

# Calcolo Numerico

## Approssimazione di dati e funzioni

Simone Rebegoldi

Corso di Laurea in Informatica  
Dipartimento di Scienze Fisiche, Informatiche e Matematiche



**UNIMORE**  
UNIVERSITÀ DEGLI STUDI DI  
MODENA E REGGIO EMILIA

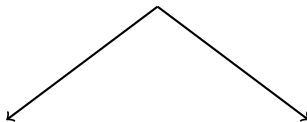


Optimization Algorithms  
and Software for  
Inverse problemS

[www.oasis.unimore.it](http://www.oasis.unimore.it)

- Nel problema dell'approssimazione, dobbiamo approssimare una funzione di cui conosciamo il valore soltanto in un insieme finito di punti dati  $(x_i, y_i)$ .
- Distinguiamo due approcci:

## Approssimazione



### Interpolazione

Se i dati non sono affetti da errore

### Approssimazione ai minimi quadrati

Se i dati sono affetti da errore

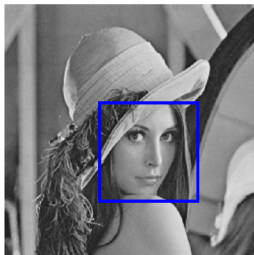
# 1. Interpolazione di dati e funzioni

# Esempio di interpolazione (stima dell'accelerazione in un videogioco)

- Supponiamo di voler programmare un gioco di guida, in cui il giocatore controlla direttamente l'accelerazione della macchina e che tale accelerazione sia descritta da una funzione  $a(t)$ .
- L'accelerazione  $a(t)$  è decisa dal giocatore durante il gioco mediante i comandi, tastiera o joystick, e non è nota a priori.
- È sensato assumere che l'interpretazione dei comandi fornisca piuttosto dei valori  $a(t_i)$ , dove  $t_i, i = 0, 1, \dots, n$  rappresentano diversi momenti del gioco.
- Per conoscere la posizione della macchina, occorre integrare due volte l'accelerazione che però non è nota in forma analitica.
- **Occorre definire una approssimazione della funzione  $a(t)$  a partire dai valori noti  $a(t_i)$ , dove  $t_i, i = 0, 1, \dots, n$ .**



- Un'immagine in formato digitale, che supponiamo per semplicità in scala di grigi, consiste in una matrice di numeri, detti pixel, che rappresentano la sfumatura di grigio presente nel corrispondente punto dell'immagine.
- Ingrandire un'immagine significa definire una matrice più grande, di cui dobbiamo determinare pixel in più rispetto a quelli dell'immagine di partenza, in modo che siano "coerenti" con essi.



## Due diverse strategie di interpolazione a confronto



## Definizione

Siano dati  $n$  punti  $(x_i, y_i)$ , con  $x_i \in [a, b]$ ,  $i = 0, \dots, n$ , fissati nel piano cartesiano, dove gli  $x_i$  sono detti nodi e  $y_i$  le ordinate dei punti.

Il problema dell'**interpolazione** consiste nel trovare una funzione  $g : [a, b] \rightarrow \mathbb{R}$  il cui grafico passa per i punti dati, ossia che soddisfa le condizioni di interpolazione

$$g(x_i) = y_i, \quad i = 0, \dots, n.$$

La funzione  $g$  si dice **interpolante** rispetto ai punti di interpolazione  $(x_i, y_i)$ .

Quando le ordinate dei punti di interpolazione sono le immagini dei nodi tramite una funzione  $f : [a, b] \rightarrow \mathbb{R}$ , allora si dice che  **$g$  interpola la funzione  $f$** .

- Nel caso in cui  $y_i = f(x_i)$ ,  $i = 0, 1, \dots, n$ , la funzione interpolante  $g$  rappresenta un'approssimazione di  $f$  nei punti in cui non è nota:

$$\begin{aligned} g(x_i) &= f(x_i), \quad i = 0, 1, \dots, n \\ g(x) &\simeq f(x), \quad \forall x \neq x_i. \end{aligned}$$

- Fissati i punti da interpolare, esistono infinite funzioni interpolanti.
- Per avere l'unicità della soluzione del problema dell'interpolazione, occorre restringere la ricerca ad una sottoclasse di funzioni più semplici, ad esempio i **polinomi**.



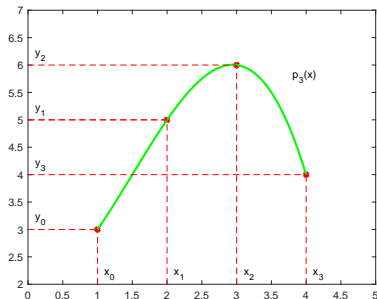
## Definizione

Dati  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ , dove gli  $x_i$  sono detti nodi e  $y_i$  le ordinate dei punti, il problema dell'**interpolazione polinomiale** consiste nel determinare un polinomio  $p_n(x) = a_0 + a_1x + \dots + a_nx^n$  di grado al più  $n$  tale che

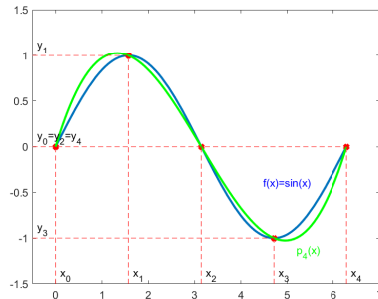
$$p_n(x_i) = y_i, \quad i = 0, 1, \dots, n.$$

Il polinomio  $p_n$  si dice **interpolante** rispetto ai punti di interpolazione  $(x_i, y_i)$ . Quando le ordinate dei punti di interpolazione sono le immagini dei nodi tramite una funzione  $f : [a, b] \rightarrow \mathbb{R}$ , allora si dice che  **$p_n$  interpola la funzione  $f$** .

Interpolazione di dati



Interpolazione di funzione



## Teorema

Dati  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ , il polinomio interpolante rispetto ai punti  $(x_i, y_i)$  esiste ed è unico se e soltanto se i nodi di interpolazione  $x_i$ ,  $i = 0, 1, \dots, n$ , sono a due a due distinti, ovvero

$$p_n(x) \text{ esiste ed è unico} \quad \Leftrightarrow \quad x_i \neq x_j, \forall i \neq j.$$

## Dimostrazione

Un generico polinomio  $p_n$  di grado  $n$  si rappresenta in forma canonica come combinazione lineare delle funzioni  $1, x, x^2, \dots, x^n$ , ovvero

$$p_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n.$$

Le condizioni di interpolazione si possono quindi scrivere come

$$p_n(x_i) = y_i, \quad i = 0, \dots, n \quad \Leftrightarrow \quad \begin{cases} a_0 + a_1x_0 + a_2x_0^2 + \dots + a_nx_0^n = y_0 \\ a_0 + a_1x_1 + a_2x_1^2 + \dots + a_nx_1^n = y_1 \\ \vdots \\ a_0 + a_1x_n + a_2x_n^2 + \dots + a_nx_n^n = y_n \end{cases}$$

dunque come un sistema lineare di  $n + 1$  equazioni in  $n + 1$  incognite della forma

$$V\alpha = y$$

$$V = \begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix} \quad \alpha = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} \quad y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

matrice dei coefficienti

incognite

termine noto

- Abbiamo riscritto in modo equivalente il problema dell'interpolazione polinomiale come il problema della risoluzione del sistema lineare

$$V\alpha = y,$$

dove  $V$  è detta **matrice di Vandermonde** e le componenti del vettore soluzione  $\alpha$  sono i coefficienti del polinomio interpolante  $p_n(x)$  in forma canonica.

- Si dimostra che

$$\det(V) = \prod_{i>j} (x_i - x_j),$$

pertanto

$$\det(V) \neq 0 \quad \Leftrightarrow \quad x_i \neq x_j, \quad i \neq j.$$

Dunque il sistema ammette una ed una sola soluzione se e solo se i nodi sono a due a due distinti.  $\square$

- Dal punto di vista numerico, “calcolare” il polinomio interpolante significa progettare un algoritmo che, dati in input i punti di interpolazione  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$  e un punto  $\bar{x} \in \mathbb{R}$ , restituisca in output il valore  $p_n(\bar{x})$ .
- Esistono diversi metodi per calcolare il polinomio interpolante, ognuno dei quali è basato su una diversa rappresentazione del polinomio rispetto ad un’opportuna **base** dello spazio dei polinomi di grado al più  $n$ .

# 1) Metodo dei coefficienti indeterminati

Dati  $(x_i, y_i)$ ,  $i = 0, 1, \dots$ , con  $x_i \neq x_j$  per  $i \neq j$ , si calcolano

$$V = \begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix}, \quad y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$

e si risolve il sistema lineare

$$V\alpha = y.$$

Le componenti del vettore soluzione  $\alpha = (a_0, a_1, \dots, a_n)$  sono i coefficienti del polinomio interpolante  $p_n(x)$  rispetto alla base canonica  $1, x, x^2, \dots, x^n$  dello spazio dei polinomi di grado al più  $n$ .

## Costo computazionale

- La soluzione diretta del sistema  $V\alpha = y$  ha un costo dell'ordine di  $n^3$  (il costo della fattorizzazione di  $V$ ).
- Una volta calcolati i coefficienti  $a_0, a_1, \dots, a_n$ , si vuole calcolare il valore del polinomio interpolante  $p_n(\bar{x})$  in un punto  $\bar{x} \neq x_i, i = 0, 1, \dots, n$ .  
L'algoritmo più efficiente è quello di Horner, che ha un costo di  $n$  prodotti e somme floating point.

## Vantaggi del metodo dei coefficienti indeterminati

- I coefficienti  $a_0, a_1, \dots, a_n$  della rappresentazione canonica non dipendono dal punto  $\bar{x}$ ; una volta calcolati, si possono utilizzare per calcolare lo stesso polinomio in punti diversi.

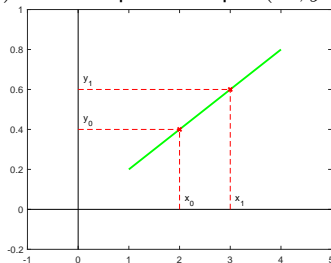
## Svantaggi del metodo dei coefficienti indeterminati

- **La matrice di Vandermonde  $V$  è in generale malcondizionata**
  - ⇒ Il numero di condizionamento  $\kappa(V)$  aumenta a dismisura al crescere di  $n$
  - ⇒ All'aumentare di  $\kappa(V)$ , aumenta la possibilità che gli algoritmi basati su di una fattorizzazione di  $V$  siano numericamente instabili, nel senso che l'errore relativo sui coefficienti può diventare enorme
  - ⇒ Il polinomio calcolato può essere molto distante da quello reale!
- **Il costo del metodo (dell'ordine di  $n^3$ ) non è ottimale**: si può fare di meglio!



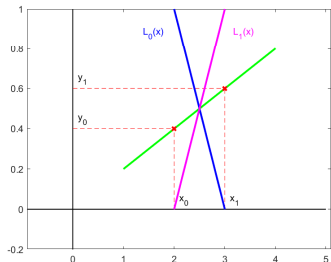
Consideriamo il caso  $n = 1$ , in cui  $p_1(x)$  è la retta passante per  $(x_0, y_0)$ ,  $(x_1, y_1)$ .

$$p_1(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0).$$



Raggruppando i termini in modo diverso, il polinomio  $p_1(x)$  si può scrivere come

$$p_1(x) = y_0 \underbrace{\frac{x - x_1}{x_0 - x_1}}_{L_0(x)} + y_1 \underbrace{\frac{x - x_0}{x_1 - x_0}}_{L_1(x)}.$$



$$p_1(x) = y_0 \underbrace{\frac{x - x_1}{x_0 - x_1}}_{L_0(x)} + y_1 \underbrace{\frac{x - x_0}{x_1 - x_0}}_{L_1(x)}.$$

- Abbiamo scritto il polinomio di interpolazione di grado 1 come combinazione lineare di polinomi di grado 1, denotati con  $L_0(x)$  e  $L_1(x)$ .
- I coefficienti della combinazione lineare sono le ordinate dei punti di interpolazione  $y_0, y_1$ .
- I polinomi  $L_0(x), L_1(x)$  soddisfano la seguente proprietà

$$\begin{cases} L_0(x_0) = 1 \\ L_0(x_1) = 0 \end{cases}, \quad \begin{cases} L_1(x_0) = 0 \\ L_1(x_1) = 1 \end{cases}$$

e sono gli unici polinomi di grado 1 a soddisfare tale proprietà.

### Teorema (rappresentazione del polinomio interpolante nella forma di Lagrange)

Dati i punti  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ , con  $x_i \neq x_j$  per  $i \neq j$ , il polinomio di interpolazione  $p_n$  rispetto ai punti  $(x_i, y_i)$  si può scrivere nella forma

$$p_n(x) = y_0 L_0(x) + y_1 L_1(x) + \dots + y_n L_n(x) = \sum_{k=0}^n y_k L_k(x),$$

dove  $L_k$ ,  $k = 0, 1, \dots, n$  sono  $n + 1$  polinomi di grado  $n$  che si annullano in tutti i nodi di interpolazione tranne il  $k$ -esimo, nel quale valgono 1:

$$\begin{cases} L_k(x_k) = 1 \\ L_k(x_j) = 0, \quad \text{se } j \neq k. \end{cases}$$

La forma esplicita di tali polinomi  $L_k$  è data da

$$\begin{aligned} L_k(x) &= \frac{(x - x_0) \cdot \dots \cdot (x - x_{k-1})(x - x_{k+1}) \cdot \dots \cdot (x - x_n)}{(x_k - x_0) \cdot \dots \cdot (x_k - x_{k-1})(x_k - x_{k+1}) \cdot \dots \cdot (x_k - x_n)} \\ &= \prod_{i=0, i \neq k}^n \frac{x - x_i}{x_k - x_i}, \quad k = 0, \dots, n. \end{aligned}$$

### Dimostrazione

Usando le proprietà che caratterizzano i polinomi  $L_k$ , si verifica che

$$p_n(x_0) = y_0 \underbrace{L_0(x_0)}_1 + y_1 \underbrace{L_1(x_0)}_0 + \dots + y_n \underbrace{L_n(x_0)}_0 = y_0$$

$$p_n(x_1) = y_0 \underbrace{L_0(x_1)}_0 + y_1 \underbrace{L_1(x_1)}_1 + \dots + y_n \underbrace{L_n(x_1)}_0 = y_1$$

$$\vdots$$

$$p_n(x_n) = y_0 \underbrace{L_0(x_n)}_0 + y_1 \underbrace{L_1(x_n)}_0 + \dots + y_n \underbrace{L_n(x_n)}_1 = y_n$$

da cui segue che  $p_n$  verifica le  $n + 1$  condizioni di interpolazione.

Dunque  $p_n$  deve l'essere l'unico polinomio interpolante i punti  $(x_i, y_i)$ .  $\square$

### Teorema

L'insieme dei polinomi  $\{L_k\}_{k=0,\dots,n}$  è un insieme linearmente indipendente e dunque una base per lo spazio dei polinomi di grado al più  $n$ .  
Tale base viene detta **base di Lagrange**.

### Dimostrazione

Siano dati  $n + 1$  coefficienti reali  $\lambda_0, \dots, \lambda_n$  tali che  $\sum_{k=0}^n \lambda_k L_k = 0$ , ovvero

$$0 = \lambda_0 L_0(x) + \lambda_1 L_1(x) + \dots + \lambda_n L_n(x), \quad \forall x \in \mathbb{R}.$$

In particolare l'uguaglianza deve valere per  $x = x_k$ , ossia per i nodi di interpolazione. Quindi vale che

$$0 = \lambda_0 L_0(x_k) + \lambda_1 L_1(x_k) + \dots + \lambda_n L_n(x_k) = \lambda_k L_k(x_k) = \lambda_k.$$

Dunque  $\lambda_k = 0$  per ogni  $k = 0, 1, \dots, n$ , da cui segue per definizione la lineare indipendenza di  $\{L_k\}_{k=0,1,\dots,n}$ .  $\square$

## 2) Metodo di Lagrange per il calcolo del polinomio interpolante

### Metodo di Lagrange per il calcolo del polinomio di interpolazione in $\bar{x}$

INPUT:  $(x_i, y_i)$ ,  $i = 0, \dots, n$  e  $\bar{x} \neq x_i$ .

1. Calcolare i valori  $L_k(\bar{x})$ ,  $k = 0, \dots, n$
2. Calcolare la combinazione lineare

$$p_n(\bar{x}) = y_0 L_0(\bar{x}) + y_1 L_1(\bar{x}) + \dots + y_n L_n(\bar{x}).$$

OUTPUT:  $p_n(\bar{x})$ .

### Costo computazionale del metodo di Lagrange

La parte computazionalmente più costosa del metodo è la valutazione della base di Lagrange in  $\bar{x}$ :

$$L_k(\bar{x}) = \prod_{i=0, i \neq k} \frac{\bar{x} - x_i}{x_k - x_i}, \quad k = 0, \dots, n.$$

Per evitare di ripetere più volte le stesse operazioni nel calcolo dei numeratori, si può procedere calcolando inizialmente la quantità

$$\omega_n(\bar{x}) = (\bar{x} - x_0) \cdot \dots \cdot (\bar{x} - x_n)$$

e osservando che

$$L_k(\bar{x}) = \frac{\omega_n(\bar{x})}{(\bar{x} - x_k) \prod_{i=0, i \neq k}^n (x_k - x_i)}.$$

## 2) Metodo di Lagrange per il calcolo del polinomio interpolante

### Costo computazionale del metodo di Lagrange

$$\begin{cases} L_k(\bar{x}) = \prod_{i=0, i \neq k}^n \frac{\bar{x} - x_i}{x_k - x_i} \\ \omega_n(\bar{x}) = (\bar{x} - x_0) \cdot \dots \cdot (\bar{x} - x_n) \end{cases} \Rightarrow L_k(\bar{x}) = \frac{\omega_n(\bar{x})}{(\bar{x} - x_k) \prod_{i=0, i \neq k}^n (x_k - x_i)}$$

$$p_n(\bar{x}) = y_0 L_0(\bar{x}) + y_1 L_1(\bar{x}) + \dots + y_n L_n(\bar{x})$$

$q \leftarrow \bar{x} - x_0$

FOR  $i = 1, \dots, n$

$q \leftarrow q * (\bar{x} - x_i)$

$n + 1$  sottrazioni

$n$  prodotti

FOR  $k = 0, \dots, n$

$d \leftarrow 1$

  FOR  $i = 0, \dots, n$

    IF  $i \neq k$

$d \leftarrow d * (x_k - x_i)$

$(n + 1)^2$  sottrazioni

$(n + 1)^2$  prodotti

$n + 1$  divisioni

$l_k \leftarrow q / ((\bar{x} - x_k) * d)$

$s \leftarrow y_0 * l_0$

FOR  $k = 1, \dots, n$

$s \leftarrow s + y_k * l_k$

$n$  somme

$n$  prodotti

Costo totale:  $\mathcal{O}(2n^2)$  operazioni.



### Vantaggi del metodo di Lagrange

- È meno costoso del metodo dei coefficienti indeterminati (da  $\mathcal{O}(n^3)$  a  $\mathcal{O}(2n^2)$  operazioni).
- È più stabile numericamente rispetto al metodo dei coefficienti indeterminati.
- I polinomi  $L_k$  dipendono soltanto dai nodi  $x_0, \dots, x_n$ , non dalle ordinate  $y_0, \dots, y_n$  dei punti di interpolazione.  
⇒ Se devo interpolare un nuovo insieme di punti  $(x_i, z_i)$ ,  $i = 0, \dots, n$ , i cui nodi sono gli stessi ma non le ordinate, allora è sufficiente cambiare gli  $n + 1$  coefficienti della combinazione lineare, senza ricalcolare tutti gli elementi della base di Lagrange:

$$(x_i, y_i), \quad i = 0, \dots, n \quad \Rightarrow \quad p_n(x) = \sum_{k=0}^n y_k L_k(x)$$
$$(x_i, z_i), \quad i = 0, \dots, n \quad \Rightarrow \quad \tilde{p}_n(x) = \sum_{k=0}^n z_k L_k(x)$$

⇒ Più conveniente rispetto al metodo dei coefficienti indeterminati, in cui il sistema lineare  $V\alpha = y$  dipende sia dai nodi che dalle ordinate.

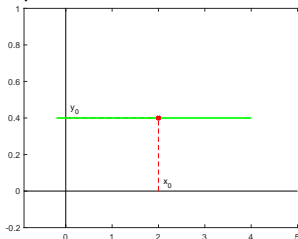
### Svantaggi del metodo di Lagrange

- Nel caso in cui si voglia aggiungere un nuovo punto  $(x_{n+1}, y_{n+1})$  all'insieme dei punti di interpolazione  $(x_i, y_i)$ ,  $i = 0, \dots, n$ , e di conseguenza aumentare il grado del polinomio interpolante, occorre aggiornare la base di Lagrange  $\Rightarrow$  si aggiornano tutti gli  $L_k$ ,  $k = 0, \dots, n$  e si calcola il nuovo elemento  $L_{n+1}$ .
- Il costo del metodo non è ottimale: si può fare ancora meglio!

- Esiste un'ulteriore rappresentazione alternativa del polinomio interpolante che permette di aggiungere e sottrarre nodi senza bisogno di dover aggiornare tutti gli elementi della base di rappresentazione del polinomio stesso: è la **rappresentazione di Newton**.

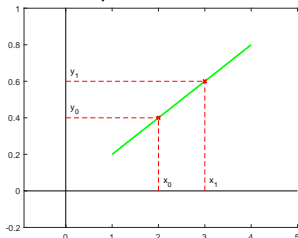
- **Caso  $n = 0$ :** dato  $(x_0, y_0)$ , il polinomio interpolante è

$$p_0(x) = y_0, \quad \forall x \in \mathbb{R}.$$



- **Caso  $n = 1$ :** dati  $(x_0, y_0)$ ,  $(x_1, y_1)$  il polinomio interpolante è

$$\begin{aligned} p_1(x) &= y_0 + \frac{y_1 - y_0}{x_1 - x_0} (x - x_0) \\ &= p_0(x) + \frac{y_1 - y_0}{x_1 - x_0} (x - x_0). \end{aligned}$$



Il polinomio di grado 1 si ottiene a partire da quello di grado 0 con l'aggiunta di un termine di grado 1 pari a  $(x - x_0)$  moltiplicato per il rapporto incrementale  $\frac{y_1 - y_0}{x_1 - x_0}$ .

## Definizione

Siano dati i punti  $(x_i, y_i)$ ,  $i = 0, \dots, n$ .

- Si definiscono le  $n + 1$  differenze divise di ordine 0 come le quantità

$$y[x_i] = y_i, \quad i = 0, 1, \dots, n.$$

- Si definiscono le  $n$  differenze divise di ordine 1 come le quantità

$$y[x_i, x_{i+1}] = \frac{y[x_{i+1}] - y[x_i]}{x_{i+1} - x_i} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \quad i = 0, \dots, n-1.$$

- Si definiscono le  $n - 1$  differenze divise di ordine 2 come le quantità

$$y[x_i, x_{i+1}, x_{i+2}] = \frac{y[x_{i+1}, x_{i+2}] - y[x_i, x_{i+1}]}{x_{i+2} - x_i}, \quad i = 0, \dots, n-2.$$

- Dato  $1 \leq m \leq n$ , le  $n - m + 1$  differenze divise di ordine  $m$  si definiscono ricorsivamente come

$$y[x_i, x_{i+1}, \dots, x_{i+m}] = \frac{y[x_{i+1}, x_{i+2}, \dots, x_{i+m}] - y[x_i, x_{i+1}, \dots, x_{i+m-1}]}{x_{i+m} - x_i},$$

per ogni  $i = 0, \dots, n - m$ .

## Tabella delle differenze divise

Permette di organizzare il calcolo delle differenze divise.  
Va letta da sinistra a destra.

$$\begin{array}{c|c|c|c} y_0 & & & \\ y_1 & y[x_0, x_1] = \frac{y_1 - y_0}{x_1 - x_0} & & \\ y_2 & y[x_1, x_2] = \frac{y_2 - y_1}{x_2 - x_1} & y[x_0, x_1, x_2] = \frac{y[x_1, x_2] - y[x_0, x_1]}{x_2 - x_0} & \\ y_3 & y[x_2, x_3] = \frac{y_3 - y_2}{x_3 - x_2} & y[x_1, x_2, x_3] = \frac{y[x_2, x_3] - y[x_1, x_2]}{x_3 - x_1} & y[x_0, x_1, x_2, x_3] = \frac{y[x_1, x_2, x_3] - y[x_0, x_1, x_2]}{x_3 - x_0} \end{array}$$

## Teorema

Siano dati i punti  $(x_i, y_i)$ ,  $i = 0, \dots, n$ , con  $x_i \neq x_j$  per  $i \neq j$ .

L'unica differenza divisa  $y[x_0, \dots, x_n]$  di ordine  $n$  coincide con il coefficiente di  $x^n$  del polinomio interpolante rispetto ai punti considerati, ovvero:

$$y[x_0, \dots, x_n] = \sum_{k=0}^n \frac{y_k}{\prod_{i=0, i \neq k}^n (x_k - x_i)}.$$

### 3) Metodo di Newton per il calcolo del polinomio interpolante

#### Teorema (rappresentazione del polinomio interpolante nella forma di Newton)

Siano dati i punti  $(x_i, y_i)$ ,  $i = 0, \dots, n$ , con  $x_i \neq x_j$  per  $i \neq j$ .

Definiamo in modo ricorsivo i seguenti polinomi:

$$p_0(x) = y[x_0]$$

$$p_{k+1}(x) = p_k(x) + y[x_0, \dots, x_{k+1}](x - x_0) \cdot \dots \cdot (x - x_k), \quad k = 0, \dots, n.$$

Allora, per ogni  $k = 0, \dots, n$ , si ha che  $p_{k+1}$  è il polinomio di interpolazione relativamente ai punti  $(x_0, y_0), \dots, (x_{k+1}, y_{k+1})$ .



#### Dimostrazione

- Per  $k = 0$ , l'asserto è banalmente vero.
- Per induzione, supponiamo che l'asserto sia vero per  $k - 1$ , ovvero: il polinomio  $p_k$  interpola i dati  $(x_0, y_0), \dots, (x_k, y_k)$ .  
Per definizione il polinomio  $p_{k+1}$  è un polinomio di grado al più  $k + 1$ ; inoltre, per ipotesi induttiva, si ha

$$p_{k+1}(x_i) = p_k(x_i) + 0 = y_i, \quad i = 0, \dots, n.$$

Sia  $g(x)$  il polinomio di interpolazione di grado al più  $k + 1$ , allora  $r(x) = g(x) - p_{k+1}(x)$  è un polinomio di grado al più  $k$ , dal momento che  $g$  e  $p_{k+1}$  hanno lo stesso coefficiente per la  $x$  di grado  $k + 1$ .

$\Rightarrow$  Abbiamo provato che  $r(x)$  è un polinomio di grado al più  $k$  che si annulla nei  $k + 1$  punti  $x_i, i = 0, \dots, k$ .

$\Rightarrow r(x) \equiv 0$ .

$\Rightarrow g(x) \equiv p_{k+1}(x)$ .  $\square$

#### Rappresentazione del polinomio interpolante nella forma di Newton

$$\begin{aligned} p_n(x) &= p_{n-1}(x) + f[x_0, \dots, x_n](x - x_0) \cdot \dots \cdot (x - x_{n-1}) \\ &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ &\quad + \dots + f[x_0, \dots, x_n](x - x_0) \cdot \dots \cdot (x - x_{n-1}) \end{aligned}$$

- Il metodo di Newton si basa sulla rappresentazione del polinomio interpolante rispetto alla seguente base dello spazio dei polinomi:

$$1, x - x_0, (x - x_0)(x - x_1), \dots, (x - x_0)(x - x_1) \dots (x - x_{n-1}).$$

- I coefficienti di tale rappresentazione sono le differenze divise presenti sulla diagonale discendente della tabella delle differenze divise.

#### Vantaggi del metodo di Newton

- Se aggiungo un punto di interpolazione  $(x_{i+1}, y_{i+1})$ , non devo aggiornare tutte le differenze divise già calcolate; è sufficiente calcolarne una nuova, quella di ordine  $n + 1$  relativa ai punti  $x_0, \dots, x_{n+1}$ . Infatti

$$\begin{aligned} p_{n+1}(x) &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ &\quad + \dots + f[x_0, \dots, x_n](x - x_0) \cdot \dots (x - x_{n-1}) \\ &\quad + f[x_0, \dots, x_{n+1}](x - x_0) \cdot \dots (x - x_n) \\ &= \underbrace{p_n(x)}_{\text{già calcolato}} + \underbrace{f[x_0, \dots, x_{n+1}](x - x_0) \cdot \dots (x - x_n)}_{\text{da calcolare}}. \end{aligned}$$

⇒ Più conveniente del metodo di Lagrange!

- La complessità computazionale è la metà rispetto a quella del metodo di Lagrange, ossia  $\mathcal{O}(n^2)$  operazioni.

## Problema

Dati  $(x_i, y_i)$ ,  $i = 0, \dots, n$ , con  $y_i = f(x_i)$ ,  $i = 0, \dots, n$ , dove  $f : [a, b] \rightarrow \mathbb{R}$  è una funzione di cui non si conosce l'espressione analitica, si vuole **definire una funzione  $g : [a, b] \rightarrow \mathbb{R}$  che “assomigli” il più possibile a  $f$ .**

- I dati del problema sono i punti del piano  $(x_i, y_i)$ ,  $i = 0, \dots, n$ .
- La soluzione del problema può essere di due tipi:
  - soluzione analitica: si definisce l'espressione esplicita della funzione  $g$ .
  - soluzione numerica: si definiscono algoritmi che permettono di calcolare la funzione  $g$  in un qualsiasi punto dell'intervallo di riferimento  $[a, b]$ .
- Il concetto di “somiglianza” o “distanza” fra funzioni si esprime mediante la definizione di opportune norme di funzioni.

## Definizione (norma infinito di una funzione)

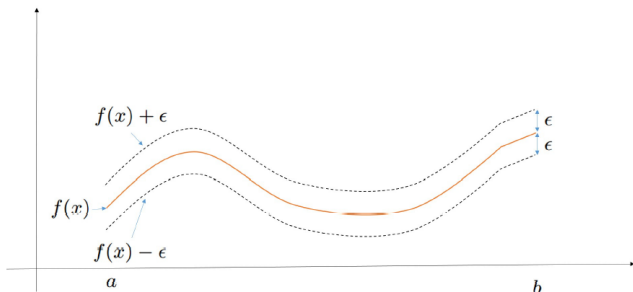
Sia  $f : [a, b] \rightarrow \mathbb{R}$  una funzione continua. La **norma infinito di  $f$**  è definita come

$$\|f\|_{\infty} = \max_{x \in [a, b]} |f(x)|.$$

La **distanza in norma infinito fra  $f$  e  $g : [a, b] \rightarrow \mathbb{R}$**  è definita come

$$d(f, g) = \|f - g\|_{\infty}.$$

- Se  $\|f - g\|_{\infty} < \epsilon$  con  $\epsilon > 0$ , allora il grafico di  $g$  si trova in un “canale” di raggio  $\epsilon$  centrato sul grafico di  $f$ .



- Nel caso dell'interpolazione, la funzione approssimante è il polinomio interpolante  $p_n(x)$  relativo ai punti  $(x_i, f(x_i))$ ,  $i = 0, \dots, n$ .
- Si definisce la funzione

$$R_n(x) = f(x) - p_n(x), \quad \forall x \in [a, b],$$

detta **resto (o errore) di interpolazione**.

## Problema dell'approssimazione di una funzione mediante il polinomio interpolante

Individuare le condizioni per cui  $p_n(x)$  si può considerare una buona approssimazione di  $f(x)$  nell'intervallo  $[a, b]$  per  $n$  sufficientemente grande, ossia

$$\lim_{n \rightarrow \infty} \|R_n\|_{\infty} = 0.$$

## Teorema (formula del resto di interpolazione)

Siano dati la funzione  $f \in C^{n+1}([a, b])$ , i nodi  $a = x_0 < x_1 < x_2 < \dots < x_n = b$  in  $[a, b]$  e il polinomio di interpolazione  $p_n(x)$  relativo a  $(x_i, f(x_i))$ ,  $i = 0, \dots, n$ .

Allora vale la seguente formula

$$R_n(x) = f(x) - p_n(x) = \frac{\omega(x)}{(n+1)!} f^{(n+1)}(\xi(x)),$$

dove  $\omega(x) = (x - x_0) \cdot (x - x_1) \cdot \dots \cdot (x - x_n)$  e  $\xi(x) \in [a, b]$ .

## Dimostrazione

- Dato che  $p_n$  soddisfa le condizioni di interpolazione, si ha

$$R_n(x_i) = 0, \quad i = 0, \dots, n,$$

cioè  $R_n(x)$  ammette gli  $n + 1$  zeri  $x_i, i = 0, \dots, n$ . Di conseguenza

$$R_n(x) = q(x)(x - x_0)(x - x_1) \dots (x - x_n) = q(x)\omega(x)$$

dove  $q(x)$  è una funzione incognita da determinare.

- Fissato  $\bar{x} \in [a, b]$ ,  $\bar{x} \neq x_i \forall i$ , si definisce la funzione  $g \in C^{n+1}([a, b])$  data da

$$g(x) = R_n(x) - q(\bar{x})\omega(x) = (q(x) - q(\bar{x}))\omega(x), \quad \forall x \in [a, b].$$

La funzione  $g(x)$  ammette gli  $n + 2$  zeri  $x_0, x_1, \dots, x_n$  e  $\bar{x}$ , ovvero

$$g(x_0) = g(x_1) = \dots = g(x_n) = g(\bar{x}) = 0.$$



## Dimostrazione

- Ricordiamo il teorema di Rolle:

Ipotesi:

$g : [a, b] \rightarrow \mathbb{R}$  continua in  $[a, b]$   
 $g$  derivabile in  $(a, b)$   
 $g(a) = g(b)$

Tesi:

$\exists c \in (a, b)$  tale che  
 $g'(c) = 0$ .

- Applicando il teorema di Rolle a  $g$  nei sottointervalli definiti dai punti  $x_0, \dots, x_n, \bar{x}$  e in sequenza alle sue derivate fino all'ordine  $n + 1$ , si ha che

$g' \in C^n([a, b])$  ha  $n + 1$  zeri distinti in  $[a, b]$

$g'' \in C^{n-1}([a, b])$  ha  $n$  zeri distinti in  $[a, b]$

$\vdots$

$g^{(n)} \in C^1([a, b])$  ha 2 zeri distinti in  $[a, b]$

$g^{(n+1)} \in C([a, b])$  ha 1 zero in  $[a, b]$ .

- Dunque esiste  $\xi \in [a, b]$  tale che

$$g^{(n+1)}(\xi) = 0.$$

Osserviamo che

$$\begin{aligned} 0 = g^{(n+1)}(\xi) &= R_n^{(n+1)}(\xi) - q(\bar{x})\omega^{(n+1)}(\xi) \\ &= f^{(n+1)}(\xi) - \underbrace{p_n^{(n+1)}(\xi)}_{=0} - q(\bar{x}) \underbrace{\omega^{(n+1)}(\xi)}_{=(n+1)!} \\ &= f^{(n+1)}(\xi) - q(\bar{x})(n+1)!. \end{aligned}$$

In conclusione

$$q(\bar{x}) = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

da cui segue la tesi.  $\square$

- La formula del resto di interpolazione appena dimostrata suggerisce alcune considerazioni di carattere qualitativo sull'accuratezza dell'approssimazione ottenibile con il polinomio di interpolazione.
- Assumendo  $|R_n(x)|$  come misura dell'accuratezza, tale misura dipende da tre fattori:
  1. il grado  $n$  del polinomio o, equivalentemente, il numero dei punti;
  2. la scelta dei punti di interpolazione, implicitamente contenuta nel termine  $\omega(x)$ ;
  3. la regolarità della funzione  $f$ .
- Dalla formula del resto di interpolazione per una funzione  $f \in C^{n+1}([a, b])$ , si può dedurre la seguente stima:

$$|R_n(x)| \leq \frac{(b-a)^{n+1}}{(n+1)!} M_n, \quad \forall x \in [a, b],$$

dove  $M_n = \max_{x \in [a, b]} |f^{(n+1)}(x)|$  è una costante che dipende dal problema e non dalla scelta dei nodi di interpolazione.

- Se  $\max_n M_n < \infty$ , ossia se le derivate della funzione sono equilimitate, allora la precedente stima garantisce la **convergenza uniforme della successione di polinomi  $\{p_n\}_{n \in \mathbb{N}}$  alla funzione  $f$** , ovvero

$$\lim_{n \rightarrow \infty} \|R_n\|_{\infty} = \lim_{n \rightarrow \infty} \max_{x \in [a, b]} |f(x) - p_n(x)| = 0.$$

La convergenza uniforme garantisce una crescente accuratezza del polinomio  $p_n$  su tutto  $[a, b]$  e non soltanto in corrispondenza dei singoli punti dell'intervallo.

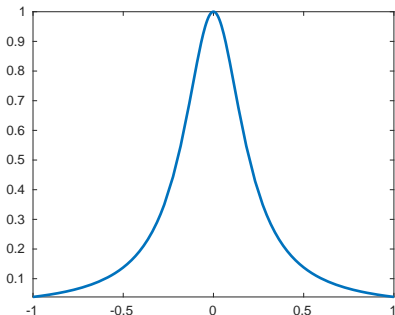
- Esempi di funzioni con derivate equilimitate sono la funzione  $e^x$  o le funzioni trigonometriche  $\sin(x)$  e  $\cos(x)$ . Tuttavia, senza l'equilimitatezza delle derivate, la convergenza uniforme della successione dei polinomi di interpolazione non è più garantita.

Sia data la **funzione di Runge**, ovvero la funzione  $f : [-1, 1] \rightarrow \mathbb{R}$  definita come

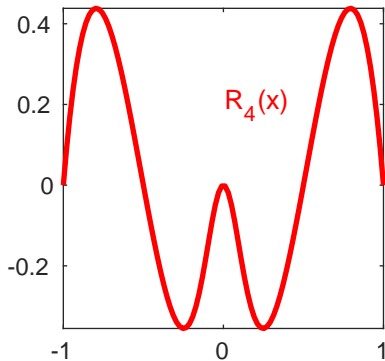
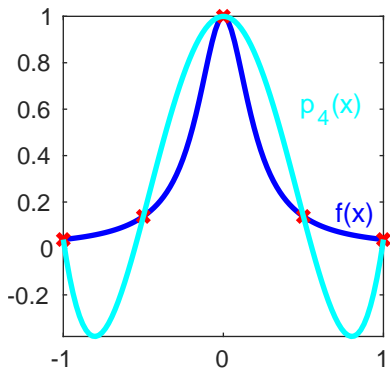
$$f(x) = \frac{1}{1 + 25x^2}, \quad \forall x \in [-1, 1].$$

Supponiamo di voler interpolare tale funzione su  $n + 1$  nodi equispaziati in  $[-1, 1]$ :

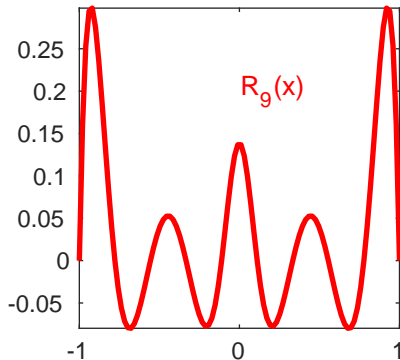
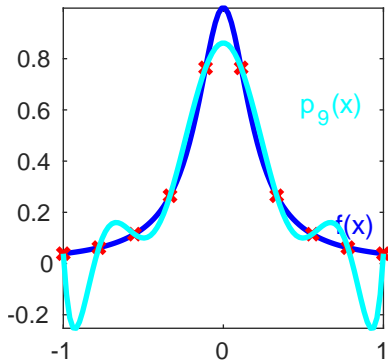
$$x_i = -1 + \frac{2i}{n}, \quad i = 0, \dots, n.$$



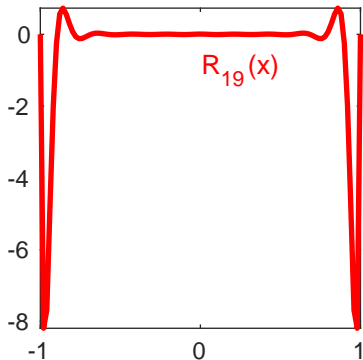
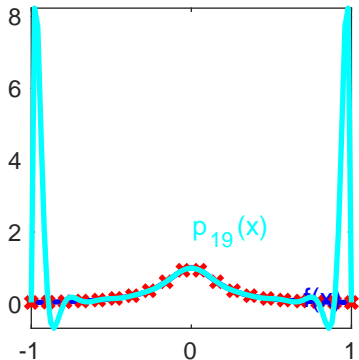
# Controesempio: la funzione di Runge



# Controesempio: la funzione di Runge



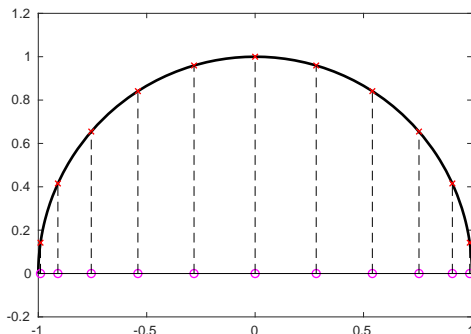
# Controesempio: la funzione di Runge





- Non sempre aumentare il numero dei punti significa aumentare l'accuratezza dell'approssimazione.
- Nell'esempio di Runge, utilizzando un numero crescente di punti equispaziati si ottiene l'effetto opposto: il polinomio di interpolazione presenta oscillazioni sempre più ampie in prossimità degli estremi dell'intervallo  $[-1, 1]$ . Questo comportamento oscillatorio è detto **fenomeno di Runge**.
- Una possibile strategia per evitare il fenomeno di Runge consiste nell'utilizzare una diversa distribuzione di nodi, più densa in prossimità degli estremi dell'intervallo dove occorre maggiore controllo.

I nodi di Chebychev si ottengono partizionando in modo uniforme la semicirconferenza goniometrica e proiettando i punti di questa partizione sul diametro.



In formule:

$$t_i = \cos \left( \frac{(2i + 1)\pi}{2(n + 1)} \right), \quad i = 0, \dots, n.$$

La precedente definizione si può estendere a qualsiasi intervallo della retta reale, tenendo conto della mappa

$$\mu(t) = \frac{b-a}{2}t + \frac{a+b}{2}$$

che ne mantiene la distribuzione.

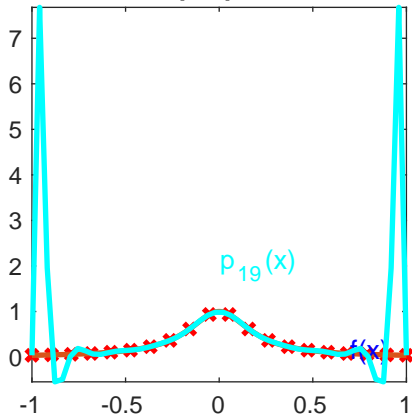
## Teorema di Bernstein

Se  $f \in C^1([a, b])$  e  $p_n$  è il polinomio di interpolazione di  $f$  relativo ai nodi di Chebychev  $t_i$ ,  $i = 0, \dots, n$ , allora

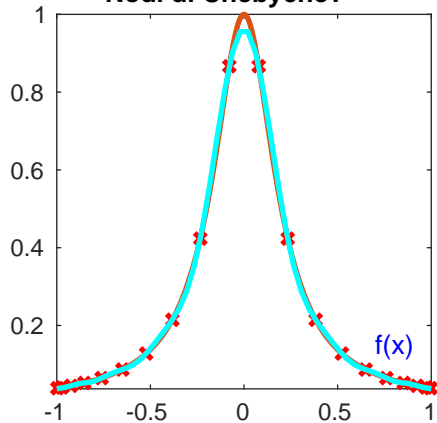
$$\lim_{n \rightarrow \infty} \|R_n\|_{\infty} = \lim_{n \rightarrow \infty} \|f - p_n\|_{\infty} = 0,$$

il che equivale a dire che  $\{p_n\}_{n \in \mathbb{N}}$  converge uniformemente a  $f$ .

## Nodi equispaziati

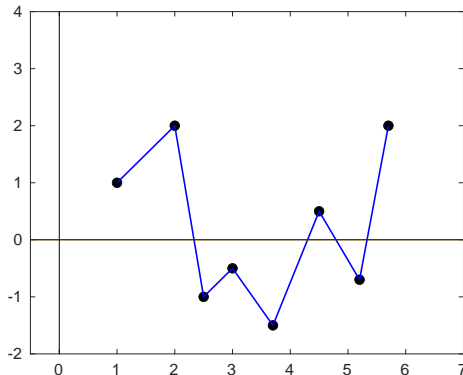


## Nodi di Chebychev



- Il polinomio di interpolazione dipende e si costruisce solo se tutti i punti di interpolazione sono noti a priori.
- Se si hanno a disposizione molti dati, si ottiene un polinomio di grado molto elevato (uno in meno del numero dei dati).
- Spesso nelle applicazioni non è possibile scegliere la distribuzione dei nodi che, nella maggior parte dei casi, è uniforme. Quindi il polinomio di interpolazione può essere soggetto al fenomeno di Runge.

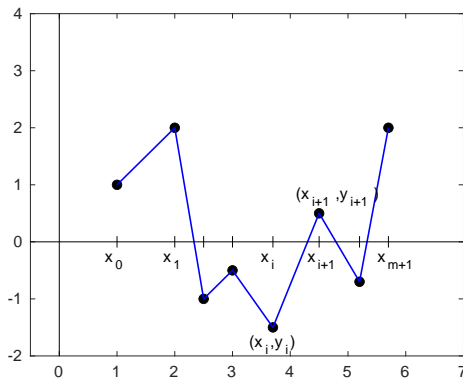
- Invece di costruire una funzione interpolante che sia un globalmente un polinomio, si considerano funzioni polinomiali a tratti, che sono polinomi di un grado fissato se ristrette a ciascun sottointervallo  $[x_i, x_{i+1}]$ .
- Ad esempio, si possono considerare le **funzioni lineari a tratti**.



## Definizione

Fissati  $m + 2$  punti di interpolazione  $(x_i, y_i)$ ,  $i = 0, 1, \dots, m + 1$ , con  $x_i \neq x_j$  per  $i \neq j$ , si definisce la **funzione lineare a tratti  $s(x)$  che interpola i punti dati** come

$$s(x) = y_i + \underbrace{\frac{y_{i+1} - y_i}{x_{i+1} - x_i}}_{s_i(x)} (x - x_i), \quad \forall x \in [x_i, x_{i+1}], \quad i = 0, \dots, m.$$



- Nel caso dell'interpolazione lineare a tratti, se  $y_i = f(x_i)$ ,  $i = 0, \dots, m+1$ , la funzione approssimante è la funzione lineare a tratti  $s(x)$  che interpola i punti  $(x_i, f(x_i))$ ,  $i = 0, \dots, m+1$ .
- Si definisce la funzione resto

$$R_s(x) = f(x) - s(x), \quad \forall x \in [a, b].$$

- Si vuole stabilire sotto quali condizioni  $s(x)$  si può considerare una buona approssimazione di  $f(x)$  in  $[a, b]$  per  $m$  sufficientemente grande, ossia

$$\lim_{m \rightarrow \infty} \|R_s\|_{\infty} = 0.$$



## Teorema

Sia  $f \in C^2([a, b])$  e siano  $a = x_0 < \dots < x_{m+1} = b$ . Indichiamo con  $h$  la quantità

$$h = \max_{i \in \{0, \dots, m\}} (x_{i+1} - x_i).$$

Se  $s(x)$  è la funzione lineare a tratti tale che  $s(x_i) = f(x_i)$ ,  $i = 0, \dots, m+1$ , allora

$$\|f - s\|_{\infty} \leq \frac{\max_{x \in [a, b]} |f''(x)|}{8} h^2.$$

## Dimostrazione

- Consideriamo la funzione resto  $R_s(x)$  su un qualunque intervallo  $[x_i, x_{i+1}]$ ,  $i = 0, \dots, m$ . Su tale intervallo,  $s(x)$  è il polinomio di primo grado interpolante i punti  $(x_i, f(x_i))$  e  $(x_{i+1}, f(x_{i+1}))$  e dunque si può applicare la formula del resto di interpolazione con  $n = 1$ , ottenendo così:

$$f(x) - s_i(x) = \frac{f''(\xi_i)}{2}(x - x_i)(x - x_{i+1}), \quad \forall x \in [a, b], \text{ con } \xi_i \in [x_i, x_{i+1}].$$

- Siccome  $f \in C^2([a, b])$ , si ha che il  $\max_{x \in [a, b]} |f''(x)|$  esiste finito. Allora

$$|f(x) - s_i(x)| \leq \frac{\max_{x \in [a, b]} |f''(x)|}{2} \max_{x \in [x_i, x_{i+1}]} |(x - x_i)(x - x_{i+1})|, \quad \forall x \in [x_i, x_{i+1}].$$

- La funzione  $\varphi(x) = |(x - x_i)(x - x_{i+1})| = (x - x_i)(x_{i+1} - x)$  che sta a destra della disuguaglianza è una parabola con concavità rivolta verso il basso, che ammette un unico punto di massimo in  $x = \frac{x_i + x_{i+1}}{2}$ . Dunque

$$\max_{x \in [x_i, x_{i+1}]} |(x - x_i)(x - x_{i+1})| = \frac{(x_{i+1} - x_i)^2}{4}.$$

- In conclusione, abbiamo provato che per ogni  $i = 0, \dots, m$  vale

$$|f(x) - s_i(x)| \leq \frac{\max_{x \in [a,b]} |f''(x)|}{8} (x_{i+1} - x_i)^2, \quad \forall x \in [x_i, x_{i+1}].$$

- Sia  $h$  l'ampiezza massima tra tutte le ampiezze dei sottointervalli, ovvero:

$$h = \max_{i \in \{0, \dots, m\}} (x_{i+1} - x_i).$$

Allora si ottiene

$$|f(x) - s_i(x)| \leq \frac{\max_{x \in [a,b]} |f''(x)|}{8} h^2, \quad \forall x \in [x_i, x_{i+1}].$$

Siccome questa maggiorazione è indipendente dallo specifico sottointervallo considerato e ricordando la definizione di  $s(x)$ , possiamo concludere che

$$|f(x) - s(x)| \leq \frac{\max_{x \in [a,b]} |f''(x)|}{8} h^2, \quad \forall x \in [a, b],$$

da cui segue la tesi.  $\square$

$$\|f - s\|_{\infty} \leq \frac{\max_{x \in [a, b]} |f''(x)|}{8} h^2.$$

## Osservazione

- Supponiamo che i nodi di interpolazione in  $[a, b]$  siano equidistanti:

$$x_i = a + \frac{b-a}{m+1}i, \quad i = 0, \dots, m+1.$$

In questo caso  $h = (b-a)/(m+1)$  e la maggiorazione sul resto diventa

$$\|f - s\|_{\infty} \leq \frac{(b-a)^2 \max_{x \in [a, b]} |f''(x)|}{8(m+1)^2}$$

$\Downarrow$

$$\lim_{m \rightarrow \infty} \|f - s\|_{\infty} = 0.$$

All'aumentare di  $m$ , l'errore commesso approssimando  $f$  con l'interpolante lineare a tratti  $s(x)$  diventa sempre più piccolo

⇒ **Se  $f \in C^2([a, b])$ , la successione delle funzioni interpolanti lineari a tratti converge sempre, dunque il fenomeno di Runge non si verifica mai.**

## Vantaggi dell'interpolazione lineare a tratti

- Qualunque sia la funzione  $f$  (purché derivabile 2 volte con continuità), la successione delle interpolanti lineari a tratti converge uniformemente ad  $f$ .
- L'interpolante lineare a tratti utilizza dei polinomi di grado basso  
⇒ La complessità computazionale per calcolare i coefficienti dell'interpolante è  $\mathcal{O}(3m)$ , dunque proporzionale al numero dei nodi.

## Svantaggi dell'interpolazione lineare a tratti

- La funzione  $s(x)$  è continua ma non derivabile  
⇒ Si perde una caratteristica importante della funzione  $f$ .

## Definizione

Sia  $[a, b]$  un intervallo e  $x_0, \dots, x_{m+1}$  nodi in  $[a, b]$  tali che  $x_0 = a$  e  $x_{m+1} = b$ . Si definisce **spline di grado  $n$  relativa alla partizione  $x_0, \dots, x_{m+1}$**  ogni funzione  $s : [a, b] \rightarrow \mathbb{R}$  tale che

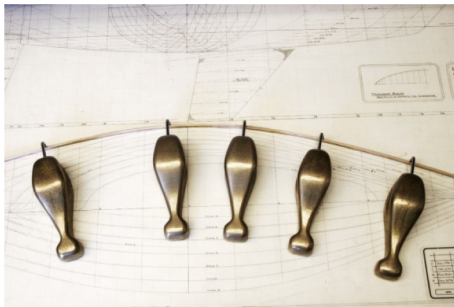
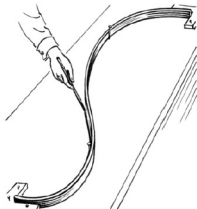
1. la funzione  $s$  ristretta all'intervallo  $[x_i, x_{i+1}]$  è un polinomio  $s_i(x)$  di grado al più  $n$ , per ogni  $i = 0, \dots, m$ ;
2. la funzione  $s$  e tutte le sue derivate fino all'ordine  $n - 1$  sono continue nell'intervallo  $[a, b]$ , ovvero

$$s \in C^{n-1}([a, b]).$$

- Per  $n = 1$ , otteniamo le funzioni lineari a tratti. Per  $n > 1$ , le spline di grado  $n$  sono funzioni polinomiali a tratti le cui derivate fino all'ordine  $n - 1$  esistono e sono continue su tutto l'intervallo  $[a, b]$ .
- Una funzione polinomiale a tratti ristretta ad ogni intervallo  $[x_i, x_{i+1}]$  è già  $C^{n-1}$  essendo un polinomio, ma non è detto che sia  $C^{n-1}$  su tutto  $[a, b]$ .
- La seconda condizione equivale a richiedere che i polinomi  $s_i$  e le loro derivate fino all'ordine  $n - 1$  si raccordino con continuità nei nodi interni:

$$s_i^{(k)}(x_{i+1}) = s_{i+1}^{(k)}(x_{i+1}), \quad k = 0, \dots, n - 1, \quad i = 0, \dots, m - 1.$$

- Il termine inglese spline (= flessibile) deriva dal nome di un particolare strumento di disegno formato da una striscia di metallo o di legno a cui veniva data la forma desiderata per mezzo di grossi pesi ai quali veniva ancorata. Serviva per realizzare disegni che non era possibile ottenere con riga e compasso, in scala anche molto grande, per esempio per la costruzione di navi.
- Il termine spline per indicare una funzione polinomiale a tratti molto regolare fu utilizzato per la prima volta nel 1946 da Schoenberg.
- Un tipo particolare di funzioni spline, dette B-spline, è alla base dei software per il disegno tecnico (CAD) e della grafica vettoriale.



## Casi particolari

- Le funzioni lineari a tratti coincidono con le splines di grado 1 ( $n = 1$ ) e si dicono anche **spline lineari**.
- L'altro caso di particolare interesse è quello delle splines cubiche ( $n = 3$ ). In questo caso le condizioni di regolarità da imporre sulla spline sono

$$s_i(x_{i+1}) = s_{i+1}(x_{i+1})$$

$$s'_i(x_{i+1}) = s'_{i+1}(x_{i+1})$$

$$s''_i(x_{i+1}) = s''_{i+1}(x_{i+1})$$

per  $i = 0, \dots, m - 1$ .



La classe delle funzioni splines di grado  $n$  forma un sottospazio vettoriale. Per individuarne la dimensione, contiamo da quanti parametri liberi dipende una generica spline di grado  $n$  relativa ad una partizione di  $m + 2$  punti.

- Ogni “pezzo” di spline  $s_i$  è un polinomio di grado  $n$ ;

$$s_i(x) = a_{0i} + a_{1i}x + \dots + a_{ni}x^n, \quad i = 0, \dots, m$$

$\Rightarrow$  ogni  $s_i$  dipende da  $n + 1$  parametri (i suoi coefficienti).

- In tutto ci sono  $m + 1$  “pezzi”, dunque  $s$  dipende da  $(n + 1)(m + 1)$  parametri.
- Le condizioni di regolarità

$$s_i^{(k)}(x_{i+1}) = s_{i+1}^{(k)}(x_{i+1}), \quad k = 0, \dots, n - 1, \quad i = 0, \dots, m - 1.$$

sono  $mn$  equazioni che gli  $(n + 1)(m + 1)$  parametri di  $s$  devono soddisfare.

## Gradi di libertà delle splines di grado $n$

Una spline di grado  $n$  dipende da  $(n + 1)(m + 1)$  parametri, di cui

$$(n + 1)(m + 1) - mn = n + m + 1$$

sono liberi. Dunque **la dimensione dello spazio delle splines di grado  $n$  relative a  $m + 2$  punti è pari a  $n + m + 1$ .**

## Definizione

Dati  $(x_i, y_i)$ ,  $i = 0, 1, \dots, m+1$ , con  $x_i \neq x_j$  per  $i \neq j$ , il problema dell'**interpolazione mediante splines di grado  $n$**  consiste nel trovare una spline  $s$  di grado  $n$  relativa ai punti  $x_0, \dots, x_{m+1}$  tale che

$$s(x_i) = y_i, \quad i = 0, \dots, m+1.$$

- Le condizioni di interpolazione sono  $m+2$ , pertanto rendono fissi (non più liberi) altrettanti parametri. I gradi di libertà rimanenti sono dunque

$$n + m + 1 - (m + 2) = n - 1.$$

- Nel caso lineare ( $n = 1$ ), una generica spline dipende da  $n + m + 1 = m + 2$  parametri, tanti quante sono le condizioni di interpolazione  
 $\Rightarrow n - 1 = 0$  e la spline lineare interpolante non ha parametri liberi  
 $\Rightarrow$  **la spline lineare interpolante è unica.**
- Per  $n > 1$  non esiste un'unica spline di interpolazione di grado  $n$ . Fissati i punti di interpolazione, esistono infinite spline di grado  $n$  che li interpolano, al variare di  $n - 1$  parametri liberi.

## Problema (costruzione delle splines cubiche di interpolazione)

Dati  $(x_i, y_i)$ ,  $i = 0, \dots, m+1$ , con  $x_i \neq x_j$  per  $i \neq j$ , costruire una spline cubica di grado  $n = 3$  relativa ai nodi  $x_0, \dots, x_{m+1}$  che interpoli i punti  $(x_i, y_i)$  dati.

- Si vogliono calcolare i coefficienti degli  $m+1$  polinomi di grado 3

$$s_i(x) = \alpha_i + \beta_i(x - x_i) + \gamma_i(x - x_i)^2 + \delta_i(x - x_i)^3, \quad i = 0, \dots, m$$

che soddisfano le condizioni di regolarità e le condizioni di interpolazione

$$\begin{cases} s_i(x_{i+1}) = s_{i+1}(x_{i+1}), & i = 0, \dots, m-1 \\ s'_i(x_{i+1}) = s'_{i+1}(x_{i+1}), & i = 0, \dots, m-1 \\ s''_i(x_{i+1}) = s''_{i+1}(x_{i+1}), & i = 0, \dots, m-1 \\ s(x_i) = y_i, & i = 0, \dots, m+1. \end{cases}$$

Sistema lineare nelle  
 $\alpha_i, \beta_i, \gamma_i, \delta_i$

- Calcolare una spline cubica di interpolazione significa quindi calcolare  $\alpha_i, \beta_i, \gamma_i, \delta_i$ , per ogni  $i = 0, \dots, m$ , in funzione dei dati  $(x_i, y_i)$ , utilizzando le condizioni di regolarità e le condizioni di interpolazione.

- Scriviamo la funzione  $s_i(x)$  e calcoliamone le sue derivate prima  $s'_i(x)$  e seconda  $s''_i(x)$ :

$$s_i(x) = \alpha_i + \beta_i(x - x_i) + \gamma_i(x - x_i)^2 + \delta_i(x - x_i)^3$$

$$s'_i(x) = \beta_i + 2\gamma_i(x - x_i) + 3\delta_i(x - x_i)^2$$

$$s''_i(x) = 2\gamma_i + 6\delta_i(x - x_i).$$

- La funzione  $s_i$  e le sue derivate nel nodo  $x_i$  valgono:

$$\begin{cases} s_i(x_i) = \alpha_i \\ s'_i(x_i) = \beta_i \\ s''_i(x_i) = 2\gamma_i \end{cases}, \quad i = 0, \dots, m.$$

- Sia  $h_i = x_{i+1} - x_i$ ,  $i = 0, \dots, m$  l'ampiezza di ogni sottointervallo. Allora la funzione  $s_i$  e le sue derivate nel nodo  $x_{i+1}$  valgono

$$\begin{cases} s_i(x_{i+1}) = \alpha_i + \beta_i h_i + \gamma_i h_i^2 + \delta_i h_i^3 \\ s'_i(x_{i+1}) = \beta_i + 2\gamma_i h_i + 3\delta_i h_i^2 \\ s''_i(x_{i+1}) = 2\gamma_i + 6\delta_i h_i \end{cases}, \quad i = 0, \dots, m.$$

- Ricordiamo che il sistema delle condizioni di regolarità e delle condizioni di interpolazione è dato da

$$\begin{cases} s_i(x_{i+1}) = s_{i+1}(x_{i+1}), & i = 0, \dots, m-1 \\ s'_i(x_{i+1}) = s'_{i+1}(x_{i+1}), & i = 0, \dots, m-1 \\ s''_i(x_{i+1}) = s''_{i+1}(x_{i+1}), & i = 0, \dots, m-1 \\ s(x_i) = y_i, & i = 0, \dots, m+1. \end{cases}.$$

- Sostituendo nelle equazioni del sistema l'ampiezza  $h_i = x_{i+1} - x_i$  e i valori delle funzioni  $s_i, s'_i, s''_i$  nei nodi  $x_i, x_{i+1}$ , e chiamando  $\alpha_{m+1} = s_m(x_{m+1})$ , otteniamo

$$\begin{cases} \alpha_i + \beta_i h_i + \gamma_i h_i^2 + \delta_i h_i^3 = \alpha_{i+1}, & i = 0, \dots, m-1 \\ \beta_i + 2\gamma_i h_i + 3\delta_i h_i^2 = \beta_{i+1}, & i = 0, \dots, m-1 \\ 2\gamma_i + 6\delta_i h_i = 2\gamma_{i+1}, & i = 0, \dots, m-1 \\ \alpha_i = y_i, & i = 0, \dots, m+1. \end{cases}$$

Notare che i coefficienti  $\alpha_i$  sono forniti dalle uguaglianze  $\alpha_i = y_i$ .

- Introducendo il nuovo parametro  $\beta_{m+1} = s'(x_{m+1})$  e sostituendo le equazioni  $\alpha_i = y_i$  nelle prime equazioni del sistema, troviamo

$$\begin{cases} \beta_i h_i + \gamma_i h_i^2 + \delta_i h_i^3 = y_{i+1} - y_i, & i = 0, \dots, m \\ \beta_i + 2\gamma_i h_i + 3\delta_i h_i^2 = \beta_{i+1}, & i = 0, \dots, m \\ \gamma_i + 3\delta_i h_i = \gamma_{i+1}, & i = 0, \dots, m-1 \end{cases} .$$

$$\begin{cases} \beta_i h_i + \gamma_i h_i^2 + \delta_i h_i^3 = y_{i+1} - y_i, & i = 0, \dots, m \\ \beta_i + 2\gamma_i h_i + 3\delta_i h_i^2 = \beta_{i+1}, & i = 0, \dots, m \\ \gamma_i + 3\delta_i h_i = \gamma_{i+1}, & i = 0, \dots, m-1 \end{cases}.$$

- Dalla prime **prime due equazioni** si ricavano  $\gamma_i, \delta_i$  in funzione di  $\beta_i, \beta_{i+1}$ :

$$\begin{cases} \gamma_i h_i^2 + \delta_i h_i^3 &= y_{i+1} - y_i - \beta_i h_i \\ 2\gamma_i h_i + 3\delta_i h_i^2 &= \beta_{i+1} - \beta_i \end{cases} \Rightarrow \begin{cases} \gamma_i h_i + \delta_i h_i^2 &= \frac{y_{i+1} - y_i}{h_i} - \beta_i \\ 2\gamma_i h_i + 3\delta_i h_i^2 &= \beta_{i+1} - \beta_i \end{cases}$$

$$\Rightarrow \begin{cases} \delta_i &= \frac{1}{h_i^2} \left( \beta_{i+1} + \beta_i - 2 \frac{y_{i+1} - y_i}{h_i} \right) \\ \gamma_i &= \frac{1}{h_i} \left( 3 \frac{y_{i+1} - y_i}{h_i} - (\beta_{i+1} + 2\beta_i) \right) \end{cases} \quad i = 0, \dots, m$$

- Infine si sostituiscono le espressioni di  $\delta_i, \gamma_i$  nella **terza**:

$$\begin{aligned} \frac{1}{h_i} \left( 3 \frac{y_{i+1} - y_i}{h_i} - (\beta_{i+1} + 2\beta_i) \right) + 3 \frac{1}{h_i^2} \left( \beta_{i+1} + \beta_i - 2 \frac{y_{i+1} - y_i}{h_i} \right) h_i &= \\ = \frac{1}{h_{i+1}} \left( 3 \frac{y_{i+2} - y_{i+1}}{h_{i+1}} - 2(\beta_{i+2} + \beta_{i+1}) \right) &= \\ i = 0, \dots, m-1 \end{aligned}$$

$$\begin{aligned} \frac{1}{h_i} \left( 3 \frac{y_{i+1} - y_i}{h_i} - (\beta_{i+1} + 2\beta_i) \right) + 3 \frac{1}{h_i^2} \left( \beta_{i+1} + \beta_i - 2 \frac{y_{i+1} - y_i}{h_i} \right) h_i &= \\ &= \frac{1}{h_{i+1}} \left( 3 \frac{y_{i+2} - y_{i+1}}{h_{i+1}} - 2(\beta_{i+2} + \beta_{i+1}) \right) \end{aligned}$$

$$i = 0, \dots, m-1$$

- Portando a sinistra tutti i termini che contengono le incognite  $\beta_i, \beta_{i+1}$  e a destra i termini noti, e moltiplicando tutto per  $h_i h_{i+1}$  si ottiene

$$h_{i+1}\beta_i + 2(h_i + h_{i+1})\beta_{i+1} + h_i\beta_{i+2} = 3h_{i+1} \left( \frac{y_{i+1} - y_i}{h_i} \right) + 3h_i \left( \frac{y_{i+2} - y_{i+1}}{h_{i+1}} \right)$$

per  $i = 0, \dots, m-1$ .

- Si tratta di un sistema di  $m$  equazioni lineari nelle  $m+2$  incognite  $\beta_i$ :

$$\begin{cases} h_1\beta_0 & +2(h_0 + h_1)\beta_1 & +h_0\beta_2 & & & = b_0 \\ h_2\beta_1 & & +2(h_1 + h_2)\beta_2 & +h_1\beta_3 & & = b_1 \\ & & \ddots & & & \\ & & h_m\beta_{m-1} & +2(h_{m-1} + h_m)\beta_m & +h_{m-1}\beta_{m+1} & = b_{m-1} \end{cases}$$

$$b_i = 3h_{i+1} \left( \frac{y_{i+1} - y_i}{h_i} \right) + 3h_i \left( \frac{y_{i+2} - y_{i+1}}{h_{i+1}} \right), i = 0, \dots, m-1$$



- Non c'è un'unica spline cubica di interpolazione, ma ne esiste una famiglia infinita che dipende da due parametri.
- Per individuare un'unica spline nella famiglia, ci si restringe ad un sottoinsieme, aggiungendo due condizioni, per esaurire i gradi di libertà.

$$h_{i+1}\beta_i + 2(h_i + h_{i+1})\beta_{i+1} + h_i\beta_{i+2} = 3h_{i+1} \left( \frac{y_{i+1} - y_i}{h_i} \right) + 3h_i \left( \frac{y_{i+2} - y_{i+1}}{h_{i+1}} \right)$$

- Si assegnano i valori  $\beta_0, \beta_{m+1}$  che corrispondono alla derivata prima della spline negli estremi dell'intervallo di interpolazione.
- Si riscrive il sistema facendo diventare i valori assegnati  $\beta_0, \beta_{m+1}$  dei termini noti.
- Dunque il sistema diventa

$$Tz = c, \quad \text{dove}$$

$$T = \begin{pmatrix} 2(h_0 + h_1) & h_0 & & & \\ h_2 & 2(h_1 + h_2) & h_1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & h_{m-2} \\ & & & h_m & 2(h_{m-1} + h_m) \end{pmatrix}$$

$$c = \begin{pmatrix} b_0 - h_1\beta_0 \\ b_1 \\ \vdots \\ b_{m-1} - h_{m-1}\beta_{m+1} \end{pmatrix}, \quad z = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}$$

$$T = \begin{pmatrix} 2(h_0 + h_1) & h_0 & & & \\ h_2 & 2(h_1 + h_2) & h_1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & h_{m-2} \\ & & & h_m & 2(h_{m-1} + h_m) \end{pmatrix}$$

- La matrice  $T \in \mathbb{R}^{m \times m}$  è strettamente diagonale dominante per righe.
- Il sistema  $Tz = c$  ha dunque un'unica soluzione.
- Dopo aver calcolato la soluzione del sistema, ottenendo i  $\beta_i$ , per sostituzione si ottengono anche i  $\gamma_i$  e i  $\delta_i$ .
- In questo modo si calcola l'unica spline cubica vincolata che interpola i punti dati.

## Esempi di splines cubiche interpolanti vincolate

- Spline cubica vincolata: si ottiene ponendo  $s'(x_0) = \beta_0$  e  $s'(x_{m+1}) = \beta_{m+1}$  (caso visto nelle slides precedenti).
- Spline cubica naturale: si ottiene ponendo  $s''(x_0) = s''(x_{m+1}) = 0$ .
- Spline cubica periodica: si utilizza solo quando  $y_0 = y_{m+1}$  e si impongono le condizioni  $s'(x_0) = s'(x_{m+1})$  e  $s''(x_0) = s''(x_{m+1})$ .
- Spline cubica not-a-knot Matlab: è un polinomio di grado 3 se ristretta agli intervalli  $[x_0, x_2]$  e  $[x_{m-1}, x_{m+1}]$ .

In tutti i precedenti casi, fissati i punti di interpolazione, esiste ed è unica la spline interpolante.

## Teorema (convergenza delle splines cubiche e delle loro derivate prime)

Sia  $s : [a, b] \rightarrow \mathbb{R}$  la spline cubica vincolata, naturale o periodica che interpola i punti  $(x_i, y_i)$ ,  $i = 0, \dots, m+1$ , con  $x_i \in [a, b]$ ,  $x_i \neq x_j$  per  $i \neq j$  e  $y_i = f(x_i)$ . Supponiamo che  $f \in C^2([a, b])$ . Allora per ogni  $x \in [a, b]$  si ha

$$|f(x) - s(x)| \leq h^{\frac{3}{2}} \left( \int_a^b (f''(x))^2 dx \right)^{\frac{1}{2}}$$
$$|f'(x) - s'(x)| \leq h^{\frac{1}{2}} \left( \int_a^b (f''(x))^2 dx \right)^{\frac{1}{2}},$$

dove  $h = \max_{i \in \{0, \dots, m\}} (x_{i+1} - x_i)$ .

- Le splines cubiche di interpolazione approssimano sempre meglio la funzione interpolata al diminuire dell'ampiezza dei sottointervalli  $[x_i, x_{i+1}]$  (non si verifica il fenomeno di Runge).
- La derivata prima della spline cubica di interpolazione approssima sempre meglio la derivata prima della funzione interpolata al diminuire dell'ampiezza dei sottointervalli  $[x_i, x_{i+1}]$ .

## Teorema (curvatura delle splines cubiche)

Tra tutte le funzioni  $g \in C^2([a, b])$  che interpolano i punti  $(x_i, y_i)$ ,  $i = 0, \dots, m+1$ , con  $x_i \neq x_j$  per  $i \neq j$ , e che soddisfano una delle seguenti condizioni:

1.  $g'(x_0) = \beta_0$ ,  $g'(x_{m+1}) = \beta_{m+1}$  con  $\beta_0, \beta_{m+1}$  valori fissati;
2.  $g''(x_0) = g''(x_{m+1}) = 0$ ;
3. (se  $y_0 = y_{m+1}$ )  $g'(x_0) = g'(x_{m+1})$  e  $g''(x_0) = g''(x_{m+1})$ ;

la spline cubica di interpolazione vincolata (o naturale o periodica) è quella per cui vale la seguente proprietà di minimo:

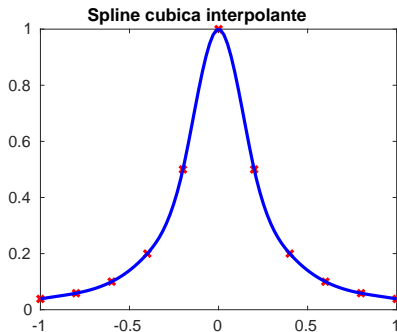
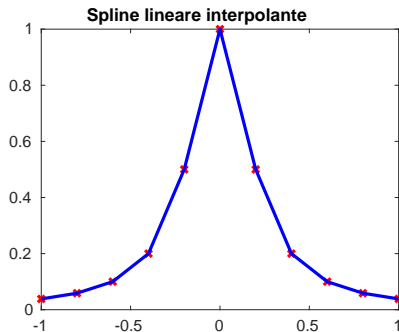
$$\int_a^b (s''(x))^2 dx \leq \int_a^b (g''(x))^2 dx$$

e l'uguaglianza vale se e solo se  $g = s$ .

- In parole povere: tra tutte le funzioni che interpolano gli stessi punti, le splines cubiche sono quelle il cui grafico presenta le curve “più dolci”, senza bruschi cambiamenti di pendenza.

## Esempio: interpolazione della funzione di Runge mediante splines

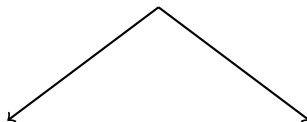
$$s(x) \approx f(x) = \frac{1}{1 + 25x^2}.$$



## 2. Approssimazione ai minimi quadrati



## Approssimazione



### Interpolazione

Se i dati non sono affetti da errore

### Approssimazione ai minimi quadrati

Se i dati sono affetti da errore

- Nelle applicazioni, le tecniche di approssimazione dei dati servono per definire il **modello matematico** che descrive un fenomeno osservabile.
- Ad esempio, nella Fisica, si vuole spesso individuare i **parametri** di una formula che lega due quantità  $x, y$  fisiche, a partire da un numero finito di misurazioni  $(x_i, y_i)$  effettuate.
- Siccome le misurazioni sono sperimentali, i dati  $(x_i, y_i)$  non sono esatti, bensì risultano affetti da errori. **Interpolare i dati con errore non fornisce un modello significativo del fenomeno osservato.**

## Esempio: la legge di Ohm

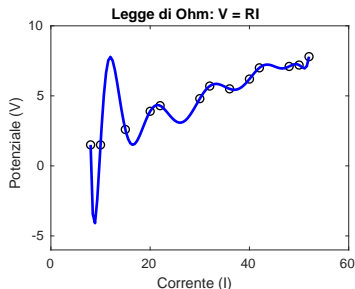
In un filo elettrico, la differenza di potenziale  $V$  ai due capi del filo e l'intensità di corrente  $I$  generata nel filo sono legate dalla **legge di Ohm**:

$$V = RI$$

dove  $R$  è la resistenza del filo.

Date le misurazioni  $(I_i, V_i)$ ,  $i = 1, \dots, m$ , calcolare  $R$ .

- La legge di Ohm è un **modello lineare**.
- Nella pratica, le misurazioni sono effettuate con un voltmetro.
  - ⇒ Le misurazioni contengono degli errori sperimentali (casuali o sistematici).
  - ⇒ I punti  $(I_i, V_i)$  non sono perfettamente allineati.
  - ⇒ Se interpoliamo i punti, **otteniamo un modello che non è più lineare!**



- Nell'esempio relativo alla legge di Ohm, è noto che il modello che descrive il fenomeno è di tipo lineare.
- Tra tutte le rette di equazione

$$y = a_0 + a_1x$$

si vuole determinare quella che si “adatta” meglio all'andamento dei dati.

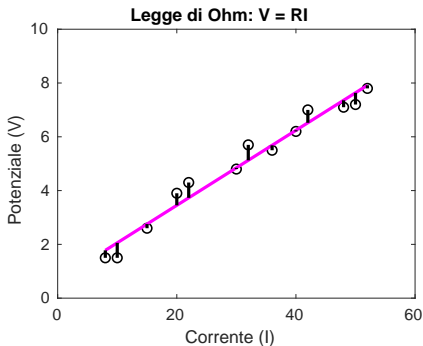
- L'approccio dell'**approssimazione ai minimi quadrati** consiste nel calcolare i coefficienti  $a_0, a_1$  in modo che la distanza della retta  $y = a_0 + a_1x$  valutata negli  $x_i, i = 1, \dots, m$ , dai valori misurati  $y_i, i = 1, \dots, m$ , sia minimo.

## Definizione

Dati  $(x_i, y_i)$ ,  $i = 1, \dots, m$ , si dice **retta ai minimi quadrati (o retta di regressione)** quella retta  $y = a_0 + a_1x$  i cui coefficienti  $a_0, a_1$  minimizzano la funzione

$$Q : \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$Q(a_0, a_1) = \sum_{i=1}^m (a_0 + a_1 x_i - y_i)^2.$$



- Introducendo i vettori

$$q(a_0, a_1) = \begin{pmatrix} a_0 + a_1 x_1 \\ a_0 + a_1 x_2 \\ \vdots \\ a_0 + a_1 x_m \end{pmatrix} \in \mathbb{R}^m, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix},$$

possiamo riscrivere l'espressione della funzione  $Q$  come segue:

$$Q(a_0, a_1) = \sum_{i=1}^m (a_0 + a_1 x_i - y_i)^2 \Rightarrow Q(a_0, a_1) = \|q(a_0, a_1) - y\|_2^2.$$

- Definendo inoltre

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \in \mathbb{R}^{m \times 2}, \quad \alpha = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \in \mathbb{R}^2,$$

si ha  $q(a_0, a_1) = A\alpha$  e dunque si riscrive ulteriormente la funzione  $Q$  come

$$Q(\alpha) = \|A\alpha - y\|_2^2.$$

- Il calcolo della retta ai minimi quadrati è stato dunque ricondotto alla soluzione del seguente problema di minimo:

$$\min_{\alpha \in \mathbb{R}^2} \|A\alpha - y\|_2^2,$$

dove

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \in \mathbb{R}^{m \times 2}.$$

Detto  $\alpha^* = (a_0^*, a_1^*)$  il vettore soluzione del problema, la retta ai minimi quadrati ha equazione  $y = a_0^* + a_1^*x$ .

- Tale problema di minimo prende il nome di **problema ai minimi quadrati**.

- L'approssimazione ai minimi quadrati porta ad individuare i due parametri di una retta

$$f(a_0, a_1; x) = a_0 + a_1 x$$

la cui “distanza cumulativa” dai dati è minima.

- In generale, la retta ai minimi quadrati non passa per i punti dati.
- Il tipo di modello scelto (in questo caso la retta) è fissato a priori, solitamente in base alla conoscenza delle caratteristiche del fenomeno che si vuole descrivere.
- L'approssimazione ai minimi quadrati si estende a modelli più generali.



## Definizione

Dati  $(x_i, y_i)$ ,  $i = 1, \dots, m$ , si dice **polinomio ai minimi quadrati di grado  $n - 1$**  quel polinomio  $f(a_0, a_1, \dots, a_{n-1}; x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1}$  i cui coefficienti  $a_0, a_1, \dots, a_{n-1}$  minimizzano la funzione

$$Q : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$Q(a_0, a_1, \dots, a_{n-1}) = \sum_{i=1}^m (a_0 + a_1x_i + \dots + a_{n-1}x_i^{n-1} - y_i)^2.$$

- Definendo

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & & & \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} \end{pmatrix}, \quad \alpha = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} \in \mathbb{R}^n$$

si ha  $Q(a_0, a_1, \dots, a_{n-1}) = \|A\alpha - y\|_2^2$  e il problema da risolvere è sempre ai minimi quadrati:

$$\min_{\alpha \in \mathbb{R}^n} \|A\alpha - y\|_2^2.$$

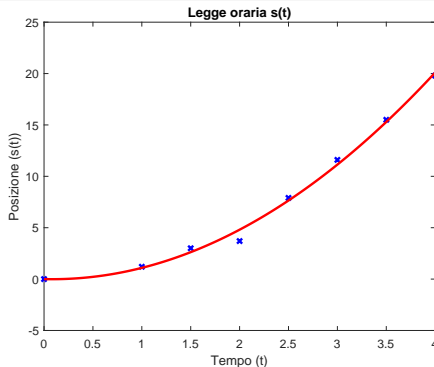
- Se  $m = n$ , il polinomio ai minimi quadrati coincide con quello interpolante.

## Esempio: legge oraria del moto uniformemente accelerato

La legge oraria di un corpo che si muove di moto rettilineo uniformemente accelerato con accelerazione  $a$  è del tipo

$$s = \frac{1}{2}at^2 + bt + c.$$

Date le misurazioni  $(t_i, s_i)$ ,  $i = 1, \dots, m$ , della posizione in funzione del tempo, calcolare  $a, b, c$ .



## Definizione

Dati  $(x_i, y_i)$ ,  $i = 1, \dots, m$ , supponiamo di aver scelto un modello con  $n$  parametri dato dalla combinazione lineare di  $n$  funzioni  $\phi_0, \phi_1, \dots, \phi_{n-1}$  formanti una base, ovvero:

$$f(a_0, a_1, \dots, a_{n-1}; x) = a_0\phi_0(x) + a_1\phi_1(x) + \dots + a_{n-1}\phi_{n-1}(x).$$

Si dice che **il modello  $f$  è ai minimi quadrati** se i suoi coefficienti  $a_0, a_1, \dots, a_{n-1}$  minimizzano la funzione

$$Q : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$Q(a_0, a_1, \dots, a_{n-1}) = \sum_{i=1}^m \underbrace{(a_0\phi_0(x_i) + a_1\phi_1(x_i) + \dots + a_{n-1}\phi_{n-1}(x_i) - y_i)^2}_{f(x_i)}.$$

- Definendo

$$A = \begin{pmatrix} \phi_0(x_1) & \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_{n-1}(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_{n-1}(x_2) \\ \vdots & \vdots & & & \\ \phi_0(x_m) & \phi_1(x_m) & \phi_2(x_m) & \cdots & \phi_{n-1}(x_m) \end{pmatrix}, \alpha = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} \in \mathbb{R}^n$$

si ha  $Q(a_0, a_1, \dots, a_{n-1}) = \|A\alpha - y\|_2^2$  e il problema da risolvere è sempre ai minimi quadrati:

$$\min_{\alpha \in \mathbb{R}^n} \|A\alpha - y\|_2^2.$$

Dati  $(x_i, y_i)$ ,  $i = 1, \dots, m$  e fissato il modello

$$f(a_0, a_1, \dots, a_{n-1}; x) = a_0 \phi_0(x) + a_1 \phi_1(x) + \dots + a_{n-1} \phi_{n-1}(x),$$

occorre risolvere il problema

$$\min_{\alpha \in \mathbb{R}^n} \|A\alpha - y\|^2.$$

- Si può affrontare usando opportune fattorizzazioni della matrice  $A \in \mathbb{R}^{m \times n}$ .
- Distinguiamo due casi:
  1. **caso non degenero:**  $A \in \mathbb{R}^{m \times n}$ , con  $n \leq m$ , ha rango  $n$  (tutte le colonne di  $A$  sono linearmente indipendenti);  
N.B. la regressione polinomiale rientra in questo caso;
  2. **caso generale:**  $A \in \mathbb{R}^{m \times n}$ , con  $n \leq m$ , ha rango  $k < n$  (non tutte le colonne di  $A$  sono linearmente indipendenti).

## Teorema

Se  $A \in \mathbb{R}^{m \times n}$  con  $n \leq m$  ha rango  $n$ , allora la soluzione del problema

$$\min_{\alpha \in \mathbb{R}^n} \|A\alpha - y\|^2$$

esiste ed è unica.

## Dimostrazione

Si inizia osservando che:

- se  $Q$  è una matrice ortogonale, allora  $\|Qx\|^2 = x^T Q^T Q x = x^T x = \|x\|^2$ ;
- se  $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{R}^m$  con  $z_1 \in \mathbb{R}^n$ ,  $z_2 \in \mathbb{R}^{m-n}$ , allora  $\|z\|^2 = \|z_1\|^2 + \|z_2\|^2$ ;
- Se  $A \in \mathbb{R}^{m \times n}$  con  $n \leq m$  ha rango  $n$ , allora esistono  $Q \in \mathbb{R}^{m \times m}$  ortogonale e  $R \in \mathbb{R}^{n \times n}$  triangolare superiore non singolare tali che

$$A = Q \begin{pmatrix} R \\ 0 \end{pmatrix} \Leftrightarrow Q^T A = \begin{pmatrix} R \\ 0 \end{pmatrix};$$

questa è la **fattorizzazione  $QR$  per matrici non quadrate aventi rango massimo per colonne**.

Dalle precedenti osservazioni, si ottiene che

$$\|A\alpha - y\|^2 = \|Q^T A\alpha - \underbrace{Q^T y}_{=\tilde{y}}\|^2 = \left\| \begin{pmatrix} R \\ 0 \end{pmatrix} \alpha - \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix} \right\|^2 = \|R\alpha - \tilde{y}_1\|^2 + \|\tilde{y}_2\|^2.$$

Il minimo di  $\|A\alpha - y\|^2$  si ha per  $R\alpha = \tilde{y}_1$ . Essendo  $R$  invertibile, la soluzione esiste ed è unica.  $\square$

La dimostrazione del teorema precedente fornisce un algoritmo per calcolare la soluzione del problema ai minimi quadrati

$$\min_{\alpha \in \mathbb{R}^n} \|A\alpha - y\|^2$$

quando  $A \in \mathbb{R}^{m \times n}$  con  $n \leq m$  ha rango  $n$ .

Algoritmo per il calcolo della soluzione del problema ai minimi quadrati (caso non degenero)

1. Calcolare la fattorizzazione

$$A = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$$

mediante trasformazioni di Householder.

2. Calcolare  $\tilde{y} = Q^T y$ .
3. Estrarre le prime  $n$  componenti di  $\tilde{y}$  e memorizzarle nel vettore  $\tilde{y}_1$ .
4. Risolvere il sistema triangolare superiore  $R\alpha = \tilde{y}_1$ .



- Se le colonne di  $A$  non sono linearmente indipendenti, ossia se il rango di  $A$  è  $k$  con  $k < n$ , allora il problema dei minimi quadrati ha infinite soluzioni, che dipendono da  $n - k$  parametri liberi.
- Nel sottospazio delle soluzioni, si individua quella che ha norma euclidea più piccola, detta **soluzione di norma minima**.
- Per la dimostrazione dell'esistenza e unicità della soluzione di norma minima, la quale definisce anche l'algoritmo per calcolarla, si utilizza una diversa fattorizzazione della matrice  $A$ , detta **decomposizione ai valori singolari (SVD)**.

## Teorema (Decomposizione ai valori singolari)

Sia  $A \in \mathbb{R}^{m \times n}$  una matrice di rango  $k$ . Allora  $A$  si può fattorizzare nella forma

$$A = U\Sigma V^T$$

dove

- $U \in \mathbb{R}^{m \times m}$ ,  $V \in \mathbb{R}^{n \times n}$  sono matrici ortogonali;
- $\Sigma \in \mathbb{R}^{m \times n}$  è una matrice rettangolare diagonale della forma

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_k & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}$$

con  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > 0$ .

I valori  $\sigma_i$ ,  $i = 1, \dots, k$ , sono detti i **valori singolari di  $A$** . La fattorizzazione  $A = U\Sigma V^T$  prende il nome di **decomposizione ai valori singolari (SVD) di  $A$** .

## Teorema

Se  $A \in \mathbb{R}^{m \times n}$  con  $n \leq m$  ha rango  $k \leq n$ , allora la soluzione di minima norma del problema

$$\min_{\alpha \in \mathbb{R}^n} \|A\alpha - y\|^2$$

esiste ed è unica.

## Dimostrazione

- Sostituendo  $A = U\Sigma V^T$  dentro al termine dei minimi quadrati, si ha

$$\|A\alpha - y\|^2 = \|U\Sigma V^T\alpha - y\|^2 = \|\Sigma \underbrace{V^T\alpha}_{\gamma} - \underbrace{U^Ty}_z\|^2 = \left\| \begin{pmatrix} \sigma_1\gamma_1 - z_1 \\ \vdots \\ \sigma_k\gamma_k - z_k \\ -z_{k+1} \\ \vdots \\ -z_m \end{pmatrix} \right\|^2.$$

Le infinite soluzioni si trovano fissando  $\gamma_i = z_i/\sigma_i$ ,  $i = 1, \dots, k$ , al variare di  $\gamma_{k+1}, \dots, \gamma_n \in \mathbb{R}^{n-k}$ .

- Siccome  $\gamma = V^T\alpha$  e  $V$  è ortogonale, si ha che

$$\alpha = V\gamma = V \begin{pmatrix} z_1/\sigma_1 \\ \vdots \\ z_k/\sigma_k \\ \gamma_{k+1} \\ \vdots \\ \gamma_n \end{pmatrix} \Rightarrow \|\alpha\|^2 = \|\gamma\|^2 = \frac{z_1^2}{\sigma_1^2} + \dots + \frac{z_k^2}{\sigma_k^2} + \gamma_{k+1}^2 + \dots + \gamma_n^2.$$

Dunque la soluzione di minima norma esiste unica e si ottiene ponendo

$$\gamma_{k+1} = \dots = \gamma_n = 0. \quad \square$$

## Soluzione del problema ai minimi quadrati (caso degenere)

$$\alpha = V\gamma = \begin{pmatrix} z_1/\sigma_1 \\ \vdots \\ z_k/\sigma_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

### Algoritmo per il calcolo della soluzione del problema ai minimi quadrati (caso degenere)

1. Calcolare la fattorizzazione  $A = U\Sigma V^T$ .

2. Calcolare  $z = U^T y$ .

3. Calcolare  $\gamma = \begin{pmatrix} z_1/\sigma_1 \\ \vdots \\ z_k/\sigma_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$

4. Calcolare  $\alpha = V\gamma$ .

## Osservazioni

- Il calcolo della SVD è molto costoso.
- Esistono diversi algoritmi per calcolare la SVD con complessità differenti. In generale procedono in due fasi: prima si riduce la matrice in una forma più semplice (bidiagonale), poi si procede con metodi iterativi.
- Nel caso non degenerare, è più conveniente utilizzare la  $QR$ , al costo di  $n^2(3m - n)/3$ .

- Benché costosa, la fattorizzazione SVD

$$A = U\Sigma V^T$$

è utilizzata in tantissime applicazioni del mondo reale, poiché fornisce uno strumento per approssimare le matrici.

- Infatti, sia  $u_i, i = 1, \dots, m$  la  $i$ -esima colonna di  $U$  e con  $v_i, i = 1, \dots, n$  la  $i$ -esima colonna di  $V$ . Dalla definizione di prodotto matriciale si ottiene

$$A = \sigma_1 u_1 v_1^T + \dots + \sigma_k u_k v_k^T$$

dove i termini  $u_i v_i^T$  sono matrici di dimensione  $m \times n$ .

Dunque  $A$  si scrive come somma pesata di matrici i cui coefficienti (i valori singolari) sono disposti in ordine decrescente.

## Esempio: regolarizzazione di sistemi lineari mal condizionati

Data  $A \in \mathbb{R}^{n \times n}$  non singolare, si vuole risolvere un sistema lineare

$$Ax = b.$$

- Si dimostra che il numero di condizionamento di  $A$  in norma 2 è

$$\kappa(A) = \frac{\sigma_1}{\sigma_n}.$$

- Nei casi di mal condizionamento si ha  $\kappa(A) \gg 1$ , ovvero  $\sigma_1 \gg \sigma_n$ .
- Una strategia per ottenere una matrice meno mal condizionata, ma non troppo diversa da  $A$ , consiste nell'individuare un  $\tilde{k} < n$ , definire

$$\tilde{A} = \sigma u_1 v_1^T + \dots + \sigma_{\tilde{k}} u_{\tilde{k}} v_{\tilde{k}}^T$$

e risolvere il sistema  $\tilde{A}x = b$ .