ALGEBRA LINEARE E ANALISI NUMERICA

Aritmetica di macchina e stabilità numerica

Andrea Franceschetti – 4357070

RELAZIONE LABORATORIO 1

Esercizio 1

L'esercizio richiede di eseguire i seguenti calcoli aritmetici:

- $(a + b) + c$;
- $a + (b + c)$.

Come da specifica, vengono assegnati alle variabili a, b, c i valori di:

- $a = (d0 + 1) * 10i -> i = 0, 1, \ldots, 6$;
- $b = (d1 + 1) * 10^{20}$;
- $c = -b$;

dove d0 e d1 corrispondono rispettivamente l'ultima e la penultima cifra del numero di matricola.

Ottenuti i risultati, si può immediatamente osservare che essi differiscono. Ciò è causato dalla mancanza di associatività dell'aritmetica in virgola mobile, nonostante comunque le due operazioni siano nettamente simili.

Nel primo caso, ovvero in $(a + b) + c$, la somma tra le variabili **a** e **b** porta a un risultato molto vicino al valore di **c**, questo causa un problema di approssimazione. Quest'ultimo, dopo aver sommato c, porta a una cancellazione numerica. Tale cancellazione da origine ad una propagazione dell'errore causando di conseguenza anche una perdita di cifre significative.

Nel secondo caso, ovvero in $a + (b + c)$, nonostante **b** e **c** siano di segno opposto, ma avendo modulo identico la cancellazione non avviene, ne consegue che non si verificano errori di approssimazione.

Infine, possiamo dire che il risultato più preciso è quello della seconda operazione.

OUTPUT:

```
---------------------
Quanto i vale: 0
La variabile a: 1
La variabile b: 8e+20
La variabile c: -8e+20

La somma (a + b) + c: 0
La somma a + (b + c): 1
L'errore assoluto della somma (a + b) + c: 1
L'errore assoluto della somma a + (b + c): 0
L'errore relativo della somma (a + b) + c: 1
L'errore relativo della somma a + (b + c): 0
---------------------
Quanto i vale: 1
La variabile a: 11
La variabile b: 8e+20
La variabile c: -8e+20

La somma (a + b) + c:0
La somma a + (b + c): 11
L'errore assoluto della somma (a + b) + c: 11
L'errore assoluto della somma a + (b + c): 0
L'errore relativo della somma (a + b) + c: 1
L'errore relativo della somma a + (b + c): 0
---------------------
Quanto i vale: 2
La variabile a: 111
La variabile b: 8e+20
La variabile c: -8e+20

La somma (a + b) + c: 0
La somma a + (b + c): 111
L'errore assoluto della somma (a + b) + c: 111
L'errore assoluto della somma a + (b + c): 0
L'errore relativo della somma (a + b) + c: 1
L'errore relativo della somma a + (b + c): 0
---------------------
Quanto i vale: 3
La variabile a: 1111
La variabile b: 8e+20
La variabile c: -8e+20

La somma (a + b) + c: 0
La somma a + (b + c): 1111
L'errore assoluto della somma (a + b) + c: 1111
L'errore assoluto della somma a + (b + c): 0
L'errore relativo della somma (a + b) + c: 1
L'errore relativo della somma a + (b + c): 0
---------------------
Quanto i vale: 4
La variabile a: 11111
La variabile b: 8e+20
La variabile c: -8e+20

La somma (a + b) + c: 0
La somma a + (b + c): 11111
L'errore assoluto della somma (a + b) + c: 11111
L'errore assoluto della somma a + (b + c): 0
L'errore relativo della somma (a + b) + c: 1
L'errore relativo della somma a + (b + c): 0
```

…. Altri output nella cartella /output

Esercizio 2

L'esercizio richiedeva di calcolare l'**errore Assoluto** e **Relativo**, di determinati punti x, tramite due algoritmi, infine di confrontare i risultati ottenuti.

Come da specifica, per i due algoritmi sono stati utilizzati il "**Polinomio di Taylor**" e la funzione **exp** della libreria "**math.h**".

Per i valori di x positivi pari a 0.5 e 30, il polinomio di Taylor, genera un valore vicino al valore di riferimento ottenuto con la funzione exp. Inoltre, per gradi alti, ovvero 100 e 150, il valore ottenuto con il Polinomio era molto vicino al valore di riferimento.

Infatti, questo era già stato definito dal testo del problema che afferma per il "Polinomio di Taylor": "al crescere del grado approssima sempre meglio la funzione di partenza".

A verifica di quest'affermazione, per esempio, prendiamo il valore $f(0.5) = 1.64872 = e^{0.5}$, con $f(0.5)$ definiamo il "Polinomio di Taylor" con $e^{0.5}$ il valore di riferimento della funzione exp. L'errore relativo per gradi alti è molto basso e pari a: $-2.69354e - 16$.

Invece per i valori negativi -0.5 e -30, sono state applicate due tipologie di calcolo differenti.

Il primo metodo, riprende l'esecuzione precedente, ovvero utilizzando il "Polinomio di Taylor". Quest'ultimo però, applicato direttamente a valori negativi evidenzia un'inefficacia dell'approssimazione effettuata. Inoltre, possiamo affermare che l'esecuzione del calcolo con il "Polinomio di Taylor", subisce una cancellazione durante la sommatoria del polinomio stesso, poiché la somma dei termini positivi è uguale in modulo alla somma dei termini negativi.

Possiamo prendere come esempio $x = -30$ e grado 100, il valore ottenuto con il "Polinomio di Taylor" è pari a $-4.82085e - 06$, mentre il valore di riferimento ottenuto dalla funzione exp è 9.35762e-14, conseguentemente l'errore relativo generato è $-5.15179e + 07$.

Il secondo metodo invece, consiste nel calcolare il reciproco del valore ottenuto dall'applicazione del "Polinomio di Taylor" al valore positivo. Con questo metodo si ottiene una buona approssimazione e un errore relativo minore.

Riprendiamo il caso di $x = -30$ e grado 100, l'errore Relativo è $-4.04647e - 16$. Confrontando questo risultato con quello del metodo precedente, si nota un netto miglioramento dell'approssimazione.

OUTPUT:

```
exp(0.5) = 1.64872
exp(30) = 1.06865e+13
exp(-0.5) = 0.606531
exp(-30) = 9.35762e-14

---- Algoritmo 1 ----
x= 0.5 N= 3
Fn(x) = 1.64583
Errore Assoluto = -0.00288794
Errore Relativo = -0.00175162

x= 0.5 N= 10
Fn(x) = 1.64872
Errore Assoluto = -1.27627e-11
Errore Relativo = -7.74096e-12

[ …. ]
x= 30 N= 100
Fn(x) = 1.06865e+13
Errore Assoluto = 0.00390625
Errore Relativo = 3.65532e-16

x= 30 N= 150
Fn(x) = 1.06865e+13
Errore Assoluto = 0.00390625
Errore Relativo = 3.65532e-16

[ … ]
x= -30 N= 100
Fn(x) = -4.82085e-06
Errore Assoluto = -4.82085e-06
Errore Relativo = -5.15179e+07

x= -30 N= 150
Fn(x) = -4.82086e-06
Errore Assoluto = -4.82086e-06
Errore Relativo = -5.1518e+07

[ … ]
---- Algoritmo 2 ----
x= -0.5 N= 3
1/Fn(x) = 0.607595
Errore Assoluto = 0.00106428
Errore Relativo = 0.0017547

x= -0.5 N= 10
1/Fn(x) = 0.606531
Errore Assoluto = 4.69513e-12
Errore Relativo = 7.74097e-12

[ … ]
x= -30 N= 100
1/Fn(x) = 9.35762e-14
Errore Assoluto = -3.78653e-29
Errore Relativo = -4.04647e-16

x= -30 N= 150
1/Fn(x) = 9.35762e-14
Errore Assoluto = -3.78653e-29
Errore Relativo = -4.04647e-16
```

…. Altri output nella cartella /output

Esercizio 3

L'esercizio richiede di calcolare la "**Precisione di Macchina**", in **singola** e **doppia** precisione.

La Precisione di Macchina ***eps***, si definisce come il valore positivo $\boldsymbol{eps} = \mathbf{2}^{-\boldsymbol{d}}$, dove d è il più grande intero positivo tale che $\mathbf{1} + \mathbf{2}^{-\boldsymbol{d}} > 1$ in aritmetica di macchina.

Quindi per trovare il valore corretto, basta diminuire i l valore di **d** di **1.**

OUTPUT:

```
Singola precisione: Valore di d: 23 => eps: 2^23 = 1.19209e-07
Doppia precisione: Valore di d: 52 => eps: 2^52 = 2.22045e-16
```