Geometria e Algebra (II), 03.12.12

Autovettori, autovalori, matrici diagonalizzabili

1. Consideriamo di nuovo la matrice (cfr. lezione IV)

$$A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix};$$

abbiamo riconosciuto che questa matrice agisce per moltiplicazione in modo particolarmente semplice sui vettori

$$u = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \; v = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

nel senso che moltiplicare A per ciascuno di questi vettori equivale a moltiplicare il vettore per uno scalare

$$Au = 1 \cdot u, \; Av = 0.5 \cdot v;$$

nel linguaggio che abbiamo sviluppato, si dice che u e' un autovettore di A, con autovalore associato 1, e che v e' un autovettore di A, con autovalore associato 0.5.

Abbiamo visto come la conoscenza di questi autovettori ed autovalori, e la circostanza che gli autovettori u e v siano linearmente indipendenti, permetta di esprimere A nella forma

$$A = PDP^{-1},$$

dove $P = [u, \; v]$ e' la matrice che ha per colonne gli autovettori, e D e' la matrice diagonale $D = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0.5 \end{smallmatrix}\right]$ che ha per elementi diagonali gli autovalori associati. Infine, abbiamo visto quanto questa espressione di A sia utile per lo studio delle potenze di A.

Sorgono naturali alcune domande:

(a) quali sono (oltre a u) gli autovettori di A cui e' associato l'autovalore 1?

(b) quali sono (oltre a v) gli autovettori di A cui e' associato l'autovalore 0.5?

(c) quali sono (oltre a 1 e 0.5) gli autovalori di A?

(a). Gli autovettori di A con autovalore associato 1 sono i vettori $x \in \mathbb{R}^2$, con $x \neq 0_2$, tali che

$$Ax = 1 \cdot x, \quad \text{cioe'} \; Ax - x = 0_2, \quad \text{cioe'} \; (A - I_2)x = 0_2.$$

Questo e' un sistema lineare omogeneo di due equazioni in due incognite, con matrice dei coefficienti

$$A - I_2 = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.2 & 0.3 \\ 0.2 & -0.3 \end{bmatrix},$$

dunque e' il sistema

$$\begin{cases} -0.2\,x_1 + 0.3\,x_2 = 0 \\ \;\;\,0.2\,x_1 - 0.3\,x_2 = 0 \end{cases}.$$

In sostanza il sistema consiste di un'unica equazione, che possiamo scrivere

$$2\,x_1 - 3\,x_2 = 0;$$

possiamo risolvere questa equazione ricavando x_1 in funzione di x_2 :

$$x_1 = 1.5\,x_2,$$

cosi' le soluzioni sono del tipo

$$\begin{bmatrix} 1.5\,t \\ t \end{bmatrix} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} t$$

dove t varia in $\mathbb{R}$; la condizione $x \neq 0_2$ impone inoltre $t \neq 0$.

Osserviamo che, ammettendo anche il valore $t = 0$, questi vettori costituiscono un sottospazio di dimensione 1 in $\mathbb{R}^2$; geometricamente, e' una retta che passa per l'origine. L'autovettore u (che si ottiene per $t = 2$) e' uno dei tanti vettori che generano questa retta.

(b). Gli autovettori di A con autovalore associato 0.5 sono i vettori $x \in \mathbb{R}^2$, con $x \neq 0_2$, tali che

$$Ax = 0.5\,x, \quad \text{cioe'}\ Ax - 0.5\,x = 0_2, \quad \text{cioe'}\ (A - 0.5\,I_2)x = 0_2.$$

Questo e' un sistema lineare omogeneo di due equazioni in due incognite, con matrice dei coefficienti

$$A - 0.5\,I_2 = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} - \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.3 \\ 0.2 & 0.2 \end{bmatrix},$$

dunque e' il sistema

$$\begin{cases} 0.3\,x_1 + 0.3\,x_2 = 0 \\ 0.2\,x_1 + 0.2\,x_2 = 0 \end{cases}.$$

In sostanza il sistema consiste di un'unica equazione, che possiamo scrivere

$$x_1 + x_2 = 0;$$

possiamo risolvere questa equazione ricavando x_1 in funzione di x_2 :

$$x_1 = -x_2,$$

cosi' le soluzioni sono del tipo

$$\begin{bmatrix} -t \\ t \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} t,$$

dove t varia in $\mathbb{R}$; la condizione $x \neq 0_2$ impone inoltre $t \neq 0$.

Osserviamo che, ammettendo anche il valore $t = 0$, questi vettori costituiscono un sottospazio di dimensione 1 in $\mathbb{R}^2$; geometricamente, e' una retta che passa per l'origine. L'autovettore v (che si ottiene per $t = 1$) e' uno dei tanti vettori che generano questa retta. Osserviamo inoltre che le due rette degli autovettori (con autovalori associati 1 e 0.5) sono distinte, e hanno in comune solo il vettore nullo.

(c). Uno scalare $\lambda \in \mathbb{R}$ e' un autovalore di A se e solo se e' autovalore associato a qualche autovettore di A, in altri termini se e solo se

$$\exists\, x \in \mathbb{R}^2,\ x \neq 0_2, \quad \text{tale che} \quad Ax = \lambda\, x,$$

cioe'

$$\exists\, x \in \mathbb{R}^2,\ x \neq 0_2, \quad \text{tale che} \quad Ax - \lambda\, x = 0_2,$$

cioe'

$$\exists\, x \in \mathbb{R}^2,\ x \neq 0_2, \quad \text{tale che} \quad (A - \lambda\, I_2)x = 0_2.$$

Questo sistema lineare omogeneo di due equazioni in due incognite ha qualche soluzione $x \neq 0_2$ se e solo se il determinate della matrice dei coefficienti e' nullo:

$$(@) \qquad \det(A - \lambda\, I_2) = 0.$$

Questa e' dunque la condizione che caratterizza gli autovalori di A. Ora,

$$A - \lambda\, I_2 = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 0.8 - \lambda & 0.3 \\ 0.2 & 0.7 - \lambda \end{bmatrix},$$

e

$$\det(A - \lambda\, I_2) = \det \begin{bmatrix} 0.8 - \lambda & 0.3 \\ 0.2 & 0.7 - \lambda \end{bmatrix}$$
$$= (0.8 - \lambda)(0.7 - \lambda) - 0.06 = \lambda^2 - 1.5\,\lambda + 0.5.$$

L'equazione (@) diviene dunque

$$\lambda^2 - 1.5\,\lambda + 0.5 = 0;$$

ha soluzioni

$$\lambda_1 = 1, \quad \lambda_2 = 0.5.$$

Abbiamo ritrovato i due autovalori di A di cui eravamo a conoscenza; possiamo ora dire che non ce ne sono altri.

2. **Autovalori, polinomio caratteristico.**

Ci chiediamo ora come si possano determinare gli autovalori di una qualsiasi matrice $A = [a_{ij}]$ quadrata di ordine n. Uno scalare $\lambda \in \mathbb{R}$ e' un autovalore di A se e solo se e' autovalore associato a qualche autovettore di A, cioe' se e solo se

$$\exists\, x \in \mathbb{R}^n,\ x \neq 0_n, \quad \text{tale che} \quad Ax = \lambda\, x,$$

cioe'

$$\exists\, x \in \mathbb{R}^n,\ x \neq 0_n, \quad \text{tale che} \quad (A - \lambda\, I_n)x = 0_n.$$

In altri termini, λ e' un autovalore di A se e solo se il sistema lineare omogeneo di n equazioni in n incognite $(A - \lambda\, I_n)x = 0_n$ ha qualche soluzione diversa dalla soluzione banale $x = 0_n$. Per il teorema su determinanti e sistemi lineari (cfr. Th. 4, lez. VI), cio' capita se e solo se il determinate della matrice dei coefficienti e' nullo:

$\det(A - \lambda I_n) = 0.$

Questa e' dunque la equazione che caratterizza gli autovalori di A; viene detta "equazione caratteristica" di A. Il primo membro

$$\det(A - \lambda I_n) = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{bmatrix}$$

e' un polinomio di grado n in λ; viene detto "polinomio caratteristico" di A.

Abbiamo dunque visto che

> *Gli autovalori di una matrice A quadrata di ordine n sono le soluzioni dell'equazione caratteristica det $(A - \lambda I_n) = 0$ di A, in altri termini sono le radici del polinomio caratteristico det $(A - \lambda I_n)$ di A.*

Cio' ha una immediata conseguenza: una matrice quadrata di ordine n ha al piu' n autovalori.

3. Puo' succedere che una matrice non abbia alcun autovalore. Cio' capita ad esempio per la matrice

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Infatti il suo polinomio caratteristico

$$\det(A - \lambda I_2) = \det \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = \lambda^2 + 1$$

non ha alcuna radice reale.

Si potrebbero considerare scalari, vettori (e matrici) ad elementi complessi; allora vedremmo che la matrice A ha due autovalori complessi coniugati $\pm i$, cui corrispondono autovettori ad elementi complessi. Noi pero' qui ci limitiamo e sempre nel nostro discorso ci limiteremo al caso reale.

4. **Autovettori, autospazi.**

Consideriamo ora una matrice $A = [a_{ij}]$ quadrata di ordine n e un suo autovalore $\lambda_0 \in \mathbb{R}$. Gli autovettori di A con autovalore associato λ_0 sono le soluzioni dell'equazione

$$Ax = \lambda_0 \, x, \quad \text{cioe'}\ (A - \lambda_0 \, I_n)x = 0_n$$

nella variabile vettoriale $x \in \mathbb{R}^n$, soggetta alla condizione $x \neq 0_n$; in altri termini, sono le soluzioni non banali del sistema lineare omogeneo

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = \lambda_0 x_1 \\ \qquad\qquad\qquad\qquad \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = \lambda_0 x_n \end{cases},$$

cioe'

$$\begin{cases} (a_{11}-\lambda_0)x_1+\cdots+ \qquad a_{1n}x_n=0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ \qquad a_{n1}x_1+\cdots+(a_{nn}-\lambda_0)x_n=0 \end{cases}.$$

Per il solo fatto che questo e' un sistema lineare omogeneo, l'insieme di tutte le sue soluzioni (compresa quella banale) e' un sottospazio di $\mathbb{R}^n$.

Dunque l'insieme degli autovettori di A con autovalore associato λ_0, piu' il vettore nullo 0_n, e' un sottospazio di $\mathbb{R}^n$; questo sottospazio viene detto "autospazio" di A con autovalore associato λ_0, e viene indicato con V_{λ_0}; la sua dimensione e' data da

$$\dim V_{\lambda_0} = n - r\,(A-\lambda_0 I_n)\,.$$

Per la definizione di autovettore ed autovalore, ogni autospazio di A contiene qualche vettore diverso dal vettore nullo 0_n.

Siano ora λ_1 e λ_2, con $\lambda_1 \neq \lambda_2$, due autovalori distinti di A. Consideriamo i due autospazi V_{λ_1} e V_{λ_2} di A con autovalori associati λ_1 e λ_2. Ci chiediamo cosa si puo' dire sulla loro intersezione

$$V_{\lambda_1} \frown V_{\lambda_2}.$$

Osserviamo che, se $v \in V_{\lambda_1} \frown V_{\lambda_2}$, allora si ha

$$Av = \lambda_1\, v, \quad \text{e} \quad Av = \lambda_2\, v,$$

da cui

$$\lambda_1\, v = \lambda_2\, v, \quad \text{cioe'} \quad (\lambda_1-\lambda_2)\, v = 0_n$$

da cui infine, essendo $\lambda_1 - \lambda_2 \neq 0$, si ottiene $v = 0_n$. Abbiamo cosi' che i due autospazi hanno in comune solo il vettore nullo:

$$V_{\lambda_1} \frown V_{\lambda_2} = \{0_n\} \qquad (\lambda_1 \neq \lambda_2).$$

5. **Matrici diagonalizzabili.**

Abbiamo visto nella lezione (IV) come il problema dello studio dell'evoluzione di un sistema possa essere espresso come lo studio delle potenze di una matrice, e questo studio possa essere affrontato usando autovettori ed autovalori della matrice. Ci sono altri problemi che possono essere affronati con l'uso di autovettori ed autovalori.

Il punto cruciale e' sempre quello di esprimere, se possibile, una matrice quadrata in funzione di una matrice diagonale ad essa collegata.

Sia A una matrice quadrata di ordine n; se A possiede n autovettori v_i con autovalori associati λ_i, cioe'

$$Av_i = \lambda_i\, v_i, \qquad i = 1, \ldots, n,$$

allora si ha

$$A\,[v_1,\ \ldots,\ v_n] = [v_1,\ \ldots,\ v_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix},$$

e in breve, indicata con P la matrice cha ha per colonne gli autovettori v_i e con D la matrice diagonale che ha per elementi diagonali gli autovalori λ_i ad essi associati,

$$AP = PD.$$

Ora, se gli autovettori $v_1, \ldots, v_n$ sono linearmente indipendenti, equivalentemente se gli autovettori $v_1, \ldots, v_n$ formano una base di $\mathbb{R}^n$, allora la matrice P e' invertibile, e possiamo ricavare A in funzione di P e di D :

$$A = PDP^{-1}.$$

Definizione 1 *Diciamo che una matrice A quadrata di ordine n e' "diagonalizzabile" se A si puo' scrivere come*

$$A = PDP^{-1},$$

dove D e' una matrice diagonale e P e' una matrice invertibile.

In base all'argomento sviluppato sopra si ha che se una matrice quadrata A possiede n autovettori che formano una base di $\mathbb{R}^n$, allora e' diagonalizzabile. E' facile verificare che vale anche il viceversa. dunque si ha

Teorema 1 *Una matrice A e' diagonalizzabile se e solo se A possiede n autovettori che formano una base di $\mathbb{R}^n$.*

6. Una matrice che non ha alcun autovalore, come quella considerata in precedenza, chiaramente non e' diagonalizzabile. Puo' succedere anche che una matrice abbia qualche autovalore, dunque qualche autovettore, ma non abbastanza per essere diagonalizzabile. Cio' capita ad esempio per la matrice

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Il suo polinomio caratteristico

$$\det(A - \lambda I_2) = \det \begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix} = \lambda^2$$

ha solo la radice 0; dunque A ha il solo autovalore $\lambda = 0$. L'autospazio V_0 e' l'insieme delle soluzioni del sistema

$$Ax = 0_2,$$

che consiste dell'unica equazione $x_2 = 0$; e dunque ha soluzioni i vettori $\left[\begin{smallmatrix} t \\ 0 \end{smallmatrix}\right]$, dove t varia in $\mathbb{R}$. Fra questi vettori se ne puo' prendere al piu' uno che sia linearmente indipendente, dunque non se ne possono prendere due che formino una base di $\mathbb{R}^2$. La matrice A non e' diagonalizzabile.

7. L'equazione caratteristica di una matrice quadrata di ordine n e' un'equazione di grado n; la determinazione delle sue soluzioni e' in generale un problema difficile. Infatti: e' ben nota una formula per le soluzioni di un'equazione di II grado, ci sono formule per le soluzioni di un'equazione di III e IV grado, ma non c'e' una formula per le soluzioni della generica equazione di grado $n \geq 5$.

   Le matrici concrete che noi considereremo avranno di regola ordine 2 o 3. Le equazioni caratteristiche di III grado saranno risolubili ricondicendosi per scomposizione al II grado. Ad esempio le equazioni del tipo

   $$a\lambda^3 + b\lambda^2 + ca\lambda + cb = 0$$

   si possono risolvere scomponendo

   $$a\lambda^3 + b\lambda^2 + ca\lambda + cb = (a\lambda + b)\lambda^2 + c(a\lambda + b) = (a\lambda + b)(\lambda^2 + c).$$

   Una nota finale. Le equazioni di grado n dispari hanno sempre almeno una soluzione reale. Infatti, se $p(\lambda) = a\lambda^n + \cdots$ e' un polinomio di grado n dispari, con $a > 0$ (un argomento analogo vale per $a < 0$), allora si ha $\lim_{\lambda \to +\infty} p(\lambda) = +\infty$ e $\lim_{\lambda \to -\infty} p(\lambda) = -\infty$, cosi' esistono $a, b \in \mathbb{R}$ con $a < b$ tali che $p(a) < 0$ e $p(b) > 0$; essendo la funzione $p(\lambda)$ continua, esiste un punto $c \in]a, b[$ tale che $p(c) = 0$.