.p.t.)$ şi deci $fg = 0 \, (a.p.t.)$ şi inegalitatea este iarăşi evidentă. Presupunem acum $A, B \in (0, \infty)$; fie $F = \frac{f}{A}$ şi $G = \frac{g}{B}$. Fie $x \in X$; deoarece $F(x) > 0$ şi $G(x) > 0$, există $s, t \in R$ astfel încât $F(x) = e^{\frac{s}{p}}$ şi $G(x) = e^{\frac{t}{q}}$. Folosind convexitatea funcţiei exponenţiale, avem:

$$F(x)G(x) = e^{\frac{1}{p} s + \frac{1}{q} t} \leq \frac{1}{p} e^s + \frac{1}{q} e^t = \frac{1}{p} F^p(x) + \frac{1}{q} G^q(x).$$

Integrând ultima inegalitate, obţinem:

$$\int_X FG \, d\mu \leq \frac{1}{p} \int_X F^p \, d\mu + \frac{1}{q} \int_X G^q \, d\mu = \frac{1}{p} + \frac{1}{q} = 1,$$

ceea ce încheie demonstraţia.

b. Se aplică inegalitatea lui Holder pentru un spaţiu de probabilitate.

22. Inegalitatea lui Minkovski

Fie $p > 1$ şi fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură; pentru orice funcţii măsurabile $f, g : X \mapsto [0, \infty)$, are loc inegalitatea:

$$\left(\int_X (f+g)^p d\mu \right)^{\frac{1}{p}} \leq \left(\int_X f^p d\mu \right)^{\frac{1}{p}} + \left(\int_X g^p d\mu \right)^{\frac{1}{p}}.$$

Cazul particular $p = 2$ este cunoscut sub numele de inegalitatea lui Schwarz.
Soluţie
Integrând egalitatea:

$$(f+g)^p = f(f+g)^{p-1} + g(f+g)^{p-1},$$

obţinem:

$$\int_X (f+g)^p = \int_X f(f+g)^{p-1} d\mu + \int_X g(f+g)^{p-1}\, d\mu.$$

Aplicând inegalitatea lui Holder celor doi termeni din membrul drept al egalităţii de mai sus, obţinem (alegem q conjugat cu p):

$$\int_X f(f+g)^{p-1} d\mu \leq \left(\int_X f^p d\mu\right)^{\frac{1}{p}} \left(\int_X (f+g)^{(p-1)q} d\mu\right)^{\frac{1}{q}}, \text{şi}$$

$$\int_X g(f+g)^{p-1} d\mu \leq \left(\int_X g^p d\mu\right)^{\frac{1}{p}} \left(\int_X (f+g)^{(p-1)q} d\mu\right)^{\frac{1}{q}}.$$

Însumând cele două inegalităţi şi folosind egalitatea $(p-1)q = p$, obţinem inegalitatea:

$$\int_X (f+g)^p d\mu \leq \left(\int_X (f+g)^p d\mu\right)^{\frac{1}{q}} \left(\left(\int_X f^p d\mu\right)^{\frac{1}{p}} + \left(\int_X g^p d\mu\right)^{\frac{1}{p}}\right) \quad (*).$$

Folosind convexitatea funcţiei putere cu exponent supraunitar, $\phi(t) = t^p$, obţinem:

$$\left(\frac{f+g}{2}\right)^p \leq \frac{1}{2}(f^p + g^p),$$

ceea ce arată că dacă membrul stâng al inegalităţii $(*)$ este ∞, atunci şi membrul drept este ∞. Putem deci presupune că $\int_X (f+g)^p d\mu < \infty$. Demonstraţia se încheie împărţind inegalitatea $(*)$ cu $\left(\int_X (f+g)^p d\mu\right)^{\frac{1}{q}}$.

5.4 MF.05.4. Serii trigonometrice

23. Definiţie
Un caz particular remarcabil de serie Fourier este **seria trigonometrică**. Considerăm spaţiul Hilbert (a se vedea cap.4) al funcţiilor periodice (de perioadă 2π) de pătrat integrabil:

$$L^2[0,2\pi] = \{f : [0,2\pi] \mapsto C \mid f \text{ măsurabilă şi } \int_0^{2\pi} |f(t)|^2 dt < \infty\}.$$

Produsul scalar este

$$< f, g >= \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}dt,$$

iar norma $\| f \|_2 = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2 dt}.$

24. Sistemul trigonometric

Pentru orice $n \in Z$, fie $\omega_n(t) = e^{int}$. Un rezultat clasic de analiză afirmă că mulţimea (sistemul trigonometric) $\mathcal{B} = \{\omega_n \mid n \in Z\}$ este bază ortonormală în $L^2[0, 2\pi]$. Pentru orice funcţie $f \in L^2[0, 2\pi]$, coeficienţii Fourier (în raport cu baza fixată mai sus), sunt

$$\widehat{f_n} =< f, \omega_n >= \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}dt, \forall n \in Z,$$

iar seria Fourier (sau seria trigonometrică) asociată funcţiei f este $\sum_{n\in Z} \widehat{f_n}\omega_n$; sumele parţiale ale seriei, $P_n = \sum_{k=-n}^{n} \widehat{f_k}\omega_k$, se numesc polinoame trigonometrice şi $\lim_{n\to\infty} P_n = f$ în spaţiul $L^2[0, 2\pi]$, sau, echivalent:

$$\lim_{n\to\infty} \| P_n - f \|_2 = 0.$$

Identitatea lui Parseval devine în acest caz:

$$\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2\, dt =\| f \|_2^2 = \sum_{n\in Z} |\widehat{f_n}|^2.$$

Folosind egalitatea $e^{int} = \cos nt + i \sin nt, \forall t \in R$, seria Fourier asociată funcţiei f se poate scrie sub forma:

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt),$$

unde coeficienţii trigonometrici (clasici) a_n şi b_n sunt:

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos ntdt,\ \forall\, n \geq 0,$$

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin ntdt,\ \forall\, n \geq 1.$$

Legătura dintre coeficienţii $\widehat{f_n}, a_n$ şi b_n este:

$$\widehat{f_0} = \frac{a_0}{2},\ \widehat{f_n} = \frac{a_n - ib_n}{2},\ \widehat{f_{-n}} = \frac{a_n + ib_n}{2},\ \forall\, n = 1, 2, ...$$

Lema lui Riemann afirmă că dacă funcţia f este integrabilă, atunci:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0.$$

În legătură cu convergenţa punctuală a seriei Fourier, are loc următorul rezultat clasic:

25. Teorema lui Dirichlet
Dacă $f : R \mapsto R$ este o funcţie periodică de perioadă 2π, măsurabilă, mărginită, având cel mult un număr finit de discontinuităţi de speţa intâi şi având derivate laterale în orice punct, atunci seria Fourier asociată funcţiei f converge în fiecare punct $x \in R$ la

$$\frac{1}{2}(f(x+0) + f(x-0)).$$

În particular, dacă funcţia f este continuă (şi verifică celelalte ipoteze din teorema lui Dirichlet), atunci are loc descompunerea:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt).$$

26. Convergenţa uniformă a seriei Fourier
Condiţii suficiente pentru convergenţa uniformă a seriei Fourier sunt date în teorema următoare:

Dacă $f : R \mapsto C$ este o funcţie continuă, de clasă $\mathcal{C}^1$ pe porţiuni şi periodică de perioadă 2π, atunci seria sa Fourier este absolut şi uniform convergentă, iar suma este f.

27. Definiţii
Numărul $\frac{a_0}{2} = \frac{1}{2\pi}\int_0^{2\pi} f(x)dx$ este media semnalului f, primul termen

$$a_1 \cos x + b_1 \sin x$$

este oscilaţia principală (în jurul valorii medii), iar termenul

$$a_n \cos nt + b_n \sin nt,\ n \geq 2$$

este armonica de ordinul n a funcţiei f. Perioada armonicei de ordinul n este $\frac{2\pi}{n}$, iar amplitudinea $A_n = \sqrt{|a_n|^2 + |b_n|^2}$; conform lemei lui Riemann rezultă $\lim\limits_{n\to\infty} A_n = 0$.

În cazul în care funcţia f are perioada $T = 2\ell$, $(\ell > 0)$, atunci toate rezultatele de mai sus sunt în continuare adevărate, cu adaptările corespunzătoare; baza ortonormală este

$$\{\epsilon_n \mid n \in Z\},\ \text{cu}\ \epsilon_n(x) = e^{i\frac{n\pi x}{\ell}},$$

iar coeficienţii Fourier sunt:

$$\widehat{f}_n = \frac{1}{2\ell}\int_0^{2\ell} f(x)e^{-i\frac{n\pi x}{\ell}}dx,\ \forall\, n \in Z,$$

$$a_n = \frac{1}{\ell}\int_0^{2\ell} f(x)\cos\frac{n\pi x}{\ell}dx,\ \forall\, n = 0,1,2,...,$$

$$b_n = \frac{1}{\ell}\int_0^{2\ell} f(x)\sin\frac{n\pi x}{\ell},\ \forall\, n = 1,2,...$$

Teorema lui Dirichlet se scrie:

$$\frac{1}{2}(f(x+0)+f(x-0)) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi x}{\ell} + b_n\sin\frac{n\pi x}{\ell}\right) =$$

$$= \sum_{n=-\infty}^{\infty}\widehat{f}_n e^{i\frac{n\pi x}{\ell}},\ \forall\, x \in R.$$

Identitatea lui Parseval devine acest caz:

$$\frac{|a_0|^2}{2} + \sum_{n\geq 1}\left(|a_n|^2 + |b_n|^2\right) = \frac{1}{\ell}\int_0^{2\ell} |f(t)|^2\, dt.$$

Evident, toate rezultatele de mai sus rămân adevărate dacă înlocuim intervalul $[0, 2\ell]$ cu orice alt interval de lungime 2ℓ, de exemplu, $[-\ell, \ell]$.

28. Serii de sinusuri şi cosinusuri

Fie $f : [0,\ell] \mapsto R$, o funcţie integrabilă şi fie $\tilde{f} : R \mapsto R$, periodică de perioadă 2ℓ, definită prin:

$$\tilde{f}(x) = \begin{cases} f(x) & ,\quad x \in [0,\ell] \\ f(-x) & ,\quad x \in (-\ell, 0) \end{cases}$$

Dacă funcţia $\tilde{f}$ satisface condiţiile teoremei lui Dirichlet, atunci, dezvoltând $\tilde{f}$ în serie Fourier, rezultă:

$$\frac{1}{2}(f(x+0)+f(x-0)) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos\frac{\pi n x}{\ell},\ \forall\, x \in (0,\ell),$$

$$f(0+0) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n,\ f(\ell-0) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(-1)^n a_n,$$

coeficienţii a_n fiind coeficienţii Fourier reali asociaţi funcţiei $\tilde{f}$.
Formula de mai sus se numeşte dezvoltarea în serie de cosinusuri a lui f.
Analog, dacă funcţia (impară):

$$\tilde{f}(x) = \begin{cases} f(x) & ,\quad x \in [0,\ell] \\ -f(-x) & ,\quad x \in (-\ell, 0) \end{cases}$$

satisface condiţiile teoremei lui Dirichlet, atunci dezvoltarea în serie de sinusuri a funcţiei f este:

$$\frac{1}{2}(f(x+0)+f(x-0)) = \sum_{n=1}^{\infty} b_n \sin\frac{\pi n x}{\ell},\ \forall\, x \in (0,\ell),$$

coeficienţii b_n fiind coeficienţii Fourier reali asociaţi funcţiei $\tilde{f}$.

29. Fenomenul Gibbs
În jurul unui punct de discontinuitate al unei funcţii date, seria Fourier asociată ei converge doar punctual (nu neapărat uniform). Acest fapt conduce la un defect de convergenţă (aparent paradox) al şirului sumelor parţiale asociat seriei trigonometrice date, numit fenomenul Gibbs. Dăm în continuare un exemplu în acest sens.

Exemplu
Considerăm restricţia funcţiei signum la intervalul $(-\pi,\pi)$,

$$\mathrm{sgn} : (-\pi,\pi) \mapsto R,\ \mathrm{sgn}(x) = \begin{cases} -1 & , \quad x \in (-\pi,0) \\ 0 & , \quad x = 0 \\ 1 & , \quad x \in (0,\pi) \end{cases}$$

Vom folosi egalitatea (pe care o propunem ca exerciţiu):

$$\mathrm{sgn}(x) = \frac{4}{\pi}\sum_{n\geq 1}\frac{\sin(2n-1)x}{2n-1},\ \forall\, x \in (-\pi,\pi).$$

Notăm cu S_n şirul sumelor parţiale:

$$S_n(x) = \frac{4}{\pi}\sum_{k=1}^{n}\frac{\sin(2k-1)x}{2k-1},\ \forall\, x \in (-\pi,\pi).$$

În punctul $x = 0$ funcţia sgn nu este continuă; seria sa Fourier converge (conform teoremei lui Dirichlet) la $\frac{1}{2}(-1+1) = 0 = \mathrm{sgn}(0)$; convergenţa $\lim_{n\to\infty} S_n(x) = \mathrm{sgn}(x),\ \forall\, x \in (-\pi,\pi)$ este punctuală, nu şi uniformă.

Exemplul se bazează pe următoarele trei afirmaţii:
a. Are loc egalitatea:

$$S_n(x) = \frac{2}{\pi}\int_0^x \frac{\sin 2nt}{\sin t}\,dt,\ \forall\, x \in (-\pi,\pi).$$

b. Funcţia S_n are un maxim în punctul $x = \frac{\pi}{2n}$ şi:

$$\lim_{n\to\infty} S_n\left(\frac{\pi}{2n}\right) = \frac{2}{\pi}\int_0^{\pi}\frac{\sin t}{t}\,dt \approx 1,1789.$$

c.

$$\lim_{n\to\infty}\left|S_n\left(\frac{\pi}{2n}\right)-\operatorname{sgn}(0+)\right|\approx 0,1789.$$

Demonstraţie

a. Calculăm mai întâi suma

$$A=\cos x+\cos 3x+...+\cos(2n-1)x,\ \forall\, x\neq k\pi,\ k\in Z.$$

Pentru aceasta, considerăm şi suma $B=\sin x+\sin 3x+...+\sin(2n-1)x$ şi calculăm:

$$A+iB=$$

$$=(\cos x+i\sin x)+(\cos 3x+i\sin 3x)+...+(\cos(2n-1)x+i\sin(2n-1)x)=$$

$$=z^2\frac{z^{2n}-1}{z^2-1},$$

unde am notat $z=\cos x+i\sin x$. După calcule, rezultă:

$$A+iB=\frac{\sin nx}{\sin x}(\cos nx+i\sin nx),$$

şi deci:

$$\cos x+\cos 3x+...+\cos(2n-1)x=\frac{\sin 2nx}{2\sin x},\ \forall\, x\neq k\pi,\ k\in Z.$$

Integrând de la 0 la x, rezultă:

$$\sum_{k=1}^{n}\frac{\sin(2k-1)x}{2k-1}=\int_0^x\frac{\sin 2nt}{2\sin t}\,dt,$$

sau, înmulţind cu $\frac{4}{\pi}$:

$$S_n(x)=\frac{2}{\pi}\int_0^x\frac{\sin 2nt}{\sin t}\,dt,\ \forall\, x\in(-\pi,\pi).$$

b. Din cele demonstrate la punctul precedent rezultă că

$$S_n'(x)=\frac{2\sin 2nx}{\pi\sin x}$$

şi deci $\frac{\pi}{2n}$ este punct critic al lui S_n; într-o vecinătate a lui $\frac{\pi}{2n}$ avem:

$$S_n'(x)=\frac{2\sin 2nx}{\pi\sin x}>0,\ x<\frac{\pi}{2n},$$

$$S_n'(x)=\frac{2\sin 2nx}{\pi\sin x}<0,\ x>\frac{\pi}{2n}.$$

Rezultă că $x = \frac{\pi}{2n}$ este punct de maxim al funcţiei S_n.
Calculăm acum:

$$S_n\left(\frac{\pi}{2n}\right) = \frac{2}{\pi}\int_0^{\frac{\pi}{2n}} \frac{\sin 2nt}{\sin t}\,dt = \frac{2}{\pi}\int_0^{\pi} \frac{\sin u}{\sin\left(\frac{u}{2n}\right)}\,\frac{du}{2n} = \frac{1}{n}\int_0^{\pi} \frac{\sin u}{\sin\left(\frac{u}{2n}\right)}\,du.$$

Rezultă:

$$\lim_{n\to\infty} S_n\left(\frac{\pi}{2n}\right) = \frac{2}{\pi}\int_0^{\pi} \frac{\sin u}{u}\,du.$$

Ultima integrală se aproximează dezvoltând funcţia sinus în serie de puteri:

$$\int_0^{\pi} \frac{\sin u}{u}\,du = \int_0^{\pi} \left(\sum_{n\geq 1} \frac{(-1)^n}{(2n-1)!}\,x^{2n-2}\right) du =$$

$$= \sum_{n\geq 1} \frac{(-1)^n}{(2n-1)!(2n-1)} x^{2n-1}\Big|_0^{\pi} = \sum_{n\geq 1} \frac{(-1)^n \pi^{2n-1}}{(2n-1)!(2n-1)}.$$

Seria fiind alternată, eroarea este mai mică decât primul termen neglijat. Cu o eroare mai mică decât 10^{-3}, se obţine

$$\lim_{n\to\infty} S_n\left(\frac{\pi}{2n}\right) \approx 1,1789.$$

c. Rezultă: $\lim\limits_{n\to\infty} \left|S_n\left(\frac{\pi}{2n}\right) - \operatorname{sgn}(0+)\right| \approx 0,1789.$

Capitolul 6

MF.06. Operatori pe spaţii Hilbert

Cuvinte cheie

operator, spectru, rază spectrală, spectru punctual, spectru, operator normal, operator autoadjunct, operator unitar, proiector, operator pozitiv, operator diagonal, operator de translaţie, operator de multiplicare, operator de convoluţie, operator integral.

În acest capitol vom studia operatorii liniari şi continui definiţi pe un spaţiu Hilbert infinit dimensional. Aşa cum am văzut în capitolul 2, dacă X este un spaţiu Banach, atunci $\mathcal{L}(X)$ este un spaţiu Banach. Dacă $X = H$ este spaţiu Hilbert, atunci pentru un operator din $\mathcal{L}(H)$ se poate defini adjunctul său, ceea ce permite un studiu mai aprofundat al unor clase speciale de operatori (normali, autoadjuncţi, unitari, pozitivi, proiectori). Reamintim că în cazul finit dimensional, $H = \mathbf{C^n}$, acest studiu a fost făcut în cap.3. De altfel ideea principală care a stat la baza expunerii ce urmează este de a pune în evidenţă felul în care dimensiunea infinită a lui H modifică (sau nu) rezultatele importante din capitolul 3. Sursele bibliografice sunt: [C02], [D02], [H01], [F01], [O01], [S02].

6.1 MF.06.1. Adjunctul unui operator

În continuare H va fi un spaţiu Hilbert (complex) infinit dimensional şi separabil, iar $\mathcal{L}(H)$ algebra Banach a operatorilor liniari şi continui pe spaţiul H. În capitolul 3 am asociat oricărui operator $T \in \mathcal{L}(\mathbf{C^n})$ un operator $T^\star$ numit adjunctul lui T, urmând ca existenţa şi proprietăţile sale să fie demonstrate în contextul unui spaţiu Hilbert infinit dimensional.

1.Propoziţie
Pentru orice operator $T \in \mathcal{L}(H)$, există şi este unic un operator $T^\star \in \mathcal{L}(H)$ cu proprietăţile:
(i) $< Tx, y >=< x, T^\star y >$, $\forall x, y \in H$.
(ii) $(T^\star)^\star = T$.
(iii) $\| T \| = \| T^\star \|$.
(iv) $\| T \|^2 = \| T^\star T \|$.
Operatorul $T^\star$ se numeşte **adjunctul** lui T; facem precizarea că pentru unicitatea sa este suficientă egalitatea **(i)**.
Demonstraţie În continuare vom folosi deseori implicaţia:
dacă $< u, v >=< u, w >, \forall u \in H$, atunci $v = w$. Fie $T \in \mathcal{L}(H)$ şi fie $y \in H$ un vector arbitrar fixat. Aplicaţia $f : H \to C$, $f(x) =< Tx, y >$ este o funcţională liniară şi continuă pe H (liniaritatea este evidentă iar continuitatea rezultă din inegalitatea: $| < Tx, y > | \leq \| T \| \| x \| \| y \|$). Conform teoremei lui Riesz de reprezentare a funcţionalelor liniare şi continue pe un spaţiu Hilbert (cap.4), există un unic vector $z \in H$ astfel încât $f(x) =< x, z >$; evident, z depinde de T şi y. Fie operatorul $T^\star : H \to H$, definit prin $T^\star y = z$. Este evident că $T^\star$ este liniar şi că verifică relaţia **(i)**. Continuitatea lui $T^\star$ rezultă din inegalitatea :

$$\| T^\star y \|^2 =< T^\star y, T^\star y >=< TT^\star y, y > \leq \| T \| \| T^\star y \| \| y \|, \forall y \in H.$$

Am demonstrat deci că $T^\star \in \mathcal{L}(H)$ şi $\| T^\star y \| \leq \| T \| \| y \|$, ceea ce arată că $\| T^\star \| \leq \| T \|$. Pentru a demonstra că $T^\star$ este unic cu proprietatea **(i)**, să presupunem prin absurd că există $S \in \mathcal{L}(H)$, $S \neq T^\star$, astfel încât $< Tx, y >=< x, Sy >$. Rezultă aşadar că $< x, Sy >=< x, T^\star y >$ pentru orice $x, y \in H$, deci $S = T^\star$, contradicţie cu ipoteza.
(ii) Pentru orice $x, y \in H$, avem:

$$< x, (T^\star)^\star y >=< T^\star x, y >= \overline{< y, T^\star x >} = \overline{< Ty, x >} =< x, Ty >,$$

şi deci $(T^\star)^\star = T$.
(iii) Din inegalitatea $\| T^\star \| \leq \| T \|$ (care a fost deja demonstrată), şi din **(ii)** rezultă imediat $\| T^\star \| = \| T \|$.
(iv) Din **(iii)**, rezultă inegalitatea $\| T^\star T \| \leq \| T^\star \| \| T \| = \| T \|^2$. Pentru a demonstra cealaltă inegalitate, fie $x \in H$; avem:

$$\| Tx \|^2 =< Tx, Tx >=< T^\star Tx, x > \leq \| T^\star Tx \| \| x \| \leq \| T^\star T \| \| x \|^2,$$

şi deci $\| T \|^2 \leq \| T^\star T \|$.

2.Propoziţie
Dacă $T, S \in \mathcal{L}(H)$, atunci:
(i) $(\alpha T + \beta S)^\star = \overline{\alpha} T^\star + \overline{\beta} S^\star$, $\forall \alpha\, \beta \in C$.
(ii) $(TS)^\star = S^\star T^\star$.

(iii) Dacă T este inversabil, atunci $T^\star$ este inversabil şi

$$(T^{-1})^\star = (T^\star)^{-1}.$$

Demonstraţie Relaţia **(i)** o propunem ca exerciţiu.
(ii) Pentru orice $x, y \in H$, avem:

$$< TSx, y >=< Sx, T^\star y >=< x, S^\star T^\star y >,$$

şi deci $(TS)^\star = S^\star T^\star$.
(iii) Dacă T este operator inversabil, atunci, folosind relaţia **(ii)** pentru operatorii $T^\star$ şi T^{-1}, obţinem:

$$T^\star(T^{-1})^\star = (T^{-1}T)^\star = I^\star = I = (TT^{-1})^\star = (T^{-1})^\star T^\star,$$

ceea ce încheie demonstraţia.

3. Definiţii
Un operator $T \in \mathcal{L}(H)$ se numeşte **autoadjunct** dacă $T = T^\star$, **normal** dacă $TT^\star = T^\star T$ şi **unitar** dacă $TT^\star = T^\star T = I$.
Spectrul unui
operator $T \in \mathcal{L}(H)$ este, prin definiţie,

$$\sigma(T) = \{\lambda \in \mathbf{C}, ; \lambda I - T \,\text{nu este inversabil}\}.$$

Se poate demonstra că spectrul oricărui operator este mulţime nevidă şi compactă în $\mathbf{C}$. Din teorema lui Banach (capitolul 2, teorema 17), rezultă că

$$\sigma(T) = \{\lambda \in ; \lambda I - T \,\text{nu este bijectiv}\}.$$

În cazul în care spaţiul Hilbert H este finit dimensional, $H = \mathbf{C^n}$, am văzut în capitolul 2 că

$$\sigma(T) = \{\lambda \in \mathbf{C}\,; \exists x \neq 0 \,\text{astfel încât}\, Tx = \lambda x\}.$$

Mulţimea $\sigma(T)$ este în acest caz finită şi elementele ei se numesc
valori proprii, iar vectorii $x \neq 0$ care verifică relaţia $Tx = \lambda x$ se numesc **vectori proprii** asociaţi lui λ. În general, pe un spaţiu Hilbert infinit dimensional, mulţimea valorilor proprii (numită şi **spectrul**
punctual şi notată $\sigma_p(T)$) este o submulţime strictă a spectrului, deoarece în acest caz există operatori care sunt injectivi dar nu sunt surjectivi. Aşa cum vom vedea în exemplele din paragraful următor, există operatori ale căror spectre punctuale sunt vide (dar, evident, spectrele lor sunt nevide). O altă submulţime remarcabilă a spectrului unui operator este **spectrul punctual aproximativ**, notat $\sigma_{pa}(T)$ şi definit prin:

$$\sigma_{pa}(T) = \{\lambda \in \mathbf{C}\,; \exists x_n \in H \text{astfel încât} \parallel x_n \parallel= 1 \text{şi} \lim_{n\to\infty}(\lambda I - T)x_n = 0\}.$$

Este evident că $\sigma_p(T) \subseteq \sigma_{pa}(T)$, deoarece dacă λ este o valoare proprie, atunci putem lua şirul constant $x_n = x$, unde vectorul x este un vector propriu de normă 1 asociat valorii proprii λ. Incluziunea $\sigma_{pa}(T) \subseteq$
$\subseteq \sigma(T)$ se demonstrează prin reducere la absurd : dacă $\lambda \in \sigma_{pa}(T)$ dar $\lambda \notin \sigma(T)$, atunci există operatorul $(\lambda I - T)^{-1} \in \mathcal{L}(H)$ şi aplicându-l relaţiei

$$\lim_{n\to\infty} (\lambda I - T)x_n = 0,$$

obţinem $\lim\limits_{n\to\infty} x_n = 0$, contradicţie cu $\| x_n \| = 1$, $\forall n \in N$. Se poate demonstra următorul rezultat ([10],p.40): spectrul unui operator normal este egal cu spectrul punctual aproximativ.
Exemple ilustrative pentru noţiunile introduse aici vor fi date în paragraful următor.
Ca şi în cazul $H = \mathbf{C^n}$, operatorii autoadjuncţi au spectrul real, iar cei unitari au spectrul inclus în cercul unitate. În plus, dacă T este un operator inversabil, atunci

$$\sigma(T^{-1}) = \{\lambda^{-1}\,;\, \lambda \in \sigma(T)\}.$$

4.Definiţie
Fie $T \in \mathcal{L}(H)$. Un subspaţiu $K \subseteq H$ se numeşte **subspaţiu invariant** pentru T dacă $T(K) \subseteq K$; subspaţiul K se numeşte **subspaţiu reducător** pentru T dacă $T(K) \subseteq K$ şi $T(K^\perp) \subseteq K^\perp$.

5.Propoziţie
Fie $T \in \mathcal{L}(H)$ şi K un subspaţiu în H; următoarele afirmaţii sunt echivalente:
(a) K este subspaţiu invariant pentru T.
(b) $K^\perp$ este subspaţiu invariant pentru $T^\star$.
Demonstraţia o repetă pe cea din cazul $H = \mathbf{C^n}$, (lema 27,cap2).
Evident, de aici rezultă că pentru un operator autoadjunct un subspaţiu invariant este şi subspaţiu reducător.
Reamintim că am notat cu $\mathrm{Ker}(T)$ şi $\mathrm{Im}(T)$ **nucleul** şi respectiv **imaginea** operatorului T. Este evident că subspaţiile $\mathrm{Ker}(T)$ şi $\mathrm{Im}(T)$ sunt invariante pentru T. In continuare demonstrăm o relaţie între nucleul unui operator şi imaginea adjunctului său.

6.Propoziţie
Pentru orice operator $T \in \mathcal{L}(H)$, avem:
(i) $\mathrm{Ker}(T) = (\mathrm{Im}(T^\star))^\perp$.
(ii) $\mathrm{Ker}(T^\star) = (\mathrm{Im}(T))^\perp$.
Demonstraţie (i) Fie $x \in \mathrm{Ker}(T)$ şi fie $y \in H$; atunci
$< T^\star y, x >=< y, Tx >= 0$ şi deci $x \perp \mathrm{Im}(T^\star)$. Pentru a demonstra incluziunea inversă fie $x \in (\mathrm{Im}(T^\star))^\perp$; atunci, pentru orice $y \in H$, avem $< Tx, y >=< x, T^\star y >= 0$ şi deci $Tx = 0$, adică $x \in \mathrm{Ker}(T)$. Egalitatea

(ii) este o consecinţă a egalităţilor **(i)** şi $(T^\star)^\star = T$.

7.Definiţie
Un operator $T \in \mathcal{L}(H)$ se numeşte **pozitiv** dacă $< Tx, x > \geq 0$, $\forall x \in$
$\in H$. Să mai observăm că dacă $T \in \mathcal{L}(H)$ este un operator arbitrar, atunci operatorul $T^\star T$ este pozitiv:

$$< T^\star Tx, x > = < Tx, Tx > = \| Tx \|^2 \geq 0 \,,\ \forall x \in H.$$

8.Propoziţie
Fie $T \in \mathcal{L}(H)$.
(i) Dacă $< Tx, x > = 0$, $\forall x \in H$, atunci $T = O$.
(ii) T este autoadjunct dacă şi numai dacă $< Tx, x > \in R$, $\forall x \in H$.
(iii) Dacă T este pozitiv, atunci $\sigma(T) \subset [0, \infty)$.
(iv) Dacă T este unitar, atunci $\sigma(T) \subseteq \{\lambda\,;\ |\lambda| = 1\}$.
(v) Dacă T este inversabil atunci $\sigma(T^{-1}) = \{\lambda^{-1}\,;\ \lambda \in \sigma(T)\}$.
Demonstraţie Vom demonstra (i) şi (ii).
(i) Pentru orice $x, y \in H$, din identitatea

$$< T(x + y), x + y > = 0,$$

rezultă :
(a) $< Tx, y > + < Ty, x > = 0$.
Înlocuind pe y cu iy în ultima egalitate şi înmulţind apoi cu i , obţinem:
(b) $< Tx, y > - < Ty, x > = 0$.
Adunând relaţiile **(a)** şi **(b)** obţinem $< Tx, y > = 0$, ceea ce încheie demonstraţia.
(ii) Dacă T este autoadjunct, atunci pentru orice $x \in H$, avem:

$$< Tx, x > = < x, T^\star x > = < x, Tx > = \overline{< Tx, x >},$$

şi deci $< Tx, x > \in R$. Reciproc, dacă $< Tx, x > \in R$, $\forall x \in H$, atunci, calculând $< Tx, y >$ cu relaţia de polarizare (cap.4), obţinem $< Tx, y > = \overline{< Ty, x >}$ şi deci $T = T^\star$. În particular, de aici rezultă că orice operator pozitiv este autoadjunct.

6.2 MF.06.2. Proiectori

O clasă importantă de operatori pozitivi, este, ca şi în cazul finit dimensional, clasa proiectorilor.

9. Definiţie
Un operator $P \in \mathcal{L}(H)$ se numeşte **proiector** (sau **proiecţie**) dacă $P^2 = P = P^\star$. Proiectorii sunt operatori pozitivi:

$$< Px, x >=< P^2x, x >=< Px, Px >\geq 0.$$

În cele ce urmează vom folosi în mod esenţial descompunerea ortogonală a lui H după un subspaţiu închis (cap. 4). Fie K un subspaţiu închis în H şi fie $P_K : H \to H$, $P_K x = y$, unde, $x = y + z$ cu $y \in K$ şi $z \in K^\perp$. În teorema următoare vom demonstra că P_K este proiecţie şi că orice proiecţie se poate construi în acest fel.

10. Teoremă
(i) Fie $K \subseteq H$ un subspaţiu închis; atunci operatorul P_K definit mai sus este un proiector.
(ii) Reciproc, dacă P este un proiector, atunci există un subspaţiu închis $K \subseteq H$ (şi anume $K = \mathrm{Im}(P)$) astfel încât $P = P_K$.
Demonstraţie (i) Fie $x_\iota = y_\iota + z_\iota$, $\iota \in \{1, 2\}$ doi vectori din H cu descompunerile ortogonale corespunzătoare ($y_\iota \in K$, $z_\iota \in K^\perp$). Dacă $\lambda_\iota \in C$, $\iota \in 1, 2$, atunci descompunerea ortogonală (unică) a vectorului $\lambda_1 x_1 + \lambda_2 x_2$ este:

$$\lambda_1 x_1 + \lambda_2 x_2 = (\lambda_1 y_1 + \lambda_2 y_2) + (\lambda_1 z_1 + \lambda_2 z_2),$$

şi deci:

$$P_K(\lambda_1 x_1 + \lambda_2 x_2) = (\lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 P_K x_1 + \lambda_2 P_K x_2,$$

ceea ce arată liniaritatea lui P_K. Continuitatea rezultă din:

$$\| P_K x_1 \|^2 = \| y_1 \|^2 \leq \| y_1 \|^2 + \| z_1 \|^2 = \| x_1 \|^2 .$$

Din relaţia de mai sus rezultă şi inegalitatea $\| P_K \| \leq 1$. Operatorul P_K este autoadjunct:

$$< P_K x_1, x_2 >=< y_1, y_2 + z_2 >=< y_1, y_2 >=< y_1 + z_1, y_2 >=< x_1, P_K x_2 > .$$

Dacă $y \in K$, atunci descompunerea sa ortogonală este $y = y + 0$, şi deci $P_K y = y$. De aici rezultă că pentru orice $x \in H$, $x = y + z$,
($y \in K$, $z \in K^\perp$), avem: $P^2 x = Py = y = Px$, ceea ce arată că P_K este o proiecţie având imaginea K.
(ii) Fie acum o proiecţie $P \in \mathcal{L}(H)$ şi fie $K = \mathrm{Im}(P)$. Evident, K este subspaţiu în H; demonstrăm că este închis. Pentru aceasta, fie $(y_n)_n \subset K$ un şir de vectori din K, convergent la $y \in H$. Din definiţia lui K rezultă că există un şir $(x_n)_n \subset H$ astfel încât $y_n = Px_n$; avem:

$$y = \lim_{n\to\infty} Px_n = \lim_{n\to\infty} P^2 x_n = P(\lim_{n\to\infty} Px_n) = Py,$$

ceea ce arată că $y \in K$ şi deci K este submulţime închisă în H; în plus, $Py = y\,,\ \forall y \in K$. Fie acum $z \in K^{\perp}$. Deoarece $P^2 z \in K$, rezultă:

$$\| Pz \|^2 = < Pz, Pz > = < z, P^2 z > = 0,$$

şi deci $Pz = 0$. Fie acum $x \in H\,,\ x = y + z$ cu $y \in K$ şi $z \in K^{\perp}$. Avem: $Px = P(y+z) = Py = y = P_K x$, ceea ce încheie demonstraţia.

11. Observaţie

Fie $P \in \mathcal{L}(H)$ o proiecţie şi fie $K = \mathrm{Im}(P)$ subspaţiul pe care ea proiectează. Atunci operatorul $I - P$ este de asemenea proiecţie şi $I - P = P_{K^{\perp}}$, sau , echivalent, $\mathrm{Im}(I-P) = K^{\perp}$. Evident, proiecţiile P şi $I - P$ satisfac relaţiile $P + (I-P) = I$ şi $P(I-P) = (I-P)P = O$. În general, suma şi diferenţa a două proiecţii P şi Q nu sunt proiecţii; propunem cititorului să demonstreze următoarele echivalenţe:

(i) $P + Q$ este proiecţie $\Longleftrightarrow PQ = QP = O$; în acest caz subspaţiul pe care proiectează $P + Q$ este suma (directă) a subspaţiilor $\mathrm{Im}(P)$ şi $\mathrm{Im}(Q)$.

(ii) $P - Q$ este proiecţie $\Longleftrightarrow PQ = QP = Q$; în acest caz, $\mathrm{Im}(P - Q)$ este suma (directă) a subspaţiilor $\mathrm{Ker}(P)$ şi $\mathrm{Im}(Q)$.

12. Definiţie

Pe mulţimea operatorilor autoadjuncţi se poate defini o relaţie de ordine (parţială) prin $T \leq S \Longleftrightarrow S - T$ este operator pozitiv, adică $< (S - T)x, x > \geq 0\,,\ \forall x \in H$. Demonstraţia este imediată. Restricţia acestei relaţii de ordine la mulţimea proiectorilor coincide cu relaţia de incluziune între subspaţiile imagine ale proiecţiilor, rezultatul fiind conţinut în următoarea teoremă. Cel mai mic proiector este operatorul identic nul, O, iar cel mai mare proiector este identitatea, I; evident, $\| O \| = 0$ şi $\| I \| = 1$. Dacă $P \in \mathcal{L}(H)$ este un proiector diferit de O şi I, atunci $\| P \| = 1$; inegalitatea $\| P \| \leq 1$ a fost deja arătată în demonstraţia teoremei 10. Pentru cealaltă inegalitate, fie $x_o \in \mathrm{Im}(P)$ astfel încât $\| x_o \| = 1$; atunci:

$$\| P \| = \sup_{\|x\|=1} \| Px \| \geq \| Px_o \| = \| x_o \| = 1.$$

13. Teoremă

Fie $P, Q \in \mathcal{L}(H)$ două proiecţii pe subspaţiile $\mathcal{X} = \mathrm{Im}(P)$ şi respectiv $\mathcal{Y} = \mathrm{Im}(Q)$. Următoarele afirmaţii sunt echivalente:

(a) $\mathcal{X} \subseteq \mathcal{Y}$.

(b) $QP = P$.

(c) $PQ = P$.

(d) $\| Px \| \leq \| Qx \|$.

(e) $P \leq Q$.

Demonstraţie (a)$\Longrightarrow$(b) Dacă $\mathcal{X} \subseteq \mathcal{Y}$, atunci pentru orice $x \in H$ avem

$Px \in \mathcal{X} \subseteq \mathcal{Y}$ şi deci $QPx = Px$.
(b)$\Longrightarrow$(c) Din $QP = P$, rezultă $PQ = P^\star Q^\star = (QP)^\star = P^\star = P$.
(c)$\Longrightarrow$(d) Din $PQ = P$ şi din inegalitatea $\| P \| \leq 1$ (arătată în demonstraţia teoremei 10), rezultă:

$$\| Px \| = \| PQx \| \leq \| Qx \| , \forall x \in H$$

.
(d)$\Longrightarrow$(e) Fie $x \in H$; folosind **(d)**, avem:

$$< Px, x > = < P^2x, x > = < Px, Px > = \| Px \|^2 \leq \| Qx \|^2 = < Qx, x > .$$

(e)$\Longrightarrow$(a) Fie $x \in \mathcal{X}$; din **(e)**, rezultă:

$$\| x \|^2 = < Px, x > \leq < Qx, x > = \| Qx \|^2 \leq \| x \|^2,$$

ceea ce arată că $\| Qx \| = \| x \| , \forall x \in \mathcal{X}$. Demonstraţia se încheie observând că din definiţia proiectorilor rezultă egalitatea:

$$\mathcal{Y} = \{ y \in H \, ; \, \| Qy \| = \| y \| \}.$$

În următoarea propoziţie dăm legătura dintre subspaţiile invariante ale unui operator şi proiectorii asociaţi acestor subspaţii.

14. Propoziţie
Fie $T \in \mathcal{L}(H)$ şi fie $P \in \mathcal{L}(H)$ un proiector; atunci:
(a) Ker(P) este invariant la $T \Longleftrightarrow PT = PTP$.
(b) Im(P) este invariant la $T \Longleftrightarrow TP = PTP$.
Înainte de a face demonstraţia, să observăm că deoarece Ker$(P)^\perp$ =Im(P), (P fiind operator autoadjunct), din afirmaţiile de mai sus rezultă:

Ker(P) şi Im(P) sunt subspaţii reducătoare pentru operatorul T dacă şi numai dacă $PT = TP$.
Demonstraţie Vom demonstra numai **(a)**, demonstraţia pentru **(b)** fiind asemănătoare; presupunem că Ker(P) este subspaţiu invariant pentru T. Fie $x \in H$ şi fie $y \in$Ker(P) şi $z \in$Im(P) =Ker$(P)^\perp$ astfel încât $x = y + z$; atunci $Px = z$ şi $PTy = 0$ şi deci avem:

$$PTx = PT(y + z) = PTy + PTz = PTz = PTPx.$$

Reciproc, dacă $x \in$Ker(P), atunci din ipoteză $PTx = PTPx = 0$, ceea ce încheie demonstraţia.

Încheiem acest paragraf cu o analogie între algebra operatorilor liniari şi continui pe un spaţiu Hilbert, $\mathcal{L}(H)$ şi corpul numerelor complexe, **C**.

corpul (comutativ) **C** $\longleftrightarrow$ algebra (necomutativă) $\mathcal{L}(H)$.

număr complex $z \in \mathbf{C} \longleftrightarrow$ operator $T \in \mathcal{L}(H)$.

conjugatul $\overline{z} \longleftrightarrow$ adjunctul $T^{\star}$.

număr real $a = \overline{a} \longleftrightarrow$ operator autoadjunct $A = A^{\star}$.

număr pozitiv $\longleftrightarrow$ operator pozitiv.

număr de modul 1 $z\overline{z} = 1 \longleftrightarrow$ operator unitar $UU^{\star} = U^{\star}U = I$.

soluţiile ecuaţiei $z^2 = z$ (adică 0 şi 1) $\longleftrightarrow$ proiector : $P^2 = P = P^{\star}$.

6.3 MF.06.3. Exemple de operatori pe spaţii Hilbert

Acest paragraf este rezervat studiului unor exemple importante de operatori, dintre care amintim operatorul diagonal, operatorii de translaţie, operatorii de multiplicare, operatorul integral şi de convoluţie.

15. Definiţie (Operatorul diagonal)
Fie $\ell^2(\mathbf{Z})$ spaţiul Hilbert al şirurilor (bilaterale) de pătrat sumabil şi fie $(\ell^{\infty}(\mathbf{Z}), \| \|_{\infty})$, mulţimea şirurilor mărginite. Pentru orice şir $\alpha \in \ell^{\infty}(\mathbf{Z})$, şi pentru orice $x \in \ell^2(\mathbf{Z})$, notăm cu αx şirul produs: $(\alpha x)(n) = \alpha(n)x(n)$, $\forall n \in \mathbf{Z}$. Inegalitatea:

$$\sum_{n=-\infty}^{\infty} |(\alpha x)(n)|^2 \leq \| \alpha \|_{\infty}^2 \sum_{n=-\infty}^{\infty} |x(n)|^2 < \infty,$$

arată că $\alpha x \in \ell^2(\mathbf{Z})$. Putem deci defini operatorul:

$$D_{\alpha} : \ell^2(Z) \longrightarrow \ell^2(Z)\,,\; D_{\alpha}x = \alpha x.$$

Vom numi D_{α} **operatorul diagonal** definit de α. Evident, se poate da o definiţie corespunzătoare şi pe spaţiul $\ell^2(\mathbf{N})$ (în acest caz $\alpha \in \ell^{\infty}(\mathbf{N})$).

16. Observaţie
Denumirea de operator diagonal este justificată de următoarea analogie cu cazul finit dimensional. Fie $(\sigma_n)_{n\in Z}$ baza canonică din $\ell^2(Z)$, adică $\sigma_n(k) = 1$ dacă $n = k$ şi 0 în rest. Fie $M = (a_{ij})_{i,j\in Z}$ "matricea infinită" a lui D_{α} în această bază, adică $a_{ij} =< D_{\alpha}\sigma_j, \sigma_i >$. Atunci $a_{ij} = \alpha(i)$, dacă $i = j$ şi $a_{ij} = 0$, dacă $i \neq j$.
Proprietăţile operatorului diagonal sunt cuprinse în următoarea teoremă.

17. Teoremă (proprietăţile operatorului diagonal)
(a) Pentru orice $\alpha\,, \beta \in \ell^{\infty}(\mathbf{Z})$ şi $a\,, b \in \mathbf{C}$, avem:

$$aD_{\alpha} + bD_{\beta} = D_{a\alpha+b\beta} \text{ şi } D_{\alpha}D_{\beta} = D_{\alpha\,\beta}.$$

(b) Operatorul D_α este liniar şi continuu.
(c) $\| D_\alpha \| = \| \alpha \|_\infty$.
(d) $(D_\alpha)^\star = D_{\overline{\alpha}}$. De aici rezultă că operatorul diagonal este normal şi în plus avem următoarele caracterizări:

$$D_\alpha \text{ este autoadjunct} \iff \alpha(n) \in \mathbf{R}\,,\ \forall n \in \mathbf{Z};$$

$$D_\alpha \text{ este pozitiv} \iff \alpha(n) \geq 0\,,\ \forall n \in \mathbf{Z}.$$

$$D_\alpha \text{ este unitar} \iff |\alpha(n)| = 1\,,\ \forall n \in \mathbf{Z}.$$

$$D_\alpha \text{ este proiector } \iff \alpha(n) = 0 \text{ sau } 1,\ \forall n \in \mathbf{Z} \text{ (sau } \alpha^2 = \alpha).$$

(e) Operatorul D_α este inversabil (în $\mathcal{L}(\ell^2(Z))$ dacă şi numai dacă şirul α este inversabil (în $\ell^\infty(Z)$), adică există $\beta \in \ell^\infty(\mathbf{Z})$ astfel încât $\alpha\beta = 1$. În consecinţă, avem:

$$\sigma(D_\alpha) = \overline{\{\alpha(n)\,;\, n \in \mathbf{Z}\}}.$$

În plus, spectrul punctual (mulţimea valorilor proprii) al operatorului diagonal este $\sigma_p(D_\alpha) = \{\alpha(n)\,;\, n \in \mathbf{Z}\}$.
Înainte de a face demonstraţia, să comparăm cu cazul finit dimensional (din capitolul 3): acolo valorile proprii erau numerele de pe diagonala principală a matricei operatorului în baza canonică (în cazul infinit dimensional, termenii şirului α); în plus, acum apar în spectru punctele limită ale "diagonalei" (şi care, în general, nu sunt valori proprii), dar care sunt în spectrul punctual aproximativ întrucât D_α este operator normal.
Demonstraţie Afirmaţiile de la punctul **(a)** sunt evidente.
(b) şi **(c)** Liniaritatea este imediată. Din inegalitatea

$$\sum_{n=-\infty}^{\infty} |(\alpha x)(n)|^2 \leq \| \alpha \|_\infty^2 \sum_{n=-\infty}^{\infty} |x(n)|^2 < \infty,$$

rezultă că $\| D_\alpha x \|_2 \leq \| \alpha \|_\infty \| x \|_2$ şi deci (folosind propoziţia 3, cap.3), obţinem continuitatea lui D_α şi inegalitatea $\| D_\alpha \| \leq \| \alpha \|_\infty$. Pentru a demonstra inegalitatea inversă, să observăm mai întâi că dacă $(\sigma_n)_{n\in Z}$ este baza canonică din $\ell^2(\mathbf{Z})$, atunci $D_\alpha \sigma_n = \alpha(n)\sigma_n$; rezultă:

$$|\alpha(n)| = \| \alpha(n)\sigma_n \|_2 = \| D_\alpha \sigma_n \|_2 \leq \| D_\alpha \| \| \sigma_n \|_2 = \| D_\alpha \|\ .$$

Luând supremum după $n \in \mathbf{Z}$ în inegalitatea de mai sus, demonstraţia se încheie.
(d) Pentru orice şiruri x şi y din $\ell^2(\mathbf{Z})$, avem:

$$< D_\alpha x, y > = \sum_{n=-\infty}^{\infty} \alpha(n)x(n)\overline{y(n)} = < x, \overline{\alpha} y > = < x, D_{\overline{\alpha}} y >,$$

ceea ce arată că $(D_\alpha)^\star = D_{\overline{\alpha}}$. Celelalte afirmaţii sunt imediate.
(e) Să presupunem mai întâi că şirul α este inversabil: există deci $\beta \in \ell^\infty(\mathbf{Z})$ astfel încât $\alpha(n)\,\beta(n) = 1\,,\,\forall n \in \mathbf{Z}$. Atunci operatorul D_β este liniar şi continuu şi $D_\alpha D_\beta = D_\beta D_\alpha = I$; deci $D_\beta = (D_\alpha)^{-1}$. Reciproc, dacă operatorul D_α este inversabil, atunci $(D_\alpha)^{-1}(\alpha(n)\sigma_n) =$
$= \sigma_n$, şi deci:

$$(D_\alpha)^{-1}\sigma_n = \frac{1}{\alpha(n)}\sigma_n.$$

De aici, trecând la norme, obţinem:

$$\| \frac{1}{\alpha(n)}\sigma_n \|_2 = \| (D_\alpha)^{-1}\sigma_n \|_2 \leq \| (D_\alpha)^{-1} \| \| \sigma_n \|_2 = \| (D_\alpha)^{-1} \| < \infty,$$

ceea ce implică $|\frac{1}{\alpha(n)}| \leq \| (D_\alpha)^{-1} \|\,,\,\forall n \in \mathbf{Z}$; în concluzie şirul $\frac{1}{\alpha}$ este mărginit şi deci α este inversabil.
Notând cu $\mathbf{1} \in \ell^\infty(\mathbf{Z})$ şirul constant 1, avem $\lambda I - D_\alpha = D_{\lambda \mathbf{1} - \alpha}\,,\,\forall \lambda \in \mathbf{C}$. Rezultă deci că operatorul $\lambda I - D_\alpha$ este inversabil dacă şi numai dacă şirul $\lambda\,\mathbf{1} - \alpha$ este inversabil; de aici rezultă simplu că $\sigma(D_\alpha) = \overline{\{\alpha(n)\,;\, n \in \mathbf{Z}\}}$.
Demonstrăm acum egalitatea $\sigma_p(D_\alpha) = \{\alpha(n)\,;\, n \in \mathbf{Z}\}$. Pentru orice n fixat în Z, fie $x_n \in \ell^2(Z)$ definit prin $x_n(k) = \alpha(n)$ pentru $k =$
$= n$ şi 0 pentru $k \neq n$. Atunci $((\alpha(n)I - D_\alpha)x_n)(k) = 0\,,\,\forall k \in \mathbf{Z}$ şi deci numărul $\alpha(n)$ este valoare proprie a operatorului D_α, iar x_n este un vector propriu asociat acestei valori proprii. Pentru a demonstra incluziunea reciprocă, fie $\lambda \in \sigma_p(D_\alpha)$ şi fie $x \in \ell^2(\mathbf{Z}])$ astfel încât x nu este şirul identic nul şi $((\lambda I - D_\alpha)x)(n) = 0\,,\,\forall n \in \mathbf{Z}$. Presupunând prin absurd că $\lambda \neq \alpha(n)\,,\,\forall n \in \mathbf{Z}$, din egalitatea

$$(\lambda - \alpha(n))x(n) = 0\,,\,\forall n \in Z,$$

obţinem $x(n) = 0\,,\,\forall n \in \mathbf{Z}$, contradicţie.

18. Definţie (operatorul de translaţie unilateral)
Pe spaţiul Hilbert $\ell^2(\mathbf{N})$ considerăm operatorul:

$$V : \ell^2(\mathbf{N}) \longrightarrow \ell^2(\mathbf{N})\,,\,(Vx)(0) = 0 \text{ şi } (Vx)(n) = x(n-1)\,,\,\forall n \geq 1.$$

Este evident că definiţia este corectă ($Vx \in \ell^2(\mathbf{N})$). Operatorul V se numeşte **operatorul de translaţie unilateral**, (unilateral shift); în teoria sistemelor V este denumit **întârziere ideală** (ideal delay).
Înainte de a enunţa şi demonstra proprietăţile operatorului V, vom face o observaţie cu caracter general în legătură cu spectrul unui operator.

19. Observaţie
Fie H un spaţiu Hilbert şi fie $T \in \mathcal{L}(H)$. Dacă $\lambda \in \sigma_p(T^\star)$, atunci , prin definiţie, există $x \in H\,,\, x \neq 0$ astfel încât $T^\star x = \lambda x$; avem deci:

$$\{0\} \neq \operatorname{Ker}(\lambda I - T^\star) = (\operatorname{Im}((\lambda I - T^\star)^\star))^\perp = (\operatorname{Im}(\overline{\lambda} I - T))^\perp,$$

ceea ce arată că $\mathrm{Im}(\overline{\lambda}I - T) \neq H$, deci operatorul $\overline{\lambda}I - T$ nu este surjectiv. Am demonstrat deci:

Dacă $\lambda \in \mathbf{C}$ este valoare proprie pentru $T^\star$, atunci $\overline{\lambda} \in \sigma(T)$.

20. Teoremă (proprietăţile operatorului de translaţie unilateral)
Fie $V : \ell^2(\mathbf{N}) \longrightarrow \ell^2(\mathbf{N})$ operatorul de translaţie unilateral; atunci:
(a) V este liniar şi continuu.
(b) V este o izometrie: $\| Vx \|_2 = \| x \|_2$, $\forall x \in \ell^2(\mathbf{N})$.
De aici rezultă, in particular, că $\| V \| = 1$.
(c) V nu este operator inversabil (nu este surjectiv).
(d) $(V^\star x)(n) = x(n+1)$, $\forall x \in \ell^2(\mathbf{N})$, $\forall n \in \mathbf{N}$ şi $\| V^\star \| = 1$.
(e) $V^\star V = I$ dar $VV^\star \neq I$.
(f) Operatorul V nu are valori proprii: $\sigma_p(V) = \emptyset$.
(g) $\sigma_p(V^\star) = \{\lambda \in \mathbf{C} \,;\, |\lambda| < 1\}$ şi $\sigma(V) = \sigma(V^\star) = \{\lambda \in \mathbf{C} \,;\, |\lambda| \leq 1\}$.
Înainte de demonstraţie, să observăm că pe spaţii finit dimensionale nu există endomorfisme injective care să nu fie surjective (de fapt operatorul V este mai mult decât injectiv, este o izometrie); dimpotrivă,în cazul finit dimensional orice izometrie este operator unitar. Este de asemenea de reţinut faptul că V nu are valori proprii, în timp ce în cazul unui operator definit pe un spaţiu finit dimensional spectrul este format numai din valori proprii.
Demonstraţie Prin definiţie, V acţionează astfel:

$$\ell^2(N) \ni x = (x(0), x(1), x(2), ..) \to (0, x(0), x(1), ..) = Vx \in \ell^2(N).$$

(a),(b),(c) Liniaritatea o lăsăm ca exerciţiu. Este evident (din schema de mai sus) că $\| Vx \|_2 = \| x \|_2$, şi deci V este izometrie. Operatorul V nu este surjectiv deoarece $\mathrm{Im}(V) = \{x \in \ell^2(\mathbf{N}) \,;\, x(0) = 0\} \neq \ell^2(\mathbf{N})$ de exemplu, nu există $x \in \ell^2(\mathbf{N})$ astfel încât $Vx = \sigma_o$.
(d) Pentru orice $x, y \in \ell^2(\mathbf{N})$, avem:

$$< Vx, y > = \sum_{n=0}^{\infty} (Vx)(n)\overline{y(n)} = \sum_{n=1}^{\infty} x(n-1)\overline{y(n)} = \sum_{n=0}^{\infty} x(n)\overline{y(n+1)},$$

ceea ce arată că adjunctul lui V este:

$$\ell^2(\mathbf{N}) \ni y = (y(0), y(1), y(2), ..) \to V^\star y = (y(1), y(2), y(3), ..) \in \ell^2(\mathbf{N}).$$

Este evident că pentru orice $x \in \ell^2(\mathbf{N})$ avem $\| V^\star x \|_2 \leq \| x \|_2$,deci $\| V^\star \| \leq 1$. Dar $\| V^\star \sigma_1 \|_2 = \| \sigma_o \|_2 = 1$, deci $\| V^\star \| = 1$.
(e) Este clar că $V^\star V = I$; dar, pentru orice $x \in \ell^2(\mathbf{N})$ cu proprietatea că $x(0) \neq 0$, avem $VV^\star x \neq x$.
(f) Arătăm acum că V nu are valori proprii. Fie , prin absurd , $\lambda \in \mathbf{C}$ astfel încât există $x \in \ell^2(N)$, $x \neq 0$, cu proprietatea $Vx = \lambda x$, adică:

$$(0, x(0), x(1), x(2), ...) = (\lambda x(0), \lambda x(1), \lambda x(2), ...).$$

De aici rezultă $x(n) = 0$, $\forall n \in \mathbf{N}$, contradicţie cu $x \neq 0$.
Am demonstrat deci că $\sigma_p(V) = \emptyset$.
(g) Deoarece $\| V \| = \| V^\star \| = 1$, rezultă că spectrele operatorilor V şi $V^\star$ sunt incluse în discul unitate închis. Vom arăta mai întâi că $\sigma_p(V^\star) = \{\lambda \in \mathbf{C}\,;\, |\lambda| < 1\}$. Din egalitatea $V^\star x = \lambda x$, rezultă:

$$(x(1), x(2), x(3), ...) = (\lambda x(0), \lambda x(1), \lambda x(2), ...),$$

şi deci $x(n+1) = \lambda^n x(0)$, $\forall n \in \mathbf{N}$. Dacă $x(0) = 0$, atunci $x = 0$. Rezultă deci că vectorii proprii x asociaţi valorii proprii λ sunt de forma:

$$x = (x(0), \lambda x(0), \lambda^2 x(0), \lambda^3 x(0), ...)\,,\ \text{cu condiţia}\ x(0) \neq 0.$$

Există însă restricţia $x \in \ell^2(\mathbf{N})$, ceea ce este echivalent cu $|\lambda| < 1$. Am demonstrat deci că:

$$\sigma_p(V^\star) = \{\lambda \in \mathbf{C}\,;\, |\lambda| < 1\}.$$

Din observaţia 21 de mai sus, rezultă că $\sigma(V) \supseteq \sigma_p(V^\star)$. În concluzie, spectrele operatorilor V şi $V^\star$ conţin discul unitate deschis şi sunt conţinute în discul unitate închis. Cum spectrul este mulţime închisă, rezultă că spectrele celor doi operatori sunt egale cu discul unitate închis.

21. Definiţie (operatorul de translaţie bilateral)
Pe spaţiul Hilbert $\ell^2(\mathbf{Z})$ considerăm aplicaţia:

$$W : \ell^2(\mathbf{Z}) \to \ell^2(\mathbf{Z})\,,\ (Wx)(n) = x(n-1)\,,\ \forall n \in \mathbf{Z}.$$

Este clar că $Wx \in \ell^2(\mathbf{Z})$, $\forall x \in \ell^2(\mathbf{Z})$. Operatorul W se numeşte **operatorul de translaţie bilateral**. Aşa cum vom vedea în teorema următoare, proprietăţile sale sunt în mod esenţial diferite de cele ale operatorului de translaţie unilateral.

22. Teoremă (proprietăţile operatorului de translaţie bilateral)
(a) W este liniar şi continuu.
(b) Adjunctul lui W este $(W^\star x)(n) = x(n+1)$, $\forall n \in \mathbf{Z}$.
(c) W este operator unitar: $WW^\star = W^\star W = I$.
În particular, $\| W \| = \| W^\star \| = 1$.
(d) Operatorii W şi $W^\star$ nu au valori proprii;
(e) $\sigma(W) = \sigma_{pa}(W) = \sigma(W^\star) = \sigma_{pa}(W^\star) = \{\lambda \in C\,;\, |\lambda| = 1\}$.
Reamintim că σ_{pa} este spectrul punctual aproximativ.
Demonstraţie (a) Liniaritatea este imediată; continuitatea rezultă din relaţia evidentă: $\| Wx \|_2 = \| x \|_2$.
(b) Pentru orice x, $y \in \ell^2(\mathbf{Z})$, avem:

$$< Wx, y >= \sum_{n \in Z} x(n-1)\overline{y(n)} = \sum_{n \in Z} x(n)\overline{y(n+1)},$$

şi deci într-adevăr $(W^\star x)(n) = x(n+1)\,,\,\forall n \in \mathbf{Z}$.
(c) Egalităţile $WW^\star = W^\star W = I$ sunt evidente.
(d) Vom demonstra că W nu are valori proprii (analog se arată şi pentru $W^\star$). Presupunem prin absurd că există $\lambda \in \mathbf{C}$ şi $x \in \ell^2(\mathbf{Z})\,,\, x \neq 0$ astfel încât $Wx = \lambda x$, adică $x(n-1) = = \lambda x(n)\,,\,\forall n \in \mathbf{Z}$. Rezultă deci că $x(n) = \lambda^{-n} x(0)\,,\,\forall n \in \mathbf{Z}$.
Dar $x \in \ell^2(\mathbf{Z})$ şi deci seriile geometrice:

$$\sum_{n=-\infty}^{0} |x(0)|^2 |\lambda|^{-2n} \text{ şi } \sum_{n=0}^{\infty} |x(0)|^2 |\lambda|^{-2n}$$

trebuie să fie simultan convergente; acest lucru este posibil numai dacă $x(0) = 0$, adică $x = 0$, contradicţie.
(e) Faptul că spectrele operatorilor W şi $W^\star$ sunt incluse în cercul unitate rezultă spectrul oricărui operator unitar este inclus în cercul unitate. Demonstraţia care urmează este totuşi independentă de această proprietate generală. Demonstrăm că spectrul punctual aproximativ al lui W este egal cu cercul unitate. Fie $\lambda = e^{it}$ şi fie $x_n \in \ell^2(\mathbf{Z})$, şirul definit prin:

$$x_n(k) = (2n+1)^{-\frac{1}{2}} e^{-ikt} \text{ pentru } |k| \leq n \text{ şi } x_n(k) = 0 \text{ în rest.}$$

Propunem cititorului să arate că $\| x_n \|_2 = 1$ şi:

$$\lim_{n \to \infty} \| (\lambda I - W) x_n \|_2 = 0.$$

Analog se calculează şi $\sigma_{pa}(W^\star)$. Cum $\sigma(W)$ şi $\sigma(W^\star)$ sunt incluse în cercul unitate, demonstraţia este încheiată.

O clasă importantă de operatori este clasa operatorilor de multiplicare; un caz particular a fost deja prezentat: operatorul diagonal pe spaţiul $\ell^2(\mathbf{Z})$. În continuare, vom da definiţia generală a acestor operatori.

23.Definiţie
Fie (Ω, μ) un spaţiu cu măsură (pozitivă); fie $L^2(\Omega, \mu)$ spaţiul Hilbert al funcţiilor de pătrat integrabil definite pe Ω şi fie $L^\infty(\Omega, \mu)$, spaţiul Banach al funcţiilor esenţial mărginite pe Ω. Pentru orice $f \in L^2(\Omega, \mu)$ şi pentru orice $\psi \in L^\infty(\Omega, \mu)$, definim funcţia produs $(\psi f)(t) = \psi(t) f(t)\,,\,\forall t \in \Omega$; avem:

$$\| \psi f \|_2 = \sqrt{\int_\Omega |\psi f|^2 d\mu} \leq \| \psi \|_\infty \| f \|_2,$$

deci $\psi f \in L^2(\Omega, \mu)$. Rezultă deci că pentru orice $\psi \in L^\infty(\Omega, \mu)$ putem defini aplicaţia:

$$M_\psi : L^2(\Omega, \mu) \to L^2(\Omega, \mu)\,,\, M_\psi f = \psi f.$$

Operatorul M_ψ se numeşte **operatorul de multiplicare** (sau **operatorul de înmulţire**) cu ψ; funcţia ψ se numeşte **multiplicator**. Operatorul diagonal D_α se obţine în cazul particular $\Omega = Z$, măsura μ este măsura de numărare şi $\psi = \alpha$.

24. Definiţie
Un alt caz particular remarcabil (pe care-l vom studia în continuare) este $\Omega = \mathcal{S}^1 = \{\lambda \in \mathbf{C}\,; |\lambda| = 1\}$ şi măsura μ măsura Lebesgue pe cercul unitate. Notăm $L^2(\mathcal{S}^1)$ spaţiul Hilbert al funcţiilor de pătrat integrabil pe cerc cu produsul scalar $< f, g >= \frac{1}{2\pi}\int\limits_0^{2\pi} f(e^{it})\overline{g(e^{it})}dt$ şi cu $L^\infty(\mathcal{S}^1)$ spaţiul Banach al funcţiilor esenţial mărginite pe cerc. Pentru orice $\psi \in L^\infty(\mathcal{S}^1)$ notăm cu M_ψ operatorul de multiplicare pe $L^2(\mathcal{S}^1)$: $(M_\psi f)(e^{it}) = \psi(e^{it})f(e^{it})$, $\forall e^{it} \in \mathcal{S}^1$.

Vom studia acum operatorii de multiplicare pe spaţiul Hilbert al funcţiilor de pătrat integrabil definite pe cercul unitate, $L^2(\mathcal{S}^1)$.
Menţionăm totuşi că rezultatele ce urmează sunt adevărate şi în cazul (general) al operatorilor de multiplicare pe spaţiul Hilbert $L^2(\Omega, \mu)$. Demonstraţiile care vor urma se adaptează (aşa cum vom preciza) uşor cazului general.

25. Propoziţie (proprietăţile operatorului de multiplicare)
Fie $\psi \in L^\infty(\mathcal{S}^1)$ şi M_ψ operatorul de multiplicare corespunzător; atunci:
(a) M_ψ este liniar şi continuu şi $\| M_\psi \| = \| \psi \|_\infty$.
(b) $M_\psi M_\phi = M_\phi M_\psi = M_{\psi\phi}, \forall \phi \in L^\infty(\mathcal{S}^1)$.
(c) $(M_\psi)^\star = M_{\overline{\psi}}$ şi M_ψ este operator normal.
(d) M_ψ este operator autoadjunct dacă şi numai dacă funcţia ψ ia valori reale: $\psi = \overline{\psi}$.
(e) M_ψ este operator pozitiv dacă şi numai dacă funcţia ψ ia valori pozitive: $\psi(e^{it}) \geq 0, \forall e^{it} \in \mathcal{S}^1$.
(f) M_ψ este operator unitar dacă şi numai dacă funcţia ψ ia valori de modul 1: $|\psi(e^{it})| = 1, \forall e^{it} \in \mathcal{S}^1$.
(g) M_ψ este proiector dacă şi numai dacă funcţia ψ este funcţia caracteristică a unei submulţimi măsurabile de pe cerc (deci ψ ia numai valorile 0 şi 1, sau, echivalent, $\psi^2 = \psi$).
Demonstraţie (a) Liniaritatea este evidentă; calculăm norma lui M_ψ. Pentru orice $f \in L^2(\mathcal{S}^1)$, avem:

$$\| M_\psi f \|_2 = \sqrt{\int_0^{2\pi} |\psi(e^{it})f(e^{it})|^2 dt} \leq \| \psi \|_\infty \| f \|_2 ,$$

şi deci am obţinut inegalitatea: $\| M_\psi \| \leq \| \psi \|_\infty$. Pentru a demonstra şi inegalitatea inversă, considerăm, pentru orice $n \in N$ mulţimea:

$$A_n = \{e^{it} \in \mathcal{S}^1 \, ; \, |\psi(e^{it})| \geq \| \, \psi \, \|_\infty - \frac{1}{n}\}.$$

Din definiţia supremumului esenţial, rezultă că măsura mulţimii A_n este nenulă (pentru orice $n \in N$). Dacă χ_n este funcţia caracteristică a mulţimii A_n , atunci $\chi_n \in L^2(\mathcal{S}^1)$ şi:

$$\| \, \chi_n \, \|_2 = \sqrt{\int_0^{2\pi} |\chi_n|^2 dt} = \sqrt{\mu(A_n)},$$

unde, $\mu(A_n)$ este măsura (Lebesgue pe cerc) a mulţimii A_n. Rezultă:

$$\| \, M_\psi \chi_n \, \|_2 = \sqrt{\int_0^{2\pi} |\psi \, \chi_n|^2 dt} \geq$$

$$\geq \sqrt{\int_0^{2\pi} (\| \, \psi \, \|_\infty - \frac{1}{n})^2 |\chi_n|^2 dt} \geq (\| \, \psi \, \|_\infty - \frac{1}{n}) \, \| \, \chi_n \, \|_2 \, .$$

Rezultă deci că:

$$\| \, M_\psi \, \| \geq \| \, \psi \, \|_\infty - \frac{1}{n} \, , \, \forall n \in N,$$

deci $\| \, M_\psi \, \| \geq \| \, \psi \, \|_\infty$.
(b) Pentru orice $\psi \, , \, \phi \in L^\infty(\mathcal{S}^1)$, avem:

$$M_\psi M_\phi f = M_\psi(\phi f) = \psi \phi f = M_{\psi\phi} f = M_{\phi\psi} f,$$

pentru orice $f \in L^2(\mathcal{S}^1)$.
(c) Dacă $f, g \in L^2(\mathcal{S}^1)$, atunci:

$$< M_\psi f, g >= \int_0^{2\pi} (\psi f) \overline{g} dt = \int_0^{2\pi} f \overline{(\overline{\psi} g)} dt =< f, M_{\overline{\psi}} g >,$$

şi deci $(M_\psi)^\star = M_{\overline{\psi}}$; operatorii de multiplicare comută toţi între ei , deci M_ψ şi $M_{\overline{\psi}}$ comută, adică M_ψ este operator normal.
Celelalte afirmaţii sunt uşor de demonstrat; de exemplu, pentru a demonstra **(f)**, să presunem mai întâi că M_ψ este unitar; atunci, din egalitatea $M_\psi M_{\overline{\psi}} = I$, rezultă $\psi \overline{\psi} = 1$. Reciproc, dacă $|\psi| = 1$, atunci, funcţia $\frac{1}{\psi} = \overline{\psi}$ este esenţial mărginită şi deci putem considera operatorul $M_{\frac{1}{\psi}}$, care este şi inversul, dar şi adjunctul lui M_ψ.

26.Observaţie
Raţionamentele de mai sus se pot reface, întocmai, şi în cazul unui spaţiu cu măsură arbitrar, (Ω, μ) cu proprietatea $\mu(\Omega) = 1$ şi, mai general, pentru un spaţiu cu măsură finită: $\mu(\Omega) < \infty$. Dacă
$\mu(\Omega) = \infty$, atunci, singura eroare în demonstraţia de mai sus ar fi aceea că

mulţimea A_n ar putea avea măsură infinită: $\mu(A_n) = \infty$ şi deci inegalitatea $\| M_\psi \| \geq \| \psi \|_\infty$ nu mai este adevărată. Pentru ca demonstraţia să fie corectă şi în acest caz, este suficient să existe o submulţime în $B_n \subset A_n$, care să fie măsurabilă de măsură finită; raţionamentul s-ar putea atunci face pentru B_n în loc de A_n (şi χ_{B_n} în loc de χ_{A_n}, etc). Pentru aceasta, trebuie ca măsura μ să aibă următorea proprietate: în orice mulţime măsurabilă (de măsură infinită) să existe o submulţime măsurabilă de măsură finită (o măsură cu această proprietate se numeşte local finită). O proprietate uzuală care implică local-finitudinea măsurii μ este ca ea să fie σ-finită, adică: există un şir de mulţimi măsurabile $(D_n)_{n\in N}$ cu proprietăţile:
$\mu(D_n) < \infty, \forall n \in \mathbf{N}$ şi $\Omega = \bigcup_{n\in N} D_n$. De exemplu, măsura Lebesgue pe $\mathbf{R}$ este σ-finită, pentru că
$\mathbf{R} = \bigcup\limits_{n\in N} (-n, n)$.

Calculul spectrului operatorului de multiplicare necesită câteva rezultate preliminare; rezultatul fundamental în această direcţie este că operatorul M_ϕ este inversabil dacă şi numai dacă funcţia ϕ este inversabilă în $L^\infty(\Omega, \mu)$, adică există $\psi \in L^\infty(\Omega, \mu)$ astfel încât $\phi\psi = 1$, sau, echivalent:
$\phi \in L^\infty(\Omega, \mu)$ este inversabilă dacă şi numai dacă există $m > 0$ astfel încât $|\phi(u)| \geq m$, $\forall u \in \Omega$ (a.p.t.).
Suficienţa conditiei de inversabilitate pentru M_ϕ admite o demonstraţie simplă; vom face demonstraţia pentru $\phi \in L^\infty(\mathcal{S}^1)$, dar ca şi mai sus, ea se poate adapta fără dificultăţi la cazul general.

27.Propoziţie
Dacă funcţia $\phi \in L^\infty(\mathcal{S}^1)$ este inversabilă, atunci operatorul de multiplicare M_ϕ este inversabil.
Demonstraţie Dacă funcţia $\phi \in L^\infty(\mathcal{S}^1)$ este inversabilă, atunci funcţia $\frac{1}{\phi}$ este în algebra $L^\infty(\mathcal{S}^1)$ şi deci operatorul de multiplicare cu $\frac{1}{\phi}$ este inversul căutat: $M_\phi M_{\frac{1}{\phi}} = I$.

28. Consecinţă
Dacă $\lambda \in \sigma(M_\phi)$, atunci $\lambda I - M_\phi$ nu este inversabil şi deci din propoziţia anterioară rezultă că funcţia $\lambda - \phi$ nu este inversabilă în $L^\infty(\mathcal{S}^1)$; concluzie: spectrul operatorului este inclus în imaginea esenţială a funcţiei ϕ: $\sigma(M_\phi) \subseteq \sigma(\phi) = \text{essran}(\phi)$. În particular, dacă ϕ este o funcţie continuă pe $\mathcal{S}^1$ atunci $\sigma(M_\phi) \subseteq \phi(\mathcal{S}^1)$.

Pentru a demonstra reciproca propoziţiei 27 (ceea ce ar rezolva pro blema calculului spectrului operatorului de multiplicare), avem nevoie de două rezultate auxiliare; ca de obicei, demonstraţiile vor fi prezentate în cazul $\Omega = \mathcal{S}^1$, dar raţionamentele sunt corecte şi pentru un spaţiu cu măsură $\sigma -$ finită, (Ω, μ).

29. Lemă
Fie $\phi : \mathcal{S}^1 \to \mathbf{C}$, o funcţie măsurabilă arbitrară.
Dacă $\phi f \in L^2(\mathcal{S}^1)$, $\forall f \in L^2(\mathcal{S}^1)$ şi dacă operatorul

$$T : L^2(\mathcal{S}^1) \to L^2(\mathcal{S}^1) \,,\, Tf = \phi f$$

este continuu, atunci $\phi \in L^\infty(\mathcal{S}^1)$ şi $\| \phi \|_\infty \leq \| T \|$.
Demonstraţie Vom demonstra că $|\phi| \leq \| T \|$ (a.p.t.); de aici va rezulta că $\| \phi \|_\infty \leq \| T \|$. Fie $M \subseteq \mathcal{S}^1$ o mulţime măsurabilă astfel încât $|\phi(z)| > \| T \|$, $\forall z \in M$; demonstraţia se încheie dacă demonstrăm că $\mu(M) = 0$ (aici am notat cu μ măsura Lebesgue pe cerc). Fie χ_M funcţia caracteristică a mulţimii M. Dacă prin absurd $\chi_M \neq 0$ a.p.t. , atunci:

$$\| T\chi_M \|_2 = \sqrt{\int_0^{2\pi} |\phi\chi_M|^2 dt} = \sqrt{\int_M |\phi|^2 dt} > \| T \| \, \| \chi_M \|_2 \,,$$

ceea ce este o contradicţie.
În demonstraţia de mai sus am folosit ipoteza de continuitate a operatorului T; se poate arăta că această ipoteză nu este necesară:

30. Lemă
Fie $\phi : \mathcal{S}^1 \to \mathbf{C}$ o funcţie măsurabilă arbitrară.
Dacă $\phi f \in L^2(\mathcal{S}^1)$, $\forall f \in L^2(\mathcal{S}^1)$, atunci $\phi \in L^\infty(\mathcal{S}^1)$.

Putem demonstra acum reciproca propoziţiei 27, şi deci avem:

31. Teoremă (spectrul operatorului de multiplicare)
Fie $\phi \in L^\infty(\mathcal{S}^1)$ şi fie M_ϕ operatorul de multiplicare asociat lui ϕ pe spaţiul $L^2(\mathcal{S}^1)$.
Atunci M_ϕ este operator inversabil dacă şi numai dacă funcţia ϕ este inversabilă (în $L^\infty(\mathcal{S}^1)$); consecinţe:
(a) Spectrul operatorului M_ϕ coincide cu imaginea esenţială a funcţiei ϕ;
(b) Raza spectrală şi norma operatorului de multiplicare sunt egale:

$$\mathrm{r}(M_\phi) = \| M_\phi \| \,.$$

Demonstraţie Implicaţia ϕ inversabilă $\Rightarrow$ M_ϕ inversabil a fost demonstrată în propoziţia 27.
Să presupunem acum că operatorul M_ϕ este inversabil, deci există $T \in \mathcal{L}(L^2(\mathcal{S}^1))$ astfel încât $TM_\phi f = M_\phi T f = f$, $\forall f \in L^2(\mathcal{S}^1)$. Rezultă deci că $\phi T f = f$, $\forall f \in L^2(\mathcal{S}^1)$, adică:

$$Tf = \frac{1}{\phi} f \,,\, \forall f \in L^2(\mathcal{S}^1).$$

Deoarece operatorul T este continuu, din lemele anterioare rezultă că funcţia $\frac{1}{\phi}$ este în $L^\infty(\mathcal{S}^1)$ şi $\| \frac{1}{\phi} \|_\infty \leq \| T \|$. În concluzie, ϕ este inversabilă în

$L^\infty(\mathcal{S}^1)$, inversa ei fiind $\frac{1}{\phi} \in L^\infty(\mathcal{S}^1)$.
(a) rezultă imediat aplicând rezultatul demonstrat mai sus operatorului $\lambda I - M_\phi = M_{\lambda-\phi}$.
(b) rezultă din **(a)** şi din definiţia razei spectrale:

$$\mathrm{r}(M_\phi) = \sup\{\, |\lambda|\, ;\, \lambda \in \sigma(M_\phi)\}.$$

32. Observaţie

Dacă funcţia ϕ este continuă pe cerc, atunci operatorul M_ϕ este inversabil dacă şi numai dacă $\phi(e^{it}) \neq 0\,, \forall e^{it} \in \mathcal{S}^1$. Rezultă deci că în acest caz spectrul operatorului M_ϕ coincide cu imaginea funcţiei ϕ. De exemplu, dacă $\phi(e^{it}) = e^{it} - 2$, atunci:

$$\sigma(M_\phi) = \mathrm{essran}(\phi) = \phi(\mathcal{S}^1) = \{\lambda \in C\, ;\, |\lambda + 2| = 1\}.$$

În particular, operatorul M_ϕ este inversabil (deoarece $0 \notin \sigma(M_\phi)$) şi inversul său este:

$$((M_\phi)^{-1} f)(e^{it}) = \frac{1}{e^{it} - 2} f(e^{it})\,, \forall f \in L^2(\mathcal{S}^1).$$

Să considerăm acum un exemplu în care funcţia multiplicator nu este continuă; fie , de exemplu:

$$\psi(e^{it}) = \frac{e^{it} - i}{e^{it} + i}\,, \text{ dacă}\, t \in [0,\, \pi) \text{ şi } \psi(e^{it}) = 1\,, \text{ dacă}\, t \in [\pi,\, 2\pi).$$

Atunci imaginea funcţiei ψ este:

$$\psi(\mathcal{S}^1) = \{1\} \bigcup \{it\, ;\, t \in [-1, 1)\}\,,$$

şi deci imaginea esenţială a lui ψ este:

$$\overline{\psi(\mathcal{S}^1)} = \{1\} \bigcup \{it\, ;\, t \in [-1, 1]\} = \sigma(M_\psi).$$

În particular, operatorul M_ψ nu este inversabil.
Tot în legătură cu acest exemplu, se observă uşor că $\lambda = 1$ este valoare proprie a operatorului M_ψ. Într-adevăr, orice funcţie neidentic nulă $f \in L^2(\mathcal{S}^1)$ care se anulează pe submulţimea $\{e^{it}\, ;\, t \in [0,\, \pi]\}$, verifică egalitatea $M_\psi f = \psi f = f$, deci este vector propriu (asociat valorii proprii 1) al operatorului M_ψ.
În general, se poate demonstra că λ este valoare proprie pentru operatorul M_ϕ dacă şi numai dacă mulţimea pe care funcţia ϕ ia valoarea λ are măsură nenulă ([10],p.40).

33. Observaţie
Aşa cum am menţionat, rezultatele demonstrate mai sus sunt adevărate şi în cazul general al unui operator de multiplicare
$M_\psi : L^2(\Omega,\mu) \to L^2(\Omega,\mu)$, cu funcţia $\psi \in L^\infty(\Omega,\mu)$, iar spaţiul cu măsură (Ω,μ) este $\sigma-$finit. Propunem cititorului să adapteze în mod corespunzător demonstraţiile.
Să considerăm acum câteva exemple pe $\mathbf{R}$. Fie deci $\Omega = \mathbf{R}$, măsura μ măsura Lebesgue (pe $\mathbf{R}$) şi fie

$$\phi(x) = \frac{1+2x^2}{1+x^2}\ , \forall x \in \mathbf{R}.$$

Atunci $\phi \in L^\infty(\mathbf{R})$ şi $\phi(\mathbf{R}) = [1,2)$. Rezultă deci că $\sigma(M_\phi) = [1,2]$; în particular, operatorul M_ϕ este pozitiv şi inversabil.
Vom da acum un exemplu de operator de multiplicare unitar pe $L^2(\mathbf{R})$; fie $\psi(x) = \frac{x+i}{ix+1}$, $\forall x \in \mathbf{R}$. Atunci $\sigma(M_\psi) = \mathcal{S}^1$, deci M_ψ este unitar.
Pentru a obţine proiectori, trebuie ca multiplicatorul ψ să fie funcţia caracteristică a unei mulţimi măsurabile. De exemplu, pentru orice $\tau \in \mathbf{R}$ (fixat) fie χ_τ funcţia caracteristică a intervalului $(-\infty,\tau]$. Atunci operatorul de multiplicare cu χ_τ este proiectorul pe subspaţiul (din $L^2(\mathbf{R})$) al funcţiilor care se anulează pe intervalul (τ,∞); el se numeşte operatorul de trunchiere la momentul τ , iar o notaţie uzuală este P_τ. Evident, $\sigma(P_\tau) = \sigma_p(P_\tau) = \{0,1\}$.

34. Definiţie
Fie $(H,<,>_H)$ şi $(K,<,>_K)$ două spaţii Hilbert şi fie $T \in \mathcal{L}(H)$ şi $S \in \mathcal{L}(K)$. Operatorii T şi S se numesc **unitar-echivalenţi** dacă există un izomorfism de spaţii Hilbert $U : H \to K$ astfel încât

$$T = U^{-1}SU.$$

De exemplu, în capitolul 3, am demonstrat că un operator normal pe $\mathbf{C^n}$ este unitar-echivalent cu un operator diagonal.

35. Propoziţie
Doi operatori unitar-echivalenţi au acelaşi spectru, aceeaşi normă, aceleaşi valori proprii, acelaşi spectru punctual aproximativ.
De asemenea, dacă S este inversabil, atunci:

$$T^\star = U^{-1}S^\star U \text{ şi } T^{-1} = U^{-1}S^{-1}U.$$

Demonstraţie Reamintim că prin izomorfism de spaţii Hilbert (sau operator unitar) se înţelege un operator liniar bijectiv cu proprietatea $< Ux, Uy >_K = < x, y >_H$, $\forall x, y \in H$. Să demonstrăm de exemplu egalitatea $\sigma(S) = \sigma(T)$; dacă $\lambda \in C$, atunci $\lambda I - S$ este inversabil $\Leftrightarrow$ $U^{-1}(\lambda I - S)U$ este inversabil $\Leftrightarrow \lambda I - U^{-1}SU$ este inversabil.

Pentru a demonstra egalitatea $T^\star = U^{-1}S^\star U$, trebuie să definim adjunctul unui operator oarecare, $A : H \to K$; procedându-se în mod analog cazului $H = K$ (propoziţia 1 din acest capitol) se demonstreză existenţa unui operator

$$A^\star : K \to H \text{, astfel încât } < Au, v >_K = < u, A^\star v >_H \ \forall u \in H \text{ şi } v \in K.$$

Proprietăţile lui $A^\star$ sunt practic identice cu cele din cazul $H = K$; de exemplu, dacă $A = U$ este unitar , atunci $U^\star = U^{-1}$. Rezultă deci $T^\star = (U^{-1}SU)^\star = U^{-1}S^\star U$. În mod analog se demonstrează celelalte proprietăţi.
O clasă de operatori care sunt unitar echivalenţi cu operatorii de multiplicare pe $L^2(\mathcal{S}^1)$ sunt **operatorii de convoluţie** pe spaţiul $\ell^2(\mathbf{Z})$.
Reamintim că dacă α şi β sunt două şiruri definite pe $\mathbf{Z}$, atunci convoluţia lor este şirul notat $\alpha \star \beta$ definit prin:

$$(\alpha \star \beta)(n) = \sum_{n=-\infty}^{\infty} \alpha(k)\beta(n-k),$$

în ipoteza că seria converge pentru orice $n \in \mathbf{Z}$ fixat. Produsul de convoluţie este comutativ, asociativ, distributiv faţă de adunare şi are element neutru şirul $\sigma_o(n) = 1$ dacă $n = 0$ şi 0 în rest.

36. Definiţie (operatorul de convoluţie pe Z)
Fie un şir $\theta \in \ell^2(\mathbf{Z})$ astfel încât:
(i) $\theta \star x \in \ell^2(\mathbf{Z})$, $\forall x \in \ell^2(\mathbf{Z})$.
Cu această condiţie îndeplinită, putem defini **operatorul de convoluţie** cu θ prin relaţia:

$$C_\theta : \ell^2(\mathbf{Z}) \to \ell^2(\mathbf{Z}) \, , \, C_\theta x = \theta \star x.$$

Menţionăm că restricţia $\theta \in \ell^2(\mathbf{Z})$ este necesară pentru existenţa transformatei Fourier inverse $\mathcal{F}^{-1}\theta$.
Liniaritatea operatorului C_θ este evidentă; continuitatea şi o explicitare a condiţiei **(i)** urmează a fi studiate în continuare.
Pentru aceasta, vom demonstra că operatorul de convoluţie este unitar echivalent cu un operator de multiplicare.

Transformarea Fourier:

$$\mathcal{F} : L^2(\mathcal{S}^1) \to \ell^2(\mathbf{Z}) \, , \, \mathcal{F}f = \hat{f},$$

unde $\hat{f}(n) = \frac{1}{2\pi}\int\limits_0^{2\pi} f(e^{it})e^{-int}dt$, este un izomorfism de spaţii Hilbert. Inversa ei este definită prin:

$$\mathcal{F}^{-1}x = \sum_{n \in Z} x(n)\omega_n,$$

unde, am notat $\omega_n(e^{it}) = e^{int}$.
Reamintim de asemenea că restricţia lui $\mathcal{F}^{-1}$ la $\ell^1(\mathbf{Z})$ admite o formulă punctuală:

$$\left(\mathcal{F}^{-1}\alpha\right)(e^{it}) = \sum_{n\in Z} \alpha(n)e^{int}, \; \forall \alpha \in \ell^1(\mathbf{Z}).$$

37. Lemă
Pentru orice $\alpha, \beta \in \ell^1(\mathbf{Z})$, are loc egalitatea:

$$\mathcal{F}^{-1}(\alpha \star \beta) = (\mathcal{F}^{-1}\alpha)(\mathcal{F}^{-1}\beta).$$

Demonstraţie Deoarece α şi β sunt în $\ell^1(\mathbf{Z})$, convoluţia $\alpha \star \beta$ există şi aparţine de asemenea spaţiului $\ell^1(\mathbf{Z})$.
Pentru orice $e^{it} \in \mathcal{S}^1$, avem:

$$[\mathcal{F}^{-1}(\alpha \star \beta)](e^{it}) = \sum_{n\in Z}(\alpha \star \beta)(n)e^{int} = \sum_{n\in Z}\sum_{k\in Z} \alpha(k)\beta(n-k)e^{int} =$$

$$= \sum_{k\in Z}[\alpha(k) \sum_{m\in Z} \beta(m)e^{i(k+m)t}] = (\sum_{k\in Z} \alpha(k)e^{ikt})(\sum_{m\in Z} \beta(m)e^{imt}) =$$

$$= [(\mathcal{F}^{-1}\alpha)(\mathcal{F}^{-1}\beta)](e^{it}),$$

schimbarea ordinei de sumare fiind posibilă deoarece seriile sunt absolut convergente.

38. Teoremă
(a) Fie $\theta \in \ell^2(\mathbf{Z})$; atunci: $(\mathcal{F}^{-1}C_\theta\mathcal{F})f = (\mathcal{F}^{-1}\theta)f$, $\forall f \in L^2(\mathcal{S}^1)$.
(b) Operatorul de convoluţie cu θ este corect definit (în sensul că îndeplineşte condiţia **(i)** din definiţia 36) dacă şi numai dacă

$$\mathcal{F}^{-1}\theta \in L^\infty(\mathcal{S}^1);$$

în acest caz, C_θ este operator continuu şi este unitar echivalent cu operatorul de multiplicare cu $\mathcal{F}^{-1}\theta$, adică $\mathcal{F}^{-1}C_\theta\mathcal{F} = M_{\mathcal{F}^{-1}\theta}$.

(c) Reciproc, fiind dat un operator de multiplicare M_ϕ pe $L^2(\mathcal{S}^1)$, (deci $\phi \in L^\infty(\mathcal{S}^1)$), există un operator de convoluţie pe $\ell^2(\mathbf{Z})$, şi anume $C_{\widehat{\phi}}$ astfel $M_\phi = \mathcal{F}^{-1}C_{\widehat{\phi}}\mathcal{F}$.
Demonstraţie (a) Vom demonstra egalitatea de la punctul **(a)** pentru funcţiile $\omega_n(e^{it}) = e^{int}$, $\forall n \in \mathbf{Z}$ în ipoteza suplimentară $\theta \in \ell^1(\mathbf{Z})$.
Pentru aceasta, calculăm mai întâi $\mathcal{F}\omega_n$:

$$(\mathcal{F}\omega_n)(m) = \frac{1}{2\pi}\int_0^{2\pi} e^{int}e^{-imt}dt = \sigma_n(m), \; \forall m \in \mathbf{Z},$$

unde, reamintim,

$$\sigma_n(m) = \begin{cases} 1 & \text{dacă} \quad m = n \\ 0 & \text{dacă} \quad m \neq n \end{cases}$$

Să mai observăm că $\sigma_n \in \ell^1(\mathbf{Z})$ şi deci:

$$\left(\mathcal{F}^{-1}\sigma_n\right)(e^{it}) = e^{int},\ \forall e^{it} \in \mathcal{S}^1.$$

Fie $\theta \in \ell^1(\mathbf{Z})$; atunci, în baza lemei anterioare, avem:

$$\mathcal{F}^{-1}C_\theta\mathcal{F}\omega_n = \mathcal{F}^{-1}\left(\theta \star \sigma_n\right) = \left(\mathcal{F}^{-1}\theta\right)\omega_n,\ \forall n \in \mathbf{Z}.$$

Deoarece subspaţiul liniar generat de funcţiile $\{\omega_n\}_{n\in Z}$ este dens în $L^2(\mathcal{S}^1)$, egalitatea **(a)** este adevărată pentru orice $f \in L^2(\mathcal{S}^1)$ pentru că $\mathcal{F}$ şi $\mathcal{F}^{-1}$, sunt aplicaţii continue (în norma $\|\ \|_2$).
Acum, egalitatea **(a)** rezultă şi pentru $\theta \in \ell^2(Z)$, deoarece $\ell^1(Z)$ este dens în $\ell^2(Z)$.
(b) Totul rezultă acum din proprietăţile operatorului de multiplicare cu $\mathcal{F}^{-1}\theta$; am demonstrat că acesta (şi deci şi C_θ) este corect definit (şi în acest caz şi continuu) dacă şi numai dacă funcţia multiplicator $\mathcal{F}^{-1}\theta$ este esenţial mărginită pe $\mathcal{S}^1$.
(c) Fie $\phi \in L^\infty(\mathcal{S}^1)$; atunci, pentru orice $n \in \mathbf{Z}$, avem:

$$\frac{1}{2\pi}\int_0^{2\pi} |\phi(e^{it})e^{-int}|dt \leq \|\ \phi\ \|_\infty < \infty,$$

şi deci funcţia ϕ are transformată Fourier:

$$\widehat{\phi} : \mathbf{Z} \mapsto \mathbf{C}\,,\ \widehat{\phi}(n) = \frac{1}{2\pi}\int_0^{2\pi} \phi(e^{it})e^{-int}dt.$$

Faptul că operatorul $C_{\widehat{\phi}}$ este corect definit şi este unitar echivalent cu M_ϕ rezultă din **(a)** şi **(b)**.
Să observăm că demonstraţia punctului **(c)** s-a bazat în mod esenţial pe faptul că măsura (Lebesgue) a cercului unitate este finită, ceea ce implică faptul că o funcţie din $L^\infty(\mathcal{S}^1)$ are transformată Fourier (coeficienţi Fourier), deoarece $L^\infty(\mathcal{S}^1) \subset L^2(\mathcal{S}^1)$.

39. Corolar (proprietăţile operatorului de convoluţie pe Z)
Fie $\theta : \mathbf{Z} \to \mathbf{C}$ astfel încât $\mathcal{F}^{-1}\theta \in L^\infty(\mathcal{S}^1)$ şi fie C_θ operatorul de convoluţie asociat; atunci, din teorema de mai sus şi din proprietăţile operatorilor de multiplicare demonstrate anterior, avem:
(a) $\|\ C_\theta\ \| = \|\ M_{\mathcal{F}^{-1}\theta}\ \| = \|\ \mathcal{F}^{-1}\theta\ \|_\infty$.
(b) $\sigma(C_\theta) = \sigma(M_{\mathcal{F}^{-1}\theta}) = \mathrm{essran}(\mathcal{F}^{-1}\theta)$.
(c) Dacă şirul $\theta \in \ell^1(Z)$, atunci funcţia $\mathcal{F}^{-1}\theta$ este continuă pe $\mathcal{S}^1$, deci (deoarece $\mathcal{S}^1$ este compact) ea este şi mărginită şi spectrul lui C_θ este imaginea funcţiei $\mathcal{F}^{-1}\theta$:

$$\sigma(C_\theta) = (\mathcal{F}^{-1}\theta)(\mathcal{S}^1).$$

În particular, în acest caz, C_θ este operator inversabil dacă şi numai dacă $\mathcal{F}^{-1}\theta$ nu se anulează pe $\mathcal{S}^1$.

(d) Dacă C_θ este operator inversabil, atunci inversul său este operatorul de convoluţie cu $\mathcal{F}\left(\frac{1}{\mathcal{F}^{-1}\theta}\right)$, adică

$$C_\theta^{-1}x = \mathcal{F}\left(\frac{1}{\mathcal{F}^{-1}\theta}\right) \star x, \ \forall x \in \ell^2(\mathbf{Z}).$$

Pentru demonstraţia punctului **(d)**, trebuie arătată mai întâi egalitatea:

$$\widehat{f\,g} = \widehat{f} \star \widehat{g}, \ \forall f, g \in L^2(\mathcal{S}^1).$$

Lăsăm detaliile ca exerciţiu.

40. Exemplu

(i) Operatorul de translaţie bilateral W (din definiţia 23) este un operator de convoluţie; într-adevăr, $W = C_{\sigma_1}$.
Mai general, W^n este operatorul de convoluţie cu şirul σ_n, pentru orice $n \in Z$.
Deoarece $(\mathcal{F}^{-1}\sigma_n)(e^{it}) = \omega_n(e^{it}) = e^{int}$, rezultă că W^n este unitar-echivalent cu M_{ω_n}:

$$\mathcal{F}^{-1}W^n\mathcal{F} = M_{\omega_n}.$$

În particular, $(\mathcal{F}^{-1}W\mathcal{F}f)(e^{it}) = e^{it}f(e^{it})$.

(ii) Fie operatorul

$$T : \ell^2(Z) \to \ell^2(Z)\,,\ (Tx)(n) = x(n+1) - 4x(n-1)\,,\ \forall n \in Z.$$

Vom demonstra că T este inversabil şi vom calcula inversul său; fie:

$$\theta : Z \to C\,,\ \theta(-1) = 1\,,\ \theta(1) = -4 \text{ şi } \theta(n) = 0 \text{ în rest.}$$

Atunci evident $T = C_\theta$ şi vom aplica rezultatele precedente; avem:

$$(\mathcal{F}^{-1}\theta)(e^{it}) = e^{it} - 4e^{-it} = \frac{e^{2it} - 4}{e^{it}},$$

şi deci C_θ este operator inversabil deoarece $\mathcal{F}^{-1}\theta$ nu se anulează pe cercul unitate.
Inversul operatorului $\mathcal{F}^{-1}C_\theta\mathcal{F} = M_{\mathcal{F}^{-1}\theta}$, este operatorul de multiplicare cu funcţia

$$\frac{1}{\mathcal{F}^{-1}\theta}(e^{it}) = \frac{e^{it}}{e^{2it} - 4}.$$

Rezultă deci că:

$$(C_\theta)^{-1}x = \left(\mathcal{F}\left(\frac{1}{\mathcal{F}^{-1}\theta}\right)\right) \star x\,,\ \forall x \in \ell^2(Z).$$

Calculăm transformata Fourier a funcţiei $\frac{1}{\mathcal{F}^{-1}\theta}$:

$$\left(\mathcal{F}\left(\frac{1}{\mathcal{F}^{-1}\theta}\right)\right)(n) = \frac{1}{2\pi}\int_0^{2\pi}\frac{e^{it}}{e^{2it}-4}e^{-int}dt = \frac{1}{2\pi i}\int_{\mathcal{S}^1}\frac{z^{-n}}{z^2-4}dz.$$

Calculând ultima integrală cu teorema reziduurilor, obţinem:

$$\left(\mathcal{F}\left(\frac{1}{\mathcal{F}^{-1}\theta}\right)\right)(n) = \begin{cases} 0, \text{ dacă}\, n \leq 0 \\ 0, \text{ dacă}\, n > 0 \,\text{şi}\, n\, \text{par} \\ -\frac{1}{2^{n-1}}, \text{ dacă}\, n > 0 \text{ şi}\, n\, \text{impar} \end{cases}$$

Rezultă deci:

$$((C_\theta)^{-1}x)(n) = \sum_{k\geq 0} -\frac{1}{2^{2k}}x(n-2k-1).$$

În continuare prezentăm **operatorul de convoluţie** pe R; rezultatele sunt similare celor din cazul Z, cu excepţia punctului **(c)** din teorema 40. Demonstraţiile vor fi omise, principalele referiri bibliografice sunt [17],p.49; [13],p.192; [6],p.949.

41. Definiţie (operatorul de convoluţie pe R)
Fie $L^1(\mathbf{R})$ spaţiul Banach al funcţiilor integrabile şi fie $k \in L^1(\mathbf{R})$. Atunci, pentru orice funcţie $f \in L^2(\mathbf{R})$, convoluţia $(k \star f)(x) = \int\limits_R k(x-y)f(y)dy$ defineşte o funcţie din $L^2(\mathbf{R})$ şi în plus $\| k \star f \|_2 \leq \| k \|_1 \| f \|_2$; pentru demonstraţie, recomandăm [6],p.951.
Rezultă deci că **operatorul de convolutie** cu funcţia $k \in L^1(\mathbf{R})$ este corect definit pe spaţiul $L^2(\mathbf{R})$:

$$C_k : L^2(\mathbf{R}) \to L^2(\mathbf{R}),\, C_k f = k \star f.$$

42. Observaţie
Reamintim câteva proprietăţi ale transformatei Fourier pe $\mathbf{R}$; ca bibliografie recomandăm [B01], [D02], [S02].
Pentru orice funcţie $k \in L^1(\mathbf{R})$, transformata sa Fourier este, prin definiţie

$$(\mathcal{F}k)(x) = \widehat{k}(x) = \int_R k(y)e^{-ixy}dy\,,\, \forall x \in R.$$

Funcţia $\widehat{k}$ este continuă şi mărginită pe $\mathbf{R}$ şi $\| \widehat{k} \|_\infty \leq \| k \|_1$. Menţionăm că există funcţii continue şi mărginite pe $\mathbf{R}$ care nu sunt transformate Fourier ale unor funcţii din $L^1(\mathbf{R})$.
Restricţia aplicaţiei $\mathcal{F}$ la subspaţiul $K = L^1(\mathbf{R})\bigcap L^2(\mathbf{R})$ ia valori în $L^2(\mathbf{R})$ şi deci, deoarece K este dens în $L^2(\mathbf{R})$, ea se poate prelungi prin

continuitate (în norma $\| \|_2$) la întregul $L^2(\mathbf{R})$; se obţine astfel transformarea Fourier (sau Fourier-Plancherel):

$$\mathcal{F} : L^2(\mathbf{R}) \to L^2(\mathbf{R}),$$

cu proprietăţile:
(a) $\mathcal{F}$ este un izomorfism de spaţii Hilbert; acest rezultat este cunoscut sub numele de **teorema lui Plancherel**.
(b) $\mathcal{F}(k \star f) = \widehat{k}\widehat{f}$.
Pentru demonstraţia teoremei lui Plancherel, recomandăm: [B01], [D02],[S02].

43. Teoremă (proprietăţile operatorului de convoluţie pe R)
Fie $k \in L^1(\mathbf{R})$, fie C_k operatorul de convoluţie cu k , $\mathcal{F}$ transformarea Fourier şi $M_{\widehat{k}}$ operatorul multiplicare cu $\widehat{k}$; atunci:
(a) $C_k = \mathcal{F}^{-1} M_{\widehat{k}} \mathcal{F}$.
(b) $\| C_k \| = \| M_{\widehat{k}} \| = \| \widehat{k} \|_\infty$.
(c) $\sigma(C_k) = \sigma(M_{\widehat{k}}) = \overline{\widehat{k}(\mathbf{R})}$.
Demonstraţie Toate afirmaţiile rezultă din lema şi observaţia anterioare şi din proprietăţile corespunzătoare ale operatorului de multiplicare.

44. Observaţie
Aşa cum am văzut, există asemănări importante între operatorii de convoluţie pe $\ell^2(\mathbf{Z})$ şi cei de pe $L^2(\mathbf{R})$. Metoda de studiu este aceeaşi: sunt unitar-echivalenţi (prin transformarea Fourier) cu operatori de multiplicare pe $L^2(\mathcal{S}^1)$ şi respectiv pe $L^2(\mathbf{R})$. În teoria sistemelor, spaţiul pe care este definit operatorul de convoluţie se numeşte domeniul timp, iar spaţiul pe care este definit operatorul de multiplicare corespunzător se numeşte domeniul frecvenţă; unitar-echivalenţa celor doi operatori prin transformarea Fourier este denumită dualitatea timp-frecvenţă. Menţionăm totuşi o deosebire remarcabilă între cele două cazuri. Deoarece orice funcţie din $L^\infty(\mathcal{S}^1)$ are transformată Fourier, rezultă că orice operator de multiplicare pe $L^2(\mathcal{S}^1)$ este unitar echivalent cu un operator de convoluţie pe $\ell^2(Z)$ (cf. teoremei 40(c)). În schimb, există funcţii $\phi \in L^\infty(R)$ care nu au transformată Fourier: integrala $\int\limits_R \phi(t) e^{-itx} dt$ nu converge (măsura Lebesgue a lui R este ∞). Concluzie: pentru un astfel de ϕ, operatorul de multiplicare M_ϕ este corect definit pe $L^2(R)$, dar nu există un operator de convoluţie pe $L^2(R)$ unitar-echivalent cu M_ϕ prin transformarea Fourier.

Dacă $M_T = (a_{ij})\, , i, j \in \{1, 2, ..., n\}$ este matricea unui operator T pe spaţiul C^n, atunci $(Tx)_i = \sum\limits_{j=1}^{n} a_{ij} x_j \, , \forall x = (x_1, x_2, .., x_n) \in C^n$. Un analog infinit dimensional al acestei definiţii este **operatorul integral**.

45. Definiţie (operatorul integral)
Să considerăm spaţiul Hilbert $(L^2(0,1)\,,\ \|\ \|_2)$ al funcţiilor de pătrat integrabil pe intervalul $[0,\ 1]$ în raport cu măsura Lebesgue. Fie $K : [0,1]\times[0,1] \to \mathbf{C}$ o funcţie de pătrat integrabil pe $[0,1]\times[0,1]$ şi fie:

$$\| K \|_2 = \sqrt{\int_0^1 \int_0^1 |K(x,y)|^2 dxdy}.$$

Demonstrăm acum că pentru orice funcţie $f \in L^2(0,1)$, funcţia g definită prin egalitatea:

$$g(x) = \int_0^1 K(x,y) f(y) dy$$

este în $L^2(0,1)$; avem (folosim inegalitatea lui Schwarz):

$$\| g \|_2^2 = \int_0^1 |\int_0^1 K(x,y) f(y) dy|^2 dx \leq$$

$$\leq \int_0^1 \left(\int_0^1 |K(x,y)|^2 dy\right) \left(\int_0^1 |f(y)|^2 dy\right) dx = \| f \|_2^2 \| K \|_2^2 .$$

Rezultă deci că putem defini aplicaţia:

$$T_K : L^2(0,1) \to L^2(0,1)\,,\ (T_K f)(x) = \int_0^1 K(x,y) f(y) dy.$$

Operatorul T_K se numeşte **operatorul integral** definit de **nucleul** K.
Din calculul de mai sus rezultă că T_K este continuu şi
$\| T_K \| \leq \| K \|_2$.

46. Propoziţie (proprietăţile operatorului integral)
Fie T_K şi T_H doi operatori integrali cu nucleele K şi respectiv H.
(a) Pentru orice $\alpha\,,\ \beta \in \mathbf{C}$, operatorul $\alpha T_K + \beta T_H$ este operator integral şi are nucleul $\alpha K + \beta H$.
(b) Operatorul $T_K T_H$ este operator integral şi are nucleul definit prin $G(x,y) = \int_0^1 K(x,z) H(z,y) dz$, deci:

$$(T_K T_H f)(x) = \int_0^1 \left(\int_0^1 K(x,z) H(z,y) dz\right) f(y) dy.$$

În cazul particular $K = H$, obţinem:

$$(T_K^2 f)(x) = \int_0^1 \left(\int_0^1 K(x,z) K(z,y) dz\right) dx.$$

(c) Dacă şirul de nuclee K_n converge în spaţiul Hilbert $L^2([0,1]\times[0,1])$ la funcţia K, atunci şirul de operatori integrali T_{K_n} converge în spaţiul

$(\mathcal{L}\left(L^2(0,1)\right), \| \|)$ la operatorul integral T_K.
(d) Adjunctul operatorului T_K este operatorul integral $T_{\widetilde{K}}$, cu nucleul $\widetilde{K}(x,y) = \overline{K(y,x)}$; în particular, T_K este autoadjunct dacă şi numai dacă nucleul K are proprietatea $K(x,y) = \overline{K(y,x)}$ (un astfel de nucleu se numeşte simetric).
Demonstraţie (a) este evident.
(b) Pentru orice $f \in L^2(0,1)$, avem:

$$(T_K T_H f)(x) = \int_0^1 K(x,y)\,(T_H f)\,(y)dy =$$

$$= \int_0^1 \int_0^1 K(x,y)H(y,z)f(z)dzdy = \int_0^1 \left(\int_0^1 K(x,y)H(y,z)dy\right) f(z)dz =$$

$$= \int_0^1 G(x,z)f(z)dz = (T_G f)(x).$$

(c) Demonstraţia este o consecinţă imediată a inegalităţii dintre normele operatorului integral şi a nucleului său:

$$\| T_{K_n} - T_K \| \leq \| K_n - K \|_2 \rightarrow 0.$$

(d) Pentru orice $f, g \in L^2(0,1)$, avem:

$$< T_K f, g >= \int_0^1 (T_K f)(x)\overline{g(x)}dx = \int_0^1 \left(f(y) \int_0^1 K(x,y)\overline{g(x)}dx\right) dy =$$

$$= \int_0^1 f(y)\overline{\left(\int_0^1 \widetilde{K}(y,x)g(x)dx\right)}dy =< f, T_{\widetilde{K}} g > .$$

Proprietăţile de mai sunt adevărate şi în cazul în care intervalul $[0,1]$ (cu măsura Lebesgue) este înlocuit de un spaţiu cu măsură $\sigma-$ finită. În particular, putem considera operatori integrali pe $\mathbf{Z}$ şi $\mathbf{R}$.
Operatorii de convoluţie sunt atunci cazuri particulare de operatori integrali, considerând nuclee de forma $K(n,m) = \theta(n-m)$, $\forall n, m \in Z$ şi respectiv $K(x-y) = k(x-y)$, $\forall x, y \in R$.
Un alt caz particular remarcabil de operator integral este **operatorul Volterra**, pe care-l vom studia în continuare.

47. Definiţie (operatorul Volterra)
Un nucleu $K \in L^2([0,1] \times [0,1])$ se numeşte **nucleu de tip Volterra** dacă are proprietatea $K(x,y) = 0$, $\forall x < y$. Rezultă deci că operatorul integral asociat unui astfel de nucleu (numit **operator Volterra**) este definit prin:

$$(T_K f)(x) = \int_0^x K(x,y)f(y)dy\,,\ \forall f \in L^2(0,1).$$

Analogia cu teoria matricelor este evidentă: operatorii de tip Volterra sunt analogul operatorilor asociaţi matricelor inferior triunghiulare. Se ştie că dacă o matrice A este strict inferior triunghiulară, atunci ea este nilpotentă, adică există $m \in N$ astfel încât $A^m = O$. Vom demonstra în continuare o proprietate asemănătoare şi pentru operatorii Volterra definiţi de nuclee mărginite; o consecinţă va fi calculul spectrului unui astfel de operator.

48. Teoremă
Fie $K \in L^2([0,1] \times [0,1])$ un nucleu de tip Volterra mărginit, deci $\| K \|_\infty < \infty$; atunci, operatorul Volterra asociat, T_K are proprietăţile:
(a) $\lim\limits_{n\to\infty} (\| T_K^n \|)^{\frac{1}{n}} = 0$.
(b) $\sigma(T_K) = \{0\}$; un operator cu această proprietate se numeşte **cvasinilpotent**.
Demonstraţie Vom demonstra mai întâi că produsul a doi operatori de tip Volterra T_K şi T_H este un operator de acelaşi tip. Ţinând cont de cele demonstrate în propoziţia 46(b), este suficient să arătăm implicaţia:

$$K(x,y) = H(x,y) = 0\,, \forall x < y \Rightarrow G(x,y) = 0\,,\ \forall x < y,$$

unde, $G(x,y) = \int\limits_0^1 K(x,z)H(z,y)dz$.
Într-adevăr, dacă $x < y$, atunci orice $z \in [0,1]$ trebuie să verifice cel puţin una din inegalităţile: $x < z$ sau $z < y$; în primul caz, avem $K(x,z) = 0$, iar în al doilea $H(z,y) = 0$, deci oricum $G(x,y) = 0$. Dacă $x \geq y$, atunci:

$$G(x,y) = \int_y^x K(x,z)H(z,y).$$

Să presupunem acum că $H = K$ şi să notăm în acest caz

$$K^{[2]}(x,y) = G(x,y) = \int_0^1 K(x,z)K(z,y)dz,$$

şi în general pentru $n \in N$:

$$K^{[n]}(x,y) = \int_0^1 K(x,z)K^{[n-1]}(z,y)dz.$$

Pentru orice $0 \leq y \leq x \leq 1$, avem:

$$|K^{[2]}(x,y)| = |\int_y^x K(x,z)K(z,y)dz| \leq \| K \|_\infty^2 (x-y).$$

Prin inducţie rezultă că pentru orice $n \in N$ şi $y \leq x$, avem:

$$|K^{[n]}(x,y)| \leq \frac{\| K \|_\infty^n}{(n-1)!}(x-y)^{n-1} \leq \frac{\| K \|_\infty^n}{(n-1)!}.$$

Rezultă deci că:

$$(\| T_K^n \|)^{\frac{1}{n}} \le \left(\| K^{[n]} \|_\infty\right)^{\frac{1}{n}} \le \frac{\| K \|_\infty}{(n-1)!^{\frac{1}{n}}} \longrightarrow 0,$$

pentru $n \longrightarrow \infty$, ceea ce încheie demonstraţia.
(b) Reamintim că raza spectrală a unui operator T este, $\mathrm{r}(T) = \sup\{|\lambda| \,;\, \lambda \in \sigma(T)\}$ şi are formula razei spectrale: $\mathrm{r}(T) = \lim_{n\to\infty} \| T^n \|^{\frac{1}{n}}$. Din definiţia razei spectrale rezultă în mod evident că dacă $\mathrm{r}(T) = 0$, atunci $\sigma(T) = \{0\}$. Din cele demonstrate la punctul **(a)**, rezultă că $\mathrm{r}(T_K) = 0$, deci $\sigma(T_K) = \{0\}$.
Mai facem observaţia că afirmaţiile din teoremă sunt adevărate şi fără ipoteză de mărginire a nucleului; demonstraţia este însă considerabil mai dificilă ([10],p.93).
Un caz particular interesant de operator Volterra se obţine considerând nucleul $V(x,y) = 1$, dacă $y \le x$, şi 0 în rest. Operatorul asociat (numit şi **operatorul Volterra integral**) este:

$$(T_V f)(x) = \int_0^x f(y)dy\,,\ \forall f \in L^2(0,1).$$

Norma lui T_V este $\| T_V \| = \frac{1}{2\pi}$; (pentru demonstraţie: [10],p.300).

6.4 MF.06.4. Operatori normali

49. Introducere
Reamintim că un operator $T \in \mathcal{L}(H)$ se numeşte normal dacă verifică egalitatea $TT^\star = T^\star T$; în paragraful precedent am studiat o clasă importantă de operatori normali: operatorii de multiplicare.
În capitolul 2 (teorema 28) am văzut că principalul rezultat referitor la structura operatorilor din $\mathcal{L}(\mathbf{C^n})$ este:

Teorema spectrală pentru operatori normali pe C^n
Un operator $T \in \mathcal{L}(\mathbf{C^n})$ este diagonalizabil în sens geometric dacă şi numai dacă T este operator normal.

Generalizarea acestui rezultat la spaţii Hilbert infinit dimensionale este o problemă fundamentală a teoriei operatorilor; ea şi câteva consecinţe ale sale constituie subiectul acestui paragraf.
În prezentarea care urmează, analogia cu cazul finit dimensional este interesantă: trebuie remarcat ce rezultate finit dimensionale au un corespondent infinit dimensional şi ce anume se schimbă în totalitate.
Să revenim acum la enunţul teoremei spectrale pentru operatori normali

pe spaţii finit dimensionale. Operator diagonalizabil (în sens geometric) însemna, în acel caz, un operator $T \in \mathcal{L}(\mathbf{C^n})$ pentru care există o bază $\mathcal{B}$ ortonormală (a lui $\mathbf{C^n}$) formată din vectori proprii ai operatorului T , sau , echivalent, există un operator unitar U astfel încât $U^{-1}TU$ să fie operator diagonal; în această situaţie matricea lui T în baza $\mathcal{B}$ are formă diagonală, (pe diagonală fiind valorile proprii ale lui T), iar coloanele matricei lui U sunt vectorii din $\mathcal{B}$. Î n cazul infinit dimensional, noţiunile de valoare proprie şi vector propriu nu mai constituie instrumente la fel de puternice ca în cazul finit dimensional; am dat exemple în paragraful precedent de operatori (chiar normali) care nu au valori proprii. Deci o "formă diagonală" pentru operatori normali pe spaţii Hilbert infinit dimensionale este puţin probabilă (aceasta nu exclude posibilitatea unei "forme diagonale" pentru clase mai restrânse de operatori).
Vom reformula acum noţiunea de operator diagonal pe C^n .
Fie $D \in \mathcal{L}(\mathbf{C^n})$ un operator diagonal, deci matricea sa în baza canonică este:

$$\begin{pmatrix} \lambda_1 & 0 & . & . & . & 0 \\ 0 & \lambda_2 & 0 & . & . & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & . & . & . & . & \lambda_n \end{pmatrix},$$

iar $(Dx)_k = \lambda_k x_k$, $\forall x = (x_1, x_2, .., x_n) \in \mathbf{C^n}$.
Să considerăm spaţiul $\mathbf{C^n}$ ca fiind mulţimea tuturor funcţiilor

$$x : \{1, 2, .., n\} \to C \, , \, x(k) = x_k,$$

vectorul $x = (x_1, x_2, .., x_n)$ identificâdu-se cu valorile funcţiei x de mai sus.
Să considerăm funcţia

$$\phi : \{1, 2, .., n\} \to \mathbf{C} \, , \, \phi(k) = \lambda_k.$$

Atunci operatorul D poate fi identificat cu operatorul de multiplicare cu ϕ :

$$(Dx)(k) = \phi(k)x(k) = (M_\phi x)(k) \, , \, \forall x \in C^n \text{ şi } k \in \{1, 2, .., n\}.$$

Spaţiul cu măsură (Ω, μ) din definiţia generală a operatorilor de multiplicare este $\Omega = \{1, 2, .., n\}$ iar măsura μ este măsura de numărare.
Cu aceste precizări, enunţul teoremei spectrale pentru operatori normali pe spaţii finit dimensionale, devine:

Teoremă
Un operator $T \in \mathcal{L}(C^n)$ este unitar-echivalent cu un operator de multiplicare M_ϕ dacă şi numai dacă T este operator normal.
Această formulare (în care **operator diagonalizabil în sens geometric** înseamnă **operator unitar- echivalent cu un operator de multiplicare**) are un analog infinit dimensional.

Reamintim că în paragraful precedent am definit şi studiat operatorii de multiplicare: dacă (Ω, μ) este un spaţiu cu măsură şi dacă $\phi \in L^\infty(\Omega, \mu)$, atunci operatorul de multiplicare cu funcţia ϕ este:

$$M_\phi : L^2(\Omega, \mu) \to L^2(\Omega, \mu)\,,\ (M_\phi f)(t) = \phi(t) f(t)\,,\ \forall t \in \Omega.$$

Am demonstrat (este de altfel evident) că operatorii de multiplicare sunt normali; reciproca acestei afirmaţii ("modulo unitar-echivalenţă") este varianta infinit dimensională a teoremei spectrale pentru operatori normali (din cazul finit dimensional).

50. Teorema spectrală pentru operatori normali pe spaţii Hilbert infinit dimensionale
Fie H un spaţiu Hilbert infinit dimensional (separabil) şi fie $T \in \mathcal{L}(H)$ un operator normal.
Atunci există un spaţiu cu măsură boreliană regulată finită (Ω, μ), un izomorfism de spaţii Hilbert (operator unitar) $U : H \to L^2(\Omega, \mu)$ şi o funcţie $\phi \in L^\infty(\Omega, \mu)$ astfel încât $T = U^{-1} M_\phi U$.

Într-o formulare concisă, teorema afirmă că **un operator este normal dacă şi numai dacă este unitar-echivalent cu un operator de multiplicare**.
Pentru alte formulări (echivalente) ale acestui rezultat cât şi pentru demonstraţie, recomandăm [10],p.61; [6],p.911; [5],p.93; [20],p.71.
Un caz particular studiat deja al acestui rezultat este unitar-echivalenţa dintre operatorii de convoluţie (care sunt normali) şi cei de multiplicare; a se vedea teoremele 40 şi 45.

Consecinţele imediate (dar remarcabile) ale acestei teoreme sunt proprietăţile operatorilor normali deduse din proprietăţile corespunzătoare ale operatorilor de multiplicare; avem deci:

51. Teoremă (proprietăţile operatorilor normali)
Fie $T \in \mathcal{L}(H)$ un operator normal; atunci:
(a) $\sigma(T^\star) = \overline{\sigma(T)} = \{\overline{\lambda}\,;\ \lambda \in \sigma(T)\}$.
(b) T este autoadjunct $\Leftrightarrow$ $\sigma(T) \subset R$.
(c) T este pozitiv $\Leftrightarrow$ $\sigma(T) \subset [0, \infty)$.
(d) T este unitar $\Leftrightarrow$ $\sigma(T) \subseteq \mathcal{S}^1$.
(e) T este proiector $\Leftrightarrow$ $\sigma(T) \subseteq \{0, 1\}$.
(f) $\mathrm{r}(T) = \| T \|$.
(g) $\| Tx \| = \| T^\star x \|\,,\ \forall x \in H$; se poate demonstra că această proprietate caracterizează operatorii normali.
Este interesant acum să ne reamintim analogia dintre numere complexe şi operatori de la sfârşitul paragrafului 1 (acest capitol). Este clar (din teorema de mai sus) că analogia este mai naturală dacă înlocuim în tabelul

respectiv "operator" cu "operator normal".
În general, dacă T nu este un operator normal, proprietăţile de mai sus nu sunt adevărate. De exemplu, operatorul integral Volterra T_V de la sfârşitul paragrafului precedent (teorema 48) are spectrul format numai din numărul 0, dar nu este autoadjunct; în schimb, dacă un operator normal are spectrul $\{0\}$, atunci, din proprietatea **(f)** de mai sus rezultă că el este operatorul identic nul.

Ca şi în cazul finit dimensional, o consecinţă importantă a teoremei spectrale este posibilitatea construirii unui **"calcul funcţional"** pentru operatori normali. Prezentăm în continuare această construcţie.

52. Definiţie
Dacă $T \in \mathcal{L}(H)$ este un operator arbitrar (fixat) şi $p \in \mathbf{C}[X]$ este un polinom, $p(z) = \sum\limits_{k=0}^{n} a_k z^k$ atunci, este natural să definim operatorul $p(T) = \sum\limits_{k=0}^{n} a_k T^k$. Am construit în felul acesta o aplicaţie (numită **"calculul funcţional polinomial al operatorului** T"):

$$\mathbf{C}[X] \ni p \to p(T) \in \mathcal{L}(H),$$

cu proprietăţile:
$(\alpha p + \beta q)(T) = \alpha p(T) + \beta q(T)$ (liniară) şi
$(pq)(T) = p(T)q(T)$ (multiplicativă),
pentru orice $\alpha, \beta \in \mathbf{C}$ şi $p, q \in \mathbf{C}[X]$; demonstraţiile sunt imediate.
Din ultima egalitate rezultă că operatorii $p(T)$ şi $q(T)$ comută.
Tot cu metode elementare, şi tot pentru un operator arbitrar, putem defini $f(T)$ şi pentru anumite funcţii raţionale.
Pentru aceasta, fie $\lambda \notin \sigma(T)$ şi fie funcţia (raţională) $g(z) = \frac{1}{\lambda - z}$; o definiţie naturală pentru operatorul $g(T)$ este $g(T) = (\lambda I - T)^{-1}$. Să observăm că definiţia este corectă deoarece operatorul $\lambda I - T$ este inversabil (λ nefiind în spectrul lui T). Să mai observăm că doi operatori de forma $(\lambda I - T)^{-1}$ şi $(\nu I - T)^{-1}$ comută între ei deoarece operatorii $\lambda I - T$ şi $\nu I - T$ comută.
Mai general, fie $q(z) = \alpha(\lambda_1 - z)(\lambda_2 - z)...(\lambda_n - z)$; observăm că putem defini operatorul:

$$\frac{1}{q}(T) = \frac{1}{\alpha}(\lambda_1 I - T)^{-1}(\lambda_2 I - T)^{-1}...(\lambda_n I - T)^{-1},$$

dacă şi numai dacă $\lambda_k \notin \sigma(T)$, $\forall k \in \{1, 2, ..n\}$. Este evident că în acest caz avem:

$$\frac{1}{q}(T) = (q(T))^{-1}.$$

În cazul general, fie $p, q \in \mathbf{C}[X]$ şi fie $f(z) = \frac{p(z)}{q(z)}$, o fracţie ireductibilă. Dacă polinomul q nu se anulează pe spectrul lui T, (sau,

echivalent , funcţia f este definită pe întreg spectrul lui T), atunci definim operatorul $f(T) = p(T)[q(T)]^{-1}$. Fie $\mathcal{R}(T)$ mulţimea funcţiilor raţionale definite pe spectrul operatorului T .
Atunci aplicaţia (numită **"calculul funcţional raţional** al lui T"):

$$\mathcal{R}(T) \ni f \to f(T) \in \mathcal{L}(H),$$

prelungeşte calculul funcţional polinomial cu păstrarea liniarităţii şi a multiplicativităţii. De fapt, aplicaţia astfel construită este un morfism de algebre. Calculul funcţional astfel definit are următoarea proprietate de **"transformare a spectrului"**.

53. Propoziţie
Fie $T \in \mathcal{L}(H)$; atunci, pentru orice $f \in \mathcal{R}(T)$, avem:

$$\sigma(f(T)) = \{f(\lambda)\,;\, \lambda \in \sigma(T)\} = f(\sigma(T)).$$

Demonstraţie Demonstrăm mai întâi incluziunea: $f(\sigma(T)) \subseteq \sigma(f(T))$. Pentru aceasta, fie $\nu \in f(\sigma(T))$; există deci $\lambda \in \sigma(T)$ astfel încât $\nu = f(\lambda)$. Fie funcţia

$$g(z) = \frac{f(\lambda) - f(z)}{\lambda - z}.$$

Demonstrăm că $g \in \mathcal{R}(T)$. Deoarece $f \in \mathcal{R}(T)$, singurul punct din $\sigma(T)$ în care funcţia g ar putea să nu fie definită este λ . Dacă $f = \frac{p}{q}$, cu $p, q \in C[X]$, atunci $q(\lambda) \neq 0$; avem:

$$g(z) = \frac{\frac{p(\lambda)}{q(\lambda)} - \frac{p(z)}{q(z)}}{\lambda - z} = \frac{p(\lambda)q(z) - q(\lambda)p(z)}{q(\lambda)q(z)(\lambda - z)}.$$

Polinomul $s(z) = p(\lambda)q(z) - q(\lambda)p(z)$ se anulează în $z = \lambda$, deci există un polinom $r \in C[X]$ astfel încât $s(z) = (\lambda - z)r(z)$; rezultă deci că funcţia g este

$$g(z) = \frac{r(z)}{q(\lambda)q(z)},$$

adică $g \in \mathcal{R}(T)$, deci are sens $g(T)$. Să presupunem prin absurd că $\nu = f(\lambda) \notin \sigma(f(T))$, deci există $[f(\lambda)I - f(T)]^{-1}$. Din relaţia

$$f(\lambda) - f(z) = (\lambda - z)g(z)$$

şi din multiplicativitatea calculului funcţional, rezultă:

$$f(\lambda)I - f(T) = (\lambda I - T)g(T).$$

Înmulţind ultima egalitate cu $[f(\lambda)I - f(T)]^{-1}$, obţinem:

$$(\lambda I - T)g(T)[f(\lambda I - f(T)]^{-1} = I,$$

şi, deoarece operatorii de mai sus comută , rezultă că $\lambda I - T$ este operator inversabil, contradicţie cu $\lambda \in \sigma(T)$.
Demonstrăm acum incluziunea inversă: $\sigma(f(T)) \subseteq f(\sigma(T))$.
Fie $\nu \in \sigma(f(T))$ şi presupunem prin absurd că $\nu \notin f(\sigma(T))$. Rezultă atunci că funcţia

$$h(z) = \frac{1}{\nu - f(z)}$$

este în $\mathcal{R}(T)$ şi din egalitatea $[\nu - f(z)]h(z) = 1$ rezultă

$$[\nu I - f(T)]h(T) = I,$$

adică $\nu I - f(T)$ este operator inversabil; contradicţie cu $\nu \in \sigma(f(T))$.

Prelungirea calculului funcţional (raţional) şi la alte clase de funcţii cu păstrarea liniarităţii, multiplicativităţii şi a proprietăţii de transformare a spectrului este o problemă importantă în teoria operatorilor. Există şi alte funcţii pentru care se pot da definiţii elementare.
De exemplu, dacă $f(z) = \sum\limits_{n=0}^{\infty} a_n z^n$, $\forall z \in \mathbf{C}$ este o funcţie întreagă atunci operatorul $f(T) = \sum\limits_{n=0}^{\infty} a_n T^n$ se poate defini pentru pentru orice $T \in \mathcal{L}(H)$ (deoarece seria de operatori converge) şi sunt păstrate proprietăţile menţionate mai sus.
Un alt exemplu este funcţia $f(z) = \overline{z}$; în acest caz, o definiţie naturală ar fi $f(T) = T^\star$. De aici rezultă că dacă $g(z) = |z|^2$, atunci $g(T) = T^\star T$ (sau $TT^\star$?) ; dacă operatorul T ar fi normal, definiţia nu ar fi ambiguă. Să alegem de exemplu $g(T) = T^\star T$, şi să luăm $T = V$ operatorul de translaţie unilaterală pe $\ell^2(\mathbf{N})$; proprietatea de transformare a spectrului nu mai este adevărată ; într-adevăr, deoarece $g(V) = V^\star V = I$, atunci

$$\sigma(g(V)) = \sigma(I) = \{1\} \text{ dar}$$

$$g(\sigma(V)) = g(\{\lambda \in \mathbf{C}\,;\, |\lambda| \leq 1\}) = [0, 1].$$

Operatorii care admit un calcul funcţional suficient de general şi cu proprietăţi remarcabile sunt operatorii normali. Pentru aceasta, definim calculul funcţional mai întâi pentru operatorii de multiplicare (ceea ce se face într-un mod natural şi simplu) şi apoi vom transfera calculul funcţional astfel construit la operatori normali arbitrari folosind unitar-echivalenţa din teorema spectrală.

54. Definiţie (calculul funcţional mărginit pentru operatorii de multiplicare)

Fie $\phi \in L^\infty(\Omega, \mu)$ şi fie M_ϕ operatorul de multiplicare cu ϕ definit pe spaţiul $L^2(\Omega, \mu)$.
Considerăm restricţia măsurii Lebesgue din plan la compactul $\sigma(M_\phi) = \text{essran}(\phi)$. Fie $F \in L^\infty(\sigma(M_\phi))$; în particular, deoarece $\sigma(M_\phi)$ este compact, F poate fi o funcţie continuă.
Deoarece $\sigma(M_\phi) = \text{essran}(\phi)$, rezultă că funcţia compunere, $F \circ \phi$ există (de fapt ea este definită a.p.t.; dacă funcţiile ϕ şi F ar fi continue, atunci $F \circ \phi$ ar fi definită peste tot şi ar fi continuă) şi este în $L^\infty(\Omega, \mu)$. Există deci operatorul de multiplicare $M_{F\circ\phi} \in \mathcal{L}(L^2(\Omega, \mu))$, ceea ce ne permite să definim operatorul $F(M_\phi) = M_{F\circ\phi}$. Aplicaţia

$$L^\infty(\sigma(M_\phi)) \ni F \to F(M_\phi) \in \mathcal{L}(L^2(\Omega, \mu))$$

se numeşte **calculul funcţional mărginit** (în sensul că F este funcţie esenţial mărginită) al operatorului M_ϕ.
Se observă că operaţia de aplicare a lui F asupra operatorului M_ϕ, înseamnă de fapt compunerea lui F cu ϕ.
Este uşor de demonstrat că aplicaţia de mai sus prelungeşte calculul funcţional cu funcţii raţionale al operatorului M_ϕ. Sunt păstrate de asemenea şi proprietăţle uzuale, inclusiv proprietatea de transformare a spectrului.

55. Teoremă
Cu notaţiile de mai sus, avem:
(a) $(\alpha F + \beta G)(M_\phi) = \alpha F(M_\phi) + \beta G(M_\phi)$,
(b) $(FG)(M_\phi) = [F(M_\phi)][G(M_\phi)]$,
(c) $\overline{F}(M_\phi) = [F(M_\phi)]^\star$,
(d) $\sigma(F(M_\phi)) = F(\sigma(M_\phi))$,
(e) Operatorii $F(M_\phi)$ şi $G(M_\phi)$ comută,
pentru orice $\alpha, \beta \in C$ şi F, G funcţii esenţial mărginite pe spectrul operatorului M_ϕ.
Demonstraţie Verificările **(a)**,**(b)** şi **(c)** sunt imediate; demonstrăm, de exemplu, **(b)**:

$$(FG)(M_\phi) = M_{(FG)\circ\phi} = M_{(F\circ\phi)(G\circ\phi)} = M_{F\circ\phi}M_{G\circ\phi} = F(M_\phi)G(M_\phi).$$

(d) Vom face demonstraţia în ipoteza că funcţia F este continuă. Propunem cititorului familiarizat cu raţionamentele specifice teoriei abstracte a măsurii să refacă demonstraţia pentru o funcţie F esenţial mărginită, ([10],p.60).
Enunţul este echivalent cu $\text{essran}(F \circ \phi) = F(\text{essran}(\phi))$; se observă că dacă şi funcţia ϕ este continuă, atunci egalitatea devine banală deoarece amândoi membri sunt egali cu închiderea imaginii funcţiei compuse $F \circ \phi$, adică $\overline{(F(\phi(\Omega))}$.
Considerăm acum cazul general şi demonstrăm incluziunea $F(\sigma(M_\phi)) \subseteq$ $\subseteq \sigma(F(M_\phi))$. Fie $\nu = F(\lambda) \in F(\sigma(M_\phi))$, cu $\lambda \in \sigma(M_\phi))$. Fie E o vecinătate a lui $F(\lambda)$; pentru a demonstra că $F(\lambda) \in \sigma(F(M_\phi))$ va trebui să

demonstrăm că măsura mulţimii $(F\circ\phi)^{-1}(E)$ este strict pozitivă (nenulă). Dar $(F\circ\phi)^{-1}(E)=\phi^{-1}(F^{-1}(E))$. Deoarece E este vecinătate a lui $F(\lambda)$ şi deoarece funcţia F este continuă, rezultă că $F^{-1}(E)$ este vecinătate a lui λ şi deci, din definiţia imaginii esenţiale a lui ϕ, rezultă că $\phi^{-1}(F^{-1}(E))$ are măsură strict pozitivă. Incluziunea inversă o demonstrăm prin trecere la complementare:

$$C-F(\sigma(M_\phi))\subseteq C-\sigma(F(M_\phi))\,.$$

Fie deci $\lambda\notin F(\sigma(M_\phi))$. Mulţimea $\sigma(M_\phi))$ este compactă şi cum F este continuă, rezultă că mulţimea $F(\sigma(M_\phi))$ este compactă. Deoarece (din ipoteză) $\lambda\notin F(\sigma(M_\phi))$, atunci există o vecinătate E a lui λ astfel încât $E\bigcap F(\sigma(M_\phi))=\emptyset$. Luând imaginile inverse prin funcţia F ale mulţimilor din egalitatea precedentă, obţinem

$$F^{-1}(E)\bigcap\sigma(M_\phi)=\emptyset\,.$$

Luând acum imaginile inverse prin funcţia ϕ, obţinem

$$\phi^{-1}(F^{-1}(E))\bigcap\phi^{-1}(\sigma(M_\phi))=\emptyset\,.$$

Dar, din definiţia imaginii esenţiale, măsura mulţimii $\Omega-\phi^{-1}(\sigma(M_\phi))$ este nulă (dacă funcţia ϕ ar fi continuă, atunci această mulţime ar fi vidă). Rezultă că şi mulţimea $(F\circ\phi^{-1})(E)$ are măsura nulă şi deci $\lambda\notin\sigma(F(M_\phi))$.

Considerăm în continuare două exemple.

(i) Fie $\phi(t)=t$, $\forall t\in[0,1]$ şi fie

$$M_\phi : L^2(0,1)\to L^2(0,1),\ (M_\phi f)(t)=tf(t).$$

Atunci, $\sigma(M_\phi)=[0,1]$ şi dacă $F\in L^\infty(0,1)$, rezultă $F(M_\phi)=M_F$.

(ii) Fie acum $\psi(t)=\frac{t+i}{it+1}$, $\forall t\in\mathbf{R}$ şi fie

$$M_\psi : L^2(\mathbf{R})\to L^2(\mathbf{R}),\ (M_\psi f)(t)=\frac{t+i}{it+1}f(t).$$

Atunci $\sigma(M_\psi)=\mathcal{S}^1$ şi pentru orice funcţie $F\in L^\infty(\mathcal{S}^1)$, avem:

$$\left(F(M_\psi)f\right)(t)=F\left(\frac{t+i}{it+1}\right)f(t),\ \forall f\in L^2(\mathbf{R}).$$

Calculul funcţional mărginit pentru un operator normal arbitrar se construieşte folosind unitar-echivalenţa din teorema spectrală.

56. Definiţie (calculul funcţional mărginit pentru operatori normali)

Fie $T\in\mathcal{L}(H)$ un operator normal şi fie M_ϕ operatorul de multiplicare unitar-echivalent cu el: $T = U^{-1}M_\phi U$. Dacă F este o funcţie esenţial mărginită pe spectrul operatorului T, (considerat ca spaţiu cu măsură cu restricţia măsurii Lebesgue din plan), atunci definim $F(T) = U^{-1}F(M_\phi)U$. Din teorema de mai sus

$$L^\infty(\sigma(T)) \ni F \to F(T) \in \mathcal{L}(H)$$

are proprietăţile:
(a) Liniară şi multiplicativă:

$$(\alpha F + \beta G)(T) = \alpha F(T) + \beta G(T),$$

$$(FG)(T) = F(T)G(T),$$

$\forall \alpha, \beta \in \mathbf{C}$, $\forall F, G \in L^\infty(\sigma(T))$; din multiplicativitate rezultă că operatorii de forma $F(T)$ comută între ei.
(b) $\overline{F}(T) = [F(T)]^\star$; de aici rezultă că toţi operatorii de forma $F(T)$ sunt normali.
(c) $\sigma(F(T)) = F(\sigma(T))$.

Să considerăm ca exemplu operatorul de convoluţie pe spaţiul $\ell^2(\mathbf{Z})$ (operatorul de convoluţie este operator normal).
Fie $\theta \in \ell^2(\mathbf{Z})$ astfel încât $\mathcal{F}^{-1}\theta \in L^\infty(\mathcal{S}^1)$; Operatorul C_θ este unitar-echivalent cu operatorul de multiplicare cu $\mathcal{F}^{-1}\theta$, mai precis $C_\theta = \mathcal{F}^{-1}M_{\mathcal{F}^{-1}\theta}\mathcal{F}$.
Pentru orice funcţie $F \in L^\infty(\sigma(C_\theta)) = L^\infty(\text{essran}(\mathcal{F}^{-1}\theta))$, avem $F(C_\theta) = \mathcal{F}^{-1}M_{F\circ(\mathcal{F}^{-1}\theta)}\mathcal{F}$.
În particular, dacă $\theta = \sigma_1$, atunci $C_{\sigma_1} = W$ (operatorul de translaţie bilateral) şi $W = \mathcal{F}^{-1}M_{\omega_1}\mathcal{F}$, unde,
$\omega_1(e^{it}) = e^{it}$. Reamintim că $\sigma(W) = \mathcal{S}^1$. Deoarece ω_1 este aplicaţia identică pe $\mathcal{S}^1$, pentru orice funcţie $F \in L^\infty(\mathcal{S}^1)$, avem:

$$F(W) = \mathcal{F}^{-1}M_F\mathcal{F}.$$

Mai general, dacă $n \in Z$, avem $F(W^n) = \mathcal{F}^{-1}M_{F\circ\omega_n}\mathcal{F}$, unde,
$\omega_n(e^{it}) = e^{int}$.

Aplicaţiile calculului funcţional sunt numeroase; indicăm în continuare câteva. Prima este existenţa rădăcinii pătrate pozitive pentru operatori pozitivi; comparaţia cu cazul finit dimensional este interesantă (a se vedea capitolul 3).

57. Teoremă (rădăcina pătrată pozitivă)
Fie $P \in \mathcal{L}(H)$ un operator pozitiv. Atunci există şi este unic un operator pozitiv $Q \in \mathcal{L}(H)$ astfel încât $Q^2 = P$; operatorul Q se numeşte rădăcina pătrată pozitivă a lui P şi se notează $\sqrt{P}$.
Demonstraţie Funcţia radical $f(t) = \sqrt{t}$ este o funcţie continuă pe spectrul operatorului P (orice operator pozitiv are spectrul în $[0, \infty)$), deci

putem defini operatorul $Q = f(P) = \sqrt{P}$. Din multiplicativitatea calculului funcţional, rezultă :

$$Q^2 = [f(P)]^2 = (f^2)(P) = \mathrm{id}(P) = P,$$

unde, am notat cu $\mathrm{id}(t) = t$ funcţia identică.
Pentru unicitate, este suficient să observăm că un operator de multiplicare pozitiv, M_ϕ, are o unică rădăciă pătrată pozitivă, $M_{\sqrt{\phi}}$.
De exemplu, dacă $\phi(t) = t\,, \forall t \in [0,1]$, atunci operatorul M_ϕ este operator pozitiv pe $L^2(0,1)$ şi

$$(\sqrt{M_\phi} f)(t) = \sqrt{t} f(t) = (M_{\sqrt{\phi}} f)(t)\,,\ \forall t \in [0,1].$$

58. Consecinţă
Fie $T \in \mathcal{L}(H)$ un operator arbitrar. Atunci T este pozitiv dacă şi numai dacă există $S \in \mathcal{L}(H)$ astfel încât $T = S^\star S$
Demonstraţie Dacă T este operator pozitiv, atunci luăm $S = \sqrt{T}$. Reciproc, orice operator de forma $S^\star S$ este pozitiv (evident).

Altă consecinţă a calculului funcţional este descompunerea polară pentru operatorii inversabili şi pentru cei normali pe spaţii Hilbert infinit dimensionale; orice număr complex nenul z admite o unică
descompunere $z = ru$ cu $r > 0$ şi $|u| = 1$. Dacă $T \in \mathcal{L}(C^n)$, atunci T admite o descompunere de forma $T = UP$, cu U operator unitar şi P pozitiv. În general, această descompunere nu este unică decât în anumite condiţii suplimentare (de exemplu dacă T este inversabil).

59. Teoremă (descompunerea polară)
Fie $T \in \mathcal{L}(H)$ un operator arbitrar.
(a) Dacă T este inversabil , atunci există $U \in \mathcal{L}(H)$ operator unitar şi $P \in \mathcal{L}(H)$ operator pozitiv astfel încât $T = UP$. În plus, această descompunere (polară) este unică.
(b) Dacă T este normal, atunci există o descompunere polară (nu neapărat unică) $T = UP$ cu U unitar şi P pozitiv; în plus, operatorii U, P, T comută între ei.
Demonstraţie (a) Fie $P = \sqrt{T^\star T}$. Operatorul T fiind inversabil, rezultă că $T^\star T$ este şi el inversabil, deci $0 \notin \sigma(T^\star T)$. Din teorema de transformare a spectrului rezultă că $0 \notin \sigma(\sqrt{T^\star T})$ deci operatorul P este inversabil. Fie $U = TP^{-1}$. Atunci U este inversabil (ceea ce este evident); U este unitar:

$$U^\star U = P^{-1} T^\star T P^{-1} = P^{-1} P^2 P^{-1} = I,$$

Unicitatea rezultă din construcţie şi din unicitatea rădăcinii pătrate pozitive.

(b) Fie funcţiile

$$p(z) = |z|\,,\ \forall z \in C \text{ şi } u(z) = \frac{z}{|z|}\,,\ \forall z \neq 0 \text{ şi } u(0) = 1.$$

Atunci p şi u sunt funcţii mărginite pe spectrul lui T. Aici se poate constata necesitatea unui calcul funcţional şi cu alte funcţii decât continue (p este continuă dar u nu este continuă). Fie $P = p(T)$ şi $U = u(T)$. Deoarece funcţia p ia valori pozitive, din teorema de transformare a spectrului rezultă că operatorul P are spectrul în $[0, \infty)$. Cum P este şi operator normal, rezultă că P este operator pozitiv. Deoarece $u(z)\overline{u(z)} = 1$, $\forall z \in$ $\mathbf{C}$, din multiplicativitatea calculului funcţional rezultă $UU^\star = U^\star U = I$, deci U este unitar. Din identitatea $u(z)p(z) = z$, $\forall z \in \mathbf{C}$, rezultă (folosind iarăşi multiplicativitatea calculului funcţional) $T = UP$.

Continuăm cu aplicaţii ale calculului funcţional şi ale formulei de descompunere polară. Orice număr real şi strict pozitiv t se poate scrie sub forma $t = e^s$, cu $s \in \mathbf{R}$; de asemenea, orice număr complex λ cu $|\lambda| = 1$, se poate scrie sub forma $\lambda = e^{i\tau}$, cu $\tau \in \mathbf{R}$. Pentru operatori liniari şi continui pe un spaţiu Hilbert, avem:

60. Propoziţie
(i) Pentru orice operator pozitiv şi inversabil $P \in \mathcal{L}(H)$, există un operator autoadjunct $S \in \mathcal{L}(H)$ astfel încât $P = \exp(S)$.
(ii) Pentru orice operator unitar $U \in \mathcal{L}(H)$, există un operator autoadjunct $A \in \mathcal{L}(H)$ astfel încât $U = \exp(iA)$.
Demonstraţie (i) Deoarece operatorul P este pozitiv şi inversabil, rezultă că $\sigma(P) \subset (0, \infty)$; rezultă deci că funcţia (continuă) logaritm natural, ln este definită pe spectrul operatorului P şi deci putem defini operatorul $S = \ln(P)$; deoarece $e^{\ln(x)} = x$, $\forall x > 0$, avem egalitatea:

$$\exp(S) = \exp(\ln(P)) = P.$$

Din proprietăţile calculului funcţional rezultă că S este operator normal, iar din teorema de transformare a spectrului rezultă incluziunea: $\sigma(S) = \ln(\sigma(P)) \subset \ln((0, \infty)) = \mathbf{R}$. Din teorema 51(b) rezultă că S este operator autoadjunct.
(ii) Deoarece U este operator unitar, rezultă că spectrul său este inclus în cercul unitate: $\sigma(U) \subseteq \mathcal{S}^1$. Fie $f \in L^\infty(\mathcal{S}^1)$ astfel încât $\exp(if(\lambda)) = \lambda$, $\forall \lambda \in \mathcal{S}^1$. Facem menţiunea că o astfel de funcţie există, ea putând fi, de exemplu, una din ramurile argumentului (care nu este continuă, dar este mărginită); dacă spectrul lui U nu este egal cu întreg cercul unitate, atunci există chiar funcţii continue cu proprietatea $\exp(if(\lambda)) = \lambda$, $\forall \lambda \in \sigma(U)$. Definim operatorul $A = f(U)$; din proprietăţile calculului funcţional rezultă că A este operator autoadjunct şi $\exp(iA) = \exp(if(U)) = U$.

61. Teoremă
(i) Orice operator inversabil $T \in \mathcal{L}(H)$ se poate scrie ca un produs de două exponenţiale, mai precis, există doi operatori autoadjuncţi $S, A \in \mathcal{L}(H)$ astfel încât $T = \exp(iA)\exp(S)$.
(ii) Mulţimea $\mathcal{G}$ a operatorilor inversabili din $\mathcal{L}(H)$ este conexă prin arce, adică pentru orice operator inversabil T, există o aplicaţie continuă $\gamma : [0,1] \to \mathcal{G}$ astfel încât $\gamma(0) = I$ şi $\gamma(1) = T$.
Demonstraţie (i) Fie $T \in \mathcal{L}(H)$ un operator inversabil şi fie, conform teoremei 59(a), descompunerea sa polară (unică): $T = UP$, unde, U este operator unitar şi P este operator pozitiv şi inversabil. Conform propoziţiei anterioare, există doi operatori autoadjuncţi $A, S \in \mathcal{L}(H)$ astfel încât $U = \exp(iA)$ şi $P = \exp(S)$, deci $T = \exp(iA)\exp(S)$.
Menţionăm că există operatori inversabili T care nu se pot scrie sub forma unei singure exponenţiale. Pe spaţii finit dimensionale, această proprietate este totuşi adevărată.
(ii) Fie $T \in \mathcal{G}$ şi fie $A, S \in \mathcal{L}(H)$, operatori autoadjuncţi astfel încât $T = \exp(iA)\exp(S)$. Fie

$$\gamma : [0,1] \to \mathcal{G},\ \gamma(t) = \exp(itA)\exp(tS).$$

Este uşor de arătat că γ satisface condiţiile cerute.

Capitolul 7

MF.07. Aplicaţii în teoria sistemelor

Cuvinte cheie

rezoluţie, sistem cauzal, sistem invariant în timp, spaţiul stărilor, controlabilitate, observabilitate.

Într-o formulare generală, un sistem este o aplicaţie între două spaţii de semnale: intrări (comenzi) şi ieşiri (răspunsuri); modelul matematic pentru semnale sunt funcţiile. Desigur, pentru a obţine rezultate interesante, sunt necesare unele condiţii restrictive. Un caz important este cel al sistemelor liniare: aici semnalele sunt elemente ale unor spaţii vectoriale, iar sistemul este o aplicaţie liniară. O altă proprietate remarcabilă este continuitatea; modelul matematic uzual pentru sistem este atunci acela al unui operator liniar şi continuu între două spaţii Banach. În acest capitol ne propunem să prezentăm câteva noţiuni din teoria sistemelor care se modelează în mod natural folosind conceptele şi rezultatele expuse în capitolele precedente. Prin sistem vom înţelege în continuare un operator liniar şi continuu pe un spaţiu Hilbert. Elementele acestui spaţiu (de obicei funcţii de timp) vor fi intrările şi ieşirile sistemului. Aşa cum am mai spus, o parte din proprietăţile intuitive ale unui sistem îşi găsesc imediat un corespondent matematic: liniaritate, continuitate. De asemenea, metodele folosite pentru studiul sistemelor sunt în mod natural rezultate de analiză funcţională. Pe lângă acestea, există şi unele constrângeri fizice, cât şi unele metode de studiu tipice teoriei sistemelor. Dintre acestea amintim cauzalitatea şi invarianţa în timp, iar ca metodă de studiu descompunerea în spaţii de stări a unui sistem. Prezentarea unor modele matematice pentru aceste noţiuni constituie obiectul acestui capitol. Pentru aprofundarea cunoştinţelor privind modelele matematice ale teoriei sistemelor, recomandăm următoarele lucrări: [F01], [S01], [D02], [B01], [O01], [S01], [S02].

7.1 MF.07.1. Sisteme cauzale

Intuitiv, un sistem este cauzal dacă ieşirea la orice moment (fixat) depinde numai de valorile intrării la momente anterioare.
De exemplu să considerăm spaţiul Hilbert $\ell^2(Z)$ şi sistemul (operatorul) de translaţie bilaterală: $(Wx)(n) = x(n-1)\,,\, \forall n \in \mathbf{Z}$. Evident că W satisface condiţia (intuitivă) de mai sus: ieşirea la momentul n este egală cu intrarea la momentul $n-1$. Să considerăm acum adjunctul (care coincide aici cu inversul) lui W, care este $(W^\star x)(n) = x(n+1)$. Evident, sistemul $W^\star$ nu este cauzal. El are chiar o proprietate duală cauzalităţii : ieşirea la un moment dat depinde numai de valorile intrării la momente posterioare; un astfel de sistem se numeşte anticauzal. Cadrul care permite o definţie pentru cauzalitate este spaţiul Hilbert cu **rezoluţie**.

1.Definiţie
Fie $(H, <,>)$ un spaţiu Hilbert (ca de obicei separabil şi complex). Reamintim că un operator $P \in \mathcal{L}(H)$ se numeşte proiector dacă $P^2 = P$. Dacă P şi Q sunt proiectori, atunci, prin definiţie, $P \leq Q$ dacă $P(H) \subseteq Q(H)$ (a se vedea paragraful 2,cap.5). Fie $\mathcal{T}$ o mulţime total ordonată având t_o şi t_∞ cel mai mic şi respectiv cel mai mare element. Prin **rezoluţie a identităţii** pe spaţiul H se înţelege orice familie de proiectori $\mathcal{P} = (P_t)_{t\in\mathcal{T}}$ cu proprietăţile:
(i) $P_t \leq P_s\,,\, \forall t \leq s, t, s \in \mathcal{T}$.
(ii) $P_{t_o} = O$ şi $P_{t_\infty} = I$.
(iii) Dacă $P_{t_n} \in \mathcal{P}$ astfel încât $\lim\limits_{n\to\infty} P_{t_n}x = Px\,,\, \forall x \in H$, atunci $P \in \mathcal{P}$.
Perechea $(H, \mathcal{P})$ se numeşte **spaţiu Hilbert cu rezoluţie**. Evident, pe acelaşi spaţiu Hilbert se pot defini mai multe rezoluţii. Interpretarea intuitivă a definiţiei este : mulţimea $\mathcal{T}$ este timpul, iar dacă $x \in H$, atunci $P_t x$ este partea (eşantionul) lui x de până la momentul t, iar $(I - P_t)x$ este partea lui x de după momentul t.

Vom introduce în continuare rezoluţiile canonice pe câteva spaţii Hilbert uzuale.

2.Exemple
(a) Pe spaţiul C^n, rezoluţia canonică este definită de mulţimea de proiectori $\{P_k\,;\, k = 0, 1, 2.., n\}$, unde $P_o = O$ şi $(P_k x)(m) = x(m)$ dacă $m \leq k$ şi 0 dacă $m > k$. Evident, în acest caz $\mathcal{T} = \{0, 1, 2, .., n\}$.
(b) Să considerăm acum spaţiul Hilbert $\ell^2(Z)$.
Fie $\mathcal{T} = \{-\infty\} \cup Z \cup \{\infty\}$, $P_{-\infty} = O\,,\, P_\infty = I$ şi pentru orice $n \in Z$ definim proiectorul $(P_n x)(m) = x(m)\,$, dacă $m \leq n$ şi 0 dacă $m > n$. Rezoluţia $\mathcal{P} = (P_n)_{n\in\mathcal{T}}$ este rezoluţia canonică pe $\ell^2(Z)$.

(c) Pe spaţiul $\ell^2(N)$ rezoluţia canonică se defineşte analog.
$\mathcal{T} = N \cup \{\infty\}$, $P_o = O$, $P_\infty = I$, iar P_n cu $n \in N$ ca mai sus.

(d) Fie acum spaţiul Hilbert al funcţiilor de pătrat integrabil pe R, $L^2(R)$. Fie $\mathcal{T} = \{-\infty\} \cup R \cup \{\infty\}$, $P_{-\infty} = O$ şi $P_\infty = I$. Pentru orice $t \in R$, definim proiectorul (numit şi trunchierea la momentul t) $(P_t f)(x) = f(x)$ dacă $x \le t$ şi 0 dacă $x > t$.

Analog se definesc rezoluţiile canonice pe spaţiile $L^2(0,\infty)$ şi $L^2[0,1]$.

3.Definiţie
Fie $(H, \mathcal{P})$ un spaţiu Hilbert cu rezoluţie şi fie $T \in \mathcal{L}(H)$. Sistemul T se numeşte **cauzal** (sau operator **subdiagonal**, **inferior triunghiular**) în raport cu rezoluţia fixată pe H dacă pentru orice $x, y \in H$ cu proprietatea $P_t x = P_t y$, $\forall P_t \in \mathcal{P}$, rezultă $P_t T x = P_t T y$, $\forall P_t \in \mathcal{P}$. Interpretarea intuitivă este evidentă: dacă intrările x şi y sunt egale până la momentul t, atunci şi ieşirile corespunzătoare, Tx şi Ty sunt egale până la momentul t. Folosind liniaritatea operatorului T, obţinem următoarele caracterizări echivalente:

4.Propoziţie
Fie $(H, \mathcal{P})$ un spaţiu Hilbert cu rezoluţie şi fie $T \in \mathcal{L}(H)$. Următoarele afirmaţii sunt echivalente:
(i) T este cauzal în raport cu rezoluţia $\mathcal{P}$.
(ii) $P_t T = P_t T P_t$, $\forall P_t \in \mathcal{P}$.
(iii) $T(I - P_t) = (I - P_t)T(I - P_t)$, $\forall P_t \in \mathcal{P}$.
(iv) Pentru orice $P_t \in \mathcal{P}$, subspaţiul $\mathrm{Ker}(P_t)$ este invariant pentru operatorul T.
Demonstraţie (i)$\Rightarrow$ (ii) Pentru $\forall x \in H$, avem $P_t[(I - P_t)x] = 0 = P_t 0$ şi deci, deoarece T este cauzal, rezultă $P_t[T(I - P_t)x] = P_t T 0 = 0$, adică $P_t T x = P_t T P_t x$.
Echivalenţa **(ii)$\Leftrightarrow$ (iii)** este evidentă.
(iii)$\Rightarrow$ (iv) Fie $x \in \mathrm{Ker}(P_t)$; din **(ii)**, avem: $P_t T x = P_t T P_t x = 0$.
(iv)$\Rightarrow$ (i) Fie $x, y \in H$ astfel încât $P_t x = P_t y$. Rezultă că $P_t(x - y) = 0$, deci $x - y \in \mathrm{Ker}(P_t)$. Din ipoteza **(iv)**, rezultă că $T(x - y) \in \mathrm{Ker}(P_t)$, adică $P_t T(x - y) = 0$, ceea ce arată că T este sistem cauzal.

5.Definiţie
Fie $(H, \mathcal{P})$ un spaţiu Hilbert cu rezoluţie. Noţiunea duală cauzalităţii este anticauzalitatea. Un sistem $T \in \mathcal{L}(H)$ se numeşte **anticauzal** în raport cu rezoluţia $\mathcal{P}$ (sau **operator supradiagonal**, **superior triunghiular**) dacă adjunctul său, $T^\star$, este cauzal.
Un sistem care este şi cauzal şi anticauzal se numeşte sistem **fără memorie**. Notăm cu $\mathcal{C}(H)$ mulţimea sistemelor cauzale pe H, cu $\mathcal{AC}(H)$ mulţimea sistemelor anticauzale şi cu $\mathcal{M}(H)$ muţimea sistemelor fără memorie.

Analogul propoziţiei anterioare pentru sisteme anticauzale este:

6.Propoziţie
Fie $(H, \mathcal{P})$ un spaţiu Hilbert cu rezoluţie şi fie $T \in \mathcal{L}(H)$; următoarele afirmaţii sunt echivalente:
(i) T este anticauzal în raport cu rezoluţia $\mathcal{P}$.
(ii) $(I - P_t)T = (I - P_t)T(I - P_t)\,, \forall P_t \in \mathcal{P}$.
(iii) $TP_t = P_tTP_t\,, \forall P_t \in \mathcal{P}$.
(iv) Pentru orice $P_t \in \mathcal{P}$, subspaţiul $\mathrm{Im}(P_t)$ este subspaţiu invariant pentru operatorul T.
Demonstraţie Totul rezultă din echivalenţa (cap. 6): subspaţiul K este invariant la $T \Leftrightarrow$ subspaţiul $K^\perp$ este invariant la $T^\star$ şi din egalitatea : $\mathrm{Ker}(T) = (\mathrm{Im}(T^\star)^\perp$.

7.Observaţie
Din propoziţiile 4 şi 6 rezultă că un sistem T este fără memorie dacă şi numai dacă pentru orice $P_t \in \mathcal{P}$, subspaţiile $\mathrm{Ker}(P_t)$ şi $\mathrm{Im}(P_t)$ sunt subspaţii reducătoare pentru T, sau, echivalent, orice proiector $P_t \in \mathcal{P}$ comută cu T, adică: $P_tT = TP_t$; (cap.5).

8.Propoziţie
Fie $(H, \mathcal{P})$ un spaţiu Hilbert cu rezoluţie. Atunci mulţimile $\mathcal{C}(H)$, $\mathcal{AC}(H)$ şi $\mathcal{M}(H)$ sunt spaţii Banach şi în plus produsul a două sisteme cauzale este cauzal (respectiv anticauzal, respectiv fără memorie.
Demonstraţie Dacă T şi S sunt doi operatori cauzali, atunci orice combinaţie liniară a lor este de asemenea operator cauzal, deoarece, conform propoziţiei 4, pentru orice $\alpha, \beta \in C$ şi $P_t \in \mathcal{P}$, avem:

$$T(\mathrm{Ker}(P_t)) \subseteq \mathrm{Ker}(P_t) \text{ şi } S(\mathrm{Ker}(P_t)) \subseteq \mathrm{Ker}(P_t) \Rightarrow$$

$$\Rightarrow (\alpha T + \beta S)(\mathrm{Ker}(P_t)) \subseteq \mathrm{Ker}(P_t).$$

Produsul: $TS(\mathrm{Ker}(P_t)) \subseteq \mathrm{Ker}(P_t)$. Pentru a demonstra completitudinea, fie T_n un şir de operatori cauzali care converge la T şi fie $x \in \mathrm{Ker}(P_t)$; atunci $P_tTx = \lim_{n\to\infty} P_tT_nx = 0$, ceea ce arată că T este operator cauzal. Analog se demonstrează şi pentru operatorii anticauzali. În cazul operatorilor fără memorie, trebuie să observăm în plus că adjunctul unui operator fără memorie este şi el fără memorie.

9.Exemple

(a) În continuare vom caracteriza operatorii cauzali pe C^n.
Fie $T \in \mathcal{L}(C^n)$ a cărui matrice în baza canonică este $A = (a_{ij})_{ij}$. Din propoziţia 6 rezultă că T este cauzal dacă şi numai dacă pentru orice $x \in C^n$

şi $k \in \{1, 2, .., n\}$ avem:

$$x(m) = 0 \text{ dacă } m \leq k \Rightarrow (Tx)(m) = 0 \text{ dacă } m \leq k.$$

Dar $(Tx)(m) = \sum\limits_{k=1}^{n} a_{mk}x(k)$ şi deci obţinem:

$$T \text{ este cauzal } \Leftrightarrow a_{ij} = 0\,,\ \forall i < j.$$

Deci un sistem pe C^n este cauzal (în raport cu rezoluţia canonică asociată bazei canonice) dacă şi numai dacă matricea sa în baza canonică este inferior triunghiulară.
Deoarece adjunctul $T^\star$ are matricea $(\overline{a_{ji}})_{ij}$, rezultă că sistemul T este anticauzal dacă şi numai dacă matricea $(a_{ij})_{ij}$ este superior triunghiulară. Din cele două caracterizări rezultă că un sistem pe C^n este fără memorie dacă şi numai dacă matricea sa (în baza canonică) este o
matrice diagonală.
Să presupunem acum că operatorul T este cauzal şi inversabil. Deoarece inversa unei matrice inferior triunghiulare este tot inferior triunghiulară, rezultă că pe spaţii finit dimensionale inversul unui sistem cauzal şi inversabil este şi el cauzal. Aşa cum vom vedea în exemplele următoare, această proprietate nu mai este adevărată pe spaţii Hilbert infinit dimensionale.

(b) Fie acum spaţiul $\ell^2(Z)$ şi fie $(\sigma_n)_{n\in Z}$ baza canonică; (a se vedea exemplul 17(ii)). Fie $T \in \mathcal{L}(\ell^2(Z))$ şi fie $(a_{ij})_{i,j\in Z}$ matricea sa (infinită), adică $a_{ij} =< T\sigma_j, \sigma_i >$. Printr-un raţionament similar cu cel din exemplul **(a)**, obţinem că T este cauzal dacă şi numai dacă
$a_{ij} = 0\,, \forall i < j\,,\, i, j \in Z$, adică matricea sa (în baza $(\sigma_n)_{n\in Z}$) este inferior triunghiulară. În particular, translaţia bilaterală W este sistem cauzal: matricea sa (în baza $(\sigma_n)_{n\in Z}$) este

$$a_{ij} = \begin{cases} 1 & \text{dacă} \quad i = j + 1 \\ 0 & \text{în rest} \end{cases}$$

Aşa cum am văzut în cap.6, W este unitar şi deci inversul său este egal cu adjunctul, care, conform definiţiei este anticauzal. Cum matricea lui $W^\star$ este superior triunghiulară, rezultă că operatorul W este cauzal şi inversabil, dar inversul său nu este cauzal. În acest caz, operatorii fără memorie sunt operatorii diagonali.

(c) Să considerăm acum cazul particular al unui operator de
convoluţie pe $\ell^2(\mathbf{Z})$, $C_\theta x = \theta \star x$, $\forall x \in \ell^2(\mathbf{Z})$, unde $\mathcal{F}^{-1}\theta \in L^\infty(\mathcal{S}^1)$. Deoarece matricea (infinită) a lui C_θ este
$a_{ij} = \theta(i - j)$, rezultă că

$$C_\theta \text{ este cauzal dacă şi numai dacă } \theta(n) = 0,\ \forall n < 0.$$

O formulare echivalentă este următoarea:

Sistemul C_θ este cauzal dacă şi numai dacă funcţia (numită funcţia de transfer a sistemului), $\mathcal{F}^{-1}\theta$ este esenţial mărginită şi analitică pe cerc, adică: $\mathcal{F}^{-1}\theta \in H^\infty(\mathcal{S}^1)$; aceasta deoarece coeficienţii săi Fourier de indici negativi sunt nuli:
$\widehat{\mathcal{F}^{-1}\theta}(n) = \theta(n) = 0, \forall n < 0$.
Interpretând spaţiul $\ell^2(Z)$ ca **domeniul timp** şi $L^2(\mathcal{S}^1)$ ca **domeniul frecvenţă**, rezultă (pentru sisteme de convoluţie), dualitatea:
cauzalitate (în domeniul timp) $\leftrightarrow$**analicitate** (în domeniul frecvenţă). Dacă sistemul de convoluţie C_θ este şi inversabil, (ceea ce este echivalent cu $0 \notin \text{essran}\left(\mathcal{F}^{-1}\theta\right)$: (cap.6), atunci inversul său este cauzal dacă şi numai dacă funcţia $\frac{1}{\mathcal{F}^{-1}\theta}$ este în subalgebra $H^\infty(\mathcal{S}^1)$. În concluzie, am obţinut:

Un sistem de convoluţie pe $\ell^2(Z)$ este **cauzal şi are un invers cauzal** dacă şi numai dacă $\mathcal{F}^{-1}\theta$ şi $\frac{1}{\mathcal{F}^{-1}\theta}$ sunt **analitice**, adică funcţia $\mathcal{F}^{-1}\theta$ este **funcţie exterioară** ("outer function").

(d) Caracterizăm acum operatorii integrali cauzali pe spaţiul $L^2(R)$. Fie deci (cap.6) $K : \mathbf{R^2} \to \mathbf{C}$ o funcţie de pătrat integrabil şi $(T_K f)(x) = \int\limits_R K(x,y)f(y)dy$, $\forall f \in L^2(\mathbf{R})$. Din exemplul 2(d) şi din propoziţia 4 rezultă că T_K este cauzal dacă şi numai dacă pentru orice $t \in \mathbf{R}$ şi pentru orice funcţie $f \in L^2(\mathbf{R})$ cu proprietatea $f(s) = 0, \forall s \leq t$ rezultă $(T_K f)(s) = 0, \forall s \leq t$. Fie $t \in \mathbf{R}$ fixat şi fie $f \in L^2(\mathbf{R})$ astfel încât $f(s) = 0, \forall s \leq t$. Pentru orice $s < t$, avem:

$$(T_K f)(s) = \int_R K(s,x)f(x)dx =$$

$$= \int_{-\infty}^{s} K(s,x)f(x)dx + \int_{s}^{\infty} K(s,x)f(x)dx = \int_{s}^{\infty} K(s,x)f(x)dx.$$

Rezultă deci că $(T_K f)(s) = 0$ dacă şi numai dacă

$$\int_{s}^{\infty} K(s,x)f(x)dx = 0.$$

Deoarece funcţia f este arbitrară pe intervalul (s, ∞), din egalitatea de mai sus rezultă $K(s,x) = 0, \forall s < x$, ceea ce încheie demonstraţia.
Pritr-un raţionament similar celui de mai sus, se poate demonstra că operatorul integral T_K este anticauzal dacă şi numai dacă nucleul K verifică egalitatea $K(x,y) = 0, \forall x > y$. În particular, rezultă că nu există operatori integrali (neidentic nuli) fără memorie pe $L^2(\mathbf{R})$.
O clasă de operatori fără memorie pe acest spaţiu este clasa operatorilor de multiplicare: pentru orice $\phi \in L^\infty(R)$, operatorul $M_\phi f = \phi f$,
$\forall f \in L^2(\mathbf{R})$ este fără memorie; lăsăm demonstraţia ca exerciţiu.

(e) În cazul particular al unui operator de convoluţie pe R, (cap.6), $(C_k f)(x) = \int\limits_R k(x-y)f(y)dy$, $\forall f \in L^2(\mathbf{R})$, din exemplul de mai sus rezultă că C_k este cauzal dacă şi numai dacă nucleul k are suportul inclus în $[0, \infty)$: $k(x) = 0$, $\forall x < 0$.

7.2 MF.07.2. Sisteme invariante în timp

O altă proprietate remarcabilă pe care o pot avea sistemele liniare este invarianţa în timp. Vom începe prin a defini noţiunea de sistem invariant în timp pe spaţiile $\ell^2(\mathbf{Z})$ şi $L^2(\mathbf{R})$ şi vom generaliza apoi definiţia la un spaţiu Hilbert abstract.

10.Definiţie
Fie W operatorul de translaţie bilateral pe spaţiul $\ell^2(\mathbf{Z})$, adică:
$(Wx)(n) = x(n-1)$, $\forall n \in \mathbf{Z}$. Un sistem $T \in \mathcal{L}(\ell^2(\mathbf{Z}))$ se numeşte **invariant în timp** dacă $TW = WT$. Evident că un sistem invariant în timp comută cu orice putere a lui W: $TW^k = W^kT$, $\forall k \in \mathbf{Z}$. Deoarece $W^kx = \sigma_k \star x$, rezultă că T este invariant în timp dacă şi numai dacă

$$\sigma_k \star (Tx) = T(\sigma_k \star x),\ \forall x \in \ell^2(\mathbf{Z}),\ \forall k \in Z.$$

Invarianţa în timp pe $L^2(\mathbf{R})$ se defineşte după cum urmează. Fie, pentru orice $s \in \mathbf{R}$ fixat, operatorul de translaţie

$$L_s : L^2(\mathbf{R}) \to L^2(\mathbf{R}),\ (L_sf)(t) = f(t-s).$$

Un sistem $T \in \mathcal{L}(L^2(\mathbf{R}))$ se numeşte **invariant în timp** dacă şi numai dacă $TL_s = L_sT$, $\forall s \in \mathbf{R}$.

Generalizare
Să observăm că în cele două definiţii date mai sus, avem de fiecare dată un grup abelian $(G, +)$, ($G = \mathbf{Z}$, respectiv $G = \mathbf{R}$), un spaţiu Hilbert H, ($H = \ell^2(\mathbf{Z})$ şi respectiv $H = L^2(\mathbf{R})$) şi o aplicaţie $\mathcal{R} : G \to \mathcal{L}(H)$, ($\mathcal{R}(k) = W^k$ şi respectiv $\mathcal{R}(s) = L_s$) cu proprietăţile:
(i) $\mathcal{R}(0) = I$.
(ii) $\mathcal{R}(u+v) = \mathcal{R}(u)\mathcal{R}(v)$, $\forall u, v \in G$.
(iii) Operatorul $\mathcal{R}(u)$ este unitar pentru orice $u \in G$.
(iv) Aplicaţia $H \times G \ni (f, u) \to (\mathcal{R}(u))\, f \in H$ este continuă.
O aplicaţie $\mathcal{R}$ cu proprietăţile de mai sus se numeşte **reprezentare continuă şi unitară** a grupului G pe spaţiul H.
În aceste condiţii, definiţia generală a sistemelor invariante în timp este: sistemul $T \in \mathcal{L}(H)$ se numeşte **invariant în timp** dacă
$\mathcal{R}(u)T = T\mathcal{R}(u)$, pentru orice $u \in G$; pentru completări, recomandăm [F01],

[S01], [O01].

11.Teoremă
Fie $T \in \mathcal{L}(\ell^2(\mathbf{Z}))$. Următoarele afirmaţii sunt echivalente:
(a) T este invariant în timp.
(b) T este operator de convoluţie, adică există $\theta : Z \to \mathbf{C}$ astfel încât $Tx = C_\theta x = \theta \star x$, $\forall x \in \ell^2(\mathbf{Z})$.
Demonstraţie Implicaţia **(b)⇒(a)** este evidentă deoarece operatorii de convoluţie comută între ei.
Fie $T \in \mathcal{L}(\ell^2(\mathbf{Z}))$ astfel încât $TW = WT$. Dacă $(a_{ij})_{i,j\in\mathbf{Z}}$ este matricea lui T (în baza canonică, $\{\sigma_k\}_{k\in Z}$), atunci din egalitatea $TW = WT$, obţinem:

$$\sum_{k\in Z} a_{ik+1}x(k) = \sum_{k\in Z} a_{i-1k}x(k),\ \forall x \in \ell^2(\mathbf{Z}),\ \forall i \in \mathbf{Z}.$$

De aici rezultă imediat că $a_{ij} = a_{i+1j+1}$, $\forall i, j \in Z$; prin inducţie (sau folosind direct egalităţile $TW^k = W^kT$, $\forall k \in \mathbf{Z}$), rezultă:

$$a_{ij} = a_{i-kj-k},\ \forall i, j, k \in \mathbf{Z}.$$

În concluzie, matricea sistemului T este constantă de-a lungul diagonalelor paralele cu diagonala principală (matrice Toeplitz), adică există $\theta : \mathbf{Z} \to \mathbf{C}$ astfel încât $a_{ij} = \theta(i - j)$, $\forall i, j \in \in Z$, deci sistemul T este un sistem de convoluţie: $Tx = C_\theta x = \theta \star x$. Funcţia $\mathcal{F}^{-1}\theta \in L^\infty(\mathcal{S}^1)$ se numeşte **funcţia de transfer a sistemului**; ea are proprietatea:

$$\frac{\mathcal{F}^{-1}(Tx)}{\mathcal{F}^{-1}x} = \mathcal{F}^{-1}\theta,$$

deci raportul dintre transformata Fourier (inversă) a ieşirii şi transformata Fourier (inversă) a intrării este constant (nu depinde de intrarea x). În general, pentru un sistem arbitrar, această proprietate constituie definiţia funcţiei de transfer a sistemului (dacă ea există).

Operatorii integrali invarianţi în timp pe $L^2(\mathbf{R})$ admit o caracterizare asemănătoare.

12.Propoziţie
Fie $K : \mathbf{R^2} \to \mathbf{C}$ o funcţie de pătrat integrabil şi fie $T_K : L^2(\mathbf{R}) \to L^2(\mathbf{R})$ operatorul integral asociat, adică: $(T_K f)(x) = \int\limits_R K(x, y)f(y)dy$.
Următoarele afirmaţii sunt echivalente:
(a) T_K este invariant în timp.
(b) T_K este operator de convoluţie, adică există $k : \mathbf{R} \to \mathbf{C}$ astfel încât $K(x, y) = k(x - y)$, deci $T_K f = C_k f = k \star f$.
Lăsăm demonstraţia ca exerciţiu.
Reiese clar din exemplele prezentate că există o legătură profundă între

noţiunea de sistem invariant în timp şi operaţia de convoluţie. Menţionăm de asemenea că noţiunea de sistem invariant în timp (şi legătura ei cu operaţia de convoluţie) se studiază şi pentru sisteme de tip distribuţie; recomandăm în acest sens [S01], [F01].

7.3 MF.07.3. Spaţiul stărilor

Intuitiv, starea (la momentul t) a unui sistem $T \in \mathcal{L}(H)$ este acea informaţie (eventual minimală) necesară pentru ca din cunoaşterea valorilor intrării posterioare momentului t să putem deduce valorile ieşirii posterioare momentului t.

13.Definiţie
Fie $(H, \mathcal{P})$ un spaţiu Hilbert cu rezoluţie (cu notaţiile din definiţia 1) şi fie $T \in \mathcal{L}(H)$. O **descompunere în spaţii de stări** (state space decomposition) a sistemului T este orice familie de triplete
$\{(X_t, \lambda_t, \theta_t); t \in \mathcal{T}\}$ cu proprietăţile:
(i) X_t este spaţiu Hilbert, $\forall t \in \mathcal{T}$.
(ii) $\lambda_t : H \to X_t$ este un operator liniar şi continuu astfel încât $\lambda_t = \lambda_t P_t, \forall t \in \mathcal{T}$.
(iii) $\theta_t : X_t \to H$ este un operator liniar şi continuu astfel încât $\theta_t = (I - P_t)\theta_t, \forall t \in \mathcal{T}$.
(iv) $(I - P_t)TP_t = \theta_t \lambda_t$.

În această definiţie, X_t se numeşte **spaţiul stărilor la momentul** t, λ_t este aplicaţia **intrare-stare**, iar θ_t aplicaţia **stare-ieşire**. Dacă $u \in H$ este o intrare arbitrară, atunci $x_t = \lambda_t u \in X_t$ se numeşte **starea sistemului** T **la momentul** t**, corespunzătoare intrării** u.
Să considerăm o intrare $u \in H$ şi fie $t \in \mathcal{T}$. Să calculăm valorile ieşirii Tu posterioare momentului t:

$$(I - P_t)Tu = (I - P_t)T[P_t + (I - P_t)]u =$$
$$= (I - P_t)TP_t u + (I - P_t)T(I - P_t)u = \theta_t \lambda_t u + (I - P_t)T(I - P_t)u =$$
$$= \theta_t x_t + (I - P_t)T(I - P_t)u.$$

De aici rezultă că, într-adevăr, cunoaşterea valorilor intrării posterioare momentului t (adică $(I - P_t)u$) şi a cuplului λ_t, θ_t (adică a stării la momentul t) permite cunoaşterea valorilor ieşirii posterioare momentului t.

14.Exemple
(i) Orice sistem $T \in \mathcal{L}(H)$ admite o descompunere "trivială" în spaţii de stări, considerând $X_t = H$, $\lambda_t = P_t$ şi $\theta_t = (I - P_t)T$, $\forall t \in \mathcal{T}$. Această de-

scompunere nu este interesantă deoarece aici spaţiul stărilor (H) este prea mare. Aşa cum vom vedea în continuare, sunt interesante acele descompuneri în care spaţiul stărilor este "mic"; o situaţie tipică în acest sens este aceea când spaţiul stărilor X_t are dimensiune finită, (deşi H are dimensiune infinită).

(ii) Sistemul diferenţial (sistem dinamic liniar)
Fie matricele $A \in \mathcal{M}_n(\mathbf{R})$, $B \in \mathcal{M}_{n,1}(\mathbf{R}), C \in \mathcal{M}_{1,n}(\mathbf{R})$ şi $D \in \mathbf{R}$. În cele ce urmează, spaţiul Hilbert $L^2(0,1)$ este considerat cu rezoluţia canonică, $(P_t)_{t\in[0,1]}$.
Fie $u \in L^2(0,1)$ şi fie $x : [0,1] \to \mathbf{R^n}$ soluţia problemei Cauchy

$$x'(t) = Ax(t) + Bu(t)\,,\; x(0) = 0.$$

Fie operatorul (numit **"sistem diferenţial"** sau **"sistem dinamic liniar"**)

$$\mathcal{D} : L^2(0,1) \to L^2(0,1),\; \mathcal{D}u = y, \text{unde}, y(t) = Cx(t) + Du(t), \forall t \in [0,1].$$

Din teoria ecuaţiilor diferenţiale rezultă ([B01],p.276):

$$x(t) = \int_0^t e^{A(t-s)}Bu(s)ds, \forall t \in [0,1],$$

şi deci $y(t) = (\mathcal{D}u)(t) = \int\limits_0^t Ce^{A(t-s)}Bu(s)ds + Du(t), \forall t \in [0,1]$. Propunem ca exerciţiu verificarea faptului că $\mathcal{D}$ este un operator liniar şi continuu pe $L^2(0,1)$. Se poate arăta (prin translaţia $y - Du = v$) că proprietăţile sistemului $\mathcal{D}$ nu se schimbă dacă presupunem că $D = 0$; vom face de aici înainte această ipoteză.
Vom construi acum o descompunere (canonică) în spaţii de stări a sistemului $\mathcal{D}$.
Fie, pentru orice $t \in [0,1]$, $X_t = \mathbf{R^n}$ şi fie aplicaţiile:

$$\lambda_t : L^2(0,1) \to \mathbf{R^n},\; \lambda_t u = x(t) = \int_0^t e^{A(t-s)}Bu(s)ds,$$

$$\theta_t : \mathbf{R^n} \to L^2(0,1),\; (\theta_t\xi)(s) = \begin{cases} Ce^{A(s-t)}\xi, & \text{dacă} \quad s \geq t \\ 0, & \text{dacă} \quad s < t \end{cases}$$

Înainte de a demonstra că descompunerea de mai sus verifică definiţia 13, să observăm că în acest caz, spaţiul stărilor ($\mathbf{R^n}$) este acelaşi la orice moment t şi este finit dimensional.
Pentru orice $t \in [0,1]$ şi $u \in L^2(0,1)$, avem:

$$\lambda_t u = \int_0^t e^{A(t-s)}Bu(s)ds = \int_0^t e^{A(t-s)}B(P_t u)(s)ds = \lambda_t P_t u.$$

$$(\theta_t \xi)(s) = \begin{cases} Ce^{A(s-t)}\xi, & \text{dacă} \quad s \geq t \\ 0, & \text{dacă} \quad s < t \end{cases} = [(I - P_t)\theta_t \xi](s).$$

Este clar că pentru orice $s < t$, avem:

$$[(I - P_t)TP_t u](s) = 0 = (\theta_t \lambda_t u)(s).$$

Fie acum $s \geq t$; avem:

$$[(I - P_t)TP_t u](s) = (TP_t u)(s) = \int_0^s Ce^{A(s-\tau)} BP_t u(\tau) d\tau =$$

$$= \int_0^t Ce^{A(s-\tau)} Bu(\tau) d\tau = Ce^{A(s-t)} \int_0^t e^{A(t-\tau)} Bu(\tau) d\tau = (\theta_t \lambda_t u)(s).$$

Vom nota (A, B, C) descompunerea (canonică) definită mai sus.

(iii) Sistemul dinamic liniar se poate defini şi pe spaţiul Hilbert $L^2(\mathbf{R})$. Pentru aceasta, fie matricele A, B, C ca în exemplul anterior; în plus, vom presupune că matricea A este **stabilă** , adică valorile proprii ale lui A sunt toate în semiplanul stâng: $\{z = a+ib \in \mathbf{C} \,;\, a < 0\}$. Pentru orice $u \in L^2(\mathbf{R})$, considerăm sistemul diferenţial:

$$x'(t) = Ax(t) + Bu(t),$$

cu condiţia iniţială $\lim\limits_{t \to -\infty} x(t) = 0$. Atunci soluţia (unică) a problemei Cauchy de mai sus este:

$$x(t) = \int_{-\infty}^t e^{A(t-\tau)} Bu(\tau) d\tau.$$

Sistemul dinamic liniar pe R este operatorul

$$\mathcal{D} : L^2(\mathbf{R}) \to L^2(\mathbf{R}),\ \mathcal{D}u = Cx.$$

Pentru demonstraţii şi completări, recomandăm [H02], S01], [S02].
Descompunerea canonică este $\{(\mathbf{R^n}, \lambda_t, \theta_t) \,;\, t \in \mathbf{R}\}$, unde:

$$\lambda_t : L^2 \to \mathbf{R^n},\ \lambda_t u = \int_{-\infty}^{\infty} e^{A(t-\tau)} Bu(\tau) d\tau,$$

$$\theta_t : \mathbf{R^n} \to L^2(\mathbf{R}),\ (\theta_t \xi)(\tau) = \begin{cases} Ce^{A(t-\tau)}\xi, & \text{dacă} \quad \tau \geq t \\ 0, & \text{dacă} \quad \tau < t \end{cases}$$

Lăsăm ca exerciţiu demonstraţia.

(iv) Sistemul discret (sistem "diferenţă")
Analogul discret al exemplului anterior este definit după cum urmează. Fie

matricele A, B, C ca mai sus şi fie $u \in \ell^2(\mathbf{N})$. Fie $x : \mathbf{N} \to \mathbf{R^n}$ soluţia recurenţei:

$$x(k+1) = Ax(k) + Bu(k),\ x(0) = 0.$$

Sistemul diferenţă este operatorul liniar şi continuu

$$\ell^2(\mathbf{N}) \ni u \to y \in \ell^2(\mathbf{N}),\ \text{unde}\,, y(k) = Cx(k), \forall k \in \mathbf{N}.$$

Este uşor de demonstrat că

$$y(k) = Cx(k) = C\sum_{j=0}^{k-1} A^j Bu(k-1-j),\ \forall k \geq 1.$$

Spaţiul stărilor este $X_k = \mathbf{R^n}$, $\forall k \in \mathbf{N}$ şi:

$$\lambda_k : \ell^2(\mathbf{N}) \to \mathbf{R^n},\ \lambda_k u = x(k),$$

$$\theta_k : \mathbf{R^n} \to \ell^2(\mathbf{N}),\ (\theta_k \xi)(j) = \begin{cases} CA^{j-k}\xi & j \geq k \\ 0 & j < k-1 \end{cases}$$

Lăsăm ca exerciţiu demonstraţia faptului că $(X_k, \lambda_k, \theta_k)_{k\in N}$ este o descompunere în sensul definiţiei 13, pe care o vom nota (A, B, C).

(v) În toate exemplele de până acum, spaţiul stărilor a fost acelaşi la fiecare moment: $X_t = \mathbf{R^n}$, $\forall t$. Dăm în continuare un exemplu în care spaţiul stărilor este variabil în timp.
Pentru aceasta, vom face observaţia că pentru a defini o descompunere în spaţii de stări a unui sistem T, este suficient să definim operatorii $(I - P_t)TP_t$, $\forall t$; menţionăm că, în general, familia de operatori $\{(I-P_t)TP_t\,;\, t \in \mathcal{T}\}$ nu determină în mod unic sistemul T. Totuşi, în exemplul care urmează, sistemul T este unic determinat în acest mod; pentru demonstraţii şi completări, recomandăm [F01].
Fie $H = \ell^2(\mathbf{N})$ cu rezoluţia canonică (cf. exemplului 2(c)) şi fie $\{\sigma_n\}_{n\in N}$ baza sa canonică. Fie:

$$(I - P_n)TP_n u = \sum_{k=0}^{n} < u, \sigma_k > \sigma_{n+1+k}.$$

Aşa cum am menţionat, familia $\{(I - P_n)TP_n\,;\, n \in \mathbf{N}\}$ determină sistemul T. O descompunere în spaţii de stări pentru sistemul T se poate obţine după cum urmează; pentru orice n=0,1,2,..., definim:

$$\lambda_n : \ell^2(\mathbf{N}) \to \mathbf{R^{n+1}},\ \lambda_n u = (< u, \sigma_o >, < u, \sigma_1 >, .., < u, \sigma_n >)\,.$$

$$\theta_n : \mathbf{R^{n+1}} \to \ell^2(\mathbf{N}),\ \theta_n(x_o, x_1, .., x_n) = \sum_{k=0}^{n} x_k \sigma_{n+1+k}.$$

Se demonstrează fără dificultate că $\{R^{n+1}, \lambda_n, \theta_n;\, n \in \mathbf{N}\}$ este o descompunere în spaţii de stări pentru T. Observăm că spaţiul stărilor este finit dimensional la orice moment, dar, dimensiunea sa creşte odată cu trecerea timpului: $\mathbf{R}$, $\mathbf{R^2}$, $\mathbf{R^3}, \ldots$.

7.4 MF.07.4. Controlabilitate şi observabilitate

15.Definiţie

Fie H un spaţiu Hilbert cu rezoluţie şi fie $T \in \mathcal{L}(H)$. O descompunere $\{X_t, \lambda_t, \theta_t\}_{t\in\mathcal{T}}$ se numeşte **complet controlabilă** dacă pentru orice $t \in \mathcal{T}$, imaginea aplicaţiei λ_t este subspaţiu dens în X_t, adică $\forall t \in \mathcal{T}$, $\forall x \in X_t$ şi $\forall \epsilon > 0$, $\exists u \in H$ astfel încât $\| \lambda_t u - x \| < \epsilon$.
Intuitiv, o descompunere este complet controlabilă dacă la orice moment, pentru orice stare dată, există o intrare care să aducă sistemul oricât de aprope de starea dată.
Descompunerea $\{X_t, \lambda_t, \theta_t\}_{t\in\mathcal{T}}$ se numeşte **complet observabilă** dacă toţi operatorii θ_t sunt mărginiţi inferior, adică $\forall t \in \mathcal{T}$, $\exists\epsilon > 0$ astfel încât $\| \theta_t x \| \geq \epsilon \| x \|$, $\forall x \in X_t$.
Intuitiv, complet observabilitate înseamnă posibilitatea determinării stării la orice moment dat dacă se cunosc valorile intrărilor şi ieşirilor posterioare momentului dat.
O descompunere care este şi complet controlabilă şi complet observabilă se numeşte **descompunere minimală**.
Propunem ca exerciţiu faptul că descompunerea (variabilă în timp), construită în exemplul 14(v) este minimală. Pentru sistemul dinamic liniar, vom demonstra mai întâi un criteriu de minimalitate pentru descompunerea sa canonică.
Un punct forte al modelului matematic prezentat aici pentru noţiunea de stare este şi următoarea teoremă de existenţă a descompunerilor minimale pentru un sistem arbitrar.

16. Teoremă

Fie H un spaţiu Hilbert cu rezoluţie şi fie $T \in \mathcal{L}(H)$. Atunci T admite o descompunere minimală.
Demonstraţie Construim mai întâi spaţiul stărilor la un moment dat. Fie deci $t \in \mathcal{T}$ fixat; definim pe spaţiul $P_t(H)$ relaţia de echivalenţă:

$$x \sim_t y \Leftrightarrow (I - P_t)TP_t x = (I - P_t)TP_t y,\ \forall x, y \in P_t(H).$$

Mulţimea claselor de echivalenţă,

$$\widehat{P_t(H)} = \{[x]_t \ ;\ x \in P_t(H)\}$$

este spaţiu vectorial cu operaţiile uzuale:

$$[x]_t + [y]_t = [x + y]_t \text{ şi } \alpha[x]_t = [\alpha x]_t, \, \forall x, y \in P_t(H), \, \forall \alpha \in \mathbf{C}.$$

Se demonstrează de asemenea că aplicaţia:

$$< [x]_t, [y]_t >_t = < (I - P_t)TP_t x, (I - P_t)TP_t y >,$$

este un produs scalar pe $\widehat{P_t(H)}$, care se organizează astfel ca un spaţiu prehilbertian. Definim spaţiul stărilor la momentul t, X_t, ca fiind completatul (închiderea) acestui spaţiu prehilbertian.
Definim acum operatorii λ_t şi θ_t.

$$\lambda_t : H \to X_t, \, \lambda_t x = [P_t x]_t.$$

Operatorul θ_t este definit iniţial pe subspaţiul (dens) $\widehat{P_t(H)}$ prin formula:

$$\theta_t [x]_t = (I - P_t)TP_t x.$$

Demonstrăm acum că θ_t este continuu, şi deci el poate fi prelungit prin continuitate la întreg spaţiul X_t:

$$\| \theta_t[x]_t \|^2 = \| (I - P_t)TP_t x \|^2 = < (I - P_t)TP_t x, (I - P_t)TP_t x > = \| [x]_t \|,$$

ultima normă fiind norma din X_t. Din relaţia de mai sus rezultă că operatorul θ_t este o izometrie şi deci, în mod evident, el este şi mărginit inferior.
Demonstrăm acum că λ_t este continuu:

$$\| \lambda_t x \| = \| \theta_t \lambda_t x \| = \| (I - P_t)TP_t x \| \leq \| T \| \| x \|, \, \forall x \in H.$$

Se verifică prin calcul direct egalităţile:

$$\lambda_t = \lambda_t P_t, \, \theta_t = (I - P_t)\theta_t \text{ şi } (I - P_t)TP_t = \theta_t \lambda_t.$$

Demonstrăm acum că descompunerea este minimală. Complet observabilitatea a fost deja demonstrată, deoarece θ_t este mărginit inferior. Pe de altă parte, din definiţie, λ_t are imagine densă în X_t, deci descompunerea este şi complet controlabilă.

17.Teoremă (Criteriile generale de controlabilitate şi observabilitate)
Fie H un spaţiu Hilbert cu rezoluţie, fie $T \in \mathcal{L}(H)$ şi fie $\{X_t, \lambda_t, \theta_t\}_{t \in \mathcal{T}}$ o descompunere a sa.

(i) Descompunerea este complet controlabilă dacă şi numai dacă pentru orice $t \in \mathcal{T}$, avem:

$$< \lambda_t \lambda_t^\star x, x > > 0, \, \forall x \in X_t, \, x \neq 0.$$

(ii) Descompunerea este complet observabilă dacă şi numai dacă pentru orice $t \in \mathcal{T}$, avem:

$$< \theta_t^\star \theta_t u, u >> 0, \forall u \in H, \ u \neq 0.$$

Demonstraţie (i) Dacă descompunerea este complet controlabilă, atunci, din definiţie, rezultă că operatorii λ_t au imagine densă. Se poate demonstra că $\lambda_t^\star$ este injectiv, $\forall t \in \mathcal{T}$ şi deci pentru orice $x \in H$, $x \neq 0$, avem:

$$< \lambda_t \lambda_t^\star x, x >=< \lambda_t^\star x, \lambda_t^\star x >=\| \lambda_t^\star x \|^2 > 0.$$

Reciproc, dacă $\lambda_t \lambda_t^\star > 0$, atunci $\lambda_t^\star$ este injectiv şi rezultă că λ_t are imagine densă.
(ii) Raţionamentul este asemănător cu cel de mai sus.

Încheiem acest paragraf cu unele particularizări şi exemplificări ale noţiunilor şi rezultatelor de până acum.

Un caz particular remarcabil se obţine aplicând criteriul general de controlabilitate şi observabilitate sistemului diferenţial.

18.Propoziţie (Criteriile lui Kalman de observabilitate şi controlabilitate pentru sistemul diferenţial)
Fie $\mathcal{D}$ sistemul diferenţial din exemplul 14(ii) şi fie (A, B, C) descompunerea sa canonică.

(i) Descompunerea (A, B, C) este complet observabilă dacă şi numai dacă următoarea matrice ("de observabilitate") are rang maxim:

$$\mathcal{Q} = \begin{pmatrix} C \\ \ldots \\ CA \\ \ldots \\ CA^2 \\ \ldots \\ \vdots \\ \ldots \\ CA^{n-1} \end{pmatrix}$$

(ii) Descompunerea (A, B, C) este complet controlabilă dacă şi numai dacă următoarea matrice ("de controlabilitate") are rang maxim:

$$\mathcal{R} = \left(B \vdots BA \vdots BA^2 \vdots \ldots \vdots BA^{n-1} \right)$$

Demonstraţie (i) Conform teoremei precedente, descompunerea este complet observabilă dacă şi numai dacă $\theta_t^\star \theta_t > 0$, $\forall t \in [0, 1]$, unde, θ_t a fost definit în exemplul 14(ii):

$$\theta_t : R^n \to L^2(0,1),\ (\theta_t\xi)(s) = \begin{cases} Ce^{A(s-t)}\xi, & \text{dacă} \quad s \geq t \\ 0, & \text{dacă} \quad s < t \end{cases}$$

Calculăm acum adjunctul lui θ_t. Reamintim că dacă ξ şi η sunt doi vectori (coloane) din $\mathbf{R^n}$, atunci produsul lor scalar este $\xi^T\eta$, unde, ξ^T este transpusul lui ξ.
Pentru orice $\xi \in \mathbf{R^n}$ şi $f \in L^2(0,1)$, avem:

$$< f, \theta_t\xi >_{L^2} = \int_t^1 f(s)Ce^{A(s-t)}\xi ds = \left[\int_t^1 f(s)Ce^{A(s-t)}ds\right]^T \xi =$$

$$= \left[\int_t^1 e^{A(s-t)}C^T f(s)ds\right]^T \xi = < \theta_t^\star f, \xi >_{\mathbf{R^n}} .$$

Rezultă deci că pentru orice $t \in [0,1]$, avem:

$$\theta_t^\star f = \int_t^1 e^{A^T(s-t)}C^T f(s)ds,\ \forall f \in L^2(0,1).$$

În concluzie, operatorul $\theta_t^\star\theta_t : \mathbf{R^n} \to \mathbf{R^n}$ este:

$$\theta_t^\star\theta_t\xi = \left[\int_t^1 e^{A^T(s-t)}C^T Ce^{A(s-t)}ds\right]\xi.$$

Deci $\theta_t^\star\theta_t > 0$ dacă şi numai dacă matricea

$$\int_t^1 e^{A^T(s-t)}C^T Ce^{A(s-t)}ds$$

este strict pozitiv definită, sau, echivalent, dacă

$$0 << \xi, \theta_t^\star\theta_t\xi >= \int_t^1 \left[Ce^{A(s-t)}\xi\right]^T \left[Ce^{A(s-t)}\xi\right] ds,\ \forall\xi \in \mathbf{R^n}, \xi \neq 0.$$

Rezultă deci că descompunerea nu este complet observabilă dacă şi numai dacă există $\xi \neq 0$ astfel încât

$$Ce^{A(s-t)}\xi = 0,\ \forall s \in [t,1],$$

sau, echivalent

$$Ce^{tA}\xi = 0,\ \forall t \in [0,1].$$

Dezvoltând e^{tA} în serie Taylor, relaţia de mai sus devine:

$$\sum_{j=0}^{\infty} \frac{CA^j\xi t^j}{j!} = 0,\ \forall t \in [0,1].$$

Din teorema de unicitate a dezvoltării în serie de puteri, rezultă:

$$CA^j\xi = 0, \ \forall j \in \mathbf{N}.$$

Deoarece puterile A^j, $\forall j \geq n$ se pot exprima în funcţie de puterile A^j, $0 \leq j \leq n-1$, (consecinţă a teoremei Hamilton-Cayley), este suficient să avem:

$$CA^j\xi = 0, \ \forall j \in \{0, 1, 2, .., n-1\}.$$

În concluzie, descompunerea este complet observabilă dacă şi numai dacă sistemul liniar şi omogen $\mathcal{Q}\xi = 0$ (cf. notaţiei din enunţ) are numai soluţia banală, adică matricea $\mathcal{Q}$ are rang maxim.
(ii) Raţionamentul este analog celui precedent; calculăm mai întâi operatorul $\lambda_t^\star : R^n \to L^2(0,1)$, pentru care obţinem:

$$(\lambda_t^\star \xi)(s) = \begin{cases} B^T e^{A^T(t-s)} & \text{dacă} \quad s \leq t \\ 0 & \text{dacă} \quad s > t \end{cases}$$

Rezultă deci că matricea (în baza canonică) a operatorului $\lambda_t \lambda_t^\star$ este:

$$\int_0^t e^{A(t-s)} B B^T e^{A^T(t-s)} ds.$$

Repetând raţionamentul de la punctul **(i)**, demonstraţia se încheie.

19.Observaţie
Pentru sistemul dinamic liniar pe $L^2(\mathbf{R})$ şi pentru sistemul discret din exemplul 14(iii) şi (iv) se poate enunţa şi demonstra un rezultat analog celui anterior; lăsăm acest fapt ca exerciţiu.

20.Definiţie
Sistemul diferenţial $\mathcal{D}$ admite următoarea generalizare vectorială.
Fie $m \in N$ şi fie

$$L^2([0,1], \mathbf{R^m}) = \left\{ u : [0,1] \to \mathbf{R^m} \, ; \, u \, \text{măsurabilă şi} \int_0^1 \| \, u(t) \, \|^2 \, dt \, < \, \infty \right\}.$$

În definiţia de mai sus $\| \ \|$ este norma euclidiană din $\mathbf{R^m}$. Cu operaţiile uzuale de adunare şi înmulţire cu scalari, $L^2([0,1], \mathbf{R^m})$ este spaţiu vectorial; se demonstrează că aplicaţia:

$$< u, v >= \int_0^1 < u(t), v(t) > dt$$

determină pe $L^2([0,1], \mathbf{R^m})$ o structură de spaţiu Hilbert; demonstraţiile acestor afirmaţii sunt adaptări ale celor din cazul scalar (m=1).

Fie $n, p \in \mathbf{N}$ şi fie $A \in \mathcal{M}_{n,n}(\mathbf{R})$, $B \in \mathcal{M}_{n,m}(\mathbf{R})$, $C \in \mathcal{M}_{p,n}(\mathbf{R})$. Pentru orice $u \in L^2([0,1], \mathbf{R^m})$, fie $x : [0,1] \to \mathbf{R^n}$ soluţia problemei Cauchy:

$$x'(t) = Ax(t) + Bu(t), \forall t \in [0,1] \ ; \ x(0) = 0.$$

Sistemul diferenţial (cazul vectorial) este aplicaţia

$$\mathcal{D} : L^2([0,1], \mathbf{R^m}) \to L^2([0,1], \mathbf{R^p}), \ \mathcal{D}u = y, \text{ unde, } y(t) = Cx(t).$$

Rezultatele demonstrate pentru cazul scalar sunt adevărate şi pentru cazul vectorial, cu adaptările corespunzătoare (care sunt evidente). Să mai observăm că în cazul vectorial matricele de observabilitate, $\mathcal{Q}$, şi controlabilitate, $\mathcal{R}$, nu mai sunt pătratice, însă criteriile lui Kalman se enunţă la fel ca în cazul scalar (rang maxim).

Propunem ca exerciţiu definirea sistemului discret vectorial, a sistemului dinamic liniar vectorial pe $L^2(\mathbf{R})$ şi a variantei vectoriale a sistemului din exemplul 14(v).

Sistemele în care intrările şi ieşirile sunt funcţii cu valori vectoriale se numesc sisteme MIMO (Multi Input, Multi Output), iar cele în care intrările şi ieşirile iau valori scalare se numesc SISO (Single Input, Single Output).

Capitolul 8

MF.08. Câmp de probabilitate

Cuvinte cheie

Evenimente, σ-algebra evenimentelor, măsura, câmp de probabilitate, probabilităţi condiţionate, independenţă , formula Bayes.

8.1 MF.08.1.Mulţimi. Funcţii.

Se presupun cunoscute noţiunile elementare ale teoriei mulţimilor. Vom recapitula unele definiţii şi vom fixa notaţiile; pentru teoria probabilităţilor o primă dificultate o constituie obişnuinţa cu "mulţimi de mulţimi", adică mulţimi ale căror elemente sunt, la rândul lor mulţimi.

Apartenenţa. Dacă A este o mulţime şi $a \in A\,(a \notin A)$ spunem că a aparţine (nu aparţine) mulţimii A (este sau nu este un element al mulţimii A). *Mulţimea vidă* $\emptyset$ este unica mulţime care nu are elemente. Două mulţimi sunt egale dacă au aceleaşi elemente.

Incluziune. Mulţimea A este inclusă în mulţimea B dacă $x \in A \Longrightarrow x \in B$ pentru orice $x \in A$ ("$\Longrightarrow$" este semnul implicaţiei logice). In acest caz scriem $A \subseteq B$.Dacă $A \subseteq B$ şi $A \neq B$ scriem $A \subset B$ (incluziune strictă). Evident: $A \subseteq A, A \subseteq B$ şi $B \subseteq A$ dacă şi numai dacă $A = B, A \subseteq B$ şi $B \subseteq C$ implică $A \subseteq C$.Dacă $A \subseteq B$ spunem că (mulţimea A este o *submulţime* (*parte*) a (mulţimii) B.Mulţimea vidă este submulţime a oricărei mulţimi (justificare?). Dată o mulţime M, mulţimea submulţimilor (părţilor) mulţimii M se notează $\mathcal{P}(M)$. Astfel: $A \in \mathcal{P}(M) \Longleftrightarrow A \subseteq M$ ("$\Longleftrightarrow$" este simbolul echivalenţei logice "dacă şi numai dacă "). Rezultă $\emptyset \in \mathcal{P}(M)$ pentru orice M. Reamintim că dacă mulţimea M are n elemente, atunci mulţimea $\mathcal{P}(M)$ are 2^n elemente. Dacă $a \in M$ se face deosebire între elementul a şi submulţimea $\{a\}$.

Combinări. Dacă A este o mulţime cu n elemente $(n \geq 1)$ şi $k \leq n$, se numeşte *combinare* de k elemente din A orice submultime a mulţimii A având k elemente. Numărul de submulţimi de k elemente ale unei multimi de n elemente este $C_n^k = \frac{n(n-1)...(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}$. Formulele principale ale calculului cu combinări au fost studiate în liceu.

Exerciţii. 1) Fie mulţimile $\emptyset, A = \{1\}, B = \{0, 1\}$. Determinaţi $P(\emptyset)$, $P(A)$, $P(B)$.

2) Justificaţi egalitatea $C_n^0 + C_n^1 + C_n^2 + ... + C_n^n = 2^n$.

Exemple. 1) Mulţimile fundamentale de numere (naturale,întregi, raţionale, reale, complexe) se vor nota, respectiv $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$. Putem considera, în mod natural, că $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

2) Mulţimea rezultatelor posibile la aruncarea unui zar este $\{1, 2, 3, 4, 5, 6\}$.

Operaţii cu mulţimi. 1) *Reuniune*: dacă $(A_i)_{i\in I}$ este o familie de mulţimi (indexată cu o mulţime de indici I), se defineşte reuniunea mulţimilor familiei $\cup A_i = \{x; \exists i \in I, x \in A_i\}$ (s-a folosit simbolul "$\exists$" pentru "există "). Reuniunea unei familii finite (infinite) de mulţimi va fi numită (impropriu) reuniune finită (infinită). Termenii "finit", "infinit" vor fi luaţi în sens intuitiv fă ˚ ă precizări suplimentare. Reuniunea a două mulţimi A, B se notează $A \cup B$ (şi analog pentru o familie finită de mulţimi).

2) *Intersecţie:* dacă $(A_i)_{i\in I}$ este o familie de mulţimi, se defineşte intersecţia mulţimilor familiei $\cap A_i = \{x; \forall i \in I, x \in A_i\}$ (s-a folosit simbolul "$\forall$" pentru "orice"). Intersecţia unei familii finite (infinite) de mulţimi va fi numită (impropriu) reuniune finită (infinită). Intersecţia a două mulţimi A, B se notează $A \cap B$ (şi analog pentru o familie finită de mulţimi).

3) *Complementara*: dacă $A \subseteq B$, complementara mulţimii A (în raport cu mulţimea B) este $C_B A = \{x; x \in B, x \notin A\}$.Dacă nu este pericol de confuzie vom omite indicele B în notaţia complementarei.

4) *Diferenţa:* pentru mulţimile A, B se defineşte diferenţa $A - B = \{x; x \in A, x \notin B\}$. Evident, $C_B A = B - A$. Proprietăţile operaţiilor cu mulţimi (asociativitatea şi comutativitatea reuniunii şi intersecţiei, distributivitatea intersecţiei faţă de reuniune şi a reuniunii faţă de intersecţie etc.) se presupun cunoscute (a se vedea şi exerciţiul de mai jos). Vom scrie legile De Morgan :

$$C(\cup A_i) = \cap(CA_i) \text{ şi } C(\cap A_i) = \cup(CA_i).$$

Exerciţii.1) Dacă A, B, C sunt mulţimi, să se arate că $A \cap (B \cup C) = (A \cup B) \cap (A \cup C)$ (distributivitate).

2) Să se determine $A \cup A, A \cap A, A - A$ pentru o mulţime A.

3) Dacă A, B sunt mulţimi, să se calculeze $A - (A - B)$.

Aranjamente. Fie A o mulţime cu $n \geq 1$ elemente şi $1 \leq k \leq n$. Numim aranjament de k elemente din A imaginea unei funcţii injective $f : \{1, 2..., k\} \to A$. Numărul funcţiilor injective definite pe $\{1, 2..., k\}$ cu valori în A este $A_n^k = n(n-1)...(n-k+1)$.

Permutări. Se notează $P_n = A_n^n$ (numărul funcţiilor bijective între două mulţimi cu n elemente).

Produs cartezian. Dacă A, B sunt mulţimi se defineşte mulţimea $A \times B = \{(x, y); x \in A, y \in B\}$ a perechilor ordonate de elemente din A, respectiv B. Reamintim că , în $A \times B, (x, y) = (u, v)$ dacă şi numai dacă $x = u, y = v$. Dacă A are m elemente, iar B are n elemente, atunci $A \times B$ are mn elemente.

O funcţie $f : A \to B$ ataşează fiecărui element $x \in A$ un element bine determinat $f(x) \in B$; A este mulţimea de definiţie a funcţiei, iar B mulţimea în care funcţia ia valori. Două funcţii f, g sunt egale dacă au aceeaşi mulţime de definiţie, aceeaşi mulţime în care iau valori şi $f(x) = g(x)$ pentru orice x . O funcţie $f : A \to B$ este injectivă dacă $f(x) = f(y)$ implică $x = y$, surjectivă dacă pentru orice $y \in B$ există $x \in A$ astfel încât $f(x) = y$, bijectivă dacă este injectivă şi surjectivă . Pentru o funcţie bijectivă $f : A \to B$ se defineşte funcţia inversă $f^{-1} : B \to A$ prin $f^{-1}(y) = xi$ dacă $f(x) = y$. Dacă $A = \emptyset$, atunci se consideră că există o unică funcţie definită pe A cu valori în B; dacă $B = \emptyset$ şi $A \neq \emptyset$ nu există funcţii definite pe A cu valori în B.

Compunerea funcţiilor. Dacă $f : A \to B$ şi $g : B \to C$ sunt două funcţii, se defineşte funcţia $g \circ f : A \to C$ prin $(g \circ f)(x) = g(f(x))$ pentru orice $x \in A$. Compunerea funcţiilor este asociativă . Dată mulţimea $A \neq \emptyset$ se defineşte funcţia $id_A : A \to A, id_A(x) = x, \forall x \in A$. Evident, dacă $f : A \to B$, atunci $f \circ id_A = f, id_B \circ f = f$.

Imagine directă , imagine reciprocă . Fie $f : M \to N$ o funcţie. Pentru $A \subseteq M, B \subseteq N$ se definesc: $f(A) = \{f(x); x \in A\}$, $f^{-1}(B) = \{x; f(x) \in B\}$.$f(A)$ este *imaginea (directă) a* mulţimii A prin funcţia f, iar $f^{-1}(B)$ *imaginea reciprocă* a mulţimii B prin funcţia f (a nu se confunda notaţia imaginii reciproce cu existenţa funcţiei inverse). Sunt importante proprietăţile: $f^{-1}(\cup B_i) = \cup f^{-1}(B_i), f^{-1}(\cap B_i) = \cap f^{-1}(B_i), f^{-1}(C_N B) = C_M f^{-1}(B)$. Pentru $y \in N$ se scrie $f^{-1}(y)$ în loc de $f^{-1}(\{y\})$; dacă y este o valoare a funcţiei f, atunci $f^{-1}(y)$ se zice mulţime de nivel a funcţiei f.

Exerciţii.4) Demonstraţi egalităţile de mai sus (din 2.2).

5) Care dintre proprietăţile 2.2 ale imaginii reciproce sunt adevărate pentru imaginea directă (evident formulate adecvat)?

6) Să se arate că , dacă $f : M \to N$ este surjectivă , atunci funcţia $g : \mathcal{P}(N) \to \mathcal{P}(M), g(B) = f^{-1}(B)$ este injectivă .

Operaţii cu funcţii. Vom studia, în acest set de lecţii, mai ales funcţii cu valori în $\mathbb{R}$ (funcţii reale). Este important ca operaţiile de baza cu aceste funcţii să fie bine întelese. Astfel:

Adunare. Fie $f, g : M \to \mathbb{R}$ două funcţii reale. Se defineşte suma $f + g : M \to \mathbb{R}$ prin:

$$(f + g)(x) = f(x) + g(x), \forall x \in M.$$

Inmulţire cu un număr. Fie $f : M \to \mathbb{R}$ şi $\alpha \in \mathbb{R}$. Se defineşte funcţia

$\alpha f : M \to \mathbb{R}$ prin:

$$(\alpha f)(x) = \alpha f(x), \forall x \in M.$$

Inmulţire. Fie $f, g : M \to \mathbb{R}$ două funcţii reale. Se defineşte produsul $fg : M \to \mathbb{R}$ prin:

$$(fg)(x) = f(x)\, g(x), \forall x \in M.$$

Se verifică uşor proprietăţile fundamentale ale acestor operaţii; nu mai insistăm. Dacă $M = \mathbb{N}$ se obţin operaţiile cu şiruri de numere reale. Cele de mai sus se extind fără modificări la funcţii cu valori în $\mathbb{C}$.

Indicatorul unei mulţimi. Fie $A \subseteq M$. Funcţia $\chi_A : M \to \mathbb{R}, \chi_A(x) = 1$ dacă $x \in A$ şi $\chi_A(x) = 0$ dacă $x \in M - A$ se numeşte *indicatorul unei mulţimi* A. Se observă usor că , reciproc, orice funcţie f definită pe M care ia (cel mult) valorile 0 sau 1 este indicatorul unei (unice) submulţimi a mulţimii M. Fie $f : M \to \mathbb{R}$ o funcţie cu un număr finit de valori $\alpha_1 \alpha_2 ... \alpha_n$ şi fie $A_1, A_2 ... A_n$ submulţimile pe care f ia valorile respective (mulţimile de nivel ale funcţiei f); desigur $A_i \cap A_j = \emptyset$ pentru $i \neq j$ şi $A_1 \cup A_2 \cup ... \cup A_n = M$ (spunem că $A_1, A_2 ... A_n$ formează o partiţie a mulţimii M. Se arată uşor (exerciţiu) că $f = \sum_{i}^{n} \alpha_i \chi_{A_i}$

Exerciţiu. 7) Să se arate că dacă $A, B \subseteq M$, atunci $\chi_{A\cup B} = \max(\chi_A, \chi_B)$ şi $\chi_{A\cap B} = \min(\chi_A, \chi_B)$, unde max(min) înseamna luarea valorii maxime (minime).

Numărabilitate. O mulţime A este numărabilă dacă există o funcţie bijectivă $f : \mathbb{N} \to A$ (elementele mulţimii se pot "aranja" într-un şir). Se arată că mulţimile $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ sunt numărabile (pentru $\mathbb{N}$ este evident), iar mulţimea $\mathbb{R}$ nu este numărabila. Pentru mai multe detalii privind acest subiect se pot consulta cursuri de Analiză Matematică .

8.2 MF.08.2. Evenimente, $\sigma-$ algebra evenimentelor.

Evenimente. Vom utiliza termenii *experienţă* şi *experiment* ca având acelaşi sens pe care îl presupunem intuitiv clar. *Rezultatele* unui experiment, vor trebui precizate pentru fiecare experiment în parte.

Exemple. 1) Experimentul constă din lansarea (aruncarea) unui zar pe o masă (goală). Se presupune că zarul se opreşte pe o anumită faţă . Rezultatul unui asemenea experiment este numărul indicat pe faţa superioară a zarului, după oprirea acestuia. Rezultatele posibile ale experimentului sunt în număr de şase.

2) O urnă conţine bile albe şi bile negre; bilele sunt identice cu excepţia culorii. Experimentul constă în extragerea, fără a privi în urnă , a unei

bile. Rezultatul unui experiment se consideră a fi culoarea bilei extrase. Rezultatele posibile sunt cele două culori.

Se observă că , în experimentele de mai sus rezultatul nu poate fi prezis cu exactitate înaintea efectuării experimentului. Vom numi asemenea experimente, *experimente aleatoare*. O realizare (concretă sau imaginată) a unui experiment aleator se numeşte *probă* . In urma unei probe unul (şi numai unul) dintre rezultatele posibile ale experimentului se va realiza. Acesta va fi rezultatul probei respective (de exemplu, la o lansare concreta a unui zar s-a obţinut numărul 6).

Evenimente. Fie Ω mulţimea rezultatelor posibile ale unui experiment aleator. Vom considera pentru început că Ω este o mulţime finită , cazul general fiind prezentat în lecţiile următoare. Vom numi *eveniment* orice submulţime $A \subseteq \Omega$. Un eveniment se realizează într-o probă dacă rezultatul probei aparţine evenimentului.

Exemple. 1) $\emptyset$ şi Ω sunt evenimente; $\emptyset$ nu se realizează în nici o probă numindu-se astfel *evenimentul imposibil*, Ω se realizează în orice probă şi se numeşte *evenimentul sigur*.

2) Dacă $\omega \in \Omega$ (deci ω este un rezultat posibil al experimentului), atunci mulţimea $\{\omega\}$ este un eveniment, numit *eveniment elementar* (notăm deosebirea dintre un rezultat posibil şi un eveniment elementar).

3) Fie o monedă cu feţele notate a, b. Considerăm experimentul aleator al aruncării de două ori a monedei cu mulţimea de rezultate posibile $\Omega = \{(a,a), (a,b), (b,a), (b,b)\}$. Evenimentul $A = \{(a,b), (b,a)\}$poate fi interpretat ca enunţul "se obţin rezultate diferite la cele două aruncări".

Aşa cum se observă din ultimul exemplu evenimentele reprezintă , intuitiv, enunţuri anticipative privind rezultatele unei probe. Am preferat definirea evenimentului ca mulţime a rezultatelor posibile favorabile adeveririi unui enunţ din motive de precizie a limbajului.

Operaţii cu evenimente. Operaţiile cu mulţimi devin, evident, operaţii cu evenimente. Aceste operaţii au următoarea interpretare "logică ":

reuniunea $A \cup B$: evenimentul $A \cup B$ se realizează dacă şi numai dacă se realizează A *sau* se realizează B

intersecţia $A \cap B$: evenimentul $A \cap B$ se realizează dacă şi numai dacă se realizează A *şi* se realizează B

complementara CA: evenimentul CA se realizează dacă şi numai dacă *nu* se realizează A

Astfel conectorii logici fundamentali (în logica propoziţiilor) au interpretare naturală în algebra submulţimilor unei mulţimi. Evenimentul CA se zice evenimentul contrar evenimentului A (mai "scurt "contrar lui A).

Evenimentele A, B se zic *incompatibile dacă* $A \cap B = \emptyset$. Interpretarea este clară : A şi B nu se pot realiza simultan. Adaugăm relaţia de incluziune între evenimente: dacă $A \subseteq B$ înseamna că orice realizare a evenimentului A este şi o realizare a evenimentului B, deci putem interpreta incluziunea ca o implicaţie.

Exemplu. In experimentul lansării zarului ($\Omega = \{1, 2, 3, 4, 5, 6\}$) evenimentele $\{1, 3, 5\}, \{2, 4, 6\}$ sunt incompatibile (un rezultat al unei probe nu poate fi, simultan, par şi impar).

Câmp de evenimente. Vom trece la modelul matematic general al evenimentelor şi operaţiilor cu acestea. Cazul discutat mai sus este un caz particular important dar limitat dacă luăm în consideraţie amploarea problemelor teoretice şi aplicative puse de noţiunea de probabilitate. Evenimentele sunt interpretate ca submulţimi şi logica combinării lor ca operaţii cu submulţimi. Atât motive matematice cat şi interpretări concrete justifică faptul că , în general, evenimentele nu vor mai fi *toate submulţimile* unei mulţimi şi printre operaţiile cu evenimente apar operaţii "infinite". De asemenea, în general, interpretarea elementelor ca rezultate posibile nu mai are aceeaşi greutate ca în discuţia anterioară (totuşi intuiţia căpătată pe cazul numărului finit de rezultate posibile işi pastrează importanţa). Teoria probabilităţilor, în forma pe care o prezentăm, este o parte (la fel ca teoria lungimilor, ariilor si volumelor, maselor etc.) a teoriei generale a măsurii având desigur problematica ei specifică şi limbajul corespunzător. Unele noţiuni vor fi date în cadrul general al teoriei măsurii şi redenumite pentru cazul teoriei probabilităţilor.

3.1 Fie Ω o mulţime (notaţia este oarecum tradiţionala şi o vom păstra). Se numeşte **σ-algebra** pe Ω o mulţime $\Delta \subseteq P(\Omega)$ astfel încât:

i) $\Omega \in \Delta$

ii) Dacă $A \in \Delta$, atunci $CA \in \Delta$

iii) Dacă $(A_n)_n$ este un şir de mulţimi $A_n \in \Delta, \forall n \in \mathbb{N}$, atunci $\cup A_n \in \Delta$

Intuitiv vorbind σ-algebra este structura mulţimilor care vor fi măsurate (evenimentelor care vor avea probabilitate). Din definiţie deducem următoarele proprietăţi pentru o σ-algebra Δ:

1) $\emptyset \in \Delta$ şi $\Delta \neq \emptyset$

2) Dacă $A, B \in \Delta$, atunci $A \cup B \in \Delta$

3) Dacă $(A_n)_n$ este un şir de mulţimi $A_n \in \Delta, \forall n \in \mathbb{N}$, atunci $\cap A_n \in \Delta$

4) Dacă $A, B \in \Delta$, atunci $A \cap B \in \Delta$

5) Dacă $A, B \in \Delta$, atunci $A - B \in \Delta$

Demonstraţia acestor proprietăţi este simplă . De exemplu, să demonstrăm 2,3) şi 5):

2) Definim şirul $A_0 = A, A_1 = B, A_n = \emptyset, n \geq 2$ şi aplicăm iii) şi iv).

3) $C(\cap A_n) = \cup(CA_n)$ şi aplicăm ii) şi iii).

5) $A - B = A \cap CB$ etc.

Exerciţiu.1) Demonstraţi 1 şi 4.

O pereche (Ω, Δ), unde Δ este o $\sigma-$algebră pe Ω se numeşte spatiu măsurabil, iar mulţimile $A \in \Delta$ mulţimi măsurabile. In teoria probabilităţilor spaţiile măsurabile vor fi numite **câmpuri de evenimente**, iar mulţimile care le aparţin vor fi numite evenimente. Denumirile: **eveniment sigur**, **eveniment imposibil**,**evenimente contrare**, **evenimente incompatibile** păstrează acelaşi sens ca mai sus.

Exemple. 1). Fie $\Delta = \mathcal{P}(\Omega)$; evident condiţiile i), ii), iii) sunt satisfăcute. Regăsim astfel cazul discutat în secţiunea 2 ca un caz particular.

2) Fie $\emptyset \neq A \subset \Omega$; atunci $\Delta = \{\emptyset, A, CA, \Omega\}$ este o $\sigma-$algebră (pe Ω)

3) Fie Ω o mulţime (nevidă) şi $f : \Omega \to \mathbb{R}$ o functie.Fie $\Delta = \{f^{-1}(B)\, ; B \subseteq \mathbb{R}\}$.Este o verificare simplă şi un bun exerciţiu (toate proprietăţile necesare se află în acest text) de a arăta că Δ este o $\sigma-$algebră . Dacă $V \subseteq \mathbb{R}$ este mulţimea valorilor funcţiei f se arată că Δ constă din reuniuni arbitrare de mulţimi de nivel $f^{-1}(v)$,$v \in V$. Acest exemplu arată că pot apărea $\sigma-$algebre de interes, diferite de $\mathcal{P}(\Omega)$ chiar în cazul mulţimlor finite.

4) Fie $\Omega = \mathbb{R}$. Există o $\sigma-$algebră pe $\mathbb{R}$ care conţine intervalele (de orice tip) şi care este cea mai mică (în sensul incluziunii) $\sigma-$algebră cu această proprietate. Această $\sigma-$algebră se numeşte $\sigma-$algebră Borel şi mulţimie care îi aparţin se numesc boreliene.

Exerciţiu. 2) Fie $A \subseteq \Omega$. Să se determine cea mai mică $\sigma-$algebră care conţine A.

8.3 MF.08.3. Câmp de probabilitate.

Probabilitate. Intr-un experiment aleator (repetabil) ordonarea anticipărilor privind realizarea evenimentelor se face prin ataşarea unui număr real între 0 şi 1 fiecărui eveniment, număr reprezentând o măsură a gradului de aşteptare privind realizarea evenimentul într-o probă . Aceste cuvinte sunt, desigur, o descriere cu scop intuitiv. Vom considera două exemple clasice şi apoi vom da definiţiile riguroase.

Exemple. 1) Am văzut că la experimentul lansării unui zar mulţimea rezultatelor posibile este $\Omega = \{1, 2, 3, 4, 5, 6\}$. Vom considera câmpul de evenimente $(\Omega, \mathcal{P}(\Omega))$; în lipsa unor informaţii suplimentare sau din raţiuni de simetrie vom ataşa rezultatelor (de fapt evenimentelor respective) aceeaşi probabilitate $\frac{1}{6}$ şi, mai general, fiecărui eveniment A probabilitatea $P(A) = \frac{cardA}{n}$, unde $cardA$ este numărul de elemente al mulţimii A.

2) Fie un experiment aleator cu o mulţime finită Ω de rezultate posibile şi ataşăm câmpul de evenimente $(\Omega, \mathcal{P}(\Omega))$. Se repetă experimentul de n ori şi se consideră drept probabilitate a unui eveniment A frecvenţa relativă de realizare a evenimentului A, $P(A) = \frac{k}{n}$, unde k este numarul de realizări ale evenimentului A. Acest model este, desigur, nesatisfăcător probabilitatea depinzând de numărul de repetări ale experimentului. Punctul de vedere frecventist în teoria probabilităţilor susţine că probabilitatea este, într-un sens care rămâne să fie definit, limita frecvenţelor relative când numărul de repetări tinde la infinit. Vom întâlni o precizare a acestor afirmaţii la legea

numerelor mari. In orice caz, intuitia probabilităţii ca frecvenţă este foarte utilă .

Câmp de probabilitate. Fie (Ω, Δ) un spaţiu măsurabil. O funcţie $\mu : \Delta \to [0, \infty]$ este o **măsură** dacă :

i) $\mu(\emptyset) = 0$

ii) Dacă $(A_n)_n$ este un şir de mulţimi din Δ, disjuncte două câte două ($A_i \cap A_j = \emptyset$ dacă $i \neq j$), atunci $\mu(\cup A_n) = \sum\limits_n \mu(A_n)$ (dacă $\mu(A_n) = \infty$ pentru cel puţin un indice, atunci $\sum \mu(A_n) = \infty$, dacă $\mu(A_n) \neq \infty, \forall n$ şi seria $\sum \mu(A_n)$ este divergentă , atunci $\sum \mu(A_n) = \infty$, dacă seria este convergentă , atunci $\sum \mu(A_n)$ este suma seriei).

Proprietatea ii) se numeşte "*numărabil aditivitate*".

O măsură pentru care $\mu(\Omega) = 1$ se numeşte **probabilitate**. O vom nota p. Vom numi **câmp de probabilitate** tripleta (Ω, Δ, P). Numărul $P(A)$ se numeşte **probabilitatea evenimentului** A. Noţiunea de câmp de probabilitate este fundamentală pentru teoria probabilităţilor.

Proprietăţile esenţiale ale unui câmp de probabilitate sunt:

1) Dacă $A, B \in \Delta, A \cap B = \emptyset$, atunci $P(A \cup B) = P(A) + P(B)$ (finit aditivitate).

2) $P(CA) = 1 - P(A)$.

3) Dacă $A, B \in \Delta$, atunci $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

4) Dacă $A, B \in \Delta, A \subseteq B$, atunci $P(A) \leq P(B)$.

5) Dacă $(A_n)_n$ este un şir crescător ($A_n \subseteq A_{n+1}, \forall n$) de evenimente, atunci $P(\cup A_n) = \lim_{n\to\infty} P(A_n)$.

6) Dacă $(A_n)_n$ este un şir descrescător ($A_n \supseteq A_{n+1}, \forall n$) de evenimente, atunci $P(\cap A_n) = \lim\limits_{n\to\infty} P(A_n)$.

Vom demonstra 1) şi 3):

3) Considerăm şirul $A_0 = A, A_1 = B, A_n = \emptyset, n \geq 2$ şi aplicăm i) şi ii).

5) $A \cup B = A \cup (B - A)$ şi aplicând 3) obţinem $P(A \cup B) = P(A) + P(B - A)$, dar $P(B - A) = P(B) - P(A \cap B)$ etc.

Remarcăm că , dacă evenimentele A, B nu sunt incompatibile, probabilitatea reuniunii nu se poate calcula numai din probabilităţile acestor evenimente (acelaşi lucru pentru intersecţie).

Se arată cu uşurinţă ca în cazul în care Δ este o mulţime finită , finit aditivitatea implică numărabil aditivitatea.

Exemple. 1) (**Egal probabilitatea**) Fie $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ şi $\Delta = \mathcal{P}(\Omega)$. Pentru $A \in \mathcal{P}(\Omega)$ definim $P(A) = \frac{cardA}{n}$, unde $cardA$ este numărul de elemente ale mulţimii A. Este un exerciţiu simplu verificarea condiţiilor i), ii), deci se obţine un câmp de probabilitate. Acest model este datorat lui Laplace. Probabilitatea unui eveniment este raportul dintre numărul de rezultate favorabile evenimentului şi numărul de rezultate posibile (se mai spune "raportul dintre numărul de cazuri favorabile evenimentului şi numărul de cazuri posibile). Modelul egal probabilităţii este justificat de lipsa de informaţie (sau de "motive") pentru a deosebi probabilistic rezultatele posibile. Desigur, acest model de câmp de probabilitate generalizează exemplul zarului prezentat mai sus. Alte modele de acest tip sunt urnele cu bile sau aruncarea unei monede.

2) Vom generaliza modelul din exemplul precedent prin descrierea tuturor câmpurilor de probabilitate pe o mulţime finită Ω şi $\Delta = \mathcal{P}(\Omega)$. Se observă că este suficient să cunoaştem probabilităţile $P(\{\omega_i\}) = p_i, i = 1, 2, ..., n$, unde $0 \leq p_i \leq 1, p_1 + p_2 + ...p_n = 1$. In adevăr, din definiţia probabilităţii rezultă $P(A) = \sum_{\omega_i \in A} p_i$. A da numerele $p_i, i = 1, 2, ..., n$ satisfacând condiţiile de mai sus înseamnă a da o **repartiţie de probabilitate** pe Ω.

3) Fie Ω o mulţime numărabilă cu elementele aşezate într-un şir $(\omega_n)_n$. Să luăm $\Delta = \mathcal{P}(\Omega)$. A da o probabilitate în acest context revine la a da un şir $(p_n)_n$ de numere între 0 şi 1 astfel încât $\sum_n p_n = 1$. Justificarea acestei afirmaţii se face ca în exemplul precedent. Şi în acest caz se foloseşte termenul **repartiţie de probabilitate.**

4) Fie $\Omega = [0, 1]$ şi Δ $\sigma-$algebra mulţimilor boreliene incluse în intervalul $[0, 1]$. Noţiunea de lungime a unui interval se extinde la mulţimile boreliene obţinând o măsură (măsura Lebesgue); nu vom intra în amănunte (vezi MF.05). Vom nota $P(A)$ măsura mulţimii $A \in \Delta$ $(P([a, b]) = b - a)$. Se obţine o probabilitate. Acest exemplu poate fi considerat ca un exemplu de probabilitate geometrică . O interpretare posibilă a acestui model este experimentul aleator "se alege, la întâmplare, un număr din intervalul $[0, 1]$"; probabilitatea ca punctul să aparţină unui interval este chiar lungimea acelui interval. Un fenomen interesant apare în legătură cu această interpretare: probabilitatea ca un punct anume să fie ales este nulă deşi, după o probă , un anume punct va fi ales; cu alte cuvinte sunt posibile evenimente de probabilitate 0. Trebuie astfel distins intre eveniment imposibil şi eveniment de probabilitate 0. Analog există evenimente de probabilitate 1 care nu sunt sigure (de exemplu, $[0, 1)$). Aceste evenimente se zic **aproape sigure**. Ar fi util de menţionat că nu orice submulţime a intervalului $[0, 1]$ este boreliană , dar nu vom intra în amanunte.

5) Fie câmpul de evenimente $(\Omega, \mathcal{P}(\Omega))$ şi $a \in \Omega$. Definim $\delta_a(A) = 1$ dacă $a \in A$ şi $\delta_a(A) = 0$ dacă $a \notin A$. Se obţine o probabilitate (exerciţiu) numită măsura Dirac concentrată în a.

6) Fie $(p_i)_{1 \leq i \leq n}$ o repartiţie de probabilitate pe Ω_1 şi $(q_j)_{1 \leq j \leq m}$ o repartiţie de probabilitate pe Ω_2 (pentru simplitate vom nota elementele celor două mulţimi cu indici). In acest caz, notând $p_{ij} = p_i q_j$, se obţine o repartiţie de probabilitate pe produsul cartezian $\Omega_1 \times \Omega_2$. Numim această repartiţie repartiţia produs a celor două repartiţii iniţiale. Să observăm că repartiţiile iniţiale se obţin din repartiţia produs ca "repartiţii marginale" $p_i = \sum_j p_{ij}, q_j =$

$\sum_i p_{ij}$.

7) Fie mulţimile finite Ω_1, Ω_2 şi $(p_{ij})_{i,j}$ o repartiţie de probabilitate pe $\Omega_1 \times \Omega_2$ (notaţii ca în exemplul precedent). Definim $p_i = \sum_j p_{ij}, q_j = \sum_i p_{ij}$. Atunci $(p_i)_i$ este o repartiţie de probabilitate pe Ω_1 şi $(q_j)_j$ o repartiţie de probabilitate pe Ω_2; aceste repartiţii se zic marginale. In general repartiţia iniţială nu este produsul repartiţiilor marginale.

Exerciţiu. Două persoane φ, ψ se vor întâlni într-un anumit loc între orele 12 şi 13. Sosit, fiecare aşteaptă 20 de minute şi dacă celălalt nu soseşte, pleacă . Care este probabilitatea ca cei doi să se întâlnească ştiind că ei sosesc aleator şi independent unul de altul?

Soluţie. Vom interpreta precizările "aleator" şi "independent" în mod intuitiv. Notând cu x, y timpii de sosire a celor două persoane, mulţimea sosirilor posibile va fi (luând unitatea de măsură a timpului minutul) mulţimea perechilor de numere reale (x, y) cu $x, y \in [0, 60]$ (deci un pătrat în plan). Mulţimea sosirilor favorabile întâlnirii este formată din perechile $(x, y), |x - y| \leq 20$. Vom considera că probabilitatea cerută este raportul dintre aria mulţimii sosirilor favorabile şi aria pătratului (o formă geometrică a egal probabilului, a se vedea şi exemplul 4) de mai sus). Un calcul simplu arată că probabilitatea respectivă este $\frac{5}{9}$.

8.4 MF.08.4. Probabilităţi condiţionate, independenţă . Formula Bayes

Probabilităţi condiţionate. In cele ce urmează vom cosidera un câmp de probabilitate fixat (Ω, Δ, P).

Fie $B \in \Delta$ astfel încât $P(B) \neq 0$. Definim, pentru fiecare eveniment $A \in \Delta$ **probabilitatea lui** A **condiţionată de** B prin $P(A|B) = \frac{P(A \cap B)}{P(B)}$. Din definiţie rezultă **formula de înmulţire** $P(A \cap B) = P(B) P(A|B)$. Interpretarea probabilităţii condiţionate $P(A|B)$ este: probabilitatea evenimentului A ştiind că evenimentul B se va realiza (sau s-a realizat). Să observam că pentru B fixat funcţia $A \longmapsto P(A|B), A \in \Delta$ este o probabilitate. In adevăr, evident, $\emptyset \longmapsto 0$, apoi dacă $(A_n)_n$ un şir de evenimente, disjuncte două câte două obţinem $P(\cup A_n|B) = \frac{P((\cup A_n) \cap B)}{P(B)} = \frac{P(\cup(A_n \cap B))}{P(B)} = \sum_n \frac{P(A_n \cap B)}{P(B)} = \sum_n P(A_n|B)$; în sfârşit este clar că $P(\Omega|B) = 1$. Intuitiv vorbind, probabilitatea condiţionată de B reprezintă o reevaluare a probabilităţilor în prezenţa unei informaţii noi.

Să presupunem că $P(A), P(B) \neq 0$. Rezultă : $P(A \cap B) = P(B) P(A|B) = P(A) P(B|A)$, probabilitatea intersecţiei este egală cu probabilitatea unuia dintre evenimente înmulţită cu probabilitatea celuilat condiţionată de primul.

Independenţa. Evenimentele A, B se zic independente dacă $P(A \cap B) = P(A) P(B)$. Independenţa este o proprietate mutuală care face să intervină în mod esenţial probabilitatea. Dacă A, B sunt independente şi $P(B) \neq 0$ obţinem $P(A|B) P(B) = P(A) P(B)$ şi deci $P(A|B) = P(A)$, astfel că realizarea evenimentului B nu influenţează probabilitatea evenimentului A (o formulare mai intuitiva a faptului că A nu depinde de B).

Exerciţiu. 1) Dacă evenimentele A, B sunt independente, atunci sunt independente A, CB şi CA, B şi CA, CB.

Exemple. 1) Fie $\Omega = \{a, b, c, d\}, \Delta = \mathcal{P}(\Omega)$ şi două probabilităţi pe $(\Omega, \Delta)\, p$ este egal probabilitatea, iar q corespunde repartiţiei $\frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{4}$. Să considerăm evenimentele $A = [a, b\}, B = [b, c\}$; avem $A \cap B = \{b\}$. In raport cu probabilitatea $p,p(A) = \frac{1}{2}, p(B) = \frac{1}{2}, p(A \cap B) = \frac{1}{4}$, deci evenimentele sunt independente. In raport cu probabilitatea $q, q(A) = \frac{7}{12}, q(B) = \frac{5}{12}, q(A \cap B) = \frac{1}{4}$ se vede că $q(A \cap B) \neq q(A) q(B)$, deci A, B nu sunt independente.

2) Intr-un câmp de probabilitate (Ω, Δ, P) evenimentele A, Ω sunt independente (oricare ar fi $A \in \Delta$);la prima vedere acest fapt poate părea ciudat.

In general independenţa evenimentelor $A_1, A_2, ..., A_n, n \geq 2$ se defineşte astfel: pentru orice $2 \leq m \leq n$ şi orice indici $1 \leq i_1 < i_2 < ... < i_r \leq n$ să rezulte $P(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_r}) = P(A_{i_1}) P(A_{i_2}) ... P(A_{i_r})$.

Formula probabilităţii totale. Fie $A_1, A_2, ..., A_n$ evenimente astfel încât:

sunt incompatibile două câte două , $P(A_i) \neq 0, i = 1, 2, ..., n$ şI $A_1 \cup A_2 \cup ... \cup A_n = \Omega$. In acest caz vom spune că $A_1, A_2, ..., A_n$ formează un **sistem complet de evenimente.** Pentru orice eveniment B are loc:

$$P(B) = \sum_{1}^{n} P(A_i) P(B|A_i) \quad (\textbf{formula probabilităţii totale}).$$

Demonstraţia este simplă : avem $B = \cup(B \cap A_i)$ şi se aplică finit aditivitatea probabilităţii şi formula de înmulţire a probabilităţilor.

Exemplu(problemă). Fie 5 urne din care: 2 urne conţin 2 bile albe şi o bilă neagră fiecare, (compoziţia A_1), o urnă conţine 10 bile negre (compoziţia A_2) şi 2 urne cu 3 bile albe şi o bilă neagră , fiecare (compoziţia A_3). Se consideră următorul experiment aleator: dintr-o urnă aleasă la întâmplare se extrage, la întâmplare, o bilă . Care este probabilitatea ca bila să fie albă ?

Soluţie. Fie B evenimentul extragerii unei bile albe. Evenimentele extragerii din urne de compoziţiile date formează un sistem complet de evenimente (notate la fel ca şi compoziţiile respective). Avem:

$P(A_1) = \frac{2}{5}, P(A_2) = \frac{1}{5}, P(A_3) = \frac{2}{5}$ (egal probabilitate pe urne), $P(B|A_1) = \frac{2}{3}, P(B|A_2) = 0, P(B|A_{31}) = \frac{3}{4}$.

Aplicând formula probabilităţii totale obţinem $p(B) = \frac{2}{5}\frac{2}{3} + \frac{1}{5}0 + \frac{2}{5}\frac{3}{4} = \frac{17}{30}$.

Această rezolvare, deşi corectă , nu pune în evidenţă clar câmpul de probabilitate pe care se lucrează ; se pare că se lucrează când cu urne când cu bile etc. Pentru lămurirea acestei situaţii este util să se considere ca rezultat al experienţei o pereche formată din compoziţia urnelor şi din culoarea bilelor. Deci $\Omega = \{(A_1, a), (A_2, a), (A_3, a), (A_1, b), (A_2, b), (A_3, b)\}$ şi luăm $\Delta = \mathcal{P}(\Omega)$. Introducem repartiţia de probabilitate $P(A_1, a) = \frac{2}{5}\frac{2}{3}$ (probabilitatea compoziţiei înmulţită cu probabilitatea extragerii tipului de bilă din compoziţia respectivă) etc. Se obţine un câmp de probabilitate în care $B = \{(A_1, a), (A_2, a), (A_3, a)\}$ şi sistemul complet de evenimente $\{(A_1, a), (A_1, b)\}$ (notat mai sus A_1) etc. Se verifică uşor că se ajunge la acelaşi rezultat ca în calculul "intuitiv" de mai sus.

Schema bilei neîntoarse. O urna conţine n bile dintre care m bile roşii şi restul albe în rest identice. Se extrage, fără a privi în urnă , o bilă ; se notează culoarea, se pune bila deoparte şi se repetă procedeul de $k \leq n$ ori. Vom nota R_i evenimentul "o bilă roşie este extrasă la pasul i" şi A_i evenimentul analog pentru o bilă albă . Ne propunem să calculăm $P(R_2|R_1)$ şi $P(R_1|R_2)$.

Avem: $P(R_1) = \frac{m}{n}, P(A_1) = \frac{n-m}{n}, P(R_1 \cap R_2) = P(R_1) P(R_2|R_1) = \frac{m}{n}\frac{m-1}{n-1}$, deci $P(R_2|R_1) = \frac{m-1}{n-1}$

$P(R_2) = P(A_1) P(R_2|A_1) + P(R_1) P(R_2|R_1) = \frac{n-m}{n}\frac{m}{n-1} + \frac{m}{n}\frac{m-1}{n-1} = \frac{m}{n}$, $P(R_1|R_2) = \frac{P(R_1 \cap R_2)}{P(R_2)} = \frac{m-1}{n-1}$.

Deci $P(R_2|R_1) = P(R_1|R_2)$. Explicaţi, intuitiv, această simetrie oarecum surprinzătoare. Construiţi câmpul de probabilitate care este la baza raţionamentului de mai sus (cu atât mai necesar cu cât rezultatul obţinut este neaşteptat).

Formula Bayes. Fie $A_1, A_2, ..., A_n$ un sistem complet de evenimente şi B un eveniment cu $P(B) \neq 0$. Avem $P(A_i|B) = \frac{P(B)P(B|A_i)}{P(B)}, i = 1, 2, ..., n$. Folosind formula probabilităţii totale obţinem :

$$P(A_i|B) = \frac{P(B)P(B|A_i)}{\sum\limits_{j=1}^{n} P(B)P(B|A_j)} \text{ formula Bayes.}$$

Formula Bayes se mai numeşte şi formula probabilităţilor ipotezelor; intuitiv vorbind, asupra rezultatului unui eveniment aleator se fac anumite ipoteze "apriori" $A_1, A_2, ..., A_n$ cu anumite probabilităţi. In urma unei probe s-a realizat evenimentul B. In aceste condiţii probabilităţile ipotezelor se "reevaluează "conform formulei Bayes. Realizarea evenimentului B a sporit informaţia disponibilă şi a permis reevaluarea ipotezelor. Vedem aici modelul procesului de învăţare atât de important în cunoaşterea umană .

Exemplu (problemă). Fie 5 urne din care: 2 urne conţin 2 bile albe şi 3 bile negre fiecare, (compoziţia A_1), 2 urne cu o bilă alba şi 4 bile negre fiecare (compoziţia A_3), o urnă conţine 10 bile negre (compoziţia A_2). Dintr-o urnă aleasă la întâmplare se extrage, la întamplare, o bilă ; bila este albă . Care este probabilitatea ca bila să fi fost extrasă din urna de compoziţie A_3?

Soluţie. Din formula Bayes deducem $P(A_3|B) = \frac{P(A_3)P(B|A_3)}{\sum_{j=1}^{3} P(A_j)P(B/A_j)}$ în care probabilităţile necesare se calculează ca în problema precedentă . Lăsăm ca exerciţiu finalizarea calculului.

Exerciţiu. 2) Vom considera o problemă de diagnostic (medical) utilizând notaţii mai des întâlnite în practică . Fie mulţimea $H = \{h, h^{'})$ mulţimea ipotezelor (h este ipoteza că pacientul are o anumită boală , iar $h^{'}$ ipoteza contrară (nu are boala)). Din statistici se obţîn probabilităţile iniţiale $P(h) = 0,008, P\left(h^{'}\right) = 0,992$. Un test de laborator Θ dă rezultatele posibile $\oplus, \ominus$ astfel încât $P(\oplus|h) = 0,98, P(\ominus|h) = 0,02, P\left(\oplus|h^{'}\right) = 0,03, P\left(\ominus|h^{'}\right) = 0,97$. Se face un test care dă rezultatul $\oplus$. Să se calculeze $P(h|\oplus)$.

Pentru acest capitol se recomandă lucrările: $[B.01], :$ $[C.01], :$ $[D.01], :$ $[F.01], :$ $[B.01], :$ $[G.01], :$ $[G.03], :$ $[B.01], :$ $[I.01], :$ $[B.01], :$ $[L.01], :$ $[B.01], :$ $[P.02], [S.01], :$ [Ş.01].

Capitolul 9

MF.09. Variabile aleatoare

Cuvinte cheie

Măsurabilitate, variabilă aleatoare, media variabilelor aleatoare, funcţie de repartiţie, densitate de repartiţie, variabile discrete, variabile continue, momente.

Acest capitol conţine şi o introducere cu caracter pregătitor în vederea definirii momentelor variabilelor aleatoare. Se prezintă succint integrarea funcţiilor pe un câmp de probabilitate. Pentru cazul general al integrării funcţiilor pe spaţii cu măsura trimitem la cursurile de Analiză Matematică şi Analiză funcţională . Vom fixa un câmp de probabilitate (Ω, Δ, p). Vom nota variabilele aleatoare conform cu standardele teoriei probabilităţilor X, Y etc. cu o excepţie relativă la funcţiile simple (funcţii cu o mulţime finită de valori).

9.1 MF.09.1. Variabile aleatoare, medie.

Funcţii măsurabile (variabile aleatoare). O funcţie $X : \Omega \to \mathbb{R}$ se zice **măsurabilă** dacă pentru orice interval I avem $X^{-1}(I) \in \Delta$. In cadrul teoriei probabilităţilor funcţiile masurabile pe campuri de probabilitate se numesc **variabile aleatoare**. Se poate arăta că , pentru o variabilă aleatoare X rezultă $X^{-1}(B) \in \Delta$ pentru orice mulţime boreliană B. Vom folosi şi notaţia $(X \in B)$ pentru $X^{-1}(B)$.

Intuitiv, putem interpreta o variabilă aleatoare ca o funcţie care ataşează fiecărui rezultat al unei experienţe aleatoare un număr real (rezultat, de exemplu, dintr-o măsurare) în aşa fel încât mulţimea acelor $\omega \in \Omega$ pentru care $X(\omega) \in B$ este un eveniment (pentru orice mulţime boreliană B). Astfel spre exemplu mulţimea "condiţiilor" care determină ca temperatura să se afle într-un anume interval constituie un eveniment.

Exemple.1) Dacă $\Delta = \mathcal{P}(\Omega)$, atunci orice funcţie $X : \Omega \to \mathbb{R}$ este o variabilă aleatoare.

2) Dacă χ_A este indicatorul mulţimii $A \subseteq \Omega$, atunci χ_A este variabilă aleatoare dacă şi numai dacă $A \in \Delta$.

Exerciţii. 1) $X : \Omega \to \mathbb{R}$ este variabilă aleatoare $\Longleftrightarrow (X \in (-\infty, \alpha)) \in \Delta, \forall \alpha \in \mathbb{R} \Longleftrightarrow (X \in (-\infty, \alpha]) \in \Delta, \forall \alpha \in \mathbb{R}$ ("$\Longleftrightarrow$" înseamnă "dacă şi numai dacă "). Analog cu intervale $(\alpha, \infty), [\alpha, \infty)$ etc.

2) O funcţie $g : \mathbb{R} \to \mathbb{R}$ se zice măsurabilă dacă $g^{-1}(B)$ este boreliană pentru orice mulţime boreliană B. (admitem, de exemplu, că dacă $g : \mathbb{R} \to \mathbb{R}$ este continuă atunci este măsurabilă). Să se arate că dacă $X : \Omega \to \mathbb{R}$ este o variabilă aleatoare şi $g : \mathbb{R} \to \mathbb{R}$ este măsurabilă atunci $g \circ X$ este o variabilă aleatoare. Admitem că dacă $g : \mathbb{R} \to \mathbb{R}$ este continuă atunci este măsurabilă .

Listăm, fără demonstraţie, proprietăţile variabilelor aleatoare (demonstraţiile nu sunt dificile şi pot fi luate ca exerciţiu):

Proprietăţile variabilelor aleatoare. Fie X, Y variabile aleatoare şi $\alpha \in \mathbb{R}$:

1) $X + Y, XY, \alpha X$ sunt variabile aleatoare.

2) Dacă $X(\omega) \neq 0, \forall \omega \in \Omega$, atunci $\frac{1}{X}$ este o variabilă aleatoare.

3) Fie $(X_n)_n$ un şir de variabile aleatoare astfel încât $X_n \xrightarrow{p} X$ (convergenţa punctuală). Atunci X este o variabilă aleatoare.

4) O funcţie simplă are, prin definiţie o mulţime finită de valori (reale). S-a arătat că funcţiile simple sunt de forma $s = \sum_{i}^{n} \alpha_i \chi_{A_i}$, unde $\alpha_1, \alpha_2, ..., \alpha_n$ sunt valorile funcţiei s, iar $A_1, A_2...A_n$ multimile de nivel; s este o variabilă aleatoare dacă şi numai dacă $A_i \in \Delta, i = 1, 2, ..., n$. Dacă X este o variabilă aleatoare pozitivă , atunci există un şir crescător $(s_n)_n$ de variabile aleatoare simple astfel încât $s_n \xrightarrow{p} X$.

5) Dacă X este o variabilă aleatoare, atunci şi $|X|$ este o variabilă aleatoare ($|X|(\omega) = |X(\omega)|$).

6 Dacă X, Y sunt variabile aleatoare, atunci $\max(X, Y)$ şi $\min(X, Y)$ sunt variabile aleatoare. Reamintim că $\max(X, Y)(\omega) = \max(X(\omega), Y(\omega)), \omega \in \Omega$ etc.

Integrarea. Fie $s = \sum_{i}^{n} \alpha_i \chi_{A_i}$ o variabilă aleatoare simplă cu valori pozitive. Definim $\int_\Omega s dP = \sum_{i}^{n} \alpha_i P(A_i)$. Fie X o variabilă aleatoare cu valori pozitive. Definim $\int_\Omega X dP = \sup_{s \leq X} \int_\Omega s dP$ (marginea superioară se ia după toate variabilele aleatoare simple mai mici ca X). In general vom avea $\int_\Omega X dP \in [0, \infty]$. Vom spune că X este integrabilă dacă $\int_\Omega X dP < \infty$. Pentru o variabilă aleatoare oarecare X definim $X^+ = \max(X, 0), X^- = -\min(X, 0)$. Evident $X = X^+ - X^-, |X| = X^+ + X^-$.

Spunem că X este integrabilă dacă $|X|$ este integrabilă ; în acest caz definim $\int_\Omega XdP = \int_\Omega X^+dP - \int_\Omega X^-dP$ (cele două integrale sunt finite aşa cum rezultă imediat).

Medie. Dacă variabila aleatoare X este integrabilă numim **medie** a sa numărul real $\int_\Omega XdP$. Vom nota media variabilei aleatoare X cu $M[X]$. Vom spune că X are medie dacă X este integrabilă .

Exemple. 1) Dacă $A \in \Delta$, atunci $M[\chi_A] = P(A)$.

2) Dacă c este variabilă aleatoare constantă ($c(\omega) = c, \forall\omega \in \Omega$), atunci $M[c] = c$ (atenţie la abuzul de notaţie în care c denotă fie o funcţie constantă fie un număr real).

3) Să reluăm exemplul $\Omega = [0,1]$, Δ mulţimile boreliene şi p extinderea lungimii intervalelor. Fie variabila aleatoare $X = \chi_{\mathbb{Q}\cap[0,1]}$ (deci funcţia care ia valoarea 1 pe $\mathbb{Q}\cap[0,1]$ şi 0 in rest). Rezultă $M[X] = 0$ (căci $p(\mathbb{Q}\cap[0,1]) = 0$ (probabilitatea unui "punct" este 0 iar $\mathbb{Q}\cap[0,1]$ este numărabilă) şi se aplică proprietatea de numărabil aditivitate a probabilităţii). Este demn de remarcat că funcţia despre care discutăm furnizează un exemplu tipic de funcţie care nu este integrabilă Riemann; cu ajutorul noii definiţii a integralei această funcţie se poate integra.

Vom reveni cu exemple dupa listarea principalelor proprietăţi ale integralei (şi în capitolul următor).

Proprietăţile mediei. Fie X, Y variabile aleatoare şi $\alpha \in \mathbb{R}$.

1) Dacă X, Y au medie, atunci $X+Y$ are medie şi $M[X+Y] = M[X] + M[Y]$

2) Dacă X are medie, atunci αX are medie şi $M[\alpha X] = \alpha M[X]$

3) Dacă X, Y au medie şi $X \leq Y$, atunci $M[X] \leq M[Y]$

4) X are medie dacă şi numai dacă $|X|$ are medie şi $|M[X]| \leq M[|X|]$

5) Orice variabilă aleatoare marginită are medie.

6) Dacă $0 \leq X_1 \leq X_2 \leq ... \leq X_n \leq ...$si $X_n \xrightarrow{p} X$ şi X are medie, atunci $\lim\limits_{n\to\infty} M[X_n] = M[X]$ (teorema de convergenţă monotonă).

7) Fie X o variabilă aleatoare pozitivă şi $(X_n)_n$ un şir crescător de variabile aleatoare simple pozitive $X_n \xrightarrow{p} X$. Sirul $(M[X_n])_n$ este mărginit dacă şi numai dacă X are medie şi în acest caz $\lim\limits_{n\to\infty} M[X_n] = M[X]$.

8) Fie Φ o variabilă aleatoare astfel încât $M[\Phi] = 1$. Definim $q : \Delta \to \mathbb{R}, q(A) = M[\Phi\chi_A] = \int_\Omega \Phi\chi_A dp = \int_A \Phi dp$ (ultima egalitate fiind o definiţie). Atunci q este o probabilitate.

Unele dintre proprietăţile de mai sus (1,2,3,4,5) au demonstraţii uşoare prin aplicarea directă a definiţiilor. Proprietatea 6 este o teorema celebră datorată matematicianului Lebesgue.

Sa revenim la exemple.

Exemple. 4) Fie Ω o mulţime finită (nevidă) cu n elemente. Considerăm câmpul de egală probabilitate asociat. Evident, în acest caz, orice variabilă aleatoare are medie şi dacă X are valorile (nu neapărat distincte)

$\alpha_1, \alpha_2, ..., \alpha_n$, atunci $M[X] = \frac{\alpha_1+\alpha_2+...+\alpha_n}{n}$, deci media aritmetică a valorilor variabilei.

5) Mai general decât în exemplul precedent fie $p_1, p_{2,...,p_n}$ o repartiţie de probabilitate pe o mulţime cu n elemente şi câmpul de probabilitate asociat. Atunci orice variabilă aleatoare are medie şi dacă X are valorile $\alpha_1, \alpha_2, ..., \alpha_n$ vom avea $M[X] = \alpha_1 p_1 + \alpha_2 p_2 + ... + \alpha_n p_n$, deci media "ponderată " a valorilor variabilei.

6) Fie $(p_n)_n$ o repartiţie de probabilitate pe o mulţime numarabilă şi câmpul de probabilitate asociat. Atunci o variabilă aleatoare X cu valorile $\alpha_0, \alpha_1, ..., \alpha_n, ...$ (un şir) are medie dacă şi numai dacă seria $\sum_0^\infty \alpha_n p_n$ este absolut convergentă şi, în acest caz, $M[X] = \sum_0^\infty \alpha_n p_n$.

Exerciţii. 3) Demonstraţi afirmaţiile din exemplele 5,6,7. Pentru exemplul 7 analizaţi, pentru inceput, cazul variabilelor aleatoare pozitive folosind proprietatea 7 (proprietăţile mediei) şi considerând şirul $(X_n)_n$, unde X_n ia valorile $\alpha_0, \alpha_1, ..., \alpha_n, 0, 0, 0....$

4) Consideraţi măsura Dirac concentrată în a. Arătaţi că orice variabilă aleatoare X are medie şi $M[X] = X(a)$ (evaluarea funcţiilor într-un punct este un caz particular de integrare).

Vom nota (oarecum ambiguu) mulţimea variabilelor aleatoare care au medie cu L^1; deci $X \in L^1$ va însemna ca X este integrabilă . Din proprietăţile mediei rezultă că L^1 este un spaţiu vectorial peste $\mathbb{R}$ şi luarea mediei este o aplicaţie liniară definită pe L^1 cu valori reale.

9.2 MF.09.2. Variabile aleatoare, funcţie de repartiţie, densitate de repartiţie, momente.

Vom continua, în acest capitol, studiul variabilelor aleatoare. Pentru simplitate vom scrie v.a ca prescurtare pentru "variabila aleatoare".(Ω, Δ, P) va fi un câmp de probabilitate fixat.

Repartiţie. Vom nota $\mathcal{B}$ $\sigma-$algebra mulţimilor boreliene pe $\mathbb{R}$; o probabilitate $p : \mathcal{B} \to [0, 1]$ va fi numită **repartiţie** sau **lege de probabilitate**. Inainte de a studia câteva legi de probabilitate des întâlnite vom arăta cum se asociază o repartiţie unei variabile aleatoare.

Fie $X : \Omega \to \mathbb{R}$ o v.a; pentru orice mulţime boreliană $S \in \mathcal{B}$ definim $p_x(S) = p(X \in S) = p\left(X^{-1}(S)\right)$. Se observă cu uşurinţă ca p_X este o probabilitate (de exemplu $p_X(\cup S_n) = \sum p_x(S_n)$ pentru orice şir $(S_n)_n$ de mulţimi boreliene disjuncte două câte două rezultă din faptul că $X^{-1}(\cup S_n) = \cup X^{-1}(S_n)$ ţinând cont că p este o probabilitate). Astfel se ataşează oricărei v.a X repartiţia p_X numită repartiţia v.a X şi implicit câmpul de probabilitate $(\mathbb{R}, \mathcal{B}, p_X)$. In general informaţiile probabilistice despre v.a X se pot obţine cunoscând repartiţia asociată ; cu alte cuvinte, două v.a cu aceeaşi repartiţie se pot considera "probabilistic" identice.

Exercitiu.1) Daţi exemplu de v.a diferite (ca funcţii) care au aceeaşi repartiţie.

Fie $g : \mathbb{R} \rightarrow \mathbb{R}$ masurabilă , X o v.a şi $Y = g \circ X = g(X)$ (ultima egalitate este o notaţie). Atunci $p_Y(S) = p_X\left(g^{-1}(S)\right)$ (verificare imediată).

Funcţie de repartiţie a unei v.a. Dacă X este o v.a definim funcţia $F_X : \mathbb{R} \rightarrow \mathbb{R}, F_X(x) = p(X < x) = p_x(-\infty, x)$.

Funcţia F_X se numeşte funcţia de repartiţie a v.a X.

Proprietăţile funcţiei de repartiţie.

1) $F_X(x) \in [0,1], \forall x \in \mathbb{R}$

2) F_X este o funcţie crescătoare.

3) F_X este continuă la stânga ($\lim\limits_{x<a} F_X(x) = F_X(a)$).

4) $\lim\limits_{x\rightarrow-\infty} F_X(x) = 0, \lim\limits_{x\rightarrow\infty} F_X(x) = 1$

5) $p(X \in [a,b)) = F_X(b) - F_X(a)$ pentru orice $a, b \in \mathbb{R}, a \leq b$

Pentru demonstraţie observăm că 1,2 sunt imediate din definiţii. Pentru proprietatea 3 să considerăm un şir $x_n \nearrow a$ (şirul este crescător cu limita a). Avem $(-\infty, a) = \cup(-\infty, x_n)$ (şirul de mulţimi este crescător); aplicând proprietatea 5 a câmpurilor de probabilitate obţinem rezultatul dorit. Analog se demonstrează 4. Proprietatea 5 este o consecinţă imediată a egalităţii $[a,b) = (-\infty,b) - (-\infty,a)$.

Se poate arăta că orice funcţie cu proprietăţile 1,2,3 de mai sus generează o repartiţie (în particular este funcţia de repartiţie a unei v.a).

Se arată , folosind proprietatea 6 a câmpurilor de probabilitate că limita la dreapta $F_X(a+)$ a funcţiei de repartiţie este $F_X(a+) = p(X \leq a)$. Deducem imediat că $p(X = a) = F_X(a+) - F_X(a)$ .Rezultă că dacă funcţia F_X este continuă , atunci $p(X = a) = 0$ pentru orice $a \in \mathbb{R}$.

Variabile aleatoare discrete. O v.a se zice **discretă** dacă mulţimea valorilor sale este finită sau numărabilă (în cazul numărabil cerem ca în fiecare interval compact să se afle o mulţime finită (eventual vidă) de valori). Dacă $\{x_i; i \in I\}$ este mulţimea valorilor notând $p_i = P(X = x_i)$ repartiţia de probabilitate corespunzătoare vom avea, pentru fiecare mulţime boreliană $S : P(X \in S) = p_X(S) = \sum\limits_{x_i \in S} p_i$. Astfel repartiţia unei v.a discrete este determinată de repartiţia de probabilitate pe valorile acesteia. Este uşor de observat că funcţia de repartiţie a unei v.a discrete este o funcţie în scară cu "salturi" în punctele corespunzătoare valorilor variabilei.

Exerciţii. 1) Demonstraţi afirmaţia din exemplul de mai sus.

2) Arătaţi că funcţia de repartiţie a v.a X care ia valorile 0 şi 1 cu probabilitate $0,7$, respectiv $0,3$ este: $F_X(x) = 0$ pentru $x \leq 0$, $F_X(x) = 0,7$ pentru $x \in (0,1]$ şi $F_X(x) = 1$ pentru $x > 1$. Reprezentaţi grafic această funcţie.

Variabile aleatoare continue. O v.a se zice **continuă** dacă există o funcţie măsurabilă $f : \mathbb{R} \to \mathbb{R}$ astfel incat:

i) $f(x) \geq 0, \forall x \in \mathbb{R}$

ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

iii) $P(X \in [a,b]) = \int_a^b f(x)\,dx$,$\forall a, b \in \mathbb{R}, a \leq b.$

In acest caz funcţia f este densitatea de probabilitate a v.a X. Se poate arăta că pentru orice mulţime boreliană S avem $P(X \in S) = p_X(S) = \int_S f(x)\,dx$. Spunem că probabilitatea p_x este dată de densitatea f în raport cu măsura Lebesgue.

Câteva precizări trebuie făcute. In definiţia de mai sus integrala este luata în sensul din capitolul V, chiar mai general, în raport cu o măsură care nu este probabilitate (acest caz se tratează similar cazului tratat în V) şi anume cu măsura Lebesgue pe $\mathbb{R}$. Nu vom intra în detalii privitor la această măsură (pentru detalii trimitem la cursurile de Analiză Matematică şi de Analiză Funcţională). Având în vedere că vom lucra cu funcţii continue (sau continue cu excepţia unei mulţimi finite de puncte) şi pozitive, este suficient să considerăm, pentru intervale, integrala Riemann (proprie sau improprie).

Reciproc, orice funcţie cu proprietăţile i, ii este densitate de repartiţie şi generează o repartiţie (şi deci este densitatea unei v.a continue).

Exemplu. Să considerăm funcţia $f(x) = 0, x < 0$ şi $f(x) = \lambda e^{-\lambda x}, x \geq 0, \lambda > 0$. Se verifică cu uşurinţă ca această funcţie satisface conditiile i, ii de mai sus. Repartiţia generată de această densitate se numeşte **exponenţială** . O v.a (continuă) cu densitatea f se zice **exponenţial repartizată** .

O v.a exponenţial repartizată are următoarea proprietate interesantă :

$$P(X > x + y / X > x) = P(X > y), x, y \geq 0.$$

In adevăr $P(X > x + y / X > x) = \frac{P((X>x+y)\cap(X>x))}{P(X>x)} = \frac{P(X>x+y)}{P(X>x)} = \frac{\lambda \int_{x+y}^{\infty} e^{-\lambda t}dt}{\lambda \int_x^{\infty} e^{-\lambda t}dt} = \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}} = e^{-\lambda y} = P(X > y)$. Se spune că o v.a exponenţial repartizată este lipsită de memorie.

Legătura dintre funcţia de repartiţie şi densitatea de repartiţie, pentru v.a continue, este dată de :

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt$$

şi (în anumite condiţii, de exemplu dacă f este continuă) $f(x) = F_X'(x)$.

Momente. Vom arăta pentru început că integralele v.a se pot calcula folosind repartiţiile asociate. Fie X o v.a şi $g : \mathbb{R} \to \mathbb{R}$ o funcţie măsurabilă . In ipoteza existenţei integralelor (nu enunţăm precis condiţiile) rezultă $M[g(X)] = \int_{\mathbb{R}} g(x)\,dp_X$ (reamintim că p_X este repartiţia asociată v.a X). Demonstraţia se face în mai mulţi paşi pentru funcţia g începând cu g indicatorul unei mulţimi boreliene, apoi g funcţie simplă măsurabilă şi apoi prin aproximarea funcţiilor măsurabile prin funcţii simple etc. Obţinem, în particular, o formulă importantă pentru medii:

$$M[X] = \int_{\mathbb{R}} x dp_X$$

Se obţin formulele corespunzătoare pentru v.a discrete sau continue:

$M[X] = \sum_i^n x_i p_i$ (cazul mulţimii finite de valori) $M[X] = \sum_i^\infty x_i p_i$(cazul numărabil) $M[X] = \int_{-\infty}^{\infty} x f(x)\, dx$ (cazul continuu) Mai general, vom defini **momentul de ordin** $k, (k \in \mathbb{N}, k \geq 1)$ al unei v.a X prin $m_k[X] = M\left(X^k\right) = \int_\Omega X^k dp$ (dacă , bineînţeles integrala există). In particular, $m_1[X] = M[X]$. Folosind rezultatele de mai sus şi luând funcţia $g(x) = x^k$ avem

$$m_k[X] = \sum_i^n x_i^k p_i \text{ sau } m_k[X] = \sum_i^\infty x_i^k p_i \text{ sau } M[X] = \textstyle\int_{-\infty}^{\infty} x^k f(x)\, dx$$

calculul mediei.

Dispersie. Momentul de ordin 2 joacă un rol important în studiul v.a.

O v.a se zice centrată dacă are media nulă . Dacă $X \in L^1$ numim v.a centrată asociată v.a $X - M[X]$ $(M[X - M[X]] = M[X] - M[M[X]] = 0$, media v.a constante este constanta însăşi).

Dispersia $D[X]$ a unei v.a este momentul de ordin 2 al variabilei centrate $X - M[X]$ (când acesta există). Astfel avem $D[X] = \int_\Omega (X - M[X])^2 dp$. Rezultă :

$D[X] = \sum_i^n (x_i - M[X])^2 p_i$ sau $D[X] = \sum_i^\infty (x_i - M[X])^2 p_i$ sau $D[X] = \int_{-\infty}^{\infty} (x - M[X])^2 f(x)\, dx$ **calculul dispersiei.**

Mulţimea v.a care au moment de ordin 2 (deci şi dispersie) va fi notată L^2.

Proprietăţile dispersiei.

1) Dacă c este o v.a constantă , atunci $D[c] = 0$

2) Dacă $X \in L^2$ şi $\alpha \in \mathbb{R}$, atunci $\alpha X \in L^2$ şi $D[\alpha X] = \alpha^2 D[X]$

3) Dacă $X, Y \in L^2$, atunci $X + Y \in L^2$ şi $D[X + Y] = D[X] + D[Y] + 2M[(X - M[X])(Y - M[Y])]$

Demonstraţia acestor proprietăţi rezultă dintr-un calcul simplu.

Vom numi **covarianţa** v.a X, Y numărul $cov[X, Y] = M[(X - M[X])(Y - M[Y])]$. Evident $cov[X, X] = D[X]$.

Calculul dispersiei: $D[X] = m_2[X] - M[X]^2$ $(D[X] = \int_\Omega (X - M[X])^2 dP = D[X] = \int_\Omega X^2 dP - 2M[X] \int_\Omega X dP + M[X]^2$etc.

Vom numi abatere pătratică medie a v.a $X \in L^2$ numărul $\sigma(X) = \sqrt{D[X]}$.

Exerciţiu. Să se determine media şi dispersia unei v.a exponenţial repartizată .

Soluţie. Densitatea de repartiţie a v.a este $f(x) = 0, x < 0$ şi $f(x) = \lambda e^{-\lambda x}, x \geq 0, \lambda > 0$. Rezultă :

$M[X] = \lambda \int_0^\infty x e^{-\lambda x} dx = \frac{1}{\lambda}$ (printr-o simplă integrare prin părţi). Apoi $m_2[X] = \lambda \int_0^\infty x^2 e^{-\lambda x} dx = \frac{2}{\lambda^2}$ (integrare prin părţi) şi deci $D[X] = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$.

Vom avea prilejul să calculăm medii şi dispersii în capitolul următor.

Pentru acest capitol se recomandă lucrările: $[B.01],:$ $[C.01],:$ $[D.01],:$ $[F.01],:$ $[B.01],:$ $[G.01],:$ $[G.03],:$ $[B.01],:$ $[I.01],:$ $[B.01],:$ $[L.01],:$ $[B.01],:$ $[P.02], [S.01],:$ [Ş.01].

Capitolul 10

MF.10. Legi de probabilitate

Cuvinte cheie

Legea binomială , legea Poisson, repartiţia uniformă , legea normală .

In acest capitol se prezintă câteva legi de probabilitate (repartiţii) importante atât teoretic cât şi pentru aplicaţii.

10.1 MF.10.1. Legea (repartiţia) binomială .

Se consideră un experiment aleator cu două rezultate posibile a şi $\overline{a}$ (experiment Bernoulli) care se realizează cu probabilităţile $\mathfrak{p}$, respectiv $\mathfrak{q}$ ($\mathfrak{p}+\mathfrak{q}=1$). Se fixează numărul natural $n \geq 1$ şi se consideră experimentul care constă în repetarea independentă , de n ori a experimentului Bernoulli. Rezultatele posibile ale acestui ultim experiment sunt n-upluri (şiruri finite de lungime n) de a şi $\overline{a}$. Intuitiv, probabilitatea de a obţine un n-uplu care conţine a de $0 \leq k \leq n$ este $\mathfrak{p}^k\mathfrak{q}^{n-k}$. Sunt C_n^k astfel de n-upluri. Probabilitatea de realizare a unui n-uplu care să conţină a de k ori va fi $p_k = C_n^k\mathfrak{p}^k\mathfrak{q}^{n-k}$. Ca urmare a acestui raţionament "intuitiv" vom trece la definirea legii binomiale: o v.a urmează o lege binomială de parametri n, k ($n \geq 1, 0 \leq k \leq n$) dacă valorile posibile sunt $0, 1, ..., n$ şi probabilităţile corespunzătoare $p_k = C_n^k\mathfrak{p}^k\mathfrak{q}^{n-k}$. Definiţia este corectă căci $\sum\limits_{k=0}^{n} C_n^k\mathfrak{p}^k\mathfrak{q}^{n-k} = (\mathfrak{p}+\mathfrak{q})^n = 1$(formula binomului). Legea binomială va fi notată $B(n, \mathfrak{p})$.

Media unei v.a binomial repartizate. Fie X o v.a care urmează legea $B(n, \mathfrak{p})$. Avem $M[X] = \sum\limits_{k=0}^{n} kC_n^k\mathfrak{p}^k\mathfrak{q}^{n-k}$; se porneşte de la identitatea (de funcţii) $(1+x)^n = \sum\limits_{k=0}^{n} C_n^k x^k$; derivând obţinem $n(1+x)^{n-1} = \sum\limits_{k=1}^{n} kC_n^k x^{k-1}$

pentru orice $x \in \mathbb{R}$. Pentru $x = \frac{p}{q}$ avem $n\left(1+\frac{p}{q}\right)^{n-1} = \sum_{k=1}^{n} kC_n^k \frac{p^{k-1}}{q^{k-1}}$ sau $\frac{n}{q^{n-1}} = \sum_{k=1}^{n} kC_n^k \frac{p^{k-1}}{q^{k-1}}$ şi înmulţind ambii membrii cu p rezultă $M[X] = n\mathfrak{p}$ (s-a presupus $0 < p < 1$).

Dispersia unei v.a binomial repartizate. In urma unui calcul similar celui de mai sus (exerciţiu) se arată că $D[X] = n\mathfrak{p}\mathfrak{q}$. Vom reveni la o alta justificare a acestui rezultat într-un capitol următor.

Exemplu. Un garaj deserveşte 70 de camioane. Probabilitatea ca, în timp de un an, un camion să intre în reparaţie este $0,3$. Să se determine numărul mediu de camioane în reparaţie într-un an.

Soluţie. Considerăm v.a "număr de autocamioane în reparaţie într-un an". Putem presupune (de ce?) că această v.a urmează o lege binomială $B(70, 0, 3)$. Rezultă media $70 \times 0,3 = 21$.

Să ne punem următoarea problemă: care sunt valorile de maximă probabilitate pentru o v.a repartizată $B(n, \mathfrak{p})$. Pentru aceasta să calculăm raportul $\frac{p_{k+1}}{p_k}$; un calcul simplu arată că $\frac{p_{k+1}}{p_k} = \frac{n-k}{k+1}\frac{\mathfrak{p}}{\mathfrak{q}}, 0 \leq k < n$. Din condiţia $\frac{p_{k+1}}{p_k} < 1$ se deduce $n\mathfrak{p} - \mathfrak{q} < k$. Obţinem concluzia: dacă $n\mathfrak{p} - \mathfrak{q} < 0$, atunci $p_0 > p_1 > ... > p_n$, dacă $n\mathfrak{p} - \mathfrak{q}$ este un întreg ≥ 0, atunci valorile de probabilitate maximă sunt $n\mathfrak{p} - \mathfrak{q}$ şi $n\mathfrak{p} - \mathfrak{q} + 1 = n\mathfrak{p} + \mathfrak{p}$ (cu aceeaşi probabilitate), iar dacă $n\mathfrak{p} - \mathfrak{q}$ nu este întreg, atunci valoarea de maximă probabilitate este întregul $k, n\mathfrak{p} - \mathfrak{q} < k < n\mathfrak{p} + \mathfrak{p}$. Valoarea cea mai probabilă a unei v.a se mai numeşte **mod** al acesteia.

Exerciţiu. Fie $n = 50$ şi $\mathfrak{p} = \frac{1}{3}$. Să se determine valorile cele mai probabile.

Soluţie. Se obţine $n\mathfrak{p} - \mathfrak{q} = 16$, deci cele mai probabile valori sunt 16 şi 17.

Diagrama în batoane. Este vorba de o reprezentare grafică, în plan, a unei v.a cu un număr finit de valori; pe axa Ox punem valorile v.a şi în fiecare asemenea punct ridicăm un segment perpendicular pe Ox de lungime proporţională cu probabilitatea valorii respective.

Exerciţiu. Construiţi diagrama în batoane pentru o v.a repartizată $B(9, 0, 4)$.

10.2 MF.10.2. Repartiţia Poisson.

Spunem că o v.a urmează o **repartiţie (lege) Poisson de parametru** $\lambda > 0$ dacă este discretă cu valori posibile $k \in \mathbb{N}$ si probabilitătile $p_k = e^{-\lambda}\frac{\lambda^k}{k!}, k \in \mathbb{N}$. Definitia este corectă fiindcă $\sum_{0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = 1$.

Media unei v.a Poisson repartizată. Dacă X urmează o lege Poisson de parametru λ avem $M[X] = \sum_0^\infty k p_k = e^{-\lambda} \sum_0^\infty k \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_1^\infty \frac{\lambda^{k-1}}{(k-1)!} = \lambda$.

Deci: $M[X] = \lambda$.

Dispersia unei v.a Poisson repartizată. Fie X o v.a Poisson repartizată cu parametrul λ. Calculăm pentru început momentul de ordin 2 al acestei v.a:

$m_2[X] = e^{-\lambda} \sum_0^\infty k^2 \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_1^\infty k \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_1^\infty [(k-1)+1] \frac{\lambda^{k-1}}{(k-1)!} = \lambda(\lambda+1)$. Obţinem: $D[X] = \lambda(\lambda+1) - \lambda^2 = \lambda$.

Deci: $D[X] = \lambda$.

Exerciţiu. Să se determine valorile de maximă probabilitate pentru o v.a Poisson repartizată.

Răspuns. Se procedează analog cazului repartiţiei binomiale. Valoarea de probabilitate maximă este întregul k satisfăcând conditia $\lambda - 1 < k < \lambda$ pentru λ neîntreg; dacă λ este întreg, atunci $\lambda - 1, \lambda$ sunt valorile de (aceeaşi) probabilitate maximă.

Aproximarea legii binomiale prin legea Poisson. Fie $B(n, \mathfrak{p}(n))_n$ un şir de repartiţii binomiale astfel încat şirul $\mathfrak{p}(n) \to 0$ si $n\mathfrak{p}(n) \to \lambda, \lambda > 0$. În aceste condiţii: pentru fiecare $k \in \mathbb{N}$ avem $\lim_{n\to\infty} C_n^k \mathfrak{p}(n)^k (1 - \mathfrak{p}(n))^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}$ (şirul de repartiţii binomiale tinde punctual către o repartiţie Poisson). Notaţia $\mathfrak{p}(n)$, oarecum neobişnuită pentru şiruri, a fost utilizată pentru a nu fi pericol de confuzie cu notaţia $\mathfrak{p}_k$ folosită pentru repartiţia unei legi binomiale fixate. Importanţa acestui rezultat stă în faptul că anumite calcule se simplifică dacă se înlocuieşte repartiţia binomială cu repartiţia Poisson. Înlocuirea legii binomiale cu o lege Poisson se recomandă în situaţia $n > 50, \mathfrak{p} < 0,1$. Nu vom demonstra relaţia de mai sus dar vom considera un exemplu (calculul limitei de mai sus poate fi luat ca exerciţiu).

Exemplu. Un sistem are 10^3 de componente. Într-o lună fiecare componentă se defectează independent de funcţionarea celorlalte iar probabilitatea de defectare este 10^{-3}. Care este probabilitatea ca nici o componentă să nu se defecteze? Putem considera că ”modelul” probabilistic al sistemului este dat de o repartiţie binomială $B(10^3, 10^{-3})$. Avem $q = 1 - 10^{-3} = 0,999$ şi probabilitatea căutată este $0,999^{1000}$. Aproximăm legea binomială cu o lege Poisson de parametru $\lambda = n\mathfrak{p} = 10^3 \cdot 10^{-3} = 1$(media legii binomiale aproximează media legii Poisson, aşa cum rezultă din condiţia de aproximare). Avem de calculat, pentru legea Poisson, $p_0 = e^{-1} = 0,368$ aşa cum rezultă cu uşurinţă.

10.3 MF.10.3. Repartiţia uniformă.

Fie $a, b \in \mathbb{R}, a < b$. O v.a este **repartizată uniform în intervalul** (a, b) dacă este o v.a continuă de densitate

$f(x) = \frac{1}{b-a}$ dacă $x \in (a, b)$ şi $f(x) = 0$ în rest. Se mai spune că v.a respectivă urmează o lege de densitate uniformă în intervalul (a, b).

Funcţia de repartiţie. Dacă X este o v.a uniform repartizată în (a, b), funcţia sa de repartiţie este $F(x) = \displaystyle\int_{-\infty}^{x} f(t)\, dt$ şi se obţine cu uşurinţă că $F(x) = 0$ dacă $x \leq a, F(x) = \frac{x-a}{b-a}$ dacă $a < x \leq b, F(x) = 1$ dacă $x > b$ (s-a notat F în loc de F_X pentru simplitate).

Exerciţiu. Să se deseneze graficul functiei F.

Media unei v.a uniform repartizată. Dacă X este o v.a uniform repartizată în (a, b), atunci $M[X] = \displaystyle\int_{-\infty}^{\infty} x f(x)\, dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$.

Deci $X \in L^1$ şi $M[X] = \frac{a+b}{2}$.

Dispersia unei v.a uniform repartizată. Dacă X este o v.a uniform repartizată în (a, b), atunci $D[X] = \displaystyle\int_{-\infty}^{\infty} \left(x - \frac{a+b}{2}\right)^2 f(x)\, dx = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a} dx = \frac{(b-a)^2}{12}$. Deci X are dispersie şi $D[X] = \frac{(b-a)^2}{12}$.

Exercitiu. Fie X ca mai sus şi $a < c < d < b$. Să se calculeze $P(c < X < d)$.

Soluţie. Avem $P(c < X < d) = \displaystyle\int_c^d \frac{dx}{b-a} = \frac{d-c}{b-a}$. Deci probabilitatea ca v.a X să ia valori în intervalul $(c, d) \subset (a, b)$ este raportul lungimilor intervalelor respective.

Exemplu. Metrourile sosesc în staţie la intervale de 2 minute. Timpul de aşteptare a unui călător sosit aleator în staţie poate fi modelat ca o v.a uniform repartizată în intervalul $(0, 2)$.

Să ne reamintim câmpul de probabilitate pe $[0, 1]$ în care evenimentele sunt mulţimile boreliene iar probabilitatea este generată de lungimea intervalelor. Legătura dintre acest câmp şi repartiţia uniformă este evidentă; îl putem considera ca un model al experimentului "alegerea la întâmplare a unui număr (punct) în intervalul $[0, 1]$". Modelul se încadrează si în ceea ce am numit "probabilitate geometrică" şi este un analog "continuu" al egal probabilităţii.

10.4 MF.10.4. Legea normală.

O v.a continuă urmează o **lege normală de parametri** $m, \sigma \in \mathbb{R}, \sigma > 0$ dacă are densitatea de repartiţie $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$. Vom folosi notaţia

$N(m,\sigma)$ pentru a ne referi la o lege normală de parametri m,σ. Astfel vom spune că v.a X urmează o lege $N(m,\sigma)$ sau că X este $N(m,\sigma)$ repartizată. Trebuie menţionat că unii autori notează $N\left(m,\sigma^2\right)$ ceea ce este notat, în acest text, $N(m,\sigma)$.

În particular densitatea de repartiţie a legii $N(0,1)$ este $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. $N(0,1)$ se mai numeşte "legea normală redusă".

Pentru a avea o definiţie corectă rămâne de verificat condiţia $\int_{-\infty}^{\infty} f(x)dx = 1$. În acest scop reamintim rezultatul fundamental $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$. Folosind acest rezultat şi o schimbare evidentă de variabile se obţine condiţia cerută.

În figură sunt reprezentate densitătile legilor $N(0,1)$ şi $N\left(0,\frac{1}{2}\right)$.

Funcţia de repartiţie. Dacă X este $N(0,1)$ repartizată, atunci funcţia sa de repartiţie va fi dată de $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$. Din cauza importanţei legii normale în probabilităţi şi statistică funcţia Φ este tabelată (a se vedea sfârşitul capitolului).

Vom folosi termenul "funcţie de repartiţie a legii normale" cu semnificaţie evidentă. Legătura dintre legile normale de diferiţi parametri este uşor de stabilit şi permite reducerea calculelor la legea normală redusă (şi deci la funcţia Φ). Astfel:

- dacă v.a X este $N(m,\sigma)$ repartizată, atunci v.a $Z = \frac{X-m}{\sigma}$ este $N(0,1)$ repartizată.

- dacă v.a X este $N(0,1)$ repartizată, atunci v.a $Y = \sigma X + m$ este $N(m,\sigma)$ repartizată.

Este suficient să dovedim una dintre afirmaţii, de exemplu a doua. Avem $F_Y(x) = P(Y < x) = P(\sigma X + m < x) = P\left(X < \frac{x-m}{\sigma}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-m}{\sigma}} e^{-\frac{t^2}{2}}dt$. În integrală facem schimbarea de variabilă $u = \sigma t + m$ şi obţinem $F_Y(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(u-m)^2}{2\sigma^2}}du$ ceea ce era de demonstrat.

Funcţia Φ are proprietatea $\Phi(x)+\Phi(-x) = 1$. Într-adevăr avem $\Phi(x) + \int_{x}^{\infty} e^{-\frac{t^2}{2}}dt = 1$ şi făcând schimbarea de variabilă $t = -u$ în integrală obţinem rezultatul. O funcţie importantă este **funcţia erorilor** care se defineşte prin $erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$. Se obţine relaţia $\Phi(x) = \frac{1}{2}+\frac{1}{2}erf\left(\frac{x}{\sqrt{2}}\right)$. Funcţia erorilor este legată de teoria erorilor accidentale de măsurare. Astfel se consideră (Gauss) că variabila aleatoare Ξ cu valori erorile accidentale de măsurare este repartizată normal $N(0,\sigma)$, în care σ este legat de precizia măsurătorii. Vom avea deci $p(a \le \Xi \le b) = \frac{1}{\sigma\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2\sigma^2}}dx =$

$\frac{1}{\sigma\sqrt{2\pi}}\left(\int_0^b e^{-\frac{x^2}{2\sigma^2}}dx - \int_0^a e^{-\frac{x^2}{2\sigma^2}}dx\right) = \frac{1}{\sqrt{\pi}}\left(\int_0^{\frac{b}{\sigma\sqrt{2}}} e^{-u^2}du - \int_0^{\frac{a}{\sigma\sqrt{2}}} e^{-u^2}du\right) =$
$\frac{1}{2}\left(erf\left(\frac{b}{\sigma\sqrt{2}}\right) - erf\left(\frac{a}{\sigma\sqrt{2}}\right)\right).$

Exerciţiu. Desenaţi graficul funcţiei erf.

Media unei v.a normal repartizate. Fie X o v.a care urmează legea $N(0,1)$. Avem $M[X] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} xe^{-\frac{x^2}{2}}dx = 0$ (integrala există şi funcţia de integrat este impară). Fie acum $Y = \sigma X + m$ o v.a care urmează legea $N(m,\sigma)$. Din proprietăţile mediei deducem $M[Y] = \sigma M[X] + m = m$.

Deci o v.a normal repartizată de parametri m, σ are media egală cu parametrul m.

Dispersia unei v.a normal repartizate. Fie X o v.a care urmează legea $N(0,1)$. Avem $D[X] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} xxe^{-\frac{x^2}{2}}dx.$ Aplicând o integrare prin părţi ($x = u, xe^{\frac{-x^2}{2}} = dv$) se obţine $D[X] = 1$. Fie acum $Y = \sigma X + m$ o v.a care urmează legea $N(m,\sigma)$. Evident, $Y - M[Y] = \sigma X$ de unde, folosind definiţia dispersiei obţinem $D[Y] = \sigma^2$. Să scriem din nou rezultatele obţinute:

Dacă o v.a urmează legea normală $N(m,\sigma)$, atunci media sa este m, iar dispersia σ^2. În acest mod parametrii legii normale capătă o semnificaţie concretă. Rezultă că σ este abaterea pătratică medie corespunzătoare.

Regula celor 3σ. Fie X o variabilă aleatoare $N(m,\sigma)$ repartizată. Atunci $P(m - 3\sigma < X < m + 3\sigma) \simeq 0,997$.

Demonstratia este imediată: se reduce la $N(0,1)$ şi se caută în tabele.

Pentru acest capitol se recomandă lucrările: $[B.01], [C.01], [D.01]\ [G.01],$ $[G.02], [G.03], [S.01]$

Capitolul 11

MF.11. Vectori aleatori

Cuvinte cheie

Vectori aleatori, repartiţie comună, densitate comună, repartiţie marginală, corelaţie, independenţa variabilelor aleatoare.

11.1 MF.11.1. Vectori aleatori, repartiţie comună.

Fie (Ω, Δ, P) un câmp de probabilitate. O funcţie $X : \Omega \to \mathbb{R}^n, X = (X_1, X_2, ..., X_n)$ pentru care componentele $X_1, X_2, ..., X_n$ sunt variabile aleatoare se numeşte vector aleator n-dimensional. Pentru a fixa ideile vom studia, cu deosebire, vectori aleatori 2-dimensionali şi vom folosi notaţia $Z = (X, Y)$ pentru această situaţie. Rezultatele se extind cu uşurinţă la vectori aleatori de dimensiune oarecare. Un dreptunghi in $\mathbb{R}^2$ este un produs cartezian de intervale; condiţia ca X, Y să fie v.a este echivalentă cu faptul că, pentru orice dreptunghi $S \subseteq \mathbb{R}^2$ rezultă $Z^{-1}(S) = (Z \in S) \in \Delta$ (egalitatea este o notaţie, conformă convenţiilor făcute anterior). Mulţimile boreliene ale planului sunt cele care aparţin celei mai mici σ-algebre care conţine dreptunghiurile şi de aici rezultă că proprietatea de a fi vector aleator este de acelaşi tip cu proprietatea de a fi v.a anume: imaginea inversă a oricărei mulţimi boreliene este eveniment (proprietatea de măsurabilitate).

Repartiţie comună. Dacă $\mathcal{B}$ este σ-algebra mulţimilor boreliene din $\mathbb{R}^2$ numim repartiţie (2-dimensională) orice probabilitate $P : \mathcal{B} \to [0, 1]$. Fie $Z = (X, Y)$ un vector aleator. Pentru fiecare mulţime boreliană $S \in \mathcal{B}$ definim $P_Z(S) = P_{X,Y}(S) = P(Z \in S)$ (prima egalitate este o notaţie); se obţine o repartiţie numită **repartiţia vectorului aleator** Z sau (mai des folosită) **repartiţia comună a v.a** X, Y. Repartiţia comună determină probabilistic vectorul aleator (acelaşi fenomen a fost descris pentru variabile aleatoare). Pentru întelegerea rolului repartiţiei comune în raport cu rolul repartiţiilor componentelor este util să ne reamintim că, în general, probabiltatea intersecţiei a două evenimente nu este determinată de probabilitătile evenimentelor respective.

Vom trata, pentru început, cazul în care componentele vectorului aleator sunt discrete şi au o mulţime finită de valori. Fie X o v.a cu mulţimea de valori $\{x_1, x_2, ..., x_m\}$ şi Y o v.a cu mulţimea de valori $\{y_1, y_2, ..., y_n\}$(definite pe o mulţime Ω şi cu mulţimea părţilor ca σ-algebră). Considerând vectorul aleator $Z = (X, Y)$ (cu mulţimea de valori $\{x_1, x_2, ..., x_m\} \times \{y_1, y_2, ..., y_n\}$) repartiţia comună este determinată de numerele $p_{ij} = p\left(X = x_i, Y = y_j\right), i = 1, 2, ..., m, j = 1, 2, ..., n, \sum_{i,j} p_{ij} = 1$.

$X\backslash Y$	y_1	.	.	.	y_j	.	.	.	y_m	marginal X
x_1	p_{11}	.	.	.	p_{1j}	.	.	.	p_{1m}	$p_{1.}$
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
x_i	p_{i1}	.	.	.	p_{ij}	.	.	.	p_{im}	$p_{i.}$
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
x_m	p_{n1}	.	.	.	p_{nj}	.	.	.	p_{nm}	$p_{n.}$
marginal Y	$p_{.1}$	.	.	.	$p_{.j}$	.	.	.	$p_{.m}$	1

În tabelul de mai sus este reprezentată repartiţia comună a v.a $X.Y$. Avem $p_{i.} = P\left(X = x_i\right) = \sum_j p_{ij}$ (prima egalitate este o notatie), $i = 1, 2, ..., n$ şi $p_{.j} = p\left(Y = y_j\right) = \sum_i p_{ij}$ repartiţiile variabilelor X, respectiv Y numite **repartiţii marginale**. Din tabel se pot obţine cu uşurinţă probabilităţi condiţionale (dacă au sens); de exemplu $p_{j/i} = p\left(Y = y_j \mid X = x_i\right) = \frac{p_{ij}}{p_{i.}}$(prima egalitate fiind o notaţie).

Asemenea tabele pot apărea, de exemplu, dintr-un recensământ în care sunt adresate două întrebări fiecare cu un număr finit de răspunsuri posibile; numerele p_{ij} reprezintă frecvenţa relativă a răspunsului (x_i, y_j) etc.

Cazul în care $p_{ij} = p_{i.} \times p.j$ pentru orice pereche (i, j) corepunde situaţiei importante când v.a X, Y sunt independente (noţiune care se va defini mai jos).

Funcţia de repartiţie comună. Dacă $Z = (X, Y)$ este un vector aleator, se defineşte **funcţia de repartiţie comună a perechii** (X, Y) prin $F_{X,Y}(x, y) = p(X < x, Y < y) = p_{X,Y}[(-\infty, x) \times (-\infty, y)]$. Funcţia $F_{X,Y}$ este definită pe $\mathbb{R}^2$ şi ia valori în intervalul $[0, 1]$.

Exerciţiu. Figuraţi, în plan, mulţimea $(-\infty, x) \times (-\infty, y)$.

Proprietăţile funcţiei de repartiţie comună. Vom nota pentru simplitate cu F funcţia de repartiţie comună.

1. Funcţia de repartiţţie comună este crescătoare în fiecare variabilă ($x_1 \leq x_2$ implică $F(x_1, y) \leq F(x_2, y)$, $\forall y$ etc).

2. $F(x, -\infty) = \lim_{y \to -\infty} F((x, y) = 0$ (prima egalitate este o notaţie) şi $F(-\infty, y) = \lim_{x \to -\infty} F((x, y) = 0$, $F(\infty, \infty) = 1$(notaţie evidentă).

3. Fie S dreptunghiul $[a, b) \times [c, d)$. Atunci $P((X, Y) \in S) = F(c, d) - F(a, d) - F(b, c) + F(a, c)$.

Lăsăm verificarea acestor proprietăţi ca exerciţiu.

Funcţiile $F_X(x) = F(x, \infty)$, $F_Y(y) = F(\infty, y)$ sunt funcţiile de repartiţţie ale v.a X, Y; sunt numite repartiţii marginale.

Densitate de repartiţie comună. Dacă $Z = (X, Y)$ este un vector aleator spunem că o funcţie (măsurabilă) $f : \mathbb{R}^2 \to \mathbb{R}$ este o **densitate a vectorului** Z sau o **densitate de repartiţie comună a v.a** X, Y dacă $f(x, y) \geq 0$ pentru orice x, y , $\int_{\mathbb{R}^2} f(x, y)\, dxdy = 1$ si $p(Z \in S) = \int_S f(x, y)\, dxdy$ pentru orice multime boreliană $S \subseteq \mathbb{R}^2$. Precizăm că pentru funcţii continue pe dreptunghiuri (şi mulţimi mai generale des întâlnite în practică) integrala generală coincide cu integrala Riemann.

Exemplu. Fie $S = (a, b) \times (c, d)$, $a < b, c < d$ un dreptunghi în plan. Funcţia $f : \mathbb{R}^2 \to \mathbb{R}, f = \frac{1}{(b-a)(d-c)} \chi_S$ (χ_S este funcţia caracteristică a mulţimii S) satisface condiţiile unei densităti. Un vector aleator cu asemenea densitate se zice uniform repartizat în dreptunghiul S.

Revenind la cazul general al existenţei unei densităti vom avea $F(x, y) = \displaystyle\int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\, dxdy$ şi (în anumite condiţii) $\frac{\partial^2 F}{\partial x \partial y}(x, y) = f(x, y)$.

Să considerăm o situaţie analoagă celei descrise la v.a. Fie $g : \mathbb{R}^2 \to \mathbb{R}$ măsurabilă, $Z : \Omega \to \mathbb{R}^2, Z = (X, Y)$ un vector aleator şi $p_{X,Y}$ repartiţia comună asociată. Notăm $g(X, Y) = g \circ Z$. Vom avea (în ipoteza existenţei mediei) $M[g(X, Y)] = \int_{\mathbb{R}^2} g(x, y)\, dp_{X,Y}$. În cazul existenţei unei densităţi f vom avea $M[g(X, Y)] = \int_{\mathbb{R}^2} g(x, y) f(x, y)\, dxdy$.

Exemplu. Dacă $Z = (X, Y)$ este un vector aleator de densitate f, atunci (în ipoteza existenţei mediei) $M[XY] = \int_{\mathbb{R}^2} xyf(x, y)\, dxdy$ şi (în ipoteza existenţei integralelor) $cov[X, Y] = \int_{\mathbb{R}^2} (x - M[X])(y - M[Y]) f(x, y)\, dxdy$. Reamintim că prin covarianţa v.a X, Y înţelegem $M[(X - M[X])(Y - M[Y])]$.

Avem $cov[X, Y] = M[XY] - M[X] M[Y]$ asa cum rezultă imediat. Se mai foloseşte notaţia $cov[X, Y] = K_{xy}$. Dacă σ_x, σ_y sunt abaterile pătratice medii (nenule) ale v.a X respectiv Y, atunci coeficientul de corelaţie al v.a X, Y se defineşte ca $r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y}$. Rezultă (exerciţiu) $|r_{xy}| \leq 1$.

11.2 MF.11.2. Independenţa variabilelor aleatoare.

Fie $X_1, X_2, ..., X_n$ v.a pe câmpul de probabilitate (Ω, Δ, p). Spunem că aceste variabile sunt **independente** dacă pentru orice mulţimi $S_1, S_2, ..., S_n$ boreliene în $\mathbb{R}$ evenimentele $(X_1 \in S_1), (X_2 \in S_2), ..., (X_n \in S_n)$ sunt independente. Astfel, de exemplu, v.a X, Y sunt independente dacă pentru orice mulţimi S,T boreliene în $\mathbb{R}$ avem $P(X\in S, Y\in T) = P(X\in S)\times P(Y\in T)$. Am notat, pentru simplitate $p(X\in S, Y\in T) = p((X\in S)\cap(Y\in T))$. Putem exprima acest fapt mai sofisticat din punct de vedere matematic spunând că independenţa v.a X,Y revine la faptul că probabilitatea $P_{X,Y}$ introdusă la începutul acestui capitol este "produsul" probabilităţilor P_X şi P_Y asociate v.a X respectiv Y. Sensul intuitiv al noţiunii de produs (sens la care ne mărginim) este ideea geometrică veche că "aria unui dreptunghi este produsul lungimilor laturilor" generalizată la măsuri.

Exemplu. Fie X o v.a şi Y o v.a constantă $Y=c$. Atunci v.a X,Y sunt independente. În adevăr fie S,T mulţimi boreliene. Sunt posibile două cazuri: $c\in T$ caz în care $(Y\in T)=\Omega$ deci $P(Y\in T)=1$ şi $c\notin T$ caz în care $(Y\in T)=\emptyset$ deci $P(Y\in T)=0$. În ambele cazuri condiţia de independenţă este verificată.

Ca o consecinţă directă a definiţiilor, în cazul v.a independente X,Y obţinem pentru funcţiile de repartiţie $F_{X,Y}(x,y)=F_X(x)F_Y(y)$.

Dacă X,Y sunt v.a continue de densităţi f respectiv g şi sunt independente, atunci densitatea lor de repartiţie comună este h cu $h(x,y)=h(x)h(y)$ pentru orice $x,y\in\mathbb{R}$.

Media produsului a două v.a independente. Dacă v.a $X,Y\in L^1$ şi sunt independente, atunci $XY\in L^1$ şi $M[XY]=M[X]M[Y]$ (reamintim că o v.a este în L^1 dacă are medie).

Demonstraţia acestui fapt este standard (în teoria integralei): se demonstrează pentru v.a simple (pozitive) şi se aproximează v.a generale cu v.a simple etc . Să arătăm doar primul pas. Fie $X=\sum\limits_{i=1}^{n}a_i\chi_{A_i}$, $Y=\sum\limits_{j=1}^{m}b_j\chi_{B_j}$ v.a simple independente. Din independenţă rezultă $P(A_i\cap B_j)=P(A_i)P(B_j)$ pentru orice $i=1,2,...,n$ si $j=1,2,...,m$.

Avem $M[XY]=M\left[(\sum\limits_{i=1}^{n}a_i\chi_{A_i})\left(\sum\limits_{j=1}^{m}b_j\chi_{B_j}\right)\right]=M\left[\sum\limits_{i,j}a_ib_j\chi_{A_i\cap B_j}\right]=$ $\sum\limits_{i,j}a_ib_jP(A_i\cap B_j)=\left(\sum\limits_{i}a_iP(A_i)\right)\left(\sum\limits_{j}b_jP(B_j)\right)=M[X]M[Y]$.

Dispersia sumei a două v.a independente. Dacă $X,Y\in L^2$ sunt v.a independente, atunci $D[X+Y]=D[X]+D[Y]$ (reamintim că notaţia L^2 este folosită pentru v.a care au dispersie).

Demonstraţia rezultă din calcul: $D[X+Y]=D[X]+D[Y]+2cov(X,Y)$ (a se vedea capitolul VI sectiunea "Momente") şi deci este suficient să arătăm că, în cazul independenţei, $cov[X,Y]=0$. Avem

$M[(X-M[X])(Y-M[Y])]=M[XY]-XM[Y]-YM[X]-M[X]M[Y]]$

şi din independenţă obţinem $M[X]M[Y]-M[X]M[Y]-M[X]M[Y]-M[X]M[Y]=0$.

Exerciţiu. i). Arătaţi că dacă $X,Y\in L^1$ sunt independente, atunci

$X-M[X], Y-M[Y]$ sunt independente şi obţineti astfel o altă demonstraţie a afirmaţiei de mai sus.

ii) Arătati că dacă $X_{1,}X_{2,}..., X_n \in L^2$ şi sunt independente, atunci $D[X+X_2+...+X_n] = D[X_1]+D[X_2]+...+D[X_n]$.

Repartiţia sumei a două v.a independente. Fie, pentru început, X, Y două v.a discrete cu mulţimi finite de valori şi independente. Notând x, y valorile "generice" ale v.a X respectiv Y un simplu calcul arată că $P(X+Y=z) = \sum\limits_{x+y=z} P(X=x)P(Y=y)$. De aici rezultă cu uşurinţă că funcţia generatoare a sumei a două v.a discrete independente este produsul funcţiilor generatoare ale termenilor. Vom da şi o formulă pentru calculul densităţii de repartiţie a sumei a două v.a continue şi independente. Anume fie v.a independente X de densitate f şi Y de densitate g. În acest caz densitatea v.a $X+Y$ este dată de $h(x) = \int_{-\infty}^{\infty} f(x-y)g(y)\,dy$. Astfel rezultă că densitatea sumei a două v.a independente este convoluţia densităţilor termenilor.

Exemplu. Suma a două v.a independente şi normal repartizate este o v.a normal repartizată.

Vom face calculul pentru cazul variabilelor reduse lăsând cazul general ca exerciţiu. Fie deci X, Y v.a $N(0,1)$ repartizate. Densitatea sumei va fi dată de $h(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2}} e^{-\frac{y^2}{2}} dy = \frac{1}{2\pi} e^{-\frac{x^2}{2}} \int_{-\infty}^{\infty} e^{-\left(y^2-xy\right)} dy =$
$= \frac{1}{2\pi} e^{-\frac{x^2}{2}} \int_{-\infty}^{\infty} e^{-\left(y^2-xy+\frac{1}{4}x^2\right)} e^{\frac{1}{4}x^2} dy = \frac{1}{2\pi} e^{-\frac{x^2}{4}} \int_{-\infty}^{\infty} e^{-\left(y-\frac{1}{4}x\right)^2} dy = \frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{4}}$.
Am obţinut densitatea legii $N\left(0,\sqrt{2}\right)$ aşa cum am afirmat (media sumei este suma mediilor, iar dispersia sumei este suma dispersiilor).

Exerciţiu. Variabilele aleatoare X, Y sunt independente şi uniform repartizate în intervalul (a,b). Să se determine densitatea sumei $X+Y$.

Răspuns. Dacă h este densitatea sumei, atunci $h(x) = 0$ dacă $x \le 2a$ sau $x > 2b$, $h(x) = \frac{x-2a}{(b-a)^2}$ dacă $2a < x \le a+b$ si $h(x) = \frac{2b-x}{(b-a)^2}$ dacă $a+b < x \le 2b$ (repartiţia corespunzătoare se numeşte Simpson).

Pentru acest capitol se recomandă lucrările:$[B.01], [C.01], [D.01]\,[G.01], [G.02], [G.03], [S.01]$.

Capitolul 12

MF.12. Legea numerelor mari

Cuvinte cheie

Şiruri de variabile aleatoare, inegalitatea Cebîşev, legea numerelor mari, teorema limită centrală, funcţie caracteristică, aproximarea variabilelor aleatoare.

12.1 MF.12.1. Inegalitatea Cebîşev.

Fie (Ω, Δ, P) un câmp de probabilitate, $X : \Omega \to \mathbb{R}$ o v.a şi $g : \mathbb{R} \to \mathbb{R}$ o funcţie măsurabilă. Să presupunem că g este crescătoare şi pozitivă pe imaginea $X(\Omega)$ a v.a X. Fie $a \in \mathbb{R}, g(a) > 0$. În aceste condiţii:

$(*)$ $P(X \geq a) \leq \frac{M[g(X)]}{g(a)}$ (reamintim că $g(X) = g \circ X$ şi presupunem existenţa mediei).

Demonstraţia este simplă: $M[g(X)] = \int_{-\infty}^{\infty} g(x)\, dP_X \geq \int_a^{\infty} g(x)\, dP_X \geq g(a) \int_a^{\infty} dP_X = g(a) P(X \geq a)$. Proprietăţile integralei utilizate în demonstraţie sunt intuitive chiar dacă se lucrează cu o integrală mai generală decât integrala Riemann. Dacă presupunem că v.a X este continuă cu densitate f, atunci demonstraţia devine:

$\int_{-\infty}^{\infty} g(x) f(x)\, dx \geq \int_a^{\infty} g(x) f(x)\, dx \geq g(a) \int_a^{\infty} f(x)\, dx = g(a) P(X \geq a)$ ceea ce arată, probabil, mai familiar.

Un caz particular deosebit de important al inegalităţii $(*)$ se obţine luând $g(x) = x^2$ şi presupunând $X \in L^2$. Anume:

$(**)$ $p(|X - M[X]| \geq a) \leq \frac{D[X]}{a^2}$, $a > 0$. (**inegalitatea Cebîşev**).

În adevăr $D[X] = M\left[(X - M[X])^2\right]$ şi se aplică $(*)$ pentru v.a $X - M[X]$.

Să luăm $a = \varepsilon\sigma[X], \varepsilon > 0$ ($\sigma[X]$ e abaterea pătratică medie presupusă > 0) în $(**)$. Vom obţine $P(|X - M[X]| \geq \varepsilon\sigma[X]) \leq \frac{1}{\varepsilon^2}$. Putem interpreta acest rezultat spunând că abaterile v.a X faţă de media sa care sunt mari în raport cu dispersia sunt foarte putin probabile. Astfel, de exemplu, luând $\varepsilon = 10$ obţinem $P(|X - M[X]| \geq \varepsilon\sigma[X]) \leq 0,01$.

Să observăm că trecând la evenimentul contrar în $(**)$ obţinem:

$(***)$ $P(|X - M[X]| < a) \geq 1 - \frac{D[X]}{a^2}$ o formă des utilizată a inegalităţii Cebîşev.

12.2 MF.12.2. Legea numerelor mari.

Fie $X_1, X_2, ..., X_n, ...$ un şir de v.a în L^2 la fel repartizate şi independente. Prin "la fel repartizate" înţelegem că au aceeaşi repartiţie (adică $p_{X_n} = p_{X_m}$ pentru orice n, m), iar prin "independente" înţelegem că $X_1, X_2, ..., X_n$ sunt independente pentru orice n. Atunci:

$\forall\varepsilon > 0 \quad \lim\limits_{n\to\infty} P\left(\left|\frac{X_1+X_2+...+X_n}{n} - M[X_1]\right| \geq \varepsilon\right) = 0$ (**legea numerelor mari** în forma Cebîşev).

Demonstraţia este foarte simplă. Fiind la fel repartizate, v.a $X_1, X_2, ..., X_n, ...$au aceeaşi medie şi aceeaşi dispersie (de ce?). Să notăm media comună cu μ şi dispersia comună cu σ^2. Avem $M[X_1 + X_2 + ... + X_n] = n\mu$ şi $D[X_1 + X_2 + ... + X_n] = n\sigma^2$(s-a folosit independenţa). Fie $\varepsilon > 0$ şi să luăm în $(**)$ $a = n\varepsilon$ şi $X = X_1 + X_2 + ... + X_n$. Obţinem $P(|X_1 + X_2 + ... + X_n - n\mu| \geq n\varepsilon) \leq \frac{n\sigma^2}{(n\varepsilon)^2}$ sau $P\left(\left|\frac{X_1+X_2+...+X_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$ de unde rezultatul, prin trecere la limită.

Înainte de a furniza un exemplu care să facă mai uşor de intuit rezultatul important obţinut să formulăm o definiţie generală care să precizeze limita de mai sus.

Convergenţa în probabilitate. Fie $(Y_n)_n$ un şir de v.a. Spunem că acest şir converge în probabilitate către v.a Y dacă

$\forall\varepsilon > 0 \quad \lim\limits_{n\to\infty} P(|Y_n - Y| \geq \varepsilon) = 0$. Cu această definiţie legea numerelor mari afirmă convergenţa în probabilitate a şirului $(\frac{X_1+X_2+...+X_n}{n})_n$ către v.a constantă μ (media comună a v.a $X_1, X_2, ..., X_n, ...$).

Nu vom intra în studiul convergenţei în probabilitate şi a legăturilor sale cu alte tipuri de convergenţă a şirurilor de v.a; vom vedea alt tip de convergenţă în formularea legii tari a numerelor mari.

Exemplu. Să considerăm un experiment Bernoulli (două rezultate posibile α cu probabilitatea $\mathfrak{p}$ şi $\overline{\alpha}$ cu probabilitatea $\mathfrak{q} = 1-\mathfrak{p}$). Ataşăm v.a X care ia doar două valori 1 dacă se realizează α şi 0 dacă se realizează $\overline{\alpha}$. Clar, X ia valoarea 1 cu probabilitatea $\mathfrak{p}$ şi valoarea 0 cu probabilitatea $\mathfrak{q}$. Rezultă $M[X] = \mathfrak{p}$ şi $D[X] = \mathfrak{pq}$ (justificaţi). Dacă $X_1, X_2, ..., X_n$ sunt independente şi având aceeaşi repartiţie ca X, atunci v.a $X_1+X_2+...+X_n$ este repartizată binomial $B(n, \mathfrak{p})$. În particular regăsim media $n\mathfrak{p}$ şi dispersia $n\mathfrak{pq}$ pentru v.a binomial repartizate. V.a $X_1+X_2+...+X_n$ numără apariţia rezultatului α în n experienţe independente. V.a $\frac{X_1+X_2+...+X_n}{n}$ reprezintă frecvenţa relativă de apariţie a rezultatului α în n experienţe independente. Legea numerelor mari aplicată în acest caz dă:

$\forall \varepsilon > 0 \ \lim\limits_{n\to\infty} P\left(\left|\frac{X_1+X_2+...+X_n}{n} - \mathfrak{p}\right| \geq \varepsilon\right) = 0$, deci şirul frecvenţelor relative converge în probabilitate către probabilitatea "teoretică" $\mathfrak{p}$.

Exerciţiu. Se poate afirma că, din 1000 de aruncări independente ale unei monede "oneste", probabilitatea de apariţie a stemei de un număr de ori cuprins între 400 şi 600 este cel puţin $0,97$?

Soluţie. Ne aflăm în contextul exemplului precedent cu $\mathfrak{p} = 0,5$ şi $n = 1000$. Aplicăm $(***)$ şi avem: $p\left(\left|\frac{X_1+X_2+...+X_n}{n} - 0,5\right| < 0,1\right) \geq 1 - \frac{0,25}{0,01\times 10^3} = 0,975$. Răspunsul este "da".

Putem generaliza legea numerelor mari renunţând la condiţia de "la fel repartizate" pentru v.a ale şirului şi păstrând condiţia de independenţă. Astfel notând μ_i, σ_i^2 media respectiv dispersia v.a X_i vom avea din $(**)$ $P\left(\left|\frac{X_1+X_2+...+X_n}{n} - \frac{\mu_1+\mu_2+...+\mu_n}{n}\right| \geq \varepsilon\right) \leq \frac{\sum\limits_{i=1}^{n}\sigma_i^2}{n^2\varepsilon^2}$. Dacă $\lim\limits_{n\to\infty}\frac{\sum\limits_{i=1}^{n}\sigma_i^2}{n^2} = 0$, atunci $\lim\limits_{n\to\infty} P\left(\left|\frac{X_1+X_2+...+X_n}{n} - \frac{\mu_1+\mu_2+...+\mu_n}{n}\right| \geq \varepsilon\right) = 0$.

Sunt posibile şi alte generalizări; de exemplu înlocuind condiţia de independenţă cu condiţia de necorelare $cov\left[X_i, X_j\right] = 0$ pentru orice $i \neq j$. Nu vom insista.

Metoda Monte Carlo. Fie $\Omega = [0,1]\times[0,1]$, $\mathcal{B}$ σ-algebra mulţimilor boreliene incluse în Ω şi p probabilitatea geometrică asociată ($P(A)$ este aria mulţimii A). Fie $g : [0,1] \to [0,1]$ o funcţie continuă. Atunci subgraficul $A_g = \{(x,y); 0 \leq x \leq 1, 0 \leq y \leq g(x)\}$ este mulţime boreliană şi ne propunem să estimăm aria acestei mulţimi prin metode probabilistice. Dacă imaginăm experimentul "un punct este ales la întamplare în Ω" atunci vom avea un experiment Bernoulli considerând rezultatele: punctul aparţine mulţimii A_g sau complementarei. În acest caz probabilitatea $\mathfrak{p}$ este chiar aria mulţimii A_g. Vom avea folosind cele de mai sus $P\left(\left|\frac{X_1+X_2+...+X_n}{n} - \mathfrak{p}\right| < \varepsilon\right) \geq 1 - \frac{\mathfrak{pq}}{n\varepsilon^2} \geq 1 - \frac{1}{4n\varepsilon^2}$ ($\mathfrak{pq} \leq \frac{1}{4}$). Dacă ne alegem o marjă de eroare $\varepsilon > 0$ si un prag $1-\delta$ de probabilitate putem determina n astfel încât $P\left(\left|\frac{X_1+X_2+...+X_n}{n} - \mathfrak{p}\right| < \varepsilon\right) \geq 1-\delta$. Practic aceasta înseamnă că generând la întâmplare n puncte în Ω, numărând câte puncte aparţin mulţimii A_g şi împărţind la n probabilitatea ca să obţinem aria $\mathfrak{p}$ cu eroarea acceptată este cât de mare am dorit.

O variantă a acestei metode se bazează pe legea de repartiţie uniformă. Dacă X este uniform repartizată în $(0,1)$, atunci $M[g(X)] = \int_0^1 g(x)\,dx$. Se consideră un şir $(X_n)_n$ de v.a independente şi uniform repartizate pe $(0,1)$ şi se atasează şirul $(Y_n)_n, Y_n = g(X_n)$. Se aplică inegalitatea Cebîşev ca mai sus observând că $M[g(X)] \in [0,1]$ şi deci şi $D[g(X)] = \int_0^1 (g(x) - M[g(x)])^2\,dx \in [0,1]$ uşurând astfel estimările. Practic, şirul $(X_n)_n$ modelează alegerea "la întâmplare" a unui şir de puncte în Ω; cu marja de eroare admisă şi cu riscul (probabilistic) acceptat se determină n. Dacă $x_1, x_2, ..., x_n$ sunt puncte obţinute (la intâmplare) se calculează $g(x_1), g(x_2), ..., g(x_n)$ şi media

$\frac{g(x_1)+g(x_2)+...+g(x_n)}{n}$ este luată drept valoarea integralei (sau, echivalent, a ariei de calculat).

Legea tare a numerelor mari. Fie $(X_n)_n$ un şir de v.a pe câmpul de probabilitate (Ω, Δ, p). Spunem că şirul tinde aproape sigur către v.a X dacă $P\{\omega; X_n(\omega)$ nu tinde la $X(\omega)\} = 0$. Deci şirul tinde punctual către X cu excepţia unui eveniment de probabilitate nulă.

Fie $X_1, X_2, ..., X_n, ...$ un şir de v.a în L^1 la fel repartizate şi independente şi μ media lor comună. În aceste condiţii şirul $(\frac{X_1+X_2+...+X_n}{n})_n$ tinde aproape sigur către (v.a constantă) μ (**legea tare a numerelor mari**). Nu vom intra în amănunte privind demonstraţia acestui rezultat. Adjectivul "tare" se referă la faptul că orice şir convergent aproape sigur este convergent în probabilitate către aceeaşi limită (reciproca nefiind adevarată).

12.3 MF.12.3. Teorema limită centrală.

Convergenţa în repartiţie. Fie (Ω, Δ, P) un câmp de probabilitate, $(X_n)_n$ de v.a pe acest câmp. Spunem că şirul $(X_n)_n$ converge în repartiţie către v.a X dacă şirul funcţiilor de repartiţie $(F_{X_n})_n$ converge către F_X în fiecare punct de continuitate al funcţiei F_X (deci dacă $x \in \mathbb{R}$ este un punct în care funcţia F_X este continuă, atunci $F_{X_n}(x) \underset{n\to\infty}{\to} F(x)$).

Se poate arăta că dacă şirul $(X_n)_n$ converge în probabilitate către X, atunci converge şi în repartiţie (catre X).

Exemplu. Fie $a_n \to a$ în $\mathbb{R}$. Considerăm v.a constante $X_n = a_n, X = a$. Se arată cu uşurinţă că şirul $(X_n)_n$ converge în repartiţie către X (arătaţi acest fapt calculând funcţiile de repartiţie corespunzătoare).

Exerciţiu. Fie $(\sigma_n)_n$ un şir crescător $\sigma_n > 0, \forall n$ şi $\sigma_n \to \sigma \in \mathbb{R}$. Fie X_n v.a $N(0, \sigma_n)$ pentru orice n şi X v.a $N(0, \sigma)$ repartizată. Converge şirul $(X_n)_n$ în repartiţie către X ?

Dacă şirul $(X_n)_n$ converge în repartiţie către X rezultă, din definiţie, că $P(X_n < x) \to P(X < x)$ în orice punct de continuitate a funcţiei F_X. De aici rezultă că pentru $a < b$ şi a, b sunt puncte de continuitate pentru F_X, avem $P(a \le X_n < b) \to P(a \le X < b)$.

Teorema limită centrală. Fie $(X_n)_n$ un şir de v.a în L^2, independente şi cu aceeaşi repartiţie pe (Ω, Δ, P). Fie $\mu = M[X_1] = M[X_2] = ...$ şi $\sigma^2 = D[X_1] = D[X_2] = ...$ $(\sigma > 0)$. În aceste condiţii:

pentru $\forall x \in \mathbb{R}$ $\lim\limits_{n\to\infty} P\left(\frac{X_1+X_2+...+X_n-n\mu}{\sigma\sqrt{n}} < x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$ (**teorema limită centrală**).

Cu alte cuvinte, dacă notăm $S_n = X_1+X_2+...+X_n$ şi $S_n^* = \frac{X_1+X_2+...+X_n-n\mu}{\sigma\sqrt{n}}$ teorema limită centrală afirmă că şirul $(S_n^*)_n$ converge în repartiţie către o v.a $N(0,1)$ repartizată. Să remarcăm că $M[S_n^*] = 0$ şi $D[S_n^*] = 1$ pentru orice n. Intuitiv putem interpreta v.a S_n ca acumulare de efecte aleatoare independente şi la fel repartizate. Deducem din teoremă şi observaţia care o precede ca pentru $a \leq b$, $\lim\limits_{n\to\infty} P\left(a \leq \frac{X_1+X_2+...+X_n-n\mu}{\sigma\sqrt{n}} < b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}dt = \Phi(b) - \Phi(a)$.

Nu vom demonstra legea numerelor mari (demonstraţia necesită dezvoltări care ies din cadrul acestui text).

Aplicarea teoremei limită centrală se face "aproximând" repartiţia variabilei S_n^* cu repartiţia normală $N(0,1)$ (mai ales în cazul estimării probabilităţii unui eveniment fixat). Rezultate mai precise în acest sens se obţin în cazul experimentului Bernoulli (şi deci în cazul repartiţiei binomiale). Să descriem, pe scurt, această situaţie. Fie $(X_n)_n$ un şir de v.a independente şi luând doar două valori: 1 cu probabilitatea $0 < \mathfrak{p} < 1$ şi 0 cu probabilitatea $\mathfrak{q} = 1 - \mathfrak{p}$. Asa cum am mai arătat, v.a $S_n = X_1 + X_2 + ... + X_n$ are o repartiţie binomială $B(n, \mathfrak{p})$. Teorema limită centrală aplicată în acest caz dă: $(*)$ $\lim_{n\to\infty} P\left(\frac{X_1+X_2+...+X_n-n\mu}{\sqrt{n\mathfrak{p}\mathfrak{q}}} < x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-\frac{t^2}{2}}dt$ (reamintim că dispersia v.a X_n este $\mathfrak{p}\mathfrak{q}$). Formula $(*)$ poartă numele de teorema DeMoivre-Laplace şi este prima formulare a teoremei limită centrală. Deducem, că mai sus, pentru $a \leq b$

$(**)$ $\lim_{n\to\infty} P\left(a \leq \frac{X_1+X_2+...+X_n-n\mathfrak{p}}{\sqrt{n\mathfrak{p}\mathfrak{q}}} < b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}dt = \Phi(b) - \Phi(a)$

În cazul acesta se poate afirma mai mult: anume că limita din $(**)$ este uniformă în raport cu a, b. Fără să intrăm în detalii privind exprimarea precisă a acestui fapt, remarcăm că înlocuirea, pentru n suficient de mare a $P\left(a \leq \frac{X_1+X_2+...+X_n-n\mathfrak{p}}{\sqrt{n\mathfrak{p}\mathfrak{q}}} < b\right)$ cu $\Phi(b) - \Phi(a)$ indiferent de a, b înseamnă că se poate lucra "global" cu repartiţia normală.

Exemplu. Folosind teorema Moivre-Laplace să estimăm probabilitatea ca din 100 de aruncări ale unei monede să apară stema de un număr de ori cuprins între 40 si 60. Pentru aceasta să notăm $\xi_n = \frac{X_1+X_2+...+X_n}{n}$ frecvenţa relativă de apariţie a stemei (v.a $X_1, X_2, ..., X_n...$ sunt independente şi la fel repartizate luând valoarea 1, în cazul apariţiei stemei cu probabilitatea $\mathfrak{p} = 0,5$). În cazul de faţă $n = 100$ şi estimăm $P(0,4 \leq \xi_n \leq 0,6)$. Aproximăm $P\left(a \leq \sqrt{\frac{n}{\mathfrak{p}\mathfrak{q}}}(\xi_n - 0,5) \leq b\right)$cu $\Phi(b) - \Phi(a)$; rezultă $a = \sqrt{\frac{100}{0,25}}(0,4 - 0,5) = -2$ şi $b = \sqrt{\frac{100}{0,25}}(0,6 - 0,5) = 2$. Probabilitatea căutată va fi $\Phi(2) - \Phi(-2)$. Din tabele rezultă $\Phi(2) = 0,9772$ iar $\Phi(-2) = 1 - \Phi(2)$. Deci $\Phi(2) - \Phi(-2) = 2\Phi(2) - 1 = 0,9544$.

Semnalăm şi un rezultat care priveşte aproximarea probabilităţilor legii binomiale cu valorile densităţii de repartiţie a legii normale reduse. Pentru a enunţa acest rezultat să considerăm repartiţia binomială $B(n, \mathfrak{p})$, $0 < \mathfrak{p} < 1$ şi să notăm $p_n(k)$ probabilitatea ca o v.a $B(n, \mathfrak{p})$ repartizată să ia valoarea k $(0 \leq k \leq n)$. Dacă pentru $x \in \mathbb{R}$ notăm $\langle x\rangle$ întregul cel mai apropiat de x (cu o conventie oarecare când sunt doi candidaţi) atunci

$(***)$ $\lim\limits_{n\to\infty} \sqrt{n\mathfrak{p}\mathfrak{q}}\,p_n(n\mathfrak{p} + x\sqrt{n\mathfrak{p}\mathfrak{q}}) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ pentru $x \in \mathbb{R}$ (în plus există anumite precizări de uniformitate a convergenţei asupra cărora nu vom insista.

Rezultatul obţinut se aplică pentru calculul aproximativ al probabilităţilor $p_n(k)$. În adevăr, scriem $k = n\mathfrak{p} + x\sqrt{n\mathfrak{p}\mathfrak{q}}$ de unde $x = \frac{k-n\mathfrak{p}}{\sqrt{n\mathfrak{p}\mathfrak{q}}}$ şi deci aproximaţia probabilităţii $p_n(k)$ cu $\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{n\mathfrak{p}\mathfrak{q}}}$.

Exemplu. Să calculăm pe baza rezultatelor de mai sus $p_{100}(55)$ pentru $\mathfrak{p} = 0.5$. Avem $x = \frac{55-50}{5} = 1$ deci valoarea aproximativă este $\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}}}{5} = 0,484$ (se poate verifica acurateţea aproximării; cel puţin trei zecimale exacte).

12.4 MF.12.4. Funcţii caracteristice.

Desi nu vom utiliza în cele ce urmează funcţiile caracteristice ale v.a vom prezenta, pe scurt, definiţia si proprietăţile principale ale acestora având în vedere importanţa subiectului în teoria generală a probabilităţilor.

Fie (Ω, Δ, P) un câmp de probabilitate şi $\Lambda : \Omega \to \mathbb{C}$, $\Lambda = U + iV$ o funcţie cu valori complexe. Vom spune că Λ este măsurabilă dacă funcţiile reale U, V sunt măsurabile; vom spune că Λ este integrabilă dacă $U, V \in L^1$ şi în acest caz vom pune $\int_\Omega \Lambda dP = \int_\Omega U dP + i \int_\Omega V dP$. Proprietăţile integralei funcţiilor cu valori complexe se deduc cu uşurinţă din proprietăţile integralei pentru v.a. Vom nota $\int_\Omega \Lambda dP = M[\Lambda]$

Fie X o v.a. Pentru $t \in \mathbb{R}$ funcţia $\omega \mapsto e^{itX(\omega)}$ va fi notată e^{itX}; este uşor de văzut că această funcţie este integrabilă ($\left|e^{itX(\omega)}\right| \leq 1$ pentru orice $\omega \in \Omega$ etc). Se defineşte funcţia caracteristică a v.a X ca integrala cu parametru $g_X(t) = M\left[e^{itX}\right]$. În cazul unei v.a discrete care ia valorile x_k cu probabilităţile p_k (k parcurgând o mulţime finită sau numărabilă de indici) vom avea $g_X(t) = \sum\limits_k e^{itx_k} p_k$ iar pentru o v.a continuă de densitate f vom avea $g_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx$. Se vede legătura dintre funcţia caracteristică şi transformata Fourier.

Exemplu. Daca X este $N(0,1)$ repartizată, atunci $g_X(t) = e^{-\frac{t^2}{2}}$ (a se vedea un curs de transformare Fourier).

Proprietăţile funcţiei caracteristice. Fie X o v.a si g_X funcţia sa caracteristică.

1) $g_X(0) = 1$, $|g_X(t)| \leq 1$.

2) g_X este o funcţie uniform continuă.

3) $g_{aX}(t) = g_X(at)$, $\forall t \in \mathbb{R}$.

4) Dacă X, Y sunt v.a independente, atunci $g_{X+Y} = g_X g_Y$.

5) Daca X are moment de ordin $k \geq 1$, atunci g_X are derivata de ordinul k şi aceasta este continuă şi $g_X^{(k)}(0) = i^k M\left[X^k\right]$.

Primele patru proprietăţi sunt uşor de demonstrat şi le propunem ca exerciţiu.

Exerciţiu. V.a X este $B(n, \mathfrak{p})$ repartizată. Să se determine funcţia sa caracteristică şi să se regăsească $M[X]$.

Soluţie. Avem $g_X(t) = \sum_{k=0}^{n} e^{itk} C_n^k \, \mathfrak{p}^k \mathfrak{q}^{n-k} = \left(\mathfrak{p}\ e^{it} + \mathfrak{q}\right)^n$. Avem $g_X^{'}(t) = n\mathfrak{p}ie^{it}\left(\mathfrak{p}\ e^{it} + \mathfrak{q}\right)^{n-1}$ deci $g_X^{'}(0) = in\mathfrak{p}$ şi regăsim $M[X] = n\mathfrak{p}$.

Pentru acest capitol se recomandă lucrarile:[C.01] , [G.01] , [G.03] , [L.01] , [S.01].

Capitolul 13

MF.13. Lanţuri Markov

Cuvinte cheie

Matrice de trecere, lanţ Markov, mers la întâmplare, proces stochastic.

În prima sectiune se introduce noţiunea de lanţ Markov, se stabilesc câteva proprietăţi importante şi se prezintă exemple. Nu se va intra în probleme de amănunt scopul nostru fiind a familiariza pe student cu primii paşi în această importanţă (atât teoretic cât şi aplicativ) parte a studiului probabilităţii. Cea de a doua secţiune va defini noţiunea generală de proces stochastic şi o va ilustra cu câteva exemple.

13.1 MF.13.1. Lanţuri Markov.

Pentru o mai bună înţelegere a celor ce urmează vom începe prin a introduce noţiunea de **sistem dinamic determinist**, în timp discret. Fie S o mulţime nevidă şi $f : S \to S$ o funcţie. Perechea (S, f) se numeşte sistem dinamic (determinist). Pentru a explica această terminologie să considerăm timpul (discret) ca fiind mulţimea $\mathbb{N}$ a numerelor naturale (momentele de timp) şi să notăm, pentru $n \in \mathbb{N}$, $f^n = id_S$ dacă $n = 0$ şi $f \circ f \circ ... \circ f, n$ termeni în rest. Interpretăm S ca **mulţimea de stări** a unui ”sistem” iar funcţia f astfel: dacă sistemul se găseşte la momentul n în starea s atunci, la momentul $n + 1$ se va găsi în starea $f(s)$. Observăm că starea la momentul $n + 1$ depinde doar de starea la momentul n şi nu de întregul ”trecut”. Dându-se o condiţie iniţială s_0 (la momentul 0) atunci se obţine o evoluţie prin aplicarea iterată a funcţiei f: $s_0, f(s_0), f^2(s_0) ... f^n(s_0) ....$ Evoluţia este deterministă pentru că starea la momentul $n+1$ este univoc determinată de starea la momentul n (si evident de f) şi tranziţia de la un moment la altul nu depinde de timp (sistemul este autonom). În ciuda acestui cadru foarte simplu se pot observa fenomene foarte interesante în studiul traiectoriilor (în sens clar) sistemelor deterministe. Asfel un fenomen foarte mult studiat în ultima vreme este haosul în sisteme deterministe (intuitiv, o sensibilă dependentă a traiectoriilor de condiţiile initiale); deja un caz aparent simplu ca $S = [0,1]$ si $f(x) = 4x(1-x)$ prezintă acest fenomen. Nu vom intra în amănunte (pentru cei interesaţi, se poate consulta cartea Robert L. Dewaney "An introduction to Chaotic Dynamical Systems" Addison-Wesley 1987).

Există situaţii când tranziţia de la o stare la alta nu este determinată cu precizie ci rezultă din starea prezentă şi din rezultatul unui fenomen (experiment) aleator (ca de exemplu aruncarea unui zar). Astfel, cunoscând starea prezentă, starea viitoare este determinată doar probabilistic adică se cunosc probabilitătile de trecere din starea prezentă în fiecare dintre stările posibile. Un model al acestei situaţii este oferit de noţiunea de lanţ Markov şi în general de noţiunea de proces stochastic. Vom începe prin a defini lanţul Markov iar la sfârşitul acestui capitol vom defini şi ilustra noţiunea de proces stochastic.

Lanţ Markov. Vom reprezenta **timpul discret** ca mulţimea $\mathbb{N}$ a numerelor naturale numite şi momente; astfel 0 este momentul iniţial etc. Fie S o **mulţime finită nevidă** ale cărei elemente vor fi numite **stări.** Vom fixa o **numerotare a stărilor** şi vom identifica stările cu numerele respective; astfel funcţiile cu valori în S vor fi considerate cu valori numere (şi deci variabile aleatoare când vor fi definite pe câmpuri de probabilitate) şi vom putea lucra cu matrice cu indici în mulţimea stărilor). Vom nota generic stările prin $i, j...$ sau $i_n..$ unde rolul indicelui este de a preciza un anume moment de timp (astfel i_n este un număr (reprezentând o stare) considerat la momentul $n \in \mathbb{N}$). Vom considera o repartiţie de probabilitate pe S, $p = (p(i))_{i\in S}$, $0 \le p(i) \le 1, \sum\limits_i p(i) = 1$ şi o matrice $\Pi = (p(i,j))_{i,j\in S}$ astfel încât $0 \le p(i,j) \le 1, \sum\limits_j p(i,j) = 1, \forall i \in S$ (o astfel de matrice se numeşte **matrice stochastică**, elementele matricei sunt ≥ 0 şi suma elementelor de pe fiecare linie este 1). Un **lanţ Markov omogen finit** cu spaţiul stărilor S, **probabilităţile iniţiale** $p(i)$ şi **matricea de trecere** Π este un şir $(X_n)_{n\in\mathbb{N}}$ de v.a definite pe un acelaşi câmp de probabilitate şi cu valori în S astfel încât:

a) $P(X_0 = i) = p(i), i \in S.$

b) $P(X_{n+1} = i_{n+1} \mid X_n = i_n, X_{n-1} = i_{n-1}, ..., X_0 = i_0) =$
$= P(X_{n+1} = i_{n+1} \mid X_n = i_n) = p(i_n, i_{n+1}), \forall n$ şi $\forall i_0, ..., i_{n-1}, i_n \in S$ (dacă prima probabilitate este definită).

Condiţia a) spune că starea iniţială este cunoscută, în general, doar probabilistic: probabilitatea ca starea iniţială să fie i este $p(i)$.

Condiţia b) spune că pentru determinarea probabilităţilor stărilor la momentul $n+1$ este nevoie doar de cunoaşterea stării la momentul n şi nu de întreg trecutul (întreaga evoluţie până la momentul n) şi aceste probabilităţi sunt date de matricea de trecere (se explică astfel necesitatea ca matricea Π să fie stochastică). Prima egalitate din b) se numeste **proprietatea**

Markov. Adjectivul ”finit” se referă la finitudinea spaţiului stărilor iar adjectivul ”omogen” se referă la faptul că, în ciuda notaţiei putin ambiguă de la b), probabilităţile de trecere nu depind de momentul trecerii (la sisteme deterministe am numit autonomie o asemenea proprietate). Deoarece nu vom considera decât lanţuri Markov omogene finite vom omite adjectivele ”omogen” şi ”finit”.

Se poate arăta că date S, p, Π există un câmp de probabilitate şi un sir de v.a îndeplinind condiţiile a),b). Nu intrăm în detalii.

Exemplu 1. Vom considera un şir de persoane astfel încât persoana n poate comunica persoanei $n+1$ doar două mesaje numerotate $1, 2$. Prima persoană din şir primeşte mesajele 1 cu probabilitatea $p\left(1\right)$ şi 2 cu probabilitatea $p\left(2\right)$, $p\left(1\right)+p\left(2\right)=1$. Dacă o persoană primeşte mesajul $i=1,2$ atunci va comunica acelaşi mesaj cu probabilitatea $1-\mathfrak{p}$ şi celălalt mesaj cu probabiliatea $\mathfrak{p}$. Evident, aceste date permit definirea unui lanţ Markov cu $S=\{1,2\}, p=\left(p\left(1\right),p\left(2\right)\right)$ şi $\Pi=\begin{pmatrix} 1-\mathfrak{p} & \mathfrak{p} \\ \mathfrak{p} & 1-\mathfrak{p} \end{pmatrix}$.

Exerciţiu. Descrieţi cel mai general lanţ Markov cu două stări.

Exemplu 2. Fie o ”particulă” care poate ocupa poziţiile $0,1,...,l$ pe dreapta reală. Vom lua $S=\{0,1,...,l\}$ şi fie p o repartiţie pe S. Dacă particula se află în poziţia $i\in\{1,2,...,l-1\}$ atunci în momentul următor se va afla în poziţia $i+1$ cu probabilitatea $\mathfrak{p}$ şi în poziţia $i-1$ cu probabilitatea $1-\mathfrak{p}$ (trecerea în alte stări este deci de probabilitate 0). Dacă $i=0$ trecerea se face, cu probabilitate 1 în starea 1 şi dacă $i=l$ trecerea se face cu probabilitate 1 în $l-1$ (poziţiile $0,l$ sunt poziţiile frontieră şi condiţia pusă se exprimă prin ”frontiere reflectante”). Astfel matricea Π este dată de $p\left(i,i+1\right)=\mathfrak{p}, p(i,i-1)=1-\mathfrak{p}, p\left(i,j\right)=0, i\neq i-1,i+1$ pentru $i\in\{1,2,...,l-1\}$si $p\left(0,1\right)=1, p\left(0,j\right)=0, j\neq 1$ şi $p\left(l,l-1\right)=1, p\left(l,j\right)=0, j\neq l-1$. Se obţine un lanţ Markov numit **mers la întâmplare** (cu frontiere reflectante).

Exerciţiu (un model de difuzie). Se consideră două urne, prima conţinând iniţial l bile albe iar cea de a doua l bile negre. Se extrage, la întâmplare câte o bilă din fiecare urnă şi bila extrasă dintr-o urnă se pune în cealaltă. Se continuă indefinit. Considerând ca mulţime a stărilor numărul de bile albe din urna care conţine iniţial doar bile albe descrieţi procesul ca un lanţ Markov. Gândind bilele ca un model pentru molecule lanţul Markov obţinut poate fi considerat ca un model de difuzie.

Repartiţia comună. Revenind la teoria generală a lanţurilor Markov să observăm că, din definiţia lanţului va rezulta posibilitatea estimării repartiţiei comune a v.a $X_0, X_1, ..., X_n$ pentru orice $n>0$. Astfel avem:

$(*)$ $P\left(X_0=i_0, X_1=i_1, ..., X_{n-1}=i_{n-1}, X_n=i_n\right)=p\left(i_0\right)p\left(i_0,i_1\right)p\left(i_1,i_2\right)...p\left(i_{n-1},i_n\right)$

Să demonstrăm această relaţie prin inducţie . Pentru $n=1$ avem $P\left(X_0=i_0, X_1=i_1\right)=P\left(X_0=i_0\right)P\left(X_1=i_1\mid X_0=i_0\right)=p\left(i_0\right)p\left(i_0,i_1\right)$. Presupunând relaţia adevarată pentru n vom avea

$P\left(X_0=i_0, X_1=i_1, ..., X_{n-1}=i_{n-1}, X_n=i_n, X_{n+1}=i_{n+1}\right)=$

$= P(X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = i_n) \cdot$
$\cdot P(X_{n+1} = i_{n+1} \mid X_0 = i_0, X_1 = i_1, ..., X_n = i_n) =$
$= P(X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = i_n) P(X_{n+1} = i_{n+1} \mid X_n = i_n) =$
$p(i_0) p(i_0, i_1) p(i_1, i_2) ... p(i_{n-1}, i_n) p(i_n, i_{n+1})$ (remarcăm folosirea esenţială a proprietăţii Markov).

Ecuaţia (relaţia) Chapman-Kolmogorov. Vom considera un lanţ Markov şi vom defini probabilităţile de trecere în n paşi astfel: punem $p(0, i, j) = \delta(i, j), i, j \in S$ (reamintim că $\delta(i, j) = 1, i = j$ şi $\delta(i, j) = 0, i \neq j$) şi $p(n, i, j) = P(X_{m+n} = j \mid X_m = i), i, j \in S\ \forall m \in \mathbb{N}$ (independenţa de m este o consecinţă a omogenităţii). Remarcăm că $p(1, i, j) = p(i, j)$ deci matricea $(p(1, i, j))_{i,j \in S} = \Pi$. Avem $p(2, i, j) = \sum\limits_k p(i, k) p(k, j)$ deci matricea $(p(2, i, j))_{i,j \in S} = \Pi^2$ şi mai general $(p(n, i, j))_{i,j \in S} = \Pi^n$ (pentru $n = 0$ avem matricea unitate, prin definiţie). În acest mod teoria lanţurilor Markov devine studiul matricelor stochastice problemele fiind cele ridicate de comportarea dinamică a lanţului. Din relaţia cunoscută $\Pi^{m+n} = \Pi^m \Pi^n$ se deduce **relaţia Chapman-Kolmogorov**:

$$(**)\ p(m + n, i, j) = \sum_k p(m, i, k) p(n, k, j),\ i, j \in S.$$

Exerciţiu. Să se demonstreze că $p(n, i, j) = \sum\limits_{i_1, ..., i_{n-1}} p(i, i_1) p(i_1, i_2) ... p(i_{n-1}, j)$, $n \geq 2$.

Vom arăta că $P(X_n = j) = (p\Pi^n)_j$ unde matricea probabilităţilor iniţiale p este de tip $1 \times cardS$ ($cardS$ este numărul stărilor). În adevăr dacă $n = 0$, Π^0 este matricea unitate deci rezultatul este clar. Dacă $n = 1$ avem $P(X_1 = j) = \sum\limits_i P(X_0 = i, X_1 = j) = \sum\limits_i P(X_0 = i) P(X_1 = j \mid X_0 = i) =$

$= \sum\limits_i p(i) p(i, j) = \left(p\Pi^1\right)_j$. Acelasi raţionament pentru cazul $n > 1$ ţinând cont că matricea de trecere în n paşi este Π^n.

Exemplu 3. Să ne întoarcem la exemplul 1 şi să luăm $0 < \mathfrak{p} < 1$. Se arată uşor prin inducţie că: $2\Pi^n = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + (1 - 2\mathfrak{p})^n \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, n \geq 1$ (în fapt formula apare ca o consecinţă a unei proprietăţi a valorilor proprii ale matricei Π dar nu insistăm; pentru cei interesaţi este vorba de descompunerea spectrală a matricei Π). Se observă cu uşurinţă că limita şirului de matrice Π^n când $n \to \infty$ este matricea $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$(limita se face pe elemente). Interpretarea acestui rezultat este că, indiferent de probabilitatea de pornire, pentru n mare probabilităţile de transmitere a celor două mesaje tind să se egalizeze. Avem un exemplu de problemă care se pune pentru lanţurile Markov: studiul comportării asimptotice a evoluţiei dinamice.

13.2 MF.13.2. Procese stochastice.

Fie (Ω, Δ, P) un câmp de probabilitate si T o mulţime (ordonată) care are în general semnificaţia de "timp". Vom nota $\mathcal{E}$ mulţimea v.a definite pe Ω. Un **proces stochastic** în timp T este o funcţie $X : T \to \mathcal{E}$. Deci pentru fiecare $t \in T$, $X(t) : \Omega \to \mathbb{R}$ este o v.a (atenţie la notaţie, X nu este o v.a ci o familie de variabile aleatoare). Să remarcăm că a da X este echivalent cu a da o funcţie $\xi : \Omega \times T \to \mathbb{R}$, $X(t)(\omega) = \xi(\omega, t)$. Astfel din funcţia ξ, fixând t obţinem v.a $X(t)$ iar fixând $\omega \in \Omega$ obţinem o funcţie $\xi(\omega,) ; T \to \mathbb{R}$ numită **traiectorie** a procesului stochastic respectiv. V.a $X(t)$ reprezintă starea procesului la momentul t, stare cunoscută doar probabilistic. În general pentru mulţimea T se ia fie $\mathbb{N}$ fie un interval al dreptei reale $\mathbb{R}$ (cel mai des $[0, \infty)$ sau $(-\infty, \infty)$). În primul caz vorbim de **procese în timp discret** iar în cel de al doilea de **procese în timp continuu**. Avem exemple de procese în timp discret: un şir de variabile aleatoare şi, în particular, un lanţ Markov. În cazul $T = \mathbb{N}$ se utilizează notaţia tipică pentru şiruri X_n pentru $X(n)$. Scopul nostru este să dăm câteva exemple importante fără a intra în detalii privind teoria generală a proceselor stochastice.

Exemplul 1 (zgomot alb). Fie procesul $X = (X_n)_{n\in\mathbb{N}}$ astfel încât v.a X_n sunt $N(0, 1)$ repartizate şi independente. Un asemenea proces se numeşte **zgomot alb**.

Exemplul 2 (mers la întamplare). Fie $(X_n)_{n\in\mathbb{N}}$ un şir de v.a independente. Procesul $\mathcal{S} = (S_n)_n$ unde $S_n = X_0 + X_1 + ... + X_n$ se numeste **mers la întamplare** (un exemplu a fost studiat la lanţuri Markov).

Exemplul 3. Orice lanţ Markov este un proces stochastic.

Exemplul 4. Fie A, B două v.a. Procesul stochastic în timp continuu $X(t) = A \cos 2\pi\nu t + B \sin 2\pi\nu t$ unde ν este fixat reprezintă o mişcare oscilatorie cu poziţie şi viteză iniţială aleatoare (exerciţiu: care sunt traiectoriile acestui proces?).

Înainte de a da următorul exemplu să definim noţiunea de proces stochastic cu creşteri independente: un proces $t \mapsto X(t), t \in [0, \infty)$ se zice cu **creşteri independente** dacă pentru orice n natural şi orice $t_1 < t_2 < ... < t_n$ v.a $X(t_1), X(t_2) - X(t_1), ..., X(t_n) - X(t_{n-1})$ sunt independente.

Exemplul 5 (mişcare browniană). Un proces stochastic $t \mapsto B(t), t \in [0, \infty)$ se numeşte mişcare browniană (standard) dacă:

- $B(0) = 0$
- dacă $0 \leq s < t$ atunci v.a $B(t) - B(s)$ este $N\left(0, \sqrt{t-s}\right)$ repartizată (deci dispersia este $\sigma^2 = t - s$)
- traiectoriile sunt funcţii continue.
- procesul este cu creşteri independente.

Intuitiv $B(t)$ este poziţia la momentul t a unei particule în mişcare browniană. Între două momente de timp mişcarea particulei este rezultatul unui mare număr de ciocniri aleatoare deci conform teoremei limită centrală deplasarea poate fi gândită ca normal repartizată etc.

Exemplul 6 (proces Poisson). Un proces stochastic $t \mapsto N(t), t \in [0,\infty)$ este un proces Poisson dacă:

- $N(0) = 0$
- procesul este cu creşteri independente.
- $\exists\lambda > 0$ astfel încât dacă $0 \leq s < t$ atunci v.a $N(t) - N(s)$ urmează o lege Poisson de parametru $\lambda(t-s)$.

Reamintim că ultima condiţie înseamnă că v.a $N(t) - N(s)$ ia valori numere naturale şi $P(N(t) - N(s) = k) = \frac{(\lambda(t-s))^k}{k!}e^{-\lambda(t-s)}, k \in \mathbb{N}$. În particular rezultă că $N(t), t > 0$ este Poisson repartizată de parametru λt.

Intuiţia este că un proces Poisson "numără" apariţiile unui eveniment aleator în timp. De exemplu: sosirile la o staţie de servicii (în teoria cozilor de aşteptare), ofertele de serviciu, defectarea unor aparate, emisia de particule radioactive etc.

Este util să facem câteva calcule privind un proces Poisson.
Astfel $P(N(1) = 0) = e^{-\lambda}$ deci $e^{-\lambda}$ este probabilitatea ca în unitatea de timp să nu aibă loc nici un eveniment. Avem $P(N(t) = 1) = \lambda te^{-\lambda t} = \lambda t(1 - \lambda t + \frac{\lambda^2 t^2}{2!} - ...) = \lambda t + o(t)$ (adică $\lim\limits_{t\to 0}\frac{o(t)}{t} = 0$) deci probabilitatea de apariţie a unui eveniment într-un interval de durată t este proporţională cu t şi aproximarea este cu atât mai bună cu cât t este mai mic. Mai departe avem $P(N(t) \geq 2) = 1 - e^{-\lambda t} - \lambda te^{-\lambda t} = \frac{\lambda^2 t^2}{2} + o(t^2)$ de unde $\lim\limits_{t\to 0}\frac{P(N(t)\geq 2)}{t} = 0$. Rezultă că pentru t "mic" probabilitatea de a avea mai mult de un eveniment într-un interval de lungime t este neglijabilă.

Fie $\mathcal{T}$ v.a "durata dintre două evenimente succesive". Avem $P(\mathcal{T} > t) = P(N(t) = 0) = e^{-\lambda t}$ şi deci $P(\mathcal{T} \leq t) = 1 - e^{-\lambda t}$. Se deduce că $\mathcal{T}$ este exponenţial repartizată deci fără memorie.

În finalul acestui capitol două definiţii:

Media unui proces stochastic. Dacă $t \mapsto X(t)$ este un proces astfel încât $X(t) \in L^1$ pentru orice t atunci funcţia $t \mapsto M[X(t)]$ se numeşte **media procesului** respectiv.

Funcţia de (auto)corelaţie. Dacă $t \mapsto X(t)$ este un proces stochastic astfel încât $X(t) \in L^2$ pentru orice t atunci funcţia $(t,s) \mapsto cov[X(t), X(s)]$ se numeşte funcţia de (auto)corelaţie a procesului.

Pentru acest capitol se recomandă lucrările:$[G.01], [I.02], [P.01], [M.02], [N.01]$.

Capitolul 14

MF.14. Statistică Matematică

Cuvinte cheie

Selecţie, repartiţie empirică, momente de selecţie, estimarea parametrilor, intervale de încredere, verificarea ipotezelor statistice, testul χ^2.

14.1 MF.14.1. Selecţie, repartiţie empirică, momente de selecţie.

În [E.01] găsim următoarea definiţie a Statisticii Matematice: "Disciplină matematică consacrată elaborării noţiunilor şi metodelor specifice studiului statistic al fenomenelor de masă". Se specifică apoi că Statistica Matematică este bazată pe Teoria Probabilităţilor. Mai precis cităm; "Statistica Matematică se ocupă cu gruparea, interpretarea şi analiza datelor referitoare la anumite fenomene precum şi cu unele previziuni privind producerea lor în viitor".

Un exemplu tipic îl constituie situaţia în care trebuie obţinute informaţii asupra unei mulţimi (finite) numeroase de obiecte (sau persoane) fără a avea posibilitatea examinării exhaustive a elementelor mulţimii (analiza de tip recensământ); în acest caz se extrage (aleator) un eşantion din mulţimea respectivă, se examinează şi se trag concluzii (cu caracter probabilist) asupra întregii mulţimi. Alegerea eşantionului trebuie să îndeplinească anumite condiţii pentru ca rezultatele obţinute să fie acceptabile (în marjă de eroare cât mai mică etc).

Terminologia utilizată în situaţia de mai sus este:

Populaţie (statistică) - mulţimea asupra căreia se efectuează studiul în cauză (termenul "populaţie" nu se referă neapărat la persoane).

Caracteristică - proprietatea studiată de statistician (de exemplu "vârsta", "venitul", "opţiunea politică", "durata de funcţionare fără reparaţie a unui tip de aparat de serie" etc).

Inferenţa statistică - tragerea de concluzii cu caracter probabilist asupra întregii populaţii (în studiu) din examinarea unui eşantion din populaţia respectivă şi justificarea acestor concluzii.

Selecţie - operaţia de extragere a unui eşantion dintr-o populaţie statistică.

Volumul unei selecţii - numărul de elemente al eşantionului.

Sondaj - prelevarea unui eşantion.

În general vom presupune că selecţia se face **prin reintroducere**, adică după extragerea unui element oarecare (şi examinarea sa) acesta este reintrodus în populaţie înainte de următoarea extragere. Vom presupune, de asemenea că elementele au şanse egale de a fi selectate (egal probabilitate la selectie). Volumul selecţiei se stabileşte conform preciziei dorite în determinarea caracteristicii studiate.

Din punct de vedere matematic orice caracteristică va fi asimilată unei variabile aleatoare (v.a).

Selecţie de volum n. Vom numi selecţie de volum n asupra v.a X (care modelează o caracteristică anume) un vector aleator n-dimensional $(X_1, X_2, ..., X_n)$ ale cărui componente sunt la fel repartizate ca X şi sunt independente. Dacă notăm (Ω, Δ, P) câmpul de probabilitate pe care sunt definite v.a $X_1, X_2, ..., X_n$ atunci pentru $\omega \in \Omega$ avem o **realizare** a selecţiei $(X_1(\omega), X_2(\omega), ..., X_n(\omega)) \in \mathbb{R}^n$ (adică valorile observate prin analiza unui eşantion).

Trebuie remarcat că noţiunea de selecţie reprezintă modelul matematic al ideii intuitive de extrageri (succesive) la întâmplare şi independent a unor elemente dintr-o populatie în vederea studiului unei caracteristici asimilate unei v.a.

Exemplu. O populaţie are elemente de două "tipuri" a şi b (de pildă, o urnă are doar bile albe şi bile negre sau un lot de conserve are conserve acceptabile şi conserve neacceptabile etc). Considerăm v.a X care ia valoarea 1 pe elementele de tip a şi valoarea 0 pe elementele de tip b. O selecţie de volum n asupra v.a X este un vector aleator n-dimensional $(X_1, X_2, ..., X_n)$ ale cărui componente sunt independente şi la fel repartizate ca X. Aceasta modelează extragerile aleatoare şi independente cu reintroducere din populaţia respectivă. Încă nu avem o "problemă" pe care să o rezolve Statistica; problema apare imediat ce presupunem, spre exemplu, că repartiţia v.a X este necunoscută şi dorim s-o determinăm analizând un eşantion (o realizare a selecţiei) . Vom reveni asupra acestei chestiuni.

Funcţie de repartiţie empirică. Fie $(X_1, X_2, ..., X_n)$ o selecţie de volum n asupra v.a X . Definim **funcţia de repartiţie empirică a selecţiei** $(X_1, X_2, ..., X_n)$ ca o funcţie de două variabile $F_n : \mathbb{R} \times \Omega \to \mathbb{R}, F_n(x, \omega) = \frac{1}{n} card\{i; X_i(\omega) < x\}$ ($cardM$ este notaţia pentru numărul de elemente al mulţimii finite M, $card\emptyset = 0$).

Pentru o mai bună înţelegere a funcţiei de repartiţie empirică să observăm că pentru $\omega \in \Omega$ se consideră realizarea $(X_1(\omega), X_2(\omega), ..., X_n(\omega))$ şi apoi pentru fiecare $x \in \mathbb{R}$ se consideră frecvenţa relativă a apartenenţei componentelor la intervalul $(-\infty, x)$.

Să considerăm câmpul de probabilitate determinat de mulţimea $\{1, 2, ..., n\}$ σ-algebra tuturor părţilor şi egal probabilitate. O v.a Y pe acest câmp poate fi privită ca un vector în $\mathbb{R}^n$, fie acesta $(y_1, y_2, ..., y_n)$ (unde $y_i = Y(i), i = 1, 2, ..., n$). Să determinăm funcţia de repartiţie a v.a Y: vom avea $F_Y(x) = P(Y < x) = \frac{1}{n} card\{i : y_i < x\}$ (un simplu exerciţiu asupra funcţiilor de repartiţie).

Am prezentat această situaţie pentru a înţelege mai bine natura funcţiei de repartiţie empirică: anume, pentru $\omega \in \Omega$ fixat putem interpreta realizarea $(X_1(\omega), X_2(\omega), ..., X_n(\omega))$ ca o v.a pe $\{1, 2, ..., n\}$ etc. Rezultă că pentru $\omega \in \Omega$ fixat funcţia $x \mapsto F_n(x, \omega)$ este o funcţie de repartiţie.

Să studiem natura funcţiei de repartiţie empirică pentru $x \in \mathbb{R}$ fixat. Pentru aceasta să notăm $A = (X < x)$ şi U v.a care ia valoarea 1 pe A şi valoarea 0 în rest. Selectia $(X_1, X_2, ..., X_n)$ induce o selecţie de volum n asupra v.a $(U_1, U_2, ..., U_n)$. Să remarcăm că $F_n(x, \omega) = \frac{U_1(\omega)+U_2(\omega)+...+U_n(\omega)}{n}$. Deci, pentru fiecare $x \in \mathbb{R}$ funcţia $\omega \mapsto F_n(x, \omega)$ este o v.a.

În acest mod am stabilit că funcţia de repartiţie empirică este "funcţie de repartiţie în x şi v.a în ω".

Fie X o v.a şi $(X_n)_n$ un şir de v.a independente şi la fel repartizate ca X. În mod evident, pentru fiecare n şirul induce o selecţie $(X_1, X_2, ..., X_n)$ de volum n asupra v.a X. Fie $(F_n)_n$ şirul de funcţii de repartiţii empirice asociat.

Pentru fiecare $x \in \mathbb{R}$şirul $(F_n(x,))_n$ converge în probabilitate către funcţia de repartiţie $F(x)$ (F este funcţia de repartiţie a v.a X).

Pentru demonstraţie este suficient să aplicăm legea numerelor mari pentru v.a U (a se vedea notaţia de mai sus).

Aplicând legea tare a numerelor mari rezultă convergenţa aproape sigură în rezultatul precedent.

Ca o observaţie importantă trebuie remarcat că există un rezultat (teorema Glivenko-Cantelli) care arată că, în situatiile de mai sus, convergenţa este chiar uniformă în raport cu $x \in \mathbb{R}$ ceea ce întăreste considerabil afirmaţiile privind convergenţa şirului de funcţii de repartiţie empirică la funcţia de repartiţie a caracteristicii. Valoarea practică a rezultatului este clară: se poate determina, aproximativ, repartiţia v.a X utilizând eşantioane şi calculând repartiţia empirică respectivă.

Accentuăm că cele expuse se bazează esenţial pe legea numerelor mari (într-o formă sau alta).

Momente de selecţie. Fie $(X_1, X_2, ..., X_n)$ o selecţie de volum n asupra v.a X. Vom defini doar două momente de selecţie ale v.a X:

Media de selecţie. Definim **media de selecţie** (n fixat) a v.a X ca fiind v.a $\overline{X_n} = \frac{X_1+X_2+...+X_n}{n}$. Avem dacă $X \in L^1$ (reamintim că această notaţie înseamnă ca v.a X are medie):

$M\left[\overline{X_n}\right] = \frac{1}{n}\left(M\left[X_1\right] + M\left[X_2\right] + ... + M\left[X_n\right]\right) = M\left[X\right]$

şi dacă $X \in L^2$ (X are moment de ordin 2) atunci

$D\left[\overline{X_n}\right] = \frac{1}{n^2}\left(D\left[X_1\right] + D\left[X_2\right] + ... + D\left[X_n\right]\right) = \frac{1}{n}D\left[X\right]$

Fie acum un şir $\left(X_n\right)_n$ de v.a independente şi la fel repartizate ca X. Rezultă definit şirul $\left(\overline{X_n}\right)_n$ al mediilor de selecţie corespunzătoare. Aplicând legea numerelor mari deducem că şirul $\left(\overline{X_n}\right)_n$ tinde în probabilitate către $M\left[X\right]$. Vom relua acest rezultat mai jos în legătură cu estimarea parametrilor: dacă considerăm o realizare $\left(X_1\left(\omega\right), X_2\left(\omega\right), ..., X_n\left(\omega\right)\right)$, pentru n suficient de mare, atunci numărul $\frac{X_1(\omega)+X_2(\omega)+...+X_n(\omega)}{n}$ constituie o "bună" aproximare pentru $M\left[X\right]$ (estimează în sensul convergenţei în probabilitate media v.a X).

Se obsevă că $D\left[\overline{X_n}\right] \to 0$.

Dispersia de selecţie. Definim **dispersia de selecţie (**n fixat**)** $S_n^2 = \frac{1}{n}\sum\limits_{i=1}^{n}\left(X_i - \overline{X_n}\right)^2$. Avem (pentru $X \in L^2$):

$M\left[S_n^2\right] = M\left[\frac{1}{n}\sum X_i^2 - \frac{2}{n}\bar{\bar{X}}_n\sum X_i + \overline{X_n}^2\right] = M\left[\frac{1}{n}\sum X_i^2 - \bar{\bar{X}}_n^2\right] = M\left[X^2\right] - M\left[\overline{X_n}^2\right]$. Un calcul analog arată că $M\left[\overline{X_n}^2\right] = \frac{1}{n}M\left[X^2\right] + \frac{n-1}{n}M\left[X\right]^2$ şi deci, în definitiv

$M\left[S_n^2\right] = M\left[X^2\right] - \frac{1}{n}M\left[X^2\right] - \frac{n-1}{n}M\left[X\right]^2 = \frac{n-1}{n}\left(M\left[X^2\right] - M\left[X\right]^2\right) = \frac{n-1}{n}D\left[X\right]$.

Să considerăm şi $S_n'^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i - \bar{\bar{X}}_n\right)^2$; rezultă $M\left[S_n'^2\right] = D\left[X\right]$. Din punct de vedere al estimării dispersiei cu ajutorul dispersiei de selecţie $S_n'^2$ are avantajul că media acestei v.a este chiar dispersia (de estimat). Vom discuta acest aspect în secţiunea următoare. Se vede că, pentru n suficient de mare S_n^2 si $S_n'^2$ sunt foarte apropiate.

Fie acum un şir $\left(X_n\right)_n$ de v.a independente şi la fel repartizate ca X şi $\left(S_n^2\right)_n$ şirul corespunzător al dispersiilor de selecţie. Din legea numerelor mari rezultă că şirul $\left(\frac{1}{n}\sum\limits_{i=1}^{n}X_i^2\right)_n$ tinde în probabilitate (chiar aproape sigur) către $M\left[X^2\right]$ iar şirul $\left(\bar{\bar{X}}_n^2\right)_n$ tinde în probabilitate (chiar aproape sigur) către $M\left[X\right]^2$. Avem $S_n^2 = \frac{1}{n}\sum\limits_{i=1}^{n}\left(X_i - \dfrac{X_1 + X_2 + ... + X_n}{n}\right)^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2 - \overline{X_n}^2$ deci şirul $\left(S_n^2\right)_n$ converge în probabilitate (aproape sigur) către $M\left[X^2\right] - M\left[X\right]^2 = D\left[X\right]$. Acest rezultat este important din acelasi motiv ca la media de selecţie: realizările selecţiei aproximează dispersia v.a X.

14.2 MF.14.2. Estimarea parametrilor. Estimaţii punctuale.

Vom relua o problemă care a mai fost discutată mai sus (secţiunea precedentă, exemplu) şi sub forma unui exerciţiu la capitolul "Legea numerelor mari". Într-o urnă sunt bile albe şi bile negre şi ne propunem ca printr-un şir de extrageri independente, cu reintroducere, să estimăm, cu o precizie dorită, proporţia de bile albe (în ipoteza că avem egal probabilitate la extragerea unei bile). Considerăm v.a X care ia valoarea 1 pe bilele albe şi valoarea 0 pe cele negre. Atunci proporţia pe care o căutăm este egală cu probabilitatea $\mathfrak{p}$ de a extrage o bilă albă şi egală cu media v.a X. Se ştie din cele de mai sus că şirul $\left(\overline{X_n}\right)_n$ tinde în probabilitate către $\mathfrak{p}$. Mai precis avem pentru, $\varepsilon > 0$, $P\left(\left|\overline{X_n} - \mathfrak{p}\right| < \varepsilon\right) \geq 1 - \frac{1}{4n\varepsilon^2}$. Alegând o abatere admisibilă ε şi probabilitate $1-\alpha$ suficient de apropiată de 1 vom putea determina n astfel încât $1 - \frac{1}{4n\varepsilon^2} \geq 1 - \alpha$. Vom efectua n extrageri şi fie $x_1, x_2, ..., x_n$ rezultatele obţinute. Atunci vom lua pentru $\mathfrak{p}$ valoarea $\frac{x_1+x_2+...+x_n}{n}$ stiind că suntem în limitele abaterii acceptate cu o mare probabilitate. Avem un exemplu de utilizare a unei selecţii pentru a estima un parametru legat de o v.a (anume media acesteia dar, în acest caz, un parametru care determină complet repartiţia acestei v.a). Reformulând putem considera că avem un model matematic al unei experienţe aleatoare sub forma unei v.a al cărui "tip" se cunoaşte dar determinarea concretă a acestei v.a necesită estimarea unui parametru.

În fond, după această discuţie tehnică bazată pe noţiuni matematice delicate şi pe rezultate fundamentale (legea numerelor mari) rămâne un "algoritm" simplu pentru estimarea mediei unei v.a: se consideră o realizare a unei selecţii de volum suficient de mare iar apoi se face media aritmetică a valorilor obtinute. De exemplu interesează venitul mediu anual al unei populaţii (un oraş o regiune etc). Se face un sondaj de volum convenabil şi se face media valorilor obţinute. Desigur sunt probleme importante privind volumul sondajului, pragul de eroare fixat, alegerea cât mai variată a indivizilor interogaţi etc, dar acestea sunt chestiuni pe care nu le discutăm în acest curs (asupra volumului selecţiei am discutat chiar mai sus).

Să mai considerăm un exemplu de estimare a unor parametri. Să presupunem că se încearcă determinarea unei valori numerice (de exemplu o constantă fizică) μ prin măsurători repetate afectate de erori aleatoare. Dacă notăm X v.a a rezultatelor măsurătorilor atunci putem propune, ca model teoretic al fenomenului de măsurare, $X = \mu + \mathcal{E}$ unde $\mathcal{E}$ este v.a a erorilor de măsurare pe care o presupunem independentă de μ. Presupunând că erorile de măsurare la diferite măsurători sunt independente şi la fel repartizate o ipoteză clasică afirmă că v.a $\mathcal{E}$ este $N(0,\sigma)$ repartizată. Rezultă că v.a X este $N(\mu,\sigma)$ repartizată. Fie $x_1, x_2, ..., x_n$ o realizare a unei selecţii $X_i = \mu + \mathcal{E}_i, i = 1, 2, ..., n$ asupra v.a X. Valorile estimate pentru μ şi σ^2 vor fi $\frac{x_1+x_2+...+x_n}{n}$ respectiv $\frac{1}{n}\sum\limits_{i=1}^{n}\left(x_i - \frac{x_1+x_2+...+x_n}{n}\right)^2$ (a se vedea dispersia de selecţie). Se notează $\bar{x}_n = \frac{x_1+x_2+...+x_n}{n}$ şi valoarea estimată pentru dispersie se scrie mai simplu $\frac{1}{n}\sum\limits_{i=1}^{n}(x_i - \bar{x}_n)^2$.

Vom formaliza procesul de estimare a parametrilor după cum urmează:

Fie $(X_1, X_2, ..., X_n)$ o selecţie de volum n asupra v.a X , $u_n : \mathbb{R}^n \to \mathbb{R}$ o funcţie continuă (mai general o funcţie măsurabilă) si $u_n(X_1, X_2, ..., X_n)$ funcţia compusă corespunzătoare (care este o v.a). Vom numi v.a $u_n(X_1, X_2, ..., X_n)$ o **statistică** (n-statistică) sau **estimator**. De exemplu fie $u_n(x_1, x_2, ..., x_n) = \frac{x_1+x_2+...+x_n}{n} = \bar{x}_n$. Fie acum $(X_n)_n$ un sir de v.a independente şi la fel repartizate ca X şi o familie de funcţii continue $u_n : \mathbb{R}^n \to \mathbb{R}$, $n \geq 1$. Atunci şirul de statistici $u_n(X_1, X_2, ..., X_n), n \geq 1$ este o estimaţie (proces de estimare). Să remarcăm că a da o familie de funcţii $u_n : \mathbb{R}^n \to \mathbb{R}, n \geq 1$ revine la a da o funcţie $U : \bigcup\limits_{n\geq 1}\mathbb{R}^n \to \mathbb{R}$, funcţiile u_n fiind restricţiile funcţiei U la $\mathbb{R}^n$; în acest mod putem scrie mai compact procesul de estimare dar nu mai insistăm.

Fie acum θ un parametru "legat" de v.a X (un parametru de care depinde repartiţia v.a X etc) şi o estimaţie ca mai sus.

Estimaţie consistentă. Spunem că estimaţia este **consistentă** (în raport cu θ) dacă şirul $(u_n(X_1, X_2, ..., X_n))_n$ tinde în probabilitate către θ. În acest caz este justificat să vorbim de o estimare a parametrului θ. În discuţia de mai sus s-a arătat că şirul $\left(\overline{X_n}\right)_n$ tinde în probabilitate către $M[X]$ deci avem o estimare (consistentă) a mediei v.a X. Analog cu şirul $\left(S_n^2\right)_n$ pentru estimarea dispersiei.

Estimaţie nedeplasată. O estimaţie a parametrului θ este **nedeplasată** dacă $M[u_n(X_1, X_2, ..., X_n)] = \theta$ pentru orice $n \geq 1$. În acest sens asa cum s-a arătat mai sus $M\left[\overline{X_n}\right] = M[X], n \geq 1$. Pe de altă parte s-a arătat că $M\left[S_n^2\right] = \frac{n-1}{n}D[X]$ deci nu avem o estimaţie nedeplasată; acesta este motivul pentru care se consideră şi estimaţia $S_n'^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i - \overline{X_n}\right)^2$ cu $M\left[S_n'^2\right] = D[X], n \geq 1$. Intuitiv noţiunea de "nedeplasată" se leagă de lipsa erorilor sistematice (bias).

Verosimilitate. Să presupunem că repartiţia unei v.a X depinde de un parametru θ necunoscut şi care trebuie estimat. De exemplu, în cazul v.a discrete repartiţia de probabilitate este $p(x,\theta)$ iar în cazul continuu densitatea de repartiţie este $f(x,\theta)$. O selecţie $(X_1, X_2, ..., X_n)$ de volum n asupra v.a X va avea repartiţia comună $p(x_1, x_2, ..., x_n, \theta) = p(x_1,\theta)\ p(x_2,\theta)...$

$p(x_n, \theta)$ în cazul discret şi $f(x_1, x_2, ..., x_n, \theta) = f(x_1, \theta) f(x_2, \theta) ... f(x_n, \theta)$ în cazul continuu. Vom folosi notaţia $L(x_1, x_2, ..., x_n, \theta)$ pentru ambele cazuri şi vom numi această funcţie **funcţia de verosimilitate** a selecţiei. Dacă există o funcţie $(x_1, x_2, ..., x_n) \mapsto u(x_1, x_2, ..., x_n)$ astfel încât $L(x_1, x_2, ..., x_n, u(x_1, x_2, ..., x_n)) \geq L(x_1, x_2, ..., x_n, \theta)$ pentru orice θ (din intervalul de variatie al parametrului) atunci numim estimatorul (statistica) $u(X_1, X_2, ..., X_n)$ se numeşte **estimator de verosimilitate maximă**.

Exemplu. Fie X v.a discretă care ia valoarea 1 cu probabilitatea θ si valoarea 0 cu probabilitatea $1 - \theta$, $\theta \in (0, 1)$. În acest caz $p(x, \theta) = \theta^x (1 - \theta)^{1-x}$ cu $x = 0$ sau $x = 1$ şi $\theta \in (0, 1)$. Rezultă $p(x_1, x_2, ..., x_n, \theta) = \theta^{\sum x_i}(1 - \theta)^{n - \sum x_i}$. Pentru a maximiza această funcţie în raport cu θ observăm că putem maximiza logaritmul funcţiei, cu acelaşi rezultat. Avem $\ln p(x_1, x_2, ..., x_n, \theta) = (\sum x_i) \ln \theta + (n - \sum x_i) \ln(1 - \theta)$. Anulând derivata obţinem $u(x_1, x_2, ..., x_n) = \frac{x_1 + x_2 + ... + x_n}{n}$ care se verifică uşor a fi maximul căutat. Rezultă estimatorul de verosimilitate maximă $u(X_1, X_2, ..., X_n) = \frac{X_1 + X_2 + ... + X_n}{n}$.

14.3 MF.14.3. Estimare prin intervale de încredere.

Vom considera o familie de repartiţii depinzând de un parametru $\theta \in \Theta$ unde Θ este un interval (deschis, nevid) în $\mathbb{R}$, (de exemplu o funcţie $F : \mathbb{R} \times \Theta \to \mathbb{R}$, presupusă continuă (ca funcţie de două variabile reale) şi astfel încât, pentru fiecare $\theta \in \Theta$ funcţia $F(, \theta) : \mathbb{R} \to \mathbb{R}$ să fie o funcţie de repartiţie sau analog cu densităţi de repartiţie). Fie $\theta_0 \in \Theta$ o valoare adevărată, presupusă necunoscută, şi X o v.a cu repartiţia corespunzatoare. Vom fixa $\alpha \in (0, 1)$ numit **coeficient de încredere** şi vom arăta cum se poate asocia, unei selecţii $(X_1, X_2, ..., X_n)$ asupra v.a X, un interval aleator care conţine parametrul θ_0 cu probabilitatea α. Acest interval va fi numit **interval de încredere** $100\%\alpha$ pentru determinarea valorii θ_0. Sensul termenului "interval aleator" este: extremităţile $\underline{\theta}, \overline{\theta}$ ale intervalului sunt v.a şi deci pentru fiecare $\omega \in \Omega$ (câmpul de probabilitate pe care lucrăm) avem un interval $\left(\underline{\theta}(\omega), \overline{\theta}(\omega)\right) \subseteq \mathbb{R}$. Sensul afirmaţiei "interval aleator $\left(\underline{\theta}, \overline{\theta}\right)$ conţine parametrul θ_0 cu probabilitatea α" este $\left\{\omega : \theta_0 \in \left(\underline{\theta}(\omega), \overline{\theta}(\omega)\right)\right\}$ este un eveniment de probabilitate α.

Fie $g : \mathbb{R}^n \times \Theta \to \mathbb{R}$ o funcţie continuă (ca funcţie de $n + 1$ variabile) şi monotonă în raport cu $\theta \in \Theta$ (adică pentru fiecare $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ funcţia $g(x,) : \Theta \to \mathbb{R}$ este monotonă). În acest caz, pentru fiecare $x \in \mathbb{R}^n, a, b \in \mathbb{R}, a < b$ mulţimea $I(x, a, b) = \{\theta; a < g(x, \theta) < b\}$ este un interval (eventual vid). În adevăr dacă $\theta_1, \theta_2 \in I(x, a, b), \theta_1 < \theta_2$ şi presupunem $g(x,)$ crescătoare atunci pentru orice $\theta_1 < \theta < \theta_2$ avem $a < g(x, \theta_1) \leq g(x, \theta) \leq g(x, \theta_2) < b$ deci $\theta \in I(x, a, b)$.

Să remarcăm că dacă X^θ este o v.a cu repartiţia corespunzătoare valorii θ (pentru cazul $\theta = \theta_0$ s-a notat $X = X^{\theta_0}$) şi $\left(X_1^\theta, X_2^\theta, ..., X_n^\theta\right)$ o selecţie asupra v.a X^{θ} atunci $g\left(X_1^{\theta}, X_2^{\theta}, ..., X_n^{\theta}, \theta\right)$ este o v.a. Pentru $\omega \in \Omega, a, b \in \mathbb{R}, a < b$ vom obţine un interval $I\left(X_1^{\theta}(\omega), X_2^{\theta}(\omega), ..., X_n^{\theta}(\omega), a, b\right)$ ca mai sus; este modul în care iau naştere intervalele aleatoare pe care le vom nota $I\left(X_1^{\theta}, X_2^{\theta}, ..., X_n^{\theta}, a, b\right)$ si $I(X_1, X_2, ..., , a, b)$ pentru $\theta = \theta_0$.

Vom face două ipoteze suplimentare:

1). V.a $g\left(X_1^{\theta}, X_2^{\theta}, ..., X_n^{\theta}, \theta\right)$ sunt **la fel repartizate** (şi, evident, la fel cu $g(X_1, X_2, ..., X_n, \theta_0)$).

2). Funcţia de repartiţie a v.a $g(X_1, X_2, ..., X_n, \theta_0)$ este continuă.

De aici rezultă că putem determina $a(\alpha), b(\alpha)$ astfel încât $P\left(a(\alpha) < g\left(X_1^{\theta}, X_2^{\theta}, ..., X_n^{\theta}, \theta\right) < b(\alpha)\right) = \alpha$ pentru orice $\theta \in \Theta$ (reamintim că $\alpha \in (0, 1)$ este coeficientul de încredere fixat). Deducem imediat că $I(X_1, X_2, ..., , a(\alpha), b(\alpha))$ este un interval de încredere $100\%\alpha$ pentru determinarea valorii θ_0.

Practic, se consideră o realizare $x = (x_1, x_2, ..., x_n)$ a selecţiei $(X_1, X_2, ..., X_n)$ şi se determină $I(x, a(\alpha), b(\alpha))$. Acest interval **nu** este aleator deci θ_0 aparţine sau nu aparţine acestui interval (deci nu acesta este sensul probabilităţii α). Putem interpreta rezultatul obţinut considerând mai multe realizări posibile ale selecţiei deci mai multe intervale: frecvenţa relativă de apariţie a intervaleleor care conţin θ_0 este α.

Exemplu. Determinarea unui interval de încredere pentru parametrul m al legii normale $N(m, 1)$.

Avem deci o familie de densitãti de repartiţie $f(u, m) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(u-m)^2}{2}}, u, m \in \mathbb{R}$. Există o "adevărată valoare" m_0 a parametrului m pentru care trebuie construit un interval de încredere. Vom lua coeficientul de încredere $\alpha = 0,95$ deci vom construi un interval de încredere 95%. Folosind notaţiile de mai sus fie X v.a normal repartizată cu media m_0. Vom nota $\overline{X} = \frac{X_1+X_2+...+X_n}{n}$ pentru n fixat (pentru simplitate am renunţat la indicele n), unde $(X_1, X_2, ..., X_n)$ este o selectie de volum n asupra v.a X.

Se ştie (a se vedea capitolul "Legi de probabilitate") că v.a $\frac{\overline{X}-m_0}{\frac{1}{\sqrt{n}}}$ este $N(0, 1)$ repartizată (construcţia făcută pe v.a corespunzătoare oricărui parametru m duce la acelaşi rezultat, independent de m). Astfel se poate lua funcţia g dată de $g(x_1, x_2, ..., x_n, m) = \frac{\frac{x_1+x_2+...+x_n}{n} - m}{\frac{1}{\sqrt{n}}}$ care satisface condiţiile 1) si 2) de mai sus.

Pentru $a, b \in \mathbb{R}, a < b$ avem $P(a < g(X_1, X_2, ..., X_n, m_0) < b) =$
$= P\left(a < \frac{\overline{X}-m_0}{\frac{1}{\sqrt{n}}} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du = \Phi(b) - \Phi(a)$ (reamintim că $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{u^2}{2}} du$ este funcţia de repartiţie a unei v,a $N(0, 1)$ repartizate şi este tabelată). Se ştie că $\Phi(t) - \Phi(-t) = 2\Phi(t) - 1$ şi deci trebuie determinat t astfel încât $2\Phi(t) - 1 = 0,95$ deci $\Phi(t) = \frac{1,95}{2} = 0,975$. Se găseşte, din tabele, ca $t = 1,96$. Avem deci:

$P\left(-1,96 < \frac{\overline{X}-m_0}{\frac{1}{\sqrt{n}}} < 1,96\right) = 0,95$, de unde rezultă

$P\left(\overline{X}-\frac{1{,}96}{\sqrt{n}}<m_0<\overline{X}+\frac{1{,}96}{\sqrt{n}}\right)$ şi intervalul aleator căutat este $\left(\overline{X}-\frac{1{,}96}{\sqrt{n}},\overline{X}+\frac{1{,}96}{\sqrt{n}}\right)$. Să presupunem că pentru $n=100$, s-a obţinut o realizare $x=(x_1,x_2,...,x_n)$ a selecţiei $(X_1,X_2,...,X_n)$ asupra v.a X şi $\overline{x}=\frac{x_1+x_2+...+x_n}{n}=20$. Vom obţine intervalul $\left(\overline{x}-\frac{1{,}96}{\sqrt{n}},\overline{x}+\frac{1{,}96}{\sqrt{n}}\right)=$ $=\left(20-\frac{1{,}96}{\sqrt{n}},20+\frac{1{,}96}{\sqrt{n}}\right)=\left(20-\frac{1{,}96}{10},20+\frac{1{,}96}{10}\right)=(19,804,20,196)$. Am dat mai sus ce interpretare se poate da acestui rezutat concret. Remarcăm şi rolul volumului selecţiei în obţinerea unui interval de lungime cât mai mică.

14.4 MF.14.4. Verificarea ipotezelor statistice.

Fie M o mulţime nevidă şi $A\subset M$ o submulţime nevidă. Pentru $x\in M$ afirmaţia (propoziţia) "$x\in A$"este, prin definiţie, o **ipoteză.** Această ipoteză poate fi acceptată (ca adevărată) sau respinsă (ca falsă). În lumea fenomenelor aleatoare ipotezele sunt acceptate sau respinse cu anumite probabilităţi şi acceptând anumite riscuri.

În cazul în care M este o mulţime de repartiţii (funcţii de repartiţie sau densităţi de repartiţie) ipotezele se numesc **statistice**. O partiţie finită a mulţimii M dă naştere la un **sistem complet de ipoteze**. Reamintim că o partiţie (finită) a mulţimii M este o familie finită de submulţimi nevide $A_1,A_2,...,A_k$ astfel încât $\bigcup_{i=1}^{k}A_i=M$ şi $A_i\cap A_j=\emptyset, i\neq j$; ipotezele corepunzătoare sunt "$x\in A_i$" $i=1,3,...,k$. O ipoteză "$x\in A$" se numeşte **simplă** dacă mulţimea A are un singur element.

Fie $H=\{H_1,H_2,...,H_k\}$ un sistem complet de ipoteze. Vom numi **test** o aplicaţie surjectivă $\varphi:\mathbb{R}^n\to H$. **Regiunea de acceptare** a ipotezei H_i este prin definiţie $\varphi^{-1}(H_i)=\{x;\varphi(x)=H_i\}, i=1,2,..,k$. Ideea aplicării unui test în vederea acceptării unei ipoteze este: se obţine în urma unor experienţe un element $x\in\mathbb{R}^n$ şi se acceptă ipoteza $\varphi(x)$. În cazul ipotezelor statistice, pentru determinarea unui element $x\in\mathbb{R}^n$ se recurge la selecţii şi la realizările acestora.

În cele ce urmează vom considera cazul în care $k=2$ deci $H=\{H_1,H_2\}$.

Ne vom limita la discuţia ipotezelor statistice considerând **cazul parametric:** avem o mulţime de funcţii de repartiţie (densităţi de repartiţie) depinzând de un parametru $\theta\in\Theta$ şi ipotezele sunt determinate de Θ_1,Θ_2 cu $\Theta_1\cap\Theta_2=\emptyset,\Theta_1\cup\Theta_2=\Theta$ cu ipotezele corespunzătoare $H,\overline{H}$. Să notăm $S=\varphi^{-1}(H)$ pentru un test dat φ.

Pentru $\theta\in\Theta$ considerăm v.a X^θ corespunzătoare valorii θ a parametrului şi fie $\left(X_1^\theta,X_2^\theta,...,X_n^\theta\right)$ o selecţie asupra v.a X^θ. Definim:

$\alpha(\theta)=P\left(\left(X_1^\theta,X_2^\theta,...,X_n^\theta\right)\in CS\right)$ pentru $\theta\in\Theta_1$ şi
$\beta(\theta)=P\left(\left(X_1^\theta,X_2^\theta,...,X_n^\theta\right)\in S\right)$ pentru $\theta\in\Theta_2$.

Funcţia α reprezintă riscul ca ipoteza H să fie respinsă deşi este adevărată (eroare de primul tip) iar funcţia β reprezintă riscul ca ipoteza H să fie acceptată deşi este falsă (risc de al doilea tip). Cele două riscuri nu pot fi micşorate simultan cât de mult dorim. În general se alege testul după preferinţele acceptării unui risc sau a celuilalt (nu intram în discuţii de amănunt). Vom presupune că preferăm să nu respingem ipoteza H în mod eronat. Vom folosi, în acest caz, notaţia $H = H_0, \overline{H} = \overline{H_0}$ şi vom numi H_0 **ipoteza nulă** iar $\overline{H_0}$ **ipoteza alternativă**. Se spune că se testează ipoteza nulă contra ipotezei alternative alegând o valoare $\gamma \in (0,1)$ (**prag de semnificaţie**) astfel încât riscul de primul tip să nu depăsească γ. În fapt dacă prin situarea în cadrul ipotezei H_0 se obţin date D astfel încât $P(D) \leq \gamma$ se poate considera ca neconcordanţa dintre ipoteza H_0 şi date este semnificativă deci ipoteza se respinge.

Exemplu (test între două ipoteze simple). Fie X o v.a normal repartizată $N(m,1)$ în care parametrul m poate lua doar două valori m_0 si m_1. Vom nota H_0 ipoteza $m = m_0$ şi o vom testa contra ipotezei alternative $m = m_1$. Vom folosi notaţiile $(X_1, X_2, ..., X_n)$ pentru o selecţie de volum n asupra v.a X, $\left(X_1^0, X_2^0, ..., X_n^0\right)$ pentru o selecţie asupra v.a X^0, $N(m_0, 1)$ repartizată şi $\left(X_1^1, X_2^1, ..., X_n^1\right)$ pentru o selecţie asupra v.a X^1, $N(m_1, 1)$ repartizată. La fel ca într-un exemplu anterior fie $\overline{X} = \frac{X_1+X_2+...+X_n}{n}$. V.a $\frac{\overline{X}-m}{\frac{1}{\sqrt{n}}}$ este $N(0,1)$ repartizată. Să alegem pragul de semnificaţie $\gamma = 0,05$. Fie $a, b \in \mathbb{R}, a < b$ astfel încât $P\left(a < \frac{\overline{X}-m}{\frac{1}{\sqrt{n}}} < b\right) = 0,95$ sau $P\left(\frac{a}{\sqrt{n}} + m < \overline{X} < \frac{b}{\sqrt{n}} + m\right) = 0,95$. Dacă ipoteza nulă este adevărată atunci rezultă
$P\left(\frac{a}{\sqrt{n}} + m_0 < \overline{X^0} < \frac{b}{\sqrt{n}} + m_0\right) = 0,95$. Pentru determinarea unui test φ este suficient de specificat mulţimea $S \subset \mathbb{R}^n, S = \varphi^{-1}(H_0)$. Având în vedere dorinţa de a minimiza riscul de primul tip putem lua
$S = \left\{(x_1, x_2, ..., x_n); \frac{a}{\sqrt{n}} + m_0 < \frac{x_1+x_2+...+x_n}{n} < \frac{b}{\sqrt{n}} + m_0\right\}$. În adevăr calculând riscul de primul tip obţinem $\alpha(m_0) = P\left(\left(X_1^0, X_2^0, ..., X_n^0\right) \in CS\right) = 1-P\left(\left(X_1^0, X_2^0, ..., X_n^0\right) \in S\right) = 1-0,95 = 0,05$ care este pragul de semnificaţie. Astfel dacă o realizare $(x_1, x_2, ..., x_n)$ a unei selecţii asupra v.a X aparţine mulţimii S ipoteza nulă se poate accepta cu riscul asumat.

Testul χ^2. Vom începe prin a reaminti definiţia funcţiei Γ: $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt, x > 0$. Avem $\Gamma(x+1) = x\Gamma(x)$ de unde, pentru $n \in \mathbb{N}, n \geq 1$: $\Gamma(n+1) = n!$. De asemenea avem $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Pentru $k \in \mathbb{N}, k \geq 1$ o v.a este $\chi^2(k)$ repartizată dacă are densitatea de repartiţie f_k dată de $f_k(x) = 0, x \leq 0$ şi $f_k(x) = \frac{x^{\frac{k}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}, x > 0$. Este uşor de verificat că f_k este o densitate de repartiţie pentru fiecare k. În acest context k se zice numărul de grade de libertate al repartiţiei: $\chi^2(k)$ este repartiţia χ^2 cu k grade de libertate (nu vom intra în detalii privind noţiunea de ”grad de libertate”).

Dacă X este $\chi^2(k)$ repartizată atunci $M[X] = k$ şi $D[X] = 2k$ (exerciţiu).

Menţionăm că funcţiile de repartiţie a v.a $\chi^2(k)$ sunt tabelate (o astfel de tabelă se află la sfârşitul acestui capitol).

Este interesant următorul rezultat: dacă $(X_1, X_2, ..., X_n)$ este o selecţie asupra unei v.a $N(m, \sigma)$ repartizate atunci (folosind notaţiile de la secţiunea precedentă) v.a $Y = \frac{1}{\sigma^2}\left[\left(X_1 - \overline{X}\right)^2 + \left(X_2 - \overline{X}\right)^2 + ... + \left(X_{k+1} - \overline{X}\right)^2\right]$ este $\chi^2(k)$ repartizată (repartiţia v.a Y nu depinde de m şi σ).

Fie X o v.a discretă care ia valorile $1, 2, ..., k$ cu probabilităţile $p_1, p_2, ..., p_k$. Să consideram o selecţie $(X_1, X_2, ..., X_n)$ asupra v.a X şi pentru o realizare $(X_1(\omega), X_2(\omega), ..., X_n(\omega))$ fie $\overline{p_n}(i, \omega) = \frac{N_i(\omega)}{n}$ unde $N_i(\omega)$ este numărul de elemente i din realizare. Definim:

$\chi_n^2(p, \overline{p_n}, k)(\omega) = n\sum_{i=1}^{k}\frac{(p_i - \overline{p_n}(i, \omega))^2}{p_i}$ unde am notat pe scurt $p = (p_1, p_2, ..., p_k)$. Se observă că $\chi_n^2(p, \overline{p_n}, k)$ este o v.a. Ea măsoară distanţa dintre repartiţia ”empirică” $\overline{p_n}$ şi repartiţia teoretică p. Înlocuind $\overline{p_n}(i, \omega)$ cu valoarea sa obţinem $\chi_n^2(p, \overline{p_n}, k)(\omega) = \sum_{i=1}^{k}\frac{(np_i - N_i(\omega))^2}{np_i}$.

Rezultatul central este: **şirul** $\left(\chi_n^2(p, \overline{p_n}, k)\right)_n$ **converge în repartiţie către** $\chi^2(k-1)$. Reamintim că prin convergenţa în repartiţie a unui şir de v.a înţelegem convergenţa punctuală în orice punct de continuitate al limitei.

Vom descrie un **test** (**testul** χ^2) bazat pe acest rezultat. Ipoteza pe care o testăm este ”repartiţia unei v.a discrete care ia valorile $1, 2, ..., k$ este

$p = (p_1, p_2, ..., p_k)$ ”. Fie α un prag de semnificaţie. Determinăm (din tabele) o valoare a astfel încât $P\left(\chi^2(k-1) > a\right) = \alpha$. Asimilăm $\chi_n^2(p, \overline{p_n}, k)$ cu $\chi^2(k-1)$ si definim regiunea de acceptare S a ipotezei prin $S = \left\{(x_1, x_2, ..., x_n); \sum_{i=1}^{k}\frac{(np_i - N_i(\omega))^2}{np_i} \le a\right\} \subseteq E^n$ unde $E = \{1, 2, ..., k\}$.

Pentru cazul unei v.a oarecare X se partiţionează mulţimea de valori $X(\Omega) = \bigcup_{i=1}^{k} B_i$ cu mulţimi boreliene (în general intervale) şi se asociază o v.a discretă Y cu valori $1, 2, ..., k$ punând $Y(\omega) = i$ dacă şi numai dacă $X(\omega) \in B_i$. Se aplică apoi testul χ^2 pentru Y pentru a testa o anumită repartiţie a v.a X (repartiţia v.a X induce, evident, o repartiţie pentru Y).

Exemplu. Vom aplica testul χ^2 pentru o lege binomială $B(40, \mathfrak{p})$ şi ipoteza $\mathfrak{p}$ =0, 04. Datele sunt $n = 100$ şi $k = 5$, rezultând din partiţionarea mulţimii valorilor posibile $0, 1, 2, ..., 40$ în $0, 1, 2, 3, \{\ge 4\}$ (este un uşor abuz de notaţie dar credem că nu există pericol de confuzie). Iată tabelul de date:

valori	N_i	np_i	$(N_i - np_i)^2$	$\frac{(N_i-np_i)^2}{np_i}$
0	28	$19,5$	$72,25$	$3,71$
1	40	$32,6$	$54,76$	$1,68$
2	21	$26,4$	$29,16$	$1,10$
3	7	14	49	$3,50$
≥ 4	4	$7,5$	$12,25$	$1,63$
total	100	100		$\chi^2 = 11,62$

Desigur $p_i, i = 0, 1, 2, 3, \geq 4$ sunt probabilităţile valorilor respective în $B(40, 0, 04)$. Din tabele rezultă că pentru $k-1 = 5-1 = 4$, $P\left(\chi^2(4) > 11,6\right)$ este aproximativ $0,02$ deci foarte mică. Suntem astfel conduşi să respingem ipoteza $\mathfrak{p} = 0,04$.

Pentru acest capitol se recomandă lucrările:$[E.01]$, $[G.02]$, $[I.01]$, $[I.03]$, $[M.01]$ $[P.02]$, $[V.01]$.

Capitolul 15

MF.15. Autoevaluare

15.1 Capitol MF.01. Spaţii metrice

15.1.1 Exerciţii şi probleme rezolvate

1. Să se caracterizeze şirurile convergente şi şirurile Cauchy într-un spaţiu metric discret. Să se demonstreze că orice spaţiu metric discret este complet.

Soluţie

Fie (X, d) un spaţiu metric discret şi fie x_n un şir în X; fie $0 < \varepsilon < 1$. Dacă $x_n \to a$, atunci există $n_\varepsilon \in N$ astfel încât $d(x_n, a) < \varepsilon < 1, \forall n \geq n_\varepsilon$, deci $d(x_n, a) = 0, \forall n \geq n_\varepsilon$; rezultă că şirul x_n este constant (începând de la un rang) ... (4 puncte)

Un raţionament similar se aplică şi în cazul şirurilor Cauchy, deci x_n este şir Cauchy dacă şi numai dacă x_n este constant (de la un rang) ... (4 puncte)

Rezultă că x_n este şir convergent dacă şi numai dacă x_n este şir Cauchy, deci X este spaţiu metric complet ... (2 puncte).

2. Mulţimea lui Cantor

Notăm cu I_0 intervalul $[0, 1]$. Eliminăm din I_0 intervalul din mijloc, $(\frac{1}{3}, \frac{2}{3})$ şi notăm

$$I_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

mulţimea astfel obţinută. Continuăm procedeul: din fiecare din intervalele $[0, \frac{1}{3}]$, $[\frac{2}{3}, 1]$ eliminăm intervalul din mijloc şi notăm cu I_2 mulţimea rezultată:

$$I_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{3}{9}\right] \cup \left[\frac{6}{9}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

Continuând procedeul, se obţine un şir de mulţimi $I_0, I_1, I_2, ...$ cu proprietăţile:

i. $I_0 \supset I_1 \supset I_2 \supset ...$

ii. I_n este reuniunea a 2^n intervale, fiecare de lungime 3^{-n}.

Prin definiţie, mulţimea lui Cantor este intersecţia: $\mathcal{C} = \bigcap_{n \in N} I_n$.

Să se demonstreze următoarele proprietăţi:
a. $\mathcal{C}$ este mulţime compactă.
b. Mulţimea $\mathcal{C}$ nu conţine intervale.
c. Mulţimea lui Cantor este perfectă (nu conţine puncte izolate); în particular, rezultă că $\mathcal{C}$ nu este mulţime numărabilă.
Soluţie
a. Mulţimea $\mathcal{C}$ este mărginită (inclusă în $[0,1]$) şi închisă (intersecţie de mulţimi închise) ... (2 puncte)
b. Din construcţie, rezultă:

$$\mathcal{C} \cap \left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m} \right) = \emptyset, \, \forall k, m \in N.$$

Dar, orice interval (α, β) conţine un interval de forma $\left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m}\right)$ dacă m este ales cu condiţia $3^{-m} < \frac{\beta-\alpha}{6}$. Rezultă că mulţimea $\mathcal{C}$ nu conţine intervale ... (4 puncte)
c. Fie $a \in \mathcal{C}$ şi fie S un interval arbitrar care-l conţine pe a; pentru orice $n \in N$, fie J_n acel interval al lui I_n care-l conţine pe a. Alegem n_0 suficient de mare astfel încât $J_{n_0} \subseteq S$; dacă notăm cu x_n acel capăt al intervalului J_n diferit de a, rezultă $x_n \in \mathcal{C} \cap S, \; x_n \neq a, \forall n \geq n_0$ ... (4 puncte)

3. Să se calculeze cu o eroare mai mică decât 10^{-3} soluţia reală a ecuaţiei $x^3 + 12x - 1 = 0$.
Soluţie
Ecuaţia are o singură soluţie reală $\xi \in (0,1)$ ... (1 punct)
Ecuaţia este echivalentă cu $x = f(x)$, unde $f(x) = \left(x^2 + 12\right)^{-1}$ ... (3 puncte)
Funcţia f este contracţie pe $[0,1]$ (spaţiu metric complet) cu factorul de contracţie $k = \frac{2}{169}$ ... (4 puncte)
Şirul aproximaţiilor succesive: $x_0 = 0, x_1 = f(0) = \frac{1}{12}, ..., x_n = f(x_{n-1})$; aproximarea cerută este x_2 ... (2 puncte).

15.1.2 Exerciţii şi probleme propuse

4. Fie $a, b \in R, a < b$.
a. Să se demonstreze că

$$d_1(f,g) = \int_a^b |f(x) - g(x)| dx$$

este distanţă pe mulţimea funcţiilor continue $\mathcal{C}[a,b]$.
b. Să se demonstreze că orice şir $f_n \in \mathcal{C}([a,b])$ convergent în raport cu distanţa d_∞ este convergent şi în raport cu distanţa d_1, dar reciproca este

falsă.

Răspunsuri

a. Se verifică direct definiţia (se folosesc proprietăţile modulului şi ale integralei).

b. Fie $f_n, f \in \mathcal{C}([a,b])$ astfel încât $d_\infty(f_n, f) \to 0$. Atunci:

$$d_1(f_n, f) = \int_a^b |f_n(x) - f(x)| dx \le (b-a) \cdot d_\infty(f_n, f) \to 0.$$

Pentru a arăta că reciproca este falsă, fie şirul $f_n(x) = \dfrac{1}{1+nx}$.
Atunci $f_n \to 0$ în raport cu distanţa d_1, dar nu converge în raport cu d_∞.

5. Să se decidă dacă următoarele funcţii sunt contracţii pe mulţimile indicate:
a. $f(x) = \sin x$, $x \in R$.
b. $f(x) = \ln x$, $x \in [e, \infty)$.
c. $f(x) = \operatorname{arctg} x$, $x \in R$.
d. $f(x) = \dfrac{1-x^2}{5(1+x^2)}$, $x \in R$.
e. $f(x) = \dfrac{2x}{1+x^2}$, $x \in R$.

Răspunsuri

a. Funcţia $f(x) = \sin x$ nu este contracţie pe R.
b. Funcţia $f(x) = \ln x$ este contracţie pe $[e, \infty)$;
c. Funcţia $f(x) = \operatorname{arctg} x$ nu este contracţie pe R.
d. Funcţia nu este contracţie pe R.
e. Funcţia este contracţie pe R.

6. In spaţiul metric (R^2, d_2) considerăm submulţimile:
$A = \{(x,y) \in R^2 \mid 0 < x^2 + y^2 \le 1\}$,
$B = \{(x,y) \in R^2 \mid 1 \le x^2 + y^2 \le 2\}$,
$D = \{(x,y) \in R^2 \mid x^2 + y^2 < 1, x > 0, y > 0\}$,
$E = \{(x,y) \in R^2 \mid ax + by + c = 0, a, b, c \in R \text{ constante fixate}\}$,
$F = \left\{\left(\frac{1}{n}, 1\right) \in R^2 \mid n \in N\right\}$, $G = F \cup \{(0,1)\}$.

Să se precizeze dacă mulţimile sunt deschise, închise, mărginite, conexe sau compacte.

Răspunsuri

A este mulţime mărginită şi conexă, dar nu este deschisă şi nici închisă. B este compactă şi conexă. D este deschisă, conexă şi mărginită. E este închisă, conexă şi nemărginită. F este mărginită, dar nu este deschisă, nici închisă şi nici conexă. G este mulţime compactă.

15.2 Capitol MF.02. Spaţii normate

15.2.1 Exerciţii şi probleme rezolvate

1. Fie $\mathcal{C}^1[a,b]$ spaţiul vectorial al funcţiilor de clasă C^1 definite pe intervalul compact $[a,b]$.
a. Să considerăm pe $\mathcal{C}^1[a,b]$ norma supremum: $\| f \|_\infty = \sup_{x\in[a,b]} |f(x)|$. Să se demonstreze că aplicaţia de derivare

$$D : \mathcal{C}^1[a,b] \mapsto \mathcal{C}[a,b],\ D(f) = f',$$

este operator liniar dar nu este şi continuu. Pe spaţiul funcţiilor continue, $\mathcal{C}[a,b]$, este considerată, ca de obicei, norma supremum.
b. Să considerăm acum pe spaţiul $\mathcal{C}^1[a,b]$ norma:

$$\| f \| = \| f \|_\infty + \| f' \|_\infty .$$

Să se demostreze că aplicaţia de derivare D este în acest caz operator continuu.
Soluţie
a. Liniaritatea este evidentă. Fie şirul $f_n(x) = \frac{1}{n}\sin nx$; atunci $\| f_n \|_\infty = \frac{1}{n}$ şi deci $f_n \to 0$ în spaţiul normat $\left(\mathcal{C}^1[a,b], \| \ \|_\infty\right)$ ... (2 puncte)
Şirul $D(f_n) = f_n'$ nu converge (la 0) în $\mathcal{C}[a,b]$ ... (3 puncte)
b. Fie $f \in \mathcal{C}^1[a,b]$; din inegalitatea:

$$\| D(f) \|_\infty = \| f' \|_\infty \le \| f \|_\infty + \| f' \|_\infty = \| f \|$$

rezultă că D este operator continuu ... (5 puncte)

2. Operatorul de înmulţire cu variabila independentă
Pe spaţiul Banach complex $(\mathcal{C}[a,b], \| \ \|_\infty)$ considerăm operatorul (de îmulţire cu variabila independentă):

$$(Mf)(x) = xf(x),\ \forall f \in \mathcal{C}[a,b],\ \forall x \in [a,b].$$

a. Să se demonstreze că M este liniar şi continuu şi $\| M \| = |b|$.
b. Să se demonstreze că spectrul lui M este :

$$\sigma(M) = [a,b].$$

c. Să se demonstreze că mulţimea valorilor proprii este vidă: $\sigma_p(M) = \emptyset$.
Soluţie
a. Pentru orice $f, g \in \mathcal{C}[a,b]$ şi $\alpha, \beta \in C$, avem:

$$(M(\alpha f + \beta g))(x) = x(\alpha f + \beta g)(x) =$$

$$= \alpha x f(x) + \beta x g(x) = (\alpha Mf + \beta Mg)(x),\ \forall x \in [a,b],$$

deci M este liniar ... (1 punct)
Continuitatea:

$$\| Mf \|_\infty = \sup_{x\in[a,b]} |xf(x)| \leq |b| \ \| f \|_\infty, \ \forall\, f \in \mathcal{C}[a,b], \forall\, x \in [a,b],$$

deci M este continuu şi în plus $\| M \| \leq |b|$ ... (2 puncte)
Notând cu $\mathbf{1}$ funcţia constantă 1, atunci $\| \mathbf{1} \|_\infty = 1$ şi deci:

$$\| M \| \geq \| M\mathbf{1} \|_\infty = \sup_{x\in[a,b]} |x| = |b|,$$

deci $\| M \| = |b|$ ... (1 punct)
b. Demonstrăm egalitatea $\sigma(M) = [a,b]$ prin dublă incluziune.
Prin definiţie, $\lambda \in \sigma(M)$ dacă şi numai dacă operatorul $\lambda I - M$ nu este inversabil (ceea ce este echivalent cu a fi bijectiv, conform teoremei lui Banach; a se vedea secţiunea teoretică a acestui capitol). Fie $\lambda_0 \in [a,b]$; din egalitatea:

$$((\lambda_0 I - M)(f))(x) = (\lambda_0 - x)f(x), \ \forall f \in \mathcal{C}[a,b], \ \forall x \in [a,b]$$

rezultă că operatorul $\lambda_0 I - M$ nu este surjectiv deoarece imaginea sa este:

$$\mathcal{I}\mathrm{m}(\lambda_0 I - M) = \{f \in \mathcal{C}[a,b] \mid f(\lambda_0) = 0\} \neq \mathcal{C}[a,b],$$

deci $[a,b] \subseteq \sigma(M)$ ... (1 punct)
În locul incluziunii inverse $\sigma(M) \subseteq [a,b]$ demonstrăm incluziunea echivalentă: $C \setminus [a,b] \subseteq C \setminus \sigma(M)$. Fie $\lambda_0 \notin [a,b]$; atunci funcţia

$$\varphi : [a,b] \mapsto C, \ \varphi_0(x) = \frac{1}{\lambda_0 - x}$$

este corect definită şi continuă, deci $\| \varphi_0 \|_\infty < \infty$ ... (1 punct)
Considerăm operatorul:

$$S : \mathcal{C}[a,b] \mapsto \mathcal{C}[a,b], \ (Sf)(x) = \varphi_0(x)f(x) = \frac{1}{\lambda_0 - x} f(x), \ \forall\, x \in [a,b].$$

Se demonstrează fără dificultate că S este operator liniar. Continuitatea rezultă din inegalitatea:

$$\| Sf \|_\infty = \sup_{x\in[a,b]} |\varphi_0(x)f(x)| \leq \| \varphi_0 \|_\infty \| f \|_\infty, \ \forall\, f \in \mathcal{C}[a,b].$$

În concluzie, operatorul $S \in \mathcal{L}(\mathcal{C}[a,b])$ ... (1 punct)
În plus, au loc egalităţile:

$$(\lambda_0 I - M)Sf = (S(\lambda_0 I - M))f = \varphi_0 \frac{1}{\varphi_0} f = f, \ \forall\, f \in \mathcal{C}[a,b],$$

deci operatorul $\lambda_0 I - M$ este inversabil şi $(\lambda_0 I - M)^{-1} = S$, ceea ce demonstrează că $\lambda_0 \notin \sigma(M)$ ... (1 punct)

c. Pentru a demonstra că M nu are valori proprii, vom arăta că pentru orice $\lambda \in [a, b]$, operatorul $\lambda I - M$ este injectiv. Fie $\lambda_0 \in [a, b]$ şi fie $f \in \mathcal{C}[a, b]$ astfel încât $(\lambda_0 I - M)f = 0$; rezultă:

$$(\lambda_0 - x)f(x) = 0, \forall\, x \in [a, b].$$

De aici rezultă

$$f(x) = 0,\ \forall x \in [a, b] \setminus \{\lambda_0\}.$$

Funcţia f fiind continuă, rezultă şi $f(\lambda_0) = 0$, deci $f(x) = 0$, $\forall\, x \in [a, b]$; in concluzie, $\lambda_0 I - M$ este injectiv ... (2 puncte)

3. Fie $(X, \|\ \ \|)$ un spaţiu Banach şi fie $T \in \mathcal{L}(X)$ un operator astfel încât $\| T \| < 1$. Să se demonstreze că $I - T$ este operator inversabil şi

$$(I - T)^{-1} = \sum_{n \geq 0} T^n.$$

În plus, are loc inegalitatea:

$$\| (I - T)^{-1} \| \leq \frac{1}{1 - \| T \|}.$$

Soluţie

Spaţiul $\mathcal{L}(X)$ este complet, deci orice serie (de operatori) absolut convergentă este şi convergentă. Fie seria $\sum\limits_{n \geq 0} T^n$ ... (2 puncte)

Seria converge absolut:

$$\sum_{n \geq 0} \| T^n \| \leq \sum_{n \geq 0} \| T \|^n = \frac{1}{1 - \| T \|},$$

deci converge în spaţiul $\mathcal{L}(X)$ ... (1 punct)

Fie $S \in \mathcal{L}(X)$ suma acestei serii şi fie $S_n = \sum\limits_{k=0}^{n} T^k$ şirul sumelor parţiale asociat; atunci:

$$(I - T)S_n = (I - T)(I + T + t^2 + ... + T^n) = I - T^{n+1} ...(2 \text{ puncte})$$

Dar şirul T^{n+1} converge la O în spaţiul $\mathcal{L}(X)$ ... (2 puncte):

$$\| T^{n+1} \| \leq \| T \|^{n+1} \to 0,\ \text{ când }\ n \to \infty,$$

deci $(I - T)S = I$. Analog se arată şi egalitatea $S(I - T) = I$, deci întradevăr $S = (I - T)^{-1}$ ... (1 punct)

În plus, dintr-un calcul făcut mai sus, rezultă:

$$\| (I - T)^{-1} \| = \| \sum_{n \geq 0} T^n \| \leq \frac{1}{1 - \| T \|},$$

ceea ce încheie demonstraţia ... (2 puncte)

15.2.2 Exerciţii şi probleme propuse

4. Fie $x_0 \in [a,b]$; pe spaţiul funcţiilor continue, $(\mathcal{C}[a,b], \| \ \|_\infty)$ considerăm aplicaţia (numită evaluarea în punctul x_0):

$$F_{x_0} : \mathcal{C}[a,b] \mapsto \mathbf{R},\ F_{x_0}(f) = f(x_0).$$

Să se demonstreze că F_{x_0} este funcţională liniară şi continuă şi să se calculeze norma $\| F_{x_0} \|$.

5. Pe spaţiul Banach $(\mathcal{C}[a,b], \| \ \|_\infty)$ al funcţiilor continue (reale) considerăm aplicaţia

$$J : \mathcal{C}[a,b] \mapsto \mathbf{R},\ J(f) = \int_a^b f(t)dt.$$

Să se demonstreze că J este funcţională liniară şi continuă şi apoi să se calculeze norma sa.

6. Operatorul de convoluţie
Fie două şiruri $x, y : \mathbf{Z} \mapsto \mathbf{C}$, cu proprietatea că pentru orice $n \in \mathbf{Z}$ seria $\sum_{k\in Z} x(n-k)y(k)$ este convergentă. În acest caz se poate defini şirul

$$x \star y : \mathbf{Z} \mapsto C,\ (x \star y)(n) = \sum_{k\in Z} x(n-k)y(k),$$

numit convoluţia (sau produsul de convoluţie) şirurilor x şi y.
a. Să se demonstreze că pentru orice $x, y \in \ell^1(\mathbf{Z})$, există convoluţia $x \star y$.
b. Să se demonstreze că pentru orice $x, y \in \ell^1(\mathbf{Z})$, convoluţia $x \star y \in \ell^1(\mathbf{Z})$, şi în plus $\| x \star y \|_1 \leq \| x \|_1 \| y \|_1$
c. Produsul de convoluţie este comutativ şi asociativ.
d. Pentru orice $m \in \mathbf{Z}$, fie şirul $\sigma_m(n) = \delta_m^n$, unde, δ_m^n este simbolul lui Kronecker. Să se demonstreze egalitatea:

$$(\sigma_m \star x)(n) = x(n-m),\ \forall\, x \in \ell^1(Z),\ \forall m, n \in \mathbf{Z}.$$

În particular, σ_0 este element neutru pentru convoluţie.
e. Fie $\theta \in \ell^1(\mathbf{Z})$ un şir fixat şi fie operatorul (de convoluţie):

$$C_\theta : \ell^1(\mathbf{Z}) \mapsto \ell^1(\mathbf{Z}),\ C_\theta x = \theta \star x.$$

Să se demonstreze că operatorul C_θ este liniar şi continuu.

15.3 Capitol MF.03. Operatori pe $\mathbf{C}^n$

15.3.1 Exerciţii şi probleme rezolvate

1. Fie $T : \mathbf{C^2} \mapsto \mathbf{C^2}$, $T(x, y) = (x + y, iy)$. Să se arate că T este diagonalizabil în sens algebric, dar nu este diagonalizabil în sens geometric.

Soluţie Matricea (în baza canonică) a lui T este: $M = \begin{pmatrix} 1 & 1 \\ 0 & i \end{pmatrix}$. Valorile proprii ale lui T sunt distincte: $\lambda_1 = 1$, $\lambda_2 = i$, deci T este diagonalizabil în sens algebric ... (4 puncte)

Adjunctul lui T are matricea $\overline{M}^t = \begin{pmatrix} 1 & 0 \\ 1 & -i \end{pmatrix}$ ... (1 punct)

Matricele M şi $\overline{M}^t$ nu comută (2 puncte), deci T nu este operator normal (2 puncte), deci T nu este diagonalizabil în sens geometric (1 punct).

2. Pentru orice $t \in \mathbf{R}$, fie $T : \mathbf{C^2} \mapsto \mathbf{C^2}$ operatorul a cărui matrice în baza canonică este:

$$M = \begin{pmatrix} 1 & e^{it} \\ e^{-it} & 1 \end{pmatrix}$$

(a) Să se demonstreze că T este operator diagonalizabil în sens geometric.
(b) Să se determine o bază în care matricea lui T are formă diagonală.
(c) Să se arate că există $\sqrt{T}$ şi să se determine matricea în baza canonică a lui $\sqrt{T}$.

Soluţie (a) Deoarece $M = \overline{M}^t$, rezultă că T este operator autoadjunct, deci este operator diagonalizabil în sens geometric ... (1 punct).
(b) Valorile proprii ale lui T sunt 0 şi 2 ... (1 punct)
Baza cerută este o bază formată din vectori proprii, de exemplu (2 puncte):

$$\mathcal{B} = \{(-\frac{\sqrt{2}}{2}e^{it}, \frac{\sqrt{2}}{2}), (\frac{\sqrt{2}}{2}e^{it}, \frac{\sqrt{2}}{2})\}$$

(c) Matricea operatorului T în baza $\mathcal{B}$ este (3 puncte):

$$D = U^{-1}MU = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix},$$

unde (1 punct):

$$U = \begin{pmatrix} -\frac{\sqrt{2}}{2}e^{it} & \frac{\sqrt{2}}{2}e^{it} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

Existenţa rădăcinii pătrate pozitive $\sqrt{T}$ rezultă din faptul că T este operator pozitiv (deoarece T este autoadjunct şi are valori proprii mai mari sau egale decât 0) ... (1 punct)
Matricea lui $\sqrt{T}$ în baza canonică este (3 puncte):

$$\sqrt{M} = U\sqrt{D}\,U^{-1},$$

unde:

$$\sqrt{D} = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2} \end{pmatrix}.$$

3. Fie $T : \mathbf{C^2} \mapsto \mathbf{C^2}$ operatorul a cărui matrice în baza canonică este:

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Să se determine $S : \mathbf{C^2} \mapsto \mathbf{C^2}$ astfel încât $e^S = T$.
Soluţie T este operator autoadjunct şi mulţimea valorilor proprii este $\{1, 3\}$ ... (2 puncte)
O bază formată din vectori proprii ai lui T este $\mathcal{B} = \{(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}), (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})\}$ ... (2 puncte)
Matricea lui T în baza $\mathcal{B}$ este (2 puncte):

$$D = U^{-1}MU = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix},$$

unde (1 punct):

$$U = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

Rezultă că operatorul căutat, $S = \ln T$, are matricea (în baza canonică) (3 puncte):

$$\ln M = U \ln D\, U^{-1},$$

unde:

$$\ln D = \begin{pmatrix} 0 & 0 \\ 0 & \ln 3 \end{pmatrix}.$$

15.3.2 Exerciţii şi probleme propuse

4. Fie $T : \mathbf{C^3} \mapsto \mathbf{C^3}$, $T(x, y, z) = (iy+z, -ix+z, x+y)$. Să se demostreze că T este operator diagonalizabil în sens geometric şi să se calculeze $\sqrt[3]{T}$.

5. Fie $r > 0$, $t \in \mathbf{R}$ şi fie $T : \mathbf{C^2} \mapsto \mathbf{C^2}$, $T(x, y) = (x + re^{it}y, re^{-it}x + y)$. Să se determine r astfel încât T să fie operator pozitiv şi în acest caz să se calculeze $\sqrt{T}$.

6. Fie $T : \mathbf{C^3} \mapsto \mathbf{C^3}$, $T(x, y, z) = (x+iy, -ix+z, y+z)$. Să se determine $S : \mathbf{C^3} \mapsto \mathbf{C^3}$ astfel încâ $e^S = T^2$.

15.4 Capitol MF.04. Spaţii Hilbert

15.4.1 Exerciţii şi probleme rezolvate

1. Fie $\mathbf{C}[X]$ spaţiul vectorial al funcţiilor polinomiale cu coeficienţi complecşi.
(a) Să se demonstreze că aplicaţia:

$$<,>: \mathbf{C}[X] \times \mathbf{C}[X] \mapsto \mathbf{C},\ < f, g >= \int_0^1 f(x)\overline{g}(x)dx$$

este produs scalar pe $\mathbf{C}[X]$.
(b) Fie $f(x) = x$ şi $g(x) = 3x - 2$. Să se demonstreze că $f \perp g$.
(c) Este $(\mathbf{C}[X], <,>)$ spaţiu Hilbert?
Soluţie (a) şi **(b)** sunt verificări directe ale definiţiilor ... (2 puncte)
(c) $(\mathbf{C}[X], <,>)$ este spaţiu prehilbertian, (1 punct) dar nu este complet. Fie $\|\ \|_2$ norma indusă de produsul scalar considerat; fie, de exemplu, şirul $f_n(x) = \sum_{k=0}^{n} \frac{1}{k!}x^k$ (2 puncte). Atunci f_n converge în norma $\|\ \|_\infty$ la $f(x) = e^x$ (1 punct) deci f_n converge la f şi în norma indusă de produsul scalar (1 punct) din cauza inegalităţii: $\| g \|_2 \leq \| g \|_\infty, \forall g \in \mathbf{C}[X]$ (1 punct). Dar $f \notin \mathbf{C}[X]$, deci $\mathbf{C}[X]$ nu este spaţiu Hilbert (2 puncte).

2. Fie, în spaţiul Banach $(\ell^p(\mathbf{N}), \|\ \|_p)$, elementele $x = (1, 1, 0, 0, ...)$ şi $y = (0, 0, 1, 1, 0, 0, ...)$.
(a) Să se calculeze $\| x \|_p$ şi $\| y \|_p$.
(b) Să se demonstreze că $(\ell^p(\mathbf{N}), \|\ \|_p)$ este spaţiu Hilbert dacă şi numai dacă $p = 2$.
Soluţie (a) $\| x \|_p = \| y \|_p = \sqrt[p]{2}$ (2 puncte).
(b) $(\ell^p(\mathbf{N}), \|\ \|_p)$ este spaţiu Hilbert dacă şi numai dacă norma $\|\ \|_p$ verifică legea paralelogramului (3 puncte):

$$\| x + y \|_p + \| x - y \|_p = 2(\| x \|_p + \| y \|_p)$$

Rezultă $2^{p+2} = 16$ (3 puncte) deci $p = 2$ ... (2 puncte).

3. Pentru fiecare $k, l \in \mathbf{Z}$, fie funcţia

$$u_{k,l} : \mathbf{R}^2 \mapsto \mathbf{C},\ u_{k,l}(x, y) = \frac{1}{2\pi} e^{i(kx+ly).}$$

(a) Să se demonstreze că $(u_{k,l})_{k,l \in Z}$ este sistem ortonormat în $L^2([0, 2\pi]^2)$ cu produsul scalar uzual: $< f, g >= \int\int_{[0,2\pi]^2} f(x, y)\overline{g(x, y)}dxdy$.
(b) Să se calculeze coeficienţii Fourier în raport cu sistemul $(u_{k,l})_{k,l \in Z}$ şi să se scrie seria Fourier pentru orice funcţie $f : \mathbf{R}^2 \mapsto \mathbf{C}$, cu proprietatea

$f(x+2\pi, y+2\pi) = f(x,y)$.
Soluţie (a) Pentru orice $k, l, p, q \in \mathbf{Z}$, avem (4 puncte):

$$\int\int_{[0,2\pi]^2} u_{k,l}(x,y)\overline{u_{p,q}(x,y)}dxdy =$$

$$= \frac{1}{4\pi^2}\int_0^{2\pi} e^{i(k-p)x}dx \int_0^{2\pi} e^{(l-q)y}dy = \begin{cases} 1 & \text{dacă} \quad p = k \text{ şi } q = l \\ 0 & \text{dacă} \quad p \neq k \text{ şi } q \neq l \end{cases}$$

(b) Coeficienţii Fourier (3 puncte):

$$\widehat{f}(k,l) = \int\int_{[0,2\pi]} f(t,s)\overline{u_{k,l}(t,s)}dtds =$$

$$= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi} f(t,s)e^{-i(kt+sl)}dtds.$$

Seria Fourier (3 puncte):

$$\sum_{k\in Z}\sum_{l\in Z}\widehat{f}(k,l)u_{k,l}(x,y) = \frac{1}{2\pi}\sum_{k\in Z}\sum_{l\in Z}\widehat{f}(k,l)e^{i(kx+ly)}.$$

15.4.2 Exerciţii şi probleme propuse

4. Să se dezvolte în serie Fourier în raport cu sistemul ortonormat din exerciţiul anterior funcţia $f(x,y) = xy$ definită pe pătratul $[-2\pi, 2\pi]^2$ şi prelungită prin periodicitate la întreg planul.
Răspuns Coeficienţii Fourier:

$$\widehat{f}(k,0) = \widehat{f}(0,l) = 0, \forall k,l \in \mathbf{Z} \text{ şi } \widehat{f}(k,l) = \frac{(-1)^k}{\pi^2 kl} \text{ în rest.}$$

Seria Fourier:

$$\frac{4}{\pi^2}\sum_{k\geq 1}\sum_{l\geq 1}\frac{(-1)^{k+l}}{kl}\sin(k\pi x)\,\sin(l\pi y)$$

5. Polinoame Legendre
Pe mulţimea funcţiilor polinoamiale cu coeficienţi reali restricţionate la intervalul $[-1,1]$, considerăm produsul scalar $< f, g >= \int_{-1}^{1} f(x)g(x)dx$. Şirul $\{1, x, x^2, x^3, ...\}$ este liniar independent; folosind procedeul Gram-Schmidt, acest şir se poate ortogonaliza şi se obţine şirul polinoamelor lui Legendre (coeficientul dominant este 1): $\{P_0, P_1, P_2, ...\}$.
(a) Să se calculeze primele 4 polinoame Legendre.

(b) Să se demonstreze formula: $P_n(x) = \frac{1}{A_{2n}^n}\left[(x^2-1)^n\right]^{(n)}$.
Răspuns (a) Prin calcul direct rezultă:

$$P_0(x) = 1, P_1(x) = x, P_2(x) = x^2 - \frac{1}{3}, P_3(x) = x^3 - \frac{3}{5}x.$$

(b) Polinoamele P_n sunt determinate (până la o constantă multiplicativă) prin relaţiile de ortogonalitate: $\int_{-1}^{1} P_m(x)P_n(x)dx = 0, \forall m \neq n$.
Fie $Q_n(x) = \left[(x^2-1)^n\right]^{(n)}, \forall n \in \mathbf{N}$; se arată că $\int_{-1}^{1} Q_m(x)Q_n(x)dx = 0, \forall m \neq n$, iar coeficientul dominant al lui Q_n este A_{2n}^n.

6. Polinoame Cebâşev
Analog cu exerciţiul anterior, polinoamele Cebâşev (notate T_n) se obţin prin ortogonalizarea sistemului $\{1, x, x^2, x^3, ...\}$ în raport cu produsul scalar

$$< f, g >= \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}}dx$$

(a) Să se calculeze primele 4 polinoame Cebâşev.
(b) Să se demonstreze că $T_n(x) = \frac{1}{2^{n-1}}\cos(n \arccos x), \forall n \in \mathbf{N}^\star$.

15.5 Capitol MF.05. Măsură şi integrală

15.5.1 Exerciţii şi probleme rezolvate

1. Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură şi fie $f : X \mapsto [0, \infty)$ o funcţie măsurabilă. Să se demonstreze că dacă $\int_X f d\mu = 0$, atunci $f = 0$ (a.p.t.).
Soluţie Pentru orice $n \in N$, fie

$$A_n = \{x \in X \ ; \ f(x) > \frac{1}{n}\}.$$

Mulţimile A_n sunt măsurabile pentru că $A_n = f^{-1}\left((\frac{1}{n}, \infty)\right)$. (1 punct). Mai mult, avem (2 puncte):

$$\bigcup_{n\in N} A_n = \{x \in X \ ; \ f(x) \neq 0\}.$$

Vom demonstra că $\mu(A_n) = 0, \forall n \in N$. Integrând pe A_n inegalitatea:

$$\frac{1}{n} < f(x), \forall x \in A_n,$$

obţinem (folosim şi $f(x) \geq 0$) (3 puncte):

$$\frac{1}{n}\,\mu(A_n) \leq \int_{A_n} f d\mu \leq \int_X f d\mu = 0,$$

deci $\mu(A_n) = 0, \forall n \in N$ (2 puncte). Avem deci (2 puncte):

$$\mu(\{x \in X \ ; \ f(x) \neq 0\}) = \mu(\bigcup_{n \in N} A_n) \leq \sum_{n \in N} \mu(A_n) = 0.$$

2. Fie $(a_{ij})_{i,j \in N}$ un şir dublu indexat astfel încât $a_{ij} \geq 0, \forall i, j \in N$. Să se demonstreze că:

$$\sum_{i \in N} \sum_{j \in N} a_{ij} = \sum_{j \in N} \sum_{i \in N} a_{ij}.$$

Facem menţiunea că membrii egalităţii pot fi şi ∞.
Soluţie Vom aplica teorema de convergenţă monotonă (1 punct). Considerăm spaţiul cu măsură (N, μ_c) şi şirul de funcţii (4 puncte):

$$f_i : N \mapsto [0, \infty), \ f_i(j) = a_{ij}.$$

Atunci, conform teoremei de convergenţă monotonă, avem (3 puncte):

$$\int_N \sum_{i \in N} f_i d\mu_c = \sum_{i \in N} \int_N f_i d\mu_c,$$

adică (2 puncte):

$$\sum_{j \in N} \sum_{i \in N} a_{ij} = \sum_{i \in N} \sum_{j \in N} a_{ij}.$$

3. Inegalitatea mediilor
Fie $n \in N$. Să se demonstreze că pentru orice numere reale nenegative $x_1, x_2, ..., x_n$, are loc inegalitatea mediilor:

$$(x_1 \cdot x_2 \cdot ...x_n)^{\frac{1}{n}} \leq \frac{x_1 + x_2 + ... + x_n}{n}.$$

Soluţie Evident, putem presupune că $x_1, x_2, ..., x_n$ sunt strict pozitive.
Fie $X = \{p_1, p_2, ..., p_n\}$ o mulţime cu n elemente şi fie $P : \mathcal{P}(X) \mapsto [0, \infty)$, măsura de probabilitate (1 punct), deci (2 puncte):

$$P(\{p_j\}) = \frac{1}{n}, \forall j = 1, 2, ..., n.$$

Fie $f : X \mapsto R, f(p_j) = \ln x_j$ şi fie $\phi(x) = e^x$. Aplicând inegalitatea lui Jensen, obţinem (5 puncte):

$$e^{\frac{1}{n} \sum_{j=1}^{n} f(p_j)} \leq \frac{1}{n} \sum_{j=1}^{n} e^{f(p_j)},$$

adică $(x_1 \cdot x_2 \cdot ...x_n)^{\frac{1}{n}} \leq \dfrac{x_1 + x_2 + ... + x_n}{n}$ (2 puncte).

15.5.2 Exerciţii şi probleme propuse

4. Fie $(X, \mathcal{A}, \mu)$ un spaţiu cu măsură astfel încât $\mu(X) = 1$, fie $a, b \in \overline{R}$, $a < b$. şi fie $f : X \mapsto R$, $g : X \mapsto [0, \infty)$, două funcţii integrabile; să se demonstreze inegalităţile:
a. $e^{\int_X f\, d\mu} \leq \int_X e^f\, d\mu.$
b. $e^{\int_X \ln g\, d\mu} \leq \int_X g\, d\mu.$
Indicaţie Se aplică inegalitatea lui Jensen funcţiei convexe $\phi(t) = e^t$.

5. Să se demonstreze formula:

$$\frac{\pi - x}{2} = \sum_{n \geq 1} \frac{\sin nx}{n},\ \forall x \in (0, 2\pi).$$

Soluţie
Fie $f(x) = \frac{\pi - x}{2}$, $x \in [0, 2\pi)$, prelungită prin periodicitate la R; calculăm coeficienţii Fourier:

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} \frac{\pi - x}{2}\, dx = \frac{1}{2\pi} \left(\pi x - \frac{x^2}{2} \right) \Big|_0^{2\pi} = 0.$$

$$a_n = \frac{1}{\pi} \int_0^{2\pi} \frac{\pi - x}{2} \cos nx\, dx =$$

$$= \frac{(\pi - x) \sin nx}{2n\pi} \Big|_0^{2\pi} - \frac{1}{2n\pi} \int_0^{2\pi} \sin nx\, dx = 0, \forall\, n \geq 1.$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} \frac{\pi - x}{2} \sin nx\, dx =$$

$$= \frac{-(\pi - x) \cos nx}{2n\pi} \Big|_0^{2\pi} - \frac{1}{2n\pi} \int_0^{2n\pi} \cos nx\, dx = \frac{1}{n}, \forall\, n \geq 1.$$

Aplicând teorema lui Dirichlet, rezultă:

$$\frac{\pi - x}{2} = \sum_{n \geq 1} \frac{\sin nx}{n},\ \forall x \in (0, 2\pi).$$

În punctele $x = 0$ şi $x = 2\pi$ funcţia f nu este continuă; în aceste puncte seria trigonometrică asociată ei are suma 0.

6. Fie $a \in R^\star$; să se dezvolte funcţia $f : [0, \pi) \mapsto R$, $f(x) = e^{ax}$:
a. în serie de cosinusuri;
b. în serie de sinusuri.

Indicaţie
Se calculează coeficienţii şi rezultă dezvoltările:

$$e^{ax} = \frac{e^{a\pi} - 1}{a\pi} + \sum_{n\geq 1} \frac{2a((-1)^n e^{a\pi} - 1)}{\pi(a^2 + n^2)} \cos nx \; \forall\, x \in [0, \pi).$$

$$e^{ax} = \sum_{n\geq 1} \frac{2n(1 - (-1)^n e^{a\pi})}{\pi(a^2 + n^2)} \sin nx, \; \forall\, x \in (0, \pi).$$

15.6 Capitol MF.06. Operatori pe spaţii Hilbert

15.6.1 Exerciţii şi probleme rezolvate

1. Fie $\alpha : \ell(\mathbf{N}) \mapsto \mathbf{C}, \alpha(n) = \frac{(-1)^n n^2}{n^2+1}$.
(a) Să se calculeze norma operatorului diagonal $D_\alpha : \ell(\mathbf{N}) \mapsto \ell(\mathbf{N})$.
(b) Să se determine valorile proprii şi spectrul operatorului D_α.
(c) Este D_α inversabil?
Soluţie (a) Şirul α este mărginit (1 punct), deci D_α este bine definit (1 punct).
Norma este $\| D_\alpha \| = \sup\{|\alpha(n)| \; ; \; n \in \mathbf{N}\} = 1$ (3 puncte).
(b) Spectrul punctual este $\sigma_p(D_\alpha) = \{\alpha(n) \; ; \; n \in \mathbf{N}\}$ (2 puncte), iar spectrul $\sigma(D_\alpha) = \sigma_p(D_\alpha) \cup \{-1, 1\}$ (2 puncte).
(c) D_α nu este inversabil deoarece $0 \in \sigma(D_\alpha)$ (1 punct).

2. Fie $\phi : \mathbf{R} \mapsto \mathbf{R}, \phi(t) = \frac{2t}{t^2+1}$.
(a) Să se calculeze norma operatorului de multiplicare $M_\phi : L^2(\mathbf{R}) \mapsto L^2(\mathbf{R})$.
(b) Să se determine spectrul lui M_ϕ.
Soluţie (a) Funcţia ϕ este mărginită, deci operatorul M_ϕ este bine definit (2 puncte).
Norma este: $\| M_\phi \| = \sup\{|\phi(t)| \; ; \; t \in \mathbf{R}\} = 1$ (4 puncte).
Spectrul este $\sigma(M_\phi) = [-1, 1]$ (4 puncte).

3. Fie $T : \ell(\mathbf{Z}) \mapsto \ell(\mathbf{Z})$, $(Tx)(n) = x(n+1) - 4x(n-1)$, $\forall n \in \mathbf{Z}$.
(a) Să se demonstreze că T este operator de convoluţie.
(b) Să se demonstreze că T este operator inversabil.
(c) Să se calculeze inversul lui T.
Soluţie (a) Fie $\alpha : \mathbf{Z} \mapsto \mathbf{R}$, $\alpha(-1) = 1$, $\alpha(1) = -4$ şi $\alpha(n) = 0$ în rest. Atunci T este operatorul de convoluţie cu şirul α, deci $T = C_\alpha$ (2 puncte).
(b) Calculăm transformata Fourier inversă a lui α (3 puncte).

$$(\mathcal{F}^{-1}(e^{it}) = e^{it} - 4e^{it}, \; \forall t \in [0, 2\pi).$$

Funcţia $\mathcal{F}^{-1}$ nu se anulează pe cercul unitate, deci spectrul lui T nu conţine pe 0, deci T este inversabil (2 puncte).

(c) Inversul lui T este operatorul de convoluţie cu şirul $\mathcal{F}(\frac{1}{\mathcal{F}^{-1}\alpha})$ (1 punct). Calculăm $\mathcal{F}(\frac{1}{\mathcal{F}^{-1}\alpha})$ (3 puncte):

$$(\mathcal{F}(\frac{1}{\mathcal{F}^{-1}\alpha}))(n) = \frac{1}{2\pi i}\int_{\mathcal{S}^1} \frac{z^{-n}}{z^2-4}dz = \begin{cases} 0, \text{ dacă } n \leq 0 \\ 0, \text{ dacă } n > 0 \text{ şi } n \text{ impar} \\ \frac{1}{2^n}, \text{ dacă } n > 0 \text{ şi } n \text{ par} \end{cases}$$

15.6.2 Exerciţii şi probleme propuse

4. Fie W operatorul de translaţie bilateral; să se demonstreze că W^2 este operator unitar şi să i se calculeze norma şi, spectrul şi spectrul punctual.

5. Fie $T : \ell(\mathbf{Z}) \mapsto \ell(\mathbf{Z})$, $(Tx)(n) = x(n+1) - x(n-1)$, $\forall n \in \mathbf{Z}$. Să se demonstreze că T este operator de convoluţie şi să i se calculeze spectrul şi norma. Este T operator inversabil?

6. Fie $\psi : \mathbf{R} \mapsto \mathbf{C}$, $\psi(x) = \frac{x+i}{ix+1}$. Să se demonstreze că operatorul de multiplicare $M_\psi : L^2(\mathbf{R}) \mapsto L^2(\mathbf{R})$ este operator unitar.

15.7 Capitol MF.07. Aplicaţii în teoria sistemelor

15.7.1 Exerciţii şi probleme rezolvate

1. Fie $T : \ell^2(\mathbf{Z}) \mapsto \ell^2(\mathbf{Z})$, $(Tx)(n) = x(n-1) + 3x(n)$.
(a) Să se demonstreze că T este sistem invariant în timp.
(b) Să se demonstreze că T este sistem cauzal.
(c) Să se demonstreze că T este sistem inversabil şi are inversul cauzal.
Soluţie (a) Fie şirul $\alpha : \mathbf{Z} \mapsto \mathbf{Z}$, $\alpha(1) = 1, \alpha(0) = 3$ şi $\alpha(n) = 0$ în rest. Atunci T este operatorul de convoluţie cu α, deci T este invariant în timp (2 puncte).
(b) T este sistem cauzal deoarece $\alpha(n) = 0$, $\forall n < 0$ (2 puncte).
(c) Calculăm (1 punct):

$$(\mathcal{F}^{-1}\alpha)(e^{it}) = \frac{1}{3+e^{it}},$$

deci T este inversabil (funcţia de transfer nu se anulează pe cerc, 2 puncte). Calculăm: (2 puncte):

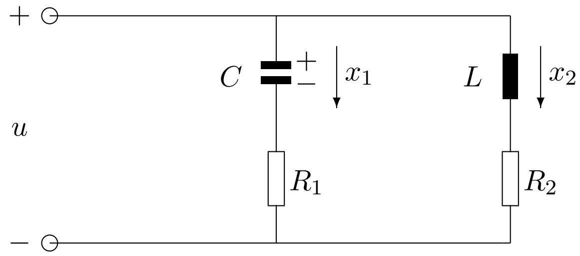
$$(\mathcal{F}^{-1}\alpha)(e^{it}) = \frac{1}{2\pi i}\int_{\mathcal{S}^1} \frac{z^{-n}}{3+z}dz = \begin{cases} 0, \text{ dacă } n \leq 0 \\ \frac{(-1)^{n-1}}{3^{n-2}}, \text{ dacă } n > 0 \end{cases},$$

deci T^{-1} este cauzal (3 puncte).

2. Fie W operatorul de translaţie bilateral şi fie $T = W^2 + W^{-2}$. Să se arate că T este sistem invariant în timp. Este T sistem cauzal? Dar anticauzal?

Soluţie T este invariant în timp deoarece este sistem de convoluţie cu şirul $\alpha : \mathbf{Z} \mapsto \mathbf{Z}$, $\alpha(2) = 1$, $\alpha(-2) = 1$ şi $\alpha(n) = 0$ în rest (4 puncte); T nu este cauzal deoarece $\alpha(n) \neq 0$, $\forall n < 0$ (3 puncte) şi nici anticauzal (3 puncte).

3. Fie R_1, R_2, C, L nenule şi să considerăm reţeaua electrică din figura alăturată.



Notăm cu u tensiunea la borne şi cu i curentul. Vom considera sistemul (intrare-ieşire) $u \to i$. Mai întâi, vom reprezenta acest sistem ca un sistem dinamic şi apoi vom studia, folosind criteriile lui Kalman, observabilitatea şi controlabilitatea reprezentării obţinute. Pentru aceasta, fie x_1 tensiunea pe condensatorul C şi x_2 curentul prin inductorul L.

Ecuaţiile (diferenţiale) ale reţelei sunt (2 puncte):

$$x_1' = -\frac{1}{R_1 C} x_1 + \frac{1}{R_1 C} u,$$

$$x_2' = -\frac{R_2}{L} x_2 + \frac{1}{L} u.$$

Curentul i este dat de formula (1 punct):

$$i = -\frac{1}{R_1} x_1 + x_2 + \frac{1}{R_1} u.$$

Fie matricele:

$$A = \begin{pmatrix} -\frac{1}{R_1 C} & 0 \\ 0 & -\frac{R_2}{L} \end{pmatrix}, \; B = \begin{pmatrix} \frac{1}{R_1 C} \\ \frac{1}{L} \end{pmatrix}, \; C = \begin{pmatrix} -\frac{1}{R_1} & 1 \end{pmatrix}, \; D = \frac{1}{R_1}$$

Notând $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, sistemul $u \to i$ se scrie:

$$x' = Ax + Bu \; , \; i = Cx + Du.$$

Pentru a decide dacă descompunerea canonică (A, B, C) este observabilă şi (sau) controlabilă, calculăm matricele de observabilitate şi controlabilitate;

obţinem (4 puncte):

$$\mathcal{Q} = \begin{pmatrix} \frac{1}{R_1C} & -\frac{1}{R_1^2C^2} \\ & \\ \frac{1}{L} & -\frac{R_2}{L^2} \end{pmatrix} \text{ şi } \mathcal{R} = \begin{pmatrix} -\frac{1}{R_1} & \frac{!}{R_1^2C} \\ & \\ 1 & -\frac{R_2}{L} \end{pmatrix}.$$

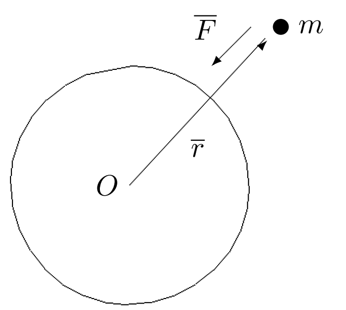
Determinanţii acestor matrice sunt (1 punct):

$$\det \mathcal{Q} = \frac{L - R_1R_2C}{R_1^2C^2L^2} \text{ şi } \det \mathcal{R} = \frac{R_1R_2C - L}{R_1^2CL},$$

şi deci în acest caz condiţia de observabilitate coincide cu cea de controlabilitate şi este: $L \neq R_1R_2C$ (2 puncte).

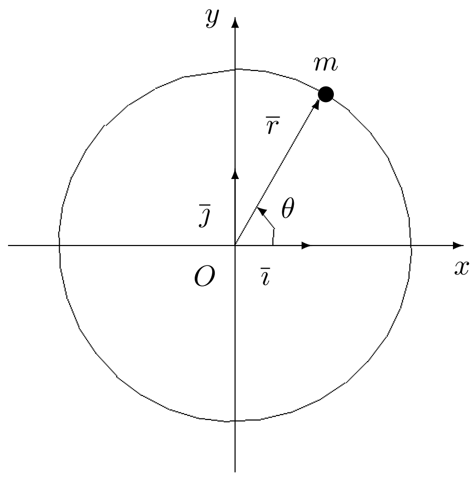
15.7.2 Exerciţii şi probleme propuse

4. Problema satelitului
Considerăm m un punct material (satelitul) care se mişcă sub acţiunea unei forţe centrale $\overline{F}$ (forţa de atracţie a Pămîntului).



Dacă $\overline{r}(t)$ este vectorul de poziţie al satelitului faţă de centrul O al Pământului la momentul t, atunci ecuaţia mişcării este $m\overline{r}''(t) = \overline{F}$. Din legea atracţiei universale, rezultă că există o constantă $k > 0$ astfel încât: $\overline{F} = -k \parallel \overline{r} \parallel^{-3} \overline{r}$. Demonstrăm acum că mişcarea este plană; pentru aceasta, este suficient să demonstrăm că produsul vectorial $\overline{r} \times \overline{r}'$ este egal cu un vector constant $\overline{v}$, (deci vectorul de poziţie $\overline{r}$ aparţine planului perpendicular pe vectorul $\overline{v}$). Într-adevăr, avem (1 punct):

$$\frac{d}{dt}\left(\overline{r} \times \overline{r}'\right) = \overline{r}' \times \overline{r}' + \overline{r} \times \overline{r}'' = \overline{r} \times \frac{1}{m}\overline{F} = -\frac{k}{m \parallel \overline{r} \parallel^3}(\overline{r} \times \overline{r}) = \overline{0}.$$



Considerăm, în planul xOy al mişcării, o bază ortononormală, $\{\overline{\imath}, \overline{\jmath}\}$; fie $r = r(t) = \| \overline{r} \|$ şi $\theta = \theta(t)$ coordonatele polare ale satelitului. Prin calcul direct, obţinem (1 punct):

$$\overline{r} = r\cos\theta\overline{\imath} + r\sin\theta\overline{\jmath},$$

$$\overline{r}' = \left(r'\cos\theta - r\theta'\sin\theta\right)\overline{\imath} + \left(r'\sin\theta + r\theta'\cos\theta\right)\overline{\jmath},$$

$$\overline{r}'' = \left(r''\cos\theta - 2r'\theta'\sin\theta - r\left(\theta'\right)^2\cos\theta - r\theta''\sin\theta\right)\overline{\imath} +$$

$$+\left(r''\sin\theta + 2r'\theta'\cos\theta - r\left(\theta''\right)^2\sin\theta + r\theta''\cos\theta\right)\overline{\jmath}.$$

Înlocuind în expresia lui $\overline{F}$, obţinem (2 puncte):

$$\overline{F} = -\frac{k}{r^3}\left(r\cos\theta\overline{\imath} + r\sin\theta\overline{\jmath}\right).$$

Înlocuind acum în ecuaţia de mişcare $m\overline{r}'' = \overline{F}$ pe $\overline{r}''$ şi $\overline{F}$ cu expresiile obţinute mai sus, obţinem relaţiile (scalare):

$$\begin{aligned} r''(t) &= r(t)\left(\theta'\right)^2(t) - \frac{k}{\left(r(t)\right)^2} \\ \theta''(t) &= -\frac{2r'(t)}{r(t)}\theta(t) \end{aligned}$$

Comenzile cu ajutorul cărora este controlată poziţia satelitului pe orbită sunt u_1 =comanda (acceleraţia) radială şi u_2 = comanda (acceleraţia) tangenţială. Rezultă deci că ecuaţiile de mişcare sunt:

$$\begin{aligned} r'' &= r\left(\theta'\right)^2 - \frac{k}{r^2} + u_1 \\ \theta'' &= -\frac{2r'\theta'}{r} + u_2 \end{aligned}$$

O soluţie particulară a acestui sistem este (1 punct)

$$r(t) = c\,,\ \theta(t) = \omega t,$$

unde, c şi ω sunt două constante ce verifică relaţia $c^3\omega^2 = k$. Se observă (din prima egalitate) că traiectoria este circulară, iar viteza unghiulară a satelitului, θ', este constantă.
Pentru a studia controlabilitatea şi observabilitatea sistemului

$$(u_1, u_2) \to (r, \theta),$$

introducem vectorul de stare (la momentul t), $x = (x_1, x_2, x_3, x_4) \in R^4$, definit prin egalităţile:

$$x_1(t) = r(t) - c\,,\ x_2(t) = r'(t)\,,\ x_3(t) = c(\theta(t) - \omega t)\,,\ x_4(t) = c(\theta'(t) - \omega).$$

Deducem acum ecuaţiile de mişcare (în spaţiul stărilor):

$$\begin{aligned}
x_1' &= r' = x_2 \\
x_2' &= r'' = r\left(\theta'\right)^2 + \frac{k}{r^2} + u_1 = (x_1 + c)\left(\frac{x_4}{c} + \omega\right)^2 + \frac{k}{(x_1 + c)^2} + u_1 \\
x_3' &= x_4 \\
x_4' &= c\,\theta'' = c\left(-\frac{2r'\theta'}{r} + u_2\right) = -c\frac{2}{x_1 + c}\,x_2\left(\frac{x_4}{c} + \omega\right) + cu_2
\end{aligned}$$

Sistemul diferenţial obţinut (în necunoscutele x_1, x_2, x_3, x_4) este neliniar; pentru a-l putea studia, liniarizăm ecuaţiile (dezvoltând în serie Taylor în jurul originii membrul drept al fiecărei ecuaţii şi păstrând termenii de gradul întâi):

$$\begin{aligned}
x_1' &= x_2 \\
x_2' &= 3\omega^2 x_1 + 2\omega x_4 + u_1 \\
x_3' &= x_4 \\
x_4' &= -2\omega x_2 + u_2
\end{aligned}$$

Fie matricele

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 3\,\omega^2 & 0 & 0 & 2\,\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\,\omega & 0 & 0 \end{pmatrix},\ B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix},\ C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Fie $y_1(t) = r(t) - t = x_1(t)$ şi $y_2(t) = c(\theta(t) - \omega\,t) = x_3(t)$.
Atunci sistemul $u = (u_1, u_2) \to (y_1, y_2 = y$ se scrie sub forma sistemului dinamic (1 punct):

$$x' = Ax + Bu\ ,\ y = Cx.$$

Pentru a studia controlabilitatea şi observabilitatea sistemului, calculăm matricele de controlabilitate şi observabilitate (2 puncte):

$$\mathcal{R} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 2\,\omega & -\omega^2 & 0 \\ 1 & 0 & 0 & 2\,\omega & -\omega^2 & 0 & 0 & -2\,\omega^3 \\ 0 & 0 & 0 & 1 & -2\,\omega & 0 & 0 & -4\,\omega^2 \\ 0 & 1 & -2\,\omega & 0 & 0 & -4\,\omega^2 & 2\,\omega^3 & 0 \end{pmatrix}$$

$$\mathcal{Q} = \begin{pmatrix} 1 & 0 & 0 & 0 & 3\,\omega^2 & 0 & 0 & -6\,\omega^3 \\ 0 & 0 & 1 & 0 & 0 & -2\,\omega & -\omega^2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2\,\omega & 0 & 0 & -4\,\omega^2 \end{pmatrix}$$

Rangurile matricelor $\mathcal{R}$ şi $\mathcal{Q}$ sunt amândouă 4 şi deci sistemul este şi controlabil şi observabil (în ipoteza că amândouă comenzile u_1 şi u_2 sunt accesibile şi, respectiv, se cunosc amândouă ieşirile y_1 şi y_2).
Să presupunem acum că una din cele două comenzi lipseşte.
Dacă $u_1 = 0$, (adică lipseşte comanda radială), atunci:

$$B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \; \mathcal{R} = \begin{pmatrix} 0 & 0 & 2\,\omega & 0 \\ 0 & 2\,\omega & 0 & -2\,\omega^3 \\ 0 & 1 & 0 & -4\,\omega^2 \\ 1 & 0 & -4\,\omega^2 & 0 \end{pmatrix}$$

Se observă că şi în acest caz rangul matricei $\mathcal{R}$ este 4, deci mişcarea satelitului poate fi controlată numai prin comandă tangenţială.
Dacă $u_2 = 0$, (deci lipseşte comanda tangenţială), atunci:

$$B = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \; \mathcal{R} = \begin{pmatrix} 0 & 1 & 0 & -\omega^2 \\ 1 & 0 & -\omega^2 & 0 \\ 0 & 0 & -2\,\omega & 0 \\ 0 & -2\,\omega & 0 & 2\,\omega^3 \end{pmatrix}$$

În acest caz, rangul matricei $\mathcal{R}$ este 3, deci satelitul nu poate fi controlat numai prin comandă radială.
Lăsăm ca exerciţiu următoarele afirmaţii:
Dacă se cunoaşte numai y_1, atunci satelitul nu este observabil (radial) (1 punct).
Dacă se cunoaşte numai y_2, atunci satelitul este observabil (tangenţial) (1 punct).

5.Sistemul dinamic liniar

Fie spaţiul Hilbert $L^2(\mathbf{R})$. şi fie matricele A, B, C ca în exemplul 14 (ii) din Cap. 7. 03. În plus, vom presupune că matricea A este stabilă, adică valorile proprii ale lui A sunt toate în semiplanul stâng: $\{z = a + ib \in \mathbf{C}\,;\, a < 0\}$. Pentru orice $u \in L^2(\mathbf{R})$, considerăm sistemul diferenţial:

$$x'(t) = Ax(t) + Bu(t),$$

cu condiţia iniţială $\lim_{t\to-\infty} x(t) = 0$. Atunci soluţia (unică) a problemei Cauchy de mai sus este:

$$x(t) = \int_{-\infty}^{t} e^{A(t-\tau)} Bu(\tau) d\tau.$$

Sistemul dinamic liniar pe R este, prin definiţie, operatorul

$$\mathcal{D} : L^2(\mathbf{R}) \to L^2(\mathbf{R}),\ \mathcal{D}u = Cx.$$

Să se demonstreze că descompunerea canonică (în sensul definiţiei 13, Cap.03) este $\{(\mathbf{R^n}, \lambda_t, \theta_t)\ ;\ t \in \mathbf{R}\}$, unde:

$$\lambda_t : L^2 \to \mathbf{R^n},\ \lambda_t u = \int_{-\infty}^{\infty} e^{A(t-\tau)} Bu(\tau) d\tau,$$

$$\theta_t : \mathbf{R^n} \to L^2(\mathbf{R}),\ (\theta_t \xi)(\tau) = \begin{cases} Ce^{A(t-\tau)}\xi, & \text{dacă} \quad \tau \geq t \\ 0, & \text{dacă} \quad \tau < t \end{cases}$$

.

Soluţie Analog cu exemplul 14(ii) Cap.7.03.

6. Sistemul discret (sistem "diferenţă")

Analogul discret al exemplului anterior este definit după cum urmează (a se vedea si exemplul 14 (ii), Cap.7.03). Fie matricele A, B, C ca mai sus şi fie $u \in \ell^2(\mathbf{N})$. Fie $x : \mathbf{N} \to \mathbf{R^n}$ soluţia recurenţei:

$$x(k+1) = Ax(k) + Bu(k),\ x(0) = 0.$$

Sistemul diferenţă este operatorul liniar şi continuu

$$\ell^2(\mathbf{N}) \ni u \to y \in \ell^2(\mathbf{N}),\ \text{unde}\,, y(k) = Cx(k), \forall k \in \mathbf{N}.$$

Este uşor de demonstrat că

$$y(k) = Cx(k) = C \sum_{j=0}^{k-1} A^j Bu(k-1-j),\ \forall k \geq 1.$$

Spaţiul stărilor este $X_k = \mathbf{R^n}$, $\forall k \in \mathbf{N}$ şi:

$$\lambda_k : \ell^2(\mathbf{N}) \to \mathbf{R^n},\ \lambda_k u = x(k),$$

$$\theta_k : \mathbf{R^n} \to \ell^2(\mathbf{N}),\ (\theta_k \xi)(j) = \begin{cases} CA^{j-k}\xi & j \geq k \\ 0 & j < k-1 \end{cases}$$

Să se demonstreze că $(X_k, \lambda_k, \theta_k)_{k \in N}$ este o descompunere în sensul definiţiei 13, Cap. 7.03.

Soluţie Analog cu exemplul 14(ii) Cap.7.03.

15.8 Capitol MF.08. Câmp de probabilitate.

15.8.1 Exerciţii şi probleme rezolvate.

1). Fie $\Omega = \{1, 2, 3\}$ şi $\Delta = \mathcal{P}(\Omega)$. Să se arate că există o probabilitate $P : \Delta \to [0, 1]$ cu $x = P\{1, 2\}$, $y = P\{2, 3\}$, $z = P\{1, 3\}$ dacă şi numai dacă $x, y, z \in [0, 1]$ şi $x + y + z = 2$.

Soluţie şi barem.

Din oficiu .... 1 punct.

Necesitatea condiţiei.

Fie P o probabilitate. Evident $x, y, z \in [0, 1]$. ... 1 punct.

Avem $x = P\{1\} + P\{2\}, y = P\{2\} + P\{3\}, z = P\{1\} + P\{3\}$ ..... 2 puncte.

Deci $x + y + z = 2(P\{1\} + P\{2\} + P\{3\}) = 2$... 1 punct.

Suficienţa condiţiei.

Fie x, y, z satisfăcând condiţiile $x, y, z \in [0, 1]$ şi $x+y+z = 2$ ... 1 punct.

Definim $P\{1\} = 1 - y, P\{2\} = 1 - z, P\{3\} = 1 - x$... 1 punct.

Rezultă $P\{1\} + P\{2\} + P\{3\} = 3 - (x + y + z) = 3 - 2 = 1$... 1 punct.

Deci avem o repartiţie de probabilitate pe Ω care generează o probabilitate P. ... 1 punct.

$P\{1, 2\} = (1 - y) + (1 - z) = 2 - (y + z) = 2 - (2 - x) = x$ etc... .... 1punct.

2). a). Fie (Ω, Δ, P) un câmp de probabilitate, $A, B, C \in \Delta$ astfel încât $P(A \cap B) \neq 0$. Să se arate că $P(A \cap B \cap C) = P(A) P(B \mid A) P(C \mid A \cap B)$.

b). O urnă conţine a bile albe şi b bile negre. Care este probabilitatea ca în trei extracţii succesive fără a pune bila înapoi să se obţină doar bile albe.

Soluţie.

a). Condiţiile de existenţă pentru probabilităţile condiţionate sunt îndeplinite. Avem :

$P(A \cap B \cap C) = P((A \cap B) \cap C) = P(A \cap B) P(C \mid A \cap B) =$
$= P(A) P(B \mid A) P(C \mid A \cap B)$.

b). Fie A, B, C evenimentele extragerii de bile albe la prima, la a doua şi la a treia extragere. Avem :

$P(A) = \frac{a}{a+b}, P(B \mid A) = \frac{a-1}{a+b-1}, P(C \mid A \cap B) = \frac{a-2}{a+b-2}$.

Folosind a) deducem $P(A \cap B \cap C) = \frac{a}{a+b} \frac{a-1}{a+b-1} \frac{a-2}{a+b-2}$.

3). Se joacă următorul joc: la o aruncare a monedei un jucător câştigă o unitate de capital dacă prevede corect rezultatul aruncării şi pierde o unitate de capital în celălalt caz. Jucătorul porneşte cu $x > 0$ unităţi de capital iar jocul se opreşte dacă jucătorul ajunge fie la 0 unităţi (se ruinează) fie la $a > x$ unităţi. Presupunând moneda "corectă" să se determine $p(x)$ probabilitatea de ruinare a jucătorului.

Soluţie.

$p(x+1)$ poate fi interpretată ca probabiltatea de ruinare condiţionată de câştigarea primului joc iar $p(x-1)$ ca probabiltatea de ruinare condiţionată de pierderea primului joc. Probabilitatea de apariţie a unei fete a monedei fiind $\frac{1}{2}$ formula probabilităţii totale (ţinând cont că evenimentele apariţiei celor două fete formează un sistem complet) ne dă relaţia $p(x) = \frac{1}{2}(p(x+1) + p(x-1))$ pentru $x = 1, 2, ..., a-1$ impreună cu condiţiile $p(0) = 1, p(a) = 0$. Avem de rezolvat recurenţa $p(x+1) = 2p(x) - p(x-1)$ în condiţiile precizate. Notând $q(x) = p(x) - p(x-1)$ obţinem relaţia $q(x+1) = q(x)$ şi deci $q(x) = q(1)$ şi deci $p(x) = p(x-1) + p(1) - p(0)$. Adunând membru cu membru găsim relaţia

$p(x) = x\ p(1) - x + 1$ şi folosind condiţia $p(a) = 0$ rezultă, în definitiv $p(x) = 1 - \frac{x}{a}$.

15.8.2 Exerciţii şi probleme propuse.

1). Fie p un număr prim. Considerăm câmpul de probabilitate (Ω, Δ, P) unde $\Omega = \{1, 2, ..., p\}, \Delta = \mathcal{P}(\Omega)$ iar P este egal probabilitatea. Să se arate că dacă evenimentele A, B sunt independente atunci cel puţin unul dintre ele este Ω sau $\emptyset$.

Răspuns. Se aplică definiţia independenţei.

2). La un examen sunt n bilete dintre care m sunt considerate "uşoare". Studenţii vin pe rând sa ia câte un bilet (care nu se mai introduce în teanc). Dintre primii doi studenti, care are "şansa" mai mare să ia un bilet uşor?

Răspuns. Sansele (exprimate ca probabilităti) sunt egale $(\frac{m}{n})$.

3). Doi arcaşi trag asupra unei ţinte câte o săgeată. Probabilitatea ca primul să lovească ţinta este $0,8$ iar pentru cel de al doilea $0,4$. După efectuarea tragerii, în ţintă se găseşte o singură săgetă. Care este probabilitatea ca aceasta să fie a primului arcaş ?

Răspuns. $\frac{6}{7}$.

4) Fie (Ω, Δ, P) un câmp de probabilitate şi $A_1, A_2, ..., A_n \in \Delta$. Să se arate că: $P(\cup A_i) = \sum\limits_{i=1}^{n} P(A_i) - \sum\limits_{i<j} P(A_i \cap A_j) + \sum\limits_{i<j<k} P(A_i \cap A_j \cap A_k) - ... + (-1)^{n-1} P(A_1 \cap A_2 \cap ... \cap A_n)$

Răspuns. Se poate utiliza inducţia.

5). O societate compusă din n perechi soţ -soţie dansează. Formarea perechilor la dans este egal probabilă. Care este limita, când $n \to \infty$ a probabilităţii ca nici o pereche care dansează să nu fie soţ -soţie?

Răspuns. $\frac{1}{e}$ (indicatie: este util a se folosi exerciţiul precedent).

15.9 Capitol MF.09 Variabile aleatoare.

15.9.1 Exerciţii şi probleme rezolvate

1) Timpul de aşteptare (în minute) la o staţie de autobuz este o v.a X cu funcţia de repartiţie F dată de: $F(x) = 0, x \leq 0$, $F(x) = \frac{x}{2}, x \in (0,1]$, $F(x) = \frac{1}{2}, x \in (1,2]$, $F(x) = \frac{x}{4}, x \in (2,4]$ si $F(x) = 1, 4 < x$. Să se arate că F este o funcţie de repartiţie şi să se calculeze:

i) probabilitatea ca un călător să aştepte mai mult de 3 minute.

ii) probabilitatea ca un călător să aştepte mai puţin de 3 minute ştiind că a aşteptat mai mult de un minut.

Soluţie şi barem.

Din oficiu.... 1 punct.

F trebuie să fie crescătoare, cu valori în $[0,1]$, continuă la stânga, cu limita 0 la $-\infty$ şi cu limita 1 la infinit... 2 puncte.

F satisface, evident aceste condiţii (chiar mai mult F este continuă şi derivabilă cu excepţia unui număr finit de puncte)... 1,5 puncte.

Pentru punctul i) trebuie calculată $P(X > 3)$... 0,5 puncte.

Funcţia F fiind continuă rezultă că $P(X > 3) = 1 - P(X \leq 3) = 1 - F(3)$ căci $P(X = 3) = 0$... 1,5 puncte.

Avem deci $P(X > 3) = 1 - \frac{3}{4} = \frac{1}{4}$... 1 punct.

Pentru punctul ii) trebuie calculată $P(X < 3 \mid X > 1)$... 0,5 puncte.

$P(X < 3 \mid X > 1) = \frac{P(X<3, X>1)}{P(X>1)} = \frac{P(1<X<3)}{P(X>1)}$... 1 punct.

În final $P(X < 3 \mid X > 1) = \frac{1}{2}$ ... 1 punct.

2). V.a discretă X ia valorile $1, 2, ..., n, ...$ cu probabilităţile $p_n = P(X = n) = e^{-\alpha}(1 - e^{-\alpha})^{n-1}$unde $\alpha > 0$ este o constantă. Să se verifice că repartiţia de probabilitate este corectă şi să se calculeze $M[X]$ şi $D[X]$.

Soluţie.

Pentru corectitudine trebuie ca $p_n \in [0,1], \forall n \geq 1$ şi $\sum_{n=1}^{\infty} p_n = 1$. Prima condiţie este clară; pentru cea de a doua avem: $\sum_{n=1}^{\infty} e^{-\alpha}\left(1 - e^{-\alpha}\right)^{n-1} = e^{-\alpha}\sum_{n=1}^{\infty}\left(1 - e^{-\alpha}\right)^{n-1} = e^{-\alpha}\dfrac{1}{1 - (1 - e^{-\alpha})} = 1$ (s-a folosit suma seriei geometrice).

Pentru existenţa şi calculul mediei trebuie considerată seria $\sum_{n=1}^{\infty} n p_n = e^{-\alpha}\sum_{n=1}^{\infty} n\left(1 - e^{-\alpha}\right)^{n-1}$. Să considerăm seria de puteri seria $\sum_{n=1}^{\infty} n x^{n-1}$ pentru $x \in (0,1)$. Se ştie că această serie este convergentă şi suma sa este derivata sumei seriei geometrice $\sum_{n=1}^{\infty} x^n = \dfrac{1}{1-x}$. Deci $\sum_{n=1}^{\infty} n x^{n-1} = \dfrac{1}{(1-x)^2}$. Deci $M[X]$ există şi $M[X] = e^{-\alpha} \frac{1}{e^{-2\alpha}} = e^{\alpha}$.

Pentru existenţa şi calculul dispersiei trebuie arătat că X are moment de ordinul 2, $m_2[X]$ şi aplicată formula $D[X] = m_2[X] - M[X]^2$. Trebuie considerată seria $e^{-\alpha} \sum_{n=1}^{\infty} n^2 \left(1 - e^{-\alpha}\right)^{n-1}$ şi la fel ca mai sus seria de puteri $\sum_{n=1}^{\infty} n^2 x^{n-1}$; suma acestei serii ($x \in (0,1)$) este derivata sumei seriei $\sum_{n=1}^{\infty} n x^n$.

Calcule analoage celor de mai sus duc la $m_2[X] = e^{-\alpha} \sum_{n=1}^{\infty} n^2 \left(1 - e^{-\alpha}\right)^{n-1} = e^{2\alpha}\left(2 - e^{-\alpha}\right)$ etc.

3). Se dă funcţia $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 1 - |1 - x|$ pentru $x \in (0,2)$ şi $f(x) = 0$ în rest. Să se arate că f este o densitate de repartiţie şi să se calculeze media şi dispersia unei v.a X continue cu densitatea f.

Soluţie.

Să calculăm $\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 x dx + \int_1^2 (2-x)\,dx = \frac{1}{2} + \frac{1}{2} = 1$; celelate condiţii pentru densitate sunt clare.

Pentru medie avem $M[X] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x^2 dx + \int_1^2 x(2-x)\,dx = 2$

Mai departe $m_2[X] = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^1 x^3 dx + \int_1^2 x^2 (2-x)\,dx = \frac{7}{6}$ etc.

15.9.2 Exerciţii şi probleme propuse.

1). V.a X are densitatea de repartiţie nulă în afara intervalului $[0,a]$ iar pe intervalul $[0,a]$ graficul densităţii este un segment de dreaptă cu o extremitate în punctul $(a,0)$. Să se determine funcţia de repartiţie, media, dispersia v.a X şi $P\left(\frac{a}{2} \leq X < a\right)$.

Răspuns. $F(x) = 0, x \leq 0$, $F(x) = \frac{x}{a}\left(2 - \frac{x}{a}\right), x \in (0,a]$ şi $F(x) = 1, x > a$; $M[X] = \frac{a}{3}$, $D[X] = \frac{a^2}{18}, P\left(\frac{a}{2} \leq X < a\right) = \frac{1}{4}$.

2). Fie $a > 0$ şi $f(x) = \frac{1}{2a} e^{-\frac{|x-a|}{a}}, x \in \mathbb{R}$. Să se arate că f este o densitate de repartiţie şi să se calculeze media şi dispersia unei v.a X cu densitatea f.

Răspuns. $M[X] = a$, $D[X] = 2a^2$.

3) V.a discretă X ia valorile $-1, 0, 1$ cu probabilităţile $P(X = -1) = 0,2, P(X = 0) = 0,3$ şi $P(X = 1) = p$. Să se determine:

p, $M[X], M[3X], M[X+1], M\left[X^2\right]$.

Răspuns. $p = 0,5$, $M[X] = 0,3$, $M[3X] = 0,9$, $M[X+1] = 1,3$, $M\left[X^2\right] = 0,7$.

4) V.a discretă X ia valori în $\mathbb{N}$ astfel încât $p_k = P(X = k) = \frac{a^k}{(a+1)^{k+1}}$ unde $a > 0$ este o constantă. V.a Y ia valorile $1, 2, ..., 99$ cu egală probabilitate. Să se calculeze $M[X + Y]$.

Răspuns. $a + 51$.

15.10 Capitol MF.10. Legi de probabilitate.

15.10.1 Exerciţii şi probleme rezolvate.

1) V.a X este uniform repartizată în intervalul $(-1, 1)$. Să se determine densitatea de repartiţie a v.a $Y = e^X$ si $M[Y]$.

Soluţie şi barem.

Din oficiu ... 1 punct.

V.a Y este definită astfel: $Y(\omega) = e^{X(w)}$ pentru orice $\omega \in \Omega$ ... 1 punct.

Densitatea de repartiţie a v.a X este $f(x) = \frac{1}{2}, x \in (-1, 1)$ si $f(x) = 0$ în rest ... 1 punct.

Funcţia de repartiţie a v.a X va fi $F_X(x) = 0, x \in (-\infty, -1]$, $F_X(x) = \frac{x+1}{2}, x \in (-1, 1]$ şi $F_X(x) = 1, 1 < x$... 1 punct.

$P\left(e^X < y\right) = P(X < \ln y)$ dacă $y > 0$ şi 0 în rest. Dar $-1 < \ln y < 1$ înseamnă $\frac{1}{e} < y < e$ ... 2 puncte.

Deci $F_Y(y) = \frac{\ln y+1}{2}$ dacă $\frac{1}{e} < y < e$ şi $F_Y(y) = 0$ în rest ... 1 punct.

Deducem densitatea v.a Y : $g(y) = \frac{1}{2y}, \frac{1}{e} < y < e$ şi $g(y) = 0$ în rest ... 2 puncte.

$M[Y] = \int_{\frac{1}{e}}^{e} \frac{1}{2} dy = \frac{1}{2}\left(e - \frac{1}{e}\right)$ ... 1 punct.

2). V.a X este $N(0, \sigma)$ repartizată şi $0 < a < b$. Să se determine σ astfel încât $P(a < X < b)$ să fie maximă.

Soluţie.

V.a $\frac{X}{\sigma}$ este $N(0, 1)$ repartizată. Deci $P(a < X < b) = P\left(\frac{a}{\sigma} < \frac{X}{\sigma} < \frac{b}{\sigma}\right) = \Phi\left(\frac{b}{\sigma}\right) - \Phi\left(\frac{a}{\sigma}\right) = \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{\frac{b}{\sigma}} e^{-\frac{x^2}{2}} dx - \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx\right)$. Considerând $P(a < X < b)$ că o funcţie de $\sigma > 0$ căutăm extremele cu ajutorul derivatei.

Avem $\frac{d}{d\sigma}(P(a < X < b)) = \frac{1}{\sqrt{2\pi}}\left(e^{-\frac{b^2}{2\sigma^2}}\left(-\frac{b}{\sigma^2}\right) - e^{-\frac{a^2}{2\sigma^2}}\left(-\frac{a}{\sigma^2}\right)\right)$. Condiţia de anulare a derivatei dă $be^{-\frac{b^2}{2\sigma^2}} = ae^{-\frac{a^2}{2\sigma^2}}$ de unde $\sigma = \sqrt{\frac{b^2-a^2}{2(\ln b - \ln a)}}$. Se verifică apoi că acesta este un punct de maxim (global).

3) Probabilitatea lovirii unei ţinte dintr-o singură tragere este 0.001. Aproximând repartiţia binomială cu o repartiţie Poisson să se determine probabilitatea lovirii ţintei de cel puţin două ori în 5000 de trageri independente.

Soluţie.

Parametrul legii Poisson va fi $\lambda = n\mathfrak{p} = 5000 \times 0,001 = 5$. Avem de calculat $1 - p_0 - p_1$ (probabilitatea ca v.a Poisson să ia valori ≥ 2).

$p_0 = e^{-5}$, $p_1 = 5e^{-5}$ deci probabilitatea căutată va fi $1 - 6e^{-5}$ care este aproximativ $0,96$.

15.10.2 Exerciţii şi probleme propuse.

1) Fie X o v.a cu funcţia de repartiţie F continuă (ca funcţie de o variabilă). Să se arate că v.a $Y = F(X)$ este uniform repartizată în $(0,1)$.

2) V.a X urmează legea $N(2,2)$. Să se exprime cu ajutorul funcţiei Φ probabilităţile $P(0 \leq X \leq 3)$ şi $P(|X| \leq 1)$ şi să se calculze folosind tabele.

Răspuns. $P(0 \leq X \leq 3) = \Phi\left(\frac{1}{2}\right) - \Phi(-1) = 0,53$ etc.

3) V.a X este uniform distribuită în $(0,1)$. Să se determine repartiţia v.a $[nX] + 1$ unde $n \geq 1$ este un număr natural fixat iar $[x]$ este partea intreagă a numărului x.

Răspuns. Egal probabilitate pe $\{1, 2, ..., n\}$.

4) Un garaj deserveşte 70 de camioane. Probabilitatea ca, într-un an, un camion să intre în reparaţie este $0,3$ şi intrarea în reparaţie a unui camion nu influenţează intrarea celorlalte. Care este media de camioane în reparaţie într-un an?

Răspuns. 21.

15.11 Capitol MF.11. Vectori aleatori.

15.11.1 Exerciţii şi probleme rezolvate.

1). O urnă conţine a bile albe şi b bile negre. Se extrage succesiv câte o bilă fără reintroducere. Fie X_1 v.a care ia valoarea 1 dacă, la prima extragere, bila este albă şi 0 dacă, la prima extragere bila este neagră şi X_2 v.a definită similar dar pentru cea de a doua extragere.

i) Să se determine repartiţiile v.a X_1, X_2 şi mediile acestor v.a.

ii) Să se calculeze repartiţia vectorului aleator (X_1, X_2) şi covarianţa $K_{x_1 x_2}$ pentru $a = 2, b = 3$.

iii) Să se calculeze $D[X_1 + X_2]$ $(a = 2, b = 3)$.

iv) Dacă se extrag două bile simultan, comparaţi probabilitatea ca acestea să fie albe cu probabilitatea să obţinem două bile albe succesiv.

Soluţie şi barem.

Din oficiu; 1 punct.

i) $P(X_1 = 1) = \frac{a}{a+b}$, $P(X_1 = 0) = \frac{b}{a+b}$ deci $M[X_1] = \frac{a}{a+b}$; 1 punct.

folosind formula probabilităţii totale avem $P(X_2=1)=\frac{a}{a+b}\frac{a-1}{a+b-1}+\frac{b}{a+b}\frac{a}{a+b-1}=\frac{a}{a+b}$ deci $M[X_2]=\frac{a}{a+b}$; 2 puncte.

ii) $P(X_1=0,X_2=0)=P(X_1=0)P(X_2=0\mid X_1=0)=\frac{3}{5}\frac{2}{4}=\frac{3}{10}$. Similar se deduc: $P(X_1=0,X_2=1)=\frac{3}{10}, P(X_1=1,X_2=0)\,\frac{3}{10}$, $P(X_1=1,X_2=1)=\frac{1}{10}$; 2 puncte.

Avem $K_{x_1x_2}=M[X_1X_2]-M[X_1]\,M[X_2]$; 0,5 puncte.

Evident $M[X_1X_2]=\frac{1}{10}$ deci $K_{x_1x_2}=\frac{1}{10}-\frac{4}{25}=-\frac{3}{50}$; 0,5 puncte.

iii) $D[X_1+X_2]=D[X_1]+D[X_2]+2K_{x_1x_2}$; 1 punct.

$D[X_1]=D[X_2]=\frac{2}{5}-\frac{4}{25}=\frac{6}{25}$; 0,5 puncte.

$D[X_1+X_2]=\frac{12}{25}-\frac{6}{50}=\frac{9}{25}$; 0,5 puncte.

iv) Dacă notăm cu p probabilitatea ca bilele extrase simultan să fie albe avem $p=\frac{C_a^2}{C_{a+b}^2}=\frac{a(a-1)}{(a+b)(a+b-1)}=P(X_1=1,X_2=1)$; 1 punct.

2) Fie (X,Y) un vector aleator cu densitatea de repartiţie comună $f(x,y)=\frac{1}{4},(x,y)\in[0,2]\times[0,2]$ şi $f(x,y)=0$ în rest. Să se calculeze:

i) $P(X\le 1,Y\le 1)$.

ii) $P(X+Y\le 1)$.

iii) $P(X+Y>2)$.

Soluţie.

Se verifică, pentru orice eventualitate, că f este o densitate (pozitivă, integrala pe $\mathbb{R}^2$ egala cu 1). Apoi avem :

$P(X\le 1,Y\le 1)=\int_{-\infty}^{1}\int_{-\infty}^{1}f(x,y)\,dxdy=\int_0^1\int_0^1\frac{1}{4}dxdy=\frac{1}{4}$, $P(X+Y\le 1)=\int\int_{x+y\le 1}f(x,y)\,dxdy=\frac{1}{4}\int_0^1dx\int_0^{1-x}dy=\frac{1}{8}$.

Pentru iii) observăm că $P(X+Y>2)=1-P(X+Y\le 2)$ şi procedăm ca mai sus. Obţinem $P(X+Y>2)=\frac{1}{2}$.

3) V.a X,Y sunt independente şi repartizate $N(2,1)$ respectiv $N(-3,2)$. Să se calculeze:

i) $P(X<2,Y<-3)$.

ii) $P(Y<X-5)$.

Soluţie.

i) $P(X<2,Y<-3)=P(X<2)\,P(Y<-3)$ (independentă). Fiindcă legea normală este simetrică în raport cu media, obţinem: $P(X<2,Y<-3)=\frac{1}{2}\frac{1}{2}=\frac{1}{4}$.

ii) $P(Y<X-5)=\frac{1}{2\pi\sqrt{2}}\int\int_{y<x-5}e^{-\frac{(x-2)^2}{2}}e^{-\frac{(y+3)^2}{8}}dxdy$ (independentă). Dreapta $y=x-5$ trece prin punctul $(2,-3)$ (coordonatele mediilor) şi din ratiuni de simetrie obţinem $P(Y<X-5)=\frac{1}{2}$.

15.11.2 Exerciţii şi probleme propuse.

1) Fie X,Y v.a independente şi la fel repartizate care iau valorile $1,2,...,n,...$ cu probabilităţile $p_n=\frac{1}{2^n}$. Să se calculeze:

i) $P(\min\{X,Y\}\le k)$.

ii) $P(Y > X), P(X = Y)$.

Răspuns.

i) $1 - \frac{1}{4^k}$.

ii) $\frac{1}{3}$, $\frac{1}{3}$.

2) V.a (X, Y) are densitatea de repartiţie comună $f(x, y) = e^{-(x+y)}, x, y \geq 0$ si $f(x, y) = 0$ în rest. Să se calculeze:

i) $P(X \leq 1, Y \leq 1)$.

ii) $P(X + Y > 2)$.

Răspuns.

i) $\left(1 - \frac{1}{e}\right)^2$.

ii) $\frac{2}{e^2}$.

3) Fie X, Y exponenţial repartizate cu parametrii λ respectiv μ $(\lambda \neq \mu)$ şi independente. Să se calculeze densitatea v.a $Z = X + Y$.

Răspuns.

$f_Z(x) = \frac{\lambda\mu}{\mu-\lambda}\left(e^{-\lambda x} - e^{-\mu x}\right), x \geq 0$.

4) V.a continuă X are densitatea f. Fie $Y = X^2$. Să se determine funcţia de repartiţie F a vectorului aleator (X, Y).

Răspuns.

$F(x, y) = 0$ dacă $y \leq 0$ sau dacă $y > 0$ şi $x \leq \sqrt{y}$; $F(x, y) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)\, dx$ dacă $y > 0$ şi $x > \sqrt{y}$; $F(x, y) = \int_{-\sqrt{y}}^{x} f(x)\, dx$ dacă $y > 0$ şi $-\sqrt{y} < x \leq \sqrt{y}$.

15.12 Capitol MF.12. Legea numerelor mari.

15.12.1 Exerciţii şi probleme rezolvate.

1) Se notează ξ_n frecvenţa relativă de apariţie, la aruncarea de n ori, a unei feţte fixate a unei monede. Să se determine n astfel încât $P\left(\left|\xi_n - \frac{1}{2}\right| < \frac{1}{100}\right) \geq 0,99$.

Soluţie şi barem.

Din oficiu; 1 punct.

Considerăm v.a X care ia doar două valori, 1 dacă apare fata respectivă si 0 în caz contrar. Clar, X ia valoarea 1 cu probabilitatea $\frac{1}{2}$. Dacă $X_1, X_2, ..., X_n$ sunt independente şi având aceeaşi repartiţie ca X atunci $\xi_n = \frac{X_1+X_2+...+X_n}{n}$: 2 puncte.

$M[X] = \frac{1}{2}$ şi $D[X] = \frac{1}{4}$; 2 puncte.

Inegalitatea Cebîşev dă $P\left(\left|\xi_n - \frac{1}{2}\right| < \frac{1}{100}\right) \geq 1 - \frac{D[X]}{n10^{-4}} = 1 - \frac{1}{4n10^{-4}} = 1 - \frac{10^4}{4n}$; 3 puncte.

Condiţia este $1 - \frac{10^4}{4n} \geq \frac{99}{100}$ care dă $n \geq 25 \cdot 10^4$; 2 puncte.

2) Fie $(X_n)_n$ un şir de v.a independente astfel încât $P(X_n = n^\alpha) = P(X_n = -n^\alpha) = \frac{1}{2}$ unde $\alpha \in \mathbb{R}$ este o constantă. Să se arate că pentru $\alpha < \frac{1}{2}$ şirul verifică legea numerelor mari.

Soluţie.

$M[X_n] = 0$ pentru orice n. Rezultă că $D[X_n] = M\left[X_n^2\right] = n^{2\alpha}$. Inegalitatea Cebîşev dă $P(|X_1 + X_2 + ... + X_n| \geq \varepsilon) \leq \frac{D[X_1+X_2+...+X_n]}{n^2\varepsilon^2} = \frac{1+2^{2\alpha}+...+n^{2\alpha}}{n^2\varepsilon^2}$. Trebuie calculată limita $\lim\limits_{n\to\infty} \frac{1+2^{2\alpha}+...+n^{2\alpha}}{n^2}$; în vederea aplicării lemei Cesaro-Stolz considerăm $\lim\limits_{n\to\infty} \frac{(n+1)^{2\alpha}}{2n+1}$.

Dacă $\alpha < \frac{1}{2}$ această limită este 0 deci $\lim\limits_{n\to\infty} P(|X_1 + X_2 + ... + X_n| \geq \varepsilon) = 0$.

3) Cu ce probabilitate putem afirma că, din 100 de aruncări ale unei monede, o faţă anume apare de un număr de ori între 40 şi 60 ?

Soluţie.

Avem $\lim\limits_{n\to\infty} P\left(a \leq \frac{X_1+X_2+...+X_n-n\mathfrak{p}}{\sqrt{n\mathfrak{p}\mathfrak{q}}} < b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}dt = \Phi(b) - \Phi(a)$ (Moivre-Laplace). În altă formă putem scrie $\lim\limits_{n\to\infty} P\left(a \leq \sqrt{\frac{n}{\mathfrak{p}\mathfrak{q}}}\left(\frac{X_1+X_2+...+X_n}{n} - \mathfrak{p}\right) < b\right) = \Phi(b) - \Phi(a)$. În cazul de faţă vom aproxima, pentru $n = 100$, $P\left(a \leq \sqrt{\frac{n}{\mathfrak{p}\mathfrak{q}}}\left(\frac{X_1+X_2+...+X_n}{n} - \mathfrak{p}\right) < b\right)$ cu $\Phi(b) - \Phi(a)$. Avem $\mathfrak{p} = \mathfrak{q} = \frac{1}{2}$ şi $\sqrt{\frac{n}{\mathfrak{p}\mathfrak{q}}} = 20$. Deducem $a = 20(0,4 - 0,5) = -2, b = 20(0,6 - 0,5) = 2$. Din tabele se deduce $\Phi(2) - \Phi(-2) = 0,954$.

15.12.2 Exerciţii şi probleme propuse.

1) Fie $(X_n)_{n\geq 2}$ un şir de v.a independente astfel încât $P(X_n = n) = P(X_n = -n) = \frac{1}{2n\ln n}, P(X_n = 0) = 1 - \frac{1}{n\ln n}$. Să se arate că acest şir satisface legea numerelor mari.

Răspuns.

$M[X_n] = 0, D[X_n] = \frac{n}{\ln n}$; $\lim\limits_{n\to\infty} \frac{\sum D[X_k]}{n^2} = 0$.

2). De câte ori este suficient să se arunce un zar astfel încât să se poată afirma că faţa 3 apare cu probabiliatea $0,99$?

Răspuns.

Aproximativ 370.

3) Fie $(X_n)_{n\geq 1}$ un sir de v.a independente astfel încât X_n ia valorile $-n, -n+1, ..., 0, ..., n-1, n$ cu probabilităţile $P(X_n = k) = P(X_n = -k) = \frac{1}{3k^3}, k \neq 0$ şi $P(X_n = 0) = 1 - \frac{2}{3}\left(1 + \frac{1}{2^3} + ... + \frac{1}{n^3}\right)$. Să se arate că şirul dat satisface legea numerelor mari.

Răspuns.

Se arată că $\lim\limits_{n\to\infty} \frac{\sum D[X_k]}{n^2} = 0$.

4) V.a discretă X ia valori naturale cu probabilităţile $p_n = P(X = n) = \frac{1}{e^2}\frac{2^n}{n!}$. Să se determine funcţia caracteristică şi $M[X]$.

Răspuns.

$g_X(t) = \sum_0^\infty e^{itn} p_n = e^{2\left(e^{it}-1\right)}$; $g_X'(t) = 2ie^{it}e^{2\left(e^{it}-1\right)}$ şi deci $g_X'(0) = 2i$ de unde $M[X] = \frac{2i}{i} = 2$.

15.13 Capitol MF.13. Lanţuri Markov.

15.13.1 Exerciţii şi probleme rezolvate.

1). Se consideră un lanţ Markov cu două stări notate 1 respectiv 2 , cu condiţia iniţială $p = (\alpha, \beta), \alpha + \beta = 1$ matricea de trecere $\Pi = \begin{pmatrix} \mathfrak{p} & \mathfrak{q} \\ \mathfrak{q} & \mathfrak{p} \end{pmatrix}$ $\mathfrak{p} + \mathfrak{q} = 1, \mathfrak{p} \in (0, 1)$.

i) Să se calculeze matricea de trecere în doi paşi.

ii) Să se calculeze probabilităţile $P_1(2), P_2(2)$ de a fi în stare 1 respectiv 2 în doi paşi pentru $\alpha = \frac{1}{3}, \beta = \frac{2}{3}$.

iii) Să se arate prin inducţie că $\Pi^n = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{(\mathfrak{p}-\mathfrak{q})^n}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ şi cu notaţii ca în ii) să se calculeze $\lim_{n\to\infty} P_1(n), \lim_{n\to\infty} P_2(n)$.

iv) Să se calculeze $P(X_0 = 1 \mid X_2 = 1)$.

Soluţie şi barem.

Din oficiu; 1 punct.

i) $\Pi^2 = \begin{pmatrix} \mathfrak{p}^2 + \mathfrak{q}^2 & 2\mathfrak{p}\mathfrak{q} \\ 2\mathfrak{p}\mathfrak{q} & \mathfrak{p}^2 + \mathfrak{q}^2 \end{pmatrix}$; 1 punct.

ii) Folosind formula probabilitătii totale avem $P_1(2) = \frac{1}{3}\left(\mathfrak{p}^2 + \mathfrak{q}^2\right) + \frac{2}{3}2\mathfrak{p}\mathfrak{q} = \frac{1}{3}\left(\mathfrak{p}^2 + \mathfrak{q}^2 + 4\mathfrak{p}\mathfrak{q}\right), P_2(2) = \frac{1}{3}2\mathfrak{p}\mathfrak{q} + \frac{2}{3}\left(\mathfrak{p}^2 + \mathfrak{q}^2\right)$ etc; 2 puncte.

iii) Se verifică pentru $n = 1$ şi apoi presupunând relaţia adevărată pentru n se înmulţeşte cu Π etc; 1 punct.

Deducem $P_1(n) = \frac{1}{2} + \frac{(\alpha-\beta)(\mathfrak{p}-\mathfrak{q})^n}{2}$, $P_2(n) = \frac{1}{2} - \frac{(\alpha-\beta)(\mathfrak{p}-\mathfrak{q})^n}{2}$; 1 punct.

$\lim_{n\to\infty} P_1(n) = \lim_{n\to\infty} P_2(n) = \frac{1}{2}$; 1 punct.

iv) Folosim formula Bayes $P(X_0 = 1 \mid X_2 = 1) = \frac{P(X_0=1)P(X_2=1|X_0=1)}{P(X_2=1)}$; 2 puncte.

Obtinem $P(X_0 = 1 \mid X_2 = 1) = \frac{\alpha+\alpha(\mathfrak{p}-\mathfrak{q})^n}{1=(\alpha-\beta)(\mathfrak{p}-\mathfrak{q})^n}$; 1 punct.

2) Se aruncă, în mod repetat, un zar.

i) Fie X_n v.a "cel mai mare număr apărut până la a n-aruncare".

ii) Fie N_n v.a "numărul de feţe 3 apărute până la a n-aruncare".

Să se studieze dacă şirurile respective generează lanţuri Markov specificându-se, în caz afirmativ, matricea de trecere.

Soluţie.

i) Dacă Y_n este v.a " numărul apărut la a n-aruncare, atunci $X_{n+1} = \max\{X_n, Y_{n+1}\}$ astfel că avem un lanţ Markov (indiferent de o condiţie iniţială fixată). Matricea de trecere va fi $p(i,j) = 0, j < i$, $p(i,j) = \frac{i}{6}, i = j$ şi $p(i,j) = \frac{1}{6}, i < j$.

ii) Şirul generează un lanţ Markov (nu cu un număr finit de stări) şi $p(i,j) = \frac{1}{6}, j = i+1$, $p(i,j) = \frac{5}{6}, i = j$, $p(i,j) = 0$,în rest. Matricea de trecere este "infinită " în acest caz.

3) Într-un lanţ Markov starea j este accesibilă din starea i dacă există $n \geq 1$ astfel încât $p(n,i,j) > 0$.

i) Să se arate că relaţia " j este accesibilă din starea i " este tranzitivă.

ii) În ce caz într-un lanţ Markov cu două stări starea 1 nu este accesibilă din ea însăsi ?

Soluţie.

i) Fie j este accesibilă din starea i şi k este accesibilă din starea j , $p(n,i,j) > 0, p(m,j,k) > 0$. Conform relaţiei Chapman - Kolmogorov avem $p(m+n,i,k) \geq p(n,i,j)\,p(m,j,k) > 0$ deci starea k este accesibilă din starea i.

ii) Matricea de trecere Π trebuie să aibă forma $\begin{pmatrix} 0 & \beta \\ \alpha & \gamma \end{pmatrix}$ cu $\beta = 1$. Dacă $\alpha \neq 0$ atunci starea 2 ar fi accesibilă din starea 1 şi starea 1 ar fi accesibilă din starea 2 deci prin tranzitivitate starea 1 ar fi accesibilă din ea însăsi. Deci necesar $\alpha = 0$ şi $\Pi = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$. Condiţia este şi suficientă.

15.13.2 Exerciţii şi probleme propuse.

1). Un lanţ Markov are matricea de trecere $\Pi = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$. Care este numărul de stări ? Să se calculeze matricea de trecere în 3 paşi.

Răspuns.

3 , $\begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$.

2). Fie $(X_n)_n$ un lanţ Markov cu mulţimea de stări S. Un eveniment este anterior momentului $n \geq 0$ dacă este de forma $A = \left\{(X_0, X_1, ..., X_n) \in M \subseteq S^{n+1}\right\}$. Să se arate că $P(X_{n+1} = i_{n+1} \mid X_n = i_n, A) = P(X_{n+1} = i_{n+1} \mid X_n = i_n)$.

Răspuns.

Indicaţie: $P(X_n = i_n, A) = \sum_{(i_0,i_1,...,i_{n-1},i_n) \in M} P(X_n = i_n, X_{n-1} = i_{n-1}, ..., X_0 = i_0)$ (i_n fixat).

3). Să se scrie matricea de trecere a mersului la întamplare cu stările $0, 1, 2, 3, 4$ şi cu frontiere (0 şi 4) absorbante.

Răspuns.

Pentru $p+q=1$,

	0	1	2	3	4
0	1	0	0	0	0
1	p	0	q	0	0
2	0	p	0	q	0
3	0	0	p	0	q
4	0	0	0	0	1

4). Fie procesul stochastic $X(t) = At + B, t \geq 0$ unde v.a A este $N(0,1)$ repartizată, v.a B este uniform repartizată în $(0,1)$ şi v.a A si B sunt independente. Să se determine media procesului şi funcţia de corelaţie.

Răspuns.

$M[X(t)] = \frac{1}{2}$, $R(t,s) = cov(At+B, As+B) = ts + \frac{1}{12}$.

15.14 Capitol MF.13. Statistică Matematică.

15.14.1 Exerciţii şi probleme rezolvate.

1) i) Să se determine un interval de încredere pentru parametrul σ^2 al legii normale $N(m,\sigma)$ cu m cunoscut folosind repartiţia χ^2.

ii) Să se determine un interval de încredere 95% pentru $N(1,\sigma)$, $n = 10$ şi datele $x_i : 1,3; 0,9; 1,1; 0,8; 1,3; 0,8; 1,1; 0,9; 0,6; 1,4$.

Soluţie şi barem.

Din oficiu; 1 punct.

i) Dacă X este $N(m,\sigma)$ repartizată atunci $\frac{X-m}{\sigma}$ este $N(0,1)$ repartizată; 1,5 puncte.

Rezultă, pentru o selecţie $(X_1, X_2, ..., X_n)$ asupra v.a X că

$Y = \frac{1}{\sigma^2}\left[(X_1-m)^2 + (X_2-m)^2 + ... + (X_n-m)^2\right]$ este $\chi^2(n)$ repartizată; 1,5 puncte.

Fie $\alpha \in (0,1)$; se pot determina a, b (utilizând tabela) astfel încât $P(Y > b) = \frac{1-\alpha}{2}, P(Y > a) = \frac{1+\alpha}{2}$; rezultă $P(a < Y < b) = \alpha$; 2 puncte.

Avem $a < \frac{1}{\sigma^2}\left[(X_1-m)^2 + (X_2-m)^2 + ... + (X_n-m)^2\right] < b$ dacă şi numai dacă $\frac{\left[(X_1-m)^2+(X_2-m)^2+...+(X_n-m)^2\right]}{a} < \sigma^2 < \frac{\left[(X_1-m)^2+(X_2-m)^2+...+(X_n-m)^2\right]}{b}$ care este un interval de încredere $100\alpha\%$ pentru σ^2; 1 punct.

ii) Se calculează $\sum(x_i-1)^2 = 0,62$; 1 punct.

Din tabele $a = 20,58, b = 3,25$; 1 punct.

În definitiv se obţine intervalul $(0,03, 0,19)$; 1 punct.

2) Se consideră experimentul aleator ce constă din aruncarea a 12 zaruri. Fie X v.a "numărul de zaruri cu feţele 4, 5 sau 6. Se consideră o selecţie

de volum $n = 4096$ asupra v.a X şi se notează N_i numărul de apariţii ale valorii $X = i = 0, 1, ..., 12$. O realizare a selecţiei este:

i	0	1	2	3	4	5	6	7	8	9	10	11	12	*total*
N_i	0	7	60	198	430	731	948	847	536	257	71	11	0	$n = 4096$

i) Să se reprezinte în plan punctele $\left(i, \frac{N_i}{n}\right), i == 0, 1, ..., 12$ si să se calculeze media de selecţie corespunzătoare realizării.

ii) Ce lege teoretică urmează v.a X ? Să se determine ε astfel încât $P(|X - m| \leq \varepsilon) = 0,998$ (unde m este media teoretică a v.a X) şi să se compare cu abaterea mediei de selecţie de la media teoretică.

Soluţie.

i) Se calculează $\frac{N_i}{n}, i == 0, 1, ..., 12$; de exemplu $\frac{N_1}{n} = 0,001, \frac{N_2}{n} = 0,0146$ etc. Media de selecţie este $6,1389$.

ii) Teoretic, v.a X este $B\left(12, \frac{1}{2}\right)$ repartizată, deci $M[X] = 12 \cdot \frac{1}{2} = 6$ iar $D[X] = 12 \cdot \frac{1}{4} = 3$. Aproximând v.a $\sqrt{\frac{4096}{3}}(\frac{X_1+X_2+...+X_{4096}}{4096} - 6)$ cu legea normală $N(0,1)$ şi folosind tabela pentru funcţia Φ găsim $\varepsilon = 0,083$. Constatând că abaterea mediei de selecţie de la media teoretică este $6,1389 - 6 = 0,1389$ rezultă că evenimentul observat (realizarea) este putin probabil.

3) (Regresie liniară). i) Fie $X, Y \in L^2$ două v.a pe un câmp de probabilitate. Să se determine $a, b \in \mathbb{R}$ astfel încât $M\left[(Y - aX - b)^2\right]$ să fie minimă (aproximare în medie patratică a v.a Y cu o v.a $aX + b$). În particular să se studieze cazul în care câmpul de probabilitate este "egal probabilitatea" pe o mulţime cu n elemente.

Soluţie. Vom presupune $D[X] \neq 0$. Pentru simplificarea calculului considerăm, la început, situaţia $M[X] = M[Y] = 0$ (variabile centrate). Avem $M\left[(Y - aX - b)^2\right] = M\left[Y^2\right] - 2aM[XY] + a^2M\left[X^2\right] + b^2$. Minimul se obţine pentru $b = 0$ şi $a = \frac{M[XY]}{M[X^2]} = \frac{M[XY]}{D[X]}$. În cazul general scriem $Y - aX - b = Y - M[Y] - a(X - M[X]) + M[Y] - aM[X] - b$ si aplicăm rezultatul precedent pentru varaibilele centrate $Y - M[Y], X - M[X]$. Obţinem $a = \frac{cov[X,Y]}{D[X]}$ şi $b = M[Y] - aM[X]$. Dreapta $y - M[Y] = \frac{cov[X,Y]}{D[X]}(x - M[X])$ se numeşte dreaptă de regresie a v.a Y pe X.

Pentru cazul egal probabil X ia valorile $x_1, x_2, ..., x_n$ cu aceeaşi probabilitate $\frac{1}{n}$ şi analog pentru Y şi $y_1, y_2, ..., y_n$. Notând $\overline{x}$ respectiv $\overline{y}$ mediile aritmetice corespunzătoare vom avea $a = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2}$ şi $b = \overline{y} - a\overline{x}$ şi dreapta de regresie corespunzătoare. Pentru că, în acest caz, determinarea numerelor a, b revine la determinarea minimului funcţiei $d^2(a,b) = \sum_{i=1}^{n}(y_i - ax_i - b)^2$ dreapta de regresie se mai numeşte şi dreapta celor mai mici pătrate.

15.14.2 Exerciţii şi probleme propuse.

1) i) Fie v.a X repartizată $N(0,1)$ şi (ca de obicei) Φ funcţia ei de repartiţie. Arătaţi că funcţia Φ aplică bijectiv $\mathbb{R}$ pe $(0,1)$ şi dacă $\alpha \in (0,1)$ dacă $t_\alpha = \Phi^{-1}\left(\frac{1+\alpha}{2}\right)$ atunci $P(|X| \leq t) = \alpha$.

ii) Să se arate că, dată o selecţie $(X_1, X_2, ..., X_n)$ asupra unei v.a X cu repartiţie normală $N(\theta, \theta)$ atunci un interval de încredere $100\alpha\%$ pentru parametrul θ are forma $\left(\frac{\overline{X}}{1+\frac{t_\alpha}{\sqrt{n}}}, \frac{\overline{X}}{1-\frac{t_\alpha}{\sqrt{n}}}\right)$ unde $\overline{X} = \frac{X_1+X_2+...+X_n}{n}$

Răspuns. i) Se foloseşte relaţia $\Phi(t) + \Phi(-t) = 1$.

ii) V.a $\sqrt{n}\frac{\overline{X}-\theta}{\theta}$ este $N(0,1)$ repartizată şi $P\left(\sqrt{n}\frac{|\overline{X}-\theta|}{\theta} \leq t_\alpha\right) = \alpha$ etc.

2) S-au efectuat $n = 4000$ de experienţe în care evenimentele A, B, C , care constituie un sistem complet de evenimente, s-au realizat de $1905, 1015, 1080$ ori. Dat pragul de semnificaţie $0,05$ să se testeze ipoteza $p_1 = P(A) = \frac{1}{2}$, $p_2 = P(B) = \frac{1}{4}$, $p_3 = P(C) = \frac{1}{4}$.

Răspuns. Se foloseşte testul χ^2. Se obţine, din date, $\chi^2 = 11,13$ şi din tabele $\chi^2(2) = 5,99$. Ipoteza se respinge.

3) Fie $(X_1, X_2, ..., X_n)$ o selecţie asupra v.a X repartizată $N(\theta, \sigma)$. Dat $\alpha \in (0,1)$ şi $a, b, a < b$ astfel încât $\Phi(b) - \Phi(a) = \alpha$, să se arate că $\left(\overline{X} - \frac{\sigma}{\sqrt{n}}b, \overline{X} - \frac{\sigma}{\sqrt{n}}a\right)$ este un interval de încredere $100\alpha\%$ pentru θ. Dacă $(X_1, X_2, ..., X_n)$ si $(Y_1, Y_2, ..., Y_m)$ sunt selecţii asupra v.a X , $N(\theta_1, \sigma_1)$ repartizată, respectiv Y , $N(\theta_2, \sigma_2)$, repartizată, iar X şi Y sunt independente, să se construiască un interval de încredere pentru parametrul $\tau = \theta_1 - \theta_2$.

Răspuns. Prima parte a mai fost discutată. Pentru partea a doua considerăm v.a $\overline{X}$ repartizată $N\left(\theta_1, \frac{\sigma_1}{\sqrt{n}}\right)$ şi v.a $\overline{Y}$ repartizată $N\left(\theta_2, \frac{\sigma_2}{\sqrt{m}}\right)$ şi apoi v.a $\frac{\overline{X}-\overline{Y}-\tau}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}$ care este $N(0,1)$ repartizată etc.

4) Să se determine un estimator de verosimilitate maximă pentru $f(x,\theta) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(x-\theta)^2}{2}}$.

Răspuns. $u(X_1, X_2, ..., X_n) = \frac{X_1+X_2+...+X_n}{n}$.

Bibliografie

BIBLIOGRAFIE - Analiză funcţională
[B.01] Brânzănescu V., Stănăşilă O. *"Matematici speciale"*, Editura All, Bucureşti,1994.
[C.01] Colojoară I. *"Analiză matematică"*, Bucureşti, Ed. didactică şi pedagogică,1983.
[C.02] Cristescu R. *"Analiză funcţională"*, Bucureşti, Ed. didactică şi pedagogică,1979.
[D.01] Davidson K., Donsig A. *Real Analysis with Real Applications*, Prentice Hall, 2002.
[D.02] Dunford N., Schwartz J.T. *"Linear operators"*, Interscience Publ.,Part I,1958;Part II,1963.
[F.01] Feintuch A., Saeks R. *"System theory; a Hilbert space approach"*, Acad.Press,1982.
[F.02] Flondor P., Stănăşilă O. *"Lecţii de analiză matematică"*, Editura All, Bucureşti,1993.
[H.01] Halmos P.R. *"A Hilbert space problem book"*, Springer-Verlag, N. Y. Inc.,1970.
[H.02] Halmos P.R. *"Measure Theory"*, Springer 1974.
[O.01] Olteanu M. *"Curs de Analiză funcţională"*, Printech, 2000.
[R.01] Rudin W. *"Real and complex analysis"*, McGraw-Hill,1962.
[R.02] Rudin W. *"Fourier analysis on groups"*, Interscience Publishers,1962.
[S.01] Stanomir D., Stănăşilă O. *"Metode matematice în teoria semnalelor"*, Ed. teh., Bucureşti,1980.
[S.02] Stănăşilă O. (coordonator), colectiv *"Enciclopedie Matematică"*, Editura AGIR, 2010.

BIBLIOGRAFIE - Probabilităţi şi statistică matematică

[B.01] V. Brânzănescu, O.Stănăşilă : *Matematici speciale*. Ed. All. 1998.

[C.01] G.Ciucu, V.Craiu, L.Săcuiu : *Culegere de probleme de teoria probabilităţilor*. Ed. Tehnică. 1967.

[D.01] M.Dumitrescu, D.Florea, C Tudor : *Probleme de teoria probabilităţilor şi statistică matematică*. Ed. Tehnică. 1985.

[E.01] *Mică Enciclopedie de Statistică* : Ed.Ştiinţifică şi pedagogică. 1985.

[F.01] P.Flondor, O.Stănăşilă : *Lecţii de Analiză Matematică şi exerciţii rezolvate*. Ed. All. 2004.

[G.01] B.V.Gnedenko: *The Theory of Probability*. Mir.1976.

[G.02]B.Grais : *Methodes Statistiques*. Dunod.1988.

[G.03] C.M.Grinstead, J.L.Snell : *Introduction to Probability*. www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/ amsbook.mac.pdf.

[I.01] M.Iosifescu, Gh. Mihoc: *Teoria Probabilităţilor şi Statistică Matematică*. Ed. Didactică şi Pedagogică. 1970.

[I.02] M.Iosifescu : *Lanţuri Markov finite şi aplicaţii*. Ed. Tehnică. 1977.

[I.03] G.Ivchenko şi alţii : *Problems in Mathematical Statistics*. http://auadno.blog.com/2011/11/22/problems-in-mathematical-statistics-e-book/

[L.01] J.Lamperti : *Probability*. Darmouth College. 1966.

[M.01] N.Micu : *Statistică Matematică*. Acad. Militară. 1979.

[M.02] Gh.Mihoc, C.Bergthaller, V.Urseanu. *Procese stochastice, elemente de teorie şi aplicaţii.* Ed.Ştiinţific ă şi pedagogică. 1978.

[N.01] D.Nualart : *Stochastic Processes.*
http://www.mat.ub.edu/˜nualart/StochProc.pdf

[P.01] E,Parzen : *Stochastic Processes.* SIAM. 1999.

[P.02] A. Papoulis : *Probability and Statistics.* Prentice Hall 1990.

[S.01] P. Sabatini : *Introduction to Probability.*
http://www.sci.utah.edu/˜gerig/CS6640-F2010/prob-tut.pdf

[S.02] R. Serfozo : *Basic probability problems.*
http://www2.isye.gatech.edu/˜rserfozo/courses/isye2027/NotesChapters1–3.pdf

[Ş.01] I.Gh.Şabac : *Matematici Speciale 2.* Ed. Tehnică. 1977.

[V.01] J. Vrbik : *Mathematical Statistics.*
http://spartan.ac.brocku.ca/˜jvrbik/MATH2P82/Statistics.PDF