Lezione 32

In questa lezione si inizia lo studio dei problemi di *approssimazione nel senso dei minimi quadrati.* Si definiscono le nozioni di *soluzione* di un sistema di equazioni lineari *nel senso dei minimi quadrati* e di *funzione che meglio approssima i dati nel senso dei minimi quadrati.* Si enuncia e risolve il *problema della migliore approssimazione in spazi con prodotto scalare* e se ne mostra il legame con la nozione di soluzione di un sistema nel senso dei minimi quadrati.

4 Approssimazione nel senso dei minimi quadrati

- *Definizione* (Soluzione di un sistema nel senso dei minimi quadrati).

  Siano $A \in \mathbb{R}^{n\times k}$ e $b \in \mathbb{R}^n$. Un vettore $x^* \in \mathbb{R}^k$ è una *soluzione del sistema* $Ax = b$ *nel senso dei minimi quadrati* se verifica una delle tre proprietà *equivalenti*:

  (A) Per ogni $y \in \mathbb{R}^k$ si ha: $\| Ax^* - b \|_2 \leqslant \| Ay - b \|_2$;

  (B) Per ogni $y \in \mathbb{R}^k$ si ha: $\| Ax^* - b \|_2^2 \leqslant \| Ay - b \|_2^2$;

  (C) Il vettore x^* è un *minimo assoluto* della funzione $n : \mathbb{R}^k \to \mathbb{R}$, *norma del residuo*, definita da:
  $$n(x) = \| Ax - b \|_2$$

  Si osservi che *se* x^* *è soluzione di* $Ax = b$ *allora è soluzione di* $Ax = b$ *nel senso dei minimi quadrati.* Infatti: $n(x^*) = 0$ e per ogni $x \in \mathbb{R}^k$ si ha $n(x) \geqslant 0$.

- *Definizione* (Funzione che meglio approssima i dati nel senso dei minimi quadrati).

  Siano I un intervallo non degenere, $\mathcal{F}$ un sottospazio vettoriale *di dimensione finita* dello spazio delle funzioni continue da I in $\mathbb{R}$, $x_0, \dots, x_k$ numeri reali in I e $y_0, \dots, y_k$ numeri reali. Un elemento f^* di $\mathcal{F}$ è una *funzione che meglio approssima i dati* $(x_0, y_0), \dots, (x_k, y_k)$ *nel senso dei minimi quadrati* se: *Per ogni* $f \in \mathcal{F}$ *si ha*:
  $$\big(f^*(x_0) - y_0 \big)^2 + \cdots + \big(f^*(x_k) - y_k \big)^2 \leqslant \big(f(x_0) - y_0 \big)^2 + \cdots + \big(f(x_k) - y_k \big)^2$$
  ovvero se f^* è un *minimo assoluto* della funzione $\mathrm{SQ} : \mathcal{F} \to \mathbb{R}$, *scarto quadratico*, definita da:
  $$\mathrm{SQ}(f) = \big(f(x_0) - y_0 \big)^2 + \cdots + \big(f(x_k) - y_k \big)^2$$

  Interpretando i dati come coordinate di punti in un piano cartesiano, lo scarto quadratico ha un semplice *significato geometrico.* Il valore $\mathrm{SQ}(f)$ è somma di $k+1$ addendi. Il j-esimo addendo è il quadrato della lunghezza del segmento individuato, sulla retta verticale di ascissa x_j, dal valore $f(x_j)$ e dall'ordinata del j-esimo dato y_j. Il valore $\mathrm{SQ}(f)$ *è una misura dello scostamento del grafico di* $f(x)$ *dai dati.*

  Si osservi che, contrariamente a quanto richiesto nel Problema dell'interpolazione polinomiale, i numeri reali $x_0, \dots, x_k$ *non* devono necessariamente essere distinti.

Alla ricerca delle soluzioni di un sistema nel senso dei minimi quadrati e delle funzioni che meglio approssimano i dati nel senso dei minimi quadrati premettiamo la nozione di migliore approssimazione in uno spazio con prodotto scalare e la sua ricerca.

Sia V uno spazio vettoriale su $\mathbb{R}$ con *prodotto scalare.* Indicato con $a \cdot b$ il prodotto scalare di a e b, si indica con $\| a \|$ la norma di a indotta dal prodotto scalare:
$$\| a \| = \sqrt{a \cdot a}$$
Siano poi W un sottospazio vettoriale di V di *dimensione finita* e v un elemento di V.

- *Definizione* (Migliore approssimazione in uno spazio con prodotto scalare).

  Un elemento w^* di W è una *migliore approssimazione di* v *in* W se verifica una delle due proprietà *equivalenti*:

(A) Per ogni $w \in W$ si ha: $\| v - w^* \| \leqslant \| v - w \|$;

(B) Il vettore w^* è un *minimo assoluto* della funzione $d : W \to \mathbb{R}$, *distanza da* v, definita da:

$$d(w) = \| v - w \|$$

Come vedremo, la nozione di migliore approssimazione in uno spazio con prodotto scalare è strettamente connessa a quella di *proiezione ortogonale*:

- *Definizione* (Proiezione ortogonale).

  Un elemento w^* di W è *proiezione ortogonale* di v su W se il vettore $v - w^*$ è ortogonale a W, ovvero se:

  $$\textit{Per ogni } w \in W \textit{ si ha: } (v - w^*) \cdot w = 0 \qquad (*)$$

Sia $w_1, \ldots, w_j$ un insieme di *generatori* di W. La condizione $(*)$ è *equivalente* a:

$$(v - w^*) \cdot w_i = 0 \qquad i = 1, \ldots, j$$

Allora una combinazione lineare $a_1 w_1 + \cdots + a_j w_j$ è proiezione ortogonale di v su W *se e solo se* per $i = 1, \ldots, j$ si ha: $\big(v - (a_1 w_1 + \cdots + a_j w_j) \big) \cdot w_i = 0$, ovvero *se e solo se*:

$$\begin{array}{rcccl} v \cdot w_1 & = & (a_1 w_1 + \cdots + a_j w_j) \cdot w_1 & = & a_1 (w_1 \cdot w_1) + \cdots + a_j (w_j \cdot w_1) \\ \vdots & & & \vdots & \\ v \cdot w_j & = & (a_1 w_1 + \cdots + a_j w_j) \cdot w_j & = & a_1 (w_1 \cdot w_j) + \cdots + a_j (w_j \cdot w_j) \end{array} \qquad (**)$$

Posto:

$$a = \begin{bmatrix} a_1 \\ \vdots \\ a_j \end{bmatrix} \quad , \quad F = \begin{bmatrix} w_1 \cdot w_1 & \cdots & w_j \cdot w_1 \\ \vdots & & \vdots \\ w_1 \cdot w_j & \cdots & w_j \cdot w_j \end{bmatrix} \quad , \quad c = \begin{bmatrix} v \cdot w_1 \\ \vdots \\ v \cdot w_j \end{bmatrix}$$

la condizione $(**)$ si riformula: *una combinazione lineare $a_1 w_1 + \cdots + a_j w_j$ è proiezione ortogonale di v su W se e solo se la colonna a è soluzione del sistema $Fx = c$* delle *equazioni normali.*

- *Teorema* (di esistenza ed unicità della proiezione ortogonale).

  Esiste un solo elemento di W proiezione ortogonale di v su W.

  Dimostrazione. Sia $w_1, \ldots, w_r$ una *base ortonormale* di W (sicuramente esistente...). Allora la matrice F del sistema delle equazioni normali è la matrice identica ed il sistema ha una sola soluzione:

  $$a = \begin{bmatrix} v \cdot w_1 \\ \vdots \\ v \cdot w_r \end{bmatrix}$$

  L'unico elemento di W proiezione ortogonale di v su W è:

  $$w^* = v \cdot w_1 \, w_1 + \cdots + v \cdot w_r \, w_r$$

Si osservi che il Teorema appena dimostrato prova che *le equazioni normali $Fx = c$ hanno sempre almeno una soluzione.* Infatti: *le componenti di ciascuna soluzione delle equazioni normali sono coordinate, rispetto ai generatori di W, della proiezione ortogonale.* L'esistenza della proiezione ortogonale w^* garantisce l'esistenza di almeno una combinazione lineare dei generatori di W tale che $a_1 w_1 + \cdots + a_j w_j = w^*$. Inoltre, se $w_1, \ldots, w_j$ sono generatori *linearmente dipendenti* (e quindi *non* una base) di W allora esistono *infinite* combinazioni lineari che generano lo stesso elemento w^*. In tal caso *il sistema $Fx = c$ ha infinite soluzioni.* Se, invece, gli elementi sono generatori *linearmente indipendenti* (e quindi *una base*) di W allora esiste *una sola* combinazione lineare che genera l'elemento w^*. In tal caso *il sistema $Fx = c$ ha una sola soluzione.*

- *Teorema* (di esistenza ed unicità della migliore approssimazione in uno spazio con prodotto scalare).

  La proiezione ortogonale w^* di v su W è l'*unica* migliore approssimazione di v in W, ovvero: w^* è *l'elemento di W più vicino a v.*

Dimostrazione. Per ogni $w \in W$ si ha:

$$\| v - w \|^2 = \| (v - w^*) + (w^* - w) \|^2$$

Il primo addendo, per definizione, è ortogonale a W ed il secondo è un elemento di W. Allora, per il Teorema di Pitagora:

$$\| (v - w^*) + (w^* - w) \|^2 = \| v - w^* \|^2 + \| w^* - w \|^2$$

e quindi:

$$\| v - w \|^2 = \| v - w^* \|^2 + \| w^* - w \|^2$$

Poiché $\| w^* - w \|^2 \geqslant 0$ e si ha $\| w^* - w \|^2 = 0$ *se e solo se* $w = w^*$, allora:

* $\| v - w \|^2 \geqslant \| v - w^* \|^2$ ovvero $\| v - w \| \geqslant \| v - w^* \|$: w^* è *una* migliore approssimazione;
* $\| v - w \|^2 = \| v - w^* \|^2$, ovvero $\| v - w \| = \| v - w^* \|$, *se e solo se* $w = w^*$: w^* è *l'unica* migliore approssimazione.

- *Esempi*

(1) Siano $V = \mathbb{R}^4$ con prodotto scalare canonico,

$$W = \text{span}\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \} \qquad \text{e} \qquad v = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix}$$

Determinare la migliore approssimazione di v in W.
Soluzione: Occorre determinare la proiezione ortogonale w^* di v su W. Poiché si ha *un solo* generatore di W, le equazioni normali si riducono ad *una* equazione in *una* incognita. Si ha:

$$F = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = 2 \qquad , \qquad c = v \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = 1$$

e le equazioni normali sono: $2x = 1$. Si ottiene l'unica soluzione $a = \frac{1}{2}$ da cui:

$$w^* = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Per il Teorema di esistenza ed unicità della migliore approssimazione in uno spazio con prodotto scalare, la migliore approssimazione di v in W è w^*.

(2) Siano $V = \mathbb{R}^4$ con prodotto scalare canonico,

$$W = \text{span}\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} , \begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \end{bmatrix} \} \qquad \text{e} \qquad v = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix}$$

Determinare la migliore approssimazione di v in W.
Soluzione: Come nel caso precedente, occorre determinare la proiezione ortogonale w^* di v su W. Poiché si hanno *due* generatori di W, le equazioni normali sono un sistema di *due* equazioni in *due* incognite. Poiché i generatori sono *dipendenti*, il sistema ha *infinite* soluzioni. Si ha:

$$F = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} \qquad , \qquad c = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

e le equazioni normali sono: $Fx = c$. L'insieme delle soluzioni è:

$$S = \left\{ \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \lambda \in \mathbb{R} \right\}$$

Tutti gli elementi di S individuano lo *stesso* elemento di W:

$$\left(\tfrac{1}{2}+2\lambda\right)\begin{bmatrix}1\\1\\0\\0\end{bmatrix}-\lambda\begin{bmatrix}2\\2\\0\\0\end{bmatrix}=\tfrac{1}{2}\begin{bmatrix}1\\1\\0\\0\end{bmatrix}=w^*$$

Si ottiene, come giusto, *lo stesso elemento di* W ottenuto nel punto precedente. Per il Teorema di esistenza ed unicità della migliore approssimazione in uno spazio con prodotto scalare, la migliore approssimazione di v in W è w^*.

(3) Siano $I=[0,2\pi]$ e $V=C(I)$ con prodotto scalare definito da:

$$f\cdot g=\frac{1}{\pi}\int_0^{2\pi}f(t)g(t)\,\mathrm{d}t$$

Siano infine:

$$W=\operatorname{span}\{\,1,\cos t,\operatorname{sen}t\,\}\quad \text{e}\quad v=t^2$$

Determinare la migliore approssimazione di v in W.

Soluzione: Come nel caso precedente, occorre determinare la proiezione ortogonale w^* di v su W. Si osservi che in questo esempio lo spazio vettoriale V ha *dimensione infinita.* Poiché si hanno *tre* generatori di W, le equazioni normali sono un sistema di *tre* equazioni in *tre* incognite. Poiché i generatori sono *indipendenti*, il sistema ha *una* soluzione. Si ha:

$$\begin{array}{llll}
1\cdot 1=\frac{1}{\pi}\int_0^{2\pi}\mathrm{d}t=2 & & & v\cdot 1=\frac{8}{3}\pi^2\\
1\cdot\cos t=0 & \cos t\cdot\cos t=1 & & v\cdot\cos t=4\\
1\cdot\operatorname{sen}t=0 & \cos t\cdot\operatorname{sen}t=0 & \operatorname{sen}t\cdot\operatorname{sen}t=1 & v\cdot\operatorname{sen}t=-4\pi
\end{array}$$

e le equazioni normali sono:

$$\begin{bmatrix}2&0&0\\0&1&0\\0&0&1\end{bmatrix}x=\begin{bmatrix}\frac{8}{3}\pi^2\\4\\-4\pi\end{bmatrix}$$

L'unica soluzione è:

$$a=\begin{bmatrix}\frac{4}{3}\pi^2\\4\\-4\pi\end{bmatrix}$$

da cui:

$$w^*=\tfrac{4}{3}\pi^2+4\cos t-4\pi\operatorname{sen}t$$

Per il Teorema di esistenza ed unicità della migliore approssimazione in uno spazio con prodotto scalare, la migliore approssimazione di v in W è w^*.

L'elemento determinato è il *minimo assoluto* della funzione $F:W\to\mathbb{R}$ definita da:

$$F(w)=\frac{1}{\pi}\int_0^{2\pi}\big(t^2-w(t)\big)^2\,\mathrm{d}t$$

* *Osservazione* (Serie di Fourier).

  Sia $g\in C(I)$. Si ricordi che, introdotto in $C(I)$ il prodotto scalare defnito nel punto (3) dell'Esercizio, e posto:

  $$a_0=g\cdot 1\quad,\quad a_k=g\cdot\cos kt\quad,\quad b_k=g\cdot\operatorname{sen}kt\qquad k=1,2,\ldots$$

  si chiama *serie di Fourier generata dalla funzione* g la *serie*:

  $$\tfrac{1}{2}a_0+\sum_{k=1}^{\infty}\big(a_k\cos kt+b_k\operatorname{sen}kt\big)$$

ovvero, introdotte le *somme parziali*:

$$s_0(t) = \tfrac{1}{2}\, a_0 \quad , \quad s_j(t) = \tfrac{1}{2}\, a_0 + \sum_{k=1}^{j} \big(a_k \, \cos kt + b_k \, \operatorname{sen} kt \big) \qquad j = 1, 2, \ldots$$

la *successione*: $s_0(t), s_1(t), \ldots$
Posto:

$$W_k = \operatorname{span} \{ 1, \cos t, \operatorname{sen} t, \ldots, \cos kt, \operatorname{sen} kt \} \qquad k = 0, 1, 2, \ldots$$

la migliore approssimazione w_k^* di g in W_k, ovvero il minimo assoluto *su* W_k della funzione $F : C(I) \to \mathbb{R}$ definita da:

$$F(f) = \frac{1}{\pi} \int_0^{2\pi} \big(g(t) - f(t)\big)^2 \, \mathrm{d}t$$

è s_k. Poiché $W_0 \subset W_1 \subset W_2 \subset \cdots$, la successione $F(w_0^*), F(w_1^*), F(w_2^*), \ldots$ è *non crescente*.
Lo studio della convergenza della successione $s_0(t), s_1(t), \ldots$ è argomento vasto e non elementare.[1]

Calcolo delle soluzioni di un sistema nel senso dei minimi quadrati

Siano $A \in \mathbb{R}^{n \times k}$ di colonne $a_1, \ldots, a_k \in \mathbb{R}^n$, $b \in \mathbb{R}^n$ e si consideri $\mathbb{R}^n$ con prodotto scalare canonico. Allora:

- Le soluzioni di $Ax = b$ nel senso dei minimi quadrati sono le *coordinate* rispetto ad $a_1, \ldots, a_k$, della migliore approssimazione di b in $\operatorname{span}\{ a_1, \ldots, a_k \}$, ovvero della proiezione ortogonale di b su $\operatorname{span}\{ a_1, \ldots, a_k \}$.

  Infatti: La migliore approssimazione di b in $\operatorname{span}\{ a_1, \ldots, a_k \}$ è il minimo assoluto y^* della funzione $d : \operatorname{span}\{ a_1, \ldots, a_k \} \to \mathbb{R}$ definita da:

  $$d(y) = \| \, b - y \, \| = \| \, y - b \, \|$$

  Allora, posto $y^* = Ax^*$, per ogni $x \in \mathbb{R}^k$ si ha $Ax \in \operatorname{span}\{ a_1, \ldots, a_k \}$ e:

  $$\| \, Ax - b \, \|_2 = \| \, Ax - b \, \| \geqslant \| y^* - b \, \| = \| \, Ax^* - b \, \|_2$$

  Dunque, per definizione, x^* è una soluzione di $Ax = b$ nel senso dei minimi quadrati. Le componenti di x^* sono coordinate rispetto ad $a_1, \ldots, a_k$ della migliore approssimazione y^*.

- Le coordinate rispetto ad $a_1, \ldots, a_k$ della migliore approssimazione di b in $\operatorname{span}\{ a_1, \ldots, a_k \}$, ovvero della proiezione ortogonale di b su $\operatorname{span}\{ a_1, \ldots, a_k \}$, sono *le soluzioni* del sistema delle equazioni normali definite dai generatori $a_1, \ldots, a_k$. Ricordando che per ogni $a, b \in \mathbb{R}^n$ si ha $a \cdot b = b^{\mathsf{T}} a$, il sistema delle equazioni normali definite dai generatori $a_1, \ldots, a_k$ si scrive:

  $$A^{\mathsf{T}} A x = A^{\mathsf{T}} b$$

Esercizi

---

1. Sia $W = \operatorname{span}\{ w_1, \ldots, w_j \}$ un sottospazio vettoriale di V, spazio vettoriale su $\mathbb{R}$ con prodotto scalare. Dimostrare che $v \in V$ è ortogonale a W se e solo se:

   $$v \cdot w_i = 0 \qquad i = 1, \ldots, j$$

[1] Si veda, ad esempio: `https://en.wikipedia.org/wiki/Convergence_of_Fourier_series#Norm_convergence`.

2. Sia V uno spazio vettoriale su $\mathbb{R}$ con prodotto scalare. Dimostrare il Teorema di Pitagora: *Siano a e b due elementi di V. Allora*:

$$a \cdot b = 0 \qquad \Rightarrow \qquad \| a + b \|^2 = \| a \|^2 + \| b \|^2$$

3. Si consideri un sistema di riferimento cartesiano nello spazio. Siano poi π il piano di equazione $3\,x_1 - x_2 = 0$ e P il punto di coordinate $(4, 1, 0)$.

   - Verificare che $P \notin \pi$.

   Posto $V = \mathbb{R}^3$ con prodotto scalare canonico, $W = \{\, x \in \mathbb{R}^3 : 3\,x_1 - x_2 = 0 \,\}$ e:

$$v = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}$$

   - Determinare una base di W;
   - Determinare la migliore approssimazione w^* di v in W;
   - Il punto P^* di coordinate le componenti di w^* è il punto di π più vicino a P. Determinare la distanza di P da π.

4. Siano:

$$A = \begin{bmatrix} 0 & -1 \\ 0 & 1 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

   Determinare le equazioni normali definite dalle colonne di A.