Ecuații Neliniare

Valentin-Ioan VINTILĂ

Facultatea de Automatică și Calculatoare - CTI
Universitatea POLITEHNICA București

11 aprilie 2023 (Lab. 7)

Cuprins

Seria Taylor

Vrem să reținem următoarea scriere a unei funcții $f : \mathbb{R} \to \mathbb{R}$ în jurul unui punct $x_0 \in \mathbb{R}$ dat:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots$$
$$= \sum_{n \geq 0} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

Dacă alegem $x_0 = 0$, dezvoltarea se numește **serie Maclaurin**.

Polinomul Taylor (1)

Cum trecem de la serie la polinom? Simplu, păstrăm doar o parte din *primii* termeni, să zicem $0 \dots n$:

$$\boxed{f(x) \approx P(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k}$$

Definim **restul** ca diferența între f și P:

$$R(x) = f(x) - P(x)$$

Polinomul Taylor (2)

Acceptăm, fără demonstrație, faptul că:

$$R(x) = \sum_{k > n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \leq \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

unde $\left|f^{(n+1)}(\xi)\right|$ este maxim, iar $\xi \in [\min\{x_0, x\}, \max\{x_0, x\}]$.

Puncte fixe (1)

Definiție (punct fix)

Fie $g : D \subseteq \mathbb{R} \to \mathbb{R}$ o funcție oarecare. Fiecare valoare $p \in D$ pentru care evaluarea funcției g întoarce tot p (adică are loc $g(p) = p$) poartă denumirea de **punct fix**.

Puncte fixe (2)

Teoremă (Banach) – sinteză

Fie $g : [a, b] \subseteq \mathbb{R} \to [a, b]$ o funcție continuă pe întreg intervalul $[a, b]$. Atunci, funcția g va avea *cel puțin un punct fix* în intervalul $[a, b]$. Suplimentar, dacă funcția g este derivabilă pe intervalul (a, b) și $|g'(x)| < 1$, $\forall x \in [a, b]$, atunci g are un **unic punct fix atractiv** în $[a, b]$.

Puncte fixe (3)

Care e legătura cu soluționarea ecuațiilor neliniare?

Fie $g : D \subseteq \mathbb{R} \to \mathbb{R}$ astfel încât să admită toate punctele fixe din mulțimea $P = \{p_1, p_2, \dots, p_n\}$, $n \in \mathbb{N}$.

Definim acum funcția $f(x) = x - g(x)$. **Ce observăm?**

$$f(x) = x - g(x) = \begin{cases} 0, & \text{dacă } x \in P \\ \text{o valoare nenulă}, & \text{dacă } x \notin P \end{cases}$$

Echivalent, studiind $g(x) = x - f(x)$:

$$g(x) = x - f(x) = \begin{cases} x, & \text{dacă } x \in P \\ \text{o valoare diferită de } x, & \text{dacă } x \notin P \end{cases}$$

Metoda aproximațiilor succesive I

Fie $g : D \subseteq \mathbb{R} \to \mathbb{R}$ o funcție ce admite toate punctele fixe din mulțimea $P = \{p_1, p_2, \dots, p_n\}$, $n \in \mathbb{N}$.

Să considerăm șirul $(x_m)_{m \in \mathbb{N}}$:

$$\boxed{x_k = \begin{cases} \text{o aproximare inițială}, & k = 0 \\ g(x_{k-1}), & k \geq 1 \end{cases}}$$

Către ce converge acest șir?

$$p = \lim_{k\to\infty} x_k = \lim_{k\to\infty} g(x_{k-1}) = g(\lim_{k\to\infty} x_{k-1}) = g(p)$$

Metoda aproximațiilor succesive I (2)

Concluzionăm că $p = \lim_{k\to\infty} x_k$ este **punct fix**.

Deci, pentru a soluționa $f(x) = 0$, ne vom crea noi funcția $g(x) = x - f(x)$ pe care aplicăm șirul anterior.

Se garantează că metoda este convergentă către un punct fix atractiv. *Puteți demonstra ca temă...*

Metoda bisecției (1)

Metoda aceasta este mult mai interesantă grafic; se aseamănă cu căutarea binară!

Pornim de la o funcție continuă $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$ astfel încât $\operatorname{sign}(f(a)) \neq \operatorname{sign}(f(b))$.

Metoda bisecției (2)

Fie c media aritmetică dintre a și b: $c = \dfrac{a+b}{2}$

Avem trei cazuri:

- $f(c) \approx 0$ – **am găsit soluția!**
- $\operatorname{sign}(f(a)) \neq \operatorname{sign}(f(c))$ – datorită continuității, știm că va exista cel puțin un $\delta \in (a, c)$ astfel încât $f(\delta) = 0$.

  Reducem așadar intervalul: $[a, b] \to [a, c]$;
- $\operatorname{sign}(f(b)) \neq \operatorname{sign}(f(c))$ – datorită continuității, știm că va exista cel puțin un $\delta \in (c, b)$ astfel încât $f(\delta) = 0$.

  Reducem așadar intervalul: $[a, b] \to [c, b]$;

Metoda bisecției – exemplu (1)

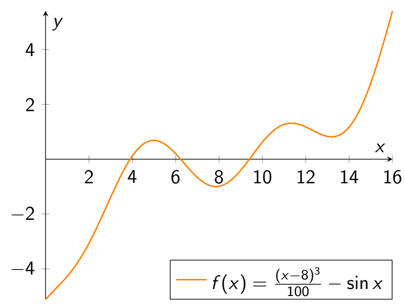
Hai să vizualizăm puțin algoritmul!

Să considerăm funcția:

$$f(x) = \frac{(x-8)^3}{100} - \sin x$$

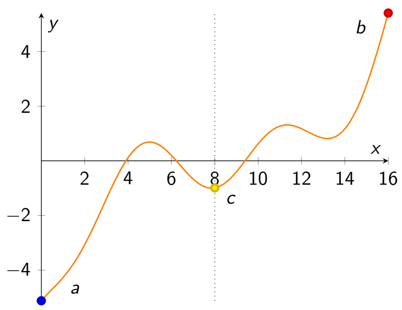
Vom fi interesați de intervalul $[a, b] = [0, 16]$.

Metoda bisecției – exemplu (2)



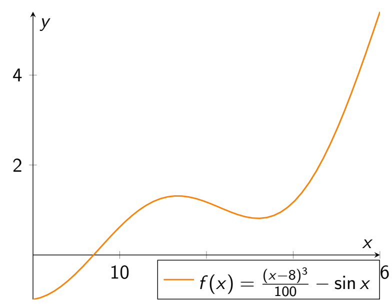
Graficul funcției $f(x)$ pe intervalul $[0, 16]$

Metoda bisecției – exemplu (3)



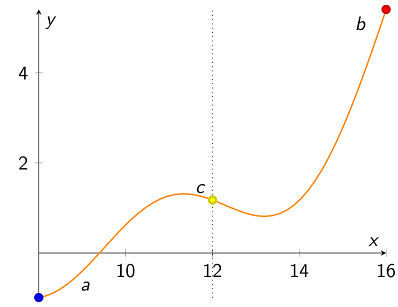
Graficul funcției $f(x)$ cu punctele de abscisă $a = 0$, $b = 16$ și $c = \frac{a+b}{2} = 8$ evidențiate

Metoda bisecției – exemplu (4)



Graficul funcției $f(x)$ pe intervalul $[8, 16]$

Metoda bisecției – exemplu (5)



Graficul funcției $f(x)$ cu punctele de abscisă $a = 8$, $b = 16$ și $c = \frac{a+b}{2} = 12$ evidențiate

Metoda bisecției – număr de iterații

Observăm cum, ușor-ușor, ne-am apropiat de $f(x) = 0$.

Câte iterații vom face? Presupunem că suntem mulțumiți cu o anumită eroare relativă dată de formula $\left|\frac{c_{nou} - c_{vechi}}{c_{nou}}\right| \leq \varepsilon$.

$$\boxed{n = \left\lceil \log_2 \left(\frac{b-a}{\varepsilon} \right) \right\rceil}$$

Ordin și rată de convergență (1)

Fie $(x_n)_n \in \mathbb{N} \in \mathbb{R}$ o secvență care converge către o valoare $x^* \in \mathbb{R}$, adică $x_n \to x^*$ atunci când $n \to +\infty$.

În această situație, vom nota cu $q \in [1, +\infty)$ **ordinul de convergență**, respectiv cu $\mu \in [0, 1]$ **rata de convergență** a secvenței (algoritmului).

Aceste valori pot fi determinate prin formula:

$$\boxed{\mu = \lim_{n \to +\infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^q}}$$

Ordin și rată de convergență (2)

Dacă ne interesează mai mult ordinul de convergență decât rata, putem folosi o aproximare practică:

$$\boxed{q \approx \frac{\ln \left| \frac{x_{k+1} - x_k}{x_k - x_{k-1}} \right|}{\ln \left| \frac{x_k - x_{k-1}}{x_{k-1} - x_{k-2}} \right|}}, \ \forall k \in \mathbb{N}$$

- Pentru $q = 1$, șirul (algoritmul) converge **liniar**;
- Pentru $q = 2$, șirul (algoritmul) converge **quadric**;
- Pentru $q = 3$, șirul (algoritmul) converge **cubic**.

Metoda bisecției – convergență

Pentru metoda bisecției, se pot calcula:

- Ordinul de convergență $q = 1$ (convergență **liniară**);
- Rata de convergență $\mu = \frac{1}{2}$.

Puteți face demonstrația ca temă...

Metoda tangentei (1)

Această metodă este cunoscută și drept **Newton-Raphson**.

Totul pleacă de la **polinoamele Taylor** de grad 2.

Fie $f : [a, b] \subseteq \mathbb{R} \to \mathbb{R}$ o funcție continuă și derivabilă pe intervalul $[a, b]$. Vom dezvolta această funcție în jurul lui $x_0 \in [a, b]$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots$$

Metoda tangentei (2)

Fie $p \in [a, b]$ astfel încât $f(p) = 0$. Dacă ne asumăm că $(p - x_0)^2 \to 0$, atunci:

$$f(p) \approx f(x_0) + f'(x_0)(p - x_0)$$

Am pornit însă de la $f(p) = 0$, deci:

$$0 \approx f(x_0) + f'(x_0)(p - x_0) \Rightarrow \boxed{p \approx x_0 - \frac{f(x_0)}{f'(x_0)}}$$

Dar nu ne forțează nimeni să folosim doar x_0, acea valoare se poate modifica în timp!

Metoda tangentei (3)

Definim așadar **recurența metodei Newton-Raphson**:

$$\boxed{x_k = \begin{cases} \text{o aproximație bună inițială}, & k = 0 \\ x_{k-1} - \dfrac{f(x_{k-1})}{f'(x_{k-1})}, & k \geq 1 \end{cases}}$$

Când $k \to \infty$, $p = x_k$.

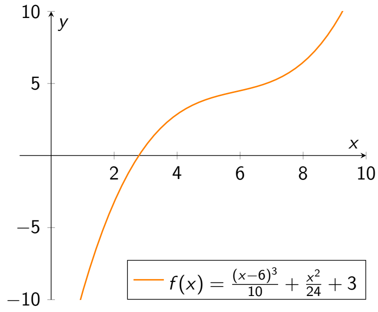
Metoda tangentei – exemplu (1)

Să considerăm funcția:

$$f(x) = \frac{(x-6)^3}{10} + \frac{x^2}{24} + 3$$

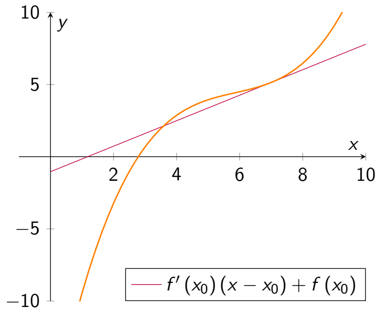
Vom începe aproximarea din $x_0 = 7$.

Metoda tangentei – exemplu (2)



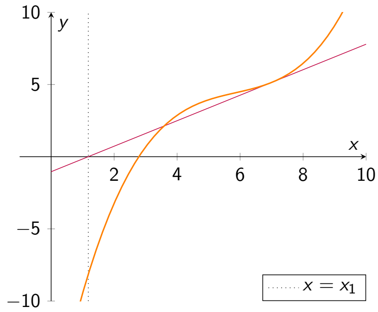
Graficul funcției $f(x)$

Metoda tangentei – exemplu (3)



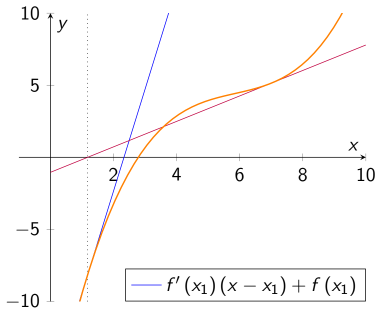
Graficul funcției $f(x)$. Tangenta în $x_0 = 7$ la graficul lui $f(x)$

Metoda tangentei – exemplu (4)



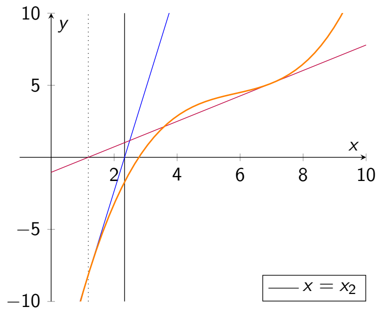
Paralela cu Oy care trece prin x_1

Metoda tangentei – exemplu (5)



Tangenta în $x_1 \approx 1.179245$ la graficul lui $f(x)$

Metoda tangentei – exemplu (6)



Paralela cu Oy care trece prin $x_2 \approx 2.331315$

Metoda tangentei – convergență

Această metodă converge în aproximativ $3 - 5$ pași, având ordinul de convergență egal cu $q = 2$ (converge **quadric**).

Cu toate acestea, metoda **poate să nu conveargă** în situația în care:

- La un anumit pas $k \in \mathbb{N}$, **derivata este (numeric)** nulă, adică $f'(x_k) \approx 0$;
- x_0 este prea departe de p.

Metoda secantei (1)

Pornim de la recurența Newton-Raphson:

$$x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})}$$

Cum putem calcula derivata aceea numeric?

Firește, vom porni de la acea formulă din liceu:

$$f'(\tau) = \lim_{h \to 0} \frac{f(\tau + h) - f(\tau)}{h}$$

Metoda secantei (2)

Putem deci să aproximăm:

$$f'(\tau) \approx \frac{f(\tau + h) - f(\tau)}{h}, \text{ unde } h \to 0$$

Ce știm însă despre recurența Newton-Raphson? Metoda converge, ceea ce înseamnă că putem accepta:

$$h = x_{k-1} - x_{k-2}$$

Rezultă așadar că, dacă $\tau = x_{k-1}$, atunci:

$$\boxed{f'(x_{k-1}) \approx \frac{f(x_{k-1}) - f(x_{k-2})}{x_{k-1} - x_{k-2}}}, \ \forall k \geq 2$$

Metoda secantei (3)

Folosind formula anterioară, putem modifica recurența din metoda Newton-Raphson pentru a ajunge la metoda secantei:

$$\boxed{x_k = \begin{cases} \text{o aproximație bună inițială}, & k \in \{0, 1\} \\ x_{k-1} - f(x_{k-1}) \cdot \dfrac{x_{k-1} - x_{k-2}}{f(x_{k-1}) - f(x_{k-2})}, & k \geq 2 \end{cases}}$$

Evident, când $k \to \infty$, $p = x_k$.

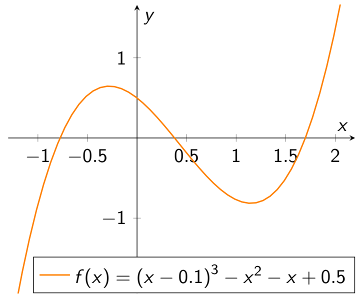
Metoda secantei – exemplu (1)

Să considerăm funcția:

$$f(x) = (x - 0.1)^3 - x^2 - x + 0.5$$

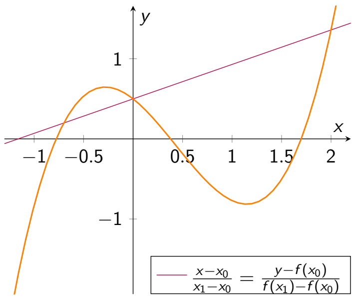
Vom începe aproximarea din $x_0 = -2$ și $x_1 = 2$.

Metoda secantei – exemplu (2)



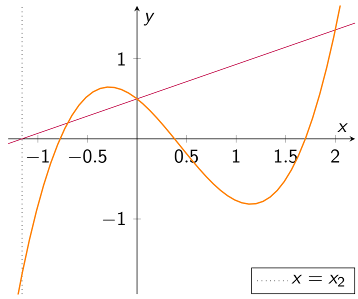
Graficul funcției $f(x)$

Metoda secantei – exemplu (3)



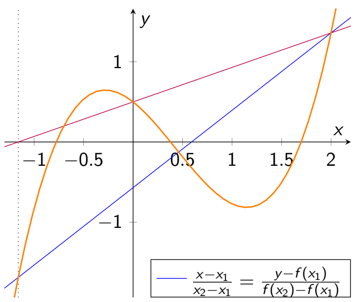
Dreapta care trece prin $(x_0, f(x_0))$ și $(x_1, f(x_1))$

Metoda secantei – exemplu (4)



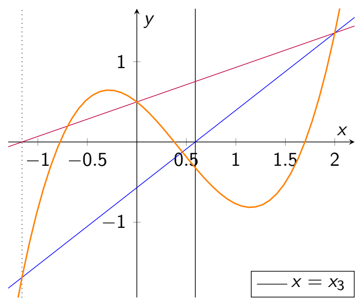
Paralela la Oy care trece prin $x_2 \approx -1.16$

Metoda secantei – exemplu (5)



Dreapta care trece prin $(x_1, f(x_1))$ și $(x_2, f(x_2))$

Metoda secantei – exemplu (6)



Paralela cu Oy care trece prin $x_3 \approx 0.590766$

Metoda secantei – convergență

Această metodă are ordinul de convergență egal cu $q = \varphi \approx 1.618$ **(proporția de aur, $\varphi = \frac{1+\sqrt{5}}{2}$)**.

Avem ordin de convergență mai slab, dar am scăpat de derivate.

Puncte fixe

Definiție (punct fix)

Fie $g : D \subseteq \mathbb{R}^n \to \mathbb{R}$ o funcție oarecare ($n \in \mathbb{N}^*$). Fiecare valoare $p \in D$ pentru care evaluarea funcției g întoarce tot p (adică are loc $g(p) = p$) poartă denumirea de **punct fix**.

Sisteme drept funcții

Cum putem vedea un sistem de ecuații neliniare drept o singură funcție de mai multe variabile?

$$\begin{cases} f_1(\boldsymbol{x}) = 0 \\ f_2(\boldsymbol{x}) = 0 \\ \vdots \\ f_n(\boldsymbol{x}) = 0 \end{cases} \Rightarrow F(\boldsymbol{x}) = \begin{bmatrix} f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \\ \vdots \\ f_n(\boldsymbol{x}) \end{bmatrix} \text{ și } F(\boldsymbol{x}) = \boldsymbol{0}$$

unde $F : \mathbb{R}^n \to \mathbb{R}^n$.

Metoda aproximațiilor succesive II

Fie $G : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ o funcție ce admite toate punctele fixe din mulțimea $P = \{\boldsymbol{p_1}, \boldsymbol{p_2}, \dots, \boldsymbol{p_n}\}$, $n \in \mathbb{N}$.

Să considerăm șirul $(\boldsymbol{x_m})_{m \in \mathbb{N}}$:

$$\boxed{\boldsymbol{x_k} = \begin{cases} \text{o aproximare inițială}, & k = 0 \\ G(\boldsymbol{x_{k-1}}), & k \geq 1 \end{cases}}$$

Asemănător, pentru a soluționa $F(\boldsymbol{x}) = \boldsymbol{0}$, ne vom crea noi funcția $G(\boldsymbol{x}) = \boldsymbol{x} - F(\boldsymbol{x})$ pe care aplicăm șirul anterior.

Matrice Jacobiană (1)

Să amintim o formulă pentru matricea Jacobiană $J(\boldsymbol{x}) \in \mathbb{R}^{n \times n}$ a unei funcții oarecare $F : \mathbb{R}^n \to \mathbb{R}$:

$$F(x_1, x_2, \dots, x_n) = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) \end{bmatrix}$$

$$\Rightarrow J(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{x}) & \dots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{x}) \\ \frac{\partial f_2}{\partial x_2}(\boldsymbol{x}) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{x}) & \dots & \frac{\partial f_2}{\partial x_n}(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_n}{\partial x_2}(\boldsymbol{x}) & \dots & \frac{\partial f_n}{\partial x_n}(\boldsymbol{x}) \end{bmatrix}, \boldsymbol{x} \in \mathbb{R}^n$$

Matrice Jacobiană (2)

La ce e bună matricea Jacobiană?

Generalizează oarecum derivatele pentru funcții de mai multe variabile care întorc vectori.

Ce metodă utiliza derivate?

Metoda **Newton-Raphson** – deci o putem generaliza acum pentru a rezolva sisteme de ecuații neliniare.

Metoda Newton (1)

Să amintim recurența de la Newton–Raphson:

$$x_k = \begin{cases} \text{o aproximație bună inițială}, & k = 0 \\ x_{k-1} - [f'(x_{k-1})]^{-1} \cdot f(x_{k-1}), & k \geq 1 \end{cases}$$

Înlocuind f cu F, respectiv f' cu $J(\boldsymbol{x})$, obținem **metoda Newton**:

$$\boxed{\boldsymbol{x_k} = \begin{cases} \text{o aproximație bună inițială}, & k = 0 \\ \boldsymbol{x_{k-1}} - [J(\boldsymbol{x_{k-1}})]^{-1} \cdot F(\boldsymbol{x_{k-1}}), & k \geq 1 \end{cases}}$$

Metoda Newton (2)

$$\boxed{\boldsymbol{x_k} = \begin{cases} \text{o aproximație bună inițială}, & k = 0 \\ \boldsymbol{x_{k-1}} - [J(\boldsymbol{x_{k-1}})]^{-1} \cdot F(\boldsymbol{x_{k-1}}), & k \geq 1 \end{cases}}$$

Ce problemă are această formulă? Vrem să scăpăm de inversa matricei Jacobiene!

Vom căuta un vector $\boldsymbol{y} \in \mathbb{R}^n$ astfel încât:

$$J(\boldsymbol{x}) \cdot \boldsymbol{y} = -F(\boldsymbol{x}) \Rightarrow \boxed{\boldsymbol{y} = -J(\boldsymbol{x})^{-1} \cdot F(\boldsymbol{x})}$$

Formula finală devine:

$$\boxed{\boldsymbol{x_k} = \begin{cases} \text{o aproximație bună inițială}, & k = 0 \\ \boldsymbol{x_{k-1}} + \boldsymbol{y}, & k \geq 1 \end{cases}}$$

Bibliografie

Metodele sunt bine descrise în curs + Wikipedia.

Drept resurse auxiliare, am utilizat:

1. Articolul lui **Paul Seeburger** despre polinoame Taylor;
2. Cartea lui **Alexandru Negrescu**, *Calcul Diferențial - O abordare prietenoasă*.

Sfârșit

Mulțumesc frumos pentru atenție!

Vă rog frumos să **completați formularul de feedback**!