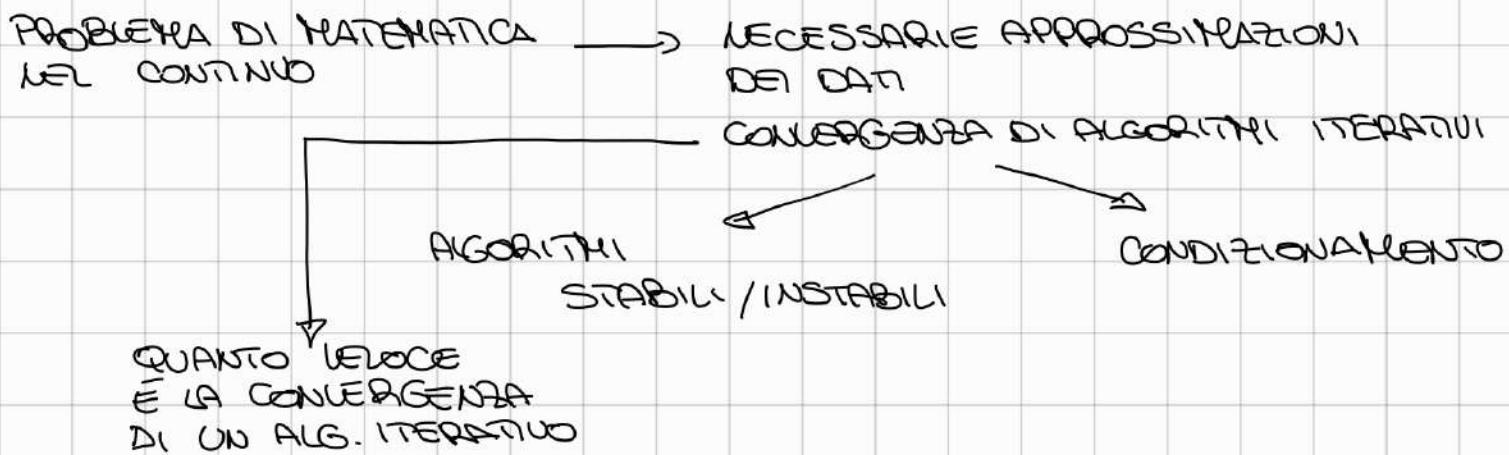


INTRODUZIONE



RAPPRESENTAZIONE DEI NUMERI IN VIRGA MOBILE

Scelto un qualunque numero intero $B > 1$, ogni numero non nullo $x \in \mathbb{R}$ ammette una rappresentazione in base B :

$$x = \text{sign}(x) B^e \sum_{j=1}^{\infty} \alpha_j B^{-j}$$

dove:

B = BASE con $B \in \mathbb{N}, B > 1$

e = ESPOLENTE con $e \in \mathbb{Z}$

α = CIFRE con $\alpha_j \in \{0, \dots, B-1\}$

$\sum_{j=1}^{\infty} \alpha_j B^{-j}$ = MANTISSA

$$\text{sign}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$$

Vale il seguente teorema:

TH Data una base intera $B > 1$ e un qualsiasi numero reale x diverso da zero, esiste un'unica rappresentazione in base B tale che:

- 1) Sia $\alpha_1 \neq 0$
- 2) Non vi sia un intero k per cui si abbia $\alpha_j = B-1, \forall j > k$

(IN SIMBOLI)

Sia $\beta \in \mathbb{N}$, $\beta > 1$, $x \in \mathbb{R} \setminus \{0\}$ allora:

$$\exists! x = \text{sign}(x) \beta^e \sum_{j=1}^{\infty} \alpha_j \beta^{-j}$$

Se:

$$1) \alpha_1 \neq 0$$

$$2) \exists k \in \mathbb{N}: \alpha_j = \beta - 1 \quad \forall j > k$$

OSS $\frac{1}{\beta} \leq \sum_{j=1}^{\infty} \alpha_j \beta^{-j} < 1 \quad \Rightarrow$

PER DHR. IN UNA BASE GENERICA SI PUÒ USARE IL SEG. RIS. SUUE SERIE GEOMETR.

$$\sum_{j=1}^{\infty} p^j = \frac{1}{1-p} - 1$$

$$p = \beta^{-1}$$

$$\sum_{j=1}^{m-1} p^j = \frac{1-p^{m+1}}{1-p} - 1$$

DEF (INSIEME DI NUMERI DI MACCHINA (FLOATING POINT NUMBERS))

Dati quattro parametri interi β, m, L, U con $\beta > 1, m > 0, L \leq U$ si definisce l'insieme finito dei numeri di macchina:

$$F(\beta, m, L, U) = \left\{ x \in \mathbb{R} : x = \text{sign}(x) \beta^e \sum_{j=1}^m \alpha_j \beta^{-j} \right\}$$

dove:

$$L \leq e \leq U$$

L = ESPOLENTE MINIMO

$$\alpha_1 \neq 0$$

U = " MASSIMO

$$\alpha_j \in \{0, \dots, \beta-1\} \cup \{0\}$$

OSS $|F(\beta, m, L, U)| = 2 \cdot \underbrace{(U-L+1)}_{\substack{\uparrow \\ \text{POSSIBILI} \\ \text{SEGANI}}} \cdot \underbrace{(\beta^m - \beta^{m-1})}_{\substack{\uparrow \\ \text{POSSIBILI} \\ \text{ESPOLENTI}}} + 1 = \text{DEI INSIGNE}$

CARDINALITÀ

POSSIBILI CIPPE

OSS Qual è il numero più grande in $F(\dots)$?

$$\beta^U \cdot (\beta-1) \sum_{j=1}^m \beta^{-j} = \boxed{\beta^U (1-\beta^{-m})}$$

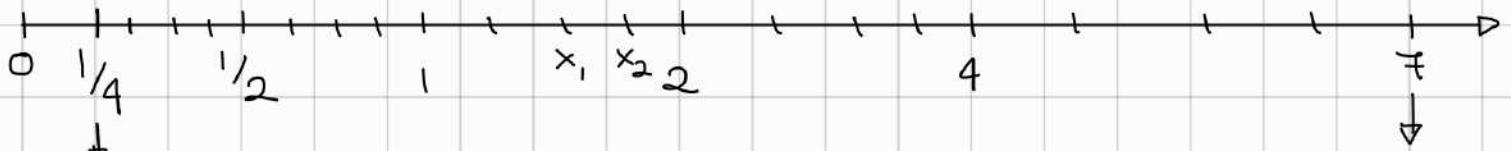
OSS Qual è il più piccolo numero positivo in $F(\dots)$?

$$\beta^L \cdot \beta^{-1} = \boxed{\beta^{L-1}}$$

Vediamo come si distribuiscono i numeri di macchina sulla retta dei reali.

ESEMPIO:

$$\beta = 2, m = 3, L = -1, U = 3$$



NUMERO PIÙ
PICCOLO RAPPRES.

$$2^{-1-1} = 2^{-2} = 1/4$$

NUMERO PIÙ GRANDE
RAPPRESENTABILE

$$2^3(1 - 2^{-3}) = 8 - 1 = 7$$

Prendiamo in considerazione intervalli $[\beta^i, \beta^{i+1}]$

$$\text{ovvero } [\frac{1}{4}, \frac{1}{2}], [\frac{1}{2}, 1], [1, 2], [2, 4], [4, 7]$$

SAREBBERE 8 res
 \neq È IL N. MAX
RAPP.

$$x = \text{sign}(x) \beta^e \sum_{i=1}^{\infty} \alpha_i \beta^{-i}$$

$\alpha_i < 1$

e il segno

Facendo variare la mantissa e tenendo fisso l'esponente si ottengono tutti i numeri di macchina che si trovano nell'intervalle $[\beta^{e-1}, \beta^e]$

I valori possibili per la mantissa equivalgono a tutti i modi in cui posso scegliere α_i . - Avremmo 3 cifre da scegliere, ma $\alpha_i = 1$ (è fissata a 1 in quanto deve essere $\neq 0$ per il teorema di cui sopra) quindi rimangono 2 cifre da scegliere (tra 0 e 1), dunque 4 combinazioni.

Ogni intervallo $[\beta^i, \beta^{i+1}]$ contiene 4 valori equispaziati all'interno del singolo intervallo; nell'intervalle successivo lo spazio tra i valori raddoppia.
La distanza tra due n. di macchina scala proporz. di β^e .

Siamo $x_1, x_2 \in F \cap [\beta^i, \beta^{i+1}]$ adiacenti

La loro distanza è:

$$|x_1 - x_2| = \beta^{i+1} \cdot \beta^{-m} = \beta^{i+1-m}$$

RAPPRESENTAZIONE IN PRECISIONE DOPPIA

β	m	L	U
2	53	-1021	1024

$$64 \text{ bit} = 8 \text{ byte}$$

Dato un numero reale, ci poniamo ora il problema di trovare un numero di macchina all'interno dell'insieme F che lo approssimi.

Sia $x \in \mathbb{R} \setminus F$

Si possono verificare 3 casi:

1) $|x| > \beta^U(1 - \beta^{-m}) \Rightarrow$ overflow: $|x|$ è maggiore del max num. rappresentabile
In genere gli si assegna $\pm\infty$

2) $|x| < \beta^{L-1} \Rightarrow$ underflow: $|x|$ è minore del min. num. rappresentabile -
In genere gli si assegna 0.

3) $\beta^{L-1} < |x| < \beta^U(1 - \beta^{-m}) \Rightarrow$ Si utilizza una funzione di ARROTONDIMENTO per scegliere con quale valore approssimarla (es. tra x_1 e x_2)



$$x_1, x_2 \in F$$

In questo caso ci sono due possibilità:

1) **TRONCAMENTO**: se x è positivo si sceglie il num. alla s.a. s_x , se è negativo il num. alla s.a. j_x

$$TR(x) = \text{sign}(x) \beta^m \sum_{j=1}^m x_j \beta^{-j}$$

Il massimo errore assoluto che può scatenare il troncamento è pari alla distanza tra due numeri di macchina:

$$|x - TR(x)| \leq \beta^{e-m}$$

2) **APPRETONDAMENTO AL PIÙ VICINO**: si valuta la prima cifra scaritata, se questa cifra è ≥ 5 (in base 10) si aggiunge un'unità all'ultima cifra troncata, altrimenti si opera come nel troncamento.

$$RD = \begin{cases} TR(x) & \text{se } x_{m+1} < \frac{\beta}{2} \\ \text{sign}(x) \beta^m \left(\sum_{j=1}^m x_j \beta^{-j} + \beta^{-m} \right) & \text{se } x_{m+1} \geq \frac{\beta}{2} \end{cases}$$

OSS Questo metodo permette di scegliere ogni volta il num. di macchina che minimizza l'errore assoluto:

$$|x - RD(x)| \leq \min |x - z| \text{ con } z \in F(\beta, m, L, U)$$

$$\leq \frac{1}{2} \beta^{e-m}$$

(cioè la metà della dist. tra due num. di macchina)

DEF (ERRORE ASSOLUTO)

$$\delta_x := |x - RD(x)|$$

DEF (ERRORE RELATIVO)

$$\varepsilon_x := \frac{|x - RD(x)|}{|x|} = \frac{\delta_x}{|x|}, x \neq 0$$

L'errore relativo dell'approssimazione di un numero reale è:

$$|\varepsilon_x| \leq \frac{\frac{1}{2} \beta^{e-m}}{\beta \cdot \beta^{-1}} = \frac{\frac{1}{2} \beta^{-m+1}}{\beta^{-m}}$$

per
 $\beta=2$

NEL CASO DELLA PREC.
 \Rightarrow DOPPIA U = $2^{-53} \approx 10^{-16}$

N.B.: L'errore relativo non dipende da e , dunque non dipende dall'intervallo in cui ci troviamo.
Dunque ε_x è uniformemente limitato dall'alto da $U = \frac{1}{2} \beta^{-m+1}$, quantità chiamata **PRECISIONE DI MACCHINA**.

IMPLEMENTAZIONE BAPP IN PRECISIONE DOPPIA

1 bit per il segno

11 bit per l'esponente

num. intero
senza segno

$$e = \hat{e} - 1022$$



52 bit (+1) per la mantissa

BIT IMPLICATO ($\alpha_1 = 1$)
NON VIENE MEMORIZZATO

VALORI SPECIALI:

- Il valore max è $\tilde{e} = 2^{l-1}$ è riservato a $\pm \infty$ e NaN (Not a Number)
- Il valore min è $\tilde{e} = 0$ è riservato ai casi di underflow che vengono gestiti tramite i **NUMERI SOTTONORMALI**

ULTERIORI RAPPRESENTAZIONI

	b	e	m
PREC. SINGOLA	2	8	23
" QUADRUPLO	2	15	112
MEZZA PRECISIONE	2	5	10
BFLOAT	2	8	8

NUMERI SOTTONORMALI (SUBNORMAL NUMBERS)

$\tilde{e} = 0$ ($e = -1022$ nel caso a precisione doppia)

Si assume implicitamente che $\alpha_1 = 0$

Di conseguenza si avranno numeri nel formato:

$$x = \text{sign}(x) b^{\sum_{j=2}^m \alpha_j b^{-j}}$$

In questo modo è possibile rappresentare numeri nell'intervallo $[b^{-m}, b^{-1}]$

In questo caso però non è più vero che l'errore relativo è sempre sup. limitato dalla precisione di macchina. All'estremo dx dell'intervallo c'è la prec. di macchina, ma più ci si avvicina all'estremo sx più i numeri hanno un errore relativo più alto. All'estremo sx l'errore relativo è 1

DEF $F \cup \{ \text{NUMERI SOTTONORMALI} \} = \text{NUM. DI MACCHINA ESTESI}$

I num. di macchina estesi presentano un underflow più graduale, però i numeri sottonormali hanno un errore relativo superiore a ϵ (in generale) -

Nell'insieme F dei numeri di macchina non tutte le proprietà delle quattro operazioni aritmetiche elementari risultano verificate, in quanto il risultato di un'operazione deve essere ricambiato a un numero di macchina.

Es. $a + b$ con $a, b \in \mathbb{R} \Rightarrow a + b \in \mathbb{R}$
 $a, b \in F \not\Rightarrow a + b \in F$

Si indicano nel seguente modo le quattro **operazioni di macchina**:

- ⊕ ADDIZIONE $a \oplus b = RD(a+b)$ secondo IEEE754
- ⊖ SOTTRAZIONE
- ⊗ MOLTIPLICAZIONE
- ⊘ DIVISIONE

L'addizione "floating point" (\oplus) non gode della proprietà associativa, mentre la moltiplicazione (\otimes) non gode della proprietà distributiva rispetto all'addizione -

Quando si sommano o sottraggono due numeri di macchina dello stesso segno che hanno lo stesso esponente e con le mantisse che differiscono di poco (quindi con le prime cifre uguali), si incorre in una perdita di cifre significative nel risultato. Tale fenomeno, detto **CANCELLAZIONE** produce una notevole amplificazione degli errori relativi. (N.B.: non si verifica nella moltiplicazione e nella divisione)

Anche nel calcolo di una funzione razionale $f(x_1, \dots, x_n) : [a_1, b_1], \dots, [a_n, b_n] \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ in un punto assegnato P_0 , in generale non si ottiene il valore $f(P_0)$ cercato, a causa delle approssimazioni che si introducono.

Tali approssimazioni producono due tipi di errore: **ASSOLUTO** e **RELATIVO**.

ERRORE ASSOLUTO

Il valore cercato $f(P_0)$ viene sostituito dal valore calcolato $f_a(P_1)$ e l'errore commesso (**ERRORE ASSOLUTO TOTALE**) risulta:

$$f_a(P_1) - f(P_0) = f_a(P_1) - f(P_1) + f(P_1) - f(P_0),$$

$$\text{ERRORE TOTALE} = \text{ERRORE ALGORITMICO} + \text{ERRORE INERENTE}$$

$$S_f = S_a + S_d$$

(ERRORE TOASHNESSO
DAI DATI)

ERRORE ALGORITMICO

Una volta fissato l'algoritmo che fornisce $f_a(P)$ risulta definito e stimabile.

Es. e^x per $x \in \mathbb{R}, x \leq 0$

Tramite la serie di Taylor approssimiamo e^x :

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \dots$$

Notiamo però che per $x \leq 0$ la serie è a segni alterni. Avviene il fenomeno della cancellazione, per cui non si ottiene un'approssimazione accorta, che si ha invece se si pone:

$$e^x = \frac{1}{e^{-x}} = \frac{1}{1 + x + \frac{x^2}{2} + \dots}$$

ERRORE INERENTE

All' errore inerente si può dare una rappresentazione generale: dalla formula di Taylor arrestata al primo termine e con punto iniziale P_0 si ottiene:

A_{xi} = COEFF. DI AMPLIFICAZIONE
(DEI' ERRORE ASSOLUTO)

$$f(P_i) = f(P_0) + \sum_{i=1}^n \underbrace{\frac{\partial f}{\partial x_i}(P_0)}_{A_{xi}} \cdot (x_i^{(i)} - x_i^{(0)}) + O(\|P_i - P_0\|^2)$$

$$\text{con } P_i = (x_1^{(i)}, \dots, x_n^{(i)})$$

$$f(P_i) - f(P_0) = \sum_{i=1}^n A_{xi} \cdot (x_i^{(i)} - x_i^{(0)})$$

$$|f(P_i) - f(P_0)| \leq \sum_{i=1}^n |A_{xi}| \cdot \underbrace{|x_i^{(i)} - x_i^{(0)}|}_{\sim x_i \cdot u}$$

\hookrightarrow PRECISIONE DI MACCHINA

- PROBLEMA DIRETTO: dato un algoritmo e delle limitazioni a $|P_i - P_0|$, determinare δ tale che $|f_a(P_i) - f(P_0)| \leq \delta$
- PROBLEMA INVERSO: dato $\delta > 0$ e $|P_i - P_0| \leq \dots$, trovare $f_a(P_i)$ tale che $|f_a(P_i) - f(P_0)| \leq \delta$

Esercizio: (Problema diretto)

Siamo:

$$f(x_1, x_2) = \frac{x_1}{x_2} \quad D = [1, 3] \times [4, 5]$$

Determinare s_f .

Calcoliamo i coefficienti di amplificazione:

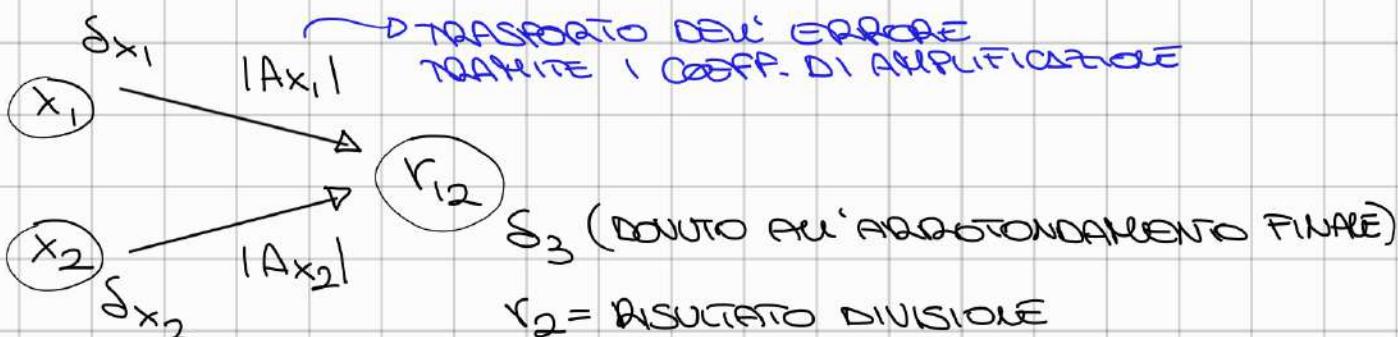
$$\left| \frac{\partial f}{\partial x_1} \right| = \frac{1}{x_2}; \quad \left| \frac{\partial f}{\partial x_2} \right| = -\frac{x_1}{x_2^2}$$

$$|A_{x_1}| = \sup_{(x_1, x_2) \in D} \left| \frac{\partial f}{\partial x_1} \right| = \frac{1}{4} \quad (\text{cioè } |A_{x_1}| \text{ ASSUME IL VALORE MASSIMO PER } x_2=4)$$

$$|A_{x_2}| = \sup_{(x_1, x_2) \in D} \left| \frac{\partial f}{\partial x_2} \right| = \frac{3}{16} \quad (\text{cioè } |A_{x_2}| \text{ ASSUME IL VALORE MASSIMO PER } x_1=3, x_2=4)$$

$$s_f \leq |A_{x_1}| \cdot s_{x_1} + |A_{x_2}| \cdot s_{x_2}$$

Calcoliamo l'errore relativo alla divisione:



Supponendo $|s_{x_i}| \leq 10^{-2}$:

$$s_f = s_d + s_a \leq \frac{1}{4} \cdot 10^{-2} + \frac{3}{16} \cdot 10^{-2} + \frac{1}{2} \cdot 10^{-2}$$

$$\leq \frac{15}{16} \cdot 10^{-2}$$

$s_a = \text{ERRORE ALGORITMICO}$
 $\text{DOVUTO ALLA FUNZIONE}$
 (DIVISIONE)

$s_d = \text{ERRORE INERENTE}$
 $\text{DERIVANTE DA PERTURBAZIONI}$
 SUI DATI INIZIALI

ERRORE RELATIVO (totale)

L'errore relativo che si commette nel calcolo di una funzione $f(P)$ in un assegnato punto P_0 è definito da:

$$\varepsilon_f = \frac{f_a(P_1) - f(P_0)}{f(P_0)} = \frac{\delta f}{f(P_0)}$$

$$\varepsilon_f = \underbrace{\frac{f_a(P_1) - f(P_1)}{f(P_0)}}_{\text{ERRORE RELATIVO TOTALE}} + \underbrace{\frac{f(P_1) - f(P_0)}{f(P_0)}}_{\varepsilon_d \text{ (ERRORE RELATIVO INERENTE)}} =$$

$$= \varepsilon_d + \underbrace{\frac{f_a(P_1) - f(P_1)}{f(P_1)} \cdot \frac{f(P_1)}{f(P_0)}}_{\varepsilon_a \text{ (ERRORE RELATIVO ALGORITMICO)}} =$$

$$= \varepsilon_d + \varepsilon_a \left(\frac{f(P_0) + f(P_1) - f(P_0)}{f(P_0)} \right) =$$

$$= \varepsilon_d + \varepsilon_a (1 + \varepsilon_d)$$

$$\varepsilon_f = \varepsilon_a + \varepsilon_d + \varepsilon_a \varepsilon_d$$

$\gamma_i = \frac{\text{COEFF. DI ANPL.}}{\text{RELATIVI}}$

$$\varepsilon_d = \frac{f(P_1) - f(P_0)}{f(P_0)} = \sum_{i=1}^n \frac{x_i^{(0)}}{f(P_0)} \frac{\partial f}{\partial x_i} (\xi_i) \cdot (x_i^{(1)} - x_i^{(0)}) =$$

TAYLOR

$$= \sum_{i=1}^n \gamma_i \varepsilon_{x_i}$$

NUERO DI CONDIZIONAMENTO

- Se ε_d è "piccolo" il problema è **BEN CONDIZIONATO**: piccole variazioni dei dati iniziali causano piccole variazioni dei dati finali - (viceversa ε_d "grande" \Rightarrow problema **MAL CONDIZIONATO**)

- Se ϵ_a è "piccolo" il metodo numerico (l' algoritmo) è **STABILE** (viceversa ϵ_a "grande" \Rightarrow **ALG-INSTABILE**)

ERRORE NELLE OPERAZIONI ARITMETICHE FLOATING POINT

OPERAZIONE	δ_f	ϵ_f
$x \oplus y$	$\delta_x + \delta_y$	$\frac{x}{x+y} \epsilon_x + \frac{y}{x+y} \epsilon_y$
$x \ominus y$	$\delta_x - \delta_y$	$\frac{x}{x-y} \epsilon_x - \frac{y}{x-y} \epsilon_y$
$x \otimes y$	$y \delta_x + x \delta_y$	$\epsilon_x + \epsilon_y$
$x \oslash y$	$\frac{1}{y} \delta_x - \frac{x}{y^2} \delta_y$	$\epsilon_x - \epsilon_y$

Si deduce che le operazioni di addizione e sottrazione non danno problemi per quanto riguarda l'errore assoluto, mentre possono rendere grande l'errore relativo nel caso in cui i due termini dell'operazione siano molto vicini in valore assoluto (con il conseguente verificarsi del fenomeno della cancellazione).

La moltiplicazione non amplifica l'errore relativo e comporta un errore assoluto che dipende dall'ordine di grandezza dei fattori.

Anche la divisione non produce amplificazione per quanto riguarda l'errore relativo, mentre l'errore assoluto diminuisce se aumenta (in valore assoluto) il divisore.

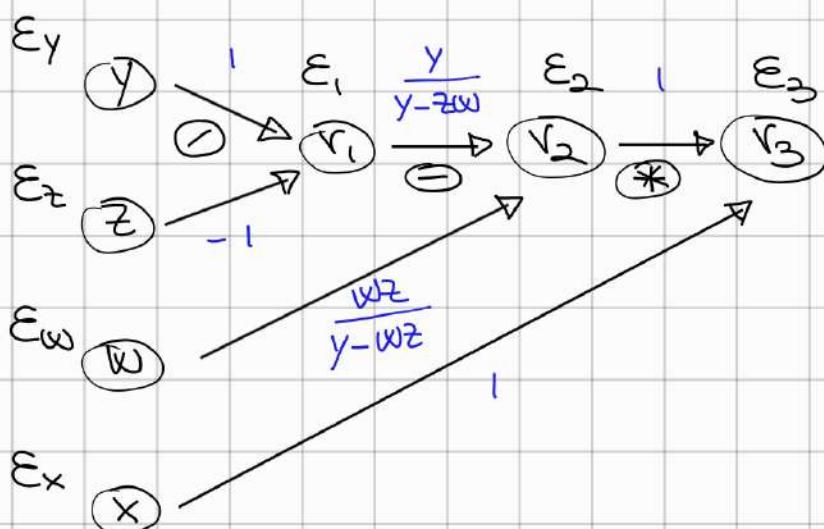
Esercizio:

Si vuole stimare l'errore relativo commesso nel calcolo della funzione $f(x, y, z, w) = x \left(\frac{y}{z} - w \right)$.

Ricordando all'uso dei grafici si calcola la funzione eseguendo un'operazione dopo l'altra e stabilendo per ciascuna di esse l'entità degli errori relativi.

SEQUENZA OPERAZIONI:

$$v_1 = \frac{y}{z}; \quad v_2 = v_1 - w; \quad v_3 = x v_2$$



■ COEFF. DI APERLIFICAZIONE RELATIVI

$$\epsilon_{v_1} = (\epsilon_y - \epsilon_z) + \epsilon_i$$

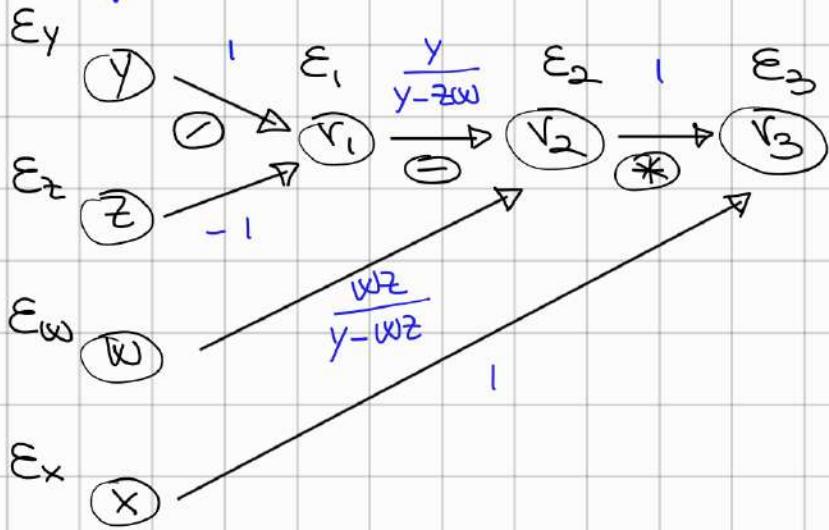
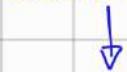
$$\epsilon_{v_2} = \frac{y}{y-zw} \epsilon_{v_1} - \frac{wz}{y-wz} \epsilon_w + \epsilon_2$$

$$\epsilon_{v_3} = \epsilon_{v_2} + \epsilon_x + \epsilon_3 =$$

$$= \frac{y}{y-zw} (\epsilon_y - \epsilon_z + \epsilon_i) - \frac{wz}{y-wz} \epsilon_w + \epsilon_2 + \epsilon_x + \epsilon_3$$

$$f(x, y, w, z) = x \cdot \left(\frac{y}{z} - w \right)$$

QTA' CON UN EVENTUALE
CORRISPONDENTE INIZIALE



ORDINE OPERAZIONI:

$$r_1 = \frac{y}{z}$$

$$r_2 = (r_1 - w)$$

$$r_3 = x \cdot r_2$$

$$\varepsilon_f = \varepsilon_a + \varepsilon_d$$

$$\textcircled{1} \quad \varepsilon_d = \sum_{i=1}^n \gamma_{x_i} \cdot \varepsilon_{x_i} \quad \text{con} \quad \gamma_{x_i} = \frac{x_i}{f(x)} \cdot \frac{\partial f}{\partial x_i}$$

$$\begin{aligned} \varepsilon_d &= \frac{x \left(\frac{y}{z} - w \right)}{x \left(\frac{y}{z} - w \right)} \varepsilon_x + \frac{y \cdot \frac{x}{z}}{x \left(\frac{y}{z} - w \right)} \varepsilon_y + \frac{w(-x)}{x \left(\frac{y}{z} - w \right)} \varepsilon_w + \\ &+ \frac{z \cdot xy \left(-\frac{1}{z^2} \right)}{x \left(\frac{y}{z} - w \right)} = \end{aligned}$$

$$= \varepsilon_x + \frac{y}{y-wz} (\varepsilon_y - \varepsilon_z) - \frac{wz}{y-wz} \varepsilon_w$$

\textcircled{2} Per calcolare ε_a si può usare il metodo del grafo
consid. i dati di partenza esatti ($\varepsilon_x = \varepsilon_y = \varepsilon_w = \varepsilon_z = 0$)

$$\varepsilon_a = \varepsilon_{r_3} = \varepsilon_3 + \varepsilon_{r_2} = \varepsilon_3 + \varepsilon_2 + \frac{y}{y-wz} \varepsilon_1$$

ESEMPIO

$$f(x, y) = x^2 - y^2$$

ALGORITMO 1:

$$z_1 = x \cdot x$$

$$z_2 = y \cdot y$$

$$z_3 = z_1 \cdot z_2$$

ALGORITMO 2:

$$z_1 = x + y$$

$$z_2 = x - y$$

$$z_3 = z_1 \cdot z_2$$

RREC. MACCHINA

↑

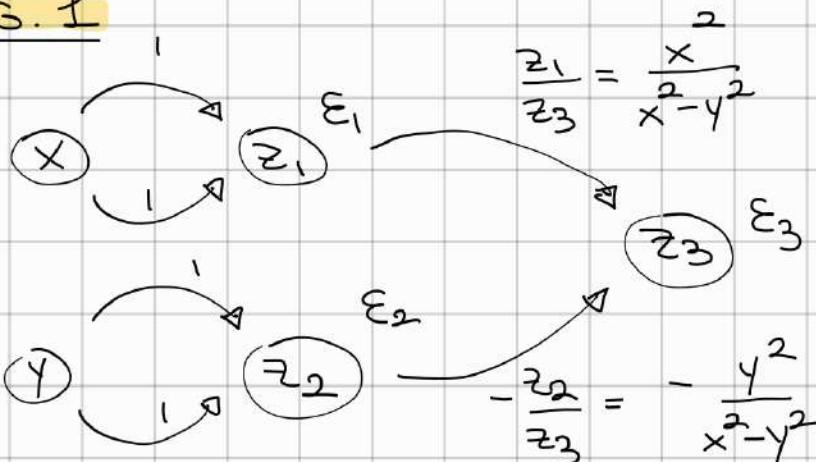
$$\varepsilon_d = \frac{x \cdot 2x}{x^2 - y^2} \varepsilon_x + \frac{y(-2y)}{x^2 - y^2} \varepsilon_y \quad |\varepsilon_x|, |\varepsilon_y| \leq u$$

↓
CUE PRESENDE DAGLI ALG.

$$|\varepsilon_d| \leq \frac{x^2 + y^2}{|x^2 - y^2|} \cdot 2u$$

TQD problema è molto condizionato quando $|x| \approx |y|$
(in questo caso l'errore potrebbe non essere piccolo)

ALG. 1



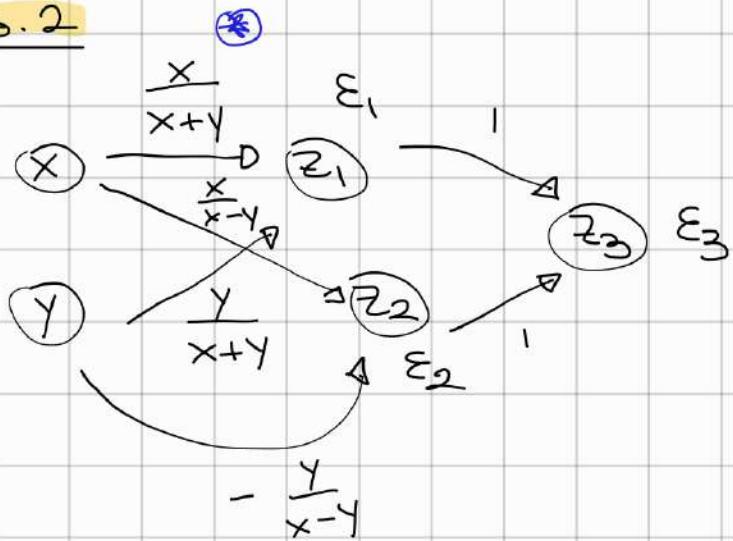
$$\varepsilon_d = \varepsilon_3 + \frac{x^2}{x^2 - y^2} \varepsilon_1 - \frac{y^2}{x^2 - y^2} \varepsilon_2$$

$$|\varepsilon_d| \leq u \left(1 - \frac{x^2 + y^2}{|x^2 - y^2|} \right)$$

\Rightarrow questo alg. è instabile quando $|x| \approx |y|$ avendo
quando x^2 e y^2 tendono a cancellarsi

dei dati
in input

ALG. 2



⊕ i coeff. di ampl.
non ci interessano
perché assumiamo
i dati di input
corretti

$$\varepsilon_d = \varepsilon_3 + \varepsilon_1 + \varepsilon_2$$

$$|\varepsilon_d| \leq 3u \Rightarrow \text{l' algoritmo è stabile}$$

ESEMPIO

Calcolare una sommatoria

$$f(x_1, \dots, x_n) = \sum_{i=1}^n x_i$$

ALG. 2

$$z_1 = x_1 + x_2$$

$$z_2 = z_1 + x_3$$

:

$$z_{n-1} = z_{n-2} + x_n$$

$$\varepsilon_d = \frac{\sum_{i=1}^n x_i \cdot \varepsilon_{x_i}}{\sum_{i=1}^n x_i}$$

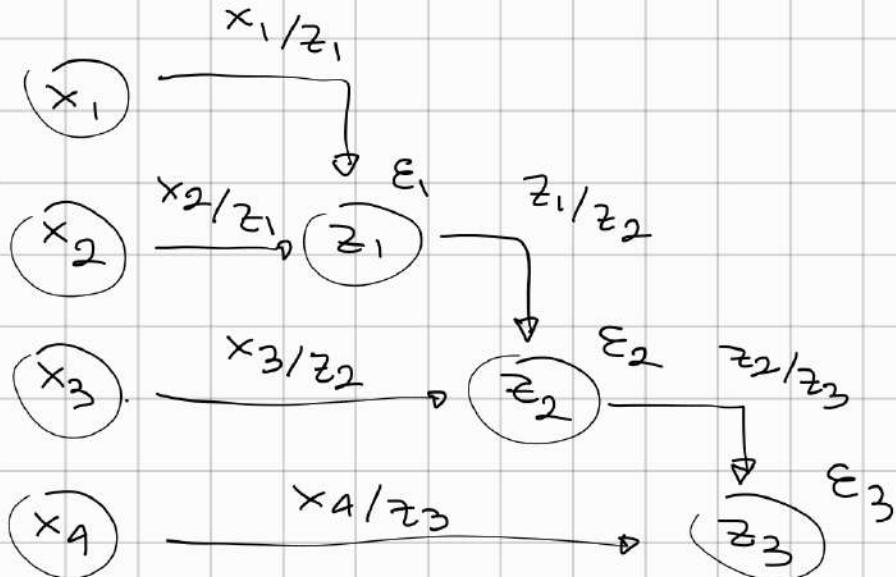
$$|\varepsilon_d| = \frac{\sum_{i=1}^n |x_i| \cdot u}{(\sum_{i=1}^n x_i)}$$

Il prod. è mol. com.
quando il den.
va a 0

Il prob. è ben cond. se x_i sono di segno concorde
(non c'è cancellazione)

$$|\varepsilon_2| \leq u$$

$$n = 4$$



$$\varepsilon_2 = \varepsilon_3 + \frac{x_1 + x_2 + x_3}{x_1 + \dots + x_4} (\varepsilon_2 + \frac{x_1 + x_2}{x_1 + x_2 + x_3} \varepsilon_1) =$$

$$= \varepsilon_3 + \frac{x_1 + x_2 + x_3}{x_1 + \dots + x_4} \varepsilon_2 + \frac{x_1 + x_2}{x_1 + \dots + x_4} \varepsilon_1$$

$$\text{Nel caso } x_j \text{ concordi} \Rightarrow |\varepsilon_2| \leq 3u$$

$$\text{Per } n \text{ qualsiasi } \varepsilon_2 = \frac{1}{\sum_{j=1}^n x_j} \cdot \sum_{i=1}^{n-1} \left(\sum_{j=1}^{i+1} x_j \right) \varepsilon_{x_i}$$

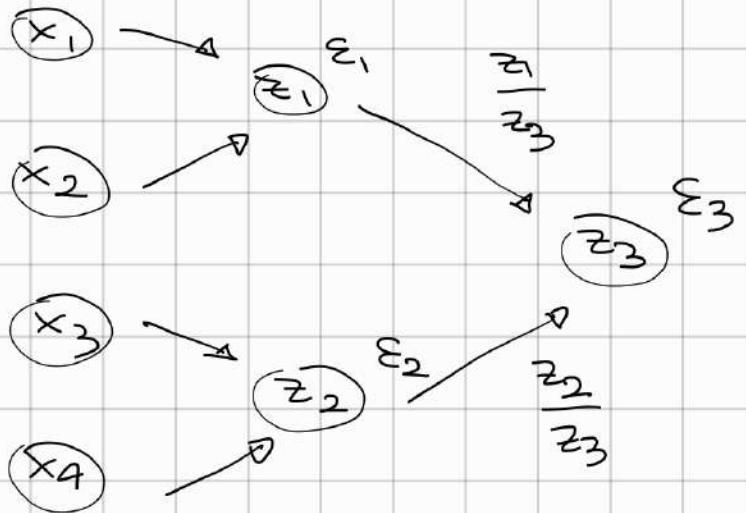
(x_j concordi)

Nel caso x_j concordi ($\text{Sign}(x_j) = \text{Sign}(x_{j+1}) \forall j$)
si può dim. che $|\varepsilon_2| \leq (n-1)u$

Si può fare meglio con alg. 2

FIG. 2 (somma in parallelo)

$n=4$



$$\epsilon_2 = \epsilon_3 + \frac{x_1 + x_2}{x_1 + \dots + x_4} \epsilon_1 + \frac{x_3 + x_4}{x_1 + \dots + x_4} \epsilon_3$$

Nel caso concorde: $|\epsilon_2| \leq 20$

Nel caso (concorde) n generico: $|\epsilon_2| \leq \log_2 n \cdot u$

ESEMPIO (valutazione di un polinomio)

$$f(x) = \sum_{j=0}^n a_j x^j \quad \epsilon_j = \frac{x \cdot f'(x)}{f(x)}$$

non abbiamo modo di limitare
 ϵ_j quando x è molto vicino
ad una radice non nulla
 $(x \neq 0, f(x) \approx 0)$

ALG. 1 (POTENZE DI x)

$$z_0 = 1 \quad p_0 = \theta_0$$

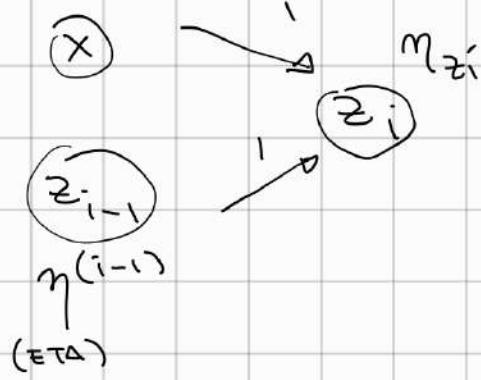
for $i = 1, \dots, n$

$$z_i = z_{i-1} \cdot x$$

$$p_i = \partial_i z_i + p_{i-1}$$

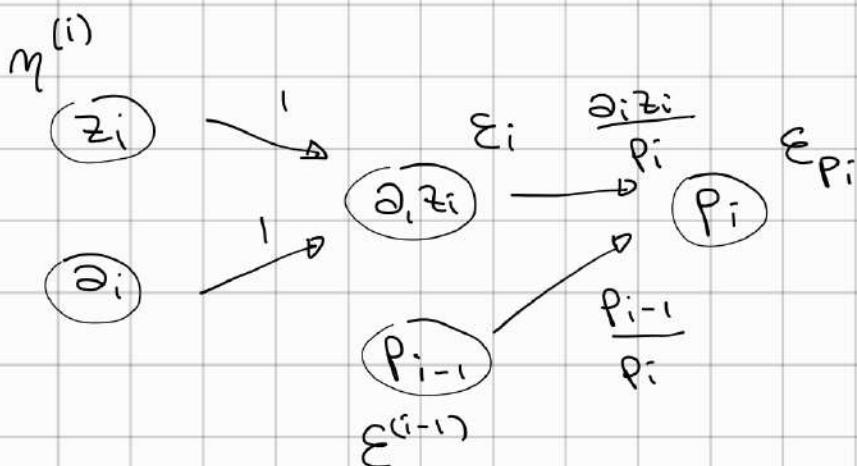
end

$|m_{z_i}| \leq v$ ERRO. DI TRONCATO
DATO DALL'ESTIMA
HOCT. PER x



$\eta^{(i-1)}$ = ERRO. ACCUMULATO SU z_{i-1}

$$\begin{cases} \eta^{(i)} = m_{z_i} + \eta^{(i-1)} \\ \eta^{(0)} = 0 \end{cases} \Rightarrow |\eta^{(i)}| \leq i \cdot v$$



$|\epsilon_i|, |\epsilon_{p_i}| \leq v$

$$\begin{aligned} \epsilon^{(i)} &= \epsilon_{p_i} + \frac{p_{i-1}}{p_i} \epsilon^{(i-1)} + \\ &+ \frac{\partial_i z_i}{p_i} (\epsilon_i + \eta^{(i)}) \end{aligned}$$

$$\begin{aligned} |\epsilon^{(i)}| &\leq v + \left| \frac{p_{i-1}}{p_i} \right| |\epsilon^{(i-1)}| + \\ &+ \left| \frac{\partial_i z_i}{p_i} \right| (i+1)v \end{aligned}$$

Nel caso ∂_i concordi ed $x > 0$

$$\frac{p_i + \partial_i z_i}{p_i} = \frac{|p_{i-1}| + |\partial_i z_i|}{|p_i|} = 1 \quad \Rightarrow |\epsilon^{(i)}| \leq (i+2)v$$

$|\varepsilon_{\text{al}}| \leq (n+2)u \Rightarrow$ l'algor. è abbastanza stabile,
meno se il polin. è di grado
molto elevato

ALG. 2 (METODO DI HORNER)

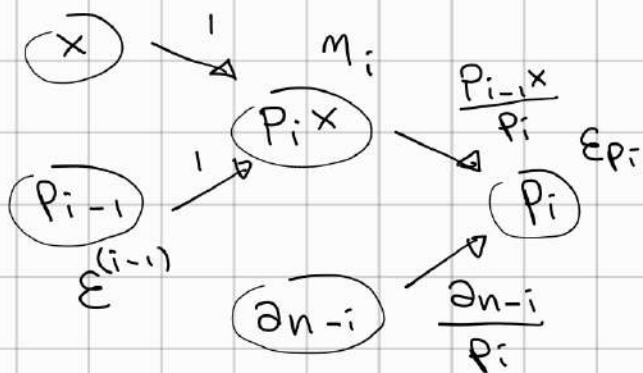
$$f(x) = \sum_{j=0}^n a_j x^j = (\dots((a_n x + a_{n-1}) x + a_{n-2}) x + \dots a_1) x + a_0$$

IN PSEUDOCODICE:

```

 $p_0 = a_n$ 
for  $i = 1, \dots, n$ 
     $p_i = p_{i-1}x + a_{n-i}$ 
end
 $f(x) = p_n$ 

```



$$\varepsilon^{(i)} = \varepsilon_{p_i} + \frac{p_{i-1}x}{p_i} (\eta_i + \varepsilon^{(i-1)}) \quad |\eta_i|, |\varepsilon_{p_i}| < u$$

$$|\varepsilon^{(i)}| \leq u + \alpha(u + |\varepsilon^{(i-1)}|)$$

$$\alpha := \max_{j=1, \dots, i} \left| \frac{p_{i-j}x}{p_i} \right|$$

- Se a_j concordi e $x > 0 \Rightarrow 0 < \alpha < 1$

$$|\varepsilon^{(i)}| \leq 2u(1 + \alpha + \alpha^2 + \dots + \alpha^{i-1}) \quad \begin{matrix} \text{si dim. per induzione} \\ \text{costante} \\ \uparrow \end{matrix}$$

$$= 2u \left(\frac{1 - \alpha^i}{1 - \alpha} \right)$$

In questo caso si ha che $|\varepsilon_{\text{al}}| \leq C \cdot u$ dove C è
indipendente dal grado del polinomio n

l'algoritmo funziona
bene anche per $<=$
 n grande

dal grado
del polinomio
 n

Il metodo di Horner è migliore quando il grado è molto alto.

FUNZIONI NON RAZIONALI

$f(x)$ non si può scrivere come una seq. finita di op. aritm. di base $\{ +, -, \cdot, / \}$

Es. $\exp(x)$

In questo caso sul computer viene valutata una f. razionale $h(x) \approx f(x)$ che approssima $f(x)$ in un intorno di x

Es. $h(x)$ = sviluppo di Taylor troncato

(relativo)

Si introduce però un errore ANALITICO dovuto all'approssim.

$$\varepsilon_{\text{an}} := \frac{h(P_0) - f(P_0)}{h(P_0)}$$

$$\text{E' erro. relativo totale } \frac{h_a(P_1) - f(P_0)}{f(P_0)} = \varepsilon_f$$

divenuta (a meno di termini trascurabili):

$$\begin{aligned} \varepsilon_f &= \varepsilon_{\text{an}} + \varepsilon_f + \varepsilon_a = \frac{h(P_0) - f(P_0)}{h(P_0)} + \frac{h(P_1) - h(P_0)}{h(P_0)} + \\ &\quad + \frac{h_a(P_1) - h(P_1)}{h(P_1)} \end{aligned}$$

ESEMPIO

$$f(x) = \log(1+x) \approx h(x) = \sum_{i=1}^n (-1)^{i-1} \cdot \frac{x^i}{i} \quad \text{con } |x| < 1$$

$$r(x) := f(x) - h(x) = \frac{(-1)^n x^{n+1}}{(n+1)(1+\eta)^{n+1}} \quad \text{con } |\eta| < |x|$$

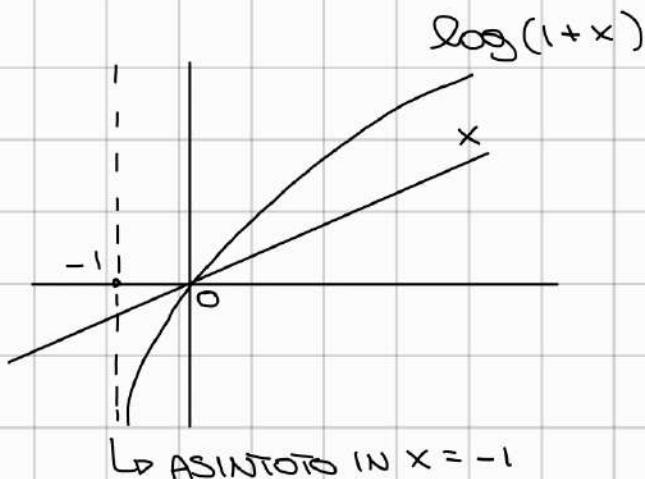
↓
PRESTO
DELL'APPROSSIMAZIONE

↓ ($\epsilon \tau \Delta$)

Dato $x \in [-1, 0)$ cosa posso dire su $|\varepsilon_{\text{an}}|$?

Vogliamo stimare $\frac{|r(x)|}{|\log(1+x)|}$

OSS



$$|\log(x)| > |x|$$

$$|\varepsilon_{\text{an}}| \leq \frac{|r(x)|}{|\log(1+x)|} \leq \frac{|r(x)|}{|x|} \leq \frac{|r(x)|}{(n+1)|1+x|^{n+1}}$$

ESEMPIO:

esponenziale $e^x \approx \sum_{k=0}^n \frac{1}{k!} x^n = h(x)$

quando $x < 0 \rightsquigarrow$ potrebbe essere instabile valutare $h(x)$ con uno degli alg. visti

SOLUZIONE: $e^{-x} = \frac{1}{e^x} \rightsquigarrow$ se $x < 0 \quad h(-x)^{-1}$

$f: \mathbb{C}^n \rightarrow \mathbb{C}^m$ con $m \neq n$

$$f(x+y) = f(x) + f(y)$$

$$f(\lambda x) = \lambda f(x) \quad \forall \lambda \in \mathbb{C}$$

$\exists! A \in \mathbb{C}^{m \times n}: f(x) = A \cdot x$

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

a_{ij} = singolo elemento

a_i = i-esima colonna di A

a^i = i-esima riga di A

$$A = [a_1 \dots | a_n]$$

$$A = \begin{bmatrix} a^1 \\ \vdots \\ a^m \end{bmatrix}$$

$A \in \mathbb{R}^{m \times n}$ se $a_{ij} \in \mathbb{R} \quad \forall i, j$ MATRICE REALE

VECTORE COLONNA $n=1 \quad \forall \in \mathbb{R}^{m \times 1} = \mathbb{R}^m$
 " RIGA $m=1 \quad \forall \in \mathbb{R}^{1 \times m}$

DEF. (MATRICE TRASPOSTA)

$A \in \mathbb{C}^{m \times n}$, si definisce $B = A^T$ come la matrice

$$b_{ij} = a_{ji}$$

le colonne di A diventano le colonne di B

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} ; \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

In particolare $B = A^T \in \mathbb{C}^{n \times m}$

DEF. (MAT. HERMITIANA)

$B = A^H$ (spesso A^*) MAT. HERM.

ottenuta come $b_{ij} = \overline{a_{ji}}$ (elemento trasposto e
coniugato)

$$\forall z \in \mathbb{C} \quad z = c + id \quad \text{no} \quad \bar{z} = c - id$$

Se A reale $A^T = A^*$ (il coniugato c'è solo nei complessi)

DEF. MAT. DIAGONALE

$A \in \mathbb{C}^{m \times m}$ (quadrata) si dice diagonale se
 $a_{ij} = 0 \quad \forall i \neq j$

es. mat. identità o $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

DEF. (MAT. TRIANGOLARE)

$A \in \mathbb{C}^{m \times n}$

A si dice triangolare inf. se ha elementi nulli sopra la diag. principale

$$j > i \Rightarrow a_{ij} = 0$$

es. $A = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 2 & 0 \\ 6 & 5 & 3 \end{bmatrix}$

Analog. A è triang. sup. se $j < i \Rightarrow a_{ij} = 0$

es. $\begin{bmatrix} 1 & 4 & 6 \\ 0 & 2 & 5 \\ 0 & 0 & 3 \end{bmatrix}$

OPERAZIONI ARITMETICHE TRA MATRICI E VETTORI

$v_1, v_2 \in \mathbb{C}^n$

PROD. SCALARE $\langle v_1, v_2 \rangle = \sum_{i=1}^n (\overline{v_2})_i \cdot (v_1)_i = v_2^* \cdot v_1$

se sono reali:

$$\langle v_1, v_2 \rangle = \sum_{i=1}^n (v_2)_i \cdot (v_1)_i = v_2^T \cdot v_1$$

COSTO IN TERMINI DI OPERAZIONI: $O(n)$ perché dobbiamo
 \downarrow
 o grande forse n multipl.
 e $n-1$ somme
 (Anche $v_1 + v_2$ costa $O(n)$)

MOLTIPL. TRA MATRICI

$A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times p}$, $C = A \cdot B \in \mathbb{C}^{m \times p}$

$$(A \cdot B)_{ij} = C_{ij} = \sum_{h=1}^n a_{ih} b_{hj} = (a^i)^T \cdot b_j$$

In particolare, calcolare 1 elemento di un prod. di matrici (1 elem. di C) costa come 1 prod. scalare tra vettori di \mathbb{C}^n

Dico calcolare $m \cdot p$ elementi (num. di entrate di C)



Il costo di $A \cdot B$ è $O(m \cdot n \cdot p)$

Nel caso di un prod. matrice - vettore

$A \in \mathbb{C}^{m \times n}$ e $v \in \mathbb{C}^n$ si ha $O(m \cdot n)$

ES. Nel caso di mat. quadrate dove $m = n = p$

$$\frac{\text{MAT} \times \text{VETT}}{O(n^3)}$$

$$\frac{\text{MAT} \times \text{VETT}}{O(n^2)}$$

PROPRIETÀ DEL PRODOTTO MATRICIALE

In generale $A \cdot B \neq B \cdot A$

$$A \cdot B = 0 \Rightarrow A = 0 \circ B = 0$$

$$\underline{\text{ES.}} \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Se $\langle v_1, v_2 \rangle = 0$ i due vett. sono ortogonali

$A \cdot B = 0 \Rightarrow$ le righe di A sono ortog. alle colonne di B

Talvolta è utile vedere $A \cdot B$ come un'operazione di una matrice sulle colonne dell'altra o dell'altra mat. sulle righe della prima.

$$\underline{\text{es.}} \quad A \in \mathbb{C}^{m \times n}$$

$$B \in \mathbb{C}^{n \times p}$$

$$A \cdot B = [Ab_1 \mid \dots \mid Ab_n] = \begin{bmatrix} \overbrace{\alpha^1 B}^{\text{COMB. LINEARE}} \\ \vdots \\ \overbrace{\alpha^n B}^{\text{COMB. LINEARE}} \end{bmatrix} \quad \begin{array}{l} \text{DEUE RIGHE DI } B \\ \text{DEUE COLONNE} \\ \text{DI } A \end{array}$$

$$Ab_j = \sum_{i=1}^n \alpha_i b_{ij}$$

Può essere utile quando risolvono una delle due mat. ha molti zero

$$\underline{\text{es.}} \quad A \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} = [a_1 + a_2 \mid a_1] \quad (\text{Analogo per righe})$$

oss Vale la prop. associativa, ovvero:

$$A \in \mathbb{C}^{m \times n}, B \in \mathbb{C}^{n \times k}, C \in \mathbb{C}^{k \times p}$$

$$\underbrace{A \cdot (B \cdot C)}_{\downarrow} = \underbrace{(A \cdot B)C}_{\downarrow}$$

$$\mathcal{O}(nkp + mnp)$$

k interviene in
entrambi i termini

costi comput. pot. differenti
se k è piccolo

IDEA Cercare di evitare di svolgere prima prenotati della forma $\boxed{\quad} \times \boxed{\quad}$

OSS

$$\begin{aligned} A(x+y) &= Ax + Ay && \text{PROPR. DISTRIBUTIVA} \\ A(B+C) &= AB + AC \\ &\downarrow && \uparrow \\ \text{COSTA MENO} & & \text{MENO DI} & \end{aligned}$$

ESEMPIO Scrivere i costi di $A(B+C)$ e di $AB+AC$ con $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times k}$, $C \in \mathbb{C}^{k \times p}$

DEF $v_1, \dots, v_s \in \mathbb{C}^n$ sono **linearmente indipendenti** se:

$$\sum_{j=1}^s \alpha_j v_j = 0 \Rightarrow \alpha_j = 0 \quad \forall j = 1, \dots, s$$

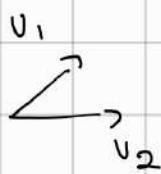
$$\text{con } \alpha_1, \dots, \alpha_s \in \mathbb{C}$$

Sono invece **linearmente dipendenti** se:

$$\exists k \in \{1, \dots, s\} : v_k = \sum_{\substack{j=1 \\ j \neq k}}^s \beta_j v_j$$

$$\text{span}(v_1, \dots, v_s) = \left\{ \sum \alpha_j v_j : \alpha_j \in \mathbb{C} \right\}$$

es. \mathbb{R}^2

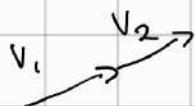


v_1, v_2 LIN-INDIP.

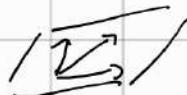
\mathbb{R}^3



vet-LIN. INDIP.



v_1, v_2 LIN. DIP.



vet. LIN. DIP.

DEF (RANGO MATRICE)

$A \in \mathbb{C}^{m \times n}$ il rango di A è il num. di colonne lin. indip. in A , che è uguale al num. max di righe indip. in A

$$\text{Im}(A) = V = \text{span}(\alpha_1, \dots, \alpha_n) = \{Ax : x \in \mathbb{C}^n\} \subseteq \mathbb{C}^m$$

$$\dim(V) = \text{rk}(A)$$

DETERMINANTE

$$\det : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}$$

SILUERO DI LAPLACE:

$$\det(A) = \begin{cases} n=1 & \alpha_{11} \\ n=2 & \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} \\ \forall i: \sum_{j=1}^n (-1)^{i+j} \alpha_{ij} \det(A_{i;j}) \end{cases}$$

MAT. OTTENUTA CANCELLANDO LA i -ESIMA RIGA E LA j -ESIMA COLONNA

$$V_j : \sum_{i=1}^n (-1)^{i+j} \alpha_{ij} \det(A_{i;j})$$

Una mat. tale che $\det(A) \neq 0$ è detta MAT. NON SINGOLARE
 $\det(A) \neq 0 \iff \text{rk}(A) = n$

DEF $A \in \mathbb{C}^{m \times n}$ una sottomatrice di dim. $k \times p$ ($k \leq m, p \leq n$) è la mat. che si ottiene prendendo gli elem. che stanno sull'inters. di k righe e p colonne di A

DEF Un minore di ordine k è il det di una sottomat. $k \times k$ di A

PROPRIETÀ $\text{rk}(A)$ corrisp. alla grandezza massima k di un minore non singolare di A .

ES. Se A ha rango 1 \Rightarrow ogni sottomat. 2×2 che si estrae da A sarà singolare ($\det(A) = 0$)

DEF Una sottomat. è detta principale se è quadrata e corrisponde agli stessi indici di riga e colonna estratti da A .

Il det di una mat. principale è un minore principale

Sottomat. di testa di ordine k

se seleziona le prime k righe e le prime k colonne

TH DI BINET - CAUCHY:

$$A \in \mathbb{C}^{m \times n}, B \in \mathbb{C}^{n \times m}, C = A \cdot B \in \mathbb{C}^{m \times m}$$

$$\det(C) = \begin{cases} 0 & \text{se } m > n \\ \sum_{[j]} \det(A_{[j]}) \det(B_{[j]}) & \text{se } m \leq n \end{cases}$$

$$A \begin{array}{|c|c|c|} \hline & \cdots & n \\ \hline m & [A_{[j]}] & \\ \hline \end{array}$$

$$B \begin{array}{|c|} \hline m \\ \hline n \\ \hline \end{array}$$

$A_{[j]}$ = sottomat. $m \times m$ dove gli indici di riga sono $\{1, \dots, m\}$ e quelli di colonna sono un sottoinsieme $[j]$ di m elementi di $\{1, \dots, n\}$

$B[\omega]$ analogo, ma dove variano gli indici di riga

COROLARIO: CASO $m=n$

$$\det(A \cdot B) = \det(A) \cdot \det(B)$$

unica scelta per $[;]$

CASI SPECIALI CALCOLO DETERMINANTE

Quando le mat. sono diagonali o triangolari

$$\det(A) = \prod_{j=1}^n a_{jj} \quad \text{con } A \in \mathbb{R}^{n \times n}$$

DEF. $A \in \mathbb{C}^{n \times n}$ si definisce l' inversa di A , $\bar{A}^{-1} \in \mathbb{C}^{n \times n}$ t.c.

$$A \cdot \bar{A}^{-1} = I = \bar{A}^{-1} \cdot A$$

$$\bar{A}^{-1} = \frac{1}{\det A} \cdot \underbrace{A_{ij}(A)}_{\substack{\text{L'MAT.} \\ \text{AGGIUNTA}}}$$

$$A_{ij}(A)_{ij} = (-1)^{i+j} \cdot \det(A_{ji})$$

\downarrow

$A_{ji} = \text{MAT. OBTENUTA}$
 $\text{DA } A$
 CANCELLANDO
 $\text{LA RIGA } j$
 $\in \text{LA COLONNA } i$

$$\text{BIRKHOFF-CAREY} \Rightarrow \det(A \cdot \bar{A}^{-1}) = \det(A) \cdot \det(\bar{A}^{-1})$$

$$\det(I) = 1$$

$$\det(\bar{A}^{-1}) = \frac{1}{\det A}$$

OSS

$$(A \cdot B)^{-1} = B^{-1} A^{-1}$$

$$(\bar{A}^{-1})^T = (\bar{A}^T)^{-1} = \bar{A}^{-T}$$

$$(A \cdot B)^T = B^T A^T$$

$$(\bar{A}^{-1})^* = (\bar{A}^*)^{-1} = \bar{A}^{-*}$$

$$(A \cdot B)^* = B^* A^*$$

DEF

$A = A^*$ MATERIA HERMITIANE \rightsquigarrow ELEM. REALI SULLA DIAGONALE

$$A = A^T$$

"

SIMMETRICHE

$$A^* A = A A^* = I \quad \text{MATERIA UNITARIA}$$

$$A^T A = A A^T = I \quad "$$

ORTOGONALI

OSS

$$\det(A^*) = \overline{\det(A)}$$

$$\det(A) = \det(A^T)$$

$$A^* A = I \Rightarrow \det(A) \cdot \overline{\det(A)} = 1 \Rightarrow \det(A) \text{ è un num. complesso di modulo } 1$$

$$A^T A = I \Rightarrow \det(A) = \pm 1$$

DEF

$$A^T = -A \quad \text{ANTISIMMETRICA}$$

$$A^* = -A \quad \text{ANTISIMMETRICA}$$

$$\begin{aligned} AA^T &= A^TA \\ AA^* &= A^*A \end{aligned} \quad \text{MATRICE NORMALE}$$

DEF

quando $A^* = A$ è importante la funzione quadratica
 $x^*Ax = f(x) \quad f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$

DEF

$$A = A^* \quad x^*Ax \geq 0 \quad \forall x \in \mathbb{C}^n \Rightarrow A \text{ si dice semidef.}$$

positiva

(autovalori non negativi)

$$x^*Ax \leq 0 \quad \forall x \in \mathbb{C}^n$$

\Downarrow A semidef. negativa

ESEMPIO DI MAT. ORTOGONALI:

$$A \in \mathbb{R}^{n \times n} \quad A^T A = I$$

DEF

$P \in \mathbb{R}^{n \times n}$ è detta MAT. DI PERMUTAZIONE se è
 ottenuta da I permutando righe o colonne

$$\text{es. } \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\begin{array}{l} 1 \rightarrow 2 \\ \sigma_C: 2 \rightarrow 3 \\ 3 \rightarrow 1 \end{array}$$

$$\begin{array}{l} 1 \rightarrow 3 \\ \text{or: } 2 \rightarrow 1 \\ 3 \rightarrow 2 \end{array}$$

Data $A \in \mathbb{C}^{n \times n}$

$AP \rightarrow$ MAT. OTTENUTA DA A PERMUTANDO LE COLONNE
 SECONDO σ

$$PA \rightarrow$$

“ “ “ ”

PERM. LE RIGHE

SECONDO σ

ES. $\begin{bmatrix} 1 & 4 & * \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \Rightarrow AP = \begin{bmatrix} * & 1 & 4 \\ 8 & 2 & 5 \\ 9 & 3 & 6 \end{bmatrix}$

$$PA = \begin{bmatrix} 2 & 5 & 8 \\ 3 & 6 & 9 \\ 1 & 4 & * \end{bmatrix}$$

$P^T P = PP^T = I \rightsquigarrow$ dipende dal fatto che P^T è associata alle permutazioni $\sigma_c^{-1}, \sigma_r^{-1}$

DEF $A \in \mathbb{C}^{n \times n}$ si dice CONVERGENTE se

$$\lim_{K \rightarrow +\infty} A^K = 0 = \text{MAT. NURO} \text{ } n \times n$$

ES. $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 3 & 0 \end{bmatrix}, A^3 = 0, A^4 = 0 \dots$

CASO PARTICOLARE PER CUI FA O A PARTE DA UNA POS. FISSATA

DEF $A \in \mathbb{C}^{n \times n}$ a predom.
si dice DIAG. FORTE se $|a_{ii}| > \sum_{j=1}^n |a_{ij}|$

ES. $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 4 \end{bmatrix}$

DEF A DIAG. DEBOLE se $|a_{ii}| \leq \sum_{j=1}^n |a_{ij}| \quad \forall i = 1, \dots, n$

\exists almeno un indice $k \in \{1, \dots, n\}$

$$|a_{kk}| > \sum_{j=1}^n |a_{kj}|$$

SISTEMI LINEARI

$$Ax = b \quad \text{con } A \in \mathbb{C}^{m \times n}, b \in \mathbb{C}^m$$

\Updownarrow

$$\begin{cases} a_{1,1}x_1 + \dots + a_{1,n}x_n = b_1 \\ \vdots \\ a_{m,1}x_1 + \dots + a_{m,n}x_n = b_m \end{cases}$$

m = num. equazioni

n = num. incognite

IN GENERALE : $\begin{array}{l} \cdot 1 \text{ soluzione} \\ \cdot \infty \text{ soluzioni} \quad (\text{sist. sottodeterminato}) \\ \cdot \emptyset \text{ soluzioni} \quad (\text{sistema inconsistente}) \end{array}$

ES. \mathbb{R}^2

$$\begin{cases} x + y = 2 \\ 3x + 3y = 6 \end{cases} \rightsquigarrow \infty \text{ soluzioni}$$



$$\begin{cases} x + y = 2 \\ x + 3y = 2 \end{cases} \rightsquigarrow 1 \text{ soluzione}$$



$$\begin{cases} x + y = -3 \\ x + y = 6 \end{cases} \rightsquigarrow \text{no soluzione}$$



TH DI GAUSS - CANCELLAZIONE

$Ax = b$ ammette soluzioni se $\text{rk}(A) = \text{rk}(A|b)$

$$A \in \mathbb{C}^{m \times n} \quad \underbrace{A|b \in \mathbb{C}^{m \times (n+1)}}$$

$\left[\begin{smallmatrix} & \dots \\ A & b \end{smallmatrix} \right]$

L'insieme delle soluzioni di $Ax = b$ è uno spazio affine $\rightsquigarrow x_0 + \text{span}(v_1, \dots, v_s)$

di dimensione $S = n - \text{rk}(A)$

CASO $m=n$: la soluzione è unica $\Leftrightarrow \text{rk}(A) = n$

Se $\text{rk}(A) = \text{rk}(A|b) < n$ allora ∞ soluzioni

CASO $m < n$: se $\text{rk}(A) = \text{rk}(A|b) \Rightarrow \infty$ soluzioni

CASO $m > n$: se $\text{rk}(A) = \text{rk}(A|b) \rightsquigarrow 1$ soluzione

se $\text{rk}(A) = n$

$\Rightarrow \infty$ soluzioni

se $\text{rk}(A) < n$

E5.

$$\begin{cases} x - 2y - 2z = 0 \\ 2x - y + 4z = 3 \\ 6x - 12y - 12z = 0 \end{cases} \Rightarrow \begin{cases} x = 2y + 2z \\ -4y - 4z - y + 4z = 3 \\ 0 = 0 \end{cases}$$

$$\Rightarrow \begin{cases} x = 2y + 2z \\ -5y = 3 \end{cases} \Rightarrow y = -\frac{3}{5} \Rightarrow x = -\frac{6}{5} + 2z$$

$$\left\{ \begin{array}{l} \text{V} \in \mathbb{R}^3 : \begin{bmatrix} -\frac{6}{5} \\ -\frac{3}{5} \\ 0 \end{bmatrix} + z \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = v, z \in \mathbb{R} \end{array} \right\}$$

DEF $A \in \mathbb{C}^{m \times m}$, $\lambda \in \mathbb{C}$, $v \in \mathbb{C}^n$ $v \neq 0$ \rightarrow AUTOM. L'EQ. È SEMPRE VERA.

(λ, v) è un AUTOCOPPIA se verificano $Av = \lambda v$

λ = AUTOVALORE DI A

v = AUTOVETTORE ASSOCIASTO A λ

$w =$ " \sim sx " \sim $A \lambda$ se $w^T A = \lambda w^T$ $w \in \mathbb{C}^n$

OSS L'insieme degli autovettori associati a un auto λ è infinito

$$Av = \lambda v \Rightarrow A(\theta v) = \theta Av = \theta \lambda v = \lambda(\theta v) \quad \theta \in \mathbb{C}$$

OSS $Av = \lambda v \rightsquigarrow (A - \lambda I)v = 0$ SISTEMA LINEARE omogeneo

Per Rouché - Capelli ci sono soluzioni $\neq 0$
 $\Leftrightarrow \det(A - \lambda I) = 0$ (EQ. CARATTERISTICA)

Gli autovalori sono quei valori scalari (in \mathbb{C}) che verificano $\{\lambda \in \mathbb{C} : \det(A - \lambda I) = 0\}$

$\{v : Av = \lambda v\} =$ SPAZIO VETTORIALE di dim. $n - \text{rk}(A - \lambda I)$

ESERCIZIO

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \quad A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = (1-\lambda)^2 - 4$$

IN GENERALE: $A \in \mathbb{C}^{n \times n}$ $\det(A - \lambda I) =$ POLINOMIO DI GRADO n IN λ

OSS $\det(A - \lambda I) = \det((A - \lambda I)^T) = \det(A - \lambda I)$
 $\Rightarrow A$ e A^T hanno gli stessi autovalori

OSS $P_A(\lambda) = \det(A - \lambda I) = (-1)^n \lambda^n + (-1)^{n-1} \sigma_1 \lambda^{n-1} + \dots + \sigma_{n-1} \lambda + \sigma_n$
 È UN POLN. MONICO (PERCHÉ HA GRADO MAX PARI A $n - 1$)

σ_j = somma dei minori principali di ordine j estratti da A

$$\sigma_1 = \sum_{i=1}^n a_{ii} = \text{TRACCIA DI } A = \text{tr}(A)$$

$$\sigma_n = \det(A)$$

$$P_A(\lambda) = (-1)^n (\lambda - \lambda_1) \dots (\lambda - \lambda_n) \Rightarrow \begin{cases} \det(A) = \prod_{i=1}^n \lambda_i \\ \text{tr}(A) = \sum_{i=1}^n \lambda_i \end{cases}$$

ESEMPIO

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\det(A) = 3 > 0$$

$$\lambda_1 \cdot \lambda_2 > 0$$

$$\text{tr}(A) = 4 > 0$$

$$\lambda_1 + \lambda_2 > 0$$

↓

$$\lambda_1, \lambda_2 > 0$$

DEF Il raggio spettrale di una matrice A è:

$$p(A) := \max_{j=1, \dots, n} |\lambda_j| \in \mathbb{R}_0^+$$

TH A è convergente ($\lim_{K \rightarrow \infty} A^K = 0$) $\Leftrightarrow p(A) < 1$

DEF Date $A, B \in \mathbb{C}^{n \times n}$ si dicono simili se $\exists S \in \mathbb{C}^{n \times n}$ non singolare t.c. $B = S^{-1}AS$ e si indica $A \sim B$

PROPRIETÀ $A \sim B \Rightarrow (\lambda, S^{-1}v)$ è autocoppia per B
 (λ, v) autocoppia per A

DIM

$$B(S^{-1}v) = S^{-1}AS(S^{-1}v) = S^{-1}Av = S^{-1}\lambda v = \lambda \cdot (S^{-1}v)$$

$$v \neq 0 \Rightarrow S^{-1}v \neq 0$$

PROPRIETÀ $A \in \mathbb{C}^{n \times n}$, (λ, v) autocoppia per A , $k \in \mathbb{N}$
 $\Rightarrow (\lambda^k, v)$ autocoppia per A^k

DIM $A^k \cdot v = A^{k-1} \cdot (Av) = A^{k-1} \lambda v = \lambda A^{k-1} v = \dots = \lambda^k v$

PROPRIETÀ $A \in \mathbb{C}^{n \times n}$, A non singolare, (λ, v) autocoppia per $A \Rightarrow (\lambda^{-1}, v)$ autocoppia per A^{-1}

DIM $Av = \lambda v \Rightarrow v = A^{-1} \cdot \lambda v = \lambda A^{-1} v$
 $\Rightarrow \lambda^{-1} v = A^{-1} v$

PROPRIETÀ A^{-k} , $k \in \mathbb{N}$ ha autocopie (λ^{-k}, v)
 \parallel
 $(A^{-1})^k = (A^k)^{-1}$

PROPRIETÀ $A = A^*$ \Rightarrow tutti gli autovalori di A sono in \mathbb{R}

DIM $r(x) = \frac{x^* Ax}{x^* x}$ QUOTIENTE DI RAY LEIGH

$$r: \mathbb{C}^n \rightarrow \mathbb{C}$$

$$r(v_j) = \frac{v_j^* A v_j}{v_j^* v_j} = \lambda_j \cdot \frac{v_j^* v_j}{v_j^* v_j} = \lambda_j$$

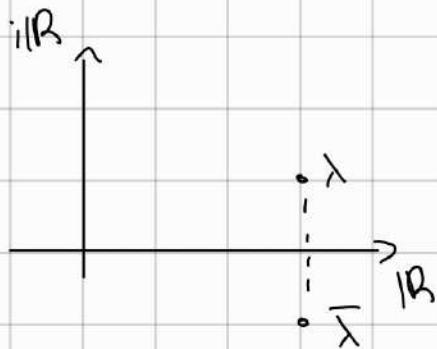
$$Av_j = \lambda_j$$

$$v_j^* Av_j = (v_j^* Av_j)^* = v_j^* A^* v_j = v_j^* A v_j$$

$$\Rightarrow v_j^* Av_j \in \mathbb{R}$$

$$\lambda_0 = r(v_j) = \frac{\text{qualcosa che } \in \mathbb{R}}{\sum_{i=1}^n |v_{j,i}|^2} \in \mathbb{R}$$

OSS $A \in \mathbb{R}^{n \times n}$, λ è autovalore di A e $\lambda \in \mathbb{C} \setminus \mathbb{R}$
 \Rightarrow anche $\bar{\lambda}$ è autovalore di A



DIM $P_A(\lambda) = \det(A - \lambda I) = \text{coeff. della}$ Polinomio A

$$0 = P_A(\lambda) = \sum_{j=0}^n \sigma_j \lambda^j \Rightarrow \bar{0} = \overline{P_A(\lambda)} = \overline{\sum_{j=0}^n \sigma_j \lambda^j} = \sum_{j=0}^n \bar{\sigma}_j \cdot \bar{\lambda}^j \\ = \sum_{j=0}^n \sigma_j \cdot \bar{\lambda}^j = P_A(\bar{\lambda}) = 0$$

COROLARIO $A \in \mathbb{R}^{3 \times 3} \Rightarrow \exists$ almeno un autovalore reale

$A \in \mathbb{R}^{(2k+1) \times (2k+1)}$ $\Rightarrow \exists$ almeno un autovalore reale

DEF λ autovalore di A si dice **MATEPUCIA ALGEBRICA** di λ ,
la moltip. come radice del polinomio caratter.

ES $p_A(\lambda) = (-1)^n (\lambda - \lambda_1)^{n_1} \cdot \dots \cdot (\lambda - \lambda_s)^{n_s}$ con $s \leq n$

$$\sum_{j=1}^s n_j = n$$

n_j è la mlt. alg. di λ_j e si indica
con $\alpha(\lambda_j)$

DEF La **MATEPUCIA GEOMETRICA** di λ , $\gamma(\lambda)$ corrisponde
alla dimensione dello spazio vettoriale
 $\{v \in \mathbb{C}^n : (A - \lambda I)v = 0\} = n - \text{rk}(A - \lambda I)$

OSS In generale se $\lambda_1, \dots, \lambda_s$ sono gli autovalori
distinti si ha $\sum_{j=1}^s \gamma(\lambda_j) \leq n$

Più precis.: $1 \leq \gamma(\lambda_j) \leq \alpha(\lambda_j) \leq n$

OSS Se gli autovalori sono tutti distinti ($s=n$)
 $\Rightarrow \gamma(\lambda_j) = \alpha(\lambda_j) = 1 \quad \forall j = 1, \dots, n$

ESEMPIO

$$A = \begin{bmatrix} 2 & 1 & & \\ & \ddots & 1 & \\ & & \ddots & \\ & & & 2 \end{bmatrix} \quad \lambda = 2 \text{ è l'unico autovalore}$$

$$\alpha(\lambda) = n$$

$$p_A(\lambda) = \det \begin{pmatrix} 2-\lambda & 1 & & \\ & \ddots & 1 & \\ & & \ddots & \\ & & & 2-\lambda \end{pmatrix} = (2-\lambda)^2$$

$$A - \lambda I = A - 2I = \begin{bmatrix} 0 & 1 & & \\ & \ddots & 1 & \\ & & \ddots & \\ & & & 0 \end{bmatrix} \Rightarrow \text{HA RANGO } n-1$$

↓

$$\gamma(\lambda) = \gamma(2) - n - (n-1) = 1$$

ESEMPIO

$$A = I \quad P_A(\lambda) = (1-\lambda)^n \Rightarrow \lambda=1, \alpha(1)=n$$

$$A - \lambda I = 0 \Rightarrow rk = 0 \Rightarrow \gamma(1) = n - 0 = n$$

↓
qualsiasi base
di autovett. va
bene per I

DEF $A \in \mathbb{C}^{n \times n}$ è **DIAGONALIZZABILE** se:

$$\exists S \text{ matr. simbolare t.c. } S^{-1}AS = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

OSS $AS = S \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

$$A[S, | \dots | s_n] = [S, | \dots | s_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$[AS, | \dots | As_n] = [\lambda_1 s_1, | \dots | \lambda_n s_n]$$

↓
LE COLONNE DI S
SONO PEROVET. DI A

Dire che una matr. è diagonalizzabile equivale a dire che \exists una base di \mathbb{C}^n fatta di autovettori di A

TH A è diagonalizzabile $\Leftrightarrow \alpha(x_j) = \gamma(x_j) \quad \forall j = 1, \dots, s$

Corollario Se ho n autovoltori distinti $\Rightarrow A$ diagonalizzabile

OSS Matrici con stessi autovoltori $\not\Rightarrow A \sim B$

OSS A e B simili \Leftrightarrow la stessa forma di Jordan

TH $A \in \mathbb{C}^{n \times n}$, $q \in \mathbb{C}$, (λ, v) autocoppia per A
 $\Rightarrow (\lambda + q, v)$ è autocoppia per $A + q \cdot I$

DIM $(A + qI)v = Av + qv = \lambda v + qv = (\lambda + q)v$

OSS $I = A^0$

COROLARIO $p(x) = \sum_{j=0}^k p_j x^j$, si consideri $p(A) = \sum_{j=0}^k p_j A^j$
 \Rightarrow se (λ, v) autocoppia di A $\Rightarrow (p(\lambda), v)$ è
autocoppia
per $p(A)$

ESEMPIO

$J = \begin{bmatrix} & & 1 \\ & \ddots & \\ 1 & \ddots & \end{bmatrix}$ = ANTIIDENTITÀ = matrice di permutazione
ottenuta da I

J è simmetrica

$I = J^T J = J \cdot J = J^2 \Rightarrow \lambda$ autovettore di J, allora $\lambda^2 = 1$

$$\alpha(1) = ? \quad \alpha(-1) = ?$$

gli autovoltori di J
possono essere
solo ± 1

$$n=2 \quad J = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow \text{tr}(J) = 0$$

$$n=3 \quad J = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \Rightarrow \text{tr}(J) = 1$$

$$\text{tr}(J) = \begin{cases} 0 & \text{se } n \text{ pari} \\ 1 & \text{se } n \text{ dispari} \end{cases} \Rightarrow \alpha(1) = \alpha(-1) = n/2$$

$(n=2k+1)$

ESEMPIO

$$A = I + \alpha \cdot b^T \quad \alpha, b \in \mathbb{C}^n$$

Gli autovalori di A saranno della forma $1 + \lambda$ dove λ è autovalore di $\alpha \cdot b^T$

$$\alpha b^T = \begin{bmatrix} \alpha_1 b_1 & \alpha_1 b_2 & \dots & \alpha_1 b_n \\ \vdots & \vdots & & \vdots \\ \alpha_n b_1 & \alpha_n b_2 & \dots & \alpha_n b_n \end{bmatrix} = \begin{bmatrix} \alpha b_1 & | & \alpha b_2 & | & \dots & | & \alpha b_n \end{bmatrix} \Rightarrow \text{rk}(\alpha b^T) = 1$$

$\sigma_j = 0 \quad \forall j \geq 1$

$$\begin{aligned} P_{\alpha b^T}(\lambda) &= (-1)^n \lambda^n + (-1)^{n-1} \sigma_1 \lambda^{n-1} + \dots + \sigma_0 = \\ &= (-1)^n \lambda^n + (-1)^{n-1} \text{tr}(\alpha b^T) \lambda^{n-1} = (-1)^n + (-1)^{n-1} (\alpha^T b) \lambda^{n-1} \end{aligned}$$

$$P_{\alpha b^T}(\lambda) = 0 \iff \lambda^n - (\alpha^T b) \lambda^{n-1} = 0$$

$\lambda = 0 \text{ CON RLT. 1}$
 $\lambda = \alpha^T b \text{ CON RLT. 1}$

\Rightarrow A ha autovalori 1 e $1 + \alpha^T b$

$$\text{Com } \alpha(1) = n-1, \alpha(1 + \alpha^T b) = 1$$

DEF $A \in \mathbb{C}^{n \times n}$ $\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$

$$F_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq p_i\} \quad p_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$$

\downarrow CERCHIO
DI GERSHGORIN
(ASSOCIAZIONE ALLE RIGA i)

TH (I TH DI GERSHGORIN)

$$\lambda \text{ autovettore di } A \Rightarrow \lambda \in \bigcup_{i=1}^n F_i$$

(tutti gli autovettori cadono all'interno dei cerchi di Gershgorin)

DIM $Ax = \lambda x, x \neq 0, x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ assumiamo che x_k sia la componente di massimo modulo $|x_n| \geq |x_j| \forall j$

$$\begin{array}{c} A \\ \hline \hline \end{array} \quad \boxed{x} = \lambda \boxed{x}$$

$$\text{La } k\text{-esima eq.} \Rightarrow \sum_{j=1}^n a_{kj} x_j = \lambda x_k = \lambda(\lambda - a_{kk}) x_k = \sum_{\substack{j=1 \\ j \neq k}}^n a_{kj} x_j$$

$$\Rightarrow |\lambda - a_{kk}| |x_k| = \left| \sum_{\substack{j=1 \\ j \neq k}}^n a_{kj} x_j \right| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |a_{kj}| |x_j| \quad \Downarrow \quad \frac{\leq 1}{|\lambda - a_{kk}| \leq \sum |a_{kj}| \frac{|x_j|}{|x_k|}}$$

[OSS $|x_k| > 0$ altrimenti $x = 0$]

$$\Rightarrow \lambda \in F_k \subseteq \bigcup_{i=1}^n F_i \quad \leq \sum_{j \neq k} |a_{kj}| = p_k$$

COROLLAIO Se matrice a predominanza diagonale forte sono mat simiglanti (\Rightarrow invertibili)

DIM Per def. $|a_{jj}| \geq p_j$

$$0 \notin \bigcup_{i=1}^n F_i$$

TH (II TH DI GERSHGORIN)

Si sono H_1 , l'unione dei k cerchi di Gershgorin ed H_2 l'unione di $n-k$ cerchi, t.c. $H_1 \cap H_2 = \emptyset$ allora k autovalori stanno in H_1 e $n-k$ autovalori stanno in H_2 (esattamente)

COROLLAIO Se tutti gli F_i sono disgiunti \Rightarrow tutti gli autovalori sono distinti (\Rightarrow A diagonale)

DEF $A \in \mathbb{C}^{n \times n}$, si considera il grafo orientato associato ad A come il grafo che ha come vertici $\{1, \dots, n\}$ e un arco dal modo i al modo j se e solo se $a_{ij} \neq 0$

ES.

$$\begin{bmatrix} 1 & 2 & \dots \\ 2 & \ddots & \vdots \\ \vdots & \ddots & 1 \\ & & 1 & 2 \end{bmatrix}$$



GRAFO FORTEMENTE CONNESSO
 \Rightarrow MAT. IRREDUCIBILE

$$\begin{bmatrix} 1 & \vdots & \ddots \\ \vdots & \ddots & \vdots \\ & \ddots & 1 \end{bmatrix}$$



DEF

Un grafo orientato si dice fortemente connesso se $\forall i, j \in \{1, \dots, n\} \exists$ un cammino orientato da i a j

DEF $A \in \mathbb{C}^{n \times n}$ si dice irriducibile se il grafo associato ad esso associato è fortemente connesso

TH (III TH DI GERSHGORIN)

Se A è irriducibile allora:

se $\lambda \in \partial F_i = \{z \in \mathbb{C} : |z - a_{ii}| = r_i\}$ per un qualche i
 \downarrow BORDO
DEL DISCONE
(FONTELLA)

Allora $\lambda \in \partial F_i \forall i = 1, \dots, n$

COROLARIO Se A è a predom. diag. debole ed è irriducibile $\Rightarrow A$ è non singolare

DIM Per def. $|a_{ii}| \geq r_i$ ma \exists almeno un indice k :

$$|a_{kk}| > r_k$$

per III Gersh.

per cui $0 \notin \partial F_k \Rightarrow 0$ non è un autovettore

ESEMPIO

$$(*) x^k + a_{k-1}x^{k-1} + \dots + a_1x + a_0 = 0 \quad \text{POL. NON CO}$$

Le soluzioni di $(*)$ coincidono con gli autovettori di:

$$F = \begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & 1 & \\ -a_0 - a_1 - \dots - a_{n-k} & & & 0 \end{bmatrix} \quad \text{MATRICE DI FABRÉNIUS (COMPANION)}$$

OSS Per raffinare la stima si può applicare Gersh anche su A^T ($0 \neq 1$ nell'es.) e si ottiene un'altra unione di cerchi $\bigcup_{i=1}^n G_i$

Per cui si può affermare che gli autovalori appartengono all'insieme

$$\left(\bigcup_{i=1}^n F_i \right) \cap \left(\bigcup_{i=1}^n G_i \right)$$

OSS $(\bigcup F_i) \cap (\bigcup G_i) \neq \bigcup (F_i \cap G_i)$

NORME VETTORIALI E METRICHE

DEF $\| \cdot \| : \mathbb{C}^n \rightarrow \mathbb{R}_+^+$ è una norma se verifica:

- 1) $\| x \| = 0 \iff x = 0$
- 2) $\| \alpha x \| = |\alpha| \cdot \| x \| \quad \alpha \in \mathbb{C}$

- 3) $\| x + y \| \leq \| x \| + \| y \|$ DISUB. TRIANGOLARE

ESEMPI

norme p , $p \in \mathbb{N}$

$$\mathbb{C}^n \ni \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \| x \|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$$

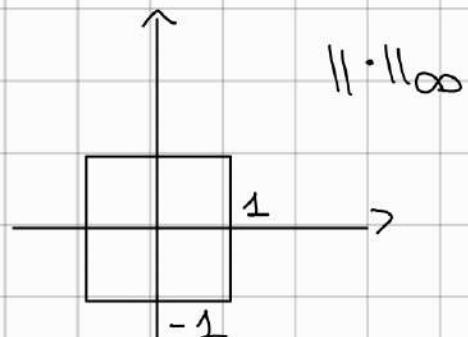
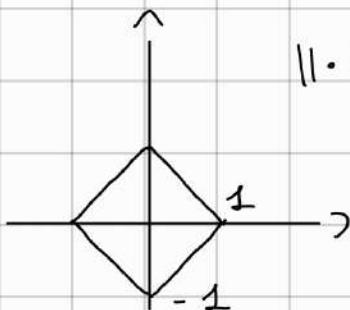
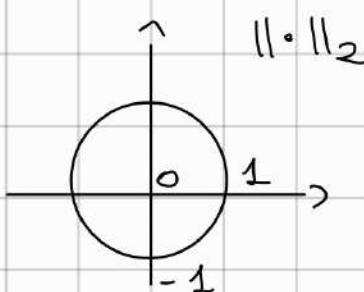
- $\| x \|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$ NORMA 2 (NORMA EUCLIDEA)

- $\| x \|_1 = \sum |x_i|$

- $\| x \|_\infty = \max_{i=1, \dots, n} |x_i|$

ESEMPIO

in \mathbb{R}^2





OSS Le norme sono funzioni continue su $\mathbb{R}^n / \mathbb{C}^m$

OSS $\|x\|_2^2 = x^* x$ su \mathbb{C}^n o $\bar{x}x$ su \mathbb{R}^n

DEF $\| \cdot \|_p$ e $\| \cdot \|_q$ si dicono equivalenti se $\exists \alpha, \beta \in \mathbb{R}^+$
t.c.:

$$\forall x \in \mathbb{C}^n, \alpha \cdot \|x\|_p \leq \|x\|_q \leq \beta \cdot \|x\|_p \quad (*)$$

OSS $(*) \Rightarrow \bar{\beta} \|x\|_q \leq \|x\|_p \leq \bar{\alpha} \|x\|_q$

Per le norme più comuni valgono:

- $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \cdot \|x\|_2$
- $\|x\|_\infty \leq \|x\|_1 \leq n \|x\|_\infty$
- $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty$

DEF $\| \cdot \| : \mathbb{C}^{n \times n} \rightarrow \mathbb{R}_0^+$ che verifica:

$$1) \|A\| = 0 \iff A = 0$$

$$2) \| \alpha A \| = |\alpha| \|A\|$$

$$3) \|A + B\| \leq \|A\| + \|B\|$$

$$4) \|A \cdot B\| \leq \|A\| \cdot \|B\| \quad \text{SUBSTITUZIONE}$$

DEF Una norma matriciale $\| \cdot \|_M$ si dice indotta da una norma vettoriale $\| \cdot \|_V$ se è definita come:

$$\|A\|_M = \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax\|_V}{\|x\|_V} = \sup_{x \in \mathbb{C}^n : \|x\|_V=1} \|Ax\|_V$$

Nel caso delle norme 1, 2, ∞ la norma matriciale indotta assume la seguente espressione:

- $\|A\|_1 = \max_{j=1, \dots, n} \sum_{i=1}^n |a_{ij}|$
- $\|A\|_2 = \sqrt{\rho(A^*A)}$
- $\|A\|_\infty = \max_{i=1, \dots, n} \sum_{j=1}^n |a_{ij}|$

DEF Data $\|\cdot\|_H$ norma matriciale e $\|\cdot\|_V$ norma vettoriale si dicono compatibili se vale che:

$$\forall A \in \mathbb{C}^{n \times n}, \forall x \in \mathbb{C}^n \quad \|Ax\|_V \leq \|A\|_H \cdot \|x\|_V$$

OSS Ogni norma matriciale indotta è compatibile con la corrispondente norma vettoriale

$$\forall x, \forall A : \frac{\|Ax\|_V}{\|x\|_V} \leq \|A\|_H \Rightarrow \|Ax\|_V \leq \|A\|_H \cdot \|x\|_V$$

ESEMPIO (di norma non indotta)

NORMA DI FROBENIUS: $\|A\|_F = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2} = \|\text{vec}(A)\|_2$

com $\text{vec}: \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n^2}$

Non è indotta perché $\|\mathbb{I}\|_F = \|\begin{bmatrix} 1 & \dots & 1 \end{bmatrix}\|_F = \sqrt{n}$
 e se fosse indotta si avrebbe $\sup_{x \neq 0} \frac{\|\mathbb{I}x\|}{\|x\|} = \sup_{x \neq 0} \frac{\|x\|}{\|x\|} = 1$

PROPRIETÀ Ogni matrice indotta $\|\cdot\|_M$ è tale per cui
 $\|I\|_M = 1$

TH DI HIRSCH $A \in \mathbb{C}^{n \times n}$ e $\|\cdot\|$ matrice matriciale
allora $\varphi(A) \leq \|A\|$

OSS Hirsch \Rightarrow gli autovalori appartengono tutti al cerchio di centro 0 e raggio $\|A\|$, per una qualsiasi norma

DIM $Ax = \lambda x, x \neq 0, \lambda \in \mathbb{C}$

e formiamo la matrice $B := [x | \overset{\circ}{0} | \dots | \overset{\circ}{0}]$

$$A \cdot B = [A \cdot x | \overset{\circ}{0} | \dots | \overset{\circ}{0}] = [\lambda x | \overset{\circ}{0} | \dots | \overset{\circ}{0}] = \lambda B$$

passando alle norme:

$$\|A\| \cdot \|B\| \geq \|A \cdot B\| = \|\lambda B\| = |\lambda| \cdot \|B\| \Rightarrow \|A\| \geq |\lambda|$$

$$\begin{array}{l} B \neq 0 \\ \|B\| \neq 0 \end{array}$$

↑
vole
autovalore

$$\Rightarrow \|A\| \geq \varphi(A)$$

TH DI HIRSCH $\| \cdot \|$ norma matriciale, $A \in \mathbb{C}^{n \times n}$

$$\rho(A) \leq \| A \|$$

COROLARIO Se $\| A \| < 1$ per una qualche norma
 $\Rightarrow A$ è convergente ($\lim_{k \rightarrow +\infty} A^k = 0$)

OSS Se $|\lambda| = \| A \|$ per una qualche norma
 $\Rightarrow |\lambda| = \rho(A)$

ESEMPIO (MATRICE STOCASTICA) $A \in \mathbb{R}^{n \times n}$, $a_{ij} \geq 0 \quad \forall i, j$

ed inoltre $\sum_{j=1}^n a_{ij} = 1 \quad \forall i = 1, \dots, n$

$$A \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} =: e \quad e \text{ è autovettore di} \\ \text{associato all'autovaleure 1}$$

$$1 = \max_{i=1, \dots, n} \sum_{j=1}^n a_{ij} = \max_{i=1, \dots, n} |a_{ij}| = \| A \|_{\infty} \Rightarrow \rho(A) = 1$$

PROPRIETÀ (INVARIANZA PER TRAVERS. CON MATRICI UNITARIE)

PER $\| \cdot \|_2 \in \| \cdot \|_F$

Siamo $x \in \mathbb{C}^n$, $A \in \mathbb{C}^{n \times n}$, $Q \in \mathbb{C}^{n \times n}$, $Q^* Q = Q Q^* = I$

Allora:

- 1) $\| x \|_2 = \| Qx \|_2$
- 2) $\| A \|_2 = \| QA \|_2 = \| AQ \|_2$
- 3) $\| A \|_F = \| QA \|_F = \| AQ \|_F$

DIM

$$1) \| x \|_2^2 = x^* x$$

$$\| Qx \|_2^2 = (Qx)^* Qx = x^* \underbrace{Q^* Q}_{I} x = x^* x$$

$$2) \|A\|_2^2 = \rho(A^*A)$$

$$\|QA\|_2^2 = P((QA)^*QA) = P(\underbrace{A^*Q^*QA}_{I}) = P(A^*A)$$

$$\|AQ\|_2^2 = P((AQ)^*AQ) = P(\underbrace{Q^*A^*AQ}_{Q^{-1}A^*AQ \sim A^*A}) = \rho(A^*A)$$

MATRICI SIMILI HANNO
GLI STESSI AUTOVALORI
 \Rightarrow HANNO LO STESSO
RAGGIO SPECTRALE

$$3) \|A\|_F = \sqrt{\sum_{i,j=1}^n |\alpha_{ij}|^2} \quad \|A\|_F^2 = \sum |\alpha_{ij}|^2$$

$$A = [\alpha_1 | \dots | \alpha_n] \Rightarrow \|A\|_F^2 = \|\alpha_1\|_F^2 + \dots + \|\alpha_n\|_F^2 = \\ = \alpha_1^* \alpha_1 + \dots + \alpha_n^* \alpha_n \\ = \text{tr}(A^*A)$$

$$A^*A = \begin{bmatrix} \alpha_1^* \alpha_1 & \dots \\ \vdots & \ddots \\ \alpha_n^* \alpha_n & \end{bmatrix}$$

$$\|A\|_F = \sqrt{\text{tr}(A^*A)}$$

$$\|QA\|_F^2 = \text{tr}((QA)^*QA) = \text{tr}(\underbrace{A^*Q^*QA}_{I}) = \text{tr}(A^*A)$$

$$\|AQ\|_F^2 = \text{tr}((AQ)^*AQ) = \text{tr}(\underbrace{Q^*A^*AQ}_{\text{SIMILE AD } A^*A}) = \text{tr}(A^*A) \\ \Rightarrow \text{STESSA TORCICA}$$

OSS Funzioni che dipendono solo dagli autovalori come ad esempio tr , \det non cambiano su matrici simili

OSS Se $A = A^* \Rightarrow \|A\|_2 = \rho(A)$

$$\|A\|_2 = \sqrt{\rho(A^*A)} = \sqrt{\rho(A^2)} = \sqrt{\rho(A)^2} = \rho(A)$$

\downarrow
 $A^* = A$

perché A^2
HA AUTOVALORI λ^2

OSS $\rho(A) \cdot \theta \geq \|A\|$? NO

↓
COSTANTE
ARBITRARIA

$\rho(A)$ può essere arbitr. piccolo e $\|A\|$ arbitr.
grande

ESEMPIO

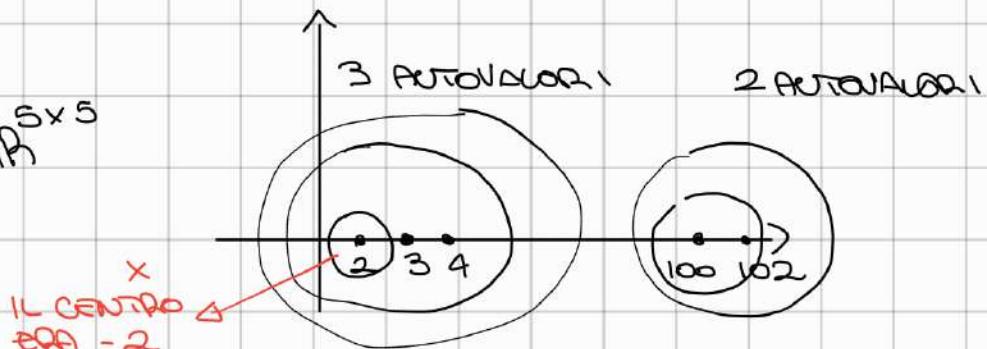
$$A = \begin{bmatrix} 0 & 0 \\ \alpha & 0 \end{bmatrix}$$

$$\rho(A) = 0$$

$$\|A\|_\infty = \|A\|_1 = |\alpha|$$

ESEMPIO (COMBINARE SIMILITUDINI E TH DI GERSHGORIN)

$$A = \begin{bmatrix} -2 & 1 & & & \\ -1 & 100 & 2 & & \\ & -2 & 3 & 3 & \\ & -3 & 102 & 4 & \\ & -4 & 4 & & \end{bmatrix} \in \mathbb{R}^{5 \times 5}$$



Se considero $S^{-1}AS$ e applico Gershgorin posso ottenere più info, idealmente vorremmo cerchi isolati e di raggio piccolo (stime più accurate)

$$S = \begin{bmatrix} 1 & S^{-1} & & \\ & 1 & S^{-1} & \\ & & 1 & \\ & & & 1 \end{bmatrix} \rightsquigarrow S^{-1} = \begin{bmatrix} 1 & S & & \\ & 1 & S & \\ & & 1 & \\ & & & 1 \end{bmatrix} \quad S \text{ PARAMETRI}$$

$$S^{-1}AS = \begin{bmatrix} 2 & S^{-1} & & & \\ -S & 100 & 2S & & \\ & -2S^{-1} & 3 & 3S^{-1} & \\ & & 102 & 4S & \\ & & & -4S^{-1} & 4 \end{bmatrix}$$

$$C(x, r)$$

$$S = 1/10$$

$$C(2, S^{-1})$$

$$C(100, 3/10)$$

$$C(100, 3S)$$

$$C(102, 2/5)$$

$$C(3, 5S^{-1})$$

$$S = 10$$

$$C(102, 4S)$$

$$C(2, 1/10)$$

$$C(4, 4S^{-1})$$

$$C(3, 1/2)$$

$$C(4, 2/5)$$

Ognuno dei cerchi \rightarrow contiene esattamente 1 autovalore

Dato che A è reale \Rightarrow tutti gli autovalori sono reali:

$$100 - \frac{3}{10} \leq \lambda_1 \leq 100 + \frac{3}{10}$$

$$102 - \frac{2}{5} \leq \lambda_2 \leq 102 + \frac{2}{5}$$

⋮

$$4 - \frac{2}{5} \leq \lambda_5 \leq 4 + \frac{2}{5}$$

piccolo

Ottimo $S: \overline{S} \cdot 3 + 5S^{-1} < 100 - 3S$

SISTEMI LINEARI

MATRICE A BLOCCI:

$A_{m \times n}$ si può vedere come matrice a blocchi $s \times t$
se:

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{t1} & A_{t2} & \dots & A_{ts} \end{bmatrix}$$

dove A_{ij} è $m_i \times n_j$

$$m_1 + m_2 + \dots + m_t = m$$

$$n_1 + n_2 + \dots + n_s = n$$

Possiamo definire somma e moltiplicazione di matrici a blocchi, perché i blocchi siano di dim. compatibili.

$$\begin{array}{c|c|c|c|c} A_{11} & & & B_{11} & \\ \hline & & & & \\ \hline & & & & \\ \hline \end{array} + \begin{array}{c|c|c|c|c} & & & B_{11} & \\ \hline & & & & \\ \hline & & & & \\ \hline \end{array} = \begin{array}{c|c} \hline A_{11} + B_{11} & A_{12} + B_{12} \\ \hline \hline A_{21} + A_{21} & A_{22} + B_{22} \\ \hline \end{array}$$

$$A \cdot B = \begin{array}{c|c} \hline A_{11}B_{11} + A_{12}B_{21} & \\ \hline \hline & \\ \hline \end{array}$$

ES

$$A = \begin{bmatrix} I & E^T \\ E & I \end{bmatrix}$$

$$\begin{bmatrix} T_{11} & & & \\ T_{21} & T_{22} & & \\ \vdots & \vdots & \ddots & \\ T_{n1} & T_{n2} & \dots & T_{nn} \end{bmatrix}$$

T_{ij} quadrati $\forall j \Rightarrow \det(T) = \det(T_1) \cdots \det(T_n)$

DEF (MATRICE RIDUCIBILE)

A $n \times n$ si dice riducibile se esiste π matrice di permutazione tale che:

$$\pi A \pi^T = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

K col (n-k) col
K righe (n-k) righe

$$Ax = b \rightsquigarrow \underbrace{\pi A x = \pi b}_{\text{SISTEMA EQUIVALENTE}} \rightsquigarrow (\pi A \pi^T)(\pi x) = \pi b$$

A Ax = b

$$x = x_1 \quad c = \pi b = c_1 \\ x_2 \quad c_2$$

$$A_{22} x_2 = c_2 \rightarrow \text{RISOLVO QUESTO SIST. } (n-k) \times (n-k) \text{ E TROVO } x_2$$

$$A_{11} x_1 + A_{12} x_2 = c_1 \rightarrow A_{11} x_1 = c_1 - A_{12} x_2 \rightarrow$$

RISOLVO
QUESTO
SIST.
 $k \times k$

In conclusione:

Indice di un sistema $n \times n$ posso risolverne 2, uno $(n-k) \times (n-k)$ e l'altro $k \times k$.

È un vantaggio perché risolvere un sistema lineare $n \times n$ ha un costo cubico $O(n^3)$.

Il caso in cui il vantaggio è più alto è: $k = n/2$.
 $O((n-k/2)^3 + (n/2)^3) = O(n^3/4)$

Non migliora la complessità asintotica, ma si guadagna un fattore 4 rispetto a risolvere il problema $n \times n$

N.B.: Π non è unica

TH Una matrice A è riducibile quando il grafo orientato associato ad A non è formalmente连通 (In particolare una matrice è riducibile quando non è irreducibile)

$G(A)$ grafo orientato associato ad A_{nxn} , $G(A)$ ha vertici $\{1, \dots, n\}$ ed ha un arco da i a j se e solo se $a_{ij} \neq 0$ (con $i \neq j$)

DIM Osserviamo che $\Pi A \Pi^T$ ed A hanno lo stesso grafo a meno di renomimare i vertici quindi $G(A)$ fort. connesso $\Rightarrow G(\Pi A \Pi^T)$ lo è
 \Rightarrow)

$$\begin{matrix} k & \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \\ n-k & \end{matrix}$$

i vertici $\{k+1, \dots, n\}$ non possono raggiungere $\{1, \dots, k\}$
 $\Rightarrow G(A)$ non è fort. connesso

\Leftarrow) $G(A)$ non fort. connesso \Rightarrow esistono i, j tali per cui non posso andare da i a j . Considero $S = \{$ insieme dei vertici raggiungibili da $i\}$, $T = \{$ insieme dei vertici non raggiungibili da $i\}$.

S è non vuoto perché contiene almeno i vertici i , T contiene i vertici j tali per cui non posso andare da i a j .

Inoltre $S \cup T = \{1, \dots, n\}$ ed inoltre $S \cap T$ non hanno vertici in comune altrimenti potrei andare da i in un vertice in T .

Considero una permutazione π che sposta tutti gli indici corrispondenti ai vertici in T ed in coda quelli raggiungibili da:

$T \quad S$

$$T \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \Rightarrow A \text{ è riducibile}$$

OSS Caso particolare di matrice riducibile

$$\pi A \pi^T = \begin{bmatrix} T & S \\ A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \quad \text{DIAGONALE A BLOCCHI}$$

Sì, fra quantità non ci sono archi da S a T ma anche: non ci sono archi da T a S

$$\pi A \pi^T = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Rightarrow \begin{cases} A_{11}x_1 = c_1 \\ A_{22}x_2 = c_2 \end{cases}$$

DEVO SEMPRE
RISOLVERE UN SISTEMA
 $K \times K$ ED UNO
 $(n-K) \times (n-K)$

IMPORTANTE: posso risolvere i due sistemi in parallelo

ESEMPIO

$$A = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 2 & 3 & -2 & 1 \\ -1 & 0 & -2 & 0 \\ 1 & -1 & 1 & 4 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -1 \\ -2 \end{bmatrix}$$

Risolvere $Ax = b$ secondo la permutazione

$$\pi = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$\pi = \pi^T$ in questo caso

$$\pi A \pi^T = \left[\begin{array}{cc|cc} 4 & -1 & 1 & 1 \\ 1 & 3 & -2 & 2 \\ \hline 0 & 0 & -2 & -1 \\ 0 & 0 & -1 & 1 \end{array} \right]$$

$$y = \pi x = \begin{bmatrix} x_4 \\ x_2 \\ x_3 \\ x_1 \end{bmatrix}$$

$$c = \pi b = \begin{bmatrix} -2 \\ -2 \\ -1 \\ 1 \end{bmatrix}$$

$$\begin{cases} \begin{bmatrix} -2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_3 \\ x_1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \end{cases} \rightsquigarrow \begin{bmatrix} x_3 \\ x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

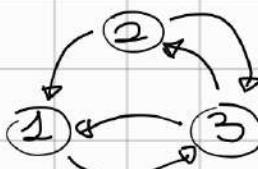
$$\begin{cases} \begin{bmatrix} 4 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_4 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} x_3 \\ x_1 \end{bmatrix} \end{cases} \rightsquigarrow \begin{bmatrix} x_4 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$\pi A \rightsquigarrow$ sposta le righe di A
 $(\pi A)\pi^T \rightsquigarrow$ sposta le colonne di πA

ESEMPIO

$$A = \begin{bmatrix} -1 & 0 & 5 \\ 2 & 12 & 13 \\ -1 & 2 & 0 \end{bmatrix}$$

$$G(A) \rightarrow$$

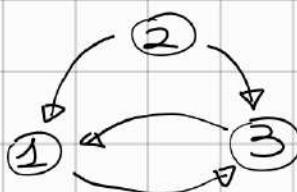


$G(A)$ FORTEMENTE CONNESSO

\Downarrow
 A è irriducibile

ESEMPIO

$$A = \begin{bmatrix} 3 & 0 & 6 \\ 2 & 4 & * \\ 3 & 0 & 0 \end{bmatrix}$$



$G(A)$ NON È FORTEMENTE CONNESSO

(DA 1 E 3 NON POSSO RAGGIUNGERE 2)

\Downarrow

A è riducibile

ESEMPIO

$$A = \begin{bmatrix} 0 & 1 & & \\ \vdots & \ddots & \vdots & 1 \\ 1 & & \ddots & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}$$



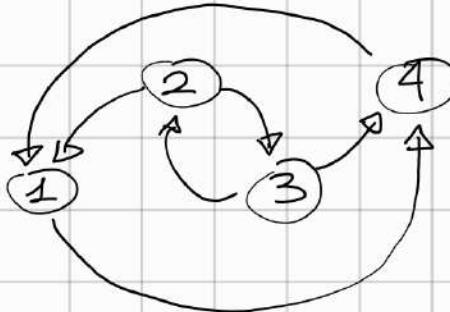
$G(A)$ È FORTEMENTE CONNESSO



A è riducibile

ESEMPIO

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 2 & 5 & 0 & -3 \\ 0 & 1 & 3 & 6 \\ 2 & 0 & 0 & -1 \end{bmatrix}$$



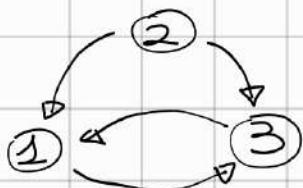
$G(A)$ NON È FORTEMENTE CONNESSO



A è riducibile

COME SI TROVA π

$$A = \begin{bmatrix} 3 & 0 & 6 \\ 2 & 4 & \neq \\ 3 & 0 & 0 \end{bmatrix}$$



NODI	NODI ABBIGLIUNGIBILI	NODI NON ABBG.
1	1, 3	2
2	1, 2, 3	\emptyset
3	1, 3	2

π permutazione che sposta nell'ordine

$\frac{2}{1}, \frac{3}{3}$

$$\pi = \begin{bmatrix} 0 & 1 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 1 \end{bmatrix} = \pi^T$$

$$\pi A \pi^T = \begin{bmatrix} \boxed{4} & \boxed{2} & \neq \\ \boxed{0} & \boxed{3} & 6 \\ 0 & 3 & 0 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

OSS Potrei continuare a ridurre i blocchi diagonali (in generale)

$$\pi_1 A \pi_1^T = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} =: A_1$$

Se A_{11} è riducibile o A_{22} è riducibile posso continuare

$$\pi_2 A_1 \pi_2^T = \pi_2 \pi_1 A \pi_1^T \pi_2^T = \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \quad \begin{bmatrix} * \\ 0 \end{bmatrix} \quad \begin{bmatrix} * & * \\ 0 & * \end{bmatrix}$$

e continuare così finché non ottengo che i blocchi diagonali sono tutti irriducibili.

$$\pi_k \pi_{k-1} \dots \pi_1 A \pi_1^T \dots \pi_{k-1}^T \pi_k^T = \begin{bmatrix} \boxed{A_{11}} & \boxed{A_{12}} & \dots & \boxed{A_{1h}} \\ \boxed{A_{21}} & \dots & \boxed{A_{2h}} \\ \vdots & & \ddots & \boxed{A_{hh}} \end{bmatrix}$$

In fondo al processo di riduzione π è triangolare sup. a blocchi con blocchi diagonali A_{jj} quadrati ed irriducibili.

$$\left\{ \begin{array}{l} A_{11}y_1 - \dots - A_{1h}y_h = c_1 \\ \vdots \\ A_{h-1,h-1}y_{h-1} + A_{h-1,h}y_h = c_{h-1} \\ A_{hh}y_h = c_h \end{array} \right. \quad Ax = b \rightsquigarrow Ty = C \rightsquigarrow$$

$$\pi = \pi_k \dots \pi_1 \text{ è una matrice di permutazione} \quad T = \pi A \pi^T, y = \pi x, c = \pi b \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_h \end{bmatrix} \quad c = \begin{bmatrix} c_1 \\ \vdots \\ c_h \end{bmatrix}$$

$\pi = \pi_k \dots \pi_1$ è una matrice di permutazione

- 1) Si risolve $A_{hh}y_h = c_h$ (si ottiene y_h)
- 2) Si risolve $A_{h-1,h-1}y_{h-1} + A_{h-1,h}y_h = c_{h-1}$ (si ottiene y_{h-1})
⋮

- h) Si risolve $A_{11}y_1 = c_1 - \dots$

METODO DI SOSTITUZIONE

ALL'INDIETRO
(BACK SUBSTITUTION)

OSS Se A non è singolare ($\det(A) \neq 0$) allora tutti i sistemi lineari visti in questo procedimento trovano soluzione unica poiché $\det(A_{jj}) \neq 0$ $\forall j = 1, \dots, n$

$$T = \pi A \pi^T \Rightarrow T \sim A \Rightarrow \det(T) = \det(A) \neq 0$$

$$\downarrow \quad \quad \quad //$$

$$\begin{bmatrix} T_{11} & \cdots & \boxed{T_{1j}} \\ \ddots & \ddots & \vdots \\ & & T_{hh} \end{bmatrix} \Rightarrow \det(T) = \prod_{j=1}^n \det(T_{jj})$$

\Downarrow

$$\det(T_{jj}) \neq 0 \quad \forall j$$

$$\left\{ \begin{array}{l} Ax = b \quad A \in \mathbb{C}^{n \times n} \quad b \in \mathbb{C}^n \quad x \in \mathbb{C}^n \\ a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{nn}x_1 + \dots + a_{nn}x_n = b_n \end{array} \right.$$

METODI DI RISOLUZIONE DI SISTEMI LINEARI

- 1) **METODI DIRETTI**: metodi che, se non ci fossero errori di arrotondament., ci darebbero la sol. esatta in un num. finito di passi (es. tramite formula esplicita per calc. la sol.)
- 2) **METODI ITERATIVI**: metodi che generano una successione di sol. approssimate t.c. $\lim_{k \rightarrow +\infty} x_k = x$

Per problemi di piccola / media dim. si usano metodi diretti (+ costosi ma + precisi), per problemi di grande dim. si usano metodi iterativi (- costosi ma - precisi)

METODO DI ELIMINAZIONE DI GAUSS (METODO DIRETTO)

(in MATLAB per risolvere $Ax = b$ $x = A \setminus b$)

Anche qui si vuole trovare un sistema equivalente $Ax = b \iff Ux = c$ dove U è triang. sup.

$$U = \begin{bmatrix} u_{11} & \dots & u_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & u_{nn} \end{bmatrix} \iff \begin{cases} u_{11}x_1 + \dots + u_{1n}x_n = c_1 \\ \vdots \\ u_{n-1,n-1}x_{n-1} + u_{n-1,n}x_n = c_{n-1} \\ u_{nn}x_n = c_n \end{cases}$$

$$\left\{ \begin{array}{l} x_n = c_n / u_{nn} \\ x_j = (c_j - \sum_{i=j+1}^n u_{ji}x_i) / u_{jj} \quad j = n-1, \dots, 1 \end{array} \right.$$

**METODO
DI SOSTITUZIONE
ALL'INDIETRO**

COSTO METODO DI SOST. AL'INDIETRO:

Al passo j : $2(n-j)+1$

$$\text{costo totale: } \sum_{j=1}^n 2(n-j)+1 = 2 \sum_{j=1}^{n-1} j + n = 2 \cdot \frac{n(n-1)}{2} + n = O(n^2)$$

COME SI ARRIVA A $Ux = C$:

Al primo passo voglio annullare gli elem. della 1° colonna che stanno sotto la diag.

$$\begin{bmatrix} * & \dots & * \\ \vdots & & \vdots \\ * & \dots & * \end{bmatrix} \rightsquigarrow \begin{bmatrix} * & * & \dots & * \\ 0 & ; & & ; \\ \vdots & ; & & \vdots \\ 0 & * & \dots & * \end{bmatrix}$$

Il sistema non cambia se riempio un'eq. con se stessa più λ volte un'altra con $\lambda \in \mathbb{C}$

$$a_{11} \dots a_{1n} \mid b_1 \quad a_{11} \neq 0$$

$a_{21} \dots a_{2n} \mid b_2$ dovrò sommare la 1° riga moltiplicata

$$a_{31} \dots a_{3n} \mid b_3 \text{ per } -\frac{a_{21}}{a_{11}} = l_{21}$$

$$a_{n1} \dots a_{nn} \mid b_n \quad -\frac{a_{31}}{a_{11}} = l_{31}$$

Moltiplicazioni

Dopo il 1° passo:

$$\begin{bmatrix} a_{11}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{2n}^{(2)} \\ \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & a_{nn}^{(2)} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(n)} \end{bmatrix}$$

$$a_{ij}^{(1)} = a_{ij} \quad b_1^{(1)} = b_1$$

$$a_{ij}^{(2)} = a_{ij}^{(1)} - a_{ij}^{(1)} \cdot l_{ii}$$

Al 2° passo, se $a_{22}^{(2)} \neq 0$

$$b_j^{(3)} = b_j^{(2)} - b_{i2} b_2^{(2)} \quad a_{ij}^{(3)} = a_{ij}^{(2)} - a_{ij}^{(2)} l_{i2} \quad l_{i2} = -\frac{a_{ij}^{(2)}}{a_{ii}^{(2)}}$$

Al passo j si modificano le righe da $j+1$ a n .

Facendo questo procedimento $n-1$ volte otengo:

$$U = \begin{bmatrix} a_{11}^{(1)} & \dots & a_{1n}^{(1)} \\ a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn}^{(n)} \end{bmatrix} \quad C = \begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(n)} \end{bmatrix}$$

COSTO DI TROVARE $U \in C$: $O\left(\frac{2}{3}n^3\right)$

OSS Per eseguire il metodo di Gauss è necessario avere al passo j $\underline{a_{jj}^{(j)}} \neq 0$ LO PIVOT

Questa condizione è equivalente ad avere nella matrice A tutti i minori principali di testa diversi da zero.

$$A = \boxed{\begin{array}{ccc|c} a_{11} & a_{12} & \dots & \\ a_{21} & a_{22} & \dots & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{array}}$$

$$a_{11} \neq 0, \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \neq 0, \dots, \det(A) \neq 0$$

OSS Questa condizione è verificata sicuramente se A è a predom. diag. forte e se

A è simmetrica e definita positiva (pos o neg)

$$A = A^* \quad x^* A x > 0 \quad \forall x \in \mathbb{C}$$

oppure

$$x^* A x < 0 \quad \forall x \in \mathbb{C}$$

In generali: non si può garantire $a_{jj}^{(j)} \neq 0$ e è problematico anche quando $|a_{jj}^{(j)}| \approx 0$

perché porta a instabilità numerica -

ESEMPIO

$$\begin{bmatrix} -2 & 1 & 1 \\ -1 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ -3 \\ 2 \end{bmatrix}$$

$$R_2 = R_2 - \frac{1}{2}R_1$$

$$\begin{array}{c|c} \left[\begin{array}{ccc|c} -2 & 1 & 1 & -2 \\ -1 & 3 & 1 & -3 \\ 1 & 1 & 2 & 2 \end{array} \right] & \xrightarrow{\text{R}_2 = R_2 - \frac{1}{2}R_1} \left[\begin{array}{ccc|c} -2 & 1 & 1 & -2 \\ 0 & 5/2 & 1/2 & -2 \\ 0 & 3/2 & 5/2 & 1 \end{array} \right] \\ \left[\begin{array}{ccc|c} -2 & 1 & 1 & -2 \\ 0 & 5/2 & 1/2 & -2 \\ 0 & 3/2 & 5/2 & 1 \end{array} \right] & \xrightarrow{\text{R}_3 = R_3 + \frac{1}{2}R_1} \left[\begin{array}{ccc|c} -2 & 1 & 1 & -2 \\ 0 & 5/2 & 1/2 & -2 \\ 0 & 0 & 11/5 & 11/5 \end{array} \right] \end{array}$$

$$\left\{ \begin{array}{l} -2x_1 = -2 + 1 - 1 = 1 \\ \frac{5}{2}x_2 = -2 - \frac{1}{2} = -\frac{5}{2} \\ \frac{11}{5}x_3 = \frac{11}{5} \Rightarrow x_3 = 1 \end{array} \right. \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

ESEMPIO 2

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 3 \\ 3 \end{bmatrix}$$

$$\begin{array}{c|c} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ -2 & 1 & 1 & 3 \\ -1 & 0 & 1 & 3 \end{array} \right] & \xrightarrow{\text{R}_2 = \text{R}_2 + 2\text{R}_1} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & -1 & 5 & 11 \\ -1 & 0 & 1 & 3 \end{array} \right] \\ & \xrightarrow{\text{R}_3 = \text{R}_3 + \text{R}_1} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & -1 & 5 & 11 \\ 0 & -1 & 3 & 7 \end{array} \right] \end{array}$$

$$R_3 = R_3 - R_2$$

$$\begin{array}{c} \xrightarrow{} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & -1 & 5 & 11 \\ 0 & 0 & -2 & -4 \end{array} \right] \end{array} \quad \left. \begin{array}{l} -2x_3 = -4 \Rightarrow x_3 = 2 \\ -x_2 + 10 = 11 \Rightarrow x_2 = -1 \\ x_1 + 1 + 2 = 4 \Rightarrow x_1 = -1 \end{array} \right\} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

OSS Si può fare anche per risolvere simultan.
 $Ax = b^{(1)}, \dots, Ax = b^{(s)}$ (stessa mat. dei coeff.
diversi termini noti)

$$Ax = B \quad B = [b^{(1)} | \dots | b^{(s)}] \in \mathbb{C}^{n \times s}$$

$$x = [x^{(1)} | \dots | x^{(s)}] \in \mathbb{C}^{n \times s}$$

$$\begin{matrix} [A|B] \\ \rightsquigarrow \\ [U|C] \end{matrix}$$

$$[U|C]$$

ed infine si applica la sost. all'indietro con
 U ed ogni colonna di C .

Ovvero $\forall j=1, \dots, n$ si risolve $Ux^{(j)} = C^{(j)}$

ESEMPIO (CALCOLO DELLA INVERSA)

$$S=n, B=I_n$$

$$Ax = I \Rightarrow x = A^{-1}$$

$$A = \left[\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 3/2 & 1/2 & -1/2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right] \rightsquigarrow$$

$$\left[\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 3/2 & 1/2 & -1/2 & 1 & 0 \\ 0 & 0 & 2/3 & 2/3 & -4/3 & 1 \end{array} \right]$$

$$\begin{cases} 2x_1 = 1+2-1 \Rightarrow x_1 = 1 \\ 3/2x_2 = -\frac{1}{2}-\frac{1}{2} \Rightarrow x_2 = -2/3 \\ \frac{2}{3}x_3 = \frac{2}{3} \Rightarrow x_3 = 1 \end{cases}$$

$$x = \begin{bmatrix} 1 & -1 & \cdot \\ -2/3 & 4/3 & \cdot \\ 1 & -2 & \cdot \end{bmatrix}$$

$$\begin{cases} 2x_1 = -4 + 2 \Rightarrow x_1 = -1 \\ 3x_2 = 1 + 1 \Rightarrow x_2 = 4/3 \\ \frac{2}{3}x_3 = -4/3 \Rightarrow x_3 = -2 \end{cases}$$

} 10EM 3. COLONNA

FATTORIZZAZIONE LU (VEDI ESSERE CASOGE 16/11/22)

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \rightarrow \begin{bmatrix} a_{11} & \dots & a_{1n} \\ 0 & * & \dots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \dots & * \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ l_{n1} & 0 & \dots & 0 & \dots & 0 \end{bmatrix}}_{H_1} \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

$$l_{21} = \frac{a_{21}}{a_{11}}$$

Trova il 1° passo di Gauss equivale a $A_1 = H_1 \cdot A$

Al passo j:

$$\underbrace{\begin{bmatrix} 1 & & & & \\ \ddots & \ddots & & & 0 \\ 0 & -l_{j+1,j} & \ddots & & \\ \vdots & & \ddots & & 1 \\ 0 & -l_{n,j} & & \ddots & \end{bmatrix}}_{H_j} \underbrace{\begin{bmatrix} * & \dots & * \\ \ddots & & 1 \\ 0 & * & \ddots & \vdots \\ & & \ddots & * \end{bmatrix}}_{A_{j-1}}$$

$$U = H_{n-1} \cdot H_{n-2} \cdot \dots \cdot H_1 \cdot A$$

$$A = H_1^{-1} \cdot \dots \cdot H_{n-1}^{-1} \cup$$

$H_j = \text{MAT. ELEM. DI GAUSS}$

$$H_i^{-1} = \begin{bmatrix} 1 & & & \\ l_{21} & \ddots & & \\ \vdots & & \ddots & \\ l_{ni} & & & 1 \end{bmatrix}, \dots, H_j^{-1} = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & l_{j+1,j} & \ddots & \\ & & l_{nj} & 1 \end{bmatrix}$$

$$L_i = H_i^{-1}, L_j H_j^{-1}$$

$$L_i = I + N_i = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 1 \end{bmatrix} + \begin{bmatrix} l_{21} & & & \\ & \ddots & & \\ & & l_{ni} & \\ & & & 0 \end{bmatrix}$$

$$L_j = I + N_j \quad N_i \cdot N_j = 0 \text{ se } i \neq j$$

$$A = L_1 \cdot \dots \cdot L_{n-1} \cdot U = (I + N_1) \cdot \dots \cdot (I + N_{n-1}) \cdot U =$$

$$= (I + N_1 + \dots + N_{n-1}) \cdot U = \begin{bmatrix} 1 & & & & 0 \\ l_{21} & 1 & & & \\ \vdots & l_{32} & \ddots & & \\ l_{ni} & l_{n2} & \dots & l_{nn-1} & 1 \end{bmatrix} U = L \cdot U$$

↑

I PRODOTTI CHE
CONTINUANO
PIÙ DI UN FATTORE
 N_j SI ANNULLANO

OSS Data A possiamo calcolare $A = LU$ senza
avere termini morti o un sistema da risolvere

Se abbiamo a disposizione la LU (precalcolata)
allora risolvere un sistema $Ax = b$ costa
 $O(n^2)$ in questo modo:

$$Ax = b \Leftrightarrow \underbrace{L \cdot U \cdot x}_{y} = b \Rightarrow \begin{cases} Ly = b & (\text{sost. A} \cdot \text{U' INDIRETTO}) \\ Ux = y & (\sim \sim \sim) \end{cases}$$

OSS Il fattore L ha solo 1 sulla diagonale per cui, per il Th di Bimat-Cauchy:

$$\det(A) = \det(LU) = \det(L) \cdot \det(U) = 1 \cdot \det(U) = \det(U)$$

11

Quindi il metodo di Gauss è anche il metodo più efficiente per calcolare $\det(A)$.

PRODOTTO
ELEMENTI
DIAGONALI
 $\prod_{j=i}^n a_{jj}^{(i)}$ (più)

Infatti, considerando lo sviluppo di Laplace costa $\mathcal{O}(n!)$ mentre con Gauss costa $\mathcal{O}(n^3)$

ES. PER $n=15$ Laplace impiega ore/giorni,
Gauss 10^{-4} / 10^{-5} secondi

ESEMPIO (METODO METODO DIRETTO)

METODO DI CRAMER

Per risolvere $Ax = b$:

$$x_j = \frac{\det(A_j)}{\det(A)}$$

A_j = mat. A a cui si sost. la colonna j con b

[IN MATLAB: $A_j = [A(:, 1:j-1), b, A(:, j+1:end)]$]

COSTO (COMBINATO CON GAUSS): $\mathcal{O}(n^4)$

=> inefficiente, ma può essere utile dal punto di vista tecnico

$$A = LU \quad L = \begin{bmatrix} 1 & & \\ \ddots & 1 & \\ * & & 1 \end{bmatrix}, \quad U = \begin{bmatrix} u_{11} & * & \\ \ddots & \ddots & \\ 0 & \ddots & u_{nn} \end{bmatrix}$$

\$u_{jj} = \partial_{jj}^{(j)}\$ = PIVOT DEL
PASSO j DELL'ALG.
DI GAUSS

OSS $\det(A) = \prod_{j=1}^n u_{jj} = \prod_{j=1}^n (\text{PIVOT}_j)$

TH $A \in \mathbb{C}^{n \times n}$ t.c. $\det(A(1:j, 1:j)) \neq 0$ per $j=1, \dots, n-1$
 ↗ MINORE DI TESTA
 DI DIM. j

Allora esiste ed è unica la fattorizzazione LU:

$$A = LU$$

OSS $\det(A(1:j, 1:j)) = \prod_{i=1}^j \partial_{ii}^{(i)}$

PIVOTING

Supponiamo che davanti l'algor. di Gauss si abbia:

PIVOTING FORZATO

$$\begin{array}{cccc|cc} * & \cdots & \cdots & * & & \\ \vdots & & & \vdots & & \\ * & \cdots & \cdots & * & & \\ 0 & * & \cdots & * & & \\ 0 & | & | & | & & \\ * & & & * & & \end{array}$$

IDEA: Si scambia la riga j con una riga tra la $(j+1)$ -esima e l' n -esima

ad esempio per r t.c. $\partial_{rj}^{(j)} \neq 0 \quad r > j$

In generale si fa sempre pivotting (scambio righe) portando in pos. (j,j) l'elem. di massimo modulo nel vettore $\begin{bmatrix} \partial_{jj}^{(j)} \\ \partial_{j+1,j}^{(j)} \\ \vdots \\ \partial_{n,j}^{(j)} \end{bmatrix}$

In questo modo sono sicuri che i moltiplicatori verifichino: $|\ell_{ji}| = \left| \frac{\partial_{ij}^{(j)}}{\partial_{jj}^{(j)}} \right| \leq 1 \quad i = j+1, \dots, n$

\Rightarrow gli elem. im L sono tutti ≤ 1

(questa proprietà garantisce stabilità numerica)

DEF Questa strategia di pivoting viene detta
PIVOTING PARZIALE (o per colonne)

Nel caso del pivoting parziale, la fattorizzazione LU che si ottiene alla fine verifica: $\Pi A = LU$

ESEMPIO

$$\begin{array}{cc} A & b \\ \left[\begin{array}{ccc|c} 0 & 3 & 9 & 50 \\ 4 & 5 & 6 & 32 \\ 1 & 2 & 3 & 14 \end{array} \right] & \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 4 & 5 & 6 & 32 \\ 0 & 8 & 9 & 50 \end{array} \right] \rightarrow \dots \end{array}$$

(interessa solo per fare conti con la fatt., per risolvere il sist. lineare non cambia: $\Pi Ax = \Pi b$)

Se $\det(A) \neq 0 \Rightarrow$ si può sempre applicare il pivoting parziale, ovvero non ci sotto mai:

$$A_j = \left[\begin{array}{cccc|c} * & * & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ * & * & * & \dots & * \\ 0 & 0 & 0 & \dots & * \\ 0 & 0 & 0 & \dots & * \\ \vdots & & & & \vdots \\ 0 & * & * & \dots & * \end{array} \right] \quad (\text{A})$$

Se $\det(A) = 0$ e si ha una situazione tipo (A)
si possono considerare sia scambi di righe che di colonne.

PERM. PER RIGHE



IDEA: considerare $\Pi_R A_j \Pi_C$ che sposta $a_{hk}^{(j)}$ $h \geq j$ in posizione (j,j)
PERM. PER COLONNE

In questo modo si scambia x_n con x_j

Più formalmente, se si applica Gauss con scambi sia di colonne che di righe, si risolve:

$$\Pi_1 A \Pi_2 \underbrace{\Pi_2^T x}_{y} = \Pi_1 b \quad \Pi_2^T y = x$$

Dal punto di vista della fattorizzazione: $\Pi_1 A \Pi_2 = LU$

PIVOTING ROTATO: ad ogni passo j si porta in pos. (j,j) l' elem. di massimo modulo im:

$$\begin{bmatrix} a_{jj}^{(j)} & \dots & a_{jn}^{(j)} \\ \vdots & & \\ a_{nj}^{(j)} & & a_{nn}^{(j)} \end{bmatrix}$$

+ costoso del pivoting parziale perché devo controllare al passo j $(n-j)^2$ entrate della matrice -

Se applico sia scambi di righe che di colonne posso sempre arrivare a:

$$\left[\begin{array}{cccc|c} * & \cdots & * & & \\ \vdots & & & & \\ & * & \cdots & * & \\ & 0 & \cdots & 0 & \\ \vdots & & & & \\ 0 & \cdots & 0 & & \end{array} \right] \xrightarrow{\text{K}} \left[\begin{array}{cccc|c} & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{array} \right] \xrightarrow{\text{S}}$$

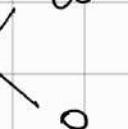
"I valori di K e S ci aiutano a capire quante soluzioni ha il sistema"

$K = \#$ pivot non nulli

$S = n - K$

$K = n \Rightarrow$ SOLUZIONE UNICA

∞ SOLUZIONI (SISTEMA SOUDETERMINATO)

$K < n \Rightarrow$  " (SISTEMA INCONSISTENTE)"

$$\left[\begin{array}{ccc|c} * & \dots & * & * \\ 0 & \dots & * & b_{k+1} \\ 0 & \dots & 0 & b_n \end{array} \right]$$

Se $\exists b_j \neq 0 \quad j \in \{k+1, \dots, n\}$
 $\Rightarrow \infty$ soluzioni

Altrimenti se

$$b_{k+1} = b_{k+2} = \dots = b_n = 0$$

$\Rightarrow \infty$ soluzioni

ESEMPIO

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 4 & 5 & 6 & 32 \\ 7 & 8 & 9 & 50 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 0 & -3 & -6 & -24 \\ 0 & -6 & -12 & -48 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 0 & -3 & -6 & -24 \\ 0 & 0 & 0 & 0 \end{array} \right] \Rightarrow \begin{array}{l} \text{Z PARATELLA} \\ \text{LIBERI} \end{array}$$

$$-3y = -24 + 6z \Rightarrow 3y = 8 - 2z$$

$$x = 14 - 2(8 - 2z) - 32 \Rightarrow x = -2 + z$$

INSIEME SOL: $\left\{ \begin{bmatrix} -2 \\ 8 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, z \in \mathbb{R} \right\}$

1 A PRENDERE
AFFINCHÉ VENGA z

ESEMPIO

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 4 & 5 & 6 & 28 \\ 3 & 6 & 9 & 42 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \quad y \in \mathbb{Z} \text{ PARATELLA LIBERI}$$

$$x = 14 - 2y - 3z$$

INSIEME SOL: $\left\{ \begin{bmatrix} 14 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, y \in \mathbb{R}, z \in \mathbb{R} \right\}$

AFFINCHÉ VENGA y

AFFINCHÉ VENGA z

Esercizio

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 4 & 5 & 6 & 32 \\ 7 & 8 & 9 & b \end{array} \right]$$

Per quali valori di a e b si hanno

- 1 sol.
- ∞ sol.
- 0 sol.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 0 & -3 & -6 & -24 \\ 0 & -6 & 2-21 & b-98 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 0 & -3 & -6 & -24 \\ 0 & 0 & a-9 & b-50 \end{array} \right]$$

$a \neq 9 \Rightarrow 1$ sol.

$a = 9, b = 50 \Rightarrow \infty$ sol.

$a = 9, b \neq 50 \Rightarrow 0$ sol.

INSIEME SOL. PER $a = 9, b = 50$:

$$\begin{aligned} -3y - 6z = -24 &\Rightarrow -3y = -24 + 6z \Rightarrow y = \frac{24 - 6z}{3} = 8 - 2z \\ x + 2y + 3z = 14 &\Rightarrow x = -2(8 - 2z) - 3z + 14 \\ &\Rightarrow x = -16 + 4z - 3z + 14 \\ &x = -2 + z \end{aligned}$$

$$\left\{ \begin{bmatrix} -2 \\ 8 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, y \in \mathbb{R}, z \in \mathbb{R} \right\}$$

$$\left[\begin{array}{ccc|c} * & \dots & * & * \\ \cdot & \ddots & \cdot & \vdots \\ \cdot & \cdot & * & * \end{array} \right] \quad \text{Se tutti i pivot sono } \neq 0 \Rightarrow 1 \text{ soluzione}$$

Altrimenti:

$$\left[\begin{array}{ccc|c} * & \dots & * & * \\ \cdot & \ddots & \cdot & \vdots \\ \cdot & \cdot & * & ? \\ 0 & \dots & 0 & ? \\ \cdot & \cdot & 0 & ? \end{array} \right]$$

Se tutti $? = 0 \Rightarrow \infty$ soluzioni
Se $\exists ? \neq 0 \Rightarrow$ SIST. IMPOSSIBILE

Possiamo sempre arrivare a uno dei casi sopra tramite pivoting (permutazioni) di colonne e righe.

SISTEMI RETTANGOLARI

$$m \neq n : A \in \mathbb{C}^{n \times n}$$

Se $m < n$ (SIST. SOUDETERMINATO)

Si può rendere il sistema quadrato aggiungendo ($n = m$) equazioni del tipo $0 = 0$

$*$ = elem $\neq 0$

APPLICO GAUSS A:

$$\left[\begin{array}{ccc|c} * & \dots & * & * \\ * & \dots & * & \vdots \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{array} \right] \quad \xrightarrow{\text{1 CASO}}$$

$$\left[\begin{array}{ccc|c} * & \dots & * & * \\ 0 & \dots & * & \vdots \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{array} \right] \quad \Rightarrow \infty \text{ SOL.}$$

2 CASO

$\left. \begin{matrix} n-m \\ \text{righe} \end{matrix} \right\}$

$$\left[\begin{array}{ccc|c} * & \dots & * & * \\ 0 & \dots & * & \vdots \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{array} \right] \quad \begin{matrix} ? = 0 \Rightarrow \infty \\ \text{SOL.} \end{matrix}$$

$\exists ? \neq 0 \Rightarrow$ SIST. IMPOSS.

CASO $m > n$ (SIST. SOTTODETERMINATO)

1 CASO

$$\left[\begin{array}{ccc|c} * & \dots & * & * \\ \cdot & \ddots & \cdot & \vdots \\ \cdot & \cdot & * & ? \\ 0 & \dots & 0 & ? \end{array} \right]$$

Tutti $? = 0 \Rightarrow 1$ soluzione

Dopo GAUSS:

Se $\exists ? \neq 0 \Rightarrow$ SIST. IMPOSSIBILE

2° CASO

$$\left[\begin{array}{ccc|c} * & \dots & * & * \\ * & \dots & * & ? \\ 0 & \dots & 0 & ? \\ 0 & \dots & 0 & ? \\ 0 & & & ? \end{array} \right]$$

Tutti $? = 0 \Rightarrow$ soluzioni
Se $3? \neq 0 \Rightarrow$ sist. IMPOSSIBILE

ESERCIZI CALCOLO LU CON PIVOTING:

ESERCIZIO

Calcolo LU con pivoting parziale:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{R_1=R_3} \begin{bmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix} \xrightarrow{R_2=R_3} \begin{bmatrix} 7 & 8 & 0 \\ 0 & 3/4 & 6 \\ 0 & 6/4 & 3 \end{bmatrix} \xrightarrow{R_3=3/4 R_3} \begin{bmatrix} 7 & 8 & 0 \\ 0 & 6/4 & 3 \\ 0 & 3/4 & 6 \end{bmatrix}$$

$$\xrightarrow{\begin{bmatrix} 7 & 8 & 0 \\ 0 & 6/4 & 3 \\ 0 & 0 & 3/4 \end{bmatrix} =: U} L = \begin{bmatrix} 1 & 0 & 0 \\ 4/7 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix}$$

Le righe della sottosistema diagonale HANNO modulo < 1

$$\pi_R \cdot A = L \cdot U \quad \pi_R = ?$$

PERMUTAZIONI PER RIGHE:

$$\begin{array}{ccc} \sigma_1 & \sigma_2 \\ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} & \rightarrow & \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} & \rightarrow & \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \\ \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} & \rightarrow & \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \end{array} \Rightarrow \pi_R = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

OSS (SENZA PIVOTING) $A = LU \Rightarrow \det(A) = \det(U)$

CON PIVOTING $\det(\pi_R) \cdot \det(A) = \det(U)$

$S = \# \text{ scambi di righe } \Rightarrow (-1)^S \cdot \det(A) = \det(U)$

ESERCIZIO

Calcolare LU con pivoting totale

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -2 & 3 & 0 \end{bmatrix} \xrightarrow[\text{col.}]{\begin{array}{l} 1 \leftrightarrow 3 \\ 1 \leftrightarrow 2 \end{array}} \begin{bmatrix} 8 & 7 & 0 \\ 5 & 4 & 6 \\ 2 & 1 & 3 \end{bmatrix} \xrightarrow{\text{RIGHTS}} \begin{bmatrix} 8 & 7 & 0 \\ 0 & -3/8 & 6 \\ 0 & -3/4 & 3 \end{bmatrix} \xrightarrow{\begin{array}{l} 2 \leftrightarrow 3 \\ \downarrow \end{array}} \begin{bmatrix} 8 & 0 & 7 \\ 0 & 6 & -3/8 \\ 0 & 3 & -3/4 \end{bmatrix}$$

PER PORTARE
L'ELEM. DI
MODULO MAX
IN POS. a_{11}

PER PORTARE
L'ELEM. DI
MODULO MAX
DESA SOUDAT.
IN POS. a_{22}

$$\begin{bmatrix} 8 & 0 & 7 \\ 0 & 6 & -3/8 \\ 0 & 0 & -9/16 \end{bmatrix} =: U, \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 5/8 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix}$$

Lo Gu elem. sotto la diagonale
HANNO MODULO < 1

$$\pi_R \cdot \pi_C = L \cdot U$$

PERM. PER ROWS

$$1 \rightarrow 3 \quad 2 \rightarrow 2 \quad 3 \rightarrow 1 \Rightarrow \pi_R = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

PERM. PER COLUMNS

$$1 \rightarrow 2 \rightarrow 2 \quad 2 \rightarrow 1 \rightarrow 3 \rightarrow 1 \Rightarrow \pi_C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\det(\pi_R) \cdot \det(A) \cdot \det(\pi_C) = \det(U)$$

$$(-1)^s \cdot \det(A) \cdot (-1)^t = \det(U)$$

$s = \# \text{ scambi righe}$
 $t = \# \text{ scambi colonne}$

METODO DI GAUSS - JORDAN

$$D = \begin{bmatrix} * & & \\ & \ddots & \\ & & *$$

IDEA: rendere il sistema diagonale ($Ax = b \Leftrightarrow Dx = c$) con trasf. analoghe a quelle del metodo di Gauss

ASSUMPTO
 $Pivot \neq 0$

$$\begin{bmatrix} * & \dots & * \\ \vdots & \ddots & \vdots \\ * & \dots & * \end{bmatrix} \rightarrow \begin{bmatrix} * & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix} \rightarrow \begin{bmatrix} * & 0 & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & \dots & * \end{bmatrix} \begin{bmatrix} * \\ * \\ \vdots \\ * \end{bmatrix}$$

Dopo n-1 passi

$$\begin{bmatrix} *, & 0 \\ 0, & * \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

$$x_j = \frac{c_j}{a_{jj}^{(n-1)}}$$

COSTO: $O(n^3)$ ma la costante davanti è più di quella dell'algoritmo di Gauss

Im più non consente di scrivere una trasf. del tipo LU

OSS Anche qui si possono applicare scambi di righe / colonne

Gauss - Jordan è utile se si vuole calcolare A^{-1} :

Si procede scrivendo $[A | I] \rightarrow \dots \rightarrow [D | B]$

$$\left[\begin{array}{c|ccc} *, & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ *, & * & \dots & * \end{array} \right] \rightarrow \left[\begin{array}{c|cc} I & A^{-1} \\ \hline D & B \end{array} \right]$$

ESEMPIO

$$A = \left[\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|ccc} 0 & 3 & 1 & 1 & 0 & 0 \\ 0 & 3/2 & 1/2 & -1/2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 2 & -2 & 0 \\ 0 & 3/2 & 1/2 & -1/2 & 1 & 0 \\ 0 & 0 & 1/3 & 2/3 & -4/3 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 0 & -2 & 0 \\ 0 & 3/2 & 0 & -3/2 & 3 & -3/2 \\ 0 & 0 & 1/3 & 2/3 & -4/3 & 1 \end{array} \right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & 2 & -4 & 3 \end{array} \right]$$

$\underbrace{\qquad\qquad}_{A^{-1}}$

CONDIZIONAMENTO DI UN SISTEMA LINEARE $Ax = b$

Supponiamo di volere risolvere $Ax = b$, ma a causa di errori di misurazione abbiamo a disposizione $(A + \delta A) \in \mathbb{C}^{n \times n}$, $(b + \delta b) \in \mathbb{C}^n$

δA relativam. piccolo rispetto ad $A \Rightarrow \frac{\|\delta A\|}{\|A\|} \leq \text{TOL}$

$$\delta b \quad \overset{n}{\sim} \quad - \quad \overset{n}{\sim} \quad b \Rightarrow \frac{\|\delta b\|}{\|b\|} \leq \text{TOL}$$

$\text{ES. TOL} = 10^{-8}$

Se matrice $\|\cdot\|$ matriciale si considera compatibile con la matrice $\|\cdot\|$ vettoriale.

Il sistema perturbato $(A + \delta A)(x + \delta x) = b + \delta b$ che x verifica $Ax = b$

Possò dire che $\frac{\|\delta x\|}{\|x\|}$ è piccolo? (idealm. \leq costante · tol)

In generale no perché cost. può essere molto grande

ESEMPI

1) $\Delta A = 0$

$$\underbrace{\begin{bmatrix} 1 & -1 \\ 1 & -1.000001 \end{bmatrix}}_A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

A

$$\begin{bmatrix} 1 & -1 \\ 1 & -1.000001 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.999999 \\ 1 \end{bmatrix} \Rightarrow x = \begin{bmatrix} -10^6 \\ -1 \end{bmatrix} \quad \|\delta x\|_2 \approx \sqrt{2}$$

2) $\Delta A \neq 0$

$$\begin{bmatrix} 3 & 4 \\ 3 & 4.00001 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 7.00001 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 4 \\ 3 & 3.99999 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 7.00004 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 22/3 \\ -4 \end{bmatrix}$$

OSS

Nel caso $\Delta A = 0$, le componenti x_i sono funzioni polinomiali delle componenti del termine noto b .
Per vederlo si considera la formula di Cramer:

$$x_i = \frac{\det(A|b)}{\det(A)}$$

Supponiamo $\Delta A = 0$, $\det(A \neq 0)$, $Ax = b$

$$A(x + \delta x) = b + \delta b \Rightarrow A\delta x = \delta b \Rightarrow \delta x = A^{-1}\delta b$$

$$\Rightarrow \|\delta x\| \leq \|A^{-1}\| \cdot \|\delta b\|$$

Noi vorremmo dire qualcosa su $\frac{\|\delta x\|}{\|x\|}$

$$b = Ax \Rightarrow \|b\| \leq \|A\| \cdot \|x\| \Rightarrow \|x\| \geq \frac{\|b\|}{\|A\|}$$

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\|A^{-1}\| \cdot \|\delta b\|}{\left(\frac{\|b\|}{\|A\|}\right)} = \underbrace{\|A\| \cdot \|A^{-1}\|}_{\text{COSTANTE CHE DIPENDE DAL PROBLEMA}} \cdot \frac{\|\delta b\|}{\|b\|}$$

ERRORE RELATIVO SULLA SOLUZIONE

COSTANTE CHE DIPENDE DAL PROBLEMA

ERRORE RELATIVO SUL TERMINE NOTO

(INDICA DI QUANTO PUÒ ESSERE AML. L'ERRORE)

DEF $\mu(A) := \|A\| \cdot \|A^{-1}\|$ NUMERO DI CONDIZIONAMENTO
(DIPENDE DALLA NORMA) $\mu_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2$

LEMMA $\mu(A) \geq 1$

DIM $\|A\| \cdot \|A^{-1}\| \geq \|A \cdot A^{-1}\| = \|I\| \stackrel{\uparrow}{\geq} \rho(I) = 1$

PER HIRSH

TH Se $A, A + SA$ sono non singolari (invertibili)
 $\Rightarrow \frac{\|\delta x\|}{\|x\|} \leq \frac{\mu(A)}{1 - \mu(A)} \cdot \frac{\|SA\|}{\|A\|} \cdot \left(\frac{\|SA\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right)$

OSS Nel caso $SA = 0$ si ottiene $\frac{\|\delta x\|}{\|x\|} \leq \mu(A) \frac{\|\delta b\|}{\|b\|}$

OSS Per calcolare o stimare il numero di condizionamento ci sono metodi che non richiedono di calcolare esplicitamente A^{-1}

ESEMPIO Se A è unitaria ($A^*A = I$) $A^{-1} = A^*$ si ha $\mu(A) = 1$

Se A è hermitiana: $\|A\|_2 = \rho(A)$, A^{-1} è ancora hermitiana $\Rightarrow \|A^{-1}\|_2 = \rho(A^{-1})$

$$\mu(A) = \max_{1 \leq j \leq n} |\lambda_j| \cdot \max_{1 \leq j \leq n} |\bar{\lambda}_j| = \frac{\max_{1 \leq j \leq n} |\lambda_j|}{\min_{1 \leq j \leq n} |\lambda_j|}$$

λ_j AUTONOMI DI A

OSS L'errore su A e su b può essere causato dagli errori di arrotondamento generati dall'algoritmo risolutivo (es. Gauss).

STIMA DEI' ERRORE A POSTERIORI

Supponiamo di aver calcolato \tilde{x} approssimazione della soluzione di $Ax = b$

$$b - A\tilde{x} = r \in \mathbb{C}^n \setminus \{0\} \quad \text{Sono interessati a } \| \tilde{x} - x \|$$

Se sottraggo dall'eq. sopra $b - Ax = 0$ ottengo
 $A(\tilde{x} - x) = -r \Rightarrow \tilde{x} - x = -A^{-1}r \Rightarrow \| \tilde{x} - x \| \leq \| A^{-1} \| \cdot \| r \|$

$$\frac{\| \tilde{x} - x \|}{\| x \|} \leq \frac{\| A^{-1} \| \cdot \| r \| \cdot \| A \|}{\| b \|} = \mu(A) \cdot \frac{\| r \|}{\| b \|}$$

$$\| x \| \geq \frac{\| b \|}{\| A \|}$$

CONCLUSIONE: Anche se il residuo è piccolo non è detto che la distanza dalla sol. vera sia piccola

RAZIONAMENTO EUSTICO: se $\mu(A) \approx 10^\alpha$ $\alpha \in \mathbb{N}$
 si perdono α cifre di accuratezza nella soluzione

DEF Se $\mu(A) \gg 1$ il problema si dice mal condizionato.

Se $\mu(A) = 1$ situazione perfetta

Se $\mu(A)$ è sotto controllo (def. che dipende dal problema) \Rightarrow problema ben condizionato

METODI ITERATIVI

Generalmente una successione $\{x^{(k)}\}_{k=0,1,2,\dots}$ con $x^{(k)} \in \mathbb{C}^n$ è data (può essere un'approssimazione non buona di x).

$\lim_{K \rightarrow +\infty} x^{(k)} = x$ inoltre $x^{(k+1)}$ si ottiene da $x^{(k)} (x^{(k+1)} = f(x^{(k)})$) in modo "efficiente".

Ad es. $x^{(k+1)}$ si ottiene effettuando prodotti matrice - vettore con A (o matrici derivate) ed il vettore $x^{(k)}$.

L'idea è che se mi bastano k iterazioni con $k < n$ allora il costo di calcolare $x^{(k)}$ è all'incirca k volte il costo di matrice - vettore → in generale $O(kn^2)$

OSS In molte applicazioni calcolare $A \cdot v$ è meno costoso di $O(n^2)$. Ad esempio se A ha molti zeri (MATRICE SPARSA).

Molto spesso nnz(A) = $O(n)$.

↳ NUM. ELEMENTI NON ZERO DI A

OSS $y = Av$ $y_j = \sum_{i=1}^n a_{ji} v_i \rightarrow$ mi basta considerare gli elementi non zero della riga j di A

Se A è sparso $A \cdot v$ costa $O(\text{nnz}(A)) = O(n)$

A è Densa se $\text{nnz}(A) = O(n^2)$

CONCLUSIONE: A sparso \Rightarrow calcolare $x^{(k)}$ costa $O(k \cdot \text{nnz}(A))$

ESEMPIO

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

$$\text{nnz}(A) = 3n - 2$$

OSS Se applico Gauss su una matrice sparsa c'è il rischio di avere a che fare con matrici dense nei passi intermedi.

ESEMPIO

$$A = \begin{bmatrix} * & \cdots & * \\ : & * & \cdots & 0 \\ \vdots & \ddots & \ddots & * \\ * & 0 & \cdots & * \end{bmatrix} \xrightarrow{\substack{\text{DOPO IL 1° PASSO} \\ \text{DI GAUSS}}} \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ : & : & \ddots & : \\ 0 & 0 & \cdots & * \end{bmatrix} \quad (\text{FILL-IN})$$

↑
CI SONO TECNICHE
DI PIVOTING PER
EVITARE FILL-IN
MA NON FUNZIONANO
SEMPRE

IDEA DI PUNTO FISSO

$$Ax - b = 0 \quad \text{lo voglio risolvere come } x = Hx + c$$

MATRICE \uparrow VETTORE \uparrow

$$x = x - G(Ax - b) \quad \text{per una matrice } G$$

$$x = (I - GA)x + Gb \Rightarrow H = I - GA = \text{MATRICE DI ITERAZIONE}$$

$$c = Gh$$

Come dovrei scegliere G ?

Se G perfetta (ma non utilizzabile sarebbe $G = A^{-1}$)

Ma calcolare $G = A^{-1}$ è troppo costoso, quindi in genere si cerca $G \approx A^{-1}$ e G facilmente costruibile.

Una volta costituita G , il metodo iterativo è:

$x^{(0)}$ dato

$$x^{(k+1)} = Hx^{(k)} + c$$

Quando converge G ?

TH Il metodo iterativo converge quando la matrice H converge ($\lim_{K \rightarrow +\infty} H^K = 0$)

DIH Se x è sol. esatta $x = Hx + c$

$$\underbrace{x^{(k+1)} - x}_{e^{(k+1)}} = Hx^{(k)} + c - Hx - c = H(\underbrace{x^{(k)} - x}_{e^{(k)}})$$

ERRORE AL
PASSO $k+1$

$$e^{(k+1)} = He^{(k)} = H e^{(k)} \xrightarrow[K \rightarrow +\infty]{} 0 \in \mathbb{C}^n$$

OSS $\rho(H) < 1$ è condizione necessaria e sufficiente affinché il metodo converga

OSS $\|H\| < 1$ è condizione sufficiente affinché il metodo converga

OSS $|\det(H)| < 1$ è condizione necessaria, ma non sufficiente affinché il metodo converga

ESEMPIO

$$H = \frac{1}{20} \begin{bmatrix} 2 & -1 & 4 \\ 3 & 1 & 7 \\ -10 & 2 & 3 \end{bmatrix} \rightsquigarrow \|H\|_\infty = \frac{3}{4} < 1 \Rightarrow \text{IL METODO È CONVERGENTE}$$

OSS $\forall \|\cdot\|$ norma matriciale (indotta) si ha

$$\lim_{K \rightarrow +\infty} \sqrt[K]{\|A^K\|} = \rho(A)$$

Quindi per k abbastanza grande $\varphi(A)^k \approx \|A^k\|_2$

In particolare se $A = A^*$ e $\|\cdot\|_2 = \|\cdot\| \Rightarrow \varphi(A)^k = \|A^k\|_2$

$$e^{(k+1)} = H^{k+1} e^{(0)} \Rightarrow \frac{\|e^{(k)}\|}{\|e^{(0)}\|} \leq \|H^{k+1}\| \approx \varphi(H)^{k+1}$$

L'errore diminuisce all'incirca di un fattore $\varphi(H)$ ad ogni iterazione.

ESEMPIO

Quanto deve essere k (quante iterazioni devo fare) affinché $\frac{\|e^{(k)}\|}{\|e^{(0)}\|} \leq 10^{-m}$

$$\varphi(H)^k \leq 10^{-m} \Leftrightarrow n \underbrace{\log_{10} \varphi(H)}_{\leq 0} \leq -m \Rightarrow k \geq -\frac{m}{\log_{10}(\varphi(H))}$$

CRITERI DI ARRESTO

1) Fissare il num. di iterazioni magari usando una stima a priori del k necessario a ridurre la distanza da x quanto vogliamo (come fatto sopra)

2) Misurare il residuo al passo k : $r^{(k)} = Ax^{(k)} - b$ e ci si ferma quando $\frac{\|r^{(k)}\|}{\|b\|} \leq \underline{\text{TOL}}$ ($\text{TOL} = 10^{-8}$)

$\|b\| = \|Ax^{(0)} - b\|$ con $x^{(0)} = 0$ (in alcuni casi può essere costoso perché c'è un prodotto mat-vec in più ad ogni iterazione)

3) Ci si ferma quando $\|x^{(k+1)} - x^{(k)}\| \leq \text{TOL}$

MOTIVAZIONE PIÙ RIGUARDA:

$$x^{(k+1)} = Hx^{(k)} + c$$

$$x = Hx + c \Rightarrow c = -(H - I)x$$

$\rho(H) < 1$ per ipotesi
 $\Rightarrow I - H$ è invertibile

$$x^{(k+1)} - x^{(k)} = Hx^{(k)} - x^{(k)} - (H - I)x = (H - I)(x^{(k)} - x)$$

$x^{(k)} - x$

$$\|e^{(k)}\| \leq \|(H - I)^{-1}\| \cdot \|x^{(k+1)} - x^{(k)}\|$$

Questa relazione ci dice che la qta $\|x^{(k+1)} - x^{(k)}\|$ è collegata direttamente a $\|e^{(k)}\|$.

In particolare se $\|(H - I)^{-1}\|$ non è gigantesca allora $\|x^{(k+1)} - x^{(k)}\| \approx \|e^{(k)}\|$

$$H = I - GA$$

$H - I = -GA \rightsquigarrow$ se G è buona appross.

$$\text{di } A^{-1} \Rightarrow \|-GA\| \approx 1$$

$$\|(-GA)^{-1}\| \approx 1$$

METODI ITERATIVI CLASSICI

METODO DI JACOBI:

G = diagonale di A , elevata alla -1

$$A = D - E - F \quad D = \begin{bmatrix} a_{11} & & \\ & \ddots & 0 \\ 0 & \cdots & a_{nn} \end{bmatrix}, \quad -E = \begin{bmatrix} 0 & & \\ & \ddots & 0 \\ -a_{11} & \cdots & 0 \end{bmatrix}, \quad -F = \begin{bmatrix} 0 & & \\ & \ddots & 0 \\ 0 & \cdots & 0 \end{bmatrix}$$

$$Ax = b \quad (D - E - F)x = b \Rightarrow Dx = (E + F)x + b \Leftrightarrow$$

$$x = D^{-1}(E + F)x + D^{-1}b$$

JACOBI

$$\left\{ \begin{array}{l} x^{(0)} \\ x^{(k+1)} = D^{-1}(E + F)x^{(k)} + D^{-1}b \end{array} \right.$$

$$H_S = \begin{bmatrix} \frac{1}{a_{11}} & & & \\ & \ddots & & \\ & & \frac{1}{a_{nn}} & \\ & & & \end{bmatrix} \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{11}} & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & -\frac{a_{n-1,n}}{a_{11}} \\ -\frac{a_{n1}}{a_{11}} & -\frac{a_{n2}}{a_{11}} & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\frac{a_{n-1,n}}{a_{11}} \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$(H_S)_{ij} = \begin{cases} -\frac{a_{ij}}{a_{ii}} & \text{se } i \neq j \\ 0 & \text{se } i = j \end{cases}$$

Le equazioni che descrivono il metodo di Jacobi sono:

$$x_j^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{i=1 \\ i \neq j}}^n a_{ij} x_j^{(k)} \right) \quad i = 1, \dots, n \\ k = 0, 1, 2, \dots$$

Oss Prima di poter scomporre (cancellare) $x^{(k)}$ è necessario aver cancellato tutto il vettore $x^{(k+1)}$

Quando converge il metodo di Jacobi?

In generale bisogna vedere se $\rho(H_S) = \rho(D^{-1}(E+F)) < 1$
Però ci sono dei casi in cui si può dire a priori.

Th Se A è a predom. diag.-forte \Rightarrow JACOBI converge

DIM

$$D^{-1}(E+F) = \begin{cases} -\frac{a_{ij}}{a_{ii}} & \text{se } i \neq j \\ 0 & \text{se } i = j \end{cases} \quad \begin{array}{l} \text{I CERCHI DI GERSHGORIN} \\ \Rightarrow \text{HANNO TUTTI CENTRO 0} \end{array}$$

Il raggio associato alla riga i :

$$\rho_i = \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| = \frac{1}{|a_{ii}|} \cdot \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < 1 \Rightarrow \rho(H_S) < 1$$

$< |a_{ii}|$ PER DEF. DI NAT. A PREDOM. DIAG. FORTE

TH Se A è a predom. diag. debole e irriducibile

\Rightarrow Jacobi converge

DIM

Se si ripete la stessa dim. sopra si ottiene che i maggiori $\rho_i \leq 1$ e ne esiste almeno uno t.c. $\rho_i < 1$

\Rightarrow per il III TH di Gershgorin non possiamo avere autovalori di modulo 1

N.B.: A irriducibile $\Rightarrow A_S$ irriducibile in quanto H_S ha lo stesso grafo orientato di A

METODO DI GAUSS-SEIDEL:

$$G = (D - E)^{-1}$$

$$A = D - E - F$$

$$Ax = b \Leftrightarrow (D - E - F)x = b \Leftrightarrow (D - E)x = Fx + b \Leftrightarrow x = (D - E)^{-1}Fx + (D - E)^{-1}b$$

$$\text{GAUSS-SEIDEL} \left\{ \begin{array}{l} x^{(0)} \\ x^{(k+1)} = (D - E)^{-1}Fx^{(k)} + (D - E)^{-1}b \end{array} \right.$$

OSS Applicare $(D - E)^{-1}$ ad un vettore si fa con il metodo di sost. all'indietro $O(n^2)$, quindi non si forma $(D - E)^{-1}$ esplicitamente

$$x^{(k+1)} = (D - E)^{-1}Fx^{(k)} + (D - E)^{-1}b \Leftrightarrow (D - E)x^{(k+1)} = Fx^{(k)} + b$$

$$\Leftrightarrow x^{(k+1)} = D^{-1}Ex^{(k+1)} + D^{-1}Fx^{(k)} + D^{-1}b$$

$x^{(k+1)}$

DIRENDEREBBE DA SÉ STESSA

$$D^{-1}E = \begin{bmatrix} 0 & \dots & 0 \\ * & \ddots & \vdots \\ \vdots & \ddots & * \\ * & \dots & * & 0 \end{bmatrix}$$

$$x^{(k+1)} = \begin{bmatrix} ① \\ ② \\ ③ \\ \vdots \\ n \end{bmatrix}$$

- ① non dipende da altri elem. di $x^{(k+1)}$
- ② dipende solo da $x_1^{(k+1)}$
- ③ dipende da $x_1^{(k+1)}$ e $x_2^{(k+1)}$

Le equazioni di Gauss-Siedel sono:

$$x_j^{(k+1)} = \frac{1}{a_{jj}} (b_j - \sum_{i=1}^{j-1} a_{ij} x_i^{(k+1)} - \sum_{i=j+1}^n a_{ij} x_i^{(k)})$$

OSS A differenza di Jacobi si possono sovrascrivere le entrate $x_j^{(k)}$ una volta che calcolo $x_j^{(k+1)}$

METODO DI GAUSS-SIEDEL

$$Ax = b \quad A \in \mathbb{C}^{n \times n}$$

$$A = D - E - F \quad H = (D - E)^{-1} F$$

$$\begin{cases} x^{(0)} \\ x^{(k+1)} = (D - E)^{-1} F x^{(k)} + (D - E)^{-1} b \end{cases}$$

$$(D - E)x^{(k+1)} = Fx^{(k)} + b \iff Dx^{(k+1)} = Ex^{(k+1)} + Fx^{(k)} + b \iff x^{(k+1)} = D^{-1}Ex^{(k+1)} + D^{-1}Fx^{(k)} + D^{-1}b$$

$$x_j^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right)$$

OSS Entrambi Jacobi e Gauss-Seidel hanno bisogno della condizione $a_{jj} \neq 0 \quad \forall j = 1, \dots, n$

Se $a_{jj} = 0$ per qualche j si può provare ad applicare Jacobi o Gauss-Seidel al sistema

$$(\Pi_R A \Pi_C)(\Pi_C^\top x) = \Pi_R b$$

TH Se A è a predom. diag. forte, oppure se A è a pred. diag. debole ma è anche irriducibile, allora il metodo di Gauss-Seidel è convergente ($\rho(H_{GS}) < 1$).

DIM

Osserviamo che le assunzioni implicano $a_{jj} \neq 0 \quad \forall j$. Supponiamo per assurdo che $\exists \lambda$ autovalore di modulo 1.

$$\begin{aligned} 0 &= \det(H_{GS} - \lambda I) = \det((D - E)^{-1} F - \lambda I) \\ &= \det((D - E)^{-1} [F - \lambda(D - E)]) = \\ &= \underbrace{\det(D - E)^{-1}}_{\neq 0} \cdot \det(F - \lambda(D - E)) \end{aligned}$$

$\neq 0$ perché è il det di una mat. triang. inferiore
 \Rightarrow tutti gli elem. diag. sono $\neq 0$

$$\Rightarrow 0 = \det(F - \lambda(D - E)) = \det(\lambda[-D + E + \lambda^{-1}F]) =$$

$$= \underbrace{\lambda^n}_{\neq 0} \cdot \det(\lambda^{-1}F - (D - E)) \Rightarrow \det(D - E - \lambda^{-1}F) = 0$$

$$D - E - \lambda^{-1}F = \begin{bmatrix} a_{11} & \frac{a_{12}}{\lambda} & \cdots & \frac{a_{1n}}{\lambda} \\ \vdots & \ddots & & \vdots \\ & & \ddots & \frac{a_{n-1n}}{\lambda} \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix}$$

$$\left| \frac{a_{ii}}{\lambda} \right| = \frac{|a_{ii}|}{|\lambda|} = |a_{ii}|$$

↓

I CERCHI DI GERSHGORIN
DI $D - E - \lambda^{-1}F$ SONO GLI
STESSI DI A

\Rightarrow PER GERSH. $D - E - \lambda^{-1}F$ NON PUÒ AVERE AUTOVALORI NULLI

OSS In generale non è vero che entrambi i metodi sono o convergenti o non convergono. Oltre alle mat. a pred. diag., c'è un'altra eccezione in cui i due metodi fanno la stessa cosa: le mat. TRIDIAGONALI.

TH Se A è tridiagonale ($A = \begin{bmatrix} a_{11} & c_1 & & & 0 \\ b_1 & \ddots & \ddots & & c_{n-1} \\ & \ddots & \ddots & \ddots & \\ 0 & \cdots & b_{n-1} & a_n \end{bmatrix}$)

Allora Jacobi e Gauss-Seidel convergono o divergono insieme.

$$\text{In particolare } \rho(H_{GS}) = \rho(H_J)^2$$

ESEMPIO: $\rho(H_J) = 0.1 \Rightarrow \rho(H_{GS}) = 0.01$

Se convergono, Gauss-Seidel va "il doppio" più veloce.

$$\frac{\|x^{(k)} - x^{(0)}\|}{\|x^{(0)} - x\|} = \frac{\|e^{(k)}\|}{\|e^{(0)}\|} \approx \rho(H)$$

ESEMPIO

$$A = \begin{bmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{bmatrix} \quad \alpha \in \mathbb{C}$$

Determinare i valori di α per cui Jacobi converge.

$$H_J = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & -\alpha & -\alpha \\ -\alpha & 0 & -\alpha \\ -\alpha & -\alpha & 0 \end{bmatrix} = \begin{bmatrix} 0 & \alpha & \alpha \\ \alpha & 0 & \alpha \\ \alpha & \alpha & 0 \end{bmatrix}$$

Quando $\rho\left(\begin{bmatrix} 0 & \alpha & \alpha \\ \alpha & 0 & \alpha \\ \alpha & \alpha & 0 \end{bmatrix}\right) < 1$?

$$\begin{bmatrix} 0 & \alpha & \alpha \\ \alpha & 0 & \alpha \\ \alpha & \alpha & 0 \end{bmatrix} = \begin{bmatrix} \alpha & \alpha & \alpha \\ \alpha & \alpha & \alpha \\ \alpha & \alpha & \alpha \end{bmatrix} - \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{bmatrix} = \alpha \left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} - I \right)$$

\rightarrow test con $\text{rk} = 1$

Se $\tilde{\lambda}$ sono gli autovalori di $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ allora $\alpha(\tilde{\lambda} - 1)$

sono quelli di $\begin{bmatrix} 0 & \alpha & \alpha \\ \alpha & 0 & \alpha \\ \alpha & \alpha & 0 \end{bmatrix}$ ($A = U V^T \Rightarrow \text{rk}(A) = 1$)

$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ \nwarrow $n-1$ (2) autovalori = 0
 1 autovalore $\neq 0$ calcolabile come traccia della matrice

$$\tilde{\lambda} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \lambda \begin{pmatrix} 2\alpha \\ -\alpha \\ -\alpha \end{pmatrix} \quad \begin{cases} |2\alpha| < 1 \\ |\alpha| < 1 \end{cases} \Rightarrow |\alpha| < \frac{1}{2}$$

Esercizio

$$A = \begin{bmatrix} 2 & 0 & 0 & \alpha \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Per quali $\alpha \in \mathbb{C}$ Gauss-Seidel converge?

$$H_{GS} = \begin{bmatrix} 2 & & & \\ 1 & 2 & & \\ & 1 & 2 & \\ & & 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} \alpha \\ & & & \\ & & & \alpha \\ & & & \\ & & & \alpha \end{bmatrix} =$$

$$= \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ & * & * & * \\ & & * & * \end{bmatrix}$$

$$\begin{bmatrix} * \\ \vdots \\ * \end{bmatrix} \text{ è la soluzione di } \begin{bmatrix} 2 & & & \\ 1 & 2 & & \\ & 1 & 2 & \\ & & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \\ z \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & & & \\ 1 & 2 & & \\ & 1 & 2 & \\ & & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \\ z \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{bmatrix} \rightsquigarrow \begin{cases} x = \alpha/2 \\ y = -\alpha/4 \\ w = \alpha/8 \\ z = -\alpha/16 \end{cases}$$

$$H_{GS} = \begin{bmatrix} & x \\ 0 & y \\ & w \\ & z \end{bmatrix} \rightsquigarrow \text{rk}(H_{GS}) = 1$$

l'unico autovalore $\neq 0$ è la traccia, ovvero $-\alpha/16$

$$\begin{bmatrix} x \\ y \\ w \\ z \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\rho(H_{GS}) < 1 \iff |\frac{\alpha}{16}| < 1 \Rightarrow |\alpha| < 16$$

Esercizio

$$A = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & 1 & \alpha \\ \alpha & 0 & 1 \end{bmatrix}$$

Per quali $\alpha \in \mathbb{C}$ i vettori GS convergono?

$$H_S = \begin{bmatrix} 0 & \alpha & 0 \\ 0 & 0 & \alpha \\ \alpha & 0 & 0 \end{bmatrix} = \alpha \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}}_{\text{MAT. DI PERM.}}$$

(a meno di segno, tutto non cambia)

$$\rho(H_S)$$

IN PARTICOLARE
È UNA MATRICI ORTOGONALE

$$H^T H = I \Rightarrow \text{HA AUTOVALORI } \lambda : |\lambda| = 1$$

H_J ha autovalori al quindi di modulo α
 $\rho(H_J) < 1 \iff |\alpha| < 1$

$$H_{GS} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & \alpha & 0 \\ 0 & 0 & \alpha \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & -\alpha & 0 \\ 0 & 0 & -\alpha \\ 0 & \alpha^2 & 0 \end{bmatrix}$$

$$0 = \det(H_{GS} - \lambda I) = -\lambda(\lambda^2 + \alpha^3) \quad \begin{cases} \lambda = 0 \\ \lambda = \pm \sqrt[3]{\alpha^3} \end{cases}$$

$$\rho(H_{GS}) < 1 \quad (\alpha^{3/2} < 1 \iff |\alpha| < 1)$$

Dalla lezione precedente:

TH Se A è a pred. diag. forte oppure se è a pred. diag. debole e irriducibile \Rightarrow GAUSS-SEIDEL per risolvere $Ax = b$ è convergente (ovvero $\rho(H_{GS}) < 1$)

$$\text{DIM } H_{GS} = (D - E)^{-1} F$$

$$D = \begin{bmatrix} a_{11} & & \\ \ddots & \ddots & \\ & & a_{nn} \end{bmatrix}; \quad -E = \begin{bmatrix} 0 & & \\ a_{21} & \ddots & 0 \\ \vdots & \ddots & \ddots \\ 0 & \dots & 0 \end{bmatrix}; \quad -F = \begin{bmatrix} 0 & & \\ & \ddots & \\ & \vdots & \ddots \\ 0 & \dots & 0 \end{bmatrix}$$

Entrambe le assunzioni $\Rightarrow a_{jj} \neq 0 \forall j$

$$\Rightarrow D - E \text{ è invertibile: } \begin{bmatrix} a_{11} & & \\ \vdots & \ddots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = D - E$$

Dimostreremo che H_{GS} non ha autovalori

$$\lambda: |\lambda| \geq 1$$

$$\begin{aligned} \text{Se } \lambda \text{ è autovalore} \Rightarrow 0 &= \det(H_{GS} - \lambda I) = \\ &= \det((D - E)^{-1} F - \lambda I) = \\ &= \det(D - E)^{-1} \cdot \det(F - \lambda(D - E)) = \\ &= \det(D - E)^{-1} \cdot (-\lambda^n) \cdot \det(D - E + \lambda^{-1} F) \end{aligned}$$

Supponiamo per assurdo $|\lambda| \geq 1$

$$\Rightarrow \det(D - E - \lambda^{-1} F) = 0$$

$$B(\lambda) := D - E - \lambda^{-1} F = \begin{bmatrix} a_{11} & a_{12}/\lambda & \dots & a_{1n}/\lambda \\ \vdots & \ddots & \ddots & a_{n-1n}/\lambda \\ a_{n1} & a_{nn} & & \end{bmatrix}$$

Se $|\lambda| \geq 1$ e A è a pred. diag. forte

$\Rightarrow B(\lambda)$ è a pred. diag. forte

$$|\alpha_{jj}| > \sum_{\substack{i \neq j \\ i=1}}^n |\alpha_{ji}| \geq \sum_{i \neq j}^n \frac{|\alpha_{ji}|}{|\lambda_i|} \Rightarrow \det(B) \neq 0$$

PER GERSHGORIN I
↓
Assurdo ↴

Lo stesso vale se A è a pred. diag.-debole e irriducibile (dipende dal fatto che A e B(x) hanno le stesse entrate ≥ 0)



PER GERSHGORIN III $\det(B(x)) \neq 0 \Rightarrow$ Assurdo
↓

EQUAZIONI NON LINEARI

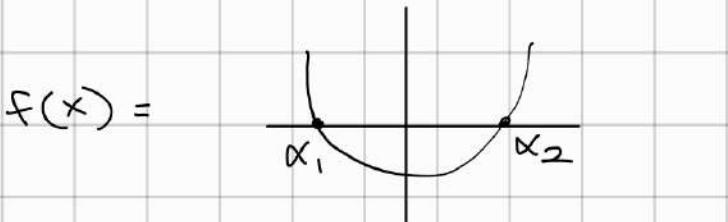
Sistemi lineari \Leftrightarrow cercare gli zeri di $f(x) = Ax - b$

Adesso vogliamo considerare $f(x) = 0$ con f non lineare.

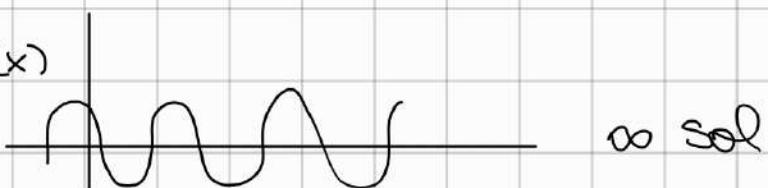
Partiamo con $f: \mathbb{R} \rightarrow \mathbb{R}$ $f \neq \alpha, x + \alpha_0$

ESEMPI

$$f(x) = x^2 + 2 \Rightarrow 0 \text{ sol.}$$



$$f(x) = \sin(x)$$



OSS In generale si cercano soluzioni vicine a un punto di partenza $x^{(0)}$ e si considerano metodi iterativi che generano una successione $\{x^{(n)}\}_{n \in \mathbb{N}}$ t.c. $\lim_{n \rightarrow +\infty} x^{(n)} = \alpha$

DEF f non lineare, $\alpha \in \mathbb{R}$: $f(\alpha) = 0$ è detta radice o zero di f

SITUAZIONE TPO: Abbiamo $x^{(0)}$ vicino ad α , oppure sappiamo che $x \in [a, b] \subset \mathbb{R}$ e vogliamo generare una successione $\{x^{(n)}\}$ che converga.

Per questi problemi non c'è un metodo diretto, si usano i metodi iterativi (per f. con polinomi di grado > 5 Galois ha dim. che non c'è un metodo diretto).

→ metodi che consideriamo prendono la forma:

$$\begin{cases} x^{(n+1)} = \phi_n(x^{(n)}, x^{(n-1)}, \dots, x^{(n-k+1)}) & n=0, 1, 2 \dots \\ x^{(0)} \text{ dato} & k \geq 1 \end{cases}$$

Nel caso di sist. lineari si aveva $\phi_n = \phi_{n+1}, \forall n \in K = 1$

DEF Quando ϕ_n è costante (non dipende da n) il metodo si dice **STAZIONARIO**

DEF Quando ϕ_n dipende da k valori noti allora si dice **METODO ITERATIVO A K PUNTI**

DEF $\{x^{(n)}\}_{n \in \mathbb{N}}$ successione che converge ad $\alpha \in \mathbb{R}$

(ovvero $\lim_{n \rightarrow +\infty} x^{(n)} = \alpha$) \Rightarrow si dice che converge:

1) **SUBLINEARMENTE** se $\lim_{n \rightarrow +\infty} \frac{|x^{(n)} - \alpha|}{|x^{(n-1)} - \alpha|} = 1$

(converge in maniera molto lenta)

2) **LINEARMENTE** se $\exists c \in (0,1)$: $\lim_{n \rightarrow +\infty} \frac{|x^{(n)} - \alpha|}{|x^{(n-1)} - \alpha|} = c$

(converge più velocemente)

3) **SUPERLINEARMENTE** se $\lim_{n \rightarrow +\infty} \frac{|x^{(n)} - \alpha|}{|x^{(n-1)} - \alpha|} = \infty$

Si parla di ordine $p \geq 1$ se $\exists c \neq 0, c > 0 : c < \infty$

$$\lim_{n \rightarrow +\infty} \frac{|x^{(n)} - \alpha|}{|x^{(n-1)} - \alpha|} = c$$

NEL CASO $p=2$
LE CIFRE SIGNIF.
DI α RADDOPP.
A OGNI PASSO

(converge molto velocemente)

DEF Se $p=2$ si parla di convergenza quadratica,
 $p=3$ cubica ecc.

N.B.: p potrebbe non essere un intero

TIROVARE DUE STIME SU DUE SI TROVANO LE RADICI:

Usare tecniche viste in analisi per dare stime
di intervalli dove si trova α

ESEMPIO

$$f(x) = x \log(x) - 1$$

(definita su $(0, +\infty)$)

Quante radici ha? Dove sono?

$$f'(x) = \log(x) + 1$$

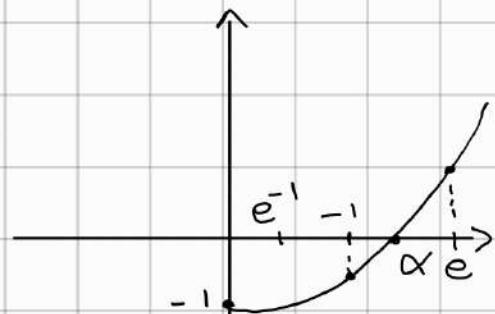
$$f' > 0 \quad x > e^{-1}$$

$$f' < 0 \quad x < e^{-1}$$

$$f(1) = -1$$

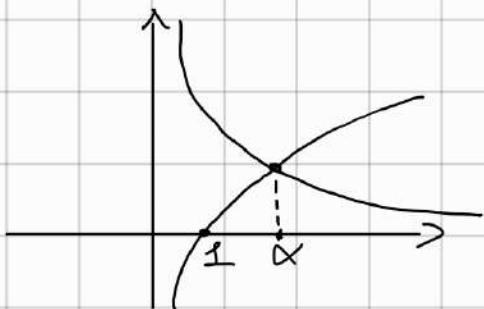
$$f(e) = e - 1 > 0$$

$$\Rightarrow \alpha \in [1, e]$$



TECNICA DI SEPARAZIONE GUIDATA

$$f(x) = 0 \iff x \log(x) - 1 = 0 \iff \log(x) = \frac{1}{x}$$



$$\begin{cases} y = \frac{1}{x} \\ y = \log(x) \end{cases}$$

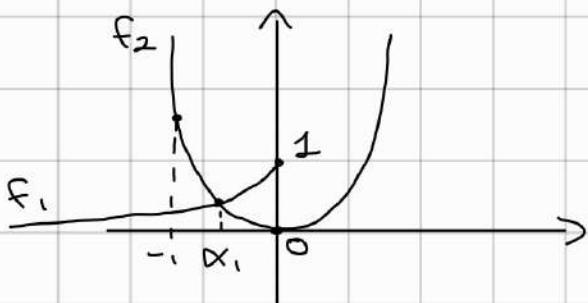
ESEMPIO

$$f(x) = 5x^2 - 2e^x$$

$$f(x) = 0 \iff f_1(x) = f_2(x)$$

$$f_1 = 5x^2$$

$$f_2 = 2e^x$$



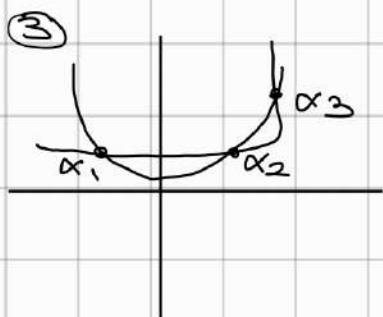
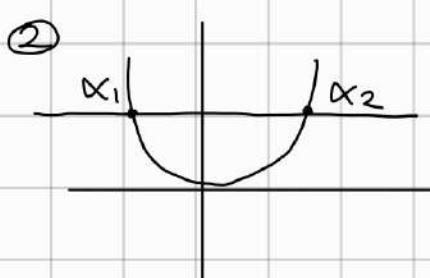
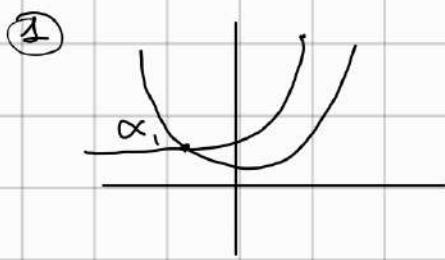
$$f_2(-\infty) < f_1(-\infty), f_1(0) = 0$$

$$f_2(0) = 2 \Rightarrow \alpha_1 < 0$$

$$f_2(-1) = \frac{2}{e} < 1, f_1(-1) = 5$$

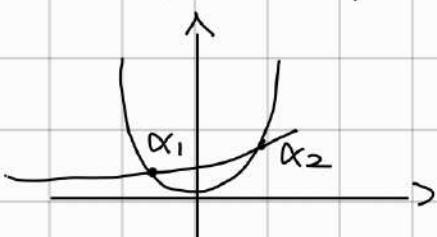
$$\downarrow \quad \alpha_1 \in (-1, 0)$$

A dx:



Dovete escludere
perché 2^x va a ∞
più veloce di x^2

$$f_1(2) = 20; f_2(2) = 2e^2 < 2 \cdot 3^2 = 18$$



$$\Rightarrow \alpha_2 \in (0, 2)$$

$$\downarrow \quad \alpha_3 \in (2, \infty)$$

non ce ne sono altre perché
in $(2, \infty)$ f_1, f_2 sono crescenti
e f_2 rimane sempre $> f_1$.

Possibile obiettivo di uno studio preliminare del problema:

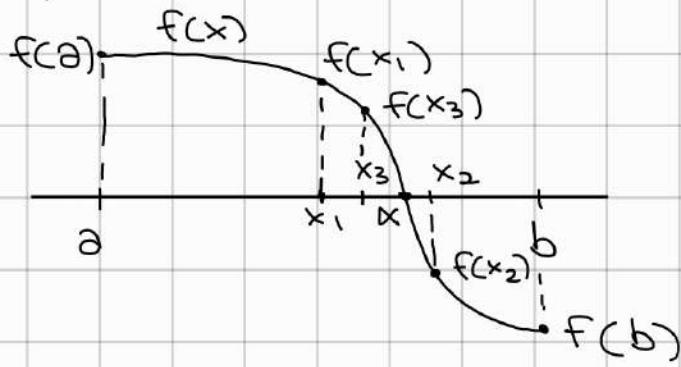
$\alpha \in [a, b] \subset \mathbb{R}$, $f(a) \cdot f(b) < 0$ con f continua



Assumiamo inoltre che α sia l' unica radice in $[a, b]$

METODO DI BISEZIONE (o METODA BINARIA)

IDEA: Utilizzare il punto medio dell'intervallo come prima approssimazione ed iterare il procedimento



$$x_1 = \frac{a+b}{2} = \text{PUNTO MEDIO}$$

Si interessa il segno di x_1
+ \Rightarrow radice alla sx
- \Rightarrow " " " " sx
IN QUESTO CASO È +

$\alpha \in [x_1, b]$ restringo l'intervallo

$$x_2 = \frac{x_1 + b}{2}, f(x_2) < 0$$

$\alpha \in [x_1, x_2]$

$$x_3 = \frac{x_1 + x_2}{2} \dots \text{ecc-} \dots \text{e restringiamo sempre di più l'intervallo}$$

$$x_0 = a, x_1 = b, x_2 = \frac{x_0 + x_1}{2}$$

$$x_3 = \frac{x_2 + \hat{x}_2}{2} \quad \hat{x}_2 = \begin{cases} x_0 & \text{se } f(x_2) f(x_0) < 0 \\ x_1 & \text{se } f(x_1) f(x_2) < 0 \end{cases}$$

Al passo n :

$$x_{n+1} = \frac{x_n + \hat{x}_n}{2} \quad \text{dove } \hat{x}_n = \begin{cases} x_n & \text{se } f(x_{n-1}) f(x_n) < 0 \\ \hat{x}_{n-1} & \text{altrimenti} \end{cases}$$

Oss La bisezione è un metodo a 2 punti ed è non stazionario (la funzione cambia con il passo)

Oss Ad ogni passo l'intervallo dove viene stimato essere α si dimezza, per cui:

$$|x_{n+1} - \alpha| \leq \frac{b-a}{2^n} \Rightarrow \text{IL METODO È CONVERGENTE}$$

cioè $\lim_{n \rightarrow +\infty} x_n = \alpha$

\downarrow
SI HA ERRORE ASSOLUTO

L'ordine di convergenza è lineare con coeff. $\frac{1}{2}$

CRITERIO DI ARRESTO: $\frac{b-a}{2^n} < \text{TOL}$ ad es. $\text{TOL} = 10^{-8}$

Esercizio

Quanti passi impiega la bisezione con questo criterio d'arresto (al massimo)?

$$\frac{b-a}{2^n} < \text{TOL} \Rightarrow \frac{b-a}{\text{TOL}} < 2^n \Rightarrow n > \log_2 \left(\frac{b-a}{\text{TOL}} \right)$$

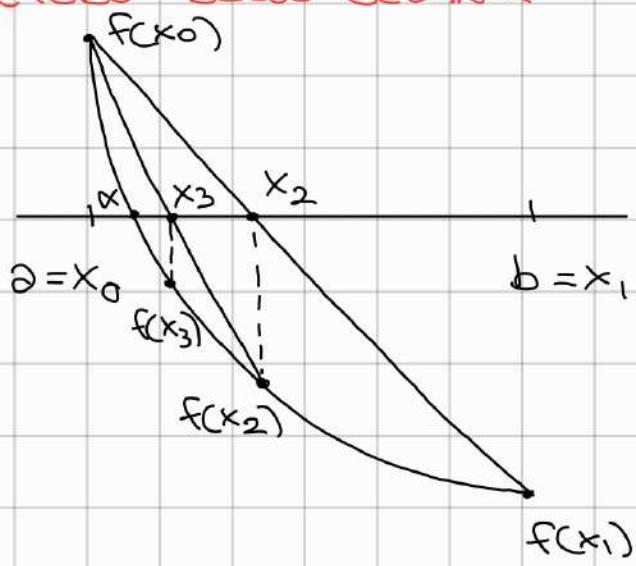
Il # di iterazioni sarà al max $\tilde{n} = \lceil \log_2 \left(\frac{b-a}{\text{TOL}} \right) \rceil$

OSS La bisezione ha bisogno di poche assunzioni:
solo la continuità e sapere il segno delle valutazioni di f .

D'altro canto converge lentamente e può essere problematico quando le valutazioni della f sono molto costose (sia in termini di tempo che denaro).

Per questo talvolta si preferiscono metodi a converg. superlineare.

METODO DELLE SECANTI



$$f(x_0) \cdot f(x_1) < 0$$

Una variante di questo metodo considera sempre l'intersezione della secante con l'asse delle x , ma i punti scelti sono sempre gli ultimi due (per fare la retta secante) indip. dal segno.

Per calcolare x_{n+1} si fa dunque la secante tra $(x_n, f(x_n))$ e $(x_{n-1}, f(x_{n-1}))$

$$\left\{ \begin{array}{l} x_0, x_1 \text{ dati} \\ x_{n+1} = x_n - f(x_n) \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \end{array} \right.$$

OSS

Il metodo delle secanti (variato[†]) è un metodo stazionario a due punti.

TH $f \in C^2([a, b])$, x_0 abbastanza vicino ad α , $f(\alpha) = 0$, $f'(\alpha) \neq 0$, allora il metodo delle secanti converge superlinearmente con ordine $p = \frac{1 + \sqrt{5}}{2} \approx 1.618 \dots$ (SEZIONE AUREA)

EQUAZIONI SCALARI NON LINEARI

$f(x) : \mathbb{R} \rightarrow \mathbb{R}$ Vogliamo risolvere $f(x) = 0$

METODI STAZIONARI A UN PASSO

$$\begin{cases} x_{n+1} = \phi(x_n) \\ x_0 \text{ DATO} \end{cases}$$

L'idea è trasformare $f(x) = 0 \Leftrightarrow x = \phi(x)$ in modo tale che α sia radice di $f \Leftrightarrow \alpha$ è punto fisso di ϕ

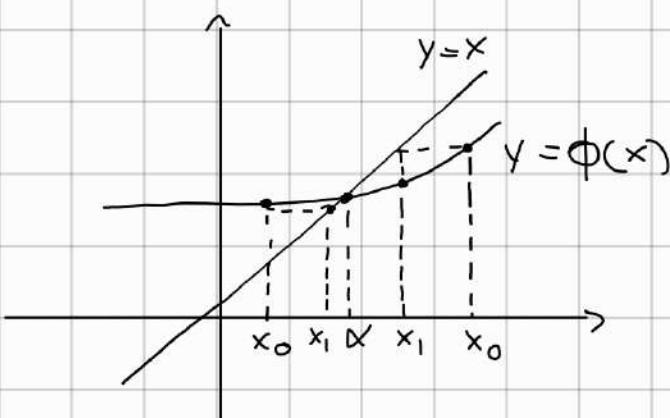
Un modo comune per costruire / scegliere $\phi(x)$ è considerare $\phi(x) = x - g(x) \cdot f(x)$ dove $g(x)$ è un'arbitraria funzione che non si annulla in un intorno di $(\alpha : f(\alpha) = 0)$

$$\phi(\alpha) = \alpha - \underbrace{g(\alpha)}_{\neq 0} \cdot \underbrace{f(\alpha)}_{\neq 0} = \alpha$$

OSS Il fatto che $g(x) \neq 0$ intorno ad α garantisce che non ci sono punti fissi di ϕ che non corrispondono a radici di f

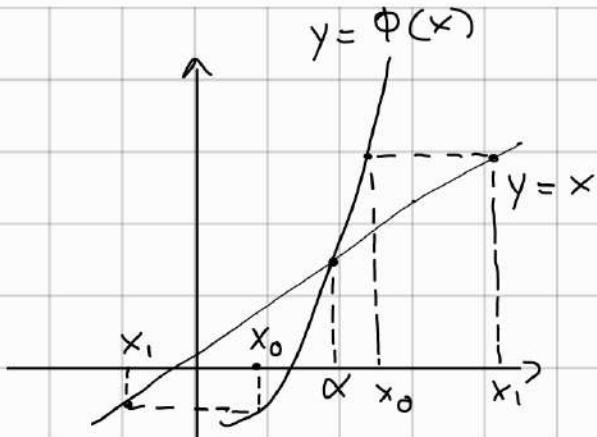
$$\begin{cases} x_{n+1} = \phi(x_n) \\ x_0 \end{cases}$$

METODO ASSOCIAZIONE A UNA ϕ CHE
VERIFICA $\phi(x) = x - g(x) \cdot f(x)$



$$\alpha = \phi(\alpha)$$

In questo caso la funzione sembra convergere ad α , ma non è sempre questo il caso.



In questo caso non funziona perché mi allontano da α . α sembra respingere la successione.

TH DI CONVERGENZA LOCALI

Sia $\alpha = \phi(\alpha)$, $\phi \in C^1(I)$, $\alpha \in I \subset \mathbb{R}$ e supponiamo che $\exists p \in \mathbb{R}^+$ e $K \in (0, 1)$ t.c.:

$$\forall x \in [\alpha - p, \alpha + p], |\phi'(x)| \leq K < 1$$

Allora:

- $$\begin{aligned} 1) \quad x_0 \in [\alpha - p, \alpha + p] &\Rightarrow x_n \in [\alpha - p, \alpha + p] \quad \forall n \\ 2) \quad \lim_{n \rightarrow +\infty} x_n &= \alpha \quad \text{se } x_0 \in [\alpha - p, \alpha + p] \end{aligned}$$

ed inoltre la converg. è almeno lineare ($\lim_{n \rightarrow +\infty} \frac{|x_n - \alpha|}{|x_{n-1} - \alpha|} = c$ con $0 \leq c < 1$)

- $$3) \quad \alpha \text{ è l'unico punto fisso di } \phi(x) \text{ in } [\alpha - p, \alpha + p]$$

DIM

TH DI LAGRANGE (O DEL VALORE MEDIO)

$$f \in C^1([\alpha, b]) \quad f(b) - f(\alpha) = (b - \alpha) f'(\varepsilon) \quad \varepsilon \in [\alpha, b]$$

- $$1) \quad \text{PER INDUZIONE: } n=0 \checkmark$$

$$|x_{n+1} - \alpha| = |\phi(x_n) - \phi(\alpha)| = |x_n - \alpha| \cdot |\phi'(\varepsilon)|$$

LAGRANGE

$$\varepsilon \in [x_n, \alpha]$$

PASSO INDUTTIVO:

$$|x_{n+1} - \alpha| \leq p \cdot \underbrace{|\phi'(\varepsilon)|}_{\leq 1} < p \Rightarrow x_{n+1} \in (\alpha - p, \alpha + p)$$

$$\begin{aligned}
 2) |x_{n+1} - \alpha| &= |\phi(x_n) - \phi(\alpha)| = |x_n - \alpha| \cdot |\phi'(\varepsilon)| \\
 &\leq k \cdot |x_n - \alpha| \leq \underbrace{k^n}_{\downarrow \text{PER } n \rightarrow +\infty} \cdot |x_0 - \alpha|
 \end{aligned}$$

$$\lim_{n \rightarrow +\infty} |x_{n+1} - \alpha| = 0$$

$\lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} \leq k \in [0, 1) \Rightarrow$ la successione
 converge linearmente
 o superlinearmente

3) Supponiamo per assurdo che $\exists \tilde{\alpha} \neq \alpha$, $\tilde{\alpha} \in [\alpha - p, \alpha + p]$
 t.c. $\phi(\tilde{\alpha}) = \tilde{\alpha}$

$$|\alpha - \tilde{\alpha}| = |\phi(\alpha) - \phi(\tilde{\alpha})| = |\alpha - \tilde{\alpha}| \cdot \phi'(\varepsilon)$$

\downarrow
LAGRANGE

$$\varepsilon \in (\alpha - p, \alpha + p)$$

$$\leq |\alpha - \tilde{\alpha}| \cdot k < |\alpha - \tilde{\alpha}|$$

\uparrow contraddizione
 $|\alpha - \tilde{\alpha}|$ non può essere < stretta di se stessa

OSS ① Se ϕ' ha segno costante in $(\alpha - p, \alpha + p)$
 allora si ha converg. monotona (crescente o decrescente) quando $\phi' > 0$

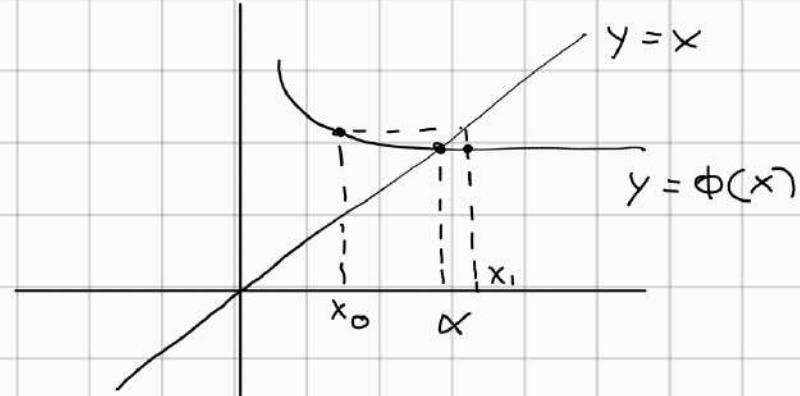
$$x_0 < \alpha \Rightarrow x_0 < x_1 < x_2 < \dots < x_n < \alpha$$

$$x_0 > \alpha \Rightarrow x_0 > x_1 > x_2 > \dots > x_n > \alpha$$

② Se $\phi' < 0$ in $(\alpha - \rho, \alpha + \rho)$ si ha convergenza alternata:

$$x_0 \in (\alpha - \rho, \alpha) \Rightarrow x_1 \in (\alpha, \alpha + \rho) \Rightarrow x_2 \in (\alpha - \rho, \alpha) \dots$$

$$x_0 \in (\alpha, \alpha + \rho) \Rightarrow x_1 \in (\alpha - \rho, \alpha) \Rightarrow x_2 \in (\alpha, \alpha + \rho) \dots$$



Formalmente:

$$x_{n+1} - \alpha = \phi(x_n) - \phi(\alpha) = (x_n - \alpha) \phi'(\varepsilon)$$

Se $\phi'(\varepsilon) > 0 \Rightarrow \text{sign}(x_{n+1} - \alpha) = \text{sign}(x_n - \alpha)$

Se $\phi'(\varepsilon) < 0 \Rightarrow \text{sign}(x_{n+1} - \alpha) = -\text{sign}(x_n - \alpha)$

$$\text{sign}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$$

Sulla velocità di convergenza si ha il seguente teorema:

α PUNTO FISSO
DI ϕ

TH $\phi \in C^P(I), \overline{\alpha = \phi(\alpha)}$ allora

$\exists p > 0 : \forall x_0 \in (\alpha - \rho, \alpha + \rho)$ la successione x_n converge con ordine $p \geq 2 \Leftrightarrow$

$$\phi'(\alpha) = \phi''(\alpha) = \dots = \phi^{(P-1)}(\alpha) = 0 \text{ e } \phi^{(P)}(\alpha) \neq 0$$

DIM

Sviluppo di TAYLOR DI $\phi(x)$: $\phi(x) = \sum_{j=0}^{+\infty} \frac{\phi^{(j)}(\alpha)}{j!} (x-\alpha)^j$

Ma poiché $\phi(x)$ è derivabile solo p volte:

$$\phi(x) = \sum_{j=0}^{p-1} \frac{\phi^{(j)}(\alpha)}{j!} (x-\alpha)^j + \underbrace{\frac{\phi^{(p)}(\varepsilon)}{(p+1)!} (x-\alpha)^p}_{\text{RESTO DI PEANO}}$$

$\Leftrightarrow \phi'(\alpha) = 0 \Rightarrow$ per continuità

$$\exists p > 0 : |\phi'(x)| \leq k \quad \forall x \in (\alpha-p, \alpha+p)$$

APERTO
CHIUSO
INDIFERENTE

In particolare, per il th. prec., la successione

$$\begin{cases} x_0 \in (\alpha-p, \alpha+p) \\ x_n = \phi(x_{n-1}) \end{cases}$$

converge ad α almeno linearmente

$$x_{n+1} = \phi(x_n) = \phi(\alpha) + \phi'(\alpha)(x_n - \alpha) + \dots + \frac{\phi^{(p-1)}(\alpha)(x_n - \alpha)^{p-1}}{(p-1)!} + \frac{\phi^{(p)}(\varepsilon)(x_n - \alpha)^p}{p!}$$

Sviluppo di TAYLOR IN α CON RESTO DI PEANO

$$= \alpha + \frac{\phi^{(p)}(\varepsilon) \cdot (x_n - \alpha)^p}{p!}$$

$\varepsilon \in [x_n, \alpha]$
SICCHE
 $x_n \rightarrow \infty \Rightarrow \lim_{n \rightarrow +\infty} \varepsilon = \alpha$

$$\lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = \lim_{n \rightarrow +\infty} \frac{|\phi(x_n) - \alpha|}{|x_n - \alpha|^p} = \lim_{n \rightarrow +\infty} \frac{|\phi^{(p)}(\varepsilon)| \cdot |x_n - \alpha|^p}{|x_n - \alpha|^p} =$$

$$= \lim_{n \rightarrow +\infty} \frac{\phi^{(p)}(\varepsilon)}{p!} = \frac{\phi^{(p)}(\alpha)}{p!} = \text{num. finito}$$

$\varepsilon \neq 0$ poiché abbiamo assunto $\phi^{(p)}(\alpha) \neq 0$

IL METODO CONVERGE SUPERUN. CON ORDINE p

\Leftarrow

\Rightarrow Preso $1 \leq r \leq p-1$ vogliamo fare vedere che
 $\phi^{(r)}(\alpha) = 0 \quad \forall r$. Procediamo per induzione:

$$\boxed{r=1} \quad 0 = \lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} = \lim_{n \rightarrow +\infty} \frac{|\phi(x_n) - \alpha|}{|x_n - \alpha|}$$

$$= \lim_{n \rightarrow +\infty} \frac{|x_n - \alpha| \cdot \phi'(\varepsilon)}{|x_n - \alpha|} = \lim_{n \rightarrow +\infty} \phi'(\varepsilon) = \phi'(\alpha)$$

$$\boxed{r \leq p-1} \quad (\text{assumendo } \phi^{(1)}(\alpha) = \dots = \phi^{(r-1)}(\alpha) = 0)$$

Il metodo converge con ordine $p > r \Rightarrow$

$$0 = \lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^r} = \lim_{n \rightarrow +\infty} \frac{|\phi(x_{n+1}) - \phi(\alpha)|}{|x_n - \alpha|^r} = ?$$

Consideriamo lo sviluppo di Taylor di ϕ in α troncato all'ordine r :

$$\phi(x_n) - \phi(\alpha) = (x_n - \alpha) \underbrace{\phi'(\alpha)}_0 + \dots + \frac{(x_n - \alpha)^{r-1}}{(r-1)!} \underbrace{\phi^{(r-1)}(\alpha)}_0 +$$

$$+ \frac{(x_n - \alpha)^r}{r!} \phi^{(r)}(\varepsilon)$$

per ipotesi iniziale

$$= \frac{(x_n - \alpha)^r}{r!} \cdot \phi^{(r)}(\varepsilon)$$

$$0 = \lim_{n \rightarrow +\infty} \frac{|\phi(x_n) - \phi(\alpha)|}{|x_n - \alpha|^r} = \lim_{n \rightarrow +\infty} \frac{\phi^{(r)}(\varepsilon)}{r!} = \frac{\phi^{(r)}(\alpha)}{r!}$$

$$\Rightarrow \phi^{(r)}(\alpha) = 0 \Rightarrow \phi^{(1)}(\alpha) = \dots = \phi^{(p-1)}(\alpha) = 0$$

CONCLUSIONE $|\phi'(\alpha)| < 1 \wedge \phi'(\alpha) \neq 0 \Rightarrow$ il metodo converge linearmente.

RIASSUNTO: Se ϕ è sufficientemente regolare (derivabile) allora:

$|\phi'(\alpha)| < 1 \Rightarrow$ converge localmente in modo almeno lineare

$\phi'(\alpha) = 0 \Rightarrow$ converge superlinealmente.

$\phi''(\alpha)$ $\begin{cases} \neq 0 \\ = 0 \end{cases} \Rightarrow$ converge quadraticamente
almeno cubica

...

Se ϕ è derivabile n volte in α allora l'ordine di convergenza è un intero

ESEMPIO

$$f(x) = x \log x - 1$$

Analizzare la convergenza dei metodi di punto fisso corrispondenti a:

$$1) \phi_1(x) = \frac{1}{\log x}$$

$$\alpha: f(\alpha) = 0$$

c'è un'unica sol-

$$\alpha \in (1, e)$$

$$f(1) = -1 < 0$$

$$f(e) = e - 1 > 0$$

$$2) \phi_2(x) = e^{1/x}$$

$$3) \phi_3(x) = \frac{x+1}{\log(x)+1}$$

$$1) x \log(x) - 1 = 0 \Leftrightarrow x = \frac{1}{\log(x)} \quad \begin{cases} x_0 \text{ DATO} \\ x_{n+1} = \frac{1}{\log(x_n)} \end{cases}$$

$$\phi'_1(x) = \frac{1}{x} \cdot \left(-\frac{1}{\log^2(x)} \right) = -\frac{1}{x \log^2(x)}$$

METODO NON CONVERGENTE

↑

$$\phi'_1(\alpha) = -\frac{1}{\alpha \log^2(\alpha)} = -\frac{1}{\log(\alpha)} \Rightarrow |\phi'_1(\alpha)| > 1$$

$$\alpha \in (1, \infty) \Rightarrow \log(\alpha) \in (0, 1)$$

$$2) x \log(x) - 1 = 0 \Leftrightarrow \log(x) = \frac{1}{x} \Leftrightarrow e^{1/x} = x$$

↓

$$\phi'_2(x) = -\frac{1}{x^2} \cdot e^{1/x}$$

$$e^{1/x} = \alpha$$

$$\phi'_2(\alpha) = -\frac{e^{1/\alpha}}{\alpha^2} = -\frac{1}{\alpha} \underset{\substack{\text{SICHIRO} \\ \alpha \in (1, e)}}{\Rightarrow} |\phi'_2(\alpha)| < 1$$

$\Rightarrow \exists$ un intervallo $(\alpha - \delta, \alpha + \delta)$ dove il metodo converge almeno linearmente

In realtà $-\frac{1}{\alpha} \neq 0$ su $(1, e)$ \Rightarrow il metodo converge linearmente

$$3) x \log(x) - 1 = 0 \Leftrightarrow x - \frac{x \log(x) - 1}{\log(x) + 1} = x \Leftrightarrow \frac{x+1}{\log(x)+1} = x$$

$$\phi'(x) = \frac{\log(x) + 1 - (x-1) \frac{1}{x}}{(\log(x) + 1)^2} = \frac{x \log(x) - 1}{x(\log(x) + 1)^2}$$

$$\phi'(\alpha) = 0$$

$$\phi''(x) = (\log(x) + 1)^3 x - (x \log(x) - 1) \left[(\log(x) + 1)^2 + \frac{2\log(x) + 1}{x} \right]$$

$$x^2 (\log(x) + 1)^4$$

$$\phi''(\alpha) = \frac{1}{\alpha(\log(\alpha) + 1)} \neq 0 \Rightarrow \text{il metodo converge superlinearmente.}$$

com ordine 2

$$x_n = \phi(x_{n-1})$$

$x = \phi(x)$ se α è soluzione

$x_n \rightarrow \alpha$ e quanto velocemente
 $n \rightarrow +\infty$

ESEMPIO

$$\phi(x) = \alpha x + \frac{b}{x} \quad a, b \in \mathbb{R}$$

Discutere per quali valori dei parametri a e b si ha:

- 1) esistenza di punti fissi ^(reali) e quanti sono
- 2) quando il metodo all'iterazione di punto fisso converge localmente e se è monotona
- 3) se ci sono così in cui converge superlinearmente e con quale ordine

Dividiamo l'analisi in due casi:

$$i) b=0 \Rightarrow \phi(x) = \alpha x$$

1 PUNTO FISSO ($\alpha=0$) PER $\alpha \neq 1$
o PUNTI FISSI (OGNI $x \in \mathbb{R}$) PER $\alpha=1$

$$ii) b \neq 0 \Rightarrow \alpha = \phi(\alpha) \Leftrightarrow \alpha = \alpha x + \frac{b}{x} \Leftrightarrow (1-\alpha)x^2 - b = 0 \Rightarrow \alpha = \pm \sqrt{\frac{b}{1-\alpha}}$$

$$\frac{b}{1-\alpha} > 0 \quad (\text{per avere i pti fissi reali}) \quad \downarrow \quad \alpha \neq 1$$

$$\begin{cases} b > 0 \\ 1-\alpha > 0 \end{cases} \Rightarrow \boxed{\alpha < 1, b > 0}$$

$$\begin{cases} b < 0 \\ 1-\alpha < 0 \end{cases} \Rightarrow \boxed{\alpha > 1, b < 0}$$

Sotto queste ip. si hanno 2 pti fissi $\alpha_{1,2} = \pm \sqrt{\frac{b}{1-\alpha}}$

DEF L' iterazione di pto fisso $\{x_n = \phi(x_{n+1})\}$

converge localmente ad α : $\alpha = \Phi(\alpha)$ se

$\exists \rho > 0$: $\forall x \in (\alpha - \rho, \alpha + \rho)$ si ha $|\phi'(x)| < 1$

OSS Se $|\phi'(\alpha)| < 1 \Rightarrow$ converge localmente

DEF L' iter. di pto fisso converge monotonicamente

ad α se $\exists \rho > 0$: $\forall x \in (\alpha - \rho, \alpha + \rho)$

$$0 \leq \phi'(x) < 1$$

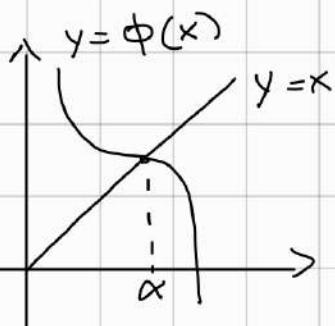
Analogamente, c'è converg. alternata (localm.)

$$\text{se } -1 \leq \phi'(x) < 0$$

OSS $\phi'(\alpha) > 0 \Rightarrow$ converge localm. im modo monot.

$\phi'(\alpha) < 0 \Rightarrow$ " " " " alternata

Se $\phi'(\alpha) = 0$ devo verificare il segno intorno ad α



$$\phi'(\alpha) = 0$$

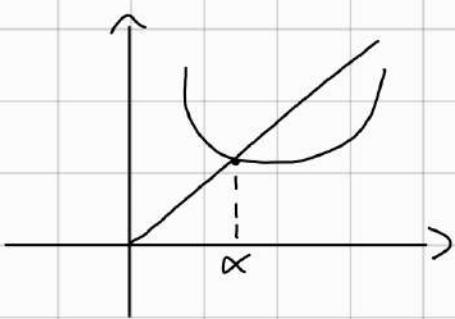
$$\phi'(x) \leq 0 \quad \forall x \in (\alpha - \rho, \alpha + \rho)$$

\Rightarrow conv. altern. e superl.

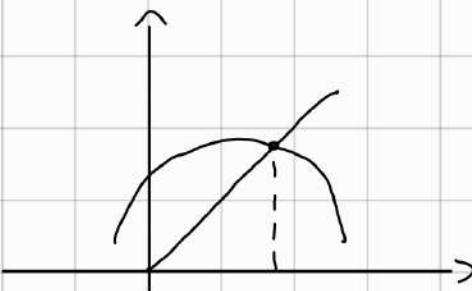


$$\phi'(\alpha) = 0 \quad \phi'(x) \geq 0 \quad \forall x \in (\alpha - \rho, \alpha + \rho)$$

\Rightarrow conv. monotona



$$\begin{aligned}\phi'(x) &= 0 \\ \phi'(x) &\leq 0 \quad \forall x \in (\alpha - \rho, \alpha) \\ \phi'(x) &\geq 0 \quad \forall x \in (\alpha, \alpha + \rho)\end{aligned}$$



$$\begin{aligned}\phi'(x) &\geq 0 \quad \forall x \in (\alpha - \rho, \alpha) \\ \phi'(x) &\leq 0 \quad \forall x \in (\alpha, \alpha + \rho) \\ \phi'(\alpha) &= 0\end{aligned}$$

2) CONVERGENZA

$$\underline{b=0} \quad \phi(x) = \alpha x \quad \phi'(x) = \alpha$$

$$\text{converg. locale} \Leftrightarrow |\alpha| < 1, b = 0$$

$$\text{converg. monotona} \Leftrightarrow 0 \leq \alpha < 1$$

$$\underline{b \neq 0} \quad \phi'(x) = \alpha - \frac{b}{x^2} \quad x_{1,2} = \pm \sqrt{\frac{b}{1-\alpha}}$$

$$\phi'(\alpha) = \alpha - \frac{b}{b} (1-\alpha) = 2\alpha - 1$$

$$|2\alpha - 1| < 1 \Leftrightarrow -1 < 2\alpha - 1 < 1 \Leftrightarrow 0 < 2\alpha < 2$$

$$\Leftrightarrow 0 < \alpha < 1$$

$$\phi'(\alpha) > 0 \quad 2\alpha - 1 \geq 0 \Rightarrow \alpha > \frac{1}{2}$$

Cosa succede per $\alpha = \frac{1}{2}$?

$$\phi(x) = \frac{x}{2} + \frac{b}{x}$$

$$\phi'(x) = \alpha - \frac{b}{x^2}$$

convexg. superlineare, ma
non monotona

....?

ESERCIZIO

$$\phi(x) = 1 + \alpha \log(x) + b \log^2(x) \quad \alpha, b \in \mathbb{R}$$

$x=1$ PRO FISSO DI $\phi(x)$

- 1) Per quali param. la convexg. è lineare e per quali è monotona
- 2) Se $\alpha=0$ si dimostri che la convexg. è superlineare e si determini l'ordine in funzione di b

$$1) \phi'(x) = \frac{\alpha}{x} + 2b \frac{\log(x)}{x} \rightsquigarrow \phi'(1) = \alpha$$

Si ha convexg. lineare per $0 < |\alpha| < 1$
(abbiamo escluso 0 perché allora la conv.
è superlineare)

Si ha conv. monotona per $0 < \alpha < 1$

$$2) \alpha = 0 \rightsquigarrow \phi'(1) = \alpha = 0$$

$$\phi''(x) = \frac{2b}{x^2} - 2b \frac{\log(x)}{x^2} = \frac{2b(1-\log(x))}{x^2} \rightsquigarrow \phi''(1) = 2b$$

Se $b \neq 0 \Rightarrow$ conv. quadratico

Se $b = 0 \Rightarrow$ conv. in 1 passo

$$\alpha = 0, \phi'(1) = 0 :$$

$$\alpha = 2$$

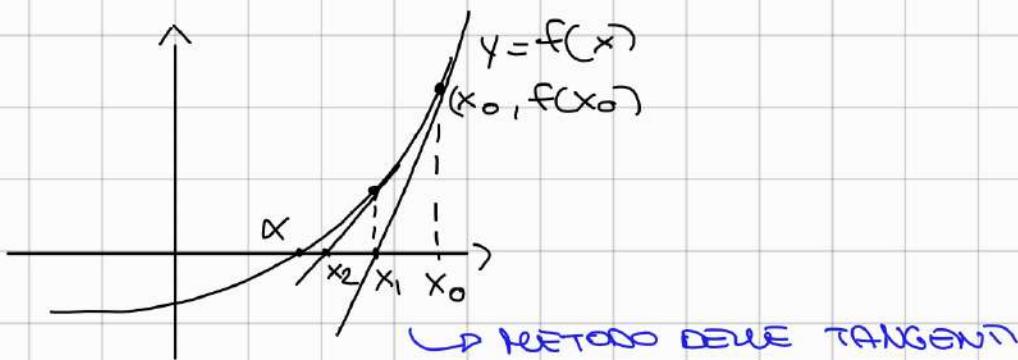
se $b > 0 \Rightarrow \phi'$ è crescente $\Rightarrow \phi' > 0 (\alpha, \alpha + \rho)$

se $b < 0 \Rightarrow \phi'$ è decrescente $\Rightarrow \phi' < 0 (\alpha - \rho, \alpha)$

METODO DI NEWTON

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

nel caso in cui vogliamo risolvere $f(x) = 0$



CRITERIO D'ARRESTO

$$|x_{n+1} - x_n| < \text{TOL} \quad \text{es. } \text{TOL} = 10^{-8}$$

OPPURE

$$\frac{|x_{n+1} - x_n|}{\min(|x_n|, |x_{n+1}|)} < \text{TOL} \rightarrow \text{piú preciso}$$

METODO DI NEWTON

$$f \in C^1([a, b]) \quad \phi(x) = x - \frac{f(x)}{f'(x)}$$

$$f(x) = 0 \quad \left\{ \begin{array}{l} x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \\ x_0 \text{ DATO} \end{array} \right.$$

DEF $f: [a, b] \rightarrow \mathbb{R}$ continua, $\alpha \in [a, b]: f(\alpha) = 0$
 α è detta **RADICE D'ORDINE r** se: $\lim_{x \rightarrow \alpha} \frac{f(x)}{(x - \alpha)^r} = c \quad c \neq 0$
 c FINITO ($\neq \pm \infty$)

OSS Vicino ad α , $f(x) \approx (x - \alpha)^r$

LEMMA Se $f \in C^r([a, b])$, $\alpha \in [a, b]$ radice di
 ordine r se e solo se:
 $0 = f(\alpha) = f'(\alpha) = \dots = f^{(r-1)}(\alpha), f^{(r)}(\alpha) \neq 0$

N.B.: non confondersi ϕ ed f e il loro comportamento (e quello delle loro derivate) in α

DEF Se $r=1 \Rightarrow \alpha$ è detta **RADICE SEMPLICE**

Se $r>1 \Rightarrow \alpha$ ^r " **RADICE MULPIA**

$f \in C^2([a, b])$

TH Se α è radice semplice di f su $[a, b]$ allora il metodo di Newton converge (localmente) in maniera superlineare.

Infine: se $f''(\alpha) = 0 \Rightarrow$ l'ordine p di converg. superlineare è ≥ 2

se $f''(\alpha) \neq 0 \Rightarrow p = 2$

$$\lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^2} = \frac{1}{2} \frac{|f''(\alpha)|}{|f'(\alpha)|}$$

DIM

$$\phi(x) = x - \frac{f(x)}{f'(x)}$$

$$\phi'(x) = 1 - \left[\frac{[f'(x)]^2 - f(x)f''(x)}{[f'(x)]^2} \right] = \frac{f(x)f''(x)}{[f'(x)]^2}$$

α RADICE SEMPLICE $\Rightarrow f'(\alpha) \neq 0, f(\alpha) = 0$

$\Rightarrow \phi'(\alpha) = 0 \Rightarrow$ converg. locale superlineare

$$\phi''(x) = \frac{(f' \cdot f'' + f \cdot f''')(f')^2 - ff'' \cdot 2f'f''}{(f')^4}$$

$$\phi''(\alpha) = \frac{(f')^3 f''}{(f')^4} = \frac{f''(\alpha)}{f'(\alpha)} \Rightarrow \begin{cases} \text{so se } f''(\alpha) = 0 \Rightarrow p > 2 \\ \neq 0 \text{ se } f''(\alpha) \neq 0 \Rightarrow p = 2 \end{cases}$$

Sviluppo di Taylor in α
è valutato in x_n

↑ (TRONCATO)

$$\lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^2} = \lim_{n \rightarrow +\infty} \frac{|\phi(x_n) - \phi(\alpha)|}{|x_n - \alpha|^2} = \text{AU'ORDINE 2 CON RESTO}$$

$$= \lim_{n \rightarrow +\infty} \frac{|\phi'(\alpha) \cdot (x_n - \alpha) + \phi''(\varepsilon)(x_n - \alpha)^2|}{|x_n - \alpha|^2}$$

$\varepsilon \in (\alpha, x_n)$

$$= \lim_{n \rightarrow +\infty} \frac{\frac{1}{2} |\phi''(\varepsilon)| |x_n - \alpha|^2}{|x_n - \alpha|^2} =$$

$$= \lim_{n \rightarrow +\infty} \frac{1}{2} |\phi''(\varepsilon)| = \frac{1}{2} |\phi''(\alpha)| = \frac{1}{2} \left| \frac{f''(\alpha)}{f'(\alpha)} \right|$$

OSS Il th assicura che Newton funziona (e funziona bene, veloce) se si parte abbastanza vicini ad α .

Se non si ha o non si è sicuri di avere un punto di partenza abbastanza vicino ad α , si può iniziare ad es. con un po' di passi di bisezione e quando si arriva a un intervallo abbast. piccolo raffinare l'approssim. con Newton.

ESERCIZIO

Calcolo del reciproco di un num. $t \in \mathbb{R}^+$ senza usare divisione.

Voglio calcolare t^{-1} : cerca la radice di $f(x) = x^{-1} - t \Rightarrow f(t^{-1}) = 0$

$$\phi(x) = x - \frac{x^{-1} - t}{-x^{-2}} = x + x - tx^2 = 2x - tx^2$$

$$x_{n+1} = 2x_n - tx_n^2$$

Come scelgo x_0 ?

$$|\phi'(x)| < 1$$

$$\phi'(x) = 2 - 2tx = 2(1 - tx)$$

$$\begin{cases} \phi'(1-tx) < 1/2 \\ \phi'(1-tx) > -1/2 \end{cases} \Leftrightarrow \begin{cases} x > 1/2t = \frac{1}{2}t^{-1} \\ x < 3/2t = \frac{3}{2}t^{-1} \end{cases}$$

\Rightarrow è necessario non sovrastimare t^{-1} più di un fattore 1.5 e non sottostimare t^{-1} più di un fattore 0.5

ESEMPIO

Calcolare $\sqrt[k]{t}$, $t \in \mathbb{R}^+$ senza estrazioni di radice
 $k \in \mathbb{N}$

Si considera $f(x) = x^k - t$

$$\text{NEWTON } \sim \phi(x) = x - \left(\frac{x^k - t}{kx^{k-1}} \right) = \frac{k-1}{k}x + \frac{t}{kx^{k-1}}$$

Come scelgo x_0 ?

$$|\phi'(x)| < 1$$

$$\phi'(x) = \frac{k-1}{k} + \frac{t}{k} \underbrace{(1-k)}_{< 0} x^{-k}$$

$$|\phi'(x)| < 1 \Leftrightarrow \begin{cases} \frac{k-1}{k} + \frac{t}{k}(1-k)x^{-k} < 1 \\ \frac{k-1}{k} + \frac{t}{k}(1-k)x^{-k} > -1 \end{cases} \Leftrightarrow \begin{cases} x^k > t(1-k) \\ x^k > \frac{(k-1)}{(2k-1)}t \end{cases} \quad \underbrace{< 0}_{\text{per } k > 1}$$

$x \geq 0 \Rightarrow$ la 1° condizione è sempre verificata
 $x > \underbrace{\sqrt[2k-1]{\frac{k-1}{2k-1}}}_{\in (0,1)} \cdot \sqrt[k]{t} \Rightarrow$ se x_0 non sottovaluta più di un fattore $\sqrt[2k-1]{\frac{k-1}{2k-1}}$ \Rightarrow il metodo converge

Se parto con una sovravalsimma invece il metodo converge sempre

TH Sotto le stesse assunzioni del th precedente (a parte che α è una radice di ordine $r \geq 1$) il metodo di Newton (modificato)

$$x_{n+1} = \begin{cases} x_n - \frac{f(x_n)}{f'(x_n)} \\ \alpha \quad \text{se } x_n = \alpha \end{cases}$$

converge linearmente e si ha $\lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} = 1 - \frac{1}{2}$

DIM

α di multpl. $r \Rightarrow f(x) = g(x)(x - \alpha)^r$

$$g(x) = \begin{cases} \frac{f(x)}{(x - \alpha)^r} & x \neq \alpha \\ \lim_{n \rightarrow \infty} \frac{f(x)}{(x - \alpha)^r} & x = \alpha \end{cases}$$

↑ il fatto che α ha mult. r implica che il lim. esiste finito.

Per costruzione $g(x)$ è continua.

$$\Phi(x) = x - \frac{g(x)(x - \alpha)^r}{g'(x)(x - \alpha)^r + g(x) \cdot r(x - \alpha)^{r-1}} =$$

$$= x - \frac{g(x)(x-\alpha)}{g'(x)(x-\alpha) + rg(x)}$$

$$\phi'(x) = 1 - \frac{[(g'(x-\alpha) + g)(r \cdot g + g'(x-\alpha)) - g(x-\alpha)[rg' + g''(x-\alpha) + g]]}{[r \cdot g + g'(x-\alpha)]^2}$$

$$\Rightarrow \underbrace{\phi'(\alpha)}_{\parallel} = 1 - \frac{r \cdot g(\alpha)^2}{r^2 \cdot g(\alpha)^2} = 1 - \frac{1}{r} < 1 \quad (\text{per } r > 1)$$

$$\lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|}$$

↓
converg. lineare
localmente

□

In generale, all'aumentare della mult. della radice, il metodo di Newton converge più lentamente.

Oss Se si conosce r (la mult. della radice) si può modificare Newton come segue:

$$x_{n+1} = x_n - r \cdot \frac{f(x_n)}{f'(x_n)}$$

Con questa scelta si ritrova $\phi'(\alpha) = 0$

IPOTESI CHE GARANTISCONO CONVERGENZA "GUARDA"

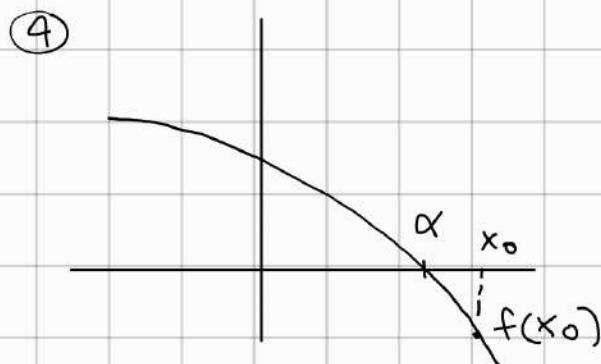
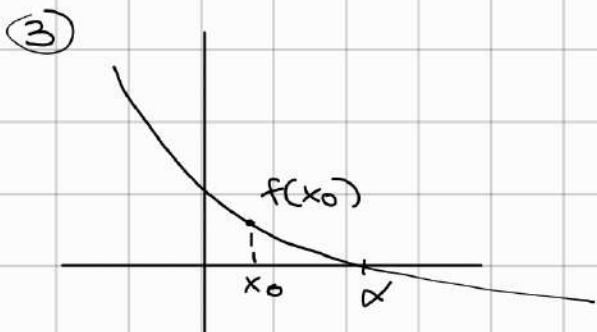
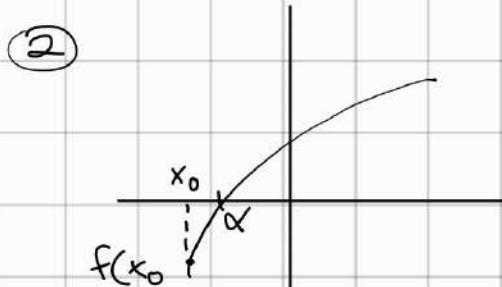
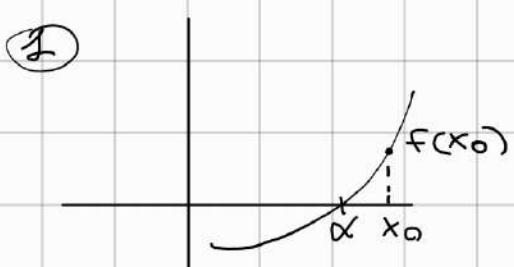
TH $[a, b] \subseteq \mathbb{R}$, $f \in C^2([a, b])$, α reale s.m. $f(a) \cdot f(b) < 0$
 $\Rightarrow \exists \alpha \in [a, b] : f(\alpha) = 0$
e supponiamo che f' e f'' hanno segno costante su $[a, b]$.

Allora, se scelgo $x_0 \in [a, b]$ t.c. $f(x_0) \cdot f''(x_0) > 0$
 \Rightarrow il metodo di Newton converge monotonicamente ad α

OSS La condizione $f(x_0) \cdot f''(x_0) > 0$ è verificata in uno dei due semi-intervalli:
 $[a, \alpha]$ o $[\alpha, b]$

OSS Si hanno 4 possibili casi:

- ① $f' > 0, f'' > 0 \Rightarrow f$ CONVessa CRESCENTE
- ② $f' > 0, f'' < 0 \Rightarrow f$ CONCAVA CRESCENTE
- ③ $f' < 0, f'' > 0 \Rightarrow f$ CONVessa DECRESCENTE
- ④ $f' < 0, f'' < 0 \Rightarrow f$ CONCAVA DECRESCENTE



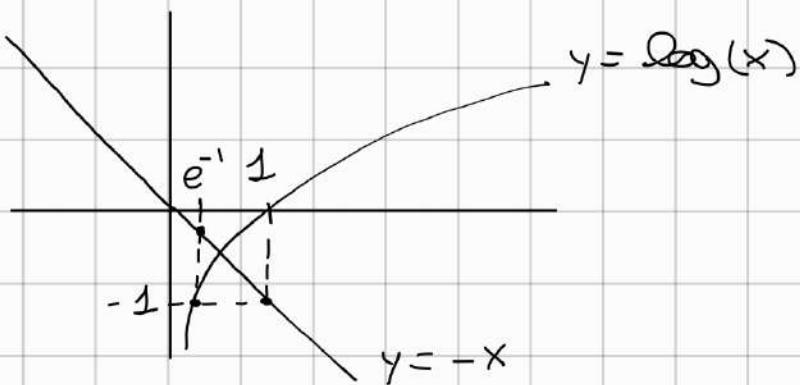
ESEMPIO

Trovare un intervallo dove si trova la radice di $f(x) = x + \log(x)$ e studiare la convergenza:

① $x_{n+1} = -\log(x_n)$

② NEWTON

$$x + \log(x) = 0 \Leftrightarrow \log(x) = -x \Leftrightarrow \begin{cases} y = -x \\ y = \log x \end{cases}$$



$$\begin{aligned} \log(e^{-1}) &= -1 < -e^{-1} \\ \log(1) &= 0 > -1 \end{aligned} \quad \Rightarrow \alpha \in (e^{-1}, 1)$$

① $\phi(x) = -\log(x) \Rightarrow \phi'(x) = -\frac{1}{x}$

ϕ' su $(e^{-1}, 1)$ è tale che $|\phi'| > 1$

\Rightarrow il metodo non converge

② $\phi(x) = x - \frac{x + \log(x)}{1 + \frac{1}{x}} = x \left(1 - \frac{x + \log(x)}{x + 1} \right) = x \left(\frac{1 - \log(x)}{x + 1} \right)$

$$f(x) = x + \log(x) \Rightarrow f' = 1 + \frac{1}{x}, f'' = -\frac{1}{x^2}$$

su $(e^{-1}, 1)$, ma anche su tutta la semiretta reale si ha $f' > 0, f'' < 0$

Per avere converg. devo scegliere

x_0 : $f(x_0) \cdot f''(x_0) > 0$, ovvero $f(x_0) < 0$

$$f(e^{-1}) = e^{-1} + (-1) < 0 \quad \checkmark \text{ in particolare ogni}$$

$x_0 \in (0, e^{-1})$ rende il
metodo di Newton converge-

ESERCIZIO

Consideriamo l'eq. $3x^2e^{-x} - 1 = 0$, $f(x) = 0$

Dire (giustificandolo) quante radici ha $f(x) = 0$ e fornire degli intervalli di separazione.

Ora se α_i , $i = 1, \dots, s$ sono le radici distinte fornire intervalli della forma $[a_i, b_i]$ +.c.

$\alpha_i \in [a_i, b_i]$ e $\alpha_j \notin [a_i, b_i]$ se $i \neq j$.

Infine studiare la convergenza locale dei metodi di punto fisso:

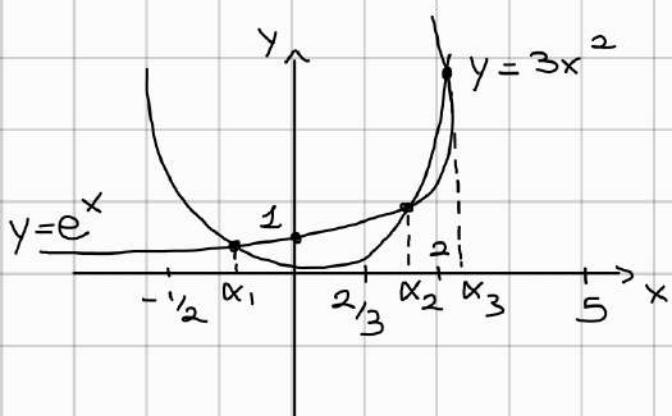
$$\textcircled{1} \quad x_{n+1} = \frac{e^{x_n}}{3x_n}$$

$$\textcircled{2} \quad x_{n+1} = \log(3x_n^2) \quad \times \text{ OSA}$$

\textcircled{3} NEWTON

(fornire $x_0^{(i)}$ +.c. $\{x_n\} \rightarrow \alpha_i$ quando parto da $x_0^{(i)}$)

$$3x^2e^{-x} - 1 = 0 \iff 3x^2 = e^x \iff \begin{cases} y = 3x^2 \\ y = e^x \end{cases}$$



$$\begin{aligned} f_1(x) &= 3x^2 \\ f_2(x) &= e^x \end{aligned}$$

$$2.5 \leq e \leq 3$$

$$f_2(0) = 1 > 0 = f_1(0)$$

$$f_1(-\frac{1}{2}) = 3(-\frac{1}{2})^2 = \frac{3}{4} > f_2(-\frac{1}{2}) = e^{-\frac{1}{2}} = \frac{1}{\sqrt{e}} < \frac{1}{\sqrt{2}}$$

$$\left(\frac{3}{4}\right)^2 = \frac{9}{16} > \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$$

$$\Rightarrow \alpha_1 \in (-\frac{1}{2}, 0)$$

$$f_1\left(\frac{2}{3}\right) = 3 \cdot \left(\frac{2}{3}\right)^2 = \frac{4}{3}$$

$$f_2\left(\frac{2}{3}\right) = e^{\frac{2}{3}} > e^{1/2} > 2^{1/2} \quad (2^{1/2})^2 = 2 \quad (4/3)^2 = \frac{16}{9} < 2$$

$$\Rightarrow f_1\left(\frac{2}{3}\right) < f_2\left(\frac{2}{3}\right)$$

$$f_1(1) = 3 > f_2(1) = e \Rightarrow \alpha_2 \in [\frac{2}{3}, 1]$$

$$f_1(2) = 3 \cdot 2^2 = 12$$

$$f_2(2) = e^2 < 3^2 = 9$$

$$\Rightarrow f_1(2) > f_2(2)$$

$$f_1(5) = 3 \cdot 5^2 = 75$$

$$f_2(5) = e^5$$

$$\Rightarrow e^5 > (2 \cdot 5)^5 = (5/2)^5 = \frac{3125}{32} \approx 97 > 75$$

$$\Rightarrow f_2(5) > f_1(5)$$

$$\Rightarrow \alpha_3 \in (2, 5)$$

$$\textcircled{1} \quad \phi(x) = \frac{e^x}{3x} \Rightarrow \phi'(x) = \frac{3xe^x - 3e^x}{9x^2} = \frac{e^x(x-1)}{3x^2}$$

$$\phi''(x) = \frac{(e^x(x-1) + e^x)3x^2 - e^x(x-1)6x}{9x^4} = \frac{e^x(x^2 - 2x + 2)}{3x^3}$$

$$x^2 - 2x + 2 \rightsquigarrow \Delta = 4 - 8 = -4 \Rightarrow \text{segno sempre positivo } \forall x \in \mathbb{R}$$

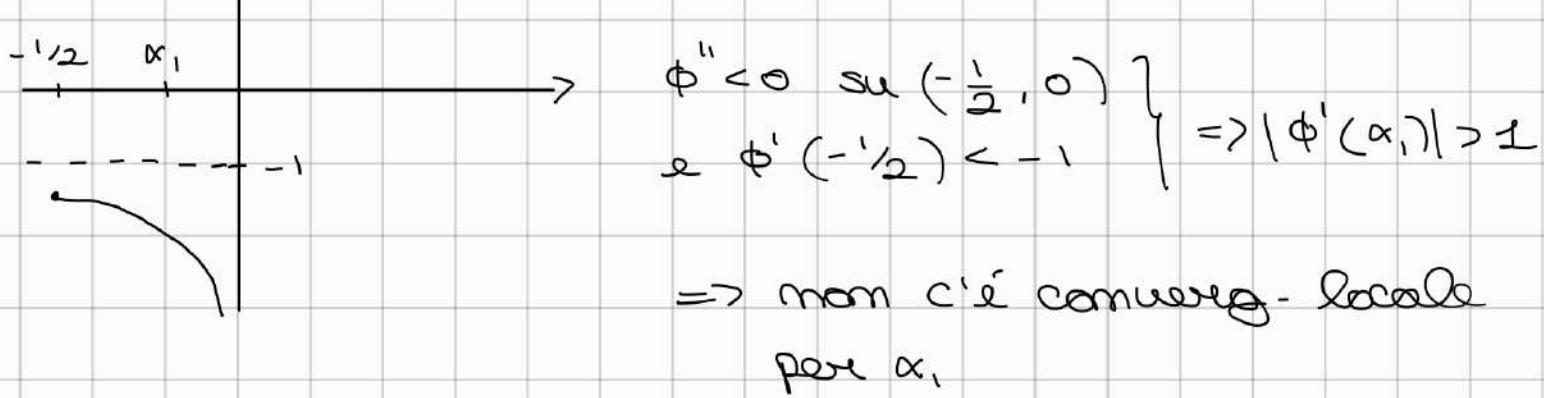
$$\text{sign}(\phi'(x)) = \begin{cases} > 0 & \text{per } \alpha_3 \\ < 0 & \text{per } \alpha_1, \alpha_2 \end{cases}$$

$$\text{sign}(\phi''(x)) = \begin{cases} < 0 & \text{per } x_1 \\ > 0 & \text{per } x_2, x_3 \end{cases}$$

Studiamo $x_1 \in (-\frac{1}{2}, 0)$

$$\phi'(-\frac{1}{2}) = \frac{e^{-\frac{1}{2}}(-\frac{3}{2})}{3/4} = -2e^{-\frac{1}{2}} < 1$$

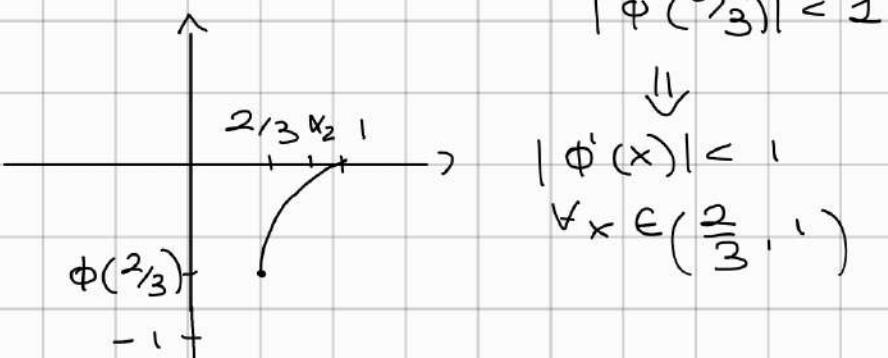
$$\text{perché } e^{-\frac{1}{2}} = \frac{1}{\sqrt{e}} > \frac{1}{\sqrt{4}} = \frac{1}{2}$$



Studiamo $x_2 \in (\frac{2}{3}, 1)$

$$\phi'(\frac{2}{3}) = \frac{e^{\frac{2}{3}}(-\frac{1}{3})}{3 \cdot \frac{4}{9}} = \frac{e^{\frac{2}{3}}(-\frac{1}{3})}{4/3} = -\frac{1}{4}e^{\frac{2}{3}} > -1$$

$$\phi'(1) = 0$$



x casa vedere cosa accade in (2, 5)

$$\textcircled{3} \quad f(x) = 3x^2 e^{-x} - 1$$

$$f'(x) = 6x e^{-x} - 3x^2 e^{-x} = 3x e^{-x}(2-x)$$

$$f'(x) = 0 \iff x = 0 \text{ oppure } x = 2 \Rightarrow f'(x_i) \neq 0 \quad i=1,2,3$$

$\Rightarrow \alpha_1, \alpha_2, \alpha_3$ sono radici semplici

\Rightarrow NEWTON converge localmente
in modo superlin. quadrattico

In questo caso si hanno segni costanti per f'' , f'
per cui usando il th di ieri si possono utilizzare
gli estremi dx o sx degli intervalli di separazione
che verificano la condizione $f(x_0) \cdot f''(x_0) > 0$

x CASA: dire chi sono $x_0^{(i)}$ $i=1,2,3$

$$x_0^{(1)} \nearrow -\frac{1}{2} \quad ?$$

$$x_0^{(2)} \nearrow \frac{2}{3} \quad ?$$

$$x_0^{(3)} \nearrow 2 \quad ?$$

EQUAZIONI NON LINEARI IN \mathbb{R}^m

Vogliamo risolvere $f(x) = 0$

$$f: \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

$$\begin{cases} f_1(x) = 0 \\ \vdots \\ f_m(x) = 0 \end{cases} \quad \begin{array}{l} \text{SISTEMA DI } m \text{ EQ.} \\ \text{IN } m \text{ INCognite} \end{array}$$

Anche qui si riformula il problema di trovare $x \in \mathbb{R}^m: f(x) = 0$ come un problema di punto fisso, ovvero si costruisce $\phi(x): \mathbb{R}^m \rightarrow \mathbb{R}^m$ tale per cui $f(x) = 0 \iff x = \phi(x)$ e si considera l'iterazione funzionale:

$$\begin{cases} x^{(n+1)} = \phi(x^{(n)}) \\ x^{(0)} \text{ DATO} \end{cases} \quad x^{(n)} \in \mathbb{R}^m \quad (\text{sono vettori di } m \text{ componenti})$$

DEF $\{x^{(n)}\}$ converge ad $x \in \mathbb{R}^m$ se per una qualche norma vettoriale $\|\cdot\|$ si ha:

$$\lim_{n \rightarrow \infty} \|x^{(n)} - x\| = 0$$

OSS Se c'è convergenza per una specifica norma (ad es. $\|\cdot\|_2$) allora c'è convergenza per ogni norma

DEF

$\{x^{(n)}\}$ converge com $\overbrace{\text{ordine}}^{\text{ad } \alpha \in \mathbb{R}^m}$ $p \geq 0$ se esistono una norma vett. $\|\cdot\|$ ed una costante $B \neq 0$ tale per cui:

$$B \neq \pm \infty$$

$$\lim_{n \rightarrow +\infty} \frac{\|x^{(n+1)} - \alpha\|}{\|x^{(n)} - \alpha\|^p} = B$$

OSS Nel caso $p=1$ si richiede $|B| < 1$

$$f(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$$

$$f(x_1, \dots, x_m) = \begin{bmatrix} f_1(x_1, \dots, x_m) \\ \vdots \\ f_m(x_1, \dots, x_m) \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

$$f(x) = 0 \iff \begin{cases} f_1(x_1, \dots, x_m) = 0 \\ \vdots \\ f_m(x_1, \dots, x_m) = 0 \end{cases}$$

ESEMPIO DI APPLICAZIONE

$$\min_{x \in \mathbb{R}^m} h(x)$$

GRADIENTE

Si devono trovare i punti stazionari t.c. $\nabla h(x) = 0$

$$\nabla h(x) := \begin{bmatrix} \frac{\partial}{\partial x_1} h(x) \\ \vdots \\ \frac{\partial}{\partial x_m} h(x) \end{bmatrix}$$

ESEMPPIO (IN DUE VARIABILI)

$$\begin{cases} x_1 - \frac{1}{4}(x_1^2 + x_2^2) = 0 \\ x_1 + x_2 - 2 = 0 \end{cases} \iff f(x) = 0$$

$$f(x) = \begin{bmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{bmatrix} \quad f_1(x_1, x_2) = x_1 - \frac{1}{4}(x_1^2 + x_2^2)$$

$$f_2(x_1, x_2) = x_1 + x_2 - 2$$

Poiché siamo in dim. 2 possiamo interpretare il problema come trovare le intersez. delle curve:

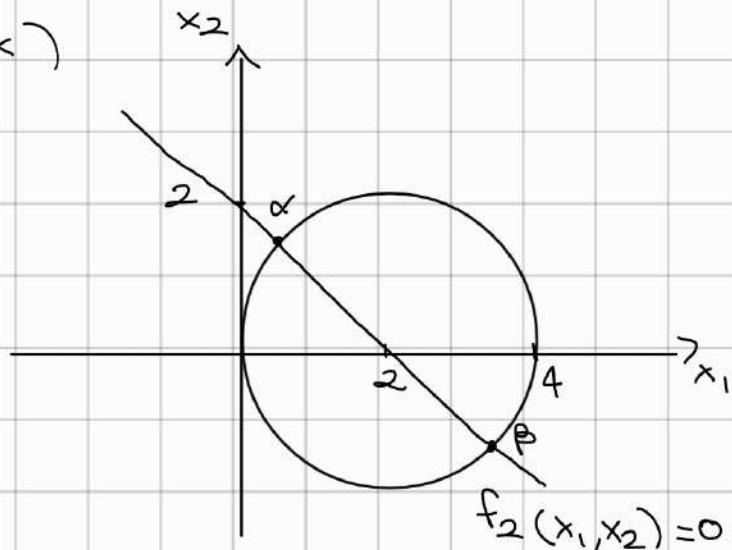
- $x_1 + x_2 - 2 = 0$ (retta $y = 2 - x$)

- $x_1 - \frac{1}{4}(x_1^2 + x_2^2) = 0$

$$\Downarrow$$

$$-\frac{(x_1 - 2)^2}{4} - \frac{x_2^2}{2} + 1 = 0$$

(cerchio di centro $(2, 0)$
 $\Leftrightarrow r = 2$)



Ci sono 2 radici $\alpha, \beta \in \mathbb{R}^2$

$$\begin{cases} x_1 - \frac{1}{4}(x_1^2 + x_2^2) = 0 \\ x_1 + x_2 - 2 = 0 \end{cases} \iff \begin{cases} x_1 = \frac{x_1^2 + x_2^2}{4} \\ x_2 = 2 - x_1 \end{cases}$$

Potremmo definire $\Phi(x_1, x_2) = \begin{bmatrix} \frac{x_1^2 + x_2^2}{4} \\ 2 - x_1 \end{bmatrix}$

Φ verifica $\Phi(\alpha) = \alpha$ e $\Phi(\beta) = \beta$

Quando $\begin{cases} x^{(n+1)} = \phi(x^{(n)}) \\ x^{(0)} \in \mathbb{R}^m \text{ DATO} \end{cases}$ converge?

DEF Nel caso in cui $\phi = \begin{bmatrix} \phi_1(x_1, \dots, x_m) \\ \vdots \\ \phi_m(x_1, \dots, x_m) \end{bmatrix}$ ha tutte

componenti derivabili rispetto ad ogni variabile
si definisce la matrice Jacobiana di ϕ come:

$$J\phi(x) = \begin{bmatrix} \frac{\partial \phi_1(x)}{\partial x_1} & \dots & \frac{\partial \phi_1(x)}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_m(x)}{\partial x_1} & \dots & \frac{\partial \phi_m(x)}{\partial x_m} \end{bmatrix} \quad (J\phi(x))_{ij} = \frac{\partial \phi_i(x)}{\partial x_j}$$

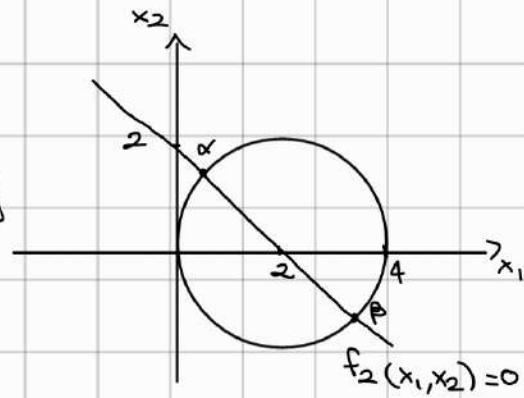
TH Se $\phi(x) \in C^1(\Omega)$ su $\Omega \subseteq \mathbb{R}^m$ t.c. $x \in \Omega$
(ed $\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} : \phi(\alpha) = \alpha$)

Se $\rho(J\phi(x)) < 1$ allora $\exists \|\cdot\|$ norma vettoriale
e $S > 0$ tale per cui $\forall x^{(0)}$ che verifica $\|x^{(0)} - \alpha\| < S$
la successione $\begin{cases} x^{(n)} = \phi(x^{(n-1)}) \\ x^{(0)} \end{cases}$ converge almeno

linearmente ad α (converge in ogni norma).

ESEMPIO

$$\Phi(x_1, x_2) = \begin{cases} \Phi_1(x_1, x_2) = \frac{1}{4}(x_1^2 + x_2^2) \\ \Phi_2(x_1, x_2) = 2 - x_1 \end{cases}$$



Converge localmente in α ?

$$\mathcal{J}\Phi(x_1, x_2) = \begin{bmatrix} x_{1/2} & x_{2/2} \\ -1 & 0 \end{bmatrix}$$

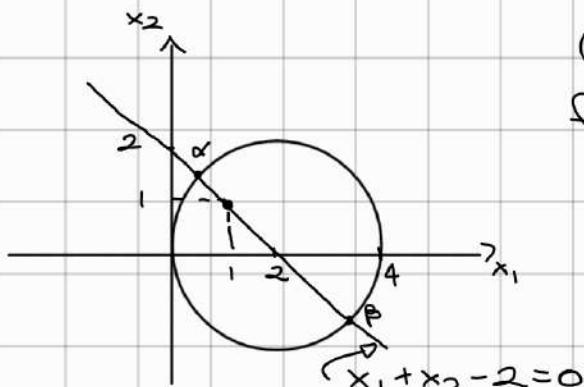
$$\mathcal{J}\Phi(\alpha_1, \alpha_2) = \begin{bmatrix} \alpha_{1/2} & \alpha_{2/2} \\ -1 & 0 \end{bmatrix} \quad \text{con} \quad \begin{cases} \frac{\alpha_1^2 + \alpha_2^2}{4} = \alpha_1 \\ 2 - \alpha_1 = \alpha_2 \end{cases}$$

Osserviamo che considerare $\|\mathcal{J}\Phi(\alpha)\|_1$ o $\|\mathcal{J}\Phi(\alpha)\|_\infty$ in questo caso non aiutano perché ci danno sempre qta ≥ 1 .

$$\det(\mathcal{J}\Phi(\alpha_1, \alpha_2) - \lambda I) = 0$$

$$\det\left(\begin{bmatrix} \frac{\alpha_1}{2} - \lambda & \frac{\alpha_2}{2} \\ -1 & -\lambda \end{bmatrix}\right) = \left(\frac{\alpha_1}{2} - \lambda\right)(-\lambda) + \frac{\alpha_2}{2} = \lambda^2 - \frac{\alpha_1}{2}\lambda + \frac{\alpha_2}{2} = 0$$

$$\Delta = \frac{\alpha_1^2}{4} - 2\alpha_2 \Rightarrow \lambda_{1,2} = \frac{\alpha_1 \pm \sqrt{\Delta}}{2}$$



$$(1, 1) \text{ sta sulla retta } x_1 + x_2 - 2 = 0$$

$$f_1(x) = x_1 + \frac{(x_1^2 + x_2^2)}{4}$$

nella parte
interna
al cerchio

$$f_1(1, 1) = 1/2 > 0$$

$$\left. \begin{array}{l} \alpha_1 < 1 \\ \alpha_2 > 1 \end{array} \right\} \Rightarrow \Delta < 0$$

$$\lambda_{1,2} = \frac{\frac{x_1}{2} \pm i\sqrt{-\Delta}}{2} = \frac{\frac{x_1}{2} \pm i\sqrt{\frac{x_1^2}{4} + 2x_2}}{2} = \frac{\frac{x_1}{4} + i\sqrt{\frac{x_1^2}{4} + 2x_2}}{2}$$

$\lambda_{1,2}$ sono due num. complessi coniugati (stesso modulo)

$$|\lambda_{1,2}| = \sqrt{\frac{x_1^2}{16} + \frac{2x_2}{4} - \frac{x_1^2}{16}} = \sqrt{\frac{x_2}{2}} = \rho(\Im\Phi(x))$$

$\Rightarrow \left| \sqrt{\frac{x_2}{2}} \right| < 1$ (perché $x_2 < 2$) \Rightarrow converg. locale in x

IDEA IN GENERALE PER COSTRUIRE $\Phi(x)$

$$x = \underbrace{x - G(x) \cdot f(x)}_{\Phi(x)}$$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}, f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}$$

dove $G(x) : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times m}$

$$G(x) = \begin{bmatrix} g_{1,1}(x) & \cdots & g_{1,m}(x) \\ \vdots & \ddots & \vdots \\ g_{m,1}(x) & \cdots & g_{mm}(x) \end{bmatrix}$$

$$x = x - G(x) f(x)$$

$$\begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} x \end{bmatrix} - \begin{bmatrix} G \end{bmatrix} \cdot \begin{bmatrix} f \end{bmatrix}$$

La scelta che si fa nel metodo di Newton in più variabili (chiamato anche metodo di NEWTON-RAPHSON) è:

$$G(x) = [\Im f(x)]^{-1}$$

$$\left\{ \begin{array}{l} x^{(n+1)} = x^{(n)} - [\Im f(x^{(n)})]^{-1} \cdot f(x^{(n)}) \\ x^{(0)} \text{ dato} \end{array} \right\} \text{METODO DI NEWTON-RAPHSON}$$

OSS Nell' implementazione pratica del metodo non si forma esplicito $[Jf(x^{(n)})]^{-1}$, ma si risolve il sistema lineare

$$Jf(x^{(n)}) \cdot d^{(n)} = f(x^{(n)})$$

e poi si calcola $x^{(n+1)} = x^{(n)} - d^{(n)}$

$$f(x) = 0 \quad x \in \mathbb{R}^m \quad f: \mathbb{R}^m \rightarrow \mathbb{R}^m$$

$$f = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix} \quad f_i: \mathbb{R}^m \rightarrow \mathbb{R}$$

$$f(x) = \begin{cases} f_1(x_1, \dots, x_m) = 0 \\ \vdots \\ f_m(x_1, \dots, x_m) = 0 \end{cases}$$

$$f(x) = 0 \iff \phi(x) = x$$

$$\begin{cases} x^{(0)} \in \mathbb{R}^m \text{ DATO} \\ x^{(n+1)} = \phi(x^{(n)}) \end{cases}$$

PER LA
CONVERG.

$$\alpha = \phi(\alpha) \rightsquigarrow \rho(\mathcal{J}\phi(\alpha)) < 1$$

$$\mathcal{J}\phi = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1}(\alpha) & \dots & \frac{\partial \phi_1}{\partial x_m}(\alpha) \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_m}{\partial x_1}(\alpha) & \dots & \frac{\partial \phi_m}{\partial x_m}(\alpha) \end{bmatrix}$$

Un modo generale per costruire $\phi(x)$ è
prendere $\phi(x) = x - G(x) \cdot f(x)$ per una qualche
funzione $G: \mathbb{R}^m \rightarrow \mathbb{R}^{m \times m}$

$$G(x) = \begin{bmatrix} g_{11}(x) & \dots & g_{1m}(x) \\ \vdots & & \vdots \\ g_{m1}(x) & \dots & g_{mm}(x) \end{bmatrix} \quad \phi(x) = [x] - [G][f(x)]$$

Idealmente $G(x)$ è non singolare ($\det(G(x)) \neq 0$)
per x vicino ad α .

Scegliendo $G(x) = \mathcal{J}f(x)^{-1}$ si ottiene il metodo di
Newton-Raphson.

$$N.R. \rightsquigarrow \phi(x) = x - \mathcal{J}f(x)^{-1} f(x)$$

$$\begin{cases} x^{(n+1)} = x^{(n)} - \mathcal{J}f(x^{(n)})^{-1} f(x^{(n)}) \\ x^{(0)} \in \mathbb{R}^m \text{ DATO} \end{cases}$$

TH Se $f(x) \in C^2(\Omega)$, $\Omega \subseteq \mathbb{R}^m$, $\alpha \in \Omega$ ($f(\alpha) = 0$),
 $\det(Jf(x)) \neq 0 \quad \forall x \in \Omega$
 $\Rightarrow \exists S$ intorno di α : $\forall x^{(0)} \in S$
 $\forall n \in \mathbb{N}$ esistente matriciale $\exists \beta > 0 : \|x^{(n+1)} - \alpha\| \leq \beta \|x^{(n)} - \alpha\|$

COROLARIO Se $\det(Jf(\alpha)) \neq 0$ $\exists \Omega$ intorno di α
per cui il metodo converge almeno quadraticamente per ogni $x^{(0)}$ in Ω .

COROLARIO Se $\det(Jf(\alpha)) = 0$ allora il metodo converge al più linearmente.

ESEMPIO

$$f(x_1, x_2) = \begin{bmatrix} \frac{1}{3}(x_1 - x_2) + x_1^2 \\ \frac{1}{3}(x_2 - x_1) + x_1 x_2 \end{bmatrix}$$

Determinare le soluzioni di $f(x_1, x_2) = [0 0]$ e dire per quali di loro il metodo di N.R. converge in maniera superlineare.

$$\begin{cases} x_1^2 + \frac{1}{3}(x_1 - x_2) = 0 \Rightarrow x_2 = x_1 + 3x_1^2 \\ \frac{1}{3}(x_2 - x_1) + x_1 x_2 = 0 \Rightarrow \frac{1}{3}(x_1 + 3x_1^2 - x_1) + x_1(x_1 + 3x_1^2) \end{cases}$$

$$\Rightarrow 2x_1^2 + 3x_1^3 = 0 \quad x_1 = 0$$

$$x_1^2(2 + 3x_1) = 0 \quad \begin{cases} x_1 = -2/3 \\ x_1 = 0 \end{cases}$$

$$\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases} \quad \begin{cases} x_1 = -2/3 \\ x_2 = 2/3 \end{cases} \quad x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \beta = \begin{pmatrix} -2/3 \\ 2/3 \end{pmatrix}$$

$$\mathcal{J}f(x_1, x_2) = \begin{bmatrix} \frac{1}{3} + 2x_1 & -\frac{1}{3} \\ -\frac{1}{3} + x_2 & \frac{1}{3} + x_1 \end{bmatrix}$$

$$\mathcal{J}f(\alpha) = \mathcal{J}f(0,0) = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

per questo punto

$\det(\mathcal{J}f(\alpha)) = \frac{1}{9} - \frac{1}{9} = 0 \Rightarrow$ il metodo non converge
superlin-, ma al
massimo linearem.

$$\mathcal{J}f(\beta) = \mathcal{J}f(-\frac{2}{3}, \frac{2}{3}) = \begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$

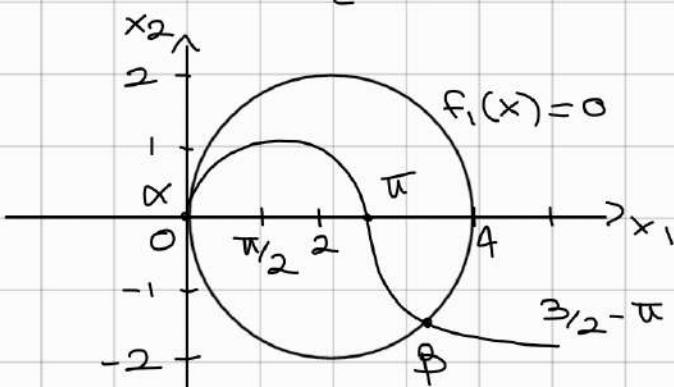
per questo punto

$\det(\mathcal{J}f(\beta)) = \frac{1}{3} + \frac{1}{9} = \frac{4}{9} \neq 0 \Rightarrow$ il metodo converge
superlinearem-

Esercizio

Determinare le soluzioni di $f(x_1, x_2) = [0 \ 0]$ e dire
per quali di loro il metodo di N.R. converge
in maniera superlineare -

$$f(x_1, x_2) = \begin{bmatrix} x_1 - \frac{(x_1^2 + x_2^2)}{4} \\ x_2 - \sin(x_1) \end{bmatrix}$$



$$\alpha = (0,0)$$

$$\beta = (\beta_1, \beta_2)$$

$$\beta_1 \in (\pi, \frac{3}{2}\pi)$$

$$\beta_2 \in (-1, 0)$$

$$\mathcal{J}f(x_1, x_2) = \begin{bmatrix} 1 - \frac{x_1}{2} & -\frac{x_2}{2} \\ -\cos(x_1) & 1 \end{bmatrix}$$

$$\mathcal{J}f(\alpha) = \mathcal{J}f(0,0) = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \Rightarrow \det(\mathcal{J}f(\alpha)) = 1 \neq 0$$

↓

N.R. converge
localm. in
modo superlin.

$$\det(\mathcal{J}f(x_1, x_2)) = 1 - \frac{x_1}{2} - \frac{x_2 \cos(x_1)}{2} = \frac{2 - x_1 - x_2 \cos(x_1)}{2} < 0$$

$$\det(\mathcal{J}f(\beta)) = \frac{2 - \beta_1 - \beta_2 \cos(\beta_1)}{2} < 0$$

$\beta_1 \in [\pi, \frac{3}{2}\pi]$
 $\beta_2 \in (-1, 0)$

$$\left. \begin{array}{l} 2 - \beta_1 < 0 \\ \cos(\beta_1) < 0 \end{array} \right\} \Rightarrow \det(\mathcal{J}f(\beta)) < 0 \neq 0 \Rightarrow \text{N.R. converge in maniera superlineare}$$

VARIANTI DEL METODO DI NEWTON-RAPHSON

Quando m cresce, il metodo di N.R. comincia ad essere troppo costoso:

- ① Si devono calcolare m^2 derivate per avere $\mathcal{J}f$
- ② Si deve risolvere un sistema lineare

NEWTON SEMPLIFICATO

Si calcola solo $\mathcal{J}f(x^{(0)})$ e si usa questo ad ogni passo:

$$x^{(n+1)} = x^{(n)} - \mathcal{J}f(x^{(0)})^{-1} f(x^{(n)})$$

Questo consente di calcolarsi la fattORIZZAZIONE LU di $Jf(x^{(0)})$ e utilizzarla per risolvere i sistemi lineari del generico passo con costo $O(m^2)$ anziché $O(m^3)$.

JACOBI - NEWTON

Si prende solo la diagonale di $Jf(x^{(n)})$

$$D(x^{(n)}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x^{(n)}) & & & \\ & \ddots & & \\ & & \ddots & \frac{\partial f_m}{\partial x_m}(x^{(n)}) \\ 0 & & 0 & \end{bmatrix}$$

$$x^{(n+1)} = x^{(n)} - D(x^{(n)})^{-1} f(x^{(n)})$$

Se num. delle derivate calcolate ad ogni passo è m e risolvere il sistema diagonale costa $O(m)$ ad ogni passo.

Per capire se questi metodi convergono localmente bisogna studiare $\varphi(J\phi(x))$:

$$\phi(x) = D(x^{(n)})^{-1} f(x^{(n)}) \text{ per JACOBI - NEWTON}$$

$$\phi(x) = \underbrace{Jf(x^{(0)})^{-1} f(x^{(n)})}_{\text{MAT. COSTANTE}} \text{ per NEWTON SEMPLIFICATO}$$

x CASA: Scrivere l'iterazione $\phi(x)$ di Jacobi - Newton nei due esempi di oggi.

ERRATA CORRIGE : CALCOLO FATTORIZZAZIONE LU
CON RIOTTING

GAUSS

$$A \rightsquigarrow \left[\begin{array}{c} \diagdown \\ \nabla \end{array} \right] = U$$

$$\pi A = LU$$

π si calcola tramite la composizione di permutazioni

$L = (l_{ij})_{ij}$ dove $-l_{ij}$ è il moltiplicatore al passo j per annullare l'elemento in posizione (i,j) $i > j$

La regola per calcolare L : quando si applica uno scambio di righe (non per colonne) si devono scambiare nello stesso modo gli elem. sotto la diagonale (non si toccano gli 1 sulla diagonale) di L calcolati fino a quel momento.

OSS Gli scambi fatti al primo passo non influiscono su L

ESEMPIO (calcolo di LU con riottino portante)

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 7 & 8 & 0 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix} \xrightarrow{\quad} \begin{bmatrix} 7 & 8 & 0 \\ 0 & 3/7 & 6 \\ 0 & 6/7 & 3 \end{bmatrix} \xrightarrow{\quad} \begin{bmatrix} 7 & 8 & 0 \\ 0 & 6/7 & 3 \\ 0 & 3/7 & 6 \end{bmatrix}$$

$$\xrightarrow{\quad} \begin{bmatrix} 7 & 8 & 0 \\ 0 & 6/7 & 3 \\ 0 & 0 & 9/14 \end{bmatrix} = U$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow L = \begin{bmatrix} 1 & 0 & 0 \\ 4/7 & 1 & 0 \\ 1/7 & 0 & 1 \end{bmatrix} \rightarrow$$

$$\rightarrow L = \begin{bmatrix} 1 & 0 & 0 \\ 1/7 & 1 & 0 \\ 4/7 & 0 & 1 \end{bmatrix} \rightarrow L = \begin{bmatrix} 1 & 0 & 0 \\ 1/7 & 1 & 0 \\ 4/7 & 1/2 & 1 \end{bmatrix}$$

$$A \rightarrow \pi_1 A \rightarrow H_1 \pi_1 A \rightarrow H_2 H_1 \pi_1 A$$

$$\pi_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, H_1 = \begin{bmatrix} 1 & 0 & 0 \\ -4/7 & 1 & 0 \\ -1/7 & 0 & 1 \end{bmatrix}, \pi_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\rightarrow H_2 \pi_2 H_1 \pi_1 A = U$$

$$H_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix}$$

$$H_2 \pi_2 H_1 \pi_1 A = U \rightarrow H_2 H_1 \pi_1 A = H_2^{-1} U$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{bmatrix}$$

$$\pi_2 H_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1/4 & 0 & 1 \\ -4/7 & 1 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1/4 & 1 & 0 \\ -4/7 & 0 & 1 \end{bmatrix}}_{\tilde{H}_1} \pi_2$$

$$\rightarrow \tilde{H}_1 H_2 \pi_i A = H_2^{-1} U \rightarrow \underbrace{\pi_i \tilde{H}_2 \pi_i}_\pi A = \underbrace{\tilde{H}_1^{-1} H_2^{-1}}_L U$$

N.B.: quando si scambiano righe al passo j si devono scambiare gli elem. di L sotto la diagonale, nelle prime $j-1$ colonne

OSS Stesso accorgimento quando si fa pivotting totale ($\pi_i \pi_j A \pi_i \pi_j = LU$)

CALCOLO DI AUTOVALORI E AUTOVETTORI

$A \in \mathbb{C}^{n \times n}$ e vogliamo trovare coppie (λ, x) $\lambda \in \mathbb{C}$, $x \neq 0$ $x \in \mathbb{C}^n$ verificando:

$$Ax = \lambda x \rightarrow \text{PROBLEMA NON LINEARE}$$



non ci sono metodi diretti, ma solo metodi iterativi:

Vedremo 2 metodi iterativi:

- 1) Per calcolare esattamente 1 coppia (λ, x)
- 2) Per calcolare tutte le coppie (λ, x)

DEFINIZIONE: Se x è autovettore $\Rightarrow \theta x \quad \forall \theta \in \mathbb{C} \quad \theta \neq 0$ è autovettore

METODO DELLE POTENZE

A diagonalizzabile

È un metodo iterativo per calcolare $(\lambda_1, v^{(1)})$ coppia autovettore-autovettore dove $|\lambda_1| \geq |\lambda_j| \forall j = 1, \dots, n$

Assumiamo che $A \in \mathbb{C}^{n \times n}$ abbia autovettori $\lambda_1, \dots, \lambda_n$ t.c. $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n| \geq 0$

e diamo $v^{(1)}, \dots, v^{(n)}$ autovettori associati

OSS Gli autovettori sono sempre in contatti con la loro molteplicità.

Per avere la proprietà che esistono n autovettori lin. indipendenti ci serve che A sia diagonalizzabile

IDEA

Prendiamo $z^{(0)} \in \mathbb{C}^n$ (scelto a caso)

$$z^{(0)} = \sum_{j=1}^n c_j v^{(j)} \quad c_j \in \mathbb{C}$$

$$Az^{(0)} = A \sum_{j=1}^n c_j v^{(j)} = \sum_{j=1}^n c_j Av^{(j)} = \sum_{j=1}^n c_j \lambda_j v^{(j)} \quad \lambda_j \in \mathbb{C}$$

$$A^2 z^{(0)} = A(Az^{(0)}) = A \sum_{j=1}^n c_j \lambda_j v^{(j)} = \sum_{j=1}^n c_j \lambda_j Av^{(j)} = \sum_{j=1}^n c_j \lambda_j^2 v^{(j)}$$

$$A^k z^{(0)} = \sum_{j=1}^n c_j \lambda_j^k v^{(j)} = \underline{c_1 \lambda_1^k v^{(1)}} + \dots + c_n \lambda_n^k v^{(n)}$$

QUANDO $k \rightarrow +\infty$ (SE $c_1 \neq 0$)

LA 1^a COMPONENTE SARÀ QUELLA DOMINANTE
(SE K È MOLTO GRANDE SI HA UN

VECTORE MOLTO SIMILE A UN MULTIPLO
SCALARE DI $v^{(1)}$ OVRNO CHE È UNA
BONNA APPROSSIMAZIONE DEL PRIMO
AUTOVETTORE)

A HERMITIANA

TH $A = A^H$, $A \in \mathbb{C}^{n \times n}$ con autovalori $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \in \mathbb{R}$
 v_1, \dots, v_n autovettori associati.

Supponiamo che $z^{(0)} \in \mathbb{C}^n$ e \mathbb{C}^n e t.e. $(v^{(1)})^H z^{(0)} \neq 0$

Allora il processo iterativo

$$\begin{cases} z^{(0)} \\ z^{(k+1)} = A z^{(k)} \end{cases} \quad k = 1, 2, 3, \dots$$

è tale per cui:

$$\lim_{k \rightarrow +\infty} \frac{z^{(k)}}{\|z^{(k)}\|} = \text{MULTIPLO SCURORE DI } v^{(1)}$$

COMPONENTE DOVE $h \in \{1, \dots, n\}$ TALE PER CUI $v_h^{(1)} \neq 0$
 IN POS. h DI
 $z^{(k)}$

$$\lim_{k \rightarrow +\infty} \frac{(z^{(k)})^H A z^{(k)}}{(z^{(k)})^H z^{(k)}} = \lambda_1$$

DEF $R_A(x) = \frac{x^H A x}{x^H x} =$ QUOTIENTE DI RAYLEIGH
 ASSOCIAZO ALLA MATRICE A

$$R : \mathbb{C}^n \rightarrow \mathbb{C}$$

DIM $A = A^H \Rightarrow A$ diagonalizzabile $\Rightarrow \exists v^{(1)}, \dots, v^{(n)}$ base di
 autovettori di A (base di \mathbb{C}^n)

Imoltre $V = [v^{(1)} | \dots | v^{(n)}]$ è unitaria, cioè $V^H V = I = VV^H$

$$z^{(0)} = \sum_{j=1}^n c_j v^{(j)} \text{ dove } c_1 \neq 0$$

$$(v^{(1)})^H z^{(0)} \neq 0 \Rightarrow c_1 \neq 0$$

$$V^H(z^{(0)}) = V^H\left(\sum_{j=1}^n c_j v^{(j)}\right) \Rightarrow V^H z^{(0)} = \sum_{j=1}^n c_j e_j \Rightarrow e_1 V^H z^{(0)} = c_1 \Leftrightarrow$$

$$\Leftrightarrow (\underbrace{V e_1}_1)^H z^{(0)} = e_1 = (V^{(1)})^H z^{(0)} \quad \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \rightarrow \text{PASO } j$$

1 COLONNA
RAT. V, ovvero $V^{(1)}$

$$A^K z^{(0)} = \sum_{j=1}^n c_j \lambda_j^K v^{(j)} = \lambda_1^K (c_1 v^{(1)} + \sum_{j=2}^n c_j \left(\frac{\lambda_j}{\lambda_1}\right)^K v^{(j)}) =$$

VECTORE CHE DIPENDE DAL PASSO K: $w^{(k)}$

$$= \lambda_1^K (c_1 v^{(1)} + w^{(k)}) \quad (*)$$

$$\text{con } \lim_{K \rightarrow +\infty} w^{(k)} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$|\lambda_1| > |\lambda_2|$$

Osservando la regola h-esima di (*):

$$\Rightarrow \lambda_1 \neq 0$$

$$z_h^{(k)} = \lambda_1^K \left(\underbrace{c_1}_{\neq 0} \cdot \underbrace{v_h^{(1)}}_{\neq 0} + \underbrace{w_h^{(k)}}_{0} \right) \Rightarrow \text{PER K ABBASTANZA GRANDE } z_h^{(k)} \neq 0$$

Quindi ha senso (è ben definito):

$$\lim_{K \rightarrow +\infty} \underbrace{\frac{z^{(k)}}{z_h^{(k)}}}_{\text{SCALARE}} = \lim_{K \rightarrow +\infty} \frac{\lambda_1^K (c_1 v^{(1)} + w^{(k)})}{\lambda_1^K (c_1 v_h^{(1)} + w_h^{(k)})} =$$

$$= \lim_{K \rightarrow +\infty} \frac{c_1 v^{(1)} + w^{(k)}}{c_1 v_h^{(1)} + w_h^{(k)}} = \frac{\cancel{c_1} v^{(1)}}{\cancel{c_1} v_h^{(1)}} =$$

$$= \underbrace{\left(\frac{1}{v_h^{(1)}} \right)}_{\in \mathbb{C}} \cdot \underbrace{v^{(1)}}_{\in \mathbb{C}^n} = \text{MULTPLI SCALARE DI } v^{(1)}$$

Infine, per quanto riguarda l'autovettore:

$$\begin{aligned}
 \lim_{k \rightarrow +\infty} \frac{(z^{(k)})^H A z^{(k)}}{(z^{(k)})^H z^{(k)}} &= \frac{\left(\begin{array}{c|c} z^{(k)} \\ \hline z_h^{(k)} \end{array}\right)^H A \left(\begin{array}{c|c} z^{(k)} \\ \hline z_h^{(k)} \end{array}\right)}{\left(\begin{array}{c|c} z^{(k)} \\ \hline z_h^{(k)} \end{array}\right)^H \cdot \left(\begin{array}{c|c} z^{(k)} \\ \hline z_h^{(k)} \end{array}\right)} = \\
 &= \frac{\left(\begin{array}{c|c} v^{(1)} \\ \hline v_h^{(1)} \end{array}\right)^H A \left(\begin{array}{c|c} v^{(1)} \\ \hline v_h^{(1)} \end{array}\right)}{\left(\begin{array}{c|c} v^{(1)} \\ \hline v_h^{(1)} \end{array}\right)^H \cdot \left(\begin{array}{c|c} v^{(1)} \\ \hline v_h^{(1)} \end{array}\right)} = \frac{(v^{(1)})^H A v^{(1)}}{(v^{(1)})^H v^{(1)}} = \\
 &= \frac{(v^{(1)})^H \lambda_1 v^{(1)}}{(v^{(1)})^H v^{(1)}} = \lambda_1 \frac{(v^{(1)})^H v^{(1)}}{(v^{(1)})^H v^{(1)}} = \lambda_1
 \end{aligned}$$

□

COMMENTO SULLE ASSUNZIONI FATTE:

i) $|\lambda_1| > |\lambda_2|$

ii) $z^{(0)}$ non sia ortogonale a $v^{(1)}$

iii) È necessario per evitare situazioni come:

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{Due autovettori } \pm i \quad (\lambda_1 = |\lambda_2|)$$

$$z^{(0)} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \rightarrow A z^{(0)} = \begin{bmatrix} -3 \\ 2 \end{bmatrix} \quad A^2 z^{(0)} = \begin{bmatrix} -2 \\ -3 \end{bmatrix} \quad A^3 z^{(0)} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

$$A^4 z^{(0)} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \dots \text{ma non si converge a nessuna delle due sol. } v^{(1)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, v^{(2)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Il metodo si blocca in un ciclo e non si trovano gli autovettori

ii) Per es. se si parte da un altro autovettore

$$\text{es. } z^{(0)} = v^{(2)} \text{ si ha } (v^{(1)})^H v^{(2)} = 0$$

$$Av^{(2)} = \lambda_2 v^{(2)}, A^2 v^{(2)} = \lambda_2^2 v^{(2)}, \dots$$

Dal punto di vista pratico, la (ii) è sempre vera quando prendo il punto di partenza $z^{(0)}$ in modo casuale dato che $c_i \neq 0$ con probabilità 1.

CONVERGENZA PER MATRICI NON HERMITIANE

Se osserviamo la dimostrazione precedente, si usa $A = A^H$ solo per:

- avere n autovettori indipendenti
- $c_i \neq 0 \Leftrightarrow (v^{(1)})^H z^{(0)} \neq 0$

Se si considera il caso $A \neq A^H$ si può avere lo stesso risultato di convergenza assumendo che A sia diagonalizzabile (\Rightarrow abbiamo n autovett. indip.) e che $(U^{(1)})^H z^{(0)} \neq 0$ dove $U^{(1)}$ è un autovett. sx associato a λ_1 :

$$(U^{(1)})^H A = \bar{\lambda}_1 (U^{(1)})^H \quad (\text{è equiv. a dire che } U^{(1)} \text{ è autovett. di } A^H)$$

Nell'implementazione pratica del metodo delle potenze si divide ad ogni iterazione per la norma del vettore in modo da evitare problemi di underflow e overflow.

```

 $z^{(0)}$  DATO
for k=1, 2, ...
     $y^{(k)} = A z^{(k-1)}$ 
     $z^{(k)} = \frac{y^{(k)}}{\|y^{(k)}\|}$ 
end

```

Criterio d'arresto:

$$|\rho(z^{(k)}) - \rho(z^{(k-1)})| < \text{TOL}$$

VARIANTI: Se volessimo l'autovettore più piccolo in modulo? Anche λ_n

IDEA A^{-1} ha autovettori $|\frac{1}{\lambda_n}| \geq |\frac{1}{\lambda_{n-1}}| \geq \dots \geq \frac{1}{|\lambda_1|}$

\Rightarrow si può applicare il metodo delle potenze ad A^{-1} assumendo che $\lambda_n < \lambda_{n-1}$

N.B.: questa volta, ad ogni passo si deve risolvere il sistema lineare $Ay^{(k)} = z^{(k-1)}$ (equivalente a calcolare $A^{-1}z^{(k-1)}$, ma non si calcola mai l'inversa, si risolve il sistema).

Più in generale, si può calcolare l'autovettore più vicino a $\sigma \in \mathbb{C}$.

Per farlo bisogna applicare il metodo delle potenze alla matrice $(A - \sigma I)^{-1}$ poiché ha autovettori $(\frac{1}{\lambda - \sigma})$. L'autovettore dominante sarà quello che

rende $|\lambda_j - \sigma|$ più piccola ovvero il λ_j più vicino a σ .

(*) $z^{(k)} = (A - \sigma I)^{-1} z^{(k-1)}$ risolvere un sist. lineare

Una buona idea è calcolare $\pi(A - \sigma I) = LU$ all'inizio e utilizzarla in tutte le iterazioni per (*).

$$Ax = \lambda x \quad \lambda \in \mathbb{C}, x \in \mathbb{C}^n \setminus \{0\}, A \in \mathbb{C}^{n \times n}$$

METODO DELLE POTENZE INVERSE (versione com σ)

Per il metodo delle potenze, scelgo $\sigma \in \mathbb{C}$ e posso cercare (x, x) t.c. λ si il più vicino a σ .

$z^{(0)}$ scelto casualmente

Calcolare fatt. LU con pivoting di $A - \sigma I$ all'inizio

$$\pi(A - \sigma I) = LU \leftarrow O(n^3)$$

- for $k = 1, 2, 3 \dots$

$$y^{(k)} = (A - \sigma I)^{-1} z^{(k-1)} \text{ (sfruttando } \pi, L, U)$$

$$z^{(k)} = y^{(k)} / \|y^{(k)}\|_2$$

end

OSS La versione del metodo delle potenze per approssimare l'autovalore dominante costa $O(n^2 \cdot \tau)$ ($\tau = \#$ iterazioni del metodo) -
mentre la versione com σ costa $O(n^3 + \tau \cdot n^2)$
(**METODO DELLE POTENZE INVERSE**)

FATTORIZZAZIONE QR

Basta $A \in \mathbb{C}^{n \times n}$ si dice fatt. QR di A una coppia di matrici (Q, R) t.c.:

$$A = Q \cdot R$$

$$Q^H Q = I \quad (\text{Q unitaria } n \times n)$$

$$R \text{ triang. sup. } n \times n \quad \begin{bmatrix} * & \cdots & * \\ 0 & \ddots & * \\ \vdots & \ddots & \vdots \end{bmatrix}$$

OSS Se si dovesse risolvere $Ax = b$ e conoscessimo Q ed R , potremmo agire come segue:

$$Ax = b \Leftrightarrow QRx = b \Leftrightarrow Rx = \overbrace{Q^H b}^X$$

e poi risolvere $Rx = y$

(sistema triangolare $O(n^2)$)

OSS Il costo della LU è $O(\frac{2}{3}n^3)$

Il costo della QR è $O(\frac{4}{3}n^3)$, il doppio

(per questo in generale in MATLAB è implement. LU)

TH Data $A \in \mathbb{C}^{n \times n}$, una sua fatt. QR esiste sempre
e non è mai unica

$$\begin{aligned} & A = QR \\ & \quad \underbrace{QD \cdot R^{-1}R}_{\substack{\text{E ANCORA} \\ \downarrow \text{UNITARIA}}} \\ & \quad D = \begin{bmatrix} * & & 0 \\ \vdots & \ddots & 0 \\ 0 & \cdots & * \end{bmatrix}, |d_{ii}| = 1 \\ & \quad d_{ii} \in \mathbb{C} \end{aligned}$$

COME CALCOLARE UNA FATT. QR

OSS Prodotto e inverso di matrici unitarie sono ancora unitarie

DEF Dato un vettore $u \in \mathbb{C}^n$ si dice **matrice di Householder** associata ad u , la matrice:

$$H(u) = I - \frac{2uu^H}{\|u\|_2^2} = I - \text{costante} \cdot \underbrace{\begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}}_{\text{MAT. } \forall k=1}$$

$$\begin{aligned} & H(u)^H = I^H - \frac{2}{\|u\|_2^2} \cdot (uu^H)^H = I - \frac{2}{\|u\|_2^2} \cdot uu^H = H(u) \end{aligned}$$

$$H(U)^H H(U) = H(U)^2 = I + \frac{4\|U\|_2^2}{\|U\|_2^4} UU^H - \frac{2UU^H}{\|U\|_2^2} - \frac{2UU^H}{\|U\|_2^2}$$

$$= I$$

$H(U)$ è hermitiana e unitaria, in particolare
è anche uguale alla sua inversa ($H(U) = H(U)^{-1}$)

LEMMA Dato $v \in \mathbb{R}^n$, se scegliamo $u = v \pm \|v\|_2 \cdot e_i =$

$$= \begin{bmatrix} v_1 \pm \|v\|_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\text{si ha che } H(u)v = \left(I - \frac{2uu^H}{\|u\|_2^2} \right)v = \underbrace{\begin{bmatrix} \mp\|v\|_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\text{MULTIPLIO DELLA BASE CANONICA}}$$

MULTIPLIO DELLA
BASE CANONICA

IDEA PER TROVARE LA FATT. QR'.

$A = [a_1 | \dots | a_n]$, al primo passo costruiamo $H(U) = H_1$,
di Householder t.c. $H_1 a_1 = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

$$H_1 A = [H_1 a_1 | H_1 a_2 | \dots | H_1 a_n] = \begin{bmatrix} * & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix}$$

Al secondo passo devo fare una cosa analoga,
ma limitandomi al vett. $\boxed{A_2}$

$$H_1 A = \begin{bmatrix} * & * & \dots & * \\ 0 & \boxed{A_2} \\ \vdots & & & \\ 0 & & & \end{bmatrix} \quad \text{scelgo } \tilde{H}_2 \in \mathbb{C}^{(n-1) \times (n-1)} \text{ t.c. } \tilde{H}_2 A_2 = \begin{bmatrix} * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix}$$

Considero $H_2 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \tilde{H}_2 & & \\ \vdots & & & \\ 0 & & & \end{bmatrix}$

$$H_2 H_1 A = \begin{bmatrix} * & \cdots & * \\ 0 & \boxed{\tilde{H}_2 A_2} \\ \vdots & & \\ 0 & & \end{bmatrix} = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & * & \cdots & * \end{bmatrix}$$

All passo k si ha:

$$H_k = \begin{bmatrix} I_{k-1} & 0 \\ 0 & \boxed{\tilde{H}_k} \end{bmatrix}$$

$$H_n H_{n-1} \cdots H_1 A = \begin{bmatrix} * & \cdots & * & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & * & \cdots & * \end{bmatrix}$$

Dopo $n-1$ passi:

$$H_{n-1} H_{n-2} \cdots H_1 A = \begin{bmatrix} * & * \\ 0 & \ddots & \vdots \\ & & * \end{bmatrix} = R$$

$$H_{n-1} \cdots H_1 A = R \Rightarrow A = \underbrace{H_1 H_2 \cdots H_{n-1}}_Q R = QR$$

ESEMPIO

$$A = \begin{bmatrix} 12 & -51 & 4 \\ 6 & 16 & -68 \\ -4 & 24 & -41 \end{bmatrix}$$

$$\alpha_1 = \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} \quad \|\alpha_1\|_2 = 14 \quad U_1 = \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} - \begin{bmatrix} 14 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \\ -4 \end{bmatrix}$$

$$H_1 = I - 2 \frac{U_1 U_1^T}{\|U_1\|_2^2} \rightarrow H_1 A = \begin{bmatrix} 14 & 21 & -14 \\ 0 & -48 & -14 \\ 0 & 168 & -77 \end{bmatrix}$$

$$\alpha_2 = \begin{bmatrix} -49 \\ 168 \end{bmatrix} \quad \|\alpha_2\|_2 = 145 \quad U_2 = \begin{bmatrix} -49 \\ 168 \end{bmatrix} - \begin{bmatrix} 145 \\ 0 \end{bmatrix} = \begin{bmatrix} -224 \\ 168 \end{bmatrix}$$

$$\tilde{H}_2 = I - \frac{2U_2 U_2^H}{\|U_2\|_2^2} \rightarrow H_2 \cdot H_A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \tilde{H}_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & -35 \end{bmatrix} = R$$

$$Q = H_1 \cdot H_2$$

Oss Il costo rimane cubico anche se si fanno $n-1$ prodotti di matrici perché le mat. di Householder hanno la struttura IDENTITÀ + RANGO 1 che consente di moltiplicarne 2 con costo quadrattico.

Se abbiamo A possiamo trovare $A = Q \cdot R$ in $O(n^3)$

METODO QR PER GLI AUTONALORI

$A \in \mathbb{C}^{n \times n}$, $S \in \mathbb{C}^{n \times n}$ invertibile, $B = S^{-1}AS$

B ha gli stessi autonvalori di A ed autovettori $S^{-1}v$ dove v è autovettore di A.

Se conosco gli autovettori e gli autovettori di B e conosco S, allora posso conoscere facilmente autovettori e autonvalori di A.

IDEA

Fare una successione di trasformazioni per similitudine per trovare una matrice T di cui riesco a trovare facilmente autonvalori ed autovettori.
Nel caso del metodo QR si cerca T triang.-sup.

$$A \xrightarrow{\quad S_1^{-1}AS_1 \quad} \xrightarrow{\quad S_2^{-1}A_2S_2 \quad} \xrightarrow{\quad S_3^{-1}A_3S_3 \quad} \cdots \rightarrow T$$

$\underline{A_1} \quad \underline{A_2} \quad \underline{A_3} \quad \underline{A_4}$

Gli autonvalori sono gli elem. diag. di T.

1' PROBLEMA: le trasf. per similitudine coinvolgono le inverse, ovvero sistemi lineari
 \Rightarrow problemi di condizionamento se S_j mancano parti colossi proprietà

SOLUZIONE: si fanno trasf. per similitudine con matrici unitarie

$$\underbrace{A_1}_{A_1} \rightarrow \underbrace{Q_1^H A_1 Q_1}_{A_2} \rightarrow \underbrace{Q_2^H A_2 Q_2}_{A_3} \rightarrow \underbrace{Q_3^H A_3 Q_3}_{A_4} \rightarrow \dots \rightarrow T$$

La successione / iterazione alla base del metodo QR è:

$$A_1 = A$$

for $K=1, 2, 3, \dots$

calcolo $A_K = Q_K R_K$ (cioè calcolo Q_K e R_K)

calcolo $A_{K+1} = R_K Q_K$

end

È una trasf. per similitudine? Sí

$$A_{K+1} = R_K Q_K = \underbrace{Q_K^H Q_K}_{I} \underbrace{R_K R_K^H}_{A_K} Q_K = Q_K^H A_K Q_K$$

$\Rightarrow A_{K+1}$ è simile ad $A_K \Rightarrow A_{K+1}$ è simile ad A

$$\begin{aligned} A_{K+1} &= Q_K^H A_K Q_K = Q_K^H Q_{K-1}^H A_{K-1} Q_{K-1} Q_K = Q_K^H \dots Q_1^H A Q_1 \dots Q_K \\ &= \tilde{Q}^H \tilde{A} \tilde{Q}, \tilde{Q} = Q_1 \dots Q_K \end{aligned}$$

TH $A \in \mathbb{R}^{n \times n}$ se $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$

allora $\lim_{K \rightarrow +\infty} A_K = T = \begin{bmatrix} t_{11} & \dots & t_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & t_{nn} \end{bmatrix}$

se invece gli autovalori sono distinti, ma possibilmente con lo stesso modulo

allora $\lim_{K \rightarrow +\infty} A_K = T = \begin{bmatrix} T_{11} & \dots & T_{1S} \\ \vdots & \ddots & \vdots \\ T_{SS} & \dots & T_{SS} \end{bmatrix}$ T_{ij} è 1×1 o 2×2

T rett. triang. a blocchi

Oss Nel 2° caso gli autovalori di T (e quindi di A) sono dati dai blocchi 1×1 (ovvero elementi sulla diag-) e dagli autovalori dei blocchi diag. 2×2 .

es: $T = \begin{bmatrix} 3 & * & * \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \rightarrow$ gli autovalori sono 3 e gli autoval. della matrice

Il metodo QR base consiste nel generare la successione $\{A_{K+1}\} = \{R_K Q_K\}$ e formarsi quando gli elem. sotto la diag. di A_K sono vicini a zero. Quindi trovarsi gli autoval. e gli autovett. di T parte triang. sup. di A_K

$$A_K = \begin{bmatrix} * & \dots & * \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix} \rightarrow T = \begin{bmatrix} * & \dots & * \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix} \text{ e calcolo autoval. e autovett. di } T \\ (\lambda_T, v_T)$$

Uso λ_T come appross. degli autoval. di A e
 \tilde{Q}_{V_T} come appross. degli autovett. di A dove
 $\tilde{Q} = Q_1 \dots Q_K$.

$$A_1 = A$$

for $j = 1, \dots, K$

$$A_j = Q_j R_j$$

$$A_{j+1} = R_j Q_j$$

end

T = parte triang. sup. di A_K

Calcolo autovett. V di T ed i suoi autoval.

Restituisco come autovett. di A la mat. $Q_1 \dots Q_K V = \tilde{Q} V$
 " " " autovalsi " " quelli di T

COME SI CALCOLANO GLI AUTOVETTORI DI UNA MAT. TRIANGOLARE CON AUTONALORI DISTINTI (NON NECESS. DISTINTI IN MODULO)

$$T = \begin{bmatrix} t_{11} & \dots & t_{1n} \\ 0 & \ddots & \vdots \\ 0 & \dots & t_{nn} \end{bmatrix} \quad t_{ii} \neq t_{jj} \quad \forall i \neq j$$

Se prendo $\lambda_j = t_{jj}$ so che l'autovett. associato è v_j
 che risolve $(T - \lambda_j I)v_j = 0$

$$\begin{bmatrix} t_{11} & t_{12} & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{j-1,j} & t_{jj} & * & \dots & * \\ 0 & \ddots & \ddots & \ddots & \vdots \\ * & \ddots & \ddots & \ddots & * \end{bmatrix} = \begin{bmatrix} T_1 & 0 & * \\ 0 & T_2 & 0 \end{bmatrix} \quad T_1, T_2 \text{ triang. sup.}$$

$$\begin{bmatrix} T_1 & 0 & * \\ 0 & T_2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \bar{x} \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \bar{0} \\ 0 \end{bmatrix} \quad \begin{aligned} x_2 &= 0 \\ \bar{x} &= 1 \\ T_1 \cdot x_1 &= -\bar{x} \cdot \bar{z} = -z \end{aligned}$$

$x_1 = T_1^{-1} z$ risolve un sistema
triang.-com la mat.
 T_1 che risulta
invertibile

OSS Per avere tutti gli autovalori devo risolvere n
sistemi lineari triangolari $\rightarrow O(n^3)$

$$Ax = \lambda x \quad A \in \mathbb{C}^{n \times n} \quad (\lambda_j, x_j) \quad \lambda_j \in \mathbb{C}, x_j \in \mathbb{C}^n$$

IDEA DEL METODO QR:

Si genera una successione A_k t.c. A_k è simile ad A $\forall k \in \mathbb{N}$, più precisamente $\exists \tilde{Q}_k$ unitaria t.c.

$$A_k = \tilde{Q}_k^H A \tilde{Q}_k$$

A_k è ben approssimabile con una mat. T triang. sup. o triang. sup. a blocchi:

\Rightarrow gli autovettori sono facili da trovare
(per T triang. sup. sono gli elem. diag., per T triang. sup. a blocchi bisogna trovare gli autovol. per ogni blocco)

Mentre gli autovettori di T si trovano con la procedura della lessione di ieri in $O(n^2)$.

$$A_1 = A$$

[for $k=1, 2, 3 \dots$

$$A_k = Q_k R_k \rightarrow O(n^3)$$

$$A_{k+1} = R_k Q_k \rightarrow O(n^3)$$

end

Si ottiene $T \approx A_K$ K suff. grande

Si calcolano (x_j, v_j) autovettori di T $\rightarrow O(n^2)$

Si restituiscono $\lambda_j, \tilde{Q} v_j$ $\tilde{Q} = Q_1 \dots Q_K$

QUESTIONI APerte

- i) Come calcolare la conveg. della succ. A_k ad una mat. triang. sup. a blocchi (con blocchi al max 2×2) -
Implm. su software convergono in $O(n)$ passi.
- ii) Gestione di autoveloci multipli e/o con lo stesso modulo -
- iii) Renderre il costo dell' iterazione $O(n^2)$ cosi che tutto l' algoritmo costi $O(n^3)$ -
- iv) **DEF** $A \in \mathbb{C}^{n \times n}$ si dice in forma di Hessenberg superiore se:

$$A = \begin{bmatrix} x & \cdots & x \\ x & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x \end{bmatrix}$$

se $i > j+1 \Rightarrow a_{ij} = 0$

(gli elem. sotto la prima sottodiag. sono = 0)

FATTO Se fatt. qe di una mat. in forma di Hessenberg sup. si può calcolare in $O(n^2)$

$$A = \begin{bmatrix} x & \cdots & x \\ x & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x \end{bmatrix} \rightarrow \begin{bmatrix} H_1 \\ I \end{bmatrix} \begin{bmatrix} x & \cdots & x \\ x & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x \end{bmatrix} = \begin{bmatrix} x & x & \cdots & x \\ 0 & x & \cdots & x \\ \vdots & \ddots & \cdots & x \\ 0 & 0 & \cdots & x \end{bmatrix}$$

Le uniche righe che cambiamo sono la 1 e la 2

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \quad A_1 \in \mathbb{C}^{2 \times n}, A_2 \in \mathbb{C}^{(n-2) \times n}$$

$$H_1 = \begin{bmatrix} \tilde{H}_1 & \\ & I \end{bmatrix} \quad \tilde{H}_1 \in \mathbb{C}^{2 \times 2}$$

$$H, A = \begin{bmatrix} \boxed{H, A_1} \\ \boxed{A_2} \end{bmatrix}$$

Dove solo calcolare

$$\tilde{H}, \cdot A_1 = \square \quad \rightarrow O(n)$$

Al passo k : $A =$

$$\begin{bmatrix} \triangle & x & \dots & x \\ \vdots & \vdots & & \vdots \\ x & x & & | \\ \ddots & \ddots & \ddots & x \end{bmatrix}$$

$$H_k = \begin{bmatrix} \boxed{H} \\ \boxed{H_k} \\ \boxed{H} \end{bmatrix}$$

H_k modifica queste due righe

$$A^{(k)} = \begin{bmatrix} \boxed{A_1^{(k)}} \\ \boxed{A_2^{(k)}} \\ \boxed{A_3^{(k)}} \end{bmatrix}$$

$$H_k \cdot A^{(k)} = \begin{bmatrix} \boxed{A_1^{(k)}} \\ \boxed{H_k A_2^{(k)}} \\ \boxed{A_3^{(k)}} \end{bmatrix}$$

Dunque: $(n-1)$ passi, ciascuno costa $O(n)$ perché
ciascuno comporta il prodotto di una mat- 2×2
con una $2 \times n$

(REMINDER: PROD. DI TAV. $m \times n$ $n \times p \rightarrow O(m \cdot n \cdot p)$)

FATTO Se $[Q, Q]$ è la fatt. QR di una mat.
in forma di Hessenberg allora:

- ① Calcolare $R \cdot Q$ costa $O(n^2)$
- ② $R \cdot Q$ è di nuovo in forma di Hessenberg
superiore

Se A_k è in forma di Hessenberg allora:

$A_{k+1} = R_k Q_k$ è ancora in forma di Hess. e lo
ottengo in $O(n^2)$

Cosa faccio se A non è di Hess?

TH $A \in \mathbb{C}^{n \times n}$ allora $\exists Q$ unitaria t.c.

$$Q^H A Q = \text{Hess}(A) \text{ dove } \text{Hess}(A) = \begin{bmatrix} * & \cdots & * \\ * & \ddots & \vdots \\ 0 & \cdots & ** \end{bmatrix}$$

DIM

Analogamente alla proc. per la fatt. QR si considera una mat. di Householder \tilde{H}_1 t.c.

$$A = \left[\begin{array}{c|cc} * & & \\ \hline 0_1 & * \end{array} \right] \quad 0_1 \in \mathbb{C}^{n-1}, \tilde{H}_1 0_1 = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$H_1 = \left[\begin{array}{c|c} I & 0 \\ \hline 0 & \tilde{H}_1 \end{array} \right] \rightarrow H_1 A = \begin{bmatrix} * & * \\ * & * \\ 0 & \vdots \\ \vdots & 0 \\ 0 & * \end{bmatrix}$$

$$H_1 A H_1^{-1} = H_1 A \tilde{H}_1 = \begin{bmatrix} * & * \\ * & * \\ 0 & \vdots \\ \vdots & * \\ 0 & * \end{bmatrix} \left[\begin{array}{c|c} I & 0 \\ \hline 0 & \tilde{H}_1 \end{array} \right] = \begin{bmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ 0 & \vdots & \ddots & \vdots \\ \vdots & 0 & \cdots & * \end{bmatrix} = A^{(2)}$$

A AL PASSO 2 ↑
(2)

Al passo k:

$$H_k = \left[\begin{array}{c|c} I & \tilde{H}_k \\ \hline \tilde{H}_k & \tilde{H}_k \end{array} \right] \rightarrow H_k A \tilde{H}_k = \begin{bmatrix} * & * & \cdots & * & \cdots & * \\ * & * & \cdots & * & \cdots & * \\ 0 & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & * & * & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & * & * & \cdots & \cdots & * \end{bmatrix}$$

Dopo $n-1$ passi ottengo una mat. in forma di Hess. con costo finale $O(n^3)$

□

METODO QR PER GLI AUTOVETTORI DI $A \in \mathbb{C}^{n \times n}$

- ① Si calcola $\text{Hess}(A)$: $A = \hat{Q} \text{Hess}(A) \hat{Q}^H \rightarrow O(n^3)$
- ② $\int A_1 = \text{Hess}(A)$
 $\left\{ \begin{array}{l} A_{k+1} = R_k Q_k \quad \text{dove } Q_k R_k = A_k \\ \end{array} \right.$
- ③ Si ottiene T triang.-sup. a blocchi e si calcolano (λ_j, v_j) autoval- ed autovett. di T
- ④ Si restituiscono λ_j e $\hat{Q} \hat{Q}^H v_j$

$\hat{Q} = Q_1 \dots Q_n$

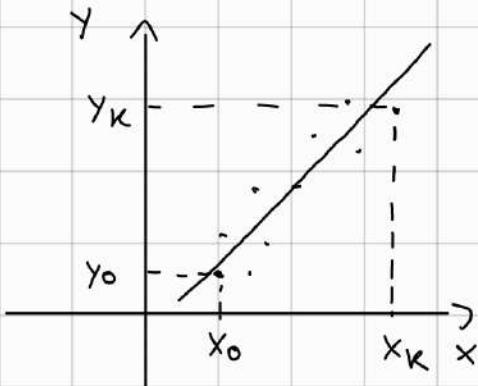
↓
TENNE CONTO DEL
FATTO CHE SI È
PASSATI ALLA
FORMA DI HESS.

APPROXIMAZIONE DI FUNZIONI

Supponiamo di conoscere in maniera empirica una funzione in alcuni punti:

$$f: \mathbb{R} \rightarrow \mathbb{R}$$

$(x_0, y_0), \dots, (x_k, y_k)$ ovvero $(k+1)$ punti, $k \in \mathbb{N}$



IDEA: Ho dei dati, cerco una funzione "semplice" che passa vicino ai dati

REGRESSIONE (o FITT)

Nell'esempio cerco una retta che passa vicino ai punti (REGRESSIONE LINEARE).

con $c_j \in \mathbb{R}$

In genere si sceglie $\Phi(x) \approx f(x)$ t.c.

$$\Phi(x) = \sum_{j=1}^s c_j e_j(x)$$

$e_0(x), \dots, e_s(x)$ funzioni di base (fornite da chi ci dà il modello)

PROBLEMA DELLA REGRESSIONE

Una volta che mi vengono forniti i dati (x_j, y_j) con $j = 0, \dots, k$ e le funzioni di base, si devono trovare i coeff. c_j (che definiscono $\Phi(x)$) t.c. $\Phi(x)$ passa "il più vicino possibile" ai dati.

↳ noi vedremo nel senso dei

MINIMI QUADRATI

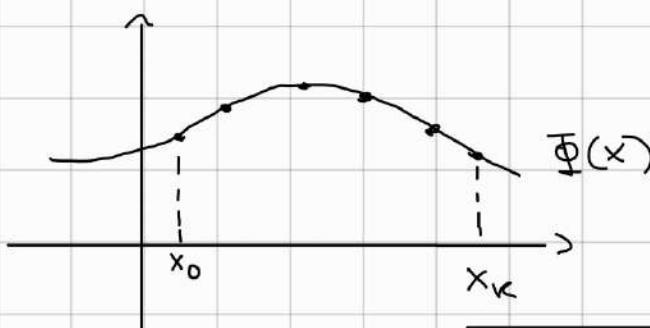
ESERCIZIO

Nel caso della regressione lineare si ha:

$$\varrho_0(x) = 1, \varrho_1(x) = x \Rightarrow \Phi(x) = \varrho_0 + \varrho_1 x$$

Se volessi $\Phi(x)$ polinomio di grado k , potrai considerare $\varrho_0(x) = 1, \varrho_1(x) = x, \dots, \varrho_k(x) = x^k \Rightarrow \Phi(x) = \sum_{j=0}^k \varrho_j x^j$

INTERPOLAZIONE



Chiediamo che $\boxed{\Phi(x_j) = y_j \quad \forall j = 0, \dots, k}$

In questo caso le qtà x_0, x_1, \dots, x_k si dicono nodi dell' interpolazione

INTERPOLAZIONE POLINOMIALE (o parabolica)

$$\Phi(x) = \sum_{j=0}^k \varrho_j x^j$$

OSS Nell' interpolazione polinomiale il num. di date ($k+1$) coincide con il num. di funzioni di base ($\{1, x, \dots, x^k\}$) -

In particolare, se ho $k+1$ punti li interpolo con un polinomio di grado al più k .

NOTAZIONE: Chiamiamo $P_k(x)$ il polinomio interpolante di grado al più k .

Il problema di interpolazione porta a considerare le equazioni:

$$\rightarrow P_k(x) = \sum_{j=0}^k a_j x^j$$

$$P_k(x_j) = y_j \quad \forall j = 0, \dots, k \quad \text{CONDIZIONE DI INTERPOLAZIONE}$$

$$a_0 + a_1 x_j + a_2 x_j^2 + \dots + a_k x_j^k = y_j \quad \forall j = 0, \dots, k$$

$$\begin{cases} a_0 + a_1 x_0 + \dots + a_k x_0^k = y_0 \\ \vdots \\ a_0 + a_1 x_k + \dots + a_k x_k^k = y_k \end{cases} \quad \Leftrightarrow V_a = y \quad \text{sistema lineare}$$

È un sist. lin. di $(k+1)$

eq. e $(k+1)$ imcognite.

$$V = \begin{bmatrix} 1 & x_0 & \dots & x_0^k \\ 1 & x_1 & \dots & x_1^k \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_k & \dots & x_k^k \end{bmatrix}$$

MATRICE DI VANDERMONDE

$$a = \begin{bmatrix} a_0 \\ \vdots \\ a_k \end{bmatrix}, \quad y = \begin{bmatrix} y_0 \\ \vdots \\ y_k \end{bmatrix}$$

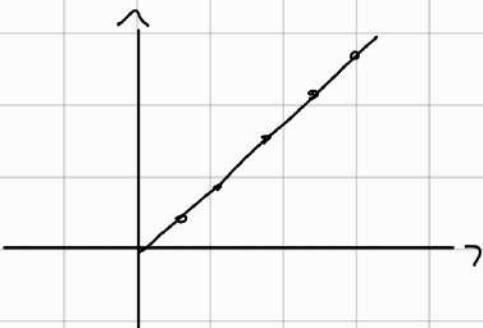
OSS Se la soluzione esiste ed è unica, allora il polinomio di interpolazione $P_k(x)$ esiste ed è unico

TH $\det(V) = \prod_{\substack{k \geq i > j \geq 0}} (x_i - x_j)$

COROLARIO Se $x_i \neq x_j$ ogni volta che $i \neq j$ allora esiste ed è unico $P_k(x)$ che interpola $(x_0, y_0), \dots, (x_k, y_k)$ indipendentemente dai valori y_0, \dots, y_k .

OSS $P_k(x)$ può avere grado minore di k , ad esempio succede se risolvendo il sist. lineare viene $a_k = 0$

ESEMPIO



5 modi per l'interpolazione
quindi si cerca $P_4(x)$
e $P_4(x)$ sarà della forma $a_0 + a_1 x$

PROBLEMA: Calcolare $P_k(x)$ risolvendo il sistema
 $Va = y$ è problematico dal punto di vista compu-
tazionale perché il numero di condizionamento
 $\|V\|_2 \cdot \|V^{-1}\|_2$ della mat. di Vandermonde cresce
in maniera esponenziale con il num. di punti
considerati, già per $k > 10$ punti i risultati non
sono più affidabili.

ALTERNATIVE PER IL CALCOLO DI $P_k(x)$

$$P_k(x) = \sum_{j=0}^n a_j x^j = \sum_{j=0}^k c_j q_j(x) \quad \text{dove } q_j(x) \text{ sono polinomi} \\ \text{di grado } \leq k$$

INTERPOLAZIONE DI LAGRANGE

$$q_j(x) := l_j(x) = \frac{(x - x_0) \dots (x - x_{j-1})(x - x_{j+1}) \dots (x - x_k)}{(x_j - x_0) \dots (x_j - x_{j-1})(x_j - x_{j+1}) \dots (x_j - x_k)} = \\ = \prod_{i \neq j} \frac{(x - x_i)}{(x_j - x_i)} = \frac{\text{POLINOMIO DI GRADO } k}{\text{SCALARE}}$$

$$P_k(x) = \sum_{j=0}^k a_j x^j = \sum_{j=0}^k c_j \underbrace{l_j(x)}_{y_j}$$

$$l_j(x_i) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$$y_j = P_k(x_j) = \sum_{i=0}^k c_i l_i(x_j) = c_j l_j(x_j) = c_j$$

ESEMPIO

	x_0	x_1	x_2	x_3
x	0	1	2	-1
y	1	1	3	3
	y_0	y_1	y_2	y_3

$$l_0(x) = \frac{(x-1)(x-2)(x+1)}{(-1) \cdot (-2) \cdot 1}$$

\uparrow \uparrow \uparrow
 $0-1$ $0-2$ $0+1$

$$l_1(x) = \frac{(x-0)(x-2)(x+1)}{1 \cdot -1 \cdot 2}$$

$$l_2(x) = \frac{(x-0)(x-1)(x+1)}{2 \cdot 1 \cdot 3}$$

$$l_3(x) = \frac{(x-0)(x-1)(x-2)}{(-1)(-2)(-3)}$$

$$\begin{aligned}
 P_3(x) &= y_0 \cdot l_0(x) + y_1 \cdot l_1(x) + y_2 \cdot l_2(x) + y_3 \cdot l_3(x) = \\
 &= l_0(x) + l_1(x) + 3l_2(x) + 3l_3(x) \\
 &\stackrel{*}{=} (\text{POLINOMIO SCONTRATO NELL BASE DI LAGRANGE}) \\
 &= x^2 - x + 1 \quad (\text{NELL BASE DEI MONOMI})
 \end{aligned}$$

OSS Nella base di Lagrange ogni funzione di base è un polinomio di grado k ($l_j(x)$ è di grado $k \forall j$) mentre nella base dei monomi no $(1, x, \dots, x^k)$

INTERPOLAZIONE DI NEWTON

Si vuole formire una base in maniera tale che se ho calcolato $P_k(x)$ e cerco $P_{k+1}(x)$ (ho appena fatto una misurazione) posso reutilizzare sia la base che i coeff. ottenuti quando ho calcolato $P_k(x)$.

Se ho x_0, \dots, x_k modi di interpolazione la base di Newton è $\{1, x - x_0, (x - x_0)(x - x_1), \dots, (x - x_0) \dots (x - x_{k-1})\}$

$$n_j(x) = \prod_{i=0}^j (x - x_i) \quad \{n_0, n_1^{(1)}, \dots, n_k^{(k)}\}$$

La cosa più complicata è trovare i coeff.

$$c_j: P_k(x) = \sum_{j=0}^k c_j n_j$$

DEF Data $f: I \rightarrow \mathbb{R}$, $I \subseteq \mathbb{R}$, $x_0, \dots, x_{k-1} \in I$

$$x_i \neq x_j \text{ se } i \neq j$$

Si definisce per ricorrenza la DIFFERENZA

DIVISA di ordine k nei punti x_0, \dots, x_{k-1} :

$$\begin{cases} f(x) & \text{SE } k=0 \\ f[x_0, x] = \frac{f(x) - f(x_0)}{x - x_0} & \text{SE } k=1 \end{cases}$$

$$f[x_0, \dots, x_{k-1}, x] = \frac{f[x_0, \dots, x_{k-2}, x] - f[x_0, \dots, x_{k-2}, x_{k-1}]}{x - x_{k-1}}$$

$k \geq 1$

ESEMPIO

D.D. di ordine 0 $\rightsquigarrow f(x)$

$$\begin{array}{c} \text{`` `` `` } \\ \text{1 } \end{array} \rightsquigarrow \frac{f(x) - f(x_0)}{x - x_0} \approx f'(x_0)$$

$$\begin{array}{c} \text{`` `` `` } \\ \text{2 } \end{array} \rightsquigarrow \frac{f[x_0, x] - f[x_0, x_1]}{x - x_1} \approx f''$$

PROPRIETÀ

i) A permutazione i_0, \dots, i_k di $\{0, 1, \dots, k\}$ si ha:

$$f[x_0, \dots, x_k] = f[x_{i_0}, x_{i_1}, \dots, x_{i_k}]$$

ESEMPIO

$$f[x_0, x] = \frac{f(x) - f(x_0)}{x - x_0} \quad f[x, x_0] = \frac{f(x_0) - f(x)}{x_0 - x}$$

ii) Sebbene $f[x_0, \dots, x_{k-1}, x]$ non è definita in x_0, \dots, x_{k-1} , se $f \in C^1(I)$ allora $f[x_0, \dots, x_{k-1}, x]$ si prolunga per continuità in x_0, \dots, x_{k-1} .

iii) Se $f \in C^k(I)$ allora $\exists \varepsilon \in [\min_{i=0, \dots, k} x_i, \max_{i=0, \dots, k} x_i]$

$$f[x_0, \dots, x_k] = \frac{f^{(k)}(\varepsilon)}{k!} \quad (\text{CONSEGUENZA DEL TH DI LAGRANGE})$$

→ coeff. dell'espansione di Newton di $P_k(x)$ sono proprio $f[x_0, \dots, x_k]$

TEOREMI DI ESPANSIONE

$f: I \rightarrow \mathbb{R}, I \subseteq \mathbb{R}, x_0, \dots, x_k \in I$ allora vale:
 $x_i \neq x_j$

POLIN. DI INTERP. $P_k(x)$
NELLA BASE DI NEWTON

$$f(x) = f(x_0) + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x_2] + \dots + (x - x_0)\dots(x - x_{k-1})f[x_0, \dots, x_k] + (x - x_0)\dots(x - x_k)f[x_0, \dots, x_k, x]$$

NON È UN POLINOMIO

DIM

Si procede per induzione su k :

$$k=0 \rightsquigarrow f(x_0) + (x-x_0) \cdot f[x_0, x] = f(x_0) + \cancel{(x-x_0)} \cdot \frac{f(x)-f(x_0)}{x-x_0}$$

$$= f(\cancel{x_0}) + f(x) - f(\cancel{x_0}) = f(x) \quad \checkmark$$

Vediamo il caso $k+1$, assumendo la tesi vera fino a k :

$$f(x_0) + \dots + (x-x_0) \dots (x-x_k) f[x_0, \dots, x_{k+1}] + (x-x_0) \dots (x-x_k) (x-x_{k+1}) f[x_0, \dots, x_{k+1}, x]$$

$$= f(x_0) + \dots + (x-x_0) \dots (x-x_n) \underbrace{\left(f[x_0, \dots, x_{k+1}] + (x-x_{k+1}) f[x_0, \dots, x_{k+1}, x] \right)}_{f[x_0, \dots, x_k, x] \text{ PER LA DEF. DI DIFF. DIVUSA}} =$$

$$= f(x_0) + \dots + (x-x_0) \dots (x-x_k) f[x_0, \dots, x_k, x] = f(x)$$

\downarrow
PASSO
INDUTTIVO

□

TH Vale l'egualanza

$$P_k(x) = f(x_0) + (x-x_0) f[x_0, x_1] + \dots + (x-x_0) \dots (x-x_{k-1}) f[x_0, \dots, x_k]$$

$\underbrace{\qquad\qquad\qquad}_{N_k(x)}$

(DIM. VERSONE SEGUENTE)



Dati $k+1$ modi x_0, \dots, x_k
 y_0, \dots, y_k

3! polinomio di interpolazione
di grado al più k :

$$P_k(x_j) = f(x_j) = y_j$$

ma non è unica la sua rapp.
(base dei monomi, base di Lagr.,
base di Newton)

RAPPRESENTAZIONE DI $P_k(x)$

BASE	COEFFICIENTI
Monomi	DEVO RISOLVERE $v_0 = y$
$l_j(x)$	$f(x_j) = y_j$
$n_j(x)$	$f[x_0, \dots, x_{j-1}]$

$l_j(x) = \prod_{\substack{i=0 \\ i \neq j}}^k \frac{(x - x_i)}{(x_j - x_i)}$ TUTTO POLINOMIO DI GRADO k
 $n_j(x) = \prod_{i=0}^{j-1} (x - x_i)$ TUTTO POLINOMIO DI GRADO $(j-1)$

Con Newton, se aumentiamo il num. di modi non dobbiamo ricalcolare tutti i coeff., ma solo quelli nuovi.

Tu Vale l'uguaglianza

$$P_k(x) = f(x_0) + (x - x_0)f[x_0, x_1] + \dots + (x - x_0) \dots (x - x_{k-1})f[x_0, \dots, x_k]$$

POLINOMIO
INTERPOLANTE

$N_k(x)$

ESPRESSIONE TRONCATA
OTTENUTA CON IL TM
DI ESPANSIONE

DIM

Verifichiamo per induzione su k che vale
 $N_k(x_j) = y_j = f(x_j) \quad \forall j = 0, \dots, k$

$$k=0 \rightsquigarrow N_0(x) = f(x_0) \Rightarrow N_0(x_0) = f(x_0) \quad \checkmark$$

$k+1$ (ASSUMENDO CHE LA TESI VA VERSO FINO A k)

Osserviamo che $N_{k+1}(x) = N_k(x) + (x - x_0) \cdots (x - x_k) f[x_0, \dots, x_k, x]$

SI ANNUIS x_j
PER $j = 0, \dots, k$

Se $j = 0, \dots, k$ allora $N_{k+1}(x_j) = N_k(x_j) = f(x_j)$

↑ PER L'IPOTESI
INDUTTIVA

Rimane da vedere se $N_{k+1}(x_{k+1}) = f(x_{k+1})$

$$N_{k+1}(x_{k+1}) = f(x_0) + (x_{k+1} - x_0) f[x_0, x_1] + \cdots + (x_{k+1} - x_k) \cdots (x_{k+1} - x_k) f[x_0, \dots, x_{k+1}]$$

$$= f(x_{k+1})$$

↑ PER IL TH DI ESPANSIONE

$$f(x) =$$

$$= f(x_0) + (x - x_0) f[x_0, x_1] + \cdots + (x - x_k) \cdots (x - x_k) f[x_0, \dots, x_k, x]$$

□

ERRORE DEL PROASS. CON IL POLIN. DI INTERPOLAZIONE

Riguardando l'enunciato del th di espansione

$$E_k(x) := f(x) - P_k(x) = \underbrace{\pi(x)}_{\hookrightarrow \pi(x) := (x - x_0) \cdots (x - x_k)} \cdot f[x_0, \dots, x_k, x]$$

$$\hookrightarrow \pi(x) := (x - x_0) \cdots (x - x_k)$$

COROLARIO Se $f \in C^{k+1}(I)$, $x_0, \dots, x_k, x \in I$

Allora (grazie alla proprietà (iii)) si ha:

$$E_k(x) = \pi(x) \cdot \frac{f^{(k+1)}(\varepsilon)}{(k+1)!} \quad \min_{j=0, \dots, k} x_j \leq x, \varepsilon \leq \max_{j=0, \dots, k} x_j$$

QUADRO DELLE DIFFERENZE DIVISE

x_0, x_1, \dots, x_k modi

$f(x_0), \dots, f(x_k)$

Per l'interp. di Newton servono le diff. divise (D.D.) di ordine k

x	$f(x)$ (D.D.0)	D.D.1	D.D.2	D.D.3	...	D.D.K
x_0	$f(x_0)$					
x_1	$f(x_1)$	$f[x_0, x_1]$				
x_2	$f(x_2)$	$f[x_0, x_2]$	$f[x_0, x_1, x_2]$			
x_3	$f(x_3)$	$f[x_0, x_3]$	$f[x_0, x_1, x_3]$	$f[x_0, x_1, x_2, x_3]$		
:						
:						

La tab. si riempie per righe e ogni riga cambia solo per l'ultimo valore -

I coeff. che ci servono sono gli elem. diagonali.

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

$$f[x_0, x_1, x_2] = \frac{f(x_2) - f[x_0, x_1]}{x_2 - x_1}$$

REGOLE:

Quando vado a calcolare la D.D. corrisp. alla riga di x_i e colonna D.D.j devo calcolare la diff. tra l' elem. alla sua sx e il primo elem. della colonna D.D.j-1 -

Infine devo dividere per $x_i - x_{j-1}$ (i modi corrisp. alle righe che sto considerando -

OSS Se tutte le entrate della colonna D.D.j sono costanti (hanno tutte lo stesso valore $\neq 0$)
 $\Rightarrow D.D.h = 0 \quad \forall h > j \Rightarrow$ il polinomio interpolante
 (tutte le colonne
 a dx sono 0) ha grado esatt. -j

ESEMPIO

	x_0	x_1	x_2	x_3	x_4
x	0	-1	2	-2	3
$f(x)$	5	3	3	-5	11
	y_0	y_1	y_2	y_3	y_4

Trovare $P_4(x)$ e dire qual'è il grado.

x	$f(x)$ (D.D.0)	D.D.1	D.D.2	D.D.3	D.D.4
0	5				
-1	3	2			
2	3	-1	-1		
-2	-9	7	-5	1	$\Rightarrow D.D.4 = 0$
3	11	2	0	1	\downarrow

$P_4(x)$ HA GRADO 3

$$P_4(x) = 5 \cdot 1 + 2x - x(x+1) + x(x+1)(x-2) = \\ = x^3 - 2x^2 - x + 5 \quad (\text{NUOVA BASE DEI MONOMI})$$

$$\left\{ 1, x, x(x+1), x(x+1)(x-2), x(x+1)(x-2)(x+2) \right\} \quad \text{BASE DI NEWTON}$$

$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$
 $n_0(k) \quad n_1(k) \quad n_2(k) \quad n_3(k) \quad n_4(k)$
 $\parallel \quad \parallel \quad \parallel \quad \parallel \quad \parallel$

ESEMPIO

x	4	8	12	16	20
$f(x)$	6	42	110	210	342

Calcolare $P_4(x)$ nella base
di Newton
(sapendo che potremmo doverci
formare prima, conviene
riempire la tab. per colonne)

x	$f(x)$	D.D.1	D.D.2	
4	6			
8	42	$(42-6)/(8-4) = 9$		
12	110	13	$(13-9)/(12-8) = 1$	
16	210	17	1	$\Rightarrow D.D.3 \text{ e } D.D.4 \text{ SONO } 0$
20	342	21	1	$\Rightarrow P_4(x) \text{ HA GRADO } 2$

$$n_0(x) = 1$$

$$n_1(x) = x - 4$$

$$n_2(x) = (x - 4)(x - 8)$$

$$n_3(x) = (x - 4)(x - 8)(x - 12) = 0$$

$$n_4(x) = (x - 4)(x - 8)(x - 12)(x - 16) = 0$$

$$P_4(x) = 6 \cdot n_0(x) + 9 \cdot n_1(x) + 1 \cdot n_2(x) =$$

$$= 6 + 9(x - 4) + (x - 4)(x - 8) \quad \text{NUOVA BASE DI NEWTON}$$

$$= x^2 - 3x + 2 \quad \text{NUOVA BASE DEI RESIDUI}$$

ESERCIZIO

x	0	1	2	3	4	5
$f(x) = y$	3	8	15	α	$4\frac{1}{2}$	β

$\alpha, \beta \in \mathbb{R}$

Dire per quali valori di α e β il polinomio interpolante $P_5(x)$ abbia grado minimo e dire qual'è il grado.

↓
grado
al più 5

x	$f(x)$	D.D.1	D.D.2	D.D.3
0	3			
1	8	5		
2	15	6	1	
4	$4\frac{1}{2}$	11	2	$\frac{1}{2}$

$$P_3(x) = 3 + 5x + x(x-1) + \frac{1}{2}x(x-1)(x-2)$$

Per fare in modo che $P_3(x)$ passi anche per $(3, \alpha)$ e $(5, \beta)$ mi basta porre:

$$\left. \begin{array}{l} \alpha = P_3(3) \Rightarrow \alpha = 27 \\ \beta = P_3(5) \Rightarrow \beta = 78 \end{array} \right\} \begin{array}{l} \text{PER QUESTI VALORI DI } \alpha \text{ E } \beta \\ \text{IL POLINOMIO HA GRADO 3} \end{array}$$

Esercizio

$$\begin{array}{c|ccccc} x & 0 & 1 & \alpha & -1 & 3 \\ \hline f(x) & 2 & 1 & 4 & 3\alpha+1 & 11 \end{array} \quad \alpha \in \mathbb{R}$$

Determinare α per cui $P_4(x)$ ha grado minimo.

Non posso fare come prima perché α compare sia in x che in $f(x)$.

x	$f(x)$	D.D.1	D.D.2	
0	2			
1	1	-1		
3	11	3	2	
-1	$3\alpha+1$	$1-3\alpha$	$(3\alpha-2)/2$	
α	4	$2/\alpha$	$(2+\alpha)/\alpha(\alpha-1)$	

\downarrow α INTONDO

PER SEMPLIF. I CALCOLI
L'ORDINE È INDIFF.

Posso rendere $(3\alpha-2)/2 = (2+\alpha)/[\alpha(\alpha-1)] = 2$?

$$2 = \frac{3\alpha-2}{2} \iff 3\alpha-2=4 \Rightarrow \alpha=2$$

$$(2+2)/[2(2-1) = 4/2 = 2 \checkmark$$

Per $\alpha=2$ si ha che $P_4(x)$ ha grado 2 (minimo)
sostit. $\alpha=2$ si trova $P_4(x) = 2x^2 - 3x + 2$

Se non fossimo riusciti con $\alpha=2$ altrimenti calcolato D.D.3 e riprovato e così via

INTERPOLAZIONE OSCULATORIA DI HERMITE

Abbiamo $k+1$ nodi x_0, \dots, x_k ed in corrispondenza dei nodi conosciamo $f(x_0), \dots, f(x_k)$ e $f'(x_0), \dots, f'(x_k)$

Vogliamo trovare un polinomio

$$H : \begin{cases} H(x_j) = f(x_j) \\ H'(x_j) = f'(x_j) \\ A_j = 0, \dots, k \end{cases}$$

Nell'interp. visto finora avevamo:

$k+1$ condizioni per determinare $(k+1)$ parametri
(i coeff. di $P_k(x)$)

Adesso abbiamo $2(k+1)$ condizioni (equazioni)

\Rightarrow possiamo determinare $2k+2$ parametri
ad esempio i coeff. di un polinomio di grado
al più $2k+1$

$H_{2k+1}(x) :=$ POLINOMIO DELL'INTERP. OSCULATORIA
DI HERMITE

OSS Determinare i coeff. di $H_{2k+1}(x)$ può essere formulato come risolvere un sist. lineare di dim. $(2k+2) \times (2k+2)$.

In maniera analoga (ma diversa) al sistema con la mat. di Vandermonde nel problema di interp. classico

Per i soliti motivi di mal condizionamento questa strada non si utilizza per trovare $H_{2k+1}(x)$.

TH $H_{2k+1}(x)$, se $x_i \neq x_j \forall i \neq j$, esiste ed è
unico nell'insieme dei polinomi di grado
al più $2k+1$

(DIM. SUA DISPENSA)

INTERPOLAZIONE OSCULATRORIA DI HERMITE

$f: \mathbb{R} \rightarrow \mathbb{R}$, $f: I \rightarrow \mathbb{R}$ $I \subseteq \mathbb{R}$

x_0, \dots, x_n nodi dell' interpolazione

$f(x_0), \dots, f(x_n)$

$f'(x_0), \dots, f'(x_n)$

Vogliamo trovare $H_{2k+1}(x)$ polinomio di grado al più $2k+1$ t.c.:

$$H_{2k+1}(x_j) = f(x_j) = y_j$$

$$H_{2k+1}'(x_j) = f'(x_j)$$

TH Se $x_i \neq x_j \quad \forall i \neq j$ ($\frac{x_0 \ x_1 \ \dots \ x_k}{x_i}$) esiste ed è unico $H_{2k+1}(x)$

IDEA
$$H_{2k+1}(x) = \sum_{j=0}^k h_{0,j}(x) \cdot f(x_j) + \sum_{j=0}^k h_{1,j} \cdot f'(x_j)$$

La base di polinomi usati per l' interp. di Hermite è:
 $\{h_{0,0}(x), h_{0,1}(x), \dots, h_{0,k}(x), h_{1,0}(x), h_{1,1}(x), \dots, h_{1,k}(x)\}$ TOTALE
POLINOMI:
 $2k+2$

$h_{0,j}(x)$ e $h_{1,j}(x)$ sono polinomi di grado $2k+1$

$h_{0,j}(x) =$ FUNZIONI FONDAMENTALI DI PRIMA SPECIE PER
L' INTERP. DI HERMITE

$$h_{0,j}(x_i) = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i=j \end{cases} \quad h'_{0,j}(x_i) = 0 \quad \forall i$$

$h_{1,j}(x) =$ FUNZIONI FONDAMENTALI DI SECONDA SPECIE

$$h_{1,j}(x_i) = 0 \quad \forall i$$

$$h'_{1,j}(x_i) = \begin{cases} 0 & i \neq j \\ 1 & i=j \end{cases}$$

$$H_{2k+1}(x) = \sum_{j=0}^k f(x_j) \cdot h_{0j}(x) + \sum_{j=0}^k f'(x_j) \cdot h_{1j}(x)$$

$$H_{2k+1}(x_i) = \sum_{j=0}^k f(x_j) h_{0j}(x_i) + \sum_{j=0}^k f'(x_j) \cancel{h_{1j}(x_i)} =$$

~~$h_{1j}(x_i)$~~ $\underset{=0}{\cancel{h_{1j}(x_i)}} = A_i$

$$= f(x_i) \underbrace{h_{0i}(x_i)}_{=1} = f(x_i)$$

$$\forall i = 0, \dots, k$$

$$H'_{2k+1}(x_i) = \sum_{j=0}^k f(x_j) \cancel{h'_{0j}(x_i)} + \sum_{j=0}^k f'(x_j) h'_{1j}(x_i) =$$

~~$h'_{0j}(x_i)$~~ $\underset{=0 \forall i}{\cancel{h'_{0j}(x_i)}} + \sum_{j=0}^k f'(x_j) h'_{1j}(x_i) =$

$$= f'(x_i) \underbrace{h'_{1i}(x_i)}_{=1} = f'(x_i)$$

$$h_{0j}(x): h_{0j}(x_i) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \quad h'_{0j}(x_i) = 0 \quad \forall i$$

$$h_{0j}(x) = (\underbrace{Ax + B}_{\text{PUN. DI GRADO } 1}) \cdot l_j^2(x)$$

$$l_j(x) = \frac{(x-x_0) \cdots (x-x_{j-1})(x-x_{j+1})(x-x_k)}{(x_j-x_0) \cdots (x_j-x_k)}$$

$$l_j(x_i) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$l_j^2(x)$ ha grado $2k$, se moltiplico per un pol. di grado 1 ottengo $h_{0j}(x)$ di grado $2k+1$

$$h_{0j}(x_i) = (Ax_i + B) l_j(x_i)$$

0 per $i \neq j \Rightarrow$ sono soddisfatte k condizioni
 $Ax_j + B$ per $i = j$

$$\cdot A x_j + B = 1$$

$$h'_{0j}(x_i) = A \ell_j^2(x_i) + 2(Ax_j + B)\ell_j(x_i)\ell_j'(x_i)$$

↑ $\circ \quad i \neq j$
 ↓
 $A + 2(Ax_j + B)\ell_j(x_j)$

$$\cdot A + 2(Ax_j + B)\ell_j'(x_j) = 0$$

Dove impostare:

$$\begin{cases} Ax_j + B = 1 \\ A + 2(Ax_j + B)\ell_j'(x_j) = 0 \end{cases} \Rightarrow \begin{cases} Ax_j + B = 1 \\ A + 2\ell_j'(x_j) = 0 \end{cases}$$

$$\Rightarrow \begin{cases} B = 1 - 2x_j\ell_j'(x_j) \\ A = -2\ell_j'(x_j) \end{cases}$$

Lo stesso per gli $h_{i,j}(x)$:

$$h_{i,j}(x) = (Cx + D) \ell_j^2(x)$$

$$h_{i,j}(x_i) = (Cx_i + D) \ell_j^2(x_i)$$

↑ $\circ \quad i \neq j$
 ↓
 $Cx_j + D \quad i = j$

$$\cdot Cx_j + D = 0$$

$$h'_{i,j}(x_i) = C \cdot \ell_j^2(x_i) + (Cx_i + D) 2\ell_j(x_i)\ell_j'(x_i)$$

↑ $\circ \quad i \neq j$
 ↓
 $C + (Cx_i + D) 2\ell_j'(x_i)$
 $i = j$

$$\cdot C + 2(Cx_j + D) l_j'(x_j) = 1$$

Dobbiamo impostare:

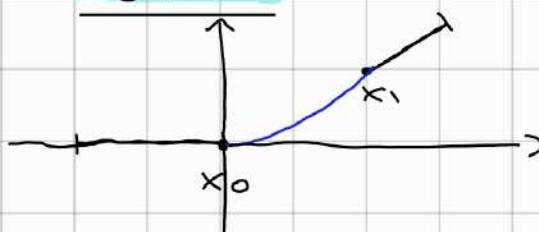
$$\begin{cases} Cx_j + D = 0 \\ C + 2(Cx_j + D) l_j'(x_j) = 1 \end{cases} \Rightarrow \begin{cases} Cx_j + D = 0 \\ C = 1 \end{cases} \Rightarrow \begin{cases} D = -x_j \\ C = 1 \end{cases}$$

Mettendo insieme le espressioni di A, B, C, D possiamo scrivere il polim. di interp. di Hermite come:

$$H_{2k+1}(x) = \sum_{j=0}^k f(x_j) [1 - 2l_j'(x_j)(x-x_j)] l_j^2(x) + \sum_{j=0}^k f'(x_j) (x-x_j) l_j^2(x)$$

OSS Una volta che sono dati x_0, \dots, x_k posso calcolare tutte le ℓ_j nell'espressione $H_{2k+1}(x)$

ESEMPIO



Vogliamo trovare il polinomio $\boxed{\text{H}}$

$$x_0 = 0, x_1 = d, f(x_1) = l, f'(x_1) = s \quad d, l, s \in \mathbb{R} \text{ parametri}$$

$$f'(0) = 0, f'(x_1) = s$$

$$\ell_0(x) = \frac{d-x}{d} \quad \ell_1(x) = \frac{x}{d}$$

$$h_{00}(x) = \left[1 + 2\frac{x}{d}\right] \frac{(x-d)^2}{d^2}$$

$$h_{01}(x) = \left[1 - \frac{2(x-d)}{d}\right] \frac{x^2}{d^2}$$

$$h_{1,0}(x) = \frac{x(x-d)^2}{d^2}$$

$$h_{1,1}(x) = (x-d) \frac{x^2}{d^2}$$

$$H_3(x) = \left[1 - 2 \frac{(x-d)}{d} \right] \frac{x^2}{d^2} L + (x-d) \frac{x^2}{d^2} S =$$

$$= f(x_0) \cdot h_{0,0}(x) + f(x_1) \cdot h_{0,1}(x) + f'(x_0) \cdot h_{1,0}(x) + f'(x_1) \cdot h_{1,1}(x)$$

$x_0, \dots, x_k \in I \subset \mathbb{R}$ (es. $I = [a, b]$)

$f(x_0), \dots, f(x_k)$

$f'(x_0), \dots, f'(x_k)$

$$H_{2k+1}(x_j) = f(x_j) \quad j=0, \dots, k$$

$$H'_{2k+1}(x_j) = f'(x_j)$$

$$H_{2k+1}(x) = \sum_{j=0}^k \left[-2\ell_j(x_j)(x-x_j) \right] \ell_j^2(x) \cdot f(x_j) + \\ + \sum_{j=0}^k (x-x_j) \ell_j^2(x) \cdot f'(x_j) = \text{POLINOMIO DI HERMITE}$$

$$\ell_j(x) = \prod_{\substack{i=0 \\ i \neq j}}^k \frac{(x-x_i)}{(x_j-x_i)} = \text{POLONMI DI LAGRANGE}$$

TH $f \in C^{2k+2}(I)$ allora:

ESPRESSIONE DELL'ERRORE

$$\forall x \in I, |f(x) - H_{2k+1}(x)| = (x-x_0)^2 \cdots (x-x_k)^2 \frac{f^{(2k+2)}(\varepsilon)}{(2k+2)!}$$

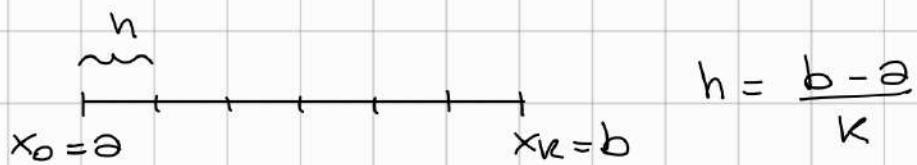
$$\min \{x_0, \dots, x_k, x\} \leq \varepsilon \leq \max \{x_0, \dots, x_k, x\}$$

OSS Nel caso di una specifica f e un dato I si puo' stimare $\max_{\varepsilon \in I} |f^{(2k+2)}(\varepsilon)|$ e di conseguenza $\max |f(x) - H_{2k+1}(x)|$

OSS L'espressione dell'errore ($|f(x) - H_{2k+1}(x)|$) suggerisce che vicino ai nodi l'approssimazione è accurata.

Quindi se sono interessato ad f su I , devo distribuire i nodi in modo da coprire tutte le parti dell'intervallo.

La scelta più intuitiva è prendere i nodi equispaziati



$$h = \frac{b-a}{k}$$

anche se non sempre funziona (es. nel caso della funz. di Runge $\frac{1}{100x^2 - 100x + 26}$ $[0, 1]$)

$I_{Pm}(x)$ POLINOMIO INTERP. di grado m su nodi equispaziati (x_0, \dots, x_k)

$$\Rightarrow \lim_{n \rightarrow +\infty} \max_{[0,1]} |f(x) - I_{Pm}(x)| = +\infty$$

OSS Questa cosa si può sistemare con scelte di nodi differenti (es. nodi di Chebychev). Se $f \in C^2(I)$ \Rightarrow i nodi di cheb. convergono

CURIOSITÀ

TH DI WEIERSTRASS

$\forall f \in C(I) \exists \{p_n\}: p_n \xrightarrow[n \rightarrow +\infty]{} f$

o equivalentemente

$\forall \epsilon > 0 \exists n \in \mathbb{N} \text{ e } p_n \text{ polim. di grado } n \max_I |f - p_n| < \epsilon$

TH DI FABER

Se si scelgono i modi del polin. intemp. di grado n , $\forall n$ allora:

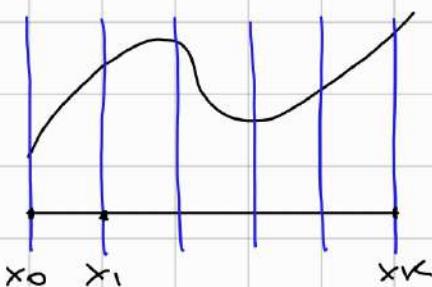
$$\exists f \in C(I) : \lim_{n \rightarrow +\infty} |f(x) - I_{P_n}(x)| \neq 0$$

CONCLUSIONE Non esiste una scelta di modi di intemp. che va bene per tutte le f. continue

INTERP. POLINOMIALE A TRAM

Si divide l' intervallo dove si vuole appross. f e in ogni intervallo si utilizza un polinomio di grado d più d com d fissato.
L' idea è che mantenendo d fisso, si possono ottenere appross. più accurate riducendo l' ampiezza degli intervalli.

ESEMPIO



INTERP. LINEARE A TRAM

$$(d=1)$$

→ F. CONTINUA, MA NON DERIVABILE NEI NODI

In generale si considera $d > 1$, e che la f. ottenuta appartenga alla classe $C^{d-1}(I)$.

Si chiede che la f. sia continua ed interpoli f, f', \dots, f^{d-1} nei nodi.

FUNZIONI SPLINE

$(k+1)$ punti x_0, \dots, x_k

$f: [a, b] \rightarrow \mathbb{R}$ con $x_0 = a$ e $x_k = b$

Si dice Spline la funzione $S_d(x)$ t.c.

- i) $S_d(x)|_{[x_{i-1}, x_i]} \text{ è un polin. di grado al più } d$ $i=1, \dots, k$
- ii) $S_d(x_i) = y_i \quad i=1, \dots, k$ (INTERPOLAZIONE)
- iii) $S_d(x) \in C^{d-1}([a, b])$

Il caso più studiato ed utilizzato (es. in computer grafica) è quello con $d=3$ (SPLINE CUBICA).

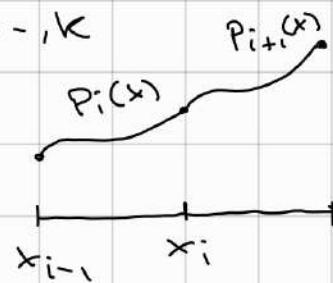
Si chiamano k intervalli $[x_{i-1}, x_i]$ e corrispondenti k polinomi di grado al più 3.

Quindi $S_3(x)$ è definita da $4k$ parametri, ovvero tutti i coeff. dei vari polinomi di grado ≤ 3 .

Le condizioni per essere Spline sono:

Sia $p_i(x) = S_3(x) |_{[x_{i-1}, x_i]} \quad i=1, \dots, k$

- $p_i(x_{i-1}) = y_{i-1} \quad i=1, \dots, k$
- $p_i(x_i) = y_i \quad i=1, \dots, k$
- $p'_i(x_i) = p'_{i+1}(x_i) \quad i=1, \dots, k-1$
- $p''_i(x_i) = p''_{i+1}(x_i) \quad i=1, \dots, k-1$



TOTALE: $4k - 2$ condizioni

Rimangono 2 condizioni, sul comportamento di $S_3(x)$ sul bordo.

Ci sono delle alternative:

① SPLINE NATURALE: $P_1''(x_0) = P_k''(x_k) = 0$

② SPLINE PERIODICA: $P_1'(x_0) = P_k'(x_k)$ $P_1''(x_0) = P_k''(x_k)$

③ SPLINE VINCOLATA: $P_1'(x_0) = y'_0$ $P_k'(x_k) = y'_k$

TH Esiste ed è unica la spline naturale con nodi equispaziati ($x_{i+1} - x_i = h = \frac{b-a}{k}$ $\forall i = 0, \dots, k$)

DIM

$$H_{2k+1}(x) = \sum_{j=0}^k \left[1 - 2\ell_j(x_j)(x-x_j) \right] \ell_j^2(x) \cdot f(x_j) + \\ + \sum_{j=0}^k (x-x_j) \ell_j^2(x) \cdot f'(x_j) = \text{POLINOMIO INTERP. DI HERMITE}$$

$$S_3'(x_i) = m_i \quad i = 1, \dots, k-1$$

Per ogni intervallo $[x_{i-1}, x_i]$ si costruisce il polin. di interp. di Hermite a 2 punti basato sulla tabella:

x	f	f'	
x_{i-1}	y_{i-1}	m_{i-1}	
x_i	y_i	m_i	

$$x_i - x_{i-1} = h \text{ (costante perché i nodi sono equispaz.)}$$

$$\rightarrow H_3(x) = P_i(x) = \left[1 + \frac{2}{h}(x-x_{i-1}) \right] \left(\frac{x-x_i}{h} \right)^2 y_{i-1}$$

$$+ \left[1 - \frac{2}{h}(x-x_i) \right] \left(\frac{x-x_{i-1}}{h} \right)^2 y_i +$$

$$+ (x-x_{i+1}) \left(\frac{x-y_i}{h} \right)^2 m_{i-1} + (x-x_i) \left(\frac{x-x_{i-1}}{h} \right)^2 m_i =$$

$$= \left[y_{i-1} + \left(m_{i-1} + \frac{2y_{i-1}}{h} \right) (x-x_{i-1}) \right] \left(\frac{x-x_i}{h} \right)^2 + \left[y_i + \left(m_i - \frac{2y_i}{h} \right) (x-x_i) \right] \left(\frac{x-x_{i-1}}{h} \right)^2$$

Per qualsiasi scelta di m_1, \dots, m_{k-1} , i polinomi $p_i(x)$, per costruzione, verificano i primi 3 insiemi di condizioni che avevamo imposto sulla spline.

Rimangono da imporre le condizioni:

- $p_i''(x_i) = p_{i+1}''(x_i) \quad i=1, \dots, k-1$
- $p_1''(x_0) = p_k''(x_k) = 0 \quad (\text{SPLINE NATURALE})$

Abbiamo l'espressione di $p_i''(x)$, facendo la derivata seconda e valutando in x_{i-1} e x_i si ottiene:

$$p_i''(x_{i-1}) = \frac{2}{h^2} [3(y_i - y_{i-1}) - h(m_i + 2m_{i-1})]$$

$$p_i''(x_i) = \frac{2}{h^2} [3(y_{i+1} - y_i) + h(2m_i + m_{i-1})]$$

Se si impone $p_i''(x_i) = p_{i+1}''(x_i)$ si ottiene

$$\frac{2}{h^2} [3(y_{i+1} - y_i) + h(2m_i + m_{i-1})] = \frac{2}{h^2} [3(y_{i+1} - y_i) - h(m_{i+1} + 2m_i)]$$

$$\Rightarrow h(m_{i+1} + 4m_i + m_{i-1}) = 3(y_{i+1} - y_{i-1})$$

$$\Rightarrow m_{i+1} + 4m_i + m_{i-1} = \frac{3}{h}(y_{i+1} - y_{i-1}) \quad i=1, \dots, k-1$$

Per quanto riguarda la spline naturale

$$p_1''(x_0) = 0 \Rightarrow 2m_0 + m_1 = \frac{3}{h}(y_1 - y_0)$$

$$p_k''(x_k) = 0 \Rightarrow m_{k-1} + 2m_k = \frac{3}{h}(y_k - y_0)$$

Abbiamo quindi $(k+1)$ equazioni in $(k+1)$ incognite (m_0, \dots, m_k) -

$$\rightarrow \left[\begin{array}{cccccc} 2 & 1 & & & & \\ 1 & 4 & 1 & & & \\ \vdots & \ddots & \ddots & & & \\ \vdots & \ddots & \ddots & & & \\ 1 & 4 & i & & & \\ 1 & 2 & & & & \end{array} \right] \left[\begin{array}{c} m_0 \\ m_1 \\ m_2 \\ \vdots \\ m_{k-1} \\ m_k \end{array} \right] = \left[\begin{array}{c} \frac{3}{h}(y_1 - y_0) \\ \frac{3}{h}(y_2 - y_0) \\ \frac{3}{h}(y_3 - y_1) \\ \vdots \\ \frac{3}{h}(y_k - y_{k-2}) \\ \frac{3}{h}(y_k - y_{k-1}) \end{array} \right] \rightarrow \text{EQUAZIONE GENERICA}$$

(*)

SE ANZI OGNI LA SPLINE NATURALE PRENDIALE SOLO LA SPLINE PERIODICA O VINCOLATA CAMBIANO SOLO LE RIGHE O E K, OVVERO LE CONDIZIONI AL BORDO

(QUELLO CHE LA F. FA IN X0 E Xk)

Identificare i coeff. m_i è equiv. a risolvere il sistema sopra dato che la matrice (*) è a predom. diag. forte \Rightarrow è invertibile \Rightarrow esiste un'unica soluzione \Rightarrow esiste un'unica scelta dei coeff. m_i che verifica tutte le condiz. per avere una spline naturale.

TJ

OSS Con più o meno lo stesso dim., si dimostra la stessa cosa per spline vincolata, periodica con modi non per forza equispaziati -

TH $f(x) \in C^4(I)$ e $S_3(x)$ spline naturale associata ad i modi equispaziati (di ampiezza h) su I . Allora $\exists k_0, k_1, k_2, k_3 > 0$ tali per cui:

$$\max_{x \in I} |f^{(j)}(x) - S_3^{(j)}(x)| \leq k_j \cdot h^{4-j}, \quad j = 0, 1, 2$$

$$\max_{x \in (x_{i-1}, x_i)} |f^{(3)}(x) - S_3^{(3)}(x)| \leq k_3 \cdot h \quad \forall i = 1, \dots, k$$

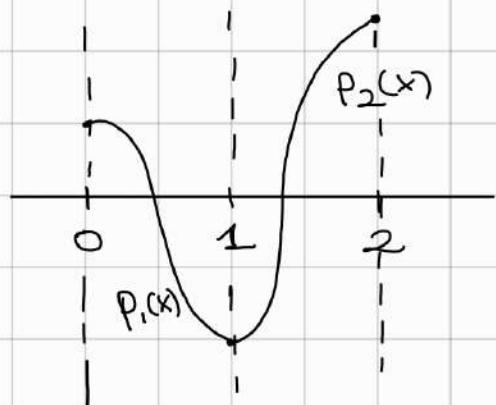
ESEMPIO

Determinare la spline cubica naturale che interpola i modi:

x	0	1	2
y	α	$-\alpha$	2α

$\alpha \in \mathbb{R}$

$$h = 1 = x_i - x_{i-1}$$



Si devono determinare i coeff. di $P_1(x) : [0, 1] \rightarrow \mathbb{R}$ e $P_2(x) : [1, 2] \rightarrow \mathbb{R}$

$$3(2\alpha - \alpha) = 3\alpha$$

Supponiamo $m_i = S'_3(x_i)$ $i = 0, 1, 2$

Allora m_0, m_1, m_2 devono risolvere:

$$\begin{cases} 2m_0 + m_1 = -6\alpha \\ m_0 + 4m_1 + m_2 = 3\alpha \\ m_1 + 2m_2 = 9\alpha \end{cases}$$

$$\begin{cases} 14m_2 - 66\alpha + 9\alpha - 2m_2 = -6\alpha \\ \Rightarrow \begin{cases} m_0 + 36\alpha - 7m_2 = 3\alpha \\ m_1 = 9\alpha - 2m_2 \end{cases} \end{cases}$$

$$\Rightarrow \begin{cases} 12m_2 = 51\alpha \\ m_0 = 7m_2 - 33\alpha \\ m_1 = 9\alpha - 2m_2 \end{cases} \Rightarrow \begin{cases} m_0 = -\frac{13}{4}\alpha \\ m_1 = \frac{\alpha}{2} \\ m_2 = \frac{17}{4}\alpha \end{cases}$$

Sostituendo questi valori nell'espressione di $p_1(x)$
($h=1, x_0=0, x_1=1, m_0 = -\frac{13}{4}\alpha, m_1 = \frac{\alpha}{2}, m_2 = \frac{17}{4}\alpha$)

$$P_1(x) = \left[y_{i-1} + \left(m_{i-1} + \frac{2y_{i-1}}{h} \right) (x - x_{i-1}) \right] \left(\frac{x - x_i}{h} \right)^2 + \left[y_i + \left(m_i - \frac{2y_i}{h} \right) (x - x_i) \right] \left(\frac{x - x_i}{h} \right)^2$$

$$P_1(x) = (1+2x)(x-1)^2\alpha + (3-2x)x^2(-\alpha) + x(x-1)^2\left(-\frac{13}{4}\alpha\right) +$$

$$+ (x-1)x^2 \cdot \frac{\alpha}{2} = \frac{5}{4}\alpha x^3 - \frac{13}{4}\alpha x + \alpha$$

$$\begin{aligned}
 P_2(x) &= \left[y_1 + \left(m_1 + \frac{2y_1}{n} \right) (x - x_1) \right] \left(\frac{x - x_2}{n} \right)^2 + \left[y_2 + \left(m_2 - \frac{2y_2}{n} \right) (x - x_2) \right] \left(\frac{x - x_1}{n} \right)^2 = \\
 &= \left[-\alpha + \left(\frac{\alpha}{2} - 2\alpha \right) (x - 1) \right] (x - 2)^2 + \left[2\alpha + \left(\frac{14}{4}\alpha - 4\alpha \right) (x - 2) \right] (x - 1)^2 = \\
 &= -\frac{5}{4}\alpha x^3 + \frac{15}{2}\alpha x^2 - \frac{43}{4}\alpha x + \frac{7}{2}\alpha
 \end{aligned}$$

$$S_3(x) = \begin{cases} P_1(x) & x \in [0, 1] \\ P_2(x) & x \in [1, 2] \end{cases}$$

↳ f. definita a tratti

APPROXIMAZIONE NEL SENSO DEI MINIMI QUADRATI

MINIMI QUADRATI vs INTERPOLAZIONE

IN CONUNO: $x_0, \dots, x_k \in \mathbb{R}$

$y_0, \dots, y_k \in \mathbb{R}$ ($f(x_0), \dots, f(x_k)$ o loro appross.)

e cerchiamo un' appross. di f :

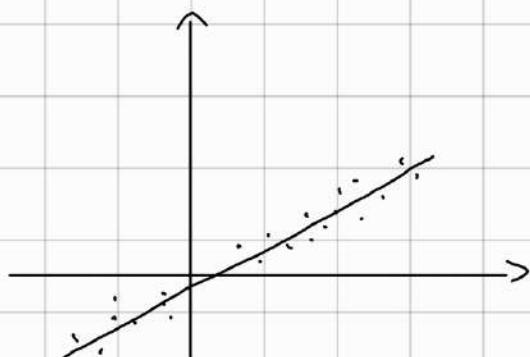
$$\Phi(x) = \sum_{j=0}^m c_j e_j(x) \quad c_j \in \mathbb{R}, e_j(x) \text{ f. notevoli (o di base)}$$

DIFERENZE: $e_j(x)$ le consideriamo anche non polino-
miali.

Assumiamo $m \leq k$ (molto spesso $m \ll k$)

ESEMPIO

REGRESSIONE LINEARE



$$e_0(x) = 1 \quad e_1(x)$$

$$\Phi(x) = c_0 + c_1 x$$

ESEMPIO

$$e_0(x) = e^{2x}, e_1(x) = \sin\left(\frac{x}{2}\right), e_2(x) = \frac{1}{3x^2+4}$$

$$\Rightarrow \Phi(x) = c_0 e^{2x} + c_1 \sin\left(\frac{x}{2}\right) + \frac{c_2}{3x^2+4}$$

La cosa importante è che $\Phi(x)$ dipenda linearmente
da c_0, \dots, c_m .

$$\text{ES. } \Phi(x) = e^{2c_0 x} + \sin\left(\frac{c_1 x}{2}\right) + \frac{c_2}{3x^2+4}$$

Non si ha dipendenza lineare da c_0, c_1

\Rightarrow non si arriva ad un problema lineare ai minimi quadrati (è un problema non lineare)

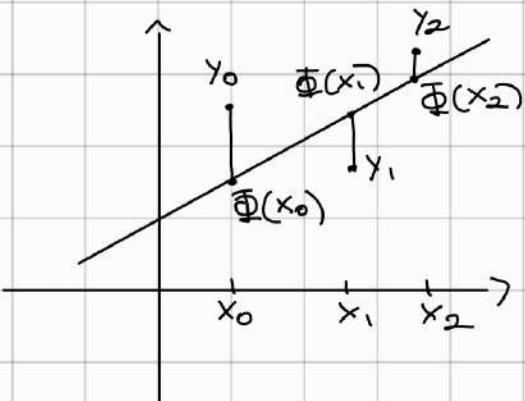
COME SI VALUTA LA QUALITÀ DI $\Phi(x) \approx f(x)$

L'idea più naturale è utilizzare le $\Phi(x_i) - y_i$ ad indicare quanto il modello si discosta dai dati reali.

Si può definire un errore tot. sui dati misurati e cercare c_0, \dots, c_m in modo tale che l'errore sia minimo:

SOMMA DEGLI SCATTI QUADRATICI

$$\Psi(c_0, \dots, c_m) = \sum_{i=0}^k (\Phi(x_i) - y_i)^2$$



LEVANDO AL QUADRATO
EVITANDO CHE L'ERRORE SI
ANNULLI A CAUSA DEI
DIVERSI SEGNI

SE INSEGUSSIMO IL MODO
ANCHE L'ERRORE AL QUADR.
I PUNTI IN CUI CAMBIA IL
SEGNO NON SARANNO PIÙ
DISTINZIONABILI

OBIETTIVO: Minimizzazione $\Psi(c_0, \dots, c_m) \rightarrow$ identifica c_0, \dots, c_m

$$\Psi(c_0, \dots, c_m) = \|v(x)\|_2^2 \quad \text{dove } v(c_0, \dots, c_m) = \begin{bmatrix} \Phi(x_0) - y_0 \\ \vdots \\ \Phi(x_k) - y_k \end{bmatrix}$$

$$\frac{\partial}{\partial c_s} \Psi(c_0, \dots, c_m) = \frac{\partial}{\partial c_s} \left(\sum_{j=0}^k (\Phi(x_j) - y_j)^2 \right) =$$

$$s \in \{0, 1, \dots, m\}$$

$$= \frac{\partial}{\partial c_s} \left(\sum_{j=0}^k \left(\sum_{h=0}^m c_h v_h(x_j) - y_j \right)^2 \right) =$$

$$\begin{aligned}
&= \sum_{j=0}^k \frac{\partial}{\partial c_s} \left(\sum c_n e_n(x_j) - y_j \right)^2 = \\
&= \sum_{j=0}^k 2 \left(\sum c_n e_n(x_j) - y_j \right) \frac{\partial}{\partial c_s} \left(\sum_{h=0}^m c_h e_h(x_j) - y_j \right) = \\
&= 2 \sum_{j=0}^k \left(\sum_{h=0}^m c_h e_h(x_j) - y_j \right) e_s(x_j) = \\
&= 2 \left[\sum_{j=0}^k \sum_{h=0}^m e_s(x_j) \cdot e_h(x_j) \cdot c_h - \sum_{j=0}^k e_s(x_j) y_j \right]
\end{aligned}$$

Dato che cerchiamo il minimo di Ψ , ci interessano i valori di c_0, \dots, c_m t.c $\nabla \Psi(c_0, \dots, c_m) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^{m+1}$

In particolare ci interessano le soluzioni di $\frac{\partial \Psi}{\partial c_s} = 0$

Ovvero vogliamo la condizione:

$$(*) \quad \sum_{j=0}^k \sum_{h=0}^m e_s(x_j) \cdot e_h(x_j) \cdot c_h = \sum_{j=0}^k e_s(x_j) y_j \quad \forall s = 0, \dots, m$$

$$A = \begin{bmatrix} e_0(x_0) & e_1(x_0) & \dots & e_m(x_0) \\ \vdots & \vdots & & \vdots \\ e_0(x_k) & e_1(x_k) & \dots & e_m(x_k) \end{bmatrix} \in \mathbb{R}^{(k+1) \times (m+1)}$$



$$C = \begin{bmatrix} c_0 \\ \vdots \\ c_m \end{bmatrix} \in \mathbb{R}^{m+1}, \quad y = \begin{bmatrix} y_0 \\ \vdots \\ y_k \end{bmatrix} \in \mathbb{R}^{k+1}$$

$$\sum_{j=0}^k e_s(x_j) y_j = (A^T y)_s = \text{COMPONENTE } s \text{ DEL PRODOTTO } A^T y$$

j-esima ENTRATA DI AC

$$\sum_{j=0}^k \sum_{h=0}^m e_s(x_j) e_h(x_j) c_h = \sum_{j=0}^k e_s(x_j) \left(\sum_{h=0}^m e_h(x_j) c_h \right) =$$

$$= \sum_{j=0}^k e_S(x_j) [A_C]_j = \\ = (A^T A_C)_{Sj}$$

(*) $\Leftrightarrow A^T A_C = A^T y$

dove $A^T A \in \mathbb{R}^{(m+1) \times (m+1)}$
 $A^T y \in \mathbb{R}^{m+1}$

$$A^T A = \begin{matrix} m+1 \\ \boxed{\quad} \end{matrix} \quad \begin{matrix} m+1 \\ \boxed{\quad} \\ m+1 \end{matrix} = \begin{matrix} m+1 \\ \boxed{\quad} \\ m+1 \end{matrix}$$

CONCLUSIONE: Per trovare i valori ottimi di c_0, \dots, c_m basta risolvere il sist. lin. $A^T A_C = A^T y$

DEF $A^T A_C = A^T y$ è detto SISTEMA DEVE EQ. NORMALI

OSS Si può dimostrare che:

$$\text{rk}(A^T A | A^T y) = \text{rk}(A^T A) \quad \forall y \Rightarrow \begin{matrix} \text{ESISTE SEMPRE ALMENO} \\ \text{UNA SOL. DEL SISTEMA} \\ \text{DEVE EQ. NORMALI} \end{matrix}$$

LEMMA Se A ha rango massimo (ovvero $\text{rk}(A) = m+1$) allora la soluzione di $A^T A_C = A^T y$ è unica.

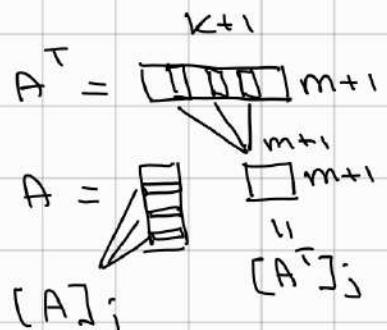
DIM Ci basta dimostrare che $\det(A^T A) \neq 0$
 Per il th di Binet-Cauchy (esteso):

$$\det(A^T A) = \sum_j \det([A^T]_j) \cdot \det([A]_j) =$$

(j sottoinsieme di $m+1$ indici in $\{0, \dots, k+1\}$)

$$= \sum_j \det([A]_j^T) \det([A]_j) =$$

$$= \sum_j \det([A]_j)^2 \geq 0 \quad \begin{matrix} \text{SICCOME} \\ \text{rk}(A) = m+1 \end{matrix} \Rightarrow \det(A^T A) > 0$$



$$\Psi(c_0, \dots, c_m) = \sum_{j=0}^m \left(\sum_{h=0}^k c_j e_j(x_h) - y_h \right)^2 = \| \underbrace{Ac - y}_2 \|_2^2 = \\ = (Ac - y)^T (Ac - y)$$

$$\nabla \Psi = A^T Ac - A^T y$$

$H\Psi(c_0, \dots, c_m) = A^T A$ = MAT. HESSIANA DELLA
FUNZIONE OGGETTO DEL
PROBLEMA AL MIN. QUADRATI

Sarà mat. $A^T A$ è simmetrica semidef. pos.:

$$(A^T A)^T = A^T A \quad \checkmark \text{ SIMMETRICA}$$

$$x \in \mathbb{R}^{m+1} \setminus \{0\} \quad x^T A^T A x = (Ax)^T Ax = \|Ax\|_2^2 \geq 0 \\ \Rightarrow A^T A \text{ SEMIDEF. POS.} \Rightarrow \text{SOLUZIONE DI } \nabla \Psi = 0 \\ \text{È PTO DI MINIMO}$$

OSS Il fatto che l' Hessianiana (che è sempre $A^T A$) è def. pos. $\forall c_0, \dots, c_m \Rightarrow \Psi$ è convessa

OSS Il valore ottimo, ovvero $\Psi(\hat{c})$ dove $\hat{c} = \begin{bmatrix} c_0 \\ \vdots \\ c_m \end{bmatrix}$

è la soluzione di $A^T Ac = A^T y$ è tipicamente ≥ 0
Se è $= 0$, questo significa che $\Phi(x) = \sum_{j=0}^m c_j e_j$
passa esattamente da $(x_0, y_0), \dots, (x_k, y_k)$
ovvero intercala la f. incognita (caso poco
probabile)

ESEMPIO

Calcoliamo la retta e la parabola che approssimano nel senso dei min-quadrati la tabella:

x	1	2	3	4	5
y	1.4	3.1	4.8	6.8	8.5

nel caso retta: $\hat{y}(x) = c_0 + c_1 x$
 $c_0(x) = 1$
 $c_1(x) = x$

$$n = 4, m = 1$$

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \in \mathbb{R}^{5 \times 2} = \mathbb{R}^{(k+1) \times (m+1)}, \quad y = \begin{bmatrix} 1.4 \\ 3.1 \\ 4.8 \\ 6.8 \\ 8.5 \end{bmatrix} \in \mathbb{R}^5 = \mathbb{R}^k$$

$$c = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} \in \mathbb{R}^{m+1} = \mathbb{R}^2$$

$$A^T A c = A^T y$$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 15 \\ 15 & 55 \end{bmatrix}$$

$$A^T y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1.4 \\ 3.1 \\ 4.8 \\ 6.8 \\ 8.5 \end{bmatrix} = \begin{bmatrix} 24.6 \\ 91.7 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 15 \\ 15 & 55 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 24.6 \\ 91.7 \end{bmatrix} \rightsquigarrow \begin{bmatrix} -0.45 \\ 1.79 \end{bmatrix}$$

$$\Phi_1(x) = 1.79x - 0.45$$

Nel caso volessimo appross. con una parabola:

$$\Phi_2(x) = c_0 + c_1 x + c_2 x^2$$

$$e_0(x) = 1, e_1(x) = x, e_2(x) = x^2$$

$$m=2, k=4$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix} \in \mathbb{R}^{(k+1) \times (m+1)} = \mathbb{R}^{5 \times 3}, \quad y = \begin{bmatrix} 1.4 \\ 3.1 \\ 4.8 \\ 6.8 \\ 8.5 \end{bmatrix} \in \mathbb{R}^{k+1} = \mathbb{R}^5$$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 9 & 16 & 25 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix} = \begin{bmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{bmatrix}$$

$$A^T y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 9 & 16 & 25 \end{bmatrix} \begin{bmatrix} 1.4 \\ 3.1 \\ 4.8 \\ 6.8 \\ 8.5 \end{bmatrix} = \begin{bmatrix} 24.6 \\ 91.7 \\ 378.3 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 945 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 24.6 \\ 91.7 \\ 378.3 \end{bmatrix} \rightsquigarrow \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -0.3 \\ 1.66 \\ 0.021\dots \end{bmatrix}$$

$$\Phi_2(x) = 0.021x^2 + 1.66x - 0.3$$

$$\Psi_1(c) = \sum_{j=0}^k (\Phi_1(x_j) - y_j)^2 = S_1 = 0.021$$

$$\Psi_2(c) = \sum_{j=0}^k (\Phi_2(x_j) - y_j)^2 = S_2 = 0.0205$$

$S_2 < S_1 \Rightarrow \Phi_2$ è un'appross. migliore di Φ_1

x_0, \dots, x_k y_0, \dots, y_k

$$\Phi(x) = \sum_{j=0}^m c_j e_j(x) \quad m \leq k$$

Problema: determinare i c_j che rendano $\Phi(x)$ una buona appross. nel senso dei minimi quadrati

Equivalentemente a risolvere

$$A = \begin{bmatrix} e_0(x_0) & \dots & e_m(x_0) \\ \vdots & & \vdots \\ e_0(x_k) & \dots & e_m(x_k) \end{bmatrix} \in \mathbb{C}^{(k+1) \times (m+1)}$$

$$\text{il sist. delle eq. matematici } A^T A c = y = \begin{bmatrix} y_0 \\ \vdots \\ y_k \end{bmatrix}$$

OSS Stiamo risolvendo $\min_{c \in \mathbb{R}^{m+1}} \|Ac - y\|_2 \iff \min_{c \in \mathbb{R}^{m+1}} \|A^T Ac - y\|_2^2$

Ci potremmo chiedere di minimizz. questa qta' anche indipendentemente dal problema di appross. di funzioni.

SISTEMI SOVRADETERMINATI

$$Ax = b \quad A \in \mathbb{R}^{m \times n} \quad m > n \quad A = \begin{array}{|c|c|} \hline & n \\ \hline & m \\ \hline \end{array}$$

È un sistema con più eq. che incognite; è molto prob. che non ci sia soluzione, ovvero: $\text{rk}(A|b) > \text{rk}(A) \Rightarrow \exists x \text{ che verifica } x \text{ esattamente}$

Quello che posso fare è cercare di minimizz. (rendere piccolo) il vettore $r = b - Ax \xrightarrow{\text{residuo}} \min \|r\|_2^2$

DEF x è una sol. di $Ax = b$ nel senso dei minimi.

quadrati se $\|Ax - b\|_2 = \min_{\tilde{x} \in \mathbb{R}^n} \|A\tilde{x} - b\|_2$

Per trovare x si deve risolvere il sist. delle eq. normali $A^T A = A^T b$

OSS La sol. nel senso dei minimi quadrati è unica quando $\text{rk}(A^T A) = n \Leftrightarrow \text{rk}(A) = n$
L'esistenza della sol. è sempre assicurata

ESEMPIO

Si calcoli la sol. di $\begin{cases} 2x_1 - x_2 = 1 \\ -x_1 + x_2 = 0 \\ x_1 + x_2 = 1 \end{cases}$ nel senso dei minimi quadrati.

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \Rightarrow A^T A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & -1 \\ -1 & 6 \end{bmatrix}$$

$$b = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \Rightarrow A^T b = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Per calcolare x , bisogna risolvere:

$$\begin{bmatrix} 6 & -1 \\ -1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \Leftrightarrow \begin{cases} 6x_1 - x_2 = 3 \\ -x_1 + 6x_2 = 1 \end{cases}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{35} \begin{bmatrix} 19 \\ 9 \end{bmatrix}$$
$$\downarrow \frac{1}{35} \begin{bmatrix} 19 \\ 9 \end{bmatrix}$$
$$\|r\|_2 = \|Ax - b\|_2 = \min_{\tilde{x} \in \mathbb{R}^n} \|A\tilde{x} - b\|_2$$

Esercizio

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \alpha & 1 \\ 0 & 0 & \alpha \\ \beta & 0 & 0 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \quad \alpha, \beta \in \mathbb{R}$$

Determinare per quali α e β la sol. del problema dei min. quadrati non è unica

Dobbiamo vedere quando $A^T A$ non ha rango max, ovvero quando A non ha rango max (3, il min. delle due dim. della mat.)

Avere rango 3 significa che esiste almeno una sottomat. 3×3 con $\det \neq 0$ (\Rightarrow non invertibile)

Im A ci sono 4 sottomat. 3×3 :

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & \alpha & 1 \\ 0 & 0 & \alpha \end{vmatrix} = \alpha(\alpha-1)$$

$$\begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & \alpha \\ \beta & 0 & 0 \end{vmatrix} = \beta\alpha$$

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & \alpha & 1 \\ \beta & 0 & 0 \end{vmatrix} = \beta(1-\alpha)$$

$$\begin{vmatrix} 1 & \alpha & 1 \\ 0 & 0 & \alpha \\ \beta & 0 & 0 \end{vmatrix} = \beta\alpha^2$$

Affinché $\text{rk}(A) \neq 3$:

$$\begin{cases} \alpha(\alpha-1) = 0 \\ \beta(1-\alpha) = 0 \\ \beta\alpha = 0 \\ \beta\alpha^2 = 0 \end{cases} \Rightarrow \begin{cases} \alpha = 0 \\ \beta = 0 \end{cases} \text{ oppure} \begin{cases} \beta = 0 \\ \alpha = 1 \end{cases}$$

ESEMPI

$$\begin{bmatrix} 1 & \alpha \\ \alpha & \alpha \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \alpha \in \mathbb{R}$$

Determinare per quali α la sol. è unica.

Dobbiamo vedere quando $\text{rk}(A) = 2$

Osserviamo i det delle sottomat. 2×2 :

$$\left| \begin{array}{cc} 1 & \alpha \\ \alpha & \alpha \end{array} \right| = \alpha(1-\alpha), \quad \left| \begin{array}{cc} 1 & \alpha \\ 1 & 1 \end{array} \right| = 1-\alpha, \quad \left| \begin{array}{cc} \alpha & \alpha \\ 1 & 1 \end{array} \right| = 0$$

$$\alpha(1-\alpha) \neq 0 \quad \text{oppure} \quad 1-\alpha \neq 0$$

\Updownarrow

$$\alpha \neq 1$$

Se $\alpha = 1$ tutti i det sono nulli $\Rightarrow \text{rk}(A) < 2$
 \Rightarrow C.N e C.S. per avere sol. unica è $\alpha \neq 1$

METODO QR PER RISOLVERE IL PROBLEMA AI MIN. QUADR.

$$A = QR \quad Q Q^H = Q^H Q = I \quad (\text{cioè } Q \text{ mat. unitaria})$$

$$Q \in \mathbb{C}^{m \times m}$$

$$A \in \mathbb{C}^{m \times n}$$

$$R \in \mathbb{C}^{m \times n}, \quad R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \quad \begin{array}{l} \text{n righe} \\ \text{m-n righe} \end{array} \quad m > n$$

$$R_1 = \begin{bmatrix} R_{11} \\ 0 \end{bmatrix}, \quad R_{11} \in \mathbb{C}^{n \times n} \text{ triang. sup.}$$

Il metodo di calcolo di Q e R è analogo al caso di $A \in \mathbb{C}^{m \times m}$: si moltiplica A a sx per n mat. di Householder.

$$A \rightarrow H_1 A = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & & \ddots & \\ 0 & * & \cdots & * \end{bmatrix}, H_2 H_1 A = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & 0 & \ddots & \\ 0 & * & \cdots & * \end{bmatrix} \rightarrow \dots \rightarrow H_n \dots H_1 A$$

Il calcolo di Q e R costa $\mathcal{O}(mn^2)$.

$$\begin{aligned} \min_{x \in \mathbb{C}^n} \|Ax - b\|_2^2 &= \min_{x \in \mathbb{C}^n} \|QRx - b\|_2^2 = \min_{x \in \mathbb{C}^n} \|Q^H(QRx - b)\|_2^2 = \\ &= \min_{x \in \mathbb{C}^n} \|Rx - \underbrace{Q^H b}_{c}\|_2^2 = \min_{x \in \mathbb{C}^n} \left\| \begin{bmatrix} R_1 \\ 0 \end{bmatrix} x - \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right\|_2^2 = \\ &= \min_{x \in \mathbb{C}^n} \left\| \begin{bmatrix} R_1 x - c_1 \\ c_2 \end{bmatrix} \right\|_2^2 = \min_{x \in \mathbb{C}^n} \underbrace{\|R_1 x - c_1\|_2^2}_{\downarrow \text{RESIDUE DI UN SIST. LIN. } n \times n} + \underbrace{\|c_2\|_2^2}_{\downarrow \text{NON DIP. DA } x} \end{aligned}$$

Se R_1 è invertibile (equivale a $\text{rk}(A) = n$)
 \Rightarrow la sol. ottima è $x = R_1^{-1}c_1$, ovvero la sol. di $R_1 x = c_1$ e il valore ottimo, cioè

$$\min_{x \in \mathbb{C}^n} \|Ax - b\|_2 = \|c_2\|_2$$

OSS $Q^H b$ garantisce il valore ottimo del problema, mentre per ottenere x si risolve il sistema $n \times n$ triangolare $R_1 x = c_1 \Rightarrow \mathcal{O}(\underbrace{mn^2}_\text{COSTO FATT. QR} + \underbrace{m^2}_\text{COSTO } Q^H b + \underbrace{n^2}_\text{COSTO PER RISOLV. IL SISTEMA } n \times n \text{ TRIANG. SUP})$

OSS

Il costo delle eq. normali è $O(\underbrace{mn^2}_{A^TA} + \underbrace{n^3}_{A^TAx = A^Tb})$
(un po' più basso, ma il
metodo è numeric. più
instabile)

Il condizionamento del sist. lin. $R_i x = c_i$ è la
radice quadrata di $A^T A x = A^T b$.

METODO QR PER PROBLEMI AI MINIMI QUADRATI

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$$

$$A \in \mathbb{C}^{m \times n}$$

$m > n \Rightarrow$ SISTEMA SOVRADETERMINATO

$Ax = b$ non ha soluzione

$$(rk(A) < rk(A(b)))$$

1° METODO DI DISSOLUZIONE: EQUAZIONI NORTELLI $A^T A x = A^T b$

2° $\sim \sim \sim \sim$: BASATO SUA FAOTORIZZ. QR

Assumiamo $rk(A) = n$ (rango massimo)

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2 \quad A \in \mathbb{C}^{m \times n}$$

$$A = QR \quad Q \in \mathbb{C}^{m \times m}$$

$$R \in \mathbb{C}^{m \times n}, \quad Q^H Q = Q Q^H = I$$

$R = \begin{cases} \begin{matrix} \star & & \\ 0 & \ddots & \\ \vdots & & \end{matrix} & n \text{ righe} \\ \begin{matrix} 0 & & \\ 0 & \ddots & \\ \vdots & & \end{matrix} & m-n \text{ righe} \end{cases}$

$$R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \quad R_1 \in \mathbb{C}^{n \times n} \text{ triang. sup.}$$

A rango massimo ($rk(A) = n$)

\Rightarrow gli elem. diagonali di
R sono tutti $\neq 0$

$$H_1 \cdot A = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{bmatrix}, H_2 \cdot H_1 \cdot A = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & * & * \end{bmatrix}, \dots H_n \cdots H_2 H_1 \cdot A = R = \begin{bmatrix} * & \cdots & * \\ 0 & \ddots & * \end{bmatrix}$$

$$\min_{x \in \mathbb{C}^n} \|Ax - b\|_2 \iff \min_{x \in \mathbb{C}} \|Ax - b\|_2^2$$

$$\|Ax - b\|_2^2 = \|QRx - b\|_2^2 = \|Q^H(QRx - b)\|_2^2 = \|Rx - Q^H b\|_2^2$$

PER L'INNANZIETÀ
DELLA $\|\cdot\|$ EUCLIDEA PER MATR. CON Q^H

$$\begin{aligned}
 &= \|R_i x - c_i\|_2^2 = \left\| \begin{bmatrix} R_{i1} \\ 0 \end{bmatrix} x - \begin{bmatrix} c_{i1} \\ c_{i2} \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} R_{i1}x \\ 0 \end{bmatrix} - \begin{bmatrix} c_{i1} \\ c_{i2} \end{bmatrix} \right\|_2^2 = \\
 &= \left\| \begin{bmatrix} R_{i1}x - c_{i1} \\ -c_{i2} \end{bmatrix} \right\|_2^2 = \underbrace{\|R_{i1}x - c_{i1}\|_2^2}_{\substack{\text{MINIMIZZARE} \\ \text{QUESTA QTA'} \\ \text{RISPETTO A } x}} + \underbrace{\|c_{i2}\|_2^2}_{\substack{\text{COSTANTE RISPETTO A } x \\ \text{E' EQUIV-A} \\ \text{RISOLVERE IL SISTEMA LIN.}}} \\
 &\quad R_{i1}x = c_{i1} \Rightarrow x = R_{i1}^{-1}c_{i1} \\
 &\quad \text{SOL. DI UN SISTEMA TRIANG.}
 \end{aligned}$$

- x si trova come soluzione di $R_i x = c_i$, dove R viene dalla fattorizzazione QR di A e c_i da $Q^H b$ (prime n componenti di $Q^H b$)
- $\|c_{i2}\|_2 = \min_{\tilde{x} \in \mathbb{R}^n} \|A \tilde{x} - b\|_2$ si ottiene considerando le ultime $m-n$ componenti di $Q^H b$

Oss Sia il metodo delle eq. normali che il metodo QR forniscano la soluzione del problema, ma per problemi grandi (es. $m = 10^4$, $n = 30$) si preferisce il metodo QR perché è più stabile numericamente.
Più precisamente, si ha che il num. di condizionamento:

$$\text{cond}(A^T A) = \text{cond}(R_i)^2$$

$$A^T A x = \overset{\swarrow}{A^T} b \qquad \qquad \qquad \downarrow R_i x = c_i$$

Se A è vicina ad essere di rk $< n$ allora il metodo delle eq. normali può dare problemi (problema mal condizionato)

ESEMPIO

$$A = \begin{bmatrix} 3 & 3 \\ 4 & 4 \\ 0 & 10^{-10} \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow A^T A = \begin{bmatrix} 25 & 25 \\ 25 & 25 + 10^{-20} \end{bmatrix}$$

$$b = \begin{bmatrix} 1 \\ 1 + 10^{-10} \end{bmatrix}$$

$\text{rk}(A) = 2$, ma è vicina ad essere di rango 1

In MATLAB, con doppia precisione ($U \approx 10^{-16}$)

$$A^T A = \begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix} \quad \text{ha } \text{rk} = 1 \text{ ed il sistema delle eq. normali non ha soluzioni}$$

ESEMPIO

Risolvere il problema ai min. quadrati

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2 \quad \text{com} \quad A = \begin{bmatrix} 3 & -6 \\ 4 & -8 \\ 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

sapendo che $A = QR$ dove $Q = \begin{bmatrix} 3/\sqrt{5} & 0 & -4/\sqrt{5} \\ 4/\sqrt{5} & 0 & 3/\sqrt{5} \\ 0 & 1 & 0 \end{bmatrix}$ $m=3$
 $n=2$

Possiamo trovare R come $R = Q^H A : \begin{bmatrix} 5 & 10 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$

$$\text{com } R_1 = \begin{bmatrix} 5 & 10 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{n \times n} = \mathbb{R}^{2 \times 2}$$

$$c = Q^H b = \begin{bmatrix} 3/\sqrt{5} & 4/\sqrt{5} & 0 \\ 0 & 0 & 1 \\ -4/\sqrt{5} & 3/\sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 5 \end{bmatrix} \Rightarrow c_1 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

$$c_2 = [5]$$

$$S = \min_{\tilde{x} \in \mathbb{R}^2} \|A \tilde{x} - b\|_2$$

Se $x \in \mathbb{R}^2$ che dà il residuo minimo è la sol. di:

$$\begin{bmatrix} 5 & 10 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix} \Leftrightarrow \begin{cases} 5x_1 + 10x_2 = 5 \\ x_2 = 2 \end{cases} \Leftrightarrow \begin{cases} x_1 = -3 \\ x_2 = 2 \end{cases}$$

INTEGRAZIONE NUMERICA

Si guarda al problema di approssimare

$$\int_a^b f(x) p(x) dx$$

$[a, b] \subset \mathbb{R}$ intervallo

$f: \mathbb{R} \rightarrow \mathbb{R}$ continua

(dopo ci servirà più regolarità, cioè l'esist. di derivate di ordine superiore)

$p(x) \geq 0$ su $[a, b]$

Nel caso $p \equiv 1$ si ritrova l'integrale definito su f .

Peché non calcolare l'integrale esattamente?

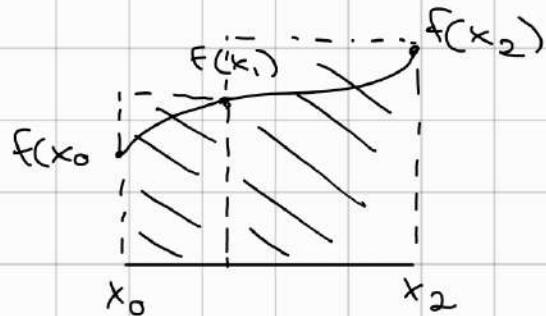
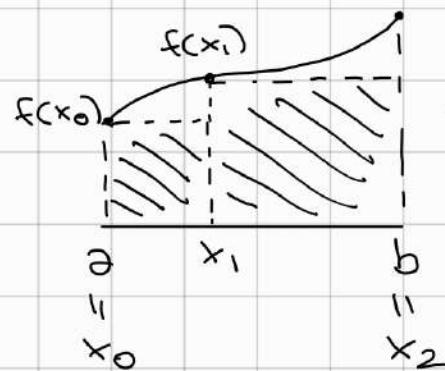
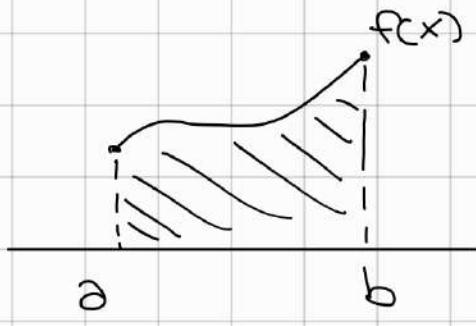
- ① Non tutte le funzioni hanno una primitiva che si scrive in termini di funzioni elementari (es. $e^{-x^2/2}$, $e^{\cos x}$ che danno origine a integrali ellittici)
- ② Anche se la primitiva esiste, può avere un'espressione molto complicata e quindi costosa da trovare e valutare
- ③ Tant'è che f è disponibile solo in maniera implicita, ovvero si possono conoscere solo un num. finito di valori $f(x_i)$ $i = 0, \dots, n$

(es. come nel problema dell'interpolazione)

$$I(\varphi f) = \int_a^b f(x) \varphi(x) dx \approx \sum_{i=0}^n \underbrace{\alpha_i}_{\text{NODI DELLA FORMULA}} f(x_i) \quad \begin{array}{l} \text{FORMULA DI QUADRATURA} \\ \downarrow \\ \text{RESI DELLA FORMULA} \end{array}$$

IDEA

DEF. DI INTEGRALE SECONDO RIEMANN



$$\begin{aligned} & f(x_0) \cdot \underbrace{\alpha_0}_{(x_1 - x_0)} + f(x_1) \cdot \underbrace{\alpha_1}_{(x_2 - x_1)} \\ & f(x_1) \cdot (x_1 - x_0) + f(x_2) \cdot (x_2 - x_1) \end{aligned}$$

APPROSS. PER ECESSO

DEF Una formula di quadratura è del tipo:

$$J_n(f) := \sum_{i=0}^n \alpha_i f(x_i)$$

$$x_0 < x_1 < \dots < x_n \in [a, b]$$

$$\alpha_0, \dots, \alpha_n \in \mathbb{R}$$

$$E_n(f) := \int_a^b f(x) \varphi(x) dx - J_n(f)$$

N.B.: gli α_i non hanno nulla a che vedere con φ

errore di appross.
estremo sx
dell'intervalle

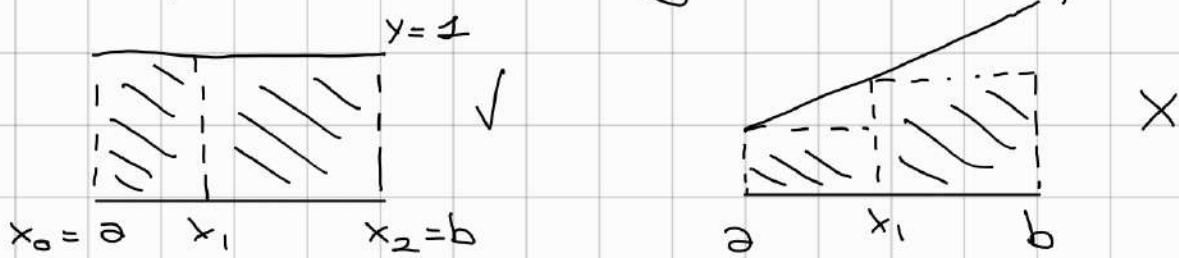
DEF

$\int_n(\cdot)$ formula di quadratura, si dice che ha precisione algebrica $m \in \mathbb{N}$ se verifica:

$$E_n(1) = E_n(x) = \dots = E_n(x^m) = 0, E_n(x^{m+1}) \neq 0$$

ESEMPIO (FORMULE CON RETTANGOLI INSCRITTI E CIRCONSCRITTI)

La precisione è di grado 0:



ESEMPIO

$$\int_{-1}^1 f(x) dx \quad a = -1, b = 1, f(x) \equiv 1$$

Fissati $x_0 = -1$ e $x_1 = 1$, qual è il grado max di una formula di quadratura con due soli punti $x_0 = -1, x_1 = 1$?

$$\int_n(f) = \sum_{i=0}^1 a_i \cdot f(x_i) = a_0 f(-1) + a_1 f(1)$$

Dobbiamo determinare a_0 e a_1 , in modo da max il grado di precisione:

$$E_1(1) = 0 \Rightarrow \underbrace{a_0 + a_1}_{\int_{-1}^1 1 dx = 2} = \int_{-1}^1 1 dx = 2 \Rightarrow a_0 + a_1 = 2$$

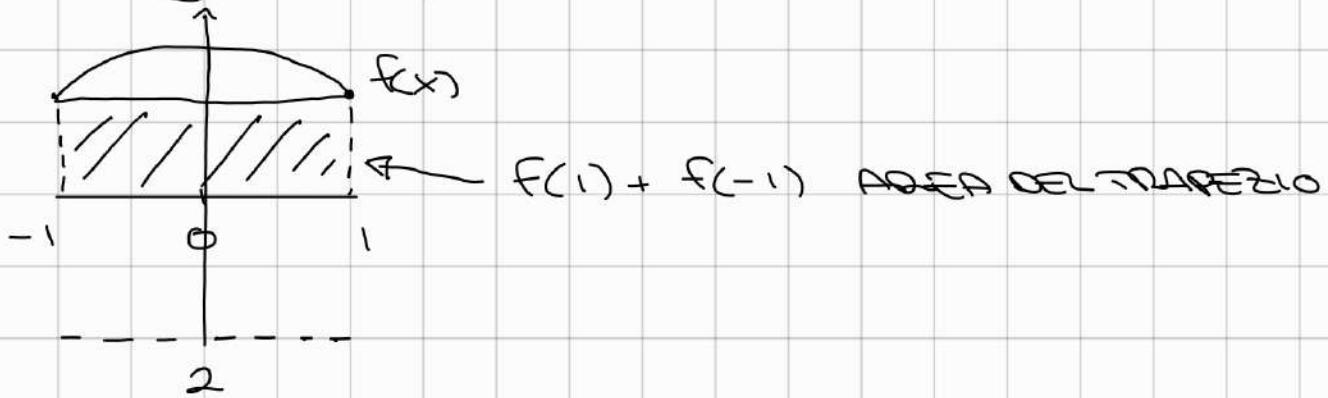
$$E_1(x) = 0 \Rightarrow \underbrace{\int_a^x}_{J_n(x)} - a_0 + a_1 = \int_a^x dx = 0 \Rightarrow -a_0 + a_1 = 0$$

$$\begin{cases} a_0 + a_1 = 2 \\ a_1 = a_0 \end{cases} \Rightarrow \begin{cases} a_0 = 1 \\ a_1 = 1 \end{cases} \Rightarrow J_1(f) = f(-1) + f(1)$$

Sia grado di precisione
almeno 1

$$E_1(x^2) = a_0 + a_1 - \int_a^1 x^2 dx = a_0 + a_1 - 2/3 = 2 - 2/3 = 4/3 \neq 0$$

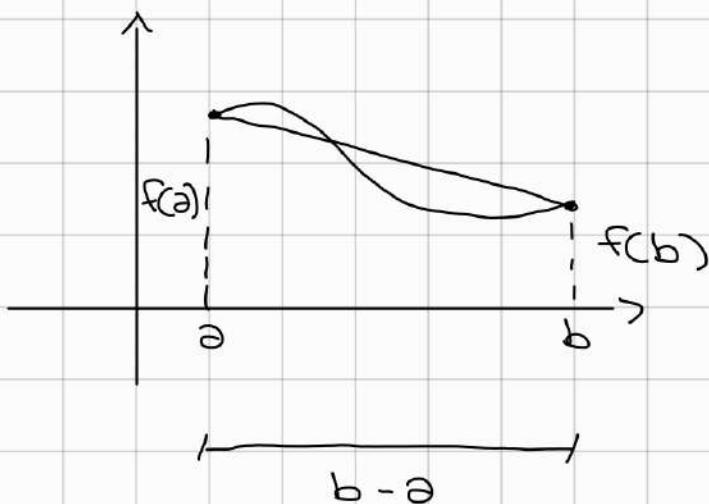
\Rightarrow Il grado di precisione è esattamente 1



Per un generico intervallo $[a, b]$ (al posto di $[-1, 1]$)
si ha:

$$\int_a^b f(x) dx \approx \frac{b-a}{2} (f(a) + f(b))$$

ALTEZZA
 ↑
 BASE MAGG
 ↑ + BASE MIN.
 FORMULA DEL TRAPEZIO



ESEMPIO

Aggiungiamo il pto medio dell' intervallo e cerchiamo di massimizzare la precisione

su $[-1, 1]$ significa $x_0 = -1, x_1 = 0, x_2 = 1$

$$J_2(f) = \alpha_0 f(-1) + \alpha_1 f(0) + \alpha_2 f(1)$$

Dobbiamo trovare $\alpha_0, \alpha_1, \alpha_2$ in modo che la precis. sia max.

Si impongono le condizioni:

$E_2(1) = E_2(x) = E_2(x^2) = 0 \rightarrow 3$ condizioni per
3 parametri

$$\begin{cases} \alpha_0 + \alpha_1 + \alpha_2 = \int_{-1}^1 1 dx = 2 \\ -\alpha_0 + \alpha_2 = \int_{-1}^1 x dx = 0 \quad \Leftrightarrow \quad \begin{cases} \alpha_0 + \alpha_1 + \alpha_2 = 2 \\ -\alpha_0 + \alpha_2 = 0 \Rightarrow \alpha_0 = \alpha_2 \\ \alpha_0 + \alpha_2 = 2/3 \Rightarrow \alpha_0 = \alpha_2 = 1/3 \end{cases} \\ \alpha_0 + \alpha_2 = \int_{-1}^1 x^2 dx = 2/3 \end{cases}$$

$$\Rightarrow \begin{cases} \alpha_1 = 2 - 2/3 = 4/3 \\ \alpha_0 = \alpha_2 = 1/3 \end{cases}$$

FORMULA DI SIMPSON PER $[-1, 1]$

$$\Rightarrow J_2(f) = \frac{1}{3} (f(-1) + 4f(0) + f(1))$$

ha precisione almeno 2

$$E_2(x^3) = 0 - \int_{-1}^1 x^3 dx = 0 - 0 = 0 \Rightarrow \text{ha precisione almeno 3}$$

$$E_2(x^4) = \frac{2}{3} - \frac{2}{5} \neq 0 \Rightarrow \text{la precisione algebrica è esattamente 3}$$

Per un generico $[a, b]$ la formula di Simpson prende la forma:

$$\boxed{S_2(f) = \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right)}$$

↓
 PRO MEIO DI $[a, b]$

ATTENZIONE: nell'esempio abbiamo calcolato $E_n(f)$ come $S_n - \int_a^b f(x) dx$ sebbene lo avessimo definito come:
 $E_n(f) := S_n - I_n(f)$

Ai fini di trovare i coeff. con una det. precisione non cambia nulla.

OSS Dire che $I_n(f)$ ha precisione $m \in \mathbb{N}$ significa che $I_n(f)$ calcola esattam. l'integrale quando f è un polinomio di al più grado m

FUNZIONE
+ PESO

$$I(\rho f) = \int_a^b f(x) \underbrace{\rho(x)}_{\geq 0} dx \quad f \in C([a, b])$$

$\rho(x) \geq 0$ su $[a, b]$ (molto spesso $\rho \equiv 1$)

$$\int_a^b x^k \rho(x) dx = m_k < \infty \quad \begin{array}{l} \text{MOMENTO} \\ \text{DI ORDINE } k \\ (k \in \mathbb{N}) \end{array}$$

$\rho(x) =$

$$J_n(f) \approx I(\rho f) \text{ dove } J_n(f) = \sum_{j=0}^n a_j \cdot \underbrace{f(x_j)}_{\substack{\text{PESI} \\ \text{(NUM. REALI)}}} \quad x_0, \dots, x_n \in [a, b]$$

$x_0 < x_1 < \dots < x_n$

Imporre che l'errore $E_n(f) = \int_a^b f \rho dx - J_n(f)$

sia uguale a zero per $f = 1, x, \dots, x^n$ identifica una ed una sola formula di quadratura.

Più precisamente: fissati x_0, \dots, x_n in $[a, b]$ esiste una formula di quadratura $J_n(f)$ con questi nodi che ha grado di precisione almeno n .

(grado di prec. 1)

$x_0 = a, x_1 = b \rightsquigarrow$ FORMULA DI TRAPEZIO

$x_0 = a, x_1 = \frac{a+b}{2}, x_2 = b \rightsquigarrow$ FORMULA DI SIMPSON

(che ha grado di prec. 3)

Imporre $E_n(1) = E_n(x) = \dots = E_n(x^n) = 0$ equivale a risolvere un sist. lin. con soluzione unica:

MAT. DI VANDERMONDE (TRASPOSTA)

$$\begin{cases} a_0 + a_1 + \dots + a_n = m_0 \\ a_0 x_0 + a_1 x_1 + \dots + a_n x_n = m_1 \\ \vdots \\ a_0 x_0^n + a_1 x_1^n + \dots + a_n x_n^n = m_n \end{cases} \Leftrightarrow \begin{bmatrix} 1 & \dots & 1 \\ x_0 & \dots & x_n \\ \vdots & & \vdots \\ x_0^n & \dots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} m_0 \\ \vdots \\ m_n \end{bmatrix}$$

$$S_n(x^k) = \int_a^b x^k dx = m_k$$



che ha prec. $\geq n$

Questa mat. è invertibile
 $(x_i \neq x_j, i \neq j) \Rightarrow$ unica sol.
 \Rightarrow unica scelta di a_0, \dots, a_n

Il problema più generale è: dato n , qual'è la scelta ottimale sia dei param. a_0, \dots, a_n che dei param. x_0, \dots, x_n per avere la max precisione.

IDEA

Devo identificare $2n+2$ parametri, mi serviranno almeno $2n+2$ equazioni.

Quindi si impongono:

$$E_n(1) = E_n(x) = \dots = E_n(x^{2n+1}) = 0$$

$$\begin{cases} a_0 + \dots + a_n = m_0 \\ a_0 x_0 + \dots + a_n x_n = m_1 \\ \vdots \\ a_0 x_0^{2n+1} + \dots + a_n x_n^{2n+1} = m_{2n+1} \end{cases}$$

Il sistema è quello di prima, ma con 2 diff.:

1) ho $2n+2$ eq. anziché $n+1$

2) non è un sistema lineare

Non è scontato, ma la sol. di questo sistema esiste ed è unica.

ESEMPIO (CASO $n=1$)

$$[a, b] = [-1, 1]$$

Con $x_0 = -1$, $x_1 = 1$ otterremo gli stessi di prec. 1.

$$\int_n(x^k) = m_k \quad k = 0, 1, 2, 3$$

$$\begin{cases} \partial_0 + \partial_1 = \int_{-1}^1 1 dx = 2 \\ \partial_0 x_0 + \partial_1 x_1 = \int_{-1}^1 x dx = 0 \\ -\partial_0 x_0^2 + \partial_1 x_1^2 = 2/3 \\ \partial_0 x_0^3 + \partial_1 x_1^3 = 0 \end{cases} \Rightarrow \begin{cases} \partial_0 = 2 - \partial_1 \\ (2 - \partial_1)x_0 + \partial_1 x_0 = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \partial_0 = 2 + \frac{2x_0}{x_1 - x_0} = \frac{2x_1}{x_1 - x_0} \\ \partial_1 = -\frac{2x_0}{x_1 - x_0} \end{cases}$$

Sostituendo ∂_0 e ∂_1 nelle ultime due eq.:

$$\begin{cases} \frac{2x_1}{x_1 - x_0} \cdot x_0^2 - \frac{2x_0}{x_1 - x_0} \cdot x_1^2 = 2/3 \\ \frac{2x_1}{x_1 - x_0} \cdot x_0^3 - \frac{2x_0}{x_1 - x_0} x_1^3 = 0 \Rightarrow \frac{2x_0 x_1}{x_1 - x_0} (x_0^2 - x_1^2) = 0 \end{cases} \Rightarrow$$

$$\begin{cases} \frac{2x_0 x_1}{x_1 - x_0} (x_0 - x_1) = 2/3 \\ x_0 = \underline{\pm x_1} \text{ oppure } x_0 = 0, x_1 \neq 0 \text{ oppure } x_0 \neq 0, x_1 = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \frac{2x_0 x_1}{x_1 - x_0} (x_0 - x_1) = 2/3 \\ x_0 = -x_1 \end{cases} \xrightarrow{\text{DO SCRITTA E PERCHÉ OTTENUTO}} 0 = \frac{2}{3}$$

$$\Rightarrow \begin{cases} -\frac{2x_1^2}{2x_1} (-2x_1) = 2/3 \\ x_0 = -x_1 \end{cases} \Rightarrow \begin{cases} 2x_1^2 = 2/3 \\ x_0 = -x_1 \end{cases} \Rightarrow \begin{cases} x_1 = \pm \sqrt{\frac{1}{3}} = \pm \frac{\sqrt{3}}{3} \\ x_0 = -x_1 \end{cases}$$

Si prende $x_0 = -\frac{\sqrt{3}}{3}$, $x_1 = \frac{\sqrt{3}}{3}$

$$\begin{cases} a_0 = \frac{2x_1}{x_1 - x_0} = 1 \\ a_1 = \frac{-2x_0}{x_1 - x_0} = 1 \end{cases}$$

$$\Rightarrow S_1(f) = 1 \cdot f\left(-\frac{\sqrt{3}}{3}\right) + 1 \cdot f\left(\frac{\sqrt{3}}{3}\right)$$

Formula di quadratura con prec. massima con 2 nodi su $[-1, 1]$

Il grado di precisione è ≥ 3 .

$$E_1(x^4) = \int_{-1}^1 x^4 dx - S_1(x^4) = \frac{2}{5} - \left(\frac{1}{9} + \frac{1}{9}\right) = \frac{2}{5} - \frac{2}{9} \neq 0$$

\Rightarrow il grado di precisione è esattamente 3

TEOREMA DI PEANO

$f \in C^{m+1}([a, b])$, $S_n(f)$ con grado di precisione m , allora:

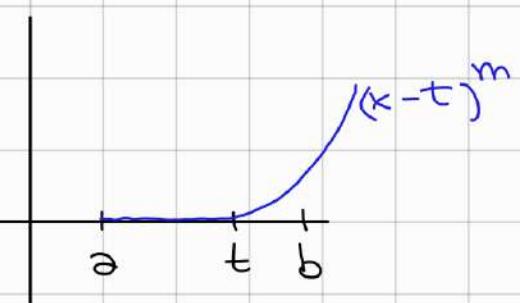
$$I(f) - S_n(f) = E_n(f) = \frac{1}{m!} \int_a^b f^{(m+1)}(t) \cdot G(t) dt$$

dove $\underbrace{G(t)}_{\downarrow} = E_n(S_m(x-t))$

NUCLEO
DI PEANO

$$S_m(x-t) = \begin{cases} (x-t)^m & x \geq t \\ 0 & \text{altrimenti} \end{cases}$$

IL VARIARE DI t DETERMINA
DOVE INIZIA L'ARCO DI PARABOLA



$$G(t) = \int_a^b s_m(x-t) dx - \mathcal{J}_n(s_m(x-t)) =$$

$$= \int_t^b (x-t)^m dx - \underbrace{\mathcal{J}_n(s_m(x-t))}_{\text{ESPRESSIONE CHE DIPENDE SOLO DA } t, \text{ NON DA } x}$$

$G: [a, b] \rightarrow \mathbb{R}$

Nel caso in cui $G(t)$ non cambia segno in $[a, b]$ (cioè sto approssimando l'integrale di $s_m(x-t)$ sempre per difetto o per eccesso, quando uso \mathcal{J}_n) allora, per il th della media integrale:

$$(*) E_n(f) = \frac{1}{m!} f^{(m+1)}(\xi) \int_a^b G(t) dt \quad \text{per qualche } \xi \in (a, b)$$

Inoltre si può dire "esattamente" chi è $\int_a^b G(t) dt$, per farlo sostituiamo in (*) $f = x^{m+1}$:

$$E_n(x^{m+1}) = \frac{(m+1)!}{m!} \cdot \int_a^b G(t) dt \Rightarrow \int_a^b G(t) dt = \frac{E_n(x^{m+1})}{m+1}$$

per cui si ha:

$$E_n(f) = \frac{f^{(m+1)}(\xi)}{(m+1)!} \cdot E_n(x^{m+1})$$

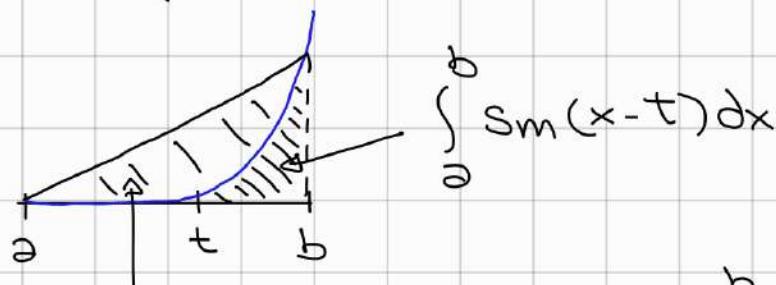
S. PUÒ CALCOLARE
STIMARE

S. PUÒ CALCOLARE
ESPLICATIVAMENTE

ESEMPIO (con la formula dei trapezi)

CASO $[a, b] = [-1, 1]$

Osserviamo che $G(t) \leq 0$ per la formula dei trapezi



$$(f(a) + f(b)) \frac{(b-a)}{2} > \int_a^b s_m(x-t) dx \text{ sempre}$$

\Rightarrow Sempre stima per eccesso \Rightarrow il segno di $G(t)$ non varia

$$\Rightarrow E_1(f) = \frac{f''(\varepsilon)}{2} \cdot E_1(x^2) = \frac{f''(\varepsilon)}{2} \cdot -\frac{4}{3} =$$

$$= -\frac{2}{3} f''(\varepsilon) \quad \text{per } \varepsilon \in (-1, 1)$$

Su un generico intervallo $[a, b]$:

da calc. su $[a, b]$

$$E_1(f) = \frac{f''(\varepsilon)}{2} \cdot \overbrace{E_1(x^2)}$$

$$E_1(x^2) = \int_a^b x^2 dx - \frac{b-a}{2}(a^2+b^2)$$

$$= \frac{b^3 - a^3}{3} - \frac{(b-a)(a^2+b^2)}{2}$$

$$= \frac{2b^3 - 2a^3 - 3b^3 + 3a^3 - 3a^2b - 3ab^2}{6}$$

$$= -\frac{(b-a)^3}{6}$$

$$\Rightarrow E_1(f) = -\frac{(b-a)^3}{12} \cdot f''(\varepsilon)$$

$$\varepsilon \in (a, b)$$

ESEMPIO (con la formula di Simpson)

$$\int_{-1}^1 f(x) dx \approx \frac{1}{3} (f(-1) + 4f(0) + f(1))$$

Anche in questo caso $G(t)$ non cambia segno
 \Rightarrow si può usare la formula per l'errore
 semplificata

su $[-1, 1]$

$$E_2(f) = \frac{f^{(4)}(\varepsilon)}{4!} \cdot E_2(x^4) \quad \text{poiché Simpson ha prec.m=3}$$

$$E_2(x^4) = \int_1^1 x^4 dx - \frac{1}{3} (1+0+1) = \frac{2}{5} - \frac{2}{3} = -\frac{4}{15}$$

$$\Rightarrow E_2(f) = -\frac{4}{15} \cdot \frac{f^{(4)}(\varepsilon)}{24} = -\frac{1}{90} f^{(4)}(\varepsilon) \quad \varepsilon \in (-1, 1)$$

Su $[a, b]$:

$$E_2(f) = -\frac{(b-a)^5}{2880} f^{(4)}(\varepsilon) \quad \varepsilon \in (a, b)$$

FORMULE INTERPOLATORIE

Per approssimare $\int_a^b f(x) \varphi(x) dx$ con certi modi

x_0, \dots, x_n potremmo considerare $P_n(x)$ polinomio
 di interpol. di f , ovvero t.c.:

$$P_n(x_j) = f(x_j) \quad \forall j = 0, 1, \dots, n$$

e considerare $\int_a^b P_n(x) \varphi(x) dx \approx I(\varphi f)$

Per quanto visto sull'interp. possiamo scrivere
 $f(x) = P_n(x) + R_n(x)$ ed inoltre $P_n(x) = \sum_{j=0}^n f(x_j) \cdot l_j(x)$

dove $l_j(x) = \frac{\prod_{i=0, i \neq j}^{n-1} (x - x_i)}{x_j - x_i}$

$$\begin{aligned} \int_a^b g(x) f(x) dx &= \int_a^b g(x) P_n(x) dx + \int_a^b g(x) R_n(x) dx = \\ &= \int_a^b \sum_{j=0}^n f(x_j) l_j(x) \cdot g(x) dx + \underbrace{\int_a^b g(x) R_n(x) dx}_{\text{RESTO}} = \\ &= \sum_{j=0}^n f(x_j) \underbrace{\int_a^b g(x) l_j(x) dx}_{\text{PESI}} + \underbrace{\int_a^b g(x) R_n(x) dx}_{\text{RESTO}} \end{aligned}$$

FORTUNA DI QUADRATURA

Integrare il polim. di interp. fornisce una formula di quadratura con modi i modi usati per l'interp. e pesi le qta $a_j = \int_a^b g(x) l_j(x) dx$

OSS Siccome $P_n(x) = f(x)$ ogni volta che f è un polim. di grado $\leq n \Rightarrow$ ogni formula interp. ha grado di precis. almeno n (se si usano $n+1$ modi) \Rightarrow è la formula di quadre. più precisa con quella scelta di modi \Rightarrow TRAPEZI e SIMSON sono formule di interp.

$$\int_a^b f(x) dx \text{ oppure } \int_a^b f(x) p(x) dx$$

(in genere $p \equiv 1$)

x_0, \dots, x_n modi dove conosciamo $f(x_0), \dots, f(x_n)$

POSSIBILE STRATEGIA: integrare il polinomio P_n di interpolazione su x_0, \dots, x_n

$$\int_a^b f(x) dx \approx \sum_{j=0}^n \underbrace{\omega_j}_{\substack{\text{RESI} \\ \uparrow}} f(x_j) \quad \text{dove } \omega_j = \int_a^b l_j(x) dx$$

FORMULA DI QUADRATURA INTERPOLATORIA

$l_i(x) = \text{POLIN. I-ES. NELLA BASE DI LAGRANGE}$

$$= l_i(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j}$$

OSS

- i) fissati x_0, \dots, x_n esiste ed è unica la formula di quadratura che ha prec. $\geq n$ (perché i coeff. ω_j risolvono un sist. di tipo Vandermonde)
- ii) In ha prec. $\geq n$ è equiv. a dire che la formula in integra esattam. polinomi con grado $\leq n$
- iii) se $P_n(x)$ è il polim. di interp. di f su x_0, \dots, x_n e $f(x)$ è un polim. di grado $\leq n$ allora $f \equiv P_n$ (perché $f - P_n$ è un polim. di grado $\leq n$ che si annulla in x_0, \dots, x_n
 $\Rightarrow f - P_n \equiv 0$)

iv) la formula interp. su x_0, \dots, x_n integra esattam. i polinomi di grado $\leq n$

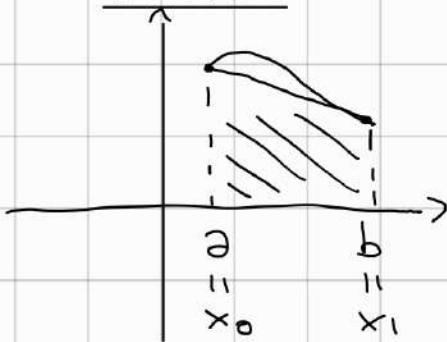
CONCLUSIONE: i) + ii) + iii) + iv):

la formula interp. su x_0, \dots, x_n è quella che massimizza la precisione una volta che i nodi sono fissati

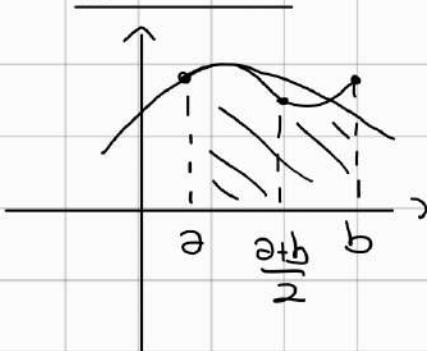
OSS

Trapezi e Simpson sono formule interp. con $n=1$ e $n=2$, rispettivamente

TRAPEZI



SIMPSON



INTEGRALE
DELLA PARABOLA
CHE INTERPOLA
IN $x_0 = a$,
 $x_1 = \frac{a+b}{2}$, $x_2 = b$

Classi di formule interp. si definiscono una volta che si sceglie la strategia con cui si scelgono i nodi.

FORMULE DI NEWTON - COTES

Si scelgono i nodi equisp. su $[a, b]$

$$x_0 = a, x_1 = a+h, x_2 = a+2h, \dots, x_j = a+jh, \dots, x_n = a+nh = b$$

$$h = \frac{b-a}{n} = \text{PASO DI INTEGRAZIONE}$$

OSS

Trapezi e Simpson sono le prime 2 formule di Newton-Cotes

TRAPEZI $\rightarrow n=1, h = b-a$, 2 nodi (x_0, x_1)

SIMPSON $\rightarrow n=2, h = \frac{b-a}{2}$, 3 nodi (x_0, x_1, x_2)

FATTO Per le formule di Newton-Cotes si può dim.
che il mulo di Peano associato

$$E_n(S_m(x-t)) = \int_a^b (x-t)^m dx - S_n(S_m(x-t))$$

non cambia segno.

$$\Rightarrow E_n(f) = \frac{f^{(m+1)}}{(m+1)!} \underbrace{E_n(x^{m+1})}_{\text{CALCABILE ESPlicitamente}} \quad \varepsilon \in (a, b)$$

In particolare, l'errore assume la forma:

$$E_n(f) = \begin{cases} c_n \cdot h^{n+2} f^{(n+1)}(\varepsilon) & n \text{ dispari} \\ c_n \cdot h^{n+3} f^{(n+2)}(\varepsilon) & n \text{ pari} \end{cases}$$

ESEMPIO

TRAPEZI $n=1, h = b-a$ $\rightarrow - \frac{(b-a)^3}{12} \cdot f''(\varepsilon) = -\frac{1}{2} \cdot h^3 \cdot f''(\varepsilon)$

SIMPSON $n=2, h = \frac{b-a}{2}$ $\rightarrow - \frac{(b-a)^5}{2880} f^{(4)}(\varepsilon) = -\frac{1}{90} h^5 \cdot f^{(4)}(\varepsilon)$

$$\varepsilon \in (a, b)$$

OSS Se n è pari, E_n ed E_{n+1} hanno la stessa dipendenza da h ed in particolare la stessa precisione \Rightarrow se si può scegliere conviene sempre la formula con n pari (se la scelta è tra n e $n+1$)

ESEMPIO

(*)

$$\text{Per } n=3 \rightarrow J_n(f) = \frac{b-a}{8} [f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)]$$

$$E_n(f) = -\frac{3}{80} h^5 f^{(4)}(\varepsilon)$$

(*) FORMULA DEL $\frac{3}{8}$ O "PULCERETTA"

Per aumentare l' accuratezza si dovrebbe rendere piccolo h ovvero ($n = \frac{b-a}{h}$) prendere più nodi
 \Rightarrow considerare formule di Newton-Cotes con n grande

FATTO Quando $n > 7$, nelle formule di Newton-Cotes appaiono coeff. negativi \Rightarrow problemi di instabilità numerica
 \Rightarrow Si preferisce seguire un'altra strada per far decrescere $E_n(f)$

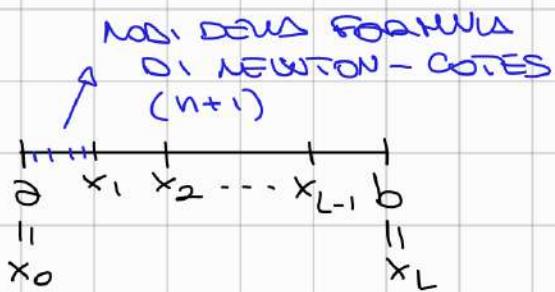
FORMULE DI NEWTON-COTES GENERALIZZATE

IDEA:

Spizzare $[a, b]$ in L sottointervalli (della stessa lungh.), ovvero si prendono x_0, \dots, x_L dove $x_j = a + j \cdot h$ con $h = \frac{b-a}{L}$ e si usa il fatto

che $\int_a^b f(x) dx = \sum_{i=1}^L \int_{x_{i-1}}^{x_i} f(x) dx$ e per ogni qd'

$\int_{x_{i-1}}^{x_i} f(x) dx$ uso una formula con $n \leq 7$, ad es. trapezi.



N. TOTALI DEI NODI:

$$(n+1)L - (L-1) = nL + 1$$

FORMULA DEL TRAPEZIO GENERALIZZATA

$$\int_a^b f(x) dx \approx \frac{b-a}{2L} \left[f(x_0) + 2 \sum_{j=1}^{L-1} f(x_j) + f(x_L) \right] = J_i^{(G)}$$

$$E_i^{(G)} = -\frac{(b-a)^3}{12L^2} f''(\varepsilon) \quad \varepsilon \in (a, b)$$

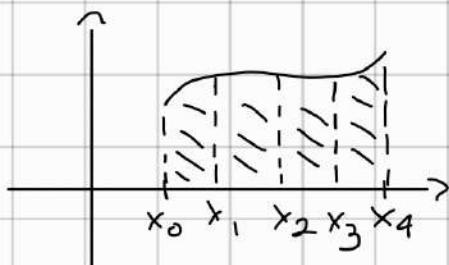
Può mandare a zero $E_i^{(G)}$ aumentando L, ovvero il num. di sottointervalli.

FORMULE DI NEWTON-COTES GENERALIZZATE

$$\int_a^b f(x) \approx J_n(f) = \sum_{i=0}^n a_i \cdot f(x_i)$$

Si divide $[a, b]$ in L sottointervalli della stessa ampiezza $\frac{b-a}{L}$; questo corrisponde a prendere i punti x_0, x_1, \dots, x_L con $x_0 = a, x_j = x_0 + j \cdot h$

$$\int_a^b f(x) dx = \sum_{i=1}^L \int_{x_{i-1}}^{x_i} f(x) dx$$



In ogni intervallo $[x_{i-1}, x_i]$ utilizzo una formula di Newton-Cotes ad $n+1$ modi ($n=1 \Rightarrow$ trapezi, $n=2 \Rightarrow$ Simpson).

$\overset{n+1 \text{ NODI}}{\text{P. ALL'INTERNO DI OGNI INTERVALLO}}$

$$x_0 \quad x_1 \quad \dots \quad x_L$$

In totale si hanno $L(n+1) - (L-1) = nL + 1$ modi

OSS Per $n=1$ (trapezi) i modi sono proprio x_0, \dots, x_L

Per $n=2$ (Simpson) i modi sono x_0, \dots, x_L ma

anche $\frac{x_{i-1} + x_i}{2}$ per $i=1, \dots, L$

$$x_{i-1} \quad \frac{x_{i-1} + x_i}{2} \quad x_i$$

ESEMPIO

$$\sum_{i=1}^L \underbrace{J_1^{(i)}(f)}_{\substack{\text{FORMULA DI TRAPEZI} \\ \text{PER } [x_{i-1}, x_i]}} = \sum_{i=1}^L \frac{b-a}{2L} (f(x_{i-1}) + f(x_i)) =$$

$$= \frac{b-a}{2L} (f(x_0) + f(x_L) + 2 \sum_{j=1}^{L-1} f(x_j)) \approx \int_a^b f(x) dx$$

$$\text{ERRORE: } \sum_{i=1}^L \frac{1}{12} \left(\frac{b-a}{L} \right)^3 \cdot f''(\xi_i) \quad \xi_i \in [x_{i-1}, x_i]$$

$$= \frac{(b-a)^3}{12L^2} \sum_{i=1}^L f''(\xi_i) = \frac{(b-a)^3}{12L^2} f''(\xi) \quad \xi \in [a, b]$$

$$= \boxed{h^2 \cdot \frac{b-a}{12} \cdot f''(\xi)} \quad \text{con } h = \frac{b-a}{L}$$

Si può fare lo stesso con la formula di Cavalieri - Simpson:

I nodi all'interno
di $[a, b]$ vengono
contati
 \uparrow
 L
2 volte

$$\int_a^b f(x) dx \approx \frac{b-a}{6L} \left[f(x_0) + f(x_L) + 2 \sum_{j=1}^{L-1} f(x_j) + 4 \sum_{j=1}^{\frac{L-1}{2}} f\left(\frac{x_j + x_{j+1}}{2}\right) \right]$$

OSS In genere si valuta il costo della formula
in base al num. di valutazioni di f e si ha
che per i trapezi sono necessarie $L+1$ valutazioni
mentre per Cavalieri - Simpson $2L+1$

$$\text{ERRORE PER CAVALIERI - SIMPSON: } \frac{(b-a)^3}{2880 L^4} f^{(IV)}(\xi) \quad \xi \in [a, b]$$

$$\frac{b-a}{2L} = h \rightarrow \frac{b-a}{180} \cdot h^4 \cdot f^{(IV)}(\xi)$$

ESEMPIO

Vogliamo calcolare $\int_0^1 \frac{1}{1+x} dx$ e supponiamo di volere un errore assoluto 10^{-4} (al massimo) in aritmetica esatta.

Quanti sottointervalli sono necessari per trapezi e Simpson? Dare una stima dal basso e confrontare il num. di valutazioni per le due formule.

Sia $E_1^{(G)}$ l'errore associato alla for. dei trapezi

$$E_1^{(G)} = \frac{1}{12L^2} \cdot f''(\varepsilon) \quad \varepsilon \in (0,1)$$

$$M_2 = \sup_{\varepsilon \in [0,1]} |f''(\varepsilon)| \Rightarrow |E_1^{(G)}| \leq \frac{M_2}{12L^2}$$

$$\frac{M_2}{12L^2} \leq 10^{-4} \Rightarrow L \geq \sqrt{\frac{10^4 M_2}{12}} = \sqrt{\frac{10^4}{6}} \approx 40.82 \dots$$

dove L è il num. di interv. da considerare
(imcoignita)

$$\boxed{L \geq 41}$$

$$f''(x) = \frac{2}{(1+x)^3} \Rightarrow M_2 = f''(0) = 2$$

41 sottointervalli $\Rightarrow 41 + 1 = 42$ valutazioni

Per Cavalieri - Simpson:

$$M_4 = \sup_{\varepsilon \in [0,1]} |f^{(IV)}(\varepsilon)|$$

$$f^{(IV)}(x) = \frac{24}{(1+x)^5} \Rightarrow M_4 = 24$$

$$\frac{M_4}{2880 \cdot L^4} = \frac{24}{2880 L^4} \leq 10^{-4} \Rightarrow L \geq \sqrt[4]{\frac{10000 \cdot 24}{2880}} \approx 3.02 \dots$$

SIMPSON
CONVERGE
PIÙ VELOCE

$\Rightarrow \boxed{L \geq 4}$ sottointervalli $\Rightarrow 2 \cdot 4 + 1 = 9$ valutazioni di f

ESTRAGGIAZIONE DI RICHARDSON

Nel caso non si sappia stimare o può l'errore (es. non abbiamo l'espressione esplicita di f) come si determina il num. di sottointervalli L ?

IDEA

Combinare valutazioni diverse della formula dei trapezi (diverse \rightarrow con num. diverso di sottointerv., cioè con passo h diverso) per ottenere un'appross. più accurata ed utilizzare quest'ultima per stimare l'errore.

TH Se $f(x) \in C^{2r+2}$ allora:

$$\text{Err}(h) = I(f) - \underbrace{\int_a^b f(x) dx}_{\substack{\text{FORTUNA} \\ \text{DEI TRAPEZI} \\ \text{CON PASSO } h}} = \alpha_1^{(1)} h^2 + \alpha_2^{(1)} h^4 + \dots + \alpha_r^{(1)} h^{2r} + O(h^{2r+2})$$

dove $\alpha_j^{(1)}$ sono i coeff. $\in \mathbb{R}$, non dipendono da h

Fissiamo $q \in \mathbb{N}$, $q > 1$ e consideriamo passi della forma $\frac{h}{q^j}$ $j = 0, 1, 2, \dots, N$ e chiamiamo $\mathcal{T}_j^{(1)}$ la

formula dei trapezi con passo $\frac{h}{q^j}$

ESEMPIO

$$\mathcal{T}_1^{(1)} = \frac{h}{2q} \left[f(x_0) + \sum_{i=1}^{Lq-1} f(x_i) + f(x_{Lq}) \right]$$

$$x_i = \underbrace{x_0}_{x_0} + i \frac{h}{q}$$

Per il q -th precedente, si puo' scrivere

$$\text{Err}\left(\frac{h}{q}\right) = \alpha_1^{(1)}\left(\frac{h}{q}\right)^2 + \alpha_2^{(1)}\left(\frac{h}{q}\right)^4 + \dots + \alpha_r^{(1)}\left(\frac{h}{q}\right)^{2r} + O(h^{2r+2})$$

Si possono combinare $\text{Err}(h)$ ed $\text{Err}\left(\frac{h}{q}\right)$ per eliminare i termini quadratici in h .

Più precisamente prendendo $\text{Err}(h) - q^2 \text{Err}\left(\frac{h}{q}\right)$ e dividendo per l'opportuna costante

$$I(f) - J_0^{(2)} = \alpha_2^{(2)} h^4 + \alpha_3^{(2)} h^6 + \dots + \alpha_r^{(2)} h^{2r} + O(h^{2r+2})$$

$$J_0^{(2)} = \frac{q^2 J_1^{(1)} - J_0^{(1)}}{q^2 - 1}$$

Formula di Romberg

Oss La formula di Romberg si ottiene combinando linearmente le formule dei trapezi generalizzate con passi h e h/q e ha un errore dell'ordine h^4

CONCLUSIONE Posso usare la valutazione della formula di Romberg come surrogato della sol. esatta per stimare l'errore della formula dei trapezi

$$|\text{Err}\left(\frac{h}{q}\right)| = |J_1^{(1)} - \int_a^b f(x) dx| \approx |J_1^{(1)} - J_0^{(2)}| \approx |J_1^{(1)} - \int_a^b f(x) dx| + |J_0^{(2)} - \int_a^b f(x) dx|$$

Il procedimento si può ripetere se abbiamo a disposizione ulteriori valutazioni dei trapezi con passi più piccoli.

Ad es. supponiamo di avere anche $\mathcal{J}_2^{(1)}$ TRAPEZI CON PASSO $\frac{h}{q^2}$

$$I(f) - \mathcal{J}_2^{(1)} = Err\left(\frac{h}{q^2}\right) = \alpha_1\left(\frac{h}{q^2}\right)^2 + \alpha_2\left(\frac{h}{q^2}\right)^4 + \dots + \alpha_r\left(\frac{h}{q^2}\right)^{2r} + O(h^{2r+2})$$

Analog. a prima, si combina $Err\left(\frac{h}{q^2}\right) \approx Err\left(\frac{h}{q}\right)$

per eliminare il termine $\alpha_1\left(\frac{h}{q}\right)^2$ -

$$\mathcal{J}_1^{(2)} = \frac{q^2 \mathcal{J}_2^{(1)} - \mathcal{J}_1^{(1)}}{q^2 - 1}$$

$$I(f) - \mathcal{J}_1^{(2)} = \alpha_2\left(\frac{h}{q}\right)^4 + \alpha_3\left(\frac{h}{q}\right)^6 + \dots + \alpha_r\left(\frac{h}{q}\right)^{2r} + O(h^{2r+2})$$

Combinando la rel. precedente con quella di $I(f) - \mathcal{J}_0^{(2)}$ si elimina il termine $\alpha^{(2)} \alpha^4$:

$$I(f) - \mathcal{J}_0^{(3)} = \alpha_3 h^6 + \alpha_4 h^8 + \dots + \alpha_r h^{2r} + O(h^{2r+2})$$

$$\mathcal{J}_0^{(3)} = \frac{q^4 \mathcal{J}_1^{(2)} - \mathcal{J}_0^{(2)}}{q^4 - 1}$$

In generale: fissato $N < r$ si calcola

$\mathcal{J}_0^{(1)}, \dots, \mathcal{J}_N^{(1)}$ (FORMULE DEI TRAPEZI CON PASSO $\frac{h}{q^j}$ $j=0, \dots, N$) RISATE COSTOSA

e poi si raffinano le stime dell'integrale

$$\text{calcolando } \mathcal{J}_i^{(k+1)} = \frac{q^{2k} \mathcal{J}_{i+1}^{(k)} - \mathcal{J}_i^{(k)}}{q^{2k} - 1} \quad k=1, \dots, N \\ i=0, 1, \dots, N-k$$

$J_i^{(k+1)}$ è un' appross. con errore $\left(\frac{h}{q_i}\right)^{2(k+1)}$

La già più accurata sarei $J_0^{(N)}$.

OSS Le formule generalizzate hanno lo stesso grado di prec. della formula usata nel singolo intervallo.

ES: trapezi general. ha grado 1

La diff. importante è che posso mandare a zero l' errore aumentando L

FORMULE DI QUADRATURA PIÙ GENERALI

Abbiamo visto fin ora formule del tipo $\sum_{j=0}^n \alpha_j f(x_j)$ ma nulla vieta di prendere qualcosa di più generico come: $\sum_{j=0}^n \alpha_j f^{(d_j)}(x_j)$ $d_j \in \mathbb{N}$

DERIVATA DI
ORDINE d_j

ESERCIZIO

Consideriamo una formula del tipo

$$\int_0^1 f(x) dx = I(f) \approx \alpha_0 f(0) + \alpha_1 f\left(\frac{1}{3}\right) + \alpha_2 f\left(\frac{2}{3}\right) + \alpha_3 f(1)$$

come devono essere scelti $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ per avere grado di prec. massimo? Qual è il grado max?

Imponiamo $E(1) = E(x) = E(x^2) = E(x^3) = 0$

$$\int_0^1 1 dx = 1 = \alpha_0 + \alpha_3$$

$$\int_0^1 x^2 dx = \frac{1}{3} = \alpha_1 \cdot \frac{2}{3} + \alpha_2 \cdot \frac{4}{3} + \alpha_3$$

$$\int_0^1 x dx = \frac{1}{2} = \alpha_1 + \alpha_2 + \alpha_3$$

$$\int_0^1 x^3 dx = \frac{1}{4} = \frac{\alpha_1}{3} + \frac{4}{3} \alpha_2 + \alpha_3$$

$$\left\{ \begin{array}{l} \partial_0 + \partial_3 = 1 \\ \partial_1 + \partial_2 + \partial_3 = \frac{1}{2} \\ \frac{2}{3}\partial_1 + \frac{4}{3}\partial_2 + \partial_3 = \frac{1}{3} \Rightarrow \\ \frac{2}{3}\partial_1 + \frac{4}{3}\partial_2 + \partial_3 = \frac{1}{4} \end{array} \right. \quad \left\{ \begin{array}{l} \partial_0 = 1 - \partial_3 \\ \partial_3 = \frac{1}{2} - \partial_1 - \partial_2 \\ \frac{2}{3}\partial_1 + \frac{4}{3}\partial_2 + \partial_3 = \frac{1}{3} \Rightarrow \\ \frac{2}{3}\partial_1 + \frac{4}{3}\partial_2 + \partial_3 = \frac{1}{4} \end{array} \right.$$

$$\left\{ \begin{array}{l} \partial_0 = 1 - \partial_3 \\ \partial_3 = \frac{1}{2} - \partial_1 - \partial_2 \\ \frac{2}{3}\partial_1 + \frac{4}{3}\partial_2 + \frac{1}{2} - \partial_1 - \partial_2 = \frac{1}{3} \Rightarrow -\frac{1}{3}\partial_1 + \frac{1}{3}\partial_2 = -\frac{1}{6} \Rightarrow \partial_2 = \partial_1 - \frac{1}{2} \\ -\frac{\partial_1}{3} = -\frac{1}{12} \Rightarrow \partial_1 = \frac{1}{4} \end{array} \right.$$

$$\left\{ \begin{array}{l} \partial_0 = \frac{1}{2} \\ \partial_1 = \frac{1}{2} \\ \partial_2 = -\frac{1}{4} \\ \partial_3 = \frac{1}{4} \end{array} \right. \quad \text{Con questa scelta il grado di proc. è esattamente 3}$$

Infine verifichiamo cos'è $E(x^4)$

$$E(x^4) = \underbrace{\int_0^1 x^4 dx}_{\frac{1}{5}} - \left(\frac{1}{2} \cdot 0 + \frac{1}{4} \cdot 4 \left(\frac{1}{3}\right)^3 - \frac{1}{4} \left(\frac{2}{3}\right)^3 + \frac{1}{2} \cdot 1 \right) = \\ = \frac{1}{5} - \left(\frac{1}{27} - \frac{2}{27} + \frac{1}{2} \right) = \frac{1}{5} - \frac{26}{54} \neq 0$$

Il grado di proc. è esattamente 3

PER VALUTARE LA FUNZIONE DEVO SOSTituire $1, x, x^2 \dots$
 A f E Poi VALUTARLO IN Δ SECONDA DEL CASO.
ESEMPIO

15/12/2022

Data la formula $a_1 f(-1) + 2f(-x_0) + a_3 f(1) + a_2 f'(0) + 2f(x_0)$
 per approssimare $f'(x)$ dx determinare i valori dei
 pesi a_1, a_2, a_3 e x_0 del modo x_0 che rendono il
 grado di precisione massimo e determinare il
 grado.

$$E(1) = 0, E(x) = 0, \dots, E(x^m) = 0, E(x^{m+1}) \neq 0$$

$$\int_{-1}^1 1 dx = 2 = a_1 + 2 + a_3 \Rightarrow a_1 + a_3 = -2$$

$$\int_{-1}^1 x dx = 0 = -a_1 - 2x_0 + a_2 + 2x_0 + a_3 \Rightarrow -a_1 + a_2 + a_3 = 0$$

$$\int_{-1}^1 x^2 dx = \frac{2}{3} = a_1 + 2x_0^2 + 2x_0^2 + a_3 \Rightarrow a_1 + a_3 + 4x_0^2 = \frac{2}{3}$$

$$\int_{-1}^1 x^3 dx = 0 = -a_1 - 2x_0^3 + 2x_0^3 + a_3 \Rightarrow -a_1 + a_3 = 0$$

$$\begin{cases} a_1 + a_3 = -2 \\ -a_1 + a_2 + a_3 = 0 \\ a_1 + a_3 + 4x_0^2 = \frac{2}{3} \\ -a_1 + a_3 = 0 \end{cases} \Rightarrow \begin{cases} 2a_1 = -2 \Rightarrow a_1 = a_3 = -1 \\ a_2 = 0 \\ x_0^2 = \frac{2}{3} \Rightarrow x_0 = \sqrt{\frac{2}{3}} \\ a_1 = a_3 \end{cases}$$

La formula diventa:

$$-f(-1) + 2(-\sqrt{\frac{2}{3}}) + 2f(\sqrt{\frac{2}{3}}) - f(1)$$

e ha grado di precisione ≥ 3

Se il sistema non fosse stato risolvibile, avremmo provato a vedere quante e quali delle 4 eq. erano verificate e si sarebbe abbassato il grado "minimo".

$$E(x^4) = \int_1^5 x^4 dx - \left(1 + 2 \cdot \frac{4}{9} + 2 \cdot \frac{4}{9} - 1\right) = \frac{2}{5} + \frac{2}{9} \neq 0$$

↓
LA FORMULA HA
GRADO DI PREC.
ESAT. 3

Esercizio

$$c^2 f(-a) + b f(0) + c f(a) \approx \int_{-a}^a f(x) dx$$

Scegliere i param. per massimizz. la precisione
con $a > 0$.

$$E(1) = 0 \iff \int_{-a}^a 1 dx = 2a = c^2 + b + c$$

$$E(x) = 0 \iff \int_{-a}^a x dx = 0 = -ac^2 + ac$$

$$E(x^2) = 0 \iff \int_{-a}^a x^2 dx = \frac{2}{3}a^3 = a^2 c^2 + a^2 c$$

$$E(x^3) = 0 \iff \int_{-a}^a x^3 dx = 0 = -a^3 c^2 + a^3 c$$

$$\begin{cases} 2a = c^2 + b + c \\ ac(1-c) = 0 \quad \begin{matrix} c=1 \\ c=0 \end{matrix} \\ \frac{2}{3}a^3 = a^2 c(c+1) \\ a^3 c (1-c) = 0 \end{cases}$$

CASO $c = 1$:

$$\begin{cases} 2a = b + 2 \Rightarrow b = 4 \Rightarrow \text{la precisione è almeno 3} \\ \frac{2}{3}a^3 = 2 \Rightarrow a = 3 \end{cases}$$

$$\int_{-3}^3 x^4 dx = \int_{-3}^3 x^4 dx = \frac{486}{5} \Rightarrow E(x^4) = \frac{486}{5} - (3^4 + 3^4) \neq 0$$

↓
LA PREC. È 3

CASO C=0:

$$\begin{cases} 2a = b \\ c = 0 \end{cases}$$

VERIFICANDO $E(.) = E(x) = 0$, MA
 $c = 0 \Rightarrow E(x^2) \neq 0 \Rightarrow$ GRADO PREC. = 1

\Rightarrow IL caso migliore è $a=3, b=9, c=1$

ESERCIZIO

$$\int_{-\alpha}^{\alpha} f(x) dx \approx \alpha [a_0 f(-\alpha) + a_1 f(0) + a_2 f(\alpha)] + \alpha^2 a_3 [f'(-\alpha) - f'(\alpha)]$$

Massimizzare la precisione per $\alpha > 0$

$$\int_{-\alpha}^{\alpha} 1 dx = 2\alpha = \alpha(a_0 + a_1 + a_2) \Leftrightarrow 2a_0 + a_2 = 2$$

$$\int_{-\alpha}^{\alpha} x dx = 0 = \underbrace{\alpha [-\alpha a_0 + \alpha a_1]}_{=0} + \underbrace{\alpha^2 a_3 [1 - 1]}_{=0} \Rightarrow 0 = 0 \quad \checkmark$$

OK, MA QUINDI
 CI SEGUONO
 UN'EQU. IN PIÙ

$$\int_{-\alpha}^{\alpha} x^2 dx = \frac{2}{3} \alpha^3 = \alpha [a_0 \alpha^2 + a_1 \alpha^2] + \alpha^2 a_3 [-2\alpha - 2\alpha]$$

$$\Rightarrow 2a_1 = 2a_0 - 4a_3$$

$$\int_{-\alpha}^{\alpha} x^3 dx = 0 = \underbrace{\alpha [-\alpha a_0^3 + \alpha a_1^3]}_{=0} + \underbrace{\alpha^2 a_3 [3\alpha^2 - 3\alpha^2]}_{=0} \Rightarrow 0 = 0 \quad \checkmark$$

NON AMO
 OTTE F
 SI ANNUN
 PER POC
 DISPONI

$$\int_{-\alpha}^{\alpha} x^4 dx = \frac{2}{5} \alpha^5 = \alpha [a_0 \alpha^4 + a_1 \alpha^4] + \alpha^2 a_3 [-4\alpha^3 - 4\alpha^3]$$

$$\Rightarrow \frac{2}{5} = 2a_0 - 8a_3$$

$$\begin{cases} 2\alpha_1 + \alpha_2 = 2 \\ 2\alpha_1 - 4\alpha_3 = \frac{2}{3} \\ 2\alpha_1 - 8\alpha_3 = \frac{2}{5} \end{cases} \Rightarrow \begin{cases} 2\alpha_1 + \alpha_2 = 2 \\ 2\alpha_1 - 2\alpha_3 = \frac{1}{3} \\ \alpha_1 - 4\alpha_3 = \frac{1}{5} \end{cases} \Rightarrow \begin{cases} \alpha_2 = 2(1 - \alpha_1) \\ \alpha_1 = \frac{1}{3} + 2\alpha_3 \\ -2\alpha_3 = \frac{1}{5} - \frac{1}{3} \end{cases}$$

$$\Rightarrow \begin{cases} \alpha_2 = \frac{16}{15} \\ \alpha_1 = \frac{7}{15} \\ \alpha_3 = \frac{1}{15} \end{cases} \Rightarrow \text{il grado di prec. è } \geq 4$$

$$E(x^5) = \int_{-\alpha}^{\alpha} x^5 dx - (\underbrace{\alpha[-\alpha, \alpha^5 + \alpha, \alpha^5] + \alpha^2 \alpha_3 [5\alpha^4 - 5\alpha^4]}_{=0}) = 0$$

$$E(x^6) = \int_{-\alpha}^{\alpha} x^6 dx - \left[\alpha \left[\frac{7}{15}\alpha^6 + \frac{7}{15}\alpha^6 \right] + \frac{\alpha^2}{15} (-6x^5 - 6x^5) \right] =$$

$$= \frac{2}{7}\alpha^7 - \frac{2}{15}\alpha^7 = \frac{2}{15}\alpha^7 \neq 0 \text{ perché } \alpha > 0$$

\Rightarrow il grado di prec. è esatt. 5

Esercizio

$$\int_{-c}^c |x| f(x) dx \approx af(x_0) + bf(x_1) \quad g(x) = |x|$$

Determinare a, b, x_0, x_1, c ($c > 0$), $x_0 \neq x_1$

$$\int_{-c}^c |x| \cdot |x| dx = \int_{-c}^0 -x dx + \int_0^c x dx = c^2 = a+b$$

$$\int_{-c}^c x \cdot |x| dx = \int_{-c}^0 -x^2 dx + \int_0^c x^2 dx = 0 = ax_0 + bx_1$$

$$\int_{-c}^c x^2 \cdot 1 \times 1 dx = \int_{-c}^0 -x^3 dx + \int_0^c x^3 dx = \frac{c^4}{2} = ax_0^2 + bx_1^2$$

$$\int_{-c}^c x^3 \cdot 1 \times 1 dx = \int_{-c}^0 -x^4 dx + \int_0^c x^4 dx = 0 = ax_0^3 + bx_1^3$$

$$\begin{cases} a+b=c^2 \\ ax_0+bx_1=0 \\ ax_0^2+bx_0^2=c^4/2 \\ ax_0^3+bx_1^3=0 \end{cases} \Rightarrow \begin{cases} x_1=-x_0 \\ a=b=c^2/2 \\ x_0^2=x_1^2=c^2/2 \Rightarrow x_0=-\frac{\sqrt{2}}{2}c \\ x_1=\frac{\sqrt{2}}{2}c \end{cases}$$

LA MOLTIPLICAZIONE PER x_0^2 È
LA SOLUZIONE ALLA 4^ EQ.

$$bx_1^3 - bx_1x_0^2 = 0 \Rightarrow bx_1(x_1^2 - x_0^2) = bx_1(x_1 + x_0)(x_1 - x_0) = 0$$

$$\cancel{b=0, x_1=0} \quad \boxed{x_1 = -x_0} \quad \cancel{x_1 = x_0}$$

Com $x_0 = -\frac{\sqrt{2}}{2}c$, $x_1 = \frac{\sqrt{2}}{2}c$, $a=b=c^2/2$ la formula

ha grado di prec. ≥ 3

$$\int_{-c}^c 1 \times 1 \cdot x^4 dx = \int_{-c}^0 -x^5 dx + \int_0^c x^5 dx = \frac{c^6}{3} \neq \frac{c^2}{2} \left(\frac{c^4}{2} + \frac{c^4}{2} \right) = \frac{c^6}{4}$$

$$\Rightarrow E(x^4) \neq 0 \Rightarrow \text{prec.} = 3$$

FORMULE GAUSSIANE

Formule interpolatorie con scelte dei nodi ottimali.
Per trovare i nodi si potrebbe risolvere un sist. non lineare di $2n+2$ eq., ma vediamo un'altra strada che si basa sui polim. ortogonali.

DEF $\Pi = \{ \text{polinomi con coeff. reali su } [\alpha, \beta] \subseteq \mathbb{R} \}$

Π è uno spazio vettoriale e si può dotare di prodotto scalare, la dim. infinita.

$p, q \in \Pi$ si def. il prod. scalare

$$\langle p, q \rangle = \int_{\alpha}^{\beta} p(x) q(x) \cdot g(x) dx$$

con $g(x)$ funz. peso ($\Rightarrow p(x) \geq 0$ su $[\alpha, \beta]$)

Due polim. sono ortogonali se $\langle p, q \rangle = 0$

DEF Si indica con $\Pi^* = \{ q_i(x), i \in \mathbb{N} : \langle q_i, q_j \rangle = \begin{cases} 0, & i \neq j \\ c_i \neq 0, & i = j \end{cases} \}$

Π^* è una base di polim. ortogonali per Π

PROPRIETÀ

(i) Prezzo $p(x)$ di grado n posso scrivere come
$$\sum_{j=0}^n \alpha_j q_j(x) = p(x)$$

ii) $\langle q_j, p \rangle = 0 \forall p(x)$ con grado $< j$

iii) $q_j(x)$ ha esatt. j zeri reali distinti in $[\alpha, \beta]$

DIM (iii)

Supponiamo per assurdo che $q_j(x)$ abbia $k < j$ zeri distinti in $[a, b]$, di cui $h \leq k$ di grado dispari.

$$q_j(x) = (x - z_1)^{d_1} \cdots (x - z_h)^{d_h} (x - z_{h+1})^{d_{h+1}} \cdots (x - z_s)^{d_s}$$

d_1, \dots, d_h dispari
 d_{h+1}, \dots, d_s pari

Considero $P(x) = (x - z_1) \cdots (x - z_h)$

$$0 = \langle q_j, P \rangle = \int_a^b q_j(x) P(x) p(x) dx - \underbrace{\int_a^b p(x) (x - z_1) \cdots (x - z_h) dx}_{\text{TUTTI GLI ESSI SONO PARI}}$$

$$0 = \int_a^b \underbrace{\cdots}_{\geq 0} \text{ed è diverso (l'integrandino) da } 0$$

✳ impossibile

□

TH Dati $[a, b]$ e la f. peso $p(x)$, la formula di quadratura interpolatoria $J_n(f)$ che usa come nodi di interp. x_0, \dots, x_n zeri del polin. ortog. $q_{n+1}(x)$ ha grado di prec. $2n+1$.
 Inoltre se $f(x) \in C^{2n+2}([a, b])$ vale.

$$E_n(f) = K_n \cdot \frac{f^{(2n+2)}(\varepsilon)}{(2n+2)!} \quad K_n > 0, \quad \varepsilon \in (a, b)$$

$$\int_a^b f(x) dx - J_n(f)$$

OSS $\mathcal{I}_n(f)$ ha come nodi gli zeri di $q_{n+1}(x)$
e come pesi $\int_a^b \ell_i(x) p(x) dx$ $\ell_i(x) = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}$

OSS Una volta fissati $[a, b]$ e $p(x)$, è fissata la base di polim. ortogonali e quindi le formule gaussiane

FATTI

i) Per calcolare i polim. ortogonali ci sono delle formule di ricorrenza a 3 termini

$$q_{n+1}(x) = \alpha_n(x) q_n(x) + \beta_n(x) q_{n-1}$$

con α_n, β_n polim. di grado 0 o 1

ii) I programmi di integrazione numerica calcolano con metodi iterativi gli zeri dei polim. ortogonali ad alta precisione

iii) Se si conoscano i nodi e i pesi delle formule Gaussiane per $[a, b] = [-1, 1]$ ed un certo $p(x)$ ci si può sempre ricordare a questo caso se si integra su un intervallo finito

$$t = \frac{b-a}{2}x + \frac{b+a}{2} : [-1, 1] \rightarrow [a, b]$$

$$\int_a^b g(t) dt = \int_{-\frac{b-a}{2}}^{\frac{b+a}{2}} \frac{b-a}{2} \cdot g\left(\frac{b-a}{2}x + \frac{b+a}{2}\right) dx = \int_{-1}^1 f(x) p(x) dx$$

$$\text{dove } f(x) := \frac{b-a}{2p(x)} \cdot g\left(\frac{b-a}{2}x + \frac{b+a}{2}\right)$$

ESEMPIO

① $[-1, 1]$, $\varphi(x) = 1 \rightarrow$ POLINOMI DI LEGENDRE

$$L_0(x) = 1, L_1(x) = x$$

$$(i+1)L_{i+1}(x) = (2i+1)xL_i(x) - iL_{i-1}(x)$$

RELATONE
DI RICORRENZA

② $[-1, 1]$, $\varphi(x) = \frac{1}{\sqrt{1-x^2}}$

$$T_0(x) = 1, T_1(x) = x \rightarrow$$
 POLINOMI DI CHEBYSHEV DI 1[°] SPECIE

$$T_{i+1}(x) = 2xT_i(x) - T_{i-1}(x)$$

③ $[-1, 1]$, $\varphi(x) = \sqrt{1-x^2} \rightarrow$ POLINOMI DI CHEBYSHEV DI 2[°] SPECIE

④ $[0, +\infty)$, $\varphi(x) = e^{-x} \rightarrow$ POLINOMI DI LAGUERRE

⑤ $(-\infty, +\infty)$, $\varphi(x) = e^{-x^2/2} \rightarrow$ POLINOMI DI HERMITE

PROPRIETÀ

- i) \rightarrow pesi delle formule Gaussiane sono sempre positivi
- ii) f continua sull' intervallo di integrazione
 $\lim_{n \rightarrow +\infty} I_n(f) = I(\varphi f)$