Metodi Numerici

Zeri di funzioni lineaZeri di equazioni non lineari

Indice di convergenza

Sia $f(\alpha) = 0$, α radice.

$$\tilde{f}(x) = f(x) + \epsilon g(x) = 0$$

(1) $|\tilde{f}(x) - f(x)| = |\epsilon g(x)|$ perturbazione sulla funzione

(2) $|\tilde{\alpha} - \alpha| = |\delta|$ perturbazione sui risultati

$$\tilde{f}(\alpha + \delta) = \tilde{f}(\alpha) + \delta\tilde{f}'(\alpha) + o(...) \approx 0$$
$$f(\alpha) + \epsilon g(\alpha) + \delta(f'(\alpha) + \epsilon g'(\alpha)) \approx 0$$

$f(\alpha) = 0$ per ipotesi

$$\epsilon g(\alpha) + \delta f'(\alpha) + \delta\epsilon g'(\alpha) \approx 0$$
$$\epsilon g(\alpha) + \delta f'(\alpha) \approx 0$$
$$(\mathbf{2})\delta = -\frac{\epsilon g(\alpha)(\mathbf{1})}{f'(\alpha)}$$
$$|\tilde{\alpha} - \alpha| = -\frac{1}{f'(\alpha)}|\tilde{f}(\alpha) - f(\alpha)|$$
$$\boldsymbol{k = -\frac{1}{f'(\alpha)}}$$

- se $k \leq 10^2$ problema **ben condizionato**
- se $10^3 < k < 10^4$ problema **mediamente mal condizionato**
- se $k > 10^4$ problema **mal condizionato**

Ordine di convergenza

Data una successione di iterati $\{x_k\}$, generata da un metodo numerico convergente ad un limite α e sia $e_k = x_k - \alpha$. Se $\exists$ due numeri reali $p \geq 1$ e $c > 0$ t.c

$$\lim_{k \to +\inf} \frac{|e_{k+1}|}{|e_k|^p} = c$$
$$|e_{k+1}| \approx c|e_k|^p$$
$$\boldsymbol{p \approx \frac{\log(\frac{|x_{k+2} - x_{k+3}|}{|x_{k+1} - x_{k+2}|})}{\log(\frac{|x_{k+1} - x_{k+2}|}{|x_k - x_{k+1}|})}}$$

- se $p = 1$ la convergenza è **lineare**
- se $1 < p < 2$ la convergenza è **superlineare**

- se $p = 2$ la convergenza è **quadratica**

Sia $e_k = x_k + \alpha$, supponiamo che $|e_k| \leq \frac{1}{2}10^{-n}$, cioè la radice x_k ha n decimali corretti

$$|e_{k+1}| \approx c|e_k|^p \leq c(\frac{1}{2}10^{-n})^p = \frac{c}{2^p}10^{-pn}$$

la radice x_k ha p-n decimali corretti.

1) Metodo bisezione

Teorema degli zeri

> Sia f(x) continua nell'intervallo [a, b] e sia tale che f(a)·f(b)<0, allora f ammette almeno uno zero in (a,b), cioè esiste almeno un punto α in (a,b) tale che f(α)=0.

Se $f(a) * f(b) < 0$ si pone $a_0 := a$ e $b_0 := b$ (teorema degli zeri)

Finché non risulta verificato il criterio di arresto:

Poni: $\mathbf{x_{k+1} := a_k + \frac{b_k - a_k}{2}}$

a) se $f(x_{k+1}) * f(a_k) < 0 \rightarrow [a_k, x_{k+1}]$

b) se $f(x_{k+1}) * f(b_k) < 0 \rightarrow [x_{k+1}, b_k]$

c) se $f(x_{k+1}) = 0 \rightarrow end$

$k = k + 1$

- **criterio d'arresto** $\rightarrow k = \lceil \log_2(\frac{b-a}{\epsilon}) - 1 \rceil = maxit$

Vantaggi:

- **convergenza globale:** La convergenza è assicurata per qualsiasi scelta del punto iniziale appartenente all'intervallo che racchiude la radice, cioè $x0$ in [a,b], che verifica il teorema.
- **convergenza lineare** $\rightarrow$ **p=1, c=1/2**

$$|e_k| = |x_k - a| \leq \frac{1}{2}|b_k - a_k| = \frac{1}{2^{k+1}}|b - a|$$
$$|e_{k+1}| = |x_{k+1} - a| \leq \frac{1}{2}|b_{k+1} - a_{k+1}| = \frac{1}{2^{k+2}}|b - a|$$
$$\frac{|e_{k+1}|}{|e_k|} \approx \frac{1}{2}$$

- **semplicità**

Svantaggi:

- **Lentezza:** convergenza lineare
- **Richiede che l'intervallo iniziale rispetti il teorema degli zeri**

2) Metodo dei falsi

A differenza della bisezione, che non trae alcun vantaggio da caratteristiche della funzione, un modo per migliorare tale metodo è quello di considerare anche i valori che la funzione assume negli estremi dell'intervallo

---

Se $f(a) * f(b) < 0$ si pone $a_0 := a$ e $b_0 := b$ (teorema degli zeri)

Finché non risulta verificato il criterio di arresto:

Poni: $\boldsymbol{x_{k+1} := a_k - f(a_k) \frac{b_k - a_k}{f(b_k) - f(a_k)}}$

a) se $f(x_{k+1}) * f(a_k) < 0 \rightarrow [a_k, x_{k+1}]$

b) se $f(x_{k+1}) * f(b_k) < 0 \rightarrow [x_{k+1}, b_k]$

c) se $f(x_{k+1}) = 0 \rightarrow end$

$k = k + 1$

return radice, numero iterazioni, vettore soluzione

Vantaggi:

- **convergenza globale:** La convergenza è assicurata per qualsiasi scelta del punto iniziale appartenente all'intervallo che racchiude la radice, cioè x0 in [a,b], che verifica il teorema.
- **convergenza superlineare, 1 < p < 2:** più veloce rispetto alla bisezione, sfrutta le caratteristiche della funzione, considerando i valori che la funzione assume negli estremi dell'intervallo

Svantaggi:

- **Rallentamento della convergenza:** l'intervallo $[a_i, b_i]$ non tende a zero

3) Metodi di linearizzazione

Data $f(x), x_0$ $f(x_0)$: si approssima la funzione con una retta per $(x_0, f(x_0))$

$$\begin{cases} y = f(x_0) + m(x - x_0) \\ y = f(x_0) + m(x - x_0) \\ y = 0 \end{cases}$$

$$x_1 = x_0 - \frac{f(x_0)}{m}$$

3.a) Metodo delle corde

m rimane costante, coincide col coefficiente angolare della retta che congiunge i punti (a,f(a)) (b,f(b)).

Finché non risulta verificato il criterio di arresto:

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

$d = \frac{f(x_k)}{m}$

Poni $x_{k+1} := x_k - d$

$k = k + 1$

return radice, numero iterazioni, vettore soluzione

Vantaggi:

- **convergenza migliore della bisezione**

Svantaggi:

- **Convergenza superlineare**
- **Dipendenza dalla scelta dell'intervallo iniziale [a,b]**

3.b) Metodo delle secanti

Assegnati i due valori iniziali x_0, x_1, al passo k l'approssimazione della funzione f nell'intervallo $[x_{k-1}, x_k]$ è la retta che passa per i punti$(x_{k-1}, f(x_{k-1}))$ $(x_k, f(x_K))$ con coefficiente angolare $m = \frac{f(x_k)-f(x_{k-1})}{x_k-_{k+1}}$

Finché non risulta verificato il criterio di arresto:

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

Poni $x_{k+1} := x_k - d$

$k = k + 1$

return radice, numero iterazioni, vettore soluzione

Vantaggi:

- **convergenza superlineare $\mathbf{p \approx 1,618}$:** più veloce della bisezione e del metodo dei falsi

Svantaggi:

- **convergenza locale:** è garantita se le approssimazioni di x_0 e x_1, si scelgono abbastanza vicine alla soluzione $\rightarrow$ convergenza locale
- **Dipendenza dalla scelta dell'intervallo iniziale [x_0, x_1]**

3.c) Metodo Newton

Ad ogni passo k, si considera la retta passante per il punto $(x_k, f(x_k))$ e tangente alla curva f(x), si sceglie x_k come punto di incontro tra la retta e l'asse x, dove $m_k = f'(x_k)$

Verifichiamo che $f'(x) \neq 0$

Finché non risulta verificato il criterio d'arresto

$d = \frac{f(x_k)}{f'(x_k)}$ (rappresenta quanto bisogna correggere l'attuale approssimazione x_k per avvicinarsi allo zero della funzione, e per ottenere la nuova approssimazione x_{k+1})

$x_{k+1} = x_k - d$

$k = k + 1$

return radice, numero iterazioni, vettore soluzione

Vantaggi:

- **convergenza quadratica:**

  Sia α radice di $f(x)$, cioè $f(\alpha) = 0$

$$f(x) = f(x_k) + (x - x_k)f'(x_l) + \frac{1}{2}(x - x_k)^2 f''(\zeta)$$
$$f(\alpha) = 0 = f(x_k) + (a - x_k)f'(x_l) + \frac{1}{2}(a - x_k)^2 f''(\zeta)$$
$$\frac{f(x_k)}{f'(x_k)} + (\alpha - x_k) + \frac{\frac{1}{2}(a - x_k)^2 f''(\zeta)}{f'(x_k)} = 0$$
$$(a - x_{k+1}) + \frac{\frac{1}{2}(a - x_k)^2 f''(\zeta)}{f'(x_k)} = 0$$

  dato che $e_{k+1} = x_{k+1} - \alpha$

$$-e_{k+1} + \frac{1e_k^2 f''(\zeta)}{2f'(x_k)} = 0$$
$$e_k = \frac{1e_k^2 f''(\zeta)}{2f'(x_k)}$$
$$\lim_{k \to +\infty} \frac{e_{k+1}}{e_k^2} = \frac{\mathbf{1}f''(\alpha)}{\mathbf{2}f'(\alpha)}$$

  Dato che $\lim_{k \to +\infty} \frac{|e_{k+1}|}{|e_k|^p} = c$ il metodo di Newton ha **ordine di convergenza p=2** e **fattore di convergenza** $\frac{\mathbf{1}f''(\alpha)}{\mathbf{2}f'(\alpha)}$

Svantaggi:

- **convergenza locale:** se il punto iniziale x0 è troppo lontano dalla radice il metodo non converge

Teorema di convergenza locale

> Se f:[a,b] $\to$ R soddisfa le ipotesi:
> i)
> $f(a)f(b) < 0$
> ii)
> f, f', f'' sono continue in [a,b]
> iii)
> $f'(x) \neq 0 \; \forall \; x \; \in [a, b]$
> Allora esiste un intorno
> $I \subset [a.b]$ dell'unica radice $\alpha \in (a, b)$ t.c., se $x \in I$, allora la successione di Newton $\{x_I\}$ converge in α

Teorema di convergenza globale (Newton)

> Sia $f(x) \in C^2[a,b]$, [a,b] intervallo chiuso e limitato. Se sono verificate le condizioni
> i)
> $f(a)f(b) < 0$
> ii)
> $f'(x) \neq 0 \; \forall \; x \in [a,b]$
> iii)
> $f''(x) < 0$ oppure $f''(x) \; \forall \; x \in [a,b]$
> iiii)
> $|\frac{f(a)}{f'(a)}| < b-a \;\; |\frac{f(b)}{f'(b)}| < b-a$
> Allora il metodo di Newton converge all'unica soluzione
> α in [a,b] $\forall \; x_0 \in [a,b]$

Sistema di equazioni non lineari

1) Metodo di Newton-Raphson

Dato $X_0 \in \mathbb{R}^n$ ed F, per ogni iterazione k:

Valutare $det(J(x_{k-1})) \neq 0$

Risolvere il sistema lineare $J(x_{k-1})s_{k-1} = -F(x_{k-1})$

Poni $X_k = X_{k-1} + s_{k-1}$

- **convergenza locale**
- **ordine di convergenza quadratico**

1.a) Approssimazione con rapporti incrementali

Variante del metodo di Newton-Raphson. Consiste nel sostituire a $J(X_{k-1})$ una sua approssimazione ottenuta mediante rapporti incrementali n-dimensioni del tipo

$$\frac{df_j}{dX_i}|_{X=X_{k-1}} \approx (J^{(k-1)})_{ij} = \frac{f_i(X_{k-1} + e_i h_{ij}) - f_l(X_{k-1})}{h_{ij}}$$

- e_i il vettore i-esimo della base canonica
- h_{ij} incremento al passo k

1.b) Metodo delle corde

Si utilizza lo stesso Jacobiano o una sua approssimazione $J(X_0)$ oppure $A(X_0)$ per tutte le iterazioni k. Si potrebbe quindi fattorizzare $J(X_0) = LU$ e utilizzare i medesimi L e U per ogni iterazione

1.c) Metodo di Shamanskii

i valuta lo Jacobiano ogni m iterazioni e quindi lo si utilizza per le m iterazioni successive
$J^{k+1} = J^i \; i = 1, ..., m$

Metodo di Newton per il calcolo del minimo di una funzione a più variabili

Dato $X_0 \in \mathbb{R}^n$ ed F, per ogni iterazione k:

Valutare $det(H(x_{k-1})) \neq 0$

Risolvere il sistema lineare $H(x_{k-1})s_{k-1} = -\nabla f(x_{k-1})$

Poni $X_k = X_{k-1} + s_{k-1}$

Norma vettoriale

Ogni applicazione $|| \cdot || : R^n a \; R_+ \cup \{0\}$ si chiama norma su R se soddisfa le proprietà:
1)
$||x|| > 0 \; \forall x \in R^n$ e $||x|| = 0 \leftrightarrow x = 0$
2)
$||\lambda x|| = |\lambda| \cdot ||x|| \; \forall \in R, \; \forall \in R^n$
3)
$||x + y|| \leq ||x|| + ||y|| \; \forall x, y \in R$

- **Norma infinito** $||x||_\infty = max|x_i|$
- **Norma 1** $||x||_1 = \sum |x_i|$
- **Norma 2** $||x||_2 = [\sum |x_i^2|]^{\frac{1}{2}}$

La norma di un vettore è definita come la radice quadrata della somma dei quadrati delle componenti del vettore, cioè $||x||_2 = \sqrt{x^T x}$

NB $||x||_\infty \leq ||x||_2 \leq ||x||_1$

Norma matriciale

Sia M(mxn) lo spazio vettoriale delle matrici mxn su R, si dice che l'applicazione $||A|| :$
$da \; M(mxn) \; a \; R_+ \cup \{0\}$ è la norma della matrice A se soddisfa le proprietà:
1)
$||A|| > 0 \; \forall A \neq 0$ e $||A|| = 0 \leftrightarrow A = 0$
2)
$||\alpha A|| = |\alpha| \cdot ||A|| \; \forall \in M(mxn), \; \forall \; \alpha \in R^n$
3)
$||A + B|| \leq ||A|| + ||B|| \; \forall A, B \in M(mxn)$
4)
$||A \cdot B|| \leq ||A|| \cdot ||B|| \; \forall \; A \in M(mxn), \; B \in M(mxn)$

- **Norma infinit**o $||A||_\infty = max \sum_{j=1}^{n} |a_{ij}|$ (la somma massima tra i vettori righe)
- **Norma 1** $||A||_\infty = max \sum_{i=1}^{m} |a_{ij}|$ (la somma massima tra i vettori colonna)
- **Norma 2** $||A||_2 = \sqrt{p(A^T A)}$ (la radice massima del polinomio caratteristico)

1) calcolo $M = A^T @ A$
2) Calcolo gli autovalori di
$M \rightarrow p(M)$
3) prendo l'autovalore massimo e ne calcolo la radice
$\rightarrow \sqrt{p(M)}$

Se A è ortogonale, cioè $A^T = A^{-1} \rightarrow ||Ax||_2 = \sqrt{(Ax)^T(Ax)} = \sqrt{x^T(A^TA)x} = \sqrt{x^Tx} = ||x||_2 \ \forall x \in R$

Se A è simmetrica allora $||A||_1 = ||A||_\infty$

Sistemi lineari

Teorema di Rouchè-Capelli

> Il sistema lineare Ax=b ammette soluzioni $\leftrightarrow$ la matrice dei coefficienti A e la matrice A|b hanno lo stesso rango: $Rank(A) = Rank(A|b)$
> Altrimenti il sistema non ha soluzioni.

Se $Rank(A) \neq Rank(A|b)$:

- m<n, cioè abbiamo più incognite che relazioni lineari tra di esse. Il sistema di dice **indeterminato.** Se indichiamo con k il rango di A e K<n allora ammette ∞^{n-k}
- m>n, cioè abbiamo meno incognite che relazioni lineari tra di esse. Il sistema si dice **sovradeterminato**. Non ammette una soluzione esatta ma una soluzione approssimata.
- m=n, cioè il numero di incognite è uguale al numero di relazioni lineari tra di esse. Il sistema si dice **normale** e sotto opportune ipotesi può ammettere una ed una sola soluzione.

Sistemi normali

Matrice non singolare

> A (nxn) è detta non singolare se soddisfa una delle seguenti condizioni equivalenti:
> 1)
> $det(A) \neq 0$
> 2)
> $\exists \ A^{-1}$
> 3)
> $rank(A) = n$ ((2) implica (3) poiché una matrice è invertibile solo se ha rango max)

Teorema soluzione di un sistema lineare

> Condizione necessaria e sufficiente affinché il sistema lineare Ax=b , $A \in M(n \times n), x, b \in Rn$ ammetta una ed una sola soluzione, comunque si scelga b, è che la matrice A sia a rango massimo (cioè che la matrice A sia invertibile); si ha perciò:
>
> $x = A^{-1}b$

Condizionamento di una sistema lineare

Indice di condizionamento $\boldsymbol{K(A) = ||A^{-1}|||A||}$

Esso dipende intrinsecamente dal problema, dalla matrice stessa, esso ci dice come le inevitabili perturbazioni che si hanno sulla matrice o sul termine noto si percuotono sulla soluzione del sistema.

- se $k \leq 10^2$ problema **ben condizionato**
- se $10^3 < k < 10^4$ problema **mediamente mal condizionato**
- se $k > 10^4$ problema **mal condizionato**

L'indice di condizionamento della matrice A rappresenta un fattore di amplificazione sulla soluzione di piccoli errori sui dati. Se A è una matrice mal condizionata, piccole perturbazioni sui dati del problema vengono amplificate nella soluzione.

Se A è ortogonale, cioè $A^TA = I \leftrightarrow A^{-1} = A^T$, allora $K(A) = 1$. Infatti

$$||A||_2 = \sqrt{p(A^TA)} = \sqrt{p(I)} = 1$$
$$||A^{-1}||_2 = ||A^T||_2 = \sqrt{p(A^TA)} = \sqrt{p(I)} = 1$$

La risoluzione di Ax=b con A ortogonale è sempre un problema ben condizionato

Esempi di matrici mal condizionate

Matrice di Vandermonde

$$A = \begin{bmatrix} 1 & x_0 & (x_0)^2 & (x_0)^3 \\ 1 & x_0 & (x_0)^2 & (x_0)^3 \\ 1 & x_0 & (x_0)^2 & (x_0)^3 \\ 1 & x_0 & (x_0)^2 & (x_0)^3 \end{bmatrix}$$

Matrice di Hilbert

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{bmatrix}$$

Metodi per la risoluzione di un sistema lineare

- **metodi diretti:** Questi metodi, in assenza di errori di arrotondamento, conducono alla soluzione esatta in un numero finito di passi. Essi sono adatti per la soluzione di sistemi con
  matrice dei coefficienti densa e di moderate dimensioni
- **metodi iterativi**: Questi metodi generano una successione di soluzioni, che, sotto opportune ipotesi, convergono alla soluzione del sistema. La matrice dei coefficienti non viene modificata durante il calcolo e quindi è più agevole sfuttarne la sparsità. Sono adatti, quindi, per la soluzione di sistemi con **matrice dei coefficienti di grandi dimensioni e sparsa**. In assenza di errori di arrotondamento conducono alla soluzione esatta in un numero infinito di passi.

1) Metodi diretti

I metodi iterativi trasformano attraverso un numero finito di passi un sistema lineare generico in un sistema equivalente che ne semplifichi la risoluzione $\rightarrow$ **fattorizzazione** della matrice dei coefficienti

$$A = B \cdot C$$
$$\begin{cases} By = b \\ Cx = y \end{cases}$$

Stabilità di un algoritmo di fattorializzazione

Poiché le operazioni di fattorizzazione vengono eseguite in aritmetica finita, alla fine dell'algoritmo si ottengono, anziché i fattori esatti B e C, matrici affetti da perturbazione $B = B + \delta B$ e $C = C + \delta C$

$$A + \delta A = (B + \delta B) \cdot (C + \delta C)$$
$$A + \delta A = BC + B\delta C + C\delta B + \delta C\delta B$$
$$\delta A = B\delta C + C\delta B + \delta C\delta B$$

Notiamo quindi che la perturbazione su A non dipende solo dalle perturbazioni su B e C, ma è tanto più grande quanto più grandi sono gli elementi di B e C

Definizione

> Data una matrice A i cui elementi sono tutti minori o uguali ad 1, si dice che un algoritmo di fattorizzazione che produce una fattorizzazione $B \cdot C$ della matrice A è
> -
> **numericamente stabile in senso forte**, se esistono delle costanti positive a e b, indipendenti dall'ordine e dagli elementi di A tali che $|b_{ij}| \leq a$ $|c_{ij}| \leq b$
> -
> **numericamente stabile in senso debole**, se le costanti a e b dipendono dall'ordine di A

1.a) Fattorizzazione Gauss (LU)

Teorema 1

> Data $A \in \ M(nxn)$, sia A_k la sottomatrice principale di testa di A ottenuta considerando le prime k righe e le prime k colonne. Se A_k è non singolare $\forall\ k \in [1, n]$ allora $\exists!$ la fattorializzazione LU di A

NB Con l'ipotesi aggiuntiva che la matrice A abbia determinante diverso da zero allora la fattorizzazione LU può essere utilizzata per risolvere un sistema lineare.

Ip
1)
$A \in \ M(nxn)$
2)
$det(A_{ii}) \neq 0$ (A non singolare)

Allora $\exists$
1)
$B = L$ matrice triangolare inferiore con 1 sulla diagonale principale
2)
$C = U$ matrice triangolare superiore

t.c $\mathbf{A = LU} \rightarrow \begin{cases} Ly = b \\ Ux = y \end{cases}$

Teorema 2

> Data una qualunque matrice A non singolare, esiste una matrice di permutazione P non singolare t.c. PA = LU.

Se A non è singolare/mal condizionata si utilizza la l'algoritmo di **Gauss con pivoting,** ovvero si utilizza una matrice PA non è singolare. La matrice PA si ottiene scambiando le righe e le colonne di A in modo tale che gli elementi sulla diagonale principale non siano nulli, cioè $a_{ij} \neq 0$.

for K = 1,...,n-1

Calcola nella colonna k-esima, a partire dall'elemento (k,k) l'indice di riga s a cui appartiene il massimo in valore assoluto.

Se s≠ k

Scambia la riga s con la riga k, memorizza lo scambio nella matrice P (viene fatto scambiando nelle matrice P la riga s con la riga k)

$l_{ik} = \frac{a_{ik}}{a_{kk}} \quad i = k+1, ..., n$

$a_{ij} = a_{ij} - l_{ik}a_{kj} \quad i, j = k-1, ..., n$

- **costo computazionale** $\rightarrow \frac{1}{3}n^3$
- **algoritmo stabile in senso debole**

- $|l_{ij}| \leq 1$ (non dipende dall'ordine della matrice)
- $|u_{ij}| \leq 2^{n-1} max|a_{ij}|$ (dipende dall'ordine di a)

1.b) Fattorizzazione Cholesky

Teorema

> Sia A una matrice di ordine n simmetrica e definita positiva, allora esiste una matrice triangolare inferiore L con elementi diagonali positivi, ($l_{ii} > 0 \quad i = 1, ...n$) tale che $\mathbf{A} = \mathbf{L} \cdot \mathbf{L^T} \rightarrow \begin{cases} Ly = b \\ L^T x = y \end{cases}$

- **costo computazionale** $\rightarrow \frac{1}{6} n^3$
- **algoritmo stabile in senso forte**
  - $max|l_{ij}| \leq \sqrt{max|a_{ij}|}$ (non dipende dall'ordine di a)

1.c) Fattorizzazione Householder (QR)

Teorema

> Sia $A \in M(m \times n)$ con $m \geq n$ e rank(A) = n (ossia le colonne di A sono linearmente indipendenti). Allora esistono una matrice $Q \in M(m \times m)$ ortogonale e una matrice $R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \in M(mxn)$ dove $R1 \in M(n \times n)$ è una matrice triangolare superiore non singolare, tale che $\mathbf{A} = \mathbf{Q} \cdot \mathbf{R} \rightarrow \begin{cases} Qy = b \\ Rx = y \end{cases}$

Essendo Q ortogonale ($Q^{-1} = Q^T$) la soluzione del sistema $Qy = b$ si riduce a $y = Q^T B$, (poiché l'inversa di una matrice ortogonale coincide con la sua inversa) $\begin{cases} y = Q^T B \\ Rx = y \end{cases}$

- **costo computazionale**
  - se $m - 1 \geq n \rightarrow mn^2 - \frac{n^3}{3}$
  - se $m = n \rightarrow \frac{2}{3} n^3$
- **algoritmo stabile in senso debole** (ma migliore di Gauss perché $\sqrt{n} < 2^{n-1}$, gli elementi di a sono maggiorati in Gauss 2^{n-1}, mentre in QR da $\sqrt{n}$)
  - $|q_{ij}| \leq 1$ (non dipende da a)
  - $|r_{ij}| \leq \sqrt{n}\, max|a_{ij}|$ (dipende da a)

2) Metodi iterativi

Tali metodi risultano particolarmente convenienti quando la matrice A del sistema è di grandi dimensioni e sparsa. Infatti, quando la matrice A è sparsa, cioè il numero degli elementi non nulli è di molto inferiore al numero degli elementi nulli, applicando i metodi diretti può accadere che vengano generati elementi non nulli in corrispondenza degli elementi nulli della matrice di partenza (fenomeno del fill-in).
Questo non avviene applicando i metodi iterativi in quanto essi si limitano ad utilizzare gli elementi non nulli della matrice senza toccare gli elementi nulli.

A differenza dei metodi diretti, i metodi iterativi trasformano la matrice A nella differenza di due matrici $\rightarrow$ **decomposizione** della matrice dei coefficienti

$$\mathbf{A} = \mathbf{M} - \mathbf{N}$$
$$(M-N)x = b \rightarrow Mx = Nx + b \rightarrow x = M^{-1}Nx + M^{-1}b$$
$$x_k = M^{-1}Nx_{k+1} + M^{-1}b$$
$$x_k = Tx_{k-1} + q$$

Dove $T = M^{-1}N$ è la matrice di iterazione, e $q = M^{-1}b$

Nei metodi successivi la matrice A è decomposta come somma di 3 matrici:

$$A = D + E + F$$

- D = matrice con elementi non nulli sulla diagonale

$$D = \begin{cases} d_{ii} = a_{ii} \;\; i = j \\ d_{ij} = 0 \quad i \neq j \end{cases}$$

- E = matrice triangolare inferiore con 0 sulla diagonale principale

$$E = \begin{cases} e_{ij} = 0 \;\; altrimenti \\ e_{ij} = a_{ij} \;\; i > j \end{cases}$$

- F = matrice triangolare superiore con 0 sulla diagonale principale

$$F = \begin{cases} f_{ij} = a_{ij} \;\; i < j \\ f_{ij} = 0 \;\; altrimenti \end{cases}$$

2.a) Metodo di Jacobi

Ip
1)
$a_{ii} \neq 0$
2)
A è non singolare
3) Ogni elemento dell'iterato (k) è indipendente dagli altri

Allora $\exists\, A = M - N$, dove $M = D$ e $N = -(E + F)$

$$x_k = -D^{-1}(E+F)x_{k-1} + D^{-1}b$$
$$x_k = D^{-1}(-(E+F)x_{k-1} + b)$$
$$x_k = M^{-1}(Nx_{k-1} + b)$$

2.b) Metodo di Gauss-Seidel

Ip
1)
$a_{ii} \neq 0$
2)
A è non singolare
3) Ogni elemento dell'iterato (k) è indipendente dagli altri

Allora $\exists\ A = M - N$, dove $M = D + E$ e $N = -F$

$$x_k = -(E+D)^{-1}Fx_{k-1} + (E+D)^{-1}b$$
$$x_k = (E+D)^{-1}(-Fx_{k-1} + b)$$
$$x_k = M^{-1}(Nx_{k-1} + b)$$

Convergenza di un sistema lineare

Definizione

> Il procedimento iterativo $x_k = Tx_{k-1} + q$ converge se $\forall\ x_0$, la successione x_k converge ad un vettore limite y:
>
> $lim_{k\to\infty}x_k = y$
> Cioè
> $\forall\ \epsilon > 0, \exists$ un indice v tc $\forall\ k > v$ si ha:
>
> $||x_k - y|| \leq \epsilon$

Teorema

> Se il sistema Ax=b ammette un'unica soluzione x e se il processo iterativo $x_k = Tx_{k-1} + q$ è convergente, allora il vettore y coincide con x:
>
> $lim_{k\to\infty}x_k = x$

Convergenza metodi iterativi

Affinché il procedimento sia convergente si deve avere che, comunque si sceglie x_0 ciascuna componente di $e_k = x_k - x$ tenda a 0 per $k \to \infty$, cioè

$$lim_{k\to\infty}T^k = 0$$
$$lim_{k\to\infty}(M^{-1}N)^k = 0$$

Teorema condizione sufficiente e necessaria per la convergenza

> Sia $A = M - N$ una matrice di ordine n con $det(A) \neq 0$, e $T = M^{-1}N$ la matrice di iterazione del procedimento iterativo $x_k = Tx_{k-1} + q$.
> Condizione **necessaria e sufficiente** per la convergenza del procedimento iterativo, comunque si scelga il vettore iniziale x_0, al vettore soluzione x del sistema $Ax = b$, è che $p(T) < 1$, ovvero che il raggio spettrale (autovalore massimo) sia minore di 1

NB L'algoritmo converge tanto più velocemente quanto è più piccolo l'autovalore di modulo massimo

Teorema 1 condizione sufficiente ma non necessaria per la convergenza

> Condizione sufficiente ma non necessaria per la convergenza, se per una qualche norma, risulta $||T|| < 1$, allora con il processo iterativo $x_k = Tx_{k-1} + q$ converge $\forall\ x_0$

Teorema 2 condizione sufficiente ma non necessaria per la convergenza

> Se la matrice A è diagonale strettamente dominante, cioè
> $|a_{ii}| > \sum_{\substack{k=1 \\ k \neq 1}}^{n} |a_{ik}|$
> Allora sia il metodo di Jacobi che quello di Gauss convergono

Teorema 3 condizione sufficiente ma non necessaria per la convergenza

> Se la matrice A è simmetrica e definita positiva, il metodo di Gauss-Seidel è convergente

2.c) Metodo Gauss-Seidel SOR

Tuttavia nonostante la matrice A soddisfi il teorema di convergenza può capitare che il metodo non converga. Infatti la convergenza del metodo è strettamente legato al condizionamento della matrice A. Infatti se A è mal condizionata, e nonostante soddisfi le ipotesi dei teoremi potrebbe comunque non convergere alla soluzione.

Per questo motivo sono stati introdotti dei **metodi di rilassamento**. L'idea è quella di far dipendere la matrice di iterazione da un parametro, detto parametro di rilassamento, e di far scegliere tale parametro in modo tale che la matrice abbia minimo raggio spettrale.

$$x_k = x_{k-1} + \omega r_k$$

Scegliendo opportunamte $\omega > 0$ si può accelerare la convergenza in modo significativo

- **under-relaxation methods,** con $0 < \omega < 1$

- **under-relaxation methods**, con $\omega > 1$, utilizzati per accelerare la convergenza in sistemi in cui il metodo di Gauss-Seidel converge ma lentamente
- **SOR (Successive Over-Relaxation**), $\omega = 1$

Metodo di Gauss-Seidel $\rightarrow x_k = -D^{-1}(Ex_k + Fx_{k-1} - b)$

Metodo Gauss-Seidel SOR $\rightarrow x_k = (1-\omega)x_{k-1} + \omega(-D^{-1}(Ex_k + Fx_{k-1} - b))$

2.d) Metodi di discesa

Teorema

> Sia A una matrice simmetrica e definita positiva, $b, x \in R^n$ allora la soluzione del sistema lineare $Ax = b$ coincide con il punto di minimo della funzione $F(x) = \frac{1}{2}x^TAx - b^Tx$.

Dimostrazione

Definiamo il vettore residuo $r = Ax - b$, se x^* è la soluzione del sistema, allora $r = Ax^* - b = 0$ $\mathbf{(1)}$

Ora consideriamo $F(x) = \frac{1}{2}x^TAx - b^Tx$ e cerchiamo il suo minimo, calcoliamo quindi il gradiente di F(x) e poniamolo = 0.
Quindi calcoliamo

$$\frac{dF}{dx_i} = \sum_{j=1}^{n} a_{ij}x_j - b_i = 0 \rightarrow \nabla F = Ax - b = \mathbf{(1)} = r = 0$$

Il vettore che annulla il gradiente coincide con la soluzione del sistema lineare, che rende nullo il residuo
Calcoliamo allora la matrice hessiana di F per verificare se il punto che annulla il gradiente è effettivamente un punto di minimo.
Sappiamo che

- determinante positivo e elemento $a_{11} > 0 \rightarrow$ punto di minimo
- determinante positivo e elemento $a_{11} < 0 \rightarrow$ punto di massimo
- determinante nullo $\rightarrow$ punto di sella

Notiamo che la matrice Hessiana risulta coincidere con la matrice A (che per ipotesi ha determinante maggiore di zero e elemento $a_{11} > 0$), quindi il punto che annulla il gradiente è il minimo della forma quadratica.

Per la risoluzione di sistemi lineari con matrice simmetrica definita positiva in generale possono essere usati i metodi per determinare il minimo di una funzione quadratica e cioè i metodi di discesa.

Metodo di discesa più ripida (Steepest Descent)

Caratterizzato dalla scelta, ad ogni passo k, della direzione p come l'antigradiente della F calcolato nell'iterato k-esimo:

$$p_k = -\nabla F(x_k) = -Ax_k + b = -r_k$$

Poiché il gradiente rappresenta la direzione di massima crescita, questo significa che ad ogni passo il vettore p, essendo l'antigradiente, coincide con la direzione di massima decrescita.

---

1) Parti con qualche $x_0,\ k = 0,\ r = Ax - b$

2) Calcola la direzione di discesa più ripida

$$p_k = -\nabla F(x_k) = -r$$

3) Scelta dello stepsize α_k

$$a_k = -\frac{< r_k, p_k >}{< Ap_k, p_k >} = -\frac{\boldsymbol{r_k^T \cdot p_k}}{\boldsymbol{p_k^T \cdot Ap_k}} = \\ t.c.\ F(x_k + a_k p_k) < F(x_k)$$

4) Aggiorna l'iterato

$$x_{k+1} = x_k + a_k p_k \\ r_{k+1} = r_k + a_k A p_k$$

- **criterio d'arresto** $\rightarrow ||r_{k+1}||_x \leq tol$
- **velocità di convergenza**

$$||x_k - x^*||_A \leq (\frac{K(A)-1}{K(A)+1})^k \cdot ||x_0 - x^*||_A \\ e_A^{(k)} = ||x_k - x^*||_A \rightarrow e_A^{(k)} \leq (\frac{K(A)-1}{K(A)+1})^k \cdot e_A^{(0)}$$

Dove $K(A) = ||A||||A^{-1}||$, tanto più K(A) è alto più il rapporto $(\frac{K(A)-1}{K(A)+1}) \approx 1$ e quindi tanto è più lenta la convergenza

Nonostante la convergenza dell'algoritmo, il metodo Steepest Descent è relativamente lento a causa del suo avanzamento a zig zig

Metodo di discesa gradiente coniugato

Caratterizzato dalla scelta, ad ogni passo k, della direzione p, non solo come l'antigradiente della F calcolato nell'iterato k-esimo, ma anche considerando le direzioni di discesa dell'iterazione precedente

---

1) Parti con qualche $x_0,\ k = 0,\ r = Ax - b$

2) Calcola la direzione di discesa più ripida

$$p_k = -\nabla F(x_k) = -r$$

3) Scelta dello stepsize α_k

$$a_k = -\frac{< r_k, p_k >}{< Ap_k, p_k >} = -\frac{\boldsymbol{r_k^T \cdot p_k}}{\boldsymbol{p_k^T \cdot Ap_k}} = \\ t.c.\ \ F(x_k + a_k p_k) < F(x_k)$$

4) Aggiorna l'iterato

$$x_{k+1} = x_k + a_k p_k \\ r_{k+1} = r_k + a_k A p_k \\ \gamma_{k+1} = \frac{< r_{k+1}, r_k >}{< r_k, r_k >} = \frac{r_{k+1}^T \cdot r_{k+1}}{r_k^T \cdot r_k} \\ p_k = -r_k + \gamma_k p_{k-1}$$

- **velocità di convergenza**

$$||x_k - x^*||_A \leq (\frac{\sqrt{K(A)} - 1}{\sqrt{K(A)} + 1})^k \cdot ||x_0 - x^*||_A \\ e_A^{(k)} = ||x_k - x^*||_A \rightarrow e_A^{(k)} \leq (\frac{\sqrt{K(A)} - 1}{\sqrt{K(A)} + 1})^k \cdot e_A^{(0)}$$

Dove $K(A) = ||A|||A^{-1}||$, tanto più K(A) è alto più il rapporto $(\frac{\sqrt{K(A)}-1}{\sqrt{K(A)}+1}) \approx 1$ e quindi tanto è più lenta la convergenza

La convergenza di questo metodo, pur rimanendo sempre legata all'indice di condizionamento di A è più veloce di quella dello Steepest Descent a parità di valori di K(A)

Teorema

> Nel metodo del gradiente coniugato le direzioni di discesa $p_k \ \ k = 0, 1, ..$ formano un sistema di direzioni coniugate $< Ap_k, p_j >= 0 \ \ k \neq j$, mentre i vettori residui $r_k \ \ k = 0, 1, ..$ formano un sistema ortogonale $< r_k, r_j >= 0 \ \ k \neq j$

Ciò significa che p_k è coniugata non solo a p_{k-1} ma a tutte le precedenti direzioni di discesa e che r_k è ortogonale a tutti i precedenti residui

Sistemi lineari sovradeterminati (m>n)

La risoluzione di un sistema lineare sovradeterminato risulta essere un **problema mal posto** in quanto potrebbe accadere che la soluzione non esista o se esiste non sia unica. infatti:

- $rank(A) = rank(A, b)$ sistema incompatibile
- $rank(A) = rank(A, b)$ sistema compatibile $\begin{cases} rank(A) < n \rightarrow \infty^{n - rank(A)}\ x^* \\ rank(A) = n \rightarrow\ \exists!\ x^* \end{cases}$

Metodo equazioni normali

Teorema equazioni normali

Per rendere il problema ben posto, cerchiamo una soluzione del sistema lineare sovradeterminato $Ax = b$ nel senso dei **minimi quadrati,** cioè, definito il vettore residuo come $r = Ax - b$ cerchiamo il vettore x^* che rende minima la norma 2 al quadrato del residuo

$$argmin_{x \in r^n} ||r(x)||_2^2 = argmin_{x \in r^n} ||Ax - b||_2^2$$
$$||Ax - b||_2^2 = (\sqrt{(Ax - b)^T (Ax - b)})^2 = (Ax - b)^T (Ax - b) = x^T A^T Ax - 2x^T A^T b + b^T b$$

Ricaviamo

$$\nabla F(x) = 2A^T Ax - 2A^T b = 0$$
$$A^T Ax = A^T b \ (*)$$

x è quindi la soluzione di un sistema nxn.

1. Se A ha rango massimo, la matrice $A^T A$ non è singolare, la soluzione nel senso dei minimi quadrati $\exists!$
2. Se A ha rango inferiore a n, la matrice $A^T A$ risulta singolare. In questo caso tra gli infiniti vettori soluzione, si assume come soluzione x quella che verifica $argmin_{x \in r^n} ||Ax - b||_2^2$

Ip
1) rk(A)=n
2) A ben condizionata

$$G = A^T A \rightarrow Gx = A^T b \ (*)$$

Dove G è una matrice simmetrica e definita positiva, abbiamo quindi trasferito il problema della risoluzione di un sistema sovradeterminato in quello della soluzione di un sistema quadrato con matrice G(nxn).

Essendo G simmetrica e definita positiva possiamo applicare la **fattorizzazione del metodo di Cholesky ($L^T L = G^T G$)**

Tuttavia $K_2(A^T A) = (K_2(A))^2$ questo implica che il sistema delle equazioni normali può risultare mal condizionato anche quando il problema nella sua forma originale non lo è.

Metodo QRLS

Proprietà

Una matrice ortogonale $Q \in R^{mxm}$ applicata ad un vettore $y \in R^m$, ne mantiene inalterata la norma 2 al quadrato, cioè

$$||y||_2^2 = ||Qy||_2^2$$

Osservazione

Sia $y \in R^m$, un vettore dove $m > n$, consideriamo la sua norma 2 al quadrato

$$||y||_2^2 = \sum_{i=1}^{m} y_i^2 = \sum_{i=1}^{n} y_i^2 + \sum_{i=n+1}^{m} y_i^2$$
$$y = \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix}$$
$$||y||_2^2 = ||\tilde{y}_1||_2^2 + ||\tilde{y}_2||_2^2$$

con y_1 avente n-componenti e y_2 avente m-n componenti

Ip
1) rk(A)=n
2) A è mediamente mal condizionata

Una matrice rettangolare $A \in R^{mxn}$, con rango n, può sempre essere fattorizzata tramite il metodo di Householder (QR)

$$A = QR = [Q] \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

Dove

- $Q(mxm)$ è ortogonale
- $R_1(nxn)$ matrice triangolare superiore con elementi $r_{ii} \neq 0$

$$||r(x)||_2^2 = ||Q^T(Ax - b)||_2^2 = ||Q^T Ax - Q^T b||_2^2 = ||Rx - h||_2^2$$

- $Q^T A = Q^{-1} A = R$
- $Q^T b = h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$

$$\begin{bmatrix} R_1 \\ 0 \end{bmatrix} x = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} R_1 x - h_1 \\ -h_2 \end{bmatrix}$$

Il minimo sarà ottenuto per x che risolve il sistema lineare $\boldsymbol{R_1 x = h_1}$

Il valore del residuo assunto per x sarà dato da $\boldsymbol{||h_2||_2^2}$

Questo metodo risulta decisamente migliore rispetto alle equazioni normali poiché:
1) lavora sempre e solo con la matrice A senza dover passare per
$A^T A$ che è molto peggio mal condizionata
2) la fattorizzazione QR nonostante sia stabile in senso debole è comunque più forte della fattorizzazione di Cholesky

Metodo SVD

Teorema

Sia $A \in R^{mxn}$ a rango $K \leq min(m, n)$. Allora esistono due matrici ortogonali $U \in R^{mxm}$ e $V \in R^{nxn}$ t.c $U^T AV = \sum$
Dove
Σ

- se A ha rango massimo

$$\Sigma = \begin{bmatrix} np.diag(A) \\ 0 \end{bmatrix}$$

- se non ha rango massimo

$$\Sigma = \begin{bmatrix} \sigma_{11}\ 0\ ....\ 0 \\ 0 \quad \sigma_{22}\ ..\ 0 \\ 0 \quad 0 \quad ...\ \ 0 \\ 0 \quad 0 \quad ...\ \ 0 \\ 0 \end{bmatrix}$$

---

lp
1) rank(A)<n
2) A è mal condizionata

Fattorizziamo la matrice A con la fattorizzazione SVD. Consideriamo inoltre che le trasformazioni ortogonali lasciano invariata la norma 2

$$||r(x)||_2^2 = ||U^T(Ax-b)||_2^2 = ||U^TAVV^Tx - U^Tb||_2^2 = ||\Sigma V^Tx - U^Tb||_2^2 = ||\Sigma c - d||_2^2$$
$$d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

Dove:

- d_1 ha k elementi
- d_2 ha m-k elementi

Il minimo sarà ottenuto per c che annulla $\Sigma c - d_1$ dove $c_i = \frac{d_i}{\sigma_i} \ \ i = 1, ..., k$

Ricaviamo poi x, dove v_i sono i vettori colonna della matrice V

$$\boldsymbol{x = Vc = \sum_{i=1}^{k} c_i v_i}$$

Il valore del residuo assunto per x sarà dato da $\boldsymbol{||d_2||_2^2}$

Retta minimi quadrati

Siano $(x_i, y_i) \ \ i = 1, ..., m-1$ si vuole determinare la retta che approssima i dati nel senso dei minimi quadrati e cioè determinare

$$P_1(x_i) = \alpha_0 + \alpha_1 x_i = y_i$$

Interpolazione polinomiale dati sperimentali

Siano (x_i, y_i) dove x_i sono detti nodo e y_i rappresentano le valutazioni di un fenomeno nelle posizioni x_i, il problema dell'interpolazione polinomiale consiste nel determinare $P_n(x) \in$

$P_n[x]$

$$P_n(x_i) = \alpha_0 + \alpha_1 x_i + \alpha_2 x_i^2 + ... + \alpha_n x_i^n = y_i$$

Risolvere il polinomio equivale a individuare i coefficienti $a_i \;\; i = 0, ..., n$, che soddisfa le condizioni P(x), sono la soluzione del sistema lineare

$$A\alpha = y$$

Tale sistema ammette una e una sola soluzione se e solo se la matrice dei coefficienti A è quadrata e ha rango massimo.
Sia A matrice di Vandermonde allora sappiamo che è quadrata perchè il numero di condizioni che imponiamo è uguale al numero delle incognite ed ha sempre rango massimo purché $x_i \neq x_k \; i \neq k$

Quindi il problema dell'interpolazione polinomiale ammette sempre soluzione e questa è unica

Osservazione

La matrice di Vandermonde è una matrice molto mal condizionata per cui la soluzione del sistema lineare sarà anch'esso un problema mal condizionato. Occorre quindi cambiare base per lo spazio $P_n[x]$, in modo tale che la matrice A coincida con la matrice identità.

Per farlo si impone come base **la base di Lagrange.**

$$L_j(x_i) = \begin{cases} 1 \;\; se \;\; i = j \\ 0 \;\; se \;\; i \neq j \end{cases}$$

Così facendo si ottiene

$$\begin{bmatrix} 1\,0\,0\,...\,0 \\ 0\,1\,0\,...\,0 \\ 0\,0\,1\,...\,0 \\ 0\,0\,0\,...\,1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ ... \\ \alpha_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{bmatrix}$$

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ ... \\ \alpha_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{bmatrix}$$

Polinomio interpolatore nella forma di Lagrange

$$L_j(x) = \prod_{k=0}^{n} \frac{(x - x_k)}{(x_j - x_k)} = 1$$

$$P_n(x) = \sum_{j=0}^{n} y_j L_j(x)$$

$$P_n(x_i) = y_j$$

L'unico polinomio di Lagrange diverso da 0 nel punto x_i è $L_i(x)$, che in x_i è 1.

- complessità polinomio di Lagrange $O(2n^2)$

Se dobbiamo effettuare la valutazione del polinomio per sistemi sovradeterminati e cioè in M>n punti la complessità computazionale sarebbe $O(2n^2 \cdot M)$. In genere M>>n quindi la complessità è molto elevata

Per migliorare l'efficienza si utilizza il polinomio di interpolazione di Newton che ha costo computazionale $O(\frac{n^2}{2} + n \cdot M)$

Teorema dell'errore

Siano (x_i, y_i) e $y_i = f(x_i) \;\; i = 0, ..., n$ siano valori assunti in quei punti da una funzione definita in [a,b] e continua insieme alle sue derivate fino a quella di ordine n+1

Sia $P_n(x)$ il polinomio di grado n che interpola tali coppie e sia $\tilde{x} \in [a, b]$, indichiamo con

$$E(\tilde{x}) = f(\tilde{x}) - P_n(\tilde{x}) = \frac{1}{(n+1)!}\omega_{n+1}(\tilde{x})f^{(n+1)}(\xi)$$

Allora $E(\tilde{x}) = 0$ quando:

- $\omega_{n+1}(x_i) = 0$
- $\frac{d^{(n+1)}f}{dx} = 0$, la derivata di ordine n+1 è nulla

Convergenza polinomio interpolatore

Al crescere del numero dei punti di interpolazione, e quindi del grado del polinomio interpolatore

- se i punti x_i sono **scelti equidistanti** nell'intervallo [a,b] $\rightarrow$ Il polinomio interpolatore **non converge** alla soluzione.
  In particolare si ha al centro dell'intervallo una buona approssimazione e delle fitte oscillazioni agli estremi, tipiche dei polinomi di grado elevato
  .
- se i punti x_i sono scelti come **zeri dei polinomi di Chebishev** $\rightarrow$ risulta minimo il termine $\omega_{n+1}(\tilde{x})$ ed all'aumentare dei punti di interpolazione si ha la **convergenza del polinomio** interpolatore

$$x_i = \cos\left(\frac{1 + 2 \cdot i}{2 \cdot (n+1)}\pi\right)$$

Costante di Lebesgue

$$\frac{||\tilde{P}_n(x) - P_n(x)||_\infty}{||P_n(x)||_\infty} \leq \Lambda \frac{||\tilde{y} - y||_\infty}{||y||_\infty}$$

Dove:

- $P_(x) = \sum_{i=0}^{n} y_i L_i(x)$ polinomio che interpola le coppie (x_i, y_i)
- $P_(x) = \sum_{i=0}^{n} y_i L_i(x)$ polinomio che interpola le coppie $(x_i, \tilde{y}_i)$

- $\Lambda_n = max_{x \in [a,b]} \sum_{i=0}^{n} |L_i(x)|$, dipende dalla scelta di nodi di interpolazione, infatti:
  - se scegliamo nodi equispaziati $\Lambda_n \approx \frac{2^{n+1}}{nlog_e(n)}$ per n grandi
  - se scegliamo nodi di Chebichev $\Lambda_n \approx \frac{2}{\pi} log_e(n)$ per n grandi

La costante di Lebesgue risulta essere il coefficiente di amplificazione degli errori relativi sui dati e pertanto identifica il numero di condizionamento del problema di interpolazione polinomiale