

Metode numerice (notițe curs)

– UPB 2020

I. Introducere

Determinarea practică a soluțiilor matematice al unor probleme științifice și ingineresti necesită, în cele mai multe cazuri, construirea unui algoritm pentru găsirea lor, adaptat posibilităților de lucru ale sistemului de calcul numeric.

Astfel, **algoritm** reprezintă totalitatea raționamentelor matematice și logice, efectuate pas cu pas pentru rezolvarea problemei, deci o rețetă generală, precis formulată, care conduce la rezultate concrete într-un interval finit de timp.

Formularea matematică (modelul) a unei probleme fizice sau ingineresti conține inevitabil aproximări datorate:

- înțelegerii incomplete a fenomenelor naturale sau a naturii aleatoare ale acestora;
- descrierii matematice exclusive a trăsăturilor esențiale ale procesului fizic, prin neglijarea deliberată a detaliilor legate de efectele secundare

Erorilor asociate acestor aproximări li se adaugă cele rezultate din prelucrarea datelor cu ajutorul unui sistem de calcul numeric. Acesta poate realiza numai un număr limitat de operații aritmetice simple (adunare, scădere, împărțire, înmulțire, ridicare la putere) cu numere finite raționale. Operații importante ca derivarea, integrarea, evaluarea seriilor infinite nu pot fi implementate direct pe un calculator numeric. Acesta este compus din registre de calcul și memorii cu capacitate limitată, încât este imposibil să se reprezinte cantitățile infinit mici sau infinit mari, ori un interval finit de numere reale.

Un algoritm adaptat posibilităților calculatorului numeric, utilizând numai operații aritmetice și unele operații logice se spune că utilizează **Metode Numerice**.

Lucrarea tratează, în continuare, prin această prismă, principalele capitole matematice utilizate în inginerie.

II. Erori

Eroarea absolută, e_x , este definită ca fiind diferența dintre valoarea adevărată (exactă), x și valoarea aproximativă, \bar{x} (determinată prin calcul sau prin măsurare):

$$e_x = x - \bar{x} \quad (\text{II.1})$$

Dintre cele 3 mărimi numai valoarea aproximativă este cunoscută de obicei, uneori cunoscându-se și marginea erorii (modulul maxim probabil).

Eroarea relativă este eroarea absolută împărțită la valoarea aproximativă (prin convenție), pentru că aceasta se cunoaște în general.

$$\varepsilon_x = e_x / \bar{x} \quad (\text{II.2})$$

II.1 Erori inițiale

Erorile inițiale sunt erori din datele inițiale, cauzate de incertitudinea măsurărilor, din greșeli (grosolane) sau din reprezentarea unei valori transcendente sau iraționale printr-un număr finit de cifre.

Exemple:

- a) o măsurătoare fizică nu poate fi exactă; instrumentul folosit transferă precizia sa mărimii măsurate; în cazul dimensiunilor liniare, un șubler indică cu certitudine 0,1 mm, un micrometru sau un comparator 0,01 mm, iar un traductor capacitiv ajunge la o rezoluție nanometrică. Numărul de cifre semnificative corecte variază deci în funcție de precizia măsurătorii, recomandându-se limitelor toleranței mărimii.
- b) multe numere nu pot fi reprezentate printr-un număr finit de cifre. Dacă exemplele lui π sau $\sqrt{2}$ sunt evidente, o simplă fracție poate să nu aibă o reprezentare exactă, în funcție de baza de numerație folosită. În sistemul zecimal, numărul $1/10$ este 0,1, în timp ce în sistemul binar el este 0,0001100110011....., deci o reprezentare infinită imposibilă în calculator. Deci, dacă se face suma a 10 numere reprezentând aproximații binare ale lui 0,1, nu se va obține 1,0.

II.2 Erori de trunchiere

O procedură numerică poate introduce ea însăși erori. De exemplu, pentru calculul sinusului unui unghi (în radiani) se poate utiliza seria Maclaurin:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

Aceasta este infinită și nu se pot utiliza toți termenii seriei în calcul. Limitarea expresiei la un număr finit de termeni conduce la o eroare reprezentată de termenii omiși, numită eroare de trunchiere.

Deși nu se poate calcula precis, eroarea de trunchiere se poate estima prin metode specifice fiecărui algoritm, limitele acestora fiind puse în evidență la fiecare problemă analizată.

II.3 Erori de calcul

II.3.1. Calculul în virgulă mobilă

La reprezentarea numerelor în dispozitive cu capacitate limitată de memorare și prelucrare este specifică tăierea cifrelor ne semnificative de la sfârșitul acestora, care poate apărea atât la introducerea numărului, cât și în urma efectuării unor operații aritmetice. În unele cazuri particulare, eroarea rezultată prin tăiere poate avea valori importante.

Rezolvarea limitării tehnice a reprezentării numerelor reale în calculator o constituie utilizarea numerelor în virgulă mobilă. Acestea sunt compuse dintr-o fracție numită mantisă și un întreg numit exponent sau caracteristică.

$$N = m \cdot b^c \quad (\text{II.3})$$

Unde: m = mantisă, c = exponent, b = baza de numerație.

Exemplu: În baza 10, numărul $N=36,82$ se va rescrie $N=0,3682 \cdot 10^2$. El este normalizat pentru că prima cifră după virgulă a mantisei este diferită de zero, adică:

$$\frac{1}{10} \leq |m| < 1 \quad (\text{II.4})$$

Procedeul de adunare a două numere reale în virgulă mobilă va fi ilustrat pe un calculator ipotetic, zecimal cu patru poziții afectate mantisei pentru numerele a) $x_1=165,2$; $x_2=0,21$ și respectiv b) $x_3=26,31$; $x_4=-19,76$.

Etapele procedurii sunt:

1) normalizarea numerelor:
a) $x_1=0,1652 \cdot 10^3$; $x_2=0,2100 \cdot 10^0$

b) $x_3=0,2631 \cdot 10^2$; $x_4=-0,1976 \cdot 10^2$

2) compararea exponenților:
a) $c_1=3 > c_2=0$

b) $c_3=c_4=2$

3) de-normalizarea numărului cu exponent mai mic prin deplasarea mantisei spre dreapta virgulei, cu atâtea poziții cât reprezintă diferența exponenților:
a) $x_2=0,2100 \cdot 10^0 = 0,0002 \cdot 10^3$

Prin de-normalizare se pot pierde cifre semnificative ale numărului.

- 4) adunare propriu-zisă
 a) $S_{12}=x_1+x_2=0,1652 \cdot 10^3+0,0002 \cdot 10^3=0,1654 \cdot 10^3$
 b) $S_{34}=x_3+x_4=0,2632 \cdot 10^2+(-0,1976) \cdot 10^2=0,0655 \cdot 10^2$
- 5) re-normalizarea rezultatului (dacă este cazul)
 b) $S_{34}=0,6550 \cdot 10^1$

În exemplul a, suma este afectată de tăierea efectuată la pasul 3, iar în exemplul b ultima cifră (0) nu mai are nici o semnificație, ea apărând în procesul deplasării. Aceste erori se propagă în următoarele operații, lor adăugându-li-se altele, specifice fiecărei etape. Tăierea ultimei cifre este evitată la unele sisteme de calcul printr-o poziție pentru așa-zisa cifră de gardă (în exemplul dat o a cincea cifră în mantisa calculatorului ipotetic), permițând salvarea unei cifre din mantisa deplasată spre dreapta.

Probleme deosebite la calculul în virgulă mobilă mai pot apare în următoarele cazuri:

- a) scăderea a două numere aproximativ egale cu exponent minim conduce la depășire inferioară în virgulă mobilă (*underflow*).

Exemplu: convenindu-se că exponentul minim reprezentabil în calculatorul ipotetic este

$$-615, \text{ atunci: } 0,1459 \cdot 10^{-615} - 0,1024 \cdot 10^{-615} = 0,0425 \cdot 10^{-615} = 0,4250 \cdot 10^{-616}.$$

Diverse sisteme de calcul rezolvă diferit această problemă, dar, de obicei, atribuie rezultatului valoarea zero în virgulă mobilă, ceea ce poate conduce la rezultate foarte ciudate cu un algoritm aparent corect.

- b) adunarea a două numere cu exponent maxim poate conduce la depășire superioară în virgulă mobilă (*overflow*).

Exemplu: Dacă exponentul maxi reprezentabil în calculatorul ipotetic este 615, atunci $0,9999 \cdot 10^{615} + 0,9999 \cdot 10^{615} = 1,9998 \cdot 10^{615} = 0,1999 \cdot 10^{616}$

De obicei, depășirea exponentului maxim reprezentabil în calculator este semnalată prin întreruperea execuției programului.

La înmulțire/împărțire mantisele normalizate se înmulțesc/împart, iar exponenții se adună/scad, normalizarea rezultatului fiind simplă. Depășirea superioară poate apare atât la înmulțire, dar și la împărțire prin zero (când împărțitorul poate fi rezultatul unei depășiri inferioare).

În mod normal, calculatoarele numerice lucrează în baza 2 sau 16, încât reprezentarea numerelor în virgulă mobilă este:

$$x = m \cdot 2^c \quad \text{cu } \frac{1}{2} \leq |m| < 1, \text{ sau} \quad (\text{II.5})$$

$$x = m \cdot 16^c \quad \text{cu } \frac{1}{16} \leq |m| < 1. \quad (\text{II.6})$$

II.3.2. Erori de rotunjire

Rezultatul oricăreia dintre cele 4 operații, într-o reprezentare a numerelor cu t cifre semnificative pentru mantisă, poate fi scris ca:

$$x = m_x \cdot 10^c + p_x \cdot 10^{c-t},$$

unde p_x este partea din mantisă asupra căreia se acționează la rotunjire, care nu este normalizată, deci $0 \leq p_x < 1$.

Unele compilatoare pentru limbaje de nivel înalt folosesc tăierea termenului $p_x \cdot 10^{c-t}$, atribuind rezultatului valoarea $m_x \cdot 10^c$.

În acest caz, modulul erorii de rotunjire maxime este:

$$\frac{e_x}{x} = \left| \frac{p_x \cdot 10^{c-t}}{m_x \cdot 10^c} \right| \leq \frac{1 \cdot 10^{c-t}}{0,1 \cdot 10^c} = 10^{-t+1} \quad (\text{II.7})$$

În consecință, maximul erorii relative de rotunjire prin tăiere nu depinde de mărimea numărului, ci de numărul de cifre semnificative reprezentate în mantisă (t).

Un alt tip de rotunjire uzuală este cea simetrică, aproximarea numărului x , fiind:

$$\bar{x} = \begin{cases} |m_x| \cdot 10^c, & \text{daca } |p_x| < \frac{1}{2} \\ |m_x| \cdot 10^c + 10^{c-t}, & \text{daca } |p_x| \geq \frac{1}{2} \end{cases} \quad (\text{II.8})$$

unde \bar{x} și m_x sunt de același semn.

Eroarea absolută maximă este:

$$|e_x| = \begin{cases} |p_x| \cdot 10^{c-t} \leq \frac{1}{2} \cdot 10^{c-t}, & \text{daca } |p_x| < \frac{1}{2} \\ |1 - p_x| \cdot 10^{c-t} \leq \frac{1}{2} \cdot 10^{c-t}, & \text{daca } |p_x| \geq \frac{1}{2} \end{cases}$$

Valoarea absolută a erorii relative maxime de rotunjire simetrică este:

$$\left| \frac{e_x}{x} \right| \leq \left| \frac{\frac{1}{2} \cdot 10^{c-t}}{m_x \cdot 10^c} \right| \leq \frac{\frac{1}{2} \cdot 10^{c-t}}{0,1 \cdot 10^c} = 5 \cdot 10^{-t},$$

deci cel mult jumătate din cea obținută la rotunjirea prin tăiere.

II.4. Propagarea erorilor

S-a constatat că, în cazuri extreme, precum scăderea a două numere foarte mici, eroarea relativă a diferenței poate fi destul de mare, propagându-se prin orice operație aritmetică ulterioară.

De aceea expresiile erorilor relative și absolute ale rezultatului unei operații în funcție de erorile operanzilor sunt foarte importante. Notându-se $x = \bar{x} + e_x$ și $y = \bar{y} + e_y$, se determină erorile absolute pentru fiecare operație:

a) adunarea (scăderea)

$$e_{x \pm y} = (x \pm y) - (\bar{x} \pm \bar{y}) = e_x \pm e_y \quad (\text{II.9})$$

b) înmulțirea

$$e_{x \cdot y} = x \cdot y - \bar{x} \cdot \bar{y} = (\bar{x} + e_x) \cdot (\bar{y} + e_y) - \bar{x} \cdot \bar{y} = \bar{x} e_y + \bar{y} e_x + e_x \cdot e_y \cong \bar{x} e_y + \bar{y} e_x \quad (\text{II.10})$$

c) împărțirea

$$e_{x/y} = x/y - \bar{x}/\bar{y} = (\bar{x} + e_x)/(\bar{y} + e_y) - \bar{x}/\bar{y} = \frac{\bar{x} + e_x}{\bar{y}} \cdot \frac{1}{1 + \frac{e_y}{\bar{y}}} - \frac{\bar{x}}{\bar{y}} =$$

$$= \frac{\bar{x} + e_x}{\bar{y}} \cdot \left[1 - \frac{e_y}{\bar{y}} + \left(\frac{e_y}{\bar{y}} \right)^2 - \dots \right] - \frac{\bar{x}}{\bar{y}} \cong \frac{e_x}{\bar{y}} - \frac{\bar{x}}{\bar{y}^2} \cdot e_y \quad (\text{II.11})$$

Relațiile deduse specifică valoarea maximă posibilă a erorii rezultatului operației respective. Cum semnul erorilor operanzilor este rareori cunoscut, rezultă că nu întotdeauna adunarea mărește eroarea, iar scăderea o micșorează așa cum s-ar putea deduce la o analiză superficială a expresiilor.

Erorile relative derivă din cele absolute astfel:

a) adunare

$$\frac{e_{x+y}}{x+y} = \frac{\bar{x}}{x+y} \cdot \frac{e_x}{x} + \frac{\bar{y}}{x+y} \cdot \frac{e_y}{y} \quad (\text{II.12})$$

b) scădere

$$\frac{e_{x-y}}{x-y} = \frac{\bar{x}}{x-y} \cdot \frac{e_x}{x} - \frac{\bar{y}}{x-y} \cdot \frac{e_y}{y} \quad (\text{II.13})$$

c) înmulțire

$$\frac{e_{x \cdot y}}{x \cdot y} = \frac{e_x}{x} + \frac{e_y}{y} \quad (\text{II.14})$$

d) împărțire

$$\frac{e_{x/y}}{x/y} = \frac{e_x}{x} - \frac{e_y}{y} \quad (\text{II.15})$$

În formulele de propagare erorile pot fi de orice tip, termenii x și y putând fi obținuți în urma unor măsurători (cu erori inițiale) sau a unor calcule anterioare (cu erori de trunchiere, tăiere sau rotunjire). Chiar și în cazul operațiilor cu numere lipsite de erori rezultatul va fi afectat de rotunjire.

Exemplu: $u = (x+y) \cdot z$, unde x, y, z nu au erori.

În urma efectuării adunării va apare o eroare de rotunjire simetrică:

$$\left| \frac{e_{x+y}}{x+y} \right| \leq \frac{1}{2} \cdot 10^{-t+1}$$

La produs se va adăuga sumei erorilor relative ale factorilor și eroarea de rotunjire r_m specifică operației.

$$\left| \frac{e_u}{u} \right| = \left| \frac{e_{x+y}}{x+y} + \frac{e_z}{z} + r_m \right| \leq \left| \frac{e_{x+y}}{x+y} \right| + |r_m| \leq \frac{1}{2} \cdot 10^{-t+1} + \frac{1}{2} \cdot 10^{-t+1} = 10^{-t+1}$$

Unde $e_z=0$.

Marginea erorii absolute a rezultatului va fi:

$$|e_u| \leq |u| \cdot 10^{-t+1}$$

Reguli pentru precizia calculelor

Sucesiunea operațiilor, adoptată prin algoritmul de calcul, poate avea o influență decisivă asupra mărimii erorii rezultatului final. Astfel:

- 1) Când se adună și/sau scad mai multe numere se va începe cu acelea mai mici
- 2) Scăderea a două numere aproximativ egale se va evita prin rescrierea expresiei (dacă este posibil)
- 3) Algoritmul trebuie conceput cu un număr minim de operații aritmetice.

III. REZOLVAREA NUMERICĂ A ECUAȚIILOR ALGEBRICE ȘI TRANSCENDENTE

Cursul urmărește studiul algoritmilor matematice cu ajutorul cărora pot fi rezolvate numeric probleme matematice uzuale. Unele limbaje de programare au deja funcții create, ce au la bază asemenea algoritme. Acest lucru nu constituie obiectul cursului deoarece aceste funcții sunt deja create fără a le putea modifica. Ideea este de a stăpâni principiile pentru a putea crea funcții adaptate oricăror situații.

- Exemplu: pentru rezolvarea unei ecuații polinomiale se poate utiliza funcția 'roots', de exemplu: $x^3 + 2x^2 - 2x + 3 = 0$
 $P = [1 \ 2 \ -2 \ 3]$
 $\text{roots}(p) \rightarrow \text{soluția}$
 Verificare $\text{polyval}(p,s)$.

Identificarea intervalelor în care se găsesc soluțiile unei ecuații.

Majoritatea metodelor de rezolvare a ecuațiilor fac apel la pornirea calculului de la un interval de unde se presupune că se găsește soluția.

Există metode de determinare a acestui interval, dar suplimentar, dacă ecuația reprezintă o funcție ce urmărește un fenomen fizic, realitatea poate și ea conduce la identificarea intervalului.

Exemplu: $2x^3 + 5x^2 + x - 2 = 0$
 $\Rightarrow a_0x^3 + a_1x^2 + a_2x + a_3 = 0$

- Regula lui Descartes:

Pentru $f(x) = 2x^3 + 5x^2 + x - 2 = 0$; 1 schimbare

$$\downarrow$$

$$1 \text{ soluție} > 0$$

$f(-x) = -2x^3 + 5x^2 - x - 2 = 0$; 2 schimbări

$$\downarrow$$

$$2 \text{ soluții} < 0$$

$$x_1^2 + x_2^2 + x_3^2 = (x_1 + x_2 + x_3)^2 - 2 \cdot (x_1 \cdot x_2 + x_2 \cdot x_3 + x_1 \cdot x_3) = \left(-\frac{a_1}{a_0}\right)^2 - 2 \cdot \frac{a_2}{a_0}$$

$$= \frac{a_1^2 - 2 \cdot a_0 \cdot a_2}{a_0^2}$$

Soluție) $x_1^2 < x_1^2 + x_2^2 + x_3^2 = \frac{a_1^2 - 2a_0 \cdot a_2}{a_0^2}$

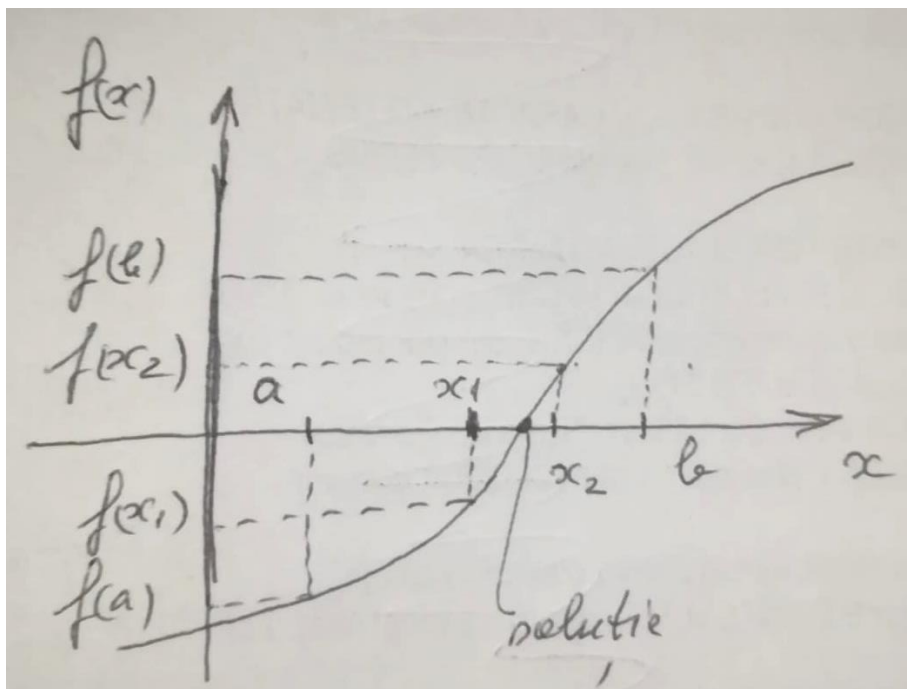
$$x_1 \in \left[-\frac{\sqrt{a_1^2 - 2 \cdot a_0 \cdot a_2}}{a_0}, +\frac{\sqrt{a_1^2 - 2 \cdot a_0 \cdot a_2}}{a_0} \right]$$

Fie x_1 soluția pozitivă $\Rightarrow x_1 \in \left[0, +\frac{\sqrt{a_1^2 - 2a_0 \cdot a_2}}{a_0} \right]$

$$\Rightarrow x_1 \in \left[0, +\frac{\sqrt{21}}{2} \right]$$

III.1. Metoda înjumătățirii intervalului (metoda biseecției).

Fie o funcție $f: \mathbb{R} \rightarrow \mathbb{R}$; pentru identificarea unei soluții în intervalul $[a, b]$, din \mathbb{R} , trebuie ca funcția să fie continuă cel puțin pe acest interval. Pentru ca în interval să se găsească cel puțin o soluție este nevoie ca $f(a) \cdot f(b) < 0$.



Algoritmul de căutare a soluției este foarte simplu și anume: se determină jumătatea intervalului $[a, b]$.

Fie $f: \mathbb{R} \rightarrow \mathbb{R}$, fie un interval $[a, b]$ în care se găsește o soluție a ecuației $f(x)=0$.

Jumătatea intervalului este $x_1 = \frac{a+b}{2}$

Se determină valoarea funcției $f(x_1)$ și se verifică dacă $f(x_1) \cdot f(a) < 0 \xrightarrow{\text{Da}}$ soluția $\in [a, x_1]$, iar dacă $\xrightarrow{\text{Nu}} f(b) \cdot f(x_1) < 0 \Rightarrow \text{soluția} \in [x_1, b]$.

În urmă acestei prime verificări, intervalul $[a, b]$ s-a înjumătățit. De exemplu, conform figurii, intervalul rămas este $[x_1, b]$. În acesta se va căuta în continuare soluția.

Se înjumătățește și acest interval determinându-se $x_2: x_2 = \frac{x_1+b}{2}$ și se verifică

$f(x_2) \cdot f(x_1) < 0$, stabilindu-se dacă soluția este în intervalul $[x_1, x_2]$ sau $[x_2, b]$. Conform figurii prezente soluția se găsește în $[x_1, x_2]$.

Procesul se continuă prin înjumătățiri succesive, până când diferența dintre capetele de interval va scădea sub o eroare acceptată stabilită inițial. În acel moment soluția poate fi considerată fie unul dintre capetele de interval ultim găsit, fie jumătatea acestuia.

III.2. Metoda poziției false

a) Metoda punctului fix.

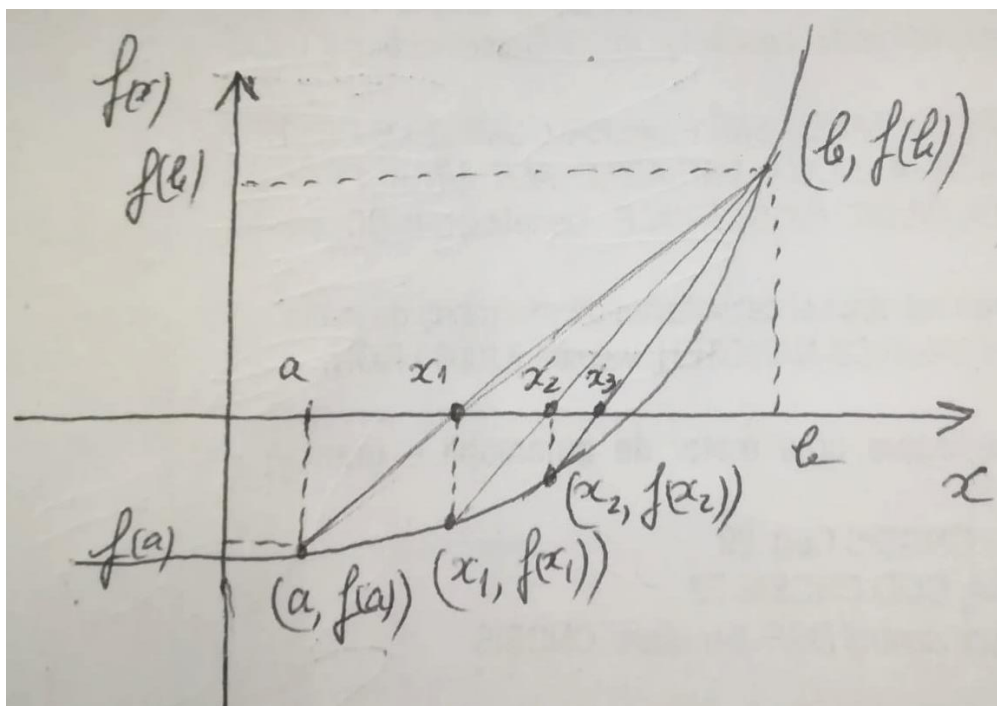
Fie funcția $f: \mathbb{R} \rightarrow \mathbb{R}$ pentru care se caută soluția ecuației $f(x)=0$ în intervalul $[a,b]$, funcția fiind continuă în acest interval.

Fie punctul $(b, f(b))$ punct fix, se construiește dreapta care trece prin punctul fix și punctul $(a, f(a))$. Punct de intersecție al acestei drepte cu axa OX îl notăm cu x_1 , iar acestei valori, îi corespunde la nivelul funcției punctul $(x_1, f(x_1))$.

Procedeul se repetă, dar de data aceasta, punctul fix $(b, f(b))$ se unește printr-o dreaptă cu noul punct $(x_1, f(x_1))$. Prin intersecția acestei noi drepte cu axa OX se determină x_2 , căruia îi corespunde la nivelul funcției punctul $(x_2, f(x_2))$.

Din nou se unește punctul fix $(b, f(b))$ cu $(x_2, f(x_2))$, iar prin intersecția cu axa OX, se determină x_3 , căruia îi corespunde la nivelul funcției $(x_3, f(x_3))$

Procedeul se continuă cu x_4, x_5, \dots, x_k , până când diferența în valoare absolută dintre două valori succesive x_k și x_{k+1} este mai mică decât o eroare acceptată.



Procesul este mai rapid decât în cazul înjumătățirii intervalului.

Prezentarea de mai sus a metodei a fost grafică / vizuală. Care sunt însă relațiile matematice care stau la baza metodei? Acestea sunt: ecuația unei drepte care trece prin două puncte date și intersecția unei drepte cu axa OX, care înseamnă de fapt că pentru ecuația dreptei, $y = 0$, pentru intersecția cu OX

- **Observație:** o dreaptă care trece prin două puncte date (x_1, y_1) și (x_2, y_2) este dată de relația: $\frac{x-x_1}{x_2-x_1} = \frac{y-y_1}{y_2-y_1}$;

- în cazul nostru pentru: $(a, f(a)), (b, f(b))$, avem $\frac{x-a}{b-a} = \frac{y-f(a)}{f(b)-f(a)} \Rightarrow y - f(a) = (x - a) \cdot \frac{f(b)-f(a)}{b-a}$

pentru intersecție cu axa OX $\Rightarrow y = 0$ avem: $-f(a) = (x - a) \cdot \frac{f(b)-f(a)}{b-a}$

$$(x - a) \cdot \frac{f(b) - f(a)}{b - a} + f(a) = 0$$

$$x \cdot \frac{f(b) - f(a)}{b - a} - \frac{a \cdot f(b) + a \cdot f(a) + b \cdot f(a) - a \cdot f(a)}{b - a}$$

$$x \cdot \frac{f(b) - f(a)}{b - a} = \frac{a \cdot f(b) - b \cdot f(a)}{b - a}$$

$$x = \frac{a \cdot f(b) - b \cdot f(a)}{f(b) - f(a)} \xrightarrow{\text{dacă punctul fix este } b} x_{k+1} = \frac{x_k \cdot f(b) - b \cdot f(x_k)}{f(b) - f(x_k)}$$

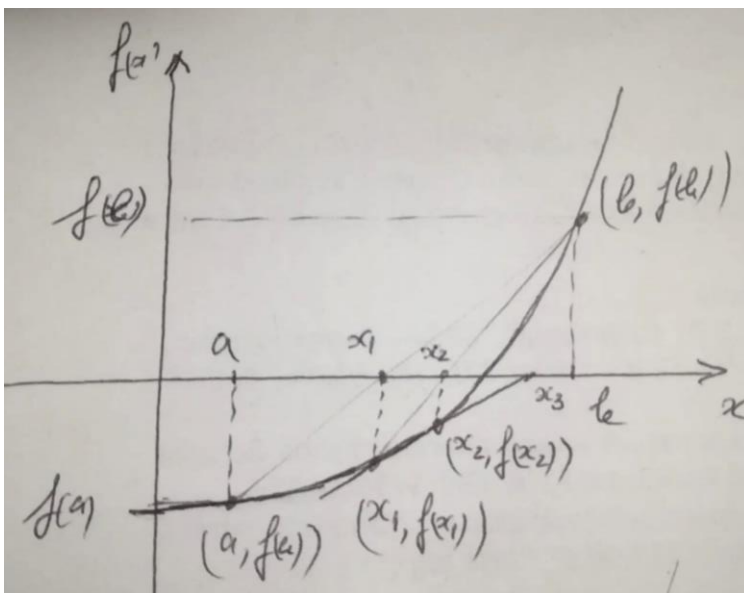
metoda punctului fix $\rightarrow |x_{k+1} - x_k| < \text{eroare}$.

b) Metoda secantei.

- Metoda secantei este similară metodei punctului fix, numai că nu se mai păstrează punctul fix, ci se operează mereu cu ultimele două puncte determinate, pentru definirea celui de al treilea.

Se construiește secanta care unește $(a, f(a))$ cu $(b, f(b))$, determinându-se x_1 ca intersecție cu OX, după care se unește $(x_1, f(x_1))$ cu $(b, f(b))$, determinându-se x_2 ca intersecție cu OX, după care se unește $(x_1, f(x_1))$ cu $(x_2, f(x_2))$ determinându-se astfel x_3 . Și așa mai departe, procedeul se continuă cu x_4, x_5, \dots, x_k , până când diferența în valoare absolută dintre două valori succesive x_k și x_{k+1} este mai mică decât o eroare acceptată.

$$x_1 = \frac{a \cdot f(b) - b \cdot f(a)}{f(b) - f(a)} \xrightarrow{\text{succesiv}} x_{k+1} = \frac{x_k \cdot f(x_{k-1}) - x_{k-1} \cdot f(x_k)}{f(x_{k-1}) - f(x_k)}$$

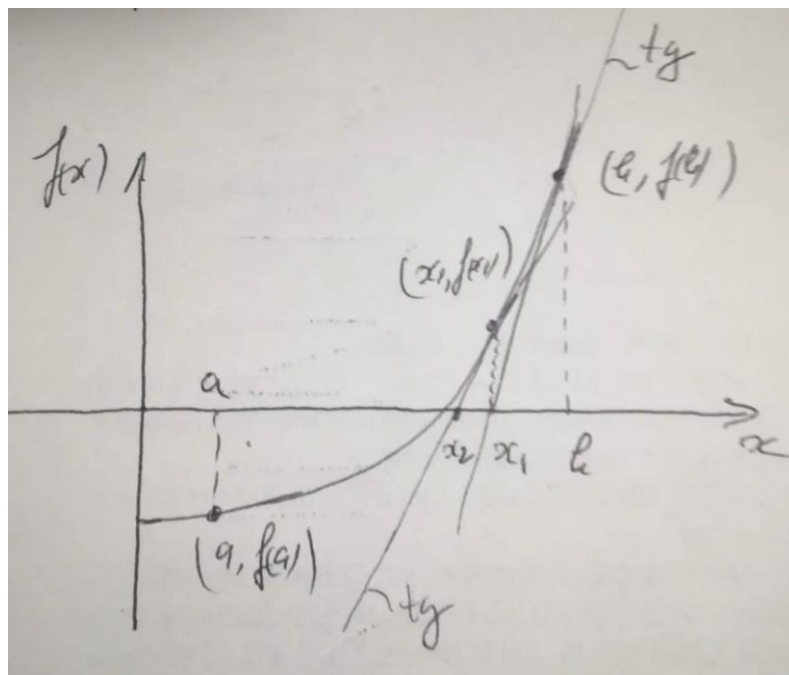


c) Metoda tangentei

Constă în mod similar din operații succesive:

- se construiește tangenta la graficul funcției într-un punct cunoscut $(x_b \cdot f(x_b))$ și intersecția acesteia cu axa OX, determinând punctul x_1 . Acestuia îi corespunde la nivelul funcției punctul $(x_1 \cdot f(x_1))$
- se construiește tangenta la graficul funcției în punctul $(x_1 \cdot f(x_1))$ și intersecția acesteia cu axa OX, determinând punctul x_2 . Acestuia îi corespunde la nivelul funcției punctul $(x_2 \cdot f(x_2))$
-
- se continuă operațiunea prin determinarea $x_3, x_4, x_5, \dots, x_k$, până când diferența în valoare absolută dintre două valori succesive x_k și x_{k+1} este mai mică decât o eroare acceptată.

(ca observație, asemănarea cu metoda secantei poate fi gândită și în felul următor: secanta construită între două puncte foarte apropiate tinde spre tangentă când punctele se suprapun)



Modalitatea de a găsi tangenta într-un punct și intersecția ei cu axa OX este următoarea:

- Dreapta care trece prin două puncte date (x_0, y_0) și (x_1, y_1) este dată de relația
$$\frac{x-x_0}{x_1-x_0} = \frac{f(x)-f(x_0)}{f(x_1)-f(x_0)}, \text{ unde } y_0=f(x_0) \text{ și } y_1=f(x_1)$$
- Când cele două puncte le „apropii”, x_1 tinde către x_0 , obținându-se ecuația tangentei într-un punct la de pe grafic:

$$\text{Astfel: } f(x) - f(x_0) = (x - x_0) \cdot \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

$\downarrow f'(x_0)$ când x_1 tinde către x_0

$$f(x) - f(x_0) = (x - x_0) \cdot f'(x_0)$$

- La intersecția tangentei cu axa OX pentru că $f(x) = 0 \Rightarrow -f(x_0) = (x - x_0) \cdot f'(x_0) \Rightarrow x - x_0 = -\frac{f(x_0)}{f'(x_0)}$

$$x = x_0 - \frac{f(x_0)}{f'(x_0)}$$

relația generală este: $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$.

III.3. METODA APROXIMĂRILOR SUCESIVE.

Metoda aproximărilor succesive este o metodă simplă, dar nu întotdeauna eficientă de rezolvare a ecuațiilor.

Potrivit acesteia, ecuația $f(x) = 0$ se rescrie sub forma $x = g(x)$, se alege o valoare de pornire x_1 și se construiește șirul iterativ $x_{k+1} = g(x_k)$, cu $k = 1, 2, 3, \dots$. Alegerea valorii inițiale ține cont de intervalul de căutare, iar în cazul unor fenomene fizice și de intervalul fizic în care se dorește identificarea soluției.

Dacă șirul este convergent el va tinde către soluția α a ecuației, pornind de la o valoare inițială stabilită în intervalul (a, b) . După cum s-a menționat anterior, identificarea intervalului (a, b) se realizează prin metode de separare sau prin limitări impuse de fenomenul fizic descris din ecuații.

Convergența șirului este asigurată dacă $|g'(x)| < 1, x \in (a, b)$.

Eroarea în cazul acestei metode este : $e = |x_n - \alpha|$.

Observație: Procesul iterativ se oprește de regulă când $|x_n - x_{n-1}| < \text{eroarea acceptată}$.

• Exemplu de aplicație:

$$f(x) = x^2 - 2 \cdot x - 3, \quad x_0 = 1;$$

(a)

$$x^2 = 2 \cdot x + 3 \Rightarrow x = \sqrt{2 \cdot x + 3}; \text{pentru } x_0 = 1 \Rightarrow$$

1; 2.2361; 2.7335; 2.9098; 2.9698; 2.9899; 2.9966; 2.9989; 2.9996; 2.9999; 3.0000;

(b) $2 \cdot x = x^2 - 3$

$$x = \frac{x^2 - 3}{2} \Rightarrow \text{nu este convergent pentru că are } x^2;$$

$$\downarrow g(x)$$

aici trebuie făcută o analiză: pentru $x_0 = 1 \Rightarrow -1, -1, -1$, salt direct în soluția -1, însă dacă s-ar porni din $x = 4$, rezultă șirul 6.5; 19.625; 19.07 \Rightarrow nu este convergent.

(c) $x \cdot (x - 2) - 3 = 0$

$$x - 2 = \frac{3}{x}$$

$$x = \frac{3}{x} + 2$$

$$\downarrow g(x)$$

$x = 1 \Rightarrow 5.26; 3.15; 2.95; 3.0165; 2.9945; 3.0018; 2.9994; 3.0002; 2.9999; 3.0000$.

Observație:

obținerea funcției g , poate porni de la: $f(x) = 0; f(x) + x = x$. Însă nu este singura variantă.

$$\downarrow g(x).$$

RELAȚIA AITKEN

Dacă se dorește îmbunătățirea rapidității procesului iterativ se poate folosi o relație de calcul numită **RELAȚIA AITKEN** \Rightarrow se bazează pe cunoașterea a cel puțin trei termeni consecutivi **determinați** în procesul iterativ, după care prin aplicarea ei, ce nu mai depinde de $g(x)$, ci doar de cele 3 valori. Astfel este accelerată convergența către soluție. Metoda se recomandă mai ales pentru expresii mari, complexe, ale funcției f inițiale și evident ale lui g . Fie funcția $g(x)$ dezvoltată în serie Taylor în jurul punctului α .

$$g(x) = g(\alpha) + (x - \alpha) \cdot g'(\alpha) + \frac{(x - \alpha)^2}{2} \cdot g''(\alpha)$$

$$\begin{aligned} x_{k+1} &= g(x_k) \\ x_{k+2} &= g(x_{k+1}) \end{aligned}$$

$$\begin{aligned} g(x_k) &= g(\alpha) + (x_k - \alpha) \cdot g'(\alpha) + \frac{(x_k - \alpha)^2}{2} \cdot g''(\alpha) \\ g(x_{k+1}) &= g(\alpha) + (x_{k+1} - \alpha) \cdot g'(\alpha) + \frac{(x_{k+1} - \alpha)^2}{2} \cdot g''(\alpha) \end{aligned}$$

$$\begin{aligned} x_{k+1} &= \alpha + (x_k - \alpha) \cdot g'(\alpha) + \frac{(x_k - \alpha)^2}{2} \cdot g''(\alpha) \\ x_{k+2} &= \alpha + (x_{k+1} - \alpha) \cdot g'(\alpha) + \frac{(x_{k+1} - \alpha)^2}{2} \cdot g''(\alpha) \end{aligned}$$

$$\begin{aligned} x_{k+1} - \alpha &= (x_k - \alpha) \cdot g'(\alpha) + \frac{(x_k - \alpha)^2}{2} \cdot g''(\alpha) \\ x_{k+2} - \alpha &= (x_{k+1} - \alpha) \cdot g'(\alpha) + \frac{(x_{k+1} - \alpha)^2}{2} \cdot g''(\alpha) \end{aligned}$$

- Observație: Dacă cele două relații se împart una la cealaltă și $g'(x_{k+1})$ se consideră apropiat de $g'(x_k)$ se obține: $\frac{x_{k+1} - \alpha}{x_{k+2} - \alpha} = \frac{x_k - \alpha}{x_{k+1} - \alpha} \Rightarrow (x_{k+1} - \alpha)^2 = (x_{k+2} - \alpha) \cdot (x_k - \alpha)$
 $(x_k - \alpha) \Rightarrow x_{k+1}^2 - 2 \cdot x_{k+1} \cdot \alpha + \alpha^2 = x_{k+2} \cdot x_k - x_{k+2} \cdot \alpha - x_k \cdot \alpha + \alpha^2$
 $\alpha^2 \Rightarrow x_{k+1}^2 - 2 \cdot x_{k+1} \cdot \alpha = x_{k+2} \cdot x_k - x_{k+2} \cdot \alpha - x_k \cdot \alpha$
- trecem termenii în stânga: $\alpha \cdot (x_{k+2} - 2 \cdot x_{k+1} + x_k) + x_{k+1}^2 - x_{k+2} \cdot x_k = 0$

$$\alpha = \frac{x_{k+2} \cdot x_k - x_{k+1}^2}{x_{k+2} - 2 \cdot x_{k+1} + x_k}$$

notăm $\alpha = x_{k+3} \Rightarrow x_{k+3} = x_{k+2} - \frac{(x_{k+2} - x_{k+1})^2}{x_{k+2} - 2 \cdot x_{k+1} + x_k} \rightarrow$ **Relația lui Aitken.**

Exemplu de aplicație:

$$f(x) = 0.5 - x + 0.2 \cdot \sin(x)$$

$$\text{soluție: } x_0 \text{ ales } = 1$$

$$\text{scriem } f(x) \text{ sub forma } x = 0.5 + 0.2 \cdot \sin(x)$$

$$x_1 = 0.5 + 0.2 \cdot \sin(x_0) = 0.6683;$$

$$x_2 = 0.5 + 0.2 \cdot \sin(x_1) = 0.6239;$$

$$x_3 = 0.5 + 0.2 \cdot \sin(x_2) = 0.6168;$$

$$x_4 = x_3 - \frac{(x_3 - x_2)^2}{x_3 - 2 \cdot x_2 + x_1} = 0.615$$

$$x_5 = x_4 - \frac{(x_4 - x_3)^2}{x_4 - 2 \cdot x_3 + x_2} = 0.6152$$

$$x_6 = 0.6151,$$

$$x_7 = 0.6150,$$

$$x_8 = 0.6150.$$