Il metodo dei minimi quadrati

Marcello Colozzo – http://www.extrabyte.info

1 Approssimazione locale

Da una serie di osservazioni sperimentali sufficientemente accurate, ricaviamo i valori assunti da una grandezza y in funzione di una variabile reale x. Precisamente:

$$y_k = f(x_k), \quad k = 0, 1, 2, ..., N \tag{1}$$

Senza perdita di generalità, si supponga il seguente ordinamento:

$$x_0 = a < x_1 < x_2 < ... < x_N = b \tag{2}$$

Ci proponiamo di approssimare la funzione f attraverso un numero assegnato di funzioni note o meglio, mediante una appropriata combinazione lineare:

$$S_n(x) = \sum_{k=0}^{n} c_k \varphi_k(x) \tag{3}$$

dove $\varphi_k(x)$ sono le predette funzioni.

1.1 Determinazione dell'errore

Per implementare un algoritmo del procedimento di approssimazione abbozzato nel § precedente, supponiamo di conoscere l'espressione elementare della funzione $f(x)$, rimanendo nell'ipotesi di voler approssimare tale grandezza con una somma del tipo (3). Appare allora naturale definire l'errore:

$$\varepsilon_n(x) = f(x) - S_n(x), \quad \forall x \in [a, b] \tag{4}$$

È chiaro che l'approssimazione è buona se $\varepsilon_n(x)$ è "trascurabilmente piccolo". Quest'ultima locuzione dipende dalla natura dell'approssimazione cercata. Ad esempio, se stiamo ricercando un'approssimazione locale, i.e. in un intorno di un punto x_k della sequenza (2), dobbiamo imporre:

$$\begin{cases} f(x_k) = S_n(x_k) \\ f'(x_k) = S'_n(x_k) \\ ... \\ f^{(n)}(x_k) = S_n^{(n)}(x_k) \end{cases} \tag{5}$$

Cioè

$$\begin{cases} c_0\varphi_0(x_k) + c_1\varphi_1(x_k) + ... + c_n\varphi_n(x_k) = f(x_k) \\ c_0\varphi'_0(x_k) + c_1\varphi'_1(x_k) + ... + c_n\varphi'_n(x_k) = f'(x_k) \\ ... \\ c_0\varphi_0^{(n)}(x_k) + c_1\varphi_1^{(n)}(x_k) + ... + c_n\varphi_n^{(n)}(x_k) = f^{(n)}(x_k) \end{cases}, \tag{6}$$

che è un sistema di $n+1$ equazioni lineari nelle $n+1$ incognite $c_0, c_1, ..., c_n$. Tale sistema sotto ovvie condizioni, permette di determinare l'unica n-pla $(\bar{c}_0, \bar{c}_1, ..., \bar{c}_n)$ che rende minimo l'errore in un intorno di x_k. Ad esempio, lo sviluppo di Taylor di punto iniziale x_k e troncato al termine di ordine $n+1$, corrisponde manifestamente alla scelta di funzioni

$$\varphi_0(x) = 1, \ \varphi_1(x) = x - x_k, \ \varphi_2(x) = (x - x_k)^2, ..., \varphi_n(x) = (x - x_k)^n$$

ovvero i polinomi di grado $0, 1, 2, ..., n$. Come esempio numerico, consideriamo la gaussiana:

$$f(x) = e^{-2x^2},$$

il cui sviluppo in serie di Taylor di punto iniziale $\frac{1}{2}$ troncato al termine di ordine 4 è la seguente combinazione lineare di polinomi:

$$\frac{1}{\sqrt{e}}\left[1 - 2\left(x - \frac{1}{2}\right)^2 + \frac{8}{3}\left(x - \frac{1}{2}\right)^3 - \frac{4}{3}\left(x - \frac{1}{2}\right)^4\right]$$

In fig. 1 riportiamo il grafico della $f(x)$ confrontato con la predetta combinazione lineare, da cui vediamo che l'approssimazione è buona in un intorno di $1/2$.

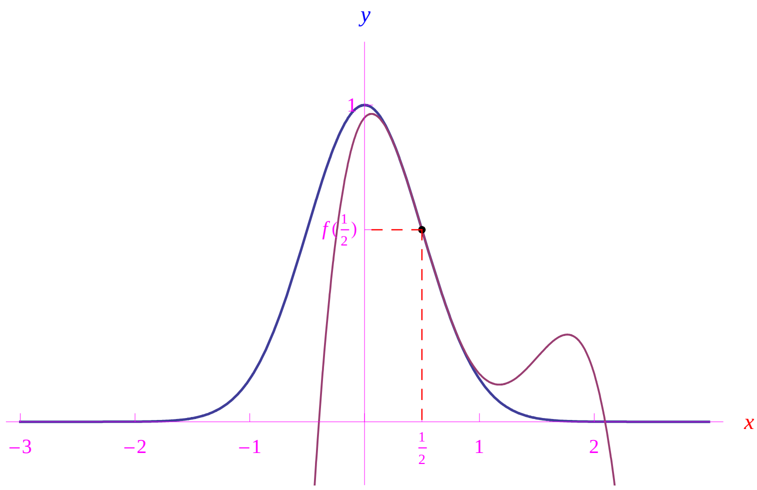

Figura 1: Lo sviluppo di Taylor troncato al quart'ordine è una buona approssimazione in un intorno di $x = \frac{1}{2}$.

2 Approssimazione globale. Il metodo dei minimi quadrati

Riferiamoci ora a una assegnata funzione $f \in C^{p\geq 1}([a, b])$, della quale ricerchiamo un'approssimazione attraverso uno sviluppo del tipo

$$S_n(x) = \sum_{k=0}^{n} c_k \varphi_k(x) \tag{7}$$

dove $\varphi_k \in C^{\omega}([a, b])$ i.e. sono funzioni analitiche in $[a, b]$. Per quanto precede, l'errore è dato da

$$\varepsilon_n(x) = f(x) - S_n(x), \quad \forall x \in [a, b] \tag{8}$$

Per una approssimazione globale, siamo tentati a rendere minimo il valore medio della funzione $\varepsilon_n(x)$ per un assegnato n. Cioè

$$\langle \varepsilon_n \rangle = \frac{1}{b-a}\int_a^b \varepsilon_n(x)\, dx, \tag{9}$$

che possiamo denominare **errore lineare medio**. Tuttavia, è facile convincersi che $\langle \varepsilon_n \rangle$ può essere trascurabilmente piccolo anche per valori molto grandi di $\varepsilon_n(x)$. Per fissare le idee supponiamo di voler approssimare in $[0,1]$ la funzione $f(x) = \sin(e^{4x})$ attraverso la combinazione lineare di polinomi:

$$\varphi_0(x) = x,\ \varphi_1(x) = x^2,\ \varphi_2(x) = x^3$$

con coefficienti

$$c_0 = 1,\ c_1 = -1,\ c_2 = \frac{1}{10}$$

Cioè con la somma

$$S_2(x) = x - x^2 + \frac{1}{10}x^3$$

Segue

$$\varepsilon_2(x) = \sin\left(e^{4x}\right) - \left(x - x^2 + \frac{1}{10}x^3\right) \tag{10}$$

da cui l'errore lineare medio

$$\langle\varepsilon_2\rangle = \int_0^1 \left[\sin\left(e^{4x}\right) - \left(x - x^2 + \frac{1}{10}x^3\right)\right] dx \simeq -3 \cdot 10^{-3}$$

Tuttavia $f(x)$ si discosta molto da $S_2(x)$ come mostrato dal grafico di fig. 2. La piccolezza dell'errore lineare medio è dovuta al fatto che la funzione $\varepsilon_2(x)$ compie in $[0,1]$ delle oscillazioni tali che il corrispondente rettangoloide ha misura relativa nulla, come si evince dal grafico di fig. 3.

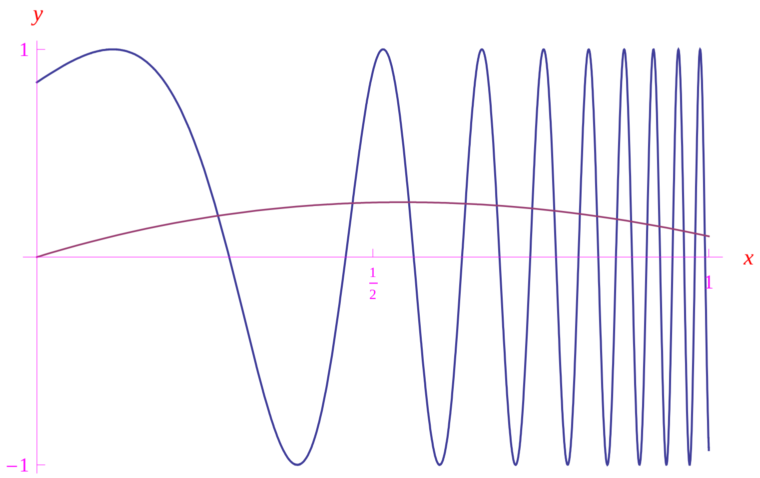

Figura 2: Grafico della funzione $f(x) = \sin(e^{4x})$ confrontato con il grafico di $S_2(x)$ in $[0,1]$. La somma S_2 non riesce ad inseguire le rapide oscillazioni di f. Pur tuttavia, l'errore lineare medio è trascurabilmente piccolo.

Una misura affidabile ma computazionalmente onerosa è data da

$$\langle|\varepsilon_2|\rangle = \int_0^1 |\varepsilon_2(x)|\, dx$$

L'onere computazionale deriva ovviamente dal valore assoluto nell'integrando che rende poco maneggevole l'integrale. Allora, proviamo con

$$\left\langle\varepsilon_2^2\right\rangle = \int_0^1 \varepsilon(x)^2\, dx \simeq 0.579$$

che è già un valore più affidabile.

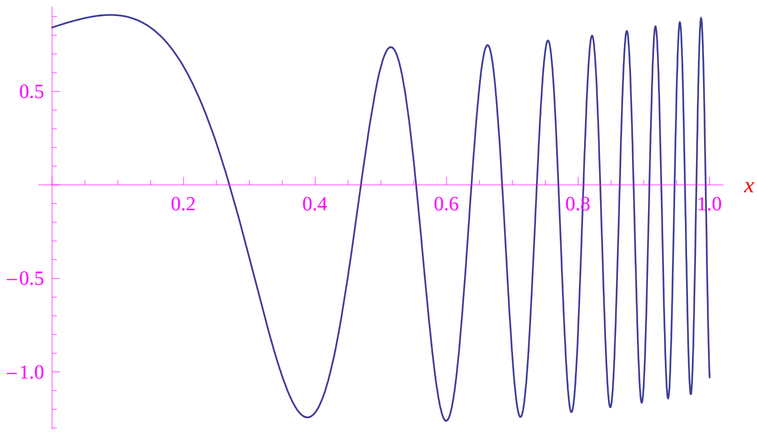

Figura 3: Andamento dell'errore (10).

Riprendendo il caso generale

$$\varepsilon_n(x) = f(x) - S_n(x) \tag{11}$$

con

$$S_n(x) = \sum_{i=0}^{n} c_i \varphi_i(x), \tag{12}$$

dove $\varphi_i(x)$ sono assegnate, e i coefficienti c_i sono da determinare in modo da rendere minimo l'**errore quadratico medio**, i.e. la media integrale in $[a, b]$ della funzione $\varepsilon_n^2(x)$

$$\langle \varepsilon_n^2 \rangle = \frac{1}{b-a} \int_a^b [f(x) - S_n(x)]^2 \, dx \tag{13}$$

Si tratta, dunque, di determinare il minimo assoluto della funzione $\langle \varepsilon_n^2 \rangle$ delle $n+1$ variabili reali. Partiamo dal caso più semplice: $n = 0$. Cioè vogliamo approssimare $f(x)$ con la somma

$$S_0(x) = c_0 \varphi_0(x),$$

per un'assegnata funzione analitica $\varphi_0(x)$. Abbiamo

$$\langle \varepsilon_0^2 \rangle = \frac{1}{b-a} \int_a^b \left[f(x)^2 - 2c_0 \varphi_0(x) + c_0^2 \varphi_0(x)^2 \right] dx,$$

da cui

$$\langle \varepsilon_0^2 \rangle = A c_0^2 + B c_0 + C \tag{14}$$

ovvero un polinomio di secondo grado in c_0 e con coefficienti

$$A = \frac{1}{b-a} \int_a^b \varphi_0(x)^2 \, dx, \quad B = -\frac{2}{b-a} \int_a^b f(x) \varphi_0(x) \, dx, \quad C = \frac{1}{b-a} \int_a^b f(x)^2 \, dx \tag{15}$$

Si noti che riesce $A > 0$ per cui la (14) è l'equazione di una parabola (nel piano cartesiano la cui ascissa riporta i valori della variabile c_0 e in ordinate l'errore quadratico medio) concava verso l'alto. Ne consegue che la predetta funzione assume un minimo assoluto nel punto

$$\bar{c}_0 = \frac{\int_a^b f(x) \varphi_0(x) \, dx}{\int_a^b \varphi_0(x)^2 \, dx} \tag{16}$$

Per cui nel caso $n=0$ la migliore approssimazione nel senso dei minimi quadrati è

$$S_0(x)=\varphi_0(x)\cdot\frac{\int_a^b f(x)\varphi_0(x)\,dx}{\int_a^b \varphi_0(x)^2\,dx}, \tag{17}$$

per una qualunque funzione analitica assegnata $\varphi_0(x)$. Passiamo al caso $n=1$:

$$\left\langle \varepsilon_1^2\right\rangle=\frac{1}{b-a}\int_a^b\left[f(x)-c_0\varphi_0(x)-c_1\varphi_1(x)\right]^2dx$$

Sviluppando il quadrato nell'integrando, otteniamo il seguente polinomio di secondo grado in due variabili c_0,c_1

$$\left\langle \varepsilon_1^2\right\rangle=Ac_0^2+Bc_1^2+Cc_0c_1+Dc_0+Ec_1+F, \tag{18}$$

di coefficienti

$$\begin{aligned}
A&=\frac{1}{b-a}\int_a^b\varphi_0(x)^2\,dx>0 \qquad (19)\\
B&=\frac{1}{b-a}\int_a^b\varphi_1(x)^2\,dx>0\\
C&=\frac{2}{b-a}\int_a^b\varphi_0(x)\varphi_1(x)\,dx\\
D&=-\frac{2}{b-a}\int_a^b f(x)\varphi_0(x)\,dx\\
E&=-\frac{2}{b-a}\int_a^b f(x)\varphi_1(x)\,dx\\
F&=\frac{1}{b-a}\int_a^b f(x)^2\,dx>0
\end{aligned}$$

Applichiamo il procedimento standard per la ricerca del minimo assoluto di una funzione reale di due variabili reali. Determiniamo innanzitutto gli eventuali punti estremali, i.e. le soluzioni del sistema lineare:

$$\begin{cases}\frac{\partial\left\langle\varepsilon_1^2\right\rangle}{\partial c_0}=0\\ \frac{\partial\left\langle\varepsilon_1^2\right\rangle}{\partial c_1}=0\end{cases}\Longleftrightarrow\begin{cases}2Ac_0+Cc_1=-D\\ Cc_0+2Bc_1=-E\end{cases},$$

da cui l'unico punto estremale $\bar{P}(\bar{c}_0,\bar{c}_1)$, ove

$$\bar{c}_0=\frac{EC-2BD}{4AB-C^2},\quad \bar{c}_1=\frac{CD-2AE}{4AB-C^2} \tag{20}$$

Per il calcolo del determinante hessiano, calcoliamo le derivate parziali seconde:

$$\frac{\partial^2\left\langle\varepsilon_1^2\right\rangle}{\partial c_0^2}=2A,\quad\frac{\partial^2\left\langle\varepsilon_1^2\right\rangle}{\partial c_0\partial c_1}=C,\quad\frac{\partial^2\left\langle\varepsilon_1^2\right\rangle}{\partial c_1\partial c_0}=C,\quad\frac{\partial^2\left\langle\varepsilon_1^2\right\rangle}{\partial c_1^2}=2B$$

per cui

$$H\left(\bar{P}\right)=\begin{vmatrix}\frac{\partial^2\left\langle\varepsilon_1^2\right\rangle}{\partial c_0^2} & \frac{\partial^2\left\langle\varepsilon_1^2\right\rangle}{\partial c_0\partial c_1}\\ \frac{\partial^2\left\langle\varepsilon_1^2\right\rangle}{\partial c_1\partial c_0} & \frac{\partial^2\left\langle\varepsilon_1^2\right\rangle}{\partial c_1^2}\end{vmatrix}=4AB>0$$

Inoltre, è

$$\forall (c_0, c_1) \in \mathbb{R}^2, \quad \frac{\partial^2 \langle \varepsilon_1^2 \rangle}{\partial c_0^2} = 2A \Longrightarrow \left.\frac{\partial^2 \langle \varepsilon_1^2 \rangle}{\partial c_0^2}\right|_{\bar{P}} = 2A > 0$$

Ne consegue che $\bar{P}$ è punto di minimo relativo. Quindi le (20) rendono minimo l'errore quadratico medio. Riprendendo l'esempio della funzione $f(x) = \sin(e^{4x})$, supponiamo di aver assegnato le funzioni:

$$\varphi_0(x) = \sin(10x), \quad \varphi_1(x) = \sin(20x)$$

e calcolando numericamente gli integrali (19) con Mathematica, otteniamo:

$$S_1(x) = 0.377 \sin(10x) - 0.008 \sin(20x), \tag{21}$$

mentre in fig. 4 confrontiamo gli andamenti della funzione e della sua approssimazione nel senso dei minimi quadrati.

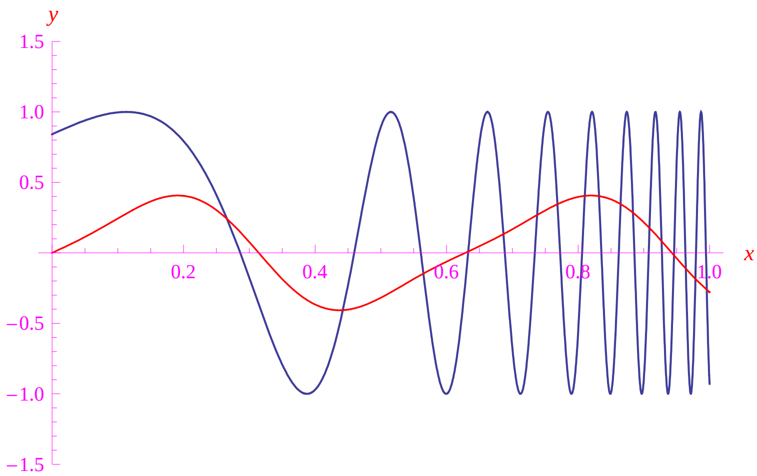

Figura 4: Andamento di $f(x) = \sin(e^{4x})$ confrontato con l'approssimazione $S_1(x)$ data dalla (21).