



Corso Luigi Einaudi, 55 - Torino

Appunti universitari

Tesi di laurea

Cartoleria e cancelleria

Stampa file e fotocopie

Print on demand

Rilegature

NUMERO : 211

DATA : 17/02/2012

A P P U N T I

STUDENTE : Blai

MATERIA : Metodi Numerici e Calcolo Scientifico,
Prof. Puppo

Il presente lavoro nasce dall'impegno dell'autore ed è distribuito in accordo con il Centro Appunti.

Tutti i diritti sono riservati. È vietata qualsiasi riproduzione, copia totale o parziale, dei contenuti inseriti nel presente volume, ivi inclusa la memorizzazione, rielaborazione, diffusione o distribuzione dei contenuti stessi mediante qualunque supporto magnetico o cartaceo, piattaforma tecnologica o rete telematica, senza previa autorizzazione scritta dell'autore.

**ATTENZIONE: QUESTI APPUNTI SONO FATTI DA STUDENTIE NON SONO STATI VISIONATI DAL DOCENTE.
IL NOME DEL PROFESSORE, SERVE SOLO PER IDENTIFICARE IL CORSO.**

Politecnico di Torino

I Facoltà di Ingegneria

Corso di Laurea Magistrale in Ingegneria Aerospaziale



Metodi numerici e calcolo scientifico

Prof. Gabriella Puppo

Anno Accademico 2010/2011

Introduzione

Esame

- Scritto 2h (con appunti, due domande, voto minimo necessario 15/30).
- Orale (solitamente due domande, fa media con lo scritto).

Bibliografia

- Calcolo numerico di base:
 - MONEGATO; *100 pagine di calcolo numerico*
 - QUARTERONI; SACCO, SALERI; *Matematica numerica* (Springer)
- MATLAB[®] – Help:
 - HIGHAM & HIGHAM; *MATLAB[®] Users Guide* (SIAM)
- Problemi di evoluzione:
 - LEVEQUE; *Finite Difference, Methods for Ordinary and Partial Differential Equations* – SIAM '07 CLAW PACK
 - LEVEQUE; *Numerical Methods for Conservation Laws*; '90 (ETH – Zurich)
 - JOHNSON; *Finite Element Methods for Partial Differential Equations*; Cambridge University press '70

$$x_{\min} = \pm 0,1000000 \cdot 10^L \sim 10^{-308}$$

$$= \pm 0,0000001 \cdot 10^L \quad \rightarrow \text{Altra rappresentazione}$$

$x \equiv$ numero che voglio rappresentare
 $\tilde{x} \equiv$ rappresentazione di macchina di x

Se $|x| < x_{\min}$ Underflow $\tilde{x} = 0$

Se $|x| > x_{\max}$ Overflow $\tilde{x} = Inf$

$0 \cdot Inf = NaN$ (Not a Number)

Se $|x| \in (x_{\min}, x_{\max})$ scrivo x in forma normalizzata:

$$x = \pm 0, a_1 a_2 a_3 \dots a_t | a_{t+1} a_{t+2} \beta^p$$

$$\tilde{x} = \pm 0, a_1 a_2 a_3 \dots a_t \beta^p \quad (\text{troncamento})$$

$$\tilde{x} = \pm 0, a_1 a_2 a_3 \dots \tilde{a}_t \beta^p \quad (\text{arrotondamento})$$

Se $a_{t+1} \cdot a_{t+2} \dots < \frac{1}{2} \beta \quad \tilde{a}_t = a_t$

else $\tilde{a}_t = a_t + 1$

Esempio:

0.123 453	0.123 512
$t = 3 \quad \beta = 10$ $a_{t+1} \cdot a_{t+2} a_{t+3} = 4.53 < \frac{1}{2} 10$ $\tilde{a}_t = a_t = 3$	$t = 3 \quad \beta = 10$ $a_{t+1} \cdot a_{t+2} a_{t+3} = 5.12 > \frac{1}{2} 10$ $\tilde{a}_t = a_t + 1 = 4$

L'errore assoluto nell'approssimazione è :

$$e = |x - \tilde{x}| = |0, a_1 a_2 a_3 \dots a_t a_{t+1} - 0, a_1 a_2 \dots \tilde{a}_t | \beta^p = |0,000 \dots a_t - \tilde{a}_t | \beta^p \leq \frac{1}{2} \beta^{-t} \beta^p$$

l'errore dipende da quanto è grande il numero x .

L'errore relativo nell'approssimazione è:

$$e_{\text{rel}} = \frac{|x - \tilde{x}|}{|x|} \leq \frac{1}{2} \frac{\beta^{-t} \beta^p}{|m| \beta^p} \leq \frac{1}{2} \beta^{1-t}$$

$$x = m \beta^p$$

$m =$ mantissa

Test: deve essere fatto su errori relativi e basandosi su eps :

Se ho un ciclo iterativo: misuratore di errore = \tilde{e}_{rel}

Esci se $|\tilde{e}_{rel}| \leq \text{toleranza} = K \cdot eps$

x : vettore

\tilde{x} : vettore dei risultati esatti

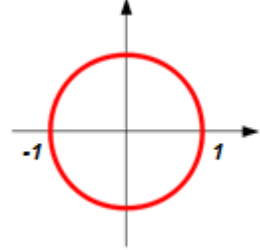
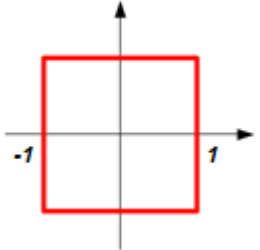
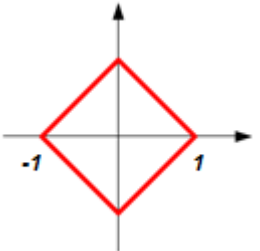
$x, \tilde{x} \in \mathbb{R}^n$

$$e = x - \tilde{x} \quad \in \mathbb{R}^n$$

distanza fra A e B :

$$|x_A - x_B| + |y_A - y_B| = \|A - B\|_1$$

$\mathbb{R}^2 \rightarrow 2$ è il numero di dimensione

Cerchio unitario	$\{x \in \mathbb{R}^2 \text{ t.c. } \ x\ _2 \leq 1\}$	
Quadrato unitario	$\{x \in \mathbb{R}^2 \text{ t.c. } \ x\ _\infty \leq 1\}$ $\max(x_1 , x_2) = 1$	
Rombo unitario	$\{x \in \mathbb{R}^2 \text{ t.c. } \ x\ _1 = 1\}$	

Autovalori e autovettori:

Un vettore $x \in \mathbb{R}^n$ è un autovettore di A se: $\exists \lambda \in \mathbb{C}$

$$\begin{aligned} Ax &= \lambda x \\ x &\neq 0 \end{aligned}$$

Ogni matrice $n \times n$ ha n autovalori $\lambda_1, \lambda_2, \dots, \lambda_n$.

Raggio spettrale di A :

$$\rho(A) = \max |\lambda_i| \quad 1 \leq i \leq n$$

Se $Ax = \lambda x$:

$$A(\alpha x) = \alpha Ax = \alpha \lambda x$$

Il vettore trasformato ha sempre la stessa direzione.

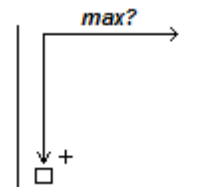
Calcolo l'effetto di A su tutti i vettori x tale che:

$$\|x\|_p = 1 \quad \text{per } p = 2, \quad \|x\|_2 = 1$$

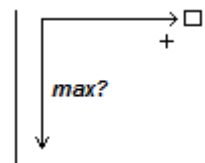
$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 \equiv \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Con MATLAB® si usa il comando: `norm(a, p)` $p = 1, p = 2, p = \infty$

$$\|A\|_1 = \max_j \left(\sum_{i=1}^n |a_{ij}| \right) = \text{max sulla somma lungo le colonne}$$



$$\|A\|_\infty = \max_i \left(\sum_{j=1}^n |a_{ij}| \right) = \text{max sulla somma lungo le righe}$$



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

Se A è simmetrica, allora A ha n autovalori $\lambda \in \mathbb{R}$ e n autovettori linealmente indipendenti

$$A^T = A \Rightarrow \|A\|_2 = \sqrt{\rho(A \cdot A)} = \sqrt{\rho(A)^2} = \rho(A)$$

$$\frac{\|x - \tilde{x}\|}{\|x\|_p} \leq K_p(A) \frac{\|b - \tilde{b}\|_p}{\|b\|_p}$$

$$Ax = b$$

$$b - \tilde{b} = Ax - A\tilde{x} = A(x - \tilde{x})$$

$$A\tilde{x} = \tilde{b}$$

$$\|x - \tilde{x}\| = \|A^{-1}(b - \tilde{b})\| \leq \|A^{-1}\| \cdot \|b - \tilde{b}\|$$

$$\|Ax\| = \|b\| \Rightarrow \|A\| \cdot \|x\| \geq \|b\|$$

Quello che ho dimostrato è:

$$K_p(A) = \|A\|_p \cdot \|A^{-1}\|_p$$

Se A è simmetrica e, come solitamente, $p = 2$:

$$\|A\| = \max |\lambda_i|$$

$$\|A^{-1}\| = \max |\text{autovalori di } A^{-1}| = \max \left| \frac{1}{\lambda} \right|$$

Infatti se w è autovettore di A con autovalori λ :

$$\|Aw\| = \lambda w$$

$$w = A^{-1}Aw = A^{-1}(\lambda w) = \lambda A^{-1}w$$

$$\lambda A^{-1}w = w \Rightarrow A^{-1}w = \frac{1}{\lambda} w$$

↳ non si può fare se $\lambda = 0$

Se A è simmetrica e $p = 2$:

$$K(A) = \frac{\max |\lambda_i|}{\min |\lambda_i|} \geq 1$$

Una matrice mal condizionata è quella che ha $r_{\text{cond}} \sim 10^{-12}$.

Su MATLAB®:

$$Ax = b \rightarrow x = A \setminus b$$

Warning: matrix is nearly singular or badly conditioned

$$r_{\text{cond}} = \frac{1}{K(A)} \sim 10^{-16}$$

Results may be inaccurate

Approssimazione di funzioni

1) Ho una una funzione f "complicata". Vorrei trovare un'approssimazione f_n t.c.:

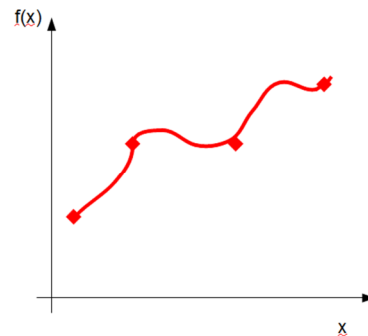
$$f - f_n \text{ piccola e } f_n \text{ sia più facile da manipolare.}$$

2) Ho dei dati, per esempio una tabella di misure sperimentali:

$$(x_1, f_1), (x_2, f_2), \dots, (x_n, f_n)$$

$$f_i = f(x_i) \text{ solo che non conosco } f.$$

Io sono chi sceglie la griglia, poiché conosco la funzione, e trasformo il sistema complicato in uno più semplice ma non esatto, solo approssimato.



Esempio caso 1:

$$\int_a^b f(x) dx$$

Cerco un polinomio di grado n , $P^n(x)$, che approssima f oppure polinomi trigonometrici di grado n , $T^n(x)$.

N.B: se funzione non ciclica, non esiste la serie di Fourier.

- **Interpolazione polinomiale globale**

È data una funzione f regolare (cioè $f \in C^\infty[a, b]$, continua tra a e b).

1) Scelgo dei punti $x_1, x_2, \dots, x_{n+1} \in [a, b]$ (griglia di interpolazione)

2) Calcolo la funzione f sulla griglia di interpolazione

$$f_1 = f(x_1)$$

$$f_{n+1} = f(x_{n+1})$$

3) Calcolo il polinomio $P^n(x)$ che interpola i dati (x_i, f_i) $i = 1, \dots, n + 1$

Per trovare P^n devo trovare i coefficienti $[a_1, a_2, \dots, a_{n+1}]$.

\Rightarrow Ho $n + 1$ incognite con $n + 1$ equazioni.

Ho il sistema:

$$a_1 x_1^n + a_2 x_1^{n-1} + \dots + a_{n+1} = f_1$$

$$a_1 x_2^n + a_2 x_2^{n-1} + \dots + a_{n+1} = f_2$$

$$a_1 x_{n+1}^n + a_2 x_{n+1}^{n-1} + \dots + a_{n+1} = f_{n+1}$$

Che in modo matriciale risulta:

$$\begin{bmatrix} x_1^n & x_1^{n-1} & x_1 & 1 \\ x_2^n & x_2^{n-1} & x_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_{n+1}^n & x_{n+1}^{n-1} & x_{n+1} & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n+1} \end{bmatrix}$$

Matrice di Vandermonde

$$V_n = V(x_1, \dots, x_{n+1})$$

$\det V_n \neq 0$ se i nodi sono distinti $\Rightarrow \exists$ una sola soluzione $[a_1, a_2, \dots, a_{n+1}]$.

Però $K(V_n) \rightarrow \infty$ per $n \rightarrow \infty$. (K è il numero di condizionamento)

(N.B.: La K determina l'amplificazione massima dell'errore. Quando n cresce, anche K cresce, però se n è troppo piccola, l'errore è alto. Quindi, si deve scegliere un valore di n che ci permetta ottenere un errore piccolo e una K non molto grande).

$P^n = \{\text{spazio dei polinomi di grado } \leq n\}$

Spazio lineare

Si possono creare polinomi mediante combinazioni lineari:

$$p \in P^n \quad \text{e} \quad \alpha p \in P^n$$

$$p_1, p_2 \in P^n \quad \rightarrow \quad p_1 + p_2 \in P^n$$

$\dim(P^n) = \text{numero gradi di libertà} = \text{nel nostro caso: } n + 1$

Posso costruire $n + 1$ funzioni linearmente indipendenti tra loro $\Phi_1(x), \Phi_2(x), \dots, \Phi_{n+1}(x)$ che formano una base di P^n cioè ogni $p \in P^n(x)$ può essere scritto come combinazione lineare delle funzioni della base.

Cioè esistono dei coefficienti $c_1, c_2, \dots, c_{n+1}(x)$ tale che:

$$p(x) = c_1 \Phi_1(x) + c_2 \Phi_2(x) + c_{n+1} \Phi_{n+1}(x)$$

$$P(x) = \sum_{i=1}^{n+1} c_i L_i(x)$$

Il polinomio di interpolazione

$$\begin{aligned} P(x_j) &= \sum_{i=1}^{n+1} c_i L_i(x_j) = f_j \\ &= c_1 L_1(x_j) + \dots + c_j L_j(x_j) + \dots = f_j \quad \text{con } L_1(x_j) = 0 \\ &= c_j L_j(x_j) = c_j = f_j \end{aligned}$$

Nella base $L_i(x)$ i coefficienti del polinomio $[c_1, c_2 \dots c_{n+1}] = [f_1, f_2 \dots f_{n+1}]$.

La matrice $[\Phi_i(x_j)] = I \rightarrow$ matrice identità

- **Base di Newton**

Un'altra base Φ_j è la base di Newton:

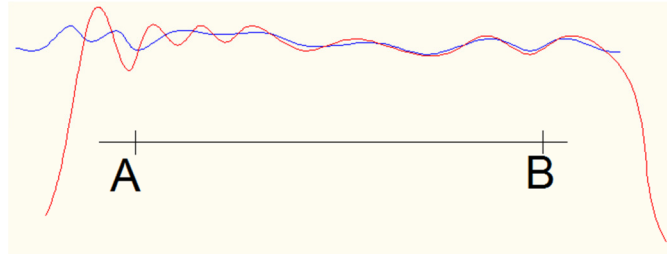
$$\begin{aligned} \Phi_1(x) &= 1 \\ \Phi_2(x) &= (x - x_1) \\ \Phi_3(x) &= (x - x_1)(x - x_2) \\ &\vdots \\ \Phi_{n+1}(x) &= (x - x_1)(x - x_2) \dots (x - x_n) \end{aligned}$$

Troviamo così una matrice triangolare

Estrapolazione

Uso il polinomio $P^n(x)$ di interpolazione per stimare f fuori dell'intervallo (a, b) .

$\exists [x_1, x_2, \dots, x_{n+1}]$



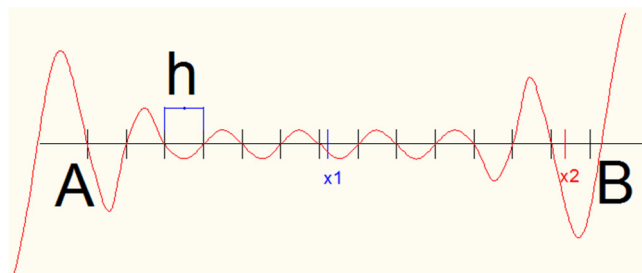
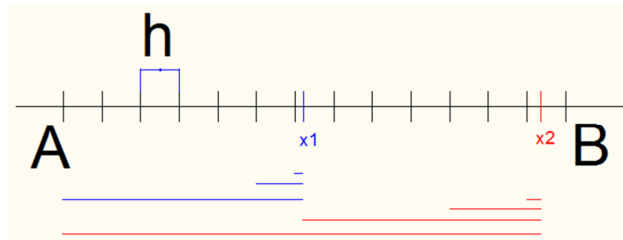
$w_{n+1}(x)$ cresce velocemente per x che si allontana da a o b .

La griglia più logica di scegliere è la griglia equispaziata:

$$a = x_1$$

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

$$b = x_{n+1}$$



Più mi sposto al bordo più cresce il contributo delle funzioni nodali, per cui devo scegliere una griglia che sia più fitta ai bordi affinché abbiamo meno errore.

Per controllare la funzione nodale ho bisogno di una griglia che sia più fitta ai bordi.

Numero di condizionamento per l'interpolazione polinomiale

Conosco la griglia x_1, x_2, \dots, x_{n+1} .

Chiamo $f_i = f(x_i)$ i dati da interpolare.

Chiamo \tilde{f}_i i dati perturbati.

Vorrei stimare l'errore, che a questo punto è $f_i - \tilde{f}_i$, si propaga sui valori del polinomio di interpolazione.

$P^n(x)$ il polinomio $P^n(x) = f_i \quad i = 1, \dots, n+1 \quad \rightarrow \quad$ dati esatti

$\tilde{P}^n(x)$ il polinomio $\tilde{P}^n(x) = \tilde{f}_i \quad i = 1, \dots, n+1 \quad \rightarrow \quad$ dati perturbati

Vorrei valutare la differenza $P^n(x) - \tilde{P}^n(x)$. Dimostrazione (*non c'è bisogno di saperla*).

$$|P^n(x) - \tilde{P}^n(x)| \leq ? |f_i - \tilde{f}_i|$$

$$P^n(x) - \tilde{P}^n(x) = \sum_{i=1}^{n+1} f_i \mathcal{L}_i(x) - \sum_{i=1}^{n+1} \tilde{f}_i \mathcal{L}_i(x)$$

$$|P^n(x) - \tilde{P}^n(x)| = \left| \sum_{i=1}^{n+1} (f_i - \tilde{f}_i) \mathcal{L}_i(x) \right|$$

$$|P^n(x) - \tilde{P}^n(x)| \leq \sum_{i=1}^{n+1} |f_i - \tilde{f}_i| \cdot |\mathcal{L}_i(x)|$$

$$|P^n(x) - \tilde{P}^n(x)| \leq \max_i |f_i - \tilde{f}_i| \sum_{i=1}^{n+1} |\mathcal{L}_i(x)|$$

Calcolo il valore massimo sia a sinistra che a destra.

$$\max_x |P^n(x) - \tilde{P}^n(x)| \leq \|f_i - \tilde{f}_i\|_\infty \cdot \max_x \sum |\mathcal{L}_i(x)|$$

$$\|P^n - \tilde{P}^n\|_\infty \leq \|f - \tilde{f}\|_\infty \cdot \left\| \sum |\mathcal{L}_i(x)| \right\|_\infty$$

Ove:

$$\left\| \sum |\mathcal{L}_i(x)| \right\|_\infty = \Lambda^n \quad \text{Costante di Lebesgue}$$

Test

- Test “buono”:

$$f(x) = (\cos x) \cdot e^x$$

- Test “cattivo”

$$f(x) = \text{gradino}$$

$$f(x) = |x|$$

$$f(x) = \text{discontinua}$$

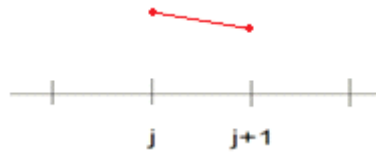
- Test di Runge

$$f(x) = \frac{1}{1+x^2} \quad x = \pm i \quad \text{polinomi}$$

La interpolazione polinomiale richiede polinomi di grado relativamente basso.

Per polinomi di grado relativamente alto ($n > 10$) usare griglie gaussiane.

2) $d = 1$: ho 2 nodi



$$x_{j,1} = x_j$$

$$x_{j,2} = x_{j+1}$$

$$P_j^1 = f_j + \frac{f_{j+1} - f_j}{x_{j+1} - x_j} (x - x_j)$$

I_h^1 è continuo su $[a, b]$.

3) $d = 2$: ho 3 nodi

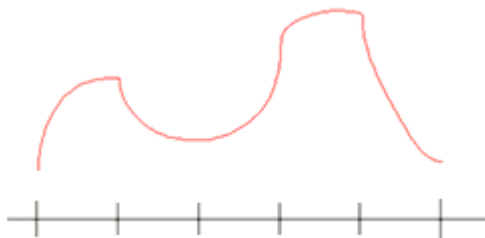


$$x_{j,1} = x_j$$

$$x_{j,2} = x + \frac{h}{2} = x_{j+\frac{1}{2}}$$

$$x_{j,3} = x_{j+1}$$

$$P_j^2 = f(x_j) \cdot \mathcal{L}_{j,1}(x) + f\left(x_{j+\frac{1}{2}}\right) \cdot \mathcal{L}_{j,2}(x) + f(x_{j+1}) \cdot \mathcal{L}_{j,3}(x)$$



$I_h^2(x)$ è continuo su $[a, b]$ ma $(I_h^2(x))'$ è discontinua.

4) $d = 3$: ho 4 nodi



Stima di errore:

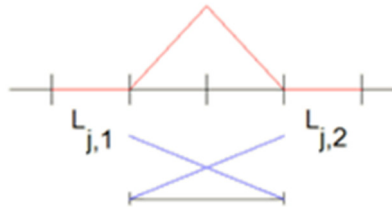
Su ogni intervallo I_j :

$$e_j(x) = f(x) - P_j^d(x) \leq \frac{1}{(d+1)!} \cdot |w_{j,d+1}(x)| \cdot |f^{d+1}(\xi)| \quad \xi \in [x_j, x_{j+1}]$$

$$\begin{aligned}
 dof &= (\text{numero per intervallo}) \cdot M - \text{numero di condizioni di raccordo} \\
 &= (d + 1) \cdot M - (M - 1) \\
 &= d M + 1
 \end{aligned}$$

Base globale:

$$d = 1$$



$$L_{j,l}(x_i) = \begin{cases} 1 & i = l, j \\ 0 & \text{else} \end{cases}$$

Questa volta $L_{j,l}(x_i)$ non è un polinomio ma una funzione polinomiale a tratti.

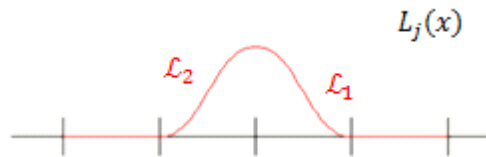
Interpolazione di Hermite:

Vorrei un interpolante che ha derivata prima continua.

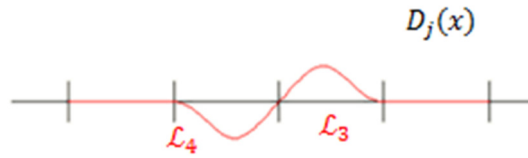
Abbiamo bisogno di polinomi di grado 3 tale che su ogni intervallo I_j :

$$\begin{array}{lll}
 I_h(x_j) = f(x_j) & I_h'(x_j) = f'(x_j) & \\
 I_h(x_{j+1}) = f(x_{j+1}) & I_h'(x_{j+1}) = f'(x_{j+1}) & \text{2 condizioni su ogni nodo}
 \end{array}$$





Trasporta il valore $f(x_j)$.



L'interpolante deve interpolare il valore della funzione.

L'interpolante di Hermite di f è:

$$I_h^3(x) = \sum_{j=0}^M f(x_j) L_j(x) + \sum_{j=0}^M f'(x_j) D_j(x)$$

Infatti se calcolo:

$$I_h^3(x) = \sum_{j=0}^M f(x_j) L_j(x_i) + 0 = f(x_i) L_i(x_i) = f(x_i)$$

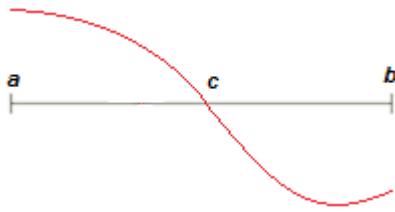
$L_i(x_i)$ fatto in maniera di valere 1 in un nodo solo

Inoltre:

$$\begin{aligned} (I_h^3)'|_{x_i} &= \sum_{j=0}^M f(x_j) L_j'(x_i) + \sum f'(x_i) D_j'(x_i) && \text{con } L_j'(x_i) = 0 \\ &= f'(x_i) \cdot D_i'(x_i) && \text{con } D_i'(x_i) = 1 \\ &= f'(x_i) \end{aligned}$$

$(M + 1)$ funzioni L_i

$(M + 1)$ funzioni D_i

Metodo di bisezione

Parto da:

$$\begin{aligned} a_0 &= a & c &= \frac{a+b}{2} \\ b_0 &= b \end{aligned}$$

$$\text{Se } f(a) * f(c) \leq 0 \quad \begin{cases} a_1 = a_0 \\ b_1 = c \end{cases}$$

$$\text{else } \begin{cases} a_1 = c \\ b_1 = b_0 \end{cases}$$

Su MATLAB®

Pongo $a_0 = a$, $b_0 = b$.

```
for k = 0,1,...
    c = (a_k + b_k)/2
```

```
    if f(a_k)f(c) ≤ 0           ⇒ x̄ ∈ [a_k, c]
```

$$\begin{aligned} a_{k+1} &= a_k \\ b_{k+1} &= c \end{aligned}$$

```
    else                       ⇒ x̄ ∈ [c, b_k]
```

$$\begin{aligned} a_{k+1} &= c \\ b_{k+1} &= b_k \end{aligned}$$

```
end
```

Test di arresto (Sulle iterazioni successive)

$$\text{Se } (b_k - a_k) < \text{tol}(b - a)$$

$$\bar{x} = c$$

Oppure (sul residuo):

$$\text{Se } |f(c)| \leq \text{tol} \cdot \frac{|f(a)+f(b)|}{2}$$

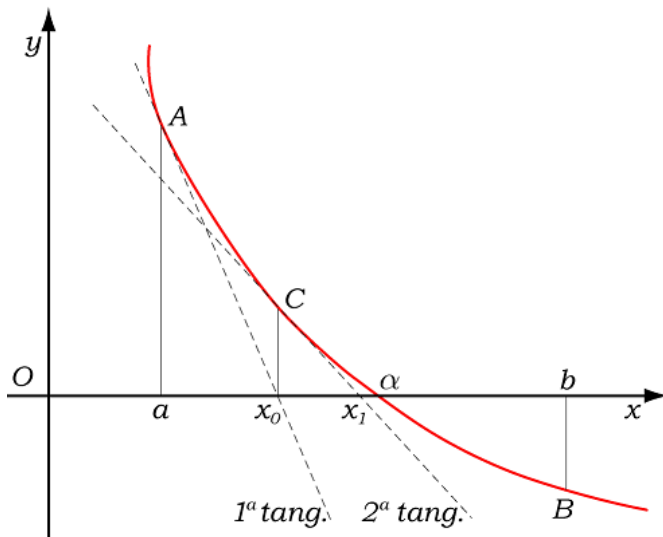
$$e_k = |\bar{x} - c_k| \leq (b_k - a_k) = \frac{1}{2}(b_{k-1} - a_{k-1}) = \frac{1}{2^k}(b - a)$$

Il metodo di bisezione converge di maniera lenta ($\frac{1}{2}$ per passo).

Metodo di Newton

Utile se c'è più di un valore \bar{x} tale che $\bar{x} = 0$

Se funziona, questo metodo converge molto velocemente.



$$y = f_i(x_k) + f'(x_k)(x - x_k)$$

$$x_{k+1} \quad \text{t.c.} \quad y = 0$$

$$0 = f(x_k) + f'(x_k)(x_{k+1} - x_k)$$

E finalmente otteniamo

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

\tilde{x} la soluzione esatta: $f(\tilde{x}) = 0$

Convergenza del metodo di Newton:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} = g(x_k)$$

Il metodo di Newton converge se

$$|g'(\tilde{x})| < 1$$

$$g'(x_k) = 1 - \frac{[f'(x_k)]^2 - [f(x_k)f''(x_k)]}{[f'(x_k)]^2} = \left| \frac{[f(x_k)f''(x_k)]}{[f'(x_k)]^2} \right| < 1$$

** Errore per il metodo di Newton (questa dimostrazione non entra all'esame):

$$0 = f(\tilde{x}) = f(x_k) + f'(x_k)(\tilde{x} - x_k) + \frac{1}{2}f''(\xi)(\tilde{x} - x_k)^2 \quad \text{con } \xi \in (\tilde{x}, x_k)$$

$$(x_k - x_{k+1}) = \frac{f(x_k)}{f'(x_k)}$$

$$0 = f(\tilde{x}) = \frac{f(x_k)}{f'(x_k)} + \frac{f'(x_k)}{f'(x_k)}(\tilde{x} - x_k) + \frac{1}{2} \frac{f''(\xi)}{f'(x_k)}(\tilde{x} - x_k)^2$$

$$0 = x_k - \tilde{x} - (x_{k+1} - \tilde{x}) + (\tilde{x} - x_k) + \frac{1}{2} \frac{f''(\xi)}{f'(x_k)}(\tilde{x} - x_k)^2$$

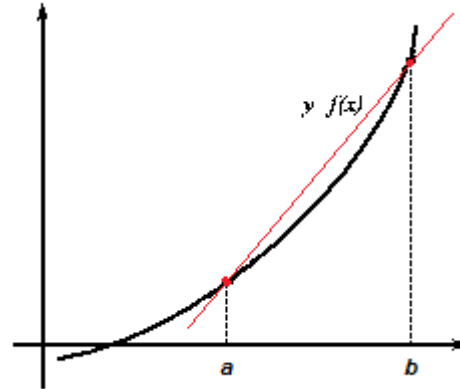
Metodo delle secanti

Per evitare che l'utente debba calcolare f' nel Metodo di Newton uso il metodo delle secanti.

Ho bisogno di 2 valori iniziali x_k, x_{k-1}

Inizialmente $x_0 = a, x_{-1} = b$

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}$$



Per il metodo delle secanti:

$$e_{k+1} \cong c(f) \cdot e_k^{1,7}$$

Esempio:

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

$$[a, b] = [1, 2]$$

$$\begin{cases} f(x_k) = f(1) = -1 \\ f(x_{k-1}) = f(2) = 2 \end{cases}$$

$$x_{k+1} = 1 - \left(\frac{-1(1-2)}{-1-2} \right) = \frac{4}{3}$$

$$\begin{cases} f(x_k) = f(4/3) = -0.22 \\ f(x_{k-1}) = f(1/3) = -1.89 \end{cases}$$

$$x_{k+2} = \frac{4}{3} - 0.22 \left(\frac{4}{3} - 2 \right) - 0,22 - 2 = 1.39$$

$$\begin{cases} f(x_k) = f(1.39) = -0.0679 \\ f(x_{k-1}) = f(4/3) = -0.22 \end{cases}$$

$$x_{k+3} = 1,39 - \left(\frac{-0.0679 \left(1.39 - \frac{4}{3} \right)}{-0.0679 - (-0.22)} \right) = 1,4153 \approx \sqrt{2}$$

Metodo di punto fisso

So che $f(\bar{x}) = 0$

Trasformo il problema della ricerca di \bar{x} , nella ricerca di un punto fisso. $g(\bar{x}) = \bar{x}$

Esempio:

$$f(x) = x^2 - 2 \quad \rightarrow \quad (\bar{x} = \sqrt{2})$$

$$g(x) = x^2 - 2 + x \quad \rightarrow \quad g(\sqrt{2}) = \sqrt{2}$$

Un'altra possibilità è:

$$g(x) = x - \frac{x^2 - 2}{2x} \quad \rightarrow \quad g(\sqrt{2}) = \sqrt{2}$$

Sistemi non lineari

$$f = \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \text{voglio } x \in \mathbb{R}^n$$

$$f(x) = 0$$

Non posso usare il metodo di bisezione.

Dati: $a, b \in \mathbb{R}^n$, $f(a), f(b) \in \mathbb{R}^n$ $f(a), f(b)$ non hanno un segno

Posso usare il metodo di Newton:

$$0 = f(x_k) + f|_{x_k} (x_{k+1} - x_k)$$

N.B: $f'|_k = Jf|_{x_k}$

Dato x_k calcolo $Jf|_{x_k}$

Risolvo il sistema lineare

$$Jf|_{x_k} \delta = -f(x_k)$$

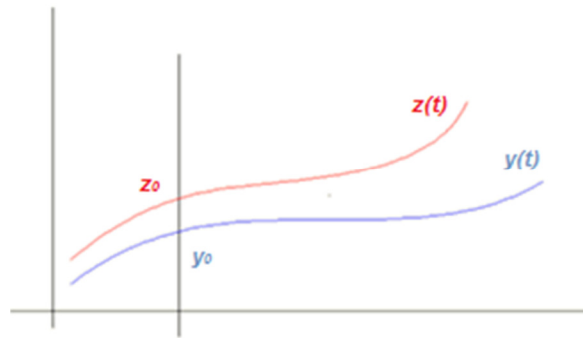
Trovo δ

$$x_{k+1} = x_k + \delta$$

Posso anche usare il metodo di punto fisso

Dato x_k : $x_{k+1} = g(x_k)$ converge se $\|Jg\| < 1$

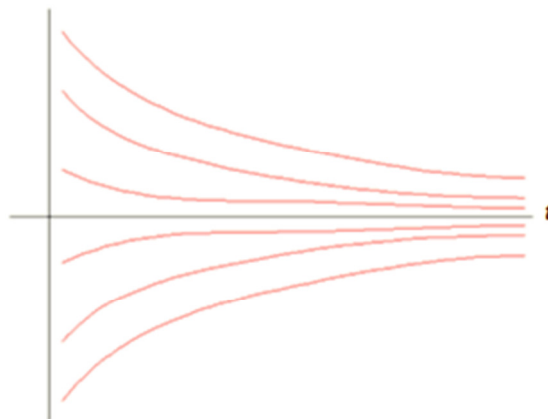
$$|y(t) - z(t)| \leq K(t)|y_0 - z_0| \quad K = \text{Costante allontanamento che dipende dal tempo}$$



Si dimostra che $|y(t) - z(t)| \leq e^{Lt}|y_0 - z_0|$

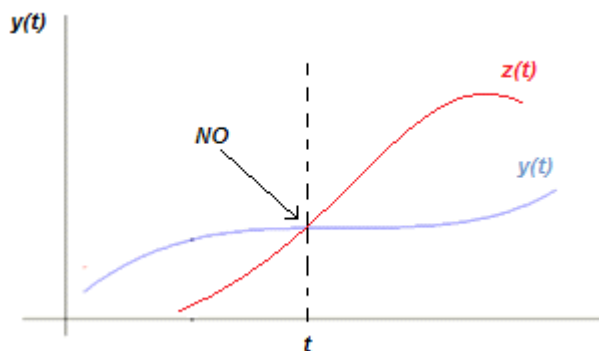
Se L è grande $\rightarrow y(t)$ e $z(t)$ si allontanano velocemente anche se $z_0 - y_0$ è piccolo.

Se la soluzione tende ad uno stato stazionario la dipendenza dai dati iniziali è meno importante.



Se vale il Teorema di $\exists!$ (esistenza e unicità)

\Rightarrow 2 soluzioni non si possono mai incrociare.



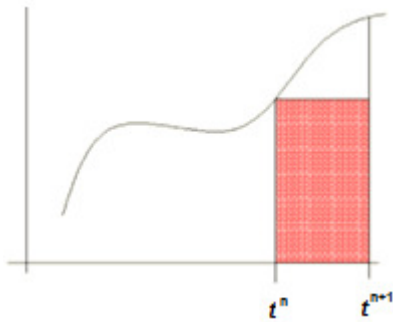
z e y avrebbero lo stesso valore iniziale

Dipendenza dei dati iniziali

Soluzione esatta:

$$\begin{cases} y'(t) = f(t, y) \\ y(t^n) = u^n \end{cases} \quad y(t^{n+1}) = u^n + \int_{t^n}^{t^n+\Delta t} f(t, y) dt$$

Costruisco la soluzione numerica approssimando l'integrale.



L'area nota è $f(t^n, u^n) \cdot \Delta t$

$$\boxed{u^{n+1} = u^n + \Delta t f(t^n, u^n)} \quad \text{Metodo di Eulero esplicito}$$

- **Eulero esplicito:**

Esempio:

```

su MATLAB®:
t^0 = t_0 , u^0 = y_0
for n = 0,1, ... n_max
    u^{n+1} = u^n + Δt f(t^n, u^n)   Metodo di Eulero esplicito
    t^{n+1} = t^n + Δt               Metodo di Eulero esplicito
end
plot(t, u)
    
```

- Risolvere l'equazione differenziale.

- Trovare $u(t)$ $0 \leq t \leq T$

$$n_{\max} = \frac{T}{\Delta t}$$

- **Eulero implicito:**

$$\int_{t^n}^{t^{n+1}} f(t, y) dt \simeq \Delta t f(t^{n+1}, u^{n+1})$$

$$\boxed{u^{n+1} = u^n + \Delta t f(t^{n+1}, u^{n+1})}$$

Per trovare u^{n+1} devo risolvere l'equazione definita sopra.

Immaginando di partire all'istante t_n dalla soluzione esata, y_n ,

$$l_{\Delta t}^n = y(t^n + \Delta t) - u^{n+1} \quad \text{nell'ipotesi } u^n = y^n \text{ (localizzazione)}$$

Errore locale di troncamento (LTE)

$$T_{\Delta t}^n = \frac{l_{\Delta t}^n}{\Delta t} = \frac{y(t^n + \Delta t) - u^{n+1}}{\Delta t}$$

Con $u^{n+1} = u^n + \Delta t f(t_n, u^n)$

$$T_{\Delta t}^n = [y(t^n + \Delta t) - [u^n + \Delta t f(t_n, u^n)]] \frac{1}{\Delta t}$$

Ove $u^n + \Delta t f(t_n, u^n)$ viene riferito a Eulero Esplicito.

Uso $u^n = y^n$.

$$T_{\Delta t}^n = \frac{1}{\Delta t} \{y(t^n + \Delta t) - [y(t^n) + \Delta t f(t_n, y^n)]\}$$

So che $y' = f$ e che y può essere sviluppata in serie.

Sviluppo y intorno a $t = t_n$.

$$y(t^n + \Delta t) = y^n + \Delta t y'|_{t^n} + \frac{1}{2} (\Delta t)^2 y''|_{t^n} + O(\Delta t)^3$$

$$T_{\Delta t}^n = \frac{1}{\Delta t} \left\{ y^n + \Delta t y'(t_n) + \frac{\Delta t^2}{2} y''(t_n) + O(\Delta t)^3 - [y^n + \Delta t f(t_n, y^n)] \right\}$$

$$T_{\Delta t}^n = \frac{1}{\Delta t} \left\{ \Delta t (y'(t_n) - f(t_n, y^n)) + \frac{1}{2} (\Delta t)^2 y''(t_n) + O(\Delta t)^3 \right\} \Rightarrow y'(t_n) - f(t_n, y^n) = 0$$

$$\boxed{T_{\Delta t}^n = \frac{1}{2} \Delta t y''(t_n) + O(\Delta t)^2} \quad \text{LTE per E.E}$$

Consistenza: un metodo numerico è consistente se:

$$T_{\Delta t} \rightarrow 0 \quad \text{per } \Delta t \rightarrow 0.$$

Accuratezza: un metodo numerico è accurato di ordine p se

$$T_{\Delta t}^n = O(\Delta t)^p \quad \text{per } \Delta t \rightarrow 0.$$

Eulero è consistente ed accurato di ordine 1.

Esempio:Calcolo di $u^{(l)}$ nel Runge-Kutta esplicito:

$$\begin{aligned} u^{(1)} &= u^n + \Delta t \sum_{l=1}^0 a_{1l} f(t + c_1 \Delta t, u^{(1)}) \\ &= u^n \end{aligned}$$

$$\begin{aligned} u^{(2)} &= u^n + \Delta t \sum_{l=1}^1 a_{2l} f(t + c_l \Delta t, u^{(l)}) \\ &= u^n + \Delta t a_{21} f(t + c_1 \Delta t, u^{(1)}) \end{aligned}$$

$$\begin{aligned} u^{(3)} &= u^n + \Delta t \sum_{l=1}^2 a_{3l} f(t + c_l \Delta t, u^{(l)}) \\ &= u^n + \Delta t a_{31} f(t + c_1 \Delta t, u^{(1)}) + \Delta t a_{32} f(t + c_2 \Delta t, u^{(2)}) \end{aligned}$$

Esempio:Calcolo di $u^{(l)}$ nel Runge-Kutta implicito:

$$\begin{aligned} u^{(1)} &= u^n + \Delta t \sum_{l=1}^1 a_{1l} f(t + c_1 \Delta t, u^{(1)}) \\ &= u^n + \Delta t a_{11} f(t_n + c_1 \Delta t, u^{(1)}) \end{aligned}$$

$$\begin{aligned} u^{(2)} &= u^n + \Delta t \sum_{l=1}^2 a_{2l} f(t + c_l \Delta t, u^{(l)}) \\ &= u^n + \Delta t a_{21} f(t + c_1 \Delta t, u^{(1)}) + \Delta t a_{22} f(t + c_2 \Delta t, u^{(2)}) \end{aligned}$$

$$\begin{aligned} u^{(3)} &= u^n + \Delta t \sum_{l=1}^3 a_{3l} f(t + c_l \Delta t, u^{(l)}) \\ &= u^n + \Delta t a_{31} f(t + c_1 \Delta t, u^{(1)}) + \Delta t a_{32} f(t + c_2 \Delta t, u^{(2)}) + \Delta t a_{33} f(t + c_3 \Delta t, u^{(3)}) \end{aligned}$$

Metodo di Heun ($v = 2$):

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & 1/2 & 1/2 \end{array}$$

$$\sum_{i=1}^v b_i = \frac{1}{2} + \frac{1}{2} = 1$$

$$c_i = \sum_{l=1}^v a_{il}$$

$$\begin{aligned} c_1 &= 0 \\ c_2 &= 1 \end{aligned}$$

$$a = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

$$b_1 = b_2 = \frac{1}{2}$$

Quelli che moltiplicano f

Con

$$u^{(1)} = u^n$$

$$\begin{aligned} u^{(2)} &= u^n + \Delta t a_{21} f(t_n + c_1 \Delta t, u^{(1)}) \\ &= u^n + \Delta t \cdot 1 \cdot f(t_n + 0 \cdot \Delta t, u^{(1)}) \\ &= u^n + \Delta t f(t_n, u^n) \end{aligned}$$

Troviamo la soluzione:

$$\begin{aligned} u^{n+1} &= u^n + \Delta t \sum_{i=1}^2 b_i f(t_n + c_i \Delta t, u^{(i)}) \\ &= u^n + \Delta t [b_1 f(t_n + c_1 \cdot \Delta t, u^{(1)}) + b_2 f(t_n + c_2 \cdot \Delta t, u^{(2)})] \\ &= u^n + \Delta t \left[\frac{1}{2} f(t_n + 0 \cdot \Delta t, u^{(1)}) + \frac{1}{2} f(t_n + \Delta t, u^{(2)}) \right] \end{aligned}$$

Esempio:

Applico Heun a $y' = \lambda y$.

$$u^{(1)} = u^n$$

$$u^{(2)} = u^n + \Delta t \lambda u^n = (1 + \Delta t \lambda) u^n$$

$$u^{n+1} = u^n + \frac{\Delta t}{2} [\lambda u^1 + \lambda u^2] = u^n + \frac{\Delta t}{2} \lambda [u^n + u^n + \Delta t + \lambda u^n]$$

$$u^{n+1} = u^n \left[1 + \Delta t \lambda + \frac{(\Delta t \lambda)^2}{2} \right]$$

$$e^{\Delta t \lambda} = [1 + \Delta t \lambda + (\Delta t \lambda)^2 + O(\Delta t^3)]$$

La soluzione esatta è $y^{n+1} = y^n \cdot e^{\Delta t \lambda}$

$$e(t) = l_{\Delta t}^{n+1} + \tilde{u}(t_{n+1}) - u^{n+1}$$

$$l_{\Delta t}^{n+1} \equiv (\text{errore commesso nel nuovo passo})$$

$$\tilde{u}(t_{n+1}) - u^{n+1} = \text{vecchio errore trasportato dall'equazione differenziale}$$

Per il metodo di Eulero Esplicito ($f(t, y) = f(y)$):

$$u^{n+1} = u^n + \Delta t f(u^n)$$

$$\tilde{u}^{n+1} = y^n + \Delta t (y^n)$$

Sostituisco e trovo:

$$e(t) = e(t_{n+1}) = l_{\Delta t}^{n+1} + y^n - u^n + \Delta t [f(y^n) - f(u^n)]$$

$$|e(t_{n+1})| \leq |l_{\Delta t}^{n+1}| + |y^n - u^n| + \Delta t |f(y^n) - f(u^n)|$$

Ricordo che f è Lipshitziana, cioè: $|f(y^n) - f(u^n)| \leq L|y^n - u^n|$

$$|e(t_{n+1})| \leq |l_{\Delta t}^{n+1}| + (1 + \Delta t \cdot L) |y^n - u^n|$$

$$|e(t_{n+1})| = |l_{\Delta t}^{n+1}| + (1 + \Delta t \cdot L) |e(t_n)|$$

$$|l_{\Delta t}^{n+1}| = \text{errore nuovo}$$

$$(1 + \Delta t \cdot L) |e(t_n)| = \text{errore vecchio trasportato dall'equazione differenziale stessa}$$

Ho costruito una relazione di ricorrenza:

$$|e(t_n)| \leq |l_{\Delta t}^n| + (1 + \Delta t \cdot L) |e(t_{n-1})|$$

Sostituendo in $|e(t_{n+1})|$:

$$|e(t_{n+1})| \leq |l_{\Delta t}^{n+1}| + (1 + \Delta t \cdot L) |l_{\Delta t}^n| + (1 + \Delta t \cdot L)^2 |e(t_{n-1})|$$

Per induzione:

$$|e(t_{n+1})| \leq (1 + \Delta t \cdot L)^{n+1} |e_0| + \sum_{k=0}^n (1 + \Delta t \cdot L)^{k+1} |l_{\Delta t}^{n-(k-1)}|$$

Tutti gli errori che io commeto vengono amplificati per il termine $(1 + \Delta t \cdot L)$:

$$(1 + \Delta t \cdot L)^k \leq (e^{\Delta t \cdot L})^k = e^{k \cdot \Delta t \cdot L}$$

$$\leq e^{(n+1) \cdot \Delta t \cdot L} = e^{t \cdot L}$$

So che:

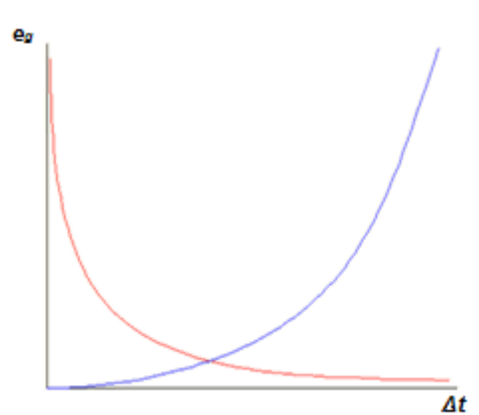
$$|e(t)| \leq C(\Delta t)^p e^{t \cdot L} \|y^{p+1}\|_{\infty} + (\text{accumulo errore di macchina})$$

accumulo errore di macchina = (n° passi) · (errore di macchina su 1 passo)

errore di macchina su 1 passo = $K_f \cdot eps$

accumulo errore di macchina $\sim \frac{t}{\Delta t} K_f \cdot eps$

$$|e(t)|_{globale} = C(\Delta t)^p e^{t \cdot L} \|y^{p+1}\|_{\infty} + \frac{t}{\Delta t} K_f \cdot eps$$



1.2.1. Se $\varepsilon \text{ tol } |u^n| < |u^{n+1} - \tilde{u}^{n+1}| < \text{tol} |u^n|$

\Rightarrow vado al passo successivo con Δt_{n+1} e $u^{n+1} = \tilde{u}^{n+1}$

1.2.2. Se $|u^{n+1} - \tilde{u}^{n+1}| > \text{tol } |u^n|$

$$\Delta t_{n+1} = \frac{1}{2} \Delta t_{n+1} \text{ e ritorno (1*)}$$

1.2.3. Se $|u^{n+1} - \tilde{u}^{n+1}| < \varepsilon \text{ tol } |u^n|$

La stima è abbastanza accurata, ma il passo può essere aumentato.

Tengo $u^{n+1} = \tilde{u}^{n+1}$, esco con $\Delta t_{n+1} = 2\Delta t_{n+1}$

Devo fissare un Δt minimo

Se $|u^n|$ è molto piccolo non posso usare il test relativo

$$\text{tol} = \max(\text{RelTol} |u^n|, \text{AbsTol})$$

Default

$$\text{RelTol} = 10^{-3}, \quad \text{AbsTol} = 10^{-6}$$

Warning:

Calcolare sempre la soluzione con almeno 2 tolleranze diverse e confrontare i risultati.

Esempio:

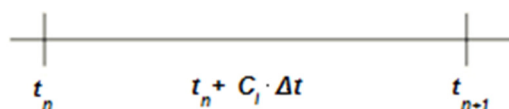
$$EE(RK 1) \quad (p = 1)$$

$$Heun (RK 2) \quad (p = 2)$$

$$\begin{aligned} u^{n+1} &= u^n + \Delta t f^n & f^n &= f(t_n, u^n) \\ u^{(1)} &= u^n \\ u^{(2)} &= u^n + \Delta t f^n \\ \tilde{u}^{n+1} &= u^n + \frac{\Delta t}{2} (f(u^{(1)}) + f(u^{(2)})) \end{aligned}$$

Embedded RK:

\tilde{u}^{n+1} utilizza informazioni che sono già state calcolate per la soluzione u^{n+1} .
ODE 45 (ode quattro cinque).



Esempio:

$$r = 1$$

$$A. u^{n+1} = u^n + \Delta t f^n \quad (EE)$$

$$(\beta_1 = 0, \beta_0 = 1) \quad (p = 1)$$

$$B. u^{n+1} = u^n + \frac{\Delta t}{2} (f^n + f^{n+1})$$

$$\left(\beta_1 = \frac{1}{2}, \beta_0 = \frac{1}{2}\right) \quad (p = 2)$$

Per costruire l'indicatore di errore uso:

1. u^{n+1} calcolato con un metodo di Adams esplicito a r passi ($p = r$).
2. \tilde{u}^{n+r} calcolato con un metodo di Adams implicito a r passi ($p = r + 1$).

$$\tilde{u}^{n+r} = \sum_{j=0}^{r-1} \alpha_j u^{n+j} + \Delta t \sum_{j=0}^{r-1} \beta_j f^{n+j} + \Delta t \beta_r (f(u^{n+j}))$$

Metodo esplicito (predictor) calcola u^{n+r} .

Metodo implicito (corrector) usa il valore u^{n+r} per calcolare \tilde{u}^{n+r} .

Lo stimatore di errore è $|u^{n+r} - \tilde{u}^{n+r}|$

ODE113 usa LMM di Adams di ordine 1 a ordine 13

Si usa per calcolare funzioni molto lunghe.

RK 1 è RK 2 permettono ottenere una soluzione più in fretta.

Stabilità

$$e(t) \leq (\Delta t)^p e^{Lt} \|y^{(p+1)}\|_{\infty} \text{ per } \Delta t \rightarrow 0$$

Poiché l'errore locale di troncamento $T_n = C(\Delta t)^p + O(\Delta t)^{p+1}$.

Consideriamo l'equazione $y' = \lambda y$ ($\lambda \in \mathbb{C}$).

Se applico il metodo di Eulero Esplicito $u^{n+1} = (1 + \Delta t \lambda) u^n$

La soluzione esatta è $y(t) = y_0 e^{\lambda t}$

Chiamo regione di stabilità la regione di \mathbb{C} tale che $|u^{n+1}| \leq |u^n|$

Per i metodi ad 1 passo:

$$u^{n+1} = R(\lambda\Delta t)u^n \qquad z = \lambda\Delta t$$

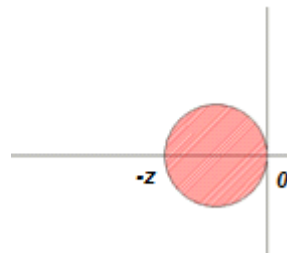
Regione di stabilità:

$$\{z \in \mathbb{C} \text{ tale che } |\mathbb{R}(z)| \leq 1\}$$

Un metodo è A-stabile se la sua regione di stabilità contiene $\mathbb{R}_e(z) < 0$.

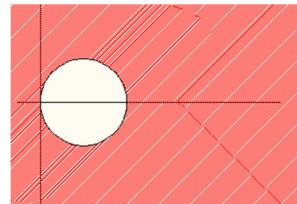
Per il metodo di Eulero Esplicito la regione di stabilità

$$(\mathbb{R}_S) = \{z: |1 + z| \leq 1\}$$



Per il metodo di Eulero Implicito la regione di stabilità

$$(\mathbb{R}_S) = \{z: |1 - z| \geq 1\}$$



Per singoli passi di integrazione il metodo implicito è "costoso" di usare.

Esempio:

$$y' = \sin(y)$$

Eulero Implicito: $u^{n+1} = u^n + \Delta t \sin(u^{n+1})$

Per trovare u^{n+1} devo risolvere un'eq non lineare \Rightarrow devo usare un metodo iterativo:

Metodi di RK espliciti:

RK 1 $u^{n+1} = (1 + z)u^n$

RK 2 $u^{n+1} = \left(1 + z + \frac{z^2}{2}\right)u^n$

RK 3 $u^{n+1} = \left(1 + z + \frac{z^2}{2} + \frac{z^3}{6}\right)u^n$

ecc...

$$u^{n+1} = (e^z + O(z)^p)u^n$$

Sia DIRK che i BDF sono costruiti in modo da essere A-stabili.

Nei DIRK:

$$u^{n+1} = R(z)u^n$$

Con $R(z)$ funzione razionale

Esempio

$$\begin{array}{c|cc} \gamma & \gamma & 0 \\ \hline 1-\gamma & 1-2\gamma & \gamma \\ & 1/2 & 1/2 \end{array} \qquad \gamma = 1 - \frac{1}{\sqrt{2}}$$

$y' = \lambda y$ In genere si lavora su sistemi differenziali

Nel caso lineare:

$$y' = Ay \qquad y(t) = y_0 e^{At}$$

Se la matrice è diagonalizzabile: $R^{-1}AR = \Lambda$

$\Lambda \equiv \text{diag}(\lambda_1, \dots, \lambda_n)$ autoval di A

$R \equiv$ matrice degli autovettori

$$e^{At} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k = R \left(\sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} \right) R^{-1} = R e^{\Lambda t} R^{-1}$$

$$A = R\Lambda R^{-1}$$

$$\Rightarrow A^2 = R\Lambda R^{-1} R\Lambda R^{-1} = R\Lambda^2 R^{-1}$$

↳ $\mathbb{I} \rightarrow$ identità

Se $\lambda(A)$ hanno parte Reale < 0

La condizioni di stabilità viene data di $\max |\lambda_i|$

$$y' = \lambda y \qquad \lambda = -10, -100, -10^7$$

$$[t, y] = \text{ode45}(f, t_0, y_0)$$

$$\text{plot}(t, y, 'o') \qquad dt_i = t_i - t_{i-1}$$

Sistema iperbolico:

Un sistema del tipo $\partial_t u + \nabla f(u) = 0$ si dice iperbolico se:

$$J_f \text{ ha } \begin{cases} n \text{ autovalori reali} \\ n \text{ autovettori linearmente indipendenti} \end{cases} \quad \begin{array}{l} n = \text{numero eq. sistema} \\ J_f = \text{Jacobiano funzione} \end{array}$$

In generale J_f è a coefficienti non costanti (anche gli autovalori e autovettori sono funzioni della u).

Anche la natura iperbolica del problema dipende da u .

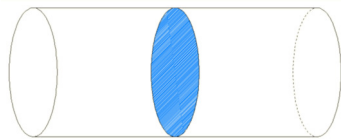
In gas dinamica ho 3 autovalori:

$$\begin{array}{l} \lambda_1 = u - c \\ \lambda_2 = u \\ \lambda_3 = u + c \end{array} \quad c = \sqrt{\gamma \frac{p}{\rho}} \quad \text{velocità del suono}$$

⇒ 3 autovettori linearmente indipendenti

⇒ gas dinamica 1D è iperbolica

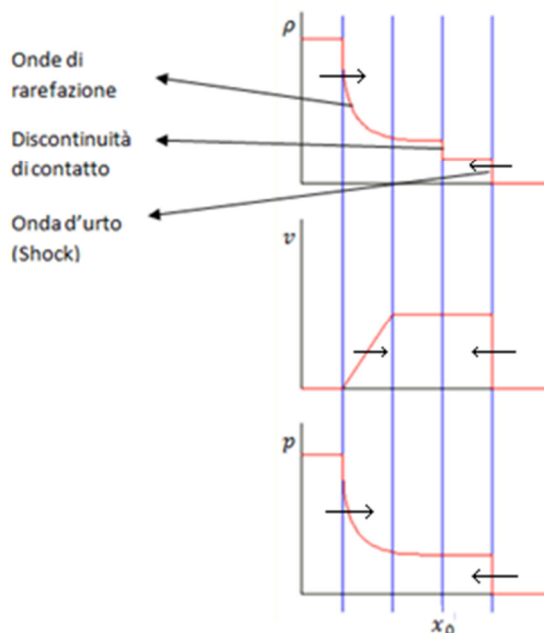
Gas dinamica – tubo d’urto

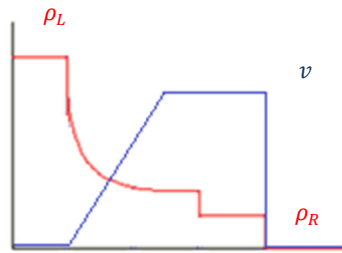


$$u(x, t = 0) = \begin{cases} u_L & x < 0 \\ u_R & x > 0 \end{cases}$$

$$u_L = \begin{pmatrix} \rho_L \\ 0 \\ E_L \end{pmatrix}$$

$$u_R = \begin{pmatrix} \rho_R \\ 0 \\ E_R \end{pmatrix}$$





• **Equazione di trasporto lineare**

$$u_t + f_x(u) = 0 \quad \Leftrightarrow \quad u_t + au_x = 0$$

$$f(u) = a \cdot u$$

$$a = \text{costante}$$



$\Gamma: (x(t), t)$
 $(\gamma(t), t) \rightarrow \text{curva}$

Cerco $\gamma(t)$ t.c $u_t + au_x = 0 \Leftrightarrow \frac{d}{dt} \Big|_j u = 0$

Voglio trovare una combinazione di x e t tale che la u rimane costante.

$$\frac{d}{dt} \Big|_j u = \frac{d}{dt} (u(\gamma(t), t^\gamma)) = \frac{du}{dx} \cdot \frac{d\gamma}{dt} + \frac{du}{dt} = 0$$

Questa relazione è vera per $\frac{d\gamma}{dt} = a$

$$\frac{du}{dt} + a \frac{du}{dx} = 0 \quad \text{è equivalente a:}$$

$$\frac{du}{dt} = 0 \quad \text{su} \quad \frac{d\gamma}{dt} = a \quad \text{curva caratteristica}$$

$$\frac{du}{dt} = 0 \quad \text{su} \quad \frac{dx}{dt} = a$$

Questo quà vuol dire:

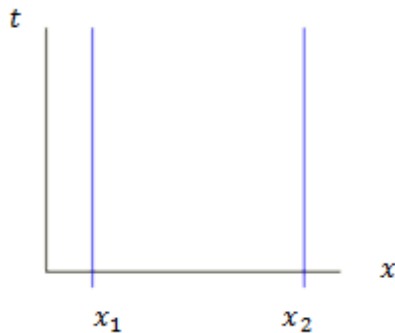
$$u(x, t) = \text{cost.} \quad \text{su} \quad \frac{dx}{dt} = a$$

Dati iniziali:

$$u_t + au_x = 0$$

$$u(x, t = 0) = u_o(x)$$

Problemi al bordo



Risolve

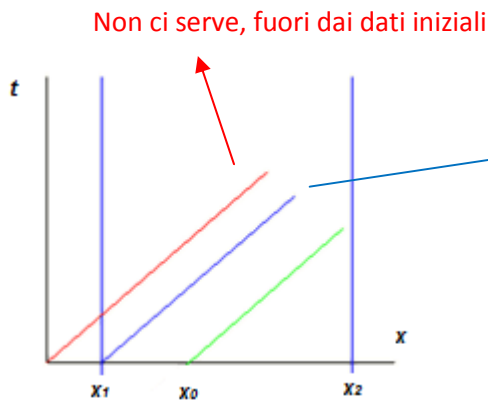
$$u_t + au_x = 0 \quad \text{per} \quad \begin{cases} x_1 \leq x \leq x_2 \\ t > 0 \end{cases}$$

Dati:

$$u(x, t = 0) = u_0(x) \quad \text{per} \quad x_1 \leq x \leq x_2$$

$$u(x = x_1, t > 0) = g_1(t)$$

$$u(x = x_2, t > 0) = g_2(t)$$



$$u(x, t) = u_0(x_0)$$

$$x_1 < x_0$$

$$x_0 - x_1 = at$$

$$u(x, t) = u_0(x_0) \quad \text{per} \quad x > at + x_1$$

$$x_0 = x - at$$

Se $x < at + x_1$ siamo sopra della caratteristica limite e non si può arrivare alle caratteristiche iniziali.

$$u(x, t) = g_1(t_0)$$

Per trovare la caratteristica per (x, t) che arriva in (x_1, t_0) :

$$x - x_1 = a(t - t_0)$$

$$t - t_0 = \frac{x - x_1}{a} \quad \rightarrow \quad t_0 = t - \frac{x - x_1}{a}$$

$$u(x, t) = \begin{cases} u_0(x - at) & \text{per} \quad x > at + x_1 \\ g_1\left(t - \frac{x - x_1}{a}\right) & \text{per} \quad x < at + x_1 \end{cases}$$

Equazioni non lineari

$$u_t + f_x(u) = 0 \quad \Leftrightarrow \quad u_t + (f(u))_x = 0 \quad \Leftrightarrow \quad u_t + a(u)u_x = 0$$

$f \equiv$ funzione non lineare di u

$$f(u) = \frac{1}{2}u^2 \quad \text{eq. di Burgers}$$

Se f è differenziabile:

$$\partial_x f(u) = f'(u) \partial_x u$$

$f'(u) = a(u)$ (velocità caratteristica)

$$u_t + a(u)u_x = 0$$

Metodo delle caratteristiche:

$$\frac{du}{dt} = 0 \quad \text{su} \quad \frac{dx}{dt} = a(u) \quad u = \text{cost.} \quad \text{su} \quad \frac{dx}{dt} = a(u)$$

$a(u)$ è costante lungo la caratteristica.

La caratteristica è una retta:

$$u(x, t) = u_0(x_0) \quad \text{su} \quad x - x_0 = a(u)t.$$

Cioè su

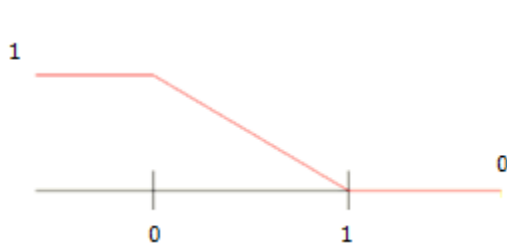
$$\frac{x - x_0}{t - t_0} = a(u_0(x_0)) \quad \rightarrow \quad x - x_0 = a(u_0(x_0)) t$$

1. Troviamo x_0 per i diversi intervalli di $u_0(x_0)$.
2. Calcolo la funzione $u(x, t)$ e gli intervalli che la definiscono.

Esercizio:

$$f(u) = \frac{1}{2}u^2 \quad \text{eq. di Burgers}$$

$$u_t + (f(u))_x = 0 \quad \Rightarrow \quad u_t + \left(\frac{1}{2}u^2\right)_x = 0 \quad \Leftrightarrow \quad u_t + u \cdot u_x = 0$$



$$u_0(x_0) = \begin{cases} 1 & x_0 < 0 \\ 1 - x_0 & 0 < x_0 < 1 \\ 0 & x_0 > 1 \end{cases} \quad 0 \leq t \leq 1$$

Questo segnale si comprime

$0 \geq t \rightarrow$ si può fare
 $t \geq 1 \rightarrow$ vedremo come si fa

per $x_0 < 0$	per $x_0 > 1$	per $0 < x_0 < 1$:
$u(x, t) = 1 \rightarrow a(u) = 1$ $\frac{x - x_0}{t} = 1$ $x_0 = x - t$	$u(x, t) = 0 \rightarrow a(u) = 0$ $\frac{x - x_0}{t} = 0$ $x_0 = x$	$u(x, t) = x \rightarrow a(u) = 1 - x_0$ $\frac{x - x_0}{t} = 1 - x_0$ $x_0 = \frac{t - x}{t - 1}$

Per definire gli intervalli di $u(x, t)$, dobbiamo sostituire i valori x_0 ottenuti negli intervalli della funzione $u_0(x, t)$. Per risolvere l'intervallo in mezzo basta assicurare che si soddisfacciano le condizioni ottenute agli entrambi lati.

$$u(x, t) = \begin{cases} 1 & x < t \\ \frac{x - 1}{t - 1} & t < x < 1 \\ 0 & x > 1 \end{cases}$$

per $x_0 < 0$:

$$u(x_0, t_0) = 1$$

$$a(u_0(x_0)) = u_0(x_0) = 1$$

$$x - x_0 = a(u_0(x_0)) \cdot t$$

$$x - x_0 = 1t$$

$$\boxed{x_0 = x - t}$$

Adesso sostituiamo x_0 per trovare l'intervallo in funzione di t :

$$x - t < 0$$

Quindi, finalmente:

$$u(x, t) = 1 \quad \text{per} \quad x < t$$

per $x_0 > 1$:

$$u_0(x_0) = 0$$

$$a(u_0(x_0)) = u_0(x_0) = 1$$

$$x - x_0 = a(u_0(x_0)) \cdot t$$

$$x - x_0 = 0$$

$$\boxed{x_0 = x}$$

Adesso sostituiamo x_0 per trovare l'intervallo in funzione di t :

$$x > 1$$

Quindi, finalmente:

$$u(x, t) = 0 \quad \text{per} \quad x > 1$$

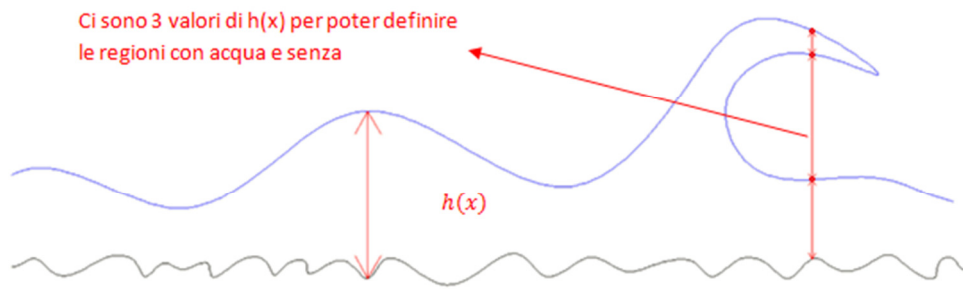
Per $0 < x_0 < 1$

$$u_0(x_0) = 1 - x_0$$

$$a(u_0(x_0)) = u_0(x_0) = 1 - x_0$$

$$x - x_0 = (1 - x_0) \cdot t$$

Onde d'acqua



Qui posso accettare soluzioni a molti valori, ma in generale questo non è possibile.

Siamo partiti da una legge integrale.

$$\frac{d}{dt} \int_a^b u(x, t) dx = f(u(a, t)) - f(u(b, t))$$

Vorrei vedere se questa equazione accetta una soluzione a gradino.

Con:

$$\lim_{x \rightarrow s^-(t)} u(x, t) = u_L$$

$$\lim_{x \rightarrow s^+(t)} u(x, t) = u_R$$

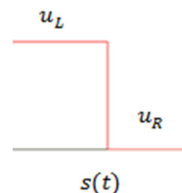
$$\lim_{x \rightarrow s^-(t)} f(u(x, t)) = f(u_L)$$

$$\lim_{x \rightarrow s^+(t)} f(u(x, t)) = f(u_R)$$

In generale:

$$u_L = u_L(t)$$

$$u_R = u_R(t)$$



$$\frac{d}{dt} \int_a^b u(x, t) = \frac{d}{dt} \int_a^{s(t)} u(x, t) + \frac{d}{dt} \int_{s(t)}^b u(x, t)$$

$$\begin{aligned} \text{Vedere } F(t) \text{ sotto} \quad \rightarrow &= \int_a^{s(t)} \partial_t u + u(s^-(t), t) + \int_{s(t)}^b \partial_t u - u(s^+(t), t) \cdot s'(t) \\ &= f(u(a, t)) - f(u(b, t)) \end{aligned}$$

$$F(t) = \int_0^t f(x) dx \quad \rightarrow \quad F'(t) = f(t)$$

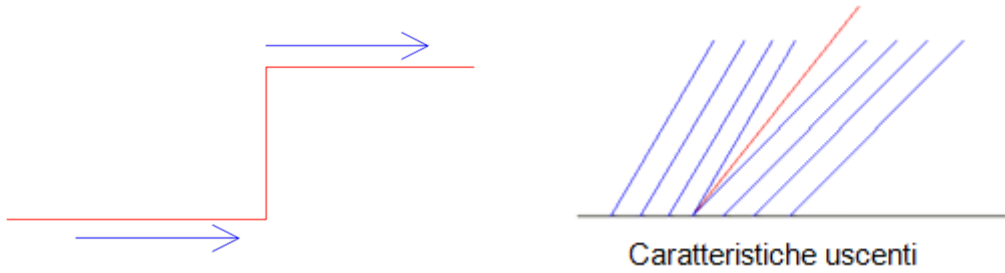
$$F(t) = \int_0^{s(t)} f(x) dx \quad \rightarrow \quad F'(t) = f(s(t)) \cdot s'(t)$$

$$F(t) = \int_a^b f(x, t) dx \quad \rightarrow \quad F'(t) = \int_a^b \partial_t f dx$$

Faccio:

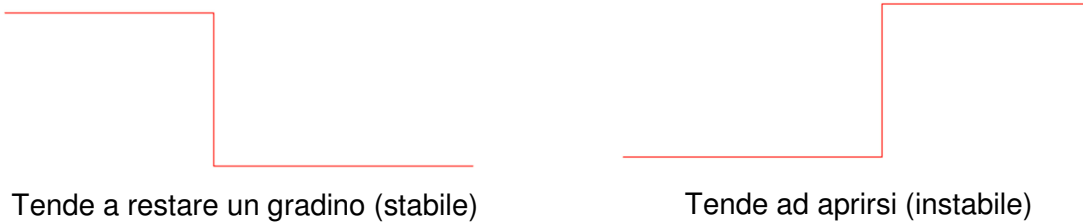
$$\begin{aligned} a &\rightarrow s^- \\ b &\rightarrow s^+ \end{aligned}$$

$$u_L > u_R \quad f'(u_L) > f'(u_R)$$



Soluzione di Burgers

$$\text{Se } s' = \frac{f(u_L) - f(u_R)}{u_L - u_R} :$$

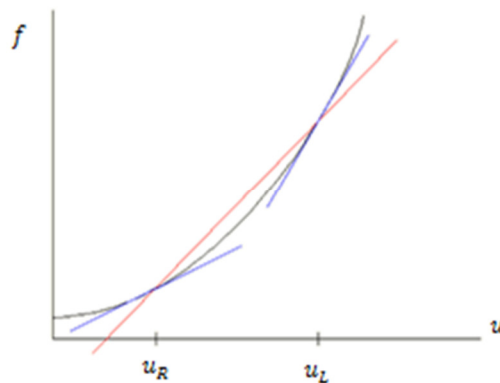


Accetto solo onde d'urto con caratteristiche entranti

$$f'(u_L) > s'(t) > f'(u_R)$$

Condizione dell'entropia

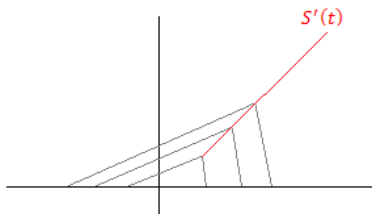
$$\text{Se } f''(u) > 0 \quad u_L > u_R$$



$$u_t + f_x(u) = 0$$

$$u(x, t = 0) = u_0$$

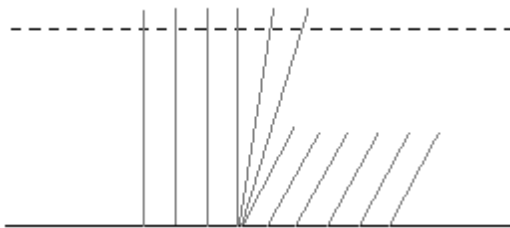
Può formare discontinuità in un tempo finito, $t = t_s$, "in quel caso" la velocità della discontinuità è data dalla relazione di R.H.



Soddisfa la condizione dell'entropia

Quando si forma l'onda d'urto entropica ottengo soluzioni che sono irreversibili nel tempo.

Rarefazione (rampa):

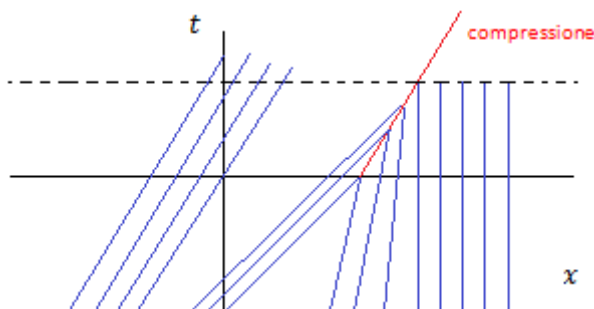


Ogni valore $u_0(\cdot)$ è contenuto nella soluzione al tempo $u(x, t)$.

(problema reversibile nel tempo)

Qui la soluzione è reversibile nel tempo.

$u_0(x)$

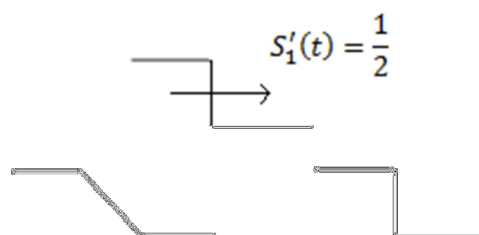


Se sei investigando qui potresti andare indietro nel tempo.

Per $t = \frac{1}{2}$ la soluzione è ancora regolare (e reversibile).

Per $t > 1$ si perde molta informazione relativa ai dati iniziali.

Per $t > 1$ ho una soluz. a gradino.



$-v_{\max}$ per $(\rho = \rho_{\max})$ v_{\max} per $(\rho = 0)$

Onde d'urto

$$S'(t) = \frac{f(\rho_L) - f(\rho_R)}{\rho_L - \rho_R} = \frac{v_{\max}}{\rho_{\max}} [\rho_{\max} - (\rho_L + \rho_R)]$$

$$S'(t) > 0 \quad \text{per} \quad \rho_{\max} > \rho_L + \rho_R$$

$$S'(t) < 0 \quad \text{per} \quad \rho_{\max} < \rho_L + \rho_R$$

In ogni caso, l'automobilista entra nella coda, perché $v_L - S' \leq 0$.

Quando si formano le code?

Condizione dell'entropia:

$$f'(\rho_L) \geq S'(t) \geq f'(\rho_R)$$

$$v_M \left(1 - \frac{2\rho_L}{\rho_{\max}}\right) \geq \frac{v_{\max}}{\rho_{\max}} (\rho_{\max} - (\rho_L + \rho_R)) \geq v_{\max} \left(1 - \frac{2\rho_R}{\rho_{\max}}\right)$$

$$-\frac{2\rho_L}{\rho_M} \geq -\frac{(\rho_L + \rho_R)}{\rho_M} \geq -\frac{2\rho_R}{\rho_M}$$

$$2\rho_L \leq \rho_L + \rho_R \leq 2\rho_R$$

Si forma onda d'urto quando c'è più intensità davanti che indietro ($\rho_L < \rho_R$).

Sistemi lineari iperbolici a coefficienti costanti

Caso scalare:

$$u_t + au_x = 0 \quad \text{ove} \quad a = \text{cost.}$$

Problema ai valori iniziali

$$\begin{cases} u_t + au_x = 0 \\ u(x, t = 0) = u_0(x) \quad (\text{nota}) \end{cases}$$

$$u(x, t) = u_0(x - at) \rightarrow \text{segnale iniziale spostato di una lunghezza } at$$

Sistema $u(x, t)$ un vettore di m incognite:

$$u(x, t) = \begin{bmatrix} u_1(x, t) \\ u_2(x, t) \\ \vdots \\ u_m(x, t) \end{bmatrix} \quad m \text{ eq.}$$

$$u_t + Au_x = 0 \quad (*) \quad A \equiv \text{Matrice } m \times m \text{ costante}$$

Definizione:

Il sistema (*) è iperbolico se A ha m autovalori reali $\lambda_1, \lambda_2, \dots, \lambda_m$ e m autovettori linearmente indipendenti r_1, \dots, r_m .

N.B.: Gli autovettori r_j sono vettori colonna con m componenti costanti. Se gli autovalori sono distinti allora è garantito che $\exists m$ autovettori linearmente indipendente. Altrimenti devo verificare che posso calcolare m autovettori linearmente indipendenti.

$$\text{Chiamo } R = [r_1, r_2, \dots, r_m] \quad \text{matrice di autovettori}$$

Poiché gli r_j sono linearmente indipendenti $\Rightarrow R$ è invertibile $\Rightarrow A$ è diagonalizzabile, cioè:

$$R^{-1}AR = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m) \quad \text{con} \quad Ar_j = \lambda_j r_j$$

Funziona come fattore di amplificazione, no ne cambia la direzione.

N.B.: poiché R è costante, anche R^{-1} è costante $\Rightarrow \partial_t R = \partial_x R = 0$

$$R^{-1}[\partial_t u + A \partial_x u] = 0$$

$$\partial_t(R^{-1}u) + R^{-1}AR R^{-1} \partial_x u = 0$$

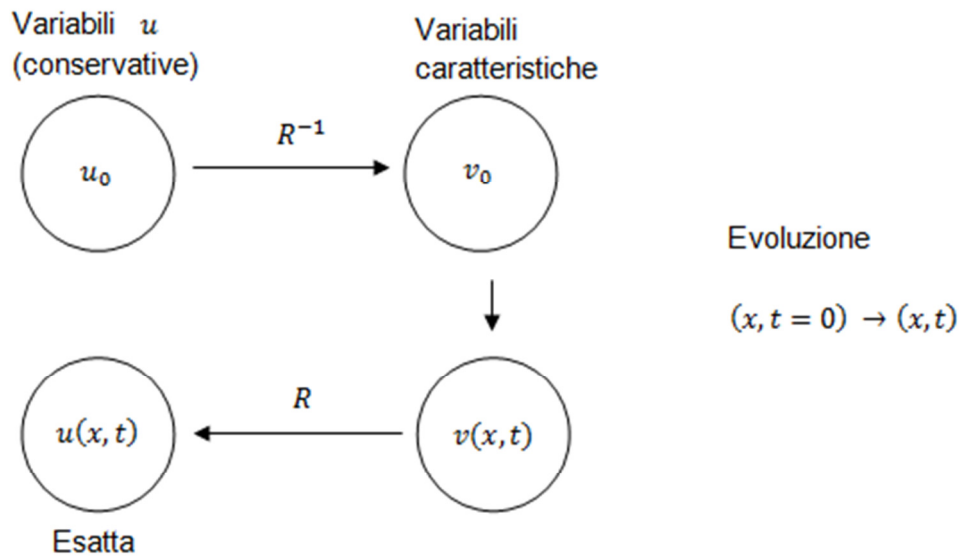
Per sistemi non lineari

$$\partial_t u + \partial_x f(u) = 0$$

In forma caratteristica:

$$\partial_t u + A(u)\partial_x u = 0 \quad \text{con} \quad A(u) = J_f$$

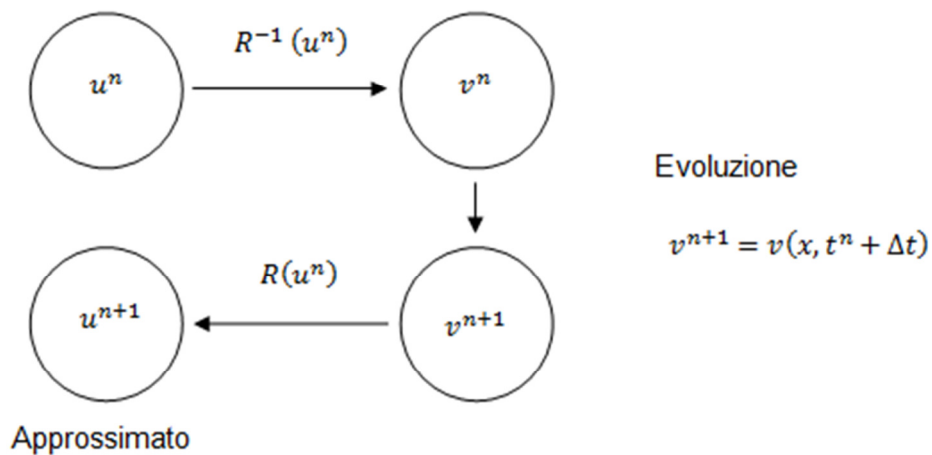
Se il sistema è lineare:



Caso non lineare:

procedo per passi Δt

chiamo $u^n(x) = u(x, t^n = n\Delta t)$



Esercizio:

$$u_t + Au_x = 0$$

Dati iniziali:

$$A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} \quad u_L = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad u_R = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$

$$u_0(x) = \begin{cases} u_L & x < 0 \\ u_R & x > 0 \end{cases} \quad u(x, t) = \begin{bmatrix} u_1(x, t) \\ u_2(x, t) \end{bmatrix}$$

Siccome la matrice A non è diagonale dobbiamo trovare prima la matrice degli autovalori:

$$\det A = \det \begin{pmatrix} 1-\lambda & 4 \\ 1 & 1-\lambda \end{pmatrix} = (1-\lambda)(1-\lambda) - 4 = \lambda^2 - 2\lambda - 3$$

$$\lambda^2 - 2\lambda - 3 \rightarrow \lambda = \frac{2 \pm \sqrt{4 + 12}}{2} = 1 \pm 2 = \begin{cases} \lambda_1 = 3 \\ \lambda_2 = -1 \end{cases}$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

Un'altra forma di trovarla è con gli autovettori d'ognuno degli autovalori. Per il primo:

$$A - \lambda_1 I = \begin{pmatrix} 1 - \lambda_1 \cdot 1 & 0 \\ 0 & 1 - \lambda_1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 - 3 & 0 \\ 0 & 1 - 3 \end{pmatrix} = \begin{pmatrix} -2 & 4 \\ 0 & -2 \end{pmatrix}$$

$$X_1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$(A - \lambda_1 I) X_1 = 0 \rightarrow \begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} -2x_1 + 4x_2 = 0 \\ x_1 - 2x_2 = 0 \end{cases}$$

$$X_1 = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} \rightarrow \boxed{X_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}}$$

Ora per il secondo autovalore, troviamo il suo autovettore:

$$A - \lambda_2 I = \begin{pmatrix} 1 - \lambda_2 \cdot 1 & 0 \\ 0 & 1 - \lambda_2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 - (-1) & 0 \\ 0 & 1 - (-1) \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 0 & 2 \end{pmatrix}$$

$$X_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Metodi numerici per equazioni lineari iperboliche

$$u_t + au_x = 0$$

$u(x, t) \rightarrow$ soluzione esatta

$$u(x_j, t^n) = u_j^n$$

$U(x, t) \rightarrow$ soluzione numerica

$$U(x_j, t^n) = U_j^n$$



$$h = x_j - x_{j-1} \quad \text{griglia uniforme}$$

$$\Delta t = t^n - t^{n-1} \quad \text{passo temporale}$$

$$\lambda = \frac{\Delta t}{h} \quad \text{parametro di griglia}$$

Devo trovare una forma di discretizzare la derivata temporale:

$$\partial_t u|_{(x_j, t^n)} \simeq \frac{1}{\Delta t} (u_j^{n+1} - u_j^n) + O(\Delta t)$$

Invece, per la discretizzazione spaziale abbiamo 3 derivate numeriche:

$$\partial_x u|_{(x_j, t^n)} \simeq \frac{u_{j+1}^n - u_j^n}{h} + O(h) \quad (AD: \text{asimmetrico destra})$$

$$\partial_x u|_{(x_j, t^n)} \simeq \frac{u_j^n - u_{j-1}^n}{h} + O(h) \quad (AS: \text{asimmetrico sinistra})$$

$$\partial_x u|_{(x_j, t^n)} \simeq \frac{u_{j+1}^n - u_{j-1}^n}{2h} + O(h^2) \quad (C: \text{centrale})$$

Per avere un metodo stabile indipendentemente del segno di a :

Metodi del 1° ordine:

- **Metodo Upwind**

$$U_j^{n+1} = U_j^n - \lambda a \begin{cases} U_j^n - U_{j-1}^n & a > 0 \\ U_{j+1}^n - U_j^n & a < 0 \end{cases}$$

Upwind \equiv sopravvento



- **Metodo di Lax - Friedrichs**

$$U_j^{n+1} = \frac{1}{2}(U_{j+1}^n + U_{j-1}^n) - \frac{\lambda a}{2}(U_{j+1}^n - U_{j-1}^n)$$

Metodi del 2° ordine:

- **Metodo di Lax - Wendroff (centrada)**

$$U_j^{n+1} = U_j^n - \frac{\lambda a}{2}(U_{j+1}^n - U_{j-1}^n) + \frac{\lambda^2 a^2}{2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n)$$

- **Metodo di Beam - Warming**

$$U_j^{n+1} = U_j^n - \lambda a \begin{cases} \frac{1}{2}U_{j-2}^n - 2U_{j-1}^n + \frac{3}{2}U_j^n & a > 0 \\ \frac{1}{2}U_{j+2}^n - 2U_{j+1}^n + \frac{3}{2}U_j^n & a < 0 \end{cases} + \frac{\lambda^2 a^2}{2} \begin{cases} U_{j-2}^n - 2U_{j-1}^n + U_j^n & a > 0 \\ U_{j+2}^n - 2U_{j+1}^n + U_j^n & a < 0 \end{cases}$$

