Interpolazione e approssimazione di funzioni

Note Title 2021-11-22

Ripasso: sistemi sovradeterminati / minimi quadrati

$A \in \mathbb{R}^{m\times n}$ $m>n$ $Ax=b$ solitamente non ha soluzione
$x \in \mathbb{R}^n$ $b \in \mathbb{R}^m$

Cerco x t.c. $\|Ax-b\|_2$ è più piccola possibile

$r = Ax-b$ $\min_{x\in\mathbb{R}^n} \|r\|_2 \iff \min_{x\in\mathbb{R}^n} \|r\|_2^2 = \min r_1^2 + r_2^2 + \dots + r_m^2$

$= \min \left(Ax-b\right)_1^2 + \left(Ax-b\right)_2^2 + \dots + \left(Ax-b\right)_m^2$

Per risolverlo

1) $A^T A x = A^T b$ <u>equazioni normali</u>

2) $A = Q_1 R_1 = \boxed{Q_1}\boxed{R_1}$, $x = R_1^{-1} Q_1^T b \iff R_1 x = Q_1^T b$

Immaginiamo un fenomeno modellato da

$$y = \phi(x) = \alpha_1 \phi_1(x) + \alpha_2 \phi_2(x) + \dots + \alpha_n \phi_n(x)$$

$\phi_i : U \to \mathbb{R}$ date, $\alpha_1, \alpha_2, \dots \alpha_n \in \mathbb{R}$ incognite

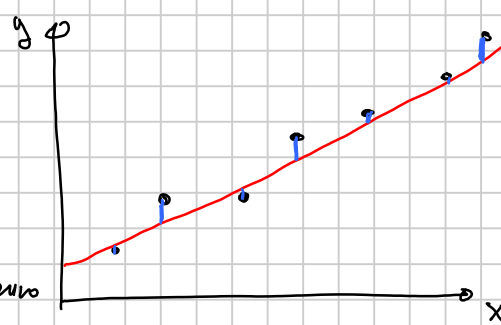
Vogliamo determinare il valore degli α_i in base ad osservazioni, cioè coppie $(x_1, y_1), (x_2, y_2), \dots (x_m, y_m)$ $x_i \in \mathbb{R}$, $y_i \in \mathbb{R}$

$x_1, \dots x_m$ <u>distinti</u>, cioè $x_i \neq x_j$ quando $i \neq j$

<u>ES</u>:

$$y = \alpha_1 \cdot 1 + \alpha_2 \cdot x$$

$\phi_1(x) = 1$ $\phi_2(x) = x$



(x_1, y_1)
$\vdots$
(x_7, y_7)

Ci vengono date coppie (x_i, y_i) che stanno sulla retta, oppure vicino ad esse

m coppie (x_i, y_i)

Voglio trovare α_1, α_2 per cui la retta approssima meglio i punti dati, cioè la somma dei quadrati delle distanze $\sum_{i=1}^{m} \left(y_i - (\alpha_1 + \alpha_2 x_i)\right)^2$ (*) sia minima

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \text{ incognite} \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \qquad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ & \vdots \\ 1 & x_m \end{bmatrix} \in \mathbb{R}^{m \times 2}$$

(*) diventa

$$\min_{\alpha \in \mathbb{R}^2} \left\| \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} - \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} y_1 - (\alpha_1 + \alpha_2 x_1) \\ y_2 - (\alpha_1 + \alpha_2 x_2) \\ \vdots \\ y_m - (\alpha_1 + \alpha_2 x_m) \end{bmatrix} \right\|^2$$

$$\min_{\alpha \in \mathbb{R}^2} \| y - X\alpha \|^2$$ problema ai minimi quadrati

con soluzione $\alpha = (X^T X)^{-1} (X^T y)$.

retta di migliore approssimazione

matrice 2×2

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \cdot \frac{1}{ad - bc}$$

Caso più generale:

$$y = \alpha_1 \phi_1(x) + \alpha_2 \phi_2(x) + \dots + \alpha_n \phi_n(x)$$

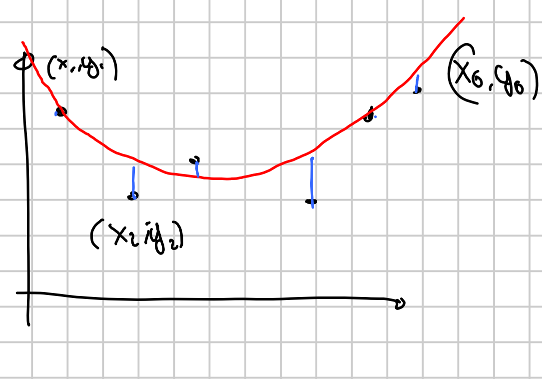
$$\min \sum_{i=1}^{m} \left(y_i - (\alpha_1 \phi_1(x_i) + \alpha_2 \phi_2(x_i) + \dots + \alpha_n \phi_n(x_m)) \right)^2 =$$

$$= \min \left\| \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{y} - \underbrace{\begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \dots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \dots & \phi_n(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(x_m) & \phi_2(x_m) & \dots & \phi_n(x_m) \end{bmatrix}}_{X} \cdot \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}}_{\alpha} \right\|^2 = \min \| y - X\alpha \|^2$$

con soluzione

$$\alpha = (X^T X)^{-1} (X^T y)$$

ES: parabola che meglio approssima dati o in generale polinomio di grado $< n$ che corrisponde a prendere



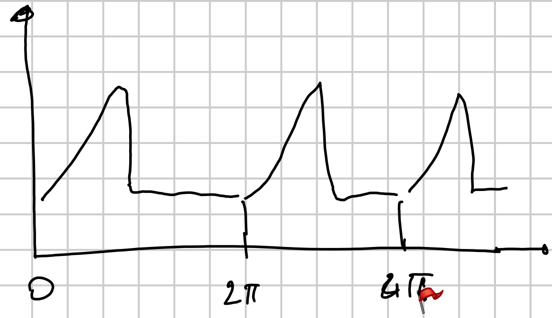
$\phi_1 = 1 \quad \phi_2 = x, \ \phi_3 = x^2, \ \dots \ \phi_n = x^{n-1}$

(approssimazione polinomiale / fit)

ES approssimazione con polinomi trigonometrici

se vogliamo approssimare una funzione che è periodica di periodo 2π,

Ha senso scegliere le funzioni $\phi_i(x)$ periodiche di periodo 2π, cioè



$$\phi_1(x) = 1$$

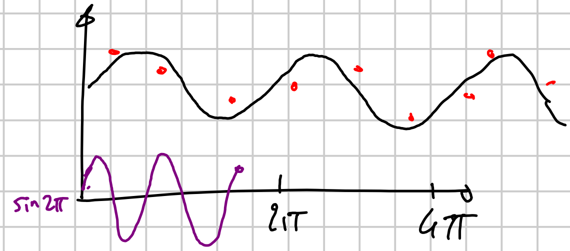
$$\phi_2(x) = \sin(x) \qquad \phi_3 = \cos(x)$$

$$\phi_4(x) = \sin(2x) \qquad \phi_5(x) = \cos(2x)$$

$$\phi_6(x) = \sin(3x) \qquad \phi_7(x) = \cos(3x)$$

$$\vdots \qquad\qquad \vdots$$

$$\phi_{2k}(x) = \sin(kx) \qquad \phi_{2k+1}(x) = \cos(kx)$$



Variante: interpolazione polinomiale

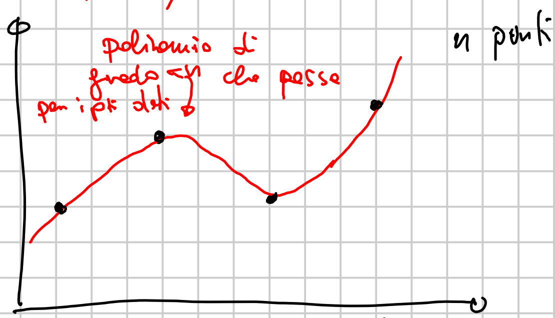
Stesso problema, ma con numero di equazioni = numero di incognite:

cerchiamo un polinomio di grado $< n$,

$$\phi(x) = \alpha_1 \cdot \underbrace{1}_{\phi_1(x)} + \alpha_2 \cdot \underbrace{x}_{\phi_2(x)} + \alpha_3 \cdot \underbrace{x^2}_{\phi_3(x)} + \dots + \alpha_n \underbrace{x^{n-1}}_{\phi_n(x)}$$

che soddisfi

$$y_1 = \phi(x_1)$$
$$y_2 = \phi(x_2)$$
$$\vdots$$
$$y_n = \phi(x_n)$$



Questa volta otteniamo un sist. lin. quadrato che possiamo risolvere esattamente.

$$\begin{cases} y_1 = \alpha_1 \cdot 1 + \alpha_2 \cdot x_1 + \dots + \alpha_n x_1^{n-1} \\ y_2 = \alpha_1 1 + \alpha_2 x_2 + \dots + \alpha_n x_2^{n-1} \\ \vdots \qquad\qquad \vdots \\ y_n = \alpha_1 1 + \alpha_2 x_n + \dots + \alpha_n x_n^{n-1} \end{cases} \iff \underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix}}_{\text{Matrice di Vandermonde associata a } (x_1, x_2, \dots x_n)} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Teo: la matrice di Vandermonde è invertibile per ogni scelta degli $(x_1, x_2, \dots x_n)$ distinti.

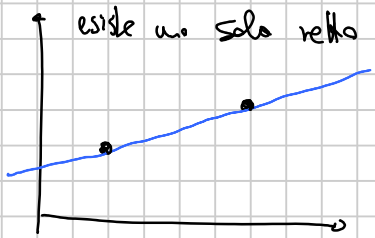
(x_i sono detti nodi)

(distinti necessario per non avere due righe uguali)

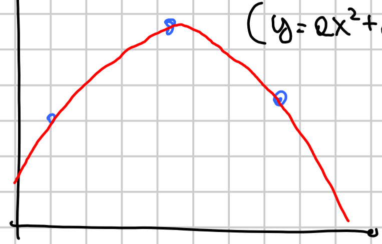
Questo implica che il problema dell'interpolazione polinomiale ha sempre soluzione unica:

Teo: per ogni n-upla di punti $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$ $x_1, x_2 \dots x_n \in \mathbb{R}$ distinti, $y_1, y_2, \dots y_n \in \mathbb{R}$, esiste uno e un solo polinomio $\phi(x) = \alpha_1 + \alpha_2 x + \dots + \alpha_n x^{n-1}$ di grado $d \leq n-1$ che passa per i punti dati, $y_i = \phi(x_i)$ $i = 1, 2, \dots n$.

ES $n=2$ esiste una sola retta per 2 punti



$n=3$ esiste una sola parabola ($y = ax^2 + bx + c$)



⚠ sul grado, ha solo $\leq$ ($\alpha_n, \alpha_{n-1}, \dots$ possono essere uguali a 0)

ES:

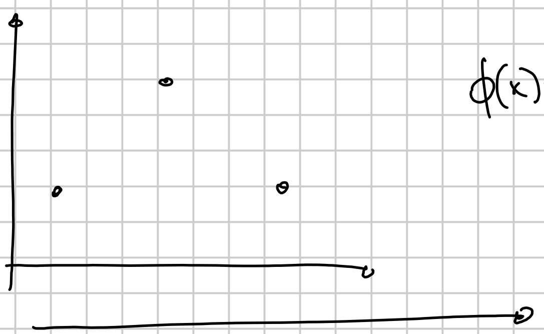
$n=3$

$(x_1, y_1) = (1,1)$

$(x_2, y_2) = (2,2)$

$(x_3, y_3) = (3,3)$

Il polinomio di interpolazione di grado ≤ 2 è $\phi(x) = 0 \cdot 1 + 1 \cdot x + 0 \cdot x^2$, che ha grado 1



$(x_1, y_1) = (1,2)$

$(x_2, y_2) = (4,5)$

$(x_3, y_3) = (8,2)$

$$\phi(x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 16 \\ 1 & 8 & 64 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix}$$

$$\phi(x) = 2.25x - 0.25x^2$$

⚠ Il problema è spesso molto mal condizionato, se il grado è abbastanza alto

---

Formula esplicita per la soluzione del problema.

Definiamo <u>polinomi di Lagrange</u>

$$L_k(x) = \frac{\prod_{j=1, j\neq k}^{n} (x - x_j)}{\prod_{j=1, j\neq k}^{n} (x_k - x_j)} \qquad k = 1, 2, \dots n$$

denominatore non si annulla mai se gli x_i sono distinti

<u>ES</u>: $x_1 = 1, \quad x_2 = 2, \quad x_3 = 4$ <u>nodi</u>

$$L_1(x) = \frac{(x-2)(x-4)}{(1-2)(1-4)} = \frac{1}{3}(x^2 - 6x + 8)$$

$$L_2 = \frac{(x-1)(x-4)}{(2-1)(2-4)} = -\frac{1}{2}(x^2 - 5x + 4)$$

$$L_3 = \frac{(x-1)(x-2)}{(4-1)(4-2)} = \frac{1}{6}(x^2 - 3x + 2) \leftarrow L_3(x_1) = 0, \; L_3(x_2) = 0, \; L_3(x_3) = 1$$

<u>Lemma</u>: I polinomi di Lagrange (costruiti su $x_1, x_2, \dots x_n$) hanno grado $n-1$ e soddisfano

$$L_k(x_i) = \begin{cases} 0 & \text{se } i \neq k \\ 1 & \text{se } i = k \end{cases}$$

<u>dim</u>: <u>grado $n-1$</u>: al numeratore, prodotto di $n-1$ termini, al denominatore solo un numero. di grado 1

<u>$L_k(x_i) = 0$ se $i \neq k$</u>: Sostituendo $x = x_i$, al numeratore ho un prodotto di $n-1$ termini, $(x_i - x_1)(x_i - x_2) \dots \dots (x_i - x_n)$

uno di questi sarà $(x_i - x_i) = 0$

<u>$L_k(x_k) = 1$</u>: sostituendo $x = x_k$, numeratore = denominatore □

<u>Teo</u>: il polinomio di interpolazione delle coppie $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$ è uguale a

$$\phi(x) = \sum_{i=1}^{n} L_i(x) \cdot y_i = L_1(x)\, y_1 + L_2(x)\, y_2 + \dots + L_n(x)\, y_n$$

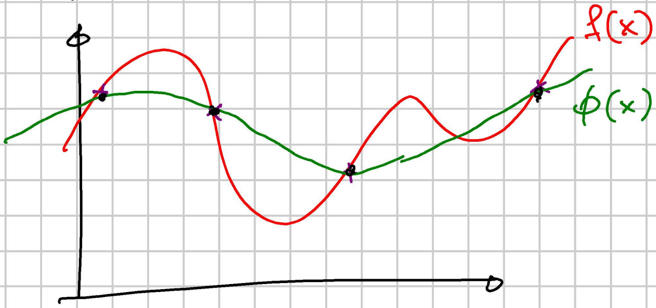
dim: dobbiamo verificare che:

(1) $\phi(x)$ ha grado $\leq n-1$: vero perché è somma pesata degli $L_i(x)$, che hanno grado $n-1$.

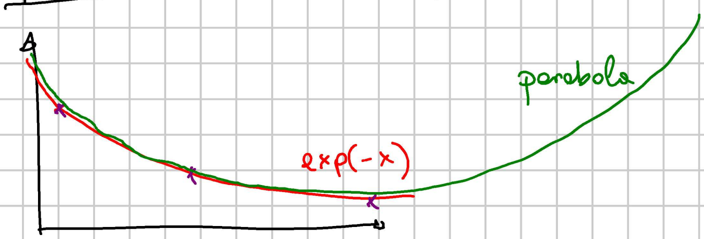
(2) Per ogni $k=1,2,\dots n$, $\phi(x_k) = y_k$: vero perché

$$\phi(x_k) = L_1(x_k)\, y_1 + L_2(x_k)\, y_2 + \dots + L_n(x_k)\, y_n = 0 \cdot y_1 + 0 \cdot y_2 + \dots + 1 \cdot y_k + \dots + 0 \cdot y_n$$
$$= y_k$$

Posso usare i polinomi di interpolazione per trovare una funzione più semplice $\phi(x)$ che ne approssima una più complicata $f(x)$:



scegliamo n nodi $x_1, \dots x_n$ e costruiamo il polinomio di interp. per $(x_i, y_i = f(x_i))$ $i=1, \dots n$



Teo: (resto dell'interpolazione)

Sia $f \in C^n([a,b])$, $x_1, x_2, \dots x_n$ nodi distinti in $[a,b]$

$\phi(x)$ polinomio di interpolazione di f nei nodi $x_1, x_2, \dots x_n$.

Allora, per ogni $x \in [a,b]$ esiste $\xi \in (a,b)$ tale che

$$f(x) - \phi(x) = \frac{f^{(n)}(\xi)}{n!} (x-x_1)(x-x_2) \cdots (x-x_n).$$

Commenti:

1) Il teorema ricorda uno sviluppo di Taylor:

$$f(x) = \underbrace{\phi(x)}_{\text{polinomio di grado} \leq n-1} + \underbrace{\frac{f^{(n)}(\xi)}{n!} (x-x_1) \cdots (x-x_n)}_{\text{resto}}$$

il termine x^n, oppure $(x-x_0)^n$
è rimpiazzato da questo prodotto.

2) Ci aspettiamo che $f(x_i) = \phi(x_i)$ per ognuno dei nodi x_i
quindi il termine di sx si annulla, e anche quello di dx, perché c'è $x_i - x_i$

3) Se $f(x)$ è essa stessa un polinomio di grado $\leq n-1$, allora
$\phi(x) = f(x)$ (il polinomio di interpolazione è la funzione stessa),
quindi $f(x) - \phi(x)$ si annulla per tutti gli x.
Si annulla anche il termine di dx, perché $f^{(n)}(x) \equiv 0$ se
f è un polinomio di grado $\leq n-1$.

<u>Integrazione numerica</u> (quadratura):

Data una funzione $f(x)$ e due estremi a, b, calcolare (o approssimare)
l'integrale definito $I = \int_a^b f(x)\,dx$

Vedremo formule che approssimano l'integrale con

$$I \approx \sum_{i=1}^{n} w_i f(x_i),$$

somma pesata della funzione valutata in certi punti

x_i <u>nodi</u>

$w_i =$ <u>pesi</u>

x_i, w_i dipenderanno da $[a,b]$, ma non da f.

Chiediamo da queste formule che valga un'uguaglianza esatta per funzioni sufficientemente semplici; ad es. $f(x) = C$ costante

$$(b-a)\,C = \int_a^b C\,dx = \sum_{i=1}^{n} w_i C = C\left(\sum_{i=1}^{n} w_i\right)$$

e perché valga vogliamo che $\sum_{i=1}^{n} w_i = b-a$