Claudio Marsan

Anche il calcolatore commette errori?

Liceo cantonale di Mendrisio

anno scolastico 1992-1993

Prima di iniziare a studiare i metodi numerici è bene chiarire agli studenti alcuni punti a proposito dell'uso del calcolatore (intendendo con calcolatore l'insieme composto dalle apparecchiature hardware e dal software) per la risoluzione di problemi matematici:

(a) Il calcolatore è solo una "stupida" macchina che contiene programmi perfettibili: sa lavorare molto in fretta e per molto tempo, senza lamentarsi troppo;

(b) Il calcolatore non è perfetto: oltre agli errori di trasmissione, statisticamente quasi nulli, ci sono gli errori dovuti alle limitazioni dei programmi e alle limitazioni hardware;

(c) Il calcolatore esegue quasi tutto ciò che gli viene ordinato, senza capire quello che fa e soprattutto senza conoscere le vere intenzioni dell'utente. È quindi necessario che l'utente sappia precisamente cosa ha ordinato al calcolatore e che sappia anche valutare se il compito è stato (o potrebbe essere stato) svolto correttamente (o eventualmente con quale grado di precisione);

(d) Il software matematico si può classificare in due grosse categorie: software per la matematica numerica (per esempio: Matlab, Mathcad, Gauss, ...) e software per la matematica simbolica (Derive, Mathematica, Maple V, Macsyma, ...).
Se per la prima categoria sono ben noti i tipi di errori (dovuti essenzialmente alla rappresentazione dei numeri reali), per la seconda categoria gli errori possono essere anche di altro tipo (per esempio nel calcolo di limiti, derivate e integrali particolari). Se possibile è meglio verificare i risultati (mediante operazioni inverse, rappresentazioni grafiche o altro).

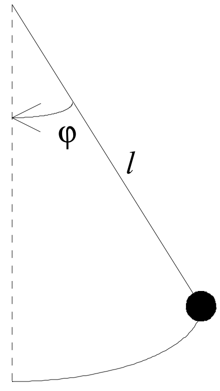
Risolvendo un problema di tipo numerico (legato per esempio a delle misurazioni di laboratorio, o a una simulazione, o a una serie di dati) si possono commettere tre tipi di errori:

- *errori inerenti*: sono errori di descrizione matematica, compresi gli errori dei dati;
- *errori di metodo*: sono dovuti ad una precisione insufficiente del metodo utilizzato (la soluzione esatta del problema posto richiede un gran numero (leggi: illimitato!) di operazioni aritmetiche, cosa che implica necessariamente delle approssimazioni);
- *errori di calcolo*: sono dovuti ad errori di arrotondamento commessi sui dati in entrata, sugli operandi e sui dati in uscita.

Esempio. Un pendolo di lunghezza l si mette in moto all'istante di tempo $t = t_0$. L'oscillazione del pendolo è descritta dall'equazione differenziale

$$l\varphi'' + g\sin\varphi + \mu\varphi' = 0 \quad \textbf{(1)}$$

dove g è l'accelerazione di gravità e μ è il coefficiente d'attrito. Qual è la posizione del pendolo nell'istante di tempo $t = t_1$?



Accettando la **(1)** come modello, la soluzione del problema comporterà un errore inerente, poiché in realtà la relazione tra l'attrito e la velocità non è lineare e inoltre le definizioni di $l, g, \mu, t_0, \varphi(t_0), \varphi'(t_0)$ sono sorgenti d'errore. Questo errore è inevitabile: solo una descrizione più rigorosa del fenomeno fisico e una definizione più precisa dei parametri permettono di ridurlo.

L'equazione differenziale **(1)** non è risolvibile esplicitamente: bisogna quindi ricorrere a dei metodi numerici, che ci daranno inevitabilmente un errore supplementare (l'errore di metodo).
Visto che i numeri con i quali si opererà avranno necessariamente un numero finito di cifre, avremo anche un errore di calcolo. ☑

Generalmente non è compito dei matematici occuparsi degli errori inerenti (di ciò si occupano prevalentemente coloro che forniscono ai matematici il modello del problema da risolvere: fisici, chimici, ingegneri, biologi, economisti, ...) ma è bene che essi conoscano perlomeno l'ordine di grandezza di tale errore: in caso contrario che senso ci sarebbe ad utilizzare nella risoluzione del problema un algoritmo che fornisce una soluzione con un errore molto più piccolo dell'errore inerente?

Il calcolatore è uno strumento in grado di risolvere problemi aritmeticamente complicati in tempi brevi, grazie soprattutto alla sua capacità di svolgere molto velocemente le operazioni aritmetiche fondamentali. Quanto velocemente può calcolare un calcolatore?

Esercizio 1. Il seguente programma, scritto in TURBO PASCAL, permette di stabilire quanti secondi impiega il calcolatore per svolgere *N* moltiplicazioni di due numeri di tipo REAL, DOUBLE, EXTENDED.

```
PROGRAM SpeedTest;  (* Claudio Marsan, 04.01.1993 *)

{-------------------------------------------------------------------------}
{ Programma per determinare il numero di moltiplicazioni al secondo eseguite }
{ dal calcolatore sul quale gira il programma                                }
{-------------------------------------------------------------------------}

{$N+,$E+}

USES
  Crt,
  Dos;

VAR
  i                    : LONGINT;
  c1,c2,c              : REAL;
  d1,d2,d              : DOUBLE;
  e1,e2,e              : EXTENDED;
  h,m,s,s100           : WORD;
  StartClock,StopClock : REAL;

  PROCEDURE ClockOn;
  { Attiva il "cronometro" }
  BEGIN
    GetTime(h,m,s,s100);
    StartClock:=3600*h+60*m+s+0.01*s100;
  END;

  PROCEDURE ClockOff;
  { Disattiva il "cronometro" }
  BEGIN
    GetTime(h,m,s,s100);
    StopClock:=3600*h+60*m+s+0.01*s100;
```

```
  END;

  PROCEDURE ElapsedTime (st : STRING);
  { Ritorna il messaggio "st" e il tempo trascorso in secondi }
  BEGIN
    Write(st,(StopClock-StartClock):5:2);
  END;

BEGIN
  ClrScr;            { pulisce lo schermo }
  Randomize;         { inizializza il generatore di numeri pseudocasuali }
  WriteLn('Numero di cicli: ',N);
  WriteLn;
  WriteLn('           sec.   molt./sec');
  WriteLn;

  c1:=1234.56*Random;           { genera un REAL casuale }
  c2:=7890.12*Random;           { genera un REAL casuale }
  ClockOn;                      { attivazione "cronometro" }
  FOR i:=1 TO N DO c:=c1*c2;    { esecuzione del ciclo di N moltiplicazioni }
  ClockOff;                     { arresto "cronometro" }
  ElapsedTime('REAL:     ');    { scrive il tempo trascorso }
  WriteLn(Round(N/(StopClock-StartClock)):10);

  d1:=123456.56*Random;         { genera un DOUBLE casuale }
  d2:=789078.12*Random;         { genera un DOUBLE casuale }
  ClockOn;                      { attivazione "cronometro" }
  FOR i:=1 TO N DO d:=d1*d2;    { esecuzione del ciclo di N moltiplicazioni }
  ClockOff;                     { arresto "cronometro" }
  ElapsedTime('DOUBLE:   ');    { scrive il tempo trascorso }
  WriteLn(Round(N/(StopClock-StartClock)):10);

  e1:=1324234.56*Random;        { genera un EXTENDED casuale }
  e2:=2789340.12*Random;        { genera un EXTENDED casuale }
  ClockOn;                      { attivazione "cronometro" }
  FOR i:=1 TO N DO e:=e1*e2;    { esecuzione del ciclo di N moltiplicazioni }
  ClockOff;                     { arresto "cronometro" }
  ElapsedTime('EXTENDED: ');    { scrive il tempo trascorso }
  WriteLn(Round(N/(StopClock-StartClock)):10);

  WriteLn;
  Write('ok!');
  ReadLn;
END.
```

(a) Dopo aver scelto un N sufficientemente grande (ossia tale da avere dei tempi dell'ordine di qualche secondo; per esempio provare con $N = 100'000$) eseguire il programma 10 volte e registrare i dati. Riportare poi tali dati su un foglio di EXCEL e ricavarne delle medie.

(b) Ripetere ciò che si è fatto nel punto precedente per l'addizione, la sottrazione e la divisione (modificare leggermente il programma). ☺

Il calcolatore è una macchina finita e quindi può rappresentare solamente numeri con un numero finito di cifre: numeri irrazionali, quali π, e, $\sqrt{3}$, non possono dunque essere rappresentati esattamente. Il calcolatore generalmente lavora in base 2 e ciò crea altri problemi poiché esistono numeri che in base 10 hanno un numero finito di cifre ma che in base 2 hanno infinite cifre, come per esempio 0,2 che in base 2 diventa periodico: $0,\overline{0011}$.

Il numero di numeri rappresentabile su un calcolatore è finito: ciò significa che non tutti i numeri sono rappresentabili esattamente, in particolare non potranno essere rappresentati numeri arbitrariamente grandi (possibile *errore di overflow*) e numeri arbitrariamente piccoli (possibile *errore di underflow*) .

Esercizio 2. Utilizzando i programmi TURBO PASCAL e QBASIC generare degli errori di overflow e degli errori di underflow mediante delle semplici operazioni aritmetiche. ☺

Utilizzando una base con un numero finito di cifre dovremo utilizzare anche un'aritmetica finita che sarà necessariamente approssimata: ciò può portare a brutte sorprese. Non bisogna quindi accettare ciecamente i risultati calcolati da una macchina (come si potrà constatare risolvendo i prossimi esercizi utilizzando del software vario, come TURBO PASCAL, QBASIC, EXCEL, DERIVE, MATHEMATICA, MAPLE V, MATLAB, MATHCAD, ...) ma avere sempre un occhio critico sia sull'input che sull'output (gli informatici hanno una regola d'oro: *garbage in, garbage out*): si eviteranno risultati a dir poco strani (probabilità maggiori di 1, varianze negative, capillari del diametro di 60 metri, ...).

Esercizio 3. Sono dati i vettori

$$\vec{a} = (2{,}718281828 \quad -3{,}141592654 \quad 1{,}414213562 \quad 0{,}5772156649 \quad 0{,}3010299957)^t$$

$$\vec{b} = (1486{,}2497 \quad 878366{,}9879 \quad -22{,}37492 \quad 4773714{,}647 \quad 0{,}000185049)^t .$$

Calcolare il prodotto scalare $\vec{a} \cdot \vec{b}$ utilizzando vari programmi. Cosa puoi notare? ☺

Esercizio 4. È dato il polinomio

$$p(x) = 8118x^4 - 11482x^3 + x^2 + 5741x - 2030.$$

Valutare il polinomio nel punto $x = 0{,}707107$ utilizzando vari programmi. Cosa puoi notare? ☺

Esercizio 5. Risolvere il seguente sistema di equazioni lineari

$$\begin{pmatrix} 64919121 & -159018721 \\ 41869520{,}5 & -102558961 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

utilizzando vari programmi. Cosa puoi notare? ☺

Esercizio 6. È noto che lo sviluppo in *serie di Taylor* di e^x è dato da

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

(a) Calcolare, utilizzando vari programmi e^{-14} arrestandosi a $k = 50$ nella serie di Taylor. Cosa puoi notare?

(b) Come sopra, ma sfruttando all'uguaglianza $e^{-14} = \dfrac{1}{e^{14}}$. Cosa puoi notare? ☺

Esercizio 7. Consideriamo un sistema di n equazioni lineari in n incognite. Utilizzando i dati dell'esercizio 1, valutare il costo computazionale necessario per risolvere con il calcolatore il sistema con il *metodo di Cramer*, per $n = 3,4,5,10,11,12,13,14,15,20,30,40,50$. Confrontare i tempi di calcolo del metodo di Cramer con quelli del *metodo di eliminazione di Gauss*, sapendo che in quest'ultimo metodo sono richieste $\frac{1}{3}n^3 + n^2 - \frac{1}{3}n$ moltiplicazioni. Cosa puoi concludere? ☺

Esercizio 8. Ci sono vari metodi per determinare il valore di π:

- *metodo di Archimede*: il valore di π è approssimato successivamente con le aree delle superfici dei poligoni regolari aventi 2^n $(n \geq 2)$ e inscritti nella circonferenza di raggio 1.

- *metodo di Vieta*: il valore di π è approssimato successivamente con la metà delle lunghezze dei perimetri dei poligoni regolari aventi 2^n $(n \geq 2)$ e inscritti nella circonferenza di raggio 1.

- *metodo di Wallis*: il valore di π è approssimato dalla relazione
$$\frac{\pi}{2} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdot \cdots \frac{2n}{2n-1} \cdot \frac{2n}{2n+1}.$$

- *metodo di Gregory e Leibniz*: il valore di π è approssimato utilizzando lo sviluppo in serie della funzione $\arctan x$:
$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots.$$
Ponendo infatti $x = 1$ si ottiene:
$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots.$$

- *metodo di Machin*: il valore di π è approssimato dalla relazione
$$\pi = 16 \arctan \frac{1}{5} - 4 \arctan \frac{1}{239} .$$

- *metodo dello sviluppo in serie di* $\arcsin x$: il valore di π è approssimato utilizzando lo sviluppo in serie della funzione $\arcsin x$:
$$\arcsin x = x + \frac{1}{2} \cdot \frac{x^3}{3} + \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{x^5}{5} + \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot \frac{x^7}{7} + \cdots.$$
Ponendo infatti $x = \dfrac{1}{2}$ si ottiene:
$$\pi = 6\left(\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2^3} + \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{1}{5} \cdot \frac{1}{2^5} + \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot \frac{1}{7} \cdot \frac{1}{2^7} + \cdots \right).$$

- *metodo dell'integrazione numerica*: il valore di π è calcolato approssimando $\frac{\pi}{4}$ con la somma s_n delle aree delle superfici dei trapezi di altezza $\frac{1}{n}$ inscritti in un quarto di cerchio di raggio unitario.

- *metodo Montecarlo*: il valore di π è approssimato nel modo seguente. Si generano n coppie di numeri (pseudo-)casuali compresi tra 0 e 2; indicando con i_n il numero di punti di coordinate (x,y) interni al cerchio inscritto nel quadrato di lato 2 avremo che
$$\pi \approx 4 \cdot \frac{i_n}{n} .$$

Studiare l'errore di metodo e l'errore di calcolo per i diversi metodi aiutandosi con vari programmi; rappresentare graficamente gli errori; studiare il costo computazionale e la velocità di convergenza per ottenere la stessa precisione (per esempio 5 cifre decimali, dieci cifre decimali) con i diversi metodi illustrati. ☺

È molto importante sapere che i numeri macchina non formano un campo, in particolare non valgono le proprietà commutativa, associativa e distributiva per addizione e moltiplicazione. Bisogna quindi prestare molta attenzione all'ordine con cui si eseguono le operazioni aritmetiche (nota: naturalmente questa osservazione non va applicata a quei programmi di matematica simbolica, quali DERIVE, MATHEMATICA E MAPLE V).

Esercizio 9. Utilizzando TURBO PASCAL e QBASIC verificare che i numeri macchina non formano un campo. ☺

In talune situazioni ci sono degli algoritmi da preferire ad altri per motivi di precisione di calcolo, di velocità di esecuzione, di stabilità rispetto a piccole perturbazioni dei dati, in altre situazioni non ci sono algoritmi sufficientemente veloci per risolvere un problema (per esempio non si conoscono algoritmi di fattorizzazione di tipo polinomiale: ciò rappresenta la fortuna della crittografia a chiave pubblica poiché per fattorizzare un intero di 200 cifre, prodotto di due primi della stessa grandezza, occorrerebbero migliaia di anni con il migliore algoritmo e con i migliori calcolatori).

Esercizio 10. Utilizzando EXCEL, TURBO PASCAL e QBASIC valutare il polinomio
$$p(x) = 3x^5 - 4x^4 + x^3 - x^2 + 6x + 1$$
nei punti $x_1 = 0{,}003$, $x_2 = -1{,}123$, $x_3 = 434{,}12$ in due modi: mediante sostituzione diretta e mediante l'applicazione della *regola di Horner-Ruffini*. Confrontare i risultati ottenuti e valutare il costo computazionale dei due metodi. ☺

Conclusione. Abbiamo visto che il calcolatore non è una macchina perfetta: è sempre compito dell'utente assicurarsi che i dati in ingresso siano corretti e coerenti e che quelli in uscita siano attendibili e accettabili. È ancora compito dell'utente scegliere un algoritmo efficiente sia dal punto di vista della velocità d'esecuzione, sia dal punto di vista della precisione di calcolo. Abbiamo anche visto che lo stesso problema, risolto con software diverso, può dare risultati diversi: oltre a conoscere i limiti di potenza e di precisione di calcolo dell'hardware a disposizione è bene conoscere anche quelli del software onde evitare spiacevoli sorprese. Dopo queste osservazioni preliminari sugli strumenti di calcolo si può passare allo studio della matematica numerica con l'ausilio del calcolatore, sempre che questo venga usato con attenzione e non con superficialità.