Un breve compendio su autovalori e autovettori (5 novembre 2019)
Corso di Matematica I per Geologia

Autovalori e autovettori. Data una matrice quadrata A, ogni vettore $\vec{v} \neq \vec{0}$ che soddisfa

$$A\vec{v} = \lambda\vec{v} \tag{1}$$

per qualche costante λ si dice *autovettore* di A. Il valore λ è il corrispondente *autovalore*. L'equazione (1) si può equivalentemente riscrivere nella forma:

$$(A - \lambda I)\vec{v} = \vec{0}, \tag{2}$$

dove I è la matrice identità. La (2) è un sistema lineare che sicuramente ammette la soluzione nulla $\vec{v} = \vec{0}$. Noi siamo interessati a trovare una soluzione non nulla $\vec{v} \neq \vec{0}$. Affinchè ciò sia possibile, dobbiamo richiedere che

$$\det(A - \lambda I) = 0, \tag{3}$$

che è l'equazione caratteristica per gli autovalori. Una volta trovata una soluzione $\lambda = \lambda^*$ alla (3), un autovettore corrispondente, $\vec{v}^*$, si trova risolvendo la (1).

Caso 2×2. Se A è una matrice 2×2 di elementi $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, l'equazione caratteristica diventa:

$$\det\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda) - bc = 0 \iff \lambda^2 - (a+d)\lambda + (ad-bc) = 0, \tag{4}$$

che è un'equazione quadratica per λ, le cui soluzioni sono

$$\lambda_\pm = \frac{1}{2}\left[a + d \pm \sqrt{(a+d)^2 - 4(ad-bc)}\right]. \tag{5}$$

Una volta identificati gli autovalori, gli autovettori si ricavano imponendo la (1) in corrispondenza di $\lambda = \lambda_\pm$. Consideriamo ad esempio l'autovalore λ_+. Un autovettore corrispondente $\vec{v}_+$ è una soluzione non nulla dell'equazione $A\vec{v}_+ = \lambda_+\vec{v}_+$. Più esplicitamente, se chiamiamo x e y le due coordinate (incognite) di $\vec{v}_+$, l'equazione per determinarle assume la forma:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda_+\begin{pmatrix} x \\ y \end{pmatrix} \iff \begin{cases} ax + by = \lambda_+ x \\ cx + dy = \lambda_+ y \end{cases}. \tag{6}$$

La condizione $\det(A - \lambda_+ I) = 0$ implica che le due equazioni del sistema sono equivalenti (ovvero: la seconda si ottiene dalla prima moltiplicandola per una costante opportuna). Quindi, ogni coppia di costanti (x, y) non simultaneamente nulle, tali che $(a - \lambda_+)x + by = 0$ è soluzione della (6), ad esempio $x = b$, $y = \lambda_+ - a$. Ogni altra soluzione è proporzionale a questa. In conclusione, ogni vettore della forma

$$\vec{v}_+ = \alpha\begin{pmatrix} b \\ \lambda_+ - a \end{pmatrix}, \tag{7}$$

con α una costante qualsiasi diversa da zero, è un autovettore di autovalore λ_+. Una procedura analoga è valida per ricavare l'autovettore $\vec{v}_-$ di autovalore λ_-.

Un esempio. Consideriamo la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$ e calcoliamone autovalori ed autovettori. Gli autovalori sono le soluzioni di

$$\det\begin{pmatrix} 1-\lambda & 2 \\ 3 & 1-\lambda \end{pmatrix} = 0 \iff (1-\lambda)^2 - 6 = 0 \iff \lambda^2 - 2\lambda - 5 = 0,$$

ovvero $\lambda_\pm = 1 \pm \sqrt{6}$, che sono quindi gli autovalori di A. Calcoliamo un autovettore $\vec{v}_+$ associato a $\lambda_+ = 1+\sqrt{6}$: l'equazione per le sue coordinate è

$$\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (1+\sqrt{6}) \begin{pmatrix} x \\ y \end{pmatrix} \quad \Longleftrightarrow \quad \begin{cases} x + 2y = (1+\sqrt{6})x \\ 3x + y = (1+\sqrt{6})y \end{cases} \quad \Longleftrightarrow \quad \begin{cases} -\sqrt{6}x + 2y = 0 \\ 3x - \sqrt{6}y = 0 \end{cases} \tag{8}$$

Con riferimento all'ultimo sistema di due equazioni, si noti che moltiplicando la prima equazione per $-\sqrt{6}$ si ottiene la seconda, che quindi è equivalente alla prima. In conclusione, ogni coppia di valori x e y tali che $-\sqrt{6}x + 2y = 0$ sono soluzioni della (8), ad esempio: $x = 2$, $y = \sqrt{6}$, od ogni altra coppia a questa proporzionale. Equivalentemente, ogni vettore della forma

$$\vec{v}_+ = \alpha \begin{pmatrix} 2 \\ \sqrt{6} \end{pmatrix}$$

con α una qualsiasi costante reale diversa da zero, è autovettore di autovalore $\lambda_+ = 1+\sqrt{6}$. Ripetendo lo stesso ragionamento per $\lambda_- = 1-\sqrt{6}$, si trova che ogni vettore della forma $\vec{v}_- = \alpha \begin{pmatrix} 2 \\ -\sqrt{6} \end{pmatrix}$, con α una qualsiasi costante reale diversa da zero, è autovettore di autovalore $\lambda_- = 1-\sqrt{6}$ (lo si verifichi per esercizio).

Osservazioni e commenti. Per semplicità ci limitiamo a commentare il caso 2×2. Commenti analoghi sono validi per matrici quadrate di dimensione maggiore.

1. Come discusso sopra, se $\vec{v}^* \neq \vec{0}$ è autovettore di A con autovalore λ^*, allora anche $\alpha\vec{v}^*$, con α una qualsiasi costante diversa da zero, è autovettore di A con lo stesso autovalore. Ogni autovettore di A identifica quindi una direzione nel piano o, equivalentemente, una retta passante per l'origine; tale direzione talvolta si chiama 'autodirezione' della matrice A. Per identificare tale direzione è sufficiente assegnare un vettore di lunghezza unitaria. Quindi, senza perdita di generalità, si può sempre suppore che gli autovettori di A siano di lunghezza 1.

2. L'interesse di identificare gli autovettori di A sta nel fatto che l'azione di A sugli autovettori è la più semplice possibile: consiste semplicemente nella moltiplicazione per una costante (il corrispondente autovalore).

3. Se una matrice A ammette due autovettori linearmente indipendenti, $\vec{v}_+$ e $\vec{v}_-$, allora tali autovettori identificano due direzioni trasversali nel piano. Ogni vettore $\vec{v}$ del piano può quindi essere scritto come somma di un vettore parallelo a $\vec{v}_+$ e di uno parallelo a $\vec{v}_-$; più precisamente, dato un qualsiasi vettore $\vec{v}$, esistono due costanti c_+ e c_- tali che

$$\vec{v} = c_+\vec{v}_+ + c_-\vec{v}_-.$$

Una volta che $\vec{v}$ viene scritto così, l'azione di A su $\vec{v}$ assume una forma particolarmente semplice:

$$A\vec{v} = A(c_+\vec{v}_+ + c_-\vec{v}_-) = c_+\, A\vec{v}_+ + c_-\, \vec{v}_- = c_+\lambda_+\vec{v}_+ + c_-\lambda_-\vec{v}_-,$$

dove $\lambda_\pm$ sono gli autovalori di $\vec{v}_\pm$.

4. Dalla formula risolutiva (5), si vede che $\lambda_+ + \lambda_- = a+d$, mentre $\lambda_+ \cdot \lambda_- = ad - bc$ (lo si verifichi per esercizio). La combinazione $a+d$ corrisponde alla somma degli elementi sulla diagonale principale di A, e si chiama 'traccia della matrice A'; per indicarla, si usa il simbolo $\mathrm{Tr}A$. La combinazione $ad - bc$ invece non è altro che il determinante della matrice A. Gli autovalori $\lambda_\pm$ sono quindi i due numeri tali che la loro somma è pari alla traccia di A, mentre il loro prodotto è il determinante di A:

$$\lambda_+ + \lambda_- = \mathrm{Tr}A, \qquad \lambda_+ \cdot \lambda_- = \det A.$$

5. Non sempre gli autovalori sono reali: con riferimento alla (5), se $(a+d)^2 - 4(ad-bc) < 0$, allora $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ammette due autovalori complessi coniugati. Ad esempio, si consideri la matrice

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

la cui azione su un vettore $\vec{v}$ del piano consiste nella rotazione di $\vec{v}$ di un angolo θ in senso antiorario. Gli autovalori si ottengono risolvendo la seguente equazione:

$$\det \begin{pmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{pmatrix} = 0 \quad \Longleftrightarrow \quad \lambda^2 - 2\cos\theta\,\lambda + 1 = 0,$$

la cui soluzione è

$$\lambda_\pm = \cos\theta \pm \sqrt{(\cos\theta)^2 - 1}.$$

Ricordando che $\cos^2\theta = 1 - \sin^2\theta$, gli autovalori si possono equivalentemente riscrivere come

$$\lambda_\pm = \cos\theta \pm i\sin\theta,$$

dove i è l'unità immaginaria. Quindi, se $\theta \neq 0, \pi$, gli autovalori non sono reali. Allo stesso tempo, gli autovettori corrispondenti non sono reali. C'è una semplice interpretazione geometrica di questo fatto: visto che l'azione di A su un qualsiasi vettore del piano consiste nel ruotarlo di un angolo θ, non esiste nessuna direzione del piano che rimane invariata sotto l'azione di A; di conseguenza, non esistono autovettori reali di A.

6. Visto che gli autovalori di una matrice 2×2 sono le soluzioni di un'equazione quadratica (vedi (5)), ci sono tre casi: o gli autovalori sono reali distinti, o reali coincidenti, o non sono reali (e in tal caso i due autovalori sono complessi, e l'uno è il complesso coniugato dell'altro).

   - Se gli autovalori sono reali distinti, allora esistono due autovettori reali linearmente indipendenti, l'uno associato al primo autovalore, e l'altro al secondo.
   - Se gli autovalori non sono reali, allora gli autovettori sono anch'essi complessi, e possono essere scelti l'uno il complesso coniugato dell'altro.
   - Se gli autovalori sono reali coincidenti, ci sono ancora due casi: o la matrice A è proporzionale alla matrice identità, ovvero $A = \begin{pmatrix} \gamma & 0 \\ 0 & \gamma \end{pmatrix}$ per qualche $\gamma \in \mathbb{R}$; in tal caso tutti i vettori non nulli del piano sono autovettori con autovalore γ. Oppure A non è proporzionale alla matrice identità. In tal caso esiste una sola autodirezione associata all'unico autovalore di A; in altre parole, non esistono due autovettori di A linearmente indipendenti.

7. Un teorema (di cui non discuteremo la dimostrazione) garantisce che, se A è una matrice *simmetrica*, se cioè i due elementi b e c fuori dalla diagonale principale sono uguali tra loro ($b = c$), allora gli autovalori di A sono reali ed esistono sempre due autovettori tra loro ortogonali (e, in particolare, linearmente indipendenti).

8. Se $\vec{v}^*$ è autovettore di A con autovalore λ^*, allora è anche autovettore di A^2, con autovettore $(\lambda^*)^2$. Infatti, $A^2\vec{v}^* = A \cdot A\vec{v}^* = A(A\vec{v}^*) = A(\lambda^*\vec{v}^*)$, dove nell'ultimo passaggio abbiamo usato il fatto che $A\vec{v}^* = \lambda^*\vec{v}^*$. Ma quindi $A^2\vec{v}^* = A(\lambda^*\vec{v}^*) = \lambda^* A\vec{v}^* = (\lambda^*)^2\vec{v}^*$, che dimostra, come promesso, che $\vec{v}^*$ è autovettore di A^2 con autovalore $(\lambda^*)^2$. Ragionando in modo analogo si trova che $\vec{v}^*$ è autovettore di A^n, qualunque sia la potenza intera $n > 0$, con autovalore $(\lambda^*)^n$. Inoltre, se A è invertibile, $\vec{v}^*$ è anche autovettore di A^{-1}, con autovalore $(\lambda^*)^{-1}$.

Matrici diagonalizzabili e diagonalizzazione. Una matrice quadrata A, di dimensione $n \times n$, che ammette n autovettori linearmente indipendenti (o, in altre parole, che ammette una *base* di autovettori) di dice **diagonalizzabile**. Due condizioni *sufficienti* che garantiscono la diagonalizzabilità di una matrice sono le seguenti: (1) A è simmetrica, oppure (2) A ammette n autovalori reali distinti. Il motivo del nome *diagonalizzabile* è che ogni matrice che ammette una base di autovettori può essere diagonalizzata (o, in altre parole, resa diagonale) da un'opportuna trasformazione: in formule, esiste una matrice Q invertibile tale che $A = Q^{-1}DQ$, dove D è una matrice diagonale con elementi diagonali $\lambda_1, \dots, \lambda_n$, gli n autovalori di A; la matrice Q può essere costruita scrivendo per colonne gli autovettori $\vec{v}_1, \dots, \vec{v}_n$ corrispondenti agli autovalori $\lambda_1, \dots, \lambda_n$; il fatto che i vettori $\vec{v}_1, \dots, \vec{v}_n$ siano linearmente indipendenti garantisce che Q sia invertibile.

Caso 2×2. Sia A una matrice diagonalizzabile 2×2 di elementi $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ e autovalori λ_+ e λ_-: questo vuol dire che esistono due vettori $\vec{v}_+, \vec{v}_-$ linearmente indipendenti tali che $A\vec{v}_+ = \lambda_+\vec{v}_+$ e $A\vec{v}_- = \lambda_-\vec{v}_-$. In componenti, queste due equazioni prendono la forma

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} v_{+,1} \\ v_{+,2} \end{pmatrix} = \lambda_+ \begin{pmatrix} v_{+,1} \\ v_{+,2} \end{pmatrix} = \begin{pmatrix} \lambda_+ v_{+,1} \\ \lambda_+ v_{+,2} \end{pmatrix}, \tag{9}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} v_{-,1} \\ v_{-,2} \end{pmatrix} = \lambda_- \begin{pmatrix} v_{-,1} \\ v_{-,2} \end{pmatrix} = \begin{pmatrix} \lambda_- v_{-,1} \\ \lambda_- v_{-,2} \end{pmatrix}, \tag{10}$$

dove abbiamo indicato con $v_{+,1}, v_{+,2}$ le due componenti di $\vec{v}_+$ e analogamente per $\vec{v}_-$.

Introduciamo la matrice Q le cui colonne siano le componenti dei vettori $\vec{v}_+$ e $\vec{v}_-$, rispettivamente:

$$Q = \begin{pmatrix} v_{+,1} & v_{-,1} \\ v_{+,2} & v_{-,2} \end{pmatrix}.$$

Moltiplicando A per Q righe per colonne e usando le (9)-(10) troviamo:

$$A \cdot Q = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} v_{+,1} & v_{-,1} \\ v_{+,2} & v_{-,2} \end{pmatrix} = \begin{pmatrix} \lambda_+ v_{+,1} & \lambda_- v_{-,1} \\ \lambda_+ v_{+,2} & \lambda_- v_{-,2} \end{pmatrix}.$$

Il membro di destra di tale equazione può essere riscritto come segue:

$$\begin{pmatrix} \lambda_+ v_{+,1} & \lambda_- v_{-,1} \\ \lambda_+ v_{+,2} & \lambda_- v_{-,2} \end{pmatrix} = \begin{pmatrix} v_{+,1} & v_{-,1} \\ v_{+,2} & v_{-,2} \end{pmatrix} \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \equiv Q \cdot D,$$

dove nell'ultimo passaggio abbiamo chiamato D la matrice diagonale di elementi diagonali λ_+ e λ_-.

In conclusione,

$$A \cdot Q = Q \cdot D$$

e quindi, moltiplicando da sinistra a membro a membro per Q^{-1} troviamo

$$Q^{-1} \cdot A \cdot Q = Q^{-1} \cdot Q \cdot D \quad \Rightarrow \quad Q^{-1} \cdot A \cdot Q = D,$$

dove nell'ultimo passaggio abbiamo usato che $Q^{-1} \cdot Q = I$ con $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ la matrice identità 2×2 e che $I \cdot D = D$. Analogamente, moltiplicando l'identità $A \cdot Q = Q \cdot D$ da destra per Q^{-1} otteniamo $A \cdot Q \cdot Q^{-1} = Q \cdot D \cdot Q^{-1}$, ovvero

$$A = Q \cdot D \cdot Q^{-1}.$$

Si noti che il fatto che Q sia invertibile (ovvero $\det Q \neq 0$) è equivalente alla condizione che $\vec{v}_+, \vec{v}_-$ siano linearmente indipendenti (o, in altre parole, che $\vec{v}_+$ e $\vec{v}_-$ non siano tra loro paralleli).