MINIMI QUADRATI

Introduzione

E' un problema di approssimazione di una funzione $f(x)$ con una funzione più semplice $g(x)$ quando, come succede negli esperimenti scientifici, non abbiamo una valutazione esatta della $f(x)$ nei nodi x_i, per cui si deve cercare di minimizzare l'errore in tali nodi. Se la funzione più semplice è un polinomio, l'obiettivo sarà quello di trovare un polinomio $p(x)$ di grado n che dia la migliore approssimazione alla funzione $f(x)$ in un senso da specificare.

In generale, diremo che $g(x)$ approssima $f(x)$, e scriveremo $g(x) \sim f(x)$,

se:

$$\|f - g\| \quad \text{è "piccola"}$$

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 $\Re$. Una *norma* $\|\cdot\|$ su V è una funzione non negativa, definita su V, che gode delle seguenti proprietà:

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

2) $\|\lambda f\| = |\lambda| \, \|f\|, \; \forall \lambda \in \Re, \; \forall f \in V$

3) $\|f + g\| \le \|f\| + \|g\| \quad \forall f, g \in V$

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

Esempio

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

$$\frac{1}{\sqrt{n}}\|\underline{v}\|_\infty \le \|\underline{v}\|_2 \le \sqrt{n}\|\underline{v}\|_\infty \qquad \forall \underline{v} \in \Re^n$$

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.

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

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

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

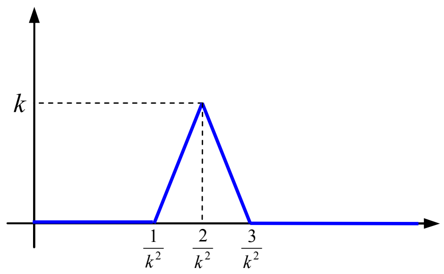
Tali norme non sono equivalenti. Consideriamo un esempio che mostra in modo pratico la non equivalenza di queste norme.

Esempio

Sia data una funzione f definita nel modo seguente:

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

Graficamente la funzione ha il seguente andamento:



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

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

Passando al limite:

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

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 φ_i, $i=1,...,n$ n funzioni linearmente indipendenti.

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

$$\left\|f-w^*\right\|\le\left\|f-w\right\| \qquad \forall w\in W$$

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 (v.par.3).

3) Dipende dalla norma scelta

La costruzione MinQuad $\| \ . \ \|_2$ si può fare in un numero finito di passi,

la costruzione MinMax $\| \ . \ \|_\infty$ dà luogo ad un procedimento iterativo.

Nonostante le differenze delle norme, c'è una caratteristica comune ai vari tipi di approssimazione, indipendentemente dalla norma scelta.

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:

<u>*Teorema di approssimazione di Weierstrass*</u>

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

Ovvero, ogni funzione continua puo' 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.

Problemi ai Minimi Quadrati

1) ***Problema discreto***: dati i punti $(x_0, y_0),...,(x_m, y_m)$ 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 $w_0,...,w_m$ costanti positive dette pesi.

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

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

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 $1, t, t^2, ..., t^n$ una base di P_n. Sia $m > n$; il polinomio cercato sarà del tipo:

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

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:

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

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

2. ***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: $\exists x \in \mathbb{R}^n : Ax = b$?

$$\underset{A}{\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}} \underset{x}{\begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}} \underset{=}{=} \underset{b}{\begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}}$$

Sia $\Re(A)$ lo spazio definito da:

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

con $rank(A) = \dim \Re(A)$. Si possono avere i seguenti casi:

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

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

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

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

$$\boxed{A^T Ax = A^T b}$$

che ha la matrice dei coefficienti quadrata e pertanto tale sistema non e' piu' sovradeterminato. Osserviamo che $\bar{x} \in \mathbb{R}^n$ 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 $r_1^2 + r_2^2 + r_3^2 =$ 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 questa quantità sia minima, 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 che essa si annulli, otteniamo:

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

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:

$$A^T A x = A^T b$$

che e' cio' 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:

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

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:

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

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

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

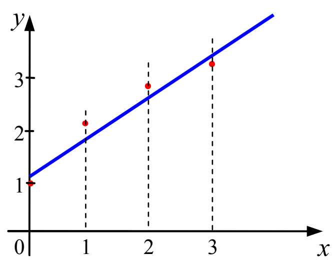
Imponiamo:

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

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 $1 \le i \le n$; determinare a e b tali che:

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

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

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

Il sistema da risolvere, dunque, sarà:

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

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

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

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.

3. 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 $<\cdot,\cdot>$ definita sul prodotto VxV, dove V è uno spazio lineare sul campo dei numeri reali, è detta *prodotto interno* se soddisfa le seguenti condizioni:

1) $<f+g,h>=<f,h>+<g,h> \quad \forall f,g,h \in V$

2) $<\lambda f,g>=\lambda<f,g> \quad \forall \lambda \in \Re,\ \forall f,g \in V$

3) $<f,g>=<g,f> \quad \forall f,g \in V$

4) $<f,f>>0 \quad f \neq 0,\ f \in V$

Esempio

$\Re^n$ è uno spazio con prodotto interno, con:

$$<\underline{x},\underline{y}>=\sum_{i=1}^{n} x_i y_i \qquad \underline{x},\underline{y} \in \Re^n$$

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:

$$<\underline{x},\underline{y}>=\underline{x}^T\underline{y}$$

Definizione

Sia V uno spazio con prodotto interno ed $f,g \in V$ tali che $<f,g>=0$. 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:

$$|\langle f, g \rangle| \le \|f\| \|g\| \qquad \forall f, g \in V$$

Teorema

Uno spazio a prodotto interno V su $\Re$ con norma indotta è uno spazio lineare normato su $\Re$.

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:

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

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 $L^2(a,b)$.

4. ***Migliore approssimazione nella norma 2***

Il problema continuo, ovvero la migliore approssimazione nella norma 2 consiste, dato $f \in L^2_w(a,b)$, 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, $W = span\{\varphi_0, \varphi_1, ..., \varphi_n\}$ ovvero: $\forall g \in W, g = \sum_{i=0}^{n} b_i \varphi_i$ cioè ogni elemento di W può essere scritto come combinazione lineare degli φ_i (si pensi, ad esempio, ai polinomi e agli elementi della base canonica x^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:

$$\sum_{i=0}^{n} a_i < \varphi_i, \varphi_k > = < f, \varphi_k > \qquad k = 0,...,n$$

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:

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

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 :

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

Imponendo che tale derivata sia nulla, otteniamo:

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

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:

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

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.

5. ***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, $j = 0,1,...$ forma un sistema di polinomi ortogonali su (a,b) rispetto a w, se ogni φ_j di grado j è tale che:

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

In particolare, se $< \varphi_j, \varphi_j > = 1$ 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:

$$a_i = < f, \varphi_i > = \int_a^b w(x) f(x)\varphi_i(x)dx \quad i = 0,...,n$$

Teorema

Se $\varphi_0, \varphi_1, ..., \varphi_n$ sono tali che $<\varphi_i, \varphi_j> = 0$ 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 $\overset{o}{I}$. Sia esso x_1 e supponiamo, per assurdo, che sia di molteplicità almeno pari a 2. Ne segue:

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

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 il numero degli zeri reali sia $m < n$; in questo caso possiamo scrivere:

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

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

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

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

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

6. ***Procedimento di ortogonalizzazione di Gram-Schmidt***

L'insieme $S = \{1, x, x^2, ..., x^n\}$ 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 se, ad esempio, [a,b]=[-1,1], si ha:

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

Per $k = 1,...,n$ poniamo:

$$q_k(x) = x^k - \sum_{i=0}^{k-1} < x^k, p_i > p_i(x) \qquad \text{e} \qquad p_k(x) = \frac{q_k(x)}{\| q_k \|} \qquad (1)$$

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 $< f, g > = \int_{-1}^{1} f(x)g(x)dx$ e sia $S = \{1, x, x^2\}$. Si ha:

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

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

$$q_1(x) = x - < x, \frac{1}{\sqrt{2}} > \frac{1}{\sqrt{2}} = x \qquad \text{e} \qquad p_1(x) = \frac{x}{\|x\|} = \sqrt{\frac{3}{2}}x$$

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

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

avendo tenuto conto dei seguenti risultati:

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

*Corollario **

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

$$< p_i, p >= 0 \qquad i \geq k$$

Dimostrazione:

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

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

da cui segue:

$$< p_n, p >= \sum_{j=0}^{r} < p, p_j >< p_n, p_j >= 0$$

7. 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:

$$< f, g >= \int_I f(x) g(x) dx$$

Siano:

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

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

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

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

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

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

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

Sostituendo nella (2) il risultato appena ottenuto, essendo nulli i λ_r per $0 \le r \le n-3$, 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:

$$0 = < p_n, p_{n-1} > = A_n < x p_{n-1}, p_{n-1} > + B_n < p_{n-1}, p_{n-1} > + C_n < p_{n-2}, p_{n-1} >$$

da cui:

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

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

$$0 = < p_n, p_{n-2} > = A_n < p_{n-1}, x p_{n-2} > + B_n < p_{n-1}, p_{n-2} > + C_n < p_{n-2}, p_{n-2} >$$

da cui:

$$C_n = -\frac{A_n < p_{n-1}, xp_{n-2} >}{\|p_{n-2}\|^2}$$

essendo $B_n < p_{n-1}, p_{n-2} >= 0$. •

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

Teorema

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

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

e tale polinomio è dato da:

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

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

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:

$$0 \le \| f - p \|_2^2 =< f - p, f - p >=< f - \sum_{j=0}^{n} \alpha_j p_j, f - \sum_{j=0}^{n} \alpha_j p_j >=< f, f > -2\sum_{j=0}^{n} \alpha_j < f, p_j > + \sum_{j=0}^{n} \alpha_j^2 +$$

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

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

$$\alpha_j = < f, p_j >$$

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

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

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

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:

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

essendo:

$$< p_j, p_k > = \delta_{jk} = \begin{cases} 1, \; se \; j = k \\ 0, \; altrimenti \end{cases}$$

8. <u>Errore dell'approssimazione ai minimi quadrati</u>

Dal teorema di Weierstrass sappiamo che:

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

ma:

$$\|f - p_n^*\|_2^2 = < f - p_n^*, f - p_n^* > = < f - p_n^*, f > - < f - p_n^*, p_n^* > = < f, f > - < p_n^*, f > =$$

$$= \|f\|^2 - \sum_{j=0}^{n} < f, p_j >< f, p_j > = \|f\|^2 - \sum_{j=0}^{n} < f, p_j >^2 \geq 0 \qquad (3)$$

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

$$\|f\|^2 = \sum_{j=0}^{\infty} < f, p_j >^2$$

nota come <u>*identità di Parseval*</u>. Quindi:

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

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

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:

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

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

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 $< f, p_j >$ con $0 \leq j \leq n$.

c) Costruire $p_n^*(x)$ utilizzando la relazione $p_n^*(x) = \sum_{j=0}^n < f, p_j > p_j(x)$

Esempio

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

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

Come esempi di polinomi ortogonali illustriamo i polinomi di Legendre e i polinomi di Chebichev.

10. Polinomi di Legendre

I polinomi di Legendre sono polinomi ortogonali sull'intervallo [-1,1], con la funzione peso $w(x) \equiv 1$:

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

Sfruttando le relazioni di ricorrenza si ha:

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

I coefficienti varranno:

$$a_k = \frac{< xq_{k-1}, q_{k-1} >}{< q_{k-1}, q_{k-1} >}$$

$$b_k = \frac{< q_{k-1}, xq_{k-2} >}{< q_{k-2}, q_{k-2} >}$$

Poniamo:

$$q_0 \equiv 1 \Rightarrow < q_0, q_0 > = \int_{-1}^{1} dx = 2$$

$$q_1(x) = (x - a_1)q_0(x);\ a_1 = \frac{< xq_0, q_0 >}{< q_0, q_0 >} = 0 \Rightarrow q_1(x) = x$$

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

Ragionando in maniera analoga troviamo:

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

11. Polinomi di Chebichev

Tali polinomi sono definiti da:

$$T_n(x) = \cos(n \arccos x) \qquad n \in N \qquad (4)$$

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:

$$T_n(\cos\vartheta) = \cos(n\vartheta) \qquad \vartheta \in [0, \pi]$$

Mostriamo che $T_n(x) \in P_n$ ed è quindi definito per $\forall\ x \in R$. Infatti:

$$\cos((k+1)\ \vartheta) = \cos k\vartheta \cos\vartheta - \mathrm{sen}\, k\vartheta\, \mathrm{sen}\vartheta$$

$$\cos((k-1)\ \vartheta) = \cos k\vartheta \cos\vartheta + \mathrm{sen}\, k\vartheta\, \mathrm{sen}\vartheta$$

a. $\cos((k+1)\ \vartheta) + \cos((k-1)\ \vartheta) = 2\cos(k\vartheta)\cos\vartheta$

da cui si ha la relazione di ricorrenza:

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x)$$

Pertanto, 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:

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

Notiamo che il coefficiente di x^k in $T_k(x)$ è 2^{k-1}.

Calcoliamo adesso gli zeri di $T_k(x)$:

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

Quindi, ritornando alla variabile *x*, otteniamo:

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

Poiché $T_k(x) \in P_k$, gli zeri trovati sono i soli zeri di $T_k(x)$ ed essi sono distinti. I punti di [-1,1] in cui $|T_k(x)| = 1$ sono di estremo relativo e sono dati da:

$$\cos(i\pi) = (-1)^i \Rightarrow \phi_i = \frac{i\pi}{k} \quad i = 0, \dots, k$$

Si ha: $$T_k(-x) = (-1)^k T_k(x).$$

Tali polinomi formano un sistema ortogonale in [-1,1] con la funzione peso:

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

Rispetto al prodotto interno:

$$\langle f, g\rangle = \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}}dx$$

e quindi si ha:

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

Dimostriamo adesso che, nell'intervallo [-1, 1], la $W(x) = \prod_{i=0}^{n}(x - x_i)$ ha $\|W\|_\infty$ minima se gli x_i sono gli zeri dei polinomi di Chebichev ovvero che 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].

Teorema. Sia: $W(x) = \prod_{i=0}^{n}(x - x_i) \in P_{n+1}$. Tra le possibili scelte dei nodi $\{x_i\}_{i=0}^{n}$, $x_i \in [-1, 1]$, $\|W\|_\infty = \max_{x\in[-1,1]} |W(x)|$ è minima se: $W(x) = \frac{T_{n+1}(x)}{2^n}$ cioè gli x_i sono gli zeri di $T_{n+1}(x)$.

Dimostrazione. $T_{n+1}(x)$ ha il coefficiente di x^{n+1} dato da 2^n. Allora, tra tutti i polinomi del tipo:

$x^{n+1} + \dots$, (noti come polinomi monici), $\frac{T_{n+1}(x)}{2^n}$ è un candidato per il minimo di $\|W\|_\infty$.

Poniamo quindi:

$$W(x) = \frac{T_{n+1}}{2^n} = (x - x_0)\dots(x - x_n) \text{ con:}$$

$$x_i = \cos\frac{2i+1}{2(n+1)}\pi$$

Si ha: $\|W\|_\infty = \frac{1}{2^n}$ poiché $\|T_{n+1}\|_\infty = 1$ se $x \in [-1,1]$.

Se poniamo: $y_i = \cos\frac{i\pi}{n+1}$ $i = 0, \ldots, n+1$ si ha:

$$W(y_i) = (-1)^i\frac{1}{2^n} \quad \text{e pertanto:} \quad W(y_{i+1}) = -W(y_i).$$

Supponiamo ora che $\exists$ $V \in P_{n+1}$ monico e tale che $\|V\|_\infty < \|W\|_\infty$. Si avra':

per i pari: $V(y_i) < W(y_i)$

per i dispari: $V(y_i) > W(y_i)$

Pertanto: $V(y_0) < W(y_0)$, $V(y_1) > W(y_1)$, …

Definendo $H(x) = V(x) - W(x)$ si ha che $H(x) \in P_n$ poiché V e W hanno lo stesso x^{n+1}. Ma $H(x)$ ha n+1 zeri, il che è assurdo e da ciò la tesi. •

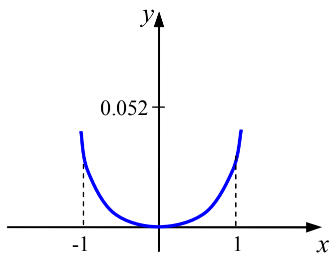
Come esempio di approssimazione, supponiamo di voler approssimare la funzione e^x e per fare ciò usiamo l'approssimazione con la serie di Taylor troncata al terzo ordine,

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

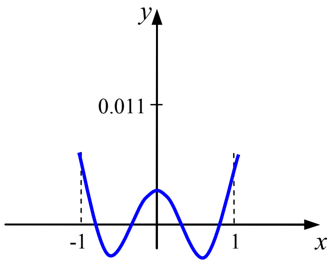
per cui l'errore sarà dato da:

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

e le approssimazioni ai minimi quadrati e quella con i polinomi di Chebichev come base, sempre con polinomi di terzo grado. Si ottengono le seguenti curve dell'errore:



Curva dell'errore di Taylor



Curva dell'errore