

**Note Aggiuntive alle Dispense di
ALGEBRA LINEARE,
c.d.l. in Informatica, Università di Firenze,
Giorgio PATRIZIO - A.A. 2011/12**

Contents

1	Prodotto di Matrici e Grafi Orientati.	2
2	Prodotto di matrici quadrate, matrice inversa.	4
3	Il Determinante.	6
4	Determinante e prodotto di Matrici.	12
5	Basi e coordinate. Cambi di base.	15
6	Applicazioni lineari e matrici.	17
7	Autovalori e Autovettori. Diagonalizzazione	19

1 Prodotto di Matrici e Grafi Orientati.

Un *grafo orientato (semplice)* $\mathcal{G} = (V, E)$ è una coppia costituita da un insieme finito V di punti, che si dicono *vertici* di \mathcal{G} , e da un insieme di linee orientate E , che si dicono *lati* di \mathcal{G} , che collegano coppie ordinate di punti (non necessariamente distinti) di V . Per ogni coppia ordinata di punti si intende che vi può essere al più un lato ossia E è in corrispondenza biunivoica con un sottoinsieme del prodotto cartesiano $V \times V$ dell'insieme dei vertici. Un esempio di grafo:

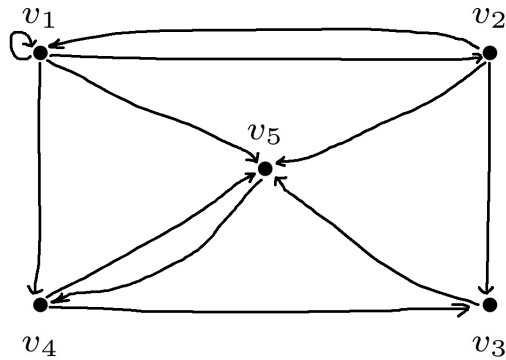


Figura 1

I grafi modellizzano molte situazioni concrete. Per esempio i vertici di un grafo possono rappresentare città e i lati linee aeree che le collegano. È possibile rappresentare un grafo mediante la sua *matrice di adiacenza* che si definisce nel modo seguente. Se $\mathcal{G} = (V, E)$ è un grafo e $V = \{v_1, \dots, v_n\}$, allora la matrice di adiacenza di \mathcal{G} è la matrice $A = (a_{ij}) \in M_{n,n}$ quadrata di ordine n definita da

$$a_{ij} = \begin{cases} 1 & \text{se esiste un lato da } v_i \text{ a } v_j \\ 0 & \text{se non esiste un lato da } v_i \text{ a } v_j. \end{cases}$$

Dunque il grafo in Figura 1 ha matrice d'adiacenza

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

Dati due vertici v_i, v_j di un grafo \mathcal{G} un *cammino di lunghezza r* da v_i, v_j è una successione di r lati di \mathcal{G} che collega v_i a v_j .

Problema: Trovare un algoritmo per calcolare il numero di cammini di lunghezza r fra due vertici di un grafo. La soluzione è data dal seguente

Teorema 1 : Sia A la matrice di adiacenza di un grafo \mathcal{G} . Se $(a_{ij}^{(r)}) = A^r = \overbrace{A \cdots A}^{r \text{ volte}}$ è la potenza r -esima della matrice A , il numero di cammini di lunghezza r da v_i a v_j è $a_{ij}^{(r)}$.

Dimostrazione. Procediamo per induzione. Se $r = 1$, l'affermazione è vera per definizione di matrice di adiacenza. Supponiamo che sia vera per cammini di lunghezza r e dimostriamola per cammini di lunghezza $r + 1$. Per ogni $i, j = 1, \dots, n$, sia c_{ij} il numero di cammini di lunghezza $r + 1$ da v_i a v_j . Per dimostrare la tesi, occorre provare che $c_{ij} = a_{ij}^{(r+1)}$. Per ogni $k = 1, \dots, n$, se esiste, un cammino di lunghezza $r + 1$ da v_i a v_j che ha per secondo vertice v_k , si decompone in un cammino di lunghezza 1 da v_i a v_k e in uno di lunghezza r da v_k a v_j . Possiamo concludere che il numero di cammini di lunghezza $r + 1$ da v_i a v_j che hanno per secondo vertice v_k è $a_{ik}a_{kj}^{(r)}$. Dunque per ottenere il numero di tutti i cammini di lunghezza $r + 1$ da v_i a v_j basta sommare per $k = 1, \dots, n$:

$$c_{ij} = \sum_{k=1}^n a_{ik}a_{kj}^{(r)} = a_{ij}^{(r+1)}$$

e quindi la tesi è dimostrata.

Q.e.d.

Ad esempio, come conseguenza del Teorema 1, per il grafo in Figura 1, le matrici

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad A^2 = \begin{pmatrix} 2 & 1 & 2 & 2 & 3 \\ 1 & 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \quad A^3 = \begin{pmatrix} 3 & 2 & 3 & 5 & 7 \\ 2 & 1 & 3 & 3 & 4 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

forniscono rispettivamente il numero dei cammini di lunghezza 1, 2, 3, 4 fra tutte le coppie di vertici.

2 Prodotto di matrici quadrate, matrice inversa.

Denotiamo con M_n lo spazio delle matrici quadrate di ordine n . Allora evidentemente se $A, B \in M_n$, allora il prodotto AB è sempre definito ed è una matrice quadrata di ordine n . Dunque il prodotto di matrici ristretta a matrici quadrate dello stesso ordine, definisce una vera e propria operazione. In particolare la matrice

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

svolge il ruolo di elemento neutro per questa operazione nel senso che per ogni matrice $A \in M_n$ si ha: $AI_n = I_nA = A$. Dunque la matrice I_n si comporta come il numero 1 per il prodotto di numeri. Per ogni numero $a \neq 0$ esiste un unico inverso a^{-1} tale che $aa^{-1} = a^{-1}a = 1$. Per il prodotto di matrici la nozione di invertibilità si esprime nel modo seguente:

Definizione. Una matrice quadrata $A \in M_n$ si dice *invertibile* se esiste una matrice $B \in M_n$ tale che

$$AB = BA = I_n. \quad (*)$$

Se $A \in M_n$ è invertibile, la matrice $B \in M_n$ tale che valga è unica, si dice *inversa* di A e si denota A^{-1} . L'unicità si dimostra facilmente. Se ci fossero due matrici $B, B' \in M_n$ tali che $AB = BA = I_n = AB' = B'A$, allora si avrebbe:

$$B' = B'I_n = B'(AB) = (B'A)B = I_nB = B.$$

È chiaro che se A è una matrice invertibile, allora $A \neq 0$, ma non è chiaro dalla definizione come si possa determinare se una matrice è invertibile. Un primo passo utile è dato dalla seguente:

Proposizione 2 : Una matrice $A \in M_n$ è invertibile se e solo se esiste una unica inversa destra di A ossia una unica matrice $B \in M_n$ tale che $AB = I_n$.

Dimostrazione. Se A è invertibile allora, come abbiamo visto esiste una unica inversa. Supponiamo che esista un'unica matrice $B \in M_n$ tale che $AB = I_n$. Allora dato che $AB - I_n = 0$, si ha

$$0 = (AB - I_n)A = ABA - A = A(BA - I_n)$$

e quindi

$$I_n = AB = AB + A(BA - I_n) = A(B + BA - I_n)$$

ossia anche $B + BA - I_n$ è un'inversa destra di A . Pertanto, per l'unicità dell'inversa destra, deve essere $B = B + BA - I_n$ e quindi $0 = BA - I_n$ ossia $BA - I_n = 0$ e quindi la tesi. *Q.e.d.*

Definizione. Una matrice quadrata $A \in M_n$ si dice *non singolare* se ammette una riduzione a scala con tutti elementi diversi da zero sulla diagonale.

Proposizione 3 : Sia $A \in M_n$ una matrice quadrata. Le seguenti affermazioni sono equivalenti:

(i) A è non singolare;

(ii) per ogni colonna $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, il sistema $ax = b$ ha soluzione unica;

(iii) ogni riduzione a scala di A ha gli elementi sulla diagonale diversi da zero;

Dimostrazione. Supponiamo che valga (i) e sia S una riduzione a scala di A on tutti elementi diversi da zero sulla diagonale. Allora per ogni colonna di termini noti b , il sistema $Ax = b$ è equivalente, per una opportuna colonna b' al sistema $Sx = b'$ che ha soluzione unica e quindi vale (ii). Supponiamo ora che valga (ii). In particolare $x = 0$ è l'unica soluzione del sistema omogeneo $Ax = 0$. Se S è una qualunque riduzione a scala di A , il sistema $Sx = 0$ ha soluzione unica $x = 0$. Supponiamo che sulla diagonale di A ci sia qualche zero. Allora, dato che $S = (s_{ij})$ è a scala, per qualche $k < n$ si deve avere

$$s_{11} \neq 0, \dots, s_{k-1,k-1} \neq 0, s_{kk} = 0, \dots, s_{nn} = 0.$$

Dunque, scrivendo per esteso il sistema $Sx = 0$, abbiamo:

$$\left\{ \begin{array}{lcl} s_{11}x_1 + \dots & & = 0 \\ & s_{22}x_2 + \dots & = 0 \\ & & \vdots \\ & s_{k-1,k-1}x_{k-1} + \dots & = 0 \\ & 0x_k + \dots & = 0 \\ & & \vdots \\ & & 0x_n = 0 \end{array} \right.$$

e quindi per ogni scelta $t_k, \dots, t_n \in \mathbf{R}$ la n -upla $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ t_k \\ \vdots \\ t_n \end{pmatrix}$ è una soluzione del sistema $Sx = 0$ che

avrebbe infinite soluzioni: contraddizione! Infine se ogni riduzione a scala di A ha elementi non nulli sulla diagonale, allora A è non singolare ossia (iii) implica (i). Q.e.d.
Siamo pronti a dimostrare che le matrici invertibili sono esattamente quelle non singolari:

Proposizione 4 Sia $A \in M_n$ una matrice quadrata. Allora A è invertibile se e solo se è non singolare.

Dimostrazione. Supponiamo che A sia invertibile e sia A^{-1} la sua inversa. allora per ogni colonna b di termini noti, il sistema $Ax = b$ ha soluzione unica $x = A^{-1}b$ e quindi A è non singolare. Sia A è non singolare. Sia I_n la matrice identità e siano

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

le sue colonne. Per $j = 1, \dots, n$, sia B^j l'unica soluzione del sistema $AB^j = e_j$. Allora la matrice $B = (B^1, \dots, B^n)$ è l'unica inversa destra di A ossia l'unica matrice tale che $AB = I_n$ e quindi A è invertibile. Q.e.d.

La Proposizione 4 indica anche come procedere per stabilire se una matrice A è invertibile e nel caso trovare l'inversa: I passo: Mediante un'eliminazione di Gauss ridurre a scala A e verificare se A è non singolare II passo: Se A è non singolare trovare le soluzioni B^1, \dots, B^n dei sistemi $AB^1 = e_1, \dots, AB^n = e_n$. L'inversa di A è la matrice $A^{-1} = (B^1, \dots, B^n)$. Questa idea si può trasformare in un algoritmo che utilizzando l'eliminazione di Gauss calcola l'inversa delle matrici invertibili.

3 Il Determinante.

Lo scopo della teoria dei determinanti è di definire una funzione, chiamata *determinante*, sullo spazio delle matrici quadrate di ordine n , per ogni n , calcolabile facendo somme e prodotti delle entrate delle matrici. In particolare si desidera che il determinante di una matrice sia non nullo se e solo se la matrice è invertibile. Per $n = 1$, ossia per matrici di ordine 1, lo scopo si ottiene semplicemente assegnando alla matrice (a) il valore $\det(a) = a$ per ogni scalare a . Si osservi che in questo caso alla matrice unità $I_1 = (1)$ si associa il numero $\det(I_1) = 1$. Nel caso $n = 2$, ossia per matrici di ordine 2, osserviamo che richiedere che una matrice sia invertibile è equivalente a richiedere che sue righe non siano proporzionali. Dunque

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

è non singolare se e solo se le righe (a_{11}, a_{12}) e (a_{21}, a_{22}) non sono proporzionali e questo succede se e solo se $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Tenendo conto di quanto detto sopra, sembra ragionevole associare alla matrice A il numero $\det(A) = a_{11}a_{22} - a_{12}a_{21}$. Si osservi che anche in questo caso questa funzione associa alla matrice unità I_2 il numero 1. Il determinante delle matrici quadrate di ordine 2 definito in questo modo ha una semplice interpretazione geometrica. Su \mathbf{R}^2 si considerino due vettori v, w . I vettori v, w individuano un parallelogramma che denotiamo $P(v, w)$. Se in termini della base canonica $\{e_1, e_2\}$ si ha $v = v_1e_1 + v_2e_2$ e $w = w_1e_1 + w_2e_2$ e

$$A = \begin{pmatrix} v_1 & v_2 \\ w_1 & w_2 \end{pmatrix},$$

allora, con una semplice verifica geometrica, si può dimostrare che

$$\det(A) = \begin{cases} \text{Area}(P(v, w)) & \text{se si passa da } v \text{ a } w \text{ ruotando in senso antiorario} \\ -\text{Area}(P(v, w)) & \text{se si passa da } v \text{ a } w \text{ ruotando in senso orario.} \end{cases}$$

In questo modo si ritrova geometricamente che se le righe della matrice di A sono proporzionali allora il determinante si annulla dato che in questo caso il parallelogramma $P(v, w)$ degenera a un segmento o a un punto. Con queste motivazioni e con l'esempio delle matrici di ordine 2 in mente, diamo ora una definizione formale. Sia $\mathbf{K} = \mathbf{R}$ o \mathbf{C} e si denoti $M_n = M_{n,n}(\mathbf{K})$ lo spazio delle matrici quadrate di ordine n con elementi in \mathbf{K} . Se $A \in M_n$ allora si può considerare A come collezione di n righe $A = (A_1, \dots, A_n)$ oppure come collezione di n colonne $A = (A^1, \dots, A^n)$

Definizione. Per ogni $n \geq 1$ un *determinante* è una funzione $d_n: M_n \rightarrow \mathbf{K}$ che ha le seguenti proprietà (A) d_n si annulla sulle matrici che hanno due righe uguali, ossia

$$d_n(A_1, \dots, A_n) = 0 \text{ se } A_i = A_j \text{ per } i \neq j.$$

(B) d_n è lineare in ciascuna riga, ossia fissato $j \in \{1, \dots, n\}$ si deve avere

$$d_n(A_1, \dots, A_j + A'_j, \dots, A_n) = d_n(A_1, \dots, A_j, \dots, A_n) + d_n(A_1, \dots, A'_j, \dots, A_n)$$

$$d_n(A_1, \dots, \lambda A_j, \dots, A_n) = \lambda d_n(A_1, \dots, A_j, \dots, A_n).$$

(C) $d_n(I_n) = 1$.

È facile controllare che il determinante che abbiamo definito per le matrici di ordine 2 verifica le condizioni (A), (B), (C) della definizione. Prima di enunciare il risultato di esistenza per matrici di ordine arbitrario, dobbiamo introdurre una notazione.

Definizione. Se $A \in M_n$, si dice *minore* (i, j) di A la matrice $A_{ij} \in M_{n-1}$ ottenuta da A togliendo la i -esima riga e la j -esima colonna.

Teorema 5 Per ogni $n \geq 1$ esiste un'unica funzione determinante $d_n: M_n \rightarrow \mathbf{K}$ che verifica le proprietà (A), (B), (C) e valgono le seguenti formula per ricorrenza per ogni matrice $A = (a_{ij}) \in M_n$:

(i) **(sviluppo di Laplace del determinante lungo la colonna j_0 -esima)** per ogni $j_0 = 1, \dots, n$ fissato, si ha:

$$d_n(A) = \sum_{i=1}^n (-1)^{i+j_0} a_{ij_0} d_{n-1}(A_{ij_0}).$$

(ii) **(sviluppo di Laplace del determinante lungo la riga i_0 -esima)** per ogni $i_0 = 1, \dots, n$ fissato, si ha:

$$d_n(A) = \sum_{j=1}^n (-1)^{i_0+j} a_{i_0j} d_{n-1}(A_{i_0j}).$$

Non è difficile provare che il risultato è vero per $n = 2$. Il caso generale segue argomentando per induzione (ometteremo la dimostrazione!). Si osservi che il determinante di matrici di qualunque ordine ha un significato geometrico che generalizza quello per matrici di ordine 2. In particolare si può dimostrare che per vettori u, v, w di \mathbf{R}^3 , se in termini della base canonica $\{e_1, e_2, e_3\}$ si ha $u = u_1e_1 + u_2e_2 + u_3e_3$, $v = v_1e_1 + v_2e_2 + v_3e_3$ e $w = w_1e_1 + w_2e_2 + w_3e_3$ e

$$A = \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix},$$

allora $|\det(A)|$ è il volume del parallelogramma definito dai vettori u, v, w

Qui ci limitiamo a fare qualche osservazione sul calcolo del determinante e sulle sue proprietà. Infatti le formule che abbiamo dato per il determinante diventano sempre più complicate al crescere di n e, d'altra parte, non abbiamo ancora verificato che la funzione determinante costruita in questo modo è in grado di "riconoscere" le matrici invertibili. Da questo momento in poi denoteremo funzione determinante con il simbolo \det e per $A \in M_n$ scriveremo $\det(A) = d_n(A)$. Abbiamo la seguente

Proposizione 6 Sia $A = (A_1, \dots, A_n) \in M_n$ una matrice quadrata di ordine n . Allora (i) Se A ha una riga nulla allora $\det(A) = 0$. (ii) Il valore del determinante non cambia se si somma a una riga un multiplo di un'altra, ossia per ogni $\lambda \in \mathbf{K}$ e per indici $i \neq j$ si ha

$$\det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) = \det(A_1, \dots, A_i + \lambda A_j, \dots, A_j, \dots, A_n).$$

(iii) Il valore del determinante cambia segno se si scambiano due righe:

$$\det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) = -\det(A_1, \dots, A_j, \dots, A_i, \dots, A_n).$$

(iv) Se S è una matrice triangolare superiore ottenuta da A mediante una riduzione di Gauss

che comporta σ scambi di righe, allora

$$\det(A) = (-1)^\sigma \det(S).$$

(v) Se A ha le righe linearmente dipendenti allora $\det(A) = 0$. (vi) Se S è una matrice triangolare superiore e siano $p_1, \dots, p_n \in \mathbf{K}$ gli elementi sulla diagonale di S , allora

$$\det(S) = p_1 \dots p_n.$$

(vii) A è invertibile se e solo se $\det(A) \neq 0$.

Dimostrazione. (i) Si ha $\det(A) = \det(A_1, \dots, A_{j-1}, 0, A_{j+1}, \dots, A_n) = 0$ dato che \det è lineare in ciascuna riga a valori in \mathbf{K} e quindi vale 0 sul vettore nullo.

(ii) Ancora utilizzando la linearità su ciascuna riga e la proprietà (A) dei determinanti, se $\lambda \in \mathbf{K}$, si ha

$$\begin{aligned} & \det(A_1, \dots, A_i + \lambda A_j, \dots, A_j, \dots, A_n) \\ &= \det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) + \lambda \det(A_1, \dots, A_j, \dots, A_j, \dots, A_n) \\ &= \det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) = \det(A). \end{aligned}$$

(iii) Si ha

$$\begin{aligned} 0 &= \det(A_1, \dots, A_i + A_j, \dots, A_i + A_j, \dots, A_n) \\ &= \det(A_1, \dots, A_i, \dots, A_i + A_j, \dots, A_n) + \det(A_1, \dots, A_j, \dots, A_i + A_j, \dots, A_n) \\ &= \det(A_1, \dots, A_i, \dots, A_i, \dots, A_n) + \det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) \\ &\quad + \det(A_1, \dots, A_j, \dots, A_i, \dots, A_n) + \det(A_1, \dots, A_j, \dots, A_j, \dots, A_n) \\ &= \det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) + \det(A_1, \dots, A_j, \dots, A_i, \dots, A_n) \end{aligned}$$

e quindi (iii) segue.

(iv) È immediata da (ii) e (iii).

(v) Se A ha le righe linearmente dipendenti, allora una sua qualunque riduzione a scala ha una riga nulla e quindi (v) segue da (i) e (iv).

(vi) Usando la linearità del determinante nella prima riga si ha

$$\begin{aligned} & \det \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix} \\ &= \det \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix} + \det \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix}. \end{aligned}$$

Dato che $\begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix}$ ha le righe linearmente dipendenti, usando (v) e, ripetendo l'argomento, si ha

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix} &= a_{11} \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix} \\ &= a_{11} \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix} + a_{11} \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix} \\ &= a_{11} a_{22} \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix} = \dots = a_{11} a_{22} \dots a_{nn} \det(I_n) = a_{11} a_{22} \dots a_{nn}. \end{aligned}$$

(vii) Una matrice A è invertibile se e solo se una sua qualunque riduzione a scala ha tutti i pivot non nulli. Per (vi) questo succede se e solo se è non nullo il determinante di una qualunque riduzione a scala di A , che per (iv) ha valore assoluto uguale al valore assoluto del determinante di A .

Q.e.d. Il punto (vi) della Proposizione 6 fornisce un metodo di calcolo molto efficace per il determinante: è sufficiente ridurre a scala la matrice mediante una riduzione di Gauss tenendo il conto degli scambi di righe! Infatti se S è un'arbitraria riduzione a scala di A ottenuta mediante σ scambi di righe e $p_1, \dots, p_n \in \mathbf{K}$ sono gli elementi sulla diagonale di S , allora, (vi) e (vii) della Proposizione 6 implicano che

$$\det(A) = (-1)^\sigma p_1 \dots p_n.$$

Come conseguenza (quasi) immediata del Teorema 5 per induzione si dimostra (noi non lo faremo!):

Teorema 7 : Sia $A \in M_n$ una matrice quadrata di ordine n . Allora $\det(A) = \det(A^T)$.

Il Teorema 7 implica che gli enunciati riguardanti il determinante espressi in termini di righe si possono riformulare in termini di colonne. In particolare il determinante è una funzione lineare in ciascuna colonna e vale la seguente:

Proposizione 8 : Sia $A = (A^1, \dots, A^n) \in M_n$ una matrice quadrata di ordine n . (i) Se A ha una colonna nulla allora $\det(A) = 0$.

(ii) Il valore del determinante non cambia se si somma a una colonna un multiplo di un'altra, ossia per ogni $\lambda \in \mathbf{K}$ e per indici $i \neq j$ si ha

$$\det(A^1, \dots, A^i, \dots, A^j, \dots, A^n) = \det(A^1, \dots, A^i + \lambda A^j, \dots, A^j, \dots, A^n).$$

(iii) Il valore del determinante cambia segno se si scambiano due colonne:

$$\det(A^1, \dots, A^i, \dots, A^j, \dots, A^i, \dots, A^j, \dots, A^n) = -\det(A^1, \dots, A^j, \dots, A^i, \dots, A^i, \dots, A^j, \dots, A^n).$$

Concludiamo queste considerazioni calcolando in due modi diversi il determinante di una matrice.

Esempio. Si consideri la matrice

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ -1 & 0 & 1 & 2 & 1 \\ 3 & 1 & 1 & -3 & 1 \\ 2 & 3 & 3 & 1 & -2 \\ 1 & 1 & 1 & 2 & 3 \end{pmatrix}$$

Calcoliamo il determinante di A riducendola a scala mediante un'eliminazione di Gauss:

$$\begin{aligned} & \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ -1 & 0 & 1 & 2 & 1 \\ 3 & 1 & 1 & -3 & 1 \\ 2 & 3 & 3 & 1 & -2 \\ 1 & 1 & 1 & 2 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 2 & 3 & 3 & 2 \\ 0 & -5 & -5 & -6 & -2 \\ 0 & -1 & -1 & -1 & -4 \\ 0 & -1 & -1 & 1 & 2 \end{pmatrix} \\ & \xrightarrow{\text{scambio di righe}} \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & -1 & -1 & 1 & 2 \\ 0 & -5 & -5 & -6 & -2 \\ 0 & -1 & -1 & -1 & -4 \\ 0 & 2 & 3 & 3 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & -1 & -1 & 1 & 2 \\ 0 & 0 & 0 & -11 & -12 \\ 0 & 0 & 0 & -2 & -6 \\ 0 & 0 & 1 & 5 & 6 \end{pmatrix} \\ & \xrightarrow{\text{scambio di righe}} \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & -1 & -1 & 1 & 2 \\ 0 & 0 & 1 & 5 & 6 \\ 0 & 0 & 0 & -2 & -6 \\ 0 & 0 & 0 & -11 & -12 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & -1 & -1 & 1 & 2 \\ 0 & 0 & 1 & 5 & 6 \\ 0 & 0 & 0 & -2 & -6 \\ 0 & 0 & 0 & 0 & 21 \end{pmatrix}. \end{aligned}$$

Dunque, utilizzando Proposizione 6 e osservando che nella riduzione a scala abbiamo operato 2 scambi di righe, possiamo concludere che

$$\det(A) = (-1)^2(1 \cdot -1 \cdot 1 \cdot -2 \cdot 21) = 42.$$

Calcoliamo ora il determinante di A combinando le altre tecniche che abbiamo introdotto:

$$\det(A) = \det \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ -1 & 0 & 1 & 2 & 1 \\ 3 & 1 & 1 & -3 & 1 \\ 2 & 3 & 3 & 1 & -2 \\ 1 & 1 & 1 & 2 & 3 \end{pmatrix} \stackrel{\text{sommando alle altre multipli prima riga}}{=} \det \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 2 & 3 & 3 & 2 \\ 0 & -5 & -5 & -6 & -2 \\ 0 & -1 & -1 & -1 & -4 \\ 0 & -1 & -1 & 1 & 2 \end{pmatrix}$$

$$\stackrel{=}{\text{sviluppo di Laplace}} \underset{\text{prima colonna}}{(-1)^{1+1} \cdot 1 \cdot \det \begin{pmatrix} 2 & 3 & 3 & 2 \\ -5 & -5 & -6 & -2 \\ -1 & -1 & -1 & -4 \\ -1 & -1 & 1 & 2 \end{pmatrix}} \stackrel{=}{\text{sottraendo prima}} \det \begin{pmatrix} 2 & 1 & 3 & 2 \\ -5 & 0 & -6 & -2 \\ -1 & 0 & -1 & -4 \\ -1 & 0 & 1 & 2 \end{pmatrix}$$

$$\stackrel{=}{\text{sviluppo di Laplace}} \underset{\text{seconda colonna}}{(-1)^{2+1} \cdot 1 \cdot \det \begin{pmatrix} -5 & -6 & -2 \\ -1 & -1 & -4 \\ -1 & 1 & 2 \end{pmatrix}} \stackrel{=}{\text{sommando terza}} -\det \begin{pmatrix} -5 & -6 & -2 \\ -2 & 0 & -2 \\ -1 & 1 & 2 \end{pmatrix}$$

$$\stackrel{=}{\text{sottraendo prima}} -\det \begin{pmatrix} -5 & -6 & 3 \\ -2 & 0 & 0 \\ -1 & 1 & 3 \end{pmatrix} \underset{\text{seconda riga}}{\text{sviluppo di Laplace}} -(-1)^{1+2}(-2) \det \begin{pmatrix} -6 & 3 \\ 1 & 3 \end{pmatrix} = (-2)(-18-3) = 42.$$

4 Determinante e prodotto di Matrici.

Il determinante rispetta la struttura moltiplicativa delle matrici. Vale infatti il seguente importante:

Teorema 9 (di Binet) *Siano $A, B \in M_n$. Allora $\det(AB) = \det(A) \det(B)$.*

Dimostrazione. Se $\det(B) = 0$ la tesi è immediata. In questo caso infatti la matrice B non è invertibile e quindi esiste un vettore $v \neq 0_n$ in \mathbf{K}^n tale che $Bv = 0_n$. Dunque si ha anche $ABv = 0_n$ e quindi la matrice AB non è invertibile e $\det(AB) = 0 = \det(A) \det(B)$. Supponiamo allora che $\det(B) \neq 0$. Si definisca la funzione $D: M_n \rightarrow \mathbf{K}$ mediante

$$D(A) = \frac{\det(AB)}{\det(B)}.$$

La tesi seguirà anche in questo caso provando che $D(A) = \det(A)$. A tal fine basterà dimostrare che la funzione D verifica le condizioni (A), (B), (C) della definizione di determinante. Dato che $D(I_n) = \frac{\det(I_n B)}{\det(B)} = \frac{\det(B)}{\det(B)} = 1$ e quindi (C) è immediata. Se la matrice A ha due righe uguali, ossia $A = (A_1, \dots, A_i, \dots, A_j, \dots, A_n)$ con $A_i = A_j$ per $i \neq j$, allora, dato che

$$AB = (A_1 B, \dots, A_i B, \dots, A_j B, \dots, A_n B),$$

anche AB ha due righe uguali e quindi $D(A) = 0$. Infine, dato che

$$\begin{aligned} \det[(A_1, \dots, (A_i + A'_i), \dots, A_n)B] &= \det[(A_1 B, \dots, (A_i + A'_i)B, \dots, A_n B)] \\ &= \det[(A_1 B, \dots, A_i B, \dots, A_n B)] + \det[(A_1 B, \dots, A'_i B, \dots, A_n B)] \\ &= \det[(A_1, \dots, A_i, \dots, A_n)B] + \det[(A_1, \dots, A'_i, \dots, A_n)B] \end{aligned}$$

e

$$\begin{aligned} \det[(A_1, \dots, (\lambda A_i), \dots, A_n)B] &= \det[(A_1 B, \dots, (\lambda A_i)B, \dots, A_n B)] \\ &= \lambda \det[(A_1 B, \dots, A_i B, \dots, A_n B)] = \lambda \det[(A_1, \dots, A_i, \dots, A_n)B], \end{aligned}$$

si ha anche che D è lineare sulla riga i -esima per ogni $i = 1, \dots, n$. *Q.e.d.*
Un'osservazione utile per le applicazioni è il seguente immediato corollario del Teorema di Binet:

Corollario 10 *Sia $A \in M_n$ e $B \in GL_n$ una matrice invertibile. Allora $\det(B^{-1}) = \frac{1}{\det(B)}$ e quindi $\det(B^{-1}AB) = \det(A)$.*

Il Teorema di Binet si applica anche alla teoria dei sistemi lineari. Tradizionalmente la teoria dei sistemi lineari veniva basata sui determinanti e lo strumento centrale era costituito dal seguente risultato che si può ricavare come conseguenza del Teorema di Binet:

Teorema 11 (di Cramer) *Siano $A = (A^1, \dots, A^n) \in GL_n$ una matrice invertibile, $b \in \mathbf{K}^n$ e $B(i) \in M_n$ le matrici ottenute sostituendo la i -esima colonna di A con b : $B(i) = (A^1, \dots, A^{i-1}, b, A^{i+1}, \dots, A^n)$. Allora l'unica soluzione $x = (x_1, \dots, x_n)$ del sistema $Ax = b$ è data da:*

$$x_i = \frac{\det(B(i))}{\det(A)} \quad \forall i = 1, \dots, n.$$

Dimostrazione. Posto per ogni $i = 1, \dots, n$

$$X(i) = (e_1, \dots, e_{i-1}, x, e_{i+1}, \dots, e_n),$$

la colonna $x = (x_1, \dots, x_n)$ la soluzione del sistema $Ax = b$ se e solo se $AX(i) = B(i)$ per ogni $i = 1, \dots, n$. Dato che $\det(X(i)) = x_i$, dal teorema di Binet ricaviamo immediatamente:

$$\det(A)x_i = \det(A)\det(X(i)) = \det(B(i))$$

ossia la tesi. *Q.e.d.*

Esempio. Si consideri il sistema $\begin{cases} 1004005x + 1002001y = 1 \\ 1002001x + 1000001y = 0 \end{cases}$

Se A è la matrice dei coefficienti del sistema si ha

$$\det(A) = \det \begin{pmatrix} 1004005 & 1002001 \\ 1002001 & 1000001 \end{pmatrix} = \det \begin{pmatrix} 2004 & 2000 \\ 1002001 & 1000001 \end{pmatrix} = \det \begin{pmatrix} 4 & 2000 \\ 2000 & 1000001 \end{pmatrix} = 4.$$

Usando Cramer, la soluzione $\begin{pmatrix} x \\ y \end{pmatrix}$ è data da $x = \frac{\det \begin{pmatrix} 1 & 1002001 \\ 0 & 1000001 \end{pmatrix}}{\det(A)} = \frac{1000001}{4}$ $y =$
 $\frac{\det \begin{pmatrix} 1004005 & 1 \\ 1002001 & 0 \end{pmatrix}}{\det(A)} = -\frac{1002001}{4}.$

Naturalmente si può risolvere il sistema utilizzando il metodo della riduzione a scala magari aiutandosi con una calcolatrice. Provare e vedere cosa succede.....e cercare di spiegare il perché.

Il Teorema di Cramer può essere utilizzato per dimostrare una formula per la matrice inversa. Anche in questo caso la formula, molto elegante e utile in molte considerazioni teoriche, non è molto efficiente dal punto di vista del calcolo pratico. Ricordiamo che se $A \in M_n$, il *minore* (i, j) di A è la matrice $A_{ij} \in M_{n-1}$ ottenuta da A togliendo la i -esima riga e la j -esima colonna. Vale allora la seguente:

Proposizione 12 : Sia $A = (a_{ij}) \in GL_n$. Se $A^{-1} = (x_{ij})$ allora vale la seguente formula:

$$x_{ij} = \frac{1}{\det(A)} (-1)^{i+j} \det(A_{ji}).$$

Dimostrazione. Come già osservato la colonna j -esima della matrice A^{-1} è la soluzione X^j del sistema $AX^j = e_j$ dove, come al solito, e_j denota la j -esimo vettore della base canonica di \mathbf{K}^n .

Dunque, se $X^j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{pmatrix}$, secondo la formula dati Cramer, si ha:

$$x_{ij} = \frac{\det(B(i))}{\det(A)}$$

dove $B(i) = (A^1, \dots, A^{i-1}, e_j, A^{i+1}, \dots, A^n)$. Sviluppando il determinante, si ha

$$\det(B(i)) = \det \begin{pmatrix} a_{11} & \dots & a_{1i-1} & 0 & a_{1i+1} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j-11} & \dots & a_{j-1i-1} & 0 & a_{j-1i+1} & \dots & a_{j-1n} \\ a_{j1} & \dots & a_{ji-1} & 1 & a_{ji+1} & \dots & a_{jn} \\ a_{j+11} & \dots & a_{j+1i-1} & 0 & a_{j+1i+1} & \dots & a_{j+1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{ni-1} & 0 & a_{ni+1} & \dots & a_{nn} \end{pmatrix} = (-1)^{i+j} \det(A_{ji})$$

e quindi la tesi.

Q.e.d.

Al crescere di n , la formula proposta nella Proposizione 12 diventa rapidamente troppo complicata e quindi per calcolare l'inversa è molto più semplice utilizzare l'eliminazione di Gauss. D'altra parte per matrici di ordine 2 la formula permette di scrivere l'inversa immediatamente.

Se $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ è invertibile, ossia se $\det(A) = ad - bc \neq 0$, allora

$$A^{-1} = \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right)^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

In molti casi risulta utile utilizzare i determinanti per calcolare il rango di una matrice. Prima un po' di nomenclatura. Sia $A = (a_{ij})$ una matrice con m righe e n colonne. Diremo che $B = (b_{kl})$ è una *sottomatrice* di ordine r di A , se esistono indici $1 \leq i_1 < \dots < i_r \leq m$ e $1 \leq j_1 < \dots < j_r \leq n$, tali che

$$b_{kl} = a_{i_k j_l} \quad \forall k, l = 1, \dots, r.$$

Vale il seguente risultato noto in letteratura come il "Teorema degli orlati":

Teorema 13 Il rango di $A \in M_{m,n}(\mathbf{K})$ è r se e solo se esiste una sottomatrice quadrata di ordine r di A con determinante diverso da 0 e tutte le sottomatrici di A di ordine maggiore di r hanno determinante uguale a zero.

Cenno della Dimostrazione per i più curiosi! Supponiamo che il rango di A sia r . Allora A ha r colonne A^{j_1}, \dots, A^{j_r} linearmente indipendenti e ogni $(r+1)$ -pla di sue colonne sono linearmente dipendenti. Dunque la matrice $\hat{A} = (A^{j_1}, \dots, A^{j_r}) \in M_{m,r}(\mathbf{K})$ ha rango r e quindi ha r righe $\hat{A}_{i_1}, \dots, \hat{A}_{i_r}$ linearmente indipendenti. Allora $\tilde{A} = (\hat{A}_{i_1}, \dots, \hat{A}_{i_r}) \in M_{r,r}(\mathbf{K})$ è una sottomatrice di A con $\det(\tilde{A}) \neq 0$. Ogni sottomatrice $B \in M_{l,l}(\mathbf{K})$ di A di ordine $l > r$ ha determinante nullo, altrimenti B sarebbe sottomatrice di una matrice in $M_{m,l}(\mathbf{K})$ costituita da $l > r = \text{rg}(A)$ colonne di A linearmente indipendenti: questo non può essere! Supponiamo ora che esista una sottomatrice C quadrata di ordine r di A con determinante diverso da 0 e tutte le sottomatrici di A di ordine maggiore di r hanno determinante uguale a zero. La matrice C è sottomatrice di una matrice in $M_{m,l}(\mathbf{K})$ costituita da r colonne di A linearmente indipendenti e quindi $r \leq \text{rg}(A)$. D'altra parte se tutte le sottomatrici $D \in M_{l,l}(\mathbf{K})$ di A di ordine $l > r$ hanno determinante nullo, allora per ogni scelta A^{j_1}, \dots, A^{j_l} di $l > r$ colonne di A costituisce un insieme di colonne linearmente dipendenti. Altrimenti la matrice $(A^{j_1}, \dots, A^{j_l})$ conterrebbe l righe $\check{A}_{i_1}, \dots, \check{A}_{i_l}$ linearmente indipendenti e quindi la matrice $\check{A} = (\check{A}_{i_1}, \dots, \check{A}_{i_l})$ sarebbe una sottomatrice quadrata di ordine $l > r$ di rango l e quindi $\det(\check{A}) \neq 0$ contro l'ipotesi Q.e.d.

5 Basi e coordinate. Cambi di base.

Ricordiamo la seguente:

Definizione. Sia V uno spazio vettoriale finitamente generato. $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base se e solo se v_1, \dots, v_n sono linearmente indipendenti e $V = \text{Span}(v_1, \dots, v_n)$. Abbiamo dimostrato

a lezione per spazi finitamente generati che le basi esistono. È facile dimostrare:

Proposizione 14 Siano V uno spazio vettoriale e $v_1, \dots, v_n \in V$ vettori linearmente indipendenti. Se per $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbf{K}$ si ha

$$\alpha_1 v_1 + \dots + \alpha_n v_n = \beta_1 v_1 + \dots + \beta_n v_n$$

allora $\alpha_j = \beta_j$ per ogni $j = 1, \dots, n$. In particolare, allora, se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V , allora ogni vettore $v \in V$ si esprime in modo unico come combinazione lineare dei vettori di \mathcal{B}

Dimostrazione Se $\alpha_1 v_1 + \dots + \alpha_n v_n = \beta_1 v_1 + \dots + \beta_n v_n$ allora $(\alpha_1 - \beta_1)v_1 + \dots + (\alpha_n - \beta_n)v_n = 0$, e quindi, dato che v_1, \dots, v_n sono vettori linearmente indipendenti, la tesi segue. Q.e.d.

Definizione. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base dello spazio vettoriale V e $v \in V$. Gli scalari $x_1, \dots, x_n \in \mathbf{K}$ tali che $v = x_1 v_1 + \dots + x_n v_n$ si dicono coordinate di v relative alla base \mathcal{B} e

il vettore colonna $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbf{K}^n$ vettore delle coordinate di v .

Vogliamo ora scoprire con che legge cambiano le coordinate un vettore quando si considerano basi diverse. Precisamente sia V uno spazio vettoriale e siano

$$\mathcal{B} = \{v_1, \dots, v_n\} \quad \mathcal{B}' = \{v'_1, \dots, v'_n\}$$

due basi per V . Vogliamo risolvere il seguente

Problema: Al variare di $v \in V$, se $x \in \mathbf{K}^n$ è il vettore delle coordinate di v alla base \mathcal{B} e $x' \in \mathbf{K}^n$ è il vettore delle coordinate di v alla base \mathcal{B}' , esprimere x in funzione di x' .

Per risolvere il problema, procediamo nel modo seguente. Per ogni $j = 1, \dots, n$, esistono, e sono unici, scalari b_{1j}, \dots, b_{nj} tali che

$$v'_j = b_{1j} v_1 + \dots + b_{nj} v_n = \sum_{i=1}^n b_{ij} v_i.$$

Dunque per $v \in V$, si deve avere:

$$\sum_{i=1}^n x_i v_i = v = \sum_{j=1}^n x'_j v'_j = \sum_{j=1}^n x'_j \left(\sum_{i=1}^n b_{ij} v_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n b_{ij} x'_j \right) v_i.$$

Per l'unicità delle coordinate, allora si deve avere:

$$x_i = \sum_{j=1}^n b_{ij} x'_j.$$

Dunque abbiamo dimostrato la seguente

Proposizione 15 Sia

$$B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} \quad (*)$$

la matrice che ha per colonne le coordinate dei vettori della base \mathcal{B}' relative alla base \mathcal{B} . Allora per ogni vettore $v \in V$, se $x \in \mathbf{K}^n$ è il vettore delle coordinate di v alla base \mathcal{B} e $x' \in \mathbf{K}^n$ è il vettore delle coordinate di v alla base \mathcal{B}' , allora

$$x = Bx'.$$

La matrice B in (*) si dice *matrice del cambio di base* dalla base \mathcal{B} alla base \mathcal{B}' (in questo ordine!).

Osservazione. La matrice B è invertibile dato che il sistema $Bx' = x$ ha soluzione unica per ogni colonna x . La sua inversa B^{-1} è esattamente la matrice del cambio di base dalla base \mathcal{B}' alla base \mathcal{B} , ossia quella che ha per colonne le coordinate dei vettori della base \mathcal{B} relative alla base \mathcal{B}' .

6 Applicazioni lineari e matrici.

Siano V e W spazi vettoriali finitamente generati e sia $T: V \rightarrow W$ un'applicazione lineare. Siano fissate una base $\mathcal{B} = \{v_1, \dots, v_n\}$ per V e una base $\mathcal{C} = \{w_1, \dots, w_m\}$ per W . Vogliamo trovare il modo di calcolare le coordinate rispetto alla base \mathcal{C} dell'immagine $T(v)$ mediante T di un vettore $v \in V$ in funzione delle coordinate di v relative alla base \mathcal{B} . Per ogni $i = 1, \dots, n$ si ha per $V_i \in \mathcal{B}$

$$T(v_i) = a_{1i}w_1 + \dots + a_{mi}w_m \quad (\bullet)$$

dove $a_{1i}w_1, \dots, a_{mi}$ sono le coordinate di $T(v_i)$ relative alla base \mathcal{C} . Chiamiamo la matrice

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

che ha per colonne le coordinate relative alla base \mathcal{C} di W delle immagini dei vettori della base \mathcal{B} di V la *matrice associata a T relativa alle basi \mathcal{B} di V e \mathcal{C} di W* . Il risultato che cerchiamo è il seguente:

Teorema 16 Per ogni $v \in V$ se $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ è la colonna delle coordinate di v relative alla base

\mathcal{B} di V e $y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ è la colonna delle coordinate di $T(v)$ relative alla base \mathcal{C} di W , allora si

ha:

$$y = Ax \quad (\bullet\bullet)$$

dove A è la matrice associata a T .

Dimostrazione Utilizzando la linearità di T e (\bullet) , abbiamo

$$\begin{aligned} y_1w_1 + \dots + y_mw_m &= T(v) = T(x_1v_1 + \dots + x_nv_n) = x_1T(v_1) + \dots + x_nT(v_n) \\ &= x_1(a_{11}w_1 + \dots + a_{1m}w_m) + \dots + x_n(a_{n1}w_1 + \dots + a_{nm}w_m) \\ &= (a_{11}x_1 + \dots + a_{1n}x_n)w_1 + \dots + (a_{m1}x_1 + \dots + a_{mn}x_n)w_m \end{aligned}$$

e quindi per ogni $j = 1, \dots, m$, per l'unicità delle coordinate, si ha

$$y_j = a_{j1}x_1 + \dots + a_{jn}x_n$$

che equivale a $(\bullet\bullet)$.

Q.e.d.

Problema: Data un'applicazione lineare $T: V \rightarrow W$ fra spazi vettoriali finitamente generati, determinare la relazione che c'è fra le matrici associate a T relativamente a diverse scelte di basi per V e W

Per semplicità diamo risposta a questa domanda nel caso di *endomorfismi* ossia di applicazioni lineari $T: V \rightarrow V$ ossia di uno spazio vettoriale in sé. In effetti questo sarà il caso che ci interesserà.

Sia allora V uno spazio vettoriale finitamente generato su \mathbf{K} e siano $\mathcal{B}, \mathcal{B}'$ basi di V e sia B la matrice del cambio di base da \mathcal{B} a \mathcal{B}' . Sia inoltre A la matrice di T rispetto alla base \mathcal{B}

(scelta nel dominio e nel codominio) e sia A' la matrice di T rispetto alla base \mathcal{B}' (scelta nel dominio e nel codominio).

Dunque, se per un vettore v rispettivamente $x \in \mathbf{K}^n$ è il vettore delle coordinate di v alla base \mathcal{B} , $x' \in \mathbf{K}^n$ è il vettore delle coordinate di v alla base \mathcal{B}' , si ha $x = Bx'$. Analogamente, se rispettivamente $y \in \mathbf{K}^n$ è il vettore delle coordinate di $T(v)$ alla base \mathcal{B} , $y' \in \mathbf{K}^n$ è il vettore delle coordinate di $T(v)$ alla base \mathcal{B}' , si ha $y = By'$. Ora, per definizione di matrice associata, abbiamo

$$y = Ax \quad \text{e} \quad y' = A'x'.$$

Ma d'altra parte:

$$y' = B^{-1}y = B^{-1}Ax = B^{-1}ABx'.$$

Dato che questo vale per ogni v , possiamo concludere:

Proposizione 17 *Le matrici A e A' associate a $T: V \rightarrow V$ rispetto a basi \mathcal{B} e \mathcal{B}' di V si ha*

$$A' = B^{-1}AB$$

dove B è la matrice del cambio di base da \mathcal{B} a \mathcal{B}' .

Definizione Due matrici A e A' quadrate di ordine n si dicono simili se esiste una matrice invertibile B di ordine n tale che $A' = B^{-1}AB$.

Concludiamo con un'osservazione sull'utilità di questa nozione. Sia $A \in M_n(\mathbf{K})$ e sia $f_A: \mathbf{K}^n \rightarrow \mathbf{K}^n$ l'applicazione lineare definita da $f_A(x) = Ax$. In molte circostanze è interessante iterare

l'applicazione f_A molte volte e studiare il comportamento dell'applicazione $f_A^k = \overbrace{f_A \circ \dots \circ f_A}^{k \text{ volte}}$ per k molto grande. Si vede subito che $f_A^k = f_{A^k}$ e quindi che per studiare questo problema occorre essere in grado di calcolare le potenze A^k di A per k molto grande. In generale questo ne è molto complicato. Per matrici diagonali però il calcolo è semplice. Si vede subito che per una matrice diagonale

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & \lambda_{n-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix} \quad \text{si ha} \quad \Lambda^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 & 0 \\ 0 & \lambda_2^k & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & \lambda_{n-1}^k & 0 \\ 0 & 0 & \dots & 0 & \lambda_n^k \end{pmatrix}.$$

D'altra parte, per matrici simili M e $A = B^{-1}MB$, si ha che anche le k -esime potenze sono simili:

$$\begin{aligned} A^k &= (B^{-1}MB)(B^{-1}MB) \dots (B^{-1}MB)(B^{-1}MB) \\ &= B^{-1}M(BB^{-1})M(B \dots B^{-1})M(BB^{-1})MB \\ &= B^{-1}MM \dots MMB = B^{-1}M^k B. \end{aligned}$$

Dunque se $A = B^{-1}\Lambda B$ per una matrice Λ diagonale, allora si ha anche $A^k = B^{-1}\Lambda^k B$ e quindi A^k è simile alla matrice diagonale Λ^k , che è molto semplice da calcolare, mediante la matrice invertibile B .

7 Autovalori e Autovettori. Diagonalizzazione

Definizione Sia V uno spazio vettoriale su \mathbf{K} e sia $T: V \rightarrow V$ un endomorfismo. Un vettore non nullo $v \in V \setminus \{O\}$ si dice *autovettore* di T relativo all'autovalore $\lambda \in \mathbf{K}$ se si ha

$$T(v) = \lambda v \quad (1)$$

Definizione bis Sia A una matrice quadrata di ordine n . Un vettore non nullo $x \in \mathbf{K}^n \setminus \{O\}$ si dice *autovettore* di A relativo all'autovalore $\lambda \in \mathbf{K}$ se si ha $Ax = \lambda x$ (dunque se e solo se x è autovettore di $f_A: \mathbf{K}^n \rightarrow \mathbf{K}^n$ relativo all'autovalore λ).

Proposizione 18 Sia V uno spazio vettoriale e sia $T: V \rightarrow V$ un endomorfismo. Allora $\lambda \in \mathbf{K}$ è autovalore di $T \iff \text{Ker}(T - \lambda Id_V) \neq \{O\}$. Inoltre $v \in V$ è un autovettore di T relativo all'autovalore $\lambda \in \mathbf{K} \iff v \in \text{Ker}(T - \lambda Id_V) \setminus \{O\}$.

Dimostrazione Basta osservare che (1) equivale a $(T - \lambda Id_V)(v) = O$. Q.e.d.

Osservazione Uno scalare $\lambda \in \mathbf{K}$ è autovalore di una matrice $A \iff \text{Ker}(f_A - \lambda Id_{\mathbf{K}^n}) \neq \{O\}$ e quindi se e solo se ci sono soluzioni non nulle per il sistema omogeneo $(A - \lambda I_n)x = O$, dunque se e solo se $A - \lambda I_n$ è una matrice singolare. In questo caso gli autovettori $x \neq O$ di A relativo all'autovalore λ sono esattamente le soluzioni non nulle del sistema omogeneo $(A - \lambda I_n)x = O$.

Definizione Sia V uno spazio vettoriale e sia $T: V \rightarrow V$ un endomorfismo. Se $\lambda \in \mathbf{K}$ è un autovalore di T , il sottospazio $V_\lambda = \text{Ker}(T - \lambda Id_V)$ si dice *autospazio* di T relativo all'autovalore λ e la sua dimensione $\dim(V_\lambda)$ si dice *molteplicità geometrica* di λ . Analogamente se $\lambda \in \mathbf{K}$ è autovalore di una matrice A , il sottospazio $V_\lambda = \text{Ker}(T - \lambda Id_V)$ si dice *autospazio* di T relativo all'autovalore λ e la sua dimensione $\dim(V_\lambda)$ si dice *molteplicità geometrica* di λ .

Proposizione 19 Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V e sia \mathcal{B} una base di V . Allora la matrice di T relativa a \mathcal{B} è diagonale $\iff \mathcal{B}$ è composta di autovettori di T .

Dimostrazione Sia $\mathcal{B} = \{v_1, \dots, v_n\}$. Se i v_i sono tutti autovettori relativi a autovalori λ_i , allora $T(v_i) = \lambda_i v_i$ per ogni i e quindi la matrice di T relativa a \mathcal{B} è diagonale con i $\lambda_1, \dots, \lambda_n$ sulla diagonale. Viceversa se la matrice di T relativa a \mathcal{B} è diagonale con i $\lambda_1, \dots, \lambda_n$ sulla diagonale, allora $T(v_i) = \lambda_i v_i$ per ogni i e quindi \mathcal{B} è composta di autovettori di T . Q.e.d.

Definizione Un endomorfismo $T: V \rightarrow V$ di uno spazio vettoriale V si dice *diagonalizzabile* se esiste una base di V rispetto alla quale T ha matrice diagonale. Questo equivale a richiedere che esiste una base di autovettori di V .

Per le matrici abbiamo

Definizione bis Una matrice A quadrata di ordine n si dice *diagonalizzabile* se e solo è simile a una matrice diagonale ossia se e solo se esiste una matrice invertibile B tale che $B^{-1}AB$ sia diagonale.

Osservazione Una matrice A quadrata di ordine n e quindi diagonalizzabile se e solo se l'endomorfismo $L_A: \mathbf{K}^n \rightarrow \mathbf{K}^n$ è diagonalizzabile. La matrice invertibile B tale che $B^{-1}AB$ sia diagonale è la matrice del cambio di base dalla base canonica alla base costituita da autovettori di A e quindi è la matrice che ha per colonne i vettori colonna che costituiscono tale base

di autovettori.

Per endomorfismi di spazi finitamente generati si ha la seguente caratterizzazione degli autovalori:

Teorema 20 Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V di dimensione n sul campo \mathbf{K} . Siano \mathcal{B} una base di V e sia A la matrice di T relativa a \mathcal{B} . Allora

(i) La funzione $p_T: \mathbf{K} \rightarrow \mathbf{K}$ definita da

$$p_T(\lambda) = \det(A - \lambda I_n)$$

è un polinomio di grado n che non dipende dalla scelta della base \mathcal{B} .

(ii) Uno scalare $\lambda_0 \in \mathbf{K}$ è un autovalore di $T \iff p_T(\lambda_0) = 0$.

Dimostrazione (i) Dimostriamo che $p_T(\lambda)$ non dipende dalla scelta della base. Sia A' la matrice di T relativa a un'altra scelta di base \mathcal{B}' di V . Se B è la matrice del cambio di base da \mathcal{B} a \mathcal{B}' allora $A' = B^{-1}AB$. Occorre provare che $\det(A' - \lambda I_n) = \det(A - \lambda I_n)$. Questo segue utilizzando il Teorema di Binet:

$$\begin{aligned} \det(A' - \lambda I_n) &= \det(B^{-1}AB - \lambda I_n) = \det(B^{-1}AB - B^{-1}\lambda I_n B) \\ &= \det(B^{-1}(A - \lambda I_n)B) = \det B^{-1} \det(A - \lambda I_n) \det B \\ &= (\det B)^{-1} \det(A - \lambda I_n) \det B = \det(A - \lambda I_n). \end{aligned}$$

Dimostriamo che $p_T(\lambda)$ è un polinomio di grado n per induzione sulla dimensione n di V . Se $n = 1$ il fatto è ovvio. Supponiamo che sia vero per $n - 1$. Allora

$$\begin{aligned} p_T(\lambda) &= \det(A - \lambda I_n) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n1} - \lambda \end{pmatrix} \\ &= (a_{11} - \lambda) \det \begin{pmatrix} a_{12} - \lambda & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{n1} - \lambda \end{pmatrix} + \text{monomi di grado } \leq n - 2 \text{ in } \lambda \end{aligned}$$

e quindi per l'ipotesi induttiva, $p_T(\lambda)$ è un polinomio di grado n .

(ii) Lo scalare $\lambda_0 \in \mathbf{K}$ è autovalore di $T \iff \text{Ker}(T - \lambda_0 Id_V) \neq \{O\} \iff (A - \lambda_0 I_n)$ non è invertibile $\iff \det(A - \lambda_0 I_n) = 0$. Q.e.d.

Definizione Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V . Il polinomio $p_T(\lambda)$ si dice *polinomio caratteristico* di T .

Definizione bis Se A è una matrice quadrata reale di ordine n allora $p_A(\lambda) = \det(A - \lambda_0 I_n)$ (che secondo la Definizione è il polinomio caratteristico dell'applicazione f_A) si dice *polinomio caratteristico* di A

Come conseguenza immediata del Teorema 20 uno scalare λ è un autovalore di un endomorfismo $T: V \rightarrow V$ (o di una matrice quadrata A) se e solo se è radice del polinomio caratteristico di T (di A). Dunque non è detto che autovalori (e conseguentemente autovettori) esistano dato che non tutti i polinomi reali hanno radici reali. L'esistenza in numero "giusto" degli autovalori e per ciascun autovalore del numero "giusto" di autovettori linearmente indipendenti, è esattamente quello

che si deve richiedere per caratterizzare la diagonalizzabilità. Illustriamo questo fatto con due esempi semplici ma particolarmente importanti per spiegare le difficoltà che si possono incontrare.

Esempio 1. Sia $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ e si consideri $f_A: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ definita come al solito da $f_A(x) = Ax$. Il polinomio caratteristico di A è $\lambda^2 + 1$ e quindi non vi sono autovalori in \mathbf{R} di A . Si conclude allora che $f_A: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ non è diagonalizzabile e che A non è simile a una matrice diagonale reale ossia non esiste una matrice invertibile B reale tale che $B^{-1}AB$ sia diagonale.

Esempio 2. Sia $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ e si consideri $f_A: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ definita come al solito da $f_A(x) = Ax$. Dato che il polinomio caratteristico di f_A (e di A) è $(\lambda - 1)^2$ Si vede subito che l'unico autovalore di f_A (e di A) è 1. D'altra parte l'autospazio corrispondente è $\{x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1 = 0\}$ che ha dimensione 1. Dunque non può esistere una base di \mathbf{R}^2 costituita da autovettori di f_A (o di A) e quindi f_A (e A) non sono diagonalizzabili.

Definizione Sia λ_0 un autovalore di un endomorfismo di uno spazio vettoriale V . La *molteplicità algebrica* di λ_0 è la molteplicità di λ_0 come radice del polinomio $p_T(\lambda)$. Sia λ_0 un autovalore di una matrice quadrata A di ordine n con entrate in \mathbf{K} . La *molteplicità algebrica* λ_0 è di λ_0 come radice del polinomio $p_A(\lambda)$

Il criterio di diagonalizzabilità in termini di molteplicità è il seguente:

Teorema 21 (i) Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V su un campo \mathbf{K} . Allora T è diagonalizzabile se e solo se tutti gli autovalori di T sono in \mathbf{K} e per ogni autovalore la molteplicità geometrica coincide con la molteplicità algebrica.

(ii) Sia A una matrice quadrata di ordine n con entrate in \mathbf{K} . Allora A è diagonalizzabile se e solo se tutti gli autovalori di A sono in \mathbf{K} e per ogni autovalore la molteplicità geometrica coincide con la molteplicità algebrica.

Non daremo dimostrazione del teorema (Ad esempio si può trovare al paragrafo 11.3 del libro Abate, *Algebra Lineare*). Per dare un'idea di quello che succede, consideriamo un caso particolare. Cominciamo con la seguente

Proposizione 22 Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale. Se v_1, \dots, v_k sono autovettori di T corrispondenti a autovalori distinti $\lambda_1, \dots, \lambda_k$, allora v_1, \dots, v_k sono linearmente indipendenti.

Dimostrazione Procediamo per induzione su k . Se $k = 1$, il risultato è banale. Supponiamo il risultato vero per $k - 1$. Siano $\alpha_1, \dots, \alpha_k$ scalari tali che

$$0 = \alpha_1 v_1 + \dots + \alpha_{k-1} v_{k-1} + \alpha_k v_k. \quad (2)$$

Allora

$$0 = T(\alpha_1 v_1 + \dots + \alpha_k v_k) = \alpha_1 T(v_1) + \dots + \alpha_k T(v_k) = \alpha_1 \lambda_1 v_1 + \dots + \alpha_{k-1} \lambda_{k-1} v_{k-1} + \alpha_k \lambda_k v_k \quad (3)$$

Moltiplicando la (2) per λ_k e sottraendo alla (3) otteniamo:

$$\begin{aligned} 0 &= \alpha_1 (\lambda_1 - \lambda_k) v_1 + \dots + \alpha_{k-1} (\lambda_{k-1} - \lambda_k) v_{k-1} + \alpha_k (\lambda_k - \lambda_k) v_k \\ &= \alpha_1 (\lambda_1 - \lambda_k) v_1 + \dots + \alpha_{k-1} (\lambda_{k-1} - \lambda_k) v_{k-1}. \end{aligned}$$

Per l'ipotesi induttiva $\alpha_1 (\lambda_1 - \lambda_k) = \dots = \alpha_{k-1} (\lambda_{k-1} - \lambda_k) = 0$ e quindi $\alpha_1 = \dots = \alpha_{k-1} = 0$, dato che $\lambda_i \neq \lambda_j$ se $i \neq j$. Ma allora da (*) segue anche $\alpha_k = 0$. Q.e.d.

Proposizione 23 Sia V spazio vettoriale di dimensione n . Se un endomorfismo $T: V \rightarrow V$ ha n autovalori distinti in \mathbf{K} , T è diagonalizzabile. Analogamente se una matrice quadrata A di ordine n con entrate in \mathbf{K} ha n autovalori distinti in \mathbf{K} , A è diagonalizzabile.

Dimostrazione Il risultato è immediata conseguenza della Proposizione 22: se $\lambda_1, \dots, \lambda_n \in \mathbf{K}$ sono autovalori a due a due distinti di T e v_1, \dots, v_n sono autovettori corrispondenti, allora v_1, \dots, v_n sono linearmente indipendenti e quindi, dato che $\dim(V) = n$, formano una base per V . Per le matrice la dimostrazione è analoga. Q.e.d.

Un altro risultato fondamentale, che non dimostreremo ma che è utile conoscere perchè caratterizza in modo semplice una intera classe di matrici diagonalizzabili, è il seguente:

Teorema 24 Sia $A \in M_{n,n}(\mathbf{R})$ una matrice reale simmetrica ossia tale che $A = A^T$. Allora tutti gli autovalori di A sono reali e A è diagonalizzabile.

In realtà per ogni matrice simmetrica reale A si può sempre trovare una matrice invertibile B tale che $B^{-1} = B^T$ (una tale matrice si dice *ortogonale*) tale che $B^{-1}AB = B^T AB$ sia diagonale con sulla diagonale gli autovalori (tutti reali!) di A . Per la dimostrazione di questo importante teorema (che è noto come Teorema Spettrale Reale) e di tutti i prerequisiti rimandiamo ad esempio ai capitoli 10 e 11 del libro Abate, *Algebra Lineare*. Qui enunciamo per comodità del lettore e dimostriamo in modo elementare il risultato per matrici di ordine 2:

Proposizione 25 (Teorema Spettrale per $n = 2$) Sia $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ una matrice reale simmetrica di ordine 2. Allora A ha due autovalori reali μ, ν . Se gli autovalori di A sono coincidenti ossia $\mu = \nu$ allora A è diagonale: $A = \begin{pmatrix} \mu & 0 \\ 0 & \mu \end{pmatrix}$. Se gli autovalori di A sono distinti $\mu \neq \nu$ allora esiste $t \in \mathbf{R}$ tale che il vettore $\begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ è autovettore relativo a μ e il vettore $\begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$ è autovettore relativo a ν . Dunque $\left\{ \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \right\}$ è una base di autovettori di \mathbf{R}^2 costituita da autovettori di A . Inoltre se $B = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$, allora B è ortogonale ossia $B^{-1} = B^T$ e $B^{-1}AB = B^T AB = \begin{pmatrix} \mu & 0 \\ 0 & \nu \end{pmatrix}$.

Dimostrazione Il polinomio caratteristico di A è

$$p(\lambda) = (a - \lambda)(c - \lambda) - b^2 = \lambda^2 - (a + c)\lambda + ac - b^2$$

che ha discriminante $\Delta = (a + c)^2 - 4ac + 4b^2 = (a - c)^2 + 4b^2 \geq 0$. Dunque le radici di $p(\lambda)$ sono reali e quindi A ha autovalori reali. Gli autovalori sono coincidenti (ossia $\mu = \nu$) se e solo se $(a - c)^2 + 4b^2 = \Delta = 0$ e questo succede se e solo se $a = c$ e $b = 0$ dunque se A è già diagonale. In questo caso deve essere $a = c = \mu$. Supponiamo ora che gli autovalori siano distinti: $\mu \neq \nu$. Osserviamo preliminarmente che si deve avere:

$$\lambda^2 - (\mu + \nu)\lambda + \mu\nu = (\lambda - \mu)(\lambda - \nu) = p(\lambda) = \lambda^2 - (a + c)\lambda + ac - b^2$$

e quindi, dato che due polinomi sono uguali se e solo se hanno coefficienti dello stesso grado uguali, si ha

$$\mu + \nu = a + c \quad \mu\nu = ac - b^2 = \text{Det}(A). \quad (4)$$

Sia $\begin{pmatrix} x \\ y \end{pmatrix}$ un autovettore relativo a μ . Allora $N = x^2 + y^2 \neq 0$ e $\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{1}{N} \begin{pmatrix} x \\ y \end{pmatrix}$ è autovettore di A relativo a μ con $X^2 + Y^2 = 1$. Dunque per qualche $t \in \mathbf{R}$ si ha $\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$.

Dimostriamo ora che $\begin{pmatrix} -Y \\ X \end{pmatrix}$ è autovettore relativo a ν . Utilizzando il fatto che $\begin{pmatrix} X \\ Y \end{pmatrix}$ è autovalore relativo a μ e la prima eguaglianza in (4), si ha:

$$\begin{pmatrix} aX + bY \\ bX + cY \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix} = \mu \begin{pmatrix} X \\ Y \end{pmatrix} = (a+c-\nu) \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} (a+c-\nu)X \\ (a+c-\nu)Y \end{pmatrix}.$$

Dunque $-cX + bY = -\nu X$ e $bX - aY = -\nu Y$, ossia

$$A \begin{pmatrix} -Y \\ X \end{pmatrix} = \begin{pmatrix} -aY + bX \\ cX - bY \end{pmatrix} = \begin{pmatrix} -\nu Y \\ \nu X \end{pmatrix} = \nu \begin{pmatrix} -Y \\ X \end{pmatrix}$$

come desiderato. Dunque $\left\{ \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \right\}$ è una base di autovettori di \mathbf{R}^2 costituita da autovettori di A e $B = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$ è una matrice invertibile. Una calcolo immediato mostra che $B^{-1} = B^T$. Le altre implicazioni sono immediate. *Q.e.d.*

Concludiamo illustrando la procedura che si deve seguire per verificare se un endomorfismo è diagonalizzabile e nel caso per trovare una base composta di autovettori.

Algoritmo di diagonalizzazione per endomorfismi: Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V su \mathbf{K} con $\dim(V) = n$

I passo: Si fissi una base \mathcal{B} di V e si trovi la matrice A associata a T relativa alla base \mathcal{B} (ricordiamo che è la matrice che ha per colonne le coordinate rispetto alla base \mathcal{B} delle immagini mediante T dei vettori della base \mathcal{B}).

II passo: Trovare il polinomio caratteristico $P_T(\lambda) = \det(A - \lambda I_n)$ e determinarne le radici. Se, contate con la loro molteplicità, ha n radici in \mathbf{K} si può proseguire, altrimenti si conclude che T non è diagonalizzabile (perché non ha abbastanza autovalori!).

III passo: Se $\lambda_1, \dots, \lambda_r$ sono gli autovalori distinti di T , per ogni $j = 1, \dots, r$ si determini una base \mathcal{C}_j per l'autospazio V_{λ_j} relativo all'autovalore λ_j .

IV passo: Si consideri $\mathcal{C} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_r$. Allora \mathcal{C} è un sistema di vettori linearmente indipendenti di V . L'endomorfismo T è diagonalizzabile se e solo se \mathcal{C} contiene n vettori perché in questo caso \mathcal{C} è una base di V (contiene il numero giusto di vettori linearmente indipendenti) costituita da autovettori di T . Se \mathcal{C} è una base, la matrice di T relativa a \mathcal{C} è la matrice diagonale Δ con sulla diagonale gli autovalori di T . La matrice B del cambio di base dalla base \mathcal{B} alla base \mathcal{C} (quella che ha per colonne le coordinate rispetto a \mathcal{B} dei vettori di \mathcal{C} , è la matrice tale che $\Delta = B^{-1}AB$.

Naturalmente l'algoritmo si può adattare per diagonalizzare matrici.

Algoritmo di diagonalizzazione matrici: Sia A una matrice quadrata con entrate in \mathbf{K} di ordine n .

I passo: Trovare il polinomio caratteristico $P_A(\lambda) = \det(A - \lambda I_n)$ e determinarne le radici. Se, contate con la loro molteplicità, ha n radici in \mathbf{K} si può proseguire, altrimenti si conclude che A non è diagonalizzabile (perché non ha abbastanza autovalori!). Ad esempio se A è simmetrica gli autovalori sono tutti reali (anche se non necessariamente si è in grado di calcolarli!).

II passo: Se $\lambda_1, \dots, \lambda_r$ sono gli autovalori distinti di A , per ogni $j = 1, \dots, r$ si determini una

base \mathcal{C}_j per l'autospazio $V_{\lambda_j} = \text{Ker}(A - \lambda_j I_n)$ relativo all'autovalore λ_j (che non è altro che l'insieme delle soluzioni del sistema omogeneo $(A - \lambda_j I_n)x = 0$).

III passo: Si consideri $\mathcal{C} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_r$. Allora \mathcal{C} è un sistema di vettori linearmente indipendenti di \mathbf{K}^n . La matrice A è diagonalizzabile se e solo se \mathcal{C} contiene n vettori perchè in questo caso \mathcal{C} è una base di \mathbf{K}^n (contiene il numero giusto di vettori linearmente indipendenti) costituita da autovettori di A . Se \mathcal{C} è una base, la matrice di f_A relativa a \mathcal{C} è la matrice diagonale Δ con sulla diagonale gli autovalori di A . Dunque se B è la matrice del cambio di base dalla base canonica alla base \mathcal{C} (quella che ha per colonne i vettori di \mathcal{C} , è la matrice tale che $\Delta = B^{-1}AB$.

Esempio. Applichiamo quanto detto per studiare la matrice $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$.

La matrice A è simmetrica reale, dunque ci aspettiamo che, contati con la molteplicità, abbia 4 autovalori reali e che sia diagonalizzabile. In effetti il polinomio caratteristico di A è $p_A(\lambda) = \det(A - \lambda I_4) = (\lambda^2 - 4\lambda + 3)(\lambda^2 - 2\lambda)$ che ha 4 radici distinte 0, 1, 2, 3 che sono gli autovalori di A . Cercare una base per gli autospazi corrispondenti a ciascun autovalore corrisponde a trovare basi rispettivamente per

$$\text{Ker}A, \text{Ker}(A - I_4), \text{Ker}(A - 2I_4), \text{Ker}(A - 3I_4).$$

Dato che

$$\text{Ker}A = \text{Span} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad \text{Ker}(A - I_4) = \text{Span} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix},$$

$$\text{Ker}(A - 2I_4) = \text{Span} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \text{Ker}(A - 3I_4) = \text{Span} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

allora una base di \mathbf{R}^4 rispetto alla quale $f_A : \mathbf{R}^4 \rightarrow \mathbf{R}^4$ ha matrice diagonale è data da

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

La matrice di f_A rispetto a questa base è

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Si osservi che la matrice

$$B = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

che ha per colonne i vettori che formano basi rispettivamente di $\text{Ker}A$, $\text{Ker}(A - I_4)$, $\text{Ker}(A - 2I_4)$, $\text{Ker}(A - 3I_4)$ ha inversa

$$B^{-1} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{pmatrix}$$

e con un semplice calcolo si verifica che $D = B^{-1}AB$.