Prodotto scalare e matrici

Matrici ortogonali

Consideriamo in $\mathbb{R}^n$ il prodotto scalare canonico

$$< X, Y >= X^T Y = x_1 y_1 + \cdots + x_n y_n.$$

Ci domandiamo se esistono matrici P che conservino il prodotto scalare, ossia che verifichino:

$$< PX, PY >=< X, Y >$$

per ogni $X, Y \in \mathbb{R}^n$.

Osserviamo subito che se $e_1, \ldots, e_n$ sono i vettori della base canonica di $\mathbb{R}^n$, le colonne di P, che sono le loro immagine, dovranno a loro volta formare una base *ortonormale* di $\mathbb{R}^n$, ossia si dovrà verificare $< P^j, P^j >= 1, < P^j, P^i >= 0$ se $i \neq j$.

Ma allora otteniamo:

- $P^T P = I_n$, in particolare il determinante di P è 1 oppure -1.

Definizione Una matrice P a n righe e a n colonne si dice *ortogonale* se verifica $P^T P = I_n$, cioè è invertibile ed ha come inversa la sua trasposta.

Ci sono altre proprietà delle matrici ortogonali.

Proposizione Sia P una matrice ortogonale e sia λ un autovalore di P. Allora $|\lambda| = 1$. Inoltre se V_λ è un autospazio e H è il sottospazio ortogonale a V_λ allora $P(H) = H$.

Prova.

Sia λ un autovalore della matrice ortogonale P e $X \in \mathbb{R}^n$ un autovettore relativo a λ.

Allora $0 \neq < X, X > = < PX, PX > = < \lambda X, \lambda X > = \lambda^2 < X, X >$. Dunque $\lambda^2 = 1$

Sia ora V_λ un autospazio e H il suo ortogonale. Allora per ogni $X \in V_\lambda$ e ogni $Y \in H$ si deve avere $< X, Y > = 0 = < PX, PY > = < \lambda X, PY > = \lambda < X, PY >$. Questo prova che P porta vettori ortogonali a V_λ in vettori ortogonali a V_λ. $\square$

Come esempio costruiamo tutte le matrici ortogonali in dimensione 2.

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Se P è ortogonale si deve avere:

$$\begin{cases} \mathrm{a}^2 + c^2 = 1 \\ \mathrm{ab+cd} = 0 \\ \mathrm{c}^2 + d^2 = 1 \end{cases}$$

Dunque possiamo scrivere $a = \cos\theta, c = \sin\theta$ e per verificare la seconda e la terza equazione abbiamo due scelte: o $b = -\sin\theta, d = \cos\theta$ oppure $b = \sin\theta, d = -\cos\theta$

Nel primo caso è facile vedere che det $(P) = 1$ e che P è una rotazione di un angolo θ in senso antiorario attorno all'origine. Nel secondo caso non si tratta di una rotazione e il determinante vale -1.

Vediamo gli autovalori nei due casi

- P ha determinante 1. Il polinomio caratteristico di P è $x^2 - 2\cos\theta x + 1$. Le radici del polinomio caratteristico sono $\cos\theta + i\sin\theta$ e $\cos\theta - i\sin\theta$ che sono numeri complessi distinti se $\sin\theta \neq 0$ cioè se θ non è un multiplo intero di π, mentre c'è una radice doppia reale che vale 1 se θ è un multiplo pari di π e -1 se θ è un multiplo dispari di π. Si noti che nel primo caso $P = I$ e nel secondo $P = -I$.

- P ha determinante -1. Il polinomio caratteristico di P è $x^2 - 1$ che ha due radici distinte 1 e -1.

In conclusione abbiamo

Una matrice ortogonale 2x2 è sempre diagonalizzabile su $\mathbb{C}$. È diagonalizzabile su $\mathbb{R}$ se è l'identità o meno l'identità nel caso che abbia determinante 1 e sempre se ha determinante -1. In qusto ultimo caso P è la riflessione rispetto ad una retta, retta che costituisce l'autospazio relativo all'autovalore 1.

A questo punto la classificazione delle matrici ortogonali in dimensione 3 diventa facile. Intanto il polinomio caratteristico, avendo grado 3 ha certamente almeno una radice reale che sarà necessariamente 1 o -1. Scelto un autovettore relativo alla radice reale, il piano ortogonale all'autovettore è invariante per P, quindi la restrizione di P a tale piano è una delle trasformazioni prima descritte (rotazione, $I, -I$, simmetria rispetto ad una retta). Otteniamo dunque la seguente situazione

Una matrice ortogonale 3x3 che non sia I con tre autovalori reali è una simmetria rispetto a un piano, rispetto ad una retta o rispetto all'origine se ha autovalore -1 rispettivamente di molteplicità $1, 2$ o 3.

Una matrice ortogonale 3x3 con un solo autovalore reale è una rotazione attorno alla retta autospazio se l'autovalore è 1. Se invece l'autovalore è -1 la trasformazione P è una rotazione attorno alla retta autospazio composta con la simmetria rispetto al piano ortogonale all'autospazio.

Le considerazioni svolte si generalizzano a matrici ortogonali di qualunque dimensione. La prova del teorema seguente è lasciata per esercizio.

Teorema

Sia P una matrice ortogonale a n righe e n colonne. Allora $\mathbb{R}^n$ ammette una decomposizione in somma diretta ortogonale

$$\mathbb{R}^n = V_1 \oplus V_{-1} \oplus H_1 \oplus \cdots \oplus H_s$$

dove V_1 e V_{-1} sono gli autospazi relativi a 1 e a -1, mentre $H_1, \ldots H_s$ sono piani invarianti per P su cui P agisce come una rotazione.

L'algoritmo di Gram-Schmidt

Vogliamo descrivere un algoritmo per trasformare una base di $\mathbb{R}^n$ in una base ortonormale.

Cominciamo con una osservazione: siano v e u due vettori non nulli di $\mathbb{R}^n$. Calcoliamo il vettore proiezione ortogonale di v sulla retta generata da u. Ricordiamo che la lunghezza di v è data da $||v|| =< v, v >^{1/2}$. Inoltre la proiezione che cerchiamo è un multiplo di u che ha come lunghezza il valore assoluto di $||v|| \cos\alpha$ (dove α è l'angolo far v e u) e che è un multiplo positivo di u se α è acuto mentre è un multiplo negativo se α è ottuso.

Dunque

$$pr(v) = ||v|| \cos\alpha \frac{u}{||u||} = \frac{< v, u >}{< u, u >} u$$

Infatti $||pr(v)||^2 =< pr(v), pr(v) >=< v, u >^2= ||v||^2(\cos\alpha)^2$.

Questo è sufficiente per darci l'algoritmo.

Definiamo induttivamente

$v'_1 = v_1, v'_2 = v_2 - \frac{<v_2,v_1>}{<v_1,v_1>} v'_1$

$v'_{i+1} = v_{i+1} - (\frac{<v_{i+1},v'_1>}{<v'_1,v'_1>} v'_1 + \cdots + \frac{<v_{i+1},v'_i>}{<v'_i,v'_i>} v'_i)$

Osserviamo che $v'_1, v'_2, \ldots v'_n$ sono per costruzione a due a due ortogonali fra loro. Inoltre per $i = 1, \ldots, n$ span$(v'_1, \ldots v'_i) =$ span $(v_1, \ldots v_i)$.

Basta a questo punto dividere ciascun v_i' per la sua norma per ottenere una base ortonormale $(u_1, \ldots, u_n)$ di $\mathbb{R}^n$ che verifica inoltre $\text{span}(u_1, \ldots u_i) = \text{span}\,(v_1, \ldots v_i)$.

Il teorema spettrale

Sia A una matrice nxn simmetrica. Calcoliamo:

$$< AX, Y >= (AX)^T Y = X^T A^T Y = X^T AY =< X, AY >$$

Per una matrice simmetrica vale il seguente

Teorema spettrale Sia A una matrice simmetrica, allora esiste una base ortonormale di $\mathbb{R}^n$ formata da autovettori di A.

Questo teorema ci dice parecchie cose.

- Una matrice simmetrica ha tutti gli autovalori reali.
- Gli autospazi di una matrice simmetrica sono tra loro ortogonali.
- Una matrice simmetrica è diagonalizzabile.

La prova del teorema spettrale si compone di tre passi principali che corrispondono proprio alle tre proprietà sopra elencate.

Gli autovalori di una matrice simmetrica sono tutti reali

Per dimostrare questo passo dobbiamo considerare lo spazio vettoriale complesso $\mathbb{C}^n$ e capire come si può calcolare la distanza di un punto di $\mathbb{C}^n$ dall'origine. Se Z ha coordinate $(z_1, \ldots, z_n)$, dovrebbe essere chiaro che la lunghezza di Z al quadrato è data da $|z_1|^2 + \cdots + |z_n|^2 = z_i\overline{z}_1 + \cdots + z_n\overline{z}_n$.

Ma allora possiamo estendere il prodotto scalare di $\mathbb{R}^n$ a una specie di prodotto scalare in $\mathbb{C}^n$ ponendo se $Z, W \in \mathbb{C}^n$

$$< Z, W >= Z^T\overline{W} = z_1\overline{w}_1 + \cdots + z_n\overline{w}_n$$

In questo modo $< Z, Z >= z_i\overline{z}_1 + \cdots + z_n\overline{z}_n = ||Z||^2$.

Bisogna osservare che questo pseudo prodotto scalare è lineare nella prima variabile ma non nella seconda, infatti $< Z, aW >= \overline{a} < Z, W >$.

A questo punto sia λ un autovalore della matrice simmerica A e sia $Z \in \mathbb{C}^n$ un autovettore. Allora $< AZ, Z >=< \lambda Z, Z >= \lambda < Z, Z >= (AZ)^T\overline{Z} = Z^TA^T\overline{Z} = Z^T\overline{AZ} =< Z, AZ >=< Z, \lambda Z >= \overline{\lambda} < Z, Z >$.

Poiché $< Z, Z >\neq 0$ si ottiene $\lambda = \overline{\lambda}$ per cui λ deve essere reale.

Gli autospazi di una matrice simmetrica sono ortogonali fra loro.

Siano $\lambda \neq \mu$ due autovalori distinti della matrice simmetrica A e siano X e Y autovettori relativi rispettivamente a λ e μ. Allora $< AX, Y >=< \lambda X, Y >= \lambda < X, Y >=< X, AY >=< X, \mu Y >= \mu < X, Y >$ Ma $\lambda \neq \mu$ per cui si può avere eguaglianza se e solo se $< X, Y >= 0$.

Una matrice simmetrica è diagonalizzabile

Siano $\lambda_1, \ldots, \lambda_k$ gli autovalori distinti di una matrice simmetrica A. Vogliamo provare che $\mathbb{R}^n = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$. Se così non fosse la somma diretta degli autospazi di A avrebbe uno spazio ortogonale H di dimensione positiva.

Proviamo che H è invariante per A. Sia $h \in H$. Allora $< v, h >= 0$ per ogni vettore $v \in V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$. Ma allora anche $Av \in V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$. Dunque $< Av, h >= 0 =< v, Ah >$. Questo prova che anche Ah è ortogonale a $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$ e quindi che H è invariante per A . Indichiamo con $B : H \to H$ l'applicazione lineare indotta da A sullo spazio H. Allora il polinomio caratteristico di A ha il polinomio caratteristico di B come fattore. Pertanto gli autovalori di B sono anche autovalori di A, in particolare sono reali e sono alcuni tra i $\lambda_1, \ldots, \lambda_k$. Dunque se dim $H > 0$, ci devono essere autovettori in H che sono anche autovettori di A. Questa è una contraddizione perché tutti gli autovettori dell'applicazione lineare A sono contenuti in $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$.

Ora basta scegliere in ogni V_{λ_i} una base ortonormale. Il Teorema spettrale è provato.