



asdf

LA DIAGONALIZZAZIONE

30 January 2012

L'intento di questo articolo è di affrontare, si spera sempre nel modo più corretto e chiaro possibile, la trattazione di un argomento importante nell'ambito dell'algebra lineare quale quello della **diagonalizzazione**.

Cercherò di affrontare l'argomento in maniera graduale percorrendo le varie tappe necessarie alla comprensione dell'argomento nel suo insieme.

Data la lunghezza dell'articolo, per orientarsi nella sua consultazione, si consiglia di cliccare sul pulsante **Mostra indice** posto in alto. Buona lettura.

Polinomi di matrici

Consideriamo, su un campo K , il polinomio seguente:

$$f(t) = a_n t^n + \cdots + a_1 t + a_0 \quad .$$

Data una matrice A quadrata, essa può essere definita come:

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I$$

dove I è la nota matrice identità.

In particolare, diremo che A è una **radice**, o uno **zero**, del polinomio $f(t)$ se $f(A) = 0$.

Esempio

Abbiamo la matrice:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; \text{ si avrà quindi che } A^2 = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} \quad .$$

Date inoltre:

$$f(t) = 2t^2 - 3t + 5 \quad g(t) = t^2 - 5t - 2,$$

avremo quanto segue.

$$f(A) = 2A^2 - 3A + 5I = \begin{bmatrix} 14 & 20 \\ 30 & 44 \end{bmatrix} + \begin{bmatrix} -3 & -6 \\ -9 & -12 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 16 & 14 \\ 21 & 37 \end{bmatrix}$$

e

$$g(A) = A^2 - 5A - 2I = \begin{bmatrix} 7 & 10 \\ 15 & -20 \end{bmatrix} + \begin{bmatrix} -5 & -10 \\ -15 & -20 \end{bmatrix} + \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Si deduce quindi che A è uno zero di $g(t)$.

Enunciamo ora il seguente **teorema**.

Siano f e g dei polinomi su K . Per ogni matrice A e per ogni scalare k ,

1. $(f + g)(A) = f(A) + g(A)$;
2. $(fg)(A) = f(A)g(A)$;
3. $(kf)(A) = kf(A)$;
4. $f(A)g(A) = g(A)f(A)$ (due polinomi in A **commutano**).

Le matrici e gli operatori lineari

Consideriamo ora un operatore lineare $T : V \rightarrow V$ su uno spazio vettoriale V . Le potenze di T sono definite mediante l'operazione di composizione:

$$T^2 = T \circ T, \quad T^3 = T^2 \circ T, \quad \dots$$

Per ogni polinomio $f(t) = a_n t^n + \dots + a_1 t + a_0$ definiamo $f(T)$ in maniera uguale a quanto fatto per le matrici, cioè:

$$f(T) = a_n T^n + \dots + a_1 T + a_0 I$$

dove I è l'**applicazione identità**. Si dirà quindi che T è una **radice** o **zero** di $f(t)$ se $f(T) = 0$.

Va precisato che quanto esposto al teorema prima enunciato vale tanto per gli operatori lineari quanto per le matrici.

Polinomio caratteristico e teorema di Cayley - Hamilton

Prendiamo ora in considerazione una matrice quadrata A di ordine n .

La matrice $M = A - tI_n$, dove I_n è la matrice quadrata identità di dimensione n e t è una incognita, può essere ottenuta se si sottrae t dalla diagonale A .

L'opposta della matrice M è la matrice $tI_n - A$ ed il suo determinante:

$$\Delta(t) = \det(tI_n - A) = (-1)^n \det(A - tI_n)$$

è un polinomio di grado n in t ed è più precisamente detto **polinomio caratteristico** di A .

Passiamo ora ad enunciare il seguente importante teorema dovuto a **Cayley** ed **Hamilton**.

Il **teorema di Cayley - Hamilton** afferma che ogni matrice A è uno zero del suo polinomio caratteristico.

Esempio

Facciamo un piccolo esempio a riguardo.

Prendiamo in considerazione la matrice A (quadrata):

$$A = \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix}.$$

Scriviamone il polinomio caratteristico:

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -3 \\ -4 & t-5 \end{vmatrix} = (t-1)(t-5) - 12 = t^2 - 6t - 7$$

Possiamo verificare quanto dichiarato dal teorema di Cayley - Hamilton e cioè che A è uno zero di $\Delta(t)$:

$$\Delta(A) = A^2 - 6A - 7I = \begin{bmatrix} 13 & 18 \\ 24 & 37 \end{bmatrix} + \begin{bmatrix} -6 & -18 \\ -24 & -30 \end{bmatrix} + \begin{bmatrix} -7 & 0 \\ 0 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Si dimostra inoltre che **matrici simili hanno lo stesso polinomio caratteristico**.

Alcune formule per il calcolo di polinomi caratteristici di matrici di ordine 2 e 3

Ordine 2

Data la matrice $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, allora :

$$\Delta(t) = t^2 - (a_{11} + a_{22})t + \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = t^2 - \text{tr}(A)t + \det(A)$$

dove $\text{tr}(A)$ è la **traccia** di A , cioè la somma degli elementi diagonali.

Ordine 3

Data la matrice $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, allora :

$$\Delta(t) = t^3 - \text{tr}(A)t^2 + (A_{11} + A_{22} + A_{33})t - \det(A)$$

dove A_{11} , A_{22} , A_{33} sono i **cofattori** degli elementi diagonali.

Possiamo enunciare quindi il seguente **teorema**.

Data una matrice quadrata di ordine n , il suo polinomio caratteristico è :

$$\Delta(t) = t^n - S_1 t^{n-1} + S_2 t^{n-2} - \dots + (-1)^n S_n$$

dove S_k è la **somma dei minori principali di ordine k** .

Esempio

Scriviamo il polinomio caratteristico della matrice :

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 3 & 2 \\ 1 & 3 & 9 \end{bmatrix}.$$

Si ha $\text{tr}(A) = 1 + 3 + 9 = 13$. I cofattori degli elementi sulla diagonale sono:

$$A_{11} = \begin{vmatrix} 3 & 2 \\ 3 & 9 \end{vmatrix} = 21, \quad A_{22} = \begin{vmatrix} 1 & 2 \\ 1 & 9 \end{vmatrix} = 7, \quad A_{33} = \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix}$$

Quindi $A_{11} + A_{22} + A_{33} = 31$. Inoltre $|A| = 27 + 2 + 0 - 6 - 6 - 0 = 17$

Conseguentemente:

$$\Delta(t) = t^3 - 13t^2 + 31t - 17$$

Polinomio caratteristico per un operatore lineare

Consideriamo ancora un operatore lineare $T : V \rightarrow V$ su uno spazio vettoriale V di dimensione finita.

Il polinomio caratteristico $\Delta(t)$ di T è definito come il polinomio caratteristico di una qualunque matrice che rappresenta T .

Se A e B sono matrici di rappresentazione di T , si ha $B = P^{-1}AP$, dove P è la matrice del **cambiamento di base**. A e B sono quindi simili e giacché due matrici simili hanno lo stesso polinomio caratteristico (lo abbiamo detto precedentemente), esse hanno lo stesso polinomio caratteristico.

Il polinomio caratteristico di T è quindi indipendente dalla base rispetto a cui è calcolata la sua matrice di rappresentazione.

$f(T) = 0$ se e solo se $f(A) = 0$, dove $f(t)$ è un qualunque polinomio ed A è una matrice qualunque che rappresenta T : fatte salve tale considerazioni si può enunciare il seguente teorema per un operatore lineare.

Un'operatore lineare T è uno zero del suo polinomio caratteristico.

La diagonalizzazione: autovalori ed autovettori

Consideriamo la nostra matrice quadrata A di ordine n .

Diremo che è possibile rappresentare A con una matrice diagonale:

$$D = \text{diag}(k_1, k_2, \dots, k_n)$$

se e solo se esiste una base S costituita da vettori (colonna) $u_1, u_2, \dots, u_n$, tali che:

$$\begin{array}{rcl} Au_1 & = & k_1 u_1 \\ Au_2 & = & k_2 u_2 \\ \dots & \dots & \dots \\ Au_n & = & k_n u_n \end{array}$$

In questo caso A si dice **diagonalizzabile**.

Inoltre: sia A una matrice qualunque e sia λ uno scalare appartenente a K . Ebbene λ si dirà **autovalore** di A se esiste un vettore (colonna) non nullo $v \in K^n$ tale che :

$$Av = \lambda v$$

Si dirà invece **autovettore** di A relativo all'autovalore λ ogni vettore che soddisfa tale relazione.

L'insieme E_λ di tutti gli autovettori relativi a λ è un sottospazio di V e si chiama **autospazio** di λ .

Enunciamo ora il seguente **teorema**.

Una matrice quadrata A di dimensione n è simile ad una matrice diagonale D se e solo se A ha n autovettori linearmente indipendenti. In tal caso gli elementi diagonali di D sono i corrispondenti autovalori e $D = P^{-1}AP$, dove P è la matrice le cui colonne sono gli autovettori.

Se una matrice A può essere diagonalizzata si dice che A ammette una **fattorizzazione diagonale**

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

Con tale fattorizzazione l'algebra di A viene quindi ad essere l'algebra della matrice diagonale D , calcolabile facilmente.

Le proprietà degli autovalori e degli autovettori

Elenchiamo nel teorema seguente le proprietà che occorreranno per trovare una rappresentazione diagonale di una matrice quadrata.

Sia A una matrice quadrata. Sono equivalenti le seguenti affermazioni:

1. lo scalare λ è un autovalore di A ;
2. la matrice $M = A - \lambda I$ è singolare ;
3. lo scalare λ è una radice del polinomio caratteristico $\Delta(t)$ di A .

L'autospazio E_λ è quindi lo spazio delle soluzioni del sistema omogeneo $MX = 0$, dove $M = A - \lambda I$, ottenuta cioè sottraendo λ dagli elementi sulla diagonale di A . Alcune matrici possono non avere autovalori (e quindi autovettori). Ricorrendo però al **Teorema fondamentale dell'algebra** (ogni polinomio su $\mathbb{C}$ ha almeno una radice) e al teorema prima esposto si può ottenere il risultato che segue.

Sia A una matrice quadrata sul campo complesso $\mathbb{C}$, allora A ha almeno un autovalore.

Consideriamo adesso la matrice A e supponiamo che λ sia un suo autovalore.

Per **molteplicità algebrica** di λ si intende la molteplicità di λ come polinomio caratteristico di A .

La **molteplicità geometrica** di λ è la dimensione del suo autospazio, in simboli : $\dim E_\lambda$.

Alla luce delle considerazioni finora fatte vale il seguente **teorema**.

Sia λ un autovalore di una matrice A : allora la molteplicità geometrica di λ **non eccede** la sua molteplicità algebrica.

Diagonalizzazione di operatori lineari

Cominciamo col dire che un operatore lineare T su uno spazio vettoriale V è detto **diagonalizzabile** se può essere rappresentato da una matrice diagonale D . Dunque T sarà diagonalizzabile **se e solo se** esiste una base di V

$$S = u_1, u_2, \dots, u_n$$

per la quale si ha:

$$\begin{array}{rcll} Tu_1 & = & k_1 u_1 & \\ Tu_2 & = & & k_2 u_2 \\ \dots & \dots & \dots & \dots \\ Tu_n & = & & k_n u_n \end{array}$$

In tal caso T è rappresentato dalla matrice diagonale:

$$D = \text{diag}(k_1, k_2, \dots, k_n)$$

relativa alla base S prima citata.

In maniera analoga a quanto fatto per le matrici procediamo col dare le seguenti definizioni.

Uno scalare $\lambda \in K$ si dice **autovalore** di T se esiste un vettore $v \in V$ non nullo per il quale

$$T(v) = \lambda v$$

Ogni vettore che soddisfa questa relazione è detto **autovettore** di T relativo all'autovalore λ .

L'insieme E_λ di tutti questi vettori è un sottospazio di V ed è detto **autospatio** di λ . Senza ripetere quanto già detto per le matrici, anche per gli operatori lineari valgono le medesime definizioni di **molteplicità algebrica** e **geometrica**.

Per completare enunciamo infine i seguenti **teoremi** che si applicano ad un operatore lineare T su uno spazio vettoriale di dimensione finita V .

T si può rappresentare con una matrice diagonale D se e solo se esiste una base S di V costituita da autovettori di T . In tal caso gli elementi diagonali di D sono i corrispondenti autovalori.

Sia T un operatore lineare. Sono equivalenti le seguenti affermazioni:

1. lo scalare λ è un autovalore di T ;
2. l'operatore lineare $\lambda I - T$ è singolare ;
3. lo scalare λ è una radice del polinomio caratteristico $\Delta(t)$ di T .

Dato uno spazio vettoriale complesso, allora T ha almeno un autovalore.

Supponiamo che $v_1, v_2, \dots, v_n$ siano gli autovettori diversi da zero di T e relativi a distinti autovalori $\lambda_1, \lambda_2, \dots, \lambda_n$, allora $v_1, v_2, \dots, v_n$ sono linearmente indipendenti.

Supponiamo che il polinomio caratteristico di T sia prodotto di n fattori distinti

$$\Delta(t) = (t - a_1)(t - a_2) \cdots (t - a_n)$$

Ebbene T può essere rappresentato dalla matrice diagonale

$$D = \text{diag}(a_1, a_2, \dots, a_3)$$

La molteplicità geometrica di un autovalore λ di T non supera la sua molteplicità algebrica.

Infine rilevante è anche il teorema che segue che **riduce la ricerca della diagonalizzazione di un operatore lineare T a quella di una matrice A** . Se A è la rappresentazione matriciale di T , T è diagonalizzabile se e solo se A è diagonalizzabile.

Calcolo di autovalori ed autovettori e diagonalizzazione delle matrici

L'algoritmo

Diamo ora un **algoritmo di diagonalizzazione** per una matrice quadrata di ordine n .

Passo 1. Trovare il polinomio caratteristico di A .

Passo 2. Trovare le radici del polinomio caratteristico di A , ossia trovarne gli autovalori.

Passo 3. Ripetere **1.** e **2.** per ogni autovalore λ di A .

1. Determinare $M = A - \lambda I$ sottraendo λ dalla diagonale di A .
2. Trovare una base per lo spazio delle soluzioni del sistema omogeneo $MX = 0$.

Passo 4. Consideriamo la collezione $S = v_1, v_2, \dots, v_n$ di tutti gli autovettori che abbiamo ottenuto nel Passo 3.

1. Se $m \neq n$ A non è diagonalizzabile.

2. Se $m = n$, A è diagonalizzabile. Detta P la matrice le cui colonne sono gli autovettori $v_1, v_2, \dots, v_n$, allora

$$D = P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

dove λ_i è l'autovalore corrispondente all'autovettore v_i .

Esempio

Applichiamo l'algoritmo di diagonalizzazione espresso ad una matrice A quadrata di ordine n :

$$A = \begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix} .$$

Calcoliamo il polinomio caratteristico:

$$\text{tr}(A) = 4 - 1 = 3, \quad |A| = -4 - 6 = -10$$

quindi

$$\Delta(t) = t^2 - 3t - 10 = (t - 5)(t + 2) .$$

Le radici λ_1 e λ_2 sono rispettivamente 5 e -2 e sono quindi gli autovalori di A .

Troviamo un autovettore v_1 di A relativo a $\lambda_1 = 5$, sottraendo quest'ultimo dalla diagonale di A ottenendo la matrice

$$M = \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix} .$$

Gli autovettori relativi a $\lambda_1 = 5$ formano la soluzione del sistema omogeneo $MX = 0$:

$$\begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ oppure } \begin{cases} -x + 2y = 0 \\ 3x - 6y = 0 \end{cases} .$$

Il sistema ha una sola variabile libera. Una soluzione diversa da zero, ad esempio $v_1 = (2, 1)$, è un autovettore che genera l'autospazio di $\lambda_1 = 5$.

Troviamo un autovettore v_2 di A relativo a $\lambda_2 = -2$, sottraendo quest'ultimo dalla diagonale di A ottenendo la matrice

$$M = \begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix} .$$

, ottenendo il sistema omogeneo:

$$\begin{cases} 6x + 2y = 0 \\ 3x + y = 0 \end{cases} .$$

Il sistema ha una sola variabile libera. Una soluzione diversa da zero, ad esempio $v_2 = (-1, 3)$, è un autovettore che genera l'autospazio di $\lambda_2 = -2$.

Sia P la matrice le cui colonne sono gli autovettori trovati. Allora:

$$P = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \text{ e quindi } P^{-1} = \begin{bmatrix} 3/7 & 1/7 \\ -1/7 & 2/7 \end{bmatrix} .$$

In definitiva $D = P^{-1}AP$ è la matrice diagonale i cui elementi sono i rispettivi autovalori:

$$D = P^{-1}AP = \begin{bmatrix} 3/7 & 1/7 \\ -1/7 & 2/7 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & -2 \end{bmatrix}$$

Diagonalizzazione di matrici reali simmetriche

Una generica matrice reale A può non essere diagonalizzabile. Se però A è una matrice simmetrica reale, valgono i seguenti **teoremi**.

Data una matrice A simmetrica reale, ogni radice λ del suo polinomio caratteristico è reale.

Data una matrice A simmetrica reale, detti u e v degli autovettori non di A , relativi ad autovalori distinti λ_1 e λ_2 , allora u e v sono ortogonali.

A partire da questi due teoremi si arriva al **teorema** che segue.

Data una matrice A simmetrica reale, esiste una matrice ortogonale P , tale che $D = P^{-1}AP$ è diagonale.

La matrice ortogonale P si ottiene normalizzando una base di autovettori ortogonali di A . A viene così detta **ortogonalmente diagonalizzabile**.

L'algoritmo

Esponiamo ora l'**algoritmo di diagonalizzazione ortonormale**:

Sia A una matrice reale simmetrica.

Passo 1. Trovare il polinomio caratteristico di A .

Passo 2. Trovare gli autovalori di A come radici del polinomio caratteristico.

Passo 3. Per ogni autovalore λ di A , trovare una base ortogonale del sottospazio relativo.

Passo 4. Normalizzare tutti gli autovalori ottenuti al Passo 3, formando quindi una base ortonormale di $\mathbb{R}^n$.

Passo 5. Sia P la matrice le cui colonne sono gli autovettori normalizzati del Passo 4.

Esempio

Data la matrice simmetrica

$$A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix},$$

troviamo una matrice ortogonale P tale che $P^{-1}AP$ è diagonale.

Procediamo come segue.

Troviamo il polinomio caratteristico di A :

$$\text{tr}(A) = 2 + 5 = 7, \quad |A| = 10 - 4 = 6, \quad \Delta(t) = t^2 - 7t + 6 = (t - 6)(t - 1)$$

Quindi $\lambda_1 = 6$ e $\lambda_2 = 1$ sono gli autovalori di A .

Sottraendo $\lambda_1 = 6$ dalla diagonale di A si ha la matrice $M = \begin{bmatrix} -4 & -2 \\ -2 & -1 \end{bmatrix}$ ed il sistema

$$\text{omogeneo} \begin{cases} -4x - 2y = 0 \\ -2x - y = 0 \end{cases} \text{ oppure } 2x + y = 0$$

Una soluzione diversa da zero è $u_1 = (1, -2)$.

Sottraendo $\lambda_2 = 1$ dalla diagonale di A si ha la matrice

$$M = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \text{ ed il sistema omogeneo } x - 2y = 0$$

La seconda equazione è eliminata in quanto multiplo della prima. Una soluzione diversa da zero è $u_2 = (2, 1)$.

u_1 e u_2 sono ortogonali. Li normalizziamo e otteniamo i vettori ortonormali:

$$\widehat{u}_1 = (1/\sqrt{5}, -2/\sqrt{5}) \text{ e } \widehat{u}_2 = (2/\sqrt{5}, 1/\sqrt{5})$$

Sia infine la matrice P di rispettive colonne u_1 e u_2 , allora:

$$P = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \text{ e } P^{-1}AP = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$$

Gli elementi diagonali di $P^{-1}AP$ sono gli autovalori corrispondenti alle colonne di P .

Bibliografia

Algebra lineare - Lipschutz, Lipson.

Estratto da "<http://www.electroyou.it/mediawiki/index.php?title=UsersPages:Asdf:la-diagonalizzazione>"