

# ANALISI NUMERICA I

## IDEE DI BASE E RICHIAMI

### 1. Quale è lo scopo dell'Analisi numerica.

Trovare gli algoritmi che risolvono un problema matematico nel minor tempo possibile e con la massima accuratezza.

### 2. Attraverso quali fasi si passa nel risolvere un problema reale.

Quando si deve risolvere un problema reale, si deve passare attraverso varie fasi che possono essere ripetute più e più volte, per ottenere l'algoritmo migliore per accuratezza e velocità di esecuzione. Le varie fasi si possono riassumere così:

1. definizione del modello reale
2. costruzione del modello matematico tramite induzione
3. formulazione del problema matematico
4. risoluzione del problema matematico (analisi numerica)
5. analisi della soluzione trovata tramite metodi deduttivi (analisi numerica)
6. verifica che la soluzione trovata risolva il problema reale (in caso negativo si ripete il processo dal punto 1 operando una modifica al modello di cui al punto 2)

### 3. Dove si colloca l'Analisi Numerica nel processo di risoluzione di un problema reale.

Nelle fasi che interessano la *risoluzione del problema matematico* (4) e l'*analisi della soluzione trovata* (5).

Partendo dal presupposto che per la maggior parte dei problemi non esiste una soluzione unica, possiamo modificare il modello, fare una nuova formulazione e risoluzione del problema e tramite l'analisi numerica fare un confronto fra le varie soluzioni ottenute, in modo da scegliere quella più appropriata per velocità di esecuzione e accuratezza.

### 4. Cosa è uno spazio vettoriale (spazio lineare).

Sia  $X = \{x, y, z, \dots\}$  un insieme di elementi di natura qualsiasi e  $K = \{\alpha, \beta, \gamma, \dots\}$  un campo di scalari.

Siano definite le seguenti operazioni:

- $x + y \quad x, y \in X$  (somma tra due elementi qualunque di X)
- $\alpha \cdot x \quad \alpha \in K$  (prodotto scalare tra ciascun elemento di K e ciascun elemento di X)

tali che soddisfino le seguenti proprietà:

1.  $x \in X, y \in X \Rightarrow (x + y) \in X$  (chiusura per la somma)
2.  $x \in X, \alpha \in K \Rightarrow (\alpha x) \in X$  (chiusura per il prodotto)
3.  $x + y = y + x$  (commutativa)
4.  $(x + y) + z = x + (y + z)$  (associativa per la somma)
5.  $\exists 0 \in X : x + 0 = x, \forall x \in X$  (elemento neutro della somma)
6.  $\forall x \in X, \exists -x \in X : x + (-x) = 0$  (elemento opposto)
7.  $\alpha(\beta x) = (\alpha\beta)x$  (associativa per il prodotto)
8.  $\alpha(x + y) = \alpha x + \alpha y$  (distributiva)
9.  $(\alpha + \beta)x = \alpha x + \beta x$  (distributiva)
10.  $1x = x$  (elemento neutro del prodotto)

Allora X è detto SPAZIO VETTORIALE (o LINEARE) sul campo K.

Se K è l'insieme dei numeri reali, allora X è detto SPAZIO LINEARE REALE.

### 5. Dai 3 esempi di spazi vettoriali.

- insieme di tutte le n-ple di numeri reali :  $R^n = \{(x^1, x^2, \dots, x^n) \mid x^1, x^2, \dots, x^n \in R\}$
- insieme di tutte le matrici  $n \times n$ :  $M^{n \times n}$
- insieme di tutti i polinomi di grado  $\leq n$  :  $P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$
- insieme di tutte le funzioni continue nell'intervallo  $[a, b]$  :  $C[a, b]$
- insieme di tutte le funzioni con derivate continue fino ad m nell'intervallo  $[a, b]$  :  $C^{(m)}[a, b]$

### 6. Dimostra che l'insieme $C[a,b]$ di tutte le funzioni continue nell'intervallo $[a,b]$ è uno spazio vettoriale sull'insieme dei numeri reali R

La somma ed il prodotto scalare sono operazioni chiuse infatti la somma di funzioni continue è ancora una funzione continua inoltre vengono soddisfatte le proprietà di uno spazio vettoriale.

### 7. Quando n elementi $x_1, x_2, \dots, x_n$ , di uno spazio vettoriale su R si dicono linearmente indipendenti.

Quando  $\forall \alpha_i \in R$ , la combinazione lineare  $\alpha_1x_1 + \alpha_2x_2 + \dots + \alpha_nx_n = 0$  se e solo se  $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$ , cioè se ogni loro combinazione lineare nulla è formata da reali tutti nulli, in caso contrario si dice che sono *linearmente dipendenti*.

### 8. Dai un esempio di 3 elementi nello spazio $C[a,b]$ di tutte le funzioni continue in $[a,b]$ che sono linearmente indipendenti.

$$P_0(x) = 5, \quad P_1(x) = 2x, \quad P_2(x) = 3x^2$$

### 9. Le 3 funzioni $y_1 = \cos(x)$ , $y_2 = \sin(x)$ e $y_3 = 2\cos(x) + 4\sin(x)$ sono linearmente indipendenti? Giustifica la risposta.

No perché  $y_3$  è combinazione lineare delle prime due infatti risulta  $y_3 = 2y_1 + 4y_2$ .

Lo spazio ha dimensione 2 perché esistono solo 2 elementi linearmente indipendenti.

### 10. Che cosa è una base di uno spazio vettoriale.

Un insieme di n elementi  $x_1, x_2, \dots, x_n$ , linearmente indipendenti di uno spazio X è chiamato base per X se  $\forall y \in X$  può essere espresso come loro combinazione lineare  $y = \alpha_1x_1 + \alpha_2x_2 + \dots + \alpha_nx_n$ , con  $\alpha_i \in R^n$

### 11. Cosa significa dire che uno spazio lineare ha dimensione finita

Significa che esistono n elementi linearmente indipendenti (dimensione dello spazio) ed ogni insieme di n+1 elementi è linearmente dipendente.

### 12. Dai 3 esempi di spazi lineari di dimensione 3.

- lo spazio  $R^3$  dei reali a 3 componenti :  $R^3 = \{(x^1, x^2, x^3) \mid x^1, x^2, x^3 \in R\}$
- l'insieme delle matrici quadrate di ordine 3 diagonali.
- l'insieme dei polinomi di grado  $\leq 2$  :  $P_2(x) = a_0 + a_1x + a_2x^2$

**13. L'insieme di tutti i polinomi di grado <4 è uno spazio lineare di dimensione finita? Dai una sua base.**

È uno spazio lineare a dimensione finita di dimensione 4. Una base è  $\{p_0(x)=1, p_1(x)=x, p_2(x)=x^2, p_3(x)=x^3\}$ .

**14. I vettori [1 2] e [2 0] costituiscono una base per lo spazio vettoriale  $\mathbb{R}^2$  ovvero con una combinazione lineare di [1 2] e [2 0] puoi formare una qualsiasi coppia di numeri reali?**

È una base di  $\mathbb{R}^2$  perché  $\alpha(1, 2) + \beta(2, 0) = 0$  se e solo se  $\alpha=\beta=0$ , ogni coppia di reali  $(x, y)$  può essere scritta come combinazione lineare della base, se così non fosse non ci troveremmo più in  $\mathbb{R}^2$  ma in  $\mathbb{R}^n$  con  $n > 2$ . Quindi per ogni coppia di reali otteniamo  $(x, y) = \alpha(1, 2) + \beta(2, 0)$  da cui  $\alpha = y/2$  e  $\beta = (2x - y)/4$ .

**15. Cosa è l'autovalore di una matrice.**

Data una matrice quadrata A di ordine n, di elementi reali complessi, si chiama autovalore di A un numero  $\lambda \in \mathbb{C}$  per il quale esiste un vettore  $u \in \mathbb{C}^n$ , non nullo, chiamato autovettore tale che:  $Au = \lambda u$ .

Una matrice quadrata  $n \times n$  ammette sempre n autovalori (in campo complesso).

**16. Cosa è il raggio spettrale di una matrice.**

L'insieme dei autovalori ( $\lambda_i$ ) di una matrice è detto *Spettro* della matrice; il massimo modulo degli autovalori è detto *raggio spettrale*.

Data una matrice quadrata A di ordine n, di elementi reali complessi, si definisce raggio spettrale di A il numero:

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$$

**17. Come si definisce la norma di uno spazio vettoriale.**

Uno spazio vettoriale X è detto normato se esiste una applicazione  $X \rightarrow \mathbb{R}^+$  (chiamata norma) ed indicata usualmente con  $\|x\|$ , tale che:

1.  $\|x\| \geq 0, \forall x \in X$ ;
2.  $\|x\| = 0$ , se e solo se  $x = 0$ ; (in assenza di questa proprietà lo spazio è detto seminormato)
3.  $\|\alpha x\| = |\alpha| \|x\|, \forall \alpha \in \mathbb{K} \text{ e } x \in X$ ;
4.  $\forall x, y \in X$  vale la disuguaglianza triangolare  $\|x + y\| \leq \|x\| + \|y\|$

**18. Dai 3 esempi di norme definite in  $\mathbb{R}^n$ , lo spazio di tutte le n-uple di numeri reali.**

Si definisce norma p con  $1 \leq p \leq \infty$  di vettori  $x \equiv (x_i)$  in  $\mathbb{R}^n$  la quantità:

$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}$$

- $p = 1$  norma di Manhattan
- $p = 2$  norma Euclidea
- $p = \infty$  norma del Massimo  $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$

**19. Cosa è la disuguaglianza triangolare in uno spazio vettoriale normato. Vale sempre?**

È la proprietà per la quale  $\forall x, y \in X, \|x + y\| \leq \|x\| + \|y\|$  vale sempre infatti fa parte della definizione di norma di uno spazio vettoriale.

**20. Come si può definire la distanza tra due elementi di uno spazio vettoriale normato.**

Dati due elementi  $x, y \in X$  si definisce distanza attraverso la norma della differenza dei due vettori  $\text{dist}(x, y) = \|x - y\|$

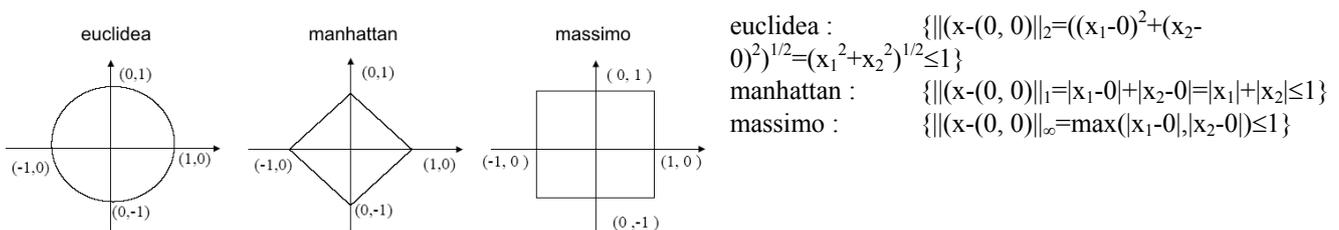
Posto  $x \equiv (x_1, x_2)$  e  $y \equiv (y_1, y_2)$  in  $\mathbb{R}^2$ :

euclidea:  $\text{dist}(x, y) = [(x_1 - y_1)^2 + (x_2 - y_2)^2]^{1/2}$

manhattan:  $\text{dist}(x, y) = |x_1 - y_1| + |x_2 - y_2|$

massimo:  $\text{dist}(x, y) = \max(|x_1 - y_1|, |x_2 - y_2|)$

**21. Considera lo spazio  $\mathbb{R}^2$ . indica graficamente la sfera unitaria  $S = \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$  considerando la distanza relativa alla norma euclidea, la norma di Manhattan e la norma del massimo.**



euclidea:  $\{\|(x-(0,0))\|_2 = ((x_1-0)^2 + (x_2-0)^2)^{1/2} = (x_1^2 + x_2^2)^{1/2} \leq 1\}$   
 manhattan:  $\{\|(x-(0,0))\|_1 = |x_1-0| + |x_2-0| = |x_1| + |x_2| \leq 1\}$   
 massimo:  $\{\|(x-(0,0))\|_\infty = \max(|x_1-0|, |x_2-0|) \leq 1\}$

**22. Dai due esempi di norme per lo spazio  $C[a,b]$  di tutte le funzioni continue in  $[a,b]$ .**

(La norma di un vettore  $x \in \mathbb{R}^n$  è una funzione continua).

$$\|f\| = \max_{a \leq x \leq b} |f(x)|, \quad \|f\| = \int_a^b |f(x)| dx$$

**23. Data la funzione  $y = \text{sen}(x)$ , dai due funzioni in  $[-2\pi, 2\pi]$  la cui distanza da y è uguale a 4 secondo la norma del massimo per funzioni.**

$$\|f\|_\infty = \max_{x \in [a,b]} |f(x)| \quad y_1 = -3\text{sen}(x) \quad y_2 = 5\text{sen}(x);$$

**24. Considera le funzioni  $y_1 = 1.5x$  e  $y_2 = x$  definite nell'intervallo  $[0,1]$ . Quale è la distanza tra le due funzioni se si considera come distanza quella derivata dalla norma di funzione  $\|f\|_\infty = \max_{x \in [0,1]} |f(x)|$**

$$\|f - g\| = \|1.5x - x\| = 0.5x$$

**25. Considera le funzioni  $y_1 = 1.5x$  e  $y_2 = x$  definite nell'intervallo  $[0,1]$ . Quale è la distanza tra le due funzioni se si considera come distanza quella derivata dalla norma di funzione:**

$$\|f\| = \left( \int_0^1 f(x)^2 dx \right)^{1/2} \rightarrow \|y_2 - y_1\| = \left\| -\frac{3}{2}x + x \right\| = \left( \int_0^1 \left(-\frac{1}{2}x\right)^2 dx \right)^{1/2} = \left( \int_0^1 \frac{1}{4}x^2 dx \right)^{1/2} = \left( \left[ \frac{1}{12}x^3 \right]_0^1 \right)^{1/2} = \left( \frac{1}{12} \right)^{1/2} = \frac{1}{2\sqrt{3}}$$

**26. Cosa significa dire che una norma di una matrice  $n \times n$  è consistente o compatibile con una norma di vettore in  $\mathbb{R}^n$**

Una norma si dice consistente o compatibile se verifica la condizione  $\|Ax\| \leq \|A\| \|x\|$  dove  $A$  è una matrice quadrata di ordine  $n$  ed  $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$  (norma naturale di matrice)

**27. Come si definisce una norma naturale indotta dalla norma di vettore.**

Sia  $\|\cdot\|$  una norma fissata di vettore e con  $A$  una matrice quadrata di ordine  $n$  si definisce norma naturale di  $A$  come l'applicazione che ad  $A$  associa la quantità  $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$  per ogni  $x \neq 0$  si può definire  $u = x/\|x\|$  (norma sfera unitaria,  $\|u\| = 1$ )  $\|A\| = \sup_{\|u\|=1} \|Au\| = \|Ay\|$  con  $\|y\| = 1$ .

**28. Quanto vale la norma della matrice identità  $n \times n$  se la norma è naturale.**

Se  $I$  è la matrice identica  $n \times n$  e  $\|\cdot\|$  è una norma naturale, allora  $\|I\| = \sup_{x \neq 0} \frac{\|Ix\|}{\|x\|} = \sup_{x \neq 0} \frac{\|x\|}{\|x\|} = 1$  è 1

**29. Dati  $A$ , matrice  $n \times n$  e  $x$  vettore in  $\mathbb{R}^n$ , verifica che per una norma naturale si ha  $\|Ax\| \leq \|A\| \|x\|$**

$$\|Ax\| = \|Axu\| \leq \|A\| \|xu\| \leq \|A\| \|x\| \|u\| = \|A\| \|x\| \text{ con } \|u\| = 1$$

**30. Quale è la norma di matrice indotta dalla norma di vettore**

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \rightarrow \|A\| = \max_i \sum_{j=1}^n |a_{ij}| \text{ cioè il massimo tra le somme dei moduli delle righe}$$

**31. Quale è la norma di matrice indotta dalla norma di vettore**

$$\|x\|_2 = \left( \sum_{i=1}^n x_i^2 \right)^{1/2} \rightarrow \|A\| = \sqrt{\rho(A^T A)} \text{ la norma spettrale, se la matrice è simmetrica coincide con il raggio Spettrale.}$$

$\rho(A^T A) = \max_s |\lambda_s(A^T A)|$ ,  $\lambda_s(A^T A)$  indica s-m autovalore della matrice  $A^T A$  (matrice simmetrica semidefinita positiva, pertanto con tutti gli autovalori  $\geq 0$ )

**32. Quale è la norma di matrice indotta dalla norma di vettore**

$$\|x\|_1 = \sum_{i=1}^n |x_i| \rightarrow \|A\| = \max_j \sum_{i=1}^n |a_{ij}| \text{ cioè il massimo tra le somme dei moduli delle colonne}$$

**33. Perché è utile definire la norma in uno spazio vettoriale.**

Poiché attraverso la norma possiamo definire la distanza  $\text{dist}(x, y) = \|x - y\|$  tra due punti qualsiasi  $x, y \in X$ . Ciò è particolarmente utile nell'ambito dell'approssimazione.

**34. Considera lo spazio vettoriale  $\mathbb{R}^2$  in cui sia definita una norma  $\|\cdot\|$ . Supponi che valga la relazione  $\lim_{x \rightarrow [2,3]} \|x\| = 12$**

**Quanto vale la norma di  $[2,3]$ . Giustifica la risposta.**

La norma è una funzione continua, ovvero  $\lim_{x_i \rightarrow \infty} x_i = x^*$  implica  $\lim_{x_i \rightarrow \infty} \|x_i\| = \|x^*\|$  quindi se  $x$  tende a  $(2, 3)$  allora  $\|x\|$  tende a  $\|(2, 3)\| = 12$

**35. Quale limitazione vale per ogni norma di matrice  $n \times n$**

Per ogni norma naturale ed ogni matrice quadrata  $A$  si ha  $\|A\| \geq \rho(A)$ . Quindi  $\rho(A)$ , raggio spettrale di  $A$ , è il limite inferiore di tutte le norme naturali, infatti se  $\lambda_s$  è un autovalore generico di  $A$  e  $\mu_s$  è il suo corrispettivo autovettore di norma 1 si ha

$$\|A\| = \max_{\|y\|=1} \|Ay\| \geq \|A\mu_s\| = \|\lambda_s \mu_s\| = |\lambda_s| \|\mu_s\| = |\lambda_s| \text{ con } \|\mu_s\| = 1$$

**36. Dai un esempio di spazio vettoriale a dimensione infinita.**

L'insieme di tutte le funzioni continue  $C[a,b]$  nell'intervallo  $[a, b]$  è uno spazio di dimensione infinita perché non esiste un insieme finito di funzioni la cui combinazione lineare produce una qualsiasi funzione.

**37. Quando un insieme  $x_1, x_2, \dots$ , di elementi di uno spazio normato è detto chiuso in  $X$ .**

Un insieme di elementi  $x_1, x_2, \dots$  di uno spazio normato  $X$ , è detto chiuso o completo in  $X$  se  $\forall x \in X$  e  $\varepsilon > 0$ ,  $\exists$  un intero  $n$  e un insieme di scalari  $\alpha_1, \alpha_2, \dots, \alpha_n$  tali che  $\|x - \sum_{i=1}^n \alpha_i x_i\| \leq \varepsilon$

**38. Quando uno spazio lineare si dice separabile.**

Quando possiede una base numerabile o finita. Per base (nel caso si intenda uno spazio a dimensione infinita) si intende un qualunque insieme  $\{x_i\}$  di elementi chiuso in  $X$  per il quale si abbia che tutti i suoi sottoinsiemi finiti sono linearmente indipendenti, in ordine col teorema di Weierstrass.

**39. Lo spazio  $C[a,b]$  di tutte le funzioni continue in  $[a,b]$  è separabile?**

Sì, infatti l'insieme  $\{1, x, x^2, \dots\}$  è una base numerabile; infatti ogni sottoinsieme finito è linearmente indipendente ed ogni funzione continua si può approssimare con la precisione voluta da un polinomio.

**40. Come si definisce il prodotto scalare in uno spazio lineare.**

Dato uno spazio lineare  $X$ , il prodotto scalare è una funzione che ad ogni coppia  $x, y \in X$  associa un numero reale indicato con  $(x, y)$  e tale che:

1.  $(x, x) \geq 0$  e se  $(x, x) = 0 \Rightarrow x = 0$
2.  $(x, y) = (y, x)$
3.  $(\alpha x, y) = \alpha (x, y)$
4.  $(x + y, z) = (x, z) + (y, z)$

Dato uno spazio in cui è definito un prodotto scalare si può passare ad uno spazio normato ponendo:  $\|x\| = (x, x)^{1/2}$

**41. Dai un esempio di prodotto scalare definito in uno spazio lineare**

$$\text{In } \mathbb{R} : p = (x, y) = x \cdot y$$

$$\text{In } \mathbb{R}^2 : p = (x, y) = x^T y = x_1 y_1 + x_2 y_2 = \|x\|_2 \cdot \|y\|_2 \cos(\theta)$$

$$\text{In } \mathbb{R}^n : p = (x, y) = x^T y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

## 42. Dai un esempio di prodotto scalare definito nello spazio $C[a,b]$ di tutte le funzioni continue in $[a,b]$

Definiamo il prodotto scalare di due funzioni continue  $\langle f(x), g(x) \rangle$  si assocerà  $\int_a^b p(x)f(x)g(x)dx$  dove  $p(x)$  è la funzione peso in quanto il prodotto di funzioni continue è ancora continua e valgono le proprietà.

## 43. Cosa è un operatore $L$ lineare.

Un operatore è detto lineare se il suo dominio è uno spazio lineare e vale la relazione  $L(\alpha x + \beta y) = \alpha Lx + \beta Ly$  per tutti gli  $x, y \in D(L)$  e tutti gli scalari  $\alpha, \beta \in K$  (campo su cui è definito  $D(L)$ )

## 44. Trasformazioni ed Operatori

Legge o regola che associa ad ogni elemento  $S_1 \subseteq X$  un unico elemento di  $S_2 \subseteq Y$  con  $X$  e  $Y$  spazi lineari.

$Tx = y, x \in X$  e  $y \in Y$

$T$  è l'operatore della trasformazione

per le trasformazioni possono definirsi le operazioni:

$(T_1 + T_2)x = T_1x + T_2x$  somma  
 $(\alpha T)x = \alpha (Tx)$  prodotto per scalare  
 $(T_1 T_2)x = T_1(T_2x)$  prodotto tra operatori

## 45. Operatore Lineare Limitato

Un operatore  $T: X \rightarrow Y$  con  $X$  e  $Y$  spazi lineari, è detto limitato se e solo se esiste una costante finita  $c$  tale che:

$$\|Tx_1 - Tx_2\| \leq c \|x_1 - x_2\|, \quad \forall x_1, x_2 \in X$$

L'insieme di tutti gli operatori lineari limitati da uno spazio lineare  $X$  in uno spazio lineare  $Y$  con le operazioni di somma e prodotto per uno scalare è uno spazio lineare indicato con  $L(X, Y)$

## 45. Matrice convergente

definizioni equivalenti:  $\lim_{m \rightarrow \infty} A^m = 0, \quad \lim_{m \rightarrow \infty} \|A^m\| = 0, \quad \rho(A) < 1$

## ERRORI E RAPPRESENTAZIONE DEI NUMERI IN UN CALCOLATORE

### 1. Cosa sono l'errore assoluto e l'errore relativo

Dato un valore di un numero reale  $a$  e un altro che lo approssima  $a'$  si definisce:

*l'errore assoluto* è la quantità  $a' - a$ , estendendo il discorso ad uno spazio Vettoriale  $\|a' - a\|$

*l'errore relativo* è la quantità  $(a' - a) / a$ , se  $a \neq 0$ , estendendo ad uno spazio Vettoriale  $\|a' - a\| / \|a\|$  (si esprime spesso in percentuale)

### 2. Indica i vari tipi di errore che possono incontrarsi nell'Analisi numerica

I risultati numerici possono essere influenzati da diversi tipi di errore. Li possiamo classificare nel modo seguente:

1. Semplificazioni introdotte nel modello. Sono errori dovuti all'adeguatezza del modello. Tra questi ci sono per esempio delle semplificazioni dovute al fatto che alcune grandezze si ritengono trascurabili;
2. Errori nei dati. Sono errori dovuti ai dati del problema, risultato di misurazioni che possono essere influenzate da errori sistematici e/o errori random, o dalla concomitanza di eventi imprevedibili;
3. Errori di arrotondamenti nei dati e nei calcoli. Sono errori dovuti alla rappresentazione dei numeri sul calcolatore;
4. Errori di troncamento. Sono errori che si introducono quando un procedimento infinito è approssimato mediante un procedimento finito.

### 3. Cosa è la rappresentazione posizionale dei numeri che utilizziamo

In generale un numero intero positivo  $N$ , viene espresso in notazione posizionale in base  $\beta$  nel seguente modo:

$N = d_n \beta^n + d_{n-1} \beta^{n-1} + \dots + d_0 \beta^0$  a rappresentato con il simbolo  $N = (d_n d_{n-1} \dots d_0)_\beta$  dove le cifre  $d_i$  sono numeri interi o simboli compresi tra 0 e  $\beta$ .

### 4. Cosa è la base di un sistema di rappresentazione dei numeri naturali.

La base è il numero di simboli tramite i quali viene rappresentato un numero, se essa è minore o uguale a 10, vengono usate le cifre intere altrimenti lettere ed altri simboli, per esempio una base binaria ha i simboli 0, 1, una base ottale ha simboli 0, 1, 2, 3, 4, 5, 6, 7, una base esadecimale ha simboli 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F, e così via.

### 5. Come può rappresentarsi in modo univoco un qualsiasi numero reale

Dato un intero  $\beta > 1$ , un numero reale  $x \neq 0$  si può esprimere in forma univoca nella seguente forma:

$$x = \text{sign}(x)(d_1 \beta^{-1} + d_2 \beta^{-2} + \dots) \beta^p = \text{sign}(x) m \beta^p$$

$m = (d_1 \beta^{-1} + d_2 \beta^{-2} + \dots)$  mantissa

$\beta^p$  esponente

$p$  caratteristica di  $x$

con:

1.  $\text{sign}(x) = 1$  se  $x > 0$ ,  $\text{sign}(x) = -1$  se  $x < 0$
2.  $p$  è un numero intero
3.  $0 \leq d_i \leq \beta - 1$ ;
4.  $d_1 \neq 0$  e le  $d_i$  non sono tutte uguali a  $\beta - 1$  a partire da un indice in poi.

### 6. Quanti sono i numeri reali definiti in matematica.

La cardinalità dei numeri reali è del continuo; sono infiniti, di un'infinità non numerabile.

### 7. Cosa è la mantissa e l'esponente nella rappresentazione normalizzata di un numero reale

Dato il numero reale rappresentato come  $x = \text{sign}(x) m \beta^p$  indichiamo con il nome di mantissa di  $x$ ,  $m = d_1 \beta^{-1} + d_2 \beta^{-2} + \dots$

La mantissa soddisfa la relazione  $1/\beta \leq m \leq 1$

## 8. Cosa è il sistema floating point, o sistema dei numeri macchina

È anche detto sistema in virgola mobile o insieme dei numeri macchina. È uno schema generale comune per la rappresentazione dei numeri reali in un calcolatore (che presenta delle limitazioni fisiche alla rappresentazione di tutti i numeri reali).

Si definisce insieme dei numeri macchina (floating-point) con  $t$  cifre significative, base  $\beta$  e range  $(L,U)$ , l'insieme dei numeri reali definito nel modo seguente:

$F(\beta, t, L,U) = \{0\} \cup \{x \in \mathbb{R} = \text{sign}(x)\beta^p \sum_{i=1}^t d_i \beta^{-i}\}$  dove  $t$  è il numero di cifre significative, intero positivo,  $\beta$  è una base di interi positivi con  $\beta \geq 2$ . Si ha inoltre  $0 \leq d_i \leq \beta - 1$ ,  $i = 1, 2, \dots$ ,  $d_1 \neq 0$ ,  $L \leq p \leq U$ . Usualmente  $U$  è positivo e  $L$  negativo. In rappresentazione posizionale un numero macchina  $x \neq 0$  viene rappresentato come:  $x = \pm .d_1 d_2 \dots d_t \beta^p$ .

$F$  è un insieme finito e non continuo.

### 9. Quanti sono i numeri di un sistema floating point.

La cardinalità dei numeri macchina è  $\#F = 2(\beta - 1)\beta^{t-1} (U + |L| + 1) + 1$  dove  $\beta$  è la base  $t$  è il numero di cifre significative  $(L,U)$  il range. Sono un numero finito.

### 10. Come si distribuiscono i numeri macchina positivi sulla retta reale.

I numeri macchina positivi si distribuiscono uniformemente tra successive potenze di 2 (in maniera non uniforme sulla retta reale, si addensano sull'origine, e si diradano man mano ci si allontana da essa):  $[2^{-2}, 2^{-1}]$ ,  $[2^{-1}, 2^0]$ ,  $[2^0, 2^1]$ ,  $[2^1, 2^2]$ .

In un sistema floating-point generico, si possono rappresentare i numeri reali positivi compresi in un intervallo, per i numeri reali negativi si ragiona in modo simmetrico,  $[\beta^{-1}\beta^L, (1 - \beta^{-t})\beta^U]$ .

### 11. Come si rappresenta un numero macchina in un calcolatore.



Una possibile rappresentazione interna ad un calcolatore dell'insieme dei numeri macchina è caratterizzato da un primo gruppo di 8 bit determina in base 2 il valore della caratteristica; il bit successivo identifica il segno, 0 per rappresentare un numero positivo, 1 per un numero negativo; il restante gruppo di 23 bit contiene le cifre della mantissa con  $d_1 = 1$  che non viene memorizzato. La lunghezza generale è di 32 bit (4 byte).

### 12. Cosa è l'underflow e overflow.

Quando la caratteristica  $p$  non appartiene a questo intervallo  $L \leq p \leq U$  il numero non può essere rappresentato esattamente, si possono avere due casi distinti così riassunti:

se  $p < L$  si verifica un underflow, e si assume come valore approssimato della  $x$  il numero 0;

se  $p > U$  si verifica un overflow, non si effettua nessun approssimazione ma si arresta il calcolo.

### 13. Come si indica la rappresentazione di un numero reale mediante un numero macchina.

Per rappresentare un numero reale positivo  $x$ , dato nella forma  $x = \text{sign}(x)m\beta^p$ , in un sistema di numeri macchina  $F(\beta, t, L, U)$  si agisce secondo il caso:

1. Il numero  $x$  è tale che  $L \leq p \leq U$  e  $d_i = 0$  per  $i > t$ ; allora  $x$  è un numero macchina ed è rappresentato esattamente;
2. La caratteristica  $p \notin [L, U]$ . Il numero non può essere rappresentato esattamente e si possono presentare i due casi:  $p < L$ , si dice che si è verificato un underflow, e solitamente si assume come valore approssimato il numero 0. In alcuni casi si tratta di un'approssimazione inaccettabile;  $p > U$ , si dice che si è verificato un overflow, solitamente non si effettua nessuna approssimazione di valore, ma si blocca il calcolo.
3. La caratteristica  $p \in [L, U]$ , ma le cifre  $d_i$  per  $i > t$ , non sono tutte nulle. In questo caso si pone il problema di scegliere un suo rappresentante in  $F$ . Tale operazione viene indicata comunemente come operazione di arrotondamento anche se in realtà possono essere utilizzate tecniche diverse. L'operazione di arrotondamento può essere fatta tramite troncamento o arrotondamento:

### 14. Cosa è il troncamento e l'arrotondamento di un numero reale

Sia  $x$  un numero reale positivo dato in forma  $x = \text{sign}(x)m\beta^p$ , diverso dallo zero e tale che  $p \in [L, U]$ . Indichiamo con  $\text{fl}(x)$  il numero macchina, cioè  $\text{fl}(x) \in F$ , che si ottiene in uno dei seguenti modi:

**troncamento:**  $\text{fl}(x) = \text{tronc}(x) = \beta^p \sum_{i=1}^t d_i \beta^{-i}$

**arrotondamento:**  $\text{fl}(x) = \beta^p \text{tronc}(\sum_{i=1}^{t+1} d_i \beta^{-i} + \frac{1}{2} \beta^{-t})$

Si agisce in maniera opposta quando  $x$  è negativo.

Facciamo notare che se  $d_{t+1} < \beta/2$ , allora le due operazioni danno i medesimi risultati, altrimenti, questi differiscono di  $\beta^{p-t}$ .

Durante l'operazione di arrotondamento si può verificare un overflow. È importante dire che quando sostituiamo un  $x$  con  $\text{fl}(x)$  si danno origine ad errori di arrotondamento, ed è importante stabilire una stima, cioè una maggiorazione della quantità  $x - \text{fl}(x)$

### 15. Quale errore si commette nel rappresentare un numero reale mediante un numero macchina.

Indichiamo con  $x$  il numero reale e con  $\text{fl}(x)$  il numero macchina corrispondente. Possiamo individuare, ora, l'errore assoluto dato da  $x - \text{fl}(x)$ , ma risulta più indicativo l'errore relativo dato da  $|(x - \text{fl}(x))/x|$ . Se non si verifica l'overflow si può dimostrare che  $|(x - \text{fl}(x))/x| \leq k\beta^{1-t} = \text{eps}$ , dove  $\text{eps}$  è una quantità detta precisione della macchina nel sistema floating point fissato. Questo numero è molto importante ed ha la seguente caratterizzazione:  $\text{eps}$  è il più piccolo numero macchina positivo tale che  $\text{fl}(1 + \text{eps}) > 1$ . Nella situazione di underflow l'errore relativo è del 100% poiché  $\text{fl}(x) = 0$ , non soddisfa  $|(x - \text{fl}(x))/x| \leq k\beta^{1-t}$ . Nel caso di troncamento, o chopping,  $k = 1$ , invece, nel caso di arrotondamento, o rounding,  $k = \frac{1}{2}$ .

## 16. Cosa è la precisione macchina in un sistema floating point.

La quantità  $\text{eps} = k\beta^{1-t}$  è detta precisione della macchina nel fissato sistema floating point. Rappresenta un maggiorante dell'errore relativo che è commesso nel rappresentare un numero reale  $x$  con un numero macchina  $\text{fl}(x)$  tale che  $|(x - \text{fl}(x))/x| \leq \text{eps}$ , dove nel caso di troncamento, o chopping,  $k = 1$ , invece,  $k = 1/2$  nel caso di arrotondamento, o rounding. La sua importanza numerica è data dalla seguente caratterizzazione:  $\text{eps}$  è il più piccolo numero macchina positivo tale che  $\text{fl}(1 + \text{eps}) > 1$ .

## 17. Che proprietà ha la precisione macchina

La precisione della macchina nel fissato sistema floating point è rappresentato da  $\text{eps}$ , che rappresenta il più piccolo numero macchina positivo che sommato ad 1 da un risultato maggiore di 1 tale che  $\text{fl}(1 + \text{eps}) > 1$ .

$\text{fl}(x) = x(1 + \varepsilon)$  dove  $|\varepsilon| < \text{eps}$ ,  $\varepsilon$  rappresenta la perturbazione  $\varepsilon = (x - \text{fl}(x))/x$

## 18. Scrivi l'algoritmo che calcola la precisione macchina

Supponendo che la base sia 2 possiamo scrivere (pseudo codice):

```
eps = 1;
eps1 = eps + 1;
while (eps1 > 1) {
    eps = 0.5 * eps;
    eps1 = eps + 1;
}
```

se la base non è 2, il valore che si ottiene è approssimato, ma nella maggior parte dei casi è sufficientemente preciso.

## 19. Che relazione esiste tra le operazioni aritmetiche che definiamo in matematica e quelle che esegue il calcolatore.

Siano  $x, y \in \mathbf{F}(\beta, t, L, U)$ , il risultato di un'operazione aritmetica con operandi  $x$  ed  $y$  sia tale che  $p \in [L, U]$ . Si fissa  $p \in [L, U]$  perché si potrebbe verificare una situazione di overflow o di underflow, oppure una situazione in cui il risultato ha un numero di cifre superiore alla precisione  $t$ .

Le operazioni macchina sono:

$x \oplus y = \text{fl}(x + y)$ ;  $x \ominus y = \text{fl}(x - y)$ ;  $x \otimes y = \text{fl}(x * y)$ ;  $x \oslash y = \text{fl}(x / y)$ .

Si potrebbe continuare a definire operazioni elementari come l'estrazione di radice, od il logaritmo, oppure ancora le funzioni trigonometriche. Poiché tra un numero reale  $x$  e la sua rappresentazione floating point esiste la relazione  $\text{fl}(x) = x(1 + \varepsilon)$  dove  $|\varepsilon| < \text{eps}$ ,  $\varepsilon$  rappresenta la perturbazione  $\varepsilon = (x - \text{fl}(x))/x$  le operazioni macchina non coincidono con le stesse definite in matematica. Si può calcolare l'errore relativo dell'operazione utilizzando la maggiorazione  $|(x - \text{fl}(x))/x| \leq k\beta^{1-t}$  ricordando che nel caso di troncamento, o chopping,  $k = 1$ , invece, nel caso di arrotondamento, o rounding,  $k = 1/2$ . Indichiamo con  $\odot$  una delle qualsiasi operazioni macchina, e la corrispondente operazione aritmetica  $\otimes$ :  $x \odot y = (x \otimes y)(1 + \varepsilon)$ , con  $|\varepsilon| < \text{eps}$

## 20. Perché in Analisi numerica ci poniamo il problema dell'affidabilità dell'errore

L'affidabilità dell'errore è importante per capire come l'errore nei dati si propaga sui risultati. Può avvenire che, un piccolo errore nei dati, produca un errore amplificato nei risultati e questo fa sì che la soluzione numerica si discosti enormemente dalla soluzione reale del problema, è utile quindi avere a priori una quantificazione dell'errore per interpretare correttamente la soluzione fornita dell'analisi numerica.

## 21. Quando un problema si dice ben posto.

Un problema si dice ben posto (Hadamard) quando ammette una ed una sola soluzione e questa dipende con continuità dai dati.

## 22. Quando un problema si dice ben condizionato

Un problema si dice ben condizionato quando è ben posto e la sua soluzione non varia di molto al variare dei dati (le perturbazioni sui dati non influenzano eccessivamente i risultati).

## 23. Quando un algoritmo si dice stabile.

Un algoritmo si dice stabile se la successione delle operazioni non amplifica eccessivamente gli errori iniziali di arrotondamento.

Un algoritmo stabile consente di trovare la soluzione di problemi ben condizionati.

## 24. Cosa è la backward analysis o analisi all'indietro degli errori.

L'analisi all'indietro deriva dall'osservazione che un'operazione macchina può essere interpretata come un'operazione esatta su operandi perturbati. In generale nella backward analysis, si cerca di trovare le perturbazioni sui dati che riproducano, operando con aritmetica esatta, i risultati calcolati. In sostanza in questo tipo di analisi gli effetti degli errori di arrotondamento commessi durante il calcolo, vengono idealmente trasferiti sui dati iniziali.

## 25. Che tipo di problemi riusciamo a risolvere in Analisi Numerica.

Problemi ben condizionati per cui si disponga di algoritmi stabili.

## 26. Cosa è la complessità computazionale di un algoritmo.

La complessità computazionale di un algoritmo stabile è la validità dello stesso secondo l'insieme dei criteri che ne definiscono la bontà, e, possono essere riassunti nel modo seguente:

1. L'insieme delle applicazioni, cioè la vastità di applicazione;
2. Semplicità di verifica delle ipotesi di applicazione;
3. Stabilità numerica;
4. Richiesta di risorse, ad esempio, il numero di operazioni e la quantità di memoria richiesta.

Si può aggiungere il grado di parallelismo dell'algoritmo, cioè la quantità di operazioni che possono essere eseguite simultaneamente da quei calcolatori in grado di operare in parallelo

## SISTEMI LINEARI

### 1. Cosa significa risolvere un sistema lineare

Dato un sistema lineare  $Ax=b$ , dove  $b=(b_i) \in \mathbb{R}^m$  (vettore dei termini noti),  $A=(a_{ij}) \in \mathbb{R}^m \times \mathbb{R}^n$  (matrice dei coefficienti) con  $i=1,2,\dots,m$  e  $j=1,2,\dots,n$  ad elementi reali, risolvere il sistema significa trovare tutte le possibili n-uple  $x=(x_1, \dots, x_n) \in \mathbb{R}^n$  stabilito che ne esistano per le quali

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad 1 \leq i \leq m.$$

### 2. Quali sono le condizioni per l'esistenza e l'unicità della soluzione di un sistema

La soluzione di sistema lineare  $Ax=b$  esiste ed è unica se  $A$  è non singolare, ovvero  $\det(A) \neq 0$

### 3. Cosa è l'inversa di una matrice A nxn

Una matrice  $B$  quadrata di ordine  $n$  è una INVERSA di  $A$  se:  $BA=AB=I$  dove  $I$  è la matrice identità, in tal caso si scrive  $B=A^{-1}$ , e la matrice  $A$  si dice invertibile se è non singolare.

### 4. Fai vedere che se si ha un metodo per risolvere un sistema lineare Ax=b, si riesce a trovare l'inversa di A

Se si sa risolvere un sistema lineare, si risolvono gli  $n$  sistemi lineari  $AX_j=e_j$  dove  $e_j$  è la colonna  $j$ -esima della matrice identità ( $I$ ) e  $X_j$  la colonna  $j$ -ma della matrice inversa ( $A^{-1}$ ).

### 5. Quale caratteristica hanno i metodi diretti per la risoluzione di un sistema lineare

I metodi diretti, in assenza di errori di arrotondamento, risolvono un sistema lineare in un numero finito di operazioni. Sfruttano sostanzialmente il metodo di eliminazione di Gauss ricavando da un'equazione una particolare incognita e sostituendola nelle equazioni rimanenti, in questo modo il sistema si riduce di un'equazione e di un'incognita, ripetendo il procedimento si arriva ad ottenere una sola equazione in una sola incognita, ricavando e sostituendo nelle diverse equazioni si ottiene la soluzione.

### 6. Quale caratteristica hanno i metodi iterativi che risolvono un sistema lineare.

Determinano la soluzione  $x^*$  come limite (punto di accumulazione) di una successione  $x^* = \lim_{i \rightarrow \infty} x_i$

La soluzione può essere sufficientemente approssimata dopo un numero piccolo di iterazioni. Se la matrice è sparsa, questa caratteristica non viene persa. Si utilizzano per affinare le soluzioni ottenute con i metodi diretti.

### 7. Cosa è lo scaling di una matrice A nxn

E' una tecnica che consiste nel moltiplicare per una costante le righe o le colonne della matrice.

#### 7b. Scaling per Ax=b

Consiste nella sostituzione del sistema con uno equivalente. Siano  $D_r$  e  $D_c$  delle matrici diagonali non singolari

scaling per righe:  $Ax=b \rightarrow D_r Ax = D_r b$ ,

scaling per colonne:  $Ax=b \rightarrow AD_c(D_c^{-1}x)=b \rightarrow (AD_c)y=b$  con  $y$  nuova incognita

### 8. Scrivi la definizione di sistema triangolare superiore e inferiore

Una matrice  $U$  si dice triangolare superiore se tutti i suoi elementi al di sotto della diagonale principale sono nulli.

Una matrice  $L$  si dice triangolare inferiore se tutti i suoi elementi al di sopra della diagonale principale sono nulli.

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \quad L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$$

### 9. Scrivi l'algoritmo che risolve il sistema triangolare inferiore Lx=b con L matrice triangolare inferiore

$y_1 = b_1/l_{11}$ ; for( $i=2, \dots, n$ ) {  $y_i = b_i$ ; for( $j=1, \dots, i-1$ ) {  $y_i = y_i - l_{ij}y_j$ ; }  $y_i = y_i/l_{ii}$ ; }

### 10. Scrivi l'algoritmo che risolve il sistema triangolare superiore Ux=b con U matrice triangolare superiore

$x_n = y_n/u_{nn}$ ; for( $i=n-1, \dots, 1$ ) {  $x_i = y_i$ ; for( $j=i+1, \dots, n$ ) {  $x_i = x_i - u_{ij}x_j$ ; }  $x_i = x_i/u_{ii}$ ; }

### 11. Quale è la complessità computazionale dell'algoritmo che risolve un sistema triangolare inferiore

Considerando le moltiplicazioni e le divisioni:

operazioni =  $1 + \sum_{i=2, n} i = \sum_{i=1, n} i = \frac{1}{2} n(n+1) \approx \frac{1}{2} n^2$  (complessità polinomiale)

### 12. Calcola esattamente il numero di moltiplicazioni eseguite dall'algoritmo che risolve un sistema triangolare inferiore

operazioni di moltiplicazione =  $\sum_{i=2, n} (i-1) = \sum_{i=1, n-1} i = \frac{1}{2} n(n-1)$

### 12. Calcola esattamente il numero di divisioni eseguite dall'algoritmo che risolve un sistema triangolare inferiore

operazioni di divisione =  $1 + n - 1 = n$

### 13. Quando si può applicare l'algoritmo di Gauss senza scambio di righe.

Si può applicare ad un sistema  $Ax=b$  se il sistema è risolvibile (matrice  $A$  dei coefficienti è non singolare, ovvero  $\det(A) \neq 0$ ) e se tutti i minori principali di  $A$  sono diversi da 0 (da cui il nome Gauss 'naive').

### 14. Scrivi l'algoritmo di Gauss senza scambio di righe

for( $k=1, 2, \dots, n-1$ ) { for( $i=k+1, \dots, n$ ) {  $m_{ik} = a_{ik}^{(k)}/a_{kk}^{(k)}$ ; for( $j=k+1, \dots, n+1$ ) {  $a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)}$ ; } } }

### 15. Sotto quale condizione si può applicare il metodo di Gauss con scambio di righe

Si può applicare ad un sistema  $Ax=b$  se il sistema è risolvibile (matrice  $A$  dei coefficienti è non singolare, ovvero  $\det(A) \neq 0$ ).

### 16. Quale è la complessità computazionale dell'algoritmo di Gauss

La complessità computazionale del metodo Gauss 'naive' è polinomiale dello stesso ordine di  $n^3$ . [operazioni =  $(1/3)n^3 + n^2 - (1/2)n \approx n^3/3$ ]

### 17. In che cosa consiste la strategia del pivot

Dato un sistema lineare  $Ax=b$ , consiste nello scambiare le righe (pivoting parziale) e le colonne (pivoting totale) in modo tale che il coefficiente (di  $x_1$  nel caso del parziale) col valore assoluto più grande (pivot) risulti nella prima riga e nella prima colonna della sottomatrice in esame.

### 18. Che vantaggi produce la strategia del pivot

Rende sempre applicabile l'algoritmo di Gauss (se  $\det(A) \neq 0$ ) e ne migliora l'accuratezza.

### 19. Cosa sono il pivoting parziale ed il pivoting totale

Vedi [17]. Nel pivoting totale si scambiano solo righe e colonne, mentre in quello parziale (o scalato) si scambiano solo le righe.

## 20. Scrivi l'algoritmo di Gauss con scambio di righe e pivoting parziale

for( $i=1, \dots, n-1$ ) { p intero tale che  $|a_{pk}^{(k)}| \geq |a_{ik}^{(k)}| \forall i=1, \dots, n+1$ ; scambio  $a_{kj}^{(k)}$  e  $a_{pj}^{(k)}$   $j=1, \dots, n+1$ ; for( $k+1, \dots, n$ ) {  $m_{ik} = a_{ik}^{(k)} / a_{kk}^{(k)}$ ;  
for( $j=k+1, \dots, n+1$ ) {  $a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)}$  } } }

## 21. Cosa è la decomposizione LU di una matrice A nxn

La decomposizione o fattorizzazione LU consiste nello scomporre una matrice A nel prodotto di due matrici, L triangolare inferiore ed U triangolare superiore con la particolarità che la matrice L ha sulla diagonale tutti 1; se esiste tale fattorizzazione è unica. Gli elementi di L corrispondono ai moltiplicatori utilizzati nel metodo di Gauss.

## 22. In quale modo il metodo di Gauss senza scambio di righe, permette di ottenere la decomposizione LU di una matrice.

$$U = M_{n-1} M_{n-2} \dots M_2 M_1 A$$

$$A = (M_1 M_2 \dots M_{n-1})^{-1} U = (M_1^{-1} M_2^{-1} \dots M_{n-1}^{-1}) U = LU$$

con:  $M_k = I - m_k e_k^T$ ;  $m_k = [0, \dots, m_{k+1,k}, m_{k+2,k}, \dots, m_{n,k}]^T$ ;  $e_k^T = [0, \dots, 1^{(k)}, \dots, 0]$

La scomposizione  $A=LU$  è ottenuta considerando la matrice  $M_k = I - m_k e_k^T$  triangolare inferiore. Moltiplicare A a sinistra per le  $M_k$  produce lo stesso risultato del metodo di Gauss.  $M_k = L^{-1}$ ; L è ancora una triangolare inferiore in quanto inversa di una inferiore, quindi ottengo  $L^{-1}A = U \rightarrow A = LU$

## 23. Nel caso si utilizzi Gauss con scambio di righe, quale fattorizzazione di A si ottiene.

$$PA = LU$$

Per lo scambio di righe si usa una matrice di permutazione P elementare ottenuta scambiando le righe della matrice identità I.

$$U = M_{n-1} P_{n-1} M_{n-2} P_{n-2} \dots M_2 P_2 M_1 P_1 A \text{ triangolare superiore}$$

$$A = (M_{n-1} P_{n-1} M_{n-2} P_{n-2} \dots M_2 P_2 M_1 P_1)^{-1} U$$

$$P = P_{n-1} P_{n-2} \dots P_2 P_1$$

$$L = P(M_{n-1} P_{n-1} M_{n-2} P_{n-2} \dots M_2 P_2 M_1 P_1)^{-1} \text{ triangolare inferiore con 1 sulla diagonale principale}$$

## 24. Che cosa è una matrice di permutazione

E' una matrice P ottenuta dalla matrice identità attraverso scambi di righe (o di colonna). PA produce uno scambio di righe di A, mentre AP produce uno scambio di colonne.

## 25. Mostrare che la matrice non singolare $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ non ha una decomposizione $A=LU$

la fattorizzazione  $A=LU$  tramite Gauss non è applicabile perchè i minori principali di A non sono diversi da zero.

## 26. Cosa significa che una matrice A nxn ha decomposizione $A=LDM^T$ .

Se i minori principali di una matrice A di ordine n sono tutti diversi dallo zero, allora esistono due matrici triangolari inferiori L e M con elementi diagonali uguali a 1 e D è una matrice diagonale tale che  $A=LDM^T$

## 27. Come ottieni con Gauss una decomposizione $A=LDM^T$ .

Supponiamo di avere A scritta secondo la fattorizzazione di Gauss, come  $A=LU$  dove L è triangolare inferiore con 1 sulla diagonale e U triangolare superiore, basta prendere come matrice diagonale D quella i cui elementi coincidono con la diagonale di U e risulta essere non singolare e invertibile, scegliendo  $M^T = D^{-1}U$ , che è triang sup. con diagonale formata tutta da uno si ha  $A=LU=L(DD^{-1})U=LD(D^{-1}U)=LDM^T$

## 28. Se di una matrice A si ha la decomposizione $A=LU$ come risolvi il sistema lineare $Ax=b$

$Ax=b$  equivale a risolvere  $LUx=b$ , se indico  $Ux=y$  ottengo due sistemi triangolari per i quali esiste un algoritmo risolutivo. Risolvo  $Ly=b$  (triangolare inferiore) e trovo y che sostituisco in  $Ux=y$  e la cui soluzione è equivalente a quella di  $Ax=b$ .

## 29. Scrivi l'algoritmo di Doolittle per trovare la fattorizzazione $A=LU$

for( $k=1, \dots, n$ ) { for( $i=k, \dots, n$ ) {  $u_{ki} = a_{ki} - \sum_{p=1, k-1}^i l_{kp} u_{pi}$  } } for( $i=k+1, \dots, n$ ) {  $l_{ik} = (a_{ik} - \sum_{p=1, k-1}^i l_{ip} u_{pk}) / u_{kk}$  } }

## 30. Nel caso di A, matrice simmetrica come si riscrive la fattorizzazione $A=LDM^T$

Nel caso particolare in cui A sia simmetrica e  $A=LDM^T$  si ha  $L=M$  cioè  $A=MDM^T$

## 31. Quando una matrice A si dice definita positiva

Una matrice A è definita positiva se  $x^T A x > 0 \forall x \neq 0, x \in \mathbb{R}^n$ . Se A è definita positiva tutti gli autovalori e i minori principali sono  $>=0$

## 32. Scrivi l'algoritmo di Cholesky per la decomposizione di una matrice $A=RR^T$

Valido se la matrice A è simmetrica e definita positiva.

for( $k=1, \dots, n$ ) {  $a_{kk} = (a_{kk} - \sum_{p=1, k-1}^k a_{kp}^2)^{1/2}$ ; for( $i=k+1, \dots, n$ ) {  $a_{ik} = (a_{ik} - \sum_{p=1, k-1}^i a_{ip} a_{kp}) / a_{kk}$  } } }

La complessità computazionale è  $n^3/6$ .

## 33. Dovendo risolvere un sistema lineare $Ax=b$ con A simmetrica definita positiva come procedi

Tramite l'algoritmo di Cholesky fattorizzo A come  $A=RR^T$  con R matrice triangolare inferiore. Riscrivo  $Ax=b$  come  $RR^T x = b$  ed indico con  $y=R^T x$ . Risolvo quindi  $Ry=b$  (sistema triangolare inferiore) e trovo y che sostituisco in  $R^T x = y$  (sistema triangolare superiore) che risolvo ottenendo x.

## 34. Supponi di risolvere il sistema lineare $Ax=b$ con A simmetrica definita positiva utilizzando l'algoritmo di Cholesky e di ottenere $x^*$ . Di quale sistema lineare puoi considerare $x^*$ soluzione esatta. Che proprietà ha quest'ultimo sistema.

$x^*$  si può interpretare come soluzione esatta del sistema perturbato  $(A+E)x^*=b$ . E è la matrice di perturbazione le cui entrate hanno un ordine di grandezza molto piccolo rispetto ad A, infatti vale  $\|E\|_2 \leq c_n \epsilon \|A\|_2$  dove  $c_n$  è una costante molto piccola dipendente da n (n = ordine della matrice A). La soluzione risulta quindi accettabile.

## 35. Cosa è una matrice a banda.

E' una matrice in cui compaiono zeri al di sopra e/o al di sotto di diagonali secondarie. Una matrice A si dice:

- a banda superiore q se  $a_{ij}=0$  per  $j>i+q$
- a banda inferiore p se  $a_{ij}=0$  per  $i>j+q$

**36. Dai una matrice 6x6 a banda superiore 2 ed inferiore 3**

$$\begin{matrix}
b_{11} & b_{12} & b_{13} & 0 & 0 & 0 \\
b_{21} & b_{22} & b_{23} & b_{24} & 0 & 0 \\
b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & 0 \\
b_{41} & b_{42} & b_{43} & b_{44} & b_{45} & b_{46} \\
0 & b_{52} & b_{53} & b_{54} & b_{55} & b_{56} \\
0 & 0 & b_{63} & b_{64} & b_{65} & b_{66}
\end{matrix}$$

**37. Supponi che per matrice A di ordine n valga A=LU e che A sia a banda superiore q e inferiore p. Quali proprietà hanno la L e la U.**

La matrice U ha banda superiore q e la matrice L ha banda inferiore p

**38. Scrivi l'algoritmo che risolve il sistema triangolare inferiore Lx=b con L matrice triangolare inferiore e a banda inferiore 4. Supponi n>4.**

Uso l'algoritmo di Gauss per matrici a banda p e q:

p=4; // banda inferiore

q=0; // banda superiore (matrice triangolare)

for(k=1,...,n-1) { for(i=k+1,...,min(k+p,n)) { a<sub>ik</sub>=a<sub>ik</sub>/a<sub>kk</sub> } for(i=k+1,...,min(k+p,n)) { for(j=k+1,...,min(k+q,n)) { a<sub>ij</sub>=a<sub>ij</sub>- a<sub>ik</sub>a<sub>kj</sub> } } }

**39. Quando una matrice si dice sparsa**

Quando presenta molti elementi uguali a zero. In generale parliamo di matrice sparsa quando gli elementi diversi da zero sono dell'ordine O(n), dove n è l'ordine della matrice

**40. Cosa è il numero di condizionamento di una matrice**

Se ||.|| indica una norma naturale di matrice il Condizionamento di una matrice A non singolare associato a tale norma e relativo alla risoluzione di sistemi lineari è dato dal seguente numero  $\mu(A)=\|A\| \|A^{-1}\|$

Si dice che una matrice è ben condizionata se il suo numero di condizionamento non è grande.

**41. Perché è importante conoscere il numero di condizionamento di una matrice**

Dato un sistema lineare del tipo Ax=b, il numero di condizionamento fornisce informazioni sull'entità della variazione della soluzione esatta al variare dei termini noti. Il numero di condizionamento ci dice quindi se il problema è ben condizionato. Se in numero di condizionamento  $\mu(A)$  è grande devo aspettarmi che piccole variazioni dei valori dei termini noti producano grandi variazioni della soluzione esatta.

**42. Quanto vale il numero di condizionamento della matrice identità**

Vale 1, perché con qualsiasi norma naturale  $\|I\|=1$  e  $I^{-1}=I$

**43. La matrice A abbia numero di condizionamento  $\mu=6$ . Quale è il numero di condizionamento della matrice  $5^*A$ . Giustifica la risposta.**

$\mu(5^*A)=\|5A\| \|(5A)^{-1}\| = |5| \|A\| |1/5| \|A^{-1}\| = \|A\| \|A^{-1}\| = \mu(A) = 6$  e infatti in generale  $\mu(\alpha A) = \mu(A)$  per ogni matrice A e ogni scalare  $\alpha \neq 0$

**44. Dai le proprietà che conosci del numero di condizionamento di una matrice**

- 1)  $\mu(\alpha A) = \mu(A)$  per ogni matrice A e ogni scalare  $\alpha \neq 0$
- 2)  $\mu(A) \geq 1$  se la norma è naturale
- 3)  $\mu_2(A) = \sigma_{\max} / \sigma_{\min}$ ; dove  $\sigma_{\max} = \sqrt{\max_{1 \leq i \leq n} \lambda_i}$  con  $\lambda_i$  autovalori di A e n ordine di A
- 4)  $\mu_2(A) = 1$  se  $A = \alpha Q$  dove  $\alpha$  è uno scalare e Q è la matrice unitaria

**45. Che relazione esiste tra il numero di condizionamento di una matrice ed il suo determinante**

Non esiste alcuna relazione. Il determinante di una matrice non da buona indicazione del malcondizionamento di una matrice.

**46. Supponi che per una matrice A il metodo di Gauss senza scambio di righe dia le matrici  $\bar{L}$  e  $\bar{U}$  come fattorizzazione di A. Cosa si può dire sull'errore commesso**

Utilizzando l'analisi all'indietro dell'errore:  $A-E = LU$ .

Supponendo di aver applicato Gauss in un calcolatore con precisione aritmetica eps, allora  $\|E\|_{\infty} \leq n^2 g_n eps \|A\|_{\infty}$   $g_n$  varia a seconda dei casi che si usi: pivoting parziale, pivoting totale, A simmetrica definita positiva.

**47. Quale è la matrice meglio condizionata**

La matrice unitaria. In tal caso  $\mu_2(A) = 1$

**48. Come operano i metodi iterativi per la risoluzione di un sistema lineare**

Determinano la soluzione  $x^*$  come limite (punto di accumulazione) di una successione  $x^* = \lim_{i \rightarrow \infty} x_i$ ;  $Ax^* = b$

La soluzione può essere sufficientemente approssimata dopo un numero piccolo di iterazioni. Se la matrice è sparsa, questa caratteristica non viene persa. Si utilizzano per affinare le soluzioni ottenute con i metodi diretti.

**49a. Schema iterativo metodi Jacobi e Gauss-Seidel**

Posto  $Ax=b$ , A matrice nxn,  $\det(A) \neq 0$ ,  $b \in R^n$ ,

A si scompone nelle matrici M ed N:  $A = M - N$ , con  $\det(M) \neq 0$

$Ax=b \Rightarrow Mx = Nx + b$ ;  $Mx^{(k+1)} = Nx^{(k)} + b$

schema delle interazioni:  $x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b$

matrice di iterazione:  $B = M^{-1}N = M^{-1}(M - A) = I - M^{-1}A$

**49. Quale è lo schema iterativo del metodo di Jacobi per la risoluzione di un sistema lineare**

Si pone:  $A=D-E-F$ ;  $M=D$ ;  $N=E+F$ ;  $B_j=(E+F)=I-D^{-1}A$

iterazione da k a k+1:  $\{x_i^{(k+1)} = (b_i - \sum_{j=1, i-1} a_{ij}x_j^{(k+1)} - \sum_{j=1+1, n} a_{ij}x_j^{(k)}) / a_{ii}\}$

**50. Quale è lo schema iterativo del metodo di Gauss-Seidel per la risoluzione di un sistema lineare.**

Si pone:  $A=D-E-F$ ;  $M=D-E$ ;  $N=E$ ;  $B_G=(D-E)^{-1} = (I-D^{-1}E)^{-1}F$

iterazione da k a k+1:  $\{x_i^{(k+1)} = (b_i - \sum_{j=1, i-1} a_{ij}x_j^{(k+1)} - \sum_{j=1+1, n} a_{ij}x_j^{(k)}) / a_{ii}\}$

**51. Quando un metodo iterativo per la risoluzione di un sistema lineare basato sullo schema  $Mx^{(k+1)} = Nx^{(k)} + b$  è convergente.**

Il metodo iterativo definito dalla matrice di iterazione B, converge per ogni punto iniziale  $x_0 \in \mathbb{R}^n$  se e solo se il raggio spettrale della matrice di iterazione  $B = M^{-1}N$  è minore di 1:  $\rho(B) < 1$ .

Infatti:  $x^* = Bx^* + M^{-1}b$  ;  $x^{(k+1)} = Bx^{(k)} + M^{-1}b$

indicando l'errore con  $e^{(k+1)} = x^* - x^{(k)}$  ;  $e^{(k+1)} = Be^{(k)} \Rightarrow e^{(k+1)} = B^k e^{(0)}$

se B è una matrice convergente allora  $\|e^{(k)}\| \rightarrow 0$  per  $k \rightarrow \infty$

**APPROSSIMAZIONE DI FUNZIONE**

**1. In che cosa consiste l'approssimazione di una funzione.**

Consiste nella sostituzione di una funzione con una che abbia una rappresentazione analitica più semplice, scelta nell'ambito di una classe fissata di funzioni a dimensione finita

**2. Dai la formulazione del problema dell'approssimazione mediante interpolazione**

Dati i punti  $(x_i, y_i)$   $i=0,1,2,\dots,n$  con  $x_i \in \mathbb{R}^m$   $m \geq 1$   $x_i \neq x_k$  con  $i \neq k$  e una famiglia di funzioni:  $\Phi(x; a_0, a_1, \dots, a_n)$  [modello matematico], si cercano i valori  $a_0, a_1, \dots, a_n$  [parametri o gradi di libertà] tali che  $\Phi(x_i; a_0, a_1, \dots, a_n) = y_i$   $i=0,1,2,\dots,n$

**3. Quando si parla di interpolazione lineare.**

Nell'ambito dell'approssimazione di una funzione dati i punti  $(x_i, y_i)$   $i=0,1,2,\dots,n$  con  $x_i \in \mathbb{R}^m$   $m \geq 1$   $x_i \neq x_k$  con  $i \neq k$  e una famiglia di funzioni:  $\Phi(x; a_0, a_1, \dots, a_n)$  si parla di interpolazione lineare se la  $\Phi$  dipende linearmente dai parametri, cioè  $\Phi(x_i; a_0, a_1, \dots, a_n) = a_0 \Phi_0(x) + a_1 \Phi_1(x) + \dots + a_n \Phi_n(x)$

**4. Dai due esempi di problemi di interpolazione.**

a) Data una tabulazione  $(s_i, t_i)$  relativa alla temperatura di una stanza  $(t_i)$  ed al tempo di rilevazione  $(s_i)$ , si vuole trovare una funzione analitica che descriva il variare della temperatura al variare del tempo.

b) Data una funzione  $f(x)$ , tramite una sua tabulazione si vuole conoscere una nuova funzione  $p(x)$  con una rappresentazione analitica semplice.

**5. In che cosa consiste l'approssimazione mediante polinomi.**

dati  $n+1$  punti (nodi)  $(x_i, y_i)$  si approssima la funzione con un polinomio di grado  $n$ :  $\Phi(x; a_0, a_1, \dots, a_n) = p(x) = a_0 + a_1 x + \dots + a_n x^n$  ; con  $p(x) \in P_n$  dove  $P_n$  è l'insieme dei polinomi di grado  $\leq n$  tali che  $p(x_i) = y_i$  con  $i=0,1,2,\dots,n$ .

**6. Dimostra che il problema dell'interpolazione polinomiale ammette soluzione e questa è unica.**

E' possibile esprimere il polinomio interpolante  $p(x) = a_0 + a_1 x + \dots + a_n x^n$  come un sistema lineare  $p(x_i) = y_i$  di  $(n+1)$  equazioni in  $(n+1)$  incognite in cui le incognite sono i parametri  $a_0, a_1, \dots, a_n$  ed il cui determinante è il determinante di Vandermonde che è  $\neq 0$  per  $x_i \neq x_k$  con  $i \neq k$ . Per cui la soluzione esiste ed è unica.

**7. Scrivi l'espressione dei polinomi interpolanti mediante la rappresentazione di Lagrange.**

$p(x) = \sum_{i=0}^n y_i L_i(x)$  dove  $L_i(x) = \prod_{j=0, j \neq i}^n ((x-x_j)/(x_i-x_j))$  con  $i=0,1,2,\dots,n$

$L_i(x_k) = \delta_{ik} = [1 \text{ se } k=i, 0 \text{ se } k \neq i]$  con  $k=0,1,2,\dots,n$

**8. In che cosa consiste l'interpolazione di Hermite.**

Dati  $k+1$  punti distinti  $x_0, x_1, \dots, x_k$  e  $k+1$  interi positivi  $\alpha_0, \alpha_1, \dots, \alpha_k$  sia  $n = k + \alpha_0 + \alpha_1 + \dots + \alpha_k$ . Per ogni insieme di numeri  $y_i^{(l)}$ ,  $i=1,2,\dots,k$  e  $l=0,1,2,\dots,\alpha_i$ ; esiste ed è unico il polinomio  $p(x) \in P_n(x)$  tale che:  $d^l p(x_i)/dx^l = y_i^{(l)}$

Ovvero si cerca il polinomio  $p(x)$  che nei punti  $x_i$  ha gli stessi valori della funzione e le stesse derivate.

**9. Dai la forma generale dell'interpolazione lineare**

Sia V uno spazio vettoriale su R. Indichiamo con  $V^*$  l'insieme dei funzionali lineari su V (operatori lineari:  $V \rightarrow R$ ). Assegnati n elementi in  $V^*$ ,  $L_1, L_2, \dots, L_n$ , e dati i numeri reali  $y_i$ ,  $i=1,2,\dots,n$  si cerca l'elemento  $v \in V$  tale che  $L_i v = y_i$   $i=1,2,\dots,n$

**10. Quali sono le condizioni sotto le quali il problema generale dell'interpolazione ammette soluzione e questa è unica.**

vedi [9]. Il problema ha una ed una sola soluzione se i funzionali lineari  $L_1, L_2, \dots, L_n$  sono linearmente indipendenti in  $V^*$ .

**11. Scrivi l'algoritmo di Neville per la valutazione di un polinomio interpolante in un punto  $x_s$ .**

$t(0) = y(0)$ ; for  $(i=1, \dots, n)$  {  $t(i) = y(i)$ ; for  $(j=i-1, -1, 0)$  {  
 $t(j) = (t(j+1) + (t(j+1) - t(j)) * (x_s - x(i)) / (x(i) - x(j)))$  } }  $fx_s = t(0)$ ;

**12. Su quale relazione si basa l'algoritmo di Neville.**

In generale vale la relazione

$P_{i(i+1)\dots k}(x) = ((x-x_i)P_{i+1\dots k}(x) - (x-x_k)P_{i+1\dots(i,k-1)}(x)) / (x_k - x_i)$

posto  $T_{i+k,k} = P_{i+1\dots i+k}$  si ha  $T_{ik} = T_{i,k-1} + (T_{i,k-1} - T_{i-1,k-1})[(x-x_i)/(x_i-x_{i-k})]$

con  $T_{i0} = y_i$  e  $1 \leq k \leq i$ ,  $i=0,1,2,\dots,n$

**13. Cosa stabilisce il teorema di Weierstrass sull'approssimazione di una funzione mediante un polinomio.**

Sia  $f(x)$  una funzione continua su un intervallo limitato e chiuso  $[a,b]$ . Allora per ogni  $\epsilon > 0$  esiste un intero  $n = n(\epsilon)$  e un polinomio  $P_n(x)$  di grado al più  $n$  tale che:  $\|f(x) - P_n(x)\|_\infty < \epsilon$  con  $\|\cdot\|_\infty$  la norma del massimo per funzioni.

Questo teorema dimostra che l'insieme delle funzioni continue  $C[a,b]$  è separabile. Infatti l'insieme dei polinomi  $1, x, x^2, x^3, \dots$  è chiuso (o completo) in  $C[a,b]$ .

**14. Cosa è una matrice di interpolazione.**

Definiamo matrice di interpolazione sull'intervallo  $[a,b]$  della retta reale una matrice triangolare del tipo:

$$\begin{bmatrix} x_0^1 & & & \\ x_0^2 & x_1^2 & & \\ x_0^3 & x_1^3 & x_2^3 & \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

con  $x_i^j \in [a,b]$  e gli elementi in ciascuna riga distinti. Gli elementi  $x_i^j$  di ciascuna riga associati ai valori  $y_i^j$  permettono la costruzione dei polinomi interpolanti.

### 15. Cosa dice il teorema di Faber sui polinomi interpolante.

Per ogni matrice di interpolazione di un intervallo limitato  $[a,b]$  della retta reale esiste una funzione continua  $f(x)$  tale che  $P_n(f)(x)$  non converge uniformemente a  $f(x)$  per  $n \rightarrow \infty$ , cioè  $\lim_{n \rightarrow \infty} \|f(x) - P_n(f)(x)\|_\infty$  non tende a zero.

### 16. Se $f(x)$ è una funzione continua in un intervallo $[a,b]$ puoi scegliere la matrice di interpolazione in modo che la successione dei polinomi converga uniformemente a $f(x)$

Si esiste un teorema che lo dimostra.

### 17. Dai l'espressione dell'errore dei polinomi interpolanti.

Sono dati  $n+1$  punti distinti  $x_0, x_1, \dots, x_n$  dell'intervallo  $[a,b]$  chiuso e limitato in  $\mathbb{R}$ . Sia  $f(x)$  una funzione assegnata nello spazio  $C^{n+1}([a,b])$  e  $p(x)$  il polinomio di grado  $n$  tale che:  $p(x_i) = f(x_i)$ ,  $i=0, 1, 2, \dots, n$

Allora per ogni  $x \in [a,b]$  esiste un valore  $\xi_x$ , con  $\xi_x \in (x, x_0, x_1, \dots, x_n)$ , tale che

$$E(x) = f(x) - p(x) = \frac{1}{(n+1)!} (x-x_0)(x-x_1)\dots(x-x_n) f^{(n+1)}(\xi_x)$$

dove  $E(x)$  è l'errore commesso nel sostituire al valore della funzione il valore del polinomio.

### 18. In quali condizioni si può dare l'espressione dell'errore dei polinomi interpolanti

Nel caso in cui  $p(x)$  sia il polinomio di interpolazione di Hermite, cioè nel caso in cui siano assegnate le condizioni anche sulle derivate, formalmente devono essere dati  $k+1$  punti distinti  $x_0, \dots, x_k$  di  $[a,b] \in \mathbb{R}$  a cui sono assegnati rispettivamente  $\alpha_0, \dots, \alpha_n$  interi se  $n = k + \alpha_0 + \dots + \alpha_k$  se  $f$  è una funzione derivabile con derivata continua fino all'ordine  $n+1$  e  $p(x)$  è il suo polinomio di interpolazione di Hermite dove  $p(x_i) = f(x_i)$  allora per ogni  $x \in [a,b]$  esiste un valore  $\xi_x$  nell'insieme  $I(x, x_0, x_1, \dots, x_k)$ , definito come il più piccolo intervallo chiuso contenente i punti  $x, x_0, x_1, \dots, x_k$  tale che  $E(x) = f(x) - p(x) = \frac{1}{(n+1)!} \pi_n(x) f^{(n+1)}(\xi_x)$  dove  $\pi_n(x) = \prod_{i=0}^k (x-x_i)^{\alpha_i+1}$

### 19. Supponi di dover necessariamente approssimare una funzione mediante polinomi interpolanti di grado crescente. Potendo scegliere la matrice di interpolazione, come la sceglieresti per avere la convergenza uniforme.

Gli elementi della matrice di interpolazione devono essere alla  $i$ -esima riga gli zeri del polinomio di Chebichev su  $[a,b]$ :

$$T_0 = 1, T_1 = x, T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

si considerano nell'intervallo  $[-1,1]$ , gli zeri di  $T_{n+1}(x)$  sono dati da:

$$z_{n+1,i} = \cos((2m-1)\pi / 2(n+1)), \text{ per } i=1, \dots, n+1$$

### 20. Come si definisce la funzione di Runge e che caratteristiche ha.

E' la funzione  $f(x) = 1/(1+x^2)$  con  $x \in [-a,a]$  con  $a > 0$   $f(x) \in C([-a,a])$ . Al crescere di  $n$  il polinomio di interpolazione oscilla vicino agli estremi discostandosi molto dalla funzione. Per  $n \rightarrow \infty$  non si ha convergenza. La scelta delle ascisse di Chebichev risolve questo problema. Le ascisse di Chebyshev sono definite nell'intervallo  $[-1,1]$ , quindi per riportare le ascisse ad un generico intervallo  $[a,b]$  si usa la formula:  $X = ((a+b)/2) + ((b-a)/2)x$

### 21. In che cosa consiste l'approssimazione mediante funzioni spline.

L'intervallo  $[a,b]$  di definizione della funzione che si vuole approssimare è suddiviso in sottointervalli  $[x_i, x_{i+1}]$  ed in ciascuno di questi si approssima la funzione con un polinomio di grado basso (1, 2 o 3). La funzione risultante composta dai polinomi di cui sopra, è chiamata funzione spline di grado 1, 2 o 3.

### 22. Dai la definizione di funzione spline.

Sia  $\Delta = \{a = x_0 < x_1 < \dots < x_n = b\}$  una suddivisione dell'intervallo  $[a,b]$ . Una funzione spline di grado  $p$  con nodi nei punti  $x_i$   $i=0, 1, 2, \dots, n$  è una funzione su  $[a,b]$  indicata con  $s_p(x)$  tale che

1. su ogni intervallo  $[x_i, x_{i+1}]$ ,  $i=0, 1, 2, \dots, n-1$ , è un polinomio di grado  $p$ .
2. la sua funzione  $s_p(x)$  e le sue prime  $p-1$  derivate sono continue su  $[a,b]$ .

### 23. Cosa sono le funzioni spline interpolanti.

Data la tabulazione di una funzione  $f(x)$ ,  $(x_i, y_i)$ ,  $i=0, 1, \dots, n$ , si chiama spline interpolante la spline  $s_p(x)$  tale che:  $s_p(x_i) = y_i$

### 24. Cosa sono le splines continue e lineari a tratti.

Sono le splines  $s_1(x)$  di grado 1. Sono dei polinomi di grado 1 di classe  $C^0$ , cioè delle rette. I gradi di libertà sono  $n+1$ . La soluzione esiste ed è unica.

### 25. Dai la formulazione del problema di migliore approssimazione.

Dato uno spazio lineare normato  $V$  e un sottospazio  $W$  di dimensione  $n+1$  e generato dagli elementi  $\Phi_0, \Phi_1, \dots, \Phi_n$  di  $V$  (costituiscono una base), ovvero ogni elemento  $w$  di  $W$  è dato da  $w = \sum_{j=0}^n a_j \Phi_j$ .

Assegnato un elemento qualsiasi  $f \in V$  occorre determinare un elemento  $w^*$  tale che:  $\|f - w^*\| < \|f - w\|$  per ogni  $w^* \in W$

### 26. Dai un esempio di problema di migliore approssimazione.

- $V = C[a,b]$ , l'insieme delle funzioni continue in  $[a,b]$
- $W = P_n(x)$ , l'insieme dei polinomi di grado  $\leq n$  avente per base  $1, x, x^2, \dots, x^n$
- la norma utilizzata è quella del massimo  $\|\cdot\|_\infty$

Data una qualsiasi funzione  $f(x)$  in  $C[a,b]$ , si cerca in  $P_n(x)$  il polinomio  $p(x)$  più vicino a  $f(x)$  secondo la distanza indotta dalla norma del massimo.

### 27. In che cosa consiste la migliore approssimazione nel senso dei minimi quadrati nel continuo

E' un caso particolare di miglior approssimazione

- $V = L^2[a,b]$ , insieme delle funzioni  $f(x)$  in  $[a,b]$  per le quali l'integrale di  $(f(x))^2$  su  $[a,b]$  esiste ed è finito
- $W$  sottospazio di  $V$  di dimensione finita. Per esempio  $W = P_n(x)$ , l'insieme dei polinomi di grado  $\leq n$
- la norma utilizzata è la norma 2, ovvero per  $f \in V$

$$\|f(x)\|_2 = \left( \int_{a,b} (f(x))^2 dx \right)^{1/2}$$

### 28. In che cosa consiste la migliore approssimazione nel senso dei minimi quadrati nel discreto.

E' un caso particolare di miglior approssimazione

- $V =$  insieme delle funzioni  $f(x)$  per cui esiste una tabulazione  $\{f(x_1), f(x_2), \dots, f(x_m)\}$  su  $S = \{x_1, x_2, \dots, x_m\}$  nell'intervallo  $[a,b]$
- $W$  sottospazio di  $V$  di dimensione finita. Per esempio  $W = P_n(x)$ , l'insieme dei polinomi di grado  $\leq n$
- la norma utilizzata è la seminorma  $\|\cdot\|_{2,s}$ , ovvero per  $f \in V$

$$\|f(x)\|_{2,s} = \left( \sum_{i=1}^m (f(x_i))^2 \right)^{1/2}$$