

Approssimazione polinomiale

Alvise Sommariva

Università degli Studi di Padova
Dipartimento di Matematica Pura e Applicata

24 aprile 2024

Problema.

Siano date

- le ascisse $x_1 < x_2 < \dots < x_N$;
 - le ordinate y_1, \dots, y_N .
- 1 Nel caso dell'interpolazione polinomiale si è cercato un polinomio p_m di grado $m = N - 1$, tale che $p_m(x_k) = y_k$ per $k = 1, \dots, N$.
 - 2 Nel caso cerchiamo di approssimare questi dati con un polinomio di grado strettamente inferiore a m , cosa risulta meglio fare?

Esempio

Date le ascisse

$$x_1 = 0, x_2 = h, x_3 = 2h, \dots, x_{20} = 2\pi, \quad \text{con } h = \frac{2\pi}{19}$$

e i valori $y_k = \sin(x_k)$ con $k = 1, \dots, 20$, cercare il polinomio di grado 3 che meglio approssima (in qualche senso da definire, che diremo poi ai minimi quadrati) i dati.

Si noti che nell'esempio, il grado è 4 e i punti non sono 5 come nell'interpolazione polinomiale bensì 20 ed è quindi improbabile soddisfare tutti i vincoli richiesti.

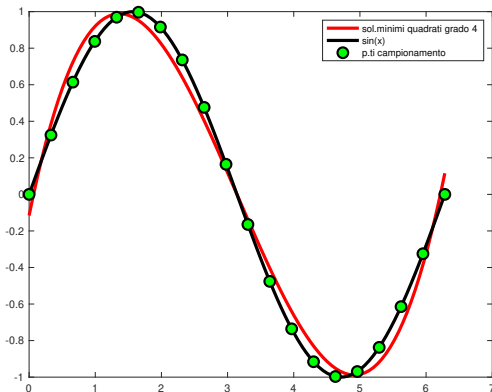


Figura: In nero il grafico di $f(x) = \sin(x)$, in rosso un polinomio approssimante (in qualche senso ancora da definire, che sarà detto in seguito ai minimi quadrati) di grado $m = 4$, in verde le coppie (x_k, y_k) , con $y_k = f(x_k)$ per $k = 1, \dots, 20$.

(Motivazione)

Dato un campionamento $\{(x_i, y_i)\}$, $i = 1, 2, \dots, N$, dove $x_i \neq x_j$ se $i \neq j$, nell'approssimazione ai **minimi quadrati** si cerca inizialmente un polinomio

$$\mathcal{L}_m(x) = \sum_{k=0}^m a_k x^k$$

di grado $m < N$, in cui tipicamente $m \ll N$, per cui i coefficienti a_0, a_1, \dots, a_m risolvano

$$\begin{cases} a_0 + a_1 \cdot x_1 + \dots + a_m \cdot x_1^m = y_1 \\ a_0 + a_1 \cdot x_2 + \dots + a_m \cdot x_2^m = y_2 \\ \dots \\ a_0 + a_1 \cdot x_N + \dots + a_m \cdot x_N^m = y_N \end{cases} \Leftrightarrow \begin{cases} a_0 + a_1 \cdot x_1 + \dots + a_m \cdot x_1^m - y_1 = 0 \\ a_0 + a_1 \cdot x_2 + \dots + a_m \cdot x_2^m - y_2 = 0 \\ \dots \\ a_0 + a_1 \cdot x_N + \dots + a_m \cdot x_N^m - y_N = 0 \end{cases}$$

Tale problema sovradeterminato (più equazioni che incognite) può non avere soluzione. Si nota

$$\begin{cases} a_0 + a_1 \cdot x_1 + \dots + a_m \cdot x_1^m - y_1 = 0 \\ a_0 + a_1 \cdot x_2 + \dots + a_m \cdot x_2^m - y_2 = 0 \\ \dots \\ a_0 + a_1 \cdot x_N + \dots + a_m \cdot x_N^m - y_N = 0 \end{cases} \Leftrightarrow \begin{cases} (a_0 + a_1 \cdot x_1 + \dots + a_m \cdot x_1^m - y_1)^2 = 0 \\ (a_0 + a_1 \cdot x_2 + \dots + a_m \cdot x_2^m - y_2)^2 = 0 \\ \dots \\ (a_0 + a_1 \cdot x_N + \dots + a_m \cdot x_N^m - y_N)^2 = 0 \end{cases}$$

$$\Leftrightarrow (a_0 + a_1 \cdot x_1 + \dots + a_m \cdot x_1^m - y_1)^2 + \dots + (a_0 + a_1 \cdot x_N + \dots + a_m \cdot x_N^m - y_N)^2 = 0$$

Così si cercano i coefficienti a_0, a_1, \dots, a_m per cui sia minima (ovvero più vicina a 0)

$$\sum_{i=1}^N (y_i - \mathcal{L}_m(x_i))^2 = (a_0 + a_1 \cdot x_1 + \dots + a_m \cdot x_1^m - y_1)^2 + \dots + (a_0 + a_1 \cdot x_N + \dots + a_m \cdot x_N^m - y_N)^2.$$

Quindi,

- dato un campionamento $\{(x_i, y_i)\}$, $i = 1, 2, \dots, N$, dove $x_i \neq x_j$ se $i \neq j$,
- ricordato che se $\mathbf{v} = (v_1, \dots, v_n)$ allora $\|\mathbf{v}\|_2 = \sqrt{\sum_{k=1}^n v_k^2}$,

nell'approssimazione ai **minimi quadrati** si cerca un polinomio $\mathcal{L}_m(x) = \sum_{k=0}^m a_k x^k$ di grado $m < N$, in cui tipicamente $m \ll N$, tale che sia minima

$$\sum_{i=1}^N (y_i - \mathcal{L}_m(x_i))^2$$

o equivalentemente, posto $\mathbf{y} = (y_i)_i$, quando lo è la somma degli scarti quadratici

$$\begin{aligned} \|\mathbf{y} - (\mathcal{L}_m(x_i))_{i=1, \dots, N}\|_2 &= \sqrt{\sum_{i=1}^N (y_i - \mathcal{L}_m(x_i))^2} \\ &= \sqrt{\sum_{i=1}^N (y_i - (a_0 + a_1 x_i + \dots + a_m x_i^m))^2} \end{aligned}$$

ovvero che risolve il problema di ottimizzazione in $m + 1$ variabili $\mathbf{a} = (a_0, \dots, a_m)$

$$\min_{\mathbf{a} \in \mathbb{R}^{m+1}} \sqrt{\sum_{i=1}^N (y_i - (a_0 + a_1 x_i + \dots + a_m x_i^m))^2}. \quad (1)$$

Teorema (Esistenza ed unicità)

- 1 Se $\{x_k\}_{k=1,\dots,N}$ sono punti a 2 a 2 distinti, allora l'approssimante ai minimi quadrati di grado $m < N$ delle coppie $\{(x_k, y_k)\}_{k=1,\dots,N}$ esiste ed è unica.
- 2 Il polinomio $\mathcal{L}_m(x) = \sum_{k=0}^m a_k^* x^k$ è soluzione ai minimi quadrati se e solo se, posto $\mathbf{y} = (y_j)$, il vettore $\mathbf{a} = (a_i^*)$ **risolve** il sistema delle **equazioni normali**

$$\mathbf{V}^T \mathbf{V} \mathbf{a} = \mathbf{V}^T \mathbf{y}$$

dove V è la matrice di Vandermonde rettangolare di dimensione $N \times (m + 1)$

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{N-1} & x_{N-1}^2 & \dots & x_{N-1}^m \\ 1 & x_N & x_N^2 & \dots & x_N^m \end{bmatrix}$$

Nota.

- La matrice V^T è la trasposta della matrice V ovvero $(V^T)_{i,j} = V_{j,i}$.
- Si noti che $V^T V \in \mathbb{R}^{m \times m}$, $V^T \mathbf{y} \in \mathbb{R}^{m \times 1}$, ovvero il sistema delle equazioni normali è di piccole dimensioni visto che tipicamente $m \ll N$.

Dimostrazione. (Facoltativa)

Sia V la matrice di Vandermonde rettangolare di dimensione $N \times (m + 1)$

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{N-1} & x_{N-1}^2 & \dots & x_{N-1}^m \\ 1 & x_N & x_N^2 & \dots & x_N^m \end{bmatrix}$$

- Si mostra che $V^T V \in \mathbb{R}^{(m+1) \times (m+1)}$ è *simmetrica e definita positiva* (e quindi *non singolare* poiché tutti gli autovalori $\lambda_i > 0$ e quindi $\det(V^T V) = \prod_{i=1}^{m+1} \lambda_i > 0$).
- si vede che il polinomio $\mathcal{L}_m(x) = \sum_{k=0}^m a_k^* x^k$ è soluzione ai minimi quadrati se e solo se il vettore $\mathbf{a} = (a_i^*)$ *risolve* il sistema delle *equazioni normali*

$$V^T V \mathbf{a} = V^T \mathbf{y}$$

per $\mathbf{y} = (y_i)$, ed essendo $V^T V$ non singolare, necessariamente la soluzione \mathbf{a} è unica.

Teorema (Minimi quadrati qualora $m = N - 1$)

Nelle ipotesi fissate, se $m = N - 1$ allora la soluzione del problema ai minimi quadrati coincide col polinomio interpolatore le coppie (x_k, y_k) , $k = 1, \dots, N - 1$.

Traccia. (Facoltativo)

Si osservi che per $m = N - 1$,

- la soluzione ai minimi quadrati $\mathcal{L}_m(x) = a_0^* + a_1^* x_i + \dots + a_m^* x_i^m$ è la soluzione del problema

$$0 \leq \min_{\mathbf{a} \in \mathbb{R}^{m+1}} \sum_{i=1}^N (y_i - (a_0 + a_1 x_i + \dots + a_m x_i^m))^2; \quad (2)$$

- essendo $m = N - 1$ possiamo determinare il polinomio interpolatore $p_m(x) = \bar{a}_0 + \bar{a}_1 x_i + \dots + \bar{a}_m x_i^m$ e visto che è tale che $p_m(x) = y_i$ per $i = 1, \dots, N$ necessariamente

$$\bar{a}_0 + \bar{a}_1 x_i + \dots + \bar{a}_m x_i^m = y_i, \quad i = 1, \dots, N,$$

e quindi

$$\sum_{i=1}^N (y_i - (\bar{a}_0 + \bar{a}_1 x_i + \dots + \bar{a}_m x_i^m))^2 = \sum_{i=1}^N 0^2 = 0.$$

Di conseguenza, da (2), per $m = N - 1$ si ha che la soluzione ai minimi quadrati $\mathcal{L}_m(x)$ coincide col polinomio p_m di grado m che interpola le coppie $(x_i, y_i)_{i=1, \dots, N}$.

Nota. (Storica)

- *Inizialmente la scoperta venne attribuita a Legendre nel 1805 che la pubblicò in *Nouvelles méthodes pour la détermination des orbites des comètes*, ed era pubblicata in molti manuali già nel 1870.*
- *Gauss, sostenne di aver sviluppato quel metodo già prima di Legendre, nel 1795. Tale discussione portò ad una aspra disputa tra i due noti matematici.*
- *In seguito, unendo tali concetti a questioni probabilistiche, Gauss permise di migliorare il calcolo delle orbite celesti.*
- *D'altra parte già nel 1801 fu in grado di predire (dopo 100 ore di calcoli!) la complicata orbita dell'asteroide Ceres, di cui aveva solo 22 osservazioni dell'astronomo italiano Piazzi (cf. [Gauss and Ceres](#)). L'orbita predetta da Gauss era particolarmente diversa da quella suggerita da altri astronomi ma si rivelò esatta, rendendolo apprezzato anche al di fuori della matematica.*

Proposito. (Curve fitting e dati inesatti)

- Forniti i dati (x_k, y_k) , $k = 1, \dots, N$, lo scopo del **curve fitting** degli scopi è quello di determinare una funzione $f \in \mathcal{F}$, con \mathcal{F} fissato, tale che f meglio approssimi i dati (x_k, y_k) , $k = 1, \dots, N$ in qualche senso da definire.

Esempio: posta $\mathcal{F} = \{f : f(x) = a + b \exp(c \cdot x), a, b, c \in \mathbb{R}\}$, determinare a, b, c tali che $f(x) = a + b \exp(c \cdot x)$ meglio approssimi i dati (x_k, y_k) , $k = 1, \dots, N$ in qualche senso da definire.

- In questa sezione consideriamo il caso in cui $\mathcal{F} = \mathbb{P}_m$, ovvero i polinomi di grado al più m , e determiniamo una funzione $f \in \mathbb{P}_m$, tale che f meglio approssimi i dati nel senso dei minimi quadrati.
- Nelle scienze applicate capita frequentemente che gli esperimenti producano set di dati $\{(x_i, y_i)\}_{i=1, \dots, N}$ dove le ascisse sono distinte.
*In particolare spesso i valori y_i sono inesatti, ad esempio per errori di misura o approssimazione. Il curve fitting permette in qualche senso di evidenziare una curva, tra un insieme di potenziali candidate, che meglio li approssima, **filtrando il rumore**.*

Definizione (Regressione lineare)

Se $\mathcal{F} = \mathbb{P}_1$, ovvero calcoliamo la soluzione ai minimi quadrati \mathcal{L}_1 , si parla di *regressione lineare*.

Esempio

Si considerino i dati della seguente tabella

| | | | | | |
|-------|------|------|------|-----|-----|
| x_i | 1 | 3 | 4 | 6 | 7 |
| y_i | -2.1 | -0.9 | -0.6 | 0.6 | 0.9 |

Si disegni il grafico della soluzione al problema di regressione lineare.

I risultati sono evidenziati in figura, dove la retta di regressione è quella che **meglio** approssima i dati nel senso dei minimi quadrati.

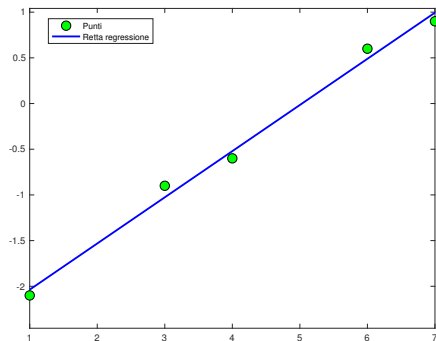


Figura: Dati e retta di regressione lineare. In questo esempio la soluzione risulta $p_1(x) = 0.505263157894737 \cdot x - 2.542105263157894$.

Esempio

Data la tabella

| | | | |
|-----|--------|-----|--------|
| 1.0 | -1.945 | 2.2 | -0.424 |
| 1.2 | -1.253 | 2.4 | -0.012 |
| 1.4 | -1.140 | 2.6 | -0.190 |
| 1.6 | -1.087 | 2.8 | +0.452 |
| 1.8 | -0.760 | 3.0 | +0.337 |
| 2.0 | -0.682 | 3.2 | +0.764 |

calcolare la retta che approssima i dati nel senso dei minimi quadrati.

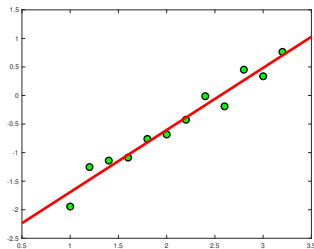


Figura: In verde i dati di campionamento, in rosso l'approssimante ai minimi quadrati di grado $m = 1$ (retta che meglio approssima i punti nel senso dei minimi quadrati).

Nota.

Negli esempi precedenti ovviamente *non esiste una retta che interpola tutti i punti*, visto che geometricamente non giacciono su una retta. Valutiamo quanto succede nell'ultimo esempio.

Se una tale retta esistesse, allora detta $y = a_0 + a_1x$ la soluzione al problema ai minimi quadrati, avremmo che per ogni coppia (x_k, y_k) della tabella sarebbe $a_0 + a_1x_k = y_k$, $k = 1, \dots, 12$ ovvero che sarebbe risolvibile il sistema in $N = 2$ incognite e $m = 12$ equazioni

$$\begin{cases} a_0 + a_1 \cdot 1 = -1.945 \\ a_0 + a_1 \cdot 1.2 = -1.253 \\ a_0 + a_1 \cdot 1.4 = -1.140 \\ \dots \\ a_0 + a_1 \cdot 3.0 = +0.337 \\ a_0 + a_1 \cdot 3.2 = +0.764 \end{cases}$$

mentre purtroppo si verifica facilmente che non ne ha (basta prendere le prime 3 equazioni e notare che non hanno una soluzione comune).

Ciò nonostante esiste la retta di miglior approssimazione ai minimi quadrati, ovvero il grafico delle coppie $(x, \mathcal{L}_1(x))$, con $\mathcal{L}_1(x) = a_0 + a_1 \cdot x$ minimizzante

$$\sqrt{(a_0 + a_1 \cdot 1 - (-1.945))^2 + (a_0 + a_1 \cdot 1.2 - (-1.253))^2 + \dots + (a_0 + a_1 \cdot 3.2 - 0.764)^2}$$

Si calcola numericamente che

$$\mathcal{L}_1(x) \approx -2.780034965034965 + 1.088111888111888 \cdot x.$$

Teorema (Qualità dell'approssimazione, (punti equispaziati))

Sia \mathcal{L}_m è la miglior approssimante polinomiale ai minimi quadrati di grado $m > 0$, relativamente alle coppie

$$(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_{N_m}, f(x_{N_m}))$$

con

- $x_j = a + (j - 1)h$, in cui $j = 1, \dots, N_m$, $h = \frac{b-a}{N_m-1}$
- $N_m \geq 2m^2 + 1$.

Se $f \in C^k([a, b])$, $k > 1$, allora esiste una costante c_k , indipendente da m , per cui

$$\|\mathcal{L}_m - f\|_\infty = \max_{x \in [a, b]} |\mathcal{L}_m(x) - f(x)| \leq c_k m^{1-k} = c_k \cdot \frac{1}{m^{k-1}}$$

- Il teorema dice che nelle condizioni fissate, più la funzione è **regolare** e più è **lecito** aspettarsi un errore più piccolo (ammesso c_k non cresca troppo).
- Si noti che nelle ipotesi indicate, per il teorema dei due carabinieri, per $n \rightarrow \infty$ la successione di polinomi \mathcal{L}_m **converge uniformemente** a f .

Esempio

Calcolare le approssimazioni \mathcal{L}_m ai minimi quadrati delle funzioni

1 $f_1(x) = x^{30} \in C^\infty([0, 1]),$

2 $f_2(x) = x^{7/2} \in C^3([0, 1]),$

3 $f_3(x) = \exp(x) \in C^\infty([0, 1]).$

in $N_m = 2m^2 + 1$ nodi equispaziati, con $m = 1, 2, \dots, 9, 10, 15, 20$.

Verificare *numericamente* la convergenza uniforme di \mathcal{L}_m a f , al crescere di m .

Come richiesto, al variare di m , sceglieremo $N_m = 2m^2 + 1$, ovvero quali ascisse di campionamento

$$x_k = a + kh, \text{ con } k = 1, \dots, N_m$$

dove

$$h = \frac{b - a}{n_m - 1} = \frac{1}{2m^2 + 1 - 1} = \frac{1}{2m^2}.$$

- Il calcolo della soluzione ai minimi quadrati viene eseguito utilizzando la funzione Matlab `polyfit`;
- se le ordinate sono valutazioni di una funzione f nei punti $x_1 < \dots < x_N$, ovvero $y_k = f(x_k)$, negli esperimenti che seguono, per avere una stima dell'errore commesso

$$E_m(f) = \|\mathcal{L}_m - f\|_\infty = \max_{x \in [0,1]} |\mathcal{L}_m(x) - f(x)|,$$

lo valutiamo in una discretizzazione Δ *sufficientemente fitta* di $[0, 1]$, ovvero

$$\Delta = \{x_j = (j-1)/9999, \quad k = 1, \dots, 10000\},$$

ottenendo

$$E_m^*(f) = \max_{x_j \in \Delta} |\mathcal{L}_m(x_j) - f(x_j)| \approx \max_{x \in [0,1]} |\mathcal{L}_m(x) - f(x)| = E_m(f).$$

I risultati sono rappresentati in tabella.

Si noti che

- Visto che $f_1(x) = x^{30} \in C^\infty([0, 1])$ ci aspettiamo che per $m \geq 30$, poiché f_1 è un polinomio di grado 30, sia in aritmetica esatta $\mathcal{L}_m = f_1$. Quindi per $m \geq 30$ ci aspettiamo un errore $E_m^*(f) = \max_{x \in [0, 1]} |\mathcal{L}_m(x) - f_1(x)|$ nullo.
- Visto che $f_2(x) = x^{7/2} \in C^3([0, 1])$ ci aspettiamo dalle stime teoriche del teorema precedente una **convergenza non troppo rapida**.
- Visto che $f_3(x) = \exp(x) \in C^\infty([0, 1])$ ci aspettiamo dalle stime teoriche del teorema precedente una **convergenza molto rapida**.

| m | $E_m^*(f_1)$ | $E_m^*(f_2)$ | $E_m^*(f_3)$ | N_m |
|-----|--------------|--------------|--------------|-------|
| 1 | $6.9e-01$ | $3.0e-01$ | $1.4e-01$ | 3 |
| 2 | $4.0e-01$ | $5.2e-02$ | $9.6e-03$ | 9 |
| 3 | $3.4e-01$ | $4.3e-03$ | $7.3e-04$ | 19 |
| 4 | $2.8e-01$ | $3.2e-04$ | $4.1e-05$ | 33 |
| 5 | $2.1e-01$ | $6.2e-05$ | $1.9e-06$ | 51 |
| 6 | $1.4e-01$ | $1.7e-05$ | $7.3e-08$ | 73 |
| 7 | $8.6e-02$ | $6.1e-06$ | $2.5e-09$ | 99 |
| 8 | $4.9e-02$ | $2.5e-06$ | $7.2e-11$ | 129 |
| 9 | $2.6e-02$ | $1.1e-06$ | $1.9e-12$ | 163 |
| 10 | $1.3e-02$ | $5.6e-07$ | $4.5e-14$ | 201 |
| 15 | $1.1e-04$ | $3.7e-08$ | $2.0e-15$ | 451 |
| 20 | $9.4e-08$ | $5.4e-09$ | $1.8e-15$ | 801 |

Tabella: Nella tabella indichiamo il grado $m = 1, \dots, 10, 15, 20$, gli errori compiuti dalle approssimazioni ai minimi quadrati $E_m^*(f_1)$, $E_m^*(f_2)$, $E_m^*(f_3)$ e il numero di punti di campionamento. Non deve sorprendere la **prima colonna**, visto che stiamo approssimando un polinomio di grado 30 con altri di grado minore. Si ricordi che $f_1(x) = x^{30}$, $f_2(x) = x^{7/2}$, $f_3(x) = \exp(x)$, $x \in [0, 1]$.

In alcuni problemi, i dati sono soggetti a rumore e uno vuole in qualche senso **togliere il rumore**.

Esempio

Consideriamo quali dati $(x_i, y_i)_{i=1, \dots, 40}$ in cui

- $x_i = -1 + (i - 1) \cdot \frac{2}{39} \in [-1, 1]$,
- $y_i = \sin(10 \cdot x_i) + 10^{-3} \mu_i$

dove μ_i è un numero random in $[-0.5, 0.5]$.

Determinare il polinomio di miglior approssimazione per $m = 5, 10, 15, 30$ e vedere come approssima $f^*(x) = \sin(10 \cdot x)$.

In qualche modo simuliamo dei dati $y_i^* = \sin(x_i)$ cui abbiamo aggiunto il rumore $10^{-3} \mu_i$.

- L'intento è approssimare la funzione $f^*(x) = \sin(10x)$ **togliendo** il rumore.
- **Non ha senso calcolare l'interpolante polinomiale**, perché si riprodurrebbero i rumori e non la funzione cercata.

In figura, mostriamo i risultati ottenuti utilizzando le approssimanti \mathcal{L}_m ai minimi quadrati di grado rispettivamente $m = 5$, $m = 10$, $m = 15$, $m = 30$.

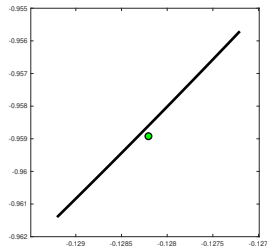
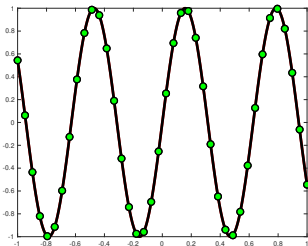


Figura: Figura a sinistra. Funzione $f^*(x) = \sin(10 \cdot x)$ e i campionamenti inesatti.

Figura a destra. Funzione $f^*(x) = \sin(10 \cdot x)$ in un intorno di $x = -0.128$ e dato inesatto.

Le pendenze non coincidono per la grafica di Matlab.

Approssimazione polinomiale e dati inesatti

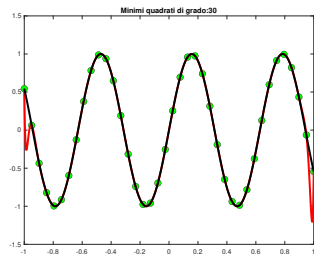
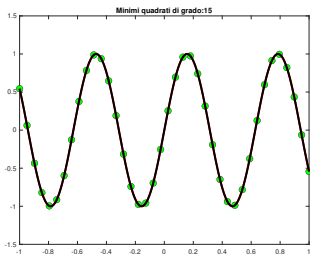
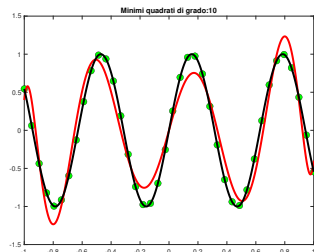
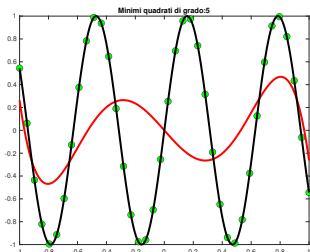


Figura: Dati inesatti (in verde) e in rosso approssimanti ai minimi quadrati di grado $m = 5$, $m = 10$, $m = 15$, $m = 30$. In nero $f(x) = \sin(10x)$.

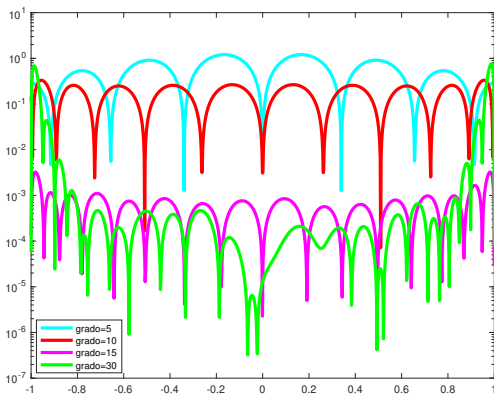


Figura: Rappresentazione grafica di $E_m(x) = |p_m(x) - f(x)|$, dove p_m é l'approssimante di grado $m = 5, m = 10, m = 15, m = 30$ (in blu) nel ricostruire $f^*(x) = \sin(10 \cdot x)$, noti dati inesatti.

- In virtù dei 6 cambi di convessità / concavità della funzione f^* , è chiaro che un **polinomio di grado basso non fornisca risultati apprezzabili**.
Ad esempio un polinomio di grado 5 può avere al più 4 cambi di convessità / concavità, e non può essere adatto allo scopo.
- Un polinomio \mathcal{L}_m di **grado m troppo alto approssimerebbe la funzione con rumore** e non f^* .
- Nel nostro esempio, $m = 15$ sembra essere un **buon compromesso**.
- Si noti come a grado $m = 30$ comincino a sorgere oscillazioni agli estremi. Inoltre per problemi dovuti a polyfit, il risultato non è quello corretto teoricamente.

La tabella che segue, mostra gli errori

$$E_m(f^*) = \max_{x \in [-1,1]} |f^*(x) - \mathcal{L}_m(x)|, \quad E_m(\mathbf{y}) = \sum_{i=1}^{40} |y_i - \mathcal{L}_m(x_i)|^2,$$

con $\mathbf{y} = (y_i)_i$ vettore dei dati perturbati, evidenziando un miglioramento di entrambe fino a grado $m = 20$ e un peggioramento nell'approssimare la funzione **senza rumore** f^* pur approssimando sempre meglio i dati (nel senso dei minimi quadrati).

| m | $E_m(f^*)$ | $E_m(\mathbf{y})$ |
|-----|-------------|-------------------|
| 5 | $1.2e + 00$ | $4.0e + 00$ |
| 10 | $3.3e - 01$ | $1.2e + 00$ |
| 15 | $4.0e - 03$ | $5.1e - 03$ |
| 20 | $5.7e - 03$ | $2.3e - 03$ |
| 25 | $3.6e - 02$ | $2.2e - 03$ |
| 30 | $1.7e + 00$ | $1.8e - 03$ |

Tabella: Grado m ed errori

- $E_m(f^*) = \max_{x \in [-1,1]} |f^*(x) - \mathcal{L}_m(x)|$,
- $E_m(\mathbf{y}) = \sum_{i=1}^{40} |y_i - \mathcal{L}_m(x_i)|^2$,

dove $f^*(x) = \sin(10x)$, $\mathbf{y} = (y_i)_i$ con $y_i = f^*(x_i) + 10^{-3}$.

Il valore $E_m(f^*)$ evidenzia un valore ottimo $m = 15$ nell'approssimare la funzione f^* non affetta da rumore.

Il valore $E_m(\mathbf{y})$, come atteso, mostra un costante miglioramento nell'approssimare i dati affetti da rumore.

Appendice facoltativa

Teorema (Errore approssimante ai minimi quadrati su opportuni campionamenti)

Se \mathcal{L}_m è la miglior approssimante polinomiale di grado $m > 0$, relativamente alle coppie $(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_N, f(x_N))$ con $a \leq x_1 < \dots < x_N \leq b$ e

- $h_m = \max_{i=1, \dots, N-1} (x_{i+1} - x_i) \leq \theta(b-a)/m^2$,
- $\theta \in (0, 1)$ (indipendente da m),

allora per ogni $f \in C^k([a, b])$, $k > 1$, esiste una costante c_k , indipendente da m , per cui

$$\max_{x \in [a, b]} |\mathcal{L}_m(x) - f(x)| \leq c_k m^{1-k}.$$

Corollario (Punti equispaziati)

Se \mathcal{L}_m è la miglior approssimante polinomiale di grado $m > 0$, relativamente alle coppie $(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_{N_m}, f(x_{N_m}))$ con

- $x_j = a + (j-1) \frac{b-a}{N_m-1}$, $j = 1, \dots, N_m$,
- $N_m \geq 2m^2 + 1$.

Allora per ogni $f \in C^k([a, b])$, $k > 1$, esiste una costante c_k , indipendente da m , per cui

$$\|\mathcal{L}_m - f\|_{\text{inf}} = \max_{x \in [a, b]} |\mathcal{L}_m(x) - f(x)| \leq c_k m^{1-k}.$$

Esempio

Sia $f \in C^5([a, b])$ (ovvero $k = 5$ nel corollario precedente). Se disponiamo dei dati $\{(x_k, f(x_k))\}_{k=1, \dots, N_m}$ con $N_m \geq 2m^2 + 1$, sappiamo che $\|\mathcal{L}_m(x) - f\|_{\text{inf}} \leq c_k m^{1-k}$.

(a) Supponiamo

- di disporre di $N_m = 20$ punti equispaziati $\{x_k\}_{k=1, \dots, 20}$ in $[a, b]$,
- di calcolare l'approssimante \mathcal{L}_3 ai minimi quadrati delle coppie $(x_k, f(x_k))$, $k = 1, \dots, 20$, ovvero $m = 3$ nel corollario precedente.

Essendo $20 = N_m \geq 2m^2 + 1 = 19$, risulta

$$\max_{x \in [a, b]} |\mathcal{L}_m(x) - f(x)| \leq c_k m^{1-k} = c_k 3^{1-5} = \frac{c_k}{81}.$$

(b) Supponiamo

- di disporre di $N_m = 300$ punti equispaziati $\{x_k\}_{k=1, \dots, 300}$ in $[a, b]$,
- di calcolare l'approssimante \mathcal{L}_{10} ai minimi quadrati delle coppie $(x_k, f(x_k))$, $k = 1, \dots, 300$, ovvero $m = 10$ nel corollario precedente.

Essendo $300 = N_m \geq 2m^2 + 1 = 201$, risulta

$$\max_{x \in [a, b]} |\mathcal{L}_m(x) - f(x)| \leq c_k m^{1-k} = c_k 10^{1-5} = \frac{c_k}{10^4}.$$

Commento

- Il teorema asserisce che la soluzione ai minimi quadrati \mathcal{L}_m di grado m , su un *set di nodi sufficientemente fitti* (ovvero l'intervallo di ampiezza maggiore è sufficientemente piccolo), darà
 - (a) *uniformemente* una buona approssimazione della funzione f da approssimare,
 - (b) tendenzialmente *migliore se la funzione è molto regolare* vista la presenza di m^{1-k} , supponendo c_k non cresca troppo.
- Non dice molto del valore della costante c_k , e quindi il termine $c_k m^{1-k}$ non è direttamente valutabile.
- Al crescere di m , il fattore $c_k m^{1-k} \rightarrow 0$ e quindi per set di nodi $\{x_k\}_{k=1}^N$ sufficientemente fitti, la successione $\{\mathcal{L}_m\}_m$ converge uniformemente a $f \in C^k([a, b])$, qualora sia $k > 1$.

Esempio (. Minimi quadrati e set di nodi equispaziati suff. fitti)

Supponiamo che al variare di m , si calcoli l'approssimante polinomiale \mathcal{L}_m delle coppie $(x_k, f(x_k))_{k=1, \dots, N_m}$, in cui

- $f \in C^k([a, b])$, $k > 1$;
- $x_j = a + (b - a) \cdot \frac{j-1}{N_m-1}$ con $j = 1, \dots, N_m = 2m^2 + 1$, ovvero N_m punti equispaziati in $[a, b]$, con passo $h_m = \frac{b-a}{N_m-1}$.

In virtù del precedente corollario si ha

$$0 \leq \max_{x \in [a, b]} |\mathcal{L}_m(x) - f(x)| \leq c_k m^{1-k}. \quad (3)$$

Se si considera in particolare una sequenza di approssimazioni $\{\mathcal{L}_m\}_{m=1, 2, \dots}$ con queste proprietà, allora da (3) abbiamo che per il teorema del confronto

$$\lim_{m \rightarrow +\infty} \|\mathcal{L}_m - f\|_\infty = \lim_{m \rightarrow +\infty} \max_{x \in [a, b]} |\mathcal{L}_m(x) - f(x)| = 0,$$

ovvero *la successione di polinomi $\{\mathcal{L}_m\}_{m=1, 2, \dots}$ converge uniformemente a f in $[a, b]$.*