

CALCOLO NUMERICO

Teoria degli errori

o RAPPRESENTAZIONE dei NUMERI

Fissato un numero intero $\beta > 1$, ogni numero $x \in \mathbb{R} - \{0\}$ ammette una rappresentazione data da
 $x = \text{Sign}(x) \beta^b \sum_{i=1}^{\infty} \alpha_i \beta^{-i}$ $\rightarrow b \in \mathbb{Z}$: ordine di grandezza del numero (esponente)
 \downarrow
 $\beta = \text{base della rappresentazione}$ $\rightarrow \text{Sign}(x) = \begin{cases} 1 & \text{se } x > 0 \\ -1 & \text{se } x < 0 \end{cases}$

TEOREMA della RAPPRESENTAZIONE: Data una base intera $\beta > 1$ ed un qualunque numero reale x diverso da 0 esiste un'unica rappresentazione in base β tale che
 $\rightarrow \alpha_1 \neq 0$
 $\rightarrow \exists k \in \mathbb{N}$ per cui si afferma $\alpha_k = \beta - 1 \quad \forall j > k$: Escluse rappresentazioni di periodo $\beta - 1$

Per numeri periodici: FRAZIONE GENERATRICE

Numero senza separatore - parte prima del periodo
Tanti 9 quante sono le cifre del periodo e tanti 0 quante sono le cifre dell'antiperiodo
↓ ESEMPIO

$$1. \overline{69} = \frac{169 - 16}{90} = \frac{135}{90}$$

Virgola mobile NORMALIZZATA: da MANTISSA (ε) è sempre $< 1 \rightarrow 1 \leq \sum_{i=1}^{\infty} \alpha_i \beta^{-i} < 1$
Minimo valore: $\alpha_1 \neq 0$ e $\alpha_i = 0 \rightarrow \sum_{i=1}^{\infty} \alpha_i \beta^{-i} = 1 \beta^{-1} = \frac{1}{\beta}$ → primo termine
Massimo superiore: $\alpha_i = \beta - 1 \rightarrow \sum_{i=1}^{\infty} \alpha_i \beta^{-i} = (\beta - 1) \sum_{i=1}^{\infty} \beta^{-i} = (\beta - 1) \frac{1}{\beta} \frac{1}{1 - \frac{1}{\beta}} = \frac{\beta - 1}{\beta} \frac{\beta}{\beta - 1} = 1$ → Serie geometrica di ragione $1/\beta$

All'interno di un calcolatore si possono rappresentare solo m (me/IN) cifre della mantissa di $x \rightarrow$ 2 modi per passare da rappresentazione ∞ a ∞

TRONCAMENTO: $tr(x) = \text{Sign}(x) \beta^b \sum_{i=1}^m \alpha_i \beta^{-i}$

ARROTONDAMENTO: $rd(x) = \begin{cases} tr(x) & \text{se } 0 \leq \alpha_{m+1} < \frac{\beta}{2} \\ \text{Sign}(x) \beta^b [\sum_{i=1}^m \alpha_i \beta^{-i} + \beta^{-m}] & \text{se } \frac{\beta}{2} \leq \alpha_{m+1} < \beta \end{cases}$

$$\downarrow \text{ESEMPIO: } 0.995 \cdot 10^0, m=2 \rightarrow 0.99 + 10^{-2} = 1.00 = 0.1 \cdot 10^1$$

Si avrà che $|tr(x) - x| < \beta^{b-m}$ $\rightarrow |rd(x) - x| \leq \frac{1}{2} \beta^{b-m}$ → Migliore approssimazione
↓ errore assoluto \rightarrow ho dovuto aggiungere l'esponente

Si indichi con M l'insieme dei numeri rappresentabili all'interno di un calcolatore → Numeri MACCHINA

L'insieme FINITO: fissati β, m e supposto $1 \leq b \leq 0$ ($b, v \in \mathbb{Z}$) la cardinalità di M risulta tale che
 $\downarrow \text{card}(M) = \omega(\beta^m - \beta^{m-1})(v - 1 + 1) \rightarrow$ per cui ci sono $\beta - 1$ possibilità, mentre per le altre

dell'insieme viene indicato con $F(\beta, m, b, v)$ cifre sono $\beta \rightarrow (\beta - 1)\beta^m = \beta^m - \beta^{m-1}$: possibili MANTISSE
 \downarrow Scelte esponente $\leftarrow (v - 1 + 1)(\beta^m - \beta^{m-1}) \cdot 2 + 1 \rightarrow$ rappresenta lo 0

Dato un qualunque numero reale $x \neq 0$ non è accurata l'esistenza di $rd(x)$ tra i numeri macchina

ESEMPIO: $F(10, 3, -99, 99)$ e $x = 0.9999 \cdot 10^{99}$, quindi $rd(x) = 0.1 \cdot 10^{100}$
che NON è rappresentabile → OVERFLOW

L'overflow viene regolato durante l'esecuzione del programma e l'esecuzione può essere avvertita oppure continuare sostituendo il numero con $rd(x) = \text{Sign}(x) \text{mant}(x)$

Analogamente per i numeri troppo piccoli si ha l'UNDERFLOW

ESEMPIO: F dell'esempio precedente → $x = 0.01 \cdot 10^{-99}$, $rd(x) = 0.1 \cdot 10^{-100}$ NON rappresentabile

↓

In generale, per qualunque z : $|rd(x) - x| \leq |z - x| \rightarrow rd(x)$ lo approssima al meglio

o d'ERRORE ASSOLUTO e RELATIVO

Si definisce errore assoluto della rappresentazione del numero reale x il valore

$$S_x = rd(x) - x \rightarrow |S_x| \leq \frac{1}{2} \beta^{b-m}$$

ESEMPIO: $0.15 \cdot 10^5$ diventa $0.2 \cdot 10^5$ con una cifra $\rightarrow S_x = 0.05 \cdot 10^5 = 0.5 \cdot 10^4 = \frac{1}{2} \cdot 10^4$

Si definisce errore relativo il valore $E_x = \frac{rd(x) - x}{x} = \frac{S_x}{x} \rightarrow$ LIMITAZIONE: $|E_x| \leq \frac{1}{2} \beta^{1-m}$
 ↳ Dipende solo della BASE e del NUMERO di CIFRE UTILIZZATE

PRECISIONE di MACCHINA: $m = \frac{1}{2} \beta^{1-m} \rightarrow$ Massimo errore relativo che si può commettere
 ↓ ponendo da x a rd(x)

Se E_x non supera m allora la rappresentazione è corretta fino alla m-esima cifra significativa

OPERAZIONI di MACCHINA

Le 4 operazioni sono chiuse \rightarrow Se eseguo un'operazione tra numeri macchina NON è detto che il risultato sia un numero macchina e le proprietà possono ALTERARE il risultato \rightarrow L'addizione NON ha proprietà associativa

ESEMPIO:

$$M = F(10, 3, -aa, aa), \quad x = 0.135 \cdot 10^{-4}, \quad y = 0.258 \cdot 10^{-2}, \quad z = -0.251 \cdot 10^{-2}$$

\oplus : somma tra elementi di M

$$a. x \oplus (y \oplus z) = 0.135 \cdot 10^{-4} \oplus (0.258 \cdot 10^{-2} \oplus -0.251 \cdot 10^{-2}) = 0.385 \cdot 10^{-4}$$

$$b. (x \oplus y) \oplus z = (0.135 \cdot 10^{-4} \oplus 0.258 \cdot 10^{-2}) \oplus -0.251 \cdot 10^{-2} = 0.380 \cdot 10^{-4}$$

Quando si straggono due numeri circa uguali si verifica il fenomeno della cancellazione \rightarrow Si perdono cifre significative nel risultato amplificando E_x

ERRORE nel CALCOLO di una FUNZIONE

Sia $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione ad n variabili e a valori reali. Si vuole calcolare $f(p) = f(x_1, x_2, \dots, x_n)$ in un punto ammesso $p_0 = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$. A seconda dei punti (p_1, p_2, \dots) NON è possibile trovare i valori esatti \rightarrow Si calcola la funzione in un punto ottenuto approssimando le cifre di p_0 . Allo stesso tempo, i calcoli vengono fatti da alcuni algoritmi approssimando le cifre non avendo quindi la funzione f reale ma fa algoritmo \rightarrow Si calcola fa (p_1)

Calcolo dell'ERRORE ASSOLUTO

Sia dato il punto $p_0 \in \mathbb{R}^n$ di coordinate $x_i^{(0)}, i=1, 2, \dots, n$ nel quale si vuole calcolare la funzione $f(p)$ \rightarrow Si considera l'insieme $D = \{p \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i, i=1, 2, \dots, n\} \rightarrow$ insieme di INDETERMINAZIONE

Il valore cercato $f(p_0)$ viene sostituito dal valore calcolato $fa(p_1)$, dove $p_1 \in D$ con coordinate $x_i^{(1)}, i=1, 2, \dots, n \rightarrow$ L'errore assoluto è dato da $\delta f = fa(p_1) - f(p_0)$: si può STIMARE

$$\text{Si namo e si ottiene } f(p_1) : \delta f = fa(p_1) - f(p_1) + f(p_1) - f(p_0) = \delta a + \delta d$$

↓ Dovuto all'ALGORITMO: δa Dovuto al PONTO: δd

Risulta definito e stimabile una volta finito l'algoritmo \rightarrow Si può dare un'espressione dell'errore assoluto δd tramite dati

$$\text{Calcolo di } \delta d: \quad f(p_1) - f(p_0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \underbrace{(x_i^{(1)} - x_i^{(0)})}_{\delta x_i} \text{ formandosi al primo termine}$$

$$= \sum_{i=1}^n p_i^{(1)} \delta x_i : p_i = \text{coefficienti di amplificazione degli errori } x_i$$

Limitazione per il modulo dell'errore assoluto $|fa(p_1) - f(p_0)| \leq \delta a + \delta d$

$$\delta d = \sum_{i=1}^n |A_i| |\delta x_i| \text{ dove } A_i \geq \sup_{p \in D} \left| \frac{\partial f}{\partial x_i} \right|, \quad i = 1, 2, \dots, n$$

Per il calcolo dell'errore assoluto vi sono due tipologie di problemi: diretto ed inverso
 → DIRETTO: Si conoscono le maggiori zioni dei due errori assoluti, quindi possiamo

rapre l'estremo superiore con cui è calcolata la funzione nel punto

→ INVERSO: Dobbiamo stabilire l'algoritmo e la precisione dei dati in modo che la precisione dei dati risulti nella soglia di precisione richiesta

ESEMPI:

problema diretto

$f(x_1, x_2) = \frac{x_1}{x_2}$ in un punto $P_0 \in D = [1,3] \times [4,5]$. Si suppone di arrotondare il risultato dell'operazione alla α^{a} cifra decimale e di introdurre i valori $x_1^{(0)}$ e $x_2^{(0)}$ con errori $|S_{x_1}| \leq 10^{-2}$, $|S_{x_2}| \leq 10^{-2}$

↓
arrotondando il risultato alla α^{a} cifra decimale si ha $|S_a| \leq \frac{1}{2} \cdot 10^{-2}$

Si avrà $\frac{\partial}{\partial x_1} f = \frac{1}{x_2}$ e $\frac{\partial}{\partial x_2} f = -\frac{x_1}{x_2^2}$

Ne consegue che $Ax_1 = \max_{P \in D} \left| \frac{1}{x_2} \right| = \frac{1}{4}$ e $Ax_2 = \max_{P \in D} \left| -\frac{x_1}{x_2^2} \right| = \frac{3}{16}$

Concludiamo che $|S_f| \leq |S_a| + |S_d| \leq \frac{1}{2} \cdot 10^{-2} + \frac{1}{4} \cdot 10^{-2} + \frac{3}{16} \cdot 10^{-2} = \frac{15}{16} \cdot 10^{-2}$

problema inverso

$x_1^{(0)} \in [1,2]$ e $x_2^{(0)} \in [-2,-1]$ e si consideri la funzione $f(x_1, x_2) = \frac{x_1}{x_2}$

Come si deve eseguire la divisione e con quale errore andato si devono introdurre x_1 e x_2 per avere $|S_f| \leq 10^{-2}$?

↓
Essendo che l'errore è diviso tra quello algoritmico tra quello algoritmico e quello truncamento dati avremo $|S_a| \leq \frac{1}{2} \cdot 10^{-2}$ e $|S_d| \leq \frac{1}{2} \cdot 10^{-2}$

Basta eseguire la divisione fino alla α^{a} cifra decimale ed arrotondare alla seconda

Ahhiamo che $\frac{\partial}{\partial x_1} f = \frac{1}{x_2}$ e $\frac{\partial}{\partial x_2} f = -\frac{x_1}{x_2^2}$ con $D = [1,2] \times [-2,-1]$

Determinando Ax_1 e Ax_2 → $Ax_1 = \max_{P \in D} \left| \frac{1}{x_2} \right| = 1$ $Ax_2 = \max_{P \in D} \left| -\frac{x_1}{x_2^2} \right| = 2$

Data la condizione: $Ax_1 |S_{x_1}| \leq \frac{1}{4} \cdot 10^{-2}$ e $Ax_2 |S_{x_2}| \leq \frac{1}{4} \cdot 10^{-2}$
da cui, notiamo
 $|S_{x_1}| \leq \frac{1}{4} \cdot 10^{-2}$ e $|S_{x_2}| \leq \frac{1}{2} \cdot 10^{-2}$

Per verificare ambedue le limitazioni si impone $|S_{x_1}| \leq 10^{-3}$ e $|S_{x_2}| \leq 10^{-3}$

CONCLUDENDO: $|S_f| \leq |S_a| + |S_d| = \frac{1}{2} \cdot 10^{-2} + 10^{-3} + 2 \cdot 10^{-3} = \frac{4}{5} \cdot 10^{-2}$

Calcolo dell'ERRORE RELATIVO

L'errore relativo che si commette nel calcolo di una funzione $f(P)$ in un adeguate punto P_0 è definito da $E_f = \frac{|f(P_1) - f(P_0)|}{|f(P_0)|}$

VERIFICA: $E_f = \frac{|f(P_1) - f(P_0)|}{|f(P_0)|} + \frac{|f(P_1) - f(P_0)|}{|f(P_1)|} \left(1 + \frac{|f(P_1) - f(P_0)|}{|f(P_0)|} \right) = E_a + E_d + \underline{E_a E_d} = E_a + E_d$
TRASCORRITO perché di ORDINE SUPERIORE

E_a : errore truncamento dai dati E_d : errore algoritmico

Calcolo di E_d : $E_d = \frac{f(p_1) - f(p_0)}{f'(p_0)} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{(x_i^{(1)} - x_i^{(0)})}{f'(p_0)}$ e definendo $\bar{x}_i = \frac{x_i^{(1)} - x_i^{(0)}}{x_i^{(0)}}$ si ottiene

$$E_d = \sum_{i=1}^n \frac{x_i^{(0)}}{f'(p_0)} \frac{\partial f}{\partial x_i} \bar{x}_i \quad \text{e ponendo } \gamma_i = \frac{x_i^{(0)}}{f'(p_0)} \frac{\partial f}{\partial x_i} \text{ ottieni come}$$

$$E_d = \sum_{i=1}^n \gamma_i \bar{x}_i \quad \text{con } \gamma_i \text{ detti COEFFICIENTI di AMPLIFICA ZONE}$$

ATT: Se i coefficienti γ_i sono tali che l'errore E_d risulta dello stesso ordine degli errori \bar{x}_i il problema si dice BEN CONDIZIONATO. Al contrario si dice MASS CONDIZIONATO (piccoli $\bar{x}_i \rightarrow$ grandi E_d)

Errore E_a : Se E_a produce errori asintoticamente limitati si dice STABILE, altrimenti viene detto INSTABILE

Errore nelle 4 OPERAZIONI ELEMENTARI

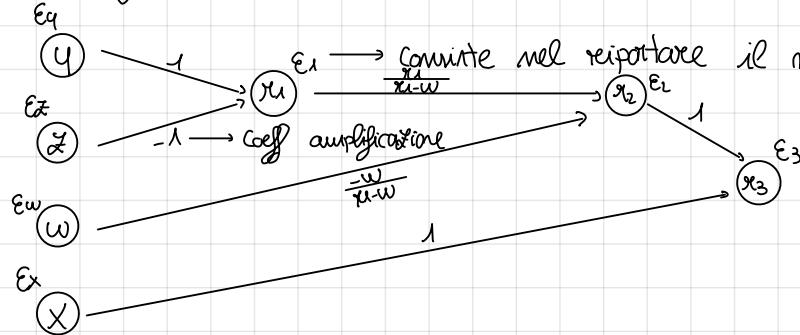
	S_d	E_d	
a. $X + Y$:	$S_x + S_y$	$\frac{x}{x+y} E_x + \frac{y}{x+y} E_y$	→ i coefficienti possono "ESPLODERE" (e con loro anche l'errore relativo) se il denominatore è prossimo allo 0
b. $X - Y$:	$S_x - S_y$	$\frac{x}{x-y} E_x - \frac{y}{x-y} E_y$	
c. $X \otimes Y$:	$y S_x + x S_y$	$E_x + E_y$	
d. $X \oslash Y$:	$\frac{1}{y} S_x - \frac{x}{y^2} S_y$	$E_x - E_y$	

Esempi di calcolo dell'ERRORE RELATIVO

ESERCIZIO 1

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

a. bisogna stabilizzare l'algoritmo di CALCOLO: $x_1 = \frac{y}{z}$, $x_2 = x_1 - w$, $x_3 = x \cdot x_2$



Per avere l'errore relativo dove possibile DAL FONDO $\rightarrow E_f = E_{x_3} = E_3 + 1 \cdot E_{x_2} + 1 \cdot E_x = E_3 + E_x + E_2 + \frac{x_1}{x_1 - w} E_{x_1} - \frac{w}{x_1 - w} E_w = E_3 + E_x + E_2 - \frac{w}{x_1 - w} E_w + \frac{x_1}{x_1 - w} (E_1 + E_2 - E_x)$

Calcoliamo le espressioni

$$a. \frac{w}{x_1 - w} = \frac{w}{\frac{y}{z} - w} = \frac{zw}{y - zw}$$

$$b. \frac{x_1}{x_1 - w} = \frac{\frac{y}{z}}{\frac{y}{z} - w} = \frac{y}{y - zw}$$

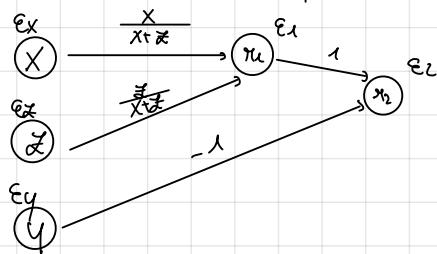
Continuando: $= \frac{y}{y - zw} E_1 + E_2 + E_3 + E_x + \frac{y}{y - zw} E_q - \frac{y}{y - zw} E_z - \frac{z}{y - zw} E_w$

errore algoritmo

errore relativo ai dati

ESERCIZIO 2

$$f(x, y, z) = \frac{x+z}{y}$$



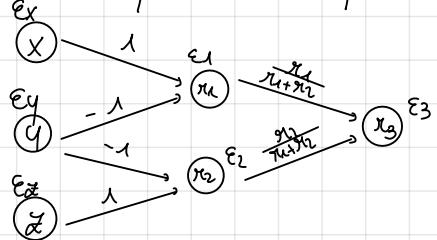
$$x_1 = \frac{x+z}{y} \quad x_2 = \frac{x_1}{y}$$

$$\begin{aligned} \dot{e}_f &= \dot{e}_2 + 1 \cdot \dot{e}_{x1} - \dot{e}_y = \dot{e}_2 - \dot{e}_y + \dot{e}_{x1} + \frac{x}{x+z} \dot{e}_x + \frac{z}{x+z} \dot{e}_z \\ &= \underbrace{\dot{e}_1 + \dot{e}_2}_{ea} + \underbrace{\frac{x}{x+z} \dot{e}_x}_{ea} - \underbrace{\dot{e}_y + \frac{z}{x+z} \dot{e}_z}_{ed} \end{aligned}$$

↓
STABILE

Secondo algoritmo

$$x_1 = \frac{x}{y} \quad x_2 = \frac{z}{y} \quad x_3 = x_1 + x_2$$



$$\dot{e}_f = \dot{e}_{x3} = \dot{e}_3 + \frac{x_1}{x_1+x_2} \dot{e}_x + \frac{x_2}{x_1+x_2} \dot{e}_z$$

CALCOLO dei COEFFICIENTI: a. $\frac{x_1}{x_1+x_2} = \frac{\frac{x}{y}}{\frac{x}{y} + \frac{z}{y}} = \frac{x}{x+z}$

$$b. \frac{x_2}{x_1+x_2} = \frac{\frac{z}{y}}{\frac{x}{y} + \frac{z}{y}} = \frac{z}{x+z}$$

$$\text{Continuo: } = \dot{e}_3 + \frac{x}{x+z} (\dot{e}_1 + \dot{e}_x - \dot{e}_y) + \frac{z}{x+z} (\dot{e}_2 + \dot{e}_z - \dot{e}_y) =$$

$$= \underbrace{\frac{x}{x+z} \dot{e}_1 + \frac{z}{x+z} \dot{e}_2}_{ea} + \underbrace{\frac{x}{x+z} \dot{e}_x - \dot{e}_y + \frac{z}{x+z} \dot{e}_z}_{ed}$$

↓
POTENZIALMENTE INSTABILE

Ricordi di algebra lineare

MATRICI: Con $A \in \mathbb{C}^{m \times n}$ si intende una matrice di m righe ed n colonne formata da $m \times n$ numeri complessi a_{ij} , $i=1,2,\dots,m$, $j=1,2,\dots,n$, detti elementi di A . i e j sono gli indici di riga e colonna dell'elemento
 m, n sono dette dimensioni di A → se $m=n$ la matrice è quadrata, altrimenti è rettangolare

Gli elementi a_{ii} , $i=1,2,\dots,n$ si dicono elementi diagonali

Con $a \in \mathbb{C}^m$ si intende un vettore con m componenti complesse indicate come a_i , $i=1,2,\dots,m$

MATRICI REALI: Se la matrice A ha tutti gli elementi reali si dice $A \in \mathbb{R}^{m \times n}$
MATRICE TRASPOSTA ed HERMITIANA

Data la matrice $A \in \mathbb{C}^{m \times n}$, la matrice $B \in \mathbb{C}^{n \times m}$ i cui elementi sono $b_{ij} = \bar{a}_{ji}$, $i=1,2,\dots,n$, $j=1,2,\dots,m$ si dice matrice TRASPOSTA di A e si indica con A^T . Si dice matrice HERMITIANA di A la matrice $B \in \mathbb{C}^{n \times m}$ i cui elementi sono $b_{ij} = \bar{a}_{ji}$ e si indica con A^H

MATRICE IDENTICA: La matrice quadrata $I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & 1 & \dots \\ \vdots & & & \ddots \\ 0 & 0 & \dots & 1 \end{pmatrix}$ si dice matrice IDENTICA

MATRICE DIAGONALE: Si dice matrice diagonale una matrice quadrata D con gli elementi non diagonali nulli → $D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ 0 & 0 & \dots & d_n \end{pmatrix}$ → $D = \text{diag}(d_1, d_2, \dots, d_n)$

MATRICI TRIANGOLARI: $L = \begin{pmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ l_{31} & l_{32} & \dots & l_{33} \end{pmatrix}$, $U = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ 0 & 0 & \dots & u_{nn} \end{pmatrix}$

Triangolare inferiore

Triangolare superiore

OPERAZIONI: **ADDITIONE:** proprietà associativa e commutativa

MULTIPLICAZIONE: proprietà associativa e distributiva rispetto all'addizione ma non voleva la proprietà commutativa e la legge di annullamento del prodotto → Se $AB = BA$ si dice che le matrici COMMUTANO. In generale per molti altri (BA) è diverso da post-moltiplicare (AB)

Per la trasposta e la coniugata: $(AB)^T = B^T A^T$ e $(AB)^H = B^H A^H$
 Il prodotto $a^H b = \sum_i a_i b_i$: prodotto SCALARE → Se $a \perp b = 0$

VETTORI LINEARMENTE INDEPENDENTI: k vettori $v^{(1)}, v^{(2)}, \dots, v^{(k)}$ ∈ \mathbb{C}^k si dicono linearmente indipendenti se $\alpha_1 v^{(1)} + \alpha_2 v^{(2)} + \dots + \alpha_k v^{(k)} = 0$ con $\alpha_i \in \mathbb{C}$, $i=1,2,\dots,k$ implica $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$

DETERMINANTE: Ad una matrice quadrata A si associa un numero complesso chiamato determinante ($\det A$) → costo computazionale MOLTO ELEVATO con conseguente INTRODUZIONE di ERROI

Se A è diagonale o rettangolare si ha che $\det(A) = \prod_{i=1}^n a_{ii}$ e se è quadrata risulta $\det(A) = \det(A^T)$. Una matrice si dice SINGOLARE se $\det(A)=0$

MINORI: Data una matrice $A \in \mathbb{C}^{m \times n}$ ed un intero $K \leq \min\{m, n\}$ si dice minore di ordine K il determinante di una matrice ottenuta da A prendendo tutti gli elementi nella intersezione di K righe e colonne comunque finite

MINORI PRINCIPALI: Data una matrice $A \in \mathbb{C}^{n \times n}$ si dicono minori principali di ordine K i determinanti delle sottomatrici di ordine K estratte da A e aventi diagonale principale composta da elementi della diagonale principale di A

MINORE PRINCIPALE di TESTA: Determinante della sottomatrice di ordine K formata dalle prime K righe e le prime K colonne

RANGO: Data una matrice $A \in \mathbb{C}^{m \times n}$, si dice di range o caratteristica della matrice A il numero $r(A)$ dato dell'ordine più alto dei suoi minori diversi da 0

TEOREMA di BINET-CACCHIY: Date due matrici $A \in C^{m \times n}$, $B \in C^{n \times m}$, il determinante della matrice prodotto $C = AB \in C^{m \times m}$ è nullo se $m > n$, altrimenti è zero alla somma di prodotti di tutti i possibili minori di ordine minimo di A per i corrispondenti minori di B

COROLARIO: Se $m = n$ si ha che $\det(AB) = \det(A) \cdot \det(B)$

MATRICE INVERSA: Una matrice $A \in C^{n \times n}$ non singolare ha la cosiddetta matrice inversa A^{-1} che verifica la condizione $AA^{-1} = A^{-1}A = I$

↓ dal corollario di Binet

$$\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det(I) = 1 \xrightarrow{\text{quindi}} \det(A^{-1}) = \frac{1}{\det(A)}$$

- CLASSI di MATRICI:**
- HERMITIANA** → $A = A^H$: deve avere TUTTI elementi REALI
 - ANTIHERMITIANA** → $A = -A^H$: gli elementi della diagonale devono essere tutti uguali a 0 o immaginari pure
 - UNITARIA** → $A^H A = A A^H = I$: la transposta coniugata coincide con l'inversa
 - NORMALE** → $A^H A = A A^H$
 - SIMMETRICA** → $A = A^T$: Non vuol dire che sia hermitiana. Coincidono solo se la matrice è REALE
 - ANTISIMMETRICA** → $A = -A^T$
 - ORTOGONALE** → Reale ed unitaria

Data una matrice hermitiana A ed un vettore $x \in C^n$, lo scalare $x^H Ax \in \mathbb{R}$ poiché $(x^H Ax)^H = x^H A^H x = x^H Ax$

↓ numero perché la matrice era hermitiana

Una matrice hermitiana A si dice **DEFINITA POSITIVA** se $x^H Ax > 0$ per ogni $x \in C^n$ con $x \neq 0$. Se invece accade che $x^H Ax \geq 0$ si dice **SEMI DEFINITA POSITIVA**

MATRICE di PERMUTAZIONE: $P \in C^{n \times n}$ è detta matrice di permutazione se è ottenuta dalla matrice identica operando una permutazione delle colonne (o delle righe) → Matrice ORTOGONALE

POSTMULTIPLICAZIONE: AP → Se nella matrice di permutazione sono scambiate la prima e la terza colonna, scambiamo scambiare anche nella matrice prodotto

PREMULTIPLICAZIONE: PA → Se nella matrice P sono scambiate la prima e la terza riga, scambieremo scambiare anche nella matrice prodotto

Se P si ottiene permutando le colonne di I , P^T si ottiene con la stessa permutazione delle righe di I

PREDOMINANZA DIAGONALE FORTE: Una matrice $A \in C^{n \times n}$ si dice a predominanza diagonale forte se $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$, $i = 1, 2, \dots, n$

PREDOMINANZA DIAGONALE DEBOLE: Una matrice $A \in C^{n \times n}$ si dice a predominanza diagonale debole se $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$, $i = 1, 2, \dots, n$ e per almeno un indice x , $1 \leq x \leq n$ si ha $|a_{xx}| > \sum_{j \neq x} |a_{xj}|$

Per una riga si deve verificare la predominanza forte

MATRICE CONVERGENTE: Una matrice $A \in C^{n \times n}$ si dice convergente se

$$\lim_{k \rightarrow \infty} A^k = 0 \rightarrow \text{matrice nulla}$$

Limite delle potenze successive

o SISTEMI LINEARI: Dati una matrice $A \in \mathbb{C}^{n \times n}$ ed un vettore $b \in \mathbb{C}^n$, un sistema di n equazioni lineari in n incognite si rappresenta nella forma $Ax=b$ dove x è il vettore delle incognite

TEOREMA di ROUHÉ - CAPELLI: Un sistema lineare $Ax=b$ ammette soluzioni se e solo se $r(A) = r(A|B)$, dove $A|B$ è la matrice A a cui abbiamo aggiunto una colonna. Se $r(A) = n$ la soluzione è UNICA mentre se $r(A) < n$ l'insieme delle soluzioni è un sottoinsieme di \mathbb{C}^n di dimensione $n-r(A)$

Se $\det(A) \neq 0$ il sistema $Ax=b$ si dice NORMALE \rightarrow Regola di Cramer: $x_i = \frac{\det(A_i)}{\det(A)}$, $i = 1, 2, \dots, n$

A_i è la matrice con la i -esima colonna sostituita dal vettore b

Se $b=0$ il sistema è OSCERVO ed ha vicine diverse soluzioni perché $r(A) = r(A|b)$.

Se risulta $\det(A) \neq 0$ l'unica soluzione è $x=0$

ESEMPIO MATEMATICO di PERMUTAZIONE

$$P = (e^{(2)} | e^{(3)} | e^{(1)}) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad P^T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad P^T A P = P^T (AP) = \begin{pmatrix} 5 & 6 & 4 \\ 3 & 9 & 7 \\ 2 & 3 & 1 \end{pmatrix}$$

$$P P^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 2 & 3 \\ 6 & 5 & 6 \\ 7 & 3 & 9 \end{pmatrix} \quad AP = \begin{pmatrix} 2 & 3 & 1 \\ 5 & 6 & 4 \\ 8 & 9 & 7 \end{pmatrix}$$

$$PA = \begin{pmatrix} 7 & 8 & 9 \\ 1 & 2 & 3 \\ 6 & 5 & 6 \end{pmatrix}$$

MATRICI RIDUCIBILI

PARTIZIONAMENTO a BLOCCHI: $A = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ \rightarrow gli elementi sono SOTTO-MATRICI di A \rightarrow il caso più importante si ha nel caso in cui i blocchi diagonali sono quadrati

Si poniamo avere matrice TRIANGOLARE a BLOCCHI o DIAGONALE a BLOCCHI

Per le matrici diagonali a blocchi si verifica $\det(A) = \prod_{i=1}^k \det(A_{ii})$

ESEMPIO

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

MATRICE RIDUCIBILE: Una matrice $A \in \mathbb{C}^{n \times n}$ si dice riducibile se esiste una matrice di permutazione P tale che la matrice $P^T A P$ sia partizionabile nella forma $B = P^T A P = \begin{pmatrix} B_{11} & 0 \\ 0 & B_{22} \end{pmatrix}$ con BLOCCHI DIAGONALI QUADRATI

Triangolare inferiore a blocchi

Se la matrice B ha qualche blocco diagonale ancora riducibile si può continuare a ridurre fino ad arrivare alla FORMA RIDOTTA non ulteriormente riducibile

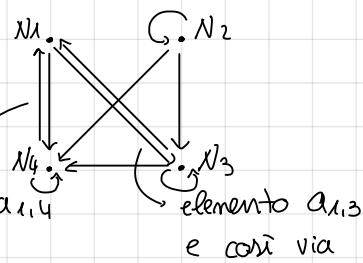
GRAFO ORIENTATO: Data una matrice $A \in \mathbb{C}^{n \times n}$, fissati n punti (modi) N_1, N_2, \dots, N_n , si dice grafo orientato associato ad A , il grafo che si ottiene congiungendo N_i a N_j con un cammino orientato da N_i a N_j per ogni arco

FORTEMENTE CONNESSO: se da ogni modo N_i è possibile raggiungere qualunque altro modo N_j seguendo un cammino orientato anche passante per altri modi

TEOREMA: $A \in \mathbb{C}^{n \times n}$ si dice IRREDUCIBILE se il grafo orientato risultante comune

ESEMPI

$$\textcircled{1} \quad A = \begin{pmatrix} 0 & 0 & 1 & 3 \\ 0 & 1 & 1 & -1 \\ 2 & 0 & 3 & 6 \\ -2 & 0 & 0 & -1 \end{pmatrix}$$



	Nc	Ns
N1	N1 N3 N4	N2
N2	N1 N2 N4	-
N3	N1 N3 N6	N2
N4	N1 N3 N4	N2

NON è forte mente comune, quindi è RIDUCIBILE

Per costruire la matrice di permutazione si mettono in testa le colonne contenenti i nodi raggiungibili $\rightarrow P = (e^{(1)} | e^{(2)} | e^{(3)} | e^{(4)})$ — Matrice di permutazione che riduce A

$$B = P^T A P = P^T \begin{pmatrix} 0 & 1 & 3 & 0 \\ 0 & 1 & -1 & 1 \\ 2 & 3 & 6 & 0 \\ -2 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 3 & 0 \\ 2 & 3 & 6 & 0 \\ -2 & 0 & -1 & 0 \\ 0 & 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{pmatrix}$$

$$\textcircled{2} \quad A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 5 & -1 & 2 & 0 \\ 0 & -5 & 1 & 7 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

	Nc	Ns
N1	N1	N2 N3 N4
N2	TUTTI	-
N3	TUTTI	-
N4	N4	N1 N2 N3

$P = (e^{(1)} | e^{(2)} | e^{(3)} | e^{(4)})$ — Matrice IDENTICA: La matrice simile era già triangolare

OPPURE

$$P = (e^{(4)} | e^{(1)} | e^{(2)} | e^{(3)}) \rightarrow B = P^T A P = P^T \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 5 & -1 & 2 \\ 7 & 0 & -5 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & 0 \\ B_{31} & B_{32} & B_{33} & 0 \\ 0 & 0 & -1 & 2 \\ 7 & 0 & -5 & 1 \end{pmatrix} = \begin{pmatrix} B_{11} & 0 & 0 \\ 0 & B_{22} & 0 \\ B_{31} & B_{32} & B_{33} \end{pmatrix}$$

$$\textcircled{3} \quad A = \begin{pmatrix} 0 & 1 & & \\ & \ddots & 0 & \\ 0 & & \ddots & \\ 1 & & & 0 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

RISOLUZIONE di un SISTEMA LINEARE

Abbiamo $\rightarrow A \in \mathbb{C}^{n \times n}$ $B = P^T A P = \begin{pmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{pmatrix}$
Vogliamo risolvere

$$\underbrace{Ax = b}_{\text{premultip}} \xrightarrow{\text{PT}} P^T A x = P^T b \xrightarrow{\text{posso scomporre}} P^T A \underbrace{I x}_{\substack{\text{I=prodotto} \\ \text{matrice per invertire}}} = P^T b \xrightarrow{\substack{\text{P}^T A \text{ e } P^T \text{ sono inverse} \\ \text{B}}} \underbrace{\frac{P^T A}{B} \underbrace{P^T x}_{q}}_{q} = \underbrace{\frac{P^T b}{B}}_{C} \xrightarrow{} B q = C$$

Scrivendo per esteso B: $\begin{pmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$

\downarrow svolgendo il prodotto

$$\begin{cases} B_{11} q_1 = c_1 \\ B_{11} q_1 + B_{21} q_2 = c_2 \end{cases} \xrightarrow{\substack{\text{trovo } q_1 \\ \text{sostituendo}}} \begin{cases} B_{11} q_1 = c_1 \\ B_{22} q_2 = c_2 - B_{21} q_1 \end{cases} \xrightarrow{\text{trovo } q_2}$$

Alla fine ho trovato $q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$ e per ottenerne x facciamo $x = Pq$

Se la matrice A NON è singolare anche la matrice B non lo è perché, per Binet, abbiamo: $\det(P^T A P) = \det(P^T) \det(A) \det(P) = \det(P^T) \det(P) \det(A) = \det(I) \det(A) = \det(B) = \det(A)$

○ AUTOVALORI e AUTOVETTORI

DATA una matrice $A \in \mathbb{C}^{n \times n}$ si dice autovettore di A ogni numero $\lambda \in \mathbb{C}$ tale che il sistema lineare $Ax = \lambda x$, $x \in \mathbb{C}^n$, abbia soluzioni $x \neq 0$. Il vettore x è detto autovettore diretto associato all'autovettore λ intendendo che x ed ogni vettore kx ($k \in \mathbb{C}$, $k \neq 0$) rappresentano lo stesso autovettore. Analogamente, è detto autovettore similire un vettore $y \in \mathbb{C}^n$ tale che $y^T A = \lambda y^T$

↓

Per il teorema di Rouché-Capelli, un sistema omogeneo ha soluzioni non nulle se e solo se la matrice dei coefficienti è singolare \rightarrow il sistema $Ax = \lambda x$ equivale a $(A - \lambda I)x = 0$ e quindi gli autovettori di A sono tutti i λ tali che $\det(A - \lambda I) = 0$ \rightarrow deriva che $\det(A - \lambda I) = \det((A - \lambda I)^T) = \det(A^T - \lambda I)$, quindi A e A^T hanno gli stessi autovalori

↓

Indichiamo con $P(\lambda) = \det(A - \lambda I)$ il polinomio caratteristico della matrice $A \in \mathbb{C}^{n \times n}$.

Prese $\lambda_1, i=1, 2, \dots, n$ le soluzioni dell'equazione caratteristica per Ruffini si ha $P(\lambda) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n) = (-1)^n \lambda^n + (-1)^{n-1} \alpha_1 \lambda^{n-1} + (-1)^{n-2} \alpha_2 \lambda^{n-2} + \dots + \alpha_{n-2} \lambda^2 - \alpha_{n-1} \lambda + \alpha_n$, dove i coefficienti $\alpha_i, i=1, 2, \dots, n$ sono, ciascuno, la somma dei minori principali di ordine i estratti dalla matrice $A \rightarrow \alpha_1 = \sum_{j=1}^n a_{1j} = \sum_{i=1}^n \lambda_i \rightarrow$ TRACCIA: $\text{tr}(A)$

$\alpha_n = \det(A) = \prod_{i=1}^n \lambda_i \rightarrow$ Dobbiamo prendere il secondo addendo della fatt.

↳ $\det = 0 \rightarrow$ almeno un autovettore è nullo

Per la traccia si ha: $(-1)^n (\lambda^{n-1}(-\lambda_1) + \lambda^{n-1}(-\lambda_2) + \lambda^{n-1}(-\lambda_3) + \dots + \lambda^{n-1}(-\lambda_n)) = (-1)^n \lambda^{n-1}(-\lambda_1 - \lambda_2 - \dots - \lambda_n) = (-1)^{n+1} \lambda^{n-1} (\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_n)$

RAGION SPECTRALE: Si dice raggio spettrale della matrice $A \in \mathbb{C}^{n \times n}$ il numero reale non negativo $\rho(A) = \max_{1 \leq i \leq n} |\lambda_i| \rightarrow$ Massimo modulo degli autovettori

Matrice convergente: Una matrice $A \in \mathbb{C}^{n \times n}$ è convergente se e solo se risulta $\rho(A) < 1$

TRASFORMAZIONE PER SIMILITUDINE: Data una matrice $A \in \mathbb{C}^{n \times n}$ ed una matrice $S \in \mathbb{C}^{n \times n}$ non singolare, si dice trasformata per similitudine della matrice A , la matrice $B = S^{-1}AS$ e le matrici A e B si dicono simili (\sim) \rightarrow Due matrici simili hanno gli stessi autovettori e, per ogni autovettore λ , se x è autovettore di A , allora $S^{-1}x$ è autovettore di B

DIMOSTRAZIONE

$B = S^{-1}AS \rightarrow \det(B - \lambda I) = \det(S^{-1}AS - \lambda I) = \det(S^{-1}AS - \lambda S^{-1}S) = \det[S^{-1}(A - \lambda I)S] = \det(S^{-1})\det(A - \lambda I)\det(S) = \det(S^{-1}S)\det(A - \lambda I) = \det(I)\det(A - \lambda I) = \det(A - \lambda I) \rightarrow$ Il polinomio caratteristico di A e B sono uguali e quindi sono uguali anche gli autovettori

↳ per gli autovettori

$Bx = \lambda x \rightarrow S^{-1}ASx = \lambda S^{-1}Sx \rightarrow ASx = \lambda Sx \rightarrow x = Sx$, quindi $x = Sx$, \square

TEOREMA: Se λ è autovettore di A , allora $\lambda^k, k \in \mathbb{N}$, è autovettore di A^k e gli autovettori di A sono anche autovettori di A^k ma NON è detto che sia vero l'opposto

↓ DIMOSTRAZIONE $\xrightarrow{\text{prem. } A} Ax = \lambda x \rightarrow A^2x = A\lambda x \rightarrow A^2x = \lambda^2 x \rightarrow A^kx = \lambda^k x : \text{verificato}$

Se A è non singolare il teorema può avere estensione a $k \in \mathbb{Z}$ e gli autovettori di A^{-1} sono i reciproci degli autovettori di A

↳ DIMOSTRAZIONE: $Ax = \lambda x \rightarrow A^{-1}Ax = \lambda A^{-1}x \rightarrow x = A^{-1}x$

TEOREMA: gli autovettori di una matrice hermitiana sono tutti reali

↓ DIMOSTRAZIONE

$A = A^H$, quindi $x^H A x \in \mathbb{R}$. Per gli autovettori si ha $x^H A x = \lambda x^H x \rightarrow \frac{x^H A x}{x^H x} = \lambda \in \mathbb{R}$: quoziente di pagherà

$\sum_{i=1}^n |x_i|^2$ che non può essere 0 se x è autovettore ed è sempre > 0

MOLTEPLICITÀ di un AUTOVALORE

Algebrica: data molteplicità algebrica $\alpha(\lambda)$ di un autovettore λ è la molteplicità di λ come radice dell'equazione caratteristica

Geometrica: data molteplicità geometrica $r(\lambda)$ di λ è la dimensione dello spazio delle soluzioni del sistema lineare omogeneo $(A - \lambda I)x = 0 \rightarrow$ numero degli autovettori linearmente indipendenti associati all'autovettore: $r(\lambda) = n - r(A - \lambda I)$

TEOREMA: $1 \leq r(\lambda) \leq \alpha(\lambda) \leq n \rightarrow A$ è diagonabile se esiste X non singolare tali che

$$X^{-1}AX = D, \quad D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

↓

Deve accadere che $\alpha(\lambda) = r(\lambda)$ per ogni autovettore

Caso particolare: Se la matrice A ha autovettori a due a due distinti è diagonabile perché $\alpha(\lambda) = r(\lambda) = 1$

TEOREMA: Sia λ autovettore di A e $q \in \mathbb{C}$, allora $B = A + qI$ ha come autovettore $\mu = \lambda + q$ con molteplicità algebrica e geometrica pari a quelle di λ e B ha gli stessi autovettori di A

DIMOSTRAZIONE

$$\det(B - \mu I) = \det(A + qI - \mu I) = \det(A - (\mu - q)I) \rightarrow \lambda = \mu - q$$

Per gli autovettori: $Bx = \mu x \rightarrow (A + qI)x = \mu x \rightarrow Ax = (\mu - q)x \rightarrow Ax = \lambda x$

ESEMPI

① $J = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ 1 & & & 0 \end{pmatrix} \in \mathbb{R}^{n \times n}$ → Calcolare gli autovettori
Matrice reale e simmetrica → Gli autovettori sono reali

↓ perché è simmetrica

$J^2 = J \cdot J = J^T \cdot J = I \rightarrow \lambda^2 = 1$, quindi $\lambda = \pm 1 \rightarrow \text{tr}(J) = \sum_{i=1}^n \lambda_i$: ha 2 possibilità, infatti
vai 0 se n è pari o 1 se n è dispari

Se n è pari: $\lambda_1 = 1 \quad \alpha(1) = n/2$
 $\lambda_2 = -1 \quad \alpha(-1) = n/2$

Se n è dispari: $\lambda_1 = 1 \quad \alpha(1) = k+1$

$n = 2k+1, k \in \mathbb{N} \quad \lambda_2 = -1 \quad \alpha(-1) = k$
 $\lambda_3 = 1$

② $A = I + \underbrace{ab^T}_{\in \mathbb{R}^{n \times n}}$

$a, b \in \mathbb{R}^m, I \in \mathbb{R}^{n \times n}$
vettori COLONNA

$B = A - I = ab^T = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} (b_1 b_2 \dots b_n) = \begin{pmatrix} a_1 b_1 + a_2 b_2 + \dots + a_n b_n \\ a_1 b_1 + a_2 b_2 + \dots + a_n b_n \\ \vdots \\ a_1 b_1 + a_2 b_2 + \dots + a_n b_n \end{pmatrix}$

2 righe DIPENDENTI: $\det = 0$ anche per successivi $\rightarrow \alpha_2 = \alpha_3 = \dots = \alpha_n = 0$

$$\varphi(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} a^T b \lambda^{n-1} = \lambda^n - a^T b \lambda^{n-1} = \lambda^{n-1} (\lambda - a^T b) = 0 \rightarrow \lambda^{n-1} \text{ relazioni} = 0 \rightarrow \lambda = a^T b$$

Quindi $\mu_1 = \mu_2 = \dots = \mu_{n-1} = 1$ e $\mu_n = 1 + a^T b$

LOCALIZZAZIONE di AUTOVALORI

cerchi di gerogorin: Data una matrice $A \in \mathbb{C}^{n \times n}$, definiamo gli insiemi

$$F_i = \{\lambda \in \mathbb{C} : |\lambda - a_{ii}| \leq p_i\}, \text{ dove } p_i = \sqrt{\sum_{j=1}^n |a_{ij}|^2}, i = 1, 2, \dots, n$$

Il cerchio F_i ha centro in a_{ii} e raggio $p_i \rightarrow$ somma dei moduli degli elementi della riga i -esima esclusi gli elementi diagonali

1° teorema di gerogorin: Se λ è autovettore di $A \in \mathbb{C}^{n \times n}$ risulta $\lambda \in F = \bigcup_{i=1}^n F_i$ (unione dei cerchi)

DIMOSTRAZIONE

$\lambda x = \lambda x$ perché λ è autovalore, x è autovettore: $x = (x_1, x_2, \dots, x_n)^T$, $x \neq 0$

Troviamo x_k tale che: $|x_k| > |x_i| \quad i=1, 2, \dots, n$

del sistema lineare considero quindi la k -esima equazione

$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = \lambda x_1 \rightarrow$ portiamo al II membro il termine che coinvolge x_k
e ottieniamo: $\sum_{j=1}^n a_{1j}x_j = \lambda x_1 - a_{1k}x_k = (\lambda - a_{1k})x_1$ quando cui moduli $\left| \sum_{j=1}^n a_{1j}x_j \right| = |\lambda - a_{1k}| |x_1|$

per la disugualanza triangolare

$$\sum_{j=1}^n |a_{1j}| |x_j| \geq |\lambda - a_{1k}| |x_1|$$

sostituendo
 x_1 a x_j
faccio una
simplificazione

$$\sum_{j=1}^n |a_{1j}| |x_j| \geq |\lambda - a_{1k}| |x_1|$$

mettendo in evidenza nella sommatoria

$\rightarrow \sum_{j=1}^n |a_{1j}| > |\lambda - a_{1k}|$ □

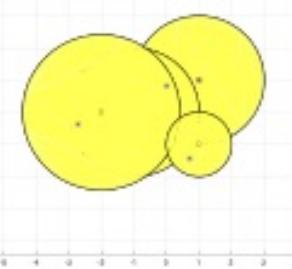
Appartenere ad A HENNO un cerchio (e quindi anche all'unione)

distanza di λ dal centro del k -esimo cerchio

raggio del k -esimo cerchio

ESEMPIO

$$A = \begin{pmatrix} 1+i & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 \\ 0 & 1+i & -2 & 1 \\ 0 & -1 & 0 & 1-i \end{pmatrix} \rightarrow$$



All'unione dei 4 cerchi appartengono gli autovettori della matrice

Corollario: Una matrice a predominanza diagonale forte non è singolare

DIMOSTRAZIONE

Per il I teorema di Gershgorin ogni cerchio ha il centro a distanza $|a_{ii}|$ dall'origine degli anni ed il raggio $\rho_i = \sum_{j \neq i} |a_{ij}|$ → Per la predominanza diagonale forte VESSONO contingenti l'origine e quindi lo 0 non è autovettore della matrice ed il determinante non può essere nullo

2° teorema di Gershgorin: Se M_1 è l'unione di K cerchi e M_2 è l'unione dei rimanenti $n-K$ e si ha $M_1 \cap M_2 = \emptyset$ allora K autovoltori appartengono a M_1 e $n-K$ a M_2

3° teorema di Gershgorin: Se A è una matrice irreducibile allora se un autovoltore appartiene alla frontiera dell'unione dei cerchi di Gershgorin esso appartiene alla frontiera di tutti i cerchi che costituiscono l'unione F

Corollario II teorema: Se una matrice presenta cerchi di Gershgorin due a due disgiunti allora gli autovoltori sono a due a due distinti

Corollario III teorema: Una matrice A a predominanza diagonale debole ed irreducibile è NON singolare

DIMOSTRAZIONE

La predominanza diagonale debole consente ai cerchi di toccare per l'origine ed almeno uno non deve farlo. Lo 0 può appartenere al massimo ad $n-1$ cerchi. Non appartenendo alla frontiera di tutti i cerchi lo 0 non è un autovettore e quindi $\det(A) \neq 0$

MATRICE DI FROBENIUS

eq. algebrica: $x^K + a_{K-1}x^{K-1} + \dots + a_1x + a_0 = 0$

Matrice di Frobenius: $F = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{K-2} & -a_{K-1} \end{pmatrix}$

→ Si può verificare che questa è l'equazione associata della matrice di Frobenius

È possibile calcolare le radici utilizzando il teorema di Gershgorin → gli autovoltori sono le soluzioni dell'equazione

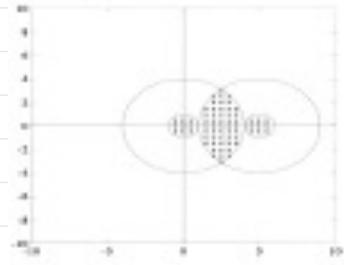
ESEMPIO

$$x^3 - 5x^2 + 3x - 1 = 0 \longrightarrow F = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 5 \end{pmatrix} \quad F_1 = F_2 = \{z \in \mathbb{C} \mid |z| \leq 1\} \text{ la regione del piano individuata contiene le soluzioni}$$

Per stringere il piano delle soluzioni posso considerare la matrice trasposta
 $\hookrightarrow F^T = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 5 \end{pmatrix} \longrightarrow G_1 = \{z \in \mathbb{C} \mid |z| \leq 1\}$
 $G_2 = \{z \in \mathbb{C} \mid |z| \leq 4\}$
 $G_3 = \{z \in \mathbb{C} \mid |z-5| \leq 1\}$

va intersecata con la regione di spazio individuata precedentemente

L'insieme individuato è più ristretto sia di F che di G



NORME

In molti casi è necessario confrontare vettori e matrici

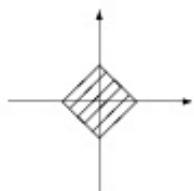
norma vettoriale: Si dice norma vettoriale, e si indica con $\|x\|$, una funzione $\|\cdot\|: \mathbb{C}^n \rightarrow \mathbb{R}^+$ che verifica le seguenti condizioni: a. $\|x\| \geq 0 \iff x=0$

varie norme

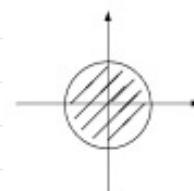
$$\text{Norma 1: } \|x\|_1 = \sum_{i=1}^n |x_i|$$

$$\text{Norma 2 (euclidea): } \|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$$

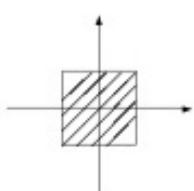
Sfera unitaria: $S = \{x \in \mathbb{R}^n; \|x\| \leq 1\}$ Norma 1



Norma 2



Norma ∞



teorema: Ogni norma vettoriale è uniformemente continua su \mathbb{C}^n

teorema di equivalenza tra norme: Date due norme vettoriali $\|x\|_p$ e $\|x\|_q$, esistono $\alpha, \beta \in \mathbb{R}^+$ tali che $\alpha \|x\|_p \leq \|x\|_q \leq \beta \|x\|_p \forall x \in \mathbb{C}^n$

norma matriciale: Si dice norma matriciale, e si indica con $\|A\|$, una funzione $\|\cdot\|: \mathbb{C}^{n \times n} \rightarrow \mathbb{R}^+$ tale che: a. $\|A\| = 0 \iff A = 0$

$$\|\alpha A\| = |\alpha| \|A\| \quad \forall A \in \mathbb{C}^{n \times n}, \forall \alpha \in \mathbb{C}$$

$$\|A+B\| \leq \|A\| + \|B\| \quad \forall A, B \in \mathbb{C}^{n \times n}$$

$$\|AB\| \leq \|A\| \|B\| \quad \forall A, B \in \mathbb{C}^{n \times n}$$

norma indotta: $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$ → Una norma matriciale si dice compatibile con una norma vettoriale se $\|Ax\| \leq \|A\| \|x\| \forall x \in \mathbb{C}^n$

Le norme indotte sono coerenti con le norme vettoriali da cui provengono: $\frac{\|Ax\|}{\|x\|} \leq \|A\| \quad \sup$

$$\text{Norma 1: } \|A\| = \max_j \sum_{i=1}^n |a_{ij}|$$

$$\text{Norma } \infty: \|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$$

$$\text{Norma 2: } \|A\|_2 = \sqrt{\rho(A^H A)}$$

$$\text{Norma di Frobenius: } \|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2} \longrightarrow \|I\|_F = \sqrt{n} \text{ mentre } \|I\| = \sup_{\|x\|=1} \|Ax\| = 1$$

Quindi: Ad una matrice $A \in \mathbb{C}^{n \times n}$ possiamo attribuire 3 numeri: determinante, raggio spettrale e norma → gli ultimi due sono numeri reali non negativi

Teorema di Hurwitz: Sia $A \in \mathbb{C}^{n \times n}$. Per ogni norma matriciale vale la relazione $\rho(A) \leq \|A\|$

DIMOSTRAZIONE

$$Ax = \lambda x \quad B = (x|0|0|...|0) \in \mathbb{C}^{n \times n}$$

$AB = \lambda B$? Vero perché solo la prima colonna è f0 \rightarrow Adesso ho due matrici
 ↳ Ponendo alla norma avremo: $\|AB\| = \|\lambda B\| = |\lambda| \|B\|$ proprietà norma $\Rightarrow \|A\| \|B\| \geq |\lambda| \|B\|$
 quindi non nulla: x autovettore

Quindi $\|A\| \geq |\lambda|$: avendo che λ è un autovettore qualsiasi si avrà $\|A\| > \rho(A)$

Corollario: Affinché una matrice sia convergente è sufficiente che una sua norma

risulti minore di 1

corollario: Se $A \in \mathbb{C}^{n \times n}$, gli autovettori di A appartengono al cerchio $\{z \in \mathbb{C} : |z| \leq \|A\|\}$
 ↳ La relazione $\rho(A) = \|A\|$ vale per le matrici hermitiane

↓
 Con la norma 2: $\|A\|_2 = \sqrt{\rho(A^H A)} = \sqrt{\rho(A^*)} = \sqrt{\rho^2(A)} = \rho(A)$

matrici di rotazione:

$$G_{rt} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & c & -s \\ & & & s & c \end{pmatrix} \leftarrow r \quad \leftarrow t$$

$G_{rt} \in \mathbb{R}^{n \times n}$ e $c^2 + s^2 = 1$
 quindi c ed s sono interpretati come seno e coseno di un
 opportuno angolo

È riducibile e si possono portare in testa le colonne contenenti (c, s) e $(-s, c)$. In questo modo, alle colonne successive avremo un blocco identico

$$\rho = (e^{(r)} | e^{(t)} | e^{(u)} | e^{(v)} | \dots | e^{(n)}) \rightarrow B = P^T G_{rt} P = \begin{pmatrix} c & -s \\ -s & c \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots & 1 \end{pmatrix} \rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_{n-2} = 1$$

Per trovare gli altri autovettori: $\lambda^2 - 2\lambda + 1 = 0 \rightarrow \lambda = c \pm \sqrt{c^2 - 1} = c \pm i\sqrt{s^2}$ con modulo $\sqrt{c^2 + s^2} = 1$

Dato che ho $n-2$ autovettori = 1 e gli altri due complessi coniugati di modulo c il determinante sarà = 1 → La matrice NON è convergente
 Inoltre, per il teorema di Hersh: $|\lambda_i| \leq \|G_{rt}\|_2 = \sqrt{\rho(G_{rt} G_{rt}^T)} = 1 \rightarrow |\lambda_i| = 1, i = 1, 2, \dots, n$

Sistemi lineari

INTRODUZIONE

Sistema lineare $Ax=b$ dato da:

$$\begin{array}{l} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{array} \rightarrow a_{ij} e b_i sono reali o complessi (caso più raro)$$

Vettore $x^T = (x_1, \dots, x_n) \in \mathbb{C}^n$ che verifica tutte le equazioni

Scrivendo il sistema lineare nella forma $Ax=b$, si ha che A è la matrice dei termini noti, b è il vettore dei termini noti e x è il vettore delle incognite

metodi diretti: se si lavora con aritmetica esatta ed un insieme finito di operazioni si giunge alla soluzione esatta

metodi iterativi: da soluzione viene approssimata dal limite di una successione di vettori

METODI DIRETTI

1. metodo di Cramer: dato il sistema lineare $Ax=b$ la soluzione si ottiene calcolando

$$x_i = \frac{\det(A_i)}{\det(A)} \text{ con } i=1, 2, \dots, n \text{ ed } A_i \text{ ottenuta sostituendo } b \text{ all'i-esima colonna}$$

Elevato costo computazionale: vanno calcolati $n+1$ determinanti

n	metodo di Cramer	metodo di Gauss
12	103 minuti	$7.1 \cdot 10^{-4}$ secondi
13	24 ore	$8.9 \cdot 10^{-4}$ secondi
15	15 giorni	$1.1 \cdot 10^{-3}$ secondi
50	$4.9 \cdot 10^{52}$ anni	$4.4 \cdot 10^{-2}$ secondi

2. metodo di Gauss: consiste nel trasformare il sistema equivalente $Rx=C$ dove $R \in \mathbb{C}^{n \times n}$ è una matrice triangolare superiore con $r_{ii} \neq 0$, $i=1, 2, \dots, n$

Sistema banalmente risolvibile tramite backwards substitution

FORMA: $\begin{array}{cccccc} r_{11}x_1 & + & r_{12}x_2 & + \dots & + & r_{1n}x_n = c_1 \\ r_{22}x_2 & + \dots & + & r_{2n}x_n = c_2 \\ \dots & & \dots & & \dots & \\ r_{nn}x_n & = & c_n \end{array} \rightarrow$ FORMULE RISOLUTIVE

$$x_n = c_n / r_{nn}$$

$$x_i = (c_i - \sum_{j=i+1}^n r_{ij}x_j) / r_{ii}, i=n-1, \dots, 1$$

Come trasformare

1. $a_{11} \neq 0$ \rightarrow Si elimina x_1 da tutte le equazioni successive sottraendo membro a membro della i -esima equazione, $i=2, 3, \dots, n$, la prima equazione i cui membri sono stati moltiplicati per il moltiplicatore $l_{i1} = \frac{a_{i1}}{a_{11}}$ \rightarrow Questo ed i successivi si dicono elementi pivotali

ESEMPIO:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$l_{i1} = \frac{a_{i1}}{a_{11}} \quad i=2, 3$$

$$l_{21} = \frac{a_{21}}{a_{11}} \quad l_{31} = \frac{a_{31}}{a_{11}}$$

$$\text{Seconda riga: } a_{21} \quad a_{22} \quad a_{23} \quad -$$

$$-\frac{a_{21}}{a_{11}} (a_{11} \quad a_{12} \quad a_{13}) = -a_{21} \quad -\frac{a_{21}a_{12}}{a_{11}} \quad -\frac{a_{21}a_{13}}{a_{11}} =$$

per andicare al siste.
lineare MODIFICATO

$$\begin{array}{ccc|c} 0 & a_{22}^{(2)} & a_{23}^{(2)} & b_2 - \frac{a_{21}}{a_{11}} b_1 = b_2^{(2)} \\ 0 & a_{32}^{(2)} & a_{33}^{(2)} & b_3 - \frac{a_{31}}{a_{11}} b_1 = b_3^{(2)} \end{array}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} \\ 0 & a_{32}^{(2)} & a_{33}^{(2)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2^{(2)} \\ b_3^{(2)} \end{pmatrix}$$

Per terminare azzeriamo la seconda incognita della terza riga con $l_{32} = \frac{a_{32}^{(2)}}{a_{22}^{(2)}}$

2. $A_{22} \neq 0 \rightarrow$ eliminazione di x_2 da tutte le equazioni seguenti utilizzando

$$l_{i2} = \frac{a_{i1}^{(2)}}{a_{22}^{(2)}}, \quad i = 3, 4, n$$

n-1: Iterazione del sistema n-1 volte:

$$\begin{aligned} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \dots + a_{1n}^{(1)}x_n &= b_1^{(1)} \\ a_{22}^{(2)}x_2 + \dots + a_{2n}^{(2)}x_n &= b_2^{(2)} \\ &\vdots \quad \dots \quad \dots \\ a_{nn}^{(n)}x_n &= b_n^{(n)} \end{aligned}$$

⚠ ATT: si può fare solo se $a_{11}^{(1)} \neq 0, a_{22}^{(2)} \neq 0, \dots, a_{nn}^{(n)} \neq 0 \rightarrow$ Altrimenti scambia colonne
MA $\det = 0$ (> 1 soluzione)

$$a_{11}^{(2)} = a_{22} - \frac{a_{11}}{a_{22}}a_{12}f_0 \rightarrow a_{11}a_{22} - a_{11}a_{12}f_0$$

↓

DETERMINANTE del blocco di testa di dimensione 2

Generalizzando: La matrice iniziale deve avere i minori principali di testa diversi da 0 → Verifichiamo questa proprietà le matrici simmetriche definite (positive o negative)

L'algoritmo in genere deve essere modificato per impedire la propagazione di errori

costo computazionale: $\frac{n^3}{3} + n^2 - \frac{n}{3}$, mentre Cramer aveva un costo di: $(n+1)!/(n-1)$

ESEMPIO di RISOLUZIONE

$$\textcircled{1} \quad \left| \begin{array}{ccc|c} -2 & 1 & 1 & x_1 \\ -1 & 3 & 1 & x_2 \\ 1 & 1 & 2 & x_3 \end{array} \right| = \left| \begin{array}{cc|c} -2 & & \\ -3 & & \\ 2 & & \end{array} \right| \rightarrow \left| \begin{array}{ccc|c} -2 & 1 & 1 & -2 \\ -1 & 3 & 1 & -3 \\ 1 & 1 & 2 & 2 \end{array} \right| \quad l_{21} = \frac{1}{2}$$

$$\left| \begin{array}{ccc|c} -2 & 1 & 1 & -2 \\ 0 & \frac{5}{2} & \frac{1}{2} & -2 \\ 0 & \frac{3}{2} & \frac{5}{2} & 1 \end{array} \right| \quad l_{31} = \frac{-1}{5}$$

$$\left| \begin{array}{ccc|c} -2 & 1 & 1 & -2 \\ 0 & \frac{5}{2} & \frac{1}{2} & -2 \\ 0 & 0 & \frac{11}{5} & 1 \end{array} \right| \quad x_3 = 1$$

$$x_2 = \left(-2 - \frac{1}{2} \right) \cdot \frac{2}{5} = \frac{-5}{2} \cdot \frac{2}{5} = -1$$

$$x_1 = \frac{\left(-2 - 1 + 1 \right)}{-2} = 1$$

$$\text{Quindi } x = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$\textcircled{2} \quad A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix} \quad \left| \begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ -2 & 1 & 1 & 3 \\ -1 & 0 & 1 & 3 \end{array} \right| \quad l_{21} = 2 \quad l_{31} = -1$$

$$\left| \begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & -1 & 5 & 1 \\ 0 & -1 & 3 & 7 \end{array} \right| \quad l_{32} = 1$$

$$\left| \begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & -1 & 5 & 1 \\ 0 & 0 & -2 & -4 \end{array} \right| \quad x_3 = 2$$

$$x_2 = \frac{(1 - 5 \cdot 2)}{-1} = -1 \quad \rightarrow x = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$$

$$x_1 = \frac{4 - 2 \cdot 2 - 1}{-2} = -1$$

Potenzialità del metodo: Supponiamo di dover risolvere K sistemi lineari con la stessa matrice dei coefficienti: $Ax^{(1)} = b^{(1)}, Ax^{(2)} = b^{(2)}, \dots, Ax^{(K)} = b^{(K)}$

Vorrebbe dire ripetere K volte il metodo → conviene prendere $AX = B$, dove abbiamo

$$X = (x^{(1)} | x^{(2)} | \dots | x^{(K)}) \quad e \quad B = (b^{(1)} | b^{(2)} | \dots | b^{(K)})$$

Caro particolare: $B = I$ (matrice identica)

In questo caso il sistema lineare diventa $AX = I$ e quindi la soluzione è $X = A^{-1}$

TECNICHE di PIVOTING

pivotting parziale: vado a calcolare $\max_{1 \leq i \leq n} |a_{ii}^{(1)}| = |a_{11}^{(1)}|$ e poi opero uno scambio tra l'r-esima equazione e la prima in modo da portarla in cima → i moltiplicatori avranno modulo compreso tra 0 e 1: riducono errori

ESEMPIO:

$$A = \begin{pmatrix} 1 -1+\varepsilon & 5 \\ 10 & 4 \\ 10003+\varepsilon & 7 \end{pmatrix}$$

$$l_{21} = 10 \quad l_{31} = 10^3$$

$$\tilde{A} = \begin{pmatrix} 10003+\varepsilon & 7 \\ 10 & 4 \\ 1 -1+\varepsilon & 5 \end{pmatrix}$$

$$\tilde{l}_{21} = 10^{-2}$$

$$\tilde{l}_{31} = 10^{-3}$$

Quando cambiano le righe gli errori vengono SMORTI, mentre nel caso precedente venivamo AMPLIFICATI

Proseguendo con la seconda eliminazione abbiamo $\max_{2 \leq i \leq n} |a_{ii}^{(2)}| = |a_{22}^{(2)}|$. Anche in questo caso si scambia la 1-esima riga con la seconda. Andiamo avanti in questo modo per tutti gli n-1 passi fino ad avere $R \times C$ dove R è triangolare superiore.

↓ In generale

Operare uno scambio tra due righe equivale a premoltiplicare per una matrice di permutazione. Se consideriamo k (k ≤ n-1) scambi avremo k matrici di permutazione ottenendo $P_k P_{k-1} \dots P_1 A = P_k P_{k-1} \dots P_1 b$ → Se avessimo avuto $\hat{A} = P_k P_{k-1} \dots P_1 A$ e $\hat{b} = P_k P_{k-1} \dots P_1 b$ e $\hat{A}x = \hat{b}$ non si sarebbero dovuti fare scambi. Abbiamo però che il prodotto di matrici di permutazione è una matrice di permutazione: $P = P_k P_{k-1} \dots P_1$ ed il sistema diventa $PAX = Pb$.

pivotting totale: $A^{(n)} = \begin{pmatrix} \text{---} \\ 0 \quad \text{|||||} \end{pmatrix}$

→ cerca il massimo modulo su tutta la parte verde della matrice (quindi di fatto su tutta la matrice)

↓

Anche in questo caso andiamo a scambiare le righe ma poi scambiamo anche le colonne per portare l'elemento individuato in posizione diagonale → **ATT:** si scambia anche l'ordine delle incognite e questi scambi vanno ANNOTATI

Sarebbe a contenere la propagazione degli errori

3. fattorizzazione LR: Nell'ipotesi che vogliano le condizioni sui minori principali di testa, l'algoritmo di eliminazione produce la fattorizzazione $A = LR$ dove R è la matrice triangolare superiore ottenuta alla fine dell'applicazione del metodo ed L ha la forma

$$L = \begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ l_{n1} & l_{n2} & \dots & \dots & l_{n,n-1} & 1 \end{pmatrix}$$

i termini sotto la diagonale principale coincidono con i moltiplicatori dell'algoritmo di eliminazione

DIMOSTRAZIONE:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

$$H_1 = \begin{pmatrix} 1 & & & \\ -l_{21} & 1 & & \\ -l_{31} & & 1 & \\ -l_{41} & & & 1 \end{pmatrix}$$

$$H_1 A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ 0 & a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \end{pmatrix}$$

→ equivalente ad aver fatto il primo passo del metodo di Gauss

$$H_2 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ -l_{32} & 1 & & \\ -l_{42} & & 1 & \end{pmatrix}$$

$$H_2 H_1 A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ 0 & 0 & a_{33}^{(2)} & a_{34}^{(2)} \\ 0 & 0 & a_{43}^{(2)} & a_{44}^{(2)} \end{pmatrix}$$

→ Secondo passo del metodo di Gauss

$$H_3 = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \\ & & l_{31} & 1 \end{pmatrix}$$

$$H_3 H_2 H_1 A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ 0 & 0 & a_{33}^{(1)} & a_{34}^{(1)} \\ 0 & 0 & 0 & a_{44}^{(1)} \end{pmatrix}$$

Matrice R: ottenuta con il metodo di Gauß

Ahhiamo quindi $H_3 H_2 H_1 A = R$ e ricavando $A: A = H_1^{-1} H_2^{-1} H_3^{-1} R$, ma chi sono queste matrici inverse? $H_1^{-1} = \begin{pmatrix} 1 & & \\ & l_{21} & 1 \\ & l_{31} & 1 \\ & l_{41} & 1 \end{pmatrix}$

$$\downarrow$$

Puoi quindi proseguire in questo modo anche per le successive. Ahhiamo poi che il prodotto di matrici triangolari imbrutti è ancora una matrice triangolare inferiore. Indichiamo quindi $H_1^{-1} H_2^{-1} H_3^{-1}$ con $L = \begin{pmatrix} 1 & & \\ & 1 & \\ & l_{31} & 1 \\ & l_{41} & l_{32} & 1 \end{pmatrix}$, ancora quindi triangolari. Ma quali sono gli elementi che compongono la matrice L?

$$H_1^{-1} H_2^{-1} = \begin{pmatrix} 1 & & & 1 \\ & 1 & & 1 \\ & l_{31} & 1 & 1 \\ & l_{41} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ 0 & l_{41} & l_{32} & 1 \end{pmatrix} \rightarrow \text{premultipliando per } H_1^{-1} \text{ ottieniamo } L = H_1^{-1} H_2^{-1} H_3^{-1} = \begin{pmatrix} 1 & & & \\ & l_{21} & 1 & 0 \\ & l_{31} & l_{32} & 1 \\ & l_{41} & l_{32} & l_{43} & 1 \end{pmatrix}$$

Ma come mai è un metodo diretto?

Consideriamo sistemi lineari con la stessa matrice di termini noti ma un termine moto diverso. Questo accade quando si presentano in tempi diversi.

$$\underbrace{Ax^{(1)}}_{A = LR} = b^{(1)} \quad Ax^{(2)} = b^{(2)} \rightarrow \begin{cases} Rx^{(1)} = q^{(1)} & : \text{Matrice triangolare superiore} \\ Lx^{(2)} = b^{(2)} & : \text{Matrice triangolare inferiore} \end{cases}$$

Non è possibile risolverli in contemporanea ma vanno risolti in serie. Va risolto PRIMA il secondo sistema e poi il primo

ESEMPIO

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 2 & 0 \\ 4 & 4 & 3 \end{pmatrix}$$

$$\begin{array}{r} 2 & 1 & 1 \\ 2 & 2 & 0 \\ \hline 4 & 4 & 3 \end{array} \quad l_{21} = 1 \quad l_{31} = 2 \\ l_{32} = 2$$

$$\begin{array}{r} 2 & 1 & 1 \\ 0 & 1 & -1 \\ \hline 0 & 2 & 1 \end{array} \quad R = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{pmatrix} \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix}$$

$$\text{Per trovare la fattorizzazione } L = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \quad R = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ 0 & x_{22} & x_{23} \\ 0 & 0 & x_{33} \end{pmatrix} \quad \text{faccio i prodotti e uguaglio} \quad R = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{metta utilizzando Gauß} \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix}$$

$$\begin{aligned} 1 + r_{12} &= 2 \rightarrow r_{12} = 1 & 2l_{21} &= 2 \rightarrow l_{21} = 1 \\ 1 + r_{13} &= 0 \rightarrow r_{13} = -1 & 2l_{31} &= 4 \rightarrow l_{31} = 2 \\ 2 - 2 + r_{23} &= 3 \rightarrow r_{23} = 3 & 2 + l_{32} &= 4 \rightarrow l_{32} = 2 \end{aligned}$$

È possibile dimostrare che utilizzando il pivoting parziale si avrà una fattorizzazione nella forma $PA = LpR$ con P = matrice neocentrica allo scambio delle righe, R ed Lp sono rispettivamente triangolari superiore/inferiore con elementi 1 nella diagonale

Dalle fattori $A = LR$ e $PA = LPP$ otteniamo $\det(A) = \det(R)$ e $\det(A) = (-1)^s \det(P)$ dove s è il numero di scambi di righe fatto nel pivoting, mentre i determinanti di R ed P sono dati dal prodotto degli elementi diagonali.

4. metodo di Gauß-Jordan: Consiste nel ridurre la matrice ad una matrice diagonale combinando le righe anche con le precedenti. I passi dell'algoritmo sono gli stessi dell'algoritmo di Gauß eccetto l'aggiunta di una combinazione per eliminare i termini sopra della diagonale \rightarrow risultato: $dx=b'$

Di nuovo, se i minori principali di testa sono non nulli, non sono richiesti scambi di righe.

$$\text{COSTO COMPUTAZIONALE: } \frac{n^3}{2} + n^2 - \frac{n}{2}$$

Trivare la matrice INVERSA: ricchezza di n sistemi lineari $Ax^{(i)} = e^{(i)}$ ove $x^{(i)}$ e $e^{(i)}, i=1,2,\dots,n$, sono l' i -esima colonna di x e di I \rightarrow Rivedendo il sistema $AX=I$ ci si riconduce ad $X=C$ ove $C=A^{-1}$

ESEMPIO

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 3 & 1 \\ 0 & 2 & 1 \end{pmatrix} \quad \begin{array}{c|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{c|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{c|ccc} 2 & 0 & 0 & 2 & -2 & 0 \\ 0 & 1 & \frac{1}{3} & -\frac{4}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} & -\frac{4}{3} & 1 \end{array}$$

$$\begin{array}{c|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & 2 & -4 & 3 \end{array} \longrightarrow A' = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 2 & -4 & 3 \end{pmatrix}$$

o MALCONDIZIONAMENTO

$$\begin{pmatrix} 1 & -1 \\ 1 & -1.000001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 \\ 1 & -1.000001 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.999999 \\ 1 \end{pmatrix} \longrightarrow \text{differenza per un solo termine noto}$$

① $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e ② $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -10^{-6} \\ -1 \end{pmatrix} \longrightarrow$ Aver cambiato una componente con un errore piccolo ha portato a variazioni di ordine molto grande nella soluzione.

MALCONDIZIONAMENTO: Perturbazioni di un certo ordine vengono amplificate ad ordini superiori. Possono essere introdotte sia similari che nello sviluppo dell'algoritmo.

Potremo avere una matrice $A + \delta A$ ove δA è la matrice delle perturbazioni. Abbiamo inoltre il vettore $b + \delta b$ ove δb è il vettore delle perturbazioni.

Vettore x perturbato a sua volta: $(A + \delta A)(x + \delta x) = b + \delta b \longrightarrow$ vorrò ritrovare una maggiorezza.

caso particolare senza perturbazioni sulla matrice A ($\delta A = 0$).
 $\hookrightarrow A(x + \delta x) = b + \delta b \longrightarrow Ax + A\delta x = b + \delta b \longrightarrow A\delta x = \delta b \longrightarrow \delta x = A^{-1}\delta b \longrightarrow \|\delta x\| = \|A^{-1}\delta b\|$

Sono uguali: $Ax = b$

$$\|Sx\| \leq \|A^{-1}\| \|Sb\|$$

per norma coerente

Inoltre, da $Ax = b \rightarrow \|Ax\| = \|b\| \rightarrow \|A\| \|x\| \geq \|b\| \rightarrow \|x\| \geq \frac{\|b\|}{\|A\|}$

↓ Unendo

$$\frac{\|Sx\|}{\|x\|} \leq \frac{\|A^{-1}\| \|Sb\|}{\|b\|} \cdot \frac{\|A\|}{\|b\|} \rightarrow \frac{\|Sx\|}{\|x\|} \leq \frac{\|A\| \|A^{-1}\| \|Sb\|}{\|b\|} \rightarrow \text{Maggiorazione errore relativo}$$

errore relativo con cui si introduce x errore relativo con cui si introduce b

→ NUMERO DI CONDIZIONAMENTO della matrice A : $\mu(A) = \|A\| \|A^{-1}\| \rightarrow$ Malcondizionamento $\propto \mu(A)$ ed il sistema si dice malcondizionato a partire dall'ordine di 10^3

↓

Il numero di condizionamento non può essere minore di 1

↪ Teorema di Hörnli: $\mu(A) = (\|A\| \|A^{-1}\|) \geq \|A A^{-1}\| = \|I\| = 1 \rightarrow \mu(A) \geq 1$

↓

IN GENERALE: $(A + SA)(x + Sx) = b + Sb$

↓ Si può dimostrare

$$\frac{\|Sx\|}{\|x\|} \leq \frac{\mu(A)}{1 - \frac{\|SA\|}{\|A\|}} \left(\frac{\|SA\|}{\|A\|} + \frac{\|Sb\|}{\|b\|} \right)$$

> 0

Il denominatore DEVE essere positivo perché la formula volga

$$1 - \frac{\|SA\|}{\|A\|} > 0$$

$$1 - \|SA\| \|A^{-1}\| > 0$$

$$\|SA\| \leq \frac{1}{\|A^{-1}\|}$$

↓

Valida solo se la posturazione non è troppo abbastanza

Maggiorazione A PRIORI

↪ Voglio introdurla a posteriori: \tilde{x} = soluzione di $Ax = b$ e $b - A\tilde{x} = r$, ove r è il vettore dei residui (o se \tilde{x} è la soluzione esatta)

↓ facendo la differenza con $b - Ax = 0$

$$A(\tilde{x} - x) = -r \text{ e quindi } \tilde{x} - x = -A^{-1}r \xrightarrow{\text{norme}} \|\tilde{x} - x\| \leq \|A^{-1}\| \|r\| \text{ e per } Ax = b \text{ si ottiene}$$

ERRORE ASSOLUTO

↓ soluzione proposta l'approssimata

$$\frac{\|\tilde{x} - x\|}{\|x\|} \leq \mu(A) \frac{\|r\|}{\|b\|} \rightarrow \text{se è piccola abbiamo un errore relativo piccolo sulla soluzione}$$

⚠ difficile da calcolare in quanto necessaria la matrice inversa

ESEMPIO

$A \in \mathbb{C}^{n \times n}$ hermitiana. Si vuole calcolare il numero di condizionamento in morma 2,

$$\text{ovvero } \mu_2(A)$$

$$\|A\|_2 = \sqrt{\rho(A^H A)} \xrightarrow{\text{matrice hermitiana}} \|A\|_2 = \rho(A)$$

$$\mu_2(A) = \|A\|_2 \|A^{-1}\|_2 = \rho(A) \rho(A^{-1}) = \max_{1 \leq i \leq n} |\lambda_i| \frac{1}{\min_{1 \leq i \leq n} |\lambda_i|} = \frac{\max_{1 \leq i \leq n} |\lambda_i|}{\min_{1 \leq i \leq n} |\lambda_i|}$$

$$\text{matrici di Hilbert} \rightarrow a_{ij} = \frac{1}{i+j-1} \quad i, j = 1, \dots, n$$

1	1/2	1/3	1.0000e+00	5.0000e-01	3.3333e-01
1/2	1/3	1/4	5.0000e-01	3.3333e-01	2.5000e-01
1/3	1/4	1/5	3.3333e-01	2.5000e-01	2.0000e-01

○ METODI ITERATIVI

Utili per matrici SPARSE (molti elementi diversi da 0). Applicando un metodo diretto si ottiene un sistema denso dunque a problemi di occupazione di memoria. Nei metodi iterativi si hanno invece Heredazioni che consentono un prodotti matrice per vettore $[O(n^2)]$

COME COSTRUIRE UNO SCHEMA ITERATIVO

metodo 1

$$Ax = b, \quad A \in \mathbb{C}^{n \times n}, \quad G \in \mathbb{C}^{n \times n}: \det(G) \neq 0$$

$$\xrightarrow{Ax - b = 0} G(Ax - b) = 0 \xrightarrow{\text{multiplico per } G^{-1}} GAx - Gb = 0 \xrightarrow{\text{sottraggo } GAx} X - GAx + Gb = X \xrightarrow{\text{sottraggo } X} X = GAx + Gb$$

$$\xrightarrow{\text{per } X = (I - GA)^{-1}X + Gb} X = Mx + C$$

L'incognita è data da un'espressione che la contiene → SCHEMA ITERATIVO del

$$\text{tipo } x^{(k)} = Mx^{(k-1)} + C, \quad k = 1, 2, 3, \dots$$

indici di iterazione

metodo 2

$$A = M - N$$

$$\det(M) \neq 0 \xrightarrow{\text{multiplico per } M^{-1}} (M - N)x = b \xrightarrow{\text{sottraggo } Nx} Mx = Xx + b \xrightarrow{\text{premultiplico per } M^{-1}} X = M^{-1}x + \frac{M^{-1}b}{C}$$

$$\xrightarrow{\text{per } X = Mx + C}$$

↓

$x^{(0)}$ si dice vettore iniziale → $\lim_{k \rightarrow \infty} x^{(k)} = a$ dove a sta in $Xa = b$ → in questo caso il sistema si dice convergente

Ahhiamo $X = Mx + C$ e $a = Ma + C$ con $a \in \mathbb{C}^n$

↓

$$x^{(k+1)} = Mx^{(k)} + C \xrightarrow{\text{sottraendo}} x^{(k+1)} - a = Mx^{(k)} - Ma \xrightarrow{\text{moltiplico per } M^{-1}} \underbrace{x^{(k+1)} - a}_{e^{(k+1)}} = M(x^{(k)} - a)$$

$$\xrightarrow{\text{per } e^{(k+1)} = Me^{(k)}} e^{(k+1)} = M(e^{(k)}) \xrightarrow{\text{moltiplico per } M^{-1}} e^{(k+1)} = M(H)e^{(k)} = H^2e^{(k)} = H^3e^{(k-1)} = \dots = H^{k+1}e^{(0)}$$

Vorrei avere che $\lim_{k \rightarrow \infty} x^{(k)} = a$ che equivale a verificare che, se $\lim_{k \rightarrow \infty} H^k = 0$, il limite è il VETTORE NUOLO

$$\lim_{k \rightarrow \infty} (x^{(k)} - a) = 0$$

da cui abbiamo

Se la matrice H è convergente la scelta del vettore iniziale non influenza la convergenza del metodo iterativo → $e^{(k+1)} = H^{k+1}e^{(0)} = \lambda^{k+1}e^{(0)}$ → λ deve essere di modulo < 1 → $p(H) < 1$, quindi la matrice è convergente

Ahhiamo quindi ottenuto che: H convergente $\iff x^{(k)} \rightarrow a$

↓

TEOREMA di CONVERGENZA: Condizione necessaria e sufficiente affinché un metodo iterativo della formula $x^{(k+1)} = Mx^{(k)} + C, k = 0, 1, \dots$ sia convergente per qualsiasi vettore iniziale $x^{(0)}$, è che la sua matrice di iterazione H sia convergente

corollario 1: Per la convergenza del metodo iterativo è necessario e sufficiente che risulti $p(H) < 1$

↓ Teorema di Mirch

corollario 2: Condizione sufficiente per la convergenza del metodo iterativo è l'esistenza di una norma matriciale per cui $\|H\| < 1$

corollario 3: Condizione necessaria per la convergenza del metodo iterativo è $|\det(H)| < 1$

ESEMPIO

$$H = \frac{1}{20} \begin{pmatrix} 2 & -1 & 4 \\ 3 & 1 & 7 \\ -10 & 2 & 3 \end{pmatrix}$$

voglio sapere se il metodo è convergente
↓ utilizzo la condiz. sufficiente con $\|H\|_\infty < 1$ in quanto tutti gli elementi sono divisi per 20 e la norma ∞ vale $\frac{3}{4}$

RISULTATO: Per una qualsiasi norma matriciale matrice si ha $\lim_{k \rightarrow \infty} \sqrt[k]{\|H^k\|} = p(H)$

$$e^{(k)} = H^k e^{(0)} \xrightarrow{\text{per } \|e^{(k)}\| = \|H^k e^{(0)}\|} \|e^{(k)}\| \leq \|H^k\| \|e^{(0)}\| \xrightarrow{\text{divido per } \|e^{(0)}\|} \frac{\|e^{(k)}\|}{\|e^{(0)}\|} \leq \|H^k\| \simeq 10^{-k}, \quad k \in \mathbb{N}$$

per il risultato precedente → $\frac{\|e^{(k)}\|}{\|e^{(0)}\|} \rightarrow 0$: sarebbe la sol. esatta

→ Dal risultato abbiamo che $\|H^k\| \approx \rho^k(H)$ per A → $\rho^k(H) = 10^{-m} \rightarrow \log_{10}^k(H) = -m$

prop logaritmi $K \log_{10} \rho(H) = -m \rightarrow m = -\frac{\log_{10} \rho(H)}{K} \in \mathbb{R}^+$ perché $\rho(H) < 1$ per convergenza
 V: velocità asintotica di convergenza

criterio di accerto: $x^{(k+1)} = Hx^{(k)} + C$ Sottraendo $x^{(k)}$ $x^{(k+1)} - x^{(k)} = Hx^{(k)} - x^{(k)} + C$

ed abbiamo che $a = \mu a + c$ in quanto verifica il sistema → $C = (I-H)a = -(H-I)a$

Sostituendo ottieniamo: $x^{(k+1)} - x^{(k)} = Hx^{(k)} - x^{(k)} - (H-I)a \rightarrow x^{(k+1)} - x^{(k)} = (H-I)x^{(k)} - (H-I)a$

$x^{(k+1)} - x^{(k)} = (H-I)(x^{(k)} - a)$ premettendo per $(H-I)^{-1}$ $x^{(k)} - a = (H-I)^{-1}(x^{(k+1)} - x^{(k)})$ ossiamo alla norma ed applichiamo la proprietà della norma costante

traslazione dello spettro: autovetori di H divisi di 1

$\|x^{(k)} - a\| \leq \| (H-I)^{-1} \| \|x^{(k+1)} - x^{(k)}\|$ ↳ $\neq 0$: autovetori di $H \neq 1$ per $\rho(H) < 1$ tolleranza ↲ mem varia al variare di K ↳ posso volutamente la norma della differenza: $\|x^{(k+1)} - x^{(k)}\|$ criterio di accerto più utilizzato

Se accade che $\rho(H)=0$ la matrice si dice nilpotente in quanto accadrà che $\exists K: \mu^K=0 \rightarrow$ il processo diventa finito in quanto avviene sulla soluzione esatta

↓

In generale, se dopo N passaggi il metodo non converge dopo ci fermiamo dopo $K > N$ iterazioni

o METODI ITERATIVI: metodi classici

metodo di Jacobi

$$A = D - E - F \rightarrow D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$$

↓ $-E, -F$: matrici triangolari inferiore/superiore con diagonale nulla

$$Ax = b \rightarrow x = Dx + C \quad \text{e} \quad g / \det(g) \neq 0$$

↪ $H = I - GA$ e $C = Gb$ → vedi il metodo 1 per la costruzione di un sistema iterativo

Decomponiamo A come $A = D - E - F$ e $G = D^{-1}$ perché possa avere D^{-1} aiuto $i=1 \dots n$ determinante di una matrice diagonale

Con le considerazioni fatte possiamo dire: $H = I - D^{-1}(D - E - F) = I - D^{-1}E + D^{-1}F = D^{-1}(E + F)$

↪ Abbiamo quindi ottenuto la matrice di Jacobi: $H_J = D^{-1}(E + F)$ e anche $C_J = D^{-1}b$

$$x^{(k+1)} = H_J x^{(k)} + C_J, \quad k=0, 1, 2, \dots \quad \text{ma come ricaviamo } H_J?$$

$$H_J = -\left(\begin{array}{cccccc} \frac{1}{a_{11}} & & & & & \\ \vdots & \ddots & & & & \\ & & \frac{1}{a_{22}} & & & \\ & & & \ddots & & \\ & & & & \frac{1}{a_{nn}} & \\ & & & & & \end{array}\right) \left(\begin{array}{c} 0 \\ a_{11} \dots a_{nn} \\ 0 \\ a_{22} \dots a_{nn} \\ \vdots \\ 0 \end{array}\right) = -\left(\begin{array}{cccccc} 0 & a_{12} & \dots & a_{1n} & & \\ a_{21} & 0 & \dots & a_{2n} & & \\ \vdots & & \ddots & & & \\ a_{n1} & a_{n2} & \dots & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{array}\right)$$

↗ $E+F$: rimane una matrice a diag = 0
 ↗ Giacomo riga che gli elementi di $E+F$ divisi per l'elemento diagonale di quella riga

Gli elementi di E ed F sono gli elementi di A cambiati di segno

metodo di Gauss-Seidel

Sistema $(D - E - F)x = b$ scritto nella forma $(D - E)x = Fx + b$ per evitare ripetere per $x^{(k+1)} = (D - E)^{-1}Fx^{(k)} + (D - E)^{-1}b$

Matrice TRIANGOLARE INFERIORE: $\det \neq 0$ se $a_{ii} \neq 0$

Abbiamo quindi che la matrice di iterazione è $H_{GS} = (D - E)^{-1}F$ che ricaviamo con un processo più complesso rispetto al metodo di Jacobi. La matrice $(D - E)^{-1}$ è triangolare inferiore come $(D - E)$. La matrice F è triangolare superiore con diagonale nulla. La prima colonna della matrice di iterazione sarà quindi nulla. Il suo determinante sarà quindi 0

differenze ed elementi comuni tra i metodi

Per entrambi i metodi si deve avere $a_{ii} \neq 0$ → Se non accade è necessario scambiare gli ordini delle equazioni, variamo di conseguenza le matrici di iterazione. Per entrambi i metodi possiamo dare delle condizioni sufficienti per la convergenza direttamente sulla matrice dei termini noti

TEOREMA: Se A è una matrice a predominanza diagonale forte, allora il metodo di Jacobi e quello di Gauss-Seidel sono convergenti

TEOREMA: Se A è irriducibile ed a predominanza diagonale debole, allora il metodo di Jacobi e di Gauß-Seidel sono convergenti.

Verifica del primo teorema per Jacobi

$$H_J = - \begin{pmatrix} 0 & a_{12}/a_{11} & a_{13}/a_{11} & \cdots & a_{1n}/a_{11} \\ a_{21}/a_{22} & 0 & a_{23}/a_{22} & \cdots & a_{2n}/a_{22} \\ a_{31}/a_{33} & a_{32}/a_{33} & 0 & \cdots & a_{3n}/a_{33} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1}/a_{nn} & a_{n2}/a_{nn} & a_{n3}/a_{nn} & \cdots & 0 \end{pmatrix}$$

matrice di iterazione di Jacobi

$$\|H_J\|_\infty : \text{utilizziamo la norma infinito}$$

$$1^{\text{a}} \text{ riga: } \sum_{i=1}^n |a_{1i}|$$

$$2^{\text{a}} \text{ riga: } \frac{|a_{11}|}{\sum_{i=1}^n |a_{1i}|}$$

In generale: $\|H_J\|_\infty = \max_{1 \leq i \leq n} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| / |a_{ii}|$ e la matrice è a predominanza diagonale forte

$$\text{quindi } \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < |a_{ii}| \quad \forall i = 1, \dots, n \quad \text{e quindi tutte le norme relative sono } < 1$$

e quindi il massimo è < 1 . Per il secondo corollario del teorema di convergenza otteniamo quindi la convergenza del metodo.

caso particolare: matrice tridiagonale \rightarrow elementi $\neq 0$ sulla diagonale principale e sulle codiagonali adiacenti (superiore ed inferiore)

$\rho(H_{GS}) = \rho^2(H_J) \rightarrow$ converge più velocemente Gauß-Seidel dato che ha raggio minore se il metodo converge: $V_{GS} = 2V_J$

versioni per componenti

metodo di Jacobi: $x^{(k+1)} = H_J x^{(k)} + c_J$ ove $H_J = D^{-1}(E+F)$ e $c_J = D^{-1}b$

\downarrow sostituisco

$$x^{(k+1)} = D^{-1}(E+F)x^{(k)} + D^{-1}b$$

$$D x^{(k+1)} = (E+F)x^{(k)} + b$$

$$\begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix} x^{(k+1)} = - \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} x^{(k)} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

prendendo adesso la i -esima componente:

$$a_{ii} x_i^{(k+1)} = b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)}$$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)} \right) \quad i = 1, 2, \dots, n$$

Poniamo trovare $x^{(k)} \rightarrow x^{(k+1)}$ e fare una sostituzione simultanea di tutte le componenti del vettore $x^{(k+1)}$ a quello del punto k

metodo di Gauß-Seidel: $x^{(k+1)} = H_{GS} x^{(k)} + c_{GS}$ ove $H_{GS} = (D-E)^{-1}F$ e $c_{GS} = (D-E)^{-1}b$

\downarrow sostituisco

$$x^{(k+1)} = (D-E)^{-1}F x^{(k)} + (D-E)^{-1}b \rightarrow (D-E)x^{(k+1)} = F x^{(k)} + b$$

$$\rightarrow D x^{(k+1)} - E x^{(k+1)} = F x^{(k)} + b \rightarrow D x^{(k+1)} = E x^{(k+1)} + F x^{(k)} + b$$

$$\begin{pmatrix} a_{11} & a_{12} & & & \\ a_{21} & a_{22} & a_{23} & & \\ & & \ddots & \ddots & \\ & & & a_{nn} & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} x^{(k+1)} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} - \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} x^{(k)} = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} x^{(k)}$$

prendiamo la i -esima componente: $a_{ii} x_i^{(k+1)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)}$

\downarrow dividendo per a_{ii}

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right)$$

Man mano che viene trovata una componente del vettore al punto $k+1$ essa viene utilizzata subito. Poco a poco si trovano anche subito le componenti del vettore al punto $k+1$ del vettore al punto K \rightarrow Metodo delle SOSTITUZIONI SUCCESSIVE

ESEMPI

1. $A = \begin{pmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{pmatrix}, \alpha \in \mathbb{C} \rightarrow$ Per quali valori di α converge il metodo di Jacobi?

⚠ Devo trovare la matrice di iterazione di Jacobi

$$H_3 = - \begin{pmatrix} 0 & \alpha & \alpha \\ \alpha & 0 & \alpha \\ \alpha & \alpha & 0 \end{pmatrix} \longrightarrow \text{CNS per convergenza: } \rho(H_3) < 1$$

$$\downarrow$$

$$H_3 - \alpha I = - \begin{pmatrix} \alpha & \alpha & \alpha \\ \alpha & \alpha & \alpha \\ \alpha & \alpha & \alpha \end{pmatrix} = -\alpha \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \longrightarrow 2 \text{ autovetori nulli ed uno} = 3 \text{ perché la somma dei } \lambda \text{ (} \leq \text{ elementi diagonali)} = 3$$

$$\sigma_1 = 0 = \sigma_2 ; \quad \sigma_3 = -3\alpha \longrightarrow \lambda_1 = \alpha = \lambda_2 ; \quad \lambda_3 = -2\alpha$$

$$\rho(H_3) = |\alpha| = 2|\alpha| \xrightarrow{\text{conv}} 2|\alpha| < 1 \longrightarrow |\alpha| < \frac{1}{2} : \text{ per questi numeri complessi il metodo converge}$$

Il risultato sarebbe lo stesso della predominanza di diagonali forte ma il metodo di inclusione sarebbe comunque sbagliato

2. $A = \begin{pmatrix} 2 & 0 & 0 & \alpha \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \quad \alpha \in \mathbb{C} \longrightarrow \text{per quale } \alpha \text{ Gauss-Seidel e Jacobi convergono?}$

$$H_3 = -\frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & \alpha \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \longrightarrow \text{Troviamo gli autovetori: Questa è una particolare matrice di Frobenius}$$

$$\lambda^4 - \alpha = 0 \longrightarrow \text{equazione caratteristica della matrice}$$

$$\lambda = \sqrt[4]{\alpha} \longrightarrow |\lambda| = \sqrt[4]{|\alpha|}$$

$$\text{moltiplicando per } -\frac{1}{2} \text{ con il modulo: } \rho(H_3) = \frac{1}{2} \sqrt[4]{|\alpha|} \longrightarrow \frac{1}{2} \sqrt[4]{|\alpha|} < 1$$

$$\longrightarrow \sqrt[4]{|\alpha|} < 2 \longrightarrow |\alpha| < 16$$

$$H_{GS} = (D - E)^{-1} F \longrightarrow D - E = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix} \quad F = \begin{pmatrix} 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

↓ calcolo inversa

$$\begin{array}{c|cccc} 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{c|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/4 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1/8 & -1/4 & 1/2 & 0 \\ 0 & 0 & 0 & 1 & -1/6 & 1/8 & -1/4 & 1/2 \end{array} \longrightarrow (D - E)^{-1} = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ -1/4 & 1/2 & 0 & 0 \\ 1/8 & -1/4 & 1/2 & 0 \\ -1/6 & 1/8 & -1/4 & 1/2 \end{pmatrix}$$

$$\begin{pmatrix} 1/2 & 0 & 0 & 0 \\ -1/4 & 1/2 & 0 & 0 \\ 1/8 & -1/4 & 1/2 & 0 \\ -1/6 & 1/8 & -1/4 & 1/2 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1/2 \alpha \\ 0 & 0 & 0 & 1/6 \alpha \\ 0 & 0 & 0 & -1/8 \alpha \\ 0 & 0 & 0 & 1/16 \alpha \end{pmatrix} \longrightarrow \text{Autovetori: 3 autovetori nulli e } \frac{1}{16} \alpha$$

⚠ ATT: è triangolare superiore, quindi gli autovetori sono gli elementi diagonali

$$\rho(H_{GS}) = \frac{1}{16} |\alpha| \longrightarrow \frac{1}{16} |\alpha| < 1 \longrightarrow |\alpha| < 16: \text{ stesso caso di prima}$$

$$\rho(H_{GS}) = \rho^4(H_3) \longrightarrow -\log_{10}(\rho(H_{GS})) = -4 \log(\rho(H_3)) \longrightarrow V_{GS} = 4V_3$$

3. $A = \begin{pmatrix} 1 & \alpha & 0 \\ 0 & 1 & \alpha \\ \alpha & 0 & 1 \end{pmatrix}$, $\alpha \in \mathbb{C}$ → determinare per quali α convergono Jacobi e Gauß-Seidel

$$H_0 = -\begin{pmatrix} 0 & \alpha & 0 \\ 0 & 0 & \alpha \\ \alpha & 0 & 0 \end{pmatrix} = -\alpha \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \rightarrow \text{Matrice di Frobenius}$$

$$\lambda^3 - \lambda = 0 \rightarrow \lambda = 1$$

$$|\lambda| = 1$$

$$\rho(H_0) = |\alpha| \rightarrow |\alpha| < 1$$

$$D - E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix} \rightarrow (D - E)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\alpha & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -\alpha & 0 \\ 0 & 0 & -\alpha \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\alpha & 0 \\ 0 & 0 & -\alpha \\ 0 & 0 & 0 \end{pmatrix} \rightarrow \lambda_1 = 0$$

$$\lambda^2 + \alpha^2 = 0 \rightarrow \lambda = \pm \sqrt{-\alpha^2}$$

$$\rho(H_{05}) = |\alpha|^{\frac{5}{2}} < 1 \rightarrow |\alpha| < 1$$

Equazioni non lineari

INTRODUZIONE

Sia $f(x) : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una funzione continua almeno su un intervallo I e si supponga che $f(x)$ sia nella forma $f(x) = a_1 x + a_0$ con a_1 e a_0 costanti.

Il problema che ci poniamo è quello di trovare $f(x) = 0$ supposto che esista.

Consideriamo un polinomio $a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0 = 0$, $a_m \neq 0$, con m intero $> 2 \rightarrow m$ radici nel campo complesso

utilizziamo un metodo iterativo: $x_{n+1} = \phi_n(x_n, x_{n-1}, \dots, x_{n-k+1})$, $n=0, 1, \dots, k-1$

funzione di iterazione del metodo

e dipende dagli argomenti e da $f(x)$

Le approssimazioni per $m=0$ devono avere note per fare i k passi iterativi ed il metodo deve accettarsi per un opportuno criterio di arresto

metodo iterativo a k punti \rightarrow se ϕ_n non varia al variare di m il metodo si dice **stazionario**. Si dice **non stazionario** altrimenti

Data una successione $\{x_n\}$ convergente ad un limite α e posto $e_n = x_n - \alpha$, vado a calcolare $\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|^p} \stackrel{0}{\longrightarrow} c$ (finito e $\neq 0$) allora $p \rightarrow$ ordine di convergenza $c \rightarrow$ fattore di convergenza

gli infinitesimi sono dello stesso ordine. p è la potenza che diamo ad e_n in modo che gli infinitesimi siano confrontabili

Se $p=1$ la convergenza si dice lineare mentre se $p=2$ la convergenza si dice quadrattica \rightarrow Se $p=1$, c deve essere sicuramente < 1 (ma comunque > 0)

Vogliamo trovare gli intervalli a cui appartengono gli zeri (di separazione). Ad un intervallo di separazione appartiene un solo 0. Una strategia per trovarli può essere lo studio del grafico (separazione grafica)

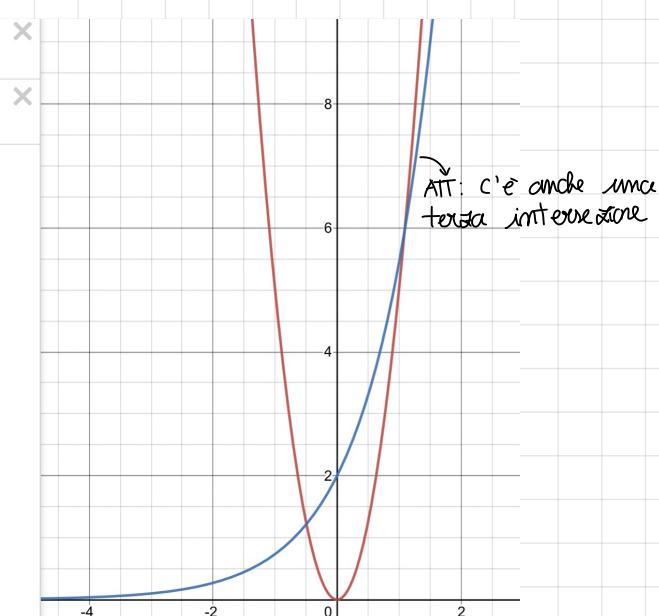
$f(x) = g(x) - h(x)$ dove $g(x)$ e $h(x)$ sono arbitrarie
 $\cancel{f(x)=0} \rightarrow g(x) = h(x)$ introduco y $\left\{ \begin{array}{l} y = g(x) \\ y = h(x) \end{array} \right. \rightarrow$ da x del sistema è soluzione dell'equazione $g(x) = h(x)$

Utilizziamo il sistema per tracciare i grafici delle due funzioni, le cui intersezioni, ed in particolare le loro arcate, costituiscono la soluzione del problema

ESEMPIO

$$5x^2 - 2e^x = 0 \rightarrow g(x) = 5x^2$$

$$h(x) = 2e^x$$



⚠ ATT: non sempre la separazione grafica fornisce risultati apprezzabili

METODO di BISEZIONE

Ad ogni passo costruisce un intervallo contenente uno z della funzione ed approssima l'ascina dello z come il punto medio di tale intervallo

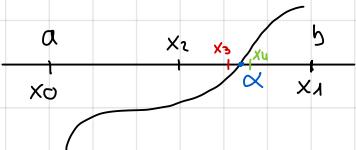
condizioni: 1. $f \in C^0([a, b])$

2. $f(a)f(b) = f(x_0)f(x_1) < 0$ → Almeno una soluzione ha molti punti dispari
↓ Molti punti soluzioni

$f(x)=0 \rightarrow f'(x)=0, f''(x)=0, \dots, f^{(n)}(x)=0, f^{(n+1)}(x) \neq 0$: Molti punti $n+1$
oppure $f(x)=(x-\alpha)^{n+1}g(x)$ / $\lim g(x)=c \neq 0$, quindi $g(x)$ non si annulla e
tal che si annulla il prodotto

$$f(x)=0 \quad x \in [a, b]$$

$$x_0=a, x_1=b \rightarrow x_2=\frac{a+b}{2}=\frac{x_0+x_1}{2}$$



$f(x_2)f(x_0) > 0$: non cambia segno quindi la funzione sta nell'altro intervallo: (x_1, x_2)

Proseguendo: $x_3=\frac{x_2+x_1}{2}$ e voluto $f(x_3)f(x_0) > 0 \rightarrow x \in (x_3, x_1)$

$$\downarrow$$

 $x_4=\frac{x_3+x_1}{2}$ e voluto $f(x_4)f(x_0) < 0 \rightarrow x \in (x_3, x_4)$

Ahiamo quindi una sola soluzione di funzione per ogni passo → modo in cui si calcola il costo computazionale di un metodo iterativo

$$\text{In generale: } x_3=\frac{x_2+\hat{x}_2}{2}, \hat{x}_2=\begin{cases} x_1 & \text{se } f(x_2)f(x_0) > 0 \text{ e quindi} \\ x_0 & \text{altrimenti} \end{cases} \quad x_n=\frac{x_{n-1}+\hat{x}_{n-1}}{2}, \hat{x}_n=\begin{cases} x_n & \text{se } f(x_n)f(x_{n-1}) < 0 \\ x_{n-1} & \text{altrimenti} \end{cases}$$

Supponiamo di voler trovare $\hat{x}: |\hat{x}-x| < \varepsilon$ → Poco importa quante iterazioni fare per precisione desiderata raggiungere la precisione desiderata

Al passo 0 l'ampiezza dell'intervallo è $b-a$. Al primo passo l'ampiezza diventa $\frac{b-a}{2}$. Svolgendo il secondo passo l'ampiezza diventa $\frac{b-a}{2^2}$. In generale quindi avremo,

dopo n passi, un'ampiezza di $\frac{b-a}{2^n}$ per il numero di iterazioni $\frac{b-a}{2^n} < \varepsilon$ da cui trovo n svolgendo:

$$\frac{b-a}{\varepsilon} < 2^n \rightarrow \log_2 \frac{b-a}{\varepsilon} < n \rightarrow \text{Basta prendere il primo intero positivo } n > \text{di } \log_2 \frac{b-a}{\varepsilon} \rightarrow n = \lceil \log_2 \frac{b-a}{\varepsilon} \rceil$$

Una volta individuato l'intervallo corretto, \hat{x} è un qualiasi punto dell'intervallo $b-a$ in quanto sicuramente distanzi da $x < \varepsilon$

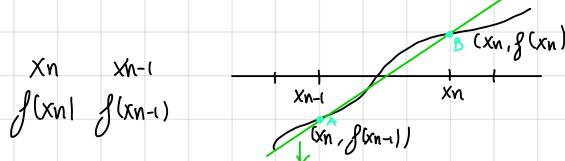
Il metodo converge linearmente. Assumendo $|x_n - x_{n-1}|$ come stima di $|x_n - x| = |e_n|$ si fa, per un abbastanza grande: $\frac{|e_n|}{|e_{n-1}|} \approx \frac{|x_n - x_{n-1}|}{|x_n - x_{n-1}|} = \frac{1}{2}$ → Per ottenere una convergenza più veloce utilizziamo quindi altri metodi

METODO delle SECANTI

Richiede due valori per partire con l'applicazione → $x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}$, $n=1, 2$

Metodo a due punti stazionario: mi servono due stime precedenti e le inserisco semplicemente nella formula

Come mai si chiama metodo delle secanti?



Retta secante al grafico passante per A e B: $\frac{x-x_n}{x_{n-1}-x_n} = \frac{y-f(x_n)}{f(x_{n-1})-f(x_n)}$
 $y=0 \rightarrow$ per trovare intersezione con l'asse delle ascisse

Risolvendo il sistema: $x = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}$ → Il metodo calcola la retta secante ad ogni iterazione e trova la sua intersezione con l'asse delle ascisse (x_{n+1})
 Reciproco del rapporto incrementale: Il metodo non richiede la derivabilità della funzione

Il metodo richiede solo la valutazione di $f(x_n)$ ad ogni passo in quanto $f(x_{n-1})$ è già nota al passo precedente. L'ordine di convergenza risulta essere $p = \frac{1+\sqrt{5}}{2}$

• METODI STAZIONARI AD UN PUNTO

Data l'equazione $f(x)=0$ si può ricavare una funzione $\phi(x)$ tale che l'equazione data sia equivalente ad un'equazione nella forma $x = \phi(x)$

Supponiamo di avere $f(x)=0$ e di cercare le soluzioni in $[a,b]$. Prendi una funzione $g(x) \neq 0 \quad \forall x \in [a,b]$. → Posso quindi moltiplicare $f(x)$ per $g(x)$ senza alterare le soluzioni. Sommo x al primo ed al secondo membro → $x + g(x)f(x) = x \rightarrow x = x - \underbrace{g(x)f(x)}_{\phi(x)}$. Se sostituisco

$f(x)=0$ ottengo $x = \phi(x)$ dove x è un punto fisso per $\phi(x)$. Sarà quindi necessario ricercare i punti fissi di ϕ (e non citeremo più la funzione f).

Possiamo utilizzare un metodo iterativo stazionario ad un punto nella forma $x_{n+1} = \phi(x_n)$, $n=0,1,2,\dots$ → Vogliamo trovare una successione convergente ad x .

TEOREMA DI CONVERGENZA LOCALE

$x = \phi(x)$, $\alpha = \phi(\alpha)$, $\alpha \in I$, $\phi \in C^1(I)$, $x_{n+1} = \phi(x_n)$ con $n=0,1,2,\dots$. Se $\exists p, K \in \mathbb{R}^+$, $K < 1$ tali che $|\phi'(x)| \leq K$, $x \in J_{\alpha-p, \alpha+p} \subset I$, allora per il metodo $x_{n+1} = \phi(x_n)$ vogliono le seguenti proposizioni:

1. Se $x_0 \in J_p$, allora $x_n \in J_p \quad \forall n \in \mathbb{N}$
2. Se $x_0 \in J_p$, allora $\lim_{n \rightarrow \infty} x_n = \alpha$
3. α è l'unico punto fisso in J_p

DIMOSTRAZIONE

proposizione 1

$x_0 \in J_p$, dimostrazione PER INDUZIONE → Supponiamo che $x_n \in J_p$, devo dimostrare che $x_{n+1} \in J_p$.

Dato che $x_n \in J_p$ accade che $|x_n - \alpha| < p$

↓ voglio dimostrare

$$|x_{n+1} - \alpha| < p \quad \text{TESI INDUTTIVA}$$

Calcoliamo $x_{n+1} - \alpha$: $\alpha = \phi(x_n) \rightarrow x_{n+1} - \alpha = \phi(x_n) - \phi(\alpha)$

↓ Teorema di Lagrange: ϕ derivabile in J_p

$$\phi(x_n) - \phi(\alpha) = \phi'(\xi_n)(x_n - \alpha)$$

Appartenere all'intervallo di estremi x_n, α , quindi in J_p
 ponendo ai valori anzidetti: $|x_{n+1} - \alpha| = |\phi'(\xi_n)| |x_n - \alpha| \leq K |x_n - \alpha| < Kp < p$

↓ ho quindi ottenuto

$$|x_{n+1} - \alpha| < p, \text{ quindi } x_{n+1} \in J_p$$

proposizione 2

$$\lim_{n \rightarrow \infty} x_n = \alpha \rightarrow \lim_{n \rightarrow \infty} (x_n - \alpha) = 0$$

→ DIMOSTRAZIONE PRECEDENTE

$$0 \leq |x_{n+1} - \alpha| \leq K |x_n - \alpha| \leq K^2 |x_{n-1} - \alpha| \leq \dots \leq K^{n+1} |x_0 - \alpha| \rightarrow 0$$

↓

O: numero positivo < 1

Teorema dei corahimieri: $\lim_{n \rightarrow \infty} |x_{n+1} - \alpha| = 0$

proposizione 3

α è l'unico punto fisso in I_p . Supponiamo per assurdo che $\exists \beta \in I_p : \phi(\beta) = \beta, \alpha \neq \beta$

\downarrow Appartiene all'intervallo di estremi α e β

$$|\alpha - \beta| = |\phi''(\xi)| |\alpha - \beta| \leq K |\alpha - \beta| < |\alpha - \beta| \rightarrow \text{ASSURDO: ottengo } |\alpha - \beta| < |\alpha - \beta|$$

IN GENERALE: L'ordine di convergenza di un metodo iterativo è un numero reale $p > 1$

TEOREMA sull'ORDINE di CONVERGENZA: Un metodo iterativo stabilisce ad un punto, la cui funzione di iterazione $\phi(x)$ sia sufficientemente derivabile, ha ordine di convergenza uguale ad un numero intero positivo p . Se il metodo converge ad α , la convergenza è di ordine p allora e solo che si abbia $\phi'(\alpha) = 0, \phi^{(i)}(\alpha) = 0$ per $1 \leq i < p, \phi^{(p)}(\alpha) \neq 0$

DIMOSTRAZIONE

$$x_{n+1} - \alpha = \phi(x_n) - \phi(\alpha) \quad \text{vedi dimostrazione precedente}$$

$$\text{Per Taylor: } \phi(x_n) = \phi(\alpha) + \phi'(\alpha)(x_n - \alpha) + \frac{\phi''(\alpha)(x_n - \alpha)^2}{2!} + \dots + \frac{\phi^{(p-1)}(\alpha)(x_n - \alpha)^{p-1}}{(p-1)!} + \frac{\phi^{(p)}(\alpha)(x_n - \alpha)^p}{p!}$$

$$\phi(x_n) - \phi(\alpha) = \phi'(\alpha)(x_n - \alpha) + \frac{\phi''(\alpha)(x_n - \alpha)^2}{2!} + \dots + \frac{\phi^{(p-1)}(\alpha)(x_n - \alpha)^{p-1}}{(p-1)!} + \frac{\phi^{(p)}(\alpha)(x_n - \alpha)^p}{p!}$$

Porto al primo membro $\phi(\alpha)$

\downarrow Da cui

$$x_{n+1} - \alpha = \frac{\phi^{(p)}(\alpha)(x_n - \alpha)^p}{p!} \quad \text{uguaglianza iniziale}$$

$$x_{n+1} - \alpha = \frac{\phi^{(p)}(\alpha)(x_n - \alpha)^p}{(x_n - \alpha)^p} = \frac{\phi^{(p)}(\alpha)}{p!}$$

\downarrow se $\phi^{(p)}(\alpha)$ continua

$$\text{Aggiungendo i valori assoluti: } \lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = \lim_{n \rightarrow \infty} \frac{|\phi^{(p)}(\alpha)|}{p!} \neq 0$$

Abbiamo verificato che l'ordine di convergenza è p

\downarrow

Verifichiamo l'implicazione inversa: Se p è l'ordine di convergenza, allora abbiamo che $\phi'(\alpha) = \phi''(\alpha) = \dots = \phi^{(p-1)}(\alpha) = 0$ e $\phi^{(p)}(\alpha) \neq 0$

$$\phi(x_n) - \phi(\alpha) = \phi'(\alpha)(x_n - \alpha) + \frac{\phi''(\alpha)(x_n - \alpha)^2}{2!} + \dots + \frac{\phi^{(p-1)}(\alpha)(x_n - \alpha)^{p-1}}{(p-1)!} + \frac{\phi^{(p)}(\alpha)(x_n - \alpha)^p}{p!}$$

Se avessi $\phi^{(i)}(\alpha) \neq 0, 1 \leq i < p$, l'ordine di convergenza sarebbe i per la dimostrazione diretta. se invece avessi $\phi^{(p)}(\alpha) = 0$, l'ordine di convergenza sarebbe maggiore di p .

criterio di arresto: $|x_{n+1} - x_n| \leq \varepsilon$

o METODO di NEWTON

Applicabile per trovare uno zero α di $f(x)$ re, in tutto un intorno di α , $f(x)$ è derivabile con continuità

\downarrow

$f(x) = 0 \rightarrow x = \phi(x)$, con $\phi(x) = x - g(x)/f'(x)$, ove $g(x) = \frac{f(x)}{f'(x)}$. Abbiamo quindi ottenuto

lo schema iterativo $\phi(x) = x - \frac{f(x)}{f'(x)}$ che prevede

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad \text{SCHEMA ITERATIVO}$$

interpretazione grafica:



tangente: $\begin{cases} y - f(x_n) = f'(x_n)(x - x_n) \\ y = 0 \end{cases}$

$$-f(x_n) = f'(x_n)(x - x_n) \rightarrow f'(x_n)x = f'(x_n)x_n - f(x_n),$$

da cui ottengo $x = x_n - \frac{f(x_n)}{f'(x_n)}$

METODO delle TANGENTI

ordine di convergenza: $\phi(x) = x - \frac{f(x)}{f'(x)}$ $\alpha = \phi(\alpha) \iff f(\alpha) = 0$

caso 1: α soluzione semplice (moltiplicità 1) $\rightarrow f(\alpha) = 0, f'(\alpha) \neq 0$

$$\phi'(x) = 1 - \frac{[f''(x)]^2 - f(x)f'''(x)}{[f'(x)]^2} = \frac{f(x)f'''(x)}{[f'(x)]^2}$$

$$\phi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2} = 0 \rightarrow \text{minimo ordine di convergenza} = 2$$

$$\phi''(x) = \frac{(f''(x)f''(x) + f(x)f'''(x))(f'(x))^2 - f(x)f''(x)2f'(x)f''(x)}{[f'(x)]^4}$$

$$\phi''(x) = \frac{[f'(x)]^3 f''(x)}{[f'(x)]^4} = \frac{f''(x)}{f'(x)} \rightarrow = 0 \text{ solo se la funzione ha un punto} \\ \rightarrow \text{d'ordine reale} > 2$$

costo computazionale: 2 calcoli di funzione per ogni passaggio

teorema di convergenza locale: Vole perché se $\phi'(\alpha) = 0$, essendo ϕ' continua, non avrebbe subito valori > 0 in un intorno di α

Per radici semplici ammette valori iniziali che lo rendono convergente

caso 2: α con moltiplicità $s > 1$

$$f(x) = g(x)(x-\alpha)^s, \text{ con } g(x) = \frac{f(x)}{(x-\alpha)^s} \quad e \quad g(x) = \lim_{x \rightarrow \alpha} g(x) \neq 0$$

$$f'(x) = g'(x)(x-\alpha)^s + sg(x)(x-\alpha)^{s-1} = (x-\alpha)^{s-1}(g'(x)(x-\alpha) + sg(x))$$

$$\phi(x) = x - \frac{f(x)}{f'(x)} = x - \frac{g(x)(x-\alpha)^s}{(x-\alpha)^{s-1}(g'(x)(x-\alpha) + sg(x))} = x - \frac{g(x)(x-\alpha)}{g'(x)(x-\alpha) + sg(x)}$$

$$\phi'(x) = 1 - \frac{(g'(x)(x-\alpha) + g(x))(g'(x)(x-\alpha) + sg(x)) - g(x)(x-\alpha)[\dots]}{[g'(x)(x-\alpha) + sg(x)]^2}$$

$$\phi'(x) = 1 - \frac{g(x)sg(x)}{s^2g^2(x)} = \frac{1 - \frac{1}{s}}{C} \rightarrow \text{è convergente in quanto possiamo sicuramente trovare valori iniziali per cui risulta convergente per il teorema di conv. locale}$$

$$s > 1 \rightarrow \phi'(x) \neq 0 \text{ e quindi } p=1$$

Modificando il metodo è possibile ottenere di nuovo la convergenza quadratica

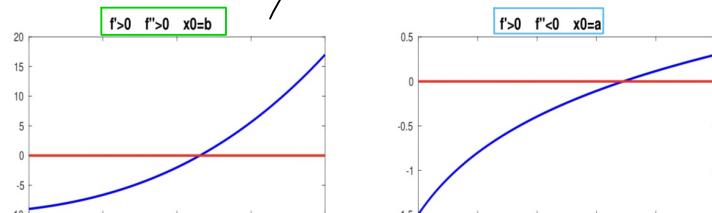
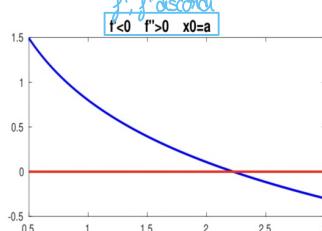
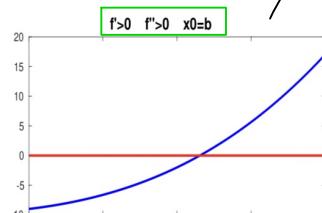
$$x_{n+1} = x_n - \frac{s f(x_n)}{f'(x_n)}, \quad n=0,1,\dots, s>1$$

condizioni sufficienti convergenza

Dato l'equazione $f(x)=0$, sia α una soluzione di moltiplicità d all'interno di un intervallo di convergenza per x e sia inoltre $f \in C^2([a,b])$. Se $f'(x)$ e $f''(x)$ sono di segno costante in $[a,b]$, allora il metodo di Newton converge scegliendo come punto iniziale un valore $x_0 \in [a,b]$ tale che $f(x_0), f''(x_0) > 0$

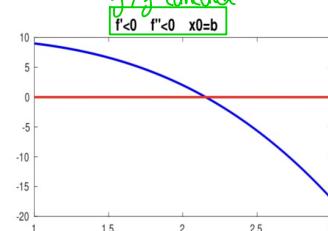
\downarrow costanza di segno

- a. $f'(x) > 0, f''(x) > 0$
- b. $f'(x) > 0, f''(x) < 0$
- c. $f'(x) < 0, f''(x) > 0$
- d. $f'(x) < 0, f''(x) < 0$



\rightarrow Genera una successione monotona decrescente con $\lim_{n \rightarrow \infty}$

\rightarrow Genera una successione monotona crescente sup. limitata da α e quindi convergente ad α



efficienza del metodo

$\alpha = \frac{p}{q}$ → Ordine di convergenza del metodo
 $\downarrow \sqrt{ }$ → Numero di valutazioni di funzione per ciascuna iterazione

bisezione: $\alpha = \frac{l}{1} = 1$

Secante: $\alpha = \frac{\lambda + \sqrt{\lambda}}{2} = \frac{1 + \sqrt{5}}{2}$

Newton per soluzioni semplici: $\alpha = \frac{2}{2} = 1$

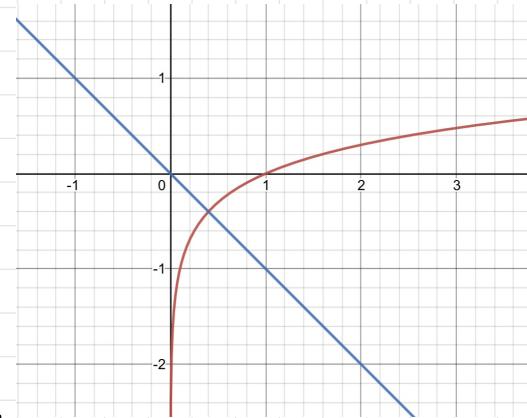
Newton per soluzioni con molteplicità > 1 : $\alpha = \frac{1}{2} = \frac{1}{2}$

ESEMPI

1. Separare le radici di $x + \log x = 0$ e studiare la convergenza dei metodi iterativi

$x_{n+1} = -\log(x_n)$ e del metodo di Newton

Punto 1: separazione grafica $\rightarrow g(x) = \log(x)$, $h(x) = -x$
 $x \in [0.5, 0.6]$



Punto 2: $x_{n+1} = -\log(x_n)$ $f(x) = -\log(x)$ $f'(x) = -\frac{1}{x}$

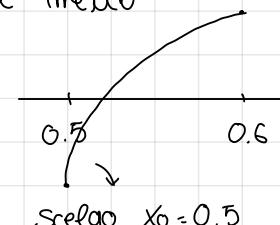
$|f'(x)| = \frac{1}{x}$ → Non accade mai che $|f'(x)| \leq k < 1$ ma è invece > 1 su tutto l'intervallo → Il metodo NON verifica le condizioni del teorema di convergenza locale e quindi conviene cambiare metodo

Punto 3: Metodo di Newton

$f(x) = x + \log(x)$ $f'(x) = 1 + \underbrace{\frac{1}{x}}$

> 0 sul nostro intervallo

$f''(x) = \underbrace{-\frac{1}{x^2}}_{< 0}$



Voglio anche trovare $\hat{x}: |\hat{x} - \alpha| < \epsilon$

$x_0 \xrightarrow{\text{ho trovato}} x_1$
 $\hat{x}_1 = x_1 + \epsilon$
 $\hat{x}_1 > x_1$

Può essere a destra di α e quindi $f(\hat{x}_1) < 0$ ①

Ho trovato un intervallo di ampiezza ϵ al quale appartiene la soluzione

Se invece accade $\hat{x}_1 < \alpha$ applicando il metodo di Newton partendo da \hat{x}_1 trovando x_2 e ripetendo il procedimento fino a trovare il zero ②

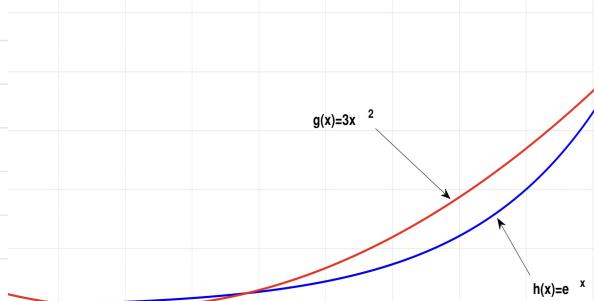
2. Separare le radici dell'equazione $3x^2 e^{-x} = 1$. Studiare la convergenza di $x_{n+1} = \frac{e^{x_n}}{3x_n}$, $x_{n+1} = \log(3x_n^2)$, metodo di Newton

Separazione grafica $\rightarrow 3x^2 = e^x$

$\alpha_1 \in [-1/2, -1/3]$,

$\alpha_2 \in [2/3, 1]$,

$\alpha_3 \in [7/2, 4]$



$x_{n+1} = \frac{e^{x_n}}{3x_n}$ ottenuta dividendo l'equazione per $3x$

$$\phi(x) = \frac{e^x}{3x} \quad \phi'(x) = \frac{3xe^x - 3e^x}{9x^2} = \frac{e^x(x-1)}{3x^2}$$

$$\phi''(x) = e^x \frac{x^2 - 2x + 2}{3x^3} \rightarrow \text{da: } \phi'' < 0 \text{ quindi } \phi' \text{ decrescente} \rightarrow \underline{\phi'(-\frac{1}{2}) = -1.2 \dots}$$

$|\phi'| > 1$, quindi non
assicura la convergenza

$$\rightarrow \alpha_2: \phi'' > 0 \text{ quindi } \phi' \text{ crescente} \rightarrow \underline{\phi'(\frac{2}{3}) > 0.69}, \phi'(1) = 0$$

$-1 < \phi' < 0$ quindi converge con
 x_0 opposto

$$\rightarrow \alpha_3: \phi'' > 0 \text{ quindi } \phi' \text{ crescente} \rightarrow \underline{\phi'(\frac{7}{2}) = 2.2 \dots > 1} \text{ quindi non è assicurata la convergenza}$$

$x_{n+1} = \log(3x_n)$ ottenuta facendo il logaritmo della funzione

$$\phi'(x) = \frac{6x}{3x^2} = \frac{2}{x} \rightarrow \text{Nel primo intervallo diverge ed anche nel secondo.}$$

Nel terzo invece converge

Metodo di Newton

$$f(x) = 3x^2 - e^x$$

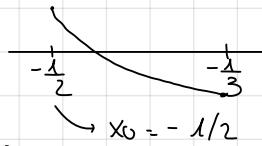
$$f'(x) = 6x - e^x$$

$$f''(x) = 6 - e^x$$

$$3. \phi(x) = \frac{x^2 - 5x + 6}{x^2 + x}$$

$$x = \phi(x)$$

Per $x_0: f' < 0, f'' > 0$



$$\alpha = \phi(\alpha)$$

$x_{n+1} = \phi(x_n) \rightarrow$ Se lo schema converge, converge ad un punto fisso

$$x = \frac{x^2 - 5x + 6}{x^2 + x} \rightarrow D: \mathbb{R} - \{-1, 0\} \rightarrow x^3 + x^2 = \cancel{x^2} - 5x + 6 \rightarrow x^3 + 5x - 6 = 0$$

Le soluzioni sono da trovarsi tra le soluzioni di 6 ed $1 \rightarrow \pm \frac{p}{q}$: termine noto
 $\frac{p}{q}$: termine di grado max

Soluzione 1: $x=1$

Metodo di Ruffini:

	1	0	5	-6
1		1	1	6
	1	1	6	/

$$(x-1)(\cancel{x^2+x+6}) = 0$$

$a > 0, \Delta > 0 \rightarrow$ non ha soluzioni reali

Unico punto fisso: $\alpha=1$

• METODI ITERATIVI IN \mathbb{R}^n

Ahhiamo $f(x): \mathbb{R}^n \rightarrow \mathbb{R}^n$ cioè un sistema lineare del tipo:

$$f_1(x_1, \dots, x_n) = 0$$

$$f_2(x_1, \dots, x_n) = 0$$

$$f_n(x_1, \dots, x_n) = 0$$

Indichiamo con $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ la soluzione. Per trovarla dobbiamo ricorrere a metodi iterativi $f(x)=0$, dobbiamo prendere una matrice $G(x) \in \mathbb{R}^{n \times n}$ tale che $\det(G(x)) \neq 0$ altrimenti non ci sarebbe un insieme di indeterminate $D = \{x \in \mathbb{R}^n: a_i < x_i < b_i, i=1, 2, \dots, n\}$

Premoltiplicando il sistema per $G(x)$: $G(x)f(x) = 0$ e le soluzioni non cambiano

Aggiungendo x : $x + G(x)f(x) = x \rightarrow \underline{\text{includendo la } x \text{ al primo membro}} \rightarrow x = x - G(x)f(x)$

Potiamo quindi costruire lo schema iterativo $x^{(n+1)} = \phi(x^{(n)})$ Sistema non lineare del tipo $x = \phi(x)$

Per utilizzare il teorema di convergenza locale:

$$\phi(x) = \begin{pmatrix} \frac{\partial \phi_1}{\partial x_1} & \dots & \frac{\partial \phi_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \phi_n}{\partial x_1} & \dots & \frac{\partial \phi_n}{\partial x_n} \end{pmatrix}$$

Matrice Jacobiana di ϕ

Se α è un punto fijo di $\phi(x)$, condizione sufficiente per la convergenza ad α del metodo iterativo $x^{(n+1)} = \phi(x^{(n)})$ è che esistano due numeri positivi K e p , $K < 1$, tali che si abbia $\|\phi(x)\| \leq K$, $\forall x \in D_p = \{x : \|x - \alpha\| \leq p\}$ perché $x^{(n)}$ sia xotto in D_p . In tal caso α è l'unico punto fijo di ϕ in D_p

metodo di Newton: si derivaibili con continuità e $J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$ non singolare

in un dominio contenente nel suo interno una soluzione α del sistema. ($f'(\alpha) = 0$ se la molteplicità è > 1 , quindi non posso nemmeno dividere)

$\phi(x) = x - J^{-1}(x)f(x)$ e quindi $x = x - J^{-1}(x)f(x)$ equivalente al sistema $f(x) = 0$ da cui $x^{(n+1)} = x^{(n)} - J^{-1}(x^{(n)})f(x^{(n)})$, $n=0, 1, \dots \rightarrow$ Metodo di Newton-Raphson

Ho m^2 funzioni da calcolare per trovare la Jacobiana e n funzioni da calcolare per trovare f . Ho poi il problema del calcolo dell'inversa. Possiamo solo evitare di fare quest'altra cosa $\rightarrow x^{(n+1)} = x^{(n)} - J^{-1}(x^{(n)})f(x^{(n)}) \rightarrow J(x^{(n)})x^{(n+1)} = J(x^{(n)})x^{(n)} - f(x^{(n)}) \rightarrow J(x^{(n)})(x^{(n+1)} - x^{(n)}) = -f(x^{(n)}) \rightarrow J(x^{(n)})d^{(n)} = -f(x^{(n)})$: Sistema lineare con matrice di coefficienti $J(x^{(n)})$

Lo schema iterativo diventa quindi: $x^{(n+1)} = x^{(n)} + d^{(n)}$ → Successione di sistemi lineari

↪ Metodo di Newton semplificato: utilizzabile quando $x^{(0)}$ è una buona approssimazione del vettore soluzione

Ho più sistemi lineari con la stessa matrice di coefficienti. L'ordine di convergenza di questo metodo è lineare

È possibile utilizzare anche altri metodi → Si considera l' i -esima equazione del sistema $f(x) = 0$ come un'equazione nella sola incognita x_i e si applica il metodo di Newton per equazioni in una sola incognita

↪ $\frac{\partial f_i(x_1, \dots, x_n)}{\partial x_i} \neq 0$, $i=1, 2, \dots, n \rightarrow x_i^{(n+1)} = x_i^{(n)} - \frac{f_i(x_1^{(n)}, \dots, x_n^{(n)})}{\frac{\partial f_i(x_1^{(n)}, \dots, x_n^{(n)})}{\partial x_i}}$ $i=1, \dots, n$, $n=0, 1, \dots$

Metodo non lineare di Jacobi-Newton

Si può utilizzare anche il metodo non lineare di Gauss-Seidel che prevede uno schema iterativo del tipo: $x_i^{(n+1)} = x_i^{(n)} - \frac{f_i(x_1^{(n)}, \dots, x_{i-1}^{(n)}, x_{i+1}^{(n)})}{\frac{\partial f_i(x_1^{(n)}, \dots, x_{i-1}^{(n)})}{\partial x_i}}$ $i=1, \dots, n$; $n=0, 1, \dots$ dove utilizziamo

le prime $i-1$ componenti già note dell'iterata $x^{(n)}$ in corso di calcolo
Possiamo anche utilizzare una versione semplificata del metodo di Newton-Raphson sostituendo le derivate della Jacobiana alle sue approssimazioni. Si utilizza quindi un'estensione del metodo delle secanti → la matrice $J(x^{(n)})$ viene sostituita dalla matrice $\Delta(x^{(n)})$ di ordine m , i cui elementi sono approssimazioni delle derivate $\frac{\partial f_i}{\partial x_k}$

• ZERI DI POLINOMI

Equazione algebrica di grado $m \geq 2$: $P(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0 = 0$ ($a_m \neq 0$). Le soluzioni possono essere tratte con i metodi già visti oppure sfruttando ulteriori proprietà della classe dei polinomi

$P(x) = 0$ e chiamiamo $P_0(x) = P(x)$, $P_1(x)$, $P_2(x)$, ..., $P_n(x)$ $n \leq m$

Successione di Sturm se verifica, su $[a, b]$:

1. $P_n(x) \neq 0 \quad \forall x \in [a, b]$
2. $P_i(x) = 0 \implies P_{i-1}(x) \cdot P_{i+1}(x) < 0 \quad i=1, 2, \dots, k-1$
3. $P_0(x) = 0 \implies P_1(x) \cdot P_0'(x) > 0$

→ Gli zeri di $P(x)$ sono tutti semplici

Introduciamo la funzione $V(x)$: VARIAZIONE. Voglio calcolarla in $V(x_0)$ (in un punto a piacere) $\rightarrow P_0(x_0) + \dots$

$P_1(x_0) + \dots$ Successione di segni: la variazione conta i cambi di segno. Il suo valore minimo è 0 ed il valore massimo è N .

teorema di Sturm: Se si dispone di una successione di Sturm relativa a $P(x)$, il numero degli zeri del polinomio $P(x)$ nell'intervallo $a < x < b$ è dato da $V(a) - V(b)$ ove $V(x)$ è il numero delle variazioni di segno presenti nella successione dei valori non nulli assunti dai polinomi della successione nel punto x .

↓ corollario

Se i polinomi della successione di Sturm sono $m+1$ ($m=k$) la successione si dice completa. In più, se tutti i polinomi hanno i coefficienti di grado massimo dello stesso segno allora l'equazione $P(x)=0$ ha m radici reali e distinte e viceversa.

$P_0(x)$	m	$-\infty$	$+\infty$
$P_1(x)$	$m-1$	-	+
$P_{m-1}(x)$	1	+	+
		-	+
		+	+

\downarrow m 0 $\underbrace{+}$

$$\text{Variazioni: } V(-\infty) - V(+\infty) = m - 0 = m \text{ soluzioni distinte}$$

Continuazione della successione di Sturm: Euclide $\rightarrow P(x), S(x) : P(x) = Q(x)S(x) + R(x)$ e continuo a fare la successione delle divisioni. Alla fine restituirà l'MCD tra $P(x)$ ed $S(x)$ che individua nell'ultimo resto $\neq 0$

↓ Metodo pratico

Si prende $P_0(x) := P(x)$ e $P_1(x) := P'(x) \rightarrow P_{r-1}(x) = P_r(x)Q_r(x) - P_{r+1}(x)$, $r=1,2,\dots$. In particolare avevamo $P_{r-1}(x) = Q_r(x)P_r(x)$ e iniziamo da $P_0(x) = Q_1(x)P_1(x) - P_2(x)$

relazione generale

resto della divisione: di massimo di grado $m-2$

Proseguendo abbiamo $P_1(x) = Q_2(x)P_2(x) - P_3(x)$

al massimo di grado $m-3$

Ahhiamo che $P_2(x) = Q_2(x)P_1(x) - P_0(x)$ e $P_3(x) = Q_3(x)P_2(x) - P_1(x)$ da cui abbiamo la relazione generale

Il processo termina, deince al crescere degli indici e quindi, per un $k \leq m$ risulta $P_{k-1}(x) = P_k(x)Q_k(x) \rightarrow$ Ci siamo ricondotti all'algoritmo di Euclide per cui abbiamo $P_k(x) = M.C.D \{ P(x), P'(x) \}$. Se la soluzione non fosse stata semplice il resto non sarebbe stata una costante. Se $P(x)$ e $P'(x)$ non hanno zeri in comune (soluzione semplice), i polinomi $P_0(x), P_1(x), \dots, P_k(x)$ formano una successione di Sturm. $P_k(x)$ non ha infatti zeri reali. Inoltre, se $P_r(x) = 0$, $1 \leq r \leq k-1$, si ha $P_{r-1}(x) = -P_{r+1}(x) \neq 0 \rightarrow P_{r-1}(x)P_{r+1}(x) < 0$. Infine, dalla definizione di $P_0(x)$ e $P_1(x)$, se $P_0(x) = 0$ segue $P_0(x)P_1(x) = (P'(x))^2 > 0$. Se invece $P(x)$ e $P'(x)$ hanno zeri reali in comune, questi sono tutti zeri di $P_k(x)$ con la stessa molteplicità che hanno come zeri di $P'(x)$. L'algoritmo di Euclide NON fornisce una successione di Sturm. Si può verificare che la successione $P_0(x), P_1(x), \dots, P_k(x)$ è una successione di Sturm.

$P_0(x) \quad P_1(x) \quad \dots \quad P_k(x)$

$P_n(x) \quad P_{n-1}(x) \quad \dots \quad P_0(x)$

di Sturm per il polinomio $P(x)/P_k(x)$ che ha tanti zeri semplici quanti sono gli zeri distinti di $P(x)$. $V(a) - V(b)$ per la nuova successione da il numero delle radici reali e distinte di $P(x)=0$ su $a < x < b \rightarrow$ Successione utilizzata per trovare $[a, b]$ contenente α , approssimato con i metodi precedenti Newton $\rightarrow x_{n+1} = x_n - \frac{P(x_n)}{P'(x_n)}$

ESEMPI

$$1. \quad x^4 - 2x^3 + 3x^2 - 4x + 1 = 0$$

$$P_0(x) = x^4 - 2x^3 + 3x^2 - 4x + 1$$

$$P_1(x) = 2x^3 - 3x^2 + 3x - 4 \rightarrow \text{Derivata prima divisa per 2}$$

$P_2(x) = -3x^2 + 9x - 2 \rightarrow$ Polinomio retto cambiato di segno (ed eventualmente moltiplicato per alcuni coefficienti)

$$P_3(x) = -8x + 3$$

$P_4(x) = -61 \rightarrow$ MCD = costante \rightarrow 2 soluzioni semplici e 2 complesse coniugate
vedo che segno assumono per $x \rightarrow +\infty$

$-\infty$	$+\infty$	0
+	+	+
-	+	-
-	-	-
+	-	+
-	-	-

+	+	+
-	+	-
-	-	-
+	-	+
-	-	-

3 variazioni 1 variazione $\rightarrow V(-\infty) - V(+\infty) = 2$ soluzioni reali

$$2. x^3 - 3x^2 + \lambda = 0, \lambda \in \mathbb{R}$$

$$P_0(x) = x^3 - 3x^2 + \lambda$$

$$P_1(x) = 3x^2 - 6x \text{ dividendo per } 3, x^2 - 2x$$

$$\begin{array}{c} P_2(x) = P_0(x) / P_1(x) \longrightarrow x^3 - 3x^2 + \lambda \\ \hline x^3 - 3x^2 & | x^2 - 2x \\ \hline // -x^2 + \lambda & | x - 1 \\ \hline -x^2 + 2x & | Q \\ \hline // \underbrace{-2x + \lambda}_{R} & \end{array}$$

$$P_2(x) = 2x - \lambda$$

$$P_3(x) = 4\lambda - \lambda^2 \quad (\text{Anche questo dato dalla divisione})$$

Quando questo è nullo la successione si forma a $P_2(x) \longrightarrow \lambda = 0, \lambda = 4$

$$\lambda \in]-\infty, 0[\quad \begin{array}{c|cc} -\infty & +\infty \\ \hline - & + \\ + & + \\ - & + \\ - & - \\ \hline 2 & 1 \end{array} \quad V(-\infty) - V(+\infty) = 1 \longrightarrow 1 \text{ soluzione reale con molteplicità 1} \text{ in quanto } P_3 \text{ è una COSTANTE}$$

$$\lambda = 0 \quad \begin{array}{c|cc} -\infty & +\infty \\ \hline - & + \\ + & + \\ - & + \\ \hline 2 & 0 \end{array} \quad P_3 \text{ non ne fa parte}$$

$V(-\infty) - V(+\infty) = 2$ radici reali ed essendo $P_2(x) = 2x$ avremo una soluzione di molteplicità 2. Dette, sono $\alpha_1 = \alpha_2 = 0$ e $\alpha_3 = 3$

$$\lambda \in]0, 4[\quad \begin{array}{c|cc} -\infty & +\infty \\ \hline - & + \\ + & + \\ - & + \\ + & + \\ \hline 3 & 0 \end{array} \quad V(-\infty) - V(+\infty) = 3 \text{ radici reali distinte (molteplicità 1) in quanto l'ultimo polinomio è una costante}$$

$$\lambda = 4$$

	$-\infty$	$+\infty$
-	+	
+	+	
-	+	
2	0	$\rightarrow \sqrt{(-\infty)} - \sqrt{(+\infty)} = 2$ radici reali $\alpha_1 = \alpha_2 = 2, \alpha_3 = -1$

$$\lambda \in]4, +\infty[$$

	$-\infty$	$+\infty$
-	+	
+	+	
-	+	
-	-	
2	1	$\rightarrow \sqrt{(-\infty)} - \sqrt{(+\infty)} = 1$ soluzione reale e 2 complesse coniugate

Quindi: $] -\infty, 0[\cup]4, +\infty[$ 1 radice reale e 2 complesse coniugate

$\lambda = 0, \lambda = 4$ 2 radici reali e una di molteplicità 2

$]0, 4[$ 3 radici reali

3. Matrice tridiagonale

$$T_n = \begin{pmatrix} a_1 & b_1 & & \\ b_1 & a_2 & b_2 & \\ & \ddots & \ddots & b_n \\ & & b_n & a_n \end{pmatrix} \in \mathbb{C}^{n \times n}$$

$$\det(T_0) = 1 \quad \det(T_1) = a_1 \quad \det(T_2) = a_1 a_2 - b_1^2 b_2$$

$$\det(T_2) = a_2 \det(T_1) - |b_2|^2 \det(T_0)$$

$$\det(T_3) = a_3 \det(T_2) - a_1 b_3 b_2 = a_3 \det(T_2) - |b_3|^2 \det(T_1)$$

\downarrow Generalizzando

$$\det(T_r) = a_r \det(T_{r-1}) - |b_r|^2 \det(T_{r-2}) \quad r = 2, \dots, n$$

Per il polinomio caratteristico: $\det(T_0 - \lambda I) = 1$

$$\det(T_1 - \lambda I) = a_1 - \lambda$$

\downarrow Generalizzando

$$\det(T_r(\lambda)) = (a_r - \lambda) \det(T_{r-1}(\lambda)) - |b_r|^2 \det(T_{r-2}(\lambda))$$

$$P(n) = (-1)^n T_n(\lambda), P_{n-1} = (-1)^{n-1} T_{n-1}(\lambda) \xrightarrow{\text{Generalize}} P_i(\lambda) = (-1)^i T_i(\lambda) \quad i = 0, \dots, n$$

Scendendo a 0 riceviamo una successione di Sturm completa

Metodi numerici per il calcolo degli autovettori

INTRODUZIONE

Non utilizziamo mai l'equazione caratteristica in quanto piccoli errori possono portare ad introdurre errori grandi.

METODO delle POTENZE

utile quando l'autovettore massimo è SEMPLICE (ha molteplicità 1)

$A \in \mathbb{C}^{n \times n}$ con 2 proprietà fondamentali \rightarrow 1. Diagonalisabilità: $S^{-1}AS = D = \begin{pmatrix} c_{11} & 0 \\ 0 & c_{nn} \end{pmatrix}$

2. $|\lambda_1| > |\lambda_2| > |\lambda_3| > \dots > |\lambda_n|$: un solo autovettore dominante

In queste ipotesi prendo $\underline{x}^{(0)} \in \mathbb{C}^n \setminus \{0\}$ e seguo l'algoritmo

Pomo verificare che $\lim_{k \rightarrow \infty} \frac{\underline{y}^{(k)}}{\underline{y}^{(0)H} \underline{A} \underline{y}^{(k)}} = \underline{x}^{(1)} \alpha$ autovettore associato a λ_1

e che $\lim_{k \rightarrow \infty} \frac{\underline{y}^{(k)H} \underline{A} \underline{y}^{(k)}}{\underline{y}^{(k)H} \underline{y}^{(k)}} = \lambda_1$

DIMOSTRAZIONE

$$\underline{x}^{(0)} = \underline{y}^{(0)} = C_1 \underline{x}^{(1)} + C_2 \underline{x}^{(2)} + \dots + C_n \underline{x}^{(n)} = \sum_{i=1}^n C_i \underline{x}^{(i)}$$

diagonalisabilità: n autovettori linearmente indipendenti \rightarrow molteplicità algebrica = geometrica

$$\underline{y}^{(k)} = \underline{A} \underline{y}^{(k-1)} = \underline{A}^2 \underline{y}^{(k-2)} = \underline{A}^3 \underline{y}^{(k-3)} = \dots = \underline{A}^k \underline{y}^{(0)}$$

\downarrow sostituendo

$$\underline{y}^{(k)} = \underline{A}^k \sum_{i=1}^n C_i \underline{x}^{(i)} = \sum_{i=1}^n C_i \underline{A}^k \underline{x}^{(i)} = \sum_{i=1}^n C_i \lambda_i^k \underline{x}^{(i)}$$

matrice per autovettore

$$= \lambda_1^k (C_1 \underline{x}^{(1)} + w^{(k)})$$

$$\begin{aligned} \underline{y}^{(k)} &= \lambda_1^k (C_1 \underline{x}^{(1)} + w^{(k)}) \\ \underline{y}_j^{(k)} &= \lambda_1^k (C_1 x_j^{(1)} + w_j^{(k)}) \end{aligned}$$

Porto fuori il primo elevato della k

$$\underline{y}^{(k)} = \lambda_1^k \left(C_1 \underline{x}^{(1)} + \sum_{i=2}^n C_i \frac{\lambda_i^k}{\lambda_1^k} \underline{x}^{(i)} \right) =$$

$$\lim_{k \rightarrow \infty} \frac{\lambda_1^k (C_1 \underline{x}^{(1)} + w^{(k)})}{\lambda_1^k (C_1 x_j^{(1)} + w_j^{(k)})} = \frac{C_1 \underline{x}^{(1)}}{C_1 x_j^{(1)}} = \frac{\underline{x}^{(1)}}{x_j^{(1)}} \rightarrow C_1 \neq 0$$

Quindi $\lim_{k \rightarrow \infty} \frac{\underline{y}^{(k)}}{\underline{y}^{(0)H} \underline{A} \underline{y}^{(k)}} = \frac{\underline{x}^{(1)}}{\underline{x}^{(1)H} \underline{x}^{(1)}}$ se $C_1 \neq 0$

Per il secondo limite: $\lim_{k \rightarrow \infty} \frac{\underline{y}^{(k)H} \underline{A} \underline{y}^{(k)}}{\underline{y}^{(k)H} \underline{y}^{(k)}}$ $\rightarrow \lim_{k \rightarrow \infty} \frac{(\underline{y}^{(k)}/\underline{y}^{(0)})^H \underline{A} (\underline{y}^{(k)}/\underline{y}^{(0)})}{(\underline{y}^{(k)}/\underline{y}^{(0)})^H (\underline{y}^{(k)}/\underline{y}^{(0)})} = \lambda_1$

Criterio di corretto

$$R(\underline{y}^{(k)}) = \frac{\underline{y}^{(k)H} \underline{A} \underline{y}^{(k)}}{\underline{y}^{(k)H} \underline{y}^{(k)}} \quad (\text{Quoziente di Rayleigh}) \rightarrow |R(\underline{y}^{(k)}) - R(\underline{y}^{(k-1)})| < \varepsilon \quad \text{ove } \varepsilon \in \mathbb{R}^+ \text{ è fissato}$$

La verifica della diagonalisabilità è difficile ma è sicuramente verificata se la matrice è normale ($\underline{A}^H \underline{A} = \underline{A} \underline{A}^H$)

Versione normalizzata del metodo delle potenze

$$\underline{x}^{(0)} \in \mathbb{C}^n \setminus \{0\}$$

$$\begin{cases} \underline{y}^{(k)} = \underline{A} \underline{x}^{(k-1)} \\ \underline{x}^{(k)} = \underline{y}^{(k)} / \| \underline{y}^{(k)} \|_\infty \end{cases} \quad k = 1, 2, \dots$$

\downarrow coefficiente di normalizzazione in generale $\alpha_k = \| \underline{y}^{(k)} \|_\infty$. Con queste scelte avremo che $\| \underline{x}^{(k)} \|_\infty = 1$

Può inoltre accadere che λ_1 abbia molteplicità $r > 1$ come: $\lambda_1 = \lambda_2 = \dots = \lambda_r$ con $|\lambda_1| > |\lambda_{r+1}| > \dots > |\lambda_n|$. Se $r > 1$, $\underline{x}^{(k)}$ converge ad un autovettore generato dagli r autovettori associati a λ_1 .

estensione

$|\lambda_1| > |\lambda_2| > |\lambda_3| > \dots > |\lambda_n|$: il metodo consente il calcolo di λ_2 e del corrispondente autovettore utilizzando la convergenza di λ_1 ed $\underline{x}^{(1)}$. Nelle matrici normali gli autovettori sono ortogonali. Supponendo di norma 1 otteniamo $\underline{x}^{(i)H} \underline{x}^{(j)} = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \quad i \geq 1, j \leq n$

Conosco λ_1 e $x^{(1)}$ e posso ricavare $A_0 = A$

$$Ax = A_0 - \lambda_1 x^{(1)} x^{(1)H}$$

↓ postmoltiplicando per $x^{(1)}$

$$Ax^{(1)} = A_0 x^{(1)} - \lambda_1 x^{(1)} \underbrace{x^{(1)H} x^{(1)}}_0 = Ax^{(1)} - \lambda_1 x^{(1)} = 0$$

d' autovettore $x^{(1)}$ sarà quindi associato all'autovettore 0. Proviamo quindi a moltiplicare per un altro autovettore: $A_1 x^{(2)} = A_0 x^{(2)} - \lambda_1 x^{(1)} \underbrace{x^{(1)H} x^{(2)}}_{\text{O}} \rightarrow A_1 x^{(2)} = A_0 x^{(2)} = \lambda_2 x^{(2)}$

Gli autovettori di A_1 saranno quindi quelli associati agli autovettori $\lambda_2, \lambda_3, \dots, \lambda_n$. gli autovettori di A_1 saranno quindi 0, $\lambda_2, \lambda_3, \dots, \lambda_n$ di cui λ_2 è quello di massimo modulo. Applicando il metodo delle potenze ad A_1 troverò λ_2 e $x^{(2)}$. Se fosse ancora verificata l'ipotesi dell'autovettore di modulo maggiore, potrei ricavare $A_2 = A_1 - \lambda_2 x^{(2)} x^{(2)H}$, i cui autovettori saranno 0, 0, $\lambda_3, \dots, \lambda_n$ → tecnica di deflazione

↓

Se ho che $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$ posso calcolare tutti gli autovettori e gli autovettori ponendo

$$A_i = A - \sum_{r=1}^i \lambda_r x^{(r)} x^{(r)H}, \quad i = 1, 2, \dots, n-1$$

• METODO DI JACOBI PER MATRICI REALI E SIMMETRICHE

$$G_{rt} = \begin{pmatrix} 1 & & & & & \\ \ddots & & & & & \\ & 1 & & & & \\ & & \cos \varphi & & -\sin \varphi & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix}$$

→ Al punto k-esimo il metodo trasforma una matrice A_k mediante una matrice ortogonale $G^{(k)}$ secondo l'algoritmo

$$A_{k+1} := A_k$$

$$A_{k+1} = G^{(k)T} A_k G^{(k)}$$

↓

Per negliere r e t guardo gli elementi non diagonali e si sceglie un elemento di massimo modulo. I suoi indici saranno gli indici

r e t che andranno a costituire la matrice di rotazione. Si sceglieranno poi $\sin \varphi$ e $\cos \varphi$ in modo che, dopo la trasformazione, l'elemento di massimo modulo precedentemente considerato sia 0. La trasformazione modifica solo gli elementi di indice r e t .

Ahhiamo le formule matematiche: $\sin^{(k+1)} = \sin^{(k)} = (-\sin^{(k)} - \cos^{(k)}) \sin \varphi + \cos^{(k)} (\cos^2 \varphi - \sin^2 \varphi)$

Niente vieta che un elemento fatto passare a 0 possa tornare ad un valore $\neq 0$ in passaggi successivi

↓

TEOREMA di JACOBI: La successione $\{A_k\}$ generata con il metodo di Jacobi converge alla matrice diagonale $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ ove $\lambda_1, \lambda_2, \dots, \lambda_n$ sono gli autovettori di A

Criterio di arresto: $\max_{1 \leq j \leq n} |a_{ij}^{(k+1)}| \leq \varepsilon$ ove $\varepsilon > 0$ è finito

La parte più dispendiosa del metodo è quella della ricerca dell'elemento di massimo modulo tra gli elementi non diagonali → Utilizziamo quindi il metodo di Jacobi circolare che prevede di annullare gli elementi al di sopra della diagonale principale rega per rega

↓

L'operazione si ripete ripetendo gli annullamenti nello stesso ordine

• METODO di GIVENS

È possibile trasformare una matrice hermitiana in una matrice tridiagonale simile. Per fare questo ragionamento poniamo utilizzare il metodo di Givens che fa utilizzo di matrici di rotazione. La differenza rispetto a Jacobi sta nel fatto che un elemento che è stato annullato non viene più modificato. Si annullano i termini non nulli

tra gli elementi a_{ij} , i-j > 2 considerati per colonna nell'ordine $a_{31}, a_{41}, \dots, a_{n1}; a_{32}, a_{42}, \dots, a_{n2}; a_{n1}, a_{n,n-1}$. Si utilizzano per questo le matrici di rotazione $G_{23}, G_{34}, \dots, G_{2n}; G_{31}, G_{32}, \dots, G_{3n}$; $G_{n-1,n}$ le matrici costruite in questo modo non modificano gli elementi già assorbiti. Nelle iterazioni successive gli elementi sono combinati linearmente e rimangono nulli se già nulli in precedenza. Al massimo dovrà assorire la norma dei primi $m-2$ interi data da $(n-2)(n-1)$. Avrà quindi una matrice tridiagonale a cui potranno applicare tutte le tecniche di calcolo degli autovalori. Se la matrice non è hermitiana/reale e simmetrica si ottiene la matrice di Hessenberg superiore $H = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \dots & h_{nn} \end{pmatrix}$

o Il metodo QR

Fattorizzazione QR teorema: Per ogni matrice $A \in \mathbb{R}^{n \times n}$ esiste una fattorizzazione data dal prodotto di una matrice Q ortogonale per una matrice R triangolare superiore. Nel caso complesso, Q è una matrice unitaria.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

Utilizzo la matrice G_{12}

↓

$G_{12}A$ con l'obiettivo che $a_{21} \rightarrow 0$

Utilizzo adesso la matrice G_{13} rendendo $a_{31}=0 \rightarrow G_{13}G_{12}A$

Potrete utilizzare adesso la matrice G_{14} rendendo $a_{41}=0 \rightarrow G_{14}G_{13}G_{12}A$

Proseguendo con gli annullamenti: $G_{14}G_{13}G_{12}A \rightarrow$ ho assorbito tutti gli elementi colorati

Dovrò quindi eseguire $\frac{(n-1)n}{2}$ premoltiplicazioni per una matrice di dimensione n

Il prodotto $G_{31}G_{21}G_{23}G_{11}G_{13}G_{12}A = R \rightarrow A = G_{12}^T \dots G_{34}^T R$

Il prodotto mi dà una matrice ortogonale: Q

Ho quindi ottenuto $A = QR$

Se ho un sistema lineare $QRx = b$, posso moltiplicare per Q^T ottenendo $Rx = Q^T b$

↓

Se ho una matrice di Hessenberg superiore, calcolare una fattorizzazione QR richiede $m-1$ moltiplicazioni per matrici di rotazione

Algoritmo del metodo QR

$$A_1 = A, A_K = Q_K R_K, A_{K+1} = R_K Q_K \text{ per } K=1,2,\dots$$

↓

$$A_1 = Q_1 R_1 \text{ da cui ricavo } A_2 = R_1 Q_1 \rightarrow A_2 = R_1 Q_1 = I R_1 Q_1 = Q_1^T Q_1 R_1 Q_1$$

↓

$A_2 = Q_1^T A_1 Q_1$: trasformazione per similitudine, quindi A_1 ed A_2 hanno gli stessi autovalori

In generale: $A_K = Q_K R_K \rightarrow A_{K+1} = R_K Q_K = Q_K^T Q_K R_K Q_K = Q_K^T A_K Q_K$. Tutte le matrici della successione hanno quindi gli stessi autovalori di A .

Teorema 1 di Schur

Per ogni matrice $A \in \mathbb{R}^{n \times n}$ esiste una matrice reale ortogonale B tale che la trasformata $S = B^{-1}AB = B^T AB$ è una matrice triangolare a blocchi con blocchi diagonali di ordine uno o due della forma

$$S = \begin{pmatrix} S_{11} & S_{12} & \dots & S_{1n} \\ S_{21} & S_{22} & \dots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n1} & S_{n2} & \dots & S_{nn} \end{pmatrix}$$

Se i blocchi di dimensione 1 danno gli autovettori reali di A , mentre ogni blocco diagonale di ordine 2 ha come autovettori una coppia di autovettori di A complessi coniugati.

Teorema 2

Se gli autovettori di $A \in \mathbb{R}^{n \times n}$ sono reali e distinti in modulo e verificano la condizione $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$ e se gli autovettori di A formano una matrice X tale che X^{-1} sia fattorizzabile LR, allora la matrice A_k , per $k \rightarrow \infty$ tende ad una matrice triangolare superiore e gli elementi diagonali $a_{ii}^{(k)}$ di A_k tendono agli autovettori λ_i di A ordinati per modulo decrescente. Se X^{-1} non è fattorizzabile LR il limite cui da sempre una matrice triangolare superiore i cui autovettori però non sono più ordinati in modulo decrescente. Se non è invece rispettata la condizione sui moduli avendo una coppia di complessi coniugati, la matrice R sarà triangolare superiore a blocchi. Se lo due autovettori distinti con lo stesso modulo si utilizza una traslazione dello spettro in modo da diaccoppiarli come modulo.

Utilizzo delle matrici di Hessenberg

$$A = \begin{pmatrix} X & X & X & X \\ X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \end{pmatrix} \longrightarrow G_{12} = \begin{pmatrix} X & X & 0 & 0 \\ X & X & 0 & 0 \\ 0 & 0 & X & 0 \\ 0 & 0 & 0 & X \end{pmatrix}, G_{23} = \begin{pmatrix} X & 0 & 0 & 0 \\ 0 & X & X & 0 \\ 0 & X & X & 0 \\ 0 & 0 & 0 & X \end{pmatrix}, G_{34} = \begin{pmatrix} X & 0 & 0 & 0 \\ 0 & X & 0 & 0 \\ 0 & 0 & X & X \\ 0 & 0 & X & X \end{pmatrix}$$

↓

$$G_{34}G_{23}G_{12}A = \begin{pmatrix} X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \\ 0 & 0 & 0 & X \end{pmatrix} \text{ e } Q = G_{12}^T G_{23}^T G_{34}^T = \begin{pmatrix} X & X & X & X \\ X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \end{pmatrix}$$

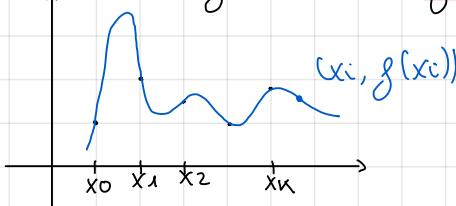
Eseguendo il prodotto RQ ottengo RA = $\begin{pmatrix} X & X & X & X \\ X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \end{pmatrix}$: Matrice di Hessenberg superiore

Interpolazione ed approssimazione di funzioni

• POLINOMI DI INTERPOLAZIONE

Di una funzione $f(x)$ che non conosco, ho i punti x_0, x_1, \dots, x_K . Vogliamo trovare la funzione interpolante. Per essere definita tale deve verificare che $g(x_i) = f(x_i)$, $i = 0, 1, \dots, K$.

→ Graficamente:



Invece di cercare una funzione qualunque, per l'interpolazione parabolica cerchiamo un polinomio di interpolazione $p(x)$ tale che verifichi $p(x_i) = f(x_i)$, $i = 0, 1, \dots, K$.

Se voglio approssimare la funzione su $K+1$ punti, cercherò un polinomio di grado K nell'insieme $\{x_i\}$. Avrò quindi $p_K(x) = a_K x^K + a_{K-1} x^{K-1} + \dots + a_1 x + a_0$ con $a_i \in \mathbb{R}$ $\forall i = 1, \dots, K$. Voglio adesso imporre che $p_K(x_i) = f(x_i)$ $\forall i = 1, \dots, K$.

$$\left. \begin{array}{l} a_K x_0^K + a_{K-1} x_0^{K-1} + \dots + a_1 x_0 + a_0 = f(x_0) \\ a_K x_1^K + a_{K-1} x_1^{K-1} + \dots + a_1 x_1 + a_0 = f(x_1) \\ \vdots \\ a_K x_N^K + a_{K-1} x_N^{K-1} + \dots + a_1 x_N + a_0 = f(x_N) \end{array} \right\} \text{Sistema lineare di incognite } a_i$$

↓ scrivendo il sistema lineare

$a = (a_K, a_{K-1}, \dots, a_1, a_0)^T$ con matrice di coefficienti

$$\begin{pmatrix} x_0^K & x_0^{K-1} & \dots & x_0 & 1 \\ x_1^K & x_1^{K-1} & \dots & x_1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_N^K & x_N^{K-1} & \dots & x_N & 1 \end{pmatrix}$$

Ottieniamo quindi $\underbrace{\begin{pmatrix} x_0^K & x_0^{K-1} & \dots & x_0 & 1 \\ x_1^K & x_1^{K-1} & \dots & x_1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_N^K & x_N^{K-1} & \dots & x_N & 1 \end{pmatrix}}_{X_i^0} \begin{pmatrix} a_K \\ a_{K-1} \\ \vdots \\ a_0 \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_N) \end{pmatrix}$ → Può accadere che il polinomio abbia grado minore di K

Matrice di vandermonde → Si può verificare che $|\det(V)| = \prod_{0 \leq i < j \leq K} |x_i - x_j|$

che non può essere 0 in quanto $x_i \neq x_j$ per $i \neq j$. Il polinomio di interpolazione è quindi unico: non conviene determinare il polinomio risolvendo il sistema lineare.

Queste matrici sono tipicamente malcondizionate:

Alternativa

$$l_r(x) = \frac{(x-x_0)(x-x_1)\dots(x-x_{r-1})(x-x_{r+1})\dots(x-x_N)}{(x_r-x_0)(x_r-x_1)\dots(x_r-x_{r-1})(x_r-x_{r+1})\dots(x_r-x_N)}$$

$r = 0, 1, \dots, K$ → polinomi fondamentali dell'interpolazione di Lagrange

Il denominatore è un NUMERO ≠ 0. Il NOMERATORE è prodotto di K fattori polinomiali di primo grado. È quindi un polinomio di grado K ESATTAMENTE.

↓ PROPRIETÀ

$$l_r(x_s) = \begin{cases} 1 & r=s \\ 0 & r \neq s \end{cases} \rightarrow l_r(x) = \sum_{i=0}^K l_i(x) f(x_i) \rightarrow l_r(x_s) = f(x_s), \quad s=0, 1, \dots, K$$

polinomio di interpolazione nella forma di Lagrange

Affiamo anche un'ulteriore proprietà → $\sum_{i=0}^K l_i(x) = 1$

ESEMPI

1.	x	0	1	2	-1
	y	1	1	3	3

$$l_0(x) = \frac{(x-1)(x-2)(x+1)}{(-1)(-2)(1)}$$

$$l_1(x) = \frac{(x-0)(x-1)(x+1)}{(1)(-1)(2)}$$

$$l_2(x) = \frac{(x-0)(x-1)(x+1)}{(2)(1)(3)}$$

$$l_3(x) = \frac{(x-0)(x-1)(x-2)}{(-1)(-2)(-3)}$$

$$L_3(x) = l_0(x) \cdot 1 + l_1(x) \cdot 1 + l_2(x) \cdot 3 + l_3(x) \cdot 3 \longrightarrow L_3(x) = x^2 - x + 1$$

↓

Per vedere se è corretto basta sostituire i valori delle x e verificare che corrispondano ai valori di y corretti

LIMITAZIONI DEL METODO DI LAGRANGE

Per aggiungere un punto è necessario rifare tutti i calcoli da 0

DIFFERENZE DIVISE: Siamo $x_0, x_1, \dots, x_{n-1} \in I$ con $x_i \neq x_j$. Abbiamo ad esempio $f[x, j] = f(x)$ che è la differenza divisa di ordine 0. La differenza divisa di ordine 1 sarà strutturata come $\frac{f[x_0, x_j] - f(x_0)}{x_j - x_0}$. Abbiamo quindi tolto l'ultimo punto.

Procedendo, la differenza divisa di ordine 2 risulta $f[x_0, x_1, x_j] = \frac{f[x_0, x_j] - f[x_0, x_1]}{x_j - x_1}$

Analogamente, la differenza divisa di ordine 3 risulta $f[x_0, x_1, x_2, x_j] = \frac{f[x_0, x_1, x_j] - f[x_0, x_1, x_2]}{x_j - x_2}$

in generale: $f[x_0, x_1, \dots, x_{n-1}, x_j] = \frac{f[x_0, x_1, \dots, x_{n-1}, x_j] - f[x_0, x_1, \dots, x_{n-1}]}{x_j - x_{n-1}}$

SIMMETRIA: Permutando l'ordine dei parametri, la differenza divisa non cambia
prolungamento per continuità: se $f \in C^1(I)$ la funzione $f[x_0, x_1, \dots, x_{n-1}, x]$ è prolungabile per continuità su tutto I .

Proprietà 3: Se $f(x) \in C^k(I)$ esiste almeno un valore ξ , con la proprietà $\min_{i \leq k} \xi < \xi < \max_{i \leq k} \xi$ tale che $f[x_0, x_1, \dots, x_k] = f^{(k)}(\xi)/k!$

TEOREMA DI ESPANSIONE: Sia $f(x) : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ e siano $x_0, x_1, \dots, x_n \in I$ con $x_i \neq x_j$

se $i \neq j \rightarrow f(x) = f(x_0) + (x-x_0)f[x_0, x_1] + (x-x_0)(x-x_1)f[x_0, x_1, x_2] + \dots + (x-x_0)(x-x_1)\dots(x-x_{n-1})f[x_0, \dots, x_{n-1}, x_n]$

↓

FUNZIONE: se tutto fare un numero $f(x)$ avrebbe un polinomio

DIMOSTRAZIONE

Per induzione: Poniamo $K=0$ $\rightarrow f(x) = f(x_0) + (x-x_0)f[x_0, x_1] =$

$$= f(x_0) + \frac{(x-x_0)}{x-x_0} f(x) - f(x_0) = f(x)$$

✓

Esplnsione al punto $K+1$: $f(x) = f(x_0) + (x-x_0)f[x_0, x_1] + (x-x_0)(x-x_1)f[x_0, x_1, x_2] + \dots$
 (supposta valida per K) $+ (x-x_0)\dots(x-x_K)f[x_0, \dots, x_{K+1}] + (x-x_0)\dots(x-x_{K+1})f[x_0, \dots, x_{K+1}, x]$

$$f[x_0, \dots, x_{K+1}, x] = \frac{f[x_0, \dots, x_K, x] - f[x_0, \dots, x_{K+1}]}{x-x_{K+1}}$$

$$f[x_0, \dots, x_K, x] = f[x_0, \dots, x_K] + (x-x_K)f[x_0, \dots, x_{K+1}, x]$$

Sostituendo: $f(x) = f(x_0) + (x-x_0)f[x_0, x_1] + (x-x_0)\dots(x-x_{K-1})f[x_0, \dots, x_K] +$
 \downarrow $+ (x-x_0)\dots(x-x_K)f[x_0, \dots, x_{K+1}] + (x-x_0)\dots(x-x_{K+1})f[x_0, \dots, x_{K+1}, x] = f(x)$ ✓

Possiamo dire che, escludendo l'ultimo termine, la sommatoria rappresenta un polinomio $\rightarrow f(x) = P_K(x) + (x-x_0)\dots(x-x_K)f[x_0, \dots, x_K, x]$

↓

Questo polinomio verifica la proprietà $P_K(x_i) = f(x_i) \quad i=0, \dots, K$

$$\hookrightarrow P_K(x) = f(x_0) + (x-x_0)f[x_0, x_1] + (x-x_0)(x-x_1)f[x_0, x_1, x_2] + \dots + (x-x_0)\dots(x-x_{K-1})f[x_0, \dots, x_K]$$

↪ polinomio di interpolazione di Newton

DIMOSTRAZIONE

Per $K=0 \rightarrow P_0(x_0) = f(x_0)$

Per un indice K avremo $P_K(x_i) = f(x_i) \quad i=0, \dots, K$ $\xrightarrow[\text{dimostrare}]{\text{vogliamo}} P_{K+1}(x_i) = f(x_i) \quad i=0, \dots, K+1$

$$P_{K+1}(x) = P_K(x) + (x-x_0) \dots (x-x_K) f[x_0, \dots, x_{K+1}]$$

$$\hookrightarrow P_{K+1}(x_i) = P_K(x_i) + 0 = P_K(x_i) = f(x_i) \quad \text{per ipotesi induttiva}$$

Dobbiamo adesso verificare che $P_{K+1}(x_{K+1}) = f(x_{K+1})$

↓

$$\text{Avremo che } P_{K+1}(x_{K+1}) = P_K(x_{K+1}) + (x_{K+1}-x_0) \dots (x_{K+1}-x_K) f[x_0, \dots, x_{K+1}]$$

Ricordando l'espansione di f : $f(x) = P_K(x) + (x-x_0) \dots (x-x_K) f[x_0, \dots, x_K, x]$

↓ calcolandola in x_{K+1}

$$f(x_{K+1}) = P_K(x_{K+1}) + (x_{K+1}-x_0) \dots (x_{K+1}-x_K) f[x_0, \dots, x_K, x_{K+1}]$$

$$\hookrightarrow \text{Ahhiamo ottenuto } P_{K+1}(x_{K+1}) = f(x_{K+1})$$

errore

$$f(x) = P_K(x) + (x-x_0) \dots (x-x_K) f[x_0, \dots, x_K, x] \rightarrow f(x) - P(x) = (x-x_0) \dots (x-x_K) f[x_0, \dots, x_K, x] = \mathcal{E}(x)$$

teorema: Se $f(x) \in C^{K+1}(I)$ avremo $f(x) - P(x) = \mathcal{E}(x) \frac{f^{(K+1)}(\xi_x)}{(K+1)!}$

Calcolo del polinomio di Newton

Le differenze finite che intervengono nel calcolo del polinomio sono $f(x_0)$, $f[x_0, x_1]$, $f[x_0, x_1, x_2]$, $f[x_0, x_1, x_2, x_3]$, ..., $f[x_0, x_1, \dots, x_K]$
↓ +0 per avere grado K

Utilizziamo il quadro delle differenze finite

x_i	$f(x_i)$	$f[x_0, x_i]$	$f[x_0, x_1, x_i]$	$f[x_0, x_1, x_2, x_i]$
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	$f(x_3)$	$f[x_0, x_3]$	$f[x_0, x_1, x_3]$	$f[x_0, x_1, x_2, x_3]$

↓ Supponiamo che i valori siano

-2 $\left\{ \begin{array}{l} \text{Le differenze finite vengono 0, anche le seguenti} \\ \downarrow \end{array} \right.$

Il grado del polinomio è pari all'ordine delle differenze finite non nulle

ESEMPIO

x	$f(x)$	DD1	DD2	DD3
0	5	—	—	—
-1	3	2	—	—
2	3	-1	-1	—
-2	-9	7	-5	1
3	11	2	0	1

Colonna costante: le prossime DD sono 0

$$P_4(x) = 5 + (x-0)\alpha + (x-0)(x+1)(-\lambda) + (x-0)(x+1)(x-2)(\lambda) = x^3 - 2x^2 - x + 5$$

ESERCIZI

1. $A \in \mathbb{C}^{3 \times 3}$ $\det(A - I) = -1$, $\det(A - 2I) = -3$, $\det(A + I)$, $\det(A) = -1$

$$P(\lambda) = \det(A - \lambda I)$$

λ	$\det(A - \lambda I)$	DD1	DD2	DD3
1	-1	—	—	—
2	-3	-2	—	—
-1	3	-2	0	—
0	-1	0	-1	-1

$$\begin{aligned}\det(A - \lambda I) &= -1 - 2(\lambda-1) + 0(\lambda-1)(\lambda-2) \\ &= -1(\lambda-1)(\lambda-2)(\lambda+1) = \\ &= -\lambda^3 + 2\lambda^2 - \lambda - 1\end{aligned}$$

x	$f(x)$	DD1	DD2
4	6	—	—
8	47	9	—
12	10	13	1
16	210	17	1
20	362	21	1

→ il polinomio di interpolazione ha grado 2

↓

$$P_2(x) = 6(x-6) + (x-6)(x-8) = x^2 - 3x + 2$$

3. x $f(x)$ α, β da determinare per avere grado minimo

$0 -1$
 -1β → Abbiamo 5 punti, quindi il grado sarà al più 4. α deve avere
 $1 -1$ valori diversi dalle x per ipotesi polinomio di interpolazione
 $2 1$
 $\alpha 5$

⚠ SOGGERIMENTO: conviene tenere in fondo le righe con i parametri sulla x e prima di queste le righe che coinvolgono parametri nel $f(x)$

x	$f(x)$
0	-1
1	-1
2	1
-1	β
α	5

do posso trovare calcolando $P_2(-1) = \beta$. Aggiungo un punto eguale
e quindi non alto il grado
Per non alto il grado dobbiamo fare $P_2(\alpha) = 5$

x	$f(x)$	DD1	DD2
0	-1	—	—
1	-1	0	—
2	1	1	1

$$P_2(x) = -1 + 0(x-0) + 1(x-0)(x-1) = x^2 - x - 1$$

$$\beta = P_2(-1) = 1 + 1 - 1 = 1$$

$\alpha^2 - \alpha - 1 = 5 \rightarrow$ ⚠ Se non ha soluzione dobbiamo trovare un polinomio di
interpolazione di grado più alto

$$\alpha^2 - \alpha - 6 = 0 \rightarrow \alpha_1 = 3, \alpha_2 = -2 : 2 polinomi di interpolazione di grado minimo$$

4. $P(x) = x^4 - 3x^3 + x^2 + x + 1$ è il polinomio di interpolazione di $x | f(x)$?



0	1
1	1
-1	5
2	-1

Non può essere perché, avendo 4 punti, il grado massimo del polinomio è 3

5. $x | f(x) \quad x \in \mathbb{R}$

$\begin{array}{|c|c|} \hline 0 & 2 \\ \hline 1 & 1 \\ \hline x & 4 \\ \hline -1 & 3x+1 \\ \hline 3 & 11 \\ \hline \end{array}$ ⚠ Adesso abbiamo lo stesso parametro, quindi non posso fare i ragiona-
-menti di prima



Tuttavia di nuovo in fondo i valori dipendenti dai parametri

$x f(x)$	DD1	DD2
0 2	—	—
1 1	-1	—
3 11	3	$\frac{2}{3x-2}$
-1 $3x+1$	$1-3x$	$\frac{2}{x(x-1)}$
$x 4$	$\frac{2}{x}$	$\frac{x+\alpha}{\alpha(\alpha-1)}$

$$x \neq 0, 1, 3, -1$$

$$\frac{3x-2}{2} = 2 \rightarrow 4 = 3x-2 \rightarrow x=2$$

sostituendola nell'ultima fa 2, quindi $x=2$ va bene

$$P_4(x) = 2 - x + 2x(x-1) = 2x^2 - 3x + 2$$

• INTERPOLAZIONE OSCULATRORIA di HERMITE

Abbiamo $k+l$ punti reali $x_0, x_1, \dots, x_k \in I$ a due a due distinti dei quali vicino molti $2k+2$ valori reali $f(x_0), f'(x_0), \dots, f(x_k), f'(x_k), \dots, f'(x_{k+l})$. L'interpolazione osculatoria di Hermite consente nel determinare un polinomio $H_{2k+l}(x)$ di grado al più $2k+l$ tale che $H_{2k+l}(x_r) = f(x_r)$, $H'(2k+l)(x_r) = f'(x_r)$, $r=0, 1, \dots, k$. Il polinomio di interpolazione di Hermite viene cercato del tipo $H_{2k+l}(x) = \sum_{r=0}^k h_{or}(x) f(x_r) + \sum_{r=0}^l h_{ir}(x) f'(x_r)$.

da dove la deriva di $H_{2k+l}(x)$ risulta: $H'(2k+l)(x) = \sum_{r=0}^k h_{or}(x) f'(x_r) + \sum_{r=0}^l h_{ir}(x) f''(x_r)$ per verificare le condizioni dette in precedenza dovremo avere che $\begin{cases} h_{or}(x_s) = S_{rs} \\ h_{ir}(x_s) = 0 \end{cases}$ e $\begin{cases} h_{or}'(x_s) = 0 \\ h_{ir}'(x_s) = S_{rs} \end{cases}$

Avevamo quindi che $h_{or}(x)$ e $h_{ir}(x)$ saranno polinomi, combinabili a piacimento in qualsiasi modo. $h_{or}(x)/h_{ir}(x)$ = polinomio di prima / seconda specie per l'interpolazione di Hermite. I polinomi $h_{or}(x)$ si cercano del tipo $h_{or}(x) = (Ax+B) l_{r^k}(x)$. Avevamo che $l_{r^k}(x_s) = S_{rs}$. Avevamo che $l_{r^k}(x)$ era di grado k , quindi $l_{r^k}'(x)$ è di grado $2k$ ed $h_{or}(x)$ sarà infine di grado $2k+l$. Avevamo che $h_{or}(x_s) = S_{rs}$ e questo è ancora verificato perché $l_{r^k}(x_s)^2 = S_{rs}$. Dobbiamo poi avere che, per $x=s$, $Ax+B=1$.

$h_{or}'(x) = Al_{r^k}'(x) + (Ax+B) l_{r^k}(x) l_{r^k}'(x)$. Vogliamo ovette che $h_{or}'(x_s) = 0$, il che è verificato per la proprietà di $l_{r^k}(x)$. $h_{or}'(x_s) = A + (Ax_s + B) 2 l_{r^k}'(x_s) = 0$. Per la 1, sostituendo nella 2, ottengo $A + 2l_{r^k}'(x_s) = 0$ da cui $A = -2l_{r^k}'(x_s)$. Della prima avremo poi $B = 1 + 2x_s l_{r^k}'(x_s)$

Ottieniamo infine $h_{or}(x) = (Cx+D) l_{r^k}'(x)$ tale che verifichi $h_{or}(x_s) = 0$, $h_{ir}(x_s) = S_{rs}$. Calcolando il polinomio in x_s ottieniamo $(Cx_s+D) l_{r^k}'(x_s) = 0 \rightarrow Cx_s + D = 0$. Calcolando la derivata prima $h_{ir}'(x) = Cl_{r^k}''(x) + (Cx+D) 2l_{r^k}'(x) l_{r^k}''(x)$. Calcolandola in x_s otterremo $C l_{r^k}''(x_s) +$

$+ (c x_s + d) 2 \ell_r(x_s) l'_r(x_s) = S_{rs}$. Per $r \neq s$ avremo che $c l'_r(x_r) + (c x_r + d) 2 \ell_r(x_r) l'_r(x_r) = 0$.
otteniamo che $c = 1$ e $d = -x_r$.

esercizio

Se $f(x) \in C^{2k+2}(\mathbb{I})$ si ha

$$f(x) - H(x) = (x - x_0)^2 \dots (x - x_k)^2 \frac{f^{(2k+2)}(x)}{(2k+2)!}$$

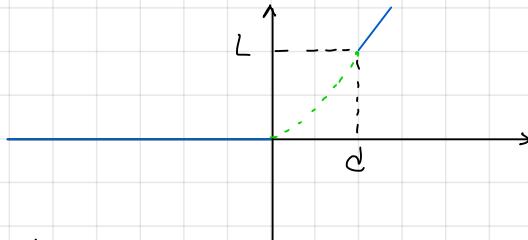
ove $\min\{x_0, x_1, \dots, x_k, x\} < x < \max\{x_0, x_1, \dots, x_k, x\}$. Se f fosse un polinomio la derivata di ordine $2k+2$ sarebbe 0 avendo il polinomio di grado massimo $2k+1$.

ESEMPIO

$$x_0 = 0, x_1 = d$$

$$\begin{aligned} f(x_0) &= 0 \\ f'(x_1) &= L \end{aligned}$$

$$\begin{aligned} f'(x_0) &= 0 \\ f'(x_1) &= S \end{aligned}$$



$$l_0(x) = \frac{x - x_1}{x_0 - x_1} = \frac{x - d}{-d} = \frac{d - x}{d}$$

$$l_1(x) = \frac{x - x_0}{x_1 - x_0} = \frac{x}{d}$$

$$l_0'(x) = -\frac{1}{d} \quad l_1'(x) = \frac{1}{d}$$

$$\begin{aligned} l_{00}(x) &= (1 - 2l_0(x_0)(x - x_0)) l_0^2(x) = \\ &= \left(1 - \frac{2}{d}(x - d)\right) \frac{d^2}{d^2} \end{aligned}$$

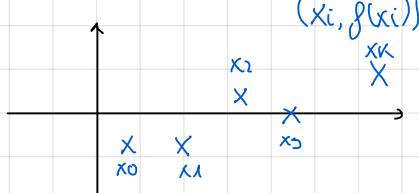
$$\begin{aligned} l_{01}(x) &= (1 - 2l_0'(x_1)(x - x_1)) l_0^2(x) = \\ &= \left(1 - \frac{2}{d}(x - d)\right) \frac{x^2}{d^2} \end{aligned}$$

$$l_{00}(x) = (x - x_0) l_0^2(x) = \frac{x(d-x)^2}{d^2}$$

$$l_{01}(x) = (x - x_1) l_0^2(x) = \frac{(x-d)x^2}{d^2}$$

osservazione: Nelle interpolazioni viste, all'aumentare dei punti può conseguire un significativo aumento del grado del polinomio, il che può portare a brevi oscillazioni. Inoltre, l'ordine $\ell_n(x)$ possono avere elevati valori agli estremi degli intervalli avendo $\lim_{n \rightarrow \infty} \max_{0 \leq x \leq 1} |\ell_n(x)| = +\infty$

funzioni spline: Fanno utilizzo di più polinomi



→ Su ciascun intervallo $[x_{i-1}, x_i]$ consideriamo un polinomio $p_i(x)$. Avremo tanti polinomi quanti sono gli intervalli. Dovremo avere che grado $(p_i(x)) = m$.

restrizione

Sceglieremo $S_m(x)$ tale che

$$S_m(x) \Big|_{[x_{i-1}, x_i]} = p_i(x) : \text{ funzione polinomiale a tratti}$$

interpolante, quindi $S_m(x_i) = f(x_i)$ e $S_m(x) \in C^{m-1}[x_0, x_n]$. Ci occupiamo del caso delle spline cubiche, ovvero quando $m=3$.

$$p_1(x_0) = f(x_0)$$

$$p_1(x_1) = f(x_1)$$

→ avendo ciascun polinomio di

$$p_2(x_1) = f(x_2)$$

grado massimo 3, avremo

$$p_3(x_2) = f(x_3)$$

6 coefficienti con 16 (L_{ik}) condizioni. Imponendo queste 6

$$p_4(x_3) = f(x_4)$$

$(2k)$ condizioni avremo che il polinomio sarà interpolante

$$\begin{matrix} \cdot & p_1(x_1) & p_2(x_2) & p_3(x_3) & p_4(x_4) \\ x_0 & x_1 & x_2 & x_3 & x_4 \end{matrix}$$

$$p_1'(x_1) = p_2'(x_2)$$

$$p_2'(x_2) = p_3'(x_3)$$

$$p_3'(x_3) = p_4'(x_4)$$

$k-1$ condizioni

e la funzione sarà C^1 . Perché la funzione sia C^1 , nei punti di raccolto le derivate rimbalza e dovrà devono essere uguali

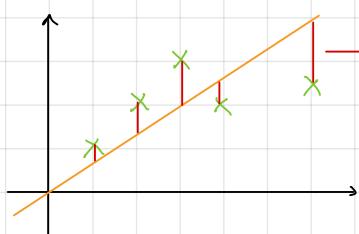
Per aumentare la classe di continuità applicherò queste condizioni sulle derivate seconde. Avrò ancora $K-1$ condizioni. Mi mancano però ancora 2 condizioni. Le due ulteriori si scelgono tra le seguenti:

- spline naturale: $q_1''(x_0) = q_K''(x_K) = 0$
- spline periodica: $q_1'(x_0) = q_1'(x_K)$, $q_1''(x_0) = q_1''(x_K)$
- spline vincolata: $q_1'(x_0) = q_0'$, $q_1''(x_K) = q_K''$ con $q_0' = f'(x_0)$ e $q_K'' = f''(x_K)$ noti

Teorema: Se i punti x_0, \dots, x_K sono tali che $x_i - x_{i-1} = h$, $i = 1, 2, \dots, K$, esiste una unica funzione spline cubica naturale $s_3(x)$.

o METODO DEI MINIMI QUADRATI

Sono dati $K+1$ punti $x_j \in I$, $j = 0, 1, \dots, K$, e si affermano i valori $f(x_j)$ per ogni x_j .



Cercavo una retta e calcolavo il distacco tra il valore dato nel punto ed il valore che aveva la retta nel punto

Considero una funzione $\phi(x)$, nel caso prendete $ax + b$ e faccio $\sum_{j=0}^K |\phi(x_j) - f(x_j)|$ → non è ideale perché non posso voler fare la derivata

$$\sum_{j=0}^K (\phi(x_j) - f(x_j))^2 = \Psi(c_0, c_1, \dots, c_m)$$

Nel caso di prima, se prendo $\phi_0(x) = x$ e $\phi_1(x) = 1$ e moltiplico lo per a e c_1 per b , ottengo la retta $ax + b$. In generale prendero $\phi_0(x), \phi_1(x), \dots, \phi_m(x)$ con $m < K$ linearmente indipendenti. Considero poi $\Phi(x) = \sum_{i=0}^m c_i \phi_i(x)$, $c_i \in \mathbb{R}$. → Il metodo dei minimi quadrati prevede di trovare i c_i tali che Ψ sia minima. → $\Psi(c_0, c_1, \dots, c_m) = \sum_{j=0}^K (\sum_{i=0}^m c_i \phi_i(x_j) - f(x_j))^2$. Se trovo un vettore tale per cui $\Psi = 0$ ho trovato una funzione interpolante in quanto, essendo una somma di quadrati, tutte le basi devono essere 0.

$$\frac{\partial \Psi}{\partial c_s} = 2 \sum_{j=0}^K \left(\sum_{i=0}^m c_i \phi_i(x_j) - f(x_j) \right) \phi_s(x_j) \quad s = 0, 1, \dots, m$$

Per trovare i punti stazionari devo porre $\nabla \Psi = 0$: $\sum_{j=0}^K \left(\sum_{i=0}^m c_i \phi_i(x_j) - f(x_j) \right) \phi_s(x_j) = 0$

Per risolvere il sistema lineare, dobbiamo trovare $c = (c_0, c_1, \dots, c_m)^T$ ed abbiamo $b = (f(x_0), f(x_1), \dots, f(x_m))^T$ e $A = \begin{pmatrix} \phi_0(x_0), \phi_1(x_0), \dots, \phi_m(x_0) \\ \phi_0(x_1), \phi_1(x_1), \dots, \phi_m(x_1) \\ \vdots \\ \phi_0(x_m), \phi_1(x_m), \dots, \phi_m(x_m) \end{pmatrix} \in \mathbb{R}^{(m+1) \times (m+1)}$

Trovare le derivate parziali = 0 significa risolvere il sistema lineare $A^T A c = A^T b$.

Un sistema lineare di questa forma ha sicuramente soluzioni: $\text{rk}(A^T A) = \text{rk}(A^T A)^T$. Il sistema viene detto sistema delle equazioni normale. Perché abbia un'unica soluzione, il determinante di $A^T A$ deve essere $\neq 0$. Per il teorema di Binet, se il rango di A è massimo, il determinante di $A^T A$ è diverso da 0. Per quanto riguarda la matrice $A^T A$, essendo che $(A^T A)^T = A^T A$, la matrice è simmetrica. La funzione Ψ come valori non negativi e non è superiormente limitata per $c_i \rightarrow +\infty$ ma è inferiormente limitata da 0. Il punto critico trovato risulta automaticamente un punto di minimo. La matrice $A^T A$ ha inoltre autovalori tutti reali non negativi. Calcolando il coefficiente di Rayleigh si ottiene $x^T A^T A x = (Ax)^T (Ax) = \lambda$ che non può essere < 0 . La funzione Ψ può avere scatti

$$x^T x \quad x^T x$$

$$\text{come } \Psi(c) = \|Ac - b\|_2^2$$

SISTEMI SORSE DETERMINATI

Vi sono più equazioni che incognite. Siano $K, m \in \mathbb{N}$ con $K > m$ ed inoltre siamo $A \in \mathbb{R}^{K \times m}$ e $b \in \mathbb{R}^K$. Il sistema lineare nonadeterminato $\bar{A}x = b$ ha soluzione se e solo se $r(A) = r(\bar{A}b)$. Nel caso in cui $r(A) < r(\bar{A}b)$, il sistema non ha soluzioni e quindi per ogni vettore $x \in \mathbb{R}^m$ si ha $b - Ax = r$, dove $r \in \mathbb{R}^K$ è detto vettore dei residui → risolvere il sistema nel senso dei minimi quadrati significa determinare $x \in \mathbb{R}^m$ che renda minimo $\psi(x) = (b - Ax)^T(b - Ax) = r^T r = \sum_{i=1}^m r_i^2 = \|r\|_2^2$. Per fare questo, dobbiamo trovare i punti per cui $\frac{\partial}{\partial x_i} ((b - Ax)^T(b - Ax)) = 0$.

Questo può essere scritto come $\bar{A}^T A x = \bar{A}^T b$, e le condizioni di esistenza sono le stesse di quelle del sistema delle equazioni normali.

ESEMPI

1. $\begin{cases} 2x_1 - x_2 = 1 \\ -x_1 + x_2 = 0 \\ x_1 + 2x_2 = 1 \end{cases}$ Si calcoli la soluzione nel senso dei minimi quadrati

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \\ 1 & 2 \end{pmatrix} \quad b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\bar{A}^T A = \begin{pmatrix} 6 & -1 \\ -1 & 6 \end{pmatrix} \quad \bar{A}^T b = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \begin{cases} 6x_1 - x_2 = 3 \\ -x_1 + 6x_2 = 1 \end{cases} \quad \begin{cases} x_1 - 6x_1 - 3 \\ -x_1 + 36x_1 - 1 = 1 \end{cases} \rightarrow x_1 = \frac{19}{35} \\ x_2 = \frac{9}{35}$$

Quindi $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{35} \begin{pmatrix} 19 \\ 9 \end{pmatrix}$

2.

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & \alpha & 1 \\ 0 & 0 & \alpha \\ \beta & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \rightarrow$$

Determinare α e β in modo che il sistema non abbia una soluzione unica nel senso dei minimi quadrati

↓

La matrice NON deve avere rango massimo, quindi da annullare i minori di grado massimo

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & \alpha & 1 \\ 0 & 0 & \alpha \end{vmatrix} = \alpha(\alpha-1)$$

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & \alpha & 1 \\ \beta & 0 & 0 \end{vmatrix} = \beta(1-\alpha) \rightarrow \alpha=1, \beta=0 \quad \rightarrow \text{Le due coppie di valori richieste sono } \alpha=1, \beta=0 \text{ oppure } \alpha=0, \beta=0$$

$$\begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & \alpha \\ \beta & 0 & 0 \end{vmatrix} = \beta \alpha \quad \left. \begin{array}{l} \beta=0 \\ \text{oppure } \alpha=0 \end{array} \right\}$$

$$\begin{vmatrix} 1 & \alpha & 1 \\ 0 & 0 & \alpha \\ \beta & 0 & 0 \end{vmatrix} = \beta \alpha^2$$

3.

$$\begin{pmatrix} 1 & \alpha \\ \alpha & \alpha \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \longrightarrow$$

Unica soluzione nel senso dei minimi quadrati: La matrice deve avere rango massimo. Dovendo quindi avere un minore di grado 2 ≠ 0

$$\begin{vmatrix} 1 & \alpha \\ \alpha & \alpha \end{vmatrix} = \alpha - \alpha^2 = \alpha(1-\alpha) \quad \alpha=0, \alpha=1 \longrightarrow$$

Per avere una soluzione unica è sufficiente che sia $\alpha \neq 1$

$$\begin{vmatrix} 1 & \alpha \\ 1 & 1 \end{vmatrix} = 1-\alpha \quad \alpha \neq 1$$

$$\begin{vmatrix} \alpha & \alpha \\ 1 & 1 \end{vmatrix} = 0$$

Integrazione Numerica

• FORMULE di QUADRATURA

Sia $f(x)$ sufficientemente regolare sull'intervallo $[a, b]$ dell'asse reale. Sia $p(x)$ una funzione polo non negativa in $[a, b]$ e tale che esistano finiti i momenti $m_n = \langle (x^k)p \rangle = \int_a^b x^k p(x) dx$, $k=0, 1, \dots$. Vogliamo approssimare $I(f) = \int_a^b p(x) f(x) dx$ con una formula di quadratura nella forma $J_n(f) = \sum_{i=0}^n a_i f(x_i)$, ove gli a_i sono detti pesi o coefficienti e sono reali e gli x_i sono detti nodi e di solito appartengono all'intervallo $[a, b]$. D'errore viene quindi definito come $\varepsilon_n(f) = |I(f) - J_n(f)|$. Si considera la base $1, x, x^2, \dots, x^m, x^{m+1}$ dello spazio vettoriale dei polinomi algebrici di grado al più $m+1$. La formula $J_n(f)$ ha grado di precisione $m \in \mathbb{N}$ se si verifica $\varepsilon_n(1) = \varepsilon_n(x) = \dots = \varepsilon_n(x^m) = 0$, $\varepsilon_n(x^{m+1}) \neq 0$. Il costo è dato dal calcolo degli x_i integrando esattamente tutti i polinomi di grado m .

ESEMPI

1.

$$\int_{-1}^1 f(x) dx = I(f) , \text{ quindi } p(x) = 1 , x_0 = -1 , x_1 = 1$$

Voglio determinare la formula di quadratura in modo che il grado di precisione sia massimo

$$J_1(f) = a_0 f(-1) + a_1 f(1) \text{ ed } \varepsilon_1(f) = |I(f) - J_1(f)|$$

$$\text{Voglio che } \varepsilon_1(1) = 0, \text{ quindi } I(1) - J_1(1) = 0 \implies a_0 - a_0 - a_1 = 0$$

$$\text{Voglio poi che } \varepsilon_1(x) = 0 \implies I(x) - J_1(x) = 0 \implies 0 + a_0 - a_1 = 0$$

$$\begin{cases} a_0 + a_1 = 2 \\ a_0 - a_1 = 0 \end{cases} \implies a_0 = a_1 = 1$$

formula trapezoidale: $J_1(f) = \frac{b-a}{2} (f(a) + f(b))$

Per determinare il grado di precisione

$$\varepsilon_1(x^2) = \frac{2}{3} - 2 = -\frac{4}{3} \neq 0$$

2. Utilizziamo adesso la formula $J_2(f) = a_0 f(-1) + a_1 f(0) + a_2 f(1)$

$$f(x)=1 \implies \begin{cases} a_0 + a_1 + a_2 = 2 \\ a_0 = a_2 \end{cases} \implies a_1 = \frac{4}{3}$$

$$f(x)=x \implies \begin{cases} a_0 + a_1 + a_2 = 2 \\ a_0 = a_2 \end{cases} \implies a_0 = a_2 = 0$$

$$f(x)=x^2 \implies \begin{cases} a_0 + a_1 + a_2 = 2 \\ a_0 = a_2 \end{cases} \implies a_0 = \frac{1}{3} = a_2$$

$$J_2(f) = \frac{1}{3} f(-1) + \frac{4}{3} f(0) + \frac{1}{3} f(1)$$

$f(x)=x^3 \implies -\frac{1}{3} + \frac{1}{3} = 0$: integra esattamente anche x^3 , per cui la precisione è diverso 3

3. $f(x)=x^4 \implies \frac{1}{3} + \frac{1}{3} = \frac{2}{3} \neq \frac{2}{5}$, quindi la precisione è 3

$$\int_{-1}^1 f(x) dx = I(f) \quad J_2(f) = a_0 f(x_0) + a_1 f(x_1)$$

$$\begin{array}{l} f(x)=1 \\ f(x)=x \\ f(x)=x^2 \\ f(x)=x^3 \end{array} \quad \begin{cases} a_0 + a_1 = 2 \\ a_0 x_0 + a_1 x_1 = 0 \\ a_0 x_0^2 + a_1 x_1^2 = \frac{2}{3} \\ a_0 x_0^3 + a_1 x_1^3 = 0 \end{cases} \quad \begin{aligned} & a_0 x_0^3 + a_1 x_1 x_0^2 = 0 \quad \text{per } x_0 \neq 0 \\ & \downarrow \text{prendo } x_0 = 0 \end{aligned}$$

$$\begin{cases} a_0 + a_1 = 2 \\ a_1 x_1 = 0 \\ a_1 x_1^2 = 2/3 \end{cases} \quad \begin{aligned} & \text{Verificando la seconda non è possibile verificare la} \\ & \text{terza, quindi per } x_0 = 0 \text{ abbiamo un grado di} \\ & \text{precisione pari ad 1.} \end{aligned}$$

Riduciamo per $x_0 \neq 0$: $a_0x_1^3 - a_1x_0x_0^2 = 0 \rightarrow a_0x_1(x_1^2 - x_0^2) = 0$, quindi $a_1 = 0 \vee x_1 = 0$
 $\vee x_1 = x_0 \vee$
 $\vee x_1 = -x_0$

Analizziamo il caso di $a_1 = 0$: $a_0 = 2$

$a_0x_0 = 0 \rightarrow x_0 = 0$ e quindi il grado di precisione è nuovamente 1

Analizziamo il caso di $x_1 = x_0$: $a_0 + a_1 = 2$

$x_0(a_0 + a_1) = 0 \rightarrow x_0 = 0$ e quindi abbiamo di nuovo il caso precedente

Analizziamo il caso di $x_1 = -x_0$: $a_0 + a_1 = 2 \rightarrow a_0 - a_1 = 1$

$x_0(a_0 - a_1) = 0 \rightarrow a_0 - a_1 = 0 \rightarrow a_0 = a_1$

$$x_0^2(a_0 + a_1) = \frac{2}{3}$$

↓

$$x_0^2 = \frac{2}{3} \rightarrow x_0 = \pm \frac{\sqrt{3}}{3} : \text{grado di precisione diverso}$$

Ho quindi trovato una formula in cui i nodi sono $\pm \sqrt{3}/3$ che è data da $J_1(f) = f\left(-\frac{\sqrt{3}}{3}\right) + f\left(\frac{\sqrt{3}}{3}\right)$

Vediamo se il grado di precisione è > 3 : $f(x) = x^4 \rightarrow \frac{1}{9} + \frac{1}{9} \neq \frac{2}{5}$ e quindi $\Delta_1(x^4) \neq 0$, perciò la formula ha grado di precisione 3 (e non superiore).

formula di Simpson

Per approssimare $\int_a^b f(x) dx$: $J_2(f) = \frac{b-a}{6} \left(f(a) + 4f\left(\frac{b+a}{2}\right) + f(b) \right)$

In generale: se ho una formula e devo determinare per quali nodi devo scrivere tante equazioni quanti sono i precedenti, ottenendo un sistema non lineare con soluzione univocamente determinata. La soluzione del sistema fornisce una formula di quadratura con grado di precisione almeno $2n+1$.

teorema di Peano: Sia $f(x) \in C^{m+1}([a,b])$ e $J_m(f)$ di grado di precisione m , allora risulta $\Delta_m(f) = \frac{1}{m!} \int_a^b f^{(m+1)}(t) G(t) dt$ e quindi $|G(t)| = \Delta_m(S_m(x-t))$

$$S_m(x-t) = \begin{cases} (x-t)^m, & t < x \\ 0, & t \geq x \end{cases}$$

Nel caso in cui $|G(t)|$ non cambia segno in $[a,b]$, per la media integrale si ottiene $\Delta_m(f) = \frac{1}{m!} f^{(m+1)}(\theta) \int_a^b G(t) dt$

Se la formula ha grado di precisione m , posso calcolare l'errore per $m+1$ facendo $\Delta_m(x^{m+1}) = \frac{1}{m!} (m+1)! \int_a^b G(t) dt \rightarrow \int_a^b G(t) dt = \frac{1}{m+1} \Delta_m(x^{m+1})$

$$\Delta_m(f) = \frac{1}{(m+1)!} f^{(m+1)}(\theta) \Delta_m(x^{m+1})$$

ESEMPI

1. Riprendiamo il precedente esempio 1

$$\int_{-1}^1 f(x) dx = f(-1) + f(1) + \Delta_1(f)$$

La formula ha grado di precisione $m=1$ ed $\Delta_1(x^2) = \frac{2}{3} - 2 = -\frac{4}{3}$

Utilizzando la formula appena descritta: $\Delta_1(f) = -\frac{4}{3} \frac{f^{(2)}(\theta)}{2} = -\frac{2}{3} f^{(2)}(\theta)$

In generale, per la formula trapezoidale abbiamo $\mathcal{E}_1(f) = -\frac{(b-a)^3}{12} f''(\theta)$

$$2. \int_{-1}^1 f(x) dx$$

$$\mathcal{I}_2(f) = \frac{1}{3} (f(-1) + 4f(0) + f(1))$$

La formula di Simpson ha il nucleo di grado 4, perciò di segno costante, quindi è possibile applicare il teorema $\mathcal{E}_2(f) = 2 + \frac{-1}{3}(1+4+1) = 0 \rightarrow$ errore della formula di Simpson

$$\mathcal{E}_2(x) = 0 - \frac{1}{3}(-1+0+1) = 0$$

$$\mathcal{E}_2(f) = -\frac{(b-a)^5}{2880} f^{(4)}(\theta)$$

$$\mathcal{E}_2(x^2) = \frac{2}{3} - \frac{1}{3}(1+0+1) = 0$$

$$\mathcal{E}_2(x^3) = 0 - \frac{1}{3}(-1+0+1) = 0$$

$$\mathcal{E}_2(x^4) = \frac{2}{5} - \frac{1}{3}(1+0+1) = \frac{2}{5} - \frac{2}{3} = -\frac{4}{15}$$

o FORMULE di TIPO INTERPOLATORIO

$$I(pf) = \int_a^b p(x) f(x) dx \quad x_0, x_1, \dots, x_n \text{ nodi tali che } x_i \neq x_j \text{ se } i \neq j$$

$$f(x) = \mathcal{L}_n(x) + \mathcal{E}_n(x)$$

$$\mathcal{L}_n(x) = \sum_{i=0}^n l_i(x) f(x_i)$$

$$\begin{aligned} I(pf) &= \int_a^b p(x) (\mathcal{L}_n(x) + \mathcal{E}_n(x)) dx \\ &= \int_a^b p(x) d\mathcal{L}_n(x) + p(x) \mathcal{E}_n(x) dx \\ &= \int_a^b p(x) \mathcal{L}_n(x) dx + \int_a^b p(x) \mathcal{E}_n(x) dx \\ &= \int_a^b p(x) \sum_{i=0}^n l_i(x) f(x_i) dx + \int_a^b p(x) \mathcal{E}_n(x) dx \\ &= \sum_{i=0}^n f(x_i) \int_a^b p(x) l_i(x) dx + \int_a^b p(x) \mathcal{E}_n(x) dx \\ &\quad \text{ai nodi dipende da } f \\ &= \sum_{i=0}^n a_i f(x_i) + \int_a^b p(x) \mathcal{E}_n(x) dx \end{aligned}$$

$\xrightarrow{\mathcal{E}_n(f)}$ Formula di tipo interpolatorio

integrale esatto del polinomio di interpolazione relativo ai nodi scelti

Se i nodi delle formule sono prefissati a due a due distinti, impone che l'errore sia nullo fino a x_n dato che i nodi si traduce in $Vx=\mu$ ove

$$V = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_0 & x_1 & \dots & x_n \\ \dots & \dots & \dots & \dots \\ x_0^n & x_1^n & \dots & x_n^n \end{pmatrix}$$

$\alpha^T = (a_0, a_1, \dots, a_n)$ e $\mu^T = (m_0, m_1, \dots, m_n)$. Abbiamo che $\det(V) \neq 0$ e quindi α esiste ed è unico ed il grado di precisione è almeno n poiché $\mathcal{E}_n(f) = \mathcal{E}_n(x) = \dots = \mathcal{E}_n(x^n) = 0$

formule di Newton-Cotes

Abbiamo $p(x) = 1$ ed i nodi sono scelti in progressione aritmetica di ragione $h = (b-a)/n$, cioè $x_i = x_0 + ih$ con $x_0 = a$ e $x_n = b \rightarrow$ la formula trapezoidale è una formula di Newton-Cotes come anche la formula di Simpson.

Si dice punto ed i pesi sono definiti con la tecnica interpolatoria $a_i = I(l_i(x))$. Fino ad $m=7$ abbiamo pesi positivi, mentre per $m \geq 8$ abbiamo pesi sia negativi che positivi, il che non porta ad una tendenza al valore esatto all'infinito. Il nucleo di α sarà invece di segno costante su $[a, b]$.

errore: $\mathcal{E}_n(f) = \begin{cases} C_p h^{p+1} f^{(p)}(\theta), & p \text{ pari} \\ C_p h^{p+2} f^{(p+1)}(\theta), & p \text{ dispari} \end{cases}$ con C_p costante dipendente da p

Passando da un numero dispari al pari seguente il grado di precisione non aumenta, mentre aumenta facendo il vice-versa.

Le formule dette si dicono chiuse in quanto gli estremi di integrazione fanno parte dell'insieme dei nodi, riceverà si dicono aperte. A parità di numero di nodi, le formule aperte sono meno precise di quelle chiuse.

$$\int_a^b f(x) dx \quad \begin{array}{c} x_0 \\ \hline a & x_1 & x_2 & x_3 & \dots & x_L & b \end{array} \longrightarrow = \int_{x_0}^{x_1} f(x) dx + \int_{x_1}^{x_2} f(x) dx + \dots + \int_{x_{L-1}}^{x_L} f(x) dx =$$

$$= \sum_{j=1}^L \int_{x_{j-1}}^{x_j} f(x) dx \quad \text{e ponendo } x_j = a + j \frac{b-a}{L}$$

formule trapezoidale

$$\frac{b-a}{2L} (f(x_0) + 2f(x_1) + 2f(x_2) + 2f(x_3) + \dots + 2f(x_{L-1}) + f(x_L))$$

$$- \frac{1}{12} \left(\frac{b-a}{L} \right)^3 [f''(\Theta_1) + f''(\Theta_2) + f''(\Theta_3) + \dots + f''(\Theta_L)]$$

$$= \frac{b-a}{2L} (f(x_0) + 2 \sum_{j=1}^{L-1} f(x_j) + f(x_L)) \quad \rightarrow \text{Formula del trapezio}$$

$$- \frac{1}{12} \left(\frac{b-a}{L} \right)^3 L f''(\Theta)$$

$\downarrow - \frac{1}{12} \frac{(b-a)^3}{L^2} f''(\Theta)$: errore connesso. La precisione è sempre 1

Se f'' è limitata per $L \rightarrow \infty$ tende a 0.

COSTO COMPUTAZIONALE

Formule dei trapezii: va calcolata la funzione in tutti i punti ($m+1$ per m intervalli)

Formula di Simpson: va calcolata nei punti medi e negli estremi: $2L+1$ valutazioni

ESEMPIO

$$\int_0^1 \frac{1}{1+x} dx = \log(1+x) \Big|_0^1 = \log 2 \quad |x| < 10^{-2}$$

$$f(x) = \frac{1}{1+x} = (1+x)^{-1}$$

$$f'(x) = -(1+x)^{-2} \quad |\Delta_1^{(6)}(f)| = \frac{b-a}{12L^2} |f''(\Theta)|$$

$$f''(x) = 2(1+x)^{-3}$$

$$f'''(x) = -6(1+x)^{-4}$$

$$f^{(IV)}(x) = 24(1+x)^{-5}$$

$$\text{Posso trovare } M_2 \geq \sup_{x \in [a,b]} |f''(x)| \longrightarrow M_2 = 2 \longrightarrow \frac{b-a}{12L^2} \cdot 2 < \frac{1}{2} \cdot 10^{-2}$$

$$\frac{1}{6L^2} < \frac{1}{200}$$

$$200 < 6L^2 \longrightarrow L^2 > \frac{100}{3}$$

il primo L è 6

$$\text{Per il metodo dei trapezii: } M_4, \sup_{x \in [a,b]} |f^{(IV)}(x)| \longrightarrow M_4 = 24$$

$$|\Delta_2^{(6)}(f)| = \frac{(b-a)^5}{2880 L^4} |f^{(IV)}(\Theta)| < \frac{(b-a)^5}{2880 L^4} M_4$$

$$\frac{1}{2880 L^4} 24 < \frac{1}{2} \cdot 10^{-2} \longrightarrow L \geq 2: 2 \text{ intervalli}$$

Si vuol dimostrare che se $f(x)$ è sufficientemente regolare in $[a, b]$, l'errore della formula del trapezio ammette uno sviluppo in potenze pari di h . Ponendo

$$J_0^{(1)} = \frac{h}{2} \left[f(x_0) + 2 \sum_{i=1}^{r-1} f(x_i) + f(x_r) \right] \rightarrow \text{Se } f(x) \in C^{2r+2}([a, b]) \text{ si può scrivere}$$

$$I(f) = J_0^{(1)} + \alpha_1^{(1)} h^2 + \alpha_2^{(1)} h^4 + \dots + \alpha_r^{(1)} h^{2r} + O(h^{2r+2}). \quad \text{Prendendo un qualsiasi } q \text{ pari ad } \frac{h}{q}, \quad q \in \mathbb{N} \text{ otteniamo } I(f) = J_0^{(1)} + \alpha_1^{(1)} \left(\frac{h}{q}\right)^2 + \alpha_2^{(1)} \left(\frac{h}{q}\right)^4 + \dots + \alpha_r^{(1)} \left(\frac{h}{q}\right)^{2r} + O\left(\frac{h}{q}\right)^{2r+2}.$$

$$\text{Avremo quindi } \alpha_1^{(1)} h^2 = I(f) - J_0^{(1)} - \alpha_2^{(1)} h^4 - \dots - \alpha_r^{(1)} h^{2r} + O(h^{2r+1}), \quad \text{che sostituendo nella seconda relazione, } I(f) = J_0^{(1)} + \frac{1}{q^2} \left(I(f) - J_0^{(1)} - \alpha_2^{(1)} h^4 - \dots - \alpha_r^{(1)} h^{2r} \right) + \alpha_2^{(1)} \left(\frac{h}{q}\right)^4 + \dots + \alpha_r^{(1)} \left(\frac{h}{q}\right)^{2r} + O\left(\frac{h}{q}\right)^{2r+2}$$

$$+ O(h^{2r+2})$$

↓

$$I(f) - \frac{1}{q^2} I(f) = J_0^{(1)} - \frac{1}{q^2} J_0^{(1)} + \beta_2 h^4 + \beta_3 h^6 + \dots + \beta_r h^{2r} + O(h^{2r+2})$$

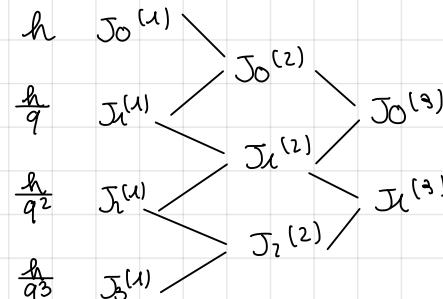
$$\frac{q^2 - 1}{q^2} I(f) = \frac{q^2 J_0^{(1)} - J_0^{(1)}}{q^2} + \beta_2 h^4 + \dots + \beta_r h^{2r} + O(h^{2r+2})$$

$$I(f) = \frac{q^2 J_0^{(1)} - J_0^{(1)}}{q^2 - 1} + \alpha_2^{(2)} h^4 + \dots + \alpha_r^{(r)} h^{2r} + O(h^{2r+2})$$

↓

Nuova approssimazione integrale: va a 0 come h^4 a differenza delle precedenti che andavano a 0 come h^2 . → Tecnica di extrapolazione

↓



• FORMULE di TIPO GAUSSIANO

polinomi ortogonali

Il spazio vettoriale polinomi algebrici a coefficienti reali e $[a, b]$ non necessariamente limitato. Per ogni coppia $p(x), q(x) \in \Pi$ si consideri il prodotto scalare $\langle p, q \rangle = \langle s, r \rangle = I(prs) = \int_a^b p(x)r(x)s(x)dx$. Si definisce Π^* l'insieme dei polinomi ortogonali cioè $\Pi^* = \{q_i(x) | \text{grad}(q_i) = i, \langle q_i, q_j \rangle = \delta_{ij}, i, j = 0, 1, 2, \dots\}$. Gli δ_{ij} sono costanti di normalizzazione ed i polinomi Π^* costituiscono una base per Π

↓ proprietà

$$a. \langle p, q_n \rangle = 0 \quad \forall p \in \Pi^{n-1}$$

$$b. q_n(x) \text{ ha } n \text{ zeri reali e distinti in } [a, b] \rightarrow \text{tutti interni}$$

↓

Per trovare i polinomi vi sono formule ricorrenti a 3 termini e sui intervalli simmetrici rispetto all'origine gli zeri dei polinomi sono simmetrici rispetto all'origine

↓

Finiti $p(x)$, $[a, b]$ ed n , scegliendo gli x_i come gli zeri del polinomio ortogonale di grado $n+1$, risulta univocamente determinata la formula di quadratura di grado di precisione almeno $2n+1$: $I(f) = \int_a^b p(x)f(x)dx \approx \sum_{i=0}^n a_i f(x_i) = J_n(f)$ detta formula di quadratura gaussiana

TEOREMA: Satti $[a, b]$ e $p(x)$, sia $f(x) \in C^{2n+2}([a, b])$ e $J_n(f)$ una formula di quadratura gaussiana, allora i nodi x_0, x_1, \dots, x_n sono gli zeri di $q_{n+1}(x) \in \Pi^*$, i a_i sono positivi e dati da $a_i = I(p_i)$, $i = 0, 1, \dots, n$ ed il grado di precisione è $2n+1$ essendo

$$\Delta n(f) = K_n \frac{f^{(2n+2)}(0)}{(2n+2)!}, \quad K_n > 0, \quad \Theta \in J_{a,b} \text{ l'errore}$$

Nodi e per i sono irrazionali e per gli integrali estesi ad $[a,b]$ si utilizzano polinomi definiti in $[-1,1]$ effettuando un cambio di variabile sull'integrale.

$$\int_a^b g(t) dt = \int_{-1}^1 p(x) g(x) dx \quad \text{ove } g(x) = \frac{b-a}{2p(x)} g\left(\frac{b-a}{2}x + \frac{b+a}{2}\right)$$

La positività dei nodi consente inoltre di dire che $\lim_{n \rightarrow \infty} J_n(f) = I(f)$ e quindi $\Delta n(f)$ tende a 0 anche se la funzione f non è derivabile

come trovare polinomi ortogonali

$$\int_{-1}^1 f(x) dx \quad g(x) = 1 \quad p(x) = 1$$

$$p(x) = ax + b \rightarrow \int_{-1}^1 p(x) x(x) dx = \langle x, 1 \rangle$$

$$\int_{-1}^1 q_0(x) p(x) dx = 0 \rightarrow \int_{-1}^1 (ax+b) dx = 0 : \quad 0 + 2b = 0 \rightarrow b=0, \quad \text{quindi } b=0 \text{ e } a \in \mathbb{R}$$

Prendo quindi $p(x) = x$.

Vogliamo trovare adesso $p_2(x) = ax^2 + bx + c$ ortogonale a $q_0(x)$ e $q_1(x)$

$$\int_{-1}^1 q_0(x) p_2(x) dx = 0 \rightarrow \int_{-1}^1 (ax^2 + bx + c) dx = 0 \quad (1)$$

$$\int_{-1}^1 p_1(x) p_2(x) dx = 0 \rightarrow \int_{-1}^1 x(ax^2 + bx + c) dx = 0 \quad (2)$$

$$(1) \frac{2}{3}a + 2c = 0 \rightarrow a + 3c = 0 \quad \text{e, scegliendo } a=1, \quad c = -\frac{1}{3}$$

$$(2) \frac{2}{3}b = 0 \rightarrow b = 0$$

Potrò quindi scrivere $p_2(x) = x^2 - \frac{1}{3}$. Se cerco dove si annulla, ovvero $p_2(x) = 0$

trovo $x = \pm \frac{\sqrt{3}}{3}$ che sono i nodi di $\int_{-1}^1 f(x) dx = a_0 f(x_0) + a_1 f(x_1) + \Delta n(f)$