

APPROSSIMAZIONE

1. Introduzione

Alcune tecniche di approssimazione sono:

Funzione nota in alcuni punti → **Interpolazione**

Approssimazione con funzione più semplice → Minimizzazione errore

(**Minimi Quadrati – MinMax**)

Nel primo caso, vogliamo che la funzione $f(x)$ da interpolare coincida con la sua funzione approssimante $g(x)$ in alcuni punti $(x_i, f(x_i))$, $i = 1, \dots, n$. In tali punti l'errore dell'approssimazione è quindi nullo e, se la $g(x)$ è un polinomio, l'interpolazione si dice polinomiale.

Nel secondo caso, invece, non abbiamo una valutazione esatta della $f(x)$ nei nodi x_i , per cui si deve cercare di minimizzare l'errore in tali nodi. L'obiettivo è quindi quello di trovare un polinomio di grado n che dia la migliore approssimazione alla funzione $f(x)$ in un senso da specificare. Diremo che $g(x)$ approssima $f(x)$, e scriveremo $g(x) \sim f(x)$, se:

$$\|f - g\| \text{ è "piccola"}$$

dove il concetto di "piccolo" è legato alla tolleranza che possiamo considerare per l'errore. Richiamiamo pertanto il concetto di norma.

Sia V uno spazio lineare su \mathfrak{R} . Una **norma** $\| \cdot \|$ su V è una funzione non negativa,

definita su V , che gode delle seguenti proprietà:

$$1) \|f\| = 0 \Leftrightarrow f = 0, f \in V$$

$$2) \|\lambda f\| = |\lambda| \|f\|, \forall \lambda \in \mathfrak{R}, \forall f \in V$$

$$\|f + g\| \leq \|f\| + \|g\| \quad \forall f, g \in V$$

Un tale spazio lineare è detto normato. Vediamo alcuni esempi.

Esempio

Sia $V = \mathfrak{R}^n$, consideriamo le norme vettoriali 1, 2 e ∞ . Se $\underline{v} = (v_1, \dots, v_n)^T \in \mathfrak{R}^n$ allora si ha:

$$\|\underline{v}\|_1 = \sum_{i=1}^n |v_i|$$

$$\|\underline{v}\|_2 = \sqrt{\sum_{i=1}^n v_i^2}$$

$$\|\underline{v}\|_\infty = \max_{i=1}^n |v_i|$$

Poiché valgono le seguenti disuguaglianze:

$$\frac{1}{\sqrt{n}} \|\underline{v}\|_\infty \leq \|\underline{v}\|_2 \leq \sqrt{n} \|\underline{v}\|_\infty \quad \forall \underline{v} \in \mathfrak{R}^n$$

tali norme sono equivalenti.

Esempio

Sia $V = C[a, b]$ e sia $w(x)$ una funzione a valori reali, non negativa, continua ed integrabile in $[a, b]$, detta funzione peso.

$$\|f\|_1 = \int_a^b |f(x)| w(x) dx$$

$$\|f\|_2 = \sqrt{\int_a^b f^2(x) w(x) dx}$$

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

Tali norme non sono equivalenti, infatti si ha il seguente:

lemma

a) $\|f\|_2 \leq W \|f\|_\infty$, $W = \sqrt{\int_a^b w(x) dx}$ $\forall f \in C[a, b]$

b) Dati due numeri positivi ε (piccolo) ed M (grande) $\Rightarrow \exists f \in C[a, b] : \|f\|_2 < \varepsilon, \|f\|_\infty > M$

Tale non equivalenza rende la teoria MinMax diversa da quella dei Minimi Quadrati. Consideriamo un esempio che mostra

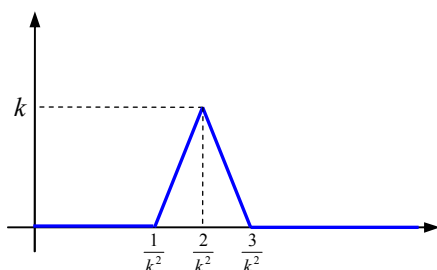
in modo pratico la non equivalenza di queste norme.

Esempio

Sia data una funzione f definita nel modo seguente:

$$f_k(x) = \begin{cases} k(k^2x - 1); & \frac{1}{k^2} \leq x \leq \frac{2}{k^2} \\ -k(k^2x - 3); & \frac{2}{k^2} \leq x \leq \frac{3}{k^2} \\ 0 & \text{altrove} \end{cases}$$

Graficamente la funzione ha il seguente andamento:



Consideriamo, poi, la funzione identicamente nulla $f(x) = 0$. Si ha:

$$\|f - f_k\|_1 = \frac{1}{k} \qquad \|f - f_k\|_2 = \frac{\sqrt{2}}{\sqrt{3}} \qquad \|f - f_k\|_\infty = k$$

Passando al limite:

$$\lim_{k \rightarrow \infty} \|f - f_k\|_1 = 0 \qquad \lim_{k \rightarrow \infty} \|f - f_k\|_2 = \frac{\sqrt{2}}{\sqrt{3}} \qquad \lim_{k \rightarrow \infty} \|f - f_k\|_\infty = \infty$$

Quindi non c'è equivalenza in $C[a,b]$ tra le tre norme.

1. Teoria generale

Sia V uno spazio lineare normato, $W \subset V$, con $\dim W = n$, e siano $\varphi_i, i=1, \dots, n$ n funzioni linearmente indipendenti.

Diremo che $w^* \in W$ è la migliore approssimazione ad $f \in V$ se:

$$\|f - w^*\| \leq \|f - w\| \qquad \forall w \in W$$

Problemi

- 1) $\exists w^*$?
- 2) E' unica?
- 3) Come si costruisce?

Si ha:

- 1) w^* esiste in uno spazio lineare normato.
- 2) L'unicità è garantita in uno spazio a prodotto interno.
- 3) La costruzione MinQuad $\| \cdot \|_2 \Rightarrow$ si può fare in un numero finito di passi,
la costruzione MinMax $\| \cdot \|_\infty \Rightarrow$ procedimento iterativo.

Nonostante le differenze delle norme, c'è una caratteristica comune dell'approssimazione indipendente dalle norme:

se non c'è un limite sul grado del polinomio approssimante p , allora l'errore $|f - p|$ può essere reso arbitrariamente

piccolo in entrambe le norme. Infatti abbiamo il:

Teorema di approssimazione di Weierstrass

$$\text{Se } f \in C[a,b], \text{ allora } \forall \varepsilon > 0 \exists p \in P_n : \|f - p\|_\infty \leq \varepsilon$$

Ovvero, ogni funzione continua può essere approssimata uniformemente da un polinomio il cui grado dipende dall'

accuratezza dell'approssimazione. Analogo risultato si ha nella norma 2.

D'altronde, se n è fissato allora il teorema non vale più. Infatti, consideriamo ad esempio la funzione:

$$f(x) = \sin x \quad x \in [0, \pi]$$

e sia $n=0$. Allora:

$$\|f - q\|_\infty \geq \frac{1}{2} \quad \forall q \in P_0$$

Problema

Data l'accuratezza dell'approssimazione, calcolare il polinomio col più basso grado possibile.

2. Minimi quadrati

1) **Problema discreto**: dati i punti $(x_0, y_0), \dots, (x_m, y_m)$ con $m > n$ ed $m, n \in \mathbb{N}$, trovare $p^*(x) \in P_n$:

$$\sum_{i=0}^m w_i [p^*(x_i) - y_i]^2 \text{ sia minimo}$$

con w_0, \dots, w_m costanti positive dette pesi.

2) **Problema continuo**: data $w(x)$ continua e positiva in $[a, b]$, trovare $p^*(x) \in P_n$:

$$\int_a^b w(x) [p^*(x) - f(x)]^2 dx \text{ sia minimo}$$

Il problema 1) è equivalente a trovare la soluzione ai minimi quadrati di un sistema lineare sovradeterminato. Vediamo qual è la connessione tra il problema discreto ai M.Q., cioè la ricerca di $p^*(x)$ che minimizzi la somma vista prima, e la soluzione ai M.Q. di un sistema sovradeterminato.

Sia $p^*(x) \in P_n$ e sia $1, t, t^2, \dots, t^n$ una base di P_n . Sia $m > n$; il polinomio cercato sarà del tipo:

$$p^*(t) = \sum_{j=0}^n x_j t^j \quad x_j \in \mathbb{R} \quad \Rightarrow \quad p^*(t_i) = \sum_{j=0}^n x_j t_i^j \quad i = 0, \dots, m$$

Ponendo $t_i^j = a_{ij} \Rightarrow p^*(t_i) = \sum_{j=0}^n a_{ij} x_j$ che è equivalente a risolvere $Ax = b$, dove:

$$A = [a_{ij}], \quad x = [x_0, \dots, x_n]^T, \quad b = [y_0, \dots, y_m]^T$$

che è un sistema lineare sovradeterminato. Risolviamo allora tale problema.

3. Soluzione ai M.Q. di un sistema lineare sovradeterminato

Sia $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ ed $m > n$. Ci chiediamo: $\exists x \in \mathbb{R}^n : Ax = b$?

$$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

$A \quad x \quad b$

Sia $\mathfrak{R}(A)$ lo spazio definito da:

$$\mathfrak{R}(A) = \{y \in \mathbf{R}^m : \exists x \in \mathbf{R}^n : Ax = y\}$$

con $rank(A) = \dim \mathfrak{R}(A)$. Si possono avere i seguenti casi:

$$\begin{array}{l} b \in \mathfrak{R}(A) \\ b \notin \mathfrak{R}(A) \end{array} \begin{cases} rank(A) = n & \exists_1 \text{ soluzione} \\ rank(A) < n & \infty \text{ soluzioni} \\ & \neg \exists \text{ soluzione} \end{cases}$$

Se $b \notin \mathfrak{R}(A)$, allora determiniamo un $\bar{x} \in \mathbf{R}^n$:

$$\|A\bar{x} - b\|_2 \leq \|Ax - b\|_2 \quad \forall x \in \mathbf{R}^n$$

Se tale $\bar{x} \in \mathbf{R}^n$ esiste, allora è la soluzione del sistema sovradeterminato nel senso dei m.q. Essa è la soluzione del sistema:

$$A^T Ax = A^T b$$

che ha la matrice dei coefficienti quadrata e pertanto tale sistema non è più sovradeterminato.

Osserviamo che $\bar{x} \in \mathbf{R}^n$ è la soluzione ai minimi quadrati di $Ax = b$, nel senso che \bar{x} rende minima la somma dei quadrati delle componenti del vettore resto $R = Ax - b$, cioè che rende minimo $R^T R$. Per dimostrare che la soluzione di tale problema di minimizzazione è data dalla soluzione del sistema lineare quadrato: $A^T Ax = A^T b$, consideriamo il seguente esempio.

Esempio

Siano $m = 3$ ed $n = 2$ e consideriamo il sistema:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \\ a_{31}x_1 + a_{32}x_2 = b_3 \end{cases}$$

Sia, inoltre, $\bar{R} = (r_1, r_2, r_3)^T$ il vettore resto. Allora il sistema si potrà scrivere:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 - b_1 = r_1 \\ a_{21}x_1 + a_{22}x_2 - b_2 = r_2 \\ a_{31}x_1 + a_{32}x_2 - b_3 = r_3 \end{cases}$$

Cerchiamo $\bar{x} = (x_1, x_2)$ tale che $r_1^2 + r_2^2 + r_3^2 = \text{minimo}$. Svolgendo i calcoli si ha che:

$$r_1^2 + r_2^2 + r_3^2 = (a_{11}^2 + a_{21}^2 + a_{31}^2)x_1^2 + (a_{12}^2 + a_{22}^2 + a_{32}^2)x_2^2 + 2(a_{11}a_{12} + a_{21}a_{22} + a_{31}a_{32})x_1x_2 + \\ - 2(a_{11}b_1 + a_{21}b_2 + a_{31}b_3)x_1 - 2(a_{12}b_1 + a_{22}b_2 + a_{32}b_3)x_2 + b_1^2 + b_2^2 + b_3^2$$

Affinché si abbia il minimo di questa quantità, dobbiamo imporre che siano nulle le derivate prime rispetto ad x_1 e x_2 . Ad esempio, calcolando la derivata prima rispetto ad x_1 ed imponendo essa che si annulli, otteniamo:

$$\frac{\partial(r_1^2 + r_2^2 + r_3^2)}{\partial x_1} = 2(a_{11}^2 + a_{21}^2 + a_{31}^2)x_1 + 2(a_{11}a_{12} + a_{21}a_{22} + a_{31}a_{32})x_2 - 2(a_{11}b_1 + a_{21}b_2 + a_{31}b_3) = 0 \quad (1)$$

Ponendo, in generale:

$$(a_i, a_j) = a_{1i} \cdot a_{1j} + \dots + a_{mi} \cdot a_{mj}$$

si potrà scrivere la (1) :

$$(a_1, a_1)x_1 + (a_1, a_2)x_2 = (a_1, b)$$

Ripetendo il ragionamento anche per la derivata rispetto ad x_2 , otterremo il sistema:

$$\begin{cases} (a_1, a_1)x_1 + (a_1, a_2)x_2 = (a_1, b) \\ (a_2, a_1)x_1 + (a_2, a_2)x_2 = (a_2, b) \end{cases}$$

Tale sistema, in forma matriciale, si scriverà:

$$A^T Ax = A^T b$$

che è ciò che volevamo ottenere.

Esempio di sistema lineare sovradeterminato

Consideriamo il sistema:

$$\begin{cases} x_1 - x_2 = 2 \\ x_1 + x_2 = 4 \\ 2x_1 + x_2 = 8 \end{cases}$$

La matrice dei coefficienti sarà:

$$A = \begin{vmatrix} 1 & -1 \\ 1 & 1 \\ 2 & 1 \end{vmatrix} \Rightarrow A^T = \begin{vmatrix} 1 & 1 & 2 \\ -1 & 1 & 1 \end{vmatrix} \Rightarrow A^T A = \begin{vmatrix} 1 & 1 & 2 \\ -1 & 1 & 1 \end{vmatrix} \begin{vmatrix} 1 & -1 \\ 1 & 1 \\ 2 & 1 \end{vmatrix} = \begin{vmatrix} 6 & 2 \\ 2 & 3 \end{vmatrix}$$

Considerando, poi, la colonna dei termini noti possiamo scrivere:

$$A^T b = \begin{vmatrix} 1 & 1 & 2 \\ -1 & 1 & 1 \end{vmatrix} \begin{vmatrix} 2 \\ 4 \\ 8 \end{vmatrix} = \begin{vmatrix} 22 \\ 10 \end{vmatrix}$$

Quindi il sistema da risolvere sarà:

$$\begin{vmatrix} 6 & 2 \\ 2 & 3 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} = \begin{vmatrix} 22 \\ 10 \end{vmatrix} \Rightarrow x_1 = \frac{23}{7} \text{ e } x_2 = \frac{8}{7}$$

Esempio di problema ai M.Q. discreto

Vediamo un esempio

Disegnare la retta che meglio approssima (M.Q.) i punti (0,1), (1,2.1), (2,2.9), (3,3.2). Dalle ipotesi del problema, dunque, si ha:

$$f(0) = 1, f(1) = 2.1, f(2) = 2.9, f(3) = 3.2$$

Sia $p(x) = a_0x + a_1$ un polinomio di primo grado di cui dobbiamo determinare i coefficienti a_0, a_1 tali che:

$$\sum_{i=0}^3 [f(x_i) - p(x_i)]^2 = \min$$

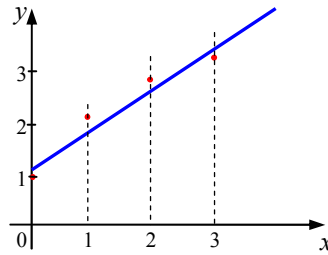
Imponiamo:

$$\begin{cases} a_1 = 1 \\ a_0 + a_1 = 2.1 \\ 2a_0 + a_1 = 2.9 \\ 3a_0 + a_1 = 3.2 \end{cases} \Leftrightarrow \begin{vmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{vmatrix} \begin{vmatrix} a_0 \\ a_1 \end{vmatrix} = \begin{vmatrix} 1 \\ 2.1 \\ 2.9 \\ 3.2 \end{vmatrix}$$

Da cui:

$$A^T A x = A^T b \Leftrightarrow \begin{vmatrix} 14 & 6 \\ 6 & 4 \end{vmatrix} \begin{vmatrix} a_0 \\ a_1 \end{vmatrix} = \begin{vmatrix} 17.5 \\ 9.2 \end{vmatrix}$$

Risolvendo il sistema otteniamo $a_0 = 0.74$ e $a_1 = 1.19$. Graficamente si ha:



Fitting lineare

Consideriamo il problema del fitting di dati ai minimi quadrati con una retta. Siano dati i punti (x_i, y_i) con $1 \leq i \leq n$; determinare a e b tali che:

$$F(a, b) = \sum_{i=1}^n [y_i - (ax_i + b)]^2 \text{ sia minima}$$

Imponiamo che le derivate prime rispetto ad a e b siano nulle:

$$\frac{\partial F}{\partial a} = -2 \sum_{i=1}^n [y_i - (ax_i + b)] x_i = 0$$

$$\frac{\partial F}{\partial b} = -2 \sum_{i=1}^n [y_i - (ax_i + b)] = 0$$

Il sistema da risolvere, dunque, sarà:

$$\begin{cases} \sum_{i=1}^n [y_i - (ax_i + b)] x_i = 0 \\ \sum_{i=1}^n [y_i - (ax_i + b)] = 0 \end{cases} \Rightarrow \begin{cases} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i \\ a \sum_{i=1}^n x_i + b \sum_{i=1}^n 1 = \sum_{i=1}^n y_i \end{cases}$$

Da cui:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}, \quad b = \frac{\left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{i=1}^n y_i \right) - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n x_i y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}$$

Esempio di fitting lineare

x	y
-----	-----

0	10
1	25
2	51
3	66
4	97
5	118

$$\sum_{i=1}^n x_i = 15, \quad \sum_{i=1}^n y_i = 367, \quad \sum_{i=1}^n x_i^2 = 55, \quad \sum_{i=1}^n x_i y_i = 1303 \quad \Rightarrow \quad a = 22.02, \quad b = 6.09$$

Esempio: fitting esponenziale

$$y_i = Ae^{rx_i}$$

$$z_i = \ln y_i = rx_i + \ln A$$

Svantaggi

Il metodo dei minimi quadrati per risolvere il problema discreto presenta i seguenti svantaggi:

- 1) la matrice $A^T A$ è spesso mal condizionata;
- 2) il metodo non è buono per risolvere il problema continuo.

Vediamo allora di trovare un'alternativa che eviti di risolvere un sistema con matrice malcondizionata e sia adatta a risolvere il problema continuo.

4. Spazi con prodotto interno

La migliore approssimazione nella norma 2 è correlata al concetto di ortogonalità che si basa sul prodotto interno.

Una funzione a valori reali $\langle \cdot, \cdot \rangle$ definita sul prodotto $V \times V$, dove V è uno spazio lineare sul campo dei numeri reali, è

detta *prodotto interno* se soddisfa le seguenti condizioni:

- 1) $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle \quad \forall f, g, h \in V$
- 2) $\langle \lambda f, g \rangle = \lambda \langle f, g \rangle \quad \forall \lambda \in \mathfrak{R}, \forall f, g \in V$
- 3) $\langle f, g \rangle = \langle g, f \rangle \quad \forall f, g \in V$
- 4) $\langle f, f \rangle > 0 \quad f \neq 0, f \in V$

Esempio

\mathfrak{R}^n è uno spazio con prodotto interno, con:

$$\langle \underline{x}, \underline{y} \rangle = \sum_{i=1}^n x_i y_i \quad \underline{x}, \underline{y} \in \mathfrak{R}^n$$

In maniera del tutto equivalente, se consideriamo $\underline{x} = (x_1, \dots, x_n)^T$ e $\underline{y} = (y_1, \dots, y_n)^T$ si può scrivere:

$$\langle \underline{x}, \underline{y} \rangle = \underline{x}^T \underline{y}$$

Definizione

Sia V uno spazio con prodotto interno ed $f, g \in V$ tali che $\langle f, g \rangle = 0$. Diremo che f è **ortogonale** a g .

Definizione

Sia V uno spazio con prodotto interno ed $f \in V$. Diremo:

$$\|f\| = \langle f, f \rangle^{1/2}$$

norma indotta.

La prova che questa espressione sia effettivamente una norma, si basa sulla disuguaglianza di Cauchy-Schwarz:

$$|\langle f, g \rangle| \leq \|f\| \|g\| \quad \forall f, g \in V$$

Teorema

Uno spazio a prodotto interno V su \mathfrak{R} con norma indotta è uno spazio lineare normato su \mathfrak{R} .

Esempio

$C[a, b]$ è uno spazio a prodotto interno, con:

$$\langle f, g \rangle = \int_a^b w(x) f(x) g(x) dx$$

con $w(x)$ funzione peso non negativa, continua ed integrabile in $[a, b]$. La norma indotta, invece, è data da:

$$\|f\|_2 = \sqrt{\int_a^b w(x)|f(x)|^2 dx}$$

N.b. non è necessaria la continuità di f in $[a,b]$ perché $\|f\|_2$ sia finita. Ad esempio:

$$f(x) = \text{sign}\left[x - \frac{1}{2}(a+b)\right] \quad \text{con } x \in [a,b]$$

ha $\|f\|_2 < \infty$, ma f è discontinua in $x = \frac{1}{2}(a+b)$.

Pertanto, per sviluppare una teoria la cui applicabilità vada oltre lo spazio lineare delle funzioni continue su (a,b) , denotiamo con $L_w^2(a,b)$ l'insieme delle funzioni f a valori reali definite su (a,b) tali che $w(x)|f(x)|^2$ sia integrabile su (a,b) . $L_w^2(a,b)$ è quindi con prodotto interno e norma 2 indotta. $C[a,b]$ è un sottoinsieme proprio di $L_w^2(a,b)$. Se $w(x) \equiv 1$ allora avremo $L^2(a,b)$.

5. Migliore approssimazione nella norma 2

Il problema continuo, ovvero la migliore approssimazione nella norma 2 consiste, dato $f \in L_w^2(a,b)$, nel trovare $p_n \in P_n$ tale che:

$$\|f - p_n\|_2 = \inf_{q \in P_n} \|f - q\|_2$$

p_n è detto **polinomio di migliore approssimazione** nella norma 2 di grado n ad f in (a,b) . Si può dimostrare che in uno spazio a prodotto interno, tale migliore approssimazione esiste ed è unica.

Vediamo adesso in che modo trovare la migliore approssimazione nel caso generale ed in un caso particolare.

Teorema

Sia V uno spazio a prodotto interno (ad esempio $V = L^2[0,1]$ oppure $V = C[0,1]$) e $W \subseteq V$ uno spazio a dimensione finita con $\dim W = n+1$ (ad esempio $W = P_n$). Sia, inoltre, $W = \text{span}\{\varphi_0, \varphi_1, \dots, \varphi_n\}$ ovvero:

$\forall f \in W, f = \sum_{i=0}^n a_i \varphi_i$; allora per $\forall f \in V$, $w^* = \sum_{i=0}^n a_i \varphi_i$ è la migliore approssimazione nel senso dei m.q. ed i coefficienti a_i sono dati dalla risoluzione del sistema:

$$\sum_{i=0}^n a_i \langle \varphi_i, \varphi_k \rangle = \langle f, \varphi_k \rangle \quad k = 0, \dots, n$$

Tale sistema è detto delle **equazioni normali**.

Dimostrazione:

Dimostriamo il teorema nel caso particolare in cui $V=C[a,b]$, $w(x) \equiv 1$ e sia $f \in C[a,b]$. Vogliamo trovare un polinomio $p(x) \in P_n$ tale che:

$$\int_a^b [f(x) - p(x)]^2 dx \quad \text{sia minimo}$$

Sia $p(x) = \sum_{i=0}^n a_i x^i$, definiamo una funzione F dei coefficienti a_i :

$$F(a_0, \dots, a_n) = \int_a^b \left[f(x) - \sum_{i=0}^n a_i x^i \right]^2 dx$$

Condizione necessaria affinché si abbia il minimo è che:

$$\frac{\partial F}{\partial a_k} = 0 \quad k = 0, \dots, n$$

Sviluppando la F otteniamo:

$$F = \int_a^b [f(x)]^2 dx - 2 \int_a^b f(x) \sum_{i=0}^n a_i x^i dx + \int_a^b \left[\sum_{i=0}^n a_i x^i \right]^2 dx$$

Calcoliamone adesso la derivata prima rispetto ad a_k :

$$\frac{\partial F}{\partial a_k} = 0 - 2 \int_a^b f(x) \delta_{ik} x^i dx + 2 \int_a^b \sum_{i=0}^n a_i x^{i+k} dx \quad \text{dove } \delta_{ik} = \begin{cases} 1, & \text{se } i = k \\ 0, & \text{altrimenti} \end{cases}$$

Imponendo che tale derivata sia nulla, otteniamo:

$$\sum_{i=0}^n a_i \int_a^b x^{i+k} dx = \int_a^b f(x) x^k dx \quad k = 0, \dots, n$$

Se $[a,b]=[0,1]$ la soluzione è data da:

$$\sum_{i=0}^n \frac{a_i}{i+k+1} = \int_0^1 f(x) x^k dx$$

La matrice $H = \left[\frac{1}{i+k+1} \right]$ con $i, k = 0, \dots, n$ è la matrice di Hilbert:

$$H = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \dots \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \dots \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Com'è noto, tale matrice è quasi singolare per n crescente, cioè è mal condizionata. Riportiamo una tabella da cui si evince la relazione tra l'ordine n e il suo numero di condizionamento K_2 .

n	$K_2(H_n)$
5	$4.8 \cdot 10^5$
10	$1.6 \cdot 10^{13}$
15	$6.1 \cdot 10^{20}$
20	$2.5 \cdot 10^{28}$
25	$1.0 \cdot 10^{36}$

dove: $K_2(H_n) = \|H_n\|_2 \cdot \|H_n^{-1}\|_2$

Pertanto la soluzione del sistema di equazioni normali perde rapidamente accuratezza a causa dell'accumularsi dell'errore di arrotondamento.

In virtù di quanto visto nel teorema appena dimostrato, è necessario trovare un metodo alternativo. Tale metodo si basa sull'uso dei polinomi ortogonali.

6. Polinomi ortogonali

Per evitare che le equazioni normali abbiano una matrice difficile da invertire, espandiamo p_n con una base differente da quella canonica, scelta in modo che la matrice risultante sia diagonale.

Definizione

Data una funzione peso $w(x)$ su (a,b) , diremo che la successione di polinomi φ_j , $j = 0, 1, \dots$ forma un sistema di polinomi ortogonali su (a,b) rispetto a w , se ogni φ_j di grado j è tale che:

$$\langle \varphi_j, \varphi_k \rangle = \int_a^b w(x) \varphi_j(x) \varphi_k(x) dx = \begin{cases} 0 & \forall k \neq j \\ \neq 0 & k = j \end{cases}$$

In particolare, se $\langle \varphi_j, \varphi_j \rangle = 1$ i polinomi si dicono **ortonormali**. L'insieme dei polinomi

ortogonali di grado n lo indicheremo con Π_n . Dalla costruzione vista in precedenza, si ha:

$$a_i = \langle f, \varphi_i \rangle = \int_a^b w(x) f(x) \varphi_i(x) dx \quad i = 0, \dots, n$$

Teorema

Se $\varphi_0, \varphi_1, \dots, \varphi_n$ sono tali che $\langle \varphi_i, \varphi_j \rangle = 0$ per $i \neq j$, allora sono linearmente indipendenti, cioè formano una base.

Viceversa, dati n polinomi linearmente indipendenti, è possibile trovare una loro combinazione lineare per cui essi risultino ortogonali.

Teorema

Se $p \in \Pi_n[a, b]$ allora esso ha n zeri reali, distinti ed interni ad $[a, b]$.

Dimostrazione.

Sia $n \in \mathbb{N}$, $n \geq 1$, $I \equiv [a, b]$ e consideriamo $\int_I p_i(x) p_j(x) dx = 0$, $i \neq j$. $p_n(x)$ non ha segno costante in I , altrimenti si avrebbe:

$$\int_I p_n(x) p_0(x) dx \neq 0$$

Per il teorema di esistenza degli zeri, esiste uno zero in $\overset{\circ}{I}$. Sia esso x_1 e

supponiamo, per assurdo, che sia di molteplicità almeno pari a 2. Ne segue:

$$r(x) = \frac{p_n(x)}{(x-x_1)^2} \in P_{n-2} \quad \Rightarrow \quad 0 = \int_I r(x) p_n(x) dx = \int \frac{p_n^2(x)}{(x-x_1)^2} dx \neq 0$$

Siamo pervenuti ad un assurdo, scaturito dall'aver supposto che x_1 sia uno zero di molteplicità almeno 2, quindi gli zeri sono semplici.

Supponiamo adesso, sempre per assurdo, che sia $m < n$; in questo caso possiamo scrivere:

$$p_n(x)[(x-x_1) \cdot \dots \cdot (x-x_m)] = r(x)[(x-x_1)^2 \cdot \dots \cdot (x-x_m)^2]$$

con $r(x) \in P_{n-m}$ di segno costante in I . Si ha:

$$0 = \int_I p_n(x) [(x-x_1) \cdot \dots \cdot (x-x_m)] = \int_I r(x) [(x-x_1)^2 \cdot \dots \cdot (x-x_m)^2] \neq 0$$

dove $(x-x_1) \cdot \dots \cdot (x-x_m) \in P_{m < n}$. Siamo pervenuti anche in questo caso ad un assurdo.

L'assurdo è scaturito dall'aver supposto $m < n$, quindi $m = n$. ■

7. Procedimento di ortogonalizzazione di Gram-Schmidt

L'insieme $S = \{1, x, x^2, \dots, x^n\}$ forma una base per P_n . I suoi elementi però non sono ortogonali rispetto a nessun prodotto interno. Poniamo $q_0(x) \equiv 1$ e definiamo:

$$p_0(x) = \frac{q_0(x)}{\|q_0\|}$$

Si osservi che $\|q_0\| \neq 1$, infatti:

$$\int_{-1}^1 q_0^2(x) dx = \|q_0\|^2 \quad \Rightarrow \quad \|q_0\| = \sqrt{2}$$

Per $k=1, \dots, n$ poniamo:

$$q_k(x) = x^k - \sum_{i=0}^{k-1} \langle x^k, p_i \rangle p_i(x) \quad \text{e} \quad p_k(x) = \frac{q_k(x)}{\|q_k\|} \quad (1)$$

Allora segue che:

$\{q_k(x)\}_{k=0}^n$ sono ortogonali

$\{p_k(x)\}_{k=0}^n$ sono ortonormali

Esempio

Consideriamo due funzioni $f, g \in L^2[-1, 1]$, con $\langle f, g \rangle = \int_{-1}^1 f(x)g(x)dx$ e sia $S = \{1, x, x^2\}$. Si ha:

$$q_0(x) = 1 \quad \text{e} \quad p_0(x) = \frac{1}{\sqrt{2}}$$

Calcoliamo q_k e p_k per $k=1, 2$. Applicando le (1) otteniamo:

$$q_1(x) = x - \langle x, \frac{1}{\sqrt{2}} \rangle \frac{1}{\sqrt{2}} = x \quad \text{e} \quad p_1(x) = \frac{x}{\|x\|} = \sqrt{\frac{3}{2}}x$$

$$q_2(x) = x^2 - \langle x^2, p_0 \rangle p_0 - \langle x^2, p_1 \rangle p_1 = x^2 - \langle x^2, \frac{1}{\sqrt{2}} \rangle \frac{1}{\sqrt{2}} - \langle x^2, \sqrt{\frac{3}{2}}x \rangle \sqrt{\frac{3}{2}}x = x^2 - \frac{1}{3}$$

$$p_2(x) = \frac{x^2 - \frac{1}{3}}{\left\| x^2 - \frac{1}{3} \right\|} = \sqrt{\frac{45}{8}} \left(x^2 - \frac{1}{3} \right)$$

avendo tenuto conto dei seguenti risultati:

$$\langle x, 1 \rangle = \int_{-1}^1 x dx = 0, \quad \|x\| = \sqrt{\int_{-1}^1 x^2 dx} = \sqrt{\frac{2}{3}}, \quad \left\| x^2 - \frac{1}{3} \right\| = \sqrt{\int_{-1}^1 \left(x^2 - \frac{1}{3} \right)^2 dx} = \sqrt{\frac{8}{45}}$$

Corollario

Sia $\{p_j\}_{j=0}^n$ un insieme di polinomi ortogonali, allora ogni $p(x)$ di grado $< k$ è tale che:

$$\langle p_i, p \rangle = 0 \quad i \geq k$$

Dimostrazione:

Se $r < n$, $p(x) \in P_r$. Quindi, poiché i $p_j(x)$ sono ortogonali, possiamo scrivere:

$$p(x) = \sum_{j=0}^r \langle p, p_j \rangle p_j(x)$$

da cui segue:

$$\langle p_n, p \rangle = \sum_{j=0}^r \langle p, p_j \rangle \langle p_n, p_j \rangle = 0$$

8. Calcolo efficiente dei polinomi ortogonali

Il procedimento di Gram-Schmidt può essere oneroso persino per piccoli n . Ad esempio, per trovare $p_{10}(x)$ è necessario il calcolo di $p_j(x)$ con $0 \leq j \leq 9$ ed ognuno di essi è usato esplicitamente. Al posto di tale procedimento, invece, si può utilizzare una formula di ricorrenza come quella dei polinomi di Chebichev.

Sia $p_n(x) \in P_n$ ortogonale rispetto a:

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

Siano:

$$\begin{aligned} p_n(x) &= a_n x^n + \dots + a_0 \\ p_{n-1}(x) &= b_{n-1} x^{n-1} + \dots + b_0 \end{aligned}$$

Poniamo $A_n = \frac{a_n}{b_{n-1}} \Rightarrow p_n(x) - A_n x p_{n-1}(x) \in P_{n-1}$. Infatti possiamo scrivere:

$$p_n(x) - A_n x p_{n-1}(x) = a_n x^n + \dots + a_0 - \frac{a_n}{b_{n-1}} x (b_{n-1} x^{n-1} + \dots + b_0) = a_n x^n + \dots + a_0 - a_n x^n - \dots - \frac{a_n b_0}{b_{n-1}} x$$

per cui semplificando il termine di grado n , rimane un polinomio di grado $n-1$. Pertanto esistono n costanti λ_i tali che:

$$p_n(x) - A_n x p_{n-1}(x) = \sum_{i=0}^{n-1} \lambda_i p_i(x) \quad (2)$$

Moltiplichiamo per $p_r(x)$ ed integriamo:

$$\int_a^b p_r(x) p_n(x) dx - A_n \int_a^b x p_r(x) p_{n-1}(x) dx = \sum_{i=0}^{n-1} \lambda_i \int_a^b p_r(x) p_i(x) dx$$

Per $0 \leq r \leq n-3$ l'uguaglianza sopra diventa:

$$0 = \lambda_r \int_a^b p_r^2(x) dx \Rightarrow \lambda_r = 0, \text{ essendo } \int_a^b p_r^2(x) dx \neq 0$$

Sostituendo nella (2) il risultato appena ottenuto, essendo nulli i λ_r per $0 \leq r \leq n-3$, nel lato destro rimangono solo i termini di grado $n-2$ ed $n-1$, cioè:

$$p_n(x) - A_n x p_{n-1}(x) = \lambda_{n-2} p_{n-2}(x) + \lambda_{n-1} p_{n-1}(x)$$

Ponendo $B_n = \lambda_{n-1}$ e $C_n = \lambda_{n-2}$ si ha:

$$p_n(x) = (A_n x + B_n) p_{n-1}(x) + C_n p_{n-2}(x)$$

Per ricavare B_n moltiplichiamo ambo i membri per $p_{n-1}(x)$ ed integriamo:

$$0 = \langle p_n, p_{n-1} \rangle = A_n \langle x p_{n-1}, p_{n-1} \rangle + B_n \langle p_{n-1}, p_{n-1} \rangle + C_n \langle p_{n-2}, p_{n-1} \rangle$$

da cui:

$$B_n = -\frac{A_n \langle x p_{n-1}, p_{n-1} \rangle}{\|p_{n-1}\|^2}$$

essendo $C_n \langle p_{n-2}, p_{n-1} \rangle = 0$. Analogamente, per ricavare C_n moltiplichiamo ambo i membri per $p_{n-2}(x)$ ed integriamo, ottenendo:

$$0 = \langle p_n, p_{n-2} \rangle = A_n \langle p_{n-1}, x p_{n-2} \rangle + B_n \langle p_{n-1}, p_{n-2} \rangle + C_n \langle p_{n-2}, p_{n-2} \rangle$$

da cui:

$$C_n = -\frac{A_n \langle p_{n-1}, x p_{n-2} \rangle}{\|p_{n-2}\|^2}$$

essendo $B_n \langle p_{n-1}, p_{n-2} \rangle = 0$.

Diamo adesso la dimostrazione del teorema (di Fourier) di esistenza ed unicità della migliore approssimazione ai minimi quadrati.

Teorema

Data $f \in L_w^2(a, b)$, $\exists_1 p_n^* \in P_n$:

$$\|f - p_n^*\|_2 = \min_{q \in P_n} \|f - q\|_2$$

e tale polinomio è dato da:

$$p_n^*(x) = \sum_{j=0}^n \langle f, p_j \rangle p_j(x)$$

dove $\{p_j(x)\}_{j=0}^n$ è l'insieme dei polinomi ortonormali definito rispetto al prodotto interno di $L_w^2(a, b)$.

Dimostrazione:

Sia $p(x) \in P_n$; poiché i $p_j(x)$ sono ortonormali, essi sono linearmente indipendenti e

quindi:

$$p(x) = \sum_{j=0}^n \alpha_j p_j(x)$$

Possiamo scrivere:

$$0 \leq \|f - p\|_2^2 = \langle f - p, f - p \rangle = \langle f - \sum_{j=0}^n \alpha_j p_j, f - \sum_{j=0}^n \alpha_j p_j \rangle = \langle f, f \rangle - 2 \sum_{j=0}^n \alpha_j \langle f, p_j \rangle + \sum_{j=0}^n \alpha_j^2 +$$

$$+ \sum_{j=0}^n \langle f, p_j \rangle^2 - \sum_{j=0}^n \langle f, p_j \rangle^2 = \langle f, f \rangle - \sum_{j=0}^n \langle f, p_j \rangle^2 + \sum_{j=0}^n (\alpha_j - \langle f, p_j \rangle)^2$$

avendo aggiunto e sottratto $\sum_{j=0}^n \langle f, p_j \rangle^2$ durante lo sviluppo dei termini. La quantità appena ottenuta è minima se:

$$\alpha_j = \langle f, p_j \rangle$$

Tali coefficienti α_j sono detti *coefficienti generalizzati di Fourier*.

Dal teorema appena dimostrato segue la *disuguaglianza di Bessel*:

$$\|f\|_2^2 \geq \sum_{j=0}^n \langle f, p_j \rangle^2$$

inoltre la funzione resto $f(x) - p_n^*(x)$ è ortogonale ad ogni $p \in P_n$. Infatti, poiché ogni $p \in P_n$ è combinazione lineare dei p_k polinomi ortonormali, si ha:

$$\langle f - p_n^*, p_k \rangle = \langle f, p_k \rangle - \sum_{j=0}^n \langle f, p_j \rangle \langle p_j, p_k \rangle = \langle f, p_k \rangle - \langle f, p_k \rangle = 0$$

essendo:

$$\langle p_j, p_k \rangle = \delta_{jk} = \begin{cases} 1, & \text{se } j = k \\ 0, & \text{altrimenti} \end{cases}$$

9. Procedura per il calcolo del polinomio di M.A. ai M.O.

La procedura per calcolare il polinomio di migliore approssimazione ai minimi quadrati è formata dai seguenti passi:

- a) Generare l'insieme $\{p_j(x)\}_{j=0}^n$ con Gram-Schmidt.
- b) Calcolare i coefficienti $\langle f, p_j \rangle$ con $0 \leq j \leq n$.
- c) Costruire $p_n^*(x)$ utilizzando la relazione $p_n^*(x) = \sum_{j=0}^n \langle f, p_j \rangle p_j(x)$

Esempio

Consideriamo la funzione $f(x) = \cos(\pi x)$. Calcolando $p_2^* \in P_2$ con la procedura appena descritta, otteniamo:

$$p_2^*(x) = -\frac{45}{2\pi^2} \left(x^2 - \frac{1}{3} \right)$$

10. Errore dell'approssimazione ai minimi quadrati

Dal teorema di Weierstrass sappiamo che:

$$\lim_{n \rightarrow \infty} \|f - p_n^*\|_2 = 0$$

ma:

$$\begin{aligned} \|f - p_n^*\|_2^2 &= \langle f - p_n^*, f - p_n^* \rangle = \langle f - p_n^*, f \rangle - \langle f - p_n^*, p_n^* \rangle = \langle f, f \rangle - \langle p_n^*, f \rangle = \\ &= \|f\|^2 - \sum_{j=0}^n \langle f, p_j \rangle \langle f, p_j \rangle = \|f\|^2 - \sum_{j=0}^n \langle f, p_j \rangle^2 \geq 0 \end{aligned} \quad (3)$$

per la disuguaglianza di Bessel. Poiché, come detto, $\lim_{n \rightarrow \infty} \|f - p_n^*\|_2 = 0$ dalla (3) otteniamo:

$$\|f\|^2 = \sum_{j=0}^{\infty} \langle f, p_j \rangle^2$$

nota come identità di Parseval. Quindi:

$$\|f - p_n^*\|_2^2 = \sum_{j=n+1}^{\infty} \langle f, p_j \rangle^2$$

che dà l'errore cercato. Tale errore dipende dai coefficienti di Fourier, per $j = n+1, n+2, \dots$

L'errore puntuale può essere calcolato applicando il seguente:

Teorema

Sia $p_n^*(x)$ il polinomio ai minimi quadrati di $f \in C^2[-1,1]$. Allora:

$$\forall \varepsilon > 0, \forall x \in [-1,1] \Rightarrow |f(x) - p_n^*(x)| \leq \frac{\varepsilon}{\sqrt{n}}$$

11. Polinomi di Legendre

I polinomi di Legendre sono polinomi ortogonali sull'intervallo $[-1,1]$, con la funzione

peso $w(x) \equiv 1$:

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x)dx$$

Sfruttando le relazioni di ricorrenza si ha:

$$q_k(x) = (x - a_k)q_{k-1}(x) - b_k q_{k-2}(x) \quad k \geq 2, b_1 = 0$$

I coefficienti varranno:

$$a_k = \frac{\langle xq_{k-1}, q_{k-1} \rangle}{\langle q_{k-1}, q_{k-1} \rangle}$$

$$b_k = \frac{\langle q_{k-1}, xq_{k-2} \rangle}{\langle q_{k-2}, q_{k-2} \rangle}$$

Poniamo:

$$q_0 \equiv 1 \Rightarrow \langle q_0, q_0 \rangle = \int_{-1}^1 dx = 2$$

$$q_1(x) = (x - a_1)q_0(x); \quad a_1 = \frac{\langle xq_0, q_0 \rangle}{\langle q_0, q_0 \rangle} = 0 \Rightarrow q_1(x) = x$$

$$q_2(x) = (x - a_2)q_1(x) - b_2q_0(x); \quad a_2 = \frac{\langle xq_1, q_1 \rangle}{\langle q_1, q_1 \rangle} = 0, \quad b_2 = \frac{\langle q_1, xq_0 \rangle}{\langle q_0, q_0 \rangle} = \frac{1}{3} \Rightarrow q_2(x) = x^2 - \frac{1}{3}$$

Ragionando in maniera analoga troviamo:

$$q_3(x) = x^3 - \frac{3}{5}x$$

12. Approssimazione MinMax

In accordo al teorema di Weierstrass, ogni funzione $f \in C[a, b]$ può essere approssimata arbitrariamente bene dall'insieme di tutti i polinomi. Se però restringiamo il campo all'insieme dei polinomi di grado n fissato, quanto detto non è più vero. Ci chiediamo quindi con quale accuratezza può essere approssimata una $f \in C[a, b]$ da polinomi di grado n fissato. Il problema è, quindi, quello di cercare, data $f \in C[a, b]$ ed $n \geq 0$, un polinomio $p_n \in P_n$ tale che:

$$\|f - p_n\|_\infty = \inf_{q \in P_n} \|f - q\|_\infty$$

Tale polinomio è detto di migliore approssimazione di grado n ad f nella norma infinita. Poiché p_n minimizza il massimo dell'errore, è detto polinomio MinMax. Tale polinomio ha una caratteristica particolare che ne permette, in alcuni casi semplici, una costruzione diretta. La suddetta caratteristica è espressa nel teorema seguente.

Teorema de La Vallée-Poussin

Sia $f \in C[a, b]$, $p_n \in P_n$ e supponiamo che esistano $n+2$ punti x_0, \dots, x_{n+1} in $[a, b]$ tali che

$f(x_i) - p_n(x_i)$ e $f(x_{i+1}) - p_n(x_{i+1})$ abbiano segno alterno, per $i = 0, \dots, n$. Allora:

$$\min_{i=0, \dots, n+1} |f(x_i) - p_n(x_i)| \leq \min_{q \in P_n} \|f - q\|_\infty$$

Il teorema seguente afferma che, se le quantità $|f(x_i) - p_n(x_i)|$ sono tutte uguali a

$\|f - p_n\|_\infty$, allora p_n è un polinomio MinMax di grado n per f in $[a, b]$.

Teorema di equioscillazione di Chebichev

Sia $f \in C[a, b]$. Allora esiste un unico polinomio $p_n^*(x) \in P_n$ di MinMax per f in $[a, b]$ ed esso è caratterizzato dalla seguente proprietà: esistono $n+2$ punti x_i , per $i = 0, \dots, n+1$, con $a \leq x_0 < \dots < x_{n+1} \leq b$ tali che:

$$|f(x_i) - p_n^*(x_i)| = \|f - p_n^*\|_\infty \quad i = 0, \dots, n+1$$

ed inoltre:

$$f(x_i) - p_n^*(x_i) = -[f(x_{i+1}) - p_n^*(x_{i+1})] \quad i = 0, \dots, n$$

cioè $f - p_n^*$ ha il suo massimo valore assoluto con segni alterni nei punti x_i detti punti critici.

Tutti questi teoremi caratterizzano $p_n^*(x)$, ma non indicano come costruirlo. Applichiamo il teorema di equioscillazione per calcolarlo in un caso semplice.

Esempio

Consideriamo la funzione $f(x) = \sqrt{x}$, con $x \in [a, b]$, $a > 0$, $n = 1$.

Sia $p_n^*(x) = a_1 + a_2x$. La funzione resto sarà:

$$r(x) = f(x) - p_n^*(x) = \sqrt{x} - a_1 - a_2x$$

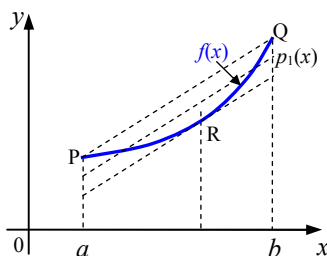
Per calcolare gli estremi di $r(x)$, annulliamone la derivata prima:

$$\frac{\partial r}{\partial x} = \frac{1}{2\sqrt{x}} - a_2 = 0 \Rightarrow x = \frac{1}{4a_2^2}$$

Poiché $r'(x)$ è monotona crescente, si annulla solo in x . Quindi gli altri due punti critici sono a e b . Si ha, allora il seguente sistema:

$$\begin{cases} \sqrt{a} - a_1 - a_2 \cdot a = D \\ \frac{1}{2a_2} - a_1 - \frac{1}{4a_2^2} = -D \\ \sqrt{b} - a_1 - a_2 \cdot b = D \end{cases}$$

che può essere risolto esplicitamente. Consideriamo la seguente figura:



Tale figura dà l'interpretazione grafica del procedimento precedente. R è il punto in cui la tangente alla curva $y = f(x)$ è parallela alla corda PQ. Il grafico di $p_1(x)$ è parallelo a queste due linee e sta a metà tra esse. Per funzioni arbitrarie, però, non esistono procedure che diano tale approssimazione in un numero finito di passi.

13. Polinomi di Chebichev

Tali polinomi sono definiti da:

$$T_n(x) = \cos(n \arccos x) \quad (4)$$

A tali polinomi è possibile dare una forma più semplice se consideriamo, nell'intervallo $[-1, 1]$, il cambiamento di variabile $x = \cos \vartheta$. Infatti, la (4) diventa:

$$T_n(\cos \vartheta) = \cos(n\vartheta)$$

Anche le formule di ricorrenza, per tali polinomi, assumono una forma piuttosto semplice:

$$\begin{cases} T_0(x) = 1, T_1(x) = x \\ T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \end{cases}$$

Sviluppandole, otteniamo:

$$T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x, \quad T_4(x) = 8x^4 - 8x^2 + 1, \dots$$

Calcoliamo adesso gli zeri:

$$\cos(k\vartheta_i) = 0 \Rightarrow k\vartheta_i = \frac{\pi}{2} + i\pi \Rightarrow \vartheta_i = \frac{2i+1}{2k}\pi$$

Quindi, ritornando alla variabile x , otteniamo:

$$x_i = \cos\left(\frac{2i+1}{2k}\pi\right) \quad i = 0, \dots, k-1$$

Tali polinomi formano un sistema ortogonale in $[-1, 1]$ rispetto alla funzione peso:

$$w(x) = (1-x^2)^{-1/2}$$

cioè:

$$\langle T_m, T_n \rangle = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} \cos(m \arccos x) \cdot \cos(n \arccos x) dx = \begin{cases} 0 & \text{se } m \neq n \\ \frac{\pi}{2} & \text{se } m = n \end{cases}$$

Tra tutti i polinomi monici di grado $n+1$, $n \geq 0$, il polinomio $\frac{T_{n+1}}{2^n}$ è quello che ha la più piccola norma del massimo su $[-1, 1]$. Inoltre, l'importanza di tali polinomi è legata alla scelta dei nodi nell'interpolazione lagrangiana.

Definiamo approssimazione quasi-MinMax, un polinomio $p_n \in P_n$ tale che $f(x) - p_n(x)$ cambi segno in $n+1$ punti ξ_j , $j = 0, \dots, n$, $a < \xi_0 < \dots < \xi_n < b$. Poiché in tal caso questa differenza ha max e min locali con segno alternante, i suddetti punti sono gli x_i , $i = 0, \dots, n+1$ richiesti dal teorema de La Vallée-Poussin. Quindi la norma infinita dell'errore del polinomio MinMax sta tra il più grande ed il più piccolo dei valori assoluti di tali max e min locali.

Poiché sappiamo che:

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\eta)}{(n+1)!} w_{n+1}(x) \quad \text{con} \quad w_{n+1}(x) = (x - \xi_0) \cdot \dots \cdot (x - \xi_n)$$

allora se $f^{(n+1)}$ è tale da non cambiare segno in $[a, b]$, ne segue che il polinomio ottenuto

interpolando la f negli ξ_i , $i = 0, \dots, n$ zeri dei polinomi di Chebichev ottenuti trasformando $[-1, 1]$ in $[a, b]$, è un polinomio di quasi MinMax.

N.b. se gli ξ_i fossero ugualmente spazati, l'errore cambierebbe segno solo negli $n-1$ punti interni ad $[a, b]$ e quindi il polinomio non sarebbe di quasi MinMax.

Esempio

Vogliamo approssimare e^x con $p \in P_3$, $x \in [-1, 1]$, $x_0 = 0$. Ricordiamo l'espansione in serie di Taylor di una funzione $f \in C^{n+1}[a, b]$ che ci dà la possibilità di approssimare la f con $p_n \in P_n$:

$$p_n(x) = f(x_0) + (x - x_0)f'(x_0) + \dots + \frac{(x - x_0)^n}{n!} f^{(n)}(x_0) \quad \text{con } x, x_0 \in [a, b]$$

Per l'esempio che stiamo considerando ($f(x) = e^x$), possiamo scrivere:

$$e^{kx} - p_n(x) = \sum_{r=n+1}^{\infty} \frac{k^r x^r}{r!}$$

Su $[0, 1]$, ad esempio, tale differenza è non negativa e monotona crescente, pertanto non cambia segno e non è quindi di quasi-MinMax. Ciononostante, basta scegliere n sufficientemente grande perché l'errore diminuisca. Questa procedura è però inefficiente.

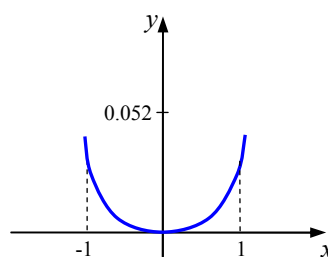
Infatti, nel nostro esempio otteniamo:

$$p_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

per cui l'errore sarà dato da:

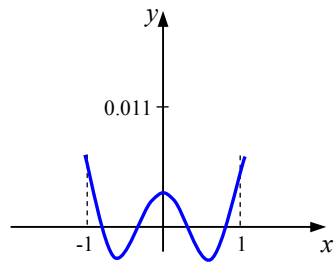
$$e^x - p_3(x) = \frac{x^4}{24} e^{\xi} \quad 0 < \xi < x \quad (x < \xi < 0)$$

Graficamente:

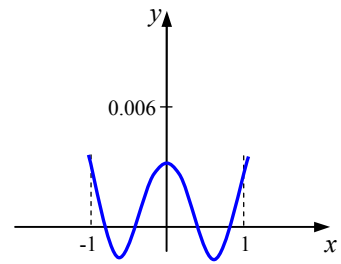


Curva dell'errore di Taylor

Altre curve dell'errore sono quelle indicate nelle seguenti figure:



Curva dell'errore minimi quadrati cubica



curva dell'errore di Chebichev cubica