Chapter 5

Autovalori e autovettori

In questo capitolo vedremo un'importante nozione dalle numerose applicazioni pratiche.

5.1 Definizione, esempi e applicazioni

Definizione 5.1. Sia V un $\mathbb{K}$-spazio vettoriale e $f : V \to V$ un endomorfismo[1]. Un vettore $v \neq \bar{0}$ di V si dice *autovettore di* f se si ha

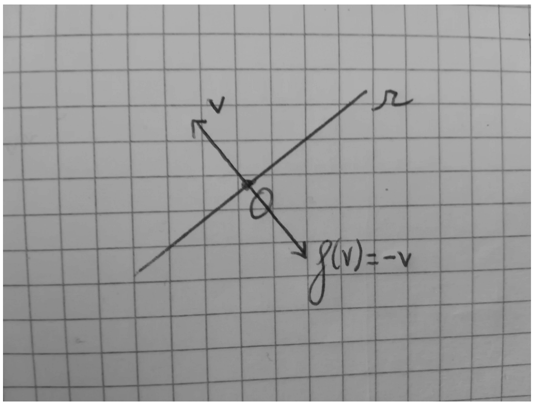
$$f(v) = \lambda v \tag{5.1}$$

per qualche $\lambda \in \mathbb{K}$.
Il coefficiente numerico λ si dice *autovalore relativo all'autovettore* v.

In altre parole, un autovettore è un vettore non nullo che viene mandato dalla funzione in un suo multiplo. Notiamo che questo è sempre banalmente vero per il vettore nullo $\bar{0}$, in quanto come sappiamo (si veda l'inizio della dimostrazione della Proposizione 4.6) se f è lineare allora $f(\bar{0}) = \bar{0}$, e quindi l'uguaglianza $f(\bar{0}) = \lambda\bar{0}$ è verificata sempre per qualunque endomorfismo e qualunque scalare λ (è per questo motivo che questo caso banale viene escluso dalla Definizione 5.1).
Vediamo subito alcuni esempi di autovettori:

Esempio 5.2. Sia $V = V_O^2$ lo spazio dei vettori geometrici applicati in un punto O del piano e sia $f : V_O^2 \to V_O^2$ la riflessione rispetto a una retta r che passa per O (si veda pagina 132). Quando riflettiamo un vettore $v = \vec{OP}$ perpendicolare alla retta r, il vettore viene mandato nel suo opposto

[1]Ricordiamo che si dicono endomorfismi le applicazioni lineari in cui dominio e codominio sono uguali.



ovvero $f(v) = -v = (-1)v$. Quindi tale vettore è un autovettore di f con autovalore associato -1.

Esempio 5.3. Nel caso di endomorfismi $f = L_A : \mathbb{K}^n \to \mathbb{K}^n$ del tipo (4.9) determinati da una matrice quadrata $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$, ovvero

$$f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{pmatrix} \tag{5.2}$$

gli autovettori e gli autovalori di L_A si dicono anche autovalori e autovettori della matrice A. Ricordando che, in base alla definizione di prodotto di matrici, si ha

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{5.3}$$

e quindi $f(x) = Ax$, un autovettore di A è allora una n-upla $x \in \mathbb{K}^n$ non nulla tale che

$$Ax = \lambda x. \tag{5.4}$$

Ad esempio, se $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$ e $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, si ha

$$Ax = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 4x$$

e quindi x è un autovettore di A con autovalore associato $\lambda = 4$.
Invece, per esempio, $x = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ non è un autovettore di A in quanto

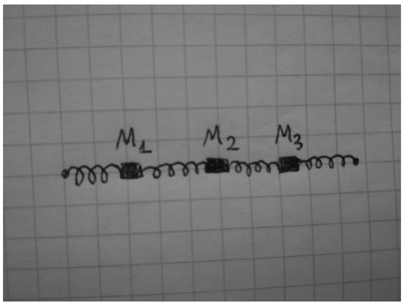
$$Ax = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}$$

e $(5, 7)$ non è multiplo di $(2, 1)$.

Prima di vedere come si calcolano autovalori e autovettori di un endomorfismo e di una matrice, citiamo alcune applicazioni pratiche di questa nozione:

Leggere

(1) Una massa collegata a una molla, se spostata dalla posizione di equilibrio, inizia a oscillare con una certa frequenza, dipendente dalla massa stessa e dalle caratteristiche fisiche della molla. Se invece abbiamo più masse $M_1, M_2, M_3 \ldots$, tutte collegate tra loro tramite delle molle,



l'oscillazione di ognuna di esse dipende anche dalla posizione e dal movimento delle altre masse del sistema e bisogna quindi tener conto di tutte le interazioni reciproche. Si mostra che tali interazioni possono essere descritte mediante una matrice quadrata, e gli autovalori di questa matrice ci danno le cosiddette *frequenze naturali di oscillazione del sistema*: il modo in cui il sistema oscilla è dato da una combinazione di queste frequenze tramite una formula che coinvolge anche gli autovettori corrispondenti.

Questo esempio spiega perché l'insieme degli autovalori di una matrice A si chiama anche il suo *spettro* (prendendo a prestito la parola dall'espressione "spettro di frequenze")

(2) Dato un corpo rigido nello spazio tridimensionale, possiamo far ruotare tale corpo attorno a un asse. A causa della forma del corpo e della distribuzione della massa in esso, eventualmente disomogenea, a seconda dell'asse che scegliamo la rotazione "non sarà stabile", ovvero ci saranno delle sollecitazioni esercitate sull'asse di rotazione che tendono a deviarlo. Vi sono tuttavia alcune direzioni con la proprietà che se facciamo ruotare il corpo attorno a esse allora tali sollecitazioni saranno nulle: gli assi con queste direzioni si dicono *assi principali d'inerzia*.

Leggere

Ebbene, si può vedere che il calcolo degli assi principali d'inerzia si riduce al calcolo degli autovettori di una opportuna matrice quadrata di ordine 3, costruita usando i dati sulla distribuzione della massa nel corpo.

(3) Abbiamo visto nel Capitolo 4 come, fissato un sistema di riferimento nello spazio tridimensionale, un piano sia descritto da una generica equazione $Ax + By + Cz = D$ di primo grado in tre incognite, nel senso che i punti appartenenti al piano sono tutti e soli quelli le cui coordinate (x, y, z) verificano l'equazione.

Ora, ci chiediamo cosa invece rappresenti una generica equazione di secondo grado nelle tre incognite date, ovvero

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + L = 0 \quad (5.5)$$

Si può dimostrare che le equazioni di questo tipo rappresentano superfici dette *quadriche* (tra le quali rientrano sfere, ellissoidi, iperboloidi, coni, cilindri)

e che, per capire quale tra queste superfici rappresenti l'equazione (5.5), bisogna calcolare gli autovalori di una matrice di ordine 4 costruita opportunamente dai coefficienti $A, B, C, \dots, L$.

Inoltre, gli autovettori forniscono informazioni sulla posizione di tale superficie nello spazio, dicendoci ad esempio quali sono i suoi assi di simmetria.

5.2 Calcolo di autovalori e autovettori

Dato un endomorfismo $f : V \to V$, allo scopo di trovare tutti i suoi autovettori, ovvero i vettori $v \neq \bar{0}$ che soddisfano la (5.1), basta lavorare in coordinate.
Più precisamente, fissiamo una base B qualunque dello spazio V e traduciamo l'uguaglianza (5.1) in coordinate rispetto a B: se x denota la n-upla delle coordinate di v (rispetto a B), disposte in colonna, allora

(1) il vettore λv a secondo membro della (5.1) ha coordinate λx (come risulta dalla proprietà (1.36) che afferma che le coordinate di λv si ottengono moltiplicando per λ le coordinate di v)

(2) il vettore $f(v)$ a primo membro della (5.1) ha coordinate (rispetto a B) che si ottengono moltiplicando la matrice $A = M_B(f)$ associata a f rispetto a B per le coordinate x di v è la matrice associata a f rispetto a B, come abbiamo visto nel Paragrafo 4.2.

Quindi la (5.1) equivale in coordinate all'uguaglianza $Ax = \lambda x$, ovvero la (5.4): in altre parole, il problema di trovare gli autovettori e gli autovalori di f si riduce a trovare gli autovettori e gli autovalori della matrice A associata a f rispetto a una base fissata.
Dobbiamo allora solo vedere come si trovano gli autovettori e gli autovalori di una matrice A.
Portando il secondo membro della $Ax = \lambda x$ a primo membro si ottiene $Ax - \lambda x = \bar{0}$: ricordando che $x = I_n x$ e riscrivendo tale uguaglianza come $Ax - \lambda I_n x = \bar{0}$, possiamo sfruttare la proprietà distributiva del prodotto di matrici e scrivere

$$(A - \lambda I_n)x = \bar{0}.$$

Abbiamo quindi trasformato il problema iniziale $f(v) = \lambda v$ in un sistema omogeneo di n equazioni in n incognite che ha $A - \lambda I_n$ come matrice dei coefficienti, nel senso che le soluzioni x di tale sistema sono proprio le coordinate (rispetto a B) dei vettori v tali che $f(v) = \lambda v$.
In particolare, noi siamo interessati all'esistenza di soluzioni non nulle di tale sistema, in quanto la soluzione nulla corrisponde al vettore nullo $v = \bar{0}$, che non è considerato un autovettore.

Ora, come sappiamo, un sistema compatibile $Mx = b$ in n incognite ha ∞^{n-r} soluzioni, essendo r il rango della matrice dei coefficienti M, e quindi ha una sola soluzione se e solo se il rango di M è uguale a n: nel caso del nostro sistema omogeneo $(A-\lambda I_n)x = \bar{0}$, esso ammetterà quindi altre soluzioni oltre quella nulla se e solo se il rango della sua matrice dei coefficienti $A - \lambda I_n$ è minore di n, condizione che, essendo la matrice quadrata, possiamo esprimere equivalentemente richiedendo che il suo determinante sia nullo.
Quindi, i $\lambda \in \mathbb{K}$ per cui esiste un vettore *non nullo* tale che $f(v) = \lambda v$ (cioè gli autovalori di f) sono esattamente quegli scalari per cui

$$\det(A - \lambda I_n) = 0. \tag{5.6}$$

L'equazione (5.6), risolvendo la quale determiniamo l'insieme di tutti gli autovalori di f si dice *equazione caratteristica*, e l'espressione $\det(A - \lambda I_n)$ a primo membro di tale equazione, che come vedremo negli esempi (si veda anche la fine del capitolo) è un polinomio di grado $n = \dim(V)$ in λ, si dice *polinomio caratteristico.*

Una volta determinati i $\lambda \in \mathbb{K}$ per cui $\det(A - \lambda I_n) = 0$, cioè tali che il sistema $(A - \lambda I_n)x = \bar{0}$ ha soluzioni non nulle, bisogna trovare tali soluzioni, che saranno esattamente le coordinate (rispetto a B) degli autovettori di f. Illustriamo subito tale procedimento con alcuni esempi.

Esempio 5.4. Sia $V = V_O^2$ lo spazio dei vettori geometrici applicati in un punto O del piano e sia $f : V_O^2 \to V_O^2$ la riflessione rispetto a una retta fissata r passante per O. Nell'Esempio 5.2 abbiamo già dato un esempio di autovettore di f, determiniamoli ora tutti usando il metodo appena descritto. A questo scopo dobbiamo prima di tutto scegliere una base B per poter lavorare in coordinate e scrivere la matrice $A = M_B(f)$ associata a f rispetto a B. Tale matrice è stata già calcolata, rispetto alla base $B = \{v_1, v_2\}$ raffigurata a pagina 143, nell'Esempio 4.3 (3), dove abbiamo visto che $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Allora

$$A - \lambda I_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} \tag{5.7}$$

e quindi il polinomio caratteristico è dato da

$$\det(A - \lambda I_2) = \lambda^2 - 1.$$

Gli autovalori di f sono quindi le radici reali di questo polinomio (V_O^2 è un $\mathbb{R}$-spazio vettoriale e gli autovalori devono stare in $\mathbb{R}$), ovvero le soluzioni dell'equazione caratteristica $\lambda^2 - 1 = 0$, che sono $\lambda_1 = 1$ e $\lambda_2 = -1$.

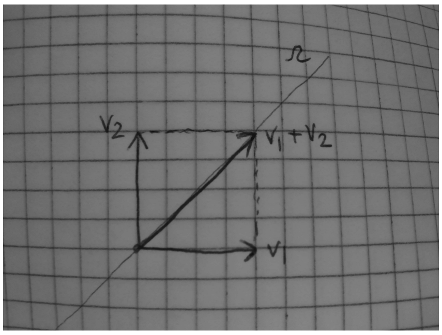
Quindi esistono dei vettori $v \in V$ non nulli per cui $f(v) = v$ e dei vettori non nulli per cui $f(v) = -v$. Per determinarli, come abbiamo detto sopra, dobbiamo risolvere il sistema $(A - \lambda I_2)x = \bar{0}$ (con $\lambda = 1$ e $\lambda = -1$), e le soluzioni di questo sistema ci daranno le coordinate di tali vettori rispetto alla base B.
Sostituendo $\lambda = 1$ nella (5.7) si ha che $(A - \lambda I_2)x = \bar{0}$ è

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ovvero

$$\begin{cases} -x_1 + x_2 = 0 \\ x_1 - x_2 = 0 \end{cases}$$

che ha come soluzioni tutte e sole le coppie del tipo (t, t), al variare di $t \in \mathbb{R}$. Tali coppie sono quindi le coordinate degli autovettori relativi a $\lambda = 1$ rispetto alla base $B = \{v_1, v_2\}$, che sono quindi tutti i vettori del tipo $tv_1 + tv_2 = t(v_1 + v_2)$, ovvero tutti i vettori proporzionali al vettore $v_1 + v_2$: è facile verificare graficamente come nel disegno seguente che tale vettore (come i suoi multipli) soddisfa proprio la proprietà $f(v) = v$, cioè la riflessione f lo lascia invariato (in quanto appartiene alla retta rispetto alla quale stiamo riflettendo).



Sostituendo invece $\lambda = -1$ nella (5.7) si ha che $(A - \lambda I_2)x = \bar{0}$ è

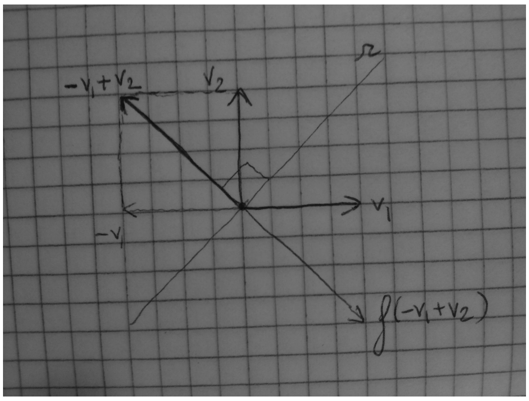
$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ovvero

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_2 = 0 \end{cases}$$

che ha come soluzioni tutte e sole le coppie del tipo $(-t, t)$, al variare di $t \in \mathbb{R}$.

Tali coppie sono le coordinate degli autovettori relativi a $\lambda = -1$ rispetto alla base $B = \{v_1, v_2\}$, che sono quindi tutti i vettori del tipo $-tv_1 + tv_2 = t(-v_1 + v_2)$, ovvero tutti i vettori proporzionali al vettore $-v_1 + v_2$: è facile verificare graficamente come nel disegno seguente che tale vettore (come i suoi multipli) soddisfa proprio la proprietà $f(v) = -v$, cioè la riflessione f lo cambia di verso (in quanto è perpendicolare alla retta di riflessione).



(avevamo già osservato nell'Esempio 5.2 che i vettori perpendicolari all'asse di riflessione sono autovettori relativi all'autovalore -1).

Per un altro esempio di tipo geometrico, si consideri ad esempio la rotazione f di angolo θ sullo spazio dei vettori applicati in un punto O nel piano. Abbiamo visto nell'Esempio 4.3 (3) che la matrice associata a f rispetto a una base B costituita da due vettori ortogonali e della stessa lunghezza è $A = M_B(f) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Prendiamo per facilità il caso $\theta = \pi/2$, nel quale ($\cos\theta = 0$, $\sin\theta = 1$) si ha $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
Il polinomio caratteristico di f è quindi

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

Poichè non esiste nessun $\lambda \in \mathbb{R}$ tale che $\lambda^2 + 1 = 0$, deduciamo che f non ha autovalori[2] (ovvero non ha autovettori).
Questo fatto poteva, in questo caso, essere dedotto direttamente dalla definizione di autovettore: poichè la rotazione cambia la direzione dei vettori, non manda nessun vettore v (diverso dal vettore nullo) in un suo multiplo, cioè non può mai succedere che $f(v) = \lambda v$.

[2]Poichè V_O^2 è un $\mathbb{R}$-spazio vettoriale, sono autovalori solo le soluzioni dell'equazione caratteristica che stanno in $\mathbb{R}$.

Osservazione 5.5. Se vediamo la matrice $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ come matrice a entrate complesse e ci proponiamo di determinare i suoi autovalori e autovettori in $\mathbb{C}$, allora il suo polinomio caratteristico $\lambda^2 + 1 = 0$ ha le due soluzioni (quindi, autovalori della matrice) $\lambda = +i$ e $\lambda = -i$.
Quindi possiamo determinarne i corrispondenti autovettori, che saranno elementi di $\mathbb{C}^2$. Più in dettaglio, per $\lambda = i$ si ha che $A - \lambda I_2 = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \underset{R_2 \to iR_2 + R_1}{\longrightarrow}$ $\begin{pmatrix} -i & -1 \\ 0 & 0 \end{pmatrix}$, quindi gli autovettori sono le soluzioni dell'unica equazione $-ix_1 - x_2 = 0$: ponendo $x_2 = t$ (dove t stavolta varia tra tutti i numeri complessi) si ha $x_1 = -\frac{1}{i}t$ e quindi gli autovettori sono dati da tutte le coppie del tipo $(-\frac{1}{i}t, t)$, al variare di $t \in \mathbb{C}$.
Analogamente, per $\lambda = -i$ si ha che $A - \lambda I_2 = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \underset{R_2 \to iR_2 - R_1}{\longrightarrow}$ $\begin{pmatrix} i & -1 \\ 0 & 0 \end{pmatrix}$, quindi gli autovettori sono le soluzioni dell'unica equazione $ix_1 - x_2 = 0$: ponendo $x_2 = t$ (dove t stavolta varia tra tutti i numeri complessi) si ha $x_1 = \frac{1}{i}t$ e quindi gli autovettori sono dati da tutte le coppie del tipo $(\frac{1}{i}t, t)$, al variare di $t \in \mathbb{C}$.

Osservazione 5.6. Si noti che in generale

$$A - \lambda I_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & \vdots & & \\ 0 & 0 & \cdots & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} - \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ & \vdots & & \\ 0 & 0 & \cdots & \lambda \end{pmatrix} = \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix}$$

ovvero $A - \lambda I_n$ è semplicemente la matrice A in cui abbiamo sottratto λ su tutta la diagonale.

Nell'esempio della riflessione, vediamo che l'insieme degli autovettori relativi a un dato autovalore coincide con l'insieme dei vettori che stanno su una retta per O (per l'autovalore $+1$, la retta di riflessione stessa, per l'autovalore -1, la retta perpendicolare all'asse di riflessione), e assieme al vettore nullo forma quindi un sottospazio vettoriale di V_O^2 (si veda l'Esempio 1.27 a pagina

47). Questo è un fatto generale, ovvero dato un endomorfismo $f : V \to V$ e dato un autovalore λ, l'insieme

$$\{v \in V \mid f(v) = \lambda v\}$$

costituito dagli autovettori relativi a λ più il vettore nullo (che soddisfa automaticamente la condizione $f(v) = \lambda v$ anche se per definizione non è un autovettore) è un sottospazio vettoriale di V, detto *autospazio relativo all'autovalore* λ e denotato $V_\lambda(f)$.
Infatti, siano $v_1, v_2 \in V_\lambda(f)$, ovvero $f(v_1) = \lambda v_1$ e $f(v_2) = \lambda v_2$. Allora

$$f(v_1 + v_2) = f(v_1) + f(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$$

(nella prima uguaglianza abbiamo sfruttato il fatto che f è un'applicazione lineare), e quindi vediamo che anche il vettore somma $v_1 + v_2$ verifica la condizione $f(v) = \lambda v$, quindi appartiene a $V_\lambda(f)$: tale sottoinsieme è quindi chiuso rispetto alla somma.
Per verificare che è chiuso rispetto al prodotto per scalari, e quindi completare la dimostrazione che si tratta di un sottospazio, consideriamo un vettore $v \in V_\lambda(f)$ (cioè $f(v) = \lambda v$) e uno scalare $c \in \mathbb{K}$. Allora

$$f(cv) = cf(v) = c(\lambda v) = (c\lambda)v = (\lambda c)v = \lambda(cv)$$

(nella prima uguaglianza abbiamo sfruttato il fatto che f è un'applicazione lineare), e quindi vediamo che anche il vettore cv verifica la condizione per appartenere a $V_\lambda(f)$, il che conclude la dimostrazione.

Vediamo ora alcuni esempi di calcolo di autovettori di matrici di ordine 3.

ESEMPIO 1

Esempio 5.7. Sia

$$A = \begin{pmatrix} 5 & 1 & -1 \\ 4 & 5 & -2 \\ 2 & 1 & 2 \end{pmatrix}.$$

Calcoliamo il polinomio caratteristico. Si ha

$$\det(A - \lambda I_3) = \det \begin{pmatrix} 5-\lambda & 1 & -1 \\ 4 & 5-\lambda & -2 \\ 2 & 1 & 2-\lambda \end{pmatrix} =$$

(sviluppando secondo Laplace rispetto alla prima riga)

$$= (5-\lambda)\det\begin{pmatrix} 5-\lambda & -2 \\ 1 & 2-\lambda \end{pmatrix} - \det\begin{pmatrix} 4 & -2 \\ 2 & 2-\lambda \end{pmatrix} - \det\begin{pmatrix} 4 & 5-\lambda \\ 2 & 1 \end{pmatrix} =$$

$$(5-\lambda)[(5-\lambda)(2-\lambda)+2]-[4(2-\lambda)+4]-[4-2(5-\lambda)]$$

ovvero, svolgendo i calcoli,

$$-\lambda^3+12\lambda^2-45\lambda+54$$

Si noti che, come abbiamo anticipato sopra, il polinomio caratteristico è un polinomio di grado uguale alla dimensione del dominio dell'endomorfismo, in questo caso uguale all'ordine della matrice A (ovvero 3) in quanto una matrice $A \in M_n(\mathbb{K})$ determina un endomorfismo $L_A : \mathbb{K}^n \to \mathbb{K}^n$ di $\mathbb{K}^n$.
Per trovare le *radici* del polinomio caratteristico, ovvero le soluzioni dell'equazione caratteristica

$$-\lambda^3+12\lambda^2-45\lambda+54=0 \tag{5.8}$$

ricordiamo che, come si dimostra in algebra, se tra le soluzioni di un'equazione polinomiale a cofficienti interi vi sono numeri razionali, questi sono sicuramente della forma $\frac{a}{b}$ dove a è un divisore del termine noto e b è un divisore del coefficiente del termine di grado più alto.
Nel caso della nostra equazione (5.8), essendo il coefficiente del termine di grado più alto uguale a -1 (e non avendo quindi altri divisori oltre a ± 1) le eventuali soluzioni razionali, se esistono, sono tra i divisori del termine noto 54, ovvero ± 1, ± 2, ± 3, ± 6, ± 9, ± 18, ± 27, ± 54.
Ad esempio, sostituendo $\lambda = 1$ nella (5.8) si trova

$$-1^3+12\cdot 1^2-45\cdot 1+54=20\neq 0$$

e quindi 1 non è soluzione dell'equazione. Anche -1, come si vede subito, non è soluzione, mentre sostituendo $\lambda = 3$ si trova

$$-3^3+12\cdot 3^2-45\cdot 3+54=-27+108-135+54=0$$

Quindi $\lambda = 3$ è soluzione dell'equazione caratteristica (ed è quindi un autovalore di A). Ora, una volta trovata una soluzione di un'equazione polinomiale, le altre possono essere trovate grazie al seguente procedimento, che ci permette di ridurre l'equazione a una di grado più basso.
Il *teorema di Ruffini* afferma che se un polinomio P in λ di grado n ammette un certo valore λ_0 come radice, allora esso è divisibile per $\lambda - \lambda_0$, ovvero si ha che P si decompone come prodotto $P = (\lambda - \lambda_0)P'$ di $\lambda - \lambda_0$ per un polinomio P' di grado $n-1$ (ovvero di 1 più basso rispetto al grado di P). A questo punto, l'equazione $P = 0$ equivale a $(\lambda - \lambda_0)P' = 0$, il che, per la legge di annullamento del prodotto, può essere vero solo in due casi: o $\lambda - \lambda_0 = 0$, cioè $\lambda = \lambda_0$ (che ci dà la soluzione che già conoscevamo) oppure $P' = 0$: ci

siamo quindi ridotti a risolvere un'equazione di grado più basso rispetto a quella iniziale.
Per applicare nella pratica questo metodo, bisogna conoscere un modo per trovare il fattore P' nella decomposizione $P = (\lambda - \lambda_0)P'$.
A questo scopo, si usa un algoritmo che descriviamo tramite l'esempio della nostra equazione (5.8), per cui $P = -\lambda^3+12\lambda^2-45\lambda+54$: si inizia riportando, come nello schema seguente, i coefficienti che moltiplicano i monomi che compongono il polinomio (da quello di grado più alto a quello di grado più basso) e in basso a sinistra la radice già trovata, ovvero $\lambda = 3$:

$$\begin{array}{c|ccc|c} & -1 & 12 & -45 & 54 \\ & & & & \\ 3 & & & & \\ \hline & & & & \end{array}$$

Ora "abbassiamo", riportandolo nella riga in basso, il primo coefficiente del polinomio

$$\begin{array}{c|ccc|c} & -1 & 12 & -45 & 54 \\ & & & & \\ 3 & & & & \\ \hline & -1 & & & \end{array}$$

Moltiplichiamo il -1 così abbassato per la radice 3 e riportiamo il risultato sotto il secondo coefficiente del polinomio

$$\begin{array}{c|ccc|c} & -1 & 12 & -45 & 54 \\ & & & & \\ 3 & & -3 & & \\ \hline & -1 & & & \end{array}$$

Ora, sommiamo 12 e -3 e riportiamo il risultato nella riga in basso

$$\begin{array}{c|ccc|c} & -1 & 12 & -45 & 54 \\ & & & & \\ 3 & & -3 & & \\ \hline & -1 & 9 & & \end{array}$$

A questo punto l'algoritmo riparte: così come prima abbiamo moltiplicato -1 per la radice $\lambda = 3$, ora moltiplichiamo il 9 appena aggiunto nella riga in basso per $\lambda = 3$, scriviamo il risultato nella colonna successiva sotto il -45, sommiamo e riportiamo nella riga in basso:

	-1	12	-45	54
3		-3	27	
	-1	9	-18	

Come prevede l'algoritmo, ripartiamo: moltiplichiamo il -18 appena aggiunto nella riga in basso per $\lambda = 3$, scriviamo il risultato nella colonna successiva sotto il 54, sommiamo e riportiamo nella riga in basso:

	-1	12	-45	54
3		-3	27	- 54
	-1	9	-18	0

L'ultimo 0 comparso è la conferma del fatto che abbiamo svolto i calcoli correttamente, e i tre coefficienti $-1, 9, -18$ ottenuti prima dello zero sono proprio i coefficienti del polinomio P' (di grado 2) tale che $P = (\lambda - 3)P'$, ordinati dal termine di grado più alto fino al termine noto, ovvero

$$-\lambda^3 + 12\lambda^2 - 45\lambda + 54 = (\lambda - 3)(-\lambda^2 + 9\lambda - 18)$$

A questo punto, come abbiamo detto sopra, l'equazione caratteristica ha come soluzioni, oltre a $\lambda = 3$, anche i λ che annullano l'altro fattore, ovvero tali che $-\lambda^2 + 9\lambda - 18 = 0$.
Risolvendo con la formula risolutiva per le equazioni di secondo grado si trova che questa equazione ha come soluzioni $\lambda = 3$ e $\lambda = 6$.
Quindi gli autovalori della matrice A sono $\lambda = 3$ e $\lambda = 6$ (con $\lambda = 3$ che si ripete due volte come soluzione dell'equazione).

Possiamo ora determinare gli autovettori corrispondenti, risolvendo il sistema omogeneo che ha $A - \lambda I_3$ come matrice dei coefficienti, prima con $\lambda = 3$ e poi con $\lambda = 6$.

Per $\lambda = 3$ si ha $A - \lambda I_3 = \begin{pmatrix} 2 & 1 & -1 \\ 4 & 2 & -2 \\ 2 & 1 & -1 \end{pmatrix}$.

Riducendo tale matrice a gradini[3] si trova subito

$$\begin{pmatrix} 2 & 1 & -1 \\ 4 & 2 & -2 \\ 2 & 1 & -1 \end{pmatrix} \underset{\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - R_1}}{\longrightarrow} \begin{pmatrix} 2 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

[3]Ma basta anche osservare che la seconda e la terza riga sono dipendenti dalla prima.

e quindi il sistema $(A-\lambda I_3)x = 0$ si riduce all'unica equazione $2x_1+x_2-x_3 = 0$. Posti allora $x_2 = t$ e $x_3 = s$, si trova $2x_1 = -t+s$, ovvero $x_1 = \frac{-t+s}{2}$. Le soluzioni del sistema, ovvero l'autospazio relativo all'autovalore $\lambda = 3$, sono quindi tutte le terne del tipo

$$\left(\frac{-t+s}{2}, t, s\right)$$

al variare di $t, s \in \mathbb{R}$.

Analogamente, per $\lambda = 6$ si ha $A - \lambda I_3 = \begin{pmatrix} -1 & 1 & -1 \\ 4 & -1 & -2 \\ 2 & 1 & -4 \end{pmatrix}$.

Riducendo tale matrice a gradini si trova

$$\begin{pmatrix} -1 & 1 & -1 \\ 4 & -1 & -2 \\ 2 & 1 & -4 \end{pmatrix} \underset{\substack{R_2 \to R_2+4R_1 \\ R_3 \to R_3+2R_1}}{\longrightarrow} \begin{pmatrix} -1 & 1 & -1 \\ 0 & 3 & -6 \\ 0 & 3 & -6 \end{pmatrix} \underset{R_3 \to R_3-R_2}{\longrightarrow} \begin{pmatrix} -1 & 1 & -1 \\ 0 & 3 & -6 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo anche l'ultima riga per 3, vediamo che il sistema $(A - \lambda I_3)x = 0$ si riduce a $\begin{cases} -x_1 + x_2 - x_3 = 0 \\ x_2 - 2x_3 = 0 \end{cases}$ Posto $x_3 = t$, dalla seconda equazione si trova $x_2 = 2t$, e sostituendo nella prima $-x_1 + 2t - t = 0$, ovvero $x_1 = t$. Le soluzioni del sistema, ovvero l'autospazio relativo all'autovalore $\lambda = 3$, sono quindi tutte le terne del tipo

$$(t, 2t, t)$$

al variare di $t \in \mathbb{R}$.

Questo esempio mostra che un autovalore λ_0 può presentarsi più volte come soluzione dell'equazione caratteristica: diamo allora la seguente, importante

Definizione 5.8. Il numero di volte che l'autovalore λ_0 di un endomorfismo f compare come soluzione del polinomio caratteristico si dice *molteplicità algebrica di* λ_0; si chiama invece *molteplicità geometrica di* λ_0 la dimensione dell'autospazio $V_{\lambda_0}(f)$ relativo a λ_0.

Osservazione 5.9. Dal momento che, come è noto dall'algebra, un polinomio in λ ammette una radice λ_0 se e solo se nella sua scomposizione in fattori compare il fattore di primo grado $(\lambda_0 - \lambda)$ (o $(\lambda - \lambda_0)$, il segno è indifferente), si ha che λ_0 compare k volte come radice del polinomio caratteristico se e solo se nella sua scomposizione in fattori compare il fattore $(\lambda_0 - \lambda)^k$.

Ad esempio, è facile verificare con un calcolo diretto che il polinomio $-\lambda^3 + 12\lambda^2 - 45\lambda + 54$, ovvero il polinomio caratteristico dell'esempio appena visto, in cui $\lambda = 3$ appare due volte come radice e $\lambda = 6$ una volta sola, si decompone come prodotto $(3-\lambda)^2(6-\lambda)$.

Per quello che riguarda la molteplicità geometrica di un autovalore λ_0, essa risulta invece essere uguale a $n - r$, dove n è la dimensione del dominio V dell'endomorfismo (ovvero l'ordine della matrice, se stiamo parlando di autovalori di una matrice) e r è il rango di $(A - \lambda_0 I_n)$: infatti, per definizione la molteplicità geometrica è la dimensione dell'autospazio relativo a λ_0, che si trova come insieme delle soluzioni del sistema omogeneo in n incognite $(A - \lambda_0 I_n)x = 0$. Ma come sappiamo dalla teoria dei sistemi (Proposizione 2.8), un tale sistema ha ∞^{n-r} soluzioni (dove r è appunto il rango di $(A - \lambda_0 I_n)$), che significa che servono $n-r$ parametri per descriverne la soluzione generale. Ad esempio, nell'esempio di sopra $n-r = 2$, e l'autospazio relativo all'autovalore $\lambda = 3$ è infatti dato da tutti e soli i vettori della forma $\left(\frac{-t+s}{2}, t, s\right)$ (due parametri). Decomponendo allora tale vettore come somma della parte che contiene t e di quella che contiene s

$$\left(\frac{-t+s}{2}, t, s\right) = \left(\frac{-t}{2}, t, 0\right) + \left(\frac{s}{2}, 0, s\right) = t\left(\frac{-1}{2}, 1, 0\right) + s\left(\frac{1}{2}, 0, 1\right)$$

vediamo che $\left(\frac{-1}{2}, 1, 0\right)$ e $\left(\frac{1}{2}, 0, 1\right)$ costituiscono una base dell'autospazio V_3, che risulta quindi avere proprio dimensione $n - r = 2$.
Analogamente, sempre nello stesso esempio, l'autospazio relativo all'autovalore $\lambda = 6$ è dato da tutte e sole le terne del tipo $(t, 2t, t)$, e quindi, poiché

$$(t, 2t, t) = t(1, 2, 1)$$

una base dell'autospazio V_6 è data dall'unico vettore $(1, 2, 1)$, e quindi la dimensione di V_6 (ovvero la molteplicità geometrica) è 1.

Vediamo ora un esempio in cui la molteplicità geometrica e geometrica sono diverse.

Esempio 5.10. Sia

ESEMPIO 2

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

Calcoliamo il polinomio caratteristico. Si ha

$$\det(A-\lambda I_3)=\det\begin{pmatrix} 3-\lambda & \frac{1}{2} & 2\\ -1 & 1-\lambda & -1\\ -1 & -\frac{1}{2} & -\lambda\end{pmatrix}=$$

(sviluppando secondo Laplace rispetto alla prima riga)

$$=(3-\lambda)\det\begin{pmatrix} 1-\lambda & -1\\ -\frac{1}{2} & -\lambda\end{pmatrix}-\frac{1}{2}\det\begin{pmatrix} -1 & -1\\ -1 & -\lambda\end{pmatrix}+2\det\begin{pmatrix} -1 & 1-\lambda\\ -1 & -\frac{1}{2}\end{pmatrix}=$$

(svolgendo i calcoli)

$$-\lambda^3+4\lambda^2-5\lambda+2$$

Come abbiamo spiegato nell'esempio precedente, le eventuali radici razionali vanno cercate tra i divisori del termine noto fratto i divisori del termine di grado più alto, quindi $\pm 1, \pm 2$.
Sostituendo $\lambda = 1$ si trova

$$-1+4+2-5=0$$

ovvero $\lambda = 1$ è soluzione dell'equazione caratteristica (ed è quindi un autovalore di A). Per determinare le altre soluzioni, applichiamo l'algoritmo descritto nell'esempio precedente: disponiamo nello schema già visto i coefficienti che moltiplicano i monomi che compongono il polinomio, da quello di grado più alto a quello di grado più basso, e in basso a sinistra la radice già trovata, ovvero $\lambda = 1$:

$$\begin{array}{c|ccc|c} & -1 & 4 & -5 & 2\\ & & & & \\ 1 & & & & \\ \hline & & & & \end{array}$$

Riportiamo nella riga in basso il primo coefficiente del polinomio

$$\begin{array}{c|ccc|c} & -1 & 4 & -5 & 2\\ & & & & \\ 1 & & & & \\ \hline & -1 & & & \end{array}$$

Moltiplichiamo il -1 così abbassato per la radice 1 e riportiamo il risultato sotto il secondo coefficiente del polinomio

$$\begin{array}{c|ccc|c} & \text{-1} & 4 & \text{-5} & 2 \\ & & & & \\ 1 & & \text{-1} & & \\ \hline & \text{-1} & & & \end{array}$$

Ora, sommiamo 4 e -1 e riportiamo il risultato nella riga in basso

$$\begin{array}{c|ccc|c} & \text{-1} & 4 & \text{-5} & 2 \\ & & & & \\ 1 & & \text{-1} & & \\ \hline & \text{-1} & 3 & & \end{array}$$

Continuiamo ad applicare l'algoritmo come descritto nell'esempio precedente:

$$\begin{array}{c|ccc|c} & \text{-1} & 4 & \text{-5} & 2 \\ & & & & \\ 1 & & \text{-1} & & \\ \hline & \text{-1} & 3 & & \end{array} \longrightarrow \begin{array}{c|ccc|c} & \text{-1} & 4 & \text{-5} & 2 \\ & & & & \\ 1 & & \text{-1} & 3 & \\ \hline & \text{-1} & 3 & & \end{array} \longrightarrow \begin{array}{c|ccc|c} & \text{-1} & 4 & \text{-5} & 2 \\ & & & & \\ 1 & & \text{-1} & 3 & \\ \hline & \text{-1} & 3 & \text{-2} & \end{array} \longrightarrow \begin{array}{c|ccc|c} & \text{-1} & 4 & \text{-5} & 2 \\ & & & & \\ 1 & & \text{-1} & 3 & \text{-2} \\ \hline & \text{-1} & 3 & \text{-2} & 0 \end{array}$$

Come abbiamo detto, l'ultimo 0 comparso è la conferma del fatto che abbiamo svolto i calcoli correttamente, e i tre coefficienti $-1, 3, -2$ ottenuti prima dello zero sono proprio i coefficienti del polinomio P' di grado 2 tale che $P = (\lambda - 1)P'$, ordinati dal termine di grado più alto fino al termine noto, ovvero

$$-\lambda^3 + 4\lambda^2 - 5\lambda + 2 = (\lambda - 1)(-\lambda^2 + 3\lambda - 2)$$

A questo punto per trovare le altre radici del polinomio dobbiamo risolvere $-\lambda^2 + 3\lambda - 2 = 0$: risolvendo con la formula risolutiva per le equazioni di secondo grado si trova che questa equazione ha come soluzioni $\lambda = 1$ e $\lambda = 2$. Concludiamo che gli autovalori della matrice A sono $\lambda = 2$ e $\lambda = 1$, con quest'ultimo che si ripete quindi due volte come soluzione dell'equazione (in base alla Definizione 5.8, esso ha quindi molteplicità algebrica 2).

Per determinare gli autovettori corrispondenti, dobbiamo risolvere il sistema omogeneo che ha $A - \lambda I_3$ come matrice dei coefficienti, prima con $\lambda = 1$ e poi con $\lambda = 2$.

Per $\lambda = 1$ si ha $A - \lambda I_3 = \begin{pmatrix} 2 & \frac{1}{2} & 2 \\ -1 & 0 & -1 \\ -1 & -\frac{1}{2} & -1 \end{pmatrix}$.

Riducendo tale matrice a gradini si trova subito

$$\begin{pmatrix} 2 & \frac{1}{2} & 2 \\ -1 & 0 & -1 \\ -1 & -\frac{1}{2} & -1 \end{pmatrix} \xrightarrow[R_3 \to 2R_3 + R_1]{R_2 \to 2R_2 + R_1} \begin{pmatrix} 2 & \frac{1}{2} & 2 \\ 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 \end{pmatrix} \xrightarrow[R_3 \to R_3 + R_2]{} \begin{pmatrix} 2 & \frac{1}{2} & 2 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

quindi il sistema $(A - \lambda I_3)x = 0$ si riduce a $\begin{cases} 2x_1 + \frac{1}{2}x_2 + 2x_3 = 0 \\ \frac{1}{2}x_2 = 0 \end{cases}$ La seconda equazione ci dice che $x_2 = 0$; posto $x_3 = t$, dalla prima equazione si trova $2x_1 + 2t = 0$, ovvero $x_1 = -t$. Le soluzioni del sistema, ovvero gli autovettori di A relativi all'autovalore $\lambda = 1$, sono quindi tutte le terne del tipo

$$(-t, 0, t)$$

al variare di $t \in \mathbb{R}$. Si noti quindi che benché la molteplicità algebrica di $\lambda = 1$ sia uguale a 2, la sua molteplicità geometrica è uguale a 1.
Completiamo l'esempio ricavando i restanti autovettori: per $\lambda = 2$ si ha invece $A - \lambda I_3 = \begin{pmatrix} 1 & \frac{1}{2} & 2 \\ -1 & -1 & -1 \\ -1 & -\frac{1}{2} & -2 \end{pmatrix}$.

Riduciamo a gradini:

$$\begin{pmatrix} 1 & \frac{1}{2} & 2 \\ -1 & -1 & -1 \\ -1 & -\frac{1}{2} & -2 \end{pmatrix} \xrightarrow[R_3 \to R_3 + R_1]{R_2 \to R_2 + R_1} \begin{pmatrix} 1 & \frac{1}{2} & 2 \\ 0 & -\frac{1}{2} & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

quindi il sistema $(A - \lambda I_3)x = 0$ si riduce a $\begin{cases} x_1 + \frac{1}{2}x_2 + 2x_3 = 0 \\ -\frac{1}{2}x_2 + x_3 = 0 \end{cases}$ Posto $x_3 = t$, dalla seconda equazione si trova $-\frac{1}{2}x_2 + t = 0$, ovvero $x_2 = 2t$; sostituendo nella prima, si ha $x_1 + t + 2t = 0$, da cui $x_1 = -3t$. Le soluzioni del sistema, ovvero l'autospazio relativo all'autovalore $\lambda = 2$, è dato quindi tutte le terne del tipo

$$(-3t, 2t, t)$$

al variare di $t \in \mathbb{R}$.

Vediamo ora un ultimo esempio di una matrice di ordine 3 in cui tutte le molteplicità algebriche sono uguali a 1.

Esempio 5.11. Sia

ESEMPIO 3

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

Calcoliamo il polinomio caratteristico. Dopo un calcolo diretto si ha

$$\det(A-\lambda I_3)=\det\begin{pmatrix} 4-\lambda & 3 & -3 \\ -3 & -4-\lambda & 3 \\ -1 & -3 & 2-\lambda \end{pmatrix}=-\lambda^3+2\lambda^2+\lambda-2$$

Come si vede subito con una sostituzione, $\lambda=1$ è radice del polinomio. Applichiamo l'algoritmo visto negli esempi precedenti per ottenere le altre radici:

$$\begin{array}{c|ccc|c} & -1 & 2 & 1 & -2 \\ & & & & \\ 1 & & & & \\ \hline & & & & \end{array} \longrightarrow \begin{array}{c|ccc|c} & -1 & 2 & 1 & -2 \\ & & & & \\ 1 & & & & \\ \hline & -1 & & & \end{array} \longrightarrow \begin{array}{c|ccc|c} & -1 & 2 & 1 & -2 \\ & & & & \\ 1 & & -1 & & \\ \hline & -1 & 1 & & \end{array}$$

$$\longrightarrow \begin{array}{c|ccc|c} & -1 & 2 & 1 & -2 \\ & & & & \\ 1 & & -1 & 1 & \\ \hline & -1 & 1 & 2 & \end{array} \longrightarrow \begin{array}{c|ccc|c} & -1 & 2 & 1 & -2 \\ & & & & \\ 1 & & -1 & 1 & 2 \\ \hline & -1 & 1 & 2 & 0 \end{array}$$

Quindi

$$-\lambda^3+2\lambda^2+\lambda-2=(\lambda-1)(-\lambda^2+\lambda+2)$$

Poiché, come si verifica subito mediante la formula risolutiva, le soluzioni di $-\lambda^2+3\lambda-2=0$ sono $\lambda=2$ e $\lambda=-1$, concludiamo che gli autovalori della matrice A sono $\lambda=1$, $\lambda=2$ e $\lambda=-1$, tutti e tre con molteplicità algebrica 1.

Concludiamo allora l'esercizio calcolando gli autovettori corrispondenti.
Per $\lambda=1$ si ha

$$A-\lambda I_3=\begin{pmatrix} 3 & 3 & -3 \\ -3 & -5 & 3 \\ -1 & -3 & 1 \end{pmatrix} \underset{\substack{R_2\to R_2+R_1 \\ R_3\to 3R_3+R_1}}{\longrightarrow} \begin{pmatrix} 3 & 3 & -3 \\ 0 & -2 & 0 \\ 0 & -6 & 0 \end{pmatrix} \underset{R_3\to R_3-3R_2}{\longrightarrow} \begin{pmatrix} 3 & 3 & -3 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo anche la prima riga per 3, si vede che il sistema $(A-\lambda I_3)x=0$ si riduce a $\begin{cases} x_1+x_2-x_3=0 \\ -2x_2=0 \end{cases}$ La seconda equazione ci dice che $x_2=0$; posto $x_3=t$, dalla prima equazione si trova $x_1-t=0$, ovvero $x_1=t$. Le soluzioni del sistema, ovvero l'autospazio relativo all'autovalore $\lambda=1$, è dato quindi tutte le terne del tipo

$$(t, 0, t)\,, \quad t \in \mathbb{R}.$$

Per $\lambda = 2$ si ha

$$A-\lambda I_3 = \begin{pmatrix} 2 & 3 & -3 \\ -3 & -6 & 3 \\ -1 & -3 & 0 \end{pmatrix} \underset{\substack{R_2 \to 2R_2+3R_1 \\ R_3 \to 2R_3+R_1}}{\longrightarrow} \begin{pmatrix} 2 & 3 & -3 \\ 0 & -3 & -3 \\ 0 & -3 & -3 \end{pmatrix} \underset{R_3 \to R_3 - R_2}{\longrightarrow} \begin{pmatrix} 2 & 3 & -3 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo anche la seconda riga per -3, si vede che il sistema $(A - \lambda I_3)x = 0$ si riduce a $\begin{cases} 2x_1 + 3x_2 - 3x_3 = 0 \\ x_2 + x_3 = 0 \end{cases}$ posto $x_3 = t$, la seconda equazione dice che $x_2 = -t$, e sostituendo nella prima si trova $2x_1 - 3t - 3t = 0$, da cui $x_1 = 3t$. Le soluzioni del sistema, ovvero l'autospazio relativo all'autovalore $\lambda = 1$, è dato quindi tutte le terne del tipo

$$(3t, -t, t)\,, \quad t \in \mathbb{R}$$

Infine, per $\lambda = -1$ si ha

$$A-\lambda I_3 = \begin{pmatrix} 5 & 3 & -3 \\ -3 & -3 & 3 \\ -1 & -3 & 3 \end{pmatrix} \underset{\substack{R_2 \to 5R_2+3R_1 \\ R_3 \to 5R_3+R_1}}{\longrightarrow} \begin{pmatrix} 5 & 3 & -3 \\ 0 & -6 & 6 \\ 0 & -12 & 12 \end{pmatrix} \underset{R_3 \to R_3 - 2R_2}{\longrightarrow} \begin{pmatrix} 5 & 3 & -3 \\ 0 & -6 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo anche la seconda riga per 6, si vede che il sistema $(A - \lambda I_3)x = 0$ si riduce a $\begin{cases} 5x_1 + 3x_2 - 3x_3 = 0 \\ -x_2 + x_3 = 0 \end{cases}$ posto $x_3 = t$, la seconda equazione dice che $x_2 = t$, e sostituendo nella prima si trova $5x_1 + 3t - 3t = 0$, da cui $x_1 = 0$. Le soluzioni del sistema, ovvero l'autospazio relativo all'autovalore $\lambda = -1$, è dato quindi tutte le terne del tipo

$$(0, t, t)\,, \quad t \in \mathbb{R}.$$

Osservazione 5.12. Notiamo che quando riduciamo a gradini la matrice $A - \lambda I_n$ dopo aver sostituito al posto di λ un autovalore λ_0 allo scopo di risolvere il sistema omogeneo $(A - \lambda_0 I_n)x = \bar{0}$ e trovare l'autospazio relativo a λ_0, in seguito alla riduzione si deve sempre annullare almeno una riga. Infatti, λ_0 è stato trovato risolvendo l'equazione caratteristica, ovvero soddisfa $\det(A - \lambda_0 I_n) = 0$, ma il determinante di una matrice si annulla se e solo se il suo rango non è massimo: quindi $(A - \lambda_0 I_n)$ non ha rango n, e questo a sua volta come sappiamo implica che nella sua riduzione a gradini si annulli almeno una riga.

Del resto, si osservi che se dopo la riduzione a gradini nella matrice $(A-\lambda_0 I_n)$ rimanessero n righe non nulle, il sistema omogeneo corrispondente (che ha n incognite) avrebbe una sola soluzione, quella nulla, il che non è possibile perché se λ_0 è un autovalore deve esistere un autovettore ad esso associato, che per definizione è un vettore *non nullo*.

5.3 Endomorfismi e matrici diagonalizzabili

Diamo la seguente, importante

Definizione 5.13. Un endomorfismo $f : V \to V$ si dice *diagonalizzabile* se esiste una base di V formata da autovettori di f (equivalentemente, se esistono n autovettori indipendenti, essendo $n = \dim(V)$).
Un matrice quadrata $A \in M_n(\mathbb{K})$ si dice diagonalizzabile se lo è l'endomorfismo $L_A : \mathbb{K}^n \to \mathbb{K}^n$ determinato da A (ovvero equivalentemente se A ammette n autovettori indipendenti).

Il motivo per cui gli endomorfismi che soddisfano la condizione della Definizione 5.13 si dicono diagonalizzabili è il seguente: se $v_1, v_2, \dots, v_n$ sono autovettori che formano una base B di V, con $f(v_1) = \lambda_1 v_1$, $f(v_2) = \lambda_2 v_2$, $\dots$, $f(v_n) = \lambda_n v_n$ allora la matrice $M_B(f)$ associata a f rispetto a B è

$$M_B(f) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

ovvero è un matrice diagonale.
Infatti, per definizione di matrice associata nelle colonne di $M_B(f)$ ci sono le coordinate delle immagini $f(v_1), f(v_2), \dots, f(v_n)$ dei vettori della base B, calcolate rispetto a B stessa: ora, $f(v_1) = \lambda_1 v_1 = \lambda_1 v_1 + 0v_2 + \cdots + 0v_n$, quindi le coordinate di $f(v_1)$ rispetto a B sono $(\lambda_1, 0, \dots, 0)$; analogamente, $f(v_2) = \lambda_2 v_2 = 0v_1 + \lambda_2 v_2 + \cdots + 0v_n$, quindi le coordinate da mettere nella seconda colonna sono $(0, \lambda_2, \dots, 0)$, e così via fino a $f(v_n) = \lambda_n v_n = 0v_1 + 0v_2 + \cdots + \lambda_n v_n$, che ci dice che nell'ultima colonna dobbiamo mettere $(0, 0, \dots, \lambda_n)$.

Vedremo alla fine del paragrafo (Esempio 5.26) l'utilità e alcune applicazioni delle matrici diagonalizzabili. Prima, vediamo come determinare se un endomorfismo $f : V \to V$ è diagonalizzabile.
Per costruire una base di V formata da autovettori di f, dobbiamo chiaramente prendere gli autovettori dagli autospazi e sperare di riuscire ad arrivare ad avere $n = \dim(V)$ autovettori indipendenti. Ora, ogni autospazio

ci fornisce al massimo un numero di autovettori indipendenti uguale alla sua dimensione: potremmo quindi ad esempio prendere una base di ogni autospazio e mettere insieme tali basi. Questo ci darà una base di V formata da autovettori se saranno soddisfatte le due seguenti condizioni:

(i) i vettori ottenuti mettendo insieme le basi degli autospazi sono indipendenti;

(ii) l'unione delle basi degli autospazi ha $n = \dim(V)$ vettori, condizione verificata se e solo se

$$\dim(V_{\lambda_1}(f)) + \cdots + \dim(V_{\lambda_k}(f)) = \dim(V) \tag{5.9}$$

dove $V_{\lambda_1}(f), V_{\lambda_2}(f), \ldots, V_{\lambda_k}(f)$ sono gli autospazi degli autovalori $\lambda_1, \ldots, \lambda_k$ di f.

Vediamo ora che la prima condizione è sempre automaticamente verificata, come affermato dalla seguente

Proposizione 5.14. Siano $\lambda_1, \ldots, \lambda_k$ gli autovalori di un endomorfismo $f : V \to V$ e siano $V_{\lambda_1}(f), V_{\lambda_2}(f), \ldots, V_{\lambda_k}(f)$ i corrispondenti autospazi.
Se $\{v_1^1, \ldots, v_{n(1)}^1\}$ denota una base di $V_{\lambda_1}(f)$, $\{v_1^2, \ldots, v_{n(2)}^2\}$ una base di $V_{\lambda_2}(f)$ e così via fino a $\{v_1^k, \ldots, v_{n(k)}^k\}$ base di $V_{\lambda_k}(f)$, allora i vettori

$$v_1^1, \ldots, v_{n(1)}^1, v_1^2, \ldots, v_{n(2)}^2, \ldots, v_1^k, \ldots, v_{n(k)}^k$$

che si ottengono mettendo insieme tali basi, sono linearmente indipendenti.

Tale proposizione risulta essere un corollario del seguente risultato, e la dimostreremo dopo di esso.

Lemma 5.15. Sia $f : V \to V$ e siano $v_1, \ldots, v_m$ autovettori relativi a autovalori $\lambda_1, \ldots, \lambda_m$ tutti diversi tra loro. Allora $v_1, \ldots, v_m$ sono linearmente indipendenti.

Proof. Dimostreremo il lemma *per induzione sul numero m dei vettori* $v_1, \ldots, v_m$. Ricordiamo che quando un teorema afferma che una certa formula o un certo fatto valgono per ogni numero naturale m (in questo caso, stiamo affermando che i vettori $v_1, \ldots, v_m$ sono indipendenti qualunque sia m), esso può essere dimostrato provando i due seguenti fatti:

(1) l'affermazione vale per almeno un numero naturale (solitamente 0 o 1)

(2) ogniqualvolta l'affermazione vale per un certo numero m, allora vale per il suo successivo $m+1$

Questi due fatti, insieme, garantiscono che l'affermazione è vera per ogni numero naturale maggiore o uguale di quello del punto (1) (quindi, se nel punto (1) il numero è 0, l'affermazione sarà vera per ogni numero naturale)[4].

Applichiamo ora tale metodo per dimostrare il Lemma 5.15: iniziamo col dimostrare la (1), e più precisamente mostriamo che l'affermazione del lemma vale per il numero naturale $m = 2$, cioè mostriamo che se ho *due* autovettori v_1, v_2 relativi a autovalori λ_1, λ_2 distinti, allora v_1 e v_2 sono indipendenti.
Ma questo è vero perchè, se v_1 e v_2 fossero dipendenti, essi sarebbero uno multiplo dell'altro, diciamo $v_2 = cv_1$ per qualche $c \in \mathbb{K}$: poichè v_1 appartiene all'autospazio $V_{\lambda_1}(f)$, e questo come abbiamo dimostrato sopra è chiuso rispetto al prodotto per scalari (in quanto sottospazio), allora anche v_2 dovrebbe appartenere a $V_{\lambda_1}(f)$, ovvero essere autovettore relativo a λ_1. Ma questo contraddice il fatto che v_2 è autovettore relativo a λ_2, che è per ipotesi diverso da λ_1.
Ora, dimostriamo la (2): supponendo che l'affermazione sia vera per un certo numero m di vettori (quindi supponendo che m autovettori relativi a autovalori distinti siano sempre indipendenti) dimostriamo che essa vale allora per $m+1$, ovvero $m+1$ autovettori $v_1, v_2, \ldots, v_{m+1}$ relativi a autovalori distinti $\lambda_1, \lambda_2, \ldots, \lambda_{m+1}$ sono indipendenti.
Per dimostrare l'indipendenza lineare di $v_1, v_2, \ldots, v_{m+1}$, dobbiamo far vedere che l'unica combinazione lineare $c_1v_1+c_2v_2+\cdots+c_{m+1}v_{m+1}$ uguale al vettore nullo $\bar{0}$ è quella con $c_1 = 0, c_2 = 0, \ldots, c_{m+1} = 0$.
In effetti, se $c_1v_1+c_2v_2+\cdots+c_{m+1}v_{m+1} = \bar{0}$, applicando a entrambi i membri l'endomorfismo f si ha $f(c_1v_1 + c_2v_2 + \cdots + c_{m+1}v_{m+1}) = f(\bar{0}) = \bar{0}$, ovvero, per la linearità di f,

$$c_1f(v_1) + c_2f(v_2) + \cdots + c_{m+1}f(v_{m+1}) = \bar{0}.$$

Ma $v_1, v_2 \ldots, v_{m+1}$ sono autovettori relativi rispettivamente agli autovalori $\lambda_1, \lambda_2, \ldots, \lambda_{m+1}$, e quindi questa uguaglianza può essere riscritta

$$c_1\lambda_1v_1 + c_2\lambda_2v_2 + \cdots + c_{m+1}\lambda_{m+1}v_{m+1} = \bar{0}. \tag{5.10}$$

[4] Intuitivamente, quello che succede è che se da (1) sappiamo che l'affermazione vale ad esempio per $m = 0$, applicando la (2) sappiamo che essa deve valere per il suo successivo, cioè $m = 1$; ma se vale per $m = 1$, sempre dalla (2) deduciamo che vale per il suo successivo $m = 2$, e così via. Rigorosamente, il fatto che questo meccanismo a catena impichi che l'affermazione è vera per l'insieme infinito di tutti i numeri naturali dipende dagli assiomi che definiscono i numeri naturali stessi, ma questo va oltre lo scopo di questo corso.

Ora, se invece moltiplichiamo primo e secondo membro della $c_1v_1 + c_2v_2 + \cdots + c_{m+1}v_{m+1} = \bar{0}$ per λ_{m+1}, otteniamo

$$c_1\lambda_{m+1}v_1 + c_2\lambda_{m+1}v_2 + \cdots + c_{m+1}\lambda_{m+1}v_{m+1} = \bar{0}. \tag{5.11}$$

Confrontiamo la (5.10) e la (5.11): se sottraiamo i rispettivi primi membri, otteniamo

$$c_1\lambda_1v_1+c_2\lambda_2v_2+\cdots+c_{m+1}\lambda_{m+1}v_{m+1}-c_1\lambda_{m+1}v_1-c_2\lambda_{m+1}v_2-\cdots-c_{m+1}\lambda_{m+1}v_{m+1} = \bar{0}$$

ovvero, semplificando il termine $c_{m+1}\lambda_{m+1}v_{m+1}$ e raccogliendo i vettori $v_1, \ldots, v_m$ negli addendi rimanenti,

$$c_1(\lambda_1 - \lambda_{m+1})v_1 + \cdots + c_m(\lambda_m - \lambda_{m+1})v_m = \bar{0}. \tag{5.12}$$

Ora, quest'ultima è una combinazione lineare degli m vettori $v_1, \ldots, v_m$ uguale al vettore nullo: ma noi stiamo supponendo che il teorema valga ogniqualvolta abbiamo m autovettori relativi a autovalori distinti, e quindi ad esempio proprio per $v_1, \ldots, v_m$: quindi $v_1, \ldots, v_m$ sono indipendenti, e la (5.12) può essere vera solo se i coefficienti $c_1(\lambda_1 - \lambda_{m+1}), \ldots, c_m(\lambda_m - \lambda_{m+1})$ sono nulli.
Ma essendo gli autovalori distinti, si ha $\lambda_1 \neq \lambda_{m+1}, \ldots, \lambda_m \neq \lambda_{m+1}$ e quindi $(\lambda_1 - \lambda_{m+1}) \neq 0, \ldots, (\lambda_m - \lambda_{m+1}) \neq 0$: ne segue che devono essere nulli $c_1, \ldots, c_m$.
Quindi la $c_1v_1 + c_2v_2 + \cdots + c_{m+1}v_{m+1} = \bar{0}$ si riduce a $c_{m+1}v_{m+1} = \bar{0}$, che, essendo $v_{m+1} \neq \bar{0}$ (gli autovettori sono per definizione non nulli), implica $c_{m+1} = 0$.
Avendo dimostrato che tutti i coefficienti $c_1, c_2, \ldots, c_{m+1} = 0$ della $c_1v_1 + c_2v_2+\cdots+c_{m+1}v_{m+1} = \bar{0}$ sono nulli, concludiamo che i vettori $v_1, v_2, \ldots, v_{m+1}$ sono indipendenti: l'affermazione del teorema vale quindi per $m + 1$ vettori, come volevamo. □

Dimostrazione della Proposizione 5.14: Dobbiamo dimostrare che i vettori

NO

$$v_1^1, \ldots, v_{n(1)}^1, v_1^2, \ldots, v_{n(2)}^2, \ldots, v_1^k, \ldots, v_{n(k)}^k$$

sono indipendenti, ovvero che se

$$c_1^1v_1^1 + \cdots + c_{n(1)}^1v_{n(1)}^1 + c_1^2v_1^2 + \cdots + c_{n(2)}^2v_{n(2)}^2 + \cdots + c_1^kv_1^k + \cdots + c_{n(k)}^kv_{n(k)}^k = \bar{0}. \tag{5.13}$$

allora i coefficienti $c_1^1, \ldots, c_{n(1)}^1, c_1^2, \ldots, c_{n(2)}^2, \ldots, c_1^k, \ldots, c_{n(k)}^k$ della combinazione sono tutti uguali a zero.

NO

Si noti che a tale scopo non possiamo usare direttamente il Lemma 5.15, in quanto non si tratta di autovettori associati ad autovalori distinti (i vettori $v_1^1, \ldots, v_{n(1)}^1$ sono tutti associati allo stesso autovalore λ_1, i vettori $v_1^2, \ldots, v_{n(2)}^2$ a λ_2, e così via). Se però denotiamo

$$v_1 = c_1^1 v_1^1 + \cdots + c_{n(1)}^1 v_{n(1)}^1,$$

$$v_2 = c_1^2 v_1^2 + \cdots + c_{n(2)}^2 v_{n(2)}^2,$$

$$\vdots$$

$$v_k = c_1^k v_1^k + \cdots + c_{n(k)}^k v_{n(k)}^k$$

e sostituendo riscriviamo quindi l'uguaglianza (5.13) come

$$v_1 + v_2 + \cdots + v_k = \bar{0}. \tag{5.14}$$

allora possiamo applicare il Lemma 5.15 a $v_1, \ldots, v_k$: infatti, da una parte, essendo $v_1^1, \ldots, v_{n(1)}^1$ una base dell'autospazio $V_{\lambda_1}(f)$, il vettore $v_1 = c_1^1 v_1^1 + \cdots + c_{n(1)}^1 v_{n(1)}^1$ appartiene anch'esso a $V_{\lambda_1}(f)$ e, se non è nullo, è un autovettore relativo a λ_1; analogamente, se v_2 non è nullo è un autovettore relativo a λ_2, e così via.

D'altra parte, l'uguaglianza (5.14) ci dice che i vettori $v_1, v_2, \ldots, v_k$ sono dipendenti, il che sarebbe assurdo se non fossero tutti nulli, in quanto in tal caso sarebbero autovettori relativi a autovalori distinti e per il Lemma 5.15 devono essere indipendenti.
Quindi si deve avere $v_1 = \bar{0}, v_2 = \bar{0}, \ldots, v_k = \bar{0}$, ovvero

$$c_1^1 v_1^1 + \cdots + c_{n(1)}^1 v_{n(1)}^1 = \bar{0}$$

$$c_1^2 v_1^2 + \cdots + c_{n(2)}^2 v_{n(2)}^2 = \bar{0}$$

$$\vdots$$

$$c_1^k v_1^k + \cdots + c_{n(k)}^k v_{n(k)}^k = \bar{0}$$

Ma la prima di queste uguaglianze implica che $c_1^1 = 0, \ldots, c_{n(1)}^1 = 0$, in quanto $v_1^1, \ldots, v_{n(1)}^1$ formano una base (dell'autospazio $V_{\lambda_1}(f)$) e quindi sono indipendenti, e analogamente le altre uguaglianze implicano l'annullarsi di tutti gli altri coefficienti $c_1^2, \ldots, c_{n(2)}^2, \ldots, c_1^k, \ldots, c_{n(k)}^k$, come volevamo. $\square$

Grazie alla Proposizione 5.14, vediamo quindi che perché f sia diagonalizzabile è necessario e sufficiente che sia verificata la (5.9): quindi la diagonalizzabilità di un endomorfismo dipende solo dalle dimensioni $\dim(V_{\lambda_i}(f))$ dei singoli autospazi, che devono fornire più autovettori possibile. Diventa allora importante stimare queste dimensioni.

Si ha il seguente, importante

Teorema 5.16. Sia $f : V \to V$ un endomorfismo. Per ogni autovalore λ_0 di f, si ha $\dim(V_{\lambda_0}(f)) \leq m_{alg}(\lambda_0)$ (dove $m_{alg}(\lambda_0)$ denota la molteplicità algebrica di λ_0, definita nella Definizione 5.8)

Proof. Supponiamo che la dimensione dell'autospazio $V_{\lambda_0}(f)$ sia uguale a k, e fissiamo una base $\{v_1, v_2, \dots, v_k\}$ di $V_{\lambda_0}(f)$.
Come garantisce il teorema del completamento (che abbiamo già usato anche nella dimostrazione del Teorema 4.8, si veda la nota a pagina 152), se $\dim(V) = n$, esistono sicuramente $n-k$ vettori, diciamo $v_{k+1}, \dots, v_n$, che assieme a $v_1, v_2, \dots, v_k$ formano una base $B = \{v_1, v_2 \dots, v_k, v_{k+1}, \dots, v_n\}$ di V. È rispetto a tale speciale base B che calcoleremo ora il polinomio caratteristico $\det(M_B(f) - \lambda I_n)$ (dal quale vedremo la molteplicità algebrica di λ_0 e potremo confrontarla con la dimensione dell'autospazio).
Per definizione, la matrice $M_B(f)$ associata a f rispetto a B ha come colonne le coordinate dei vettori

$$f(v_1), f(v_2), \dots, f(v_k), f(v_{k+1}), \dots, f(v_n)$$

rispetto a B. Poichè i vettori $v_1, v_2, \dots, v_k$ appartengono all'autospazio $V_{\lambda_0}(f)$, si ha

$$\begin{aligned} f(v_1) &= \lambda_0 v_1 = \lambda_0 v_1 + 0v_2 + \cdots + 0v_k + 0v_{k+1} + \cdots + 0v_n \\ f(v_2) &= \lambda_0 v_2 = 0v_1 + \lambda_0 v_2 + \cdots + 0v_k + 0v_{k+1} + \cdots + 0v_n \\ &\dots \\ f(v_k) &= \lambda_0 v_k = 0v_1 + 0v_2 + \cdots + \lambda_0 v_k + 0v_{k+1} + \cdots + 0v_n \end{aligned}$$

mentre sulle immagini $f(v_{k+1}), \dots, f(v_n)$ (e quindi sulle loro coordinate) non sappiamo dire nulla. Quindi la matrice $M_B(f)$ avrà la seguente forma

$$M_B(f) = \begin{pmatrix} \lambda_0 & 0 & \dots & 0 & * & \dots & * \\ 0 & \lambda_0 & \dots & 0 & * & \dots & * \\ & & & \dots & & & \\ 0 & 0 & \dots & \lambda_0 & * & \dots & * \\ 0 & 0 & \dots & 0 & * & \dots & * \\ & & & \dots & & & \\ 0 & 0 & \dots & 0 & * & \dots & * \end{pmatrix} \tag{5.15}$$

dove, nelle colonne successive alla k-esima, stiamo indicando con gli asterischi le entrate che non conosciamo.
Quindi

$$M_B(f) - \lambda I_n = \begin{pmatrix} \lambda_0 - \lambda & 0 & \dots & 0 & * & \dots & * \\ 0 & \lambda_0 - \lambda & \dots & 0 & * & \dots & * \\ & & & \dots & & & \\ 0 & 0 & \dots & \lambda_0 - \lambda & * & \dots & * \\ 0 & 0 & \dots & 0 & * - \lambda & \dots & * \\ & & & \dots & & & \\ 0 & 0 & \dots & 0 & * & \dots & * - \lambda \end{pmatrix} \tag{5.16}$$

Ora, il polinomio caratteristico è per definizione il determinante di questa matrice. Sviluppando secondo Laplace rispetto alla prima colonna, si trova

$$\det(M_B(f) - \lambda I_n) = (\lambda_0 - \lambda) \det \begin{pmatrix} \lambda_0 - \lambda & \dots & 0 & * & \dots & * \\ & & \dots & & & \\ 0 & \dots & \lambda_0 - \lambda & * & \dots & * \\ 0 & \dots & 0 & * - \lambda & \dots & * \\ & & \dots & & & \\ 0 & \dots & 0 & * & \dots & * - \lambda \end{pmatrix}. \tag{5.17}$$

Sviluppando ancora rispetto alla prima colonna, e continuando così fino a esaurire le prime k colonne, si arriva a

$$\det(M_B(f) - \lambda I_n) = (\lambda_0 - \lambda)^k \det \begin{pmatrix} * - \lambda & \dots & * \\ & \dots & \\ * & \dots & * - \lambda \end{pmatrix} \tag{5.18}$$

il che dimostra che nel polinomio caratteristico il fattore $(\lambda_0 - \lambda)$ compare *almeno* con un esponente k (non sappiamo se nel determinante della matrice che resta, della quale non conosciamo le entrate, tale fattore compaia ancora o no). Quindi la molteplicità algebrica di λ_0 è maggiore o uguale a k, che era la dimensione dell'autospazio $V_{\lambda_0}(f)$, ovvero la molteplicità geometrica di λ_0. La dimostrazione è conclusa. $\square$

Osservazione 5.17. Nel caso in cui un autovalore λ_0 abbia molteplicità algebrica $m_{alg}(\lambda_0) = 1$, il Teorema 5.16 implica che $\dim(V_{\lambda_0}(f)) \leq 1$. D'altra parte, la dimensione di $V_{\lambda_0}(f)$ non può essere minore di 1, cioè zero, in quanto un autospazio non può consistere del solo vettore nullo (λ_0 è un autovalore, quindi deve avere un autovettore non nullo a lui associato). Si conclude che $\dim(V_{\lambda_0}(f))$ deve in tal caso anch'essa essere uguale a 1.

Combinando il fatto che per ogni autovalore la dimensione dell'autospazio corrispondente non possa essere maggiore della molteplicità algebrica con la condizione (5.9), si ottiene il seguente criterio necessario e sufficiente per la diagonalizzabilità di un endomorfismo:

Teorema 5.18. Un endomorfismo $f : V \to V$ di uno spazio vettoriale V definito su un campo $\mathbb{K}$ è diagonalizzabile se e solo se sono soddisfatte entrambe le seguenti condizioni:

Criterio di diagonalizzabilità

(i) tutte le radici del polinomio caratteristico appartengono al campo $\mathbb{K}$

(ii) per ogni autovalore λ_0, si ha $\dim(V_{\lambda_0}(f)) = m_{alg}(\lambda_0)$

Proof. Come sappiamo dalla (5.9), la diagonalizzabilità di f equivale ad avere $\dim(V_{\lambda_1}(f))+\cdots+\dim(V_{\lambda_k}(f)) = n$, dove $\lambda_1, \ldots, \lambda_k$ sono gli autovalori (cioè le radici del polinomio caratteristico che stanno in $\mathbb{K}$) e n è la dimensione di V. Dal Teorema 5.16 segue che

$$\dim(V_{\lambda_1}(f)) + \cdots + \dim(V_{\lambda_k}(f)) \leq m_{alg}(\lambda_1) + \cdots + m_{alg}(\lambda_k). \qquad (5.19)$$

NO

Ora, usando questa disuguaglianza mostriamo che se una tra le condizioni (i) o (ii) non è verificata, allora non può essere $\dim(V_{\lambda_1}(f))+\cdots+\dim(V_{\lambda_k}(f)) = n$.

Infatti, supponiamo che la (i) non sia verificata, ovvero che non tutte le radici del polinomio caratteristico stiano nel campo $\mathbb{K}$: questo significa che il polinomio caratteristico non sarà prodotto dei fattori $(\lambda-\lambda_1)^{m_{alg}(\lambda_1)}\cdots(\lambda-\lambda_k)^{m_{alg}(\lambda_k)}$ e basta, ma ci saranno altri fattori (che si annullano per valori che sono fuori dal campo $\mathbb{K}$), quindi il grado totale n del polinomio (che coincide con la dimensione di V) sarà maggiore del contributo dato dai fattori $(\lambda-\lambda_1)^{m_{alg}(\lambda_1)}\cdots(\lambda-\lambda_k)^{m_{alg}(\lambda_k)}$, ovvero maggiore di $m_{alg}(\lambda_1)+\cdots+m_{alg}(\lambda_k)$ (si veda anche l'Esempio 5.19 per alcuni casi illustrativi).

Ma allora, $m_{alg}(\lambda_1) + \cdots + m_{alg}(\lambda_k) < n$ combinato con la (5.19) ci dice che $\dim(V_{\lambda_1}(f)) + \cdots + \dim(V_{\lambda_k}(f)) < n$, e quindi f non è diagonalizzabile.

Supponiamo ora invece che la (i) sia verificata, ovvero che $m_{alg}(\lambda_1) + \cdots + m_{alg}(\lambda_k) = n$, ma che non sia verificata la (ii), ovvero che per almeno un autovalore λ si abbia $\dim(V_\lambda(f)) < m_{alg}(\lambda)$. Questo implica che nella disuguaglianza (5.19) ci sia il segno $<$ di minore stretto, e quindi

$$\dim(V_{\lambda_1}(f)) + \cdots + \dim(V_{\lambda_k}(f)) < m_{alg}(\lambda_1) + \cdots + m_{alg}(\lambda_k) = n$$

NO

e di nuovo abbiamo che la condizione (5.9) non può essere verificata, e f non è diagonalizzabile.

Viceversa, se le condizioni (i) e (ii) sono verificate, allora $\dim(V_{\lambda_1}(f)) + \cdots + \dim(V_{\lambda_k}(f)) = m_{alg}(\lambda_1) + \cdots + m_{alg}(\lambda_k)$ e $m_{alg}(\lambda_1) + \cdots + m_{alg}(\lambda_k) = n$, che combinate ci dicono che la (5.9) è soddisfatta e f è diagonalizzabile. □

Esempio 5.19. Come esempio illustrativo della dimostrazione appena vista, se avessimo un $\mathbb{R}$-spazio vettoriale V di dimensione $n = \dim(V) = 5$ e polinomio caratteristico

$$\lambda^5 - 5\lambda^4 + 9\lambda^3 - 9\lambda^2 + 8\lambda - 4 = (\lambda - 1)(\lambda - 2)^2(\lambda^2 + 1)$$

(si ricordi che il polinomio caratteristico ha grado n uguale alla dimensione dello spazio) avremmo solo 3 radici reali contate con le loro molteplicità ($\lambda_1 = 1$ contata una volta e $\lambda_2 = 2$ contata due volte, mentre il fattore $\lambda^2 + 1$ non ha radici reali). Questo è un esempio della situazione, vista nella dimostrazione, in cui la condizione (i) non è verificata: infatti, in tal caso si ha

$$m_{alg}(\lambda_1) + m_{alg}(\lambda_2) = 1 + 2 = 3 < 5$$

Poiché ogni autovalore (ovvero ogni radice reale) può dare al massimo tanti autovettori indipendenti quanti la propria molteplicità algebrica, abbiamo che $\lambda_1 = 1$ darà un autovettore e $\lambda_2 = 2$ darà al massimo 2 autovettori: quindi non avremo mai 5 autovettori indipendenti necessari a formare una base, e l'applicazione non potrà essere diagonalizzabile.

Come esempio di situazione in cui invece la (i) è verificata, supponiamo che (sempre con $\dim(V) = n = 5$) il polinomio caratteristico sia

$$\lambda^5 - 5\lambda^4 + 7\lambda^3 + \lambda^2 - 8\lambda + 4 = (\lambda + 1)(\lambda - 1)^2(\lambda - 2)^2$$

ovvero $\lambda_1 = 1$ contata due volte, $\lambda_2 = 2$ contata due volte, $\lambda_3 = -1$ contata una volta, e quindi

$$m_{alg}(\lambda_1) + m_{alg}(\lambda_2) + m_{alg}(\lambda_3) = 2 + 2 + 1 = 5$$

Sempre dal fatto che nessun autovalore può dare più autovettori indipendenti della sua molteplicità algebrica segue che per avere una base di 5 autovettori indipendenti *ogni autovalore deve dare esattamente tanti autovettori indipendenti quanti la molteplicità algebrica* (ovvero il massimo possibile):

in altre parole, concordemente con il Teorema 5.18, la dimensione di ogni autospazio (la molteplicità geometrica) deve essere uguale alla molteplicità algebrica.
Se per esempio $\lambda_1 = 1$ ci desse un solo autovettore (meno della sua molteplicità algebrica) non riusciremmo a ottenere 5 autovettori indipendenti in tutto in quanto i restanti autovalori ci daranno al massimo tanti autovettori indipendenti quanti le loro molteplicità algebriche, cioè 2+1=3.

Osservazione 5.20. Poiché nell'Osservazione 5.17 abbiamo visto che nel caso di molteplicità algebrica uguale a 1 la condizione $m_{geom}(\lambda) = m_{alg}(\lambda)$ è automaticamente verificata, questo suggerisce che quando calcoliamo gli autospazi di un endomorfismo per capire se è diagonalizzabile conviene iniziare da quelli con molteplicità algebrica maggiore di 1, in quanto se già per uno di questi la molteplicità geometrica non è uguale a quella algebrica allora possiamo fermarci evitando di calcolare gli altri autospazi: l'endomorfismo sicuramente non sarà diagonalizzabile.

Osservazione 5.21. Un'altra conseguenza di quanto osservato sopra è il seguente criterio (sufficiente ma non necessario) di diagonalizzabilità:

Se un endomorfismo $f : V \to V$ su un $\mathbb{K}$-spazio vettoriale V di dimensione n ha esattamente n autovalori distinti tra loro, allora f è diagonalizzabile.

Infatti, dal momento che gli n autovalori sono distinti, nessuno di loro compare più di una volta nella scomposizione del polinomio, quindi le molteplicità algebriche sono tutte uguali a 1 e, come abbiamo visto nell'Osservazione 5.17, anche le molteplicità geometriche devono essere tutte uguali a 1: quindi ognuno degli n autovalori distinti darà un autovettore, e in tutto avremo n autovettori indipendenti con i quali formare una base per lo spazio V.

Esempio 5.22. La matrice A dell'Esempio 5.7 è diagonalizzabile, e una base di $\mathbb{R}^3$ formata da autovettori di A è $(-\frac{1}{2}, 1, 0)$, $(\frac{1}{2}, 0, 1), (1, 2, 1)$, ottenuta mettendo insieme la base $(-\frac{1}{2}, 1, 0)$ e $(\frac{1}{2}, 0, 1)$ dell'autospazio relativo a 3 e la base $(1, 2, 1)$ dell'autospazio relativo a 6 (si veda anche l'Osservazione 5.9)

Invece, la matrice A dell'Esempio 5.10, non è diagonalizzabile in quanto uno degli autovalori ($\lambda = 1$) ha molteplicità algebrica uguale a due mentre il suo autospazio V_1, ha base formata dal vettore $(-1, 0, 1)$, e quindi $\dim(V_1) = 1$: la molteplicità geometrica di $\lambda = 1$ è minore della sua molteplicità algebrica. L'altro autospazio V_2 relativo all'autovalore 2 ha come base il vettore $(-3, 2, 1)$ e mettendo insieme le basi degli autospazi otteniamo solo due vettori indipendenti, insufficienti a formare una base di $\mathbb{R}^3$.

Infine, nell'Esempio 5.11, invece, abbiamo trovato i tre autovalori distinti $\lambda = 1$, $\lambda = 2$, $\lambda = -1$, ognuno con molteplicità algebrica uno, e in base all'Osservazione 5.21 sappiamo già che la matrice è diagonalizzabile: l'autospazio V_1 ha come base il vettore $(1,0,1)$, l'autospazio V_2 il vettore $(3,-1,1)$ e l'autospazio V_{-1} il vettore $(0,1,1)$, quindi una base di $\mathbb{R}^3$ formata da autovettori di A è $(1,0,1),(3,-1,1),(0,1,1)$, ottenuta mettendo insieme le basi degli autospazi.

Ora, mostriamo che la diagonalizzabilità di una matrice A (ovvero dell'endomorfismo $L_A : \mathbb{K}^n \to \mathbb{K}^n$ da lei determinato) si può equivalentemente enunciare nel modo seguente, molto utile in alcune applicazioni:

Proposizione 5.23. Una matrice $A \in M_n(\mathbb{K})$ è diagonalizzabile se e solo se esiste una matrice M invertibile di ordine n tale che il prodotto $M^{-1}AM$ è una matrice diagonale[5].

Proof. Supponiamo che A sia diagonalizzabile, cioè che ammetta n autovettori indipendenti

$$v_1 = (v_{11}, v_{21}, \dots, v_{n1}),\ v_2 = (v_{12}, v_{22}, \dots, v_{n2}), \dots, v_n = (v_{1n}, v_{2n}, \dots, v_{nn})$$

Disponendo tali autovettori in colonna, si ottiene una matrice

$$M = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ & \vdots & & \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}$$

quadrata di ordine n che è sicuramente invertibile in quanto le sue colonne, che sono gli autovettori trovati, sono per ipotesi indipendenti, e quindi il suo rango (ricordiamo che il rango è il numero di righe o colonne indipendenti di una matrice) è n (si veda il Teorema 4.19).
Mostriamo ora che la matrice M così costruita verifica la condizione dell'enunciato, ovvero che $M^{-1}AM$ è una matrice diagonale: più precisamente, mostriamo che vale l'uguaglianza

$$\begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ & \vdots & & \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}^{-1} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ & \vdots & & \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & \vdots & & \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{5.20}$$

[5]Ricordiamo che una matrice A si dice diagonale se ha le entrate A_{ij}, con $i \neq j$ (cioè quelle fuori dalla diagonale) uguali a zero.

dove λ_1 è l'autovalore di v_1, λ_2 è l'autovalore di v_2, e così via.
Infatti, moltiplicando entrambi i membri di tale uguaglianza a sinistra per M, vediamo che essa equivale a

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ & \vdots & & \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ & \vdots & & \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & \vdots & & \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{5.21}$$

Per mostrare l'uguaglianza scritta in questa forma, osserviamo che, per definizione di prodotto righe per colonne, la prima colonna del prodotto di matrici a primo membro della (5.21) si ottiene moltiplicando le righe di A per la prima colonna di V, ovvero

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix}$$

Ma, essendo la prima colonna di V autovettore di A, si ha

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix} = \lambda_1 \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11} \\ \lambda_1 v_{21} \\ \vdots \\ \lambda_1 v_{n1} \end{pmatrix}$$

Facendo lo stesso ragionamento sulle altre colonne del prodotto di matrici a primo membro della (5.21), si vede allora che

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ & \vdots & & \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} & \cdots & \lambda_n v_{1n} \\ \lambda_1 v_{21} & \lambda_2 v_{22} & \cdots & \lambda_n v_{2n} \\ & \vdots & & \\ \lambda_1 v_{n1} & \lambda_2 v_{n2} & \cdots & \lambda_n v_{nn} \end{pmatrix} \tag{5.22}$$

Ma come si verifica subito sempre svolgendo un prodotto righe per colonne, questo è uguale al secondo membro della (5.21).

Per mostrare viceversa che l'esistenza di una matrice invertibile M tale che $M^{-1}AM$ sia una matrice diagonale implica che A è diagonalizzabile, basta ripercorrere la dimostrazione a ritroso: se vale un'uguaglianza come la (5.20), che come abbiamo detto equivale alla (5.21), allora ogni colonna v_i della matrice M soddisfa l'uguaglianza $Av_i = \lambda_i v_i$, ed è quindi un autovettore di A. Quindi, abbiamo che A ammette n autovettori, che sono indipendenti essendo le colonne di una matrice invertibile, e A è diagonalizzabile. $\square$

Esempio 5.24. Consideriamo la matrice $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$. Si ha

$$\det(A - \lambda I_2) = \det \begin{pmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 9 = \lambda^2 - 2\lambda - 8.$$

e si trova subito che le soluzioni dell'equazione caratteristica sono $\lambda = 4$ e $\lambda = -2$, che sono quindi gli autovalori di A.
Per trovare i corrispondenti autovettori, basta risolvere il sistema omogeneo che ha $A - \lambda I_2$ come matrice dei coefficienti, prima con $\lambda = 4$ e poi con $\lambda = -2$. Per $\lambda = 4$ si ha $A - \lambda I_2 = \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}$. Come si vede, le righe rappresentano le equazioni dipendenti (una opposta dell'altra) $-3x_1 + 3x_2 = 0$ e $3x_1 - 3x_2 = 0$. Dividendo per 3, si ha quindi l'unica equazione $-x_1 + x_2 = 0$: posto $x_2 = t$, si trova $x_1 = x_2 = t$. Quindi le soluzioni del sistema, ovvero gli autovettori di A relativi all'autovalore 4, sono esattamente tutti i vettori del tipo (t, t), al variare di $t \in \mathbb{R}$ (ad esempio, per $t = 1$ si ritrova proprio l'autovettore visto subito dopo la Definizione 5.1). Analogamente, per $\lambda = -2$ si ha $A - \lambda I_2 = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$, quindi le righe rappresentano la stessa equazione $3x_1 + 3x_2 = 0$. Dividendo per 3 e posto $x_2 = t$, si trova $x_1 = -x_2 = -t$. Quindi le soluzioni del sistema, ovvero gli autovettori di A relativi all'autovalore -2, sono esattamente tutti i vettori del tipo $(-t, t)$, al variare di $t \in \mathbb{R}$.
Poiché $(t, t) = t(1, 1)$ e $(-t, t) = t(-1, 1)$, abbiamo due autovettori indipendenti $(1, 1)$ e $(-1, 1)$ e la matrice è diagonalizzabile. Per verificare la validità della (5.20), disponiamo tali autovettori in colonna, ottenendo $V = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$; come si verifica subito, l'inversa di tale matrice è $V^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$ e si ha

$$V^{-1}AV \quad \cancel{M^{-1}AM} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 4 & -2 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$$

cioè, come previsto dalla (5.20), otteniamo la matrice diagonale che ha sulla diagonale gli autovalori di A.

Osservazione 5.25. Un importante teorema sulla diagonalizzabilità afferma che le matrici A le cui entrate soddisfano la condizione $A_{ij} = A_{ji}$ (tali matrici si dicono *simmetriche*) sono sempre automaticamente diagonalizzabili. Torneremo su tale fatto nel prossimo capitolo, dove vedremo anche che per tali matrici la matrice M la cui esistenza è garantita dalla Proposizione 5.23 ha una forma particolare, importante nelle applicazioni.

Leggere

Esempio 5.26. La diagonalizzabilità di matrici nella forma enunciata nella Proposizione 5.23 si usa ad esempio quando si vogliono effettuare dei cambi di coordinate per rendere più semplice un problema su cui si sta lavorando: per fare un esempio relativo all'Analisi matematica, supponiamo di dover risolvere il sistema di equazioni differenziali

$$\begin{cases} \dot{x}_1 = x_1 + 3x_2 \\ \dot{x}_2 = 3x_1 + x_2 \end{cases}$$

dove $x_1 = x_1(t)$ e $x_2 = x_2(t)$ sono funzioni della variabile reale t e con $\dot{x}_1, \dot{x}_2$ stiamo indicando le derivate prime rispetto a t.
Tale sistema, sfruttando la definizione di prodotto di matrici, può essere riscritto nella forma

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

ovvero, posto $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ e $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$, possiamo scrivere

$$\dot{x} = Ax \tag{5.23}$$

Ora, A è esattamente la matrice vista nell'Esempio 5.24 e per la quale sappiamo che $M^{-1}AM = D = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$, dove $M = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Allora, usiamo la M per definire le due funzioni ausiliarie y_1 e y_2 tramite la relazione $y = M^{-1}x$, dove $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, ovvero

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x_1 + \frac{1}{2}x_2 \\ -\frac{1}{2}x_1 + \frac{1}{2}x_2 \end{pmatrix} \tag{5.24}$$

Riformulando il problema in termini di y_1 e y_2, vediamo subito che il problema si semplifica: poiché la $y = M^{-1}x$ equivale (moltiplicando entrambi i membri a sinistra per M) a $My = x$, sostituendo My al posto di x nella (5.23) troviamo

$$(\dot{My}) = AMy$$

ovvero (la matrice M ha entrate numeriche costanti, quindi quando deriviamo My si deriva solo y)

$$M\dot{y} = AMy$$

da cui, moltiplicando entrambi i membri a sinistra per l'inversa M^{-1} otteniamo

$$\dot{y} = M^{-1}AMy = Dy$$

ovvero

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4y_1 \\ 2y_2 \end{pmatrix}$$

Abbiamo quindi il sistema

$$\begin{cases} \dot{y}_1 = 4y_1 \\ \dot{y}_2 = 2y_2 \end{cases}$$

che rispetto al sistema iniziale ha il vantaggio di consistere di due equazioni differenziali ciascuna in una sola variabile, che si risolvono subito (la prima ci dà $y_1 = c_1e^{4t}$, la seconda $y_2 = c_2e^{2t}$, dove c_1, c_2 sono due costanti). A questo punto, per trovare le x_1 e x_2 basta ricordare che $x = My$, ovvero $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 - y_2 \\ y_1 + y_2 \end{pmatrix}$, e quindi $x_1 = c_1e^{4t} - c_2e^{2t}$ e $x_2 = c_1e^{4t} + c_2e^{2t}$.

In poche parole, per risolvere il problema abbiamo trovato un cambio di variabili che ha reso il problema più semplice separando le incognite del sistema, e la matrice M giusta da usare per realizzare questo cambio ci è stata suggerita dalla Proposizione 5.23.

Questa stessa idea si usa in numerosi altri contesti, ad esempio in geometria nel piano, se una curva è data da un'equazione di secondo grado $Ax^2+Bxy+Cy^2+Ex+Fy+G=0$ (tali curve si dicono coniche, e comprendono ellissi, parabole, iperboli) rispetto a un certo sistema di riferimento fissato, si può cambiare il sistema di riferimento in modo la semplificare l'equazione dela curva: non mostriamo i dettagli, ma si può vedere che il cambio di riferimento corretto da effettuare è determinato da un calcolo di autovettori e autovalori di una matrice di ordine 2 che si costruisce usando i coefficienti A, B, C dei termini di secondo grado dell'equazione.

La Proposizione 5.23 si può riformulare usando la seguente

Definizione 5.27. Due matrici quadrate $A, B \in M_n(\mathbb{K})$ si dicono *simili* se esiste una matrice invertibile M di ordine n tale che $B = M^{-1}AM$.

Alla luce di tale definizione, possiamo rienunciare la Proposizione 5.23 dicendo che *una matrice è diagonalizzabile se e solo se è simile a una matrice diagonale.*

Il concetto di similitudine di matrici è molto importante quando si affronta il *problema del cambiamento di coordinate*: come sappiamo, lavorare in coordinate rispetto a una base ci permette di risolvere numericamente problemi relativi ai vettori di qualunque spazio vettoriale di dimensione finita: abbiamo visto che in coordinate ogni vettore si rappresenta mediante una n-upla, e un'applicazione lineare si rappresenta mediante una matrice, ovvero più precisamente si costruisce la cosiddetta matrice $M_B(f)$ associata a un'applicazione lineare rispetto a una base, con la proprietà che se x è la n-upla delle coordinate di un vettore, allora $M_B(f)x$ è la m-upla delle coordinate della sua immagine. Questo è vero qualunque sia la base scelta, tuttavia a seconda del problema specifico che dobbiamo risolvere, alcune basi possono essere più convenienti di altre: saper cambiare base e quindi coordinate si rivela quindi fondamentale per poter risolvere nel modo più efficiente alcuni problemi, come vedremo anche nel prossimo capitolo.
Quello che dimostreremo ora, limitandoci agli endomorfismi, è il seguente

Teorema 5.28. Dato un endomorfismo $f : V \to V$, se $A = M_B(f)$ e $A' = M_{B'}(f)$ sono le matrici che rappresentano f rispetto a due basi B e B' di V, allora A e A' sono simili.

Per fare ciò abbiamo preliminarmente bisogno di risolvere il problema seguente: date due basi B e B' di uno spazio vettoriale V, come sono legate le coordinate di un vettore v rispetto alla base B a quelle rispetto alla base B'? come si passa dalle une alle altre?

Vediamo i dettagli: sia $B = \{v_1, \dots, v_n\}$ e $B' = \{v'_1, \dots, v'_n\}$.
Dato un generico vettore $v \in V$, siano $x_1, \dots, x_n$ le sue coordinate rispetto alla base B e $x'_1, \dots, x'_n$ le sue coordinate rispetto alla base B'. Per definizione di coordinate, questo significa che

$$v = x_1 v_1 + \cdots + x_n v_n \tag{5.25}$$

e analogamente,

$$v = x'_1 v'_1 + \cdots + x'_n v'_n \tag{5.26}$$

Ora, allo scopo di confrontare le due espressioni e ricavare la relazione cercata tra $x'_1, \dots, x'_n$ e $x_1, \dots, x_n$, sostituiamo nella (5.26) i vettori $v'_1, \dots, v'_n$ con le loro espressioni come combinazioni lineari dei vettori della base B (questo si può fare perché, essendo B una base, ogni vettore di V si può scrivere come combinazione lineare dei vettori di B, e in particolare $v'_1, \dots, v'_n$). Più precisamente, si avranno espressioni del tipo

$$v'_1 = a_{11}v_1 + \cdots + a_{n1}v_n \tag{5.27}$$

$$\vdots$$

$$v'_n = a_{1n}v_1 + \cdots + a_{nn}v_n \tag{5.28}$$

Sostituendo queste uguaglianze al secondo membro della (5.26), si trova

$$v = x'_1(a_{11}v_1 + \cdots + a_{n1}v_n) + \cdots + x'_n(a_{1n}v_1 + \cdots + a_{nn}v_n)$$

cioè, moltiplicando e mettendo in evidenza $v_1, \dots, v_n$,

$$v = (a_{11}x'_1 + \cdots + a_{1n}x'_n)v_1 + \cdots + (a_{n1}x'_1 + \cdots + a_{nn}x'_n)v_n.$$

Ora, confrontando questa uguaglianza con la (5.25), concludiamo che

$$\begin{array}{c} x_1 = a_{11}x'_1 + \cdots + a_{1n}x'_n \\ \vdots \\ x_n = a_{n1}x'_1 + \cdots + a_{nn}x'_n \end{array} \tag{5.29}$$

in quanto, come abbiamo visto quando abbiamo dimostrato l'unicità delle coordinate (Proposizione 1.21) non si può scrivere il vettore v in due modi diversi come combinazione lineare dei vettori $v_1, \dots, v_n$ della base B.

Le (5.29) rispondono al problema che ci siamo posti su come siano legate le coordinate rispetto a una base alle coordinate rispetto a un'altra: da una trasformazione di tipo lineare. Sfruttando la definizione di prodotto righe per colonne, le uguaglianze (5.29) possono essere raccolte nell'unica uguaglianza tra matrici

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

che, posto $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $x' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ e $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$, possiamo riscrivere semplicemente come $x = Ax'$.

La matrice A, le cui colonne sono definite dalle espressioni (5.27), ..., (5.28), si chiama *matrice di cambiamento di coordinate dalla base* B' *alla base* B, e si denota $M_{BB'}$. In base alle (5.27), ..., (5.28), essa può essere quindi essere definita come *la matrice che ha sulle sue colonne le coordinate dei vettori della base* B'*, calcolate rispetto alla base* B, e soddisfa la proprietà fondamentale

$$x = M_{BB'}x' \tag{5.30}$$

Osservazione 5.29. Si noti che la matrice di cambiamento di coordinate è sempre una matrice quadrata di ordine uguale alla dimensione n dello spazio V: infatti, dal momento che per definizione sulle colonne ha le coordinate dei vettori di B' rispetto a B, avrà tante colonne quanti sono i vettori di B', il cui numero è proprio la dimensione di V; e d'altra parte ogni colonna ha tante entrate quante sono le coordinate del vettore, cioè n, quindi ha anche n righe.
Inoltre, è facile vedere che la matrice di cambiamento di coordinate è una matrice invertibile: a questo scopo, si moltiplichino entrambi i membri della (5.30) per la matrice $M_{B'B}$ di cambiamento di coordinate dalla base B alla base B':

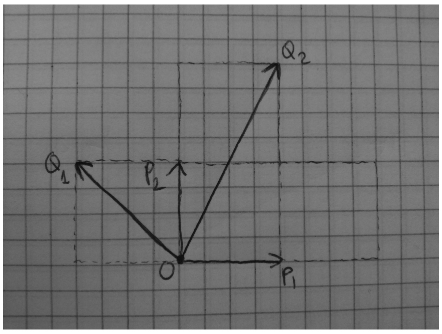
$$M_{B'B}x = M_{B'B}M_{BB'}x'$$

Il primo membro di tale uguaglianza coincide con la n-upla x' delle coordinate del vettore rispetto alla base B' (in quanto $M_{B'B}$ fa passare da coordinate rispetto a B a coordinate rispetto a B'), ovvero

$$x' = M_{B'B}M_{BB'}x'$$

Tale uguaglianza può essere vera per ogni n-upla x' se e solo se la matrice prodotto $M_{B'B}M_{BB'}$ è uguale alla matrice identica I_n (che è l'elemento neutro per il prodotto): ma $M_{B'B}M_{BB'} = I_n$, che ci dice appunto che $M_{BB'}$ è invertibile e la sua inversa è la matrice $M_{B'B}$ (sempre di cambiamento di coordinate, ma con le basi invertite).

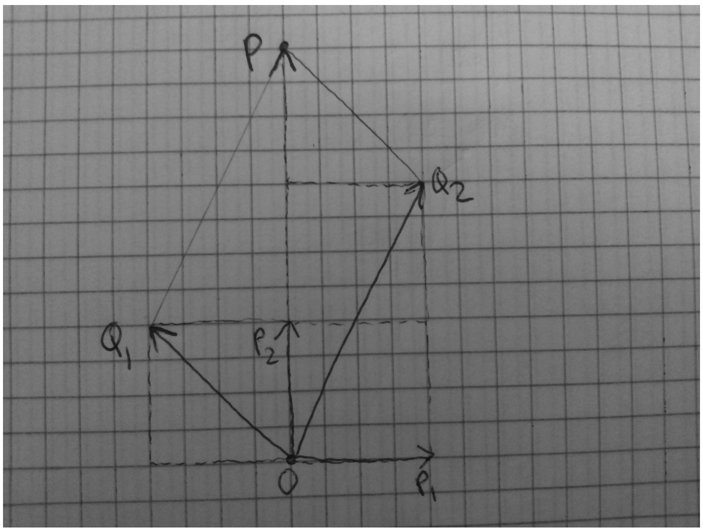
Esempio 5.30. Consideriamo le basi $B = \{v_1 = \vec{OQ}_1, v_2 = \vec{OQ}_2\}$ e $B' = \{v'_1 = \vec{OP}_1, v'_2 = \vec{OP}_2\}$ rappresentate nel disegno seguente



Vogliamo calcolare la matrice $M_{B'B}$ di cambiamento di coordinate da B a B'.
In base alla definizione, questa matrice avrà nella prima colonna le coordinate del primo vettore $v_1 = \vec{OQ}_1$ della base B calcolate rispetto alla base B': essendo , come si vede graficamente dal disegno, $\vec{OQ}_1 = -\vec{OP}_1 + \vec{OP}_2$, queste coordinate sono -1 e 1.
Analogamente, la matrice avrà nella seconda colonna le coordinate del secondo vettore $v_2 = \vec{OQ}_2$ della base B calcolate rispetto alla base B': essendo, come si vede graficamente dal disegno, $\vec{OQ}_2 = \vec{OP}_1 + 2\vec{OP}_2$, queste coordinate sono 1 e 2.
La matrice è quindi

$$M_{B'B} = \begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix}$$

Per illustrare come funziona la formula $x' = M_{B'B}x$, consideriamo il vettore $v = \vec{OP}$ in figura



per il quale si vede che $\vec{OP} = \vec{OQ}_1 + \vec{OQ}_2$, e quindi $(1, 1)$ sono le sue coordinate rispetto a B. Moltiplicando per la matrice di cambiamento di coordinate $M_{B'B}$ si trova

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

che ci dice che $(0, 3)$ sono le coordinate di $\vec{OP}$ rispetto a $B' = \{\vec{OP}_1, \vec{OP}_2\}$, ovvero deve essere $\vec{OP} = 0\vec{OP}_1 + 3\vec{OP}_2 = 3\vec{OP}_2$, conformemente a quello che si deduce guardando il disegno.

Esempio 5.31. A illustrazione della $M_{B'B}^{-1} = M_{BB'}$, siano $B = \{\vec{OQ}_1, \vec{OQ}_2\}$ e $B' = \{\vec{OP}_1, \vec{OP}_2\}$ le stesse basi dell'esempio precedente, per le quali abbiamo già calcolato

$$M_{B'B} = \begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix}.$$

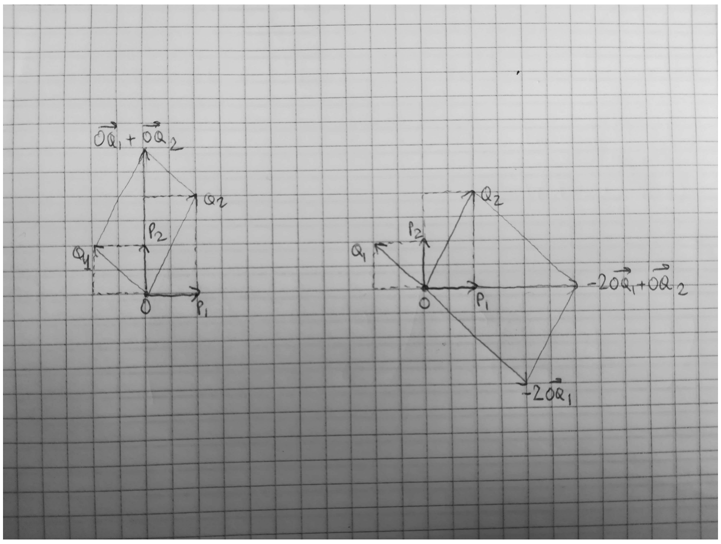
Si ha allora, usando la formula (4.40),

$$M_{BB'} = M_{B'B}^{-1} = \frac{1}{-3} \begin{pmatrix} 2 & -1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Quindi deve essere vero che le coordinate del primo vettore di B', calcolate rispetto a B sono $-\frac{2}{3}, \frac{1}{3}$, e che le coordinate del secondo vettore di B', calcolate sempre rispetto a B sono $\frac{1}{3}, \frac{1}{3}$, ovvero

$$\vec{OP}_1 = -\frac{2}{3}\vec{OQ}_1 + \frac{1}{3}\vec{OQ}_2 = \frac{1}{3}(-2\vec{OQ}_1 + \vec{OQ}_2)$$
$$\vec{OP}_2 = \frac{1}{3}\vec{OQ}_1 + \frac{1}{3}\vec{OQ}_2 = \frac{1}{3}(\vec{OQ}_1 + \vec{OQ}_2)$$

Queste uguaglianze sono confermate da un'analisi grafica:



Siamo ora pronti a dimostrare il Teorema 5.28.
Se $A = M_B(f)$ e $A' = M_{B'}(f)$ sono le matrici associate all'endomorfismo f rispetto alle basi B e B', allora dato un vettore v, di coordinate x rispetto a B e coordinate x' rispetto a B', si hanno le seguenti

$$y = Ax, \quad y' = A'x' \tag{5.31}$$

dove y denota la n-upla delle coordinate di $f(v)$ rispetto a B e y' la n-upla delle coordinate di $f(v)$ rispetto a B'.
D'altra parte, se M denota la matrice $M_{BB'}$ di cambiamento di coordinate dalla base B' alla base B, si hanno le

$$x = Mx', \quad y = My'$$

Sostituendo queste due uguaglianze nella prima delle (5.31), si trova

$$My' = AMx'$$

Moltiplicando ~~Dividendo~~ entrambi i membri a sinistra per l'inversa M^{-1} di M (abbiamo visto che la matrice di cambiamento di coordinate è invertibile) si trova

$$y' = M^{-1}AMx'$$

Ora, questa uguaglianza ci dice che la matrice $M^{-1}AM$, moltiplicata per le coordinate x' di un vettore v rispetto a B', ci dà le coordinate y' dell'immagine $f(v)$ (sempre rispetto a B'): ma questa è esattamente la proprietà che caratterizza la matrice A' associata a f rispetto a B' (la seconda delle (5.31), e quindi si deve avere

$$A' = M^{-1}AM$$

che ci dice proprio che A e A' sono simili, che è quanto volevamo dimostrare.

Il fatto che le matrici $A = M_B(f)$ e $A' = M_{B'}(f)$ associate a uno stesso endomorfismo rispetto a basi diverse siano simili ci permette di mostrare un fatto che abbiamo trascurato finora, ovvero il fatto che *il polinomio caratteristico* $\det(M_B(f) - \lambda I_n)$ *di un endomorfismo non dipende dalla base scelta per calcolarlo.*
Infatti, osserviamo che da $A' = M^{-1}AM$ discende che

$$A' - \lambda I_n = M^{-1}(A - \lambda I_n)M \tag{5.32}$$

Per verificare questa uguaglianza, basta svolgere i calcoli a secondo membro:

$$M^{-1}(A-\lambda I_n)M = M^{-1}(AM-\lambda M) = M^{-1}AM-M^{-1}(\lambda M) = M^{-1}AM-\lambda(M^{-1}M) = A'-\lambda I_n$$

(nella penultima uguaglianza abbiamo usato la proprietà (4.49)).
Ora, calcoliamo il determinante di entrambi i membri della (5.32):

$$\det(A' - \lambda I_n) = \det[M^{-1}(A - \lambda I_n)M] =$$

(usando il Teorema 4.28 di Binet che afferma che il determinante del prodotto di matrici è uguale al prodotto dei determinanti)

$$= \det(M^{-1})\det(A - \lambda I_n)\det(M) = \det(A - \lambda I_n)$$

dove abbiamo semplificato $\det(M^{-1}) = \frac{1}{\det(M)}$ (Proposizione 4.46) con $\det(M)$.
Quindi $\det(M_B(f)-\lambda I_n) = \det(M_{B'}(f)-\lambda I_n)$ e la dimostrazione è conclusa.

Concludiamo questo capitolo con alcune osservazioni sul polinomio caratteristico che ci saranno utili nel capitolo successivo.

Per definizione, il polinomio caratteristico è

$$\det\begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{pmatrix}$$

Quindi, per definizione di determinante, esso è formato dalla somma di tutti i possibili prodotti di n entrate di tale matrice scelte in modo da stare sempre su righe e colonne diverse[6] In particolare, l'unico di tali prodotti in cui λ appare n volte è

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda)$$

Se sviluppassimo tale prodotto, vedremmo allora che esso ci dà il monomio $(-\lambda)^n = (-1)^n \lambda^n$.
Questo conferma che il polinomio caratteristico ha effettivamente grado n, e ci dice che il coefficiente del monomio di grado massimo è sempre $(-1)^n$ (quindi esso inizia con $+\lambda^n$ se n è pari e con $-\lambda^n$ se n è dispari, si vedano anche tutti gli esempi sopra).

Per quello che riguarda invece il suo termine noto, basta osservare che se

$$\det(A - \lambda I_n) = (-1)^n \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

allora ponendo $\lambda = 0$ a entrambi i membri si ha $\det(A) = c_0$: quindi il termine noto del polinomio caratteristico coincide sempre con il determinante della matrice A.
Infine, come è noto dall'algebra, nel campo $\mathbb{C}$ dei numeri complessi ogni polinomio si decompone come prodotto di fattori di primo grado, e più precisamente per il polinomio caratteristico avremo

$$\det(A - \lambda I_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda)$$

dove $\lambda_1, \lambda_2, \ldots, \lambda_n$ sono le radici del polinomio (anche quelle complesse e eventualmente ripetute più volte).
Ponendo $\lambda = 0$ anche in quest'ultima espressione otteniamo

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$$

ovvero concludiamo che *il determinante di una matrice è uguale al prodotto delle radici del suo polinomio caratteristico.*
Ad esempio, se $A = \det \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, come abbiamo visto nell'Osservazione 5.5, si ha che il polinomio caratteristico ha le radici complesse $\lambda_1 = +i$ e $\lambda_2 = -i$, il cui prodotto $\lambda_1\lambda_2 = (+i)(-i) = -i^2 = +1$ è proprio uguale al determinante di A.

[6] Questo conferma che il polinomio caratteristico è effettivamente un polinomio in λ, perché costruito usando solo somme e prodotti.