

## Avviso

I contenuti di queste slide non sono esaustivi ai fini del buon esito dell'esame. Fare riferimento anche alle lezioni tenute in aula ed ai testi consigliati:

- G. Monegato, Fondamenti di Calcolo Numerico. Ed. CLUT
- A. Quarteroni, F. Saleri, 'Calcolo Scientifico, esercizi e problemi risolti con matlab e octave' Springer.

## Condizionamento di Lagrange

Si studia il condizionamento (anche la stabilità in questo caso) del problema di interpolazione di Lagrange. Data una funzione  $f(x)$  e una sequenza di nodi  $x_j$  si trova un polinomio interpolante  $P_n(x)$ . Con  $\tilde{f}_j$  e  $\tilde{P}_n$  indichiamo le versioni perturbate dei dati e del polinomio. Si ha:

$$P_n(x) - \tilde{P}_n(x) = \sum_j L_j(x)(f(x_j) - \tilde{f}_j)$$

prendendo il modulo diventa:

$$|P_n(x) - \tilde{P}_n(x)| \leq \sum_j |L_j(x)| \max_j |f(x_j) - \tilde{f}_j|$$

## Condizionamento di Lagrange 2

- $\Lambda_n = \|\sum_j |L_j(x)|\|_\infty$  è la costante di Lebesgue,
- $\max_j |f(x_j)| = \max_j |P_n(x_j)| \leq \|P_n(x)\|_\infty$ ,
- si trova:

$$\frac{\|P_n - \tilde{P}_n\|_\infty}{\|P_n\|_\infty} \leq \Lambda_n \frac{\max_j |f(x_j) - \tilde{f}_j|}{\max_j |f(x_j)|}$$

La costante di Lebesgue rappresenta l'indice di condizionamento del problema.

- Nodi di Tchebichev:  $\Lambda_n \approx \frac{2}{\pi} \log n$
- Nodi Equispaziati:  $\Lambda_n \sim \frac{2^{n+1}}{en \log n}$

## Caratterizzazione errore interpolazione

**Teorema** Se  $f \in C^0([a, b])$  e  $P_n$  è il suo polinomio di interpolazione relativo ai punti  $x_0, \dots, x_n$ , si ha:

$$\|f - P_n\|_\infty \leq (1 + \Lambda_n) E_n(f)$$

dove  $E_n(f) = \inf_{q \in \mathcal{P}_n} \|f - q\|_\infty$ .

L'errore di interpolazione dipende:

- dalla costante di Lebesgue che contiene l'informazione sulla scelta dei nodi  $x_0, \dots, x_n$ .
- da  $E_n(f)$  che contiene l'informazione della sola funzione  $f$  da interpolare.

## Problema dei Minimi quadrati

- Sono dati  $m + 1$  punti sperimentali  $\{(x_i, y_i)\} \quad i = 0, \dots, m$  (alcuni possono avere anche la stessa ascissa).
- E' nota la dipendenza funzionale  $f_n(x)$  che approssima i dati sperimentali. La  $f_n(x)$  viene rappresentata come elemento di uno spazio lineare dimensione  $n + 1$ .
- La dimensione dello spazio lineare è inferiore al numero di punti sperimentali ( $n \ll m$ ).
- Si vogliono determinare gli  $n + 1$  coefficienti della funzione approssimante.

## $f_n(x)$ approssimante dei minimi quadrati

La funzione approssimante si scrive come combinazione lineare di funzioni base  $\{\phi_j(x)\}$  dello spazio lineare:

$$f_n(x) = \sum_{j=0}^n z_j \phi_j(x). \quad (1)$$

Per determinare i coefficienti non si possono assegnare le condizioni  $f_n(x_i) = y_i \quad (i = 0 \dots m)$ , perché il numero di equazioni che si otterrebbero sarebbe superiore al numero dei coefficienti incogniti.

## Condizione dei minimi quadrati

Si impone che la somma dei quadrati degli errori commessi sulle ordinate sia minima:

$$r^2(z) = \sum_{i=0}^m (y_i - f(x_i))^2 = \|y - Az\|_2^2$$

dove  $A$  è la matrice (rettangolare) delle funzioni della base, valutate sulle ascisse dei dati:

$$A = \begin{pmatrix} \phi_0(x_0) & \cdots & \phi_n(x_0) \\ \vdots & & \vdots \\ \phi_0(x_m) & \cdots & \phi_n(x_m) \end{pmatrix},$$

e  $z$  è il vettore dei coefficienti incogniti.

## Matrice $A$ nella base polinomiale

Nel caso polinomiale le funzioni base sono:

$$\phi_0(x) = 1, \quad \phi_1(x) = x, \quad \phi_2(x) = x^2, \dots, \phi_n(x) = x^n$$

e la matrice  $A$  è:

$$A^T = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ x_0 & x_1 & x_2 & \dots & x_m \\ x_0^2 & x_1^2 & x_2^2 & \dots & x_m^2 \\ \dots & & & & \\ x_0^n & x_1^n & x_2^n & \dots & x_m^n \end{pmatrix}$$

## Sistema Equazioni Normali

La condizione di minimo si trova imponendo che le componenti del gradiente dell'errore siano nulle:

$$\frac{\partial \|y - Az\|_2^2}{\partial z_i} = 0.$$

per cui il vettore dei coefficienti incogniti è dato dalla soluzione del seguente sistema quadrato (*sistema delle equazioni normali*):

$$A^T A z = A^T y$$

## Risoluzione numerica - Cholesky

Considerato che  $A^T A$  è una matrice simmetrica definita positiva si potrebbe usare la fattorizzazione di Cholesky  $A^T A = H H^T$  con  $H$  matrice triangolare inferiore e risolvere due sistemi triangolari  $H z' = A^T y$  e  $H^T z = z'$ . Tuttavia questo metodo risulta essere sensibile al condizionamento della matrice.

## Sistema Equazioni Normali 2

Il sistema ammette soluzione unica se la matrice  $A$  possiede rango massimo (le colonne sono linearmente indipendenti); la matrice dell'equazione  $A^T A$  è simmetrica e definita positiva. L'equazione normale ha la seguente forma:

$$\begin{pmatrix} \sum_i \phi_0^2(x_i) & \cdots & \sum_i \phi_0(x_i)\phi_n(x_i) \\ \vdots & & \vdots \\ \sum_i \phi_n(x_i)\phi_0(x_i) & \cdots & \sum_i \phi_n^2(x_i) \end{pmatrix} \begin{pmatrix} z_0 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} \sum_i y_i \phi_0(x_i) \\ \vdots \\ \sum_i y_i \phi_n(x_i) \end{pmatrix}$$

## Fattorizzazione QR

In alternativa si lavora direttamente sul problema della minimizzazione. Utilizzando la fattorizzazione  $QR$ :

- $Q$  è una matrice  $((m+1) \times (m+1))$  ortogonale:  $QQ^T = I$ .
- $R$  è rettangolare  $((m+1) \times (n+1))$ , composta da una matrice triangolare superiore  $((n+1) \times (n+1))$  sovrapposta ad una rettangolare  $((m-n) \times (n+1))$  di elementi nulli.

# Utilizzo Fattorizzazione QR

Questa fattorizzazione ci aiuta a trovare il vettore di minimo, infatti:

$$\|y - Az\|_2^2 = \|Q^T y - Q^T Q R z\|_2^2 = \|Q^T y - R z\|_2^2$$

la parte della matrice  $R$  con elementi nulli può essere scorporata per cui (ricorda che la norma-2 è una somma di quadrati):

$$\|Q^T y - R z\|_2^2 = \|\tilde{Q}^T y - \tilde{R} z\|_2^2 + \sum_{i=n+1}^m [(Q^T y)_i]^2$$

ed il minimo viene raggiunto quando  $z = \tilde{R}^{-1} \tilde{Q}^T y$ , che annulla il primo membro della somma; il secondo membro fornisce una stima del residuo.

# Regressione lineare 1

E' noto a priori che la relazione tra le ascisse e le ordinate dei punti è di tipo lineare. Si utilizza lo spazio dei polinomi di grado 1, ossia si approssimano i dati  $\{(x_i, y_i)\} i = 1, \dots, m$  tramite rette:

$$f_1(x) = z_1 x + z_0.$$

Tuttavia per una corretta applicazione del modello di regressione lineare si deve aggiungere l'ipotesi che tutti i valori  $y_i$  abbiano lo "stesso peso statistico". Infatti in un modello più generale si può ipotizzare che i singoli valori  $y_i$  siano soggetti ad un errore casuale la cui intensità può dipendere dal valore stesso.

# Regressione lineare 2

Sullo spazio delle rette la matrice  $A$  assume la forma:

$$A^T = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_m \end{pmatrix}$$

Ed il corrispondente sistema di equazioni normali diventa (verificare);

$$\begin{pmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{pmatrix} \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m y_i x_i \end{pmatrix}$$

(Attenzione: ora i punti sono  $m$ : da 1 a  $m$  !)

# Soluzioni regressione lineare

Le cui soluzioni sono date da:

- $z_1 = \frac{S_{xy}}{S_{xx}}$
- $z_0 = \langle y \rangle - z_1 \langle x \rangle$

avendo utilizzato i simboli statistici:

- valor medio:  $\langle x \rangle = \frac{1}{m} \sum_{i=1}^m x_i$
- $S_{xy} = \sum_{i=1}^m x_i y_i - m \langle x \rangle \langle y \rangle$  e  $S_{xx} = \sum_{i=1}^m x_i^2 - m \langle x \rangle^2$  per i termini proporzionali alla covarianza.