Università degli Studi di Trieste
Laurea Triennale in Intelligenza Artificiale e Data Analytics
A. A. 2024–2025
Analisi Numerica

ESERCIZI SULL'ARITMETICA DI MACCHINA

Esercizio 1

1. Si ricordi che, fissata una base (un numero naturale $B > 1$), ogni numero $x \in \mathbb{R}$ si può scrivere (rappresentazione a virgola fissa) come

$$\begin{aligned} x &= \text{sign}(x)\ (d_m \ldots d_1 d_0 . d_{-1} \ldots d_{-n} \ldots)_B \\ &= \text{sign}(x) \left(\sum_{j=0}^{m} d_j B^j + \sum_{j=1}^{\infty} d_{-j} B^{-j} \right) \end{aligned}$$

   dove $d_j, d_{-j} \in \{0, 1, \ldots, B-1\}$ sono le cifre della rappresentazione in base B (ad esempio $\{0,1\}$ in base 2, $\{0, \ldots, 9\}$ in base 10); chiamiamo $\sum_{j=0}^{n} d_j B^j$ parte intera del numero e $\sum_{j=1}^{\infty} d_{-j} B^{-j}$ parte frazionaria del numero

2. Perché la serie che rappresenta la parte frazionaria converge?
   traccia: si utilizzi il confronto con la serie geometrica di ragione $a = 1/B$, osservando che $d_{-j} \leq B - 1$ (si tratta del criterio di confronto tra serie a termini non negativi);

3. La parte frazionaria di un numero irrazionale è infinita (perché?); la parte frazionaria di un numero razionale può essere finita o infinita a seconda della base: $1/3 = (0.333\ldots)_{10}$ ma come si scrive $1/3$ in base 3?

Esercizio 2 Si ricordi che un numero reale può essere scritto in virgola mobile normalizzata in base B come:

$$x = \text{sign}(x) B^e (0.d_1 \ldots d_t \ldots)_B = \text{sign}(x) B^e \sum_{j=1}^{\infty} d_j B^{-j}$$

$d_j \in \{0, 1, \ldots B-1\}$, $d_1 \neq 0$, dove chiamiamo $\sum_{j=1}^{\infty} d_j B^{-j}$ *mantissa* e $e \in \mathbb{Z}$ *esponente* della rappresentazione; a cosa serve la normalizzazione $d_1 \neq 0$?

Si facciano esempi di numeri reali rappresentati in virgola mobile normalizzata in base $B = 2$ e $B = 10$

Esercizio 3 L'insieme dei numeri macchina si definisce (modello teorico)

$$\mathbb{F}(B, t, L, U) = \{x = \pm (0.d_1 d_2 \ldots d_t) B^e\,,\ d_j \in \{0, 1, \ldots, B-1\}\,,\ d_1 \neq 0\,, e \in [L, U] \subset \mathbb{Z}\} \cup \{0\}$$

Si studi (considerando il modello teorico):

1. $\text{card}(\mathbb{F}) = 1 + 2(B-1)B^{t-1}(U-L+1)$ (sugg.: $\mathbb{F}$ è simmetrico, $\mathbb{F}^- = -\mathbb{F}^+$; si contino le possibili mantisse e i possibili esponenti)
2. $\min \mathbb{F}^+ = B^{L-1}$ (sugg.: chi è la minima mantissa?)
3. $\max \mathbb{F}^+ = B^U(1 - B^{-t})$ (sugg.: utilizzare la somma geometrica per calcolare la massima mantissa)
4. Si rifletta sul fatto che la densità dei numeri macchina è variabile calcolando la distanza tra numeri macchina consecutivi; dove e come cambia tale densità?

Esercizio 4

1. Si maggiori, usando le serie, il massimo errore di troncamento a t cifre che si commette nell'approssimazione di un numero reale x
2. Si maggiori, usando le serie, il massimo errore di arrotondamento a t cifre che si commette nell'approssimazione di un numero reale x
3. La precisione di macchina, $u = B^{1-t}/2$, *non* è il più piccolo numero floating-point positivo (che invece è ...)

Esercizio 5 Calcolare il minimo e il massimo numero rappresentabile e la cardinalità dell'insieme di numeri macchina $\mathbb{F}(2, 4, -2, 3)$. Quanto vale la distanza relativa massima tra due numeri consecutivi ϵ_M e quanto la precisione di macchina u?

Esercizio 6 Supponendo di avere a disposizione 3 bit per l'esponente e 5 bit per la mantissa, e usando lo standard IEEE-754r

1. Si dica qual è l'insieme dei numeri macchina.
2. Si determini il minimo e il massimo numero rappresentabile, la distanza relativa massima tra due numeri consecutivi ϵ_M e la precisione di macchina u.
3. Si dica quale numero rappresenta la configurazione di bit 0 101 1011.
4. Si dia la rappresentazione binaria in virgola mobile normalizzata dei seguenti numeri reali: (a) 6.5 (b) -7.3 (c) 16
5. Quantificare gli errori di rappresentazione commessi ai punti 4.a) e 4.b).
6. Quanto vale la distanza assoluta tra $x = 6.5$ e il suo successivo numero macchina x_+?

Esercizio 7 Cosa succede se si prova a calcolare $a^2 - b^2$ con $a = 1.4 \cdot 10^{154}$ e $b = 1.3 \cdot 10^{154}$ in un calcolatore con aritmetica IEEE-754r in doppia precisione? Come si potrebbe realizzare questo calcolo in maniera stabile?

Esercizio 8 Si dimostri con un esempio che la proprietà associativa del prodotto non è verificata in aritmetica di macchina. (Suggerimento: ricordare che in $\mathbb{F}(2, 53, -1022, 1023)$ il massimo numero rappresentabile è $\approx 1.7977 \cdot 10^{308}$).

Esercizio 9 Si faccia un esempio in cui la proprietà associativa della somma in aritmetica di macchina non sia verificata per effetto di cancellazione numerica in $\mathbb{F}(10, 6, L, U)$.

Si faccia un esempio in cui la proprietà associativa della somma non è valida in aritmetica di macchina per overflow.

Esercizio 10 Si definisca $a_n = n\left(\sqrt{n^2+1} - n\right)$. Sapendo che $\lim_{n\to\infty} a_n = \frac{1}{2}$, quale sarà il valore fornito da MATLAB/OCTAVE per a_n quando $n = 10^8$? Rispondere al quesito senza calcolare a_n, e quantificare gli errori assoluto e relativo commessi assumendo come valore vero quello del limite.

Esercizio 11 Sia $x = 10^{-15}$. Si calcoli l'espressione

$$\frac{(1+x)-1}{x}.$$

Perché il risultato è meno accurato che prendendo $x = 8.88178419700125 \cdot 10^{-16}$?
Si noti che $x = 4\epsilon_M$.

Esercizio 12 Si scelga la risposta esatta per ognuno dei seguenti quesiti:

1. Si dica quanto vale la precisione di macchina u in $\mathbb{F}(2, 8, -6, 7)$, assumendo lo standard IEEE-754r.

   $\boxed{\text{A}}$ $\frac{1}{2} \cdot 2^{-8}$ $\boxed{\text{B}}$ 2^{-7} $\boxed{\text{C}}$ 2^{-8}

2. Il numero $1 + eps$ nell'aritmetica IEEE 754-r viene rappresentato come:

   $\boxed{\text{A}}$ $\boxed{0|01111111111|1\overbrace{00\ldots0000}^{\text{50 zeri}}1}$

   $\boxed{\text{B}}$ $\boxed{0|11111111111|\overbrace{00\ldots0000}^{\text{51 zeri}}1}$

   $\boxed{\text{C}}$ $\boxed{0|01111111111|\overbrace{000\ldots0000}^{\text{51 zeri}}1}$

3. Si dica quanto vale la distanza relativa massima tra due numeri consecutivi ϵ_M in $\mathbb{F}(2, 4, -2, 3)$, assumendo lo standard IEEE-754r.

   $\boxed{\text{A}}$ 2^{-3} $\boxed{\text{B}}$ $\frac{1}{2} \cdot 2^{-3}$ $\boxed{\text{C}}$ 2^{-2}

Esercizio 13 Per alcuni valori di x la funzione reale di variabile reale $f(x) = \sqrt{x^2+1} - x$ non può essere calcolata in maniera accurata in un calcolatore; quali? Si spieghi il perchè con un esempio.
Come si potrebbe risolvere il problema?

Esercizio 14 Si dica per quali valori di x la formula $f(x) = \frac{1}{\sqrt{x+2}-\sqrt{x}}$ soffre di cancellazione numerica. Si scriva una formula alternativa stabile.

Esercizio 15 Data la seguente successione di integrali definiti

$$I_n = \int_0^1 \frac{x^n}{x+5}\,dx,$$

1. Si calcoli I_0.
2. Si consideri la seguente formula ricorsiva per il calcolo di I_n:

$$\begin{aligned} s_0 &= \ln(1.2) \\ s_n &= \frac{1}{n} - 5s_{n-1} \qquad n > 0 \end{aligned}$$

   Si studi la stabilità di tale formula esprimendo l'errore al passo n, $|e_n|$ in funzione dell'errore iniziale $|e_0|$.
3. Si ricavi una formula stabile per il calcolo della successione I_n giustificando la risposta.

Esercizio 16 Sia data una funzione (derivabile) $y = f(x)$. Se il dato di ingresso x è perturbato di una quantità Δx, e detto $\Delta y = f(x + \Delta x) - f(x)$ l'errore assoluto sul valore della funzione, si dimostri che quello relativo verifica, per Δx "piccolo"

$$\left|\frac{\Delta y}{y}\right| = K(f,x)\frac{|\Delta x|}{|x|}$$

dove $K(f,x) = \dfrac{|x \cdot f'(x)|}{|y|}$ è detto Numero di condizionamento.

Esercizio 17 Dimostrare che il numero di condizionamento del problema di calcolare la funzione $f(x) = \sqrt{x}$ è $K(f,x) = \frac{1}{2}$.

Esercizio 18 Si dica per quali valori di x risulta malcondizionato il problema di calcolare

a) $f(x) = 4x^3 + 2x^2 - 4x$

b) $f(x) = 3x^2 + 10x$.

Esercizio 19 Date le funzioni

$$f_1(x) = 1 - \sqrt{1 - x^2}, \qquad f_2(x) = 1 - x$$

se ne calcoli analiticamente il condizionamento.

- Per quali valori di x le rispettive funzioni saranno malcondizionate?
- Per quali valori di x la funzione f_1 presenta invece problemi di cancellazione numerica?