Lezione 4

Siano β un numero intero maggiore o uguale a due ed m un numero intero positivo. Il calcolatore usa i numeri di macchina $F(\beta, m)$ per *approssimare numeri reali*, tramite la funzione arrotondamento. Studiamo *quantitativamente* l'approssimazione.

- *Definizione* (funzioni errore).

  Sia $\mathrm{rd} : \mathbb{R} \to F(\beta, m)$ la funzione arrotondamento.

  - La funzione $\delta : \mathbb{R} \to \mathbb{R}$ tale che:

    $$\delta(x) = \mathrm{rd}(x) - x$$

    si chiama *funzione errore assoluto*;
  - La funzione $\epsilon : \mathbb{R} \setminus \{ 0 \} \to \mathbb{R}$ tale che:

    $$\epsilon(x) = \frac{\mathrm{rd}(x) - x}{x} = \frac{\delta(x)}{x}$$

    si chiama *funzione errore relativo*.

  La funzione errore assoluto è *dispari*, quella errore relativo *pari*.

- *Esercizio*.

  Sia $x = 1/3$. Determinare l'errore assoluto $\delta(x)$ e l'errore relativo $\epsilon(x)$ commessi approssimando x con l'arrotondato di x in $F(10, 3)$.

- *Teorema* (stime delle funzioni errore).

  Siano $\mathrm{rd} : \mathbb{R} \to F(\beta, m)$ la funzione arrotondamento e $x = \beta^b\, g$ un numero reale positivo. Si ha:

  $$|\,\delta(x)\,| \leqslant \tfrac{1}{2}\,\beta^{b-m} \qquad , \qquad |\,\epsilon(x)\,| \leqslant \tfrac{1}{2}\,\beta^{1-m}$$

  (Dim. ...)

  - La validità delle stime si estende per simmetria al caso $x < 0$. Cosa accade per $x = 0$?

- *Definizione* (precisione di macchina).

  Si chiama *precisione di macchina* la quantità (*caratteristica dell'insieme* $F(\beta, m)$):

  $$u = \tfrac{1}{2}\,\beta^{1-m}$$

  In termini di precisione di macchina si ha:

  $$|\epsilon(x)| \leqslant u \qquad , \qquad |\delta(x)| \leqslant u\,|\,x\,|$$

- *Esempi*.

  In *Scilab* si ha $u = 2^{-53} = 10^{-15}\, 0.11102230246251565404236316680908203125$, infatti (vedere anche l'Esercitazione 1):

  ```
  -->number_properties('eps')
   ans  =

      1.110D-16
  ```

  Per la calcolatrice tascabile *HP 49G*: $u = \frac{1}{2}\, 10^{-11}$.

  Nessuno dei due insiemi $F(2, 53)$ e $F(10, 12)$ include l'altro (vedere l'Esercizio 4). Un *criterio di confronto* sensato nel contesto dell'uso dei numeri in virgola mobile per approssimare numeri reali è quello di *confrontare la precisione di macchina*: tanto più è *piccola* quanto più *piccolo* è *il massimo* errore relativo commesso arrotondando numeri reali. Per i due insiemi in esame si ha:

  $$\text{precisione di } F(2, 53) < \text{precisione in } F(10, 12)$$

  dunque: il massimo errore relativo commesso arrotondando numeri reali in $F(2, 53)$ è minore del massimo errore relativo commesso arrotondando numeri reali in $F(10, 12)$.

- La funzione errore relativo è *limitata*: per ogni numero reale x non nullo l'errore relativo commesso approssimando x con $\mathrm{rd}(x)$ non supera la precisione di macchina, quantità *indipendente da* x. La funzione errore assoluto, invece, *non è limitata*. Questa differenza (importante!) è conseguenza della struttura dell'insieme dei numeri in virgola mobile e rende *naturale* misurare l'errore commesso approssimando un numero reale con un elemento di un insieme di numeri *in virgola mobile* con la funzione errore *relativo*.

  – *Esercizio.*
    Disegnare il grafico delle funzioni $x \mapsto u$ e $x \mapsto u\,|\,x\,|$. Discutere il legame tra i grafici disegnati e quelli delle funzioni $x \mapsto |\,\epsilon(x)\,|$ e $x \mapsto |\,\delta(x)\,|$.

- *Teorema* (riscrittura dell'arrotondato di un numero reale).

  Sia $\mathrm{rd} : \mathbb{R} \to F(\beta, m)$ la funzione arrotondamento. Per ogni numero reale x esistono numeri reali d ed e tali che:

$$\mathrm{rd}(x) = x + d \qquad \text{con} \qquad |\,d\,| \leqslant u\,|\,x\,|$$

  e:

$$\mathrm{rd}(x) = x\,(1 + e) \qquad \text{con} \qquad |\,e\,| \leqslant u$$

  (Infatti: $d = \delta(x)$, $e = \epsilon(x)$ per $x \neq 0$, $e = 0$ per $x = 0$.)

  Queste *uguaglianze* significano che $\mathrm{rd}(x)$ si può *interpretare* come (piccola) *perturbazione additiva* di x oppure, rispettivamente, come (piccola) *perturbazione moltiplicativa* di x.

(B) Funzioni predefinite

Siano β un numero intero maggiore o uguale a due ed m un numero intero positivo. Si ricordi che le funzioni predefinite sono le operazioni che il nucleo interno sa fare con i numeri di macchina $F(\beta, m)$.

- *Definizione* (funzioni predefinite).

  L'insieme delle *funzioni predefinite* del nucleo interno è l'unione dei tre sottoinsiemi seguenti:

  – *Pseudo-operazioni aritmetiche*

$$\oplus, \ominus, \otimes : F(\beta, m) \times F(\beta, m) \to F(\beta, m) \qquad \text{tali che} \qquad \xi_1 \oplus \xi_2 = \mathrm{rd}(\xi_1 + \xi_2) \quad , \quad \ldots$$

    e:

$$\oslash : F(\beta, m) \times F(\beta, m) \setminus \{\,0\,\} \to F(\beta, m) \qquad \text{tale che} \qquad \xi_1 \oslash \xi_2 = \mathrm{rd}(\xi_1 / \xi_2)$$

  – *Funzioni predefinite corrispondenti alle funzioni elementari*
    Sia $f : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}$, una *funzione elementare* (ovvero: una funzione trigonometrica, esponenziale, logaritmica, radice n–esima,...). *La* funzione predefinita corrispondente ad f è la funzione ϕ da $F(\beta, m) \cap \Omega$ in $F(\beta, m)$ definita da:

$$\phi(\xi) = \mathrm{rd}(\,f(\xi)\,)$$

  – *Confronti*

$$<, \leqslant, =, \neq, \geqslant, >: F(\beta, m) \times F(\beta, m) \to \{\,\mathtt{T}, \mathtt{F}\,\}$$

    Sono le *restrizioni* ad $F(\beta, m) \times F(\beta, m)$ delle corrispondenti funzioni sui numeri reali.

  Le funzioni predefinite sono definite *nel modo migliore possibile* nel senso che 'il valore di una funzione predefinita è il numero di macchina che *dista meno* dal risultato esatto.'

- *Esempi.*

  – *Funzionamento dell'interfaccia, caso elementare*
    In *Scilab* si ottiene:

```
-->sin(86.3)
 ans  =

    -0.9956042
```

Detta `SEN` la funzione predefinita del nucleo interno di *Scilab* corrispondente alla funzione elementare sen e rd_* la funzione arrotondamento nell'insieme dei numeri reali che ammettono, in base dieci, una scrittura composta di al più dieci simboli, il valore finale mostrato da *Scilab* è così ottenuto:

(1) L'interfaccia interpreta la stringa immessa dall'utilizzatore come richiesta del calcolo del valore $\xi = \mathtt{SEN}(\,\mathrm{rd}(86.3)\,)$ – dunque interpreta la stringa `86.3` come scrittura in base dieci di un numero reale – e 'passa' la richiesta al nucleo interno;
(2) Il nucleo interno calcola il valore richiesto e lo assegna alla variabile `ans`;
(3) L'interfaccia mostra all'utilizzatore una stringa che rappresenta in base dieci il valore $\mathrm{rd}_*(\xi)$.

Questo meccanismo mette a disposizione dell'utilizzatore la funzione `sin`, definita per ogni numero reale ed a valori nell'insieme dei numeri reali che ammettono (in base dieci) una scrittura composta di al più dieci simboli, tale che:

$$\mathtt{sin}(x) = \mathrm{rd}_*(\,\mathtt{SEN}(\,\mathrm{rd}(x)\,)\,)$$

– *Proprietà delle pseudo-operazioni aritmetiche.*
Sia $M = F(10, 2)$. Si ha allora:
(1) $\oplus$ *è simmetrica* (per ogni $\xi_1, \xi_2 \in M$ si ha $\xi_1 \oplus \xi_2 = \xi_2 \oplus \xi_1$)
(2) $\oplus$ *non è associativa*: con $\xi_1 = 10^2\, 0.10$ e $\xi_2 = \xi_3 = 10^0\, 0.38$ si ha

$$(\xi_1 \oplus \xi_2) \oplus \xi_3 \neq \xi_1 \oplus (\xi_2 \oplus \xi_3)$$

(3) $\oplus$ *è debolmente monotona* (per ogni $\xi_1, \xi_2, \alpha \in M$ si ha $\xi_1 > \xi_2 \Rightarrow \xi_1 \oplus \alpha \geqslant \xi_2 \oplus \alpha$)
(4) 'lo zero non è unico': esiste *un* solo elemento $\alpha \in M$ tale che per ogni $\xi \in M$ si ha $\xi \oplus \alpha = \xi$, precisamente $\alpha = 0$. Ma: per ogni $\xi \neq 0$ esiste $\alpha \neq 0$ tale che $\xi \oplus \alpha = \xi$ (ad esempio: $10^2\, 0.67 \oplus 10^{-2}\, 0.11 = 10^2\, 0.67$).

Esercizi

---

1. Sia $x = 1/3$. Determinare l'errore assoluto $\delta(x)$ e l'errore relativo $\epsilon(x)$ commessi approssimando x con l'arrotondato di x in $F(2, 3)$. Verificare che i risultati ottenuti sono coerenti con le stime riportate nella Definizione di precisione di macchina.

2. Dimostrare, utilizzando le proprietà della funzione rd che:
   (1) Per ogni ξ in $F(10, 2)$ si ha: $\xi \oplus (-\xi) = 0$
   (2) Per ogni ξ in $F(10, 2)$ esiste un solo α in $F(10, 2)$ tale che: $\xi \oplus \alpha = 0$

3. Sia $M = F(\beta, m)$. Discutere ciascuno dei seguenti asserti:
   (1) l'errore relativo commesso approssimando $x \in \mathbb{R}$ con $\mathrm{rd}(x)$ è minore o uguale ad u;
   (2) l'errore assoluto commesso approssimando $x \in \mathbb{R}$ con $\mathrm{rd}(x)$ è minore o uguale ad 1;
   (3) se $\xi \in M$, anche $\beta^2 \xi \in M$;
   (4) se $x \in \mathbb{R}$ e $\mathrm{rd}(x) = 0$, allora $x = 0$;
   (5) gli intervalli $[\beta, \beta^2]$ e $[\beta^{10}, \beta^{11}]$ contengono lo stesso numero di elementi di M;
   (6) se ξ ed α sono due elementi positivi di M allora $\xi \oplus \alpha > \xi$;
   (7) se $x \in \mathbb{R}$ e $\xi \in M$ sono tali che $\mathrm{rd}(x) = \xi$ allora $\mathrm{rd}(\beta^{12} x) = \beta^{12} \xi$.

4. Posto $M_2 = F(2, 53)$ e $M_{10} = F(10, 12)$, dimostrare che:
   (1) Un decimo è contenuto in M_{10} ma non in M_2;
   (2) Detta u la precisione di macchina in M_2 (il cui valore espresso in base dieci è riportato sopra) si ha: $u \in M_2$ ma $u \notin M_{10}$.

   Dedurre da quanto provato che $M_2 \not\subset M_{10}$ e che $M_{10} \not\subset M_2$.