Miglior approssimazione in spazi euclidei

15 gennaio 2009

1 Introduzione astratta

Sia E uno spazio vettoriale dotato di un prodotto interno $(\cdot,\cdot)$ (talvolta un tale spazio è detto *euclideo*, cf. [7, p.148]), cioè una funzione reale definita sulle coppie $x, y \in E$ con le seguenti proprietà

1. $(x,x) \geq 0$ dove $(x,x) = 0$ se e solo se $x = 0$;
2. $(x,y) = (y,x)$;
3. $(\lambda x, y) = \lambda(x,y)$
4. $(x, y+z) = (x,y) + (x,z)$

per ogni $x, y, z \in E$ e $\lambda \in \mathbb{R}$.

A partire dal prodotto interno si può definire lo spazio normato $(E, \|\cdot\|)$ ponendo $\|f\| = \sqrt{(f,f)}$.

Vediamo alcuni esempi di spazi euclidei:

1. $\mathbb{R}^n$ dotato dell'usuale prodotto scalare, è uno spazio euclideo; se $e_1, \ldots, e_n$ è una base ortonormale, cioè per cui $(\phi_j, \phi_k) = \delta_{j,k}$ (dove al solito $\delta_{j,k}$ è il delta di Kronecker), allora ogni vettore $x \in \mathbb{R}^n$ si può scrivere come

$$x = \sum_{k=1}^{n} c_n e_n, \;\; c_k = (x, e_k).$$

   Infatti, moltiplicando ambo i membri di x per e_k si ha per la bilinearità del prodotto scalare

$$(x, e_k) = \left(\sum_{k=1}^{n} c_n e_n, e_k \right) = \sum_{k=1}^{n} c_n (e_n, e_k) = c_k (e_k, e_k) = c_k.$$

2. lo spazio $C([a,b])$ delle funzioni continue nel compatto $[a,b]$, dotato del prodotto scalare
$$(f,g) = \int_a^b f(x)g(x)\,dx$$
è uno spazio euclideo, cf. [7, p.145]. Una sua base ortogonale (cioè per cui $(\phi_j,\phi_k) = c_{j,j}\delta_{j,k}$ con $c_{j,j} \neq 0$ per ogni j), facilmente ortonormalizzabile, è il sistema di funzioni trigonometriche
$$1, \cos\left(\frac{2\pi nt}{b-a}\right), \sin\left(\frac{2\pi nt}{b-a}\right),\ \ n = 1,2,\ldots.$$

1.1 Sull'elemento di miglior approssimazione in spazi euclidei

Nella ricerca dell'elemento di miglior approssimazione in spazi euclidei partiamo da un teorema in un sottospazio di dimensione finita. Cominciamo introducendo una generalizzazione del noto teorema di Pitagora.

Teorema 1.1 *Sia E uno spazio euclideo, e siano $f,g \in E$ tali che $(f,g) = 0$ (cioè f e g sono ortogonali). Allora $\|f+g\|^2 = \|f\|^2 + \|g\|^2$.*

Dimostrazione (cf. [3, p.90]). Essendo $(f,g) = 0$, dalla bilinearità del prodotto interno,
$$\begin{aligned} \|f+g\|^2 &= (f+g, f+g) = (f,f) + (g,f) + (f,g) + (g,g) \\ &= (f,f) + 0 + 0 + (g,g) \\ &= \|f\|^2 + \|g\|^2 \end{aligned} \tag{1}$$
Il teorema di Pitagora servirà per dimostrare il seguente teorema (della proiezione ortogonale).

Teorema 1.2 *Sia $f \in E$, E spazio euclideo e $\{\phi\}_{1,\ldots,N}$ un sistema finito di elementi di E linearmente indipendenti. Allora la soluzione del problema*
$$\|f - f^*\|_2 = \min_{g \in span\{\phi\}_{1,\ldots,N}} \|f-g\|_2$$
è
$$f^* = \sum_{1,\ldots,N} c_j^* \phi_j$$
dove i coefficienti c_j^* verificano le cosidette equazioni normali
$$\sum_{k=1}^{N} (\phi_j,\phi_k) c_k^* = (\phi_j, f),\ j = 1,\ldots,N.$$
La soluzione è caratterizzata dalla proprietà di ortogonalità cioè che $f^* - f$ è ortogonale a tutti gli ϕ_k, con $k = 1,\ldots,n$.

Un caso importante è quello in cui $\{\phi\}_{1,\ldots,N}$ è un sistema ortogonale, cioè
$$(\phi_j,\phi_k) = c_{j,k}\delta_{j,k},\ \ c_{j,j} \neq 0,$$

dove al solito $\delta_{j,k}$ denota il delta di Kronecker; allora i coefficienti c_k (detti in questo caso di Fourier) sono calcolabili più semplicemente con la formula

$$c_j = \frac{(f,\phi_j)}{(\phi_j,\phi_j)},\ j = 1,\dots,N.$$

Dimostrazione (cf. [3, p.92]). Sia $(c_k)_{1,\dots,N}$, una sequenza di coefficienti e supponiamo che per almeno un indice j sia $c_j \neq c_j^*$. Allora

$$\begin{aligned}\sum_{k=1}^{N} c_j\phi_j - f &= \left(\sum_{k=1}^{N} c_j\phi_j - f^*\right) + \left(f^* - f\right)\\ &= \sum_{k=1}^{N}(c_j - c_j^*)\phi_j + \left(f^* - f\right) \qquad (2)\end{aligned}$$

Se $u = f^* - f$ è ortogonale a tutti i ϕ_j, allora è ortogonale pure alla combinazione lineare di ϕ_j come ad esempio $v = \sum_{k=1}^{N}(c_j - c_j^*)\phi_j$. Dal teorema di Pitagora, poichè $(u,v) = 0$ implica $\|u+v\|^2 = \|u\|^2 + \|v\|^2$, abbiamo

$$\begin{aligned}\|\sum_{k=1}^{N} c_j\phi_j - f\|^2 &= \|\left(\sum_{k=1}^{N} c_j\phi_j - f^*\right) + (f^* - f)\|^2\\ &= \|\sum_{k=1}^{N} c_j\phi_j - f^*\|^2 + \|f^* - f\|^2\\ &= \|\sum_{k=1}^{N}(c_j - c_j^*)\phi_j\|^2 + \|f^* - f\|^2\\ &> \|f^* - f\|^2 \qquad (3)\end{aligned}$$

Di conseguenza se $f^* - f$ è ortogonale a tutti i ϕ_j allora f^* è la miglior approssimazione di f in $\mathrm{span}\{\phi\}_{1,\dots,N}$. Rimane allora da mostrare che le condizioni di ortogonalità

$$\left(\sum_{k=1}^{N} c_j^*\phi_j - f, \phi_k\right) = 0,\ k = 1,\dots,N$$

possano essere soddisfatte. Questo problema è equivalente alla soluzione del sistema di equazioni normali

$$\sum_{k=1}^{N}(\phi_j,\phi_k)c_k^* = (\phi_j, f),\ j = 1,\dots,N \qquad (4)$$

Ma questo è vero. Se $\phi_1,\dots,\phi_N$ sono non solo N vettori linearmente indipendenti ma formano perfino un sistema ortogonale, si avrebbe direttamente da (4) che

$$(\phi_k,\phi_k)c_k^* = (\phi_k, f)$$

e visto che $(\phi_k,\phi_k) \neq 0$ si vede subito che $(c_k^*)_k$ esistono unici e uguali a $c_k^* = \frac{(\phi_k,f)}{(\phi_k,\phi_k)}$.

Se invece $\phi_1, \ldots, \phi_N$ non formano un sistema ortogonale, il sistema di equazioni normali ha una e una sola soluzione se il sistema omogeneo di equazioni

$$\sum_{k=1}^{N} (\phi_j, \phi_k) c_k^* = 0, \ j = 1, \ldots, N \tag{5}$$

ha la sola soluzione nulla. Se così non fosse, da (5)

$$\begin{aligned} \| \sum_{j=1}^{N} c_j \phi_j \|^2 &= \left(\sum_{j=1}^{N} c_j \phi_j, \sum_{k=1}^{N} c_k \phi_k \right) \\ &= \sum_{k=1}^{N} c_k \sum_{j=1}^{N} c_j \left(\phi_j, \phi_k \right) \\ &= \sum_{k=1}^{N} c_k \cdot 0 \\ &= 0 \end{aligned} \tag{6}$$

per cui essendo $\| \cdot \|$ una norma, necessariamente $\sum_{j=1}^{N} c_j \phi_j = 0$ il che contraddice il fatto che i ϕ_k erano linearmente indipendenti. ∎

Osserviamo subito che se una base ortogonale è a disposizione allora il calcolo della miglior approssimazione non richiede la soluzione del sistema delle equazioni normali bensi' il solo calcolo di alcuni prodotti interni e N divisioni. Inoltre si noti che se $\{\phi_k\}_{k=1,\ldots,N}$ è un sistema ortogonale, allora i coefficienti di Fourier c_j^* sono indipendenti da N col vantaggio che se è necessario aumentare il numero totale di parametri c_j^*, non è necessario ricalcolare quelli precedentemente ottenuti.

1.2 Facoltativo: Sui sistemi ortogonali

Al momento non abbiamo detto nulla riguardo una possibile base dello spazio euclideo E. E cosa serve richiedere perchè abbiano cardinalità numerabile? In tal caso, esistono delle basi da preferire, come per esempio quelle ortonormali?

Riguardo a queste questioni, si può provare che ogni spazio euclideo *separabile* (cioè che contiene un sottinsieme $S \subseteq X$ denso e numerabile, cf. [7, p.48])), ha una base ortonormale finita o numerabile.

Inoltre vale il seguente teorema detto di *ortogonalizzazione*, basato sull'algoritmo di *Gram-Schmidt*

Teorema 1.3 *Siano $f_1, \ldots, f_n, \ldots$ un insieme numerabile di elementi linearmente indipendenti di uno spazio euclideo E. Allora E contiene un insieme di elementi $\{\phi_k,\}_{k=1,\ldots,n,\ldots}$ tale che*

1. *il sistema $\{\phi_n\}$ è ortonormale (cioè (ϕ_m, ϕ_n)=$\delta_{m,n}$, dove $\delta_{m,n}$ è il delta di Kronecker);*
2. *ogni elemento ϕ_n è una combinazione lineare di $f_1, \ldots, f_n$;*
3. *ogni elemento f_n è una combinazione lineare di $\phi_1, \ldots, \phi_n$.*

Si osservi che

- l'insieme di partenza $f_1, \ldots, f_n, \ldots$ non deve essere necessariamente finito, come di solito viene spesso richiesto nell'algoritmo di ortogonalizzazione di matrici;
- l'insieme $\phi_1, \ldots, \phi_n, \ldots$ non deve essere necessariamente finito;
- se lo spazio euclideo ha una base numerabile formata da elementi linearmente indipendenti $f_1, \ldots, f_n, \ldots$, allora ha pure una base ortonormale.

Alcune definizioni:

- I numeri
$$c_k = (x, \phi_k), \ k = 1, 2, \ldots$$
sono detti *coefficienti di Fourier* rispetto il sistema $\{\phi_k\}$.
- Se $\{\phi_k\}$ è un sistema ortonormale di E, $f \in E$, la serie (formale)
$$\sum_{k=1}^{+\infty} c_k \phi_k$$
è chiamata *serie di Fourier* di f.

1.3 Facoltativo: Sull'elemento di miglior approssimazione in spazi euclidei

Viene naturale chiedersi quali proprietà ha l'elemento di miglior approssimazione, e cosa bisogna assumere perchè la serie di Fourier di f converga a f. In parte risponde il seguente teorema, detto di *Bessel*

Teorema 1.4 *Dato un sistema ortonormale*

$$\phi_1, \ldots, \phi_n, \ldots$$

in uno spazio euclideo E, sia $f \in E$. Allora l'espressione

$$\| f - \sum_{k=1}^{n} a_k \phi_k \|$$

ha il minimo per

$$a_k = c_k = (f, \phi_k), \ k = 1, 2, \ldots, n$$

ed è uguale a

$$\| f \|^2 - \sum_{k=1}^{n} c_k^2.$$

Inoltre vale la disuguaglianza di Bessel

$$\sum_{k=1}^{\infty} c_k^2 \leq \| f \|^2.$$

Osserviamo che

- la serie nella disuguaglianza di Bessel ha un insieme numerabile di termini;
- la soluzione al problema di miglior approssimazione in norma $\|\cdot\|$ esiste ed è unica: per ottenerla basta calcolare i coefficienti di Fourier; questo punto è fondamentale perchè dice costruttivamente come calcolare l'elemento di miglior approssimazione.

Definizione. Supponiamo che valga l'uguaglianza di Parseval

$$\sum_{k=1}^{\infty} c_k^2 = \|f\|^2$$

per ogni f nello spazio euclideo E. Allora il sistema $\{\phi_k\}$ si dice *chiuso*.

Definizione. Un sistema ortogonale (o ortonormale) $\{\phi_k\}_{k=1,\ldots,n,\ldots}$ è completo quando il più piccolo sottospazio di E contenente $\{\phi_k\}_{k=1,\ldots,n,\ldots}$ è l'intero spazio E. Un tale sistema è detto *base ortogonale (ortonormale)* .

Si possono dimostrare i seguenti ed importanti teoremi

Teorema 1.5 *Un sistema ortonormale $\{\phi_k\}_{k=1,\ldots,n,\ldots}$ in uno spazio euclideo è chiuso se e solo se ogni elemento $f \in E$ è la somma della sua serie di Fourier.*

Teorema 1.6 *Se un sistema ortonormale $\{\phi_k\}_{k=1,\ldots,n,\ldots}$ è completo allora $\{\phi_k\}_{k=1,\ldots,n,\ldots}$ è chiuso e viceversa.*

A questo punto abbiamo capito che se uno spazio euclideo E ha un sistema ortonormale chiuso (o equivalentemente completo) allora la serie di Fourier di un elemento f di E coincide con f stesso.

References

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

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

[3] G. Gilardi *Analisi Due, seconda edizione*, McGraw-Hill, (1996).

[4] D.H. Griffel, *Applied functional analysis*, Dover publications, 2002.

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

[6] A. Quarteroni, R. Sacco e F. Saleri *Matematica Numerica*, Springer, (1998).