

UNIVERSITATEA TEHNICĂ A MOLDOVEI
Facultatea de Calculatoare, Informatică și Microelectronică
Departamentul
INFORMATICĂ ȘI INGINERIA SISTEMELOR

Disciplina
Metode și modele de calcul
Modulul I (Metode numerice)

Programul de licență
CALCULATOARE și REȚELE, TEHNOLOGIA INFORMAȚIEI
(frecvența redusă)

Prof. univ. dr. Vasile MORARU
Anul universitar: 2021-2022

A grayscale photograph of a person wearing large headphones, looking intently at a laptop screen. The person's hands are clasped together near their chin. The background is blurred, suggesting an office or studio environment. The text is overlaid on the image within a white rectangular border.

ȘEDINȚA NR.1

*NUMERE CU VIRGULĂ
MOBILĂ*

Sumar

**Introducere.
Clasificarea problemelor**

Erori absolute și erori relative

Propagarea și sursele erorilor

Numere cu virgulă mobilă

**Aritmetica virgulei mobile
și erorile de rotunjire**

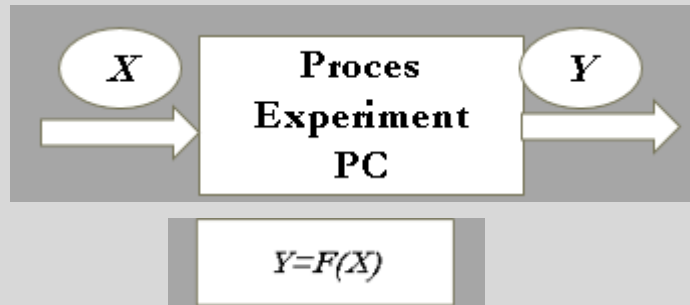
**Determinarea parametrilor unui
sistem de calcul**

Efectul erorilor de rotunjire

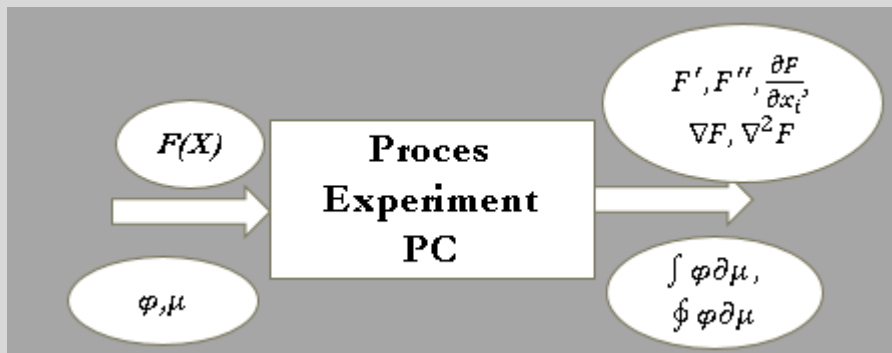
Clasificarea problemelor

Probleme directe

I. Modelarea matematică (calculul funcției)

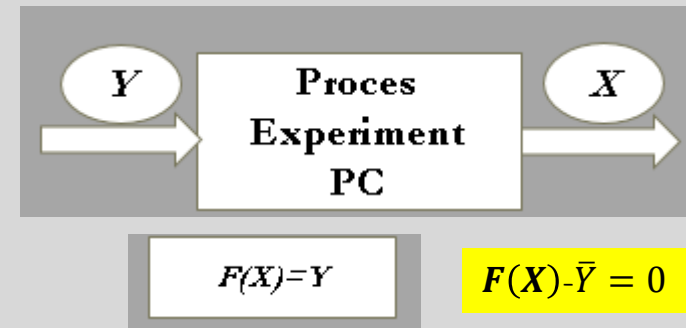


II. Calculul diferențial și integral (numeric !)

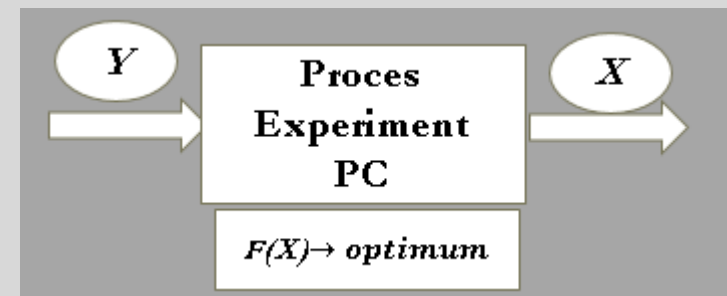


Probleme inverse

I. Rezolvarea ecuațiilor



II. Optimizări



Scurt istoric

Științele moderne încep cu **René Descartes**. Ideile sale au făcut ca științele să facă progrese uimitoare și astfel să apară în mod necesar ideea de **metodă**. El se găsește la originea filosofiei și a culturii europene, eliberând-o din chingile scolasticii a lui Aristotel

De la Descartes ne-a rămas **metoda de cunoaștere** și acel lucru excepțional care este exprimarea în termeni matematici (algebrici) , cunoscut ca prima unificare din științele matematice dintre algebră și geometrie.

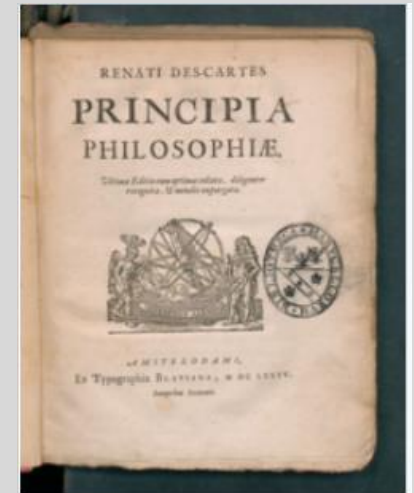
Descartes și contemporanii lui (**Kepler, Galileo**) au inițiat trecerea de la filosofie la sistemele științifice exprimând cu convingere pentru prima dată ideea de **model matematic**.

Newton a realizat primul succes notabil în reprezentarea matematică a fenomenelor naturale prin conceperea, împreună cu **Leibniz**, a calcului diferențial, un instrument excepțional de analiză locală pe care se bazează toate dezvoltările științifice a zilelor noastre



René Descartes

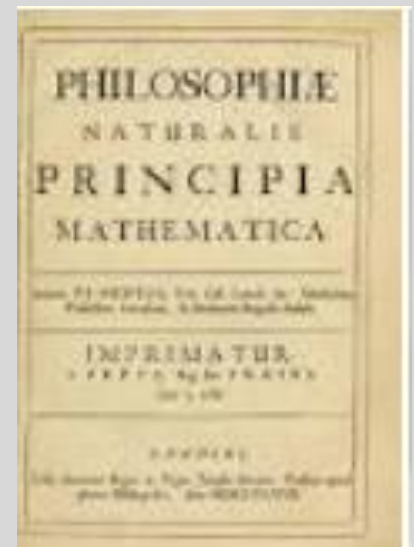
(1596-1650)



Principia philosophiæ, 1685



(1642-1726)



Erori absolute și erori relative

Notăm cu x_* valoarea aproximativă pentru numărul exact x . Expresia

$$\Delta(x_*) = |x - x_*|$$

poartă numele de ***eroare absolută***.

Eroarea relativă $\delta(x_*)$ reprezintă raportul dintre eroarea absolută și valoarea aproximativă, adică

$$\delta(x_*) = \frac{|x - x_*|}{|x_*|}, \quad x_* \neq 0.$$

Numărul pozitiv ε este o ***margină*** (sau o ***limită***) a ***erorii absolute*** a numărului aproximativ x_* dacă

$$|x - x_*| \leq \varepsilon,$$

iar numărul pozitiv r este o ***limită a erorii relative*** dacă

$$\frac{|x - x_*|}{|x_*|} \leq r.$$

Notăția $x = x_* \pm \varepsilon$ semnifică întotdeauna faptul, că $|x - x_*| \leq \varepsilon$, adică

$$x_* - \varepsilon \leq x \leq x_* + \varepsilon.$$

Zerourile de la începutul numărului servesc numai pentru a fixa poziția virgulei zecimale. Cifrele cuprinse între prima și ultima cifră diferită de zero sau care indică ordinele păstrate în calcule se numesc *cifre semnificative*.

Exemplu. Numărul aproximativ $x_* = 3 \cdot 10^1 + 6 \cdot 10^0 + 0 \cdot 10^{-1} + 5 \cdot 10^{-2} + 8 \cdot 10^{-3}$ are cinci cifre semnificative, iar numărul $y_* = -(2 \cdot 10^{-3} + 8 \cdot 10^{-4} + 0 \cdot 10^{-5}) = -0.00280$ are trei cifre semnificative (primele trei zerouri sunt nesemnificative).

Dacă mărimea erorii x_* nu depășește $0.5 \cdot 10^{-t}$ se spune că numărul aproximativ x_* are *t cifre zecimale corecte*.

Numărul de zecimale corecte ne permite să ne facem o idee despre mărimea erorii absolute în timp ce numărul de cifre semnificative ne dă o idee sumară despre mărimea erorii relative

Propagarea erorilor

La adunare sau scădere, marginea erorii absolute a rezultatului este dată de suma marginilor pentru erorile absolute ale termenilor.

La înmulțire și împărțire marginile erorilor relative ale factorilor se adună.

Eroarea relativă a diferenței poate fi destul de mare, dacă diferența $|x_* - y_*|$ este foarte mica:

$$\delta (x_* - y_*) = \frac{\Delta(x_* - y_*)}{|x_* - y_*|} \leq \frac{\Delta(x_*) + \Delta(y_*)}{|x_* - y_*|}$$

Exemplu. Fie $x=0.1234\pm 0.5\cdot 10^{-4}$ și $y=0.1233\pm 0.5\cdot 10^{-4}$ atunci $x-y = 0.0001\pm 0.0001$ și marginea erorii este tot atât de mare ca și estimarea rezultatului.

Cele mai serioase erori care apar în calculele efectuate cu ajutorul calculatorului electronic sunt datorate acestui fenomen care poartă denumirea de anulare prin scădere sau de neutralizare a termenilor .

De câte ori este posibil neutralizarea termenilor se evită prin rescrierea formulelor de calcul sau prin alte schimbări în algoritm. De exemplu, o expresie de forma

$$(\alpha + \gamma)^2 - \alpha^2$$

poate fi scrisă sub forma

$$\gamma(\gamma + 2\alpha)$$

Sursele erorilor

1. **Erori provenite din simplificarea modelului fizic**, pentru a fi descris într-un model matematic; erori din măsurările inițiale sau din soluții aproximative ale altor probleme etc. Aceste tipuri de erori se numesc **erori inerente**. Ele nu pot fi influențate de metoda de calcul.
2. **Erori de metodă sau de trunchiere**. Majoritatea metodelor numerice necesită un număr infinit de operații aritmetice pentru a ajunge la soluția exactă a problemei. De aceea suntem nevoiți să trunchiem metoda după un număr finit de operații. Ceea ce omitem constituie eroarea de trunchiere.
3. **Erori de rotunjire** în datele de intrare în calcule și în datele de ieșire. Multe numere nu pot fi reprezentate exact printr-un număr dat de cifre. Dacă în calcule numerice trebuie să folosim numărul π , îl putem scrie 3.14, 3.14159 sau 3.1415926 etc.

Nici un număr irațional nu poate fi reprezentat printr-un număr finit de cifre.

$$\sqrt{2} \cong 1.4142\dots$$

$$\pi = 3.14159265\dots$$

$$\sqrt{17} = 4.123105\dots$$

Chiar și unele numere raționale nu au o reprezentare exactă.

$$\frac{1}{3} \cong 0.333333 \dots$$

Numere cu virgulă mobilă

Un număr scris în virgulă mobilă este compus dintr-o fracție, numită **mantisă** și un întreg, numit **exponent**:

$$x = \pm m \cdot \beta^e,$$

β este baza sistemului de numerație (binar, octal sau hexazecimal),

m este mantisa numărului și e este exponentul, afectat de semn.

$$\frac{1}{\beta} \leq m < 1$$

m are forma

$$m = \frac{d_1}{\beta^1} + \frac{d_2}{\beta^2} + \dots + \frac{d_t}{\beta^t},$$

unde numerele întregi d_1, d_2, \dots, d_t , numite cifre, verifică inegalitățile

$$0 \leq d_i \leq \beta - 1, \quad i = 1, 2, \dots, t; \quad L \leq e \leq U.$$

Dacă prima cifră din mantisă este diferită de zero, atunci numărul reprezentat în virgulă mobilă se numește **normalizat**.

Numere cu virgulă mobilă

Sistemul de calcul cu numere cu virgulă mobilă este o mulțime

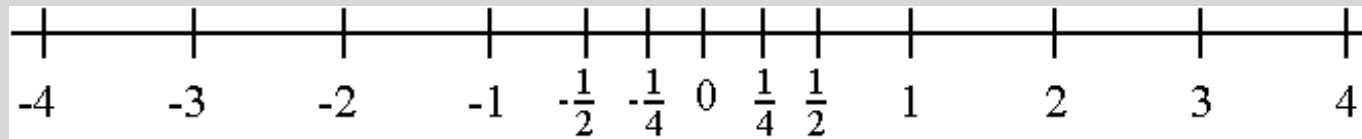
$$F(\beta, t, L, U)$$

β - baza sistemului de numerație

t - precizia mașinii

$[L, U]$ - intervalul exponenților

Exemplu. Sistemul de calcul $F(2,3,-1,-2)$ conține 33 puncte:



Mulțimea F nu poate reproduce oricât de detaliat structura continuă a numerelor reale.

Nu putem reprezenta în calculator numerele al căror modul depășește cel mai mare element al lui F sau care sunt mai mici în modul decât cel mai mic număr din F .

Unitatea de rotunjire a mașinii ε_M

În calculator numărul real $x \neq 0$ este reprezentat de numărul cu virgulă mobilă $fl(x)$, a cărei mantisă m_* se obține din mantisa m a lui x rotunjind-o la t cifre (de aceea spunem că precizia mașinii este t).

Dacă se efectuează rotunjirea corectă atunci

$$|m - m_*| \leq \frac{1}{2} \beta^{-t}.$$

Eroarea relativă în $fl(x)$ este

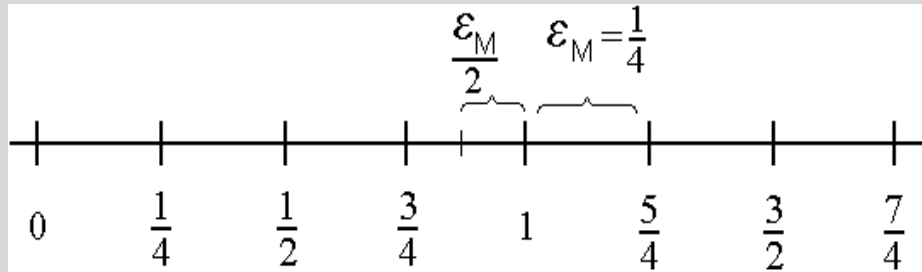
$$\frac{|fl(x) - x|}{|x|} \leq \frac{1}{2} \beta^{1-t},$$

Numărul $\varepsilon_M = \frac{1}{2} \beta^{1-t}$ se numește ***unitatea (de rotunjire a) mașinii***. Efectuând rotunjirea corectă, numărul $fl(x)$ este cel mai apropiat element de x , care aparține lui F . Dacă se folosește rotunjirea prin tăiere (se elimină compararea primei cifre neglijate), atunci $\varepsilon_M = \beta^{1-t}$ și $fl(x)$ este cel mai apropiat element din F , inferior lui x .

Parametrii unui sistem de calcul

În afară de parametrul ε_M în practică sunt larg răspândiți încă doi parametri σ și λ : cel mai mic element pozitiv și elementul maximal al lui F .

Exemplu. În sistemul de calcul $F(2,3,-1,1)$ cel mai mic element pozitiv este egal cu $\sigma = 1/2 \cdot 2^{-1} = 1/4$, iar cel mai mare element $\lambda = 7/8 \times 2^1 = 7/4$.



Unitatea de rotunjire a mașinii ε_M se mai numește *epsilon al mașinii* și este cel mai utilizat parametru ce caracterizează un sistem de calcul dat. Acest parametru ne dă măsura de “discretizare” a sistemului F care are loc pentru tot intervalul numerelor nenule în virgulă mobilă.

Distanta dintre numărul $x \in F$ și numărul cel mai apropiat de el în sistemul dat nu e mai mică decât $\varepsilon_M/x//\beta$ și nu e mai mare decât $\varepsilon_M/x/$ (numai dacă numărul x nu este situat în vecinătatea lui zero).

Aritmetica virgulei mobile și erorile de rotunjire

Majoritatea sumelor $x+y$ cu $x \in F$, $y \in F$, aparțin ele însele lui F , foarte rar produsul (exact) obișnuit $x \times y$ aparține lui F , deoarece, de regulă, el are $2t$ sau $(2t-1)$ cifre semnificative .

Depășirea este mai probabilă la înmulțire. În cazul înmulțirii cu virgulă mobilă, este posibilă apariția unui **zero-mașină**, atunci când pentru $x \neq 0$ și $y \neq 0$, produsul $x \times y$ este nenul dar este mai mic, în modul, decât cel mai mic element pozitiv al lui F (apariția unui zero-mașină este posibilă și în cazul scăderii deși aceasta se întâmplă foarte rar).

Pentru $x \neq 0$ eroarea relativă în $fl(x)$

$$\delta(x) = \frac{|fl(x) - x|}{|x|}$$

îndeplinește condiția

$$\delta(x) \leq \varepsilon_M = \begin{cases} \beta^{1-t} & , \text{dacă are loc rotunjirea prin trunc hier} \\ \frac{1}{2} \beta^{1-t} & , \text{dacă are loc rotunjirea corectă .} \end{cases}$$

Determinarea parametrilor unui sistem de calcul

Parametrul ε_M poate fi definit ca cel mai mic număr pozitiv care adăugat în sistemul de calcul F la unitate dă în rezultat un număr cu virgulă mobilă ce aparține din nou lui F și este strict mai mare ca 1 , adică

$$fl(1+\varepsilon_M) > 1 .$$

Parametrii σ și λ se reprezintă prin parametrii β , t , L și U după cum urmează

$$\sigma = \beta^{L-1}, \lambda = \beta^U (1 - \beta^{-t})$$

Operațiile aritmetice $+$, $-$, $*$, $/$ se realizează la calculatoarele numerice după cum urmează (prin x este notat rezultatul exact al operației aritmetice):

- 1) dacă $\sigma \leq |x| \leq \lambda$, atunci rezultatul operației se **rotunjește**;
- 2) dacă $|x| < \sigma$, atunci rezultatul se anulează, adică apare un **zero-mașină**;
- 3) dacă $|x| > \lambda$, atunci calculele se întrerup și apare un **semnal de depășire**.

Calculul parametrului ε_M

```
typedef float real;
real calceps()
{ static volatile real eps, eps2, eps21; eps = 1.0;
  eps2 = eps * 0.5;
  eps21 = eps2 + 1.0;
  while( eps21 > 1.0 ){
  eps = eps2;
  eps2 = eps * 0.5;
  eps21 = eps2 + 1.0;
  }
  return eps;
}
```


Efectul erorilor de rotunjire

Calculul mediei aritmetice

Evaluarea recurentă a unei integrale

Exemple de sisteme rău condiționate

Calculul mediei aritmetice

Media aritmetică dintre două numere reale a și b poate fi calculată prin formulele :

$$c = \frac{a+b}{2} \quad (1)$$

$$c = a + \frac{b-a}{2} \quad (2)$$

Formula (1) necesită cu o operație de adunare mai puțin decât formula (2) dar din punct de vedere al exactității nu este întotdeauna cea mai bună.

Exemplu 1. $\beta=10$, $t=3$, $a=0.596$ și $b=0.600$, rotunjirea corectă

$$c = \frac{0.596+0.600}{2} = \frac{1.20}{2} = 0.600, \text{ (valoarea corectă a lui } c \text{ este egală cu } 0.598)$$

$$c = 0.596 + \frac{0.600 - 0.596}{2} = 0.596 + \frac{0.004}{2} = 0.598$$

Exemplu 2. $\beta=10$, $t=4$, $a=-3.483$ iar $b=8.765$ (rotunjirea prin tăiere)

$$c = \frac{-3.483+8.765}{2} = \frac{5.282}{2} = 2.641 \text{ (rezultat corect)}$$

$$c = -3.483 + \frac{8.765+3.483}{2} = -3.483 + \frac{12.24}{2} = -3.483 + 6.120 = 2.637 .$$

Este necesar de a ne folosi de una din formulele (1) sau (2), în dependență de semnele lui a și b :

dacă $sign(a) \neq sign(b)$ atunci $c=(a+b)/2$; în caz contrar $c=a+(b-a)/2$.

Exemple de sisteme rău condiționate

Exemplu 1.

$$\begin{cases} 5x - 331y = 5, \\ 6x - 397y = 7. \end{cases}$$

Soluția exactă a sistemului este $x=332$, $y=5$. Dacă se reia sistemul de mai sus, admițând însă o variație mică a coeficientului 5 de pe lângă x din prima ecuație

$$\begin{cases} 5.01x - 331y = 5, \\ 6x - 397y = 7, \end{cases}$$

atunci avem soluția $x = -111.7845\dots$, $y = -1.7070\dots$

S-a produs o catastrofă! Despre un astfel de sistem se spune că este **rău** (sau **prost**) **condiționat**.

Exemplu 2.

$$\begin{cases} 14x + 13y - 66z = 1, \\ 12x + 11y - 13z = 1, \\ 11x + 10y + 4z = 1, \end{cases}$$

Soluție unică $x=1$, $y=-1$, $z=0$. Înlocuim elementele din partea dreaptă a sistemului $(1, 1, 1)$ cu 1.001 , 0.999 și 1.001 respectiv. Atunci, lucrând doar cu trei cifre semnificative, vom obține

$$x = -0.683, y = 0.843, z = 0.006.$$

Concluzii

Rezolvarea numerică a unei probleme nu este o chestiune simplă, adică nu este suficient să facem niște calcule pentru a ajunge la soluția problemei.

Mai este necesar de a studia și de a face aprecieri, privind condiționarea problemei și stabilitatea numerică a algoritmilor de calcul.



ÎNTREBĂRI!

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Maecenas porttitor congue massa.

