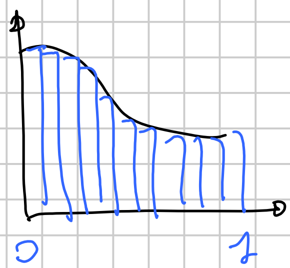
CALCOLO NUMERICO

Note Title 2021-09-27

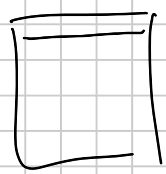
Algoritmi per problemi di analisi e algebra lineare

- problemi senza soluzione:

<u>es</u>: $\int_0^1 e^{-x^2}\,dx$



- algoritmi lenti <u>es</u>: regola di Laplace



$$\det(A) = \sum_i (-1)^{i+1} a_{i1} \det(A^{(i\,1)})$$

- algoritmi inesatti:

$$\tfrac{1}{3}+\tfrac{1}{3}+\tfrac{1}{3} = 0.333\cdots3 + 0.333\cdots3 + 0.333\cdots3 = 0.999\cdots9$$

Risolvendo sistemi lineari, errori di questo tipo vengono "amplificati"

+ programmare (Matlab)

no Mistellerlo

no cercare di portare computer giovedì

Aritmetica di macchina

Rappresentazione in base di un numero

<u>Teo</u>: Fissata una base $\beta > 1$, ogni reale $\neq 0$ si può scrivere come

$$x = \pm \beta^p \sum_{i=1}^{\infty} c_i \beta^{-i}$$

dove i c_i (cifre) sono interi $0 \le c_i < \beta$.

La scrittura è unica se aggiungo le condizioni: $c_1 \neq 0$ e che i c_i non siano tutti uguali a $\beta - 1$ da un certo punto in poi

$\beta = 10$

<u>ES</u>: $3572 = 10'000 \cdot 0{,}3572 = 10^{4}(\underbrace{3}_{c_1}\cdot 10^{-1} + \underbrace{5}_{c_2}\cdot 10^{-2} + \underbrace{7}_{c_3}\cdot 10^{-3} + \underbrace{2}_{c_4}\cdot 10^{-4})$, $p = 4$

$c_5 = c_6 = \dots = 0$

esponente

<u>ES</u>: $3.1415926\ldots = 10 \cdot 0{,}3141592\ldots = 10^{1}\underbrace{(3\cdot 10^{-1} + 1\cdot 10^{-2} + 4 \cdot 10^{-3} + \ldots}_{\text{mantissa}}$

$= 1000 \cdot 0{,}003141592\ldots = 10^{3}(\boxed{0}\cdot 10^{-1} + 0 \cdot 10^{-2} + 3\cdot 10^{-3} + 1 \cdot 10^{-4} + \ldots)$

no! prima cifra $= 0$ "non normalizzata"

$\underbrace{3571{,}99999}_{\text{nd, } \infty\ 9}\ldots = 3572$. periodico

---

<u>Def</u>: l'insieme dei numeri di macchina normalizzati (o floating-point) è l'insieme dei numeri della forma

$$\pm \beta^{p} \sum_{i=1}^{t} c_i \beta^{-i} \qquad p \in \{m, m+1, \ldots M\}$$

$$\mathbb{F}(\beta, t, m, M)$$

<u>ES</u>: $\beta = 10$, 3 cifre, esponenti $m = -2$, $M = +2$

Sono numeri di macchina ad es,

$-10^{-1} \cdot 0{,}357 = 0{,}0357$

$962 = 10^{1}(9\cdot 10^{-1} + 6 \cdot 10^{-2} + 2 \cdot 10^{-3})$

<u>non</u> lo sono:

$962 = 10^{3}(9 \cdot 10^{-1} + 6 \cdot 10^{-2} + 2 \cdot 10^{-3})$

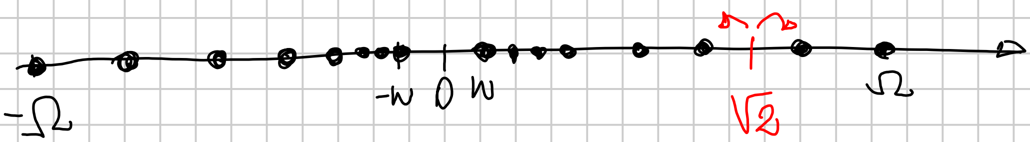
$0{,}3579 = 10^{0}(3 \cdot 10^{-1} + 5 \cdot 10^{-2} + 7 \cdot 10^{-3} + 9 \cdot 10^{-4})$

Il numero positivo più piccolo rappresentabile è

$$\omega = \beta^m\left(1\cdot\beta^{-1}+0\cdot\beta^{-2}+0\cdot\beta^{-3}+\dots+0\cdot\beta^{-t}\right) = \beta^{m-1}$$

Il num. più grande rappresentabile

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



Un computer rappresenta questi numeri di macchina + valori speciali:

- 0
- $+\infty$, $-\infty$, -0 $\quad 1/-\infty = -0$
- NaN → "codice di errore" es. $0/0$
- "numeri denormalizzati" tra $[0,\omega)$

In un computer solitamente $\beta=2$.

double, float64, binary64: $\quad$ (64 bit)

$$\beta=2 \quad t=53 \quad m=-1022 \quad M=+1023$$

$$\omega \approx 2.2\cdot 10^{-308}, \quad \Omega \approx 1.8\cdot 10^{308}$$

Numero di macchina successivo a uno dato (positivo) $(\neq \Omega)$

$$x = \beta^p\left(c_1\beta^{-1}+c_2\beta^{-2}+\dots c_t\beta^{-t}\right)$$

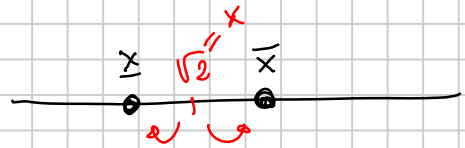
il successivo si ottiene aggiungendo 1 all'ultimo posto:

$$x+\beta^p\cdot\beta^{-t} = x+\beta^{p-t}$$

(ovvio se non lo riporti; se lo riporti, potrebbe succedere di passare a t+1 cifre è nullo ma in questo caso ultima cifra 0)

Teo: Dato un reale $x \in [-\Omega, -\omega] \cup [\omega, \Omega]$, esiste un numero di macchina $\tilde{x}$ tale che

$$\frac{|\tilde{x} - x|}{|x|} < \beta^{1-t}$$



Dim: assumiamo per semplicità x positivo, e x non numero di macchina

x compreso tra due num. di macchina successivi $\underline{x}$ $\overline{x}$

$x \in (\underline{x}, \overline{x})$

Posso prendere indifferentemente $\tilde{x} = \underline{x}$ (arrot. verso 0, troncamento)

$\tilde{x} = \overline{x}$ (arrot. verso ∞)

$\tilde{x} =$ il più vicino (round to nearest)

esponente di $\underline{x}$

$$|\tilde{x} - x| < |\overline{x} - \underline{x}| = \beta^{\boxed{p}-t}$$

stesso p

$$x = \beta^{\boxed{p}} \sum_{i=1}^{\infty} c_i \beta^{-i} \geq \beta^{p}\left(1 \cdot \beta^{-1} + 0 \cdot \beta^{-2} + 0 \cdot \beta^{-3} + \dots\right) = \beta^{p-1}$$

$$\frac{|\tilde{x} - x|}{|x|} < \frac{\beta^{p-t}}{\beta^{p-1}} = \underbrace{\beta^{1-t}} \qquad \square$$

per i double, $u \approx 2{,}2 \cdot 10^{-16}$ ("precisione di macchina")

$\frac{|\tilde{x} - x|}{|x|}$ $\quad x \quad$ $\tilde{x}$ $\quad \tilde{x} - x \quad$ 1.5 cm

1,00000

1,00001 errore sulla 6^a cifra significativa $\Leftrightarrow$ $\frac{|\tilde{x} - x|}{|x|} = 10^{-6}$

$$\varepsilon = \underbrace{\frac{\tilde{x} - x}{x}}_{\text{errore come differenza}} \Leftrightarrow \tilde{x} = x + \varepsilon x = x\underbrace{(1+\varepsilon)}_{\text{errore come prodotto}}$$

ES: In base-2, $\frac{3}{10} = 0,0\overline{1001}$

$\beta = 2, \; t = 53$

$$x = 3/10 \approx \tilde{x} = 2^{-1}\left(1\cdot 2^{-1} + 0\cdot 2^{-2} + 0\cdot 2^{-3} + 1\cdot 2^{-4} + 1\cdot 2^{-5} + 0\cdot 2^{-6} + \dots + 1\cdot 2^{-53}\right)$$

$\tilde{x} = 0,29999999999999999889 \dots$ $\qquad \frac{\tilde{x} - x}{x} < 2.2 \cdot 10^{-16} = u$

ES in Matlab

$a = 0.3$ $\qquad \tilde{a} = a(1+\varepsilon_1)$ $\qquad |\varepsilon_1| < u$

$b = 0.4$ $\qquad \tilde{b} = b(1+\varepsilon_2)$ $\qquad |\varepsilon_2| < u$

$c = a + b$

$\tilde{a} + \tilde{b}$ non è per forza un numero di macchina!

$\tilde{c} = (\tilde{a} + \tilde{b})(1+\varepsilon_3)$ = num. di macchina più vicino ad $\tilde{a} + \tilde{b}$

$= \tilde{a} \oplus \tilde{b}$

$a \ominus b \qquad a \otimes b \qquad a \oslash b$

Altro possibile tipo di errori:

overflow: $\Omega \oplus \Omega = \text{inf}$

$\Omega \circledast \Omega = \text{inf}$

underflow $\qquad \omega \oslash 1000 = 0$

Analisi dell'errore

f funzione razionale (definita solo componendo $+, -, *, /$), ad es.

$f(x) = x^2 + 3/x$

Cosa succede se provo a calcolarla con Matlab, ad es.

```
>> x = 0.3
>> y = x * x - 3 / x
```

$\rightsquigarrow \; y = x \otimes x \ominus 3 \oslash x$

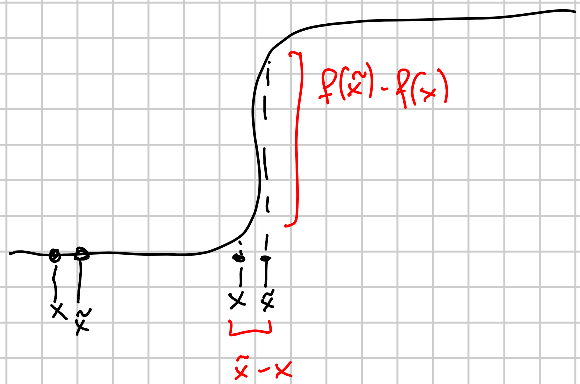
Errore inerente $\qquad f(0.3) \rightsquigarrow$ numero non rappresentabile

Computer può provare solo a calcolare $f(\tilde{x})$, dove $\tilde{x}$ è l'approssimazione di x che abbiamo a disposizione

→ numero di macchina più vicino
→ possibilmente anche altri errori (misurazioni, calcoli precedenti, ...)

$$\frac{|\tilde{x}-x|}{|x|} = \varepsilon \leq u \quad \text{(precisione di macchina)}$$

$$\boxed{\varepsilon_{in} = \frac{f(\tilde{x}) - f(x)}{f(x)}}$$



Errore algoritmico:

Nel calcolare materialmente f, devo rimpiazzare le operazioni con operazioni di macchina, ad es.

$$f(x) = x^2 + 3/x \qquad g(\tilde{x}) = \tilde{x} \circledast \tilde{x} \oplus 3 \oslash \tilde{x}$$

$$\varepsilon_{alg} = \frac{g(\tilde{x}) - f(\tilde{x})}{f(\tilde{x})}$$

$$\varepsilon_{tot} = \frac{g(\tilde{x}) - f(x)}{f(x)}$$

$f(x)$ = funzione esatta su x esatto

$f(\tilde{x})$ = funzione esatta su x approssimato

$g(\tilde{x})$ = f. approssimata su x approssimato

Nella prossima lezione:

$$\varepsilon_{tot} = \varepsilon_{in} + \varepsilon_{alg} + O(\varepsilon^2)$$