TEOREMA DE PUNCT FIX A LUI BANACH

Ion MUNTEANU[1]
munteanuion74@gmail.com

ABSTRACT: This article presents Banach's Fixed Point Theorem (also known as the contraction mapping theorem or contraction mapping principle) is an important tool in the theory of metric spaces; it guarantees the existence and uniqueness of fixed points of certain self-maps of metric spaces, and provides a constructive method to find those fixed points.

KEYWORDS: Banach, Theorem, Cauchy sequence, contraction, fixed point.

Teorema de punct fix a lui Banach, cunoscută și sub denumirea de *principiul contracțiilor*, este un instrument important în teoria spațiilor metrice; ea garantează existența și unicitatea soluțiilor ecuațiilor de forma $f(x) = x$ $f(x) = x$, pentru o clasa largă de aplicații *f*, și furnizează totodată o metodă constructivă de determinare a acestor soluții. Teorema a fost formulată și demonstrată, în 1922, de fondatorul *analizei funcționale*, Stefan Banach (1892-1945), și reprezintă o abstractizare a *metodei aproximațiilor succesive*, metodă utilizată în mod empiric încă din antichitate pentru rezolvarea ecuațiilor numerice, și, în cazul ecuațiilor diferențiale, introdusă de Joseph Liouville în 1837 și dezvoltată sistematic de Emile Picard începând cu anul 1890.

Iată, pentru început, câteva noțiuni și rezultate de bază din teoria spațiilor metrice.

Definiția 1. Numim *spațiu metric* o mulțime nevidă *X* dotată cu o metrica *d*, adică cu o funcție $d: X \times X \to \mathbb{R}$ $d: X \times X \to \mathbb{R}$ care satisface următoarele axiome:

i) $\forall x \in X, \forall y \in X \forall x \in X, \forall y \in X$, $d(x, y) \geq \geq 0$
și $d(x,y) = 0 \Leftrightarrow x = y d(x,y) = 0 \Leftrightarrow x = y$;

ii) $\forall x \in X, \forall y \in X \forall x \in X, \forall y \in X$, $d(x,y) = d(y,x)$;

1 Profesor de matematică la Colegiul Tehnic „Dimitrie Ghika”, Comănești, județul Bacău.

iii) $\forall x \in X, \forall y \in X, \forall z \in X \forall x \in X, \forall y \in X, \forall z \in X$, $d(x, z) \leq\leq d(x, y) + d(y, z)$.

Metrica *d* din definiția de mai sus mai este numită și *distanță* pe *X*, axiomele ei reținând proprietățile esențiale ale noțiunii comune de distanță. Pentru a desemna în același timp și mulțimea suport și metrica considerată, vom folosi notația *(X, d)*.

Dintre spațiile metrice uzuale amintim doar următoarele: mulțimea numerelor reale $\mathbb{R}\mathbb{R}$ cu distanța obișnuită $d(x,y) = |x - y|$; mulțimea numerelor complexe $\mathbb{C}\mathbb{C}$ cu $d(z,w) = |z - w|$; spațiile $\mathbb{R}^n$ $\mathbb{R}^n$ cu metrica euclidiană

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2},$$

și spațiul $C_{[a,b]}$ $C_{[a,b]}$ al funcțiilor continue $x : [a, b] \to \mathbb{R} \to \mathbb{R}$ dotat cu *metrica convergenței uniforme*

$$d(x,y) = \sup_{t \in [a,b]} |x(t) - y(t)| \sup_{t \in [a,b]} |x(t) - y(t)|$$

Pentru orice punct x_o al unui spațiu metric *X* definim *sfera* de rază $r > 0$ și centru x_o prin

$S(x_o, r) = \{x \in\in X : d(x_o, x) < r\}$.

Fie $A \subset\subset X$. Punctul $x_o \in\in X$ este *punct interior* mulțimii *A* dacă există $r_o > 0$ astfel încât $S(x_o, r_o) \subset\subset A$. Mulțimea punctelor interioare lui *A* formează *interiorul* lui A, notat cu $\dot{A}\dot{A}$. Punctul $x_o \in\in X$ este *punct aderent* lui *A* dacă, pentru orice $r > 0, S(x_0, r) \cap A = 0$ $r > 0, S(x_0, r) \cap A = 0$. Mulțimea punctelor aderente lui *A* formează *închiderea* lui A, notată $\bar{A}\bar{A}$. Un punct aderent lui *A* dar care nu este și punct interior lui *A* se numește *punct de frontieră*, mulțimea lor formând *frontiera* lui A, notată ½*A*. Au loc egalitățile:

½$A = \bar{A}\bar{A} \setminus \dot{A}\dot{A} = \bar{A} \cap \bar{A} \cap \overline{X \setminus A X \setminus A}$.

În cazul particular când *A* este o sferă, avem:

$\dot{S}(x_0, r_0) = S(x_0, r_0)$,

$\bar{S}\bar{S}(x_o, r_o) = \{x \in X : \in X : d(x_o, x) \leq\leq r_o\}$

$$\partial S(x_0, r_0) = \{x \in X: d(x_0, x) = r_0\}.$$

O submulțime $A \subset\subset X$ se numește *deschisă* dacă A = A, si se numește *’închisa* dacă $\dot{A}\dot{A}$= A. Se poate arăta că A este deschisă dacă și numai dacă complementara sa, $X \setminus A$, este închisă.

Clasa submulțimilor deschise definește o topologie pe X, numită *topologia indusa de metrică*. În această topologie, $V \subset \subset X$ este o *vecinătate* a punctului $x_o \in\in X$ dacă conține o sferă centrată în x_o. Prin definiție, un șir (x_n) din X este *convergent* la $x^* \in\in X$ dacă în orice vecinătate V a lui x^* se găsesc toți termenii șirului începând de la un rang $n_V \in \mathbb{N} \in \mathbb{N}$ încolo. În acest caz x^* se numește *limita* șirului și notăm $\lim_{n\to\infty} x_n \lim_{n\to\infty} x_n = x^*$. Convergența șirului de puncte (x_n) în spațiul metric X este caracterizată de șirul numeric al distanțelor dintre termenii șirului și limita sa. Mai precis, avem echivalența

$$\lim_{n\to\infty} x_n = x^* \Leftrightarrow \lim_{n\to\infty} d(x_n, x^*) = 0.$$

Un șir (x_n) din X se numește *șir fundamental* sau *șir Cauchy* dacă pentru orice $\varepsilon > 0$ există un rang $n_\varepsilon \in \mathbb{N}\; n_\varepsilon \in \mathbb{N}$ de la care începând distanța dintre oricare doi termeni ai șirului este mai mică decât ε. Pe scurt:

$$(x_n)\ \textit{șir Cauchy} \Leftrightarrow \varepsilon > 0\ \exists n_\varepsilon \in \mathbb{N}\ \textit{a.î.}\ n \geq n_\varepsilon \text{ și } m \geq n_\varepsilon \Rightarrow d(x_n, x_m) < \varepsilon.$$

În orice spațiu metric, dacă un șir este convergent este și fundamental, dar reciproca nu este valabilă în general. De exemplu, în mulțimea numerelor raționale dotată cu distanța obișnuită, șirul

$$x_n = \sum_{k=0}^{n} \frac{1}{10^n}$$

este fundamental fară să fie convergent.

Definiția 2. Numim *spațiu metric complet* un spațiu metric în care orice șir Cauchy este convergent.

Toate exemplele de spații metrice uzuale amintite mai sus, ($\mathbb{R}\mathbb{R}$, d), ($\mathbb{C}\mathbb{C}$,d), ($\mathbb{R}^n\mathbb{R}^n$,d) și ($C_{[a,b]}C_{[a,b]}$, d) sunt spații metrice complete.

Fie (X_1,d_1) și (X_2,d_2) două spații metrice. O funcție $f : X_1 — X_2$ este *continuă* dacă din $x_n \to\to x^*$ în X_1 rezultă că $f(x_n) \to\to f(x^*)$ în X_2.

Definiția 3. Funcția $f : X_1 \to\to X_2$ este *lipschitziana* cu *constanta Lipschitz* $L \geq\geq 0$ dacă pentru orice x și y din X_1 are loc majorarea

$$d_1(x,y) \leq L d_2(x,y).$$

O funcție lipschitziană cu constanta Lipschitz $L < 1$ se numește *contracție.*

Orice funcție lipschitziană este continuă, reciproca nu are loc în general.

Definiția 4. Fie $f : X \to\to X$ o funcție oarecare. Elementul $x^* \in\in X$ se numește *punct fix* al aplicației f dacă satisface egalitatea

$$f(x^*) = x^*.$$

Pentru a ușura înțelegerea Teoremei de punct fix a lui Banach, vom prezenta mai întâi metoda generală de rezolvare a ecuațiilor numerice prin aproximații succesive, așa cum apare ea pentru prima oară în scrierile rămase de la Heron din Alexandria (circa 10 - 70 d.H.) și anume chiar în *Metrica,* o culegere de formule și metode de calcul pentru lungimi, arii și volume, multe dintre ele preluate de la babilonieni. Printre acestea, și următoarea metodă de extragere a rădăcinei pătrate, formulată, în exemplul următor, în limbajul matematic actual.

Exemplul 1 (Heron). Ecuația

$$x^2 = a$$

poate fi rezolvată (în mulțimea numerelor reale strict pozitive) astfel: o scriem sub forma echivalentă

$$x = \frac{1}{2}\left(x + \frac{a}{x}\right),$$

și, pentru funcția $f : (0, +\infty\infty) \to\to (0, +\infty\infty)$ dată de

$$f(x) = \frac{1}{2}\left(x + \frac{a}{x}\right) \quad (1)$$

calculăm în mod recurent, pentru $n = 0, 1, 2, \ldots$, aproximațiile

$$x_{n+1} = f(x_n) x_{n+1} = f(x_n),$$

unde termenul initial $x_0 > 0\, x_0 > 0$ este ales arbitrar (dar cu cât mai aproape de $\sqrt{a}\, \sqrt{a}$, cu atât mai bine). Obținem un șir convergent la un număr $x^* \geq\geq 0$ care verifică ecuația $x^* = f(x^*)$, adică $x^* = \sqrt{a}\sqrt{a}$. Vom opri calculul efectiv al aproximațiilor x_n când observăm că s-au „stabilizat” un număr suficient de zecimale.

Justificarea convergentei șirului x_n este relativ simplă, ea poate fi stabilită prin studierea mărginirii și a monotoniei șirului. Noi revenim la principiul contracțiilor:

Teorema 1 (Banach). *Fie (X, d) un spațiu metric complet și **f** o contracție pe X cu constanta Lipschitz q. Atunci **f** are un singur un punct fix x^* în X. Mai mult, pentru orice punct inițial $x_o \in\in X$, șirul aproximațiilor succesive*

$$\begin{cases} x_0 \in X \\ x_{n+1} = f(x_n), n \in \mathbb{N}, \end{cases} \quad (2)$$

*este convergent la punctul fix x^* al lui **f**, viteza de convergență fiind dată de estimarea*

$$d(x_n, x^*) \leq \frac{d(x_0, x_1)}{1-q} q^n. \qquad (3)$$

Demonstrație. Considerăm un punct $x_o \in\in X$ fixat arbitrar și definim șirul (x_n) prin relația (2). Vom arăta, pentru început, că (x_n) este un șir Cauchy. Deoarece *f* este o contracție, avem, pentru orice $n \in \mathbb{N} n \in \mathbb{N}$,

$$d(x_{n+1}, x_{n+2}) = d(f(x_n), f(x_{n+1})) \leq q d(x_n, x_{n+1}),$$

de unde rezultă

$$d(x_n, x_{n+1}) \leq q^n d(x_0, x_1),$$

pentru orice $n \in \mathbb{N}. n \in \mathbb{N}.$ De aici obținem imediat, pentru orice $n, m \in \mathbb{N} \in \mathbb{N}$ cu $n \leq\leq m$,

$$d(x_n, x_m) \leq \sum_{i=n}^{m-1} d(x_i, x_{i+1}) \leq d(x_0, x_1) \sum_{i=n}^{m-1} q^i \leq \frac{d(x_0, x_1)}{1-q} q^n. \ (4)$$

Am folosit majorarea dată de suma seriei geometrice

$$\sum_{i=0}^{m-n-1} q^i \leq \sum_{i=0}^{\infty} q^i = \frac{1}{1-q},$$

care este convergentă deoarece constanta Lipschitz *q* este în intervalul [0,1).

Fie $\varepsilon\varepsilon > 0$ fixat arbitrar. Deoarece $\lim_{n\to\infty} \frac{d(x_0, x_1)}{1-q} q^n = 0$ rezultă că există un $n_\varepsilon \in \mathbb{N}\ n_\varepsilon \in \mathbb{N}$ astfel încât $n \geq n_\varepsilon n \geq n_\varepsilon$ implică $\frac{d(x_0, x_1)}{1-q} q^n < \varepsilon.$. Din (4) urmează că, pentru orice $n \geq n_\varepsilon n \geq n_\varepsilon$ și $m \geq n_\varepsilon m \geq n_\varepsilon$, avem $d(x_n, x_m) < \varepsilon\varepsilon$, și deci (x_n) este șir Cauchy. Spațiul metric *X* fiind complet, rezultă că (x_n) este convergent, adică exista $x^* \in\in X$ astfel încât

$$\lim_{n\to\infty} x_n = x^*.$$

În sfârșit, deoarece f este o contracție, este continuă, și trecând la limită în relația de recurență (2), obținem egalitățile

$$x^* = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(x^*),$$

care arată ca x^* este punct fix pentru f.

Pentru a demonstra unicitatea punctului fix, fie $x^{**} \in\in X$, cu $x^* \neq\neq$ x**, un alt punct fix al lui f. Atunci $d(x^*, x^{**}) > 0$ și obținem imediat că

$$d(x^*, x^{**}) = d(f(x^*), f(x^{**})) \quad q\, d(x^*, x^{**}) < d(x^*, x^{**}),$$

de unde rezultă o contradicție.

Estimarea (3) se obține din (4) prin trecere la limita cu $m \to \infty \to \infty$. □

Observația 1. Pentru șirul aproximaților succesive sunt valabile egalitățile:

$$x_1 = f(x_0),\ x_2 = f(f(x_0)) = (f \circ f)(x_0),\ x3 = f\left(f\left(f(x_0)\right)\right) = (f o f o f)(x_0)$$

ș.a.m.d. Definim *șirul iteratelor* functiei f ca fiind

$f^{\circ n} = f \circ f \circ \cdots \circ f$ (de n ori),

și avem

$$x_n = f^{\circ n}(x_0) x_n = f^{\circ n}(x_0),$$

pentru orice $n \geq\geq 1$, adică x_n este șirul valorilor în x_o ale iteratelor funcției.

Exemplul 2. Definim $f : \mathbb{C}\mathbb{C} \to \mathbb{C} \to \mathbb{C}$ prin

$$f(z) = az + i,$$

unde

$$a = \frac{9}{10}\left(\cos\frac{\pi}{7} + i \sin\frac{\pi}{7}\right).$$

Deoarece

$$|f(u) - f(v)| = |a||u - v| \leq \frac{9}{10}|u - v|,$$

pentru orice u și v din $\mathbb{C}\mathbb{C}$, f este o contracție. Unicul său punct fix este solutia ecuatiei $z = az + iz = az + i$, adică $z^* = \frac{i}{1-a}$

Găsirea unei contracții potrivite pentru rezolvarea unei anumite ecuații este, în general, o chestiune dificilă, de multe ori urmatoarea *variantă locală* este salvatoare: dacă știm că o aplicație $f : X \to X$ are un punct fix x^* undeva în spațiul metric complet (X, d), și dacă reușim să arătăm că există $r_o > 0$ și $q \in [0,1)$ astfel încât

$$d(f(x), f(y)) \leq qd(x,y),$$

pentru orice x și y din $\bar{S}(x^*, r_o)$, atunci rezultă imediat că f este o contracție pe $X_o = \bar{S}(x^*, r_o)$ și, prin urmare, pentru orice x_o suficient de aproape de x^* (adica $x_o \in X_o$) șirul valorilor în x_o ale iteratelor lui f este convergent la x^*.

În cazurile numerice $X = \mathbb{R}$ sau $X = \mathbb{C}$ avem chiar un rezultat mai precis:

Teorema 2. *Fie funcția* $f : \mathbb{R} \to \mathbb{R}$ (sau $f : \mathbb{C} \to \mathbb{C}$) *derivabilă cu derivata continuă și fie* $x^* \in \mathbb{R} \in \mathbb{R}$ *(respectiv* $x^* \in \mathbb{C} \in \mathbb{C}$) *un punct fix al său. Dacă*

$$|f'(x^*)| < 1,$$

atunci există $r_o > 0$ *astfel încât, pentru orice* x_o *cu* $|x_0 - x^*| \leq r_0$ *avem*

$$\lim_{n\to\infty} f^{\circ n}(x_o) = x^*.$$

Demonstrație. Fixăm o constantă q astfel încât $|f'(x^*)|$ $|f'(x^*)| < q < 1$. Din continuitatea derivatei în x^*, rezultă că există $r_o > 0$ astfel încât $|f'(x)| \leq |f'(x)| \leq q$ pentru orice x din $\bar{S}(x^*, r_o)$. Din teorema creșterilor finite urmează că $|f(x) - f(y)| < q|x - y||f(x) - f(y)| < q|x - y|$ pentru orice x și y din $\bar{S}(x^*, r_o)$, și aplicăm în continuare varianta locală a principiului contracțiilor. □

În cazul unei contracții $f : X \to X$, spunem că punctul său fix, $x^* \in X$, este un *atractor global,* deoarece are proprietatea că pentru orice $x_o \in X$ șirul iteratelor $(f^{\circ n}(x_o))$ converge la x^*. Se mai spune că x^* este *atractorul* contractiei f.

În cazul Teoremei 2 spunem că punctul fix x^* este un *atractor local,* deoarece convergența lui $(f^{\circ n}(x_o))$ la x^* este asigurată numai pentru x_o dintr-o vecinătate a punctului fix. Tot în cazul funcțiilor numerice, un punct fix x^* este numit *repulsor* dacă $|f'(x^*)| > 1|f'(x^*)| > 1$ și *neutru* dacă $|f'(x^*)| = 1|f'(x^*)| = 1$. În cazul când x^* este repulsor se poate arăta că pentru orice $x \neq\neq x^*$ dar suficient de apropriat de acesta, $d(x, x^*) < d(f(x), x^*, d(x, x^*) < d(f(x), x^*$, și, prin urmare, un șir de iterate $x_n = x_n = f^{\circ n}(x_o)$ poate să conveargă la x^* *numai dacă* $x_n = x_n =$ *x^* de la un loc încolo.*

Observația 2. *Din definiția derivatei rezultă imediat că, dacă o funcție lipschițiană definită pe* $\mathbb{R}\mathbb{R}$ sau $\mathbb{C}\mathbb{C}$ este derivabilă, atunci derivata sa este mărginită în modul de constanta Lipschitz a funcției. Reluând Exemplul 1, constatăm că funcția f dată de (1) nu este o contracție pe $X=(0,+\infty\infty)$, deoarece derivata sa

$$f'(x) = \frac{1}{2}\left(1 - \frac{a}{x^2}\right),$$

este nemărginită pentru $x \to 0x \to 0$. Totuși deoarece în punctul fix $x^* = \sqrt{a}\sqrt{a}$ avem $f'(\sqrt{a}) = 0f'(\sqrt{a}) = 0$, Teorema 2 este aplicabilă și justificăm foarte ușor convergența metodei lui Heron, dar numai pentru datele inițiale x_o „suficient de apropiate“ de $\sqrt{a}\sqrt{a}$.

Bibliografie:

[1] Postolică,V., *Eficiență prin Matematică aplicată*, Editura Matrix Rom, București, 2006

[2] Postolică, V., *Optimizare și aplicații.Spații abstracte,eficiență*, proiecții *imediate și de perspectivă*, Editura Matrix Rom, București, 2009

[3] Precupanu, A., *Bazele Analizei Matematice,* Editura Canova, Iași,1993