

METODI NUMERICI TEORIA

Numeri Finiti

$F(\beta, t, L, U)$ è l'insieme dei numeri floating point (senza numero fisso per parte intera e decimale) (β è la base, t cifre della mantissa, L e U esponenti minimi e massimi). in Python: $U=1023$, $L=-1022$, $t=53$, $\beta=2$

- Rounding to even:
 - Regola per arrotondare, riduce la distorsione di molte operazioni effettuate arrotondando per eccesso/difetto.
 - Se ultima cifra della mantissa è pari: non si incrementa
 - Se ultima cifra della mantissa è dispari: si incrementa di 1

$$\beta = 10, t = 2, \alpha = 0.185, fl(\alpha) = 0.18$$

$$\beta = 10, t = 4, \alpha = 0.37975, fl(\alpha) = 0.3798$$

- Numero macchina più piccolo rappresentabile:

$$\alpha_{\min} = \beta^{-1} \beta^L = \beta^{L-1}$$

- Numero macchina più grande rappresentabile:

$$\alpha_{\max} = (1 - \beta^{-t}) \beta^U = \beta^U - \beta^{U-t}$$

- Cardinalità dei numeri macchina:

$$\#F = 2 \cdot (\beta - 1) \beta^{t-1} (U - L + 1) + 1$$

- Spacing tra β^p e β^{p+1} (distanza tra due numeri macchina) dove p è l'esponente dell'estremo a sinistra dell'intervallo:

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

- Precisione di macchina: eps (spacing tra β^0 e β^1):

$$\text{eps} = \beta^{1-t}$$

- u (roundoff unit) differenza relativa tra 1 e il numero macchina più piccolo > di 1:

$$u := \frac{1}{2} \text{eps} = \frac{1}{2} \beta^{1-t}$$

- Errore Assoluto di approssimazione:

$$E_a = |\alpha - fl(\alpha)|$$

- Errore Relativo di arrotondamento:

$$E_{rel} = \frac{|\alpha - fl(\alpha)|}{|\alpha|}$$

Condizionamento e stabilità

- Un problema è ben posto se la sua soluzione esiste, è unica e dipende in modo continuo dai dati
- Il condizionamento di un problema permette di misurare quanto la soluzione cambia in base a un cambiamento nei dati
- $x \sim x + \sigma_x$ è il dato effetto da errore
- Problema ben Condizionato:
 - Piccole variazioni nei dati x portano a variazioni relativamente piccole nei risultati $f(x)$
- Problema mal condizionato:
 - Piccole variazioni nei dati x portano a variazioni molto grandi nei risultati $f(x)$
- Indice di Condizionamento K
 - Quantifica l'entità con cui l'errore relativo sui dati si amplifica sull'errore relativo sui risultati del problema.
 - Se K è piccolo il problema è ben condizionato
 - Se $K \leq 1$, allora è ben condizionato

$$K = \left| \frac{f'(x)x}{f(x)} \right|$$

Indice di cond. della valutazione di una funzione in un punto x

$$\frac{\|f(x) - f(\tilde{x})\|}{\|f(x)\|} \leq K \frac{\|x - \tilde{x}\|}{\|x\|}$$

- La stabilità degli algoritmi (gli indichiamo con Ψ (psi)) esprime il comportamento dell'algoritmo considerato rispetto alla propagazione degli errori
- Errore Algoritmico (deriva dalle operazioni aritmetiche in aritmetica finita):

$$E_{alg} = \frac{\Psi(\tilde{x}) - f(\tilde{x})}{f(\tilde{x})}$$

- Errore Inerente (deriva dalla rappresentazione finita dei numeri):

$$E_{in} = \frac{f(\tilde{x}) - f(x)}{f(x)}$$

- Errore totale:

$$E_{tot} = \frac{\Psi(\tilde{x}) - f(x)}{f(x)} \quad E_{tot} \approx E_{in} + E_{alg}$$

- Modulo dell'Errore Algoritmico:
 - Se $g(n)$ è lineare, algoritmo stabile
 - Se $g(n)$ è esponenziale, algoritmo instabile

$$|E_{alg}| \approx g(n) \cdot \varepsilon \quad |\varepsilon| \leq u$$

- Se ho somma di n numeri finiti, per rendere minimi gli errori, bisogna ordinarli in ordine crescente
- Se invece ho il prodotto non ho questi problemi

Norme Vettoriali

- Lunghezza di un vettore
- Proprietà:

$$1) \|x\| > 0 \quad \forall x \in R^n \quad e \quad \|x\| = 0 \text{ se e solo se } x=0$$

La norma di un vettore è sempre non negativa, è nulla se e solo se il vettore è nullo.

$$2) \|\lambda x\| = |\lambda| \cdot \|x\| \quad \forall \lambda \in R, \forall x \in R^n$$

La norma di un vettore scalato è uguale al valore assoluto dello scalare per la norma del vettore.

$$3) \|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in R^n$$

La norma di un vettore somma è minore o uguale alla somma delle norme dei due vettori.

- Norma Infinito:

$$\|x\|_{\infty} = \max_i |x_i|$$

- Norma 1:

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

- Norma 2: ($\sqrt{x^T x}$)

$$\|x\|_2 = \left[\sum_{i=1}^n x_i^2 \right]^{\frac{1}{2}}$$

- Matrice è ortogonale se $AA^T = A^TA = I$ ossia se la trasposta coincide con l'inversa
- Se la matrice è ortogonale: $\|Ax\|_2 = \|x\|_2$
- Quando si calcola l'errore relativo, nel numeratore e nel denominatore devono esserci lo stesso tipo di norme
- Relazione tra i diversi tipi di norma:

$$\|x\|_{\infty} \leq \|x\|_2 \leq \|x\|_1$$

Norme Matriciali

- Proprietà:

$$1) \|A\| > 0 \text{ per tutte } A \neq 0 \text{ e } \|A\| = 0 \Leftrightarrow A = 0$$

$$2) \|\alpha A\| = |\alpha| \|A\| \quad \forall A \in M(m \times n), \forall \alpha \in R$$

$$3) \|A + B\| \leq \|A\| + \|B\| \quad \forall A, B \in M(m \times n)$$

$$4) \|A \cdot B\| \leq \|A\| \cdot \|B\| \quad \forall A \in M(m \times p), B \in M(p \times n)$$

- Norme compatibili: (A matrice e x vettore $\rightarrow Ax$ è un vettore risultato dal prodotto)

$$\|Ax\|_v \leq \|A\|_M \cdot \|x\|_v$$

- Norma indotta ($\| \cdot \|_N$):

$$\|Ax\|_v \leq C \cdot \|x\|_v$$

- Più piccola costante C per cui $\rightarrow \|Ax\|_v \leq \|A\|_N \cdot \|x\|_v$

$$\|A\|_N = \sup_{\|x\|_v \neq 0} \frac{\|Ax\|_v}{\|x\|_v} = \max_{\|x\|_v=1} \|Ax\|_v$$

○ In conclusione \rightarrow

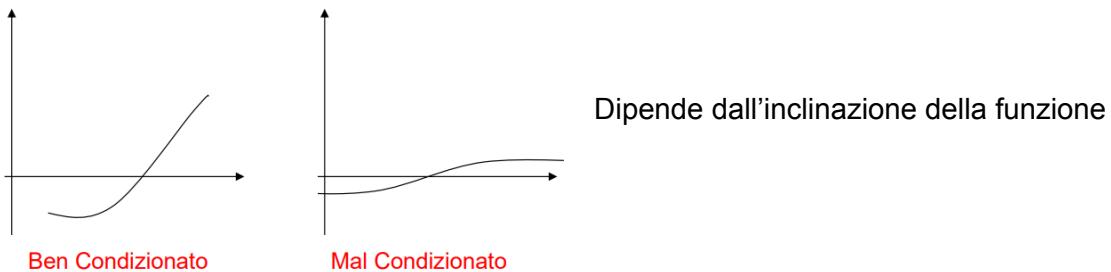
- Formule delle norme matriciali:

- $\|A\|_1 \rightarrow$ Si considera la norma 1 di tutte le colonne e si prende il valore max.
- $\|A\|_\infty \rightarrow$ Si considera la norma 1 di tutte le righe e si prende il valore max.
- $\|A\|_2 = \sqrt{\rho(A^T A)}$ dove ρ è il raggio spettrale, ossia l'autovalore di modulo massimo della matrice $A^T A$

- A è simmetrica se $A^T = A$
- λ è un autovalore di A se $\det(A - \lambda I) = 0$
- A è semidefinita positiva se $x^T A x \geq 0$
- A è definita positiva se $x^T A x > 0$
- Se A simmetrica e semidefinita positiva \rightarrow gli autovalori di A sono reali ≥ 0
- Se A simmetrica e definita positiva \rightarrow gli autovalori di A sono reali > 0
- $A^T A$ è simmetrica e semidefinita positiva

Soluzione Numerica di Equazioni non Lineari

- Obiettivo: trovare gli zeri di una funzione
- Bisogna rendere il problema ben posto:
 - Individuare un intervallo I contenente una sola radice e poi applicare il metodo iterativo fino a convergenza alla soluzione (radice).
- Radice multipla e molteplicità:
 - Se $f(\alpha) = 0$ e $f'(\alpha) = 0$ e $f''(\alpha) \neq 0 \rightarrow \alpha$ è una radice multipla con molteplicità 2
 - Esempio:
 - $f(x) = (x + 2)^2$ $x = -2$ è radice multipla di molteplicità 2
 - $f(x) = (x - 1)^6$ $x = 1$ è radice multipla di molteplicità 6
- Errore assoluto sui dati:
 - $|\tilde{\alpha} - \alpha| = |\delta|$ poiché $\tilde{\alpha} = \alpha + \delta$
- Errore assoluto sulla soluzione:
 - $|\tilde{f}(x) - f(x)| = |\varepsilon g(x)|$ poiché $\tilde{f}(x) = f(x) + \varepsilon g(x) = 0$
- Condizionamento del problema del calcolo degli zeri:
 - $|\tilde{\alpha} - \alpha| = |\delta| \approx K |\varepsilon g(\alpha)|$
 - $K = \frac{1}{|f'(\alpha)|}$
 - Se il denominatore è molto piccolo (vicino allo zero), il problema è mal condizionato, viceversa ben condizionato
- Se la molteplicità della soluzione è > 1 , è un problema numericamente molto difficile



- Per risolvere questi problemi con metodo iterativo bisogna:
 - Scegliere il valore iniziale x_0 (di innesco) e Convergenza ad α
 - Ordine di Convergenza
 - Scegliere il criterio di arresto

Ordine di Convergenza:

- Sia data una successione di iterati $\{x_k\}$ convergente ad un limite α e sia $e_k = x_k - \alpha$

Se esistono due numeri reali $p \geq 1$ e $c > 0$ tali che $\lim_{k \rightarrow +\infty} \frac{|e_{k+1}|}{|e_k|^p} = c$ si dice che la successione ha **ordine di convergenza p e fattore di convergenza c**

- Se $p = 1 \rightarrow$ convergenza lineare
- Se $1 < p < 2 \rightarrow$ convergenza superlineare
- Se $p = 2 \rightarrow$ convergenza quadratica
- Per k tendente all'infinito, la radice x_k ha $p * n$ decimali corretti

- **Metodo di Bisezione:**

$$c_k = \frac{1}{2}(a_{k-1} + b_{k-1})$$

- Si calcola: punto medio: operazione pericolosa numericamente

- Si calcola $f(c_k)$
 - Se $= 0 \rightarrow c_k$ è la radice cercata
 - Se $f(a_{k-1}) * f(c_k) < 0 \rightarrow$ si pone $b_k = c_k ; a_k = a_{k-1}$
 - Se $f(a_{k-1}) * f(c_k) > 0 \rightarrow$ si pone $a_k = c_k ; b_k = b_{k-1}$

- Poiché la formula del punto medio è pericolosa, si usa:

$$c_k = a_{k-1} + \frac{b_{k-1} - a_{k-1}}{2}$$

1. Se $f(a) * f(b) < 0$ si pone $a_0 := a$ e $b_0 := b$

2. Finchè non risulta verificato il criterio di arresto:

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

a) Se $f(x_{k+1}) * f(a_k) < 0 \Rightarrow a_{k+1} := a_k; b_{k+1} := x_{k+1}$

b) Altrimenti se $f(x_{k+1}) * f(b_k) < 0 \Rightarrow a_{k+1} := x_{k+1}; b_{k+1} := b_k$

c) Altrimenti se $f(x_{k+1}) = 0 \Rightarrow x_{k+1} := \alpha$

d) $k = k + 1$

- **Metodo Regula Falsi:** (uguale a bisezione tranne che nel porre x_{k+1})

1. Se $f(a) * f(b) < 0$ si pone $a_0 := a$ e $b_0 := b$

2. Finchè non risulta verificato il criterio di arresto:

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

a) Se $f(x_{k+1}) * f(a_k) < 0 \Rightarrow a_{k+1} := a_k; b_{k+1} := x_{k+1}$

b) Altrimenti se $f(x_{k+1}) * f(b_k) < 0 \Rightarrow a_{k+1} := x_{k+1}; b_{k+1} := b_k$

c) Altrimenti se $f(x_{k+1}) = 0 \Rightarrow x_{k+1} := \alpha$

d) $k = k + 1$

METODI DI LINEARIZZAZIONE:

- Si approssima la funzione con una retta per $(x_0, f(x_0))$:

- $y = f(x_0) + m(x - x_0)$

- In generale:

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

- **Metodo delle Corde:**

- Usa valore costante $m \neq 0$

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

- Si collegano $F(a)$ e $F(b)$ con una retta e nel punto di intersezione con l'asse x , si prende il valore della funzione in quel punto e si traccia una retta parallela a quella iniziale e passante per quel punto ecc...

- **Metodo delle Secanti:**

- Dati due valori iniziali x_0 e 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 $(x_{k-1}, f(x_{k-1})), (x_k, f(x_k))$

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

Il metodo delle secanti può essere più veloce ma non converge sempre (**convergenza locale**). Non c'è più la certezza di avere sempre il punto cercato all'interno dell'intervallo

- **Metodo di Newton:**

- A ogni passo k, si considera la retta passante per il punto $(x_k, f(x_k))$ e tangente alla funzione f(x). Questa retta rappresenta il polinomio di Taylor di grado 1, che è la retta che approssima meglio la funzione f(x) in un intorno del punto $(x_k, f(x_k))$
- Il nuovo iterato sarà l'intersezione tra la retta e l'asse delle x
- $y = f(x_k) + f'(x_k)(x - x_k)$

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

- Se lo zero ha molteplicità m \rightarrow non ha più convergenza quadratica e si usa il modificato
- Metodo di Newton Modificato (m è la molteplicità):

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

- Metodo a convergenza globale:

- La convergenza è assicurata per qualsiasi scelta del punto iniziale appartenente all'intervallo che racchiude la radice, cioè x_0 in $[a,b]$
- Bisezione, Regula Falsi

- Metodo a convergenza locale:

- La convergenza è assicurata per x_0 appartenente ad un intorno della soluzione.
- Secanti, Newton

- Teorema di convergenza locale:

Se $f : [a: b] \rightarrow R$ soddisfa le seguenti ipotesi

- i) $f(a)f(b) < 0$
- ii) f, f', f'' sono continue in $[a; b]$, ossia $f \in C^2[a; b]$
- iii) $f'(x) \neq 0 \forall x \in [a, b]$

allora esiste un intorno $I \subset [a, b]$ dell'unica radice $\alpha \in (a, b)$ tale che, se $x \in I$, allora la successione di Newton $\{x_i\}_{i \geq 1}$ converge ad α .

- Teorema di convergenza globale del metodo di Newton:

Sia $f(x) \in C^2[a, b]$, $[a, b]$ intervallo chiuso e limitato. Se sono verificate le seguenti condizioni

1. $f(a)f(b) < 0$
2. $f'(x) \neq 0 \quad \forall x \in [a, b]$
3. $f''(x) > 0 \quad \text{oppure} \quad f''(x) < 0 \quad \forall x \in [a, b]$
4. $\left| \frac{f(a)}{f'(a)} \right| < b - a \quad \left| \frac{f(b)}{f'(b)} \right| < b - a$

allora il metodo di Newton **converge** all'unica soluzione α in $[a, b]$, per ogni scelta di x_0 in $[a, b]$.

- La soluzione ottimale è utilizzare un metodo globale e dopo alcuni passi usarne uno locale

Sistemi di Equazioni non Lineari

- Jacobiano: in ogni riga c'è il gradiente di ogni f

$$J(X) = \begin{bmatrix} \frac{\partial f_1(X)}{\partial x_1} & \frac{\partial f_1(X)}{\partial x_2} & \dots & \dots & \frac{\partial f_1(X)}{\partial x_n} \\ \frac{\partial f_2(X)}{\partial x_1} & \frac{\partial f_2(X)}{\partial x_2} & \dots & \dots & \frac{\partial f_2(X)}{\partial x_n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial f_n(X)}{\partial x_1} & \frac{\partial f_n(X)}{\partial x_2} & \dots & \dots & \frac{\partial f_n(X)}{\partial x_n} \end{bmatrix} \rightarrow J(X) = \begin{bmatrix} \frac{\partial F(X)}{\partial x_1} & \frac{\partial F(X)}{\partial x_2} & \dots & \dots & \frac{\partial F(X)}{\partial x_n} \end{bmatrix}$$

$$\nabla F(X) = J^T(X)$$

- Metodo di Newton-Raphson con $n = 2$:

- Individuare un punto del piano in cui le due funzioni si annullano contemporaneamente
- Svolgo taylor nelle f e faccio il sistema dove pongo i due polinomi = 0

$$\begin{cases} 0 = f_1(x_1^{(k)}, x_2^{(k)}) + \frac{\partial f_1}{\partial x_1}(x_1^{(k)}, x_2^{(k)})(x_1 - x_1^{(k)}) + \frac{\partial f_1}{\partial x_2}(x_1^{(k)}, x_2^{(k)})(x_2 - x_2^{(k)}) \\ 0 = f_2(x_1^{(k)}, x_2^{(k)}) + \frac{\partial f_2}{\partial x_1}(x_1^{(k)}, x_2^{(k)})(x_1 - x_1^{(k)}) + \frac{\partial f_2}{\partial x_2}(x_1^{(k)}, x_2^{(k)})(x_2 - x_2^{(k)}) \end{cases} \quad (2)$$

- Definiamo:

$$F(X_k) = [f_1(x_1^{(k)}), f_2(x_2^{(k)})]^T$$

$$X - X_k = [x_1 - x_1^{(k)}, x_2 - x_2^{(k)}]^T$$

- Il sistema (2) si può esprimere in forma matriciale come:

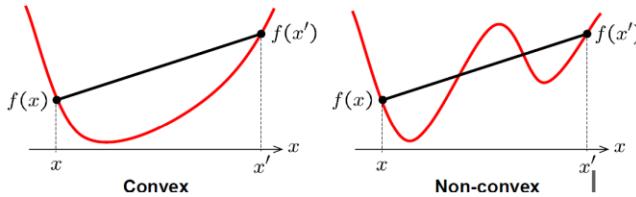
$$0 = F(X_k) + J(X_k)(X - X_k)$$

$$J(X_k)(X - X_k) = -F(X_k)$$

- (s_{k-1} è il termine noto) metodo a convergenza locale e ordine quadrattico

- | |
|---|
| 1. Valutare $J(X_{k-1})$ |
| 2. Risolvere il sistema lineare $J(X_{k-1})s_{k-1} = -F(X_{k-1})$ |
| 3. Porre $X_k = X_{k-1} + s_{k-1}$ |

- Metodo delle Corde:
 - Si utilizza sempre lo stesso Jacobiano $J(X_0)$
- Metodo di Shamanskii:
 - Si valuta lo Jacobiano ogni m iterazioni e quindi lo si utilizza per le m iterazioni successive
- Graficamente, per risolvere questi sistemi, si interseca il grafico delle figure con il piano $z = 0$, creando così delle figure 2d
- Massimi e minimi di funzioni in 2 variabili:
 - I punti in cui si annulla il gradiente si chiama punto critico
 - Si calcola l'Hessiana in quel punto
 - Si calcola il determinante dell'Hessiana
 - Minimo locale: $\det > 0$ e l'elemento in $(1, 1) > 0$
 - Massimo locale: $\det > 0$ e l'elemento in $(1, 1) < 0$
 - Sella: $\det < 0$
 - Se funzione è convessa e differenziabile: (x_0, y_0) è minimo se e solo se gradiente in quel punto = 0



- Metodo di Newton-Raphson per calcolo del minimo di una f in più variabili:

Dato $x_0 \in R^n$ ed F , per ogni iterazione k

1. Valutare $H(X_{k-1})$
2. Risolvere il sistema lineare $H(X_{k-1})s_{k-1} = -\nabla f(X_{k-1})$
3. Porre $X_k = X_{k-1} + s_{k-1}$

s_{k-1} definisce una direzione di discesa da X_{k-1} ad X_k

- Calcolo gradiente
- Calcolo Hessiana
- Scelgo un x_0 (in R^3 è un punto con due coordinate) e valuto l'Hessiana e il gradiente in quel punto

Sistemi Lineari

- Teorema di Rouchè-Capelli:
 - $Ax = b$ ammette soluzioni se e solo se la matrice A e la matrice $[A \ b]$ hanno lo stesso rango
 - Se $m < n$: se k è rango matrice A e $k < n$, sistema ammette infinite soluzioni
 - Se $m > n$: sistema ammette soluzione approssimata
 - Se $m = n$: può ammettere 1 ed 1 sola soluzione
- A è NON Singolare se soddisfa una delle seguenti:
 - $\det(A) \neq 0$
 - \exists l'inversa A^{-1} di A
 - $\text{rank}(A) = n$
- Condizione necessaria e sufficiente affinchè $Ax=b$ ammetta una ed una sola soluzione è che A si di rango massimo (A invertibile) $\rightarrow x = A^{-1}b$
 - Se $b = 0$ e A non è singolare \rightarrow esiste solo $x = 0$
 - Metodo dell'inversa poco efficiente e accurato e anche Cramer ($n+1$)!
- Indice di condizionamento: $K(A) = \|A^{-1}\| * \|A\|$
- Se A ortogonale ($A^T = A^{-1} \rightarrow AA^T = A^TA = I$) $Ax=b$ è sempre ben condizionato
- $K(A)$ piccolo ($K \leq 10^3$): problema ben condizionato
- $K(A)$ grande ($K > 10^3$): problema mal condizionato
- Matrici mal condizionate:
 - Matrice di Vandermonde: matrice $(n+1) \times (n+1)$ dove $a_{ij} = (x_i)^j$
 - Matrice di Hilbert:
$$h_{ij} = \frac{1}{i+j-1}$$
$$A = \begin{bmatrix} 1 & x_0 & (x_0)^2 & (x_0)^3 \\ 1 & x_1 & (x_1)^2 & (x_1)^3 \\ 1 & x_2 & (x_2)^2 & (x_2)^3 \\ 1 & x_3 & (x_3)^2 & (x_3)^3 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{bmatrix}$$
- Matrice densa: numero di elementi diversi da 0 è maggiore del 33%
- Matrice sparsa: numero di elementi diversi da 0 è minore del 33%
- Matrice triangolare superiore: tutti gli elementi sotto la diagonale sono 0
- Matrice triangolare inferiore: tutti gli elementi sopra la diagonale sono 0

Metodi per soluzione di sistemi lineari:

- Metodi diretti: con matrice dei coefficienti densa e di dimensioni moderate
 - Trasformare A in $A = BC \rightarrow BCx = b$
 - Spezzo problema in: $\begin{cases} By = b \\ Cx = y \end{cases}$
 - **Metodo di eliminazione gaussiana:**
 - $A = LU$ dove L (Lower) è triangolare inferiore con elementi 1, U (Upper) triangolare superiore
 - $\det(A) = \det(LU) = \det(L) * \det(U)$
 - **Metodo di Cholesky:**
 - $A = LL^T = R^T R$ dove L triangolare inferiore con elem diagonali positivi e R triangolare superiore con elem diag positivi
 - **Metodo di Householder:**
 - $A = QR$ dove Q ortogonale ($Q^{-1} = Q^T$) e R triangolare superiore
 - QR esiste sempre, LL^T esiste per matrici simmetriche e def.pos., QR no unica
- Metodi iterativi: generano successione di vettori che tendono alla soluzione, matrice dei coefficienti non viene modificata. Per matrici di grandi dimensione e sparsa

Soluzione di sistemi con matrici triangolari:

- Sostituzione in avanti:

```

for i=1,2,..,n
  x_i=b_i
  for j=1,2,...,i-1
    x_i=x_i-l_ij*x_j
  end for j
  x_i=x_i/l_ii
end for i
  
```

- Sostituzione all'indietro:

```

for i=1,2,..,n
  x_i=b_i
  for j=i,2,...,1
    x_i=x_i-l_ij*x_j
  end for j
  x_i=x_i/l_ii
end for i
  
```

- Complessità computazionale: $O(n^2)$

Fattorizzazione LU di Gauss:

- Teorema 1:

- Sia A_k la sottomatrice principale di testa di A ottenuta considerando le prime k righe e colonne di A
- Se A_k è non singolare per ogni k allora esiste ed è unica la fattorizz. LU di A
- Per risolvere sistema deve essere anche $\det(A) \neq 0$

for $k = 1, \dots, n-1$

for $i = k + 1, \dots, n$

$$l_{ik} = a_{ik} / a_{kk}$$

for $j = k + 1, \dots, n$

$$a_{ij} = a_{ij} - l_{ik} * a_{kj}$$

- Matrice di Permutazione P :

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

- Ottenuta scambiando due righe tra loro
- Se moltiplico $P * A$ equivale a scambiare le stesse due righe di A
- Se moltiplico $A * P$ scambio le colonne

- Teorema 2:

- Data A non singolare, esiste una matrice di permutazione P non singolare tc **$PA=LU$**
- Ly = Pb e Ux = y risolvere questo sistema

Inizializza la matrice P all'identità

for $k = 1, \dots, n-1$

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 $s \neq k$

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)

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

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

Fattorizzazione di Cholesky:

- Solo se matrici simmetriche e definite positive
- Se A simmetriche e def. pos. allora esiste una matrice triangolare inferiore L con elem. diagonali positivi tc **$A = LL^T$**
- $$\begin{cases} Ly = b \\ L^T x = y \end{cases}$$

Fattorizzazione Housholder:

- $\begin{cases} Qy = b \\ Rx = y \end{cases}$

Stabilità di un algoritmo di fattorizzazione:

- Sia $A = BC$ la fattorizzazione di A
- Siano $\mathcal{B} = B + \delta B$ $\mathcal{C} = C + \delta C$ dovute alle operazioni della fattorizzazione
- $\delta A = B \cdot \delta C + \delta B \cdot C + \delta B \cdot \delta C$
 - La perturbazione A , non solo dipende dalle piccole perturbazioni di B e C , ma è tanto più grande quanto più grandi sono gli elementi dei fattori B ed C .
- A , con elem. minori di 1, è stabile se esistono a e b indipendenti dall'ordine e dagli elementi di A tc: $|b_{ij}| \leq a$ $|c_{ij}| \leq b$ costanti che limitano la crescita dei fattori

Metodi Iterativi per Sistemi Lineari

- Usati per matrici sparse e di grandi dimensioni perché complessità $O(kn^2)$ k iterazioni

Metodi iterativi basati sulla decomposizione di A :

- Sia $Ax=b$ con $\det(A) \neq 0$, decomponiamo $A = M - N$ con $\det(M) \neq 0$
- Sostituendo si ha
$$x^{(k)} = Tx^{(k-1)} + q$$
 per $k = 1, 2, \dots$
 - $T = M^{-1}N$ è la matrice di iterazione e $q = M^{-1}b$
 - $x^{(k)}$ successione di iterati
- $A = D + E + F$ con D diagonale, E tr. inferiore F tr. superiore
- **Metodo di Jacobi:**
 - $A = M - N$ si ottiene con $M = D$ e $N = -(E + F)$
 - Matrice di iterazione: $T_J = M^{-1}N = -D^{-1}(E+F)$
 - Dato un sistema, dalla riga i -esima ricavo la x_i i -esima. La x_i i -esima è = al termine noto + le altri componenti del sistema tranne il termine diagonale diviso il termine diagonale
 - Si può usare questo algoritmo se gli elementi diagonali di A sono diversi da 0. Se non lo sono e A è non singolare, si possono scambiare le colonne e rispettivamente le incognite. Inoltre il metodo è programmabile in parallelo
- **Metodo di Gauss-Seidel:**
 - $A = M - N$ si ottiene con $M = E + D$ e $N = -F$
 - Matrice di iterazione: $T_G = M^{-1}N = -(E+D)^{-1}F$
 - Uguale a Jacobi ma utilizzando i risultati intermedi per aggiornare i risultati mano a mano
 - Non si presta a essere parallelizzato perché dipende da ris. precedenti

Convergenza:

- Condizione necessaria e sufficiente per la convergenza del metodo iterativo è che il raggio spettrale (autovalore di modulo massimo) della matrice di iterazione T sia minore di 1

$$\rho(T) < 1$$

$$\frac{\|e^{(k+1)}\|}{\|e^{(k)}\|} \approx \rho(T) \quad e^{(k)} = x^{(k)} - x, \quad k = 0, 1, \dots \quad \text{e è l'errore}$$

•

Calcolo Differenziale in più Variabili

- Derivata direzionale: (= prodotto scalare del gradiente per il vettore direzione)

$$D_v f(x_0, y_0) = \nabla f(x_0, y_0) \cdot v = \frac{\partial f(x_0, y_0)}{\partial x} \cdot v_x + \frac{\partial f(x_0, y_0)}{\partial y} \cdot v_y$$

- Gradiente è direzione di massima crescita di f e -gradiente di massima decrescita
- Teorema di Schwarz: derivata seconda di f prima rispetto a x poi a y = al contrario
- Forma quadratica: $q(h_1, h_2) = a_{11}h_1^2 + a_{12}h_1h_2 + a_{21}h_2h_1 + a_{22}h_2^2$
- Forma quadratica definita positiva è convessa ed ammette un unico punto di minimo

Metodi di Discesa

- Se A simmetrica e def. pos., la soluzione del sistema Ax=b è il punto di minimo della

$$\text{funzione quadratica: } \frac{1}{2}x^T Ax - b^T x = \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j - \sum_{i=1}^n b_i x_i$$

1. Parti con qualche $x^{(0)}$, $k = 0$

2. Determina la direzione di discesa $p^{(k)}$

3. Scegli lo step-size $\alpha^{(k)}$ tale che

$$F(x^k + \alpha^{(k)} p^{(k)}) < F(x^{(k)})$$

4. Aggiorna l'iterato

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} p^{(k)}$$

5. Incrementa il contatore $k=k+1$

generico alg. si ripete fino a convergenza

$$\bullet \quad \alpha^{(k)} = -\frac{\langle r^{(k)}, p^{(k)} \rangle}{\langle Ap^{(k)}, p^{(k)} \rangle} = -\frac{(r^{(k)})^T p^{(k)}}{(p^{(k)})^T A p^{(k)}} \quad \text{è lo step-size}$$

$$\bullet \quad \boxed{\langle \nabla F(x^{(k)}), p^{(k)} \rangle < 0} \quad \text{condizione di ammissibilità di } p^{(k)}$$

- p è la direzione

- Metodo Steepest Descent:
 - Scelgo a ogni passo k , p come l'antigradiente di F , quindi la direzione di max. decrescita

1. Parti con qualche $x^{(0)}$, $k = 0$

2. Calcola la direzione di discesa più ripida

$$p^{(k)} = -\nabla f(x^{(k)})$$

3. Scegli lo stepsize $\alpha^{(k)}$ tale che

$$F(x^k + \alpha^{(k)} p^{(k)}) < F(x^{(k)})$$

$$\alpha^{(k)} = -\frac{\langle r^{(k)}, p^{(k)} \rangle}{\langle Ap^{(k)}, p^{(k)} \rangle}$$

4. Aggiorna l'iterato

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} p^{(k)}$$

$$r^{(k+1)} = r^{(k)} + \alpha^{(k)} A p^{(k)}$$

5. Incrementa il contatore $k=k+1$

si fa fino a convergenza

si raggiunge convergenza quando

$$\|r^{(k+1)}\|_2 \leq tolleranza$$

metodo a convergenza lineare

$$e^{(k)} = x^{(k)} - x^*$$

e è l'errore e ρ è il fattore di convergenza

$$\frac{\|e^{(k+1)}\|_A}{\|e^{(k)}\|_A} \approx \rho$$

Metodi di Discesa (Gradiente Coniugato)

- Due vettori p e q sono coniugati rispetto ad A se $p^T A q = 0$ dove A è la matrice dell'ellisse. cioè sono ortogonali ma in una sorta di altro spazio
- Questo metodo tiene conto sia del gradiente r (il residuo coincide con il gradiente) che della direzione di discesa dell'iterazione precedente

```

Scelto x(0) arbitrario, si calcola r(0)=Ax(0) - b, si prende
p(0) = -r(0)
k=0;
while arresto>=ε
    αk = -<r(k), p(k)>
    x(k+1) = x(k) + αkp(k)
    r(k+1) = r(k) + αkAp(k)
    arresto = ||r(k+1)||22
    γk+1 = <r(k+1), r(k+1)>
    p(k+1) = -r(k+1) + γk+1p(k)
    k=k+1
end while

```

Osservazione: l'algoritmo del gradiente coniugato così ottimizzato necessita di un'unica moltiplicazione matrice per vettore per ogni iterazione.

ordine di convergenza
lineare

$$e^{(k)} = x^{(k)} - x^*$$

$$\frac{\|e^{(k+1)}\|_A}{\|e^{(k)}\|_A} \approx \rho$$

Sistemi Lineari Sovradeterminati

- Hanno num. di equazioni superiore al num. di incognite: possono non avere soluzioni
- Risoluzione di questi sistemi è problema mal posto (può non esistere soluzione o averne +)

Risoluzione di sistemi lineari sovradeterminati Least Squares:

- Cerchiamo la soluzione che minimizza la norma 2 al quadrato del residuo:

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$$

- Per risolverlo uso equazioni normali: $Gx = A^Tb$ dove $G = A^TA$ matrice nxn
- Ha soluzione unica se $\text{rank}(A) = n$ ($\det(A)$ diverso da 0)
- La soluzione è $x = (G)^{-1} A^Tb$
- Il calcolo di A^TA può amplificare errori numerici, per approccio più stabile si usa la fattorizzazione QR
- Teorema:

Dato il sistema lineare sovradeterminato $Ax = b$

$$x^* = \arg \min ||Ax - b||_2^2 \Leftrightarrow x^* \text{ è la soluzione di } A^T Ax = A^T b$$

La soluzione è unica se e solo se la matrice A ha rango massimo $\text{rank}(A) = n$

La soluzione del problema dei minimi quadrati mediante equazioni normali richiede solo che la matrice A del sistema sovradeterminato $Ax=b$ abbia rango massimo.

La matrice $G = A^TA$ è simmetrica e definita positiva e quindi il sistema può essere risolto utilizzando il metodo di Cholesky

Metodo QR:

- Fattorizzo $A = QR$ dove Q ortogonale e R triangolare superiore
- Il problema diventa: $\min_x \|Ax - b\| = \min_x \|QRx - b\| = \min_x \|Rx - Q^T b\|$
- Si risolve un sistema triangolare $Rx = Q^T b$ più stabile ed efficiente

Decomposizione in Valori Singolari di una matrice: SVD

- Il metodo delle equazioni normali e il metodo QR richiedono rango massimo. Se A non ha rango massimo si usa SVD
- $A = U\Sigma V^T$ U ortogonale, V ortogonale, S matrice diagonale con valori singolari $>= 0$
- Valori singolari sono i valori sulla diagonale di S:
 - σ_1 è il massimo valore singolare
 - $\sigma_{\max} / \sigma_{\min}$ dà l'indice di condizionamento di A
 - Numero di valori singolari non nulli è il rango di A
- Il problema dei minimi quadrati diventa $x = A^+b = V\Sigma^+U^T b$
 - A^+ è la pseudo-inversa di A
 - S^+ si ottiene invertendo i valori singolari non nulli e trasponendo

Interpolazione Polinomiale di Dati Sperimentali

- Ho una serie di coppie di punti noti $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$
 - Gli x_i sono detti nodi e devono essere tutti diversi
 - Gli y_i sono i valori del fenomeno che osservi nei punti x_i
- Si vuole trovare un polinomio di grado al più n tc: $P_n(x_i) = y_i$ per ogni i
 - Ossia voglio costruire un polinomio che passa per tutti i punti dati
 - $P_n(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots + \alpha_n x^n$
- Equivale a individuare i coefficienti α di $P_n(x)$ che soddisfi le condizioni $P_n(x_i) = y_i$ ossia la **condizione di interpolazione**
- $\xi_1 = \min\{x_0, x_1, \dots, x_n\}$ e $\xi_2 = \max\{x_0, x_1, \dots, x_n\}$:
 - Interpolazione se $x \in [\xi_1, \xi_2]$
 - Estrapolazione se $x \notin [\xi_1, \xi_2]$
- Tutto ciò equivale a un sistema $A\alpha = y$ con A matrice di Vandermonde
- Il polinomio interpolatore esiste ed è sempre unico perché matrice di Vandermonde
- La matrice dei coefficienti è quadrata perché il numero delle condizioni che imponiamo è uguale al numero delle incognite. E' sempre a rango massimo perché la matrice di Vandermonde ha sempre rango massimo se $x_i \neq x_k$ se $i \neq k$
- Quindi siamo sicuri che, dal punto di vista teorico, il problema dell'interpolazione polinomiale ammette sempre soluzione e questa è unica.
- A (matrice di Vandermonde) matrice molto mal condizionata, si usa la base di Lagrange per lo spazio vettoriale
- Date le coppie di punti il polinomio che le interpola nella forma di Lagrange è:

$$P_n(x) = \sum_{j=0}^n y_j L_j^{(n)}(x) \quad L_j^{(n)}(x) = \frac{(x - x_0)(x - x_1)\dots(x - x_{j-1})(x - x_{j+1})\dots(x - x_n)}{(x_j - x_0)(x_j - x_1)\dots(x_j - x_{j-1})(x_j - x_{j+1})\dots(x_j - x_n)}$$

- Costo computazionale polinomio nella forma di Lagrange: $O(n^2)$
- Presenta difficoltà di applicazione
- Non si ha la convergenza del polinomio interpolatore alla funzione che ha generato i dati
- **Definizione** $\lambda_n(x) := \sum_{i=0}^n |L_i(x)|$ funzione di Lebesgue.
-