CALCOLO

CAPITOLO 1

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

RAPPRESENTAZIONE DEI NUMERI

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

Con $\beta > 1$ ogni numero non nullo $x \in \mathbb{R}$ ammette una rappresentazione in base β.

β = base

b = esponente

α_i = cifre della rappresentazione $0 \le \alpha_i \le \beta - 1$

$$\text{sign}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$$

Es. $(123)_{10} = 0,1230000\ldots \cdot 10^3$

$= 10^3 \cdot (1 \cdot 10^{-1} + 2 \cdot 10^{-2} + 3 \cdot 10^{-3} + 0 \cdot 10^{-4} + \ldots)$

$\alpha_1 = 1 \quad \alpha_2 = 2 \quad \alpha_3 = 3 \quad \alpha_4 = \alpha_5 = \ldots = 0 \quad i = 4 \ldots$

TEOREMA DI RAPPRESENTAZIONE

Data una base $\beta > 1$ e un numero reale $x \neq 0$ esiste una unica rappresentazione tale che:

1. $\alpha_1 \neq 0$
2. non vi sia un intero $k \Rightarrow \alpha_j = \beta - 1 \quad \forall j > k$

Si chiama RAPPRESENTAZIONE IN VIRGOLA MOBILE NORMALIZZATA.

CONDIZIONE 1:

$\alpha_1 \neq 0$ altrimenti avrei più di una rappresentazione.

es. $0,124 = 0,0124 \cdot 10 = 0,00124 \cdot 10^2 \ldots$

(α_1) (α_1)

CONDIZIONE 2:

ad esempio $0,\overline{9} = 1$. Se non ci fosse la seconda condizione avrei 2 rappresentazioni per 1. Invece non si possono ripetere α_j con $j > k$. Non si considerano i numeri periodici di periodo $\beta - 1$

INSIEME DI VARIAZIONE DELLA MANTISSA:

$$\frac{1}{\beta} \leq \sum_{i=1}^{\infty} \alpha_i \beta^{-i} < 1$$

↑ MANTISSA

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

MINIMO ⇒ se tutti gli α_i sono $\neq 0$ e $\alpha_1 \neq 0$
Allora $\alpha_1 = 1$ quindi minimo $= \beta^{-1} = \frac{1}{\beta}$

MASSIMO ⇒ $\sum_{i=1}^{\infty} (\beta-1)\cdot \beta^{-i} = (\beta-1) \sum_{1}^{\infty} \beta^{-i} =$

$$= (\beta-1)\left(\frac{1}{\beta} + \frac{1}{\beta^2} + \dots\right) = \frac{\beta-1}{\beta}\cdot\left(1 + \frac{1}{\beta} + \frac{1}{\beta^2}\dots\right)$$

$$= \frac{\beta-1}{\beta}\cdot\frac{1}{1-\frac{1}{\beta}} = \frac{\beta-1}{\beta}\cdot\frac{\beta}{\beta-1} = 1$$

Questo perché $\left(1 + \frac{1}{\beta} + \frac{1}{\beta^2}\dots\right)$ è una progressione geometrica che converge a $\frac{1}{1-\frac{1}{\beta}}$.

RAPPRESENTAZIONE NEL CALCOLATORE

All'interno del calcolatore si può rappresentare sono un certo numero m di cifre della mantissa di x.

La rappresentazione corrisponde a $y = tr(x)$ o a $y = rd(x)$

$y = tr(x)$ TRONCAMENTO a m cifre significative

$$y = tr(x) = sign(x)\cdot \beta^b \cdot \sum_{1}^{m} \alpha_i \beta^{-i}$$

$y = rd(x)$ ARROTONDAMENTO

$$y = rd(x) = \begin{cases} tr(x) & \text{se } 0 \leq \alpha_{m+1} < \frac{\beta}{2} \\ sign(x)\cdot\beta^b\left[\sum_{1}^{m} \alpha_i \beta^{-i} + \beta^{-m}\right] & \text{se } \frac{\beta}{2} \leq \alpha < \beta \end{cases}$$

Aggiunge uno al coefficiente α_m se α_{m+1} è maggiore di $\beta/2$ o uguale

L'ARROTONDAMENTO è in generale una migliore rappresentazione

$$|tr(x) - x| \leq \beta^{b-m} \qquad |rd(x) - x| \leq \frac{1}{2}\beta^{b-m}$$

L'unico numero non rappresentabile è lo $\emptyset$. Ne esiste una funzione $sign_0(x)$ che fa questo, cioè lo aggiunge.

$$sign_0(x) = \begin{cases} sign(x) & \text{se } x \neq 0 \\ \emptyset & \text{se } x = \emptyset \end{cases}$$

NUMERI DI MACCHINA

Si indica con M l'insieme dei numeri z rappresentabili all'interno del calcolatore.

M = insieme finito.

Fissati β, m e supposto $L \leq b \leq U$

la CARDINALITÀ di M è

$$CARD(M) = 2(\beta^m - \beta^{m-1})(U - L + 1) + 1$$

M si indica con $F = (\beta, m, L, U)$

(precisione = numero degli x_i)

Quindi non tutti i numeri sono rappresentabili fra i numeri di macchina.

Se il numero risulta troppo grande e non appartiene a M abbiamo OVERFLOW

Se è troppo piccolo abbiamo UNDERFLOW

Tutti i calcolatori segnalano overflow, alcuni fermano programma altri proseguono con

$$rd(x) = sign(x) \cdot \max_{y \in M} |y|$$

L'underflow non è sempre segnalato, a volte $rd(x) = 0$.

ERRORE ASSOLUTO della rappresentazione x

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

↑ quello che ho rappresentato — ↑ numero reale che volevo rappresentare

ERRORE RELATIVO

$$\varepsilon_x = \frac{rd(x) - x}{x} = \frac{\delta_x}{x}$$

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

$$|\delta x| = |x - rd(x)| \leq \frac{1}{2}\beta^{b-m}$$

$$|\varepsilon_x| = \frac{|\delta x|}{|x|} < \frac{\frac{1}{2}\beta^{b-m}}{\beta^{b-1}} = \frac{1}{2}\beta^{1-m}$$

(minore stretto perché il più piccolo numero lo rappresento senza errori)

$$u = \frac{1}{2}\beta^{1-m} \Rightarrow \text{PRECISIONE DI MACCHINA}$$

ERRORI CALCOLO FUNZIONE

Una funzione non razionale φ all'interno di un calcolatore viene sempre sostituita con una razionale f che comporta un errore detto di TRONCAMENTO

$$f \sim \varphi \qquad \text{ERRORE} = f - \varphi$$

Ma anche nel calcolo delle funzioni razionali in un punto non si ottiene il valore cercato a causa delle approssimazioni

$$f(P0) \Rightarrow f(P_1)$$

Abbiamo 2 tipi di errori:

- OPERAZIONI ARITMETICHE nella $f(P)$ (reale) devono essere sostituite con operazioni di macchina. $f(P)$ diventa $f_a(P)$

$$f_a \sim f \qquad \boxed{f_a \text{ approssima } f}$$

- QUANDO IL PUNTO P_0 ha coordinate non rappresentabili con i numeri di macchina e si approssima

$$\boxed{P_1 \text{ che approssima } P_0}$$

ERRORE ASSOLUTO

Si vuole calcolare ~~xxxxxxxxxx~~ la funzione $f(P)$ nel punto $P_0 \in \mathbb{R}^n$ di coordinate $x_i^{(0)}$ con $i = 1, 2, \dots, n$

$$P_0 = (x_1^{(0)}, x_2^{(2)}, \dots, x_n^{(0)})$$

Consideriamo un insieme D (detto insieme di indeterminazione del punto P_0)

$$D = \{P \in \mathbb{R}^n \mid a_i \le x_i \le b_i,\ i = 1, \dots, n\}$$

Il valore $f(P_0)$ cercato viene sostituito con $f_a(P_1)$ con P_1 di coordinate $x_i^{(1)}$, $i = 1, \dots, n$ che appartiene a D

$$P_0 \in D \qquad P_1 \in D$$

Quindi l'errore commesso è

$$f_a(P_1) - f(P_0) \quad \text{ERRORE TOTALE}$$

$$\underbrace{f_a(P_1) - f(P_0)}_{\text{ERRORE TOTALE } (\delta_f)} = \underbrace{f_a(P_1) - f(P_1)}_{\text{ERRORE ALGORITMICO } (\delta_a)} + \underbrace{f(P_1) - f(P_0)}_{\text{ERRORE TRASMESSO DAI DATI } (\delta_d)}$$

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

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

L'errore algoritmico risulta definito e stimabile, quello trasmesso possiamo dare una rappresentazione generale, se $f(P) \in C^1(D)$, dalla formula di Taylor al primo ordine.

$$\delta_d = f_a(P_1) - f(P_0) = \sum_1^n \underbrace{\frac{\partial f}{\partial x_i}}_{\gamma_i} \underbrace{(x_i^{(1)} - x_i^{(0)})}_{\delta x_i}$$

$$\delta_d = \sum_1^n \gamma_i \, \delta x_i$$

$\gamma_i \Rightarrow$ COEFFICIENTI DI AMPLIFICAZIONE di δx_i

STIMA MODULO ERRORE ASSOLUTO

$$|f_a(P_1) - f(P_0)| \le E_a + E_d \quad (E_d = \max \delta_d)$$

(E_a: max modulo δ_a)