Lezioni del 21/1/2014 e del 28/1/2014

Ricordiamo che si denota con

$$GL_n(\mathbb{K}) := \{ M \in \mathcal{M}_n(\mathbb{K}) \mid M \text{ è non singolare } [\Leftrightarrow \det M \neq 0] \}.$$

<u>NOTA</u> $GL_n(\mathbb{K})$ non è un gruppo rispetto alla somma di matrici (la matrice nulla non è invertibile; in particolare, non è sottospazio vettoriale di $\mathcal{M}_n(\mathbb{K})$); però $GL_n(\mathbb{K})$ è un gruppo rispetto al prodotto righe x colonne. Infatti, se A, B sono invertibili anche AB lo è (l'identità del gruppo è la matrice $I = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$).

Ricordiamo che la matrice P di un cambiamento di base sta in $GL_n(\mathbb{K})$.

Abbiamo visto che la traccia, il determinante e il rango sono invarianti per similitudine, cioè se $A \in M_n(\mathbb{K})$, $\forall P \in GL_n(\mathbb{K})$ si ha:

$$tr(A) = tr(P^{-1}AP)$$
$$det(A) = det(P^{-1}AP)$$
$$rg(A) = rg(P^{-1}AP)$$

<u>Osservazione</u>. Ogni invariante per similitudine si può associare alle applicazioni lineari: se $f \in \mathcal{L}(V)$ e i è un invariante per similitudine, possiamo ben definire:

$$i(f) = i(M_{\mathcal{B}}(f))$$

dove $M_{\mathcal{B}}(f)$ è la matrice associata a f rispetto a una qualunque base $\mathcal{B}$ (uguale in partenza e in arrivo). La definizione non dipende dalla base che scegliamo, perché la matrice associata a un'altra base cambierà per similitudine.

Quindi sono ben definite: $tr(f)$, $det(f)$, $rg(f)$ (quest'ultimo lo avevamo già definito in altro modo equivalente]

Problema: Bastano i 3 invarianti scritti sopra per caratterizzare le classi di matrici simili?
Cioè, è vero che se A, B hanno uguali i 3 invarianti detti, allora A è simile a B?

Se $n > 1$ (il caso $n=1$ è banale) la risposta è no:

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

si ha $\operatorname{rg} A = \operatorname{rg} B = 2$, $\operatorname{tr} A = \operatorname{tr} B = 2$, $\det A = \det B = 1$;
ma $A = I$ è simile solo a sé stessa: $P^{-1} I P = I$, $\forall$
$P \in GL_n(\mathbb{K})$.

Al problema: <u>stabilire se due date matrici A e B sono simili</u>
non risponderemo in generale [si dovrebbe fare la <u>forma canonica di Jordan</u>, ma non c'è tempo].
Risolveremo il problema nel caso particolare (importante) in cui una delle due matrici è diagonale. In particolare, risponderemo alla domanda: quando una data matrice $A \in M_n(\mathbb{K})$ è diagonalizzabile? [<u>def</u>: A si dice diagonalizzabile se è

simile a una matrice diagonale].

Una matrice diagonale $A = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ avente sulla diagonale numeri $\lambda_1, \lambda_2, \dots, \lambda_n$ (che possono ripetersi), se la vediamo come applicazione lineare $f: \mathbb{K}^n \to \mathbb{K}^n$, $\underline{x} \to A\underline{x}$, $\underline{x} \in \mathbb{K}^n$, agisce geometricamente in modo molto particolare:

- Ogni asse cartesiano $\text{Span}\{\underline{e}_i\} = V_i$ viene mandato in sé, perché

$$A\,\underline{e}_i = \lambda_i\,\underline{e}_i \qquad , i = 1, \dots, n;$$

quindi $f(V_i) \subset V_i$ [e se $\lambda_i \neq 0$, $f(V_i) = V_i$].

Def. Sia $f: V \to V$ endomorfismo. Un sottospazio $W \subset V$ si dirà *f-invariante* (o semplicemente invariante, quando f è sottintesa) se $f(W) \subset W$.

Quindi se A è diagonale, gli assi cartesiani sono sottospazi A-invarianti di dim 1.

Siccome se A e B sono simili, esse si possono vedere come matrici associate alla stessa $f: K^n \to K^n$, rispetto a diverse basi, le proprietà geometriche di f devono poter "essere lette" nello stesso modo usando A o B.
Ad esempio se $W \subset K^n$ è A-invariante e $B = P^{-1} A P$, allora

$$B(P^{-1}(W)) = P^{-1}APP^{-1}(W) = P^{-1}A(W) \subset P^{-1}(W)$$

(perché $A(W) \subset W$) quindi $P^{-1}(W)$ [che ha la stessa dimensione di W perché P^{-1} è di rango massimo] è B-invariante.

<u>Nota</u>. Sia $f \in \mathcal{L}(V)$ (endomorfismo). Allora un sottospazio $W \subset V$ di dimensione 1 è f-invariante $\Leftrightarrow \exists\, \lambda \in \mathbb{K}$ t.c. $\forall\, \underline{v} \in W$, $f(\underline{v}) = \lambda \underline{v}$.

[dimostrare questa affermazione come semplice esercizio!]

Def. Un numero $\lambda \in \mathbb{K}$ per il quale ci sono vettori $\underline{v} \in V$, $\underline{v} \neq \underline{0}$, t.c.

$$f(\underline{v}) = \lambda \underline{v} \qquad (*)$$

si dice AUTOVALORE di f. Un vettore non nullo $\underline{v}$ che soddisfa $(*)$ si dice AUTOVETTORE di f, relativo all'autovalore λ.

Per quanto detto sopra, $\underline{v}$ è autovettore per $f \iff$ la retta $\mathrm{Span}\{\underline{v}\}$ è f-invariante.

La definizione di autovalori e autovettori si dà nello stesso modo per le matrici:

Def se $\quad A\underline{x} = \lambda \underline{x}; \quad \underline{x} \in \mathbb{K}^n \setminus \{\underline{0}\}, \quad \lambda \in \mathbb{K},$

allora λ si dice autovalore di A e $\underline{x}$ autovettore relativo all'autovalore λ.

<u>NOTA</u>. Se $f \in \mathcal{L}(V)$ e $f(\underline{v}) = \lambda \underline{v}$ per qualche $\lambda \in \mathbb{K}$, $\underline{v} \in V \setminus \{\underline{0}\}$, allora passando in coordinate rispetto a una base $\mathcal{B}$ di V, si ha, detti $A = M_{\mathcal{B}}(f)$, $\underline{x} = [\underline{v}]_{\mathcal{B}} \in \mathbb{K}^n$:

$$A\,\underline{x} = \lambda\,\underline{x}$$

cioè $\underline{x}$ è autovettore di A rispetto all'autovalore λ. Quindi gli autovalori di f sono tutti e soli gli autovalori di $M_{\mathcal{B}}(f)$, e gli autovettori di f relativi a λ hanno come coordinate, relativamente a $\mathcal{B}$, gli autovettori di $M_{\mathcal{B}}(f)$ relativi a λ.

Vediamo quindi come si trovano autovalori e autovettori di una matrice $A \in M_n(\mathbb{K})$ (e di un endomorfismo f).

Possiamo riscrivere

$$A \underline{v} = \lambda \underline{v} \Leftrightarrow (A - \lambda I) \underline{v} = \underline{0} \qquad (1)$$

(I matrice identità di ordine n). Siccome stiamo cercando soluzioni $\underline{x} \neq \underline{0}$, i $\lambda \in \mathbb{K}$ per cui queste esistono sono precisamente quei valori per cui

$$\dim(\ker(A - \lambda I)) > 0 \Leftrightarrow rg(A - \lambda I) < n \Leftrightarrow \det(A - \lambda I) = 0.$$

Quindi:

<u>Teorema</u>. 1) Gli autovalori di A sono le soluzioni dell'equazione algebrica: $\boxed{\det(A-\lambda I) = 0.}$ (2)

2) Gli autovettori relativi all'autovalore λ sono tutte le soluzioni non nulle del sistema lineare omogeneo:

$$\boxed{(A-\lambda I)\underline{x} = \underline{0}} \quad (3)$$

<u>Def</u> Il polinomio di grado n in λ: $\det(A-\lambda I)$ si dice POLINOMIO CARATTERISTICO di A, e si indica con $p_A(\lambda) := \det(A-\lambda I)$.

Ricordiamo il

<u>Teorema fondamentale dell'algebra.</u> Ogni polinomio $p(x) \in \mathbb{C}[x]$

ha una radice.

Per il teorema di Ruffini, se $p(x) \in \mathbb{K}[X]$ ha una radice $\alpha \in \mathbb{K}$, allora $p(x)$ è divisibile per $x-\alpha$. Esiste cioè un polinomio $q(x)$, con $gr(q(x)) = gr(p(x)) - 1$, tale che $p(x) = (x-\alpha)\, q(x)$.

[Coroll. grado $p(x) = n \Rightarrow p(x)$ ha al max n radici (in $\mathbb{K}$)]. Applicando ricorsivamente il teorema fondamentale dell'algebra si ottiene

Corollario Se $p(x) \in \mathbb{C}[X]$ ha grado n, allora $p(x)$ si fattorizza in n fattori lineari: $p(x) = c(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$, dove $c \in \mathbb{K}$ è una costante (il coefficiente del monomio di grado n di $p(x)$) e possono esserci ripetizioni tra le α_i. Raggruppando radici uguali si può scrivere $p(x) = c(x-\alpha_1)^{n_1}(x-\alpha_2)^{n_2}\cdots(x-\alpha_k)^{n_k}$

dove le $\alpha_1, \dots, \alpha_k$ sono tutte le radici distinte di $p(x)$, e $n_1, \dots, n_k \in \mathbb{N}$, $n_1 + \dots + n_k = n$.

Def. Si dice **MOLTEPLICITÀ ALGEBRICA** di una radice $\alpha \in \mathbb{K}$ di $p(x) \in \mathbb{K}[x]$ il massimo esponente m per cui $(x-\alpha)^m$ divide $p(x)$ [cioè $p(x) = (x-\alpha)^m q(x)$, $q(x) \in \mathbb{K}[x]$].

La molteplicità algebrica di un autovalore λ della matrice A è la molteplicità di λ come radice del polinomio caratteristico $p_A(\lambda)$. Si indicherà con

$\mu_a(\lambda)$. Quindi se $\lambda_1, \dots, \lambda_k$ sono i diversi autovalori di $A \in M_n(\mathbb{C})$, si avrà:

$$p_A(\lambda) = \pm (\lambda - \lambda_1)^{m_a(\lambda_1)} (\lambda - \lambda_2)^{m_a(\lambda_2)} \cdots (\lambda - \lambda_k)^{m_a(\lambda_n)}$$

NOTA: il segno sarà + se n è pari, - se n è dispari.

Nel caso in cui $\mathbb{K} = \mathbb{R}$, ci sono polinomi $p(x) \in \mathbb{R}[x]$ che non hanno radici (es.: $p(x) = x^2 + 1$). A differenza del caso complesso, in cui ogni matrice $A \in M_n(\mathbb{C})$ ha autovalori (in $\mathbb{C}$), esistono matrici reali $A \in M_n(\mathbb{R})$ che non hanno autovalori reali

Ad esempio $A=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$ (corrispondente a una rotazione di $\pi/2$) ha polinomio caratteristico $\det\begin{pmatrix}-\lambda&-1\\1&-\lambda\end{pmatrix}=\lambda^2+1$.

La cosa è geometricamente evidente, perché eventuali autovettori reali corrisponderebbero a rette invarianti, mentre una rotazione di $\pi/2$ nel piano non lascia invariata nessuna retta. Similmente, la rotazione di angolo θ data dalla matrice $A=\begin{bmatrix}\cos\theta&-\operatorname{sen}\theta\\\operatorname{sen}\theta&\cos\theta\end{bmatrix}$ ha polinomio caratteristico $p_A(\lambda)=\lambda^2-2\cos\theta\cdot\lambda+1$

che ha radici $\lambda=\cos\theta\pm\sqrt{\cos^2\theta-1}=\cos\theta\pm\sqrt{-\operatorname{sen}^2\theta}$.

A meno che non sia $\theta=0$ o $\theta=\pi$, le due radici scritte sono complesse ma non reali: $\lambda_{1,2}=\cos\theta\pm i\operatorname{sen}\theta$.

Proposizione. Un polinomio $p(x) \in \mathbb{R}[x]$ che abbia la radice complessa α ha anche la radice complessa coniugata $\bar{\alpha}$. Inoltre le molteplicità algebriche di α e di $\bar{\alpha}$ coincidono.

---

dim. Sia $p(x) = c_0 x^n + c_1 x^{n-1} + \dots + c_{n-1} x + c_n$, $c_i \in \mathbb{R}$.
Si ha $p(\alpha) = 0 \iff \overline{p(\alpha)} = \bar{0} = 0$. Ma siccome $p(x)$ è a coefficienti reali si ottiene: $p(\alpha) = 0 \iff$

$$\overline{p(\alpha)} = \overline{(c_0 \alpha^n + c_1 \alpha^{n-1} + \dots + c_n)} = \bar{c}_0 (\bar{\alpha})^n + \bar{c}_1 (\bar{\alpha})^{n-1} + \dots + \bar{c}_n =$$

$$= c_0 (\bar{\alpha})^n + c_1 (\bar{\alpha})^{n-1} + \dots + c_n = 0,$$ cioè $\bar{\alpha}$ è radice di $p(x)$.

Per le molteplicità, si osservi che se $\alpha \in \mathbb{C}$ ma $\alpha \notin \mathbb{R}$, il polinomio $g(x) = (x-\alpha)(x-\bar{\alpha}) = x^2 - 2\mathrm{Re}\,\alpha \cdot x + |\alpha|^2$

è a coefficienti reali e divide $p(x)$, cioè

$$p(x) = g(x)\, q(x) \quad , \quad q(x) \in \mathbb{R}[x] , \quad gr(q(x)) = n-2.$$

Allora si conclude per induzione su n _

Def. Sia $\lambda \in \mathbb{K}$ un autovalore di $A \in M_n(\mathbb{K})$.
Lo spazio $V_{A,\lambda}$ delle soluzioni del sistema (3) $[(A-\lambda I)\underline{x}=\underline{0}]$ (consistente di tutti gli autovettori relativi a λ più il vettore nullo) si dice AUTOSPAZIO relativo all'autovalore λ.

Se $f: V \longrightarrow V$ è un endomorfismo e $\lambda \in \mathbb{K}$ è un autovalore di f, si definisce (come per una matrice) AUTOSPAZIO relativo all'autovalore λ il sottospazio $V_{f,\lambda} = \{ \underline{v} \in V \mid f(\underline{v}) = \lambda \underline{v} \}$ (tutti gli autovettori relativi a λ più il vettore nullo).

Se $A = M_{\mathcal{B}}(f)$ è la matrice associata a f rispetto alla base $\mathcal{B}$, si ha un isomorfismo tra

$$V_{f,\lambda} = \{ \underline{v} \in V \mid f(\underline{v}) = \lambda \underline{v} \} \xrightarrow{\simeq} V_{A,\lambda} = \{ \underline{x} \in \mathbb{K}^n \mid A \underline{x} = \lambda \underline{x} \}$$

dato da

$$\underline{v} \longrightarrow [\underline{v}]_{\mathcal{B}}$$

Quando non ci sia possibilità di confusione denoteremo un autospazio solo con V_λ.

La MOLTEPLICITÀ GEOMETRICA di λ, indicata con $\mu_g(\lambda)$, è la dimensione dell'autospazio relativo a λ.

Esempio 1. Sia $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Si ha $p_A(\lambda) = (\lambda - 1)^2$, quindi A ha come unico autovalore $\lambda = 1$ con $\mu_a(1) = 2$.

$$\dim \{ \underline{x} \mid (A - 1 \cdot I)\underline{x} = \underline{0} \} = \dim \left\{ \underline{x} \mid \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = 1$$

quindi $\mu_g(1) = 1 \ (< \mu_a(1))$.

2. Se A è una matrice diagonale, $A = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}$ allora

$p_A(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda)$, che ha radici gli elementi diagonali $a_{11}, \dots, a_{nn}$ (che possono ripetersi). Quindi gli autovalori di A si trovano già sulla diagonale.

Se
$$A = \begin{bmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_1 & & & & \text{\Large 0} & \\ & & & \lambda_2 & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_2 & & \\ & \text{\Large 0} & & & & & \ddots & \\ & & & & & & & \lambda_n \ddots \lambda_n \end{bmatrix}$$
dove $\lambda_i \neq \lambda_j$, e λ_i compare μ_i volte sulla diagonale, allora $\mu_a(\lambda_i) = \mu_i$. Inoltre anche $\mu_g(\lambda_i) = \mu_i$; verifichiamolo per λ_1 (gli altri casi sono analoghi).

$$rg(A - \lambda_1 I) = rg \begin{bmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & \lambda_2 - \lambda_1 & & & \\ & & & & \ddots & & \\ & & & & & \lambda_2 - \lambda_1 & \\ & & & & & & \ddots \\ & & & & & & & \lambda_k - \lambda_1 \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_k - \lambda_1 \end{bmatrix} = n - \mu_a(\lambda_1)$$

quindi $\mu_a(\lambda_1) = \mu_g(\lambda_1)$

3. Siccome il determinante di una matrice triangolare è il prodotto degli elementi diagonali, si ha che se A è triangolare (inferiore o superiore) $p_A(\lambda) = (a_{11}-\lambda)\cdots(a_{nn}-\lambda)$ e quindi gli autovalori di A sono gli elementi diagonali a_{ii}, con molteplicità algebrica pari al numero di volte che si ripetono. La molteplicità geometrica può essere inferiore, come nell'esempio 1.

<u>NOTA</u>. Se λ è autovalore di $f: V \to V$, si definisce il polinomio caratteristico di f: $p_f(\lambda)$ come il polinomio $p_A(\lambda)$, con $A = M_{\mathcal{B}}(f)$, per una qualche base $\mathcal{B}$ di V.

Occorre dimostrare che il polinomio caratteristico di f non dipende dalla base usata per calcolarlo.

Più esplicitamente si ha:

<u>Proposizione</u>. Se $B = P^{-1}AP$, allora $p_A(\lambda) = p_B(\lambda)$.

<u>dim</u>.

$p_B(\lambda) = \det(P^{-1}AP - \lambda I) = \det(P^{-1}(A - \lambda I)P)$. Per il teorema di Binet, comunque si sostituisca a λ un numero $\alpha \in \mathbb{K}$, si trova $p_B(\alpha) = p_A(\alpha)$. Ma due polinomi (a coefficienti in $\mathbb{C}$ o in $\mathbb{R}$) che coincidono per infiniti valori sono uguali (cioè hanno gli stessi coefficienti)*.

Quindi $p_A(\lambda) = p_B(\lambda)$.

[<u>oss.</u> Si potrebbe dimostrare che il teo. di Binet vale anche su anelli (non necessariamente campi) e quindi fare a meno dell'argomento della dimostrazione. La proposizione vale $\forall$ campo $\mathbb{K}$].

<u>Corollario</u> Tutti i coefficienti del polinomio caratteristico sono invarianti per similitudine.

(*) Sia $h(x) = f(x) - g(x)$. $h(\alpha) = f(\alpha) - g(\alpha) = 0$ per infiniti α. Ma un polinomio non può avere più radici del suo grado. Quindi $h = 0$.

Oss. Se A è simile alla matrice diagonale D, siccome hanno gli stessi autovalori e gli autovalori di D sono gli elementi diagonali, se ne ricava che la forma diagonale di A (se diagonalizzabile) è data dagli autovalori di A.

Sia $p_A(\lambda) = (-1)^n(\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + a_3\lambda^{n-3} + \dots + a_n)$.

Allora:

- $a_1 = -\operatorname{tr}(A)$
- $a_n = (-1)^n \det(A)$
- in generale $a_k = (-1)^k \sum_{i_1 < \dots < i_k} \det\left[A^{i_1,\dots,i_k}_{i_1,\dots,i_k}\right]$

dove $A^{i_1\dots i_k}_{i_1\dots i_k}$ è la sottomatrice "diagonale" costruita prendendo le righe di indici $i_1,\dots,i_k$ e le colonne con gli stessi indici.

<u>cenno di dim</u>. Nella formula del determinante, i monomi che contengono λ^k si ottengono scegliendo k termini uguali a λ dalla diagonale, in posizione $(i_1, i_1), \dots, (i_k, i_k)$, e tutti gli altri nella sotto-matrice quadrata di ordine $n-k$ ottenuta cancellando le righe e le colonne degli stessi indici.

<u>Es</u>. Se $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ allora

$$p_A(\lambda) = -\lambda^3 + I_1 \cdot \lambda^2 - I_2 \cdot \lambda + I_3 \qquad \text{con}$$

$I_1 = \operatorname{tr} A$, $I_3 = \det A$,

$$I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

Teorema. Sia $f \in \mathcal{L}(V)$. Per ogni autovalore α di f, vale

$$\boxed{\mu_g(\alpha) \leq \mu_a(\alpha)}$$

dim.

Ricordiamo che $\mu_g(\alpha) = \dim V_\alpha$, dove V_α è l'autospazio relativo all'autovalore α. Prendiamo una base $\mathcal{B}$ di V costruita estendendone una di V_α: $\underline{v}_1, \dots, \underline{v}_k$, con $k = \mu_g(\alpha)$. Poiché $f(\underline{v}_i) = \alpha \underline{v}_i$ per $i = 1, \dots, k$, si ha che la matrice associata a f rispetto a $\mathcal{B}$ ha la forma:

$$M_{\mathcal{B}}(f) = \left[\begin{array}{c|c} \alpha I_k & B \\ \hline 0 & C \end{array}\right]$$

dove I_k è la matrice identità di ordine k, e $C \in M_{n-k}(\mathbb{K})$.

Allora

$$M_{\mathcal{B}}(f) - \lambda I = \left[\begin{array}{c|c} (\alpha-\lambda) I_k & B \\ \hline 0 & C - \lambda I_{n-k} \end{array}\right]$$

Allora $\det(M_B(f)-\lambda I) = \det[(\alpha-\lambda)I_k]\det(C-\lambda I_{n-k}) =$

$$= (\alpha-\lambda)^k \, p_C(\lambda).$$

Segue: $k \leq \mu_a(\alpha)$, che è quanto volevamo.

Corollario. La stessa disuguaglianza vale per gli autovalori di una matrice A.

dim. Basta vedere A come l'applicazione lineare

$$f : \mathbb{K}^n \to \mathbb{K}^n : \underline{x} \to A\underline{x}.$$

Allora A risulta la matrice associata a f rispetto alla base canonica, quindi f ed A hanno lo stesso polinomio caratteristico. Gli autospazi di f sono per definizione quelli di A.

Torniamo al problema di decidere quando una matrice è diagonalizzabile.

Definizione

Per un endomorfismo $f: V \to V$, diremo che f è diagonalizzabile se $\exists$ base $\mathcal{B}$ di V t.c. $M_{\mathcal{B}}(f)$ sia diagonale.

Abbiamo visto che se D è una matrice diagonale, allora la base canonica di $\mathbb{K}^n$ è una base di autovettori (gli assi cartesiani sono invarianti). Si ha anche il viceversa.

Osservazione: $f: V \to V$ è diagonalizzabile $\Leftrightarrow$ V ha una base di autovettori.

---

dim. $\Rightarrow$ Se $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ e $A = M_{\mathcal{B}}(f)$ è diagonale, allora $f(\underline{v}_j) = A_{jj}\, \underline{v}_j$, quindi i $\underline{v}_j$ sono autovettori.

$\Leftarrow$ Se $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ è base di autovettori, allora $f(\underline{v}_j) = \lambda_j \underline{v}_j$, $j = 1, \dots, n$, da cui

$$M_{\mathcal{B}}(f) = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

—

Per quanto sopra detto, si ha che f è diagonalizzabile $\Leftrightarrow$ la matrice associata $M_{\mathcal{B}}(f)$ è diagonalizzabile, per una qualunque base $\mathcal{B}$.

Infatti esiste una base, $\mathcal{B}'$, tale che $M_{\mathcal{B}'}(f)$ è diagonale, ma $M_{\mathcal{B}}(f)$ è simile a $M_{\mathcal{B}'}(f)$ per la formula di cambiamento di base. Viceversa, se $A = M_{\mathcal{B}}(f)$ è diagonalizzabile, $P^{-1}AP = D$, $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$, allora $AP = PD \Rightarrow AP^j = \lambda_j P^j$, $j = 1, \dots, n$

cioè i vettori $\underline{v}_j \in V$ con coordinate $[\underline{v}_j]_{\mathcal{B}} = P^j$ sono autovettori, e sono una base perché P è invertibile. Quindi per il teo. precedente f è diagonalizzabile (la matrice associata ad f rispetto ai $\underline{v}_j$ è diagonale).

Osserviamo anche che la matrice A è diagonalizzabile ($\Leftrightarrow$ è simile a una matrice diagonale) $\Leftrightarrow$ l'applicazione lineare $f_A : \underline{x} \to A\underline{x} : \mathbb{K}^n \to \mathbb{K}^n$ è diagonalizzabile (cioè $\mathbb{K}^n$ ha base di autovettori di f_A). Basta applicare la precedente osservazione alla $A = M_{\mathcal{B}}(f_A)$, con $\mathcal{B}$ base canonica di $\mathbb{K}^n$.

Teorema. Sia $f \in L(V)$, e siano $\lambda_1, \dots, \lambda_k$ autovalori distinti di f e $V_{\lambda_1}, \dots, V_{\lambda_k}$ i corrispondenti autospazi. Allora qualunque k-pla di autovettori $\underline{v}_1 \in V_{\lambda_1}, \dots, \underline{v}_k \in V_{\lambda_k}$ (relativi ad autovalori distinti) è indipendente.
[in altri termini: autovettori relativi ad autovalori distinti sono linearmente indipendenti].

---

dim. Per induzione su k. Il caso $k=1$ è ovvio.

Sia $\alpha_1 \underline{v}_1 + \dots + \alpha_k \underline{v}_k = \underline{0}$ (1) una relazione tra i $\underline{v}_j$.

Applicando f troviamo per linearità:

$$f(\alpha_1 \underline{v}_1 + \dots + \alpha_k \underline{v}_k) = \alpha_1 f(\underline{v}_1) + \dots + \alpha_k f(\underline{v}_k) = \lambda_1 \alpha_1 \underline{v}_1 + \dots + \lambda_k \alpha_k \underline{v}_k = \underline{0} \quad (2)$$

Sottraendo dalla (2) la (1) moltiplicata per λ_k si trova:

$$(\lambda_1-\lambda_n)\alpha_1 \underline{v}_1+\cdots+(\lambda_{n-1}-\lambda_n)\alpha_{k-1}\underline{v}_{k-1} = \underline{0}$$

Per induzione su k, gli autovettori $\underline{v}_1, \cdots, \underline{v}_{n-1}$ sono linearmente indipendenti: quindi i coefficienti $(\lambda_1-\lambda_k)\alpha_1, \cdots, (\lambda_{k-1}-\lambda_k)\alpha_{n-1}$ sono nulli. Essendo $\lambda_j \neq \lambda_n$ per $j<n$, ne deduciamo $\alpha_1 = \cdots = \alpha_{k-1} = 0$. Sostituendo in (1) otteniamo quindi $\alpha_k \underline{v}_n = \underline{0}$ e pertanto anche $\alpha_n = 0$.

<u>Corollario</u> Autospazi $V_{\lambda_1}, \cdots, V_{\lambda_k}$ relativi ad autovalori distinti sono in somma diretta.

<u>Corollario</u>. Se $f \in L(V)$, $\dim V = n$, ha n autovalori distinti allora f è diagonalizzabile.

dim Infatti, V ha una base di autovettori e si applica il teorema visto in precedenza.

Sia $f \in \mathcal{L}(V)$, $\dim V = n$, e siano $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tutti gli autovalori distinti di f, e $V_{\lambda_1}, \dots, V_{\lambda_k}$ i corrispondenti autospazi, di dimensione rispett. $m_1 = \mu_g(\lambda_1)$, $m_2 = \mu_g(\lambda_2), \dots, m_k = \mu_g(\lambda_k)$. Possiamo scegliere una base $\underline{v}_1^{(1)}, \dots, \underline{v}_{m_1}^{(1)}$ di V_{λ_1}; $v_1^{(2)}, \dots, v_{m_2}^{(2)}$ di V_{λ_2}, $\dots$, $v_1^{(k)}, \dots, v_{m_k}^{(k)}$ di V_{λ_k} e l'unione di tutti i vettori:

$\underline{v}_1^{(1)}, \dots, \underline{v}_{m_1}^{(1)}, \underline{v}_1^{(2)}, \dots, v_{m_2}^{(2)}, \dots\dots, v_1^{(k)}, \dots, v_{m_k}^{(k)}$ costituisce una base di $V_{\lambda_1} + \dots + V_{\lambda_k} = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}$ (sono in somma diretta per il coroll. preced.).

Quindi V ha una base di autovettori ($\Leftrightarrow$ è diagonalizzabile) se e solo se

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$$

Ricordiamo che $\forall$ autovalore $\lambda \in \mathbb{K}$ vale $\mu_g(\lambda) \leq \mu_a(\lambda)$ e che $\sum_{j=1}^{k} \mu_a(\lambda_j) \leq \deg p_A(\lambda) = n$. Si ha l'uguaglianza se $\mathbb{K} = \mathbb{C}$, ma se $\mathbb{K} \neq \mathbb{C}$ (ad es. $\mathbb{K} = \mathbb{R}$) può valere la disuguaglianza stretta (addirittura abbiamo visto che ci sono endomorfismi di uno spazio vettoriale reale che non hanno autovalori reali).

Abbiamo così dimostrato il seguente teorema.

TEOREMA. Sia $f \in \mathcal{L}(V)$ e siano $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tutti gli autovalori distinti di f. Allora f è diagonalizzabile se e solo se:

1) $\sum_{j=1}^{k} \mu_a(\lambda_j) = n \quad (= \dim V)$;

2) $\mu_g(\lambda_j) = \mu_a(\lambda_j) \,, \qquad \forall j = 1, \dots, k.$

<u>Oss.</u> Sia $f: V \longrightarrow V$ un endomorfismo di uno spazio vettoriale reale. Può succedere che f non è diagonalizzabile, ma lo diventa se si "estendono" gli scalari da $\mathbb{R}$ a $\mathbb{C}$. Ad esempio, la matrice di una rotazione di $\mathbb{R}^2$ $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ abbiamo visto

che non ha autovalori reali, quindi non può essere diagonalizzabile sul campo $\mathbb{R}$. Se però consideriamo la stessa matrice A in $M_2(\mathbb{C})$, allora A ha autovalori distinti $\cos\theta \pm i \operatorname{sen}\theta$ e quindi è simile alla matrice $\begin{bmatrix} \cos\theta + i \operatorname{sen}\theta & 0 \\ 0 & \cos\theta - i \operatorname{sen}\theta \end{bmatrix}$.