

## LEGENDA

- Polinomio di Chebyshev
- Th. di Minimax
- Funzioni spline
- Approssimazione ai minimi quadrati
- Retta di regressione

### 1. NUMERI MACCHINA

- Rappresentazione numerica
- Distanza
- Underflow
- Overflow
- Troncamento
- Arrotondamento
- Operazioni tra numeri macchina
- Proprietà operazioni macchina

### 2. EQUAZIONI NON LINEARI

- Th. degli zeri
- Metodo di bisezione
- Metodo della falsa posizione
- Metodi di iterazione funzionale
- Interpretazione geometrica
- Convergenza locale
- Criterio di stop
- Ordine di convergenza
- Metodo di Newton-Raphson
- Th. convergenza
- Molteplicità di una radice
- Metodo della direzione costante
- Metodo della secante

### 3. SISTEMI DI EQUAZIONI LINEARI

- Regola di Cramer
- Metodo di eliminazione di Gauss
- Strategie di pivoting
- Pivoting parziale
- Pivoting totale
- Fattorizzazione LU
- Equivalenza tra Gauss e LU
- Condizionamento sistemi lineari

### 4. INTERPOLAZIONE

- Interpolazione polinomiale
- Polinomio interpolante di Lagrange
- Errori di interpolazione
- Th. di rappresentazione dell'errore di interpolazione
- Fenomeno di Runge
- Minimizzare il resto nel problema di interpolazione

### 5. QUADRATURA

- Formule di quadratura di tipo interpolatorio
- Grado di precisione
- Formula di Newton-Cotes
- Formula dei Trapezi
- Resto nella formula dei trapezi
- Th. della media generalizzato
- Formula di Simpson
- Formula del punto di mezzo
- Formule di quadratura composta
- Formula dei trapezi composta
- Formula di Simpson composta
- Formula del punto di mezzo composta

### 6. DERIVAZIONE NUMERICA

- Differenze in avanti
- Differenze all'indietro

### 7. EQ. DIFFERENZIALI

- Metodo di Eulero esplicito
- Metodo di Eulero implicito
- Metodo dei trapezi
- Metodo del midpoint
- Classificazione metodi numerici
- Studio della convergenza

## Cos'è il Calcolo Numerico?

È una branca delle matematiche che cerca di risolvere problemi matematici utilizzando un uso algoritmico. Il computer ha una memoria finita e questo porta ad approssimazioni di numeri e risultati.

Le approssimazioni sono sottesi di errori.

Gli errori possono essere

- condizionati (dovuti dall'esigenza di calcolare somme o differenze in un tempo finito)
- di approssimazione (dovuti all'esigenza di rappresentare un numero con infinite cifre in una memoria finita)

È possibile misurare l'errore commesso in 2 modi:

$$1) \text{ ERRORE ASSOCIAZIO} \quad e_A(x) = \| x - \tilde{x} \|$$

$$2) \text{ ERRORE RELATIVO} \quad e_R(x) = \frac{\| x - \tilde{x} \|}{\| x \|} \quad \text{con } x \neq 0$$

## Rappresentazione numerica

Rappresentare  $\mathbb{R}$  diventa impossibile. Viene utilizzata un insieme  $\mathbb{F}$  dei floating point numbers.

Si utilizza la rappresentazione mantissa cioè

- parte intera nulla
- prime cifre dopo la virgola non nulle.

$$x = (-1)^s \cdot 0. d_1 d_2 d_3 \dots d_t \cdot \beta^e \quad \text{con } d_1 \neq 0$$

$$\begin{matrix} \text{CIFRE MANTISSA} \\ \uparrow \\ \text{ESP. MASSIMO} \end{matrix}$$

$$\mathbb{F}(\beta, t, m, M) = \left\{ x \in \mathbb{R} : x = (-1)^s \cdot 0. d_1 d_2 \dots d_t \cdot \beta^e \right\} \cup \{0\}$$

$$\text{con } t \geq 1, \beta \geq 2, m > 0, d_1 \neq 0, d_i \in [0, \beta - 1],$$

$$e \in [-m, M]$$

1 bit per il segno ( $s$ )

$t$  bit per le mantisse ( $d_1, \dots, d_t$ )

$M + m + 1$  bit per l'esponente ( $E$ )

## NUMERI MACCHINA

1) numeri macchine hanno vere proporzioni.

2)  $F$  è discreto

3) I numeri rappresentabili sono solo una piccola parte di  $\mathbb{R}$

4) La distanza tra 2 nn. consecutivi non è costante.

## Distanze

$$x_i = \beta^E \cdot 0.d_1d_2 \dots d_{t-1}d_t$$

$$x_{i+1} = \beta^E \cdot 0.d_1d_2 \dots d_{t-1}\tilde{d}_t \quad \text{con } \tilde{d}_t = d_t + 1$$

$$\text{La distanza è } x_{i+1} - x_i = \beta^E \cdot \underbrace{0.000\dots 1}_{t-1} = \boxed{\beta^{E-t}}$$

Se  $E = -m$  allora la distanza è  $\beta^{-m-t}$

Se  $E = M$  la distanza vale  $\beta^{n-t}$

$$\text{Il rapporto fra le distanze è } \frac{\beta^{n-t}}{\beta^{-m-t}} = \beta^{n+m} \quad (\text{esame})$$

## Underflow ed overflow

Ci sono numeri troppo vicini allo zero per essere rappresentati e numeri troppo grandi.

$$\text{Se } x = \pm \beta^E \sum_{i=1}^n \text{di } \beta^{-i}$$

il numero è più piccolo del più piccolo elemento di  $F$  e si verifica il fenomeno dell' UNDERFLOW.

Se  $x < \min \mathbb{F}^+$  allora  $x$  sarà calcolato come 0

Se  $x = \pm \beta^E \sum_{i=1}^n d_i \beta^{-i}$  con  $n \leq t$  e  $E > M$

il numero  $i$  più grande del più grande el. d. IF è  
si verifica il fenomeno dell' OVERFLOW

Se  $x > \max \mathbb{F}$  allora  $x$  sarà considerato Inf

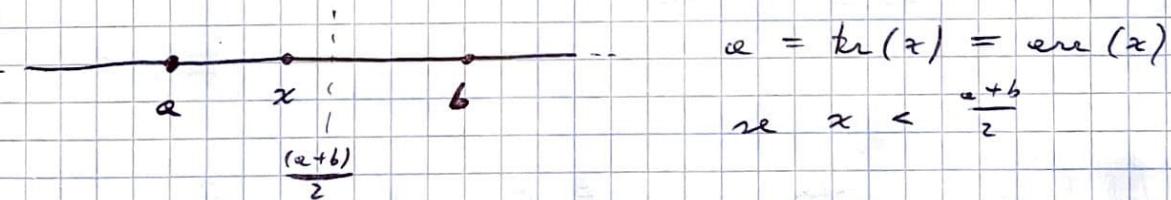
### Troncamento ed arrotondamento

Se  $x = \pm \beta^E \sum_{i=1}^n d_i \beta^{-i}$  con  $n > t$  dovranno essere

arrotondare o troncare il numero per poterlo rappresentare.

- Troncamento  $\tilde{x} = \text{tr}(x) = \pm \beta^E \sum_{i=1}^t d_i \beta^{-i}$
  - Arrotondamento  $\tilde{x} = \text{err}(x) = \pm \left( \beta^E \sum_{i=1}^t d_i \beta^{-i} + \underbrace{\sqrt{\beta^{t-t}}}_{\text{err}(x)} \right)$
- con  $\sqrt{\beta} = \begin{cases} 0 & d_{t+1} < \beta/2 \\ 1 & d_{t+1} \geq \beta/2 \end{cases}$

Se  $x \notin \mathbb{F}(\beta, t, m, M)$   $\exists a, b \in \mathbb{F} \Rightarrow a \leq x < b$



$$a = \text{tr}(x) = \text{err}(x)$$

$$\text{se } x < \frac{a+b}{2}$$

$$a = \text{tr}(x) \wedge b = \text{err}(x) \text{ se } x \geq \frac{a+b}{2} \quad \text{anche underflow}$$

L'arrotondimento può portare ad overflow - quando

$$x = \pm \beta^{E''} \sum_{i=1}^{t+1} d_i \beta^{-i} \text{ con } d_i = \beta - 1 \quad \forall i \in [1, t]$$

$$\text{e } d_{t+1} \geq \frac{\beta}{2} \quad \text{quindi } \text{err}(x) = \tilde{x} = \pm \beta'' = \pm 0.1 \beta^{\frac{M}{t+1}}$$

Il troncamento e l'arrotondimento producono ovviamente degli errori.

Calcolo con l'errore di trasformazione

$$|\text{tr}(x) - x| \leq b - \epsilon = \underline{\beta^{E-t}} \geq e_1(x)$$

$$|x| = \beta^E \cdot 0.2, 0.3, \dots, 0.7 \geq \beta^E \cdot 0.1 = \beta^{E-1} \Rightarrow$$

$$\Rightarrow \frac{1}{|x|} \leq \beta^{1-E} \Rightarrow e_n(x) = \frac{|x - \text{tr}(x)|}{|x|} \leq \frac{\beta^{E-t}}{\beta^{n+E-1}} = \underline{\beta^{t-n}}$$

$$|\text{arr}(x) - x| \leq \frac{1}{2}(b - \epsilon) = \beta^{E-t} \cdot \frac{1}{2} \geq e_1(x)$$

$$\frac{1}{2} e_n(x) = \frac{|\text{arr}(x) - x|}{|x|} = \frac{1}{2} \beta^{E-t} \cdot \beta^{1-E} = \frac{1}{2} \beta^{1-t}$$

$\forall u < \beta^{1-t}$  sono letti zero macchine per il moltiplicatore

( $\forall u < \frac{1}{2} \beta^{1-t}$  sono letti zero macchine per l'addizione)

Dunque  $\varepsilon_x = \frac{\tilde{x} - x}{x}$  con  $|\varepsilon_x| \leq u \Rightarrow \tilde{x} = x(1 + \varepsilon_x)$

$\tilde{x} = x(1 + \varepsilon_x)$  è la ~~approssimazione~~ <sup>relazione</sup> tra  $x \in \mathbb{R}$  ed

$\tilde{x} \in \mathbb{F} \Rightarrow \tilde{x} = \text{fl}(x)$  ed è una rappresentazione in  $\mathbb{F}$

Operazioni con i numeri macchine (operazioni macchine)

\* ADDIZIONE

$$\text{Se } \tilde{x}, \tilde{y} \in \mathbb{F}(\beta, t, m, M) \Rightarrow x + y \in \mathbb{F} \quad \forall \tilde{x}, \tilde{y} \in \mathbb{F}$$

$$\text{Siano } x, y \in \mathbb{F} \quad x \odot y = (x \cdot y)(1 + \varepsilon) \text{ con } |\varepsilon| \leq u \quad (1)$$

$$\text{e } x \cdot y = x \odot y \Leftrightarrow \varepsilon = 0 \Rightarrow x \cdot y = \text{fl}(x \cdot y)$$

$x \odot y = \text{tr}(x \cdot y)$  e  $x \odot y = \text{arr}(x \cdot y)$  soddisfano le relazioni (1) e quindi sono formali ~~nataturali~~ macchine operazioni.

Ogni operazione ha un metodo per essere effettuate in maniera corretta.

• ADDIZIONE E SOTTRAZIONE (SOMMA ALGEBRICA)

Sei  $x, y \in \mathbb{F}^t(\beta, t, m, M)$

$$x \oplus y = \text{err} \left( \sum_{i=1}^{\max(E_x, E_y)} (d_i x + d_i y) \beta^{-i} \right)$$

3 passaggi sono

1. Prendere l'esponente maggiore e modificare il num.  
con esponente minore affinché  $E_y = E_x$
2. Sommare le mantisse
3. Normalizzare le somme tra mantisse
4. Rimuovere / troncare il risultato.

E<sub>x</sub>  $x = 0.78546 \cdot 10^2, y = 0.61332 \cdot 10^{-1}$

$$x \oplus y = 0.78607 \cdot 10^2$$

$$\tilde{x} \tilde{y} = 0.00061 \cdot 10^2$$

$$\text{mantissa}(x) + \text{mantissa}(\tilde{y}) = 0.78607 \text{ // } 11$$

E<sub>y</sub> Sei  $x \approx y \wedge, x = 0.75868531 \cdot 10^2$

$$y = 0.75868100 \cdot 10^2$$

$$\tilde{x} = \text{fl}(x) = 0.75869 \cdot 10^2 \quad \tilde{y} = \text{fl}(y) = 0.75868 \cdot 10^2$$

$$\text{erf}(x)(\tilde{x} - \tilde{y}) = 0.00001 \cdot 10^2 = 10.1 \cdot 10^{-2}$$

$$x - y = 0.00000431 \cdot 10^2 = 4.31 \cdot 10^{-4} = 0.431 \cdot 10^{-3}$$

$$\text{er}(x-y) = \frac{|\text{fl}(x) - (x-y)|}{|x-y|} = \frac{(0.1 - 0.0431) \cdot 10^{-2}}{0.0431 \cdot 10^{-2}} =$$

$\approx 1.32$  (molto alto) → questo è il caso delle cancellazioni numeriche si copre significativa quanto  $x \approx y$

La cancellazione numerica di cifre significative rappresenta un caso di problema nel calcolo.

- PRODOTTO

$$x \odot y = \text{mantissa}(x) \cdot \text{mantissa}(y) \cdot \beta^{E_x + E_y}$$

3 passaggi sono:

1. eseguire il prodotto fra le mantisse
2. sommare / trarre le t cifre dopo aver moltiplicato
3. moltiplicare per  $\beta^{E_x + E_y}$  (sommare gli esponenti)

Esempio:  $x = 0.11111 \cdot 10^3, y = 0.52521 \cdot 10^2$

$$\text{mantissa}(x) \cdot \text{mantissa}(y) = 0.05835608 =$$

$$= 0.58356 \cdot 10^{-1}$$

$$0.58356 \cdot 10^{3+2-1} = 0.58356 \cdot 10^4 = x \odot y$$

- DIVISIONE

3 passaggi sono:

1. Scegli l'esponente di  $x$  affinché  $\text{mantissa}(x) < \text{mantissa}(y)$
2. Dividi le mantisse
3. Normalizza e sottrai / tratti il numero
4. Sottrai gli esponenti

Esempio:  $x = 0.11100 \cdot 10^5, y = 0.11000 \cdot 10^2$

$$\tilde{x} = 0.01110 \cdot 10^6$$

$$\text{mant}(\tilde{x}) / \text{mant}(y) = 0.11000$$

$$x \oslash y = 0.11000 \cdot 10^{6-2+0} = 0.11000 \cdot 10^4$$

## Proprietà operazioni matricie

1. L'insieme  $\mathbb{F}(\beta, \epsilon, m, M)$  non è chiuso rispetto alle operazioni matricie (e come nell'overflow)
2. L'elemento neutro  $\forall x$  non è unico (molti molti peccati vengono considerati / mettiti come 0)
3. L'elemento neutro per il prodotto non è unico (molti molti vicini all'1)
4. Non ha le proprietà associative di somma e prodotto  $(a + b) + c \neq a + (b + c)$
5. Non ha le proprietà distributiva delle somme rispetto al prodotto.  $a(b + c) \neq ab + ac$

## Conoscimento dei problemi

Un problema si dice ben conosciuto se ne ha pochi errori nei dati e risposte poco avviate dai risultati.

Un problema è mal conosciuto se non è ben conosciuto.

Il conoscerneamento è una caratteristica intrinseca del problema e non dipende dal metodo di risoluzione adottato.

Se si incontra un problema mal conosciuto bisogna ristudiarlo con un meglio conoscerneamento tramite tecniche di spaccosconoscimento.

## EQ. NON LINEARI

Ricerca degli zeri

Sia  $f: [a, b] \rightarrow \mathbb{R}$ ,  $f \in C([a, b])$  e  $f(a)f(b) < 0$   
 $\Rightarrow \exists c \in (a, b) \text{ s.t. } f(c) = 0$

Metodo di Bisezione

Consiste nel costituire una successione di intervalli  $\{I_k\}_{k=0}^{\infty}$  tali

1.  $I_0 = [a_0, b_0]$

2.  $I_{n+1} \subset I_n$

3.  $a \in I_n \quad \forall n \geq 0$

4.  $|I_k| \rightarrow 0$  quando  $k \rightarrow +\infty$

$I_0 = [a_0, b_0] \quad c_1 = \frac{a_0 + b_0}{2}$

Se  $f(a)f(c) < 0 \Rightarrow I_1 = [a_1, b_1] \quad a_1 = a_0$   
 $b_1 = c_1$

Se  $f(c)f(b) < 0 \Rightarrow I_1 = [a_1, b_1] \quad a_1 = c_1$   
 $b_1 = b_0$

$c_{k+1} = \frac{a_k + b_k}{2}$

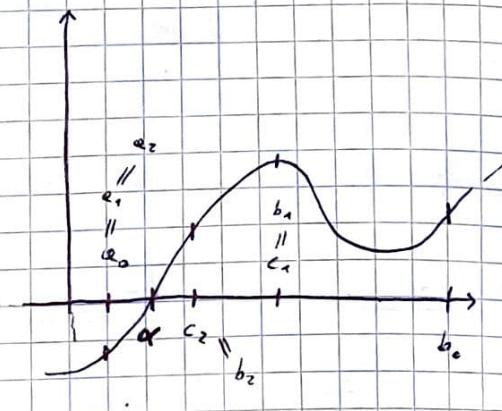
Se  $f(c_k) = 0 \Rightarrow c_k = \alpha \xrightarrow{\text{zero di } f}$

Criterio di Stop

Ora come si evita di sovradimensionare la condizione  $f(c_k) = 0$   
non viene quasi mai verificata.

Bisogna stabilire un criterio di stop.

Esempio  $b_k - a_k \leq \varepsilon$  con  $\varepsilon$  tolleranza prefissata



Se viene scelta un eventuale tolleranza preferita  
è possibile dedurne al numero di passi:

$$b_k - e_k \leq \varepsilon \Rightarrow \frac{b_{k-1} - e_{k-1}}{z} = \frac{b_0 - e_0}{z^k} \leq \varepsilon \Rightarrow$$

$$\Rightarrow \varepsilon \cdot z^k \geq b_0 - e_0 \Rightarrow$$

$$\Rightarrow k \geq \log_z \left( \frac{b_0 - e_0}{\varepsilon} \right)$$

Se  $0 \notin [a, b]$  si può stabilire un altro criterio di stop:

$$\frac{b_k - e_k}{\min(|e_k|, |b_k|)} \leq \varepsilon \quad \leftarrow \text{errore relativo } n \leq \varepsilon$$

$$b_k - e_k \leq \varepsilon \quad \leftarrow \text{errore assoluto} \leq \varepsilon$$

Un altro criterio potrebbe essere  $|f(c_n)| \leq \varepsilon$

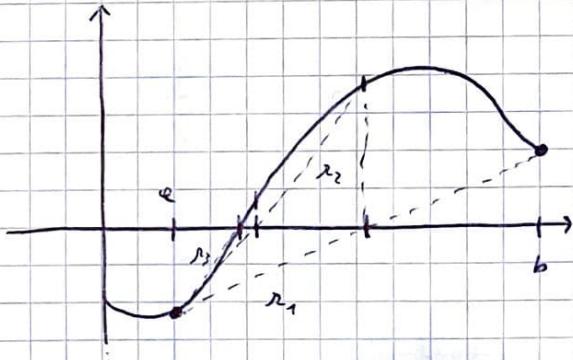
### Metodo della Falsa Posizione

$$y = f(a) + \frac{f(b) - f(a)}{b - a} (z - a)$$

$$\Rightarrow y = 0 \Leftrightarrow$$

$$c_1 = a - \frac{b - a}{f(b) - f(a)}$$

$$I_{n+1} = [a_1, b_1] = \begin{cases} a_1 = a_0 & b_1 = c_1 & f(a_0) f(c_1) < 0 \\ a_1 = c_1 & b_1 = b_0 & f(a_0) f(c_1) > 0 \end{cases}$$



Ugualmente al metodo degli zeri ma con un diverso  
metodo per il calcolo di  $c_n$

Entrambi questi metodi hanno vantaggi e svantaggi simili  
ma ben..

Vantaggi:

- richiedono esclusivamente le conoscenze di  $f$ ;
- possono ricevere buone approssimazioni di  $\alpha$  che possono essere usate come starting point per metodi iterativi.

Svantaggi:

- la convergenza è lenta (si guadagna una precisione in base a  $\frac{1}{2}$  per ogni passo)
- la velocità di convergenza non dipende da  $f$

### Metodi di Intersezione Funzionale

Sono metodi del tipo  $x_{k+1} = g(x_k)$  per  $k = 0, 1, 2, \dots$   
dove  $x_0$  è un starting point

Si può fare l'operazione per trovare  $\alpha$  se  $f(\alpha) = 0$  è chiamata  
per fare  $g(x) = f(x) - \alpha$ .  $g(x)$  è detta f. interattiva.

Per trovare gli zeri di una funzione è possibile definire  
una funzione interattiva come segue:

$$g(x) = x - \frac{f(x)}{h(x)} \quad \text{con } h(x) : [a, b] \rightarrow \mathbb{R} \setminus \{0\}$$

e  $\alpha$  è il punto per cui  $g(\alpha) = 0$  è la zero della  $f(x)$

$$\text{Avvicinamente } g(\alpha) = \alpha - \frac{f(\alpha)}{h(\alpha)} = \alpha \Leftrightarrow f(\alpha) = 0.$$

Rh. Sei  $g \in C([a, b])$  e la successione  $\{x_k\}$  tali  
 $x_{k+1} = g(x_k)$  sia contenuta in  $[a, b]$ . Se  $\{x_k\}$  converge  
per  $k \rightarrow \infty \Rightarrow \{x_k\} \rightarrow \alpha = g(\alpha)$

Dam  $\alpha = \lim_{k \rightarrow \infty} x_{k+1} = \lim_{k \rightarrow \infty} g(x_k) = g\left(\underbrace{\lim_{k \rightarrow \infty} x_k}_{\alpha}\right)$

||

$$\lim_{k \rightarrow \infty} x_k$$

Questa è la condizione necessaria ma non è sufficiente.

Th (Cond. suff.)

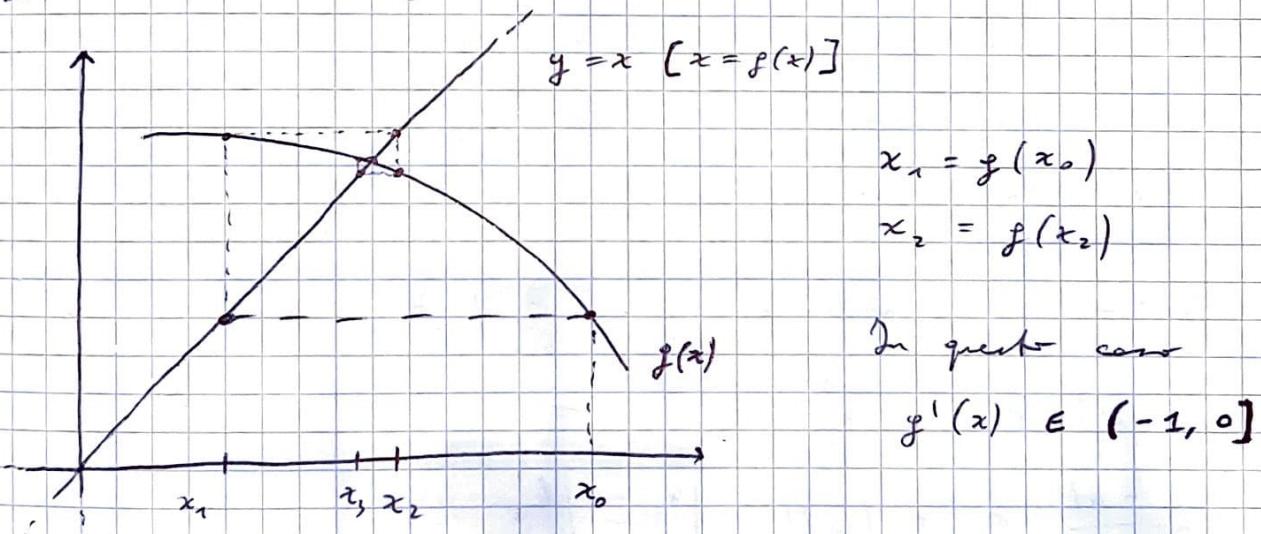
Se  $\alpha$  p.t. fissa s.t.  $f(x) = g \in C^1([a-\rho, a+\rho])$   
per  $\rho > 0$

se  $|f'(x)| < 1 \quad \forall x \in [a-\rho, a+\rho] \Rightarrow$

- $\Rightarrow$  1) se  $x_0 \in [a-\rho, a+\rho] \Rightarrow x_k \in [a-\rho, a+\rho] \quad \forall k$   
2)  $\{x_k\} \rightarrow \alpha$   
3)  $\alpha$  è l'unico p.t. fissa s.t.  $f(x) = g(x)$  in  $[a-\rho, a+\rho]$

Questa condizione ( $|f'(x)| < 1$ ) garantisce che c'è  
un unico p.t. fissa, che la successione converge in +  
e che la successione rimane sempre nell'intervallo  
 $[a-\rho, a+\rho]$  per qualche  $\rho > 0$

Interpretazione geometrica



Convergenza locale Un metodo iterativo è localmente  
convergente se per  $\bar{x} = f(\bar{x}) \in [a, b]$  esiste  $\exists [a, b] \ni x_0 \in$   
 $[a, b] : \{x_k\} \rightarrow \bar{x}$

È possibile che spostando anche di poco  $x_0$  si trovi  
un p.t. fissa severo (o che non si trovi)

## Criterio di stop

$e_k(\alpha) = |\alpha - x_k| < \varepsilon \Leftrightarrow$  non conoscendo  $\alpha$  non posso calcolare  $e_k$  e quindi questo criterio di stop non è valido.

Un criterio valido potrebbe essere  $|x_{k+1} - x_k| < \varepsilon$  o

$$\frac{|x_{k+1} - x_k|}{\max(|x_{k+1}|, |x_k|)} \leq \varepsilon \quad \text{o al posto} \quad |f(x_k)| \leq \varepsilon$$

## Ordine di convergenza

Se  $\{x_n\}_{k=0}^{+\infty} \rightarrow \{\alpha\}$  ed  $x_k \neq \alpha \forall k$ .

$$\text{Se } \exists p \geq 1 \text{ s.t. } \lim_{k \rightarrow +\infty} \frac{|x_{k+1} - \alpha|}{|x_k - \alpha|^p} = \gamma \text{ con } \begin{cases} 0 \leq \gamma \leq 1 & p=1 \\ \gamma > 0 & p>1 \end{cases}$$

$\Rightarrow$  la successione ha ordine di convergenza  $p$  e  $\gamma$  prende il nome di costante assoluta di convergenza.

Se  $p=1$  e  $\gamma \in (0, 1)$  la conv. è lineare;

Se  $p>1$  la conv. è superlineare;

Se  $p=2$  la conv. è quadratica;

Se  $p=3$  la conv. è cubica.

La relazione (2) implica che  $\exists \beta \approx \gamma \Rightarrow$

$$|x_{k+1} - \alpha| \leq \beta |x_k - \alpha|^p \Rightarrow$$

$$\Rightarrow \frac{|x_{k+1} - \alpha|}{|\alpha|} \leq \beta |\alpha|^{p-1} \left| \frac{x_k - \alpha}{\alpha} \right|^p$$

Queste 2 espressioni indicano quanto si riducono per ogni passo gli errori assoluti e relativi.

$$e_{k+1}^A(\alpha) \leq \beta e_k^A(\alpha)^p$$

$$e_{k+1}^R(\alpha) \leq \beta |\alpha|^{p-1} e_k^R(\alpha)^p$$

Rh See  $\{x_k\}_{k=0}^{+\infty}$  con  $x_{k+1} = f(x_k)$ ,  $\{x_n\} \rightarrow \alpha = f(\alpha)$   
 $\exists f(x) \in C^{\infty}(I(\alpha, \rho))$

Le successione  $\{x_n\}$  ha ordine di convergenza  $n \geq 1 \iff$

$$\Leftrightarrow f'(\alpha) = f''(\alpha) = \dots = f^{(n-1)}(\alpha) = 0 \wedge f^{(n)}(\alpha) \neq 0$$

$$\wedge \gamma = \frac{|f^{(n)}(\alpha)|}{n!}$$

### Metodo di Newton - Raphson

Anche chiamato metodo delle tangenti. Unica ipotesi  
 per cui può essere applicato è  $f \in C^1([a, b])$

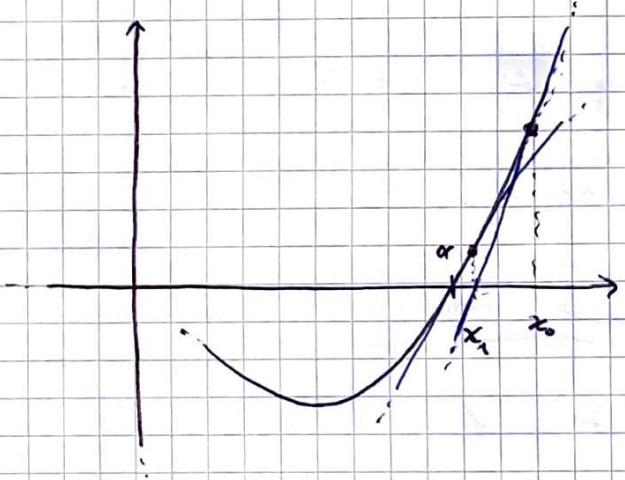
$$t_k : y - f(x_k) = f'(x_k)(x - x_k)$$

$$\text{Ponendo } y = 0$$

$$-f(x_k) = f'(x_k)(x - x_k)$$

$$\Rightarrow x = x_k - \frac{f(x_k)}{f'(x_k)}$$

$$\quad // \quad x_{k+1}$$



Che ha la forma molto simile a  $g(x) = x_k - \frac{f(x_k)}{h(x_k)}$

$$\text{dove } x_{k+1} = g(x_k), h(x_k) = f'(x_k)$$

Nel metodo di Newton - Raphson  $h(x) = f'(x)$

Rh. convergenza Sei  $f \in C^3([a, b]) \wedge f'(x) \neq 0$

$\forall x \in [a, b] \subset [a, b] \ni \alpha \Rightarrow$

$$\Rightarrow 1) \exists [\alpha - \rho, \alpha + \rho] \ni \forall x_0 \in [\alpha - \rho, \alpha + \rho]$$

$\{x_n\} \rightarrow \alpha \wedge 2) p \geq 2$  ( $p$  ordine di convergenza)

$$\underline{\text{Osservazione}} \quad g'(x) = 1 - \frac{[f'(x)]^2 - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2}$$

Visto che  $f'(\alpha) \neq 0$

$$g'(\alpha) = \frac{f(\alpha) \cdot f''(\alpha)}{[f'(\alpha)]^2} = 0$$

quindi  $g^{(1)}(\alpha) = 0 \Rightarrow \mu \geq 2$

$$\exists |g'(x)| < 1 \quad \forall x \in [\alpha - p, \alpha + p]$$

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

$$g''(\alpha) = \frac{f'(\alpha)f''(\alpha) \cdot [f'(\alpha)]^2}{[f'(\alpha)]^4} = \frac{f''(\alpha)}{f'(\alpha)} = \text{non zero}$$

quindi  $\mu = 2$  se  $f''(\alpha) \neq 0$ ,  $\mu \geq 3$  se  $f''(\alpha) = 0$

$$\text{Se } \mu = 2 \quad \gamma = \frac{1}{2!} \left| \frac{f''(\alpha)}{f'(\alpha)} \right|$$

Multiplicità di una radice

Se  $f \in C^r([a, b])$  per  $r \in \mathbb{N} \Rightarrow r > 0$

Una radice si dice di multiplicità  $r$  se

$$\lim_{x \rightarrow \alpha} \frac{f(x)}{(x - \alpha)^r} = \gamma \quad \text{con } \gamma \neq 0 \wedge \gamma \neq \pm\infty$$

Se  $\alpha$  è una radice di multiplicità  $r \Rightarrow$

$$\Rightarrow f(\alpha) = f'(\alpha) = \dots = f^{(r-1)}(\alpha) = 0 \wedge f^{(r)}(\alpha) \neq 0$$

Note Se  $r > 1 \Rightarrow f'(\alpha) = 0 \wedge \mu = 1$

$$f(x) = q(x)(x - \alpha)^r \quad \text{con } q(\alpha) \neq 0$$

$$g(x) = x - \frac{q(x)(x - \alpha)^r}{rq(x) + q'(x)(x - \alpha)} \Rightarrow$$

$$\Rightarrow g'(\alpha) = 1 - \frac{1}{n}$$

Se  $n > 1$   $|g'(x)| < 1$  quindi il metodo è convergente  
ma l'ordine di convergenza è 1. ( $g'(x) \neq 0$ )

Se si conosce la multiplicità delle radice si può  
preferire il metodo di Newton - Raphson.

$$g(x_k) = x_k - n \frac{f(x_k)}{f'(x_k)} \quad \text{con } k = 0, 1, 2, \dots$$

In questo modo  $g'(\alpha) = 1 - n \frac{1}{n} = 0$  quindi  
l'ordine di convergenza rimane  $n \geq 2$

### Metodo delle derivate costante

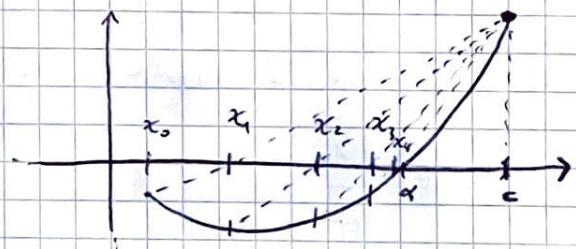
Se le derivate prime di  $f$  si mantengono sensibilmente  
costante allora  $M = f'(x_k)$  quindi

$$g(x_k) = x_k - \frac{f(x_k)}{M} \quad \leftarrow \text{metodo delle der. cost.}$$

Il metodo è convergente se  $|g'(\alpha)| = |1 - \frac{f'(\alpha)}{M}| < 1$

quindi  $\frac{f'(\alpha)}{M} > 0$  ( $f'(\alpha)$  ed  $M$  devono avere stesso  
segno)

### Metodo delle secante



$$x_{k+1} = x_k - \frac{x_k - c}{f(x_k) - f(c)} \quad \text{con } c \in [x_k, x_{k+1}]$$

$$\Rightarrow g(x) = x - \frac{x - c}{f(x) - f(c)}$$

# SISTEMI DI EQ. LINEARI

VETTORI  
SOLUZIONE  
↓

Risolvere un sistema lineare significa trovare un vettore  $\bar{x} \in \mathbb{R}^n$

$A\bar{x} = \bar{b}$  con  $A$  matrice dei coeff. e  $\bar{b}$  vettore dei termini noti.

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + \dots = b_2 \\ \dots \dots \dots \dots \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right.$$

Determinante  $|A| = a_{11}$  per  $A \in \mathbb{R}^{n \times n}$

$$|A| = \sum_{j=1}^n a_{ij} (-1)^{i+j} |A_{ij}| \quad \text{Regole di Laplace}$$

per  $A_{ij}$  v. complementare.

Ch. Binet

$$\text{Proprietà } |I| = 1 \quad |A^\top| = |A| \quad |A \cdot B| = |A| \cdot |B|$$

$$|\alpha A| = \alpha^n |A| \quad \text{con } A \in \mathbb{R}^{n \times n} \quad \alpha \in \mathbb{R}$$

$|A| = 0$  se almeno una riga (colonna) è comb. lineare di altre righe (colonne)

Una matrice è detta singolare se  $|A| = 0$

$$|A^{-1}| = \frac{1}{|A|}, \quad A^{-1} \cdot A = A \cdot A^{-1} = I_n \quad \text{con } A^{-1} \text{ matrice inversa di } A$$

Inversa di  $A$

Osserva una matrice  $n \times n$  si rendere non sono presenti soluzioni (c'è un eq. deg. delle altre e quindi n il numero di incognite) i maggiori di n eq. ind.

## Regole di Cramer

$$x_i = \frac{|A_i|}{|A|} \quad \text{con } i = 1 \dots n$$

dove  $A_i$  è la matrice che si ottiene sostituendo  
le colonne  $i$ -esime con le colonne dei term. noti  $b$

Per calcolare  $\bar{x}$  fornito da  $n$  elementi è necessario  
calcolare  $x_i$   $\forall i \in [1, n]_N$  e quindi calcolare  
 $|A_i|$   $\forall i \in [1, n]_N$  e  $|A|$  (calcolare, quindi,  
 $n+1$  determinanti)

Quel<sup>4</sup> è il num. di operazioni da fare per calcolare  
tutte le  $x_i$  e quindi, risolvere il sistema?

Indichiamo con  $f(n)$  il num. di operazioni da svolgere  
per calcolare  $|A|$  con  $A \in \mathbb{R}^{n \times n}$  Ese  $f(2) = 3$

$$f(1) = 0 \quad f(2) = 3 \quad \text{e generalmente L'Hopital seppure}$$

che  $f(n) = \underbrace{n f(n-1)}_{\substack{\text{n det di matrici} \\ \text{di dim } n-1}} + \underbrace{2n-1}_{\substack{\text{n moltip.} \\ \text{n-1 somme alg.}}}$

Per semplificare poniamo  $f(n) \approx n f(n-1)$  ed allora

$$f(n) = n f(n-1) = n (n-1) f(n-2) = n (n-1) \dots 3 f(2) =$$

$$= \frac{n!}{2} \cdot f(2) = \frac{3}{2} n!$$

Se per calcolare  $|A|$  con  $A \in \mathbb{R}^{n \times n}$  ci vogliono  
 $\frac{3}{2} n!$  operazioni, per trovare  $\bar{x}$  dobbiamo calcolare  
 $n+1$  det. quindi  $\frac{3}{2} n! (n+1) = \frac{3}{2} (n+1)!$

Un numero enorme che non ci consente di calcolare questi val.

Notiamo però che risolvere sistemi tridiagonali  
è decisamente più semplice ed efficiente.

Osservi  $A \in \mathbb{R}^{n \times n}$  triangolare superiore

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{nn}x_n = b_n \end{array} \right.$$

$$Questa mtr. \quad x_n = \frac{b_n}{a_{nn}}$$

L'eq. precedente ha come sol.  $x_n$  come incognita  
e così via.

$$\left\{ \begin{array}{l} x_n = \frac{b_n}{a_{nn}} \\ x_i = \frac{b_i - \sum_{j=i+1}^n a_{ij}x_j}{a_{ii}} \quad i = n-1 \dots 1 \end{array} \right.$$

Questo prende il nome di metodo di sostituzione all'indietro.

Se  $A \in \mathbb{R}^{n \times n}$  è triangolare inferiore

$$\left\{ \begin{array}{l} x_1 = \frac{b_1}{a_{11}} \\ x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j}{a_{ii}} \quad i = 2 \dots n \end{array} \right.$$

Questo, invece, si chiama sostituzione avanti.

Vediamo il costo computazionale di questi 2 metodi.  
(considerandoli equivalenti)

Indico con  $C(n)$  il nr. di operazioni necessarie

$$C(n) = \sum_{i=1}^n (2i-1) = 2 \sum_{i=1}^n i - \sum_{i=1}^n 1 = \frac{2n(n+1)}{2} - n = n^2 + n - n = n^2$$

Rovare  $\bar{x}$  in una matrice traspolare i decessante non c'è (n! < n! per n grandi).

Si potrebbe trovare un metodo non troppo costoso per trasformare la matrice A in una matrice traspolare ed applicare il metodo di sost. in avanti (ell' metodo).

### Metodo di Eliminazione di Gauss

Def I sistemi non eq. se ammettono le stesse soluzioni.

Se noncens in eq. del sistema ed in' altre al sistema sembra equivalente.

Tutte le stesse se anche queste non sono tutte le cont. lineare di n'altre regole del sistema ed in' altre regole. E' in questo principio che si basa il metodo di eliminazione di Gauss.

$$\text{Bisogna } A^{(1)} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & \dots & e_{1n} \\ e_{21} & e_{22} & e_{23} & \dots & e_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e_{m1} & e_{m2} & e_{m3} & \dots & e_{mn} \end{bmatrix} \quad b^{(1)} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Perro ed estremare l'el.  $e_{21}^{(1)}$ .

$$l_{21} = -\frac{e_{21}}{e_{11}}$$

è il moltiplicatore che consente alle prime regole  $i \neq 1$  di eliminare l'elemento  $e_{ij}$  alle prime regole  $i \neq 1$ .

Calcolando questi  $l_{21}, l_{31}, \dots$  ha così le nuove A<sup>(1)</sup> e b<sup>(1)</sup>

$$A^{(2)} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & \dots & e_{1n} \\ 0 & e_{22}^{(1)} & \dots & \dots & e_{2n}^{(1)} \\ 0 & e_{32}^{(1)} & \dots & \dots & e_{3n}^{(1)} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & e_{m2}^{(1)} & \dots & \dots & e_{mn}^{(1)} \end{bmatrix}, \quad b^{(2)} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \vdots \\ b_m^{(1)} \end{bmatrix}$$

Calcolando poi  $l_{32}, l_{42}, \dots, l_{n2}$  per ottenere A<sup>(2)</sup>.

fanché non troverò  $A^{(n)}$  che è incongruo.

Cerchiamo quel  $i$  la formula che descrive questo algoritmo per calcolare il vettore computazionale.

$$(3) \quad \alpha_{11}^{(1)} x_1 + \sum_{j=2}^n \alpha_{ij}^{(1)} x_j = b_i^{(1)} \quad \text{nel caso delle prime regole}$$

$$(4) \quad \alpha_{ii}^{(1)} x_i + \sum_{j=2}^n \alpha_{ij}^{(1)} x_j = b_i^{(1)} \quad \text{per } i = 2 \dots n$$

Moltiplico (3) per  $\lambda_{i1} = -\frac{\alpha_{i1}^{(1)}}{\alpha_{11}^{(1)}}$  ed ottengo

$$-\alpha_{i1}^{(1)} x_1 + \sum_{j=2}^n \left( -\alpha_{ij}^{(1)} \frac{\alpha_{i1}^{(1)}}{\alpha_{11}^{(1)}} \right) x_j = b_i^{(1)} \cdot \left( -\frac{\alpha_{i1}^{(1)}}{\alpha_{11}^{(1)}} \right)$$

Sommare queste e (4) ecco ottengo  $n-1$  eq.

$$\sum_{j=2}^n \left( \alpha_{ij} - \alpha_{ij} \frac{\alpha_{i1}^{(1)}}{\alpha_{11}^{(1)}} \right) x_j = b_i - b_i \frac{\alpha_{i1}^{(1)}}{\alpha_{11}^{(1)}} \quad \text{per } i = 2 \dots n$$

$$\Rightarrow \begin{cases} \alpha_{11}^{(1)} x_1 + \sum_{j=2}^n \alpha_{ij}^{(1)} x_j = b_i^{(1)} & \leftarrow \text{rimane inviolata} \\ \sum_{j=2}^n \alpha_{ij}^{(1)} x_j = b_i^{(1)} & \text{per } i = 2 \dots n \end{cases}$$

$$\text{con } \alpha_{ij}^{(2)} = \alpha_{ij}^{(1)} - \frac{\alpha_{i1}^{(1)}}{\alpha_{11}^{(1)}} \alpha_{ij}^{(1)} \quad \text{per } j = 2 \dots n$$

$$\text{e } b_i^{(2)} = b_i^{(1)} - \frac{\alpha_{i1}^{(1)}}{\alpha_{11}^{(1)}} b_1^{(1)} \quad \text{per } i = 2 \dots n$$

Possiamo ripetere i passi precedenti pertanto le

$$\sum_{j=2}^n \alpha_{ij}^{(2)} x_j = b_i^{(2)} \quad \text{con } i = 2 \dots n$$

$$\alpha_{22}^{(2)} x_2 + \sum_{j=3}^n \alpha_{2j}^{(2)} x_j = b_2^{(2)}$$

Moltiplico per  $\lambda_{i2} = -\frac{\alpha_{i2}^{(2)}}{\alpha_{22}^{(2)}}$  ecco ottengo

$$-\alpha_{i2}^{(2)} x_2 + \sum_{j=3}^n \left( -\alpha_{ij}^{(2)} \frac{\alpha_{i2}^{(2)}}{\alpha_{22}^{(2)}} \right) x_j = b_2^{(2)} - b_2^{(2)} \frac{\alpha_{i2}^{(2)}}{\alpha_{22}^{(2)}}$$

che scrivere  $a_{i2}^{(2)} + \sum_{j=3}^n a_{ij}^{(2)} x_j = b_i^{(1)}$  ottenendo

$$\sum_{j=3}^n \left( a_{ij}^{(2)} - a_{2j}^{(2)} \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} \right) x_j = b_i^{(2)} - b_2^{(1)} \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} \quad \text{per } i = 3 \dots n$$

$$\Rightarrow \begin{cases} a_{22}^{(2)} x_2 + \sum_{j=2}^n a_{2j}^{(2)} x_j = b_2^{(1)} & \leftarrow \text{mane sovraetate} \\ \sum_{j=3}^n a_{ij}^{(2)} x_j = b_i^{(2)} \end{cases}$$

$$\text{con } a_{ij}^{(3)} = a_{ij}^{(2)} - a_{2j}^{(2)} \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} \quad \wedge \quad b_{ij}^{(3)} = b_i^{(2)} - b_2^{(1)} \frac{a_{i2}^{(2)}}{a_{22}^{(2)}}$$

Procedendo così fino ad ottenere in  $A^{(n)}$  avremo sempre le stesse formule:

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - a_{kj}^{(k)} \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \quad \wedge \quad b_{ij}^{(k+1)} = b_i^{(k)} - b_k^{(k)} \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$$

Il metodo di eliminazione di Gauss ha un'importante limitazione: calcolare  $a_{ij}$  è possibile non quando  $a_{jj} \neq 0$ .

Se  $a_{ij}$  (detto el. parziale) è nullo non possiamo trasformare la matrice.

E' questo ed un metodo di protezione è possibile evitare questo problema.

Calcolando il costo computazionale del metodo di Gauss dividendo in 4 fasi il processo risolutivo:

1. N. operazioni per calcolare  $a_{ij}^{(k+1)}$  e  $b_i^{(k+1)}$
2. N. operazioni per calcolare  $A^{(k)}$  e  $b^{(k)}$
3. N. operazioni per passare da  $A^{(1)}$  ad  $A^{(n)}$
4. N. operazioni per risolvere  $A^{(n)}$   $\rightarrow n^2 n$

$$1) \quad a_{ij}^{(k+1)} = a_{ij}^{(k)} - a_{kj}^{(k)} \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \quad \leftarrow \begin{matrix} 3 \text{ op.} \\ \uparrow \text{ invertibili} \end{matrix}$$

$$2) \quad b_i^{(k+1)} = b_i^{(k)} - b_k^{(k)} \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \quad \leftarrow \begin{matrix} 2 \text{ op.} \\ \left( \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \text{ già calcol.} \right) \end{matrix}$$

$$1) 5 \text{ operazioni } (3 + 2)$$

$$2) \left| \left\{ a_{ij} \right\}_{i,j \in [1,n-k]} \right| = 2(n^2 - k)^2$$

$$\left| \left\{ b_i \right\}_{i \in [1, n-k]} \right| = (n-k)$$

Summ. Bernoulli

$$\sum_{i=1}^m i^2 = \frac{m(m+1)(2m+1)}{6}$$

$$1) 2 \cdot 3(n-k) + 2(n-k)^2$$

$$3) f(n) = 2 \sum_{k=1}^{n-1} (n-k)^2 + 3 \sum_{k=1}^{n-1} (n-k) =$$

$$= 2 \sum_{k=1}^{n-1} m^2 - 4 \sum_{k=1}^{n-1} nk + 2 \sum_{k=1}^{n-1} k^2 + 3 \sum_{k=1}^{n-1} n - 3 \sum_{k=1}^{n-1} k :$$

$$= 2m^2(n-1) - 2n \sum_{k=1}^{n-1} (n-k) + 2 \sum_{k=1}^{n-1} k^2 + 3 \sum_{k=1}^{n-1} n - 3 \sum_{k=1}^{n-1} k :$$

$$+ 3n(n-1) - 3 \sum_{k=1}^{n-1} k =$$

$$= 2/m(n-1) [ 2m^2 - 2m(n-2) + \frac{1}{3}m(2m-1) +$$

$$+ 3n - \frac{3}{2}(n-2) ] =$$

$$= (n-1) [ 2m^2 - 2m(n-2) + 4n + \frac{2}{3}m^2 - \frac{1}{3}m + 3n - \frac{3}{2}n + 3 ] =$$

$$= (n-1) \left( \frac{2}{3}m^2 + \frac{35}{6}n + 3 \right) \approx \frac{2}{3}m^3$$

<sup>4</sup> Al quarta operazione  $\frac{2}{3}m^3$  bisogna sommare  $n^2$  op.

$\Rightarrow$  le op. sono  $O(n^3)$

### Strategie di deviazione

Una strategia utile per evitare di muovere el. parziali nulli e, quindi, interrompere il metodo di Gauss.

Le strategie di Bowdeng si basa sui 2 teoremi:

1. scomponere 2 regole non cambia assolutamente il sistema
2. in una regola è sempre presente un el. non nullo se  $|A| \neq 0$

Bowdeng parziale

Prevede che el.  $\kappa$ -esimo passo si scambi le righe  $k$

$$\text{con una regola } r \geq k \Rightarrow |a_{rk}^{(k)}| = \max_{i \in [k, n]} |a_{ik}^{(k)}|$$

Bowdeng totale

Prevede lo scambio tra le righe che si chiamano con la regola  $k$  e le regole  $r \neq k$  e le colonne  $k$  con le colonne  $r \geq k \Rightarrow |e_{rr}^{(k)}| = \max_{\substack{i \in [k, n] \\ j \in [k, n]}} |e_{ij}^{(k)}|$

Le strategie di Bowdeng delle garantiscono che l'el. percorso non sia molto piccolo che, quindi, il moltiplicatore non sia grande. Come aspetti negativi ha le necessarie permutazioni di righe e colonna.

Fattorizzazione LU

Per risolvere un sistema con il metodo di Gauss  $A\bar{x} = \bar{b}$  è necessario prendere  $A^{(n)}\bar{x} = \bar{b}^{(n)}$ .

Se cambiano  $\bar{b}^{(n)}$  in  $\bar{c}$ , ci bisogna di rieffettuare di nuovo tutta il metodo.

Nella fattorizzazione LU esistono L ed U matrici triangolari inferiori e superiori  $\Rightarrow A = L \cdot U$

$$\Rightarrow A\bar{x} = \bar{b} \Rightarrow L\bar{U}\bar{x} = \bar{b}$$

$$\text{Perche } U\bar{x} = \bar{y} \Rightarrow L\bar{y} = \bar{b}$$

Considero  $\bar{b}$  componibile solo  $L$  e non  $U$ .

Suppongo  $A = LU$

$$\begin{bmatrix} e_{11} & \cdots & e_{1j} & \cdots & e_{1n} \\ e_{21} & \cdots & e_{2j} & \cdots & e_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ e_{n1} & \cdots & e_{nj} & \cdots & e_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ l_{11} & 1 & \cdots & 0 & 0 \\ l_{12} & l_{22} & \ddots & \ddots & 0 \\ l_{1n} & l_{2n} & \cdots & l_{nn} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{1j} & u_{1n} \\ 0 & u_{22} & u_{2n} \\ \vdots & \ddots & \vdots \\ 0 & 0 & u_{nn} \end{bmatrix}$$

$$e_{ij} = \sum_{k=1}^n l_{ik} u_{kj} = \sum_{k=1}^{min(i,j)} l_{ik} u_{kj} \leftarrow \text{escludere index above } l_{ij} = 0 \text{ or } u_{ij} = 0$$

Considero il caso  $i \leq j$

$$l_{ij} = \sum_{k=1}^i l_{ik} u_{kj} = \sum_{k=1}^{i-1} l_{ik} u_{kj} + l_{ii} u_{ij} = \sum_{k=1}^{i-1} l_{ik} u_{kj} + u_{ij}$$

$$\Rightarrow u_{ij} = e_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \wedge u_{1j} = e_{1j} \quad \forall j \in [1, n]$$

Considero ora il caso  $j \leq i$

$$e_{ij} = \sum_{k=1}^j l_{ik} u_{kj} = \sum_{k=1}^{j-1} l_{ik} u_{kj} + l_{jj} u_{jj}$$

$$\Rightarrow l_{ij} = \frac{1}{u_{jj}} \left( e_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \right)$$

Esercizio 2 tecniche per calcolare  $L$  ed  $U$

- Recette di Crout



- Recette di Doolittle

Diagram showing a flow from the text 'Recette di Doolittle' to two 2x2 matrix diagrams:

1	3
2	4
5	6
7	

1	
2	3
4	5
6	7

Quint:

1. regole 1 di U
  2. regole 2 di L
  3. regole 2 di U
  4. regole 3 di L
  5. regole 3 di U
- ...

Distributore:

1. regole 1 di U
  2. ~~colonne~~ ~~regole~~ 1 di L
  3. regole 2 di U
  4. colonne 2 di L
  5. regole 2 di U
- ...

Le matrici U coincidono con  $A^{(m)}$  mentre in L avremo i -~~eff~~ moltiplicatori.

Nelle tecniche di fatt. LU non ci rimane che trovare una strategia di permutazione.

Eg. se non è fatt. LU

$$A\bar{x} = \bar{b} \Rightarrow A^{(1)}\bar{x} = \bar{b}^{(1)}$$

$$L^{(1)} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ l_{31} & 0 & \cdots & 0 \\ \vdots & & & \ddots \\ l_{n1} & 0 & \cdots & 1 \end{bmatrix} \Rightarrow A^{(1)} = L^{(1)} A^{(1)}$$
$$\bar{b}^{(1)} = L^{(1)} \bar{b}^{(1)}$$

Si dimostra in generale che  $L^{(k)} A^{(k)} = A^{(n-k)}$

$$\text{con } L^{(k)} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad L^{(k)} \bar{b}^{(k)} = \bar{b}^{(n-k)}$$

Ora anche  $A^{(m)} A^{(m)} = L^{(m-1)} A^{(m-1)} \quad \bar{b}^{(m)} = L^{(m-1)} \bar{b}^{(m)}$

$$A^{(m)} = L^{(m-1)} A^{(m-1)} = L^{(m-1)} L^{(m-2)} \cdots L^{(1)} A^{(1)}$$

$$(L^{(k)})^{-1} = \cancel{\cancel{\cancel{\cancel{L^{(k)}}}}} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \rightarrow -l_{ij} \text{ al posto di } l_{ij}$$

$$(L^{(1)})^{-1} (L^{(2)})^{-1} \cdots (L^{(m-1)})^{-1} = \# L$$

$$(L^{(n)})^{-1} (L^{(n)})^{-1} (L^{(n)})^{-1} \dots (L^{(n-n)})^{-1} = \text{fl L}$$

$$(L^{(n-1)} L^{(n-2)} L^{(n-3)} \dots L^{(1)} L^{(0)})^{-1} = \text{fl L}$$

$$\underbrace{(L^{(n)})^{-1} (L^{(n)})^{-1} \dots (L^{(n-1)})^{-1}}_{\text{L } A^{(n)}} A^{(n)} = A$$

$$\text{L } A^{(n)} = A \Rightarrow A^{(n)} = U \Rightarrow \text{L } U = A \quad \text{com}$$

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -l_{21} & 1 & 0 & 0 \\ -l_{31} - l_{21} & 1 & 0 & 0 \\ -l_{41} - l_{31} - l_{21} & 1 & 0 & 0 \end{pmatrix} \quad \text{A } U = A^{(n)}$$

Condizionamento di un sistema lineare

$$\begin{cases} x + y = 2 \\ 1000x + 1001y = 2001 \end{cases} \Rightarrow \begin{cases} x = 1 \\ y = 1 \end{cases}$$

Perturbazione  $x$  (di  $1\%$ ).  $\tilde{x} = x(1+0.01)$

$$\begin{cases} \tilde{x} + y = 2 \\ 1000\tilde{x} + 1001y = 2001 \end{cases} \Rightarrow \begin{cases} \tilde{x} = -\frac{1}{3} \\ y = \frac{1991}{300} \end{cases}$$

$$\|\tilde{x} - \bar{x}\| = 1.57 = \epsilon_A(\bar{x}) = \epsilon_A(\tilde{x}) \quad \text{ALERT! SSM!}$$

$$\bar{A} \bar{x} = \bar{b} \Rightarrow (A + \delta A)(\bar{x} + \delta \bar{x}) = (\bar{b} + \delta \bar{b})$$

$$\Rightarrow \frac{\|\delta \bar{x}\|}{\|\bar{x}\|} \leq \|A\| \|A^{-1}\| \left( \frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right) k$$

K è detto indice di condizionamento del sistema.

## INTERPOLAZIONE

È un metodo che consente nell' approssimazione di una funzione di conoscere solo alcuni punti.

L'insieme dei punti  $x_i$  noti sono chiamati nodi.

Certe volte è possibile interpolare una f. anche se quest'ultima è nota ma è troppo complessa.

$$f(x) \approx g(x; \alpha_0, \alpha_1, \dots, \alpha_n) \text{ con parametri } \alpha_0, \dots, \alpha_n$$

$$f(x_i) = g(x_i) \quad \forall i = 0 \dots n$$

Il procedimento per usare è quello nel quale

$$g(x; \alpha_0, \alpha_1, \dots, \alpha_n) = \sum_{i=0}^n \alpha_i \Phi_i(x) \text{ dove}$$

$$f(x_i) = \sum_{i=0}^n \alpha_i \Phi_i(x_i) \quad \leftarrow g deve passare dai nodi$$

Risolvere quest'ultima eq. trovando i coeff.  $\alpha_i$ . Vi è il metodo dei coeff. indeterminati.

## Interpolazione polinomiale

$$\text{con } \Phi_i(x) = x^i \quad \forall i$$

$$g(x; \alpha_0, \alpha_1, \dots, \alpha_n) = \sum_{i=0}^n \alpha_i x^i \leftarrow \text{polinomio}$$

$$\Rightarrow \begin{cases} \alpha_0 + \alpha_1 x_0 + \alpha_2 x_0^2 + \dots + \alpha_n x_0^n = f(x_0) \\ \alpha_0 + \alpha_1 x_1 + \alpha_2 x_1^2 + \dots + \alpha_n x_1^n = f(x_1) \\ \alpha_0 + \alpha_1 x_m + \alpha_2 x_m^2 + \dots + \alpha_n x_m^n = f(x_m) \end{cases}$$

$$\text{risolvibile come } V \bar{\alpha} = \bar{y} \quad \text{con} \quad \bar{y} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_m) \end{bmatrix}, \quad \bar{\alpha} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$$e \quad V = \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^n \end{pmatrix} \in \mathbb{R}^{(m+1) \times (n+1)}$$

Se  $V$  (matrice di Vandermonde) è non singolare il problema dell'interpolazione avrà al più una soluzione ( $V$  non è singolare se i valori sono distinti a 2 a 2)

Risolvere il metodo dei vett. indeterminati richiede la soluzione di un sistema che potrebbe essere enorme e anche nel caso normale.

### Polinomio Interpolante di Lagrange

$$L_n(x) = \sum_{k=0}^n l_{nk}(x) f(x_k) \quad \text{con } l_{nk}(x) \text{ polinomio di grado } n.$$

Ricavando le cosiddette di interpolazione  $L_n(x_i) = f(x_i)$  quindi  $L_n(x_i) = \sum_{k=0}^n l_{nk}(x_i) f(x_k) = f(x_i)$

$$\Rightarrow l_{nk}(x_i) = \begin{cases} 0 & k \neq i \\ 1 & k = i \end{cases}$$

Per far sì che il polinomio  $l_{nk}(x)$  sia esattamente zero per  $x_i$  con  $i \neq k$  proviamo:

$$l_{nk}(x) = c_k \prod_{\substack{i=0 \\ i \neq k}}^n (x - x_i) \quad \text{con } c_k \text{ cost.}$$

$$l_{nk}(x_k) = c_k \prod_{\substack{i=0 \\ i \neq k}}^n (x_k - x_i) = 1 \Rightarrow$$

$$\Rightarrow c_k = \frac{1}{\prod_{\substack{i=0 \\ i \neq k}}^n (x_k - x_i)} \quad \text{quindi}$$

$$L_n(x) = \sum_{k=0}^n \left( \prod_{\substack{i=0 \\ i \neq k}}^n \frac{x - x_i}{x_k - x_i} \right) f(x_k) \quad \leftarrow \text{POLINOMIO DI LAGRANGE}$$

con  $l_{nk}(x)$  polinomi fondamentali di Lagrange.

## Errore di Interpolazione

$$e(x) = f(x) - P_m(x)$$

$f(x)$  non è mai grande non è necessariamente celebabile  
 $e(x)$  errore (o resto) dell'interpolazione.

che si rappresenta con l'errore di interpolazione

Se  $f \in C^{n+1}([a, b])$  e  $x_i \in [a, b]$ ,  $i = 0 \dots n+1$

Se  $P_m(x)$  è il pol. di interpolazione di  $f(x)$  nei punti

$x_i \Rightarrow \forall x \in [a, b] \exists \xi = \xi(x) \ni$

$$e(x) = f(x) - P_m(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \underbrace{(x-x_0)(x-x_1) \dots (x-x_n)}_{w_{n+1}(x)}$$

$$\text{con } w_{n+1}(x) = \prod_{i=0}^n (x-x_i)$$

$$\Rightarrow e(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} w_{n+1}(x)$$

polinomio notale

$$\underline{\text{Ese}} \quad e(x) = |f(x) - L_1(x)| = \frac{f''(\xi_x)}{2!} (x-x_0)(x-x_1)$$

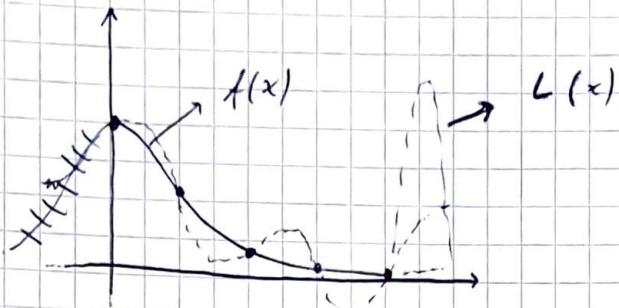
che ponendo  $x = x_0 + h \cdot z$  con  $h = x_1 - x_0$  e  $z \in [0, 1]$

$$\begin{aligned} \text{dunque } e(x) &= \frac{f''(\xi_x)}{2} h z (x_0 + h z - x_1) = \\ &= \frac{f''(\xi_x)}{2} h z (h z - h) = \frac{f''(\xi_x)}{2} h^2 z(z-1) = \\ &= \frac{f''(x_0 + zh)}{2} h^2 z(z-1) \end{aligned}$$

$$\max_{z \in [0, 1]} |z(z-1)| = \frac{1}{4} \Rightarrow e(x) \leq \frac{\boxed{f''(\xi_x)}}{8} h^2 \max_{x \in [x_0, x_1]} |f''(x)|$$

## Fenomeno di Runge

$$f(x) = \frac{1}{1+x^2} \text{ in } [-5, 5] \text{ con } x_i = -5 + \frac{10i}{m}$$



$L(x)$  si allontana sempre più da  $f(x)$  con  $m \rightarrow +\infty$

Qui riducendo l'intervallo  $[a, b]$  più eamente l'oscillazione e l'ampiezza.

$$\lim_{n \rightarrow \infty} \max_{x \in [-5, 5]} |f(x) - L_n(x)| = +\infty$$

Ovvero fenomeno i chiamato fenomeno di Runge ed è un grande problema dell'interpolazione di Lagrange.

## Minimizzare il resto nel problema di Interpolazione

$$e(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} w_{n+1}(x)$$

Non posso mai fare  
rendere  $w_{n+1}(x)$  in polinomio di grado al più  $n+1$  e'

$$\max_{x \in [a, b]} |\tilde{p}(x)| = \min_{p \in P_{n+1}} \max_{x \in [a, b]} |p(x)|$$

fattibile utilizzando i polinomi di Chebyshev.

## Polinomio di Chebyshev

$$T_n(x) = \cos(n \arccos x)$$

per  $x \in [-1, 1]$

$$\text{Ese} \quad T_0(x) = \cos(0) = 1, \quad T_1(x) = x$$

dove vale la seguente formula ricorsiva:

$$T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x)$$

Dato  $\arccos x = \theta \Rightarrow \cos \theta = x$

$$T_n(x) = \cos(n\theta)$$

$$T_{n+1}(x) = \cos[(n+1)\theta] = \cos(n\theta) \overbrace{\cos \theta}^x - \sin(n\theta) \sin \theta$$

$$T_{n-1}(x) = \cos[(n-1)\theta] = \cos(n\theta) \overbrace{\cos \theta}^x + \sin(n\theta) \sin \theta$$

$$\begin{aligned} T_{n+1}(x) &= 2 \cancel{\cos n\theta} \overbrace{\cos \theta}^x - T_{n-1}(x) \\ &= 2x T_n(x) - T_{n-1}(x) \end{aligned}$$

$$T_0(x) = 1, \quad T_1(x) = x$$

$$T_2(x) = 2x \cdot x - 1 = 2x^2 - 1$$

$$T_3(x) = 2x(2x^2 - 1) - x = 4x^3 - 3x$$

$$\Rightarrow T_n(x) \approx 2^{n-1} x^{n+1}$$

Proiettare  $T_n(x)$

$$1. \max_{x \in [-1, 1]} |T_n(x)| = 1$$

$$2. T_{2k}(x-x) = T_{2k}(x) \quad \wedge \quad T_{2k+1}(x-x) = -T_{2k+1}(x)$$

$$3. |T_n(x)| = 1 \quad \text{per } x = \cos\left(\frac{k\pi}{n}\right)$$

$$\text{In particolare } T_n(x_k) = (-1)^k \cos x_k = \cos\left(\frac{k\pi}{n}\right)$$

$$\text{con } k = 0 \dots n$$

$$4. T_n(x_k) = 0 \quad \text{per } x_k = \cos \frac{(2k+1)\pi}{2n}$$

$$\text{con } k = 0 \dots n-1$$

## Ri. di minimax

Sia  $p_n(x)$  un qualsiasi polinomio monico di grado  $n$

$$\frac{1}{2^{n+1}} = \max_{x \in [-1,1]} |\tilde{T}_n(x)| \leq \max_{x \in [-1,1]} |p_n(x)|$$

$$\text{con } \tilde{T}_n(x) = \frac{1}{2^{n+1}} T_n(x) \quad \leftarrow \text{rl. Chebyshev monico}$$

$$(\text{In realtà} \quad \max_{x \in [-1,1]} |p_n(x)| > \max_{x \in [-1,1]} |\tilde{T}_n(x)| \text{ quando il pol. si Chebyshev monico è il polinomio di grado } n \text{ con il più piccolo dei massimi.})$$

$$\text{Ponendo } w_{n+1}(x) = \tilde{T}_{n+1}(x) \text{ avremo}$$

$$e(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \cdot \tilde{T}_{n+1}(x) \leq \frac{\overline{|f^{(n+1)}(\xi_x)|}}{(n+1)!} \cdot \frac{1}{2^{n+2}} \max_{x \in [-1,1]} |f^{(n+1)}(x)|$$

Se  $[a,b] \neq [-1,1]$  bisogna effettuare una trasformazione lineare fra i 2 intervalli.

$$t = \frac{b-a}{2}x + \frac{a+b}{2} \quad \text{con } x \in [-1,1]$$

Dove dati  $x_k$  zeri del pol. di Chebyshev  $T_{n+1}(x)$

$$\tau_k = \frac{b-a}{2}x_k + \frac{a+b}{2} \quad \text{con } k = 0 \dots n$$

↑ NODI

$$\text{con } x_k = \cos \frac{(2k+1)\pi}{2(n+1)}$$

Usando gli zeri del polinomio di Chebyshev come nodi avremo un il resto minimo (errore minimo)

$$\text{I nodi da usare sono } \tau_k = \frac{b-a}{2}x_k + \frac{a+b}{2} \quad \text{con}$$

$$x_k = \cos \frac{(2k+1)\pi}{2(n+1)} \text{ zeri di } T_{n+1}(x)$$

Se i punti di nodi sono suff. elevati ~~abbastanza~~  
 la f. interpolante potrebbe mostrare un comportamento  
 fortemente oscillante.

A questo punto è preferibile usare l'approccio Spline.

### Funzione Spline

$$\Delta =: \alpha \equiv x_0 < x_1 < x_2 < \dots < x_m < x_{m+1} \equiv b$$

è una decomposizione dell'intervallo  $[a, b]$

$s(x)$  è una funzione spline di grado  $m \geq 1$  relativa  
 alla decomposizione  $\Delta$  che ha le seguenti proprietà:

- 1)  $s(x)$  in  $[x_i, x_{i+1}] \in \mathcal{C}_m$
- 2)  $s^{(k)}(x)$  è continua per  $k \in [0, m-1]$   
 $(s(x) \in \mathcal{C}^{m-1})$

### A approssimazione ai minimi quadrati

$$\varepsilon_i = \phi(\alpha_0, \alpha_1, \dots, \alpha_m; x_i) - y_i, \quad i = 0, 1, \dots, n$$

è la differenza fra il valore assunto dalla funzione  
 interpolante e la f. effettiva nei nodi  $y_i = f(x_i)$

$$\text{Pertanto } \bar{\alpha} = [\alpha_0, \alpha_1, \dots, \alpha_m] \Rightarrow \bar{\varepsilon} = [\varepsilon_0, \varepsilon_1, \dots, \varepsilon_m]$$

abbie le minime norme euclidiene al quadrato

$$\|\bar{\varepsilon}\|_2^2 = \sum_{i=0}^n \varepsilon_i^2 = \sum_{i=0}^n (\phi(\alpha_0, \alpha_1, \dots, \alpha_m; x_i) - y_i)^2$$

$$\Rightarrow Q(\alpha_0^*, \alpha_1^*, \dots, \alpha_m^*) = \min_{\alpha_0, \alpha_1, \dots, \alpha_m \in \mathbb{R}} Q(\alpha_0, \alpha_1, \dots, \alpha_m)$$

$$\text{con } Q(\alpha_0, \alpha_1, \dots, \alpha_m) = \|\bar{\varepsilon}\|_2^2 \quad \text{retta}$$

$$\text{Analiticamente il caso di } \phi(\alpha_0, \alpha_1; x) = \alpha_0 x + \beta$$

## Rechte Linie regression

$$\phi(\alpha, \beta, x) = \alpha x + \beta \quad \text{mit } \alpha, \beta \in \mathbb{R}$$

$$\phi(\alpha, \beta, x_i) - y_i = \alpha x_i + \beta - y_i$$

$$\psi(\alpha, \beta) = \sum_{i=0}^n (\alpha x_i + \beta - y_i)^2 \quad \leftarrow \text{die minimieren}$$

$$\begin{aligned} \frac{\partial \psi}{\partial \alpha} &= \sum_{i=0}^n 2x_i (\alpha x_i + \beta - y_i) = 0 \\ \frac{\partial \psi}{\partial \beta} &= \sum_{i=0}^n 2(\alpha x_i + \beta - y_i) = 0 \end{aligned} \quad \left. \right\} \Rightarrow$$

$$\begin{aligned} \Rightarrow & \left\{ \begin{array}{l} \alpha \sum_{i=0}^n x_i^2 + \beta \sum_{i=0}^n x_i - \sum_{i=0}^n x_i y_i = 0 \\ \alpha \sum_{i=0}^n x_i + \beta \sum_{i=0}^n 1 - \sum_{i=0}^n y_i = 0 \end{array} \right. \Rightarrow \\ \Rightarrow & \left\{ \begin{array}{l} \alpha \sum_{i=0}^n x_i^2 + \beta \sum_{i=0}^n x_i = \sum_{i=0}^n x_i y_i \\ \alpha \sum_{i=0}^n x_i + \beta (n+1) = \sum_{i=0}^n y_i \end{array} \right. \Rightarrow \end{aligned}$$

$$\begin{aligned} \Rightarrow & \left\{ \begin{array}{l} S_{xx} \alpha + S_x \beta = S_{xy} \\ S_x \alpha + (n+1) \beta = S_y \end{array} \right. \quad \begin{array}{l} S_{xx} = \sum x_i^2 \\ S_x = \sum x_i \\ S_y = \sum y_i \\ S_{xy} = \sum x_i y_i \end{array} \end{aligned}$$

Usando Cramer

$$\alpha = \frac{\begin{vmatrix} S_{xy} & S_x \\ S_y & n+1 \end{vmatrix}}{\begin{vmatrix} S_{xx} & S_x \\ S_x & n+1 \end{vmatrix}} = \frac{(n+1)S_{xy} - S_x S_y}{S_{xx}(n+1) - S_x^2}$$

$$\beta = \frac{\begin{vmatrix} S_{xx} & S_{xy} \\ S_x & S_y \end{vmatrix}}{\begin{vmatrix} S_{xx} & S_x \\ S_x & n+1 \end{vmatrix}} = \frac{S_{xx} S_y - S_x S_{xy}}{S_{xx}(n+1) - S_x^2}$$

## QUADRATURA

Le quadrature sono tecniche che consentono di calcolare

$$\int_a^b f(x) dx = I(f) \quad \text{conoscendo } f(x) \text{ nell'int. } [a, b]$$

quando è desiderato calcolare o non conoscere  $F(x)$   
o quando si conoscono solo in certi punti i punti  
appartenenti ad  $f(x)$

Formule di Quadratura di tipo Interpolazione

$$\sum_{k=0}^n w_k f(x_k) \approx I(f)$$

Formule di quadratura con  $x_k$  nodi e  $w_k$  pesi

$$Ritroviamo f(x) = L_n(x) + e(x)$$

$$\int_a^b f(x) dx = \int_a^b L_n(x) dx + \int_a^b e(x) dx =$$

$$= \int_a^b \sum_{k=0}^n l_{mk}(x) f(x_k) dx + \int_a^b e(x) dx =$$

$$= \sum_{k=0}^n \left[ \int_a^b l_{mk}(x) \right] f(x_k) + \int_a^b e(x) dx$$

$$\text{Ponendo } w_k = \int_a^b l_{mk}(x) dx \text{ e } R_{m+1}^{(f)} = \int_a^b e(x) dx$$

formula interpolatrice

$$\Rightarrow I(f) = \sum_{k=0}^n w_k f(x_k) + R_{m+1}^{(f)}$$

$$(I(f) \approx \sum_{k=0}^n w_k f(x_k))$$

↑  
RESTO DELLA FORMULA DI  
QUADRATURA

Grado di precisione

Una formula ha grado di precisione  $q$  se fornisce il  
valore esatto di  $I(f)$  quando  $f$  è un polinomio di grado  
al più  $q$  e  $\exists p(x) \in P_{q+1} \setminus P_q \Rightarrow R_{q+1}(f) \neq 0$

Prop. Una formula di quadratura di tipo interpolazione ha grado di precisione  $q \geq n$  ( $n = \min. \deg. \text{ dei nodi} - 1$ )

$$\underline{\text{Dim}} \quad \int_a^b p_m(x) dx = \sum_{i=0}^n w_i p_m(x_i) + R_{m+1}(f)$$

$$R_{m+1}(f) = \int_a^b e(x) dx = \int_a^b w_{m+1}(x) \frac{p_m^{(m+1)}(x)}{(m+1)!} dx \equiv 0$$

$$(p_m^{(m+1)}(x) = 0 \quad \forall p_m \in \mathcal{P}_m)$$

### Formule di Newton - Cotes

Intervalllo  $[a, b]$  in  $n$  subintervalli di ampiezza  $h$

$$\text{con } h = \frac{b-a}{n}$$

$$x_i = a + ih \quad \text{con } i = 0, 1, \dots, n$$

$$\int_a^b f(x) dx = \sum_{i=0}^n w_i f(x_i) + R_{m+1}(f) - \epsilon$$

le formule di Newton - Cotes

Proprietà  $w_i = w_{n-i}$  (proprietà di simmetria)

$$\underline{\text{Dim}} \quad c = \frac{x_n + x_{n-i}}{2} = \frac{b+a}{2} \Rightarrow \begin{cases} x_i = 2c - x_{n-i} \\ x_{n-i} = 2c - x_i \end{cases}$$

$$w_k = \int_a^b \prod_{\substack{i=0 \\ i \neq k}}^m \frac{x - x_i}{x_k - x_i} dx = \int_a^b \prod_{\substack{i=0 \\ i \neq k}}^m \frac{x - 2c + x_{n-i}}{2c - x_{n-k} - 2c + x_{n-i}} dx = \\ = \int_a^b \prod_{\substack{i=0 \\ i \neq k}}^m \frac{2c - x - x_{n-i}}{x_{n-k} - x_{n-i}} dx$$

$$\text{Posto } t = 2c - x : \quad w_k = \int_a^b \prod_{\substack{i=0 \\ i \neq k}}^m \frac{t - x_{n-i}}{x_{n-k} - x_{n-i}} dt = \\ \left( \begin{array}{l} t = 2c - a \Rightarrow t = b \\ t = 2c - b = a, \quad dt = -dx \end{array} \right)$$

riconoscere gli estremi di integrazione

$$\int_a^b f(x) dx = - \int_b^a f(t) dt = \\ = \int_a^b f(t) dt$$

$$\text{con } j = m - i$$

✓

## Formule dei Regoli

$$x_0 = a, \quad x_1 = b \quad \text{e} \quad h = b - a$$

$$T_2 = w_0 f(x_0) + w_1 f(x_1)$$

$$\begin{aligned} w_0 &= \int_a^b l_{x_0}(x) dx = \int_a^b \frac{x-x_1}{x_0-x_1} dx = \int_a^b \frac{x-b}{a-b} dx = \\ &= \frac{1}{a-b} \left[ \frac{(x-b)^2}{2} \right]_{x=a}^{x=b} = -\frac{h^2}{2} \cdot \frac{1}{-h} = \frac{h}{2} \end{aligned}$$

$a$  e  $b$  non sono simmetrici rispetto a  $c \Rightarrow$

$$\Rightarrow w_0 = w_1 = \frac{h}{2}$$

$$T_2 = \frac{h}{2} (f(x_0) + f(x_1)) \quad \leftarrow \text{Formule dei Regoli}$$

Resta nelle formule dei Regoli

$$R_2(f) = \frac{1}{2!} \int_a^b (x-a)(x-b) f''(\xi_x) dx$$

Per manipolare quest'espressione abbiamo bisogno di conoscere ilh. delle medie generalizzate.

ih. delle medie generalizzate

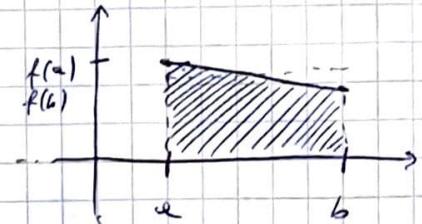
Siano  $f, g : [a, b] \rightarrow \mathbb{R}$  f. contiene con  $g(x) \neq 0$  e segno costante e  $g(x) \neq 0 \quad \forall x \in (a, b) \Rightarrow$

$$\Rightarrow \int_a^b f(x) g(x) dx = f(\xi_g) \int_a^b g(x) dx, \quad \xi_g \in [a, b]$$

Ossia  $(x-a)(x-b)$  è e segno costante

$$\int_a^b (x-a)(x-b) f''(\xi_x) dx = f''(\eta) \int_a^b (x-a)(x-b) dx$$

$$\Rightarrow R_2(f) = \frac{f''(\eta)}{2!} \int_a^b (x-a)(x-b) dx$$



$$\text{Ansatze } h = b - a \quad \wedge \quad x = a + ht \quad \text{mit } t \in [0, 1]$$

$$(dx = h \cdot dt)$$

$$R_2(f) = \frac{f''(y)}{2!} \int_0^1 ht \cdot (-h + ht) h \, dt =$$

$$= \frac{f''(y)}{2!} \int_0^1 h^3 t(t-1) \, dt = \frac{f''(y)}{2!} h^3 \left[ \frac{t^3}{3} - \frac{t^2}{2} \right]_0^1$$

$$= -\frac{h^3}{12} f''(y)$$

$$\exists R_2(f) = 0 \iff f \in \mathcal{P}_1$$

Il grado di precisione  $q = 1$   
 (con  $f$  si perde 2  $R_2(f) \neq 0$ )

### Formule di Simpson

$$x_0 = a, \quad x_2 = b \quad \wedge \quad x_1 = c = \frac{a+b}{2}$$

$$S_3 = w_0 f(a) + w_1 f(c) + w_2 f(b)$$

$$h = \frac{b-a}{2}$$

$$w_0 = \int_a^b l_{2,0}(x) \, dx = \int_a^b \frac{(x-b)(x-c)}{(a-b)(a-c)} \, dx \quad \xrightarrow{x=c+ht, \, dx=h \, dt}$$

$$\Rightarrow w_0 = \int_{-1}^1 \frac{(ht-h)(ht)}{(-2h)(-h)} \cdot h \, dt = \int_{-1}^1 \frac{(t-1)ht}{2} \, dt =$$

$$= \frac{h}{2} \left[ \frac{t^3}{3} - \frac{t^2}{2} \right]_{-1}^1 = \frac{h}{2} \left( -\frac{1}{6} + \frac{5}{6} \right) = \frac{2}{3} \cdot \frac{h}{2} = \frac{h}{3}$$

$$w_0 = w_2 = \frac{h}{3} \quad \text{per symmetrie}$$

$$w_1 = \int_a^b l_{2,1}(x) \, dx = \int_a^b \frac{(x-a)(x-b)}{(c-a)(c-b)} \, dx =$$

$$= \int_{-1}^1 \frac{\cancel{h}(1+ht)(ht-h)}{\cancel{h}(-h)} \, dt = h \int_{-1}^1 1-t^2 \, dt =$$

$$= h \left[ t - \frac{t^3}{3} \right]_{-1}^1 = h \left( \frac{2}{3} + \frac{2}{3} \right) = \frac{4}{3} h$$

$$S_3 = \frac{h}{3} (f(a) + 4f(c) + f(b)) \leftarrow \text{formula di Simpson}$$

$$\text{errore } R_3(f) = -\frac{h^5}{90} \frac{f''(\sigma)}{\sigma} \text{ con } \sigma \in (a, b)$$

$$\Rightarrow \text{grado di precisione } q = 3$$

$$\text{NOTA: } h = \frac{b-a}{2}$$

### Formule del p.t. di metà

Sia  $c = \frac{a+b}{2}$ . Sviluppo la regola di Taylor con c p.t. medie.

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(\xi_x)}{2!}(x-c)^2$$

$$\int_a^b f(x) dx = \int_a^b \cdot f(c) dx + f'(c) \underbrace{\int_a^b x-c dx}_{=0} + \\ + \int_a^b \frac{f''(\xi)}{2} (x-c)^2 dx = f(c) (b-a) + \int_a^b \frac{f''(\xi)}{2} (x-c)^2 dx$$

$$\text{Ponendo } h = b-a \Rightarrow \int_a^b f(x) dx \approx h f(c)$$

$$R(f) = \int_a^b \frac{f''(\xi)}{2} (x-c)^2 dx = \frac{f''(\eta)}{2} \int_a^b (x-c)^2 dx = \\ = \frac{f''(\eta)}{2} \left. \frac{(x-c)^3}{3} \right|_a^b = \frac{f''(\eta)}{2} \cdot \left( \frac{h^3}{24} + \frac{h^3}{24} \right) = \frac{f''(\eta)}{324} h^3 \approx$$

$$R(f) = \frac{h^3}{24} f''(\eta) \quad M = h f(c)$$

$$\Rightarrow \text{grado di precisione } q = 1$$

Il grado di precisione di questa regola è uguale a quello del trapezio ma ha qualche formula necessaria solo per il p.t.

### Formule di quadratura composte

Consiste nel suddividere l'intervallo  $[a, b]$  in  $N$  sottointervalli con ampiezza uguale, aumentando la precisione.

$$\int_a^b f(x) dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x) dx \text{ con } x_i = a + ih$$

$$e h = \frac{b-a}{N}$$

### Formule di Regra Composite

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{i=0}^{N-1} \left[ \frac{h}{2} (f(x_i) + f(x_{i+1})) - \frac{1}{12} h^3 f''(\eta_i) \right] = \\ &= \frac{h}{2} \left( \sum_{i=0}^{N-1} (f(x_i) + f(x_{i+1})) \right) - \frac{1}{12} h^3 \sum_{i=0}^{N-1} f''(\eta_i) = \\ &= \frac{h}{2} (f(x_0) + f(x_N)) + h \sum_{i=1}^{N-1} f(x_i) - \frac{h^3}{12} \cdot N f''(\eta) \end{aligned}$$

$\eta \in (x_i, x_{i+1})$

$$T_C(h) = \frac{h}{2} (f(x_0) + f(x_N)) + h \sum_{i=1}^{N-1} f(x_i)$$

$\eta \in (a, b)$

$$R_T = -\frac{h^3}{12} N f''(\eta) = -\frac{(b-a)^3}{12 N^3} N f''(\eta) = -\frac{(b-a)^3}{12 N^2} f''(\eta)$$

$$\sum_{i=1}^N f(x_i) = N f(x) \quad \leftarrow \text{Rh. delle medie nel discreto}$$

$$\text{se } x_i \in [a, b] \quad x \in (a, b)$$

$$\begin{aligned} \text{Supponiamo che } |R_T| &\leq \frac{1}{12} \frac{(b-a)^3}{N^2} \max_{x \in (a,b)} |f''(x)| \\ \Rightarrow \varepsilon &\geq \frac{1}{12} \frac{(b-a)^3}{N_E^2} M \Rightarrow N_\varepsilon \geq \sqrt{\frac{(b-a)^3 \cdot M}{12 \varepsilon}} \end{aligned}$$

Possiamo stabilire in quanti passi occorrono  $f[a, b]$  per ottenere una precisione  $\varepsilon \geq |R_T|$

### Formule di Simpson Composite

Scomporre  $[a, b]$  in  $N$  intervalli con  $N$  passi.

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{i=0}^{\frac{N}{2}-1} \int_{x_{2i}}^{x_{2i+2}} f(x) dx = \\ &= \sum_{i=0}^{\frac{N}{2}-1} \left[ \frac{h}{3} (f(x_{2i}) + 4f(x_{2i+1}) + f(x_{2i+2})) - \frac{h^3}{90} f'''(\eta_i) \right] = \end{aligned}$$

$$= \frac{h}{3} \sum_{i=0}^{\frac{N}{2}-1} [f(x_{2i}) + 4f(x_{2i+1}) + f(x_{2i+2})] - \frac{h^5 N}{180} f'''(h)$$

$$S_C(h) = \frac{h}{3} \left[ f(x_0) + f(x_N) + 2 \sum_{i=1}^{\frac{N}{2}-1} f(x_{2i}) + 4 \sum_{i=1}^{\frac{N}{2}-1} f(x_{2i+1}) \right]$$

$$R_S = -\frac{h^5 N}{180} f'''(h) = -\frac{(b-a)^5}{180 N^4} f'''(h)$$

$$|R_S| \leq \frac{(b-a)^5}{180 N^4} \max_{x \in (a,b)} |f'''(x)|$$

$$\epsilon \geq \frac{(b-a)^5}{180 N_\epsilon^4} M \Rightarrow N_\epsilon^* \geq \sqrt[4]{\frac{(b-a)^5 M}{180 \epsilon}}$$

Formule del punto di massimo compresa

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{i=0}^{\frac{N}{2}-1} \int_{x_{2i}}^{x_{2i+2}} f(x) dx = \\ &= \sum_{i=0}^{\frac{N}{2}-1} \frac{2h}{N} f(x_{2i+1}) + \sum_{i=0}^{\frac{N}{2}-1} \frac{(3h)^3}{24} f''(\eta_i) = \\ &= 2h \sum_{i=0}^{\frac{N}{2}-1} f(x_{2i+1}) + \frac{h^3}{24} N f''(y) \end{aligned}$$

$$M_C(h) = 2h \sum_{i=0}^{\frac{N}{2}-1} f(x_{2i+1})$$

$$R_M = \frac{(b-a)^3}{6 N^2} f''(y) \rightarrow \epsilon N_\epsilon \geq \sqrt{\frac{(b-a)^3 M}{6 \epsilon}}$$

## DERIVAZIONE NUMERICA

È una tecnica per trovare le derivate di una funzione senza calcolare effettivamente  $f^{(m)}(x)$

Supponiamo che  $f \in C^k([a,b])$  con  $k$  suff. grande per i nostri scopi e consideriamo  $[a,b]$  un

$$J \Delta \equiv t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n = b$$

Considerer 3 punti consecutivi  $t_{n-1}, t_n, t_{n+1}$

$$t_{n-1} = t_n - h \quad \text{e} \quad t_{n+1} = t_n + h$$

$$f(t_{n+1}) = f(t_n) + f'(t_n)h + \frac{f''(t_n)}{2}h^2 + \frac{f^{(3)}(t_n)}{6}h^3 + \frac{f^{(4)}(\xi_1)}{24}h^4$$

$$f(t_{n-1}) = f(t_n) - f'(t_n)h + \frac{f''(t_n)}{2}h^2 - \frac{f^{(3)}(t_n)}{6}h^3 + \frac{f^{(4)}(t_n)}{24}h^4$$

$$f(t_{n+1}) + f(t_{n-1}) = 2f(t_n) + f''(t_n)h^2 + \frac{h^4}{24}(f^{(4)}(\xi_1) + f^{(4)}(\xi_2))$$

$$\Rightarrow f(t_{n+1}) + f(t_{n-1}) \approx 2f(t_n) + f''(t_n)h^2$$

$$\Rightarrow f''(t_n) \approx \frac{f(t_{n+1}) + f(t_{n-1}) - 2f(t_n)}{h^2}$$

$$E(f''(t_n)) = -\frac{h^2}{24}(f^{(4)}(\xi_1) + f^{(4)}(\xi_2)) = -\frac{h^2}{72}f^{(4)}(x)$$

Breve di approssimare le derivate prime

$$f(t_{n+1}) = f(t_n) + f'(t_n)h + \frac{f''(t_n)}{2}h^2 + \frac{f^{(3)}(\xi_1)}{6}h^3$$

$$f(t_{n-1}) = f(t_n) - f'(t_n)h + \frac{f''(t_n)}{2}h^2 - \frac{f^{(3)}(\xi_2)}{6}h^3$$

$$f(t_{n+1}) - f(t_{n-1}) = 2f'(t_n)h + \frac{f^{(3)}(\omega)}{3}h^3$$

$$\Rightarrow f'(t_n) \approx \frac{f(t_{n+1}) - f(t_{n-1})}{2h}$$



FORMULA DELLE  
DIFF. CENTRALI

$$E(f'(t_n)) = -\frac{f^{(3)}(\delta)}{6}h^2$$

con  $\delta \in [t_{n-1}, t_{n+1}]$

## Differenze in avanti

Sviluppo con Taylor  $f(t_{m+1})$

$$f(t_{m+1}) = f(t_m) + h f'(t_m) + h^2 \frac{f''(\sigma)}{2} \Rightarrow$$

$$\Rightarrow f'(t_m) \approx \frac{f(t_{m+1}) - f(t_m)}{h}$$

$$E(f'(t_m)) = -\frac{h}{2} f''(\sigma) \quad \text{con } \sigma \in [t_m, t_{m+1}]$$

## Differenze all'indietro

Sviluppo con Taylor  $f(t_{m-1})$

$$f(t_{m-1}) = f(t_m) - h f'(t_m) + h^2 \frac{f''(\sigma)}{2} \Rightarrow$$

$$\Rightarrow f'(t_m) \approx \frac{f(t_m) - f(t_{m-1})}{h}$$

$$E(f'(t_m)) = \frac{h}{2} f''(\sigma) \quad \text{con } \sigma \in [t_{m-1}, t_m]$$

Note Queste formule hanno ordine 1 ( $E \propto h$ ) mentre le formule delle differenze centrali ha ordine 2 ( $E \propto h^2$ ) quindi è più accurate.

## EQ. DIFFERENZIALI

Differenze in problemi di Cauchy nei valori iniziali:

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases} \quad \text{dove } f : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R} \text{ con} \\ f \in C([t_0, T]) \text{ e Lipschitziana}$$

Lipschitzianità:  $f$  è Lipschitziana rispetto a  $y$  se

$$\exists L > 0 \Rightarrow \forall x, y \in \mathbb{R} : |f(t, x) - f(t, y)| \leq L |x - y|$$

$$\forall t \in [t_0, T]$$

$y(t)$  è soluzione dell'eq. diff. se rispetta le condizioni del sistema ed è  $\in C^1([t_0, T])$

Si discuterà l'intervalle di integrazione  $[t_0, T]$

$$t_{n+1} = t_n + h, \quad t_m = t_0 + mh \quad \text{con } m = 0, 1, \dots, N$$

$$\text{con } h = \frac{T - t_0}{N}$$

$$y'(t) = f(yt, y(t)) \Rightarrow$$

$$\Rightarrow \int_{t_n}^{t_{n+1}} y'(t) dt = \int_{t_n}^{t_{n+1}} f(t, y(t)) dt \Rightarrow$$

$$\Rightarrow y(t_{n+1}) - y(t_n) = \int_{t_n}^{t_{n+1}} f(t, y(t)) dt$$

### Metodo di Euler Esplicito

$$y(t_{n+1}) - y(t_n) = h f(t_n, y(t_n)) + \text{errore}$$

$$\Rightarrow y(t_{n+1}) - y(t_n) \approx h f(t_n, y(t_n))$$

$$\text{Approssimo } y_{n+1} \approx y(t_{n+1}) \quad \text{e} \quad y_n \approx y(t_n)$$

$$\Rightarrow y_{n+1} - y_n = h f(t_n, y_n) \Leftrightarrow y_{n+1} = y_n + h f(t_n, y_n)$$

È chiamato metodo di Euler esplicito in quanto, noto  $y_n$ , è possibile calcolare esplicitamente  $y_{n+1}$ .

### Metodo di Euler Implicito

$$y(t_{n+1}) - y(t_n) = h f(t_{n+1}, y(t_{n+1})) + \text{errore} \Rightarrow$$

$$\Rightarrow y(t_{n+1}) - y(t_n) \approx h f(t_{n+1}, y(t_{n+1})) \Rightarrow$$

$$\Rightarrow y_{n+1} - y_n \approx h f(t_{n+1}, y_{n+1}) \Leftrightarrow$$

$$\Rightarrow y_{n+1} = y_n + h f(t_{n+1}, y_{n+1})$$

## Metodo dei Regressi

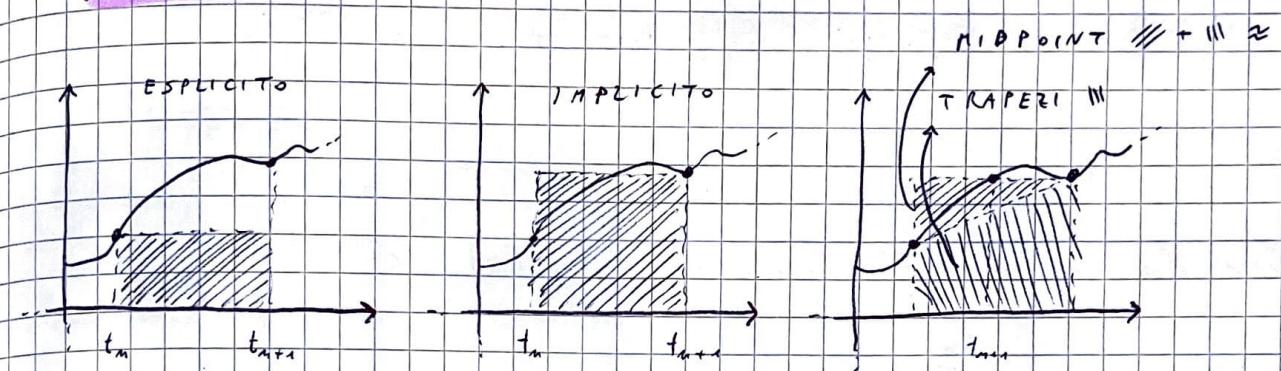
$$y(t_{m+1}) - y(t_m) \approx \frac{h}{2} [f(t_m, y_m) + f(t_{m+1}, y_{m+1})]$$

$$\Rightarrow y_{m+1} = y_m + \frac{h}{2} [f(t_m, y_m) + f(t_{m+1}, y_{m+1})]$$

## Metodo del Midpoint

$$y(t_{m+1}) - y(t_m) \approx 2h f(t_{m+1}, y(t_{m+1}))$$

$$\Rightarrow y_{m+1} = y_m + 2h f(t_{m+1}, y_{m+1})$$



## Classificazione metodi numerici

Se la soluzione nell'istante  $t_{n+k}$  dipende da  $k$  valori arbitrari in precedente  $t_{n+k-1}, \dots, t_{n+1}, t_n$   
il metodo si dice ~~di~~ a  $k$  passi ( $k$ -step)

Il metodo di Euler ~~è~~ il metodo dei regressi  
non ~~è~~ passo 1 (è un passo)

Il metodo del Midpoint è a due passi.

$$y_{n+k} = y_n + h \Phi(t_n, \dots, t_{n+k}, y_n, \dots, y_{n+k})$$

$$y_{n+1} = y_n + h \Phi(t_n, t_{n+1}, y_n, y_{n+1})$$

Se  $y_{n+1} = y_n + h \Phi(t_n, y_n)$  [o più in generale  $y_{n+k} = y_n + h \Phi(t_n, \dots, t_{n+k-1}, y_n, \dots, y_{n+k-1})$ ] il metodo è esplicito (euler esplicito + midpoint)

Se poi non fosse il metodo è implicito  
(Euler implicito + trappola)

Per calcolare  $y_{m+1}$  in un metodo implicito dovremo risolvere un'eq. che potrebbe non essere lineare se  $f(t, y)$  non è lineare.

$$y_{m+1} = y_m + h f(t_{m+1}, y_{m+1})$$

$$g(x) = y_m + h f(t_{m+1}, x)$$

$$y^{(m+1)} = g(y^{(m)}) = y_m + h f(t_{m+1}, y^{(m)})$$

$$y^{(0)} = y_m$$

Affinché il metodo converga  $|g'(x)| < 1$

$$g'(x) = h \frac{\partial f}{\partial y}(t_{m+1}, x)$$

$$\Rightarrow \left| h \frac{\partial f}{\partial y}(t_{m+1}, x) \right| < 1 \Rightarrow h \left| \frac{\partial f}{\partial y} \right| < 1 \Rightarrow$$

$$\Rightarrow \left| \frac{\partial f}{\partial y} \right| < \frac{1}{h}$$

Questa condizione è verificata sicuramente perché

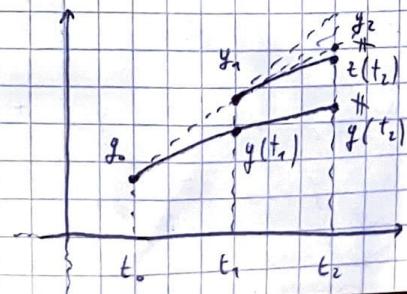
$f(t, y)$  è Lipschitziana quindi  $h \cdot L < 1$

Nel metodo di Euler Implicito

$$y = y_0 + f(t_0, y_0)(t - t_0) \quad \text{che con } t = t_0 + h$$

$$\Rightarrow y = y_0 + h f(t_0, y_0)$$

$$\Rightarrow y_{m+1} = y_m + h f(t_m, y_m)$$



Un altro metodo che consente con il metodo di Euler esplicito di calcolare le soluzioni in  $N$  intervalli:

$$t_0 < t_1 < t_2 < \dots < t_n = T$$

$$y'(t_m) = f(t_m, y_m) \approx \frac{y_{m+1} - y_m}{h} \quad \leftarrow \text{METODO DIFF. IN AVANTI}$$

Mentre  $\epsilon = h$ , non precise sarà la soluzione.

Il metodo è convergente se migliora l'approssimazione diminuendo opportunamente il passo  $h$ .

### Studio della convergenza

$e(t_m, h) := y(t_m) - y_m$  è l'errore globale

Le convergenze concrete nell'ammirazione di  $e(t_m, h)$  quando  $h \rightarrow 0$

Un schema è convergente nel p.t.  $T$  quando

$$\lim_{h \rightarrow 0} e(t_m, h) = 0 \iff \lim_{n \rightarrow +\infty} e(t_n, y_n) = g(T)$$

Supponiamo di aver applicato il metodo di Euler esplicito e indici con  $y_m$  la soluzione approssimata.

Che ne  $z(t)$  la soluzione dell'eq. differenziale che si ottiene prendendo  $y_m$  come cond. iniziale.

Nelle soluzioni globali ci sono 2 errori:

1. Errore di truncamento locale

$$I_n(h, f) := z(t_{m+1}) - [y_m + h f(t_m, y_m)]$$

Quest'errore è dovuto alla sostituzione delle soluzioni con le tangenti.

$$\tau_n(h, f) = \frac{I_n(h, f)}{h} = \frac{z(t_n + h) - z(t_n)}{h} - f(t_n, z(t_n))$$

è chiamato errore di discretizzazione locale

Osserviamo  $\tau_n \rightarrow 0$  per  $h \rightarrow 0$ ,  $\Rightarrow$  il metodo è conservante.

$\frac{z(t_n + h) - z(t_n)}{h}$  con  $h \rightarrow 0$  non è altro che la

$$\text{derivate} \Rightarrow \lim_{h \rightarrow 0} \frac{z(t_n + h) - z(t_n)}{h} = z'(t_n)$$

$$\Rightarrow \tau_n(h, f) = z'(t_n) - z'(t_n, z(t_n)) = 0$$

(perché  $f(t_n, z_n) = z'(t_n)$ )

quindi il metodo di Euler Explicito è conservante.

Inoltre scrivo, espandendo  $z(t_n + h)$  con Taylor:

$$z(t_n + h) = z(t_n) + zh z'(t_n) + \frac{h^2}{2} z''(t_n + \theta h)$$

$$\text{con } \theta \in (0, 1)$$

Perciò che  $z'(t_n) = f(t_n, y_n) = f(t_n, z(t_n))$

$$\text{si ha } I_n(h, f) = z(t_{n+1}) - [y_n + h f(t_n, \frac{y_n}{2} + \theta h)] =$$

$$= z(t_n + h) - z(t_n) - h z'(t_n) = \frac{h^2}{2} z''(t_n + \theta h)$$

$$\Rightarrow \tau_n(h, f) = \frac{I_n(h, f)}{h} = \frac{h}{2} z''(t_n + \theta h)$$

$$\Rightarrow |\tau_n(h, f)| \leq M \cdot h \text{ con } M = \frac{1}{2} \max_{\theta \in (0,1)} |z''(t_n + \theta h)|$$

$\tau(f, h) = O(h^2) \Rightarrow$  metodo di ordine n

In questo caso  $\tau(f, h) \propto h \Rightarrow \tau = O(h^2) \Rightarrow$

$\Rightarrow$  il metodo di Euler esplica l'ordine 1

2. Considero la tangente passante per il p.t.  
 $(t_n, y_n)$  e non per  $(t_n, y(t_n))$

Questo contributo non è locale in quanto rappresenta l'ecciduo degli errori di truncamento locale connessi in precedente.

Non è garantito che questi errori tendano a 0 quando  $h \rightarrow 0$

In sostanzia: non è detto che la convergenza a 0 degli errori locali porti alla convergenza a 0 dell'errore globale.

$$e(x) = 0 \Leftrightarrow \tau_n(f, h) \rightarrow 0 \text{ con } h \rightarrow 0 \forall n$$

Il metodo è stabile se l'accumulo degli errori.

Li truncamento locale si mantiene limitato per  $h \rightarrow 0$

$$\text{CONVERGENZA} \Leftrightarrow \text{STABILITÀ} + \text{CONSISTENZA}$$