

◆ **Esercizio 6.16** Dati  $(m+1)(n+1)$  valori  $f_{i,j}$ , per  $i = 1, 2, \dots, m+1$  e  $j = 1, 2, \dots, n+1$ , posto

$$X_{m,i}(x) := \prod_{k=1, k \neq i}^{m+1} \frac{x - x_k}{x_i - x_k}, \quad i = 1, 2, \dots, m+1$$

$$Y_{n,i}(y) := \prod_{k=1, k \neq i}^{n+1} \frac{y - y_k}{y_i - y_k}, \quad i = 1, 2, \dots, n+1$$

mostrare che

$$p_{m,n}(x, y) := \sum_{i=1}^{m+1} \sum_{j=1}^{n+1} X_{m,i}(x) Y_{n,j}(y) f_{i,j}$$

è un polinomio di grado  $m$  in  $x$  e di grado  $n$  in  $y$  della forma  $\sum_{i=1}^{m+1} \sum_{j=1}^{n+1} a_{i,j} x^i y^j$ , che soddisfa alle seguenti condizioni di interpolazione

$$p_{m,n}(x_i, y_j) = f_{i,j}, \quad i = 1, 2, \dots, m+1, \quad j = 1, 2, \dots, n+1$$

## 6.2 Problema generale di approssimazione

Nell'approssimazione mediante interpolazione si cerca un polinomio, più in generale un elemento di una famiglia appropriata di funzioni (ad esempio, funzioni razionali fratte, funzioni trigonometriche, esponenziali ecc.), che *coincide* con la funzione da approssimare in punti prefissati.

Nelle applicazioni, questo tipo di approssimazione non è sempre possibile o conveniente. Supponiamo, ad esempio, che la funzione  $f(x)$  esprima una relazione tra quantità fisiche, o chimiche. I valori  $f_i$  sono allora determinati mediante misurazioni, e in generale si ha  $f_i = f(x_i) + \epsilon_i$ , ove gli errori  $\epsilon_i$  sono incogniti; pure i punti di osservazione  $x_i$  possono essere affetti da errore. A meno che gli errori siano piccoli e le misurazioni siano in numero basso, non è allora ragionevole descrivere  $f(x)$  mediante un polinomio che passi esattamente attraverso tali punti. Può essere, invece, più opportuna una approssimazione nella quale l'influenza degli errori di misurazione sia minimizzata. In effetti, questo risultato può essere ottenuto utilizzando il *metodo dei minimi quadrati*, che esamineremo nel seguito.

Una differente situazione, ma per la quale è ancora opportuna una approssimazione di tipo diverso dalla interpolazione, si presenta quando una funzione assegnata in forma analitica, ad esempio  $e^x$ ,  $\log x$ ,  $\sin x$ , ecc., è approssimata da un polinomio (o da una funzione razionale fratta) allo scopo di costruire una procedura di calcolo della funzione da utilizzare in un calcolatore. In questo caso il tipo di approssimazione più conveniente corrisponde a cercare il *polinomio di minimo grado* (e quindi con il minimo costo di valutazione) che approssima la funzione assegnata con un errore minore di una tolleranza assegnata su tutto un intervallo prefissato.

Le due situazioni precedenti sono casi particolari della seguente situazione generale. Supponiamo che la funzione  $f(x)$ , i cui valori possono essere assegnati su tutto un intervallo (caso *continuo*) o solo in corrispondenza ad un numero finito di punti  $x_0, x_1, \dots, x_m$  (caso *discreto*), appartenga ad uno spazio lineare  $V$ .

**Spazi lineari normati** Ricordiamo che uno spazio  $V$  è detto *lineare*, se per ogni coppia di elementi  $f$  e  $g$  in  $V$ , e un numero arbitrario reale  $\alpha$ ,  $\alpha f$  e  $f + g$  appartengono a  $V$ . Inoltre, un sottoinsieme non vuoto  $U$  di uno spazio lineare  $V$  è un *sottospazio* di  $V$  se per ogni coppia arbitraria  $f_1, f_2$  di elementi in  $U$ , e un numero arbitrario reale  $\alpha$ ,  $\alpha f_1$  e  $f_1 + f_2$  appartengono a  $U$ .

Vi sono diversi tipi di spazi lineari che hanno interesse nelle applicazioni; per il seguito, tuttavia, ci limiteremo a considerare lo spazio lineare  $V = C^0([a, b])$  delle funzioni a valori reali e continue sull'intervallo  $[a, b]$  per il caso continuo, e lo spazio dei vettori  $\mathbb{R}^n$  nel caso discreto.

Per precisare il senso dell'approssimazione occorre introdurre nello spazio  $V$  una *distanza*. Essa può essere definita a partire dalla nozione di *norma*.

**Definizione 6.1 (norma)** Dato uno spazio lineare  $V$ , si dice *norma* una trasformazione  $V \rightarrow \mathbb{R}$ , indicata usualmente con  $\|\cdot\|$ , con le seguenti proprietà, ove  $f$  e  $g$  sono elementi arbitrari in  $V$

$$\begin{aligned} \|f\| &\geq 0; \quad \|f\| = 0 \iff f = 0 \\ \|\alpha f\| &= |\alpha| \cdot \|f\| \quad \forall \alpha \in \mathbb{R} \\ \|f + g\| &\leq \|f\| + \|g\| \quad (\text{disuguaglianza triangolare}) \end{aligned}$$

Si definisce allora *distanza* tra due elementi  $f$  e  $g$  in  $V$  la quantità  $\|f - g\|$ .

Nello spazio  $C^0([a, b])$  si possono, ad esempio, definire le seguenti norme

$$\begin{aligned} \|f\|_2 &= \left( \int_a^b |f(x)|^2 dx \right)^{1/2} && \text{norma 2 (euclidea)} \\ \|f\|_\infty &= \max_{a \leq x \leq b} |f(x)| && \text{norma del massimo (di Chebichev)} \end{aligned}$$

a cui corrispondono altrettante distanze e tipi diversi di approssimazione.

Nel caso discreto, cioè quando sono assegnati i valori  $f(x_i)$  in un insieme di punti  $S = \{x_i\}_{i=0}^m$ , si hanno definizioni analoghe (corrispondenti alla definizione di norma introdotta per  $\mathbb{R}^n$  nell'Appendice A) Ad esempio, si ha

$$\begin{aligned} \|f\|_{2,S} &= \left( \sum_{i=0}^m |f(x_i)|^2 \right)^{1/2} \\ \|f\|_{\infty,S} &= \max_{x_i \in S} |f(x_i)| \end{aligned}$$

Osserviamo che nel caso in cui  $f(x)$  sia definita su tutto un intervallo  $[a, b]$  contenente l'insieme di punti  $S$ , le quantità ora definite possono essere nulle senza che  $f(x)$  sia identicamente nulla su tutto  $[a, b]$ ; per tale motivo esse vengono dette *seminorme*, in quanto non verificano tutte le condizioni della Definizione 6.1.

Nelle applicazioni è talvolta opportuno generalizzare le definizioni precedenti introducendo una *funzione peso*  $w(x) > 0$ , con la quale assegnare una differente importanza ai valori  $f(x)$  nel calcolo della distanza. Si ha allora, ad esempio

$$\begin{aligned} \|f\|_{2,w} &= \left( \int_a^b w(x) |f(x)|^2 dx \right)^{1/2} \\ \|f\|_{2,S,w} &= \left( \sum_{i=0}^m w(x_i) |f(x_i)|^2 dx \right)^{1/2} \end{aligned}$$

con analogia definizione nel caso della norma del massimo. Naturalmente, la funzione  $w(x)$  deve essere tale da assicurare nel caso continuo l'esistenza dell'integrale per ogni  $f(x)$  nello spazio  $V$ .

Terminiamo questi brevi richiami sugli spazi lineari, ricordando la nozione importante di *prodotto scalare* e la conseguente nozione di *sistema ortogonale*, generalizzando le nozioni introdotte nel caso di  $V = \mathbb{R}^n$  in Appendice A. Per una generica coppia di funzioni  $f(x), g(x)$  nello spazio  $V = C^0([a, b])$  e per una opportuna funzione peso  $w(x)$  si può definire un *prodotto scalare* ponendo

$$(f, g) := \begin{cases} \int_a^b w(x)f(x)g(x) dx & \text{(caso continuo)} \\ \sum_{i=0}^m w(x_i)f(x_i)g(x_i) & \text{(caso discreto)} \end{cases}$$

da cui la seguente definizione di ortogonalità.

**Definizione 6.2 (sistema ortogonale)** *Le funzioni  $f(x), g(x) \in V$  si dicono ortogonali se  $(f, g) = 0$ . Un insieme di funzioni  $\phi_0, \phi_1, \dots, \phi_n$  appartenenti a  $V$  è chiamato un sistema ortogonale se  $(\phi_i, \phi_j) = 0$  per  $i \neq j$ , e  $(\phi_i, \phi_i) \neq 0$  per ogni  $i = 0, \dots, n$ . Quando  $(\phi_i, \phi_i) = 1$ , per ogni  $i$ , il sistema è detto ortonormale.*

Nel seguito (cfr. Paragrafo 6.2.2) considereremo differenti esempi di sistemi ortogonali nell'ambito dei polinomi. Osserviamo che per la norma, o seminorma, euclidea si ha  $\|f\|_2 = \sqrt{(f, f)}$ .

Si può allora dimostrare facilmente la seguente generalizzazione della uguaglianza di Pitagora. Se  $\{\phi_0, \phi_1, \dots, \phi_n\}$  è un sistema ortogonale, si ha

$$\left\| \sum_{j=0}^n c_j \phi_j \right\|_2^2 = \sum_{j=0}^n c_j^2 \|\phi_j\|_2^2 \quad (6.20)$$

Da tale uguaglianza si ricava, in particolare, che se  $\{\phi_0, \phi_1, \dots, \phi_n\}$  è un sistema ortogonale, si ha

$$\left\| \sum_{j=0}^n c_j \phi_j \right\| = 0 \quad \text{se e solo se } c_j = 0, j = 0, 1, \dots, n$$

ossia le funzioni  $\phi_0, \phi_1, \dots, \phi_n$  sono *linearmente indipendenti*.

**Approssimazione lineare** Ritornando alla definizione del problema di approssimazione, siano  $\phi_0, \phi_1, \dots, \phi_n$   $n+1$  funzioni assegnate nello spazio  $V = C^0([a, b])$  e  $U$  il sottospazio di  $V$  costituito dalle combinazioni lineari

$$g_n(x) = c_0\phi_0(x) + c_1\phi_1(x) + \dots + c_n\phi_n(x) \quad (6.21)$$

al variare delle costanti  $c_0, c_1, \dots, c_n$  in  $\mathbb{R}$ . Se, ad esempio, si sceglie  $\phi_i(x) = x^i, i = 0, 1, \dots, n$ , allora  $g_n(x)$  è un polinomio di grado al più  $n$  e l'insieme  $U$  è lo spazio dei polinomi di grado minore o uguale a  $n$ . Pur rimanendo nell'ambito dei polinomi, vi possono essere, tuttavia, per  $\phi(x)$  scelte più convenienti per le applicazioni (in particolare, come vedremo nel seguito i sistemi di polinomi ortogonali).

Fissata allora in  $V$  una particolare norma  $\|\cdot\|$ , il problema dell'approssimazione di una funzione  $f(x) \in V$  mediante funzioni del sottospazio  $U$  può essere formulato come un *problema di minimo*, corrispondente alla ricerca della combinazione  $g_n^*(x) \in U$ , tale che

$$\|f - g_n^*\| \leq \|f - g_n\| \quad \forall g_n \in U \quad (6.22)$$

Più precisamente, tale problema è chiamato un *problema di approssimazione lineare*, in quanto le funzioni approssimanti  $g_n(x)$  dipendono linearmente dai parametri incogniti  $c_0, c_1, \dots, c_n$ . In questo capitolo tratteremo in particolare i problemi di approssimazione di tipo lineare. I problemi di approssimazione di tipo non lineare, importanti per la modellistica matematica, saranno considerati sotto vari aspetti nei successivi Capitoli.

Nel seguito del capitolo esamineremo più in dettaglio l'approssimazione lineare corrispondente rispettivamente alla norma *euclidea* e alla norma del *massimo*.

### 6.2.1 Norma euclidea. Minimi quadrati

Si cerca  $g_n^*$  della forma (6.21) che verifica per tutte le funzioni  $g_n(x)$  dello stesso tipo le seguenti disuguaglianze

$$\text{(caso continuo)} \quad \|f - g_n^*\|_{2,w}^2 \leq \|f - g_n\|_{2,w}^2, \quad \text{ove } \|f\|_{2,w}^2 = \int_a^b w(x) |f(x)|^2 dx$$

$$\text{(caso discreto)} \quad \|f - g_n^*\|_{2,S,w}^2 \leq \|f - g_n\|_{2,S,w}^2, \quad \text{ove } \|f\|_{2,S,w}^2 = \sum_{i=0}^m w(x_i) |f(x_i)|^2$$

La funzione  $g_n^*$  viene detta *elemento di migliore approssimazione nel senso dei minimi quadrati*; dal punto di vista geometrico, la funzione  $g_n^*$  rappresenta la proiezione ortogonale di  $f$  sul sottospazio  $U$  (cfr. Figura 6.16).

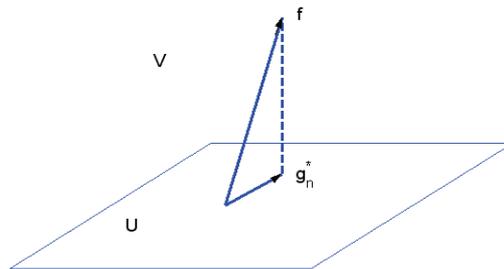


Figura 6.16 Interpretazione geometrica del problema dei minimi quadrati.

Dal punto di vista teorico, ossia esistenza ed unicità dell'elemento di migliore approssimazione, si ha il seguente risultato, valido sia nel caso discreto che continuo.

**Teorema 6.6** (minimi quadrati) *Supponiamo che le funzioni  $\phi_0, \phi_1, \dots, \phi_n$  siano linearmente indipendenti. Allora, esiste una ed una sola funzione*

$$g_n^*(x) = \sum_{j=0}^n c_j^* \phi_j(x) \quad (6.23)$$

tale che

$$\|f(x) - g_n^*(x)\|_2 \leq \|f(x) - g_n(x)\|_2 \quad \text{per ogni } g_n(x) = \sum_{j=0}^n c_j \phi_j(x)$$

ove  $\|\cdot\|_2$  indica una delle due norme  $\|\cdot\|_{2,w}$ ,  $\|\cdot\|_{2,S,w}$ . Inoltre, la funzione  $g_n^*(x)$  è la soluzione del seguente sistema lineare (detto sistema delle equazioni normali)

$$(f(x) - g_n^*(x), \phi_k(x)) = 0, \quad k = 0, 1, 2, \dots, n \quad (6.24)$$

ove  $(\cdot, \cdot)$  indica il prodotto scalare in  $V$  corrispondente alla norma considerata.

**DIMOSTRAZIONE.** Consideriamo come illustrazione grafica del teorema la Figura 6.16. Il sottospazio  $U$  di  $V$  è generato dalle combinazioni lineari di  $\phi_0, \phi_1, \dots, \phi_n$ . Il teorema afferma che  $f - g_n^*$  è ortogonale alle funzioni  $\phi_i$ , e quindi a tutti gli elementi del sottospazio  $U$ . Si ha pertanto che  $g_n^*$  è la proiezione ortogonale di  $f$  su  $U$  e, in sostanza, il teorema afferma che la proiezione ortogonale  $g_n^*$  di  $f$  su  $U$  è l'elemento in  $U$  che ha la minima distanza euclidea da  $f$ . Osserviamo che le equazioni normali (6.24) possono essere scritte nel seguente modo

$$(g_n^*, \phi_k) = (f, \phi_k), \quad k = 0, 1, \dots, n$$

da cui, tenendo conto della rappresentazione (6.23)

$$\sum_{j=0}^n c_j^* (\phi_j, \phi_k) = (f, \phi_k), \quad k = 0, 1, \dots, n \quad (6.25)$$

Si ha, pertanto, che i coefficienti  $c_j^*$  sono soluzioni di un sistema lineare con matrice dei coefficienti  $[(\phi_j, \phi_k)]$ ,  $j, k = 0, 1, \dots, n$  e termine noto  $[(f, \phi_k)]$ ,  $k = 0, 1, \dots, n$ . Dimostriamo che la matrice dei coefficienti è non singolare e che quindi il sistema lineare (6.25) ammette una ed una sola soluzione. Ragionando per assurdo, se la matrice fosse singolare, il sistema omogeneo avrebbe una soluzione  $c_0, c_1, \dots, c_n$  non identicamente nulla, cioè tale che

$$\sum_{j=0}^n c_j (\phi_j, \phi_k) = 0, \quad k = 0, 1, \dots, n$$

Ma, allora, si avrebbe

$$\left\| \sum_{j=0}^n c_j \phi_j \right\|_2^2 = \sum_{k=0}^n c_k \left( \sum_{j=0}^n c_j (\phi_j, \phi_k) \right) = 0$$

e, contrariamente all'ipotesi, le funzioni  $\phi_j$  sarebbero linearmente dipendenti.

Dimostriamo ora che ogni funzione  $g_n = \sum_{j=0}^n c_j \phi_j$ , con  $c_j \neq c_j^*$  per almeno un indice  $j$ , ha una distanza da  $f$  maggiore che  $g_n^*$ . In effetti, dalla seguente identità

$$f - \sum_{j=0}^n c_j \phi_j = (f - g_n^*) + \sum_{j=0}^n (c_j^* - c_j) \phi_j$$

tenendo conto (cfr. (6.25)) che  $(f - g_n^*, \phi_j) = 0$  per  $j = 0, 1, \dots, n$ , si ha

$$(f - g_n^*, \sum_{j=0}^n (c_j^* - c_j) \phi_j) = 0$$

ossia gli elementi  $f - g_n^*$  e  $\sum_{j=0}^n (c_j^* - c_j) \phi_j$  sono ortogonali. Applicando allora l'uguaglianza di Pitagora (6.20), si ha

$$\|f - \sum_{j=0}^n c_j \phi_j\|_2^2 = \|f - g_n^*\|_2^2 + \|\sum_{j=0}^n (c_j^* - c_j) \phi_j\|_2^2$$

da cui

$$\|f - \sum_{j=0}^n c_j \phi_j\|_2^2 \geq \|f - g_n^*\|_2^2$$

L'uguaglianza si ottiene soltanto per  $c_j^* = c_j$ , dal momento che gli elementi  $\phi_j$  sono linearmente indipendenti.

**Esempio 6.9** (*Caso continuo*) Determinare il polinomio  $g_2^*(x) = c_0 + c_1x + c_2x^2$  di migliore approssimazione nel senso dei minimi quadrati, con  $w(x) \equiv 1$ , della funzione  $f(x) = e^x$  su tutto l'intervallo  $[0, 1]$ .

In questo caso abbiamo  $\phi_0 = 1$ ,  $\phi_1 = x$ ,  $\phi_2 = x^2$ , e il sistema lineare delle equazioni normali (6.25) diventa il seguente

$$\begin{aligned} c_0 + \frac{1}{2}c_1 + \frac{1}{3}c_2 &= e - 1 \\ \frac{1}{2}c_0 + \frac{1}{3}c_1 + \frac{1}{4}c_2 &= 1 \\ \frac{1}{3}c_0 + \frac{1}{4}c_1 + \frac{1}{5}c_2 &= e - 2 \end{aligned}$$

Risolvendo il sistema in doppia precisione ( $\approx 16$  cifre decimali), si ottiene come soluzione il vettore

$$c^* = [1.01299130990276, 0.85112505284626, 0.83918397639947]$$

In Figura 6.17 è rappresentato l'errore  $f(x) - g_2^*(x)$ . Osserviamo che se arrotondiamo il termine noto del sistema delle equazioni normali a sei cifre, cioè poniamo  $[e - 1, 1, e - 2] \approx [1.71828, 1, 0.71828]$  e risolviamo il corrispondente sistema ancora in doppia precisione, si ottiene come soluzione il vettore

$$\tilde{c} = [1.01292000000000, 0.85152000000000, 0.83880000000000]$$

Si vede, quindi, che gli errori relativi sui dati si amplificano sui risultati, indicando la presenza di malcondizionamento nella matrice del sistema. In effetti, tale matrice è la matrice di Hilbert che abbiamo già considerato in precedenza dal punto di vista del condizionamento.

Una osservazione importante, comunque, è che il malcondizionamento del sistema delle equazioni normali è legato alla scelta particolare della base  $\phi_j$ . Come vedremo nel seguito, rispetto ad altre basi il sistema può essere bencondizionato.

**Esempio 6.10** (*Caso discreto*) Consideriamo un modo alternativo di ottenere le equazioni normali (6.25). I coefficienti  $c_0^*$ ,  $c_1^*$ ,  $\dots$ ,  $c_n^*$  possono essere determinati minimizzando la funzione errore

$$d(c_0, c_1, \dots, c_n) = \sum_{i=0}^m w(x_i) (f(x_i) - g_n^*(x_i))^2$$

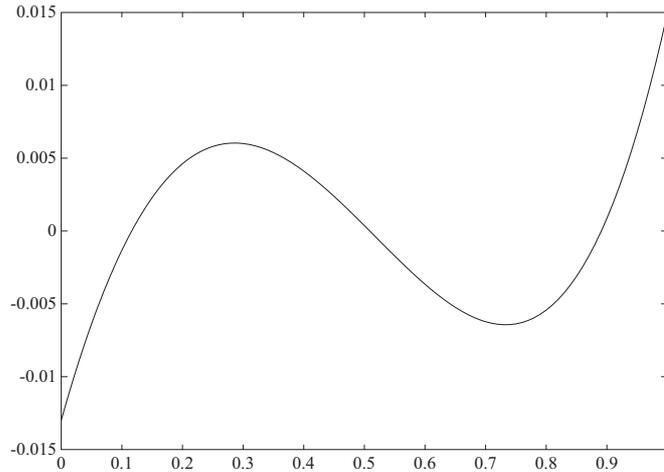


Figura 6.17 Errore  $f(x) - g_n^*(x)$  su  $[0, 1]$  per  $f(x) = e^x$  e  $g_n^*$  polinomio di secondo grado di migliore approssimazione nel senso dei minimi quadrati.

Ricordiamo allora che in un punto di minimo si ha

$$\frac{\partial d}{\partial c_k} = 0, \quad k = 0, 1, \dots, n$$

da cui

$$c_0 \sum_{i=0}^n w_i \phi_0(x_i) \phi_k(x_i) + \dots + c_n \sum_{i=0}^n w_i \phi_n(x_i) \phi_k(x_i) = \sum_{i=0}^n w_i \phi_k(x_i) f(x_i)$$

che coincide con il sistema (6.24). In termini matriciali il sistema precedente può essere scritto nel seguente modo. Introducendo per brevità la notazione

$$\mathbf{tab} f = [f(x_0), f(x_1), \dots, f(x_m)]^T$$

e indicando con  $\mathbf{A}$  la matrice di colonne  $\mathbf{tab} \phi_i, i = 0, 1, \dots, n$ , si ha

$$\mathbf{tab} g_n^* = [\mathbf{tab} \phi_0, \mathbf{tab} \phi_1, \dots, \mathbf{tab} \phi_n] \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{A} \mathbf{c}$$

Consideriamo, quindi, il seguente sistema di  $m + 1$  equazioni e  $n + 1$  incognite

$$\mathbf{tab} g_n^* = \mathbf{tab} f, \quad \iff \quad \mathbf{A} \mathbf{c} = \mathbf{b}$$

ove si è posto  $\mathbf{b} := \mathbf{tab} f$ . Il sistema delle equazioni normali corrisponde allora al seguente sistema di  $n$  equazioni e  $n$  incognite

$$\mathbf{A}^T \mathbf{A} \mathbf{c} = \mathbf{A}^T \mathbf{b} \tag{6.26}$$

In base al Teorema 6.6 la matrice  $\mathbf{A}^T \mathbf{A}$  è non singolare quando le funzioni  $\phi_i, i = 0, 1, \dots, n$  sono linearmente indipendenti sui punti  $x_0, x_1, \dots, x_m$ , cioè quando i vettori  $\mathbf{tab} \phi_i, i = 0, 1, \dots, n$  sono linearmente indipendenti. Ad esempio, nel caso in cui  $\phi_i = x^i$  una condizione sufficiente per la indipendenza lineare è che sia  $m \geq n$  e i punti  $x_i$  siano distinti. Infatti, se i vettori  $\mathbf{tab} \phi_i$  fossero linearmente dipendenti,

si avrebbe  $\sum_{j=0}^n c_j \mathbf{t} \mathbf{a} \phi_j = 0$  per  $c_0, c_1, \dots, c_n$  non tutti nulli; ma questo significa  $\sum_{j=0}^n c_j x_i^j = 0$  per  $i = 0, 1, \dots, m$ . Si avrebbe, quindi, un polinomio di grado  $\leq n$ , non identicamente nullo con un numero di zeri  $\geq n + 1$ .

Osserviamo anche che la matrice  $\mathbf{A}^T \mathbf{A}$  è una matrice *simmetrica definita positiva*. Per la risoluzione numerica del sistema (6.26) è, quindi, possibile utilizzare il *metodo di Choleski*. In caso di malcondizionamento, sono, tuttavia, opportune procedure più stabili, quali il metodo **QR** o la *decomposizione in valori singolari*.

Ricordiamo, infine, che il calcolo del polinomio di migliore approssimazione nel senso dei minimi quadrati è noto in statistica come *problema di regressione polinomiale*. In tale contesto, e quando la dispersione dei punti è descrivibile tramite una distribuzione normale, si introduce la seguente misura della dispersione dei dati intorno al polinomio di regressione

$$s_{y/x} = \sqrt{\frac{S_r}{m+1-(n+1)}}$$

ove  $S_r = \sum_{i=0}^m (f(x_i) - c_0 - c_1 x_i - c_2 x_i^2 - \dots - c_n x_i^n)^2$ . Il denominatore  $m+1-(n+1)$  indica il numero dei gradi di libertà, tenendo conto dei coefficienti  $c_i$  del polinomio. Oltre alla quantità  $s_{y/x}$ , nota come *errore standard della stima*, si può calcolare il *coefficiente di correlazione*  $r$  ponendo

$$r^2 = \frac{S_t - S_r}{S_t}$$

ove  $S_t$  indica la somma dei quadrati delle differenze tra la variabile dipendente (la  $y$ ) e il suo valore medio. La differenza  $S_t - S_r$  quantifica la riduzione dell'errore ottenuta assumendo come modello un polinomio di grado  $n$ . Nel caso di un'approssimazione ideale si ha  $S_r = 0$  e  $r^2 = 1$ , ossia il polinomio di regressione rappresenta perfettamente i dati. Al contrario, se  $r^2 = 0$  l'approssimazione data dal polinomio non porta alcun vantaggio.

## 6.2.2 Polinomi ortogonali

Nel paragrafo precedente abbiamo visto che per particolari scelte delle funzioni  $\phi_j$ , ad esempio per  $\phi_j(x) = x^j$ ,  $j = 0, 1, \dots, n$ , il sistema delle equazioni normali può essere *malcondizionato*, con conseguenti difficoltà numeriche per la sua risoluzione. Tali difficoltà possono essere evitate assumendo  $\phi_j$  come elementi di una base di polinomi ortogonali, cioè tali che per una fissata funzione peso  $w$  si abbia  $(\phi_j, \phi_k) = 0$  per  $j \neq k$ . In questo caso, infatti, le equazioni normali diventano

$$c_k (\phi_k, \phi_k) = (f, \phi_k), \quad k = 0, 1, \dots, n$$

e i coefficienti  $c_k$ , chiamati anche *coefficienti di Fourier*, possono essere ottenuti direttamente

$$c_k = \frac{(f, \phi_k)}{(\phi_k, \phi_k)}, \quad k = 0, 1, \dots, n$$

Rileviamo un altro vantaggio nell'uso dei polinomi ortogonali. Se  $p_n^*$  è il polinomio di migliore approssimazione di grado  $\leq n$ , per ottenere il polinomio di migliore approssimazione di grado  $\leq (n+1)$  è sufficiente calcolare  $c_{n+1}^*$  e porre  $p_{n+1}^* = p_n^* + c_{n+1}^* \phi_{n+1}$ .

**Esempio 6.11** Come esempio introduttivo, calcoliamo i polinomi  $P_i$  ortogonali rispetto al seguente prodotto scalare

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