

Polinomi ortogonali

Alvise Sommariva

Università degli Studi di Padova
Dipartimento di Matematica

19 marzo 2014

Il problema ai minimi quadrati

Consideriamo lo spazio normato delle funzioni reali misurabili (cf. [6, p.284]) quadrato integrabili ($L^2(a, b), \|\cdot\|_2$) dove a, b è un intervallo della retta reale, non necessariamente limitato (cf. [6, p.386], ricordando [6, p.308]), e

$$\|g\|_2 = (g, g), \quad (f, g)_2 = \int_a^b f(x) \cdot g(x) dx.$$

Si dimostra (non facile!) che questo spazio euclideo è un esempio di **spazio di Hilbert** (cf. [6, p.388]), cioè uno spazio euclideo che è completo, separabile e infinito dimensionale (cf. [6, p.155]).

Il problema ai minimi quadrati

Più in generale se $w : (a, b) \rightarrow \mathbb{R}$ è una funzione positiva allora lo spazio $(L_w^2(a, b), \|\cdot\|_{2,w})$ definito come

$$L_w^2(a, b) = \left\{ f \text{ misurabili t.c. } \int_a^b |f(x)|^2 w(x) dx < \infty \right\}$$

è uno **spazio di Hilbert** dotato del prodotto scalare

$$(f, g)_{2,w} = \int_a^b f(x) \cdot g(x) w(x) dx$$

(cf. [2, p.23]).

Il problema ai minimi quadrati

Supponiamo di seguito che in particolare sia $w : (a, b) \rightarrow \mathbb{R}$ una funzione **nonnegativa**, con (a, b) non necessariamente limitato, tale che

1. $\int_a^b |x|^n w(x) dx < +\infty$ per tutti gli $n \in \mathbb{N}$;
2. $\int_a^b g(x) w(x) dx = 0$ per qualche funzione continua e non negativa g implica $g \equiv 0$ in (a, b) .

Il problema ai minimi quadrati

Teorema

Lo spazio vettoriale $L_w^2(a, b)$ contiene lo spazio dei polinomi $\mathcal{P}_n$ di grado n (con $n \in \mathbb{N}$ arbitrario).

Dimostrazione.

Infatti, se $p_n(x) = \sum_{k=0}^n a_k x^k$ allora per la disuguaglianza triangolare e il fatto che per ogni k si ha

$$\|x^k\|_{2,w}^2 = \int_a^b |x|^{2k} w(x) dx < +\infty$$

necessariamente

$$\|p_n\|_{2,w} = \left\| \sum_{k=0}^n a_k x^k \right\|_{2,w} \leq \sum_{k=0}^n |a_k| \|x^k\|_{2,w} < +\infty,$$

e quindi $p_n \in L_w^2(a, b)$. $\square$

Il problema ai minimi quadrati

Nota.

Si osservi che se a e b sono finiti, essendo per il teorema di Weistrass $\|x^n\| = \max_{x \in [a,b]} |x|^n < +\infty$, abbiamo

$\int_a^b |x|^n w(x) dx \leq \|x^n\| \int_a^b w(x) dx$ da cui se $\int_a^b w(x) dx < +\infty$ automaticamente $\int_a^b |x|^n w(x) dx < +\infty$ per tutti gli $n \in \mathbb{N}$.

Fissata $f \in L_w^2([a, b])$ ed $n \in \mathbb{N}$, il problema ai minimi quadrati (nel continuo) consiste nel determinare il polinomio p_n di grado n tale che sia minima la quantità (cf. [1, p.204-207])

$$\|f - p_n\|_{2,w} = \int_a^b |f(x) - p_n(x)|^2 w(x) dx.$$

Il problema ai minimi quadrati

Essendo $L_w^2([a, b])$ uno spazio euclideo, e $(\phi_k)_{k=0,\dots,n}$ una base di $\mathcal{P}_n$, abbiamo visto che la soluzione del problema

$$\|f - f^*\|_{2,w} = \min_{g \in \text{span}\{\phi_0, \dots, \phi_n\}} \|f - g\|_{2,w}$$

è

$$f^* = \sum_{j=0}^n \gamma_j^* \phi_j$$

dove i coefficienti γ_j^* verificano le cosiddette equazioni normali

$$\sum_{k=0}^n (\phi_j, \phi_k)_{2,w} \gamma_k^* = (\phi_j, f)_{2,w}, \quad j = 0, \dots, n.$$

La soluzione è caratterizzata dalla proprietà di ortogonalità cioè che $f^* - f$ è ortogonale a tutti gli ϕ_k , con $k = 1, \dots, n$, ovvero

$$(f, \phi_k)_{2,w} = (f^*, \phi_k)_{2,w}, \quad k = 0, \dots, n$$

Il problema ai minimi quadrati

Nel caso $\phi_k(x) = x^k$ per $k = 0, \dots, n$, le equazioni normali si riscrivono quindi come

$$\sum_{k=0}^n \gamma_k^* \left(\int_a^b x^{j+k} w(x) dx \right) = \int_a^b x^j f(x) w(x) dx, \quad j = 0, \dots, n.$$

o in forma matriciale, posto per $j, k = 0, \dots, n$

$$A_{j,k} := \int_a^b x^{j+k} w(x) dx, \quad \beta_j = \int_a^b x^j f(x) w(x) dx,$$

come $A\gamma = \beta$.

Il problema ai minimi quadrati

Vediamo quale caso particolare $w(x) \equiv 1$, $(a, b) = (0, 1)$. Poichè

$$A_{j,k} = \int_0^1 x^{j+k} dx = \frac{1}{j+k+1}$$

si nota che la matrice A coincide con la **matrice di Hilbert** (cf. [7])

$$H_{s,t} = \frac{1}{(s-1) + (t-1) + 1}, \quad s, t = 1, \dots, n+1$$

La matrice di Hilbert H è estremamente sensibile a piccoli cambiamenti nei coefficienti o dei valori del termine noto (si dice **malcondizionata**).

Il problema ai minimi quadrati

Per vederlo (cf. [1, p.530-535]), sia $\mathcal{A}$ una matrice non singolare di ordine N , $\|\cdot\|$ una norma indotta di matrice [8] cioè tale che

$$\|\mathcal{A}x\| \leq \|\mathcal{A}\| \|x\|, \forall x \in \mathbb{R}^N$$

e indichiamo con

$$\kappa(\mathcal{A}) = \|\mathcal{A}\| \|\mathcal{A}^{-1}\|$$

il **condizionamento della matrice** rispetto alla norma $\|\cdot\|$.

Si dimostra [3, p.137] che se

$$(\mathcal{A} + \delta\mathcal{A})(x + \delta x) = b + \delta b,$$

allora

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\kappa(\mathcal{A})}{1 - \kappa(\mathcal{A}) \frac{\|\delta\mathcal{A}\|}{\|\mathcal{A}\|}} \cdot \left(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta\mathcal{A}\|}{\|\mathcal{A}\|} \right).$$

Il problema ai minimi quadrati

Per capirne il significato, si supponga per semplicità $\delta\mathcal{A} = 0$, Allora

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\kappa(\mathcal{A})}{1 - \kappa(\mathcal{A}) \frac{\|\delta\mathcal{A}\|}{\|\mathcal{A}\|}} \cdot \left(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta\mathcal{A}\|}{\|\mathcal{A}\|} \right) = \kappa(\mathcal{A}) \frac{\|\delta b\|}{\|b\|}$$

per cui più grande è $\kappa(\mathcal{A})$ più grande è il rischio che piccole perturbazioni sui termini noto b provochino grandi effetti sul calcolo della soluzione x .

- ▶ Si noti che questa stima è solo una maggiorazione.
- ▶ In molti casi risulta un buon indicatore di quanto un sistema lineare sia *difficile* da risolvere al calcolatore, indipendentemente dal metodo utilizzato.

Il problema ai minimi quadrati

Tornando alla matrice di Hilbert H , vediamo in Matlab valori del numero di condizionamento in norma 2. Se λ_k sono gli autovalori di A ordinati in modulo, cioè

$$|\lambda_1| \leq \dots \leq |\lambda_l| \leq \dots \leq |\lambda_n|$$

si dimostra che essendo la matrice H **simmetrica e definita positiva** (cf. [3, p.140])

$$\mathcal{K}_2(A) = \frac{|\lambda_n|}{|\lambda_1|}.$$

Curiosamente, quest'ultima uguaglianza lega due grandi nomi dell'informatica moderna, essendo la corrente definizione di condizionamento introdotta da Turing mentre il rapporto degli autovalori a secondo membro da Von Neumann [3, p.226].

Il problema ai minimi quadrati

Per quanto anticipato, **calcoliamo il condizionamento della matrice di Hilbert** e i corrispettivi autovalori massimi e minimi in modulo. Notiamo che Matlab, non approssimando gli autovalori di una matrice per ottenere il numero di condizionamento (cf. [3, p.139]), per valori grandi dello stesso offre solo un'approssimazione.

D'altra parte si può dimostrare che il condizionamento (in norma 2) della matrice di Hilbert di ordine n è approssimativamente **$\exp(3.5 n)$** (cf. [3, p.139]).

Il problema ai minimi quadrati

```
>> help cond
```

COND Condition number with respect to inversion.

COND(X) returns the 2-norm condition number (the ratio of the largest singular value of X to the smallest).

Large condition numbers indicate a nearly singular matrix.

COND(X,P) returns the condition number of X in P-norm:

$$\text{NORM}(X,P) * \text{NORM}(\text{INV}(X),P).$$

where P = 1, 2, inf, or 'fro.'

See also RCOND, CONDEST, CONDEIG, NORM, NORMEST.

Il problema ai minimi quadrati

```
>> for n=5:5:40
    fprintf('\n \t [n]: %2.0f [cond (norma 2)]: %2.2
           e [det]: %2.2e', n, cond(hilb(n)), ...
           det(hilb(n))); end
```

[n]: 5	[cond (norma 2)]: 4.77 e+005	[det]: 3.75 e-012
[n]: 10	[cond (norma 2)]: 1.60 e+013	[det]: 2.16 e-053
[n]: 15	[cond (norma 2)]: 8.49 e+017	[det]: -2.19 e-120
[n]: 20	[cond (norma 2)]: 1.91 e+018	[det]: -1.12 e-195
[n]: 25	[cond (norma 2)]: 1.46 e+019	[det]: 8.14 e-275
[n]: 30	[cond (norma 2)]: 6.19 e+018	[det]: 0.00 e+000
[n]: 35	[cond (norma 2)]: 1.03 e+019	[det]: 0.00 e+000
[n]: 40	[cond (norma 2)]: 4.55 e+019	[det]: 0.00 e+000

Il problema ai minimi quadrati

```
>> for n=5:5:40
    fprintf('\n \t [n]: %2.0f [min autov.]: %2.2e [
        max autov.]: %2.2e', ...
        n, min(abs(eig(hilb(n)))), max(abs(eig(hilb(n)))))
    ;
end

[n]: 5 [min autov.]:3.29e-006 [max autov.]:1.57e+000
[n]:10 [min autov.]:1.09e-013 [max autov.]:1.75e+000
[n]:15 [min autov.]:7.58e-018 [max autov.]:1.85e+000
[n]:20 [min autov.]:3.05e-018 [max autov.]:1.91e+000
[n]:25 [min autov.]:1.70e-018 [max autov.]:1.95e+000
[n]:30 [min autov.]:9.49e-019 [max autov.]:1.99e+000
[n]:35 [min autov.]:1.68e-017 [max autov.]:2.01e+000
[n]:40 [min autov.]:2.22e-017 [max autov.]:2.04e+000
>>
```


Il problema ai minimi quadrati

Ricapitolando

1. il problema ai minimi quadrati continui ha un'unica soluzione se e solo se un certo sistema lineare $A\gamma = \beta$ ha una e una sola soluzione;
2. nel caso $w \equiv 1$, $(a, b) = (-1, 1)$ la matrice A coincide con la matrice di Hilbert che ha un numero di condizionamento estremamente grande, il che comporta che il sistema lineare $A\gamma = \beta$ sarà estremamente sensibile a piccole perturbazioni (cioè malcondizionato);
3. la matrice di Hilbert sembra essere singolare, cioè avente determinante nullo, per $n \geq 0$; in realtà, come già detto, si dimostra che è simmetrica e definita positiva, e quindi tutti gli autovalori sono positivi da cui il determinante è non nullo essendo il prodotto degli autovalori.

Il problema ai minimi quadrati

Notiamo subito che il sistema è difficile da risolvere poichè abbiamo rappresentato il polinomio p_n^* nella base $1, x, \dots, x^n$. Cosa succede se lo scriviamo come combinazione lineare di elementi di un'altra base $\{\phi_n\}$? Se ad esempio scegliessimo la base $\{\phi_n\}$ cosicchè

$$(\phi_k, \phi_m) = a_k \delta_{k,m}, \quad a_k > 0, \quad k, m \in 0, 1, \dots, n$$

con $\delta_{k,m}$ delta di Kronecker, finiremmo per risolvere un problema $A\gamma = \beta$ dove A è una matrice diagonale e quindi senza alcuna difficoltà troveremmo

$$\gamma_k = \frac{\beta_k}{a_k}$$

con $a_k > 0$.

Risulta quindi di importanza fondamentale cercare basi di polinomi $\{\phi_j\}$ per cui

$$(\phi_k, \phi_m) = a_k \delta_{k,m}, \quad a_k > 0, \quad k, m \in 0, 1, \dots, n.$$

Sia $w : (a, b) \rightarrow \mathbb{R}$ una funzione nonnegativa, con (a, b) non necessariamente limitato, tale che

1. $\int_a^b |x|^n w(x) dx < +\infty$ per tutti gli $n \in \mathbb{N}$;
2. $\int_a^b f(x) w(x) dx = 0$ per qualche funzione continua e non negativa g implica $g \equiv 0$ in (a, b) .

La funzione w è detta **peso**.

Le funzioni peso più comuni sono (cf. [1, p.206], [10], [11],[9])

1. $w(x) = 1$ con $x \in [-1, 1]$ (peso di Legendre) [14];
2. $w(x) = \frac{1}{\sqrt{1-x^2}}$ con $x \in (-1, 1)$ (peso di Chebyshev);
3. $w(x) = (1-x^2)^{\gamma-(1/2)}$ con $x \in (-1, 1)$, $\gamma > (-1/2)$ (peso di Gegenbauer);
4. $w(x) = (1-x)^{\alpha} \cdot (1+x)^{\beta}$ con $x \in (-1, 1)$, $\alpha > -1$, $\beta > -1$ (peso di Jacobi);
5. $w(x) = \exp(-x)$ con $x \in (0, +\infty)$ (peso di Laguerre) [13];
6. $w(x) = \exp(-x^2)$ con $x \in (-\infty, +\infty)$ (peso di Hermite) [12];

Vediamo ora che proprietà ha un insieme di polinomi $\{\phi_k\}_{k=0,\dots,n}$ in cui il grado $\deg$ di ogni polinomio sia $k - 1 < \deg(\phi_k) = k$ e $(\phi_i, \phi_j)_{2,w} = c_i \delta_{i,j}$ (con $\delta_{i,j}$ il delta di Kronecker, $c_i > 0$, $i, j = 0, \dots, n$).

Una tal famiglia **triangolare** di polinomi (cioè tale che $\deg(\phi_k) = k$) si dice **ortogonale** rispetto alla funzione peso w nell'intervallo di riferimento.

Si può dimostrare

- ▶ usando la procedura di Gram-Schmidt che una tal famiglia triangolare di polinomi esiste e con la stessa procedura costruirla direttamente;
- ▶ inoltre è immediato osservare che ogni polinomio di grado n si può scrivere univocamente come combinazione lineare di $\phi_0, \dots, \phi_n$.

Di conseguenza se $p_n = \sum_{j=0}^n a_j \phi_j$, allora per la bilinearità del prodotto scalare $(\cdot, \cdot)_{2,w}$

$$(\phi_{n+1}, p_n)_{2,w} = (\phi_{n+1}, \sum_{j=0}^n a_j \phi_j)_{2,w} = \sum_{j=0}^n a_j (\phi_{n+1}, \phi_j)_{2,w} = 0.$$

Inoltre (cf. [4, p.978], [1, p.213])

Teorema

*Sia $\{\phi_k\}_{k=0,\dots,n}$ una famiglia triangolare di polinomi ortogonali in (a, b) rispetto ad una funzione peso w . Allora il polinomio ϕ_n ha **esattamente n radici reali e distinte nell'intervallo aperto (a, b)** .*

Dimostrazione.

Siano $x_1, \dots, x_m$ (con $m \leq n$) tutti e soli gli zeri di ϕ_n interni ad (a, b) con molteplicità rispettivamente $\alpha_1, \dots, \alpha_m$. Di conseguenza, per qualche numero a_n abbiamo

$$\phi_n(x) = a_n \left(\prod_{k=1}^m (x - x_k)^{\alpha_k} \right) r(x)$$

avendo supposto $\prod_{k=1}^m (x - x_k)^{\alpha_k} \equiv 1$ se non ci sono zeri interni ad (a, b) .

*Il polinomio r per costruzione non ha zeri in (a, b) e quindi non si annulla mai ed essendo una funzione continua ha **segno costante**.*

Dimostrazione.

Consideriamo il polinomio

$$q(x) = \left(\prod_{k=1}^m (x - x_k)^{\text{mod}_2(\alpha_k)} \right).$$

Se uno zero di ϕ_n ha molteplicità dispari ma maggiore di 1 o uno almeno ha molteplicità pari o esiste uno zero complesso non in (a, b) , è facile osservare che il grado di q è minore di n .

Osserviamo ora che qualsiasi sia un numero naturale, $\alpha_k + \text{mod}_2(\alpha_k)$ è un numero pari.

Dimostrazione.

Di conseguenza avendo $\left(\prod_{k=1}^m (x - x_k)^{\alpha_k + \text{mod}_2(\alpha_k)}\right) r_{n-m}(x)$ segno costante e non coincidente col polinomio nullo,

$$\begin{aligned} 0 &= (\phi_n, q) \\ &= \int_a^b \phi_n(x) q(x) w(x) dx \\ &= \int_a^b a_n \prod_{k=1}^m (x - x_k)^{\alpha_k} r_{n-m}(x) \prod_{k=1}^m (x - x_k)^{\text{mod}_2(\alpha_k)} w(x) dx \\ &= \int_a^b a_n \left(\prod_{k=1}^m (x - x_k)^{\alpha_k + \text{mod}_2(\alpha_k)} \right) r_{n-m}(x) w(x) dx \neq 0 \end{aligned}$$

da cui la contraddizione. (1)

Dimostrazione.

Potrebbe venire il dubbio su perchè qualche zero non possa essere a o b . Nella dimostrazione avrebbe quale unico effetto che r_{n-m} si annulla in a o b , rimanendo di segno costante in (a, b) .

La conclusione è che il polinomio ortogonale p_n ha n radici distinte e semplici, interne ad (a, b) . $\square$

Polinomi ortogonali

Vediamo ora il Teorema di Clenshaw ([1, p.214]), che stabilisce a partire dai polinomi ortogonali ϕ_0, ϕ_1 di grado 0 e 1 risp. come calcolare ricorsivamente la famiglia di pol. ortog..

Teorema

Sia $\{\phi_k\}_{k=0,\dots,n}$ una famiglia triangolare di polinomi ortogonali in $[a, b]$ rispetto ad una funzione peso w . Allora per $n \geq 1$

$$\phi_{n+1}(x) = \alpha_n(x - \beta_n)\phi_n(x) - \gamma_n\phi_{n-1}(x)$$

dove, detto a_n il coefficiente di grado massimo di ϕ_n , si ha

$$\alpha_n = \frac{a_{n+1}}{a_n} \quad (2)$$

$$\beta_n = \frac{(x\phi_n, \phi_n)_{2,w}}{(\phi_n, \phi_n)_{2,w}} \quad (3)$$

$$\gamma_n = \frac{\alpha_n(x\phi_{n-1}, \phi_n)_{2,w}}{(\phi_{n-1}, \phi_{n-1})_{2,w}} \quad (4)$$

Nota.

Notiamo che

- ▶ *scelti i polinomi ortogonali ϕ_0 e ϕ_1 , la procedura determina la famiglia triangolare di polinomi ortogonali di grado superiore, non appena sono disponibili i coefficienti $\alpha_k, \beta_k, \gamma_k$.*
- ▶ *Se ϕ_n è tale che $(\phi_n, \phi_k) = 0$ per $k = 0, \dots, n-1$ allora per $\tau \neq 0$ pure $\tilde{\phi}_n = \tau\phi_n$ è tale che*

$$(\tilde{\phi}_n, \phi_k) = (\tau\phi_n, \phi_k) = \tau(\phi_n, \phi_k) = 0, \quad k = 0, \dots, n-1$$

e quindi potrebbe essere considerato quale polinomio ortogonale di grado n .

Nota.

- *In pratica spesso si sceglie $\alpha_n = 1$ e i polinomi $\phi_0(x) = 1$, $\phi_1(x) = x - \int_a^b xw(x)dx / \int_a^b w(x)dx$ cosicchè i polinomi ortogonali siano monici cioè con coefficiente di grado massimo uguale a 1. In molti altri casi, come nelle routines Matlab di W. Gautschi, si pone $p_{-1}(x) = 0$, $p_0(x) = 1$ e quindi si applica la formula ricorsiva di Clenshaw con $\alpha_k = 1$ per $k > 0$.*

Polinomi ortogonali



K. Atkinson, *An Introduction to Numerical Analysis*, Wiley, (1989).



K. Atkinson e W. Han, *Theoretical Numerical Analysis, A Functional Analysis Framework*, Springer, (2001).



D. Bini, M. Capovani e O. Menchi, *Metodi numerici per l'algebra lineare*, Zanichelli, (1993).



V. Comincioli, *Analisi Numerica, metodi modelli applicazioni*, McGraw-Hill, (1990).



G. Dahlquist e A. Bjorck, *Numerical methods*, Dover, (2003).



A.N. Kolmogorov e S.V. Fomin, *Introductory Real Analysis*, Dover publications, 1970.



Wikipedia, (Hilbert Matrix), http://en.wikipedia.org/wiki/Hilbert_matrix.



Wikipedia, (Norma matriciale), http://it.wikipedia.org/wiki/Norma_matriciale.



Wikipedia, (Orthogonal polynomials), http://en.wikipedia.org/wiki/Orthogonal_polynomials.



Wikipedia, (Polinomi di Chebyshev), http://it.wikipedia.org/wiki/Polinomi_di_Chebyshev.



Wikipedia, (Polinomi Ortogonali), http://it.wikipedia.org/wiki/Polinomi_ortogonali.



Wikipedia, (Polinomi di Hermite), http://it.wikipedia.org/wiki/Polinomi_di_Hermite.



Wikipedia, (Polinomi di Laguerre), http://it.wikipedia.org/wiki/Polinomi_di_Laguerre.



Wikipedia, (Polinomi di Legendre), http://it.wikipedia.org/wiki/Polinomi_di_Legendre.