CALCOLO NUMERICO

1 - DISPENSE (Prof. Ghelardoni)

2 - BINI - CAPOVANI - MENCHI = Metodi Numerici per l'algebra lineare, Zanichelli

3 - BEVILACQUA, BINI, ↑ = Metodi Numerici, Zanichelli

Ricevimento: Lunedì Pomeriggio 14.30 - 17.00

25/09/2012

$\beta > 1$, $\beta \in \mathbb{N}$ (escluso lo zero)

$x \in \mathbb{R}$, $b \in \mathbb{Z}$

$0 \le \alpha_i \le \beta - 1 \quad i = 1, \dots, t$

$\alpha_i \in \mathbb{N}_0$

Allora x si può rappresentare come $x = \text{sign}(x)\, \beta^{b} \underbrace{\sum_{i=1}^{t} \alpha_i \beta^{-i}}_{\text{Mantissa}}$

dove $\text{sign}(x) = \begin{cases} +1 & \text{se } x \ge 0 \\ -1 & \text{se } x < 0 \end{cases}$

Es. $87,22$

$\beta = 10$

$87,22 = 10^{2} \cdot 0,8722$ ⟸ $b = 2$, $\alpha_1 = 8$, $\alpha_2 = 7$, $\alpha_3 = \alpha_4 = 2$, $\alpha_i = 0 \ \forall i \ge 5$

MASTER COPY

Tel. 388/9837745

oppure $10^{3} \cdot 0,08722$ $\alpha_1 = 0$, $\alpha_2 = 8$, $\alpha_3 = 7$, $\alpha_4 = \alpha_5 = 2$ e $b = 3$

l'esponente lo posso scegliere come voglio io.

Unicità della rappresentazione

1) $\alpha_1 \neq 0$ quindi la seconda rappresentazione non va bene

2) $\exists k \in \mathbb{N}$ t.c. $\alpha_i = \beta - 1 \ \forall i \ge k$

Non ci deve essere periodo uguale alla base meno 1. Es. $0,\bar{9}$

Es. $1,1\bar{9} = \frac{119 - 11}{90} = \frac{108}{90} = 1,2$ le due rappresentazioni sono uguali, quindi per avere unicità si esclude quella periodica, perché infinita.

La rappresentazione ottenuta si dice in virgola mobile NORMALIZZATA

NORMALIZZATA perché la mantissa $\beta^{-1} \le \sum_{i=1}^{\infty} \alpha_i \beta^{-i} < \sum_{i=1}^{\infty} (\beta - 1) \beta^{-i}$

Cioè maggiorato gli d_i con $\beta-1$.

Quella è una serie geometrica di ragione $\frac{1}{\beta}$

Questa è uguale a $(\beta-1)\beta^{-1}\cdot\frac{\beta}{\beta-1}=1$

Arrivando a dire che $\beta^{-1} \leq \sum_{i=1}^{\infty} d_i \beta^{-i} < 1$

↑ Questo perché pur l'uguale deve valere $d_1 \neq 0$ (Meglio $d_1=1$ $d_i=0\ \forall i \geq 2$)

27/09/2012

Rappresentazione cifre binarie

$\tilde{x} = \text{sign}(x)\beta^{\tilde{b}} \sum_{i=1}^{m} \tilde{d}_i \beta^{-i}$ la mantisse è finita

Metodi per rappresentare

TRONCAMENTO $\tilde{x} = tr(x) = \text{sign}(x)\beta^{b} \sum_{i=1}^{m} d_i \beta^{-i}$

ARROTONDAMENTO $\tilde{x} = rd(x) = \text{sign}(x)\beta^{b} \cdot \begin{cases} \sum_{i=1}^{m} d_i \beta^{-i} & \text{se } 0 \leq d_{m+1} < \beta/2 \\ \sum_{i=1}^{m} d_i \beta^{-i} + \beta^{-m} & \text{se } \beta/2 \leq d_{m+1} \leq \beta-1 \end{cases}$

Es. arrotondamento $12.896 \to 10^2 \cdot 0,12896$ (Usiamo 4 cifre)

$\downarrow$

$10^2 \cdot 0,1290$

$10^1 \cdot 0,99997 \to 10^2 \cdot 0,1000$

INSIEME dei numeri di macchina M

$F(\beta, m, L, U)$ dove $L \leq b \leq U$ m = numero di cifre per la mantissa

(valori minimo e massimo per l'esponente)

Qual è la cardinalità di questo insieme?

Partiamo dalle mantisse: possibili valori $\to (\beta-1)\cdot\beta^{m-1}$

Possibili valori di esponenti $= U-L+1$

Y possibili valori sono dunque

$\text{card}(M) = 2\cdot(U-L+1)(\beta-1)\cdot\beta^{m-1} + (1)$ → Rappresentazione dello zero

ERRORE ASSOLUTO

$$x = \text{sign}(x)\, \beta^{b} \sum_{i=1}^{\omega} d_i \beta^{-i}$$

$$\tilde{x} = tr(x) = \text{sign}(x)\, \beta^{b} \sum_{i=1}^{m} d_i \beta^{-i}$$

$$\delta_x = tr(x) - x \qquad \delta_x = rd(x) - x$$

$$|\delta_x| < \beta^{b-m} \qquad |\delta_x| \leq \frac{1}{2}\beta^{b-m}$$

ERRORE RELATIVO (o precisione di macchina)

$$\varepsilon_x = \frac{tr(x) - x}{x} \longrightarrow |\varepsilon_x| < \frac{\beta^{b-m}}{\beta^{b-1}} = \beta^{1-m}$$

← Numeratore più grande possibile, Denominatore più piccolo

$$\varepsilon_x = \frac{rd(x) - x}{x} \longrightarrow |\varepsilon_x| < \frac{\frac{1}{2}\beta^{b-m}}{\beta^{b-1}} = \frac{1}{2}\beta^{1-m}$$

L'errore per arrotondamento è sempre minore, quindi ora useremo solo rappresentazioni per arrotondamento

CALCOLO DI UNA FUNZIONE IN UN PUNTO

$$f(P): \mathbb{R}^n \longrightarrow \mathbb{R} \qquad f(P_0)$$

$P_0 = (x_1^0, x_2^0, \ldots, x_n^0)$ Che errore compio nel calcolo del valore?

$P_0 \in D = \{P \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i \ \forall i = 1 \ldots n\}$ (Indeterminazione)

ex. $1 < \sqrt{2} < 2$, $3 < \pi < 4$, $2 < e < 3$

$$f(P_0) \longrightarrow f_e(P_1)$$

ERRORE

$$f_e(P_1) - f(P_0) = \underbrace{f_e(P_1) - f(P_1)}_{\text{Qui c'è errore viene fatto dalle scelte di } f_e} + \underbrace{f(P_1) - f(P_0)}_{\text{L'errore viene fatto cambiare primo in } f}$$

f_e → δ_A ERRORE ASSOLUTO ALGORITMICO

ERRORE ASSOLUTO TRASMESSO DAI DATI: δ_d

$$\delta_f = \delta_a + \delta_d$$

MASTER COPY

Tel. 388/9837745

$\delta_d = f(P_2) - f(P_0)$

$f(P_1) = f(P_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_i^{(1)} - x_i^{(0)})$ Polinomio di Taylor troncato al primo grado

Quindi $\delta_d = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\underbrace{(x_i^{(1)} - x_i^{(0)})}_{\text{Errore assoluto delle componenti}} = \sum_{i=1}^{n} \underbrace{\frac{\partial f}{\partial x_i}}_{\gamma_i \to \text{coefficienti di amplificazione}} \delta_{x_i}$

$$|\delta_d| = \left|\sum_{i=1}^{n} \frac{\partial}{f_i} \delta_{x_i}\right| \leq \sum_{i=1}^{n} |\gamma_i \delta_{x_i}|$$

Facciamo un'ulteriore stima Prendiamo $\forall P \in D \ \sup \left|\frac{\partial f}{\partial x_i}(P)\right| = A_{x_i}$

Per cui $|\delta_d| \leq \sum_{i=1}^{n} A_{x_i} |\delta_{x_i}|$

ESEMPIO

$f(x_1, x_2) = x_1 \cdot x_2$

$P_0 = (\sqrt{2}, \pi)$ $D = [1,2] \times [3,4]$ $P_1 (1{,}4, 3{,}14)$

↑ Approssimazione a 2 cifre

$\frac{\partial f}{\partial x_1} = x_2 \longrightarrow A_{x_1} = 4$

$\frac{\partial f}{\partial x_2} = x_1 \longrightarrow A_{x_2} = 2$

$|\delta_{x_1}| \leq \frac{1}{2} \cdot 10^{-2}$ $|\delta_{x_2}| \leq \frac{1}{2} \cdot 10^{-2}$

$|\delta_d| \leq 4 \cdot \frac{1}{2} \cdot 10^{-2} + 2 \cdot \frac{1}{2} \cdot 10^{-2} = 3 \cdot 10^{-2}$

$|\delta_A| \leq \frac{1}{2} \cdot 10^{-2}$ ← Faccio il prodotto dei valori di P_1, e arrotondo a due cifre

$|\delta_f| \leq \frac{1}{2} \cdot 10^{-2}$ Questo si chiama PROBLEMA DIRETTO

PROBLEMA INVERSO: Come devo scegliere i dati, effettivi, scelto E

$|f_a(P_1) - f(P_0)| < E$?

ESEMPIO

$f(x_1, x_2) = \frac{x_1}{x_2}$ $\quad \varepsilon = 10^{-3}$

$P_0(\sqrt{2}, \overline{11})$ $\quad D = [1,2] \times [3,4]$

- $\frac{\partial f}{\partial x_1} = \frac{1}{x_2} \rightarrow A_{x_1} = \frac{1}{3}$

$\frac{\partial f}{\partial x_2} = -\frac{x_1}{x_2^2} \rightarrow A_{x_2} = \frac{2}{3}$

$|\delta_f| \leq |\delta_d| + |\delta_\theta| \leq 10^{-3}$ Supponiamo che $|\delta_A| \leq \frac{1}{2} \cdot 10^{-3}$ basta arrotondare alle divisione fatte a 3 cifre.

$|\delta_d| \leq \frac{1}{2} \cdot 10^{-3}$

$|\delta_d| \leq \sum_{i=1}^{2} A_{x_i} |\delta_{x_i}| \leq \frac{1}{2} \cdot 10^{-3} = A_{x_1}|\delta_{x_1}| + A_{x_2}|\delta_{x_2}| \leq \frac{1}{2} \cdot 10^{-3}$

Supponiamo $A_{x_1}|\delta_{x_1}| \leq \frac{1}{4} \cdot 10^{-3} \rightarrow \frac{1}{3}|\delta_{x_1}| \leq \frac{1}{4} \cdot 10^{-3} \rightarrow |\delta_{x_1}| \leq \frac{3}{4} 10^{-3}$

$A_{x_2}|\delta_{x_2}| \leq \frac{1}{4} \cdot 10^{-3} \rightarrow \frac{2}{9}|\delta_{x_2}| \leq \frac{1}{4} \cdot 10^{-3} \rightarrow |\delta_{x_2}| \leq \frac{9}{8} \cdot 10^{-3}$

Avrei in questo caso basta arrotondare alla 3^ cifra decimale

$|\delta_{x_1}| \leq \frac{1}{2} \cdot 10^{-3} < \frac{3}{4} \cdot 10^{-3}$

$|\delta_{x_2}| \leq \frac{1}{2} \cdot 10^{-3} < \frac{9}{8} \cdot 10^{-3}$ → Qui potremmo avere fare il troncamento $\left(\frac{8}{8} > 1\right)$

MASTER COPY

Tel. 388/9837745