Appunti per Geometria 1 - secondo modulo

Monica Idà

May 25, 2016

Contents

1 Applicazioni lineari **3**
- 1.1 Le prime definizioni . . . 3
- 1.2 La dualità . . . 10
- 1.3 Applicazioni lineari e matrici . . . 13
- 1.4 Equazioni di un sottospazio vettoriale . . . 19

2 Autovalori ed autovettori **21**
- 2.1 Radici di un polinomio . . . 21
- 2.2 Matrici simili - Endomorfimi diagonalizzabili . . . 22
- 2.3 Condizioni di diagonalizzabilità . . . 25

3 La forma di Jordan **37**
- 3.1 Spazi quoziente . . . 37
- 3.2 Matrici a blocchi . . . 39
- 3.3 Endomorfismi nilpotenti . . . 42
- 3.4 La forma di Jordan . . . 48

4 Spazi vettoriali euclidei **53**
- 4.1 Forme bilineari . . . 53
- 4.2 Prodotti scalari . . . 55
- 4.3 Esistenza di basi ortonormali . . . 58
- 4.4 Endomorfismi unitari . . . 62
- 4.5 Il teorema spettrale . . . 65

5 Spazi affini **68**
- 5.1 La struttura di spazio affine su uno spazio vettoriale . . . 68
- 5.2 Affinità . . . 73
- 5.3 Equazioni di una affinità . . . 77

6 Spazi euclidei **80**
- 6.1 La struttura di spazio euclideo su uno spazio vettoriale reale . . . 80
- 6.2 Isometrie . . . 82
- 6.3 Le isometrie di E^2 . . . 83

Queste sono le note del secondo modulo del corso di Geometria 1, di cui il primo modulo è stato tenuto dalla prof.M.Manaresi; quindi qui assumiamo noto tutto quanto visto in tale modulo.

Useremo le seguenti notazioni:

Se A e B sono due insiemi, $A \subseteq B$ vuol dire che A è contenuto in o uguale a B, mentre $A \subset B$ vuol dire che A è incluso strettamente in B.

Siano $f : X \to Y$ e $g : Y \to Z$ due funzioni; la composizione di f e g è la funzione

$$\begin{array}{rccc} g \circ f : & X & \to & Z \\ & x & \mapsto & g(f(x)) \end{array}$$

Siano $f : X \to Y$ una funzione, sia $A \subseteq X$, e sia $B \subseteq Y$ tale che $f(A) \subseteq B$; allora possiamo considerare f *ristretta ad* A *e coristretta a* B, che è la funzione:

$$\begin{array}{rccc} f_{|A}^{|B} : & A & \to & B \\ & x & \mapsto & f(x) \end{array}$$

Sia V un K-spazio vettoriale, e U, W due sottospazi di V; se la somma $U + W$ è diretta, diremo che U e W formano una somma diretta; se $V = U \oplus W$, diremo U è un supplementare di W.

1 Applicazioni lineari

1.1 Le prime definizioni

Nel seguito K denota un campo fissato.

Definizione 1.1 Siano V, W due K-spazi vettoriali; una applicazione $f : V \to W$ è detta *lineare* se $\forall v, v' \in V$, $\forall a, b \in K$, si ha:

$$f(av + bv') = af(v) + bf(v')$$

Osserviamo che, in particolare, $f(0) = f(0v) = 0f(v) = 0$.

Nel seguito la frase: " Sia $f : V \to W$ lineare" sottointende che V e W sono spazi vettoriali su uno stesso campo.

Osservazione 1.2 Ricordiamo che se V è un K-spazio vettoriale, $(V, +)$ è un gruppo abeliano; quindi una applicazione lineare $f : V \to W$ è un morfismo dei gruppi additivi che soddisfi la condizione $f(av) = af(v)$ $\forall v \in V$, $\forall a \in K$.

Esempi 1.3 Sono lineari le seguenti applicazioni:

1) $f: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & 3x \end{array}$

2) $g: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x, y) & \mapsto & (x + y, 3x - y) \end{array}$

3) $d: \begin{array}{ccc} \mathbb{R}[x] & \to & \mathbb{R}[x] \\ a_0 + a_1x + \ldots + a_nx^n & \mapsto & a_1 + 2a_2x + \ldots + na_nx^{n-1} \end{array}$

4) $f: \begin{array}{ccc} M_2(\mathbb{R}) & \to & M_2(\mathbb{R}) \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \mapsto & \begin{pmatrix} a-d & b \\ c & 0 \end{pmatrix} \end{array}$

Esempi 1.4 Non sono lineari le seguenti applicazioni:

1) $g: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x, y) & \mapsto & (1 + x, y) \end{array}$

2) $f: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^2 \end{array}$

Infatti: in 1) $g(0) \neq 0$; in 2) $f(1 + 2) = 9, f(1) + f(2) = 5$.

Definizione 1.5 Siano V, W due K-spazi vettoriali.

Una applicazione lineare $f : V \to V$ è detta *endomorfismo di* V.

Una applicazione lineare e biettiva $f : V \to W$ è detta *isomorfismo* (di spazi vettoriali). Se esiste un isomorfismo $f : V \to W$, scriviamo $V \cong W$.

Un isomorfismo $f : V \to V$ è detto *automorfismo di* V.

Proposizione 1.6 *a) Siano* V, W, U K*-spazi vettoriali e* $f : V \to W$*,* $g : W \to U$ *applicazioni lineari. Allora* $g \circ f : V \to U$ *è lineare.*

b) Sia $f : V \to W$ *un isomorfismo. Allora l'applicazione inversa* f^{-1} *è lineare e quindi è un isomorfismo.*

c) Si ha: $V \cong V$*;*

$V \cong W \Rightarrow W \cong V$*;*

$V \cong W, W \cong U \Rightarrow V \cong U$*.*

Dimostrazione *a)* $\forall a, b \in K, \forall v, v' \in V$, si ha:
$(g \circ f)(av+bv') = g(f(av+bv')) = g(af(v)+bf(v')) = ag(f(v))+b(g(f(v')) = a(g \circ f)(v)+b(g \circ f)(v')$.

b) La $f^{-1} : W \to V$ esiste perché f è biettiva. Siano $w, w' \in W$, $a, b \in K$; allora esistono, e sono unici, $v, v' \in V$ tali che $f(v) = w$, $f(v') = w'$, e si ha

$$f^{-1}(aw + bw') = f^{-1}(af(v) + bf(v')) = f^{-1}(f(av + bv')) = av + bv' = af^{-1}(w) + bf^{-1}(w')$$

c) L'identità su V è lineare e biettiva, quindi $V \overset{id}{\cong} V$;

$V \overset{f}{\cong} W \Rightarrow W \overset{f^{-1}}{\cong} V$ per il punto *b)*;

$V \overset{f}{\cong} W, W \overset{g}{\cong} U \Rightarrow V \overset{g \circ f}{\cong} U$ per il punto *a)*, tenendo conto che la composizione di due biezioni è una biezione. ■

Notazione 1.7 Sia V un K-spazio vettoriale finitamente generato (d'ora in poi scriviamo f.g.), e sia $\mathcal{B} = (v_1, \ldots, v_n)$ una base. Il vettore di coordinate $(a_1, \ldots, a_n)$ rispetto alla base $\mathcal{B}$ verrà denotato così:

$$(a_1, \ldots, a_n)_{\mathcal{B}} := a_1v_1 + \ldots + a_nv_n$$

Proposizione-Definizione 1.8 *Sia* V *un* K*-spazio vettoriale f.g., e sia* $\mathcal{B} = (v_1, \ldots, v_n)$ *una base. L'applicazione:*

$$\begin{array}{cccc} \psi_{\mathcal{B}} : & K^n & \to & V \\ & (a_1, \ldots, a_n) & \mapsto & (a_1, \ldots, a_n)_{\mathcal{B}} \end{array}$$

è un isomorfismo di K*-spazi vettoriali, detto* l'isomorfismo tra V e K^n definito dalla base $\mathcal{B}$.

Dimostrazione Sappiamo che $\psi_{\mathcal{B}}$ è biunivoca. Linearità:

$$\psi_{\mathcal{B}}(a(a_1, \ldots, a_n)+b(b_1, \ldots, b_n)) = \psi_{\mathcal{B}}((aa_1+bb_1, \ldots, aa_n+bb_n)) = (aa_1+bb_1)v_1+\ldots+(aa_n+bb_n)v_n =$$

$$= a(a_1v_1 + \ldots + a_nv_n) + b(b_1v_1 + \ldots b_nv_n) = a\psi_{\mathcal{B}}((a_1, \ldots, a_n)) + b\psi_{\mathcal{B}}((b_1, \ldots, b_n))$$

■

Notazione 1.9 Nel seguito denotiamo sempre con $\phi_{\mathcal{B}}$ l'isomorfismo inverso di $\psi_{\mathcal{B}}$, che associa ad ogni vettore $v \in V$, $v = (a_1, \ldots, a_n)_{\mathcal{B}}$, le sue coordinate rispetto alla base $\mathcal{B}$:

$$\begin{array}{cccc} \phi_{\mathcal{B}} : & V & \to & K^n \\ & v & \mapsto & (a_1, \ldots, a_n) \end{array}$$

Esempi 1.10 1) Se $V = K^n$ e $\mathcal{E} = (e_1, \dots, e_n)$ è la base canonica, $\forall (a_1, \dots, a_n) \in K^n$ si ha $(a_1, \dots, a_n) = (a_1, \dots, a_n)_\mathcal{E}$, quindi $\psi_\mathcal{E} = id_{K^n}$.

2) Siano $v_1 = (1, 2), v_2 = (3, 1) \in \mathbb{R}^2$; allora $\mathcal{B} = (v_1, v_2)$ è una base per $\mathbb{R}^2$. Consideriamo ad esempio il vettore $v = (0, 5)$; si ha $v = 3v_1 - v_2$, quindi $v = (3, -1)_\mathcal{B}$ e $\phi_\mathcal{B}(v) = (3, -1)$.

Abbiamo quindi dimostrato il:

Teorema 1.11 *Se V è un K-spazio vettoriale di dimensione finita $n > 0$, allora*

$$V \cong K^n.$$

Teorema 1.12 *Siano V e W K-spazi vettoriali, sia V f.g. e sia $\mathcal{B} = (v_1, \dots, v_n)$ una sua base; siano $w_1, \dots, w_n$ vettori di W. Allora esiste un'unica applicazione lineare $f : V \to W$ tale che*

$$f(v_i) = w_i, \quad i = 1, \dots, n \tag{$*$}$$

Dimostrazione $\exists$): $\forall v \in V$, siano $a_1, \dots, a_n \in K$ tali che $v = \sum a_i v_i$; poniamo

$$f(\sum a_i v_i) := \sum a_i w_i$$

La f così definita è lineare: $\forall v, v' \in V$, $v = \sum a_i v_i, v' = \sum b_i v_i$, $\forall a, b \in K$, si ha:

$$f(av + bv') = f(a \sum a_i v_i + b \sum b_i v_i) = f(\sum (aa_i + bb_i) v_i) =$$

$$= \sum (aa_i + bb_i) w_i = a \sum a_i w_i + b \sum b_i w_i = af(v) + bf(v')$$

!): Siano f, g applicazioni lineari tali che $f(v_i) = w_i$, $g(v_i) = w_i \quad i = 1, \dots, n$; allora $\forall v \in V$, $v = \sum a_i v_i$, si ha:

$$f(v) = f(\sum a_i v_i) \underbrace{=}_{f \text{ lineare}} \sum a_i f(v_i) = \sum a_i w_i = \sum a_i g(v_i) \underbrace{=}_{g \text{ lineare}} g(\sum a_i v_i) = g(v)$$

■

Notazione 1.13 Nelle notazioni di Teorema 1.12 diciamo che l'applicazione lineare f è ottenuta estendendo per linearità le $(*)$

Esempio 1.14 L'applicazione definita in 1.8 è ottenuta estendendo per linearità le $e_i \mapsto v_i$.

Teorema 1.15 *Sia $f : V \to W$ lineare. Allora valgono le seguenti affermazioni:*

a) Se $v_1, \dots, v_n \in V$ sono l.d., allora $f(v_1), \dots, f(v_n)$ sono l.d.

b) Se U è sottospazio di V, allora $f(U) := \{f(u) \mid u \in U\}$ è sottospazio di W.

c) Se T è sottospazio di W, allora $f^{-1}(T) := \{v \in V \mid f(v) \in T\}$ è sottospazio di V.

d) Se $U =< u_1, \dots, u_n >$, allora $f(U) =< f(u_1), \dots, f(u_n) >$.

Dimostrazione *a*) Per ipotesi esistono $a_1, \dots, a_n \in K$ non tutti 0 tali che: $a_1v_1 + \ldots + a_nv_n = 0$; allora

$$0 = f(a_1v_1 + \ldots + a_nv_n) = a_1f(v_1) + \ldots + a_nf(v_n)$$

b) $\forall a, b \in K$, $\forall w, w' \in f(U)$, esistono u, u' tali che $w = f(u), w' = f(u')$, e quindi:

$$aw + bw' = af(u) + bf(u') = f(\underbrace{au + bu'}_{\in U}) \in f(U)$$

c) $\forall a, b \in K$, $\forall v, v' \in f^{-1}(T)$, si ha $f(v), f(v') \in T$, e quindi:

$$f(av + bv') = af(v) + bf(v') \in T \Rightarrow av + bv' \in f^{-1}(T)$$

d) Si ha

$$f(U) = \{w \in W \mid \exists u \in U, w = f(u)\} =$$
$$= \{w \in W \mid \exists a_1, \dots, a_n \in K, w = f(a_1u_1 + \ldots + a_nu_n) = a_1f(u_1) + \ldots + a_nf(u_n)\} =$$
$$=< f(u_1), \dots, f(u_n) >$$

■

Proposizione-Definizione 1.16 *Sia $f : V \to W$ lineare.* Il nucleo di f *è il nucleo di f come morfismo dei gruppi additivi:*

$$Kerf := \{v \in V, f(v) = 0\}$$

e l'immagine di f *è l'immagine insiemistica di f:*

$$Imf := f(V) = \{w \in W, \exists v \in V, w = f(v)\}.$$

Si ha che $Kerf$ è un sottospazio di V, e Imf è un sottospazio di W.

Inoltre:

per definizione f è suriettiva se e solo se $Imf = W$;

f è iniettiva se e solo se $Kerf = \{0\}$;

se $(v_1, \dots, v_n)$ sono generatori per V, $Imf =< f(v_1), \dots, f(v_n) >$.

Dimostrazione $Kerf = f^{-1}(< 0 >)$ e $Imf = f(V)$ sono sottospazi rispettivamente di V e W per Teorema 1.15.

f iniettiva $\Rightarrow Kerf = \{0\}$ è ovvio; viceversa, sia $Kerf = \{0\}$; allora $f(v) = f(v') \Rightarrow f(v - v') = 0 \Rightarrow v - v' \in Kerf \Rightarrow v = v'$. ■

Proposizione 1.17 *Siano V e W K-spazi vettorali f.g., $f : V \to W$ un isomorfismo, $v_1, \dots, v_n \in V$, e $U \subseteq V$, $T \subseteq W$ sottospazi. Allora si ha:*

a) *$v_1, \dots, v_n$ generatori per $V \iff f(v_1), \dots, f(v_n)$ generatori per W;*

b) *$v_1, \dots, v_n$ l.i. $\iff f(v_1), \dots, f(v_n)$ l.i.;*

c) *$(v_1, \dots, v_n)$ base per $V \iff (f(v_1), \dots, f(v_n))$ base per W;*

d) *$dimV = dimW$;*

e) *$dimU = dimf(U)$, $dimT = dimf^{-1}(T)$.*

Dimostrazione $a)$ segue da Teorema 1.15 $d)$; $b)$ segue da Teorema 1.15 $a)$; $c)$ segue da $a)+b)$; $d)$ segue da $c)$; $e)$ si vede così: $f(U)$ e $f^{-1}(T)$ sono sottospazi per Teorema 1.15, e le applicazioni:

$$f_{|U}^{|f(U)}: \begin{array}{ccc} U & \rightarrow & f(U) \\ v & \mapsto & f(v) \end{array}, \qquad f_{|T}^{-1\,|f^{-1}(T)}: \begin{array}{ccc} T & \rightarrow & f^{-1}(T) \\ w & \mapsto & f^{-1}(w) \end{array}$$

sono lineari e biettive, quindi isomorfismi di spazi vettoriali, e si conclude usando $d)$. ■

Proposizione-Definizione 1.18 *Sia V di dimensione finita n, $\mathcal{B}$ base per V, e siano $v_1, \ldots, v_k \in V$, di coordinate rispetto a $\mathcal{B}$:*

$$v_1 = (v_{11}, \ldots, v_{n1})_{\mathcal{B}}, \ldots, v_k = (v_{1k}, \ldots, v_{nk})_{\mathcal{B}}.$$

Allora

$$rg(v_1, \ldots, v_k) := dim < v_1, \ldots, v_k >= rg \begin{pmatrix} v_{11} & \ldots & v_{1k} \\ & \ldots & \\ v_{n1} & \ldots & v_{nk} \end{pmatrix}$$

Dimostrazione Consideriamo l'isomorfismo che associa ad ogni vettore $v \in V$, $v = (a_1, \ldots, a_n)_{\mathcal{B}}$, le sue coordinate rispetto alla base $\mathcal{B}$ (vedi 1.9):

$$\phi_{\mathcal{B}}: \begin{array}{ccc} V & \rightarrow & K^n \\ v & \mapsto & (a_1, \ldots, a_n) \end{array}$$

allora $\phi_{\mathcal{B}}(v_i) = (v_{1i}, \ldots, v_{ni})$; per Prop. 1.17 $e)$,

$$dim < v_1, \ldots, v_k >= dim < (v_{11}, \ldots, v_{n1}), \ldots, (v_{1k}, \ldots, v_{nk}) >= rg \begin{pmatrix} v_{11} & \ldots & v_{1k} \\ v_{n1} & \ldots & v_{nk} \end{pmatrix}.$$

■

Teorema 1.19 *Sia $f : V \rightarrow W$ lineare, e sia V f.g.. Allora $Kerf$ e Imf hanno dimensione finita e si ha:*

$$dimV = dim\,Kerf + dim\,Imf$$

Dimostrazione Poiché $Kerf$ è un sottospazio di V, anche $Kerf$ è f.g.; sia quindi $(v_1, \ldots, v_s)$ una sua base, e completiamola ad una base $(v_1, \ldots, v_s, v_{s+1}, \ldots, v_n)$ di V.

Basta quindi provare $dim\,Imf = n - s$; facciamo vedere che $f(v_{s+1}), \ldots, f(v_n)$ è una base per Imf:

$f(v_{s+1}), \ldots, f(v_n)$ sono generatori: $Imf =< \underbrace{f(v_1)}_{=0}, \ldots, \underbrace{f(v_s)}_{=0}, f(v_{s+1}), \ldots, f(v_n) >$

$f(v_{s+1}), \ldots, f(v_n)$ sono l.i.: siano $a_{s+1}, \ldots, a_n \in K$ tali che:

$$0 = a_{s+1}f(v_{s+1}) + \ldots + a_nf(v_n) = f(a_{s+1}v_{s+1} + \ldots + a_nv_n) \Rightarrow a_{s+1}v_{s+1} + \ldots + a_nv_n \in Kerf$$

$$\Rightarrow \exists a_1, \ldots, a_s \text{ tali che } a_{s+1}v_{s+1} + \ldots + a_nv_n = a_1v_1 + \ldots + a_sv_s \Rightarrow$$

$$a_1v_1 + \ldots + a_sv_s - a_{s+1}v_{s+1} - \ldots - a_nv_n = 0 \underbrace{\Rightarrow}_{v_1,\ldots,v_n\ l.i.} a_1 = \ldots = a_n = 0$$

Definizione 1.20 *Sia $f : V \to W$ lineare, sia V f.g., e sia $(v_1, \ldots, v_n)$ una base per V; si chiama* rango di f *l'intero*

$$rg(f) := dim\, Imf = rg(f(v_1), \ldots, f(v_n))$$

Corollario 1.21 *Siano V, W spazi vettoriali f.g. con $dimV = dimW = n$, e sia $f : V \to W$ lineare. Sono equivalenti:*

a) $Kerf =< 0 >$

b) $Imf = W$

c) f *isomorfismo.*

Dimostrazione $c) \Rightarrow a)$ e $c) \Rightarrow b)$ sono ovvi.
$a) \Rightarrow c)$: da Teorema 1.19 si ha

$$dimV = 0 + dim\, Imf \Rightarrow dim\, Imf = n \Rightarrow Imf = W$$

quindi f è $1-1$ e su, cioé f isomorfismo; $b) \Rightarrow c)$ è analogo.

Teorema 1.22 *Due K-spazi vettoriali V e W f.g. sono isomorfi se e solo se $dimV = dimW$.*

Dimostrazione Prop. 1.17 dice che $V \cong W \Rightarrow dimV = dimW$. Viceversa:

$$dimV = dimW = n \underbrace{\Rightarrow}_{1.11} V \cong K^n, W \cong K^n \underbrace{\Rightarrow}_{1.6} V \cong W.$$

Esempio 1.23 Siano V e W K-spazi vettoriali con $dimV = dimW$, e siano $\mathcal{B} = (v_1, \ldots, v_n)$, $\mathcal{C} = (w_1, \ldots, w_n)$ basi rispettivamente per V e W. Per costruire un isomorfismo tra V e W basta considerare l'applicazione lineare (vedi 1.12):

$$\begin{array}{cccc} f: & V & \to & W \\ & v_1 & \mapsto & w_1 \\ & \vdots & \vdots & \vdots \\ & v_n & \mapsto & w_n \end{array}$$

che è un isomorfismo; infatti $Imf =< f(v_1), \ldots, f(v_n) >=< w_1, \ldots, w_n >= W$.

Esempio 1.24 Consideriamo l'applicazione lineare:

$$\begin{array}{cccc} g: & M_2(\mathbb{R}) & \to & \mathbb{R}^4 \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \mapsto & (a, b, c, d) \end{array}$$

$Kerg =< 0 >$, quindi per Cor. 1.21 è un isomorfismo di $\mathbb{R}$-spazi vettoriali; si ha $dimM_2(\mathbb{R}) = dim\mathbb{R}^4 = 4$.

Quindi i sottospazi di $M_2(\mathbb{R})$ sono tutti e soli i sottoinsiemi di $M_2(\mathbb{R})$ della forma

$$\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}), \quad h_1(a, b, c, d) = 0, \ldots h_s(a, b, c, d) = 0, \ h_i \text{ lineare omogeneo}\}$$

Per esempio,

$$U := \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}), \quad a-b+2c=0,\, 3a+b=0\} = \{\begin{pmatrix} a & -3a \\ -2a & d \end{pmatrix},\ a,d \in \mathbb{R}\}$$

è un sottospazio di dimensione 2, perché $rg\begin{pmatrix} 1 & -1 & 2 & 0 \\ 3 & 1 & 0 & 0 \end{pmatrix} = 2$, quindi il sottospazio di $\mathbb{R}^4$ $g(U) = \{(a,b,c,d) \in \mathbb{R}^4 \mid a-b+2c=0,\, 3a+b=0\}$ ha dimensione 2.

Esempio 1.25 Consideriamo l'applicazione lineare:

$$\begin{array}{cccc} f: & M_2(\mathbb{R}) & \to & M_2(\mathbb{R}) \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \mapsto & \begin{pmatrix} a-d & b \\ c & 0 \end{pmatrix} \end{array}$$

Ci chiediamo se f è un isomorfismo; studiamo ad esempio il nucleo di f:

$$Kerf = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}), \quad a-d=0,\, b=0, c=0\} = \{\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix},\ a \in \mathbb{R}\} =< I_2 > \neq 0$$

quindi f non è un isomorfismo, e $dimKerf = 1 \Rightarrow dimImf = 3$.

Se vogliamo conoscere una base di Imf, possiamo fare ad esempio così: utilizziamo l'isomorfismo g visto nell'esempio precedente; allora

$$Imf = \{\begin{pmatrix} a-d & b \\ c & 0 \end{pmatrix} \in M_2(\mathbb{R}),\ a,b,c,d \in \mathbb{R}\} \ \Rightarrow g(Imf) = \{(a-d,b,c,0) \in \mathbb{R}^4,\ a,b,c,d \in \mathbb{R}\}$$

e si ha

$$\{(a-d,b,c,0) \in \mathbb{R}^4,\ a,b,c,d \in \mathbb{R}\} = \{(a-d)(1,0,0,0) + b(0,1,0,0) + c(0,0,1,0),\ a,b,c,d \in \mathbb{R}\} =$$

$$=< (1,0,0,0),(0,1,0,0),(0,0,1,0) >$$

essendo i vettori e_1, e_2, e_3 l.i., essi formano una base per $g(Imf)$, quindi una base per Imf è data da:

$$(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix})$$

Osservazione 1.26 Dalle precedenti proposizioni ed esempi si capisce che la filosofia è questa: per ogni questione attinente all'algebra lineare si può sostituire V con W se se V e W sono spazi vettoriali isomorfi.

Adesso che abbiamo definito le applicazioni lineari, passiamo a studiare gli insiemi di applicazioni lineari, che come vedremo hanno anch'essi una struttura di K-spazio vettoriale.

Definizione 1.27 Siano V, W K-spazi vettoriali; poniamo:

$$Hom(V,W) := \{f, \quad f: V \to W \text{ lineare}\}$$

$$End(V) := Hom(V,V) = \{f, \quad f: V \to V \text{ lineare}\}$$

$$V^* := Hom(V,K) = \{f, \quad f: V \to K \text{ lineare}\}$$

$$GL(V) := \{f, \quad f: V \to V \text{ automorfismo}\} \subset \text{End(V)}$$

Proposizione-Definizione 1.28 *Siano V e W due K-spazi vettoriali; per ogni $f, g \in Hom(V, W)$, per ogni $\lambda \in K$, definiamo le applicazioni lineari $f + g$ e λf così:*

$$\begin{array}{rccc} f+g: & V & \to & W \\ & v & \mapsto & f(v)+g(v) \end{array} \qquad \begin{array}{rccc} \lambda f: & V & \to & W \\ & v & \mapsto & \lambda f(v) \end{array}$$

Questo dà due operazioni somma e prodotto per scalari:

$$\begin{array}{rccc} +: & Hom(V,W)\times Hom(V,W) & \to & Hom(V,W) \\ & (f,g) & \mapsto & f+g \end{array} \qquad \begin{array}{rccc} \cdot: & K\times Hom(V,W) & \to & Hom(V,W) \\ & (\lambda,f) & \mapsto & \lambda f \end{array}$$

che rendono $Hom(V, W)$ un K-spazio vettoriale.

Dimostrazione Basta verificare che le operazioni sono ben definite, cioé che $f+g$ e λf sono lineari, e che $Hom(V, W)$ verifica le proprietà richieste ad un K-spazio vettoriale (esercizio).
Notiamo che, in particolare, lo 0 di $Hom(V, W)$ è l'applicazione nulla:

$$\begin{array}{rccc} 0: & V & \to & W \\ & v & \mapsto & 0 \end{array}$$

■

Osservazione 1.29 In particolare quindi, $End(V)$ e V^* sono K-spazi vettoriali; V^* è detto *lo spazio duale di V*.

Si osservi che $GL(V)$ non è un K-spazio vettoriale con queste operazioni: per esempio, se $f \in GL(V)$, anche $-f \in GL(V)$, ma $f - f = 0 \notin GL(V)$.

1.2 La dualità

Proposizione-Definizione 1.30 *Sia V un K-spazio vettoriale di dimensione finita, e sia $\mathcal{B} = (v_1, \ldots, v_n)$ una base di V. Per $i = 1, \ldots, n$ sia $\eta_i : V \to K$ l'applicazione lineare così definita:*

$$\eta_i(v_j) := \begin{cases} 0 & \text{se i} \neq \text{j} \\ 1 & \text{se i} = \text{j} \end{cases}$$

Allora $\mathcal{B}^* = (\eta_1, \ldots, \eta_n)$ è una base per V^*, detta la base duale di $\mathcal{B}$.

In particolare, $dimV = dimV^*$ e quindi $V \cong V^*$.

Dimostrazione Le η_i sono ben definite (1.12).

Proviamo $V^* =< \eta_1, \ldots, \eta_n >$:
sia $f \in V^*$; allora risulta $f = f(v_1)\eta_1 + \ldots + f(v_n)\eta_n$, infatti queste due applicazioni lineari prendono gli stessi valori sui vettori della base $\mathcal{B}$:

$$(f(v_1)\eta_1 + \ldots + f(v_n)\eta_n)(v_i) = f(v_1)\eta_1(v_i) + \ldots + f(v_n)\eta_n(v_i) = f(v_i)\eta_i(v_i) = f(v_i) \quad i = 1, \ldots, n$$

Proviamo $\eta_1, \ldots, \eta_n$ l.i.: se $a_1, \ldots, a_n \in K$, e $a_1\eta_1 \ldots + a_n\eta_n = 0$, si ha:

$$a_i = (a_1\eta_1 \ldots + a_n\eta_n)(v_i) = 0(v_i) = 0 \quad i = 1, \ldots, n$$

■

Osservazione 1.31 Nelle notazioni precedenti, per quanto visto nella dimostrazione, se $f \in V^*$ si ha:

$$f = (f(v_1), \dots, f(v_n))_{\mathcal{B}^*}$$

cioé le coordinate di una applicazione lineare $f : V \to K$ rispetto alla base duale di $\mathcal{B}$ sono i suoi valori sui vettori di $\mathcal{B}$.

Osservazione 1.32 Nelle notazioni precedenti, si ha

$$\eta_i((x_1, \dots, x_n)_{\mathcal{B}}) = x_i$$

infatti $\eta_i((x_1, \dots, x_n)_{\mathcal{B}}) = \eta_i(x_1 v_1 + \dots + x_n v_n) = x_1 \eta_i(v_1) + \dots + x_n \eta_i(v_n) = x_i$, cioé l'applicazione lineare $\eta_i : V \to K$ associa ad un vettore la sua i-esima coordinata rispetto alla base $\mathcal{B}$.

Quindi, se $a_1, \dots, a_n, b_1, \dots, b_n \in K$, si ha:

$$(a_1\eta_1 + \dots + a_n\eta_n)(b_1 v_1 + \dots + b_n v_n) = a_1 b_1 + \dots + a_n b_n.$$

Esempi 1.33 *a*) Sia $V = \mathbb{R}^3$, sia $\mathcal{E} = (e_1, e_2, e_3)$ la base canonica e sia $\mathcal{E}^* = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ la base duale; allora

$$\varepsilon_1((x, y, z)) = x, \quad \varepsilon_2((x, y, z)) = y, \quad \varepsilon_3((x, y, z)) = z$$

Consideriamo l'applicazione lineare

$$\begin{array}{rccc} f: & \mathbb{R}^3 & \to & \mathbb{R} \\ & (x, y, z) & \mapsto & 2x + z \end{array}$$

allora $f(e_1) = 2, f(e_2) = 0, f(e_3) = 1$, quindi in V^* si ha $f = (2, 0, 1)_{\mathcal{E}^*} = 2\varepsilon_1 + \varepsilon_3$.

b) Sia $V = \mathbb{R}^3$, con la base $\mathcal{B} = \underbrace{(1, 1, 1)}_{v_1}, \underbrace{(1, 0, 1)}_{v_2}, \underbrace{(0, 0, 2)}_{v_3}$, e sia $\mathcal{B}^* = (\eta_1, \eta_2, \eta_3)$ la base duale. Allora se $v = (5, 2, 7)$, si ha $v = 2v_1 + 3v_2 + v_3$, da cui per esempio

$$(3\eta_1 - \eta_2 + 2\eta_3)(v) = (3\eta_1 - \eta_2 + 2\eta_3)(2v_1 + 3v_2 + v_3) = 3 \cdot 2 + (-1) \cdot 3 + 2 \cdot 1 = 5$$

Definizione 1.34 Sia V un K-spazio vettoriale; lo spazio vettoriale $(V^*)^*$ è detto lo spazio biduale di V, e viene denotato con V^{**}.

Osservazione 1.35 Se V ha dimensione finita n, allora

$$V \cong V^* \cong V^{**}$$

e $dimV = dimV^{**}$ per 1.30.

Sia $\mathcal{B} = (v_1, \dots, v_n)$ una base di V, e sia $\mathcal{B}^* = (\eta_1, \dots, \eta_n)$; allora un possibile isomorfismo è dato da (vedi 1.23):

$$\begin{array}{rccc} \phi: & V & \to & V^* \\ & v_1 & \mapsto & \eta_1 \\ & \vdots & \vdots & \vdots \\ & v_n & \mapsto & \eta_n \end{array}$$

Questa ϕ dipende però dalla base scelta, come mostra il seguente esempio.

Esempio 1.36 Sia $V = \mathbb{R}^2$, sia $\mathcal{E} = (e_1, e_2)$ la base canonica e sia $\mathcal{E}^* = (\varepsilon_1, \varepsilon_2)$ la base duale; sia $v_2 = (1, 1)$ e consideriamo anche la base $\mathcal{B} = (e_1, v_2)$, e la sua duale $\mathcal{B}^* = (\eta_1, \eta_2)$.

Osserviamo che $\varepsilon_1 \neq \eta_1$, infatti per esempio:

$$\varepsilon_1(e_2) = 0, \quad \eta_1(e_2) = \eta_1(-e_1 + v_2) = -1.$$

Consideriamo i due isomorfismi

$$\begin{array}{rccc} \phi: & V & \to & V^* \\ & e_1 & \mapsto & \varepsilon_1 \\ & e_2 & \mapsto & \varepsilon_2 \end{array} \qquad \begin{array}{rccc} \psi: & V & \to & V^* \\ & e_1 & \mapsto & \eta_1 \\ & v_2 & \mapsto & \eta_2 \end{array}$$

Si ha

$$\phi(e_1) = \varepsilon_1 \neq \eta_1 = \psi(e_1) \ \Rightarrow \ \phi \neq \psi.$$

Teorema 1.37 *Sia V un K-spazio vettoriale di dimensione finita; esiste un isomorfismo canonico (nel senso che dipende solo da V, e non dalla scelta di una base di V):*

$$\begin{array}{rccc} \beta: & V & \stackrel{\cong}{\to} & V^{**} \\ & v & \mapsto & v^{**} \end{array}$$

dove

$$\begin{array}{rccc} v^{**}: & V^* & \to & K \\ & f & \mapsto & f(v) \end{array}$$

Dimostrazione - β è ben definita perché $v^{**} : V^* \to K$ è lineare, infatti:

$$v^{**}(af + bg) = (af + bg)(v) = af(v) + bg(v) = av^{**}(f) + bv^{**}(g), \qquad \forall a, b \in K, \ \forall f, g \in V^*$$

- β è lineare:

$$\beta(av + bw) = (av + bw)^{**} \stackrel{?}{=} a\beta(v) + b\beta(w) = av^{**} + bw^{**} \qquad \forall a, b \in K, \ \forall v, w \in V$$

Questo è vero poiché si ha, $\forall f \in V^*$:

$$(av + bw)^{**}(f) = f(av + bw) = af(v) + bf(w) = av^{**}(f) + bw^{**}(f) = (av^{**} + bw^{**})(f)$$

- β è $1-1$ perché $Ker\beta =< 0 >$, infatti:

$$v \in Ker\beta \iff v^{**} = 0 \iff v^{**}(f) = 0 \ \ \forall f \in V^* \iff f(v) = 0 \ \ \forall f \in V^* \stackrel{(\heartsuit)}{\iff} v = 0$$

proviamo $(\heartsuit)$:
se $v = 0 \Rightarrow f(v) = 0 \ \ \forall f \in V^*$ è banalmente vero; viceversa, supponiamo $f(v) = 0 \ \ \forall f \in V^*$ e $v \neq 0$, e completiamo v ad una base $(v, v_2, \ldots, v_n)$ di V; allora l'applicazione lineare

$$\begin{array}{rccc} f: & V & \to & K \\ & v & \mapsto & 1 \\ & v_2 & \mapsto & 0 \\ & \vdots & \vdots & \vdots \\ & v_n & \mapsto & 0 \end{array}$$

non si annulla in v, contraddizione.

- β è su perché è una applicazione lineare $1-1$ tra spazi vettoriali f.g. della stessa dimensione. ∎

1.3 Applicazioni lineari e matrici

Definizione 1.38 Siano V, W due K-spazi vettoriali f.g., e siano $\mathcal{B} = (v_1, \ldots, v_n)$ base per V, $\mathcal{C} = (w_1, \ldots, w_s)$ base per W.

Sia $f : V \to W$ una applicazione lineare; la matrice $s \times n$ la cui i-esima colonna è data dalle coordinate del vettore $f(v_i)$ rispetto alla base $\mathcal{C}$ si chiama *la matrice associata ad f rispetto alle basi $\mathcal{B}$ nel dominio e $\mathcal{C}$ nel codominio*, e si denota con $M_{\mathcal{C},\mathcal{B}}(f)$:

$$M_{\mathcal{C},\mathcal{B}}(f) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{s1} & \ldots & a_{sn} \end{pmatrix}$$

dove

$$f(v_i) = a_{1i}w_1 + \ldots + a_{si}w_s = (a_{1i}, \ldots, a_{si})_{\mathcal{C}}$$

Proposizione 1.39 *Siano V, W due K-spazi vettoriali f.g., siano $\mathcal{B} = (v_1, \ldots, v_n)$ base per V, $\mathcal{C} = (w_1, \ldots, w_s)$ base per W, e sia $f : V \to W$ una applicazione lineare.*
Allora se $v = (x_1, \ldots, x_n)_{\mathcal{B}}$ e $f(v) = (y_1, \ldots, y_s)_{\mathcal{C}}$, risulta:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix} = M_{\mathcal{C},\mathcal{B}}(f) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Dimostrazione Sia

$$M_{\mathcal{C},\mathcal{B}}(f) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{s1} & \ldots & a_{sn} \end{pmatrix}.$$

Allora

$$f(v) = f(x_1v_1 + \ldots + x_nv_n) = x_1f(v_1) + \ldots + x_nf(v_n) =$$

$$= x_1(a_{11}w_1+\ldots+a_{s1}w_s)+\ldots+x_n(a_{1n}w_1+\ldots+a_{sn}w_s) = (a_{11}x_1+\ldots+a_{1n}x_n)w_1+\ldots+(a_{s1}x_1+\ldots+a_{sn}x_n)w_s$$

cioé

$$\begin{cases} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ y_s = a_{s1}x_1 + \ldots + a_{sn}x_n \end{cases} \qquad (*)$$

Le $(*)$ si possono scrivere in forma matriciale così:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{s1} & \ldots & a_{sn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad (**)$$

■

Definizione 1.40 Le $(*)$ o equivalentemente le $(**)$ vengono dette *le equazioni della f rispetto alle basi $\mathcal{B}$ nel dominio e $\mathcal{C}$ nel codominio.*

Ponendo $X := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $Y := \begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix}$, le $(**)$ si scrivono:

$$Y = M_{\mathcal{C},\mathcal{B}}(f)\, X$$

Esempio 1.41 Consideriamo l'applicazione lineare:

$$\begin{array}{cccc} f: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & e_1 & \mapsto & (1,3) \\ & e_2 & \mapsto & (-2,7) \end{array}$$

quindi f agisce sul generico vettore $(x,y) \in \mathbb{R}^2$ così:

$$f((x,y)) = f(xe_1 + ye_2) = xf(e_1) + yf(e_2) = x(1,3) + y(-2,7) = (x-2y, 3x+7y)$$

e si ha

$$M_{\mathcal{E},\mathcal{E}}(f) = \begin{pmatrix} 1 & -2 \\ 3 & 7 \end{pmatrix}$$

Siano $v_1 = (1,1), v_2 = (2,-1)$; poiché $det \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \neq 0$, v_1, v_2 sono l.i.; sia $\mathcal{B} = (v_1, v_2)$; cerchiamo $M_{\mathcal{E},\mathcal{B}}(f)$. Si ha:

$$f(v_1) = f((1,1)) = (-1,10), \quad f(v_2) = f((2,-1)) = (4,-1)$$

per cui

$$M_{\mathcal{E},\mathcal{B}}(f) = \begin{pmatrix} -1 & 4 \\ 10 & -1 \end{pmatrix}$$

Quindi se $v = (1,3)_{\mathcal{B}}$, si ha $f(v) = (11,7)$ perché

$$\begin{pmatrix} -1 & 4 \\ 10 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 7 \end{pmatrix}$$

Definizione 1.42 Siano V, W due K-spazi vettoriali f.g. e siano $\mathcal{B} = (v_1, \ldots, v_n)$ base per V, $\mathcal{C} = (w_1, \ldots, w_s)$ base per W. Sia $A = (a_{ij}) \in M_{s,n}(K)$; si dice *applicazione lineare associata ad A rispetto alle basi $\mathcal{B}$ nel dominio e $\mathcal{C}$ nel codominio* l'applicazione, che si verifica essere lineare,

$$\begin{array}{cccc} F^{\mathcal{C},\mathcal{B}}(A): & V & \to & W \\ & (x_1, \ldots, x_n)_{\mathcal{B}} & \mapsto & (y_1, \ldots, y_s)_{\mathcal{C}} \end{array}$$

dove

$$\begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

cioé $F^{\mathcal{C},\mathcal{B}}(A)(v_i) := a_{1i}w_1 + \ldots + a_{si}w_s$.

Proposizione 1.43 *Siano V, W due K-spazi vettoriali f.g. e siano $\mathcal{B} = (v_1, \dots, v_n)$ base per V, $\mathcal{C} = (w_1, \dots, w_s)$ base per W. L'applicazione*

$$\begin{array}{rccc} M_{\mathcal{C},\mathcal{B}}: & Hom(V,W) & \rightarrow & M_{s,n}(K) \\ & f & \mapsto & M_{\mathcal{C},\mathcal{B}}(f) \end{array}$$

è un isomorfismo di K-spazi vettoriali, con inversa

$$\begin{array}{rccc} F^{\mathcal{C},\mathcal{B}}: & M_{s,n}(K) & \rightarrow & Hom(V,W) \\ & A & \mapsto & F^{\mathcal{C},\mathcal{B}}(A) \end{array}$$

In particolare, $dim Hom(V,W) = ns$.

Dimostrazione - *linearità di $M_{\mathcal{C},\mathcal{B}}$*: bisogna provare che $\forall f, g \in Hom(V,W)$, $\forall a, b \in K$,

$$M_{\mathcal{C},\mathcal{B}}(af + bg) = aM_{\mathcal{C},\mathcal{B}}(f) + bM_{\mathcal{C},\mathcal{B}}(g)$$

questo è vero perché vale $(af + bg)(v_i) = af(v_i) + bg(v_i)$, quindi queste uguaglianze si mantengono scrivendo i vettori in coordinate rispetto a $\mathcal{C}$; basta ora ricordare la definizione di somma di due matrici e di prodotto di uno scalare per una matrice.

- *biettività di $M_{\mathcal{C},\mathcal{B}}$*: risulta dalla definizione

$$M_{\mathcal{C},\mathcal{B}}(F^{\mathcal{C},\mathcal{B}}(A)) = A \quad \forall A \in M_{s,n}(K),$$

$$F^{\mathcal{C},\mathcal{B}}(M_{\mathcal{C},\mathcal{B}}(f)) = f \quad \forall f \in Hom(V,W)$$

Quindi $M_{\mathcal{C},\mathcal{B}}$ è biettiva con inversa $F^{\mathcal{C},\mathcal{B}}$. ■

Osservazione 1.44 Siano V, W due K-spazi vettoriali f.g., siano $\mathcal{B} = (v_1, \dots, v_n)$ base per V, $\mathcal{C} = (w_1, \dots, w_s)$ base per W, e sia $f : V \rightarrow W$ una applicazione lineare.
Da 1.39 si ha che le coordinate di $f(v)$ rispetto a $\mathcal{C}$ si scrivono come polinomi lineari omogenei nelle coordinate di v rispetto a $\mathcal{B}$; infatti, come visto nella dimostrazione di 1.39, se $v = (x_1, \dots, x_n)_{\mathcal{B}}$, $f(v) = (y_1, \dots, y_s)_{\mathcal{C}}$, e $M_{\mathcal{C},\mathcal{B}}(f) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{s1} & \dots & a_{sn} \end{pmatrix}$, si ha:

$$\begin{cases} y_1 = a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ y_s = a_{s1}x_1 + \dots + a_{sn}x_n \end{cases} \qquad (*)$$

Viceversa, si verifica facilmente che se $f : V \rightarrow W$ è una applicazione tale che le coordinate di $f(v)$ rispetto a $\mathcal{C}$ si scrivono come polinomi lineari omogenei nelle coordinate di v rispetto a $\mathcal{B}$, cioé una applicazione della forma:

$$\begin{array}{rccc} f: & V & \rightarrow & W \\ & (x_1, \dots, x_n)_{\mathcal{B}} & \mapsto & (b_{11}x_1 + \dots + b_{1n}x_n, \dots, b_{s1}x_1 + \dots + b_{sn}x_n)_{\mathcal{C}} \end{array}$$

dove i $b_{ij} \in K$, allora f è lineare. Infatti, posto

$$B := \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \vdots & \vdots \\ b_{s1} & \dots & b_{sn} \end{pmatrix}$$

basta osservare che si ha

$$f = F^{\mathcal{C},\mathcal{B}}(B).$$

Siamo quindi in grado di riconoscere subito se una applicazione data mediante le coordinate sia lineare o meno, e di scrivere facilmente esempi di applicazioni lineari.

Proposizione 1.45 *Siano U, V, W K-spazi vettoriali f.g. di basi rispettivamente $\mathcal{A} = (u_1, \ldots, u_t)$, $\mathcal{B} = (v_1, \ldots, v_n)$, $\mathcal{C} = (w_1, \ldots, w_s)$. Siano $g : U \to V$ e $f : V \to W$ applicazioni lineari. Allora*

$$M_{\mathcal{C},\mathcal{A}}(f \circ g) = M_{\mathcal{C},\mathcal{B}}(f) M_{\mathcal{B},\mathcal{A}}(g)$$

dove il prodotto tra matrici è il prodotto righe per colonne.

Dimostrazione

$$\begin{array}{ccccc} & g & & f & \\ U & \to & V & \to & W \\ \mathcal{A} & & \mathcal{B} & & \mathcal{C} \end{array}$$

Siano

$$u = (z_1, \ldots, z_t)_{\mathcal{A}}, \quad v = g(u) = (x_1, \ldots, x_n)_{\mathcal{B}}, \quad w = f(v) = f(g(u)) = (y_1, \ldots, y_s)_{\mathcal{C}}$$

allora

$$\begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix} = M_{\mathcal{C},\mathcal{B}}(f) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = M_{\mathcal{C},\mathcal{B}}(f) \left(M_{\mathcal{B},\mathcal{A}}(g) \begin{pmatrix} z_1 \\ \vdots \\ z_t \end{pmatrix} \right) = (M_{\mathcal{C},\mathcal{B}}(f) M_{\mathcal{B},\mathcal{A}}(g)) \begin{pmatrix} z_1 \\ \vdots \\ z_t \end{pmatrix}$$

■

Osservazione 1.46 a) Siano V, W due K-spazi vettoriali f.g., siano $\mathcal{B} = (v_1, \ldots, v_n)$ base per V, $\mathcal{C} = (w_1, \ldots, w_s)$ base per W, e sia $f : V \to W$ una applicazione lineare; poniamo

$$A := M_{\mathcal{C},\mathcal{B}}(f) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{s1} & \ldots & a_{sn} \end{pmatrix}.$$

Allora si ha

$$Ker f = \{(x_1, \ldots, x_n)_{\mathcal{B}} \in V \mid A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0\}, \qquad Im f = < (a_{11} \ldots a_{s1})_{\mathcal{C}}, \ldots, (a_{1n} \ldots a_{sn})_{\mathcal{C}} >$$

Ne segue

$$rg(f) = dim\, Im f = rg(f(v_1), \ldots, f(v_n)) = rg(A),$$

e

$$dim Ker f = n - rg(A).$$

b) Nelle notazioni precedenti supponiamo inoltre $dim W = dim V = n$. Da a) segue:

$$f \text{ isomorfismo} \iff M_{\mathcal{C},\mathcal{B}}(f) \in GL_n(K) \qquad (\diamond)$$

Notazione 1.47 Sia V uno spazio vettoriale f.g., sia $B = (v_1, \dots, v_n)$ una base di V, sia $f \in End(V)$ e sia $A \in M_n(K)$. Poniamo:

$M_{\mathcal{B}}(f) := M_{\mathcal{B},\mathcal{B}}(f)$ (detta *la matrice di f rispetto a $\mathcal{B}$*)

$F^{\mathcal{B}}(A) := F^{\mathcal{B},\mathcal{B}}(A)$ (detto *l'endomorfismo di V associato ad A rispetto a $\mathcal{B}$*)

Proposizione 1.48 *Sia V uno spazio vettoriale f.g., sia $\mathcal{B} = (v_1, \dots, v_n)$ una base di V e sia $f \in End(V)$. Allora valgono i seguenti fatti:*

$$f = id_V \iff M_{\mathcal{B}}(f) = I_n;$$
$$f \in GL(V) \iff M_{\mathcal{B}}(f) \in GL_n(K);$$

se $f \in GL(V)$ o equivalentemente se $M_{\mathcal{B}}(f) \in GL_n(K)$, allora

$$M_{\mathcal{B}}(f^{-1}) = M_{\mathcal{B}}(f)^{-1}.$$

Osserviamo che se $\mathcal{B}$ e $\mathcal{C}$ sono due basi, allora $M_{\mathcal{C},\mathcal{B}}(id_V) \neq I_n \iff \mathcal{B} \neq \mathcal{C}$.

Dimostrazione La prima affermazione segue dalla definizione di matrice associata, la seconda è un caso particolare di 1.46 *b*), la terza si vede così:

$$f \circ f^{-1} = id_V \iff M_{\mathcal{B}}(f \circ f^{-1}) = I_n \iff M_{\mathcal{B}}(f)M_{\mathcal{B}}(f^{-1}) = I_n.$$

Corollario 1.49 *Nelle notazioni precedenti, l'isomorfismo di K-spazi vettoriali*

$$\begin{array}{rccc} M_{\mathcal{B}}: & End(V) & \to & M_n(K) \\ & f & \mapsto & M_{\mathcal{B}}(f) \end{array}$$

è anche un isomorfismo di anelli, prendendo come operazioni su $End(V)$ la somma usuale e la composizione di funzioni come prodotto, e su $M_n(K)$ la somma usuale e il prodotto righe per colonne.

Dimostrazione Essendo un isomorfismo di spazi vettoriali lo è in particolare di gruppi additivi; poiché per 1.45 $M_{\mathcal{B}}(f \circ g) = M_{\mathcal{B}}(f)M_{\mathcal{B}}(g)$, e $M_{\mathcal{B}}(id_V) = I_n$, è anche un morfismo di anelli, dunque essendo biettiva è un isomorfismo. ■

Definizione 1.50 Sia V un K-spazio vettoriale f.g. e siano $\mathcal{B} = (v_1, \dots, v_n)$, $\mathcal{C} = (w_1, \dots, w_n)$ due sue basi.

La matrice $M_{\mathcal{C},\mathcal{B}}(id_V)$ è detta *la matrice del cambiamento di coordinate dalla base $\mathcal{B}$ alla base $\mathcal{C}$*, o anche *la matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{C}$*; si ha

$$M_{\mathcal{C},\mathcal{B}}(id_V) = \begin{pmatrix} v_{11} & \dots & v_{1n} \\ \vdots & \vdots & \vdots \\ v_{n1} & \dots & v_{nn} \end{pmatrix}$$

dove la colonna i-esima è formata dalle coordinate di v_i rispetto alla base $\mathcal{C}$.

Risulta, se $v = (x_1, \dots, x_n)_{\mathcal{B}} = (y_1, \dots, y_n)_{\mathcal{C}}$:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = M_{\mathcal{C},\mathcal{B}}(id_V) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Proposizione 1.51 *Nelle notazioni di 1.50 si ha*

$$M_{\mathcal{C},\mathcal{B}}(id_V)^{-1} = M_{\mathcal{B},\mathcal{C}}(id_V)$$

Dimostrazione

$$M_{\mathcal{C},\mathcal{B}}(id_V)M_{\mathcal{B},\mathcal{C}}(id_V) = M_{\mathcal{C},\mathcal{C}}(id_V) = I_n, \qquad M_{\mathcal{B},\mathcal{C}}(id_V)M_{\mathcal{C},\mathcal{B}}(id_V) = M_{\mathcal{B},\mathcal{B}}(id_V) = I_n.$$

1.4 Equazioni di un sottospazio vettoriale

Nel seguito sono fissate le seguenti notazioni: V è un K-spazio vettoriale, $dimV = n$, $\mathcal{B} = (v_1, .., v_n)$ è una base per V.

È facile provare le seguenti affermazioni:

(equazioni cartesiane di un sottospazio vettoriale)

a) Siano $M = \begin{pmatrix} m_{11} & \dots & m_{1n} \\ \vdots & & \vdots \\ m_{s1} & \dots & m_{sn} \end{pmatrix} \in M_{s,n}(K), \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \; 0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$

Allora

$$W := \{(x_1, .., x_n)_{\mathcal{B}} \in V \mid MX = 0\}$$

è un sottospazio vettoriale di dimensione $n - r(M)$.
b) Sia W un sottospazio vettoriale di dimensione t; allora esiste un sistema lineare omogeneo $MX = 0$ di rango $n - t$ in n incognite tale che

$$W = \{(x_1, .., x_n)_{\mathcal{B}} \in V \mid MX = 0\}.$$

Si dice che $MX = 0$ è *un sistema di equazioni cartesiane di W rispetto alla base $\mathcal{B}$.*
Dunque, un sottospazio vettoriale di dimensione t può venire rappresentato da $n - t$ equazioni indipendenti, cioé come intersezione di $n - t$ iperpiani.

Si osservi che le equazioni cartesiane non sono univocamente determinate: tutti e soli i sistemi lineari equivalenti a $MX = 0$ danno equazioni cartesiane per W rispetto a $\mathcal{B}$.

(equazioni parametriche di un sottospazio vettoriale)

a) Date espressioni della forma:

$$\begin{cases} x_1 = b_{11}\lambda_1 + ... + b_{1s}\lambda_s \\ ...... \\ x_n = b_{n1}\lambda_1 + ... + b_{ns}\lambda_s \qquad\qquad \lambda_1, ..., \lambda_s \in K, \end{cases}$$

l'insieme $W := \{(x_1, .., x_n)_{\mathcal{B}} \in V \mid \exists \lambda_1, ..., \lambda_s \in K, x_i = b_{i1}\lambda_1 + ... + b_{is}\lambda_s, \quad i = 1, .., n\}$ è il sottospazio vettoriale di dimensione $r(b_{ij})$:

$$W =< (b_{11}, .., b_{n1})_{\mathcal{B}}, .., (b_{1s}, .., b_{ns})_{\mathcal{B}} > .$$

b) Sia $W =< w_1, ..., w_s >$ un sottospazio vettoriale di V, con $w_1 = (b_{11}, .., b_{n1})_{\mathcal{B}}, .., w_s = (b_{1s}, .., b_{ns})_{\mathcal{B}}$. Allora W si può descrivere come l'insieme dei vettori $(x_1, .., x_n)_{\mathcal{B}}$ tali che esistono $\lambda_1, ..., \lambda_s \in K$ tali che:

$$\begin{cases} x_1 = b_{11}\lambda_1 + ... + b_{1s}\lambda_s \\ ...... \\ x_n = b_{n1}\lambda_1 + ... + b_{ns}\lambda_s, \qquad\qquad \lambda_1, ..., \lambda_s \in K, \end{cases} \qquad (*)$$

Si osservi che le $(*)$ si possono anche scrivere in forma matriciale così:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda_1 \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix} + \ldots + \lambda_s \begin{pmatrix} b_{1s} \\ \vdots \\ b_{ns} \end{pmatrix}$$

Si dice che le $(*)$ sono *un sistema di equazioni parametriche di* W *rispetto ad* $\mathcal{B}$, e $\lambda_1, ..., \lambda_s$ sono detti parametri; si ha $dimW = r(b_{ij})$.
Dunque un sottospazio vettoriale di dimensione t può venire rappresentato con esattamente t parametri; basta scegliere i generatori w_i in modo che siano linearmente indipendenti. Si osservi che le equazioni parametriche non sono univocamente determinate: cambiando sistema di generatori per W si hanno altre equazioni parametriche.

Osservazione a) Sia W un sottospazio vettoriale di dimensione t e di equazioni cartesiane $MX = 0$ rispetto a $\mathcal{B}$ (sarà quindi $r(M) = n - t$); per passare ad equazioni parametriche sempre rispetto a $\mathcal{B}$ basta scrivere la soluzione generale del sistema che dipenderà da t incognite libere, cioè da t parametri.

b) Sia W un sottospazio vettoriale di dimensione t e di equazioni parametriche $(*)$ rispetto a $\mathcal{B}$; per passare ad equazioni cartesiane sempre rispetto a $\mathcal{B}$ basta imporre

$$r \begin{pmatrix} x_1 & b_{11} & \ldots & b_{1s} \\ ... & & & \\ x_n & b_{n1} & \ldots & b_{ns} \end{pmatrix} = t$$

scegliemdo un minore non nullo di ordine t della matrice (b_{ij}), ed imponendo che i suoi orlati di ordine $t + 1$ siano tutti nulli..

2 Autovalori ed autovettori

Iniziamo richiamando qualcosa sulle somme dirette di sottospazi:

Proposizione 2.1 *Sia V un K-spazio vettoriale, siano $E_1, \dots, E_s$ sottospazi di V di dimensione finita, e sia $n_j := dim E_j > 0$, $j = 1, \dots, s$. Sono equivalenti:*

a) *$E_1 + \dots + E_s$ è somma diretta;*

b) *Se $\mathcal{B}_j = (v_{j,1}, \dots, v_{j,n_j})$ è base per E_j, $j = 1, \dots, s$, allora*

$$\mathcal{B} = (v_{1,1}, \dots, v_{1,n_1}, \dots, v_{s,1}, \dots, v_{s,n_s})$$

è base per $E_1 + \dots + E_s$ (in altre parole, gli insiemi $\mathcal{B}_1, \dots, \mathcal{B}_s$ sono a due a due disgiunti, e la loro unione dotata di un ordine è una base per $E_1 + \dots + E_s$);

c) *$dim(E_1 + \dots + E_s) = dim E_1 + \dots + dim E_s$.*

Avremo anche bisogno di qualche richiamo sulle radici di un polinomio.

2.1 Radici di un polinomio

Lavoriamo con polinomi a coefficienti in un campo K; $K[x]$ denota l'anello di tali polinomi.

Per maggiori dettagli e per le dimostrazioni si rimanda per esempio alle Note di Algebra 2: http://www.dm.unibo.it/ ida/tutto11-12-14.pdf

Definizione 2.2 Sia $f = a_0 + a_1 x + \dots + a_n x^n$ un polinomio a coefficienti in K; diciamo che un elemento u di K è una radice, o uno zero, di f se $f(u) = a_0 + a_1 u + \dots + a_n u^n = 0$ (se cioé la funzione polinomiale associata ad f si annulla in u).

Definizione 2.3 Siano $f, g \in K[x]$; diciamo che g divide f e scriviamo $g \mid f$ se esiste un polinomio h tale che $f = gh$.

Proposizione 2.4 *Sia K campo, $u \in K$, $f \in K[x]$; allora*

$$x - u \mid f \iff f(u) = 0$$

Definizione 2.5 Sia K campo, $u \in K$, $f \in K[x]$, $f \neq 0$; *la molteplicità $\mu(u)$ di u come radice di* f è l'unico intero $m \in \mathbb{N}$ tale che

$$f = (x - u)^m g, \quad g(u) \neq 0$$

Si dimostra che la definizione è ben posta e che si ha

$$\mu(u) = max\{k \in \mathbb{N}, \quad (x - u)^k \mid f\}$$

Quindi se $\mu(u) = 0$ allora $f(u) \neq 0$.

Una *radice semplice, risp. multipla, di f* è una radice di molteplicità 1, risp. > 1.

Proposizione 2.6 *Sia K campo, $f \in K[x]$, f di grado n (cioé $f = a_0+a_1x+\ldots+a_nx^n$ con $a_n \neq 0$); allora se $u_1, \ldots, u_k$ sono le radici (a due a due distinte) di f in K, si ha $\mu(u_1) + \ldots + \mu(u_k) \leq n$. Se $\mu(u_1) + \ldots + \mu(u_k) = n$, diciamo che f ha n radici contate con molteplicità in K, e in tal caso f fattorizza nel prodotto di n fattori lineari:*

$$f = a_n(x - u_1)^{\mu(u_1)} \cdot \ldots \cdot (x - u_k)^{\mu(u_k)}$$

Esempio 2.7 *a*) Sia $f = x^2 - 6x - 5$; se pensiamo ad f come polinomio reale, le radici in $\mathbb{R}$ sono $3 \pm \sqrt{9+5}$ cioé $u_1 = 3 + \sqrt{14}, u_2 = 3 - \sqrt{14}$. Quindi $x - u_1 \mid f$ e $x - u_2 \mid f$; è facile verificare che

$$x^2 - 6x - 5 = (x - (3 + \sqrt{14}))(x - (3 - \sqrt{14}))$$

D'altra parte f ha coefficienti interi, quindi in $\mathbb{Q}$, ma le sue radici non stanno in $\mathbb{Q}$, quindi come polinomio razionale f non si fattorizza in fattori lineari.

b) Sia $f = 2x^2 - 6x - 3$; se pensiamo ad f come polinomio reale, le radici in $\mathbb{R}$ sono $\frac{3\pm\sqrt{9+6}}{2}$ cioé $u_1 = \frac{3+\sqrt{15}}{2}, u_2 = \frac{3-\sqrt{15}}{2}$. Quindi $x - u_1 \mid f$ e $x - u_2 \mid f$; è facile verificare che

$$2x^2 - 6x - 3 = 2(x - \frac{3+\sqrt{15}}{2})(x - \frac{3-\sqrt{15}}{2}).$$

La differenza con il caso precedente è che $x^2 - 6x - 5$ è un polinomio *monico*, cioé con coefficiente direttore $= 1$, mentre $2x^2 - 6x - 3$ ha coefficiente direttore $= 2$.

2.2 Matrici simili - Endomorfimi diagonalizzabili

In questa sessione V denota sempre un K-spazio vettoriale, se di dimensione finita o no verrà precisato di volta in volta, e id denota id_V, mentre I denota la matrice identità dell'ordine opportuno.

Proposizione 2.8 *Sia $dimV = n$, $\mathcal{A}$, $\mathcal{B}$ basi per V, $f \in EndV$. Si ha:*

$$M_{\mathcal{A}}(f) = M_{\mathcal{B},\mathcal{A}}(id)^{-1}\, M_{\mathcal{B}}(f)\, M_{\mathcal{B},\mathcal{A}}(id)$$

Dimostrazione

$$\begin{array}{ccccccc} & id & & f & & id & \\ V & \rightarrow & V & \rightarrow & V & \rightarrow & V \\ \mathcal{A} & & \mathcal{B} & & \mathcal{B} & & \mathcal{A} \end{array}$$

Dato che $id \circ f \circ id = f$, per 1.45 si ha:

$$M_{\mathcal{A}}(f) = M_{\mathcal{A},\mathcal{A}}(id \circ f \circ id) = M_{\mathcal{A},\mathcal{B}}(id)\, M_{\mathcal{B},\mathcal{B}}(f)\, M_{\mathcal{B},\mathcal{A}}(id) = M_{\mathcal{B},\mathcal{A}}(id)^{-1}\, M_{\mathcal{B}}(f)\, M_{\mathcal{B},\mathcal{A}}(id)$$

■

Corollario 2.9 *Nelle notazioni di 2.8, si ha*

$$detM_{\mathcal{A}}(f) = detM_{\mathcal{B}}(f)$$

Dimostrazione Per il Teorema di Binet si ha:

$$detM_{\mathcal{A}}(f) = det\left(M_{\mathcal{B},\mathcal{A}}(id)^{-1}\right)\, detM_{\mathcal{B}}(f)\, detM_{\mathcal{B},\mathcal{A}}(id) = \left(det\left(M_{\mathcal{B},\mathcal{A}}(id)\right)\right)^{-1}\, detM_{\mathcal{B}}(f)\, detM_{\mathcal{B},\mathcal{A}}(id)$$

Definizione 2.10 Sia V f.g., $\mathcal{B}$ una sua base e sia $f \in EndV$. Poniamo

$$detf := detM_{\mathcal{B}}(f).$$

Si osservi che la definizione è ben posta grazie a 2.9

Definizione 2.11 Due matrici $A, B \in M_n(K)$ si dicono *simili* se esiste $P \in GL_n(K)$ tale che

$$B = P^{-1}AP$$

Osservazione 2.12 La similitudine è una relazione di equivalenza in $M_n(K)$. Infatti:
- $A = I_n^{-1}AI_n$
- $B = P^{-1}AP \Rightarrow PBP^{-1} = A \Rightarrow A = (P^{-1})^{-1}BP^{-1}$
- $B = P^{-1}AP,\ C = Q^{-1}BQ \Rightarrow C = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$ dove P, Q sono invertibili, quindi anche P^{-1} e PQ lo sono.

Proposizione 2.13 *Sia $dimV = n$, e siano $A, B \in M_n(K)$. Allora A e B sono simili $\Longleftrightarrow$ rappresentano uno stesso endomorfismo di V rispetto a due basi diverse. Più precisamente,*

$$A, B \text{ simili } \Longleftrightarrow \exists f \in EndV,\ \mathcal{C}, \mathcal{D} \text{ basi di } V \text{ tali che } A = M_{\mathcal{C}}(f), B = M_{\mathcal{D}}(f)$$

Dimostrazione $\Leftarrow$) : segue da 2.8

$\Rightarrow$) : Sia $P = (p_{ij}) \in GL_n(K)$ tale che $B = P^{-1}AP$. Sia $\mathcal{C} = (v_1, \ldots, v_n)$ una base arbitraria e sia $f = F^{\mathcal{C}}(A)$. Sia $\mathcal{D} = (w_1, \ldots, w_n)$ la base di V tale che le coordinate di w_i rispetto a $\mathcal{C}$ siano la colonna i-esima di P: $w_i = (p_{1i}, \ldots, p_{ni})_{\mathcal{C}}$. Allora $P = M_{\mathcal{C},\mathcal{D}}(id)$, e si ha:

$$B = (M_{\mathcal{C},\mathcal{D}}(id))^{-1}M_{\mathcal{C}}(f)M_{\mathcal{C},\mathcal{D}}(id) \underbrace{=}_{Prop.2.8} M_{\mathcal{D}}(f)$$

■

Definizione 2.14 Una matrice $A \in M_n(K)$ si dice *diagonalizzabile* (abbreviato in dz) se è simile ad una matrice diagonale.

Diagonalizzare una matrice $A \in M_n(K)$ che sia dz vuol dire trovare una $P \in GL_n(K)$ tale che $P^{-1}AP$ sia diagonale.

Sia $dimV = n$; un endomorfismo $f \in EndV$ si dice *diagonalizzabile* se esiste una base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}(f)$ sia una matrice diagonale; in tal caso, $\mathcal{B}$ si dice base diagonalizzante.

Per Prop. 4.35, f è dz $\Longleftrightarrow$ se $\mathcal{C}$ è una qualsiasi base di V, $M_{\mathcal{C}}(f)$ è dz.

Osservazione 2.15 Se f è dz e $\mathcal{B} = (v_1, \ldots, v_n)$ è una base diagonalizzante, si ha

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

e quindi

$$f(v_1) = \lambda_1 v_1, \ldots, f(v_n) = \lambda_n v_n.$$

Viceversa, se f è tale che esista una base $\mathcal{B} = (v_1, \ldots, v_n)$ con $f(v_1) = \lambda_1 v_1, \ldots, f(v_n) = \lambda_n v_n$, allora f è dz e $\mathcal{B}$ è base diagonalizzante.

Esempio 2.16 L'applicazione lineare:

$$\begin{array}{cccc} f: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & e_1 & \mapsto & (2,0) \\ & e_2 & \mapsto & (0,\frac{1}{3}) \end{array}$$

è dz, $\mathcal{E}$ è una base diagonalizzante, e si ha

$$M_{\mathcal{E}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}.$$

Esempio 2.17 L'applicazione lineare:

$$\begin{array}{cccc} f: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (1,1) & \mapsto & (2,2) \\ & (1,-1) & \mapsto & (3,-3) \end{array}$$

è dz: infatti, se $v_1 = (1,1), v_2 = (1,-1)$, e $\mathcal{B}$ è la base $\mathcal{B} = (v_1, v_2)$, si ha:

$$M_{\mathcal{B}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Quindi anche $M_{\mathcal{E}}(f)$ deve essere dz e simile a $M_{\mathcal{B}}(f)$; verifichiamolo. Si ha:

$$\begin{array}{c} (1,0) = \frac{1}{2}(1,1) + \frac{1}{2}(1,-1) \\ (0,1) = \frac{1}{2}(1,1) - \frac{1}{2}(1,-1) \end{array}$$

da cui

$$f(e_1) = f(\tfrac{1}{2}(1,1) + \tfrac{1}{2}(1,-1)) = \tfrac{1}{2}f((1,1)) + \tfrac{1}{2}f((1,-1)) = \tfrac{1}{2}(2,2) + \tfrac{1}{2}(3,-3) = (\tfrac{5}{2}, -\tfrac{1}{2})$$

$$f(e_2) = f(\tfrac{1}{2}(1,1) - \tfrac{1}{2}(1,-1)) = \tfrac{1}{2}f((1,1)) - \tfrac{1}{2}f((1,-1)) = \tfrac{1}{2}(2,2) - \tfrac{1}{2}(3,-3) = (-\tfrac{1}{2}, \tfrac{5}{2})$$

cioé

$$M_{\mathcal{E}}(f) = \begin{pmatrix} \frac{5}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{2} \end{pmatrix}.$$

Per Prop. 2.8 deve essere

$$M_{\mathcal{B}}(f) = M_{\mathcal{B},\mathcal{E}}(id)\, M_{\mathcal{E}}(f)\, M_{\mathcal{E},\mathcal{B}}(id) \tag{$\diamond$}$$

dove

$$M_{\mathcal{E},\mathcal{B}}(id) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad M_{\mathcal{B},\mathcal{E}}(id) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix};$$

si verifica subito che, come già sappiamo da 1.51,

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

e che vale ($\diamond$):

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

$$= \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 4 & 6 \\ 4 & -6 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 8 & 0 \\ 0 & 12 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

2.3 Condizioni di diagonalizzabilità

Sotto quali condizioni una matrice, o, equivalentemente, un endomorfismo è dz?
Le definizioni e proposizioni che seguono servono a rispondere a questa domanda.

Definizione 2.18 Siano V un K-spazio vettoriale ed $f \in EndV$.

Un vettore $v \in V$ si dice *autovettore di* f se $v \neq 0$ e se $\exists \lambda \in K$ tale che

$$f(v) = \lambda v.$$

Uno scalare $\lambda \in K$ si dice *autovalore di* f se $\exists v \in V, v \neq 0$ tale che

$$f(v) = \lambda v.$$

Se $\lambda \in K$ e $v \in V, v \neq 0$ sono tali che $f(v) = \lambda v$, si dice che λ *è autovalore di* f *relativo all'autovettore* v, o anche che v *è autovettore di* f *relativo all'autovalore* λ.

Chiamiamo spettro di f l'insieme degli autovalori di f:

$$\Lambda_f := \{\lambda \in K \mid \lambda \text{ autovalore di} f\}.$$

Osservazione 2.19 L'autovalore relativo ad un autovettore è unico: infatti, sia $v \neq 0$ e $f(v) = \lambda v$, $f(v) = \mu v$; f è una applicazione, quindi $\lambda = \mu$.

Proposizione 2.20 *Sia V un K-spazio vettoriale di dimensione finita n, e siano $f \in EndV, \lambda \in K$. Allora:*

$$\lambda \text{ è autovalore per } f \iff Ker(f - \lambda id) \neq <0> \iff det(f - \lambda\, id) = 0$$

Dimostrazione Consideriamo l'applicazione lineare:

$$\begin{array}{rccc} f - \lambda\, id: & V & \to & V \\ & v & \mapsto & f(v) - \lambda v \end{array}$$

La prima equivalenza è quindi semplicemente la definizione di autovalore. La seconda si vede così: sia $\mathcal{B}$ una base per V; essendo $(f - \lambda id) \in EndV$, da 1.46

$$Ker(f - \lambda\, id) \neq <0> \iff 0 < dimKer(f - \lambda\, id) = n - rg(M_{\mathcal{B}}(f - \lambda\, id)) \iff$$

$$\iff rg(M_{\mathcal{B}}(f - \lambda\, id)) < n \iff detM_{\mathcal{B}}(f - \lambda\, id) = 0 \iff det(f - \lambda\, id) = 0$$

Definizione 2.21 Siano V un K-spazio vettoriale, $f \in EndV$, e $\lambda \in K$ autovalore per f. Allora

$$V_\lambda := \{v \in V \mid v \text{ autovettore di } f \text{ relativo a } \lambda\} \cup \{0\}$$

è un sottospazio di V detto *l'autospazio di* f *relativo all'autovalore* λ. Si ha:

$$V_\lambda = \{v \in V \mid f(v) = \lambda v\} = \{v \in V \mid (f - \lambda\, id)(v) = 0\} = Ker(f - \lambda\, id)$$

Esempio 2.22 In Esempio 2.16:

$$f: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ e_1 & \mapsto & (2,0)\,, \\ e_2 & \mapsto & (0,\frac{1}{3}) \end{array} \qquad M_{\mathcal{E}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}.$$

2 e $\frac{1}{3}$ sono autovalori, e_1 è un autovettore relativo all'autovalore 2, ed e_2 è un autovettore relativo all'autovalore $\frac{1}{3}$.

Cerchiamo V_2:

$$V_2 = \{v \in \mathbb{R}^2 \mid v \text{ autovettore di } f \text{ relativo a } 2\} \cup \{0\} = \{v \in \mathbb{R}^2 \mid v \neq 0, f(v) = 2v\} \cup \{0\} =$$

$$= \{v \in \mathbb{R}^2 \mid f(v) = 2v\} = \{(x,y) \in \mathbb{R}^2 \mid \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}\}$$

Risolviamo quindi il sistema $\begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$, cioé

$$\begin{cases} 2x = 2x \\ \frac{1}{3}y = 2y \end{cases}$$

Quindi

$$V_2 = \{(x,y) \in \mathbb{R}^2 \mid y = 0\}.$$

In Esempio 2.17 :

$$f: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (1,1) & \mapsto & (2,2)\,, \\ (1,-1) & \mapsto & (3,-3) \end{array} \qquad M_{\mathcal{B}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

dove $v_1 = (1,1), v_2 = (1,-1)$, e $\mathcal{B}$ è la base $\mathcal{B} = (v_1, v_2)$, si ha analogamente che 2 e 3 sono autovalori, v_1 è un autovettore relativo all'autovalore 2, e v_2 è un autovettore relativo all'autovalore 3.

Proposizione 2.23 *Sia $f \in EndV$. Se $v_1, \dots, v_k \in V$ sono autovettori relativi rispettivamente agli autovalori $\lambda_1, \dots, \lambda_k$, e se i λ_i sono a due a due distinti, allora $v_1, \dots, v_k$ sono l.i.*

Dimostrazione Per induzione: $v_1 \neq 0$ per definizione; siano $v_1, \dots, v_i$ l.i. e proviamo che lo sono $v_1, \dots, v_{i+1}$.

Se $v_1, \dots, v_{i+1}$ sono l.d., allora esistono $a_1, \dots, a_i \in K$ tali che

$$v_{i+1} = a_1 v_1 + \ldots + a_i v_i \qquad (*)$$

per cui

$$f(v_{i+1}) \underbrace{=}_{\text{per } (*)} f(a_1 v_1 + \ldots + a_i v_i) = a_1 f(v_1) + \ldots + a_i f(v_i) = a_1 \lambda_1 v_1 + \ldots + a_i \lambda_i v_i$$

e anche

$$f(v_{i+1}) = \lambda_{i+1} v_{i+1} \underbrace{=}_{\text{per } (*)} \lambda_{i+1}(a_1 v_1 + \ldots + a_i v_i)$$

da cui si ha

$$a_1\lambda_1 v_1 + \ldots + a_i\lambda_i v_i = \lambda_{i+1}(a_1 v_1 + \ldots + a_i v_i) \Rightarrow a_1(\lambda_1 - \lambda_{i+1})v_1 + \ldots + a_i(\lambda_i - \lambda_{i+1})v_i = 0$$

ed essendo $v_1, \ldots, v_i$ l.i., si ha

$$\begin{cases} a_1(\lambda_1 - \lambda_{i+1}) = 0 \\ \ldots \\ a_i(\lambda_i - \lambda_{i+1}) = 0 \end{cases}$$

e poiché $\lambda_1 \neq \lambda_{i+1}, \ldots, \lambda_i \neq \lambda_{i+1}$, si ha $a_1 = \ldots = a_i = 0$, da cui $v_{i+1} = 0$; siamo arrivati ad una contraddizione. ■

Corollario 2.24 *Sia $f \in EndV$, e siano $\lambda_1, \ldots, \lambda_k$ autovalori a due a due distinti; allora la somma di sottospazi*

$$V_{\lambda_1} + \ldots + V_{\lambda_k}$$

è una somma diretta.

In particolare, se $dimV$ è finita, si ha anche che

$$dim(V_{\lambda_1} + \ldots + V_{\lambda_k}) = dimV_{\lambda_1} + \ldots + dimV_{\lambda_k},$$

e se $\mathcal{B}_i = (v_{i,1}, \ldots, v_{i,n_i})$ è una base per V_{λ_i}, $i = 1, \ldots, k$, allora

$$(v_{1,1}, \ldots, v_{1,n_1}, \ldots, v_{k,1}, \ldots, v_{k,n_k})$$

è una base per $V_{\lambda_1} + \ldots + V_{\lambda_k}$.

Dimostrazione Sappiamo che la somma $V_{\lambda_1} + \ldots + V_{\lambda_k}$ è diretta se 0 si scrive in modo unico come somma $v_1 + \ldots + v_k$ con $v_i \in V_{\lambda_i}$, quindi necessariamente come $0 = 0 + \ldots + 0$. Sia $0 = v_1 + \ldots + v_k$ con $v_i \in V_{\lambda_i}$; se qualcuno dei v_i fosse diverso da 0, diciamo $v_{i_1} \neq 0, \ldots, v_{i_t} \neq 0$, $v_{i_1}, \ldots, v_{i_t}$ sarebbero autovettori relativi ad autovalori a due a due distinti, una cui combinazione lineare a coefficienti $\neq 0$ sarebbe nulla, il che contraddice 2.23.

Definizione 2.25 Sia $A = (a_{ij}) \in M_n(K)$ e sia t una indeterminata. Si pone:

$$p_A(t) := det(A - tI_n) = det\begin{pmatrix} a_{11} - t & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - t & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - t \end{pmatrix}$$

Si ha che $p_A(t)$ è un polinomio di grado n in t, detto *il polinomio caratteristico di A*.

Definizione 2.26 Sia V K-spazio vettoriale di dimensione n, $f \in EndV$, e sia t una indeterminata. Si pone:

$$p_f(t) := det(f - t\, id_V)$$

Ricordiamo che $detf$ è il determinante di una matrice associata ad f rispetto ad una qualunque base (2.10).

Sia $\mathcal{B}$ una base per V, e $A = (a_{ij}) := M_{\mathcal{B}}(f)$; poiché se $a \in K$, allora $M_{\mathcal{B}}(a\, id_V) = aI_n$, ne segue $M_{\mathcal{B}}(f - a\, id_V) = A - aI_n$, cioé:

$$p_f(t) = p_A(t) = det \begin{pmatrix} a_{11}-t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-t & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-t \end{pmatrix};$$

$p_f(t)$ è un polinomio di grado n in t, detto *il polinomio caratteristico di* f; $(-1)^n p_f(t)$ è un polinomio monico di grado n.

Osservazione 2.27 Se $A, B \in M_n(K)$ sono matrici simili allora hanno lo stesso polinomio caratteristico: infatti A e B rappresentano, rispetto a due opportune basi, lo stesso $f \in End\,\mathbb{R}^2 \Rightarrow p_A(t) = p_f(t) = p_B(t)$.

Attenzione: NON vale il viceversa.

Proposizione 2.28 *Sia V K-spazio vettoriale di dimensione n, $f \in EndV$. Allora $\lambda \in K$ è autovalore di $f \iff p_f(\lambda) = 0$, cioé λ radice di $p_f(t)$.*

In particolare, f possiede al più n autovalori distinti.

Dimostrazione Per 2.20 $\lambda \in K$ è autovalore di $f \iff det(f - \lambda\, id_V) = 0 \iff p_f(\lambda) = 0$. Poiché f ha grado n, sappiamo dall'algebra che ha al più n radici. ■

Proposizione 2.29 *Sia V K-spazio vettoriale, $dimV = n$. Allora $f \in EndV$ è dz $\iff$ V possiede una base costituita da autovettori di f.*

Se f è dz, e $\mathcal{B} = (v_1, \dots, v_n)$ è una base di autovettori, allora $M_{\mathcal{B}}(f)$ è una matrice diagonale della forma

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & 0\dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori (non necessariamente distinti!) relativi rispettivamente a $v_1, \dots, v_n$.

Dimostrazione È Osservazione 2.15 riformulata nel linguaggio degli autovettori; la base diagonalizzante di 2.15 è formata da autovettori.

Esempio 2.30 Riprendiamo la f di Esempio 2.16 e 2.22:

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & e_1 & \mapsto & (2,0), \\ & e_2 & \mapsto & (0, \frac{1}{3}) \end{array} \qquad M_{\mathcal{E}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}$$

quindi $p_f(t) = det \begin{pmatrix} 2-t & 0 \\ 0 & \frac{1}{3}-t \end{pmatrix} = (2-t)(\frac{1}{3}-t)$; le radici di $p_f(t)$ sono, come devono essere, gli autovalori 2 e $\frac{1}{3}$.

Esempio 2.31 Sia $V = \mathbb{R}^3$, ed $f \in EndV$ con

$$M_{\mathcal{E}}(f) = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & 3 & -1 \end{pmatrix}$$

Cerchiamo gli autovalori:

$$p_f(t) = det(M_{\mathcal{E}}(f) - tI_3) = det \begin{pmatrix} 1-t & 1 & 0 \\ -1 & 1-t & 0 \\ -1 & 3 & -1-t \end{pmatrix} = (-1-t)((1-t)^2+1) = -(t+1)(t^2-2t+2)$$

poiché $\Delta(t^2 - 2t + 2) = 4 - 8 < 0$, l'unica radice reale di p_f è -1: $\quad \Lambda_f = \{-1\}$.

Se cambiamo campo, se cioé $V = \mathbb{C}^3$ ed $f \in EndV$ con la stessa matrice, allora f deve avere 3 radici contate con molteplicità in $\mathbb{C}$:

$$t^2 - 2t + 2 = 0 \Rightarrow t = 1 \pm i$$

per cui $\quad \Lambda_f = \{-1, 1-i, 1+i\}$.

Esempio 2.32 Sia $V = \mathbb{R}^3$, ed $f \in EndV$ con

$$M_{\mathcal{E}}(f) = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 1 & 0 \\ -1 & 3 & -1 \end{pmatrix}$$

Cerchiamo gli autovalori:

$$p_f(t) = det \begin{pmatrix} 1-t & -2 & 0 \\ -2 & 1-t & 0 \\ -1 & 3 & -1-t \end{pmatrix} = (-1-t)((1-t)^2-4) = -(t+1)(t^2-2t-3) = -(t+1)^2(t-3)$$

essendo -1 e 3 le radici di $t^2 - 2t - 3$; in questo caso $\Lambda_f = \{-1, 3\}$, e diciamo che l'autovalore -1 ha molteplicità algebrica 2 (Definizione 2.34).

Teorema 2.33 *Sia V K-spazio vettoriale, $dimV = n$, $f \in EndV$, $\Lambda_f = \{\lambda_1, \dots, \lambda_s\} \subset K$ dove i λ_i sono a due a due distinti. Sono equivalenti:*

a) f è dz

b) $V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_s}$

c) $n = dimV_{\lambda_1} + \dots + dimV_{\lambda_s}$.

Dimostrazione *a)* $\Rightarrow$ *b)*: per hp V ha una base costituita da autovettori per f: $(v_1, \dots, v_n)$; ogni v_i appartiene ad un qualche autospazio V_{λ_j}, quindi

$$V = < v_1, \dots, v_n > \subseteq V_{\lambda_1} + \dots + V_{\lambda_s}$$

e l'altra inclusione è ovvia. La somma degli autospazi è diretta per 2.24, quindi si ha b).

$b) \Rightarrow a)$: una base di V è unione (disgiunta) di una base di ciascun V_{λ_j}, quindi V ha una base costituita da autovettori per f.

$b) \iff c)$: sappiamo che $V_{\lambda_1} + \ldots + V_{\lambda_s}$ è somma diretta per 2.24, quindi $dim(V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_s}) = dimV_{\lambda_1} + \ldots + dimV_{\lambda_s}$. ∎

Definizione 2.34 Sia V K-spazio vettoriale, $dimV = n$, $f \in EndV$ e $\lambda \in K$ un autovalore per f; chiamiamo *molteplicità algebrica* $\mu(\lambda)$ *dell'autovalore* λ *per* f la molteplicità di λ come radice del polinomio caratteristico, cioé l'intero $\mu(\lambda)$, $1 \leq \mu(\lambda) \leq n$, tale che

$$p_f(t) = (t-\lambda)^{\mu(\lambda)} g(t), \quad g(\lambda) \neq 0.$$

Chiamiamo *molteplicità geometrica di* λ *per* f il numero naturale

$$dimV_\lambda$$

Proposizione 2.35 *Sia V K-spazio vettoriale, $dimV = n$, $f \in EndV$ e $\lambda \in K$ un autovalore per f. Allora:*

$$1 \leq dimV_\lambda \leq \mu(\lambda)$$

In particolare, se un autovalore è radice semplice di p_f, cioé se $\mu(\lambda) = 1$, allora $dimV_\lambda = \mu(\lambda) = 1$.

Dimostrazione La prima disuguaglianza è ovvia. Per dimostrare la seconda, poniamo $d := dimV_\lambda$; sia $(v_1, \ldots, v_d)$ una base per V_λ, e completiamo ad una base $\mathcal{B} = (v_1, \ldots, v_d, \ldots, v_n)$ per V. Allora, utilizzando la scrittura a blocchi (si veda Definizione 3.7),

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda I_d & D \\ 0 & C \end{pmatrix}$$

dove $D \in M_{d,n-d}(K)$ e $C \in M_{n-d}(K)$ sono opportune matrici; quindi

$$p_f(t) = det(M_{\mathcal{B}}(f) - tI_n) = det \begin{pmatrix} \lambda I_d - t I_d & D \\ 0 & C - t I_{n-d} \end{pmatrix} = (\lambda - t)^d det(C - t I_{n-d}) = (\lambda - t)^d p_C(t)$$

dove $p_C(t)$ è un polinomio di grado $n-d$, che può o meno avere λ come radice; quindi $d \leq \mu(\lambda)$.

Si osservi che il calcolo del determinante della matrice a blocchi viene fatto come se fosse una matrice 2×2; questo si può fare perché la matrice è una matrice "triangolare a blocchi"; vediamolo in questo particolare caso:

$$det \begin{pmatrix} \lambda I_d - t I_d & D \\ 0 & C - t I_{n-d} \end{pmatrix} = det \begin{pmatrix} \lambda - t & 0 & \ldots & 0 & d_{11} & \ldots & d_{1,n-d} \\ 0 & \lambda - t & \ldots & 0 & \vdots & \ldots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \vdots & \lambda - t & d_{d,1} & \ldots & c_{d,n-d} \\ 0 & 0 & \vdots & 0 & c_{11} - t & \ldots & c_{1,n-d} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 0 & c_{n-d,1} & \ldots & c_{n-d,n-d} - t \end{pmatrix}$$

e sviluppando sempre secondo la prima colonna e la prima riga vediamo che il determinante vale $(\lambda - t)(\lambda - t)\ldots(\lambda - t) det(C - t I_{n-d})$. ∎

Proposizione 2.36 *Sia V K-spazio vettoriale, $dimV = n$, $f \in EndV$ e $\Lambda_f = \{\lambda_1, \dots, \lambda_s\} \subset K$, dove i λ_i sono a due a due distinti. Allora f è dz se e solo se valgono le seguenti due condizioni:*

a) p_f ha n radici contate con molteplicità in K, cioé vale $\mu(\lambda_1) + \ldots + \mu(\lambda_s) = n$;

b) $dimV_{\lambda_j} = \mu(\lambda_j)\ \ \forall j = 1, \ldots, s$.

Dimostrazione Se valgono $a)$ e $b)$, allora $n = \mu(\lambda_1) + \ldots + \mu(\lambda_s) = dimV_{\lambda_1} + \ldots + dimV_{\lambda_s}$ e per 2.33 f è dz.

Viceversa se f è dz allora (2.33)

$$n = dimV_{\lambda_1} + \ldots + dimV_{\lambda_s} \leq \mu(\lambda_1) + \ldots + \mu(\lambda_s) \leq n$$

per cui questi sono tutti $=$ e quindi valgono $a)$ e $b)$.

Corollario 2.37 *Se $dimV = n$ ed $f \in EndV$ ha n autovalori distinti, allora f è dz.*

Dimostrazione Se $f \in EndV$ ha n autovalori distinti $\lambda_1, \ldots, \lambda_n$, allora $\mu(\lambda_j) = 1\ \ \forall j = 1, \ldots, n$, quindi per 2.35 $dimV_{\lambda_j} = \mu(\lambda_j) = 1\ \ \forall j = 1, \ldots, n$, e quindi valgono $a)$ e $b)$ di 2.36.

Attenzione! Non è vero il viceversa.

Corollario 2.38 *Se $K = \mathbb{C}$, allora f è dz $\iff$ $dimV_\lambda = \mu(\lambda)\ \ \forall \lambda \in \Lambda_f$.*

Dimostrazione Poiché $\mathbb{C}$ è un campo algebricamente chiuso, la condizione $a)$ in 2.36 è sempre verificata.

Osservazione 2.39 Se $K = \mathbb{C}$ e $dimV = n > 0$ allora ogni $f \in EndV$ ha almeno un autovalore.

Se $K = \mathbb{R}$ e $dimV = n$ è dispari allora ogni $f \in EndV$ ha almeno un autovalore.

Esempio 2.40 In Esempio 2.16 e 2.30:

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & e_1 & \mapsto & (2,0)\,, \\ & e_2 & \mapsto & (0, \frac{1}{3}) \end{array} \qquad M_{\mathcal{E}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}.$$

avevamo visto che $p_f(t) = (2-t)(\frac{1}{3} - t)$, e gli autovalori sono 2 e $\frac{1}{3}$.

Avevamo anche determinato l'autospazio $V_2 = \{(x,y) \in \mathbb{R}^2 \mid y = 0\} =< (1,0) >$ cercando le soluzioni del sistema $f(v) = 2v$, o, equivalentemente, $(f - 2id_V)(v) = 0$; infatti, come già sappiamo, $V_2 = Ker(f - 2id_V)$.
Vediamo ora chi è l'altro autospazio:

$$V_{\frac{1}{3}} = Ker(f - \frac{1}{3}id_V)$$

Ora si ha:

$$M_{\mathcal{E}}(f - \frac{1}{3}id_V) = \begin{pmatrix} 2 - \frac{1}{3} & 0 \\ 0 & \frac{1}{3} - \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{5}{3} & 0 \\ 0 & 0 \end{pmatrix}$$

Quindi $Ker(f - \frac{1}{3}id_V)$ è dato dalle soluzioni del sistema

$$\begin{pmatrix} \frac{5}{3} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

cioé

$$\begin{cases} \frac{5}{3}x = 0 \\ 0y = 0 \end{cases}$$

da cui

$$V_{\frac{1}{3}} = \{(x, y) \in \mathbb{R}^2 \mid x = 0\} =< (0, 1) >$$

Si ha $V = V_2 \oplus V_{\frac{1}{3}}$ e infatti f è dz.

Esempio 2.41 Riprendiamo Esempio 2.31: $V = \mathbb{R}^3$, $f \in EndV$ con

$$M_{\mathcal{E}}(f) = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & 3 & -1 \end{pmatrix}$$

Sappiamo che c'è un unico autovalore, -1, con $\mu(-1) = 1$, perché abbiamo visto che:

$$p_f(t) = -(t+1)(t^2 - 2t + 2), \quad \Delta(t^2 - 2t + 2) < 0$$

Quindi f non è dz. Cerchiamo l'autospazio V_{-1}:

$$V_{-1} = Ker(f - (-1)id_V) = Ker(f + id_V)$$

Ora si ha:

$$M_{\mathcal{E}}(f + id_V) = \begin{pmatrix} 1+1 & 1 & 0 \\ -1 & 1+1 & 0 \\ -1 & 3 & -1+1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 2 & 0 \\ -1 & 3 & 0 \end{pmatrix}$$

Quindi $Ker(f + id_V)$ è dato dalle soluzioni $(x, y, z) \in \mathbb{R}^3$ del sistema

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

cioé

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

da cui

$$V_{-1} = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0, y = 0\} =< (0, 0, 1) >= asse\ z$$

Abbiamo visto che se cambiamo campo, se cioé $V = \mathbb{C}^3$ ed $f \in EndV$ con la stessa matrice, allora:

$$p_f(t) = -(t+1)(t - (1-i))(t - (1+i))$$

quindi f ha 3 autovalori distinti in $\mathbb{C}$: $-1, 1-i, 1+i$, con $\mu(-1) = \mu(1-i) = \mu(1+i) = 1$.

In questo caso si avrà $\mu(-1) = dimV_{-1} = 1$, $\mu(1-i) = dimV_{1-i} = 1$, $\mu(1+i) = dimV_{1+i} = 1$ per 2.35, quindi per 2.38 f è dz come endomorfismo di $\mathbb{C}^3$.

Esempio 2.42 Riprendiamo Esempio 2.32: $V = \mathbb{R}^3$, $f \in EndV$ con

$$M_{\mathcal{E}}(f) = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 1 & 0 \\ -1 & 3 & -1 \end{pmatrix}$$

Abbiamo visto che $p_f(t) = -(t+1)^2(t-3)$ e quindi gli autovalori sono -1 e 3, con $\mu(-1) = 2$, $\mu(3) = 1$.

Cerchiamo l'autospazio $V_{-1} = Ker(f + id_V)$:

$$M_{\mathcal{E}}(f + id_V) = \begin{pmatrix} 1+1 & -2 & 0 \\ -2 & 1+1 & 0 \\ -1 & 3 & -1+1 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ -1 & 3 & 0 \end{pmatrix}$$

Quindi V_{-1} è dato dalle soluzioni $(x, y, z) \in \mathbb{R}^3$ del sistema

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

cioé

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

da cui

$$V_{-1} = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0, y = 0\} =< (0, 0, 1) >= asse\ z$$

quindi $dimV_{-1} < \mu(-1)$, per cui f non è dz.

Se non avessimo voluto sapere chi è V_{-1}, ma solo se f è dz o meno, bastava osservare che

$$rgM_{\mathcal{E}}(f + id_V) = rg \begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ -1 & 3 & 0 \end{pmatrix} = 2 \Rightarrow dimKer(f + id_V) = 1 < \mu(-1) = 2$$

Esercizio 2.43 $a)$ Sia

$$\begin{array}{rccc} f: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x, y, z) & \longmapsto & (y - z, -x + 2y - z, x - y + 2z) \end{array}$$

È f dz? Se sì, si determini una base diagonalizzante.

$b)$ Oppure la domanda potrebbe essere formulata così:

Sia $A = M_{\mathcal{E}}(f)$. È A dz? Se sì, si diagonalizzi A.

$a)$ Semplifichiamo le notazioni: $I = I_3$, $id = id_{\mathbb{R}^3}$. Si ha $M_{\mathcal{E}}(f) = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$.

Cominciamo con il cercare gli autovalori:

$$p_f(t) = det(M_{\mathcal{E}}(f) - tI) = det \begin{pmatrix} 0-t & 1 & -1 \\ -1 & 2-t & -1 \\ 1 & -1 & 2-t \end{pmatrix} \underbrace{=}_{\text{sviluppo secondo riga 3}}$$

$$= (-1 + 2 - t) + (t - 1) + (2 - t)(-t(2 - t) + 1) =$$

$$= 1 - t + t - 1 + (2 - t)(t^2 - 2t + 1) = (2 - t)(t - 1)^2$$

quindi gli autovalori sono $\lambda_1 = 1$ con $\mu(\lambda_1) = 2$ e $\lambda_2 = 2$ con $\mu(\lambda_2) = 1$.

Vediamo se f è dz:

la condizione $a)$ di 2.36 vale perché p_f ha 3 radici contate con molteplicità in $\mathbb{R}$, cioé $\mu(\lambda_1)+\mu(\lambda_2) = 3$;

la condizione $b)$ di 2.36 vale per λ_2 perché $\mu(\lambda_2) = 1 \Rightarrow dimV_{\lambda_2} = 1$; vediamo se vale anche per λ_1:

$$rg(M_{\mathcal{E}}(f) - \lambda_1 I) = rg \begin{pmatrix} 0-1 & 1 & -1 \\ -1 & 2-1 & -1 \\ 1 & -1 & 2-1 \end{pmatrix} = rg \begin{pmatrix} -1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} = 1$$

quindi $dimKer(f - \lambda_1 id_V) = 2 \Rightarrow dimV_{\lambda_1} = 2 = \mu(\lambda_1)$ ed f è dz.

Cerchiamo una base diagonalizzante, cioé di autovettori:

$V_{\lambda_1} = Ker(f - \lambda_1 id)$, quindi è lo spazio delle soluzioni del sistema

$$\begin{pmatrix} -1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Abbiamo già visto che il rango della matrice del sistema è 1, quindi il sistema è equivalente a $-x + y - z = 0$ e il suo spazio delle soluzioni è

$$V_{\lambda_1} = \{(x, y, z) \in \mathbb{R}^3 \mid x = y - z\} = \{(y - z, y, z) \mid y, z \in \mathbb{R}\} =< \underbrace{(1, 1, 0), (-1, 0, 1)}_{\text{base per } V_{\lambda_1}} >$$

$V_{\lambda_2} = Ker(f - \lambda_2 id)$, quindi è lo spazio delle soluzioni del sistema

$$(M_{\mathcal{E}}(f) - \lambda_2 I) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cioé

$$\begin{pmatrix} -2 & 1 & -1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

il rango della matrice del sistema è necessariamente 2, dato che $\mu(\lambda_2) = 1 \Rightarrow dimV_{\lambda_2} = 1$, e il sistema è equivalente a

$$\begin{cases} -x - z = 0 \\ x - y = 0 \end{cases}$$

e il suo spazio delle soluzioni è

$$V_{\lambda_2} = \{(x, y, z) \in \mathbb{R}^3 \mid x = y = -z\} = \{(x, x, -x) \mid x \in \mathbb{R}\} =< \underbrace{(1, 1, -1)}_{\text{base per } V_{\lambda_2}} >$$

Sappiamo che $V = V_{\lambda_1} \oplus V_{\lambda_2}$, e in effetti l'unione della base trovata per V_{λ_1} e della base trovata per V_{λ_2} è una base per V:

$$det \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \neq 0$$

e una base diagonalizzante è quindi $\mathcal{B} = ((1, 1, 0), (-1, 0, 1), (1, 1, -1)$; si ha

$$M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Osserviamo che se prendiamo un ordine diverso sui vettori della base $\mathcal{B}$, considerando per esempio la base $\mathcal{C} = ((1, 1, -1), (1, 1, 0), (-1, 0, 1)$, troviamo

$$M_{\mathcal{C}}(f) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

b) So che f è dz $\iff$ A lo è. Inoltre abbiamo già trovato una base diagonalizzante, $\mathcal{B}$. Quindi da 2.8 sappiamo che

$$M_{\mathcal{B}}(f) = M_{\mathcal{E},\mathcal{B}}(id)^{-1} \, M_{\mathcal{E}}(f) \, M_{\mathcal{E},\mathcal{B}}(id)$$

dove

$$M_{\mathcal{E},\mathcal{B}}(id) = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

Quindi abbiamo diagonalizzato A, cioé abbiamo trovato una matrice invertibile $P := M_{\mathcal{E},\mathcal{B}}(id)$ tale che

$$P^{-1} \, A \, P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Osservazione 2.44 Facciamo il punto della situazione:

partendo da un endomorfismo:

Dato V K-spazio vettoriale con $dimV = n$, $\mathcal{B}$ base di V, e $f \in EndV$, poniamo $A := M_{\mathcal{B}}(f)$; $I = I_n$ e $id = id_V$;

a) ricerca autovalori:

- calcoliamo $p_f(t)$ come $det(A - tI)$

- calcoliamo le radici in K di p_f (quando ne siamo capaci!)

b) ricerca autospazi:

per ogni autovalore λ trovato in **a)**, risolviamo il sistema lineare omogeneo $(A - \lambda I)X = 0$ e lo spazio delle soluzioni è V_λ, in coordinate rispetto a $\mathcal{B}$;

c) f è dz?

Per rispondere non è necessario passare da **b)**;

- basta eseguire **a)**, cioé calcolare gli autovalori $\lambda_1, \ldots, \lambda_s$ con le relative molteplicità algebriche $\mu(\lambda_1), \ldots, \mu(\lambda_s)$;

- se $\mu(\lambda_1) + \ldots + \mu(\lambda_s) = n$, e se, per ogni i tale che $\mu(\lambda_i) > 1$, si ha

$$\underbrace{n - rg(A - \lambda_i I)}_{=dim V_{\lambda_i}} = \mu(\lambda_i),$$

allora f è dz.

d) ricerca di una base $\mathcal{D}$ di autovettori quando f è dz:

- eseguo **b)** e costruisco una base per ogni V_{λ_i} poi unisco le basi:

$$\mathcal{D} = (\underbrace{v_{11}, \ldots, v_{1,\mu(\lambda_1)}}_{base\ per\ V_{\lambda_1}}, \underbrace{v_{21}, \ldots, v_{2,\mu(\lambda_2)}}_{base\ per\ V_{\lambda_2}}, \ldots, \underbrace{v_{s1}, \ldots, v_{s,\mu(\lambda_s)}}_{base\ per\ V_{\lambda_s}})$$

e si ha:

$$M_{\mathcal{D}}(f) = M_{\mathcal{B},\mathcal{D}}(id)^{-1}\, M_{\mathcal{B}}(f)\, M_{\mathcal{B},\mathcal{D}}(id)$$

dove $M_{\mathcal{B},\mathcal{D}}(id)$ ha sulle colonne le componenti dei vettori di $\mathcal{D}$ rispetto alla base $\mathcal{B}$, e

$$M_{\mathcal{D}}(f) = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_s & & \\ & & & & & \ddots & \\ & & & & & & \lambda_s \end{pmatrix}$$

dove λ_j compare sulla diagonale $\mu(\lambda_j)$ volte, per $j = 1, \ldots, s$.

partendo da una matrice:

Sia data una matrice $A \in M_n(K)$

a) A è dz?

penso ad A come matrice di un endomorfismo f di uno spazio vettoriale V di dim n rispetto ad una base $\mathcal{B}$, e procedo come sopra. Di solito, si prende $V = K^n$ e $\mathcal{B} = \mathcal{E}$, quindi $f = F^{\mathcal{E}}(A)$.

b) diagonalizzare A quando A è dz:

come sopra, penso ad A come matrice di un endomorfismo f di uno spazio vettoriale V di dim n rispetto ad una base $\mathcal{B}$, per esempio $V = K^n$ e $\mathcal{B} = \mathcal{E}$, e costruisco una base di autovettori per f, arrivando quindi a trovare una matrice $P \in GL_n(K)$ tale che $P^{-1}\, A\, P$ sia diagonale.

3 La forma di Jordan

3.1 Spazi quoziente

Definizione 3.1 Sia V un K-spazio vettoriale f.g. e W un suo sottospazio. Allora $(W,+)$ in particolare è un sottogruppo del gruppo abeliano $(V,+)$, quindi come avete visto in Algebra 1 possiamo fare il gruppo quoziente $(V/W,+)$; ricordiamo che la relazione di equivalenza è $v \sim w \iff v-w \in W$, quindi, denotando la classe di equivalenza di v con $\bar{v}$, si ha $\bar{v} = \{v+w,\ w \in W\} = v+W$, e

$$\bar{v} + \bar{w} := \overline{v+w}$$

Anche il prodotto per scalari risulta compatibile con la relazione, infatti se $a \in K$ e $v, w \in V$, con $v \sim w$, allora $av - aw = a(v-w) \in W$, quindi possiamo definire un prodotto per scalari sul quoziente nel modo seguente:

$$\forall a \in K, \forall v \in V, \quad a\bar{v} = \overline{av}$$

ed è facile verificare (esercizio) che con questo prodotto per scalari V/W è un K-spazio vettoriale; si osservi che se V è f.g. anche V/W lo è.

La proiezione canonica sul quoziente verrà denotata con

$$\begin{array}{rccc} \pi: & V & \to & V/W \\ & v & \mapsto & \bar{v} \end{array}$$

La π è una applicazione lineare, e $Ker\pi = W$.

Definizione 3.2 a) Diremo che i vettori $v_1, \ldots, v_k \in V$ sono l.i. modulo W se $v_1, \ldots, v_k$ sono l.i. e la somma $< v_1, \ldots, v_k > +W$ è diretta;

b) Diremo che i vettori $v_1, \ldots, v_k \in V$ sono una base di V modulo W se $v_1, \ldots, v_k$ sono l.i. e $V =< v_1, \ldots, v_k > \oplus W$.

Osservazione 3.3 Sia $W \neq < 0 >$. I vettori $v_1, \ldots, v_k \in V$ sono l.i. modulo W, risp. una base modulo W, se e solo se presa una qualunque base $w_1, \ldots, w_s$ di W i vettori $v_1, \ldots, v_k, w_1, \ldots, w_s$ sono l.i., risp. una base di V.
Basta infatti ricordare che la somma di due sottospazi è diretta $\iff$ l'unione disgiunta delle basi è una base per la somma dei due sottospazi.
I vettori $v_1, \ldots, v_k \in V$ sono l.i. modulo $< 0 >$, risp. una base modulo $< 0 >$, se e solo se sono l.i., risp. una base per V.

Proposizione 3.4 (a) I vettori $v_1, \ldots, v_k \in V$ sono l.i. modulo $W \iff \bar{v}_1, \ldots, \bar{v}_k$ sono l.i. in $V/W \iff (\forall a_1, \ldots, a_k \in K, \sum a_i v_i \in W \Rightarrow a_i = 0\ \forall i)$.

(b) I vettori $v_1, \ldots, v_k \in V$ sono una base di V modulo $W \iff \bar{v}_1, \ldots, \bar{v}_k$ sono una base di V/W.

(c) $v_1, \ldots, v_k$ l.i.modulo $W \Rightarrow \exists v_{k+1}, \ldots, v_t$ tali che $v_1, \ldots, v_t$ siano una base modulo W.

Dimostrazione Sia $w_1, \ldots, w_s$ una base per W. La seconda equivalenza in (a) è la definizione di indipendenza lineare nel quoziente, infatti $\sum a_i \bar{v}_i = 0$ in $V/W \iff \sum a_i v_i \in W$.

(a) $\Rightarrow$) : Sia $< v_1, \ldots, v_k > +W =< v_1, \ldots, v_k > \oplus W$; devo provare $\bar{v}_1, \ldots, \bar{v}_k$ l.i. in V/W. Si ha:

$$\sum a_i \bar{v}_i = 0 \Rightarrow \sum a_i v_i \in W \Rightarrow \sum a_i v_i \in < v_1, \ldots, v_k > \cap W \Rightarrow \sum a_i v_i = 0 \Rightarrow a_i = 0\ \forall i = 1, \ldots, k$$

$(a) \Leftarrow)$: Siano $\bar{v}_1, \ldots, \bar{v}_k$ l.i. in V/W, e proviamo che $v_1, \ldots, v_k, w_1, \ldots, w_s$ sono l.i.
Siano $a_1, \ldots, a_k, b_1, \ldots, b_s \in K$ tali che $\sum a_i v_i + \sum b_j w_j = 0$; allora

$$\sum a_i v_i = -\sum b_j w_j \in W \Rightarrow \sum a_i \bar{v}_i = 0 \Rightarrow a_1 = \ldots = a_k = 0 \Rightarrow \sum b_j w_j = 0 \Rightarrow b_1 = \ldots = b_s = 0$$

$(b) \Rightarrow)$: Sia $V =< v_1, \ldots, v_k > \oplus W$; devo provare $V/W =< \bar{v}_1, \ldots, \bar{v}_k >$; poiché $v_1, \ldots, v_k, w_1, \ldots, w_s$ è base per V, $\forall \bar{v} \in V/W$ esistono $a_1, \ldots, a_k, b_1, \ldots, b_s \in K$ tali che $v = \sum a_i v_i + \sum b_j w_j \Rightarrow \bar{v} = \sum a_i \bar{v}_i$.

$(b) \Leftarrow)$: Siano $\bar{v}_1, \ldots, \bar{v}_k$ base in V/W; devo provare che $v_1, \ldots, v_k, w_1, \ldots, w_s$ sono generatori per V:

sia $v \in V$; allora $\exists a_1, \ldots, a_k$ tali che $\bar{v} = \sum ai\bar{v}_i \Rightarrow \bar{v} - \sum a_i \bar{v}_i = 0 \Rightarrow v - \sum a_i v_i \in W \Rightarrow \exists b_1, \ldots, b_s$ tali che $v - \sum a_i v_i = \sum b_j w_j \Rightarrow v = \sum a_i v_i + \sum b_j w_j$

Esempi 3.5 (a) Sia $V = \mathbb{R}^4$ e $W =< e_1, e_2 >$; i vettori e_3, e_4 sono una base di V modulo W.

(b) Sia $V = \mathbb{R}^3$ e $W =< e_1 >$; i vettori $(1, 1, 1)$ e $(3, -1, -1)$ sono l.i. in V ma non sono l.i. modulo W perché $(1, 1, 1) + (3, -1, -1) \in W$.

Lemma 3.6 *Sia V un K-spazio vettoriale f.g. e sia*

$$0 = W_0 \subset W_1 \subset \ldots \subset W_k = V$$

una catena di sottospazi. Sia $(v_{i,1}, \ldots, v_{i,n_i})$ una base di W_i modulo W_{i-1} per $i = 1, \ldots, k$. Allora

$$(v_{1,1}, \ldots, v_{1,n_1}, \ldots, v_{k,1}, \ldots, v_{k,n_k})$$

è una base per V.

Dimostrazione Immediata, osservando che $(v_{1,1}, \ldots, v_{1,n_1})$ è una base per W_1, dunque $(v_{1,1}, \ldots, v_{1,n_1}, v_{2,1}, \ldots, v_{2,n_2})$ è una base per W_2,...

3.2 Matrici a blocchi

Definizione 3.7 Sia $A \in M_{n,m}(K)$ e siano $A_{ij} \in M_{a_i,b_j}(K)$ sue sottomatrici tali che A si scriva:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1r} \\ A_{21} & A_{22} & \dots & A_{2r} \\ \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{sr} \end{pmatrix}$$

Si osservi che ad esempio le matrici della prima riga $A_{11}, A_{12}, \dots, A_{1r}$ sono matrici $a_1 \times b_j$ quindi hanno tutte lo stesso numero di righe, e analogamente le matrici della prima colonna $A_{11}, A_{21}, \dots, A_{s1}$ sono matrici $a_i \times b_1$ quindi hanno tutte lo stesso numero di colonne.

Deve essere inoltre $\sum a_i = n$, $\sum b_j = m$.

Diremo che abbiamo scritto A come *una matrice a blocchi.*

Se $A \in M_n(K)$ è una matrice quadrata a blocchi della forma seguente, dove $A_1 \in M_{a_1}, \dots, A_r \in M_{a_r}$ sono matrici quadrate la cui diagonale principale è contenuta nella diagonale principale di A:

$$A = \begin{pmatrix} A_1 & * & & \dots & * \\ 0 & A_2 & & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_{r-1} & * \\ 0 & 0 & \dots & 0 & A_r \end{pmatrix} \qquad (\spadesuit)$$

diciamo che A è *triangolare superiore a blocchi.*

Se $A \in M_n(K)$ è una matrice quadrata a blocchi della forma seguente, dove $A_1 \in M_{a_1}, \dots, A_r \in M_{a_r}$ sono matrici quadrate la cui diagonale principale è contenuta nella diagonale principale di A:

$$A = \begin{pmatrix} A_1 & 0 & & \dots & 0 \\ 0 & A_2 & & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_{r-1} & \\ 0 & 0 & \dots & 0 & A_r \end{pmatrix} \qquad (\heartsuit)$$

diciamo che A è *diagonale a blocchi.*

Proposizione 3.8 *Sia A una matrice triangolare a blocchi come in ($\spadesuit$); allora*

$$det A = \prod_{i=1}^{r} det(A_i)$$

Dimostrazione Per la dimostrazione si veda ad esempio [C]:

http://mate.unipv.it/cornalba/dispense/alglin/block.pdf

Definizione 3.9 Sia V un K-spazio vettoriale, e sia $f \in End(V)$; se T è un sottospazio di V tale che $f(T) \subseteq T$ diremo che T è un *sottospazio invariante per f o f-invariante.*

Se T è un sottospazio invariante per f, è possibile considerare $f_{|T}^{|T} \in End(T)$.

Proposizione 3.10 *Sia V un K-spazio vettoriale finitamente generato, $dimV = n$, e sia $f \in End(V)$.*
Supponiamo che esistano sottospazi f-invarianti $T_1, \ldots, T_q$ tali che $V = T_1 \oplus \ldots \oplus T_q$; sia $\mathcal{B}_i = (v_{i,1}, \ldots, v_{i,n_i})$ una base di T_i per $i = 1, \ldots, q$, dunque $\mathcal{B} = (v_{1,1}, \ldots, v_{1,n_1}, \ldots, v_{q,1}, \ldots, v_{q,n_q})$ base per V. Infine poniamo

$$f_i := f_{|T_i}^{|T_i}$$

Allora $M_{\mathcal{B}}(f)$ si scrive come matrice diagonale a blocchi così:

$$M_{\mathcal{B}}(f) = \begin{pmatrix} M_{\mathcal{B}_1}(f_1) & 0 & \ldots & 0 \\ 0 & M_{\mathcal{B}_2}(f_2) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & M_{\mathcal{B}_q}(f_q) \end{pmatrix}$$

Dimostrazione Vediamo ad esempio come nasce il primo blocco $M_{\mathcal{B}_1}(f_1)$: sia $j = 1, \ldots, n_1$, $v_{1,j} \in \mathcal{B}_1$; poiché T_1 è invariante, si ha $f(v_{1,j}) = f_1(v_{1,j}) \in T_1$; quindi se $f(v_{1,j}) = (a_{1,j}, \ldots, a_{n_1,j})_{\mathcal{B}_1}$, si ha

$$f(v_{1,j}) = f_1(v_{1,j}) = \sum_{k=1}^{n_1} a_{k,j} v_{1,k} = (a_{1,j}, \ldots, a_{n_1,j}, 0, \ldots, 0)_{\mathcal{B}}$$

quindi le prime n_1 colonne della matrice $M_{\mathcal{B}}(f)$ sono:

$$M_{\mathcal{B}}(f) = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n_1} & \ldots & \ldots \\ a_{2,1} & \ldots & a_{2,n_1} & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n_1,1} & \ldots & a_{n_1,n_1} & \ldots & \ldots \\ 0 & \ldots & 0 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \ldots & \ldots \end{pmatrix}$$

e si ha appunto

$$\begin{pmatrix} a_{1,1} & \ldots & a_{1,n_1} \\ a_{2,1} & \ldots & a_{2,n_1} \\ \ldots & \ldots & \ldots \\ a_{n_1,1} & \ldots & a_{n_1,n_1} \end{pmatrix} = M_{\mathcal{B}_1}(f_1)$$

■

Corollario 3.11 *Nelle ipotesi e notazioni di 3.10 si ha*

$$p_f(t) = \prod_{i=1}^{q} p_{f_i}(t)$$

Osservazione 3.12 Naturalmente la 3.10 si inverte, nel senso che se $f \in End(V)$ ed esiste una base $\mathcal{B}$ per V tale che $M_{\mathcal{B}}(f)$ sia diagonale a blocchi, con blocchi di ordini $n_1, \ldots, n_q$, considerando la partizione di $\mathcal{B}$ data da:

$$\mathcal{B} = (\underbrace{v_{1,1}, \ldots, v_{1,n_1}}_{\textit{primi } n_1 \textit{ vettori}}, \ldots, \underbrace{v_{q,1}, \ldots, v_{q,n_q}}_{\textit{ultimi } n_q \textit{ vettori}})$$

e ponendo $T_i :=< v_{i,1}, \ldots, v_{i,n_i} >$, si ha che i T_i sono f-invarianti e il blocco i-esimo è $M_{\mathcal{B}_i}(f_{|T_i}^{|T_i})$.

Ad esempio, sia $V = \mathbb{R}^5$, e consideriamo la matrice $A = M_{\mathcal{E}}(f)$:

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

qui $T_1 =< e_1, e_2 >$, $T_2 =< e_3, e_4, e_5 >$, e $f(T_1) = T_1$, $f(T_2) =< e_3 + e_4, -e_3 + e_4, 3e_3 >\subset T_2$.

3.3 Endomorfismi nilpotenti

In tutta questa sezione, V è un K-spazio vettoriale, $dimV = n$, ed $f \in EndV$; id denota id_V, e I denota la matrice identità dell'ordine opportuno per il contesto.

Definizione 3.13 Se $k \in \mathbb{N}$ poniamo

$$f^0 := id, \quad f^k := \underbrace{f \circ f \circ \ldots \circ f}_{k \text{ volte}} \quad \text{se } k > 0$$

Se $a \in K$, la notazione af è già stata introdotta così come la somma di endomorfismi di V; quindi se $q = a_m x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0 \in K[x]$, poniamo

$$q(f) := a_m f^m + a_{m-1} f^{m-1} + \ldots + a_1 f + a_0\, id;$$

$q(f)$ sta in $EndV$.

Analogamente, se $A \in M_n(K)$, e $k \in \mathbb{N}$, il prodotto righe per colonne di A con se stessa k volte è denotato con

$$A^0 := I, \quad A^k := \underbrace{A \cdot A \ldots \cdot A}_{k \text{ volte}} \quad \text{se } k > 0$$

Se $a \in K$, la notazione aA è già stata introdotta così come la somma di matrici di $M_n(K)$; quindi se $q = a_m x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0 \in K[x]$ è un polinomio a coefficienti in K, poniamo

$$q(A) := a_m A^m + a_{m-1} A^{m-1} + \ldots + a_1 A + a_0 I;$$

$q(A)$ sta in $M_n(K)$.

Sia $\mathcal{B}$ una base per V e sia $A = M_{\mathcal{B}}(f)$; da 1.49 segue

$$M_{\mathcal{B}}(q(f)) = q(A)$$

Un endomorfismo f *è detto nilpotente* se esiste un intero $m > 0$ tale che $f^m = 0$; *una matrice* $A \in M_n(K)$ *è nilpotente* se esiste un intero $m > 0$ tale che $A^m = 0$.

Ovviamente, se f, risp. A è nilpotente, una sua qualsiasi potenza f^k, risp. A^k, è nilpotente.

Osservazione 3.14 Un endomorfismo f nilpotente può avere solo l'autovalore 0; se infatti v è autovettore relativo all'autovalore $\lambda \neq 0$, per ogni $m > 0$ si ha $f^m(v) = f(f(\ldots f(v)\ldots)) = \lambda(\lambda(\ldots(\lambda v)\ldots)) = \lambda^m v \neq 0$.

Osservazione 3.15 Se $f \in End(V)$ è nilpotente, deve essere $kerf \neq < 0 >$, altrimenti f è un automorfismo e quindi, dato che la composizione di due automorfismi è un automorfismo, qualunque potenza di f è un automorfismo. Inoltre, se $Imf \neq 0$ e $Imf \cap Kerf = < 0 >$, f non può essere nilpotente, cioé $f^k \neq 0$ per tutti i $k > 0$; questo si vede per induzione su k; per $k = 1$ è vero; supponiamo che $f^k \neq 0$; allora $f^{k+1} \neq 0$, infatti esiste un $v \in V$ tale che $f^k(v) \neq 0$, dunque $f^k(v) \in Imf \setminus \{0\} \Rightarrow f^k(v) \notin Kerf \Rightarrow f^{k+1}(v) = f(f^k(v)) \neq 0$.

Quindi: se $f \in End(V)$ è nilpotente ogni sua potenza f^t è nilpotente e quindi per ogni $t > 0$ per cui $Imf^t \neq 0$ deve essere $Imf^t \cap Kerf^t \neq < 0 >$

Esempio 3.16 Costruiamo un endomorfismo nilpotente $f : \mathbb{R}^4 \to \mathbb{R}^4$; daremo f assegnando i suoi valori sulla base canonica. Scegliamo per esempio $Kerf =< e_1 >$ e poniamo $f(e_2) = e_1 \in Kerf$. Andando avanti così poniamo $f(e_k) = e_{k-1}$ per $k = 2, 3, 4$. Si ha: $f(e_1) = 0, \quad f^2(e_2) = f(f(e_2)) = f(e_1) = 0, \quad f^3(e_3) = f^2(f(e_3)) = f^2(e_2) = 0, \quad f^4(e_4) = f^3(f(e_4)) = f^3(e_3) = 0$. Quindi f così costruito è nilpotente e $min\{m > 0,\ f^m = 0\} = 4$.

Esempio 3.17 Consideriamo la matrice

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

si verifica subito che

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

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

quindi se $V := \mathbb{R}^4$, $g \in EndV$, $g = F^{\mathcal{E}}(N)$, si ha che N e g sono nilpotenti, e il $min\{m > 0, g^m = 0\} = 3$.

Consideriamo gli endomorfismi g, g^2 e $g^3 = 0$; si ha $rgN = 2$, $rgN^2 = 1$ e ovviamente $rgN^3 = 0$, quindi $dimKer\, g = 2$, $dimKer\, g^2 = 3$ e $dimKer\, g^3 = 4$; osserviamo che il rango decresce e ad un certo punto diventa 0, e quindi la dimensione dei nuclei cresce ad un certo punto diventa $dimV = 4$; calcoliamo ora questi sottospazi.

$W_1 := Kerg$ è lo spazio delle soluzioni del sistema $NX = 0$:

$$\begin{pmatrix} 2 & 1 & -3 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{equivalente a } \begin{cases} x_2 - x_3 = 0 \\ x_1 - x_3 = 0 \end{cases}$$

$W_2 := Kerg^2$ è lo spazio delle soluzioni del sistema $N^2X = 0$:

$$\begin{pmatrix} 1 & 1 & -2 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{equivalente a} \quad x_1 + x_2 - 2x_3 = 0$$

e $W_3 := Ker\, g^3 = V$; abbiamo quindi costruito una successione crescente di sottospazi:

$$W_0 = 0 \subset \underbrace{W_1}_{\substack{Ker\, g \\ dim\, 2}} \subset \underbrace{W_2}_{\substack{Ker\, g^2 \\ dim\, 3}} \subset \underbrace{W_3 = V}_{\substack{Ker\, g^3 \\ dim\, 4}}$$

Osserviamo che $v \in W_i \Rightarrow \underbrace{g^i(v)}_{g^{i-1}(g(v))} = 0 \Rightarrow g(v) \in W_{i-1}$. Adesso costruiamo:

una base di $W_3 = V$ modulo W_2: $\quad e_1 \quad$ (infatti $e_1 \notin W_2 \Rightarrow < e_1 > \oplus W_2 = V$)

una base di W_2 modulo W_1: $\quad g(e_1) = (2, 0, 1, 0) \quad$ (infatti $e_1 \in W_3 \Rightarrow g(e_1) \in W_2$, e poiché $g(e_1) \notin W_1 \Rightarrow < g(e_1) > \oplus W_1 = W_2$)

una base di W_1 modulo W_0: $\quad g^2(e_1) = (1, 1, 1, 0),\ e_4 \quad$ (infatti $g^2(e_1) = g(g(e_1)) \in W_1$, ed essendo $W_0 = 0$, basta completare ad una base di W_1)

si ha quindi che $\mathcal{B} := (g^2(e_1), g(e_1), e_1, e_4)$ è una base di V,

$$M_{\mathcal{B}}(g) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{$*$}$$

e $M_{\mathcal{B}}(g) = M_{\mathcal{E},\mathcal{B}}(id)^{-1}\, M_{\mathcal{E}}(g)\, M_{\mathcal{E},\mathcal{B}}(id)$, cioé, se

$$P := M_{\mathcal{E},\mathcal{B}}(id) = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{e quindi} \qquad P^{-1} = M_{\mathcal{B},\mathcal{E}}(id) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

si ha

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = P^{-1} N P$$

Diremo che abbiamo messo la matrice N in forma di Jordan, perché abbiamo trovato una matrice della forma $(*)$ simile ad N; questo è un caso molto particolare perché essendo l'endomorfismo di partenza nilpotente, può avere solo l'autovalore nullo.

Si osservi che g non è dz perché ha solo l'autovalore 0 con $\mu(0) = 4$, ma l'autospazio V_0 è $Kerg$ e abbiamo visto che $dimKerg = 2$.

Il prossimo teorema generalizza quanto visto in questo esempio.

Definizione 3.18 Siano $a \in K$ e $k \in \mathbb{N}$, $k > 0$; *il blocco di Jordan di ordine k e di autovalore a* è la matrice quadrata di ordine k:

$$J_k(a) = \begin{pmatrix} a & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & a & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & a & 1 & 0 & \dots & 0 \\ & & & \ddots & & & \\ 0 & \dots & 0 & \dots & \dots & a & 1 \\ 0 & \dots & \dots & \dots & \dots & 0 & a \end{pmatrix}$$

il cui polinomio caratteristico è $(a - t)^k$. Si osservi che 1 compare $k - 1$ volte in $J_k(a)$.

Definizione 3.19 Diciamo che una matrice $A \in M_n(K)$ è *in forma (canonica) di Jordan* se è una matrice diagonale a blocchi, e i blocchi sulla diagonale sono blocchi di Jordan:

$$A = \begin{pmatrix} J_{m_{1,1}}(\lambda_1) & & & & & & \\ & \ddots & & & & & \\ & & J_{m_{1,q_1}}(\lambda_1) & & & & \\ & & & \ddots & & & \\ & & & & J_{m_{s,1}}(\lambda_s) & & \\ & & & & & \ddots & \\ & & & & & & J_{m_{s,q_s}}(\lambda_s) \end{pmatrix}$$

Si osservi che una tale matrice ha ordine

$$n = m_{1,1} + \ldots + m_{1,q_1} + \ldots + m_{s,1} + \ldots + m_{s,q_s}$$

e polinomio caratteristico

$$p_A = (\lambda_1 - t)^{m_{1,1}+\ldots+m_{1,q_1}} \cdot \ldots \cdot (\lambda_s - t)^{m_{s,1}+\ldots+m_{s,q_s}}$$

Teorema 3.20 *Sia $A \in M_n(K)$ una matrice nilpotente; allora A è simile ad una matrice in forma di Jordan, quindi, dato che l'unico autovalore è 0, ad una matrice del tipo:*

$$\begin{pmatrix} J_{m_1}(0) & & & \\ & \ddots & & \\ & & \ddots & \\ & & & J_{m_q}(0) \end{pmatrix} \qquad (\star)$$

Equivalentemente, sia V un K-spazio vettoriale, $dimV = n$, e $g \in EndV$ un endomorfismo nilpotente; allora esiste una base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}(g)$ sia in forma di Jordan $(\star)$.

Esempio 3.21 Vediamo che aspetto ha $(\star)$ se $n = 6$, quando $q = 2$ e $m_1 = 3, m_2 = 3$:

$$\begin{pmatrix} \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} & \\ & \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} \end{pmatrix}$$

quando $q = 2$ e $m_1 = 4$, $m_2 = 2$:

$$\begin{pmatrix} \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} & \\ & \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} \end{pmatrix}$$

e quando $q=3$ e $m_1=m_2=m_3=2$:

$$\begin{pmatrix} \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} & & \\ & \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} & \\ & & \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} \end{pmatrix}$$

Dimostrazione di Teorema 3.20 Se $g=0$ abbiamo finito. Se $g \neq 0$, sia

$$k := min\{m > 0,\ g^m = 0\}$$

e poniamo

$$W_i := ker\, g^i, \quad i = 0, \ldots, k;$$

quindi $W_0 = 0, W_k = V$ e valgono i seguenti fatti:

(a) $W_{i-1} \subseteq W_i$, infatti $v \in W_{i-1} \Rightarrow g^{i-1}(v) = 0 \Rightarrow g^i(v) = g(g^{i-1}(v)) = g(0) = 0 \Rightarrow v \in W_i$

(b) $g(W_i) \subseteq W_{i-1}$, infatti $v \in W_i \Rightarrow \underbrace{g^i(v)}_{g^{i-1}(g(v))} = 0 \Rightarrow g(v) \in W_{i-1}$

(c) $g^{-1}(W_{i-1}) \subseteq W_i$, infatti $v \in g^{-1}(W_{i-1}) \Rightarrow g(v) \in W_{i-1} \Rightarrow \underbrace{g^{i-1}(g(v))}_{g^i(v)} = 0 \Rightarrow v \in W_i$

(d) $W_i \underbrace{\subseteq}_{\text{vale per qualsiasi funzione}} g^{-1}g(W_i) \underbrace{\subseteq}_{\text{vale per (b)}} g^{-1}(W_{i-1})$

(e) $g^{-1}(W_{i-1}) = W_i$ per (c)+(d)

(f) $W_{i-1} \subset W_i \quad \forall i$, infatti se esiste j, $0 \leq j \leq k$, tale che $W_{j-1} = W_j \Rightarrow W_j = g^{-1}(W_{j-1}) = g^{-1}(W_j) = W_{j+1}$ e andando avanti così $W_{j-1} = W_j = W_{j+1} = \ldots = W_k = V \Rightarrow g^{j-1} = 0$ ma $j-1 < k$ e quindi contraddiciamo l'ipotesi $k = min\{m > 0, g^m = 0\}$.

Riassumendo, abbiamo:

$$\begin{array}{cccccccc} & Ker\, g & Ker\, g^2 & Ker\, g^{i-1} & Ker\, g^i & Ker\, g^k & & \\ 0 = W_0 \subset & W_1 \subset & W_2 \subset & \ldots W_{i-1} \subset & W_i \subset \ldots \subset & W_k = V & & \\ & \downarrow g & \downarrow g & \downarrow g & \downarrow g & \downarrow g & & \\ & 0 = W_0 \subset & W_1 & \ldots \subset W_{i-2} \subset & W_{i-1} \subset & \ldots W_{k-1} & \subset & W_k = V \end{array}$$

Adesso costruiamo una base nel modo seguente:

Prendiamo vettori $v_{k,1}, \ldots, v_{k,m_k} \in W_k$ che siano una base modulo W_{k-1}. Allora in particolare $v_{k,1}, \ldots, v_{k,m_k}$ sono vettori l.i. modulo W_{k-1}; per (b)

$$g(v_{k,1}), \ldots, g(v_{k,m_k}) \in W_{k-1}$$

e questi vettori sono l.i. modulo W_{k-2}, infatti

$$a_1 g(v_{k,1}) + \ldots + a_{m_k} g(v_{k,m_k}) = g(a_1 v_{k,1} + \ldots + a_{m_k} v_{k,m_k}) \in W_{k-2} \Rightarrow$$
$$g^{k-2} g(a_1 v_{k,1} + \ldots + a_{m_k} v_{k,m_k}) = 0 \Rightarrow a_1 v_{k,1} + \ldots + a_{m_k} v_{k,m_k} \in W_{k-1} \Rightarrow a_1 = \ldots = a_{m_k} = 0$$

Quindi esistono vettori $v_{k-1,1}, \ldots, v_{k-1,m_{k-1}} \in W_{k-1}$ tali che

$$g(v_{k,1}), \ldots, g(v_{k,m_k}), v_{k-1,1}, \ldots, v_{k-1,m_{k-1}} \text{ base di } W_{k-1} \text{ modulo } W_{k-2}$$

e andando avanti così esistono vettori $v_{k-2,1}, \ldots, v_{k-2,m_{k-2}} \in W_{k-2}$ tali che

$g^2(v_{k,1}), \ldots, g^2(v_{k,m_k}), g(v_{k-1,1}), \ldots, g(v_{k-1,m_{k-1}}), v_{k-2,1}, \ldots, v_{k-2,m_{k-2}}$ base di W_{k-2} modulo W_{k-3}

$\ldots$

$g^i(v_{k,1}), \ldots, g^i(v_{k,m_k}), g^{i-1}(v_{k-1,1}), \ldots, g^{i-1}(v_{k-1,m_{k-1}}), \ldots, v_{k-i,1}, \ldots, v_{k-i,m_{k-i}}$ base di W_{k-i} modulo W_{k-i-1}

$\ldots$

$g^{k-1}(v_{k,1}), \ldots, g^{k-1}(v_{k,m_k}), g^{k-2}(v_{k-1,1}), \ldots, g^{k-2}(v_{k-1,m_{k-1}}), \ldots, v_{1,1}, \ldots, v_{1,m_1}$ base di W_1 modulo W_0

Facendo l'unione di tutte queste basi abbiamo una base $\mathcal{B}$ di $W_k = V$; riordinando i vettori tale base si può scrivere:

$$\begin{array}{ccccc}
g^{k-1}(v_{k,1}) & g^{k-2}(v_{k,1}) & \ldots & g(v_{k,1}) & v_{k,1} \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
g^{k-1}(v_{k,m_k}) & g^{k-2}(v_{k,m_k}) & \ldots & g(v_{k,m_k}) & v_{k,m_k} \\
 & g^{k-2}(v_{k-1,1}) & \ldots & g(v_{k-1,1}) & v_{k-1,1} \\
 & \ldots & \ldots & \ldots & \ldots \\
 & g^{k-2}(v_{k-1,m_{k-1}}) & \ldots & g(v_{k-1,m_{k-1}}) & v_{k-1,m_{k-1}} \\
 & & \ldots & \ldots & \ldots \\
 & & & & v_{1,1} \\
 & & & & \ldots \\
 & & & & v_{1,m_1}
\end{array} \qquad (\diamondsuit)$$

Sia $v = v_{i,j}$ uno dei vettori della base e sia $s \geq 0$ tale che $g^s(v), g^{s-1}(v), \ldots, g(v), v$ compaiono in $\mathcal{B}$ ma $g^{s+1}(v)$ non compare in $\mathcal{B}$; questo significa che v è apparso per la prima volta come vettore di una base di W_{s+1} modulo W_s, quindi $g^{s+1} = 0$.

Sia $T := < g^s(v), g^{s-1}(v), \ldots, g(v), v >$; allora $g(T) = < g(g^s(v)), g(g^{s-1}(v)), \ldots, g(g(v)), g(v) > \subseteq T$, quindi T è invariante per g. Se denotiamo con $\mathcal{C}$ la base $(g^s(v), g^{s-1}(v), \ldots, g(v), v)$ di T, risulta:

$$M_{\mathcal{C}}(g|_{T}^{T}) = \begin{pmatrix}
0 & 1 & 0 & \ldots & \ldots & \ldots & 0 \\
0 & 0 & 1 & 0 & \ldots & \ldots & 0 \\
0 & 0 & 0 & 1 & 0 & \ldots & 0 \\
 & & & \ddots & & & \\
0 & \ldots & 0 & \ldots & \ldots & 0 & 1 \\
0 & \ldots & \ldots & \ldots & \ldots & 0 & 0
\end{pmatrix} = J_{s+1}(0)$$

Da 3.10 si ha quindi la tesi; i sottospazi invarianti in cui è decomposto V sono i sottospazi generati dai vettori su una singola riga di $(\diamondsuit)$.

■

3.4 La forma di Jordan

Esempio 3.22 Consideriamo la matrice

$$A = \begin{pmatrix} 4 & 1 & -3 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

e calcoliamone il polinomio caratteristico:

$$det(A - \lambda I) = \begin{vmatrix} 4-\lambda & 1 & -3 & 0 \\ 0 & 1-\lambda & 1 & 0 \\ 1 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda) \begin{vmatrix} 4-\lambda & 1 & -3 \\ 0 & 1-\lambda & 1 \\ 1 & 0 & 1-\lambda \end{vmatrix} =$$

$$= (2-\lambda)\left(\begin{vmatrix} 1 & -3 \\ 1-\lambda & 1 \end{vmatrix} + (1-\lambda) \begin{vmatrix} 4-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} \right) = (2-\lambda)(1 - 3 + 3\lambda + (1-\lambda)^2(4-\lambda) =$$

$$= (\lambda - 2)(\lambda^3 - 6\lambda^2 + 12\lambda - 8) = (\lambda - 2)^4$$

quindi c'è un solo autovalore, 2, di molteplicità algebrica 4.

La matrice $A - 2I = \begin{pmatrix} 2 & 1 & -3 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ è la matrice N di esempio 3.17.

Sia $V := \mathbb{R}^4$, $f \in EndV$, $f = F^{\mathcal{E}}(A)$ e usiamo le notazioni di 3.17: si ha quindi

$$f = 2id + g, \quad A = 2I + N = M_{\mathcal{E}}(2id + g)$$

Avevamo visto che

$$M_{\mathcal{B}}(g) = M_{\mathcal{E},\mathcal{B}}(id)^{-1} \, M_{\mathcal{E}}(g) \, M_{\mathcal{E},\mathcal{B}}(id)$$

si ha quindi, posto $P := M_{\mathcal{E},\mathcal{B}}(id)$:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = P^{-1} N P$$

da cui

$$2I + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 2I + P^{-1}NP = P^{-1}(2I + N)P = P^{-1}AP$$

cioé

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = P^{-1}AP \qquad (*)$$

Abbiamo quindi messo la matrice A in forma di Jordan; in $(*)$ compaiono due blocchi di Jordan, $J_3(2)$ e $J_1(2)$.

Teorema 3.23 *Sia $A \in M_n(K)$ una matrice tale che valga la condizione*

(a) il polinomio caratteristico ha n radici contate con molteplicità in K, cioé, se $\lambda_1, \ldots, \lambda_s$ sono i suoi autovalori a due a due distinti, vale $\mu(\lambda_1) + \ldots + \mu(\lambda_s) = n$.

Allora A è simile ad una matrice in forma di Jordan:

$$\begin{pmatrix} J_{m_{1,1}}(\lambda_1) & & & & & & \\ & \ddots & & & & & \\ & & J_{m_{1,q_1}}(\lambda_1) & & & & \\ & & & \ddots & & & \\ & & & & J_{m_{s,1}}(\lambda_s) & & \\ & & & & & \ddots & \\ & & & & & & J_{m_{s,q_s}}(\lambda_s) \end{pmatrix} \qquad (\star\star)$$

dove $m_{i,1} + \ldots + m_{i,q_i} = \mu(\lambda_i)$ per $i = 1, \ldots, s$.

Equivalentemente, sia V un K-spazio vettoriale, $dimV = n$, e $f \in EndV$ un endomorfismo tale che valga la condizione (a); allora esiste una base $\mathcal{B}$ di V, che chiameremo base di Jordan per f, tale che $M_\mathcal{B}(f)$ sia in forma di Jordan $(\star\star)$.

Se $K = \mathbb{C}$, la condizione (a) è sempre soddisfatta, quindi una matrice in $M_n(\mathbb{C})$ è sempre simile ad una matrice in forma di Jordan.

Per la dimostrazione faremo uso del teorema che segue, per la cui dimostrazione si rimanda a [C], (5.8):

Teorema 3.24 *Sia V un K-spazio vettoriale, $dimV = n$, e $f \in EndV$ un endomorfismo tale che valga la condizione*

(a) il polinomio caratteristico ha n radici contate con molteplicità in K, cioé, se $\lambda_1, \ldots, \lambda_s$ sono i suoi autovalori a due a due distinti, vale $\mu(\lambda_1) + \ldots + \mu(\lambda_s) = n$.

Allora, ponendo $U_i := Ker(f - \lambda_i\, id)^{\mu(\lambda_i)}$, $i = 1, \ldots, s$, si ha:

$$dimU_i = \mu(\lambda_i), \qquad V = U_1 \oplus \ldots \oplus U_s$$

Dimostrazione di 3.23.
Passo 1: I sottospazi U_i sono f-invarianti: proviamolo:

$$v \in U_i \overset{?}{\Rightarrow} f(v) \in U_i$$

Sia $\lambda = \lambda_i, \mu = \mu(\lambda_i)$; dobbiamo provare

$$(f - \lambda\, id)^\mu(v) = 0 \Rightarrow (f - \lambda\, id)^\mu(f(v)) = 0$$

Si vede subito che per ogni $w \in V$ si ha $(f \circ (f - \lambda id))(w) = ((f - \lambda id) \circ f)(w)$ da cui segue che $f \circ (f - \lambda id) = (f - \lambda id) \circ f$, e da qui per induzione si vede che per ogni $t > 0$:

$$(f - \lambda\, id)^t \circ f = f \circ (f - \lambda\, id)^t$$

Quindi

$$(f - \lambda\, id)^{\mu}(f(v)) = ((f - \lambda\, id)^{\mu} \circ f)(v) = (f \circ (f - \lambda\, id)^{\mu})(v) = f((f - \lambda\, id)^{\mu}(v)) = f(0) = 0.$$

Passo 2: Poiché $V = U_1 \oplus \ldots \oplus U_s$ e gli U_i sono f-invarianti, per 3.10 se $\mathcal{B}_i$ è una base di U_i per $i = 1, \ldots, s$, $\mathcal{B}$ è la base di V unione delle $\mathcal{B}_i$, e $f_i := f_{|U_i}^{|U_i}$, allora $M_{\mathcal{B}}(f)$ si scrive come matrice diagonale a blocchi così:

$$M_{\mathcal{B}}(f) = \begin{pmatrix} M_{\mathcal{B}_1}(f_1) & 0 & \ldots & 0 \\ 0 & M_{\mathcal{B}_2}(f_2) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & M_{\mathcal{B}_s}(f_s) \end{pmatrix}$$

Passo 3: Studiamo un singolo blocco $M_{\mathcal{B}_i}(f_i)$: poniamo

$$g_i := f_i - \lambda_i\, id$$

allora $g_i \in End(U_i)$ è nilpotente, quindi per il Teorema 3.20 U_i ha una base $\mathcal{B}_i$ tale che $M_{\mathcal{B}_i}(g_i)$ sia in forma di Jordan, con tutti 0 sulla diagonale principale:

$$\begin{pmatrix} J_{m_{i,1}}(0) & & & \\ & \ddots & & \\ & & \ddots & \\ & & & J_{m_{i,q_i}}(0) \end{pmatrix} \qquad (\star)$$

D'altra parte $f_i = g_i + \lambda_i\, id$ per cui

$$M_{\mathcal{B}_i}(f_i) = M_{\mathcal{B}_i}(g_i + \lambda_i\, id) = M_{\mathcal{B}_i}(g_i) + \lambda_i I = \begin{pmatrix} J_{m_{i,1}}(\lambda_i) & & & \\ & \ddots & & \\ & & \ddots & \\ & & & J_{m_{i,q_i}}(\lambda_i) \end{pmatrix}$$

e mettendo insieme i blocchi otteniamo $(\star\star)$. ■

Proposizione 3.25 *Sia $A \in M_n(\mathbb{C})$; la matrice in forma di Jordan a cui è simile A è univocamente determinata a meno dell'ordine con cui compaiono i blocchi di Jordan.*
Quindi possiamo dire che $(\star\star)$ è la forma canonica di Jordan di A.

Per la dimostrazione si veda ad esempio [C2].

Proposizione 3.26 *Due matrici $A, B \in M_n(\mathbb{C})$ sono simili se e solo se hanno la stessa forma di Jordan (a meno dell'ordine con cui compaiono i blocchi di Jordan).*

Dimostrazione Basta usare la proprietà transitiva della similitudine tra matrici.

Osservazione 3.27 Facciamo il punto della situazione: come si fa a mettere un endomorfismo o una matrice in forma di Jordan, nel caso questo sia possibile? Se ci è data una matrice $A \in M_n(K)$ pensiamo ad A come matrice di un endomorfismo f dello spazio vettoriale K^n rispetto alla base canonica $\mathcal{E}$, quindi $A = M_{\mathcal{E}}(f)$.

Vediamolo quindi per gli endomorfismi. Useremo le notazioni di Teorema 3.24. Sia quindi V un K-spazio vettoriale, $dimV = n$, $\mathcal{B}$ base di V, $f \in EndV$ un endomorfismo e $A = M_{\mathcal{B}}(f)$; ci chiediamo se esiste una matrice in forma di Jordan che rappresenta f e in tal caso qual è; per finire, ci possiamo chiedere qual è una sua base di Jordan.

a) esiste una forma di Jordan? Cerchiamo gli autovalori di f e le loro molteplicità algebriche; siano $\lambda_1, \dots, \lambda_s$ i suoi autovalori a due a due distinti. Se non vale la condizione (a) di Teorema 3.24: $\mu(\lambda_1) + \ldots + \mu(\lambda_s) = n$ la risposta è no e ci fermiamo, altrimenti proseguiamo. Si osservi infatti che la condizione (a) è non solo sufficiente, come ci dice Teorema 3.24, ma anche necessaria all'esistenza di una matrice in forma di Jordan che rappresenta f, perché se tale matrice esiste, il suo polinomio caratteristico ha n radici contate con molteplicità.

b) costruzione di una base di Jordan: Se da a) sappiamo che una forma di Jordan esiste, cerchiamo i sottospazi f-invarianti

$$U_i = Ker\,(f - \lambda_i\, id)^{\mu(\lambda_i)}, \qquad i = 1, \dots, s$$

risolvendo i sistemi lineari:

$$U_i : \quad (A - \lambda_i I)^{\mu(\lambda_i)} X = 0, \qquad i = 1, \dots, s$$

Ponendo

$$f_i := f_{|U_i}^{|U_i}, \qquad g_i := f_i - \lambda_i id_{U_i}$$

per ogni $i = 1, \dots, s$ l'endomorfismo $g_i \in End(U_i)$ è nilpotente, quindi possiamo fare come nella dimostrazione di Teorema 3.20, che è una dimostrazione costruttiva, e trovare una base di Jordan per ogni g_i, sia $\mathcal{C}_i = (v_{i,1}, \dots, v_{i,\mu(\lambda_i)})$. L'unione di tutte queste basi è una base di Jordan per f. Infatti sappiamo che si ha:

$$dimU_i = \mu(\lambda_i), \qquad V = U_1 \oplus \ldots \oplus U_s,$$

l'unione delle basi $\mathcal{C} = (v_{1,1}, \dots, v_{1,\mu(\lambda_1)}, \dots, v_{s,1}, \dots, v_{s,\mu(\lambda_s)})$ è una base per V e $M_{\mathcal{C}}(f)$ si scrive come matrice diagonale a blocchi così:

$$M_{\mathcal{C}}(f) = \begin{pmatrix} M_{\mathcal{C}_1}(f_1) & 0 & \ldots & 0 \\ 0 & M_{\mathcal{C}_2}(f_2) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & M_{\mathcal{C}_s}(f_s) \end{pmatrix}$$

dove l'i-esimo blocco sulla diagonale è una matrice di ordine $\mu(\lambda_i)$ in forma di Jordan.

d) trovare la forma di Jordan per f senza trovare una base di Jordan:

Rispondiamo al problema per un $g \in End(V)$ che sia nilpotente, perché fatto questo si procede come in b). Sia $k := min\{m > 0,\ g^m = 0\}$; ponendo $W_i := Ker(g^i)$, se B è una matrice che rappresenta g, si ha:

$$dimW_i = n - rg(B^i), \qquad i = 0, \dots, k$$

quindi

$$dimW_t - dimW_{t-1} = (n - rgB^t) - (n - rgB^{t-1}) = rgB^{t-1} - rgB^t \qquad (*)$$

Seguendo la dimostrazione di Teorema 3.20, vediamo che nella forma di Jordan di g ci sono

m_k blocchi di ordine k, m_{k-1} blocchi di ordine $k-1$, ..., m_1 blocchi di ordine 1 dove

m_k = numero dei vettori che formano una base di $W_k = V$ modulo W_{k-1},

m_{k-1}=numero dei vettori che, insieme ad altri m_k vettori, formano una base di W_{k-1} modulo W_{k-2},

m_{k-2}=numero dei vettori che, insieme ad altri $m_k + m_{k-1}$ vettori, formano una base di W_{k-2} modulo W_{k-3},

...

m_1=numero dei vettori che, insieme ad altri $m_k + m_{k-1} + \ldots + m_2$ vettori, formano una base di W_1 modulo $W_0 =< 0 >$

e per scrivere la matrice in forma di Jordan ci basta conoscere gli m_i. Si ha quindi, tenendo conto di $(*)$:

$m_k = dimW_k - dimW_{k-1} = rgB^{k-1} - rgB^k = rgB^{k-1}$

$m_{k-1} + m_k = dimW_{k-1} - dimW_{k-2} = rgB^{k-2} - rgB^{k-1}$

$m_{k-2} + m_{k-1} + m_k = dimW_{k-2} - dimW_{k-3} = rgB^{k-3} - rgB^{k-2}$

...

$m_1 + \ldots + m_k = dimW_1 - dimW_0 = rgB^0 - rgB^1 = n - rgB^1$

cioé

$m_j + \ldots + m_k = rgB^{j-1} - rgB^j$ per $j = 1, \ldots, k$; se $j \leq k-1$ si ha quindi

$$m_j = rgB^{j-1} - rgB^j - (m_{j+1} + \ldots + m_k) = rgB^{j-1} - rgB^j - (rgB^j - rgB^{j+1}) = rgB^{j-1} - 2rgB^j + rgB^{j+1}$$

e poiché $B^k = B^{k+1} = 0$ questa uguaglianza vale anche per $j = k$.

Si osservi che in particolare ne segue che il numero dei blocchi di Jordan per g è $m_1 + \ldots + m_k = dimKerg$; quindi se $f \in End(V)$ ha un unico autovalore λ di molteplicità algebrica n, allora il numero dei blocchi di Jordan di una matrice di Jordan che rappresenta f è la molteplicità geometrica di λ.

4 Spazi vettoriali euclidei

4.1 Forme bilineari

Definizione 4.1 Sia V un K-spazio vettoriale. Una applicazione di insiemi

$$f : V \times V \to K$$

si dice forma bilineare su V se è lineare su entrambi i fattori, cioé se $\forall a, b \in K, \forall v, v', w, w' \in V$:

$$f(av + bv', w) = af(v, w) + bf(v', w), \quad f(v, aw + bw') = af(v, w) + bf(v, w')$$

Una forma bilineare f su V si dice simmetrica se

$$f(v, w) = f(w, v) \quad \forall v, w \in V$$

Si osservi che $\forall v, w \in V$ si ha $f(0, w) = f(v, 0) = 0$.

Definizione 4.2 Sia V un K-spazio vettoriale, $dimV = n$, sia $\mathcal{B} = (v_1, \dots, v_n)$ una base di V e sia f una forma bilineare su V; *la matrice di f rispetto a $\mathcal{B}$* è la matrice

$$Mat(f, \mathcal{B}) := (a_{ij}) \qquad \text{dove} \quad a_{ij} = f(v_i, v_j), \quad i, j = 1, \dots, n$$

Esempio 4.3 $V = \mathbb{R}^2$,

$$\begin{array}{rccc} f: & V \times V & \to & \mathbb{R} \\ & ((x_1, x_2), (y_1, y_2)) & \mapsto & x_1y_1 - x_1y_2 + 3x_2y_1 + 4x_2y_2 \end{array}$$

è una forma bilineare su V; non è simmetrica, perché ad esempio $f(e_1, e_2) = -1$, $f(e_2, e_1) = 3$. Si ha

$$Mat(f, \mathcal{E}) = \begin{pmatrix} 1 & -1 \\ 3 & 4 \end{pmatrix}$$

Proposizione 4.4 *Nelle notazioni di 4.2 si ha, se $v = (x_1, \dots, x_n)_{\mathcal{B}}$ e $w = (y_1, \dots, y_n)_{\mathcal{B}}$, ponendo $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ risulta:*

$$f(v, w) = {}^tX \, Mat(f, \mathcal{B}) \, Y$$

Dimostrazione Sia $Mat(f, \mathcal{B}) = (a_{ij})$. Risulta:

$$f(v, w) = f(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j) = \sum_{i=1}^n x_i f(v_i, \sum_{j=1}^n y_j v_j) = \sum_{i,j=1}^n x_i y_j f(v_i, v_j) = \sum_{i,j=1}^n x_i y_j a_{ij};$$

$${}^tX \, Mat(f, \mathcal{B}) \, Y = (\, x_1 \quad \dots \quad x_n \,) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (\, x_1 \quad \dots \quad x_n \,) \begin{pmatrix} a_{11}y_1 + \dots + a_{1n}y_n \\ \vdots \\ a_{n1}y_1 + \dots + a_{nn}y_n \end{pmatrix} =$$

$$= \sum_{i=1}^n x_i(a_{i1}y_1 + \dots + a_{in}y_n) = \sum_{i,j=1}^n a_{ij} x_i y_j.$$

■

Osservazione 4.5 Usiamo le notazioni di 4.4. Data una matrice $A \in M_n(K)$ e fissata la base $\mathcal{B}$, è possibile associarle una forma bilineare f su V nel modo seguente:

$$f(v, w) := {}^tX\,A\,Y$$

Si vede subito che la forma bilineare è simmetrica se e solo se lo è la matrice associata ad f rispetto a $\mathcal{B}$.
In questo modo abbiamo stabilito una corrispondenza biunivoca, che dipende dalla scelta della base $\mathcal{B}$, tra forme bilineari su V e matrici di $M_n(K)$, e questa induce una corrispondenza biunivoca tra forme bilineari simmetriche su V e matrici simmetriche di $M_n(K)$.

4.2 Prodotti scalari

D'ora in poi fino alla fine del capitolo $K = \mathbb{R}$.

Definizione 4.6 Sia V un $\mathbb{R}$-spazio vettoriale, $dimV = n$; una forma bilineare simmetrica f su V si dice *definita positiva* se

$$f(v, v) > 0 \quad \forall v \in V, v \neq 0$$

Una forma bilineare simmetrica definita positiva su V è detta *prodotto scalare* .

Un $\mathbb{R}$-spazio vettoriale con un assegnato prodotto scalare f è detto *spazio vettoriale euclideo* (abbreviato in spazio v.e.); se $v, w \in V$, il prodotto scalare dei due vettori verrà denotato, anziché con $f(v, w)$, con $v \bullet w$.

Si osservi che se W è un sottospazio di V, allora W con il prodotto scalare indotto da quello di V è ancora uno spazio v.e.

Definizione 4.7 $\mathbb{R}^n$ con il seguente prodotto scalare:

$$(x_1, \ldots, x_n) \bullet (y_1, \ldots, y_n) = x_1y_1 + \ldots + x_ny_n \tag{$*$}$$

è uno spazio vettoriale euclideo; il prodotto scalare $(*)$ è detto il prodotto scalare standard.

Definizione 4.8 Sia V un spazio v.e., $dimV = n$. La *norma di un vettore* v è

$$||v|| := \sqrt{v \bullet v}$$

Un vettore di norma 1 è detto *versore.*
Diciamo che due vettori v, w sono *ortogonali* se

$$v \bullet w = 0$$

0 è ortogonale ad ogni $v \in V$.

Un insieme finito di vettori non nulli $\{v_1, \ldots, v_t\}$ è un *insieme ortogonale di vettori* se $v_i \bullet v_j = 0$ $\forall i \neq j, 1 \leq i, j \leq t$; è un *insieme ortonormale di vettori* se inoltre i v_i sono versori; analoghe definizioni si danno per una base.

Quindi una base $\mathcal{B} = (v_1, \ldots, v_n)$ è una base ortonormale se e solo se

$$v_i \bullet v_j = \delta_{ij} \ = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Osservazione 4.9 *a*) Sia $\{v_1, \ldots, v_t\}$ un insieme ortogonale di vettori; allora

a)

$$\{\frac{v_1}{||v_1||}, \ldots, \frac{v_t}{||v_t||}\}$$

è un insieme ortonormale di vettori (possiamo dividere per $||v_i||$ perché per ipotesi $v_i \neq 0$ $\forall i$), infatti:

$$\frac{v_i}{||v_i||} \bullet \frac{v_i}{||v_i||} = \frac{v_i \bullet v_i}{||v_i||^2} = 1$$

b) i $v_1, \dots, v_t$ sono l.i., infatti: sia $a_1 v_1 + \ldots + a_t v_t = 0$; allora per ogni i

$$0 = (a_1 v_1 + \ldots + a_t v_t) \bullet v_i = a_i (\underbrace{v_i \bullet v_i}_{>0}) \Rightarrow a_i = 0$$

Esempi 4.10 a) In $\mathbb{R}^n$ con il prodotto scalare standard la base canonica è una base ortonormale.

b) In $\mathbb{R}^2$ con il prodotto scalare standard ogni coppia di vettori ortogonali sulla circonferenza unitaria è una base ortonormale.

c) In $\mathbb{R}^3$ con il prodotto scalare standard tagliando la sfera unitaria S^2 con un piano per O, prendendo una coppia di vettori ortogonali sulla circonferenza intersezione che si ottiene, e poi prendendo un versore sulla retta (vettoriale) ortogonale al piano, si ha una base ortonormale.

Proposizione-Definizione 4.11 *Sia V uno spazio v.e., $dimV = n$, e sia S un sottoinsieme di V; allora $S^\perp$ denota il seguente sottospazio di V:*

$$S^\perp := \{v \in V, v \bullet w = 0 \quad \forall w \in S\}$$

Se $v \in V$, si pone $v^\perp := \{v\}^\perp$.

Se W è un sottospazio di V, e $(w_1, \dots, w_t)$ è una base di W, si ha:

$$W^\perp = w_1^\perp \cap \ldots \cap w_t^\perp$$

Dimostrazione Provare che $S^\perp$ è un sottospazio è un facile esercizio.

Proviamo $W^\perp \subseteq w_1^\perp \cap \ldots \cap w_t^\perp$: $u \in W^\perp \Rightarrow u \bullet w_i = 0 \quad \forall i = 1, \dots, t$ (perché $w_i \in W$)$\Rightarrow u \in w_i^\perp \quad \forall i = 1, \dots, t \Rightarrow u \in \cap_{i=1}^t w_i^\perp$.

Proviamo ora $W^\perp \supseteq w_1^\perp \cap \ldots \cap w_t^\perp$: sia $u \in \cap_{i=1}^t w_i^\perp \Rightarrow \forall v \in W, v = \sum_{i=1}^t a_i w_i$, risulta: $u \bullet v = u \bullet (\sum_{i=1}^t a_i w_i) = \sum_{i=1}^t a_i (u \bullet w_i) = 0 \Rightarrow u \in W^\perp$. ∎

Definizione 4.12 Due sottospazi U e W si dicono ortogonali se $\forall u \in U, \forall w \in W,\ u \bullet w = 0$.
Si osservi che U e W sono ortogonali $\iff U \subseteq W^\perp \iff W \subseteq U^\perp$.

Definizione 4.13 Sia V uno spazio v.e., $dimV = n$; l'n-sfera unitaria è definita come l'insieme dei vettori di norma 1:

$$S^n := \{v \in V, ||v|| = 1\}$$

Teorema 4.14 (Disuguaglianza di Schwarz)
Sia V uno spazio v.e.; allora

$$|v \bullet w| \le ||v|| \cdot ||w||, \qquad \forall v, w \in V \tag{$\star$}$$

e vale l'= se e solo se v e w sono l.d.

Dimostrazione Se v, w sono l.d. vale $(\star)$ con l'$=$, infatti: se $v = 0$ si ha $|v \bullet w| = ||v|| \cdot ||w|| = 0$; se $v \neq 0$, e $w = av$, $a \in \mathbb{R}$, allora

$$|v \bullet av| = |a| \cdot ||v||^2 = ||v|| \cdot (|a| \cdot ||v||) = ||v|| \cdot ||w||$$

Caso generale: $\forall t \in \mathbb{R}$ si ha:

$$0 \leq ||tv + w||^2 = (tv + w) \bullet (tv + w) = t^2||v||^2 + 2t(v \bullet w) + ||w||^2 \qquad (\dagger)$$

Se v e w sono l.i., $||tv + w||^2 \neq 0$, e $(\dagger)$ è una disuguaglianza che deve essere verificata per ogni t, quindi il discriminante del polinomio deve essere < 0:

$$0 > \frac{1}{4}\Delta = (v \bullet w)^2 - ||v||^2 \cdot ||w||^2$$

cioé vale $(\star)$. Se vale l'$=$ in $(\star)$, allora il discriminante di $(\dagger)$ si annulla, quindi esiste una radice doppia t_0 di $(\dagger)$, cioé $||t_0 v + w||^2 = 0$, da cui $t_0 v + w = 0$, cioé v e w sono l.d.. ■

Proposizione 4.15 *Sia V uno spazio v.e., $dimV = n$. L'applicazione norma:*

$$\begin{array}{rccc} || \; || : & V & \to & \mathbb{R} \\ & v & \mapsto & ||v|| \end{array}$$

gode delle seguenti proprietà:

a) $||v|| \geq 0 \quad \forall v \in V$, $||v|| = 0 \iff v = 0$

b) $||av|| = |a| \cdot ||v|| \quad \forall a \in \mathbb{R}, \forall v \in V$

c) $||v + w|| \leq ||v|| + ||w|| \quad \forall v, w \in V$ *(***disuguaglianza triangolare***),*
con $=$ se e solo se $v = \rho w$ per un $\rho \geq 0$.

Dimostrazione *a)*: il prodotto scalare è una forma definita positiva, e $||v|| = \sqrt{v \bullet v}$.

b): $||av|| = \sqrt{av \bullet av} = \sqrt{a^2(v \bullet v)} = |a| \cdot \sqrt{v \bullet v} = |a| \cdot ||v||$

c) da 4.14 si ha $v \bullet w \leq |v \bullet w| \leq ||v|| \cdot ||w||$. Quindi:

$$||v + w||^2 = (v + w) \bullet (v + w) = v \bullet v + 2v \bullet w + w \bullet w \leq ||v||^2 + 2||v|| \cdot ||w|| + ||w||^2 = (||v|| + ||w||)^2$$

Se vale l'$=$, allora deve essere $v \bullet w = |v \bullet w| = ||v|| \cdot ||w||$, quindi da 4.14 v e w sono l.d., con $v = \rho w$ per un $\rho \geq 0$.

Esempio 4.16 In $\mathbb{R}^n$ con il prodotto scalare standard si ha

$$||(x_1, \ldots, x_n)|| = \sqrt{(x_1^2 + \ldots + x_n^2)}$$

Quindi in $\mathbb{R}$ si ha $||x|| = \sqrt{x^2} = |x|$, in $\mathbb{R}^2$ si ha $||(x, y)|| = \sqrt{x^2 + y^2}$, e in $\mathbb{R}^3$ si ha $||(x, y, z)|| = \sqrt{x^2 + y^2 + z^2}$

4.3 Esistenza di basi ortonormali

Osservazione 4.17 In questo paragrafo mostriamo che uno spazio v.e. di dimensione finita ha sempre una base ortonormale. Questo fatto è interessante perché, come ci mostra il lemma seguente, lavorare in uno spazio v.e. con $dimV = n$ e una base ortonormale equivale a lavorare in $\mathbb{R}^n$ con la base canonica $\mathcal{E}$:

Lemma 4.18 *Sia V spazio v.e., $dimV = n$, e $\mathcal{N} = (\varepsilon_1, \dots, \varepsilon_n)$ una base ortonormale per V. Allora si ha:*

a)
$$Mat(\bullet, \mathcal{N}) = I_n;$$

b)
$$(x_1, \dots, x_n)_{\mathcal{N}} \bullet (y_1, \dots, y_n)_{\mathcal{N}} = {}^tXI_nY = {}^tXY = x_1y_1 + \dots + x_ny_n$$

e in particolare
$$||(x_1, \dots, x_n)_{\mathcal{N}}|| = \sqrt{x_1^2 + \dots + x_n^2};$$

c) Se $v = (x_1, \dots, x_n)_{\mathcal{N}}$, si ha $x_i = v \bullet \varepsilon_i$, e quindi
$$v = (v \bullet \varepsilon_1)\varepsilon_1 + \dots + (v \bullet \varepsilon_n)\varepsilon_n$$

Dimostrazione Per *a*) e *b*) basta usare la definizione di base ortonormale e 4.4; per *c*), basta osservare che $v \bullet \varepsilon_i = (x_1\varepsilon_1 + \dots + x_n\varepsilon_n) \bullet \varepsilon_i = x_i$.

Il teorema seguente permette di costruire, a partire da una base qualsiasi, una base ortonormale:

Teorema 4.19 (Procedimento di ortonormalizzazione di Gram -Schmidt)
Siano V uno spazio v.e., $dimV = n$, $\mathcal{B} = (v_1, \dots, v_n)$ una base di V. Poniamo :
$$w_1 := v_1$$
$$w_{t+1} := v_{t+1} - \frac{v_{t+1} \bullet w_1}{w_1 \bullet w_1} w_1 - \dots - \frac{v_{t+1} \bullet w_t}{w_t \bullet w_t} w_t \qquad 1 \leq t \leq n-1$$

Allora: la definizione è ben posta perché $w_i \bullet w_i \neq 0 \quad \forall i$, e si ha:

a)
$$< v_1, \dots, v_k > = < w_1, \dots, w_k > \qquad \forall k \leq n;$$

b)
$$\left(\frac{w_1}{||w_1||}, \dots, \frac{w_n}{||w_n||} \right)$$

è una base ortonormale per V.

Dimostrazione Si parte costruendo una base ortogonale $(w_1, \ldots, w_n)$, e poi si normalizzano i vettori.

Primo passo: proviamo contemporaneamente che i vettori w_k sono ben definiti (w_k è ben definito se $w_1 \bullet w_1 \neq 0, \ldots, w_{k-1} \bullet w_{k-1} \neq 0$), e che vale a). Per induzione:

w_2 è ben definito perché $w_1 = v_1 \neq 0$; inoltre $< w_1 >=< v_1 >$.

supponiamo w_k ben definito, e $< w_1, \ldots, w_{k-1} >=< v_1, \ldots, v_{k-1} >$. Allora

$$< v_1, \ldots, v_k >\subseteq< w_1, \ldots, w_k >$$

perché v_k è combinazione lineare di $w_1, \ldots, w_k$;

$$< w_1, \ldots, w_k > \underbrace{\subseteq}_{w_k \; c.l. \; di \; v_k, w_1, \ldots,, w_{k-1}} < v_1, \ldots, v_{k-1}, v_k, w_1, \ldots, , w_{k-1} > \underbrace{=}_{\text{ipotesi induttiva}} < v_1, \ldots, v_k >$$

cioé $< v_1, \ldots, v_k >=< w_1, \ldots, w_k >$; essendo $dim < v_1, \ldots, v_k >= k$, ne segue che $w_1, \ldots, w_k$ sono l.i., da cui $w_k \neq 0 \Rightarrow w_{k+1}$ è ben definito.

Secondo passo: proviamo b) per induzione su k:

Si ha: $w_1 \bullet w_2 = w_1 \bullet v_2 - \frac{v_2 \bullet w_1}{w_1 \bullet w_1} w_1 \bullet w_1 = 0$.

Supponiamo che valga $w_i \bullet w_j = 0$ per $1 \leq i < j \leq k$, e proviamo che vale $w_j \bullet w_{k+1} = 0$ per $1 \leq j \leq k$. Si ha:

$$w_j \bullet w_{k+1} = w_j \bullet (v_{k+1} - \sum_{i=1}^{k} \frac{v_{k+1} \bullet w_i}{w_i \bullet w_i} w_i) = w_j \bullet v_{k+1} - \frac{v_{k+1} \bullet w_j}{w_j \bullet w_j} w_j \bullet w_j = 0$$

■

Esempio 4.20 Si consideri lo spazio v.e. $\mathbb{R}^3$ con il prodotto scalare standard e la base

$$\mathcal{B} = (\underbrace{(1,0,1)}_{v_1}, \underbrace{(0,1,-3)}_{v_2}, \underbrace{(0,0,1)}_{v_3})$$

Costruiamo una base ortonormale per $\mathbb{R}^3$ applicando Gram-Schmidt alla base $\mathcal{B}$:

$$w_1 = v_1 = (1,0,1)$$

$$w_2 = (0,1,-3) - \frac{(0,1,-3) \bullet (1,0,1)}{(1,0,1) \bullet (1,0,1)}(1,0,1) = (0,1,-3) - \frac{-3}{2}(1,0,1) = (0+\frac{3}{2}, 1, -3+\frac{3}{2}) = (\frac{3}{2}, 1, -\frac{3}{2})$$

$$w_3 = (0,0,1) - \frac{(0,0,1) \bullet (1,0,1)}{(1,0,1) \bullet (1,0,1)}(1,0,1) - \frac{(0,0,1) \bullet (\frac{3}{2}, 1, \frac{-3}{2})}{(\frac{3}{2}, 1, \frac{-3}{2}) \bullet (\frac{3}{2}, 1, \frac{-3}{2})}(\frac{3}{2}, 1, \frac{-3}{2}) =$$

$$= (0,0,1) - \frac{1}{2}(1,0,1) - \frac{\frac{-3}{2}}{\frac{9}{4}+1+\frac{9}{4}}(\frac{3}{2}, 1, \frac{-3}{2}) = (-\frac{1}{2}, 0, \frac{1}{2}) + \frac{3}{2} \cdot \frac{4}{22}(\frac{3}{2}, 1, \frac{-3}{2}) =$$

$$= (-\frac{1}{2}, 0, \frac{1}{2}) + \frac{3}{11}(\frac{3}{2}, 1, \frac{-3}{2}) = (-\frac{1}{2}, 0, \frac{1}{2}) + (\frac{9}{22}, \frac{6}{22}, \frac{-9}{22}) = (-\frac{1}{11}, \frac{3}{11}, \frac{1}{11})$$

Verifichiamo che w_1, w_2, w_3 sono a due a due ortogonali:

$$w_1 \bullet w_2 = (1,0,1) \bullet (\frac{3}{2}, 1, -\frac{3}{2}) = 0$$

$$w_1 \bullet w_3 = (1,0,1) \bullet (-\frac{1}{11}, \frac{3}{11}, \frac{1}{11}) = 0$$

$$w_2 \bullet w_3 = (\frac{3}{2}, 1, -\frac{3}{2}) \bullet (-\frac{1}{11}, \frac{3}{11}, \frac{1}{11}) = 0$$

Adesso basta normalizzare:

$$||w_1|| = \sqrt{w_1 \bullet w_1} = \sqrt{(1,0,1) \bullet (1,0,1)} = \sqrt{2}$$

$$||w_2|| = \sqrt{w_2 \bullet w_2} = \sqrt{(\frac{3}{2}, 1, \frac{-3}{2}) \bullet (\frac{3}{2}, 1, \frac{-3}{2})} = \sqrt{\frac{11}{2}}$$

$$||w_3|| = \sqrt{w_3 \bullet w_3} = \sqrt{(-\frac{1}{11}, \frac{3}{11}, \frac{1}{11}) \bullet (-\frac{1}{11}, \frac{3}{11}, \frac{1}{11})} = \sqrt{\frac{1}{11^2} + \frac{9}{11^2} + \frac{1}{11^2}} = \sqrt{\frac{11}{11^2}} = \frac{1}{\sqrt{11}}$$

Quindi la base ortonormale cercata è:

$$(\frac{1}{\sqrt{2}}(1,0,1),\ \sqrt{\frac{2}{11}}(\frac{3}{2}, 1, -\frac{3}{2}),\ \sqrt{11}(-\frac{1}{11}, \frac{3}{11}, \frac{1}{11}))$$

Proposizione-Definizione 4.21 *Sia V spazio v.e., $dimV = n$, e W un sottospazio. Si ha:*

a)

$$V = W \oplus W^{\perp}$$

Quindi un vettore $v \in V$ si scrive in modo unico come $\quad v = w + w'$, $w \in W$, $w' \in W^{\perp}$*;*
w, risp. w', è detto la proiezione ortogonale di v su W, risp. su $W^{\perp}$.

b) Sia $\mathcal{N} = (\varepsilon_1, \ldots, \varepsilon_t)$ una base ortonormale per W, e sia $v \in V$; allora

$$w = (v \bullet \varepsilon_1)\varepsilon_1 + \ldots + (v \bullet \varepsilon_t)\varepsilon_t$$

c) Vale la relazione:

$$||v||^2 = ||w||^2 + ||w'||^2$$

detta identità pitagorica.

Dimostrazione Sia $\mathcal{N} = (\varepsilon_1, \ldots, \varepsilon_t)$ una base ortonormale per W (per il Teorema 4.19 sappiamo che ne esiste almeno una); per 4.11 $W^{\perp} = \varepsilon_1^{\perp} \cap \ldots \cap \varepsilon_t^{\perp}$.

$a), b)$: Si ha $W \cap W^{\perp} = 0$ perché $u \bullet u = 0 \Rightarrow u = 0$. Ora basta provare che se $v \in V$, e w definito come in $b)$, allora $w' := v - w \in W^{\perp}$: infatti se questo è vero, allora $v = w + w'$ si scrive come somma di un vettore di W e uno di $W^{\perp}$.

Si ha

$$v - w \in W^{\perp} \iff v - w \in \varepsilon_1^{\perp} \cap \ldots \cap \varepsilon_t^{\perp} \iff (v - w) \bullet \varepsilon_j = 0 \ \forall j = 1, \ldots, t$$

e in effetti:

$$(v - w) \bullet \varepsilon_j = (v - \sum_{i=1}^{t} (v \bullet \varepsilon_i)\varepsilon_i) \bullet \varepsilon_j = v \bullet \varepsilon_j - v \bullet \varepsilon_j = 0$$

c) Da $v = w + w'$ segue

$$||v||^2 = ||w + w'||^2 = (w + w') \bullet (w + w') = ||w||^2 + ||w'||^2 + 2 \underbrace{w \bullet w'}_{=0}$$

■

Osservazione 4.22 Sia $u \neq 0$; 4.21 ci dice che la proiezione ortogonale w di un vettore v sulla retta $< u >$ è:

$$w = \left(v \bullet \frac{u}{||u||} \right) \frac{u}{||u||} = \frac{v \bullet u}{u \bullet u} u$$

e la proiezione ortogonale w' di v sull'iperpiano $< u >^{\perp}$ è:

$$w' = v - \frac{v \bullet u}{u \bullet u} u$$

Quindi quello che succede nella dimostrazione di Gram -Schmidt è questo: cominciamo a porre $w_1 := v_1$, quindi

$$w_2 = v_2 - \frac{v_2 \bullet w_1}{w_1 \bullet w_1} w_1$$

è la proiezione ortogonale di v_2 su $< w_1 >^{\perp} = < v_1 >^{\perp}$; ora w_1 e w_2 sono ortogonali, quindi $(\frac{w_1}{||w_1||}, \frac{w_2}{||w_2||})$ è una base ortonormale per $< w_1, w_2 >$, e quindi 4.21 ci dice che la proiezione ortogonale di v_3 su $< w_1, w_2 >^{\perp} = < v_1, v_2 >^{\perp}$ è:

$$w_3 = v_3 - \frac{v_3 \bullet w_1}{w_1 \bullet w_1} w_1 - \frac{v_3 \bullet w_2}{w_2 \bullet w_2} w_2$$

e così via.

4.4 Endomorfismi unitari

Proposizione-Definizione 4.23 Una matrice $A \in M_n(\mathbb{R})$ si dice *ortogonale* se è invertibile e se

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

Si osservi che $A^{-1} = {}^tA \iff {}^tAA = I_n \iff A\,{}^tA = I_n$ (vedi Lemma 4.25).
Poniamo

$$O(n) := \{A \in GL_n(\mathbb{R}), \quad A \text{ ortogonale}\}$$

Si ha che $O(n)$ con il prodotto righe per colonne è un sottogruppo di $GL_n(\mathbb{R})$, detto *il gruppo ortogonale di ordine n.*

Dimostrazione Si ha $O(n) \neq \emptyset$ perché $I_n \in O(n)$. Se $A, B \in O(n)$, allora $(AB)^{-1} = B^{-1}A^{-1} = {}^tB\,{}^tA = {}^t(AB)$. Se $A \in O(n)$, allora $(A^{-1})^{-1} = ({}^tA)^{-1} = {}^t(A^{-1})$, quindi $A^{-1} \in O(n)$. ∎

Notazione 4.24 Sia A una matrice quadrata di ordine n; indicheremo con A^i la sua riga i-esima e con A_i la sua colonna i-esima.

Lemma 4.25 *Se $A, B \in M_n(K)$, si ha*

$$AB = I \iff BA = I$$

Dimostrazione Siano $f, g \in End(K^n)$, $A = M_{\mathcal{E}}(f), B = M_{\mathcal{E}}(g)$, e supponiamo ad esempio $AB = I$. Allora:

$$AB = I \Rightarrow f \circ g = id \Rightarrow \forall v \in V, f(g(v)) = v \Rightarrow f \ su$$

ed essendo f un endomorfismo di uno spazio vettoriale di dimensione finita, se è su è anche $1-1$, quindi esiste f^{-1} che necessariamente coincide con l'inversa destra di f, cioé g; allora $B = M_{\mathcal{E}}(f^{-1})$ e $f^{-1} \circ f = id \Rightarrow BA = I$. ∎

Lemma 4.26

$$A \in O(n) \iff {}^t(A_i)A_j = \delta_{ij} \iff A^i\,{}^t(A^j) = \delta_{ij}$$

Dimostrazione Dimostriamo la prima equivalenza:

$$\underbrace{{}^t(A_i)}_{=({}^tA)^i} A_j = \delta_{ij} \iff {}^tAA = I_n$$

Analogamente per la seconda:

$$A^i \underbrace{{}^t(A^j)}_{=({}^tA)_j} = \delta_{ij} \iff A\,{}^tA = I_n$$

∎

Proposizione 4.27 *Siano V uno spazio v.e., $dimV = n$, $\mathcal{B} = (v_1, \dots, v_n)$ una base e $\mathcal{N} = (\varepsilon_1, \dots, \varepsilon_n)$ una base ortonormale di V; allora $\mathcal{B}$ è ortonormale $\iff M_{\mathcal{N},\mathcal{B}}(id_V) \in O(n)$.*

Dimostrazione Sia $A = (a_{ij}) = M_{\mathcal{N},\mathcal{B}}(id_V)$; la colonna A_i è data dalle coordinate di v_i rispetto ad $\mathcal{N}$, che è ortonormale, quindi da 4.18

$$v_i \bullet v_j = {}^tA_iA_j$$

Ne segue

$$v_i \bullet v_j = \delta_{ij} \iff {}^tA_iA_j = \delta_{ij} \underbrace{\iff}_{lemma\ 4.26} A \in O(n)$$

■

Osservazione 4.28 $A \in O(n) \Rightarrow detA = \pm 1$, infatti

$$detA = det({}^tA) = det(A^{-1}) = (detA)^{-1} \Rightarrow (\det A)^2 = 1$$

Proposizione-Definizione 4.29 Poniamo

$$SO(n) := \{A \in O(n), \quad detA = 1\}$$

Si ha che $SO(n)$ con il prodotto righe per colonne è un sottogruppo di $O(n)$(e quindi di $GL_n(\mathbb{R})$), detto *il gruppo ortogonale speciale di ordine n*.

Proposizione-Definizione 4.30 *Siano V uno spazio v.e., $dimV = n$, $f \in EndV$; f è detto* unitario *se soddisfa una delle seguenti condizioni equivalenti:*

a) $f(v) \bullet f(w) = v \bullet w \ \forall v, w \in V$ (cioé f preserva il prodotto scalare);

b) $||f(v)|| = ||v|| \ \forall v \in V$ (cioé f preserva la norma);

c) se $\mathcal{N} = (\varepsilon_1, \dots, \varepsilon_n)$ è una base ortonormale allora $(f(\varepsilon_1), \dots, f(\varepsilon_n))$ è una base ortonormale;

d) esiste una base ortonormale $\mathcal{N} = (\varepsilon_1, \dots, \varepsilon_n)$ tale che $(f(\varepsilon_1), \dots, f(\varepsilon_n))$ sia una base ortonormale;

e) se $\mathcal{N} = (\varepsilon_1, \dots, \varepsilon_n)$ è una base ortonormale allora $M_{\mathcal{N}}(f) \in O(n)$.

In particolare da tutto ciò segue che un endomorfismo unitario è un automorfismo; inoltre da e) e da 4.28 segue che se f è unitario allora $detf = \pm 1$.

Dimostrazione $a) \Rightarrow b)$ è ovvio;

$b) \Rightarrow a)$: $\forall v.w \in V$,

$$2v \bullet w + ||v||^2 + ||w||^2 = ||v + w||^2 = ||f(v + w)||^2 = ||f(v) + f(w)||^2 =$$
$$= 2f(v) \bullet f(w) + ||f(v)||^2 + ||f(w)||^2 \Rightarrow f(v) \bullet f(w) = v \bullet w$$

$a) \Rightarrow c) \Rightarrow d)$ sono ovvie;

$d) \Rightarrow a)$: siano $v = \sum a_i\varepsilon_i, w = \sum b_i\varepsilon_i$; allora

$$v \bullet w = \sum a_ib_i = (\sum a_if(\varepsilon_i)) \bullet (\sum b_jf(\varepsilon_j)) = f(v) \bullet f(w)$$

$a) \iff e)$: sia $\mathcal{N} = (\varepsilon_1, \dots, \varepsilon_n)$ una base ortonormale e $A := M_{\mathcal{N}}(f)$. Allora per quanto appena dimostrato $a) \iff (f(\varepsilon_1), \dots, f(\varepsilon_n))$ base ortonormale $\iff \delta_{ij} = f(\varepsilon_i) \bullet f(\varepsilon_j) = {}^tA_iA_j \iff A \in O(n)$. ■

Proposizione-Definizione 4.31 Sia V uno spazio v.e., $dimV = n$.
Poniamo
$$O(V) := \{f \in GL(V), \quad f \text{ unitario}\}$$
Si ha che $O(V)$ con la composizione è un sottogruppo di $GL(V)$, detto *il gruppo ortogonale di* V.
Poniamo inoltre
$$SO(V) := \{f \in O(V), \quad detf = 1\}$$
Si ha che $SO(V)$ con la composizione è un sottogruppo di $O(V)$(e quindi di $GL(V)$), detto *il gruppo ortogonale speciale di* V; gli elementi di $SO(V)$ sono detti *rotazioni di* V.

Proposizione 4.32 *Sia V uno spazio v.e., $dimV = n$, e sia $\mathcal{N} = (\varepsilon_1, \dots, \varepsilon_n)$ una base ortonormale. L'isomorfismo di gruppi*
$$\begin{array}{cccc} M_{\mathcal{N}}: & GL(V) & \stackrel{\cong}{\rightarrow} & GL_n(\mathbb{R}) \\ & f & \mapsto & M_{\mathcal{N}}(f) \end{array}$$
induce isomorfismi tra i sottogruppi:
$$\begin{array}{ccc} O(V) & \stackrel{\cong}{\rightarrow} & O(n) \\ \cup & & \cup \\ SO(V) & \stackrel{\cong}{\rightarrow} & SO(n) \end{array}$$

Dimostrazione Immediata.

Proposizione 4.33 *Sia V uno spazio v.e., $dimV = n$, e sia $f \in O(V)$; se λ è autovalore per f allora $\lambda = \pm 1$.*
Equivalentemente, se $A \in O(n)$ e λ è autovalore per A allora $\lambda = \pm 1$.

Dimostrazione Sia $v \neq 0$, v autovettore per f relativo a λ; allora
$$||v|| = ||f(v)|| = ||\lambda v|| = |\lambda| \cdot ||v|| \Rightarrow |\lambda| = 1$$
Se $A \in O(n)$, basta prendere una base $\mathcal{N}$ ortonormale di V e considerare $f \in EndV$, con $M_{\mathcal{N}}(f) = A$; per 4.30, $f \in O(V)$. ■

4.5 Il teorema spettrale

Definizione 4.34 Sia V uno spazio v.e., $dimV = n$, e sia $f \in EndV$; f è detto *simmetrico* o *autoaggiunto* se

$$f(v) \bullet w = v \bullet f(w) \qquad \forall v, w \in V$$

Proposizione 4.35 *Sia V uno spazio v.e., $dimV = n$, e sia $\mathcal{N} = (\varepsilon_1, \dots, \varepsilon_n)$ una base ortonormale. Allora $f \in EndV$ è simmetrico $\iff$ $M_{\mathcal{N}}(f)$ è una matrice simmetrica.*

Dimostrazione Sia $A = M_{\mathcal{N}}(f)$, $v = (x_1, \dots, x_n)_{\mathcal{N}}$ e $w = (y_1, \dots, y_n)_{\mathcal{N}}$; ponendo $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ risulta:

$$v \bullet f(w) = {}^tX(AY), \qquad f(v) \bullet w = {}^t(AX)Y = {}^tX{}^tAY$$

quindi f è simmetrico $\iff$ ${}^tXAY = {}^tX{}^tAY \quad \forall X, Y \iff A = {}^tA$. ■

Notazione 4.36 Se $B = (b_{ij}) \in M_{n,s}(\mathbb{C})$, nel seguito indichiamo la matrice complessa coniugata di B con

$$\bar{B} := (\bar{b}_{ij})$$

Lemma 4.37 *Sia $A \in M_n(\mathbb{R})$, A simmetrica; allora p_A ammette solo radici reali, e ne ammette almeno una.*

Dimostrazione Pensiamo A come matrice in $M_n(\mathbb{C})$; in $\mathbb{C}$ certamente esiste almeno una radice λ di p_A.

L'endomorfismo $f = F^{\mathcal{E}}(A)$ (qui il vettore $(x_1, \dots, x_n)$ è identificato con la matrice colonna $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$):

$$\begin{array}{rccc} f: & \mathbb{C}^n & \to & \mathbb{C}^n \\ & X & \mapsto & AX \end{array}$$

ha quindi λ come autovalore; sia $(z_1, \dots, z_n)$ un autovettore relativo a λ, e poniamo $Z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$.

Si osservi che essendo $(z_1, \dots, z_n) \neq 0$ si ha

$${}^tZ\bar{Z} = z_1\bar{z}_1 + \ldots + z_n\bar{z}_n = |z_1|^2 + \ldots + |z_n|^2 > 0 \qquad (*)$$

Inoltre si ha $AZ = \lambda Z$, da cui, tenendo conto che A è reale e quindi $\bar{A} = A$, segue:

$$A\bar{Z} = \bar{\lambda}\bar{Z}$$

da cui, moltiplicando a sinistra per ${}^tZ$, si ha

$${}^tZA\bar{Z} = {}^tZ(\bar{\lambda}\bar{Z}) = \bar{\lambda}({}^tZ\bar{Z})$$

ed essendo A simmetrica:

$$ {}^tZA\bar{Z} \underbrace{=}_{A={}^tA} {}^tZ{}^tA\bar{Z} = {}^t(AZ)\bar{Z} = {}^t(\lambda Z)\bar{Z} = \lambda({}^tZ\bar{Z}) $$

quindi

$$ \bar{\lambda}({}^tZ\bar{Z}) = \lambda({}^tZ\bar{Z}) \underbrace{\Rightarrow}_{da\ (*)} \lambda = \bar{\lambda} $$

■

Teorema 4.38 (Teorema Spettrale) *Sia V uno spazio v.e., $dimV = n \geq 1$, e sia $f \in EndV$ simmetrico; allora esiste una base ortonormale $\mathcal{L}$ di autovettori per f.*

Dimostrazione Per induzione su n. Se $n = 1$ l'enunciato è banalmente vero. Sia $n \geq 2$ e supponiamo che l'enunciato valga in dimensione $n-1$. Sia $\mathcal{N}$ una base ortonormale per V; per 4.35 $M_{\mathcal{N}}(f)$ è simmetrica e per 4.37 ha almeno un autovalore reale λ.

Sia v un autovettore associato a λ; quindi $v_1 := \frac{v}{||v||}$ è un autovettore di norma 1. Sia

$$ U := v_1^{\perp} $$

allora $dimU = n-1$ e U è f-invariante, infatti:

$$ \forall u \in U, \qquad f(u) \bullet v_1 = u \bullet f(v_1) = u \bullet \lambda v_1 = 0 \Rightarrow f(u) \in U $$

Sia $g := f_{|U}^{|U}$ l'endomorfismo indotto da f; g è simmetrico, e per ipotesi induttiva ha una base ortonormale di autovettori $(v_2, \ldots, v_n)$.

Quindi $\mathcal{L} = (v_1, v_2, \ldots, v_n)$ è una base ortonormale di autovettori per f.

Teorema 4.39 (Teorema Spettrale - seconda formulazione) *Sia $A \in M_n(\mathbb{R})$, A simmetrica; allora esiste una matrice $M \in O(n)$ tale che*

$$ M^{-1}AM = {}^tMAM $$

sia diagonale.

Dimostrazione Sia $V = \mathbb{R}^n$ con il prodotto scalare standard, e $f \in EndV$ tale che $A = M_{\mathcal{E}}(f)$; sia $\mathcal{L}$ una base ortonormale di autovettori per f, che esiste per 4.38. Allora la matrice $M_{\mathcal{L}}(f)$ è diagonale, e ponendo $M := M_{\mathcal{E},\mathcal{L}}(id_V)$ si ha

$$ M_{\mathcal{L}}(f) = M_{\mathcal{L},\mathcal{E}}(id_V)AM_{\mathcal{E},\mathcal{L}}(id_V) = M^{-1}AM $$

Da 4.27 inoltre $M_{\mathcal{E},\mathcal{L}}(id_V) \in O(n) \Rightarrow M^{-1} = {}^tM$. ■

Corollario 4.40 *Gli autospazi di un endomorfismo simmetrico sono ortogonali tra loro.*

Dimostrazione Per il Teorema Spettrale esiste una base ortonormale $\mathcal{L} = (v_1, \ldots, v_n)$ di autovettori per f; se V_λ e V_μ sono due autospazi distinti, possiamo supporre, riordinando i vettori della base, che sia ad esempio $V_\lambda =< v_1, \ldots, v_h >$ e $V_\mu =< v_{h+1}, \ldots, v_t >$. Se $v \in V_\lambda$ e $w \in V_\mu$, esistono $a_1, \ldots, a_h, b_{h+1}, \ldots, b_t \in \mathbb{R}$ tali che $v = \sum_{i=1}^{h} a_i v_i, \quad w = \sum_{j=h+1}^{t} b_j v_j$, e quindi, dato che $v_i \bullet v_j = \delta_{ij}$, si ha:

$$v \bullet w = (\sum_{i=1}^{h} a_i v_i) \bullet (\sum_{j=h+1}^{t} b_j v_j) = \sum_{i,j}^{h} a_i b_j v_i \bullet v_j = 0 \Rightarrow V_\lambda \subseteq V_\mu^\perp$$

■

5 Spazi affini

5.1 La struttura di spazio affine su uno spazio vettoriale

Nel primo modulo di questo corso avete parlato di rette di $\mathbb{R}^2$ e rette e piani di $\mathbb{R}^3$, non necessariamente passanti per l'origine O dei vettori, cioé avete iniziato a studiare la struttura affine del piano e dello spazio 3-dimensionale reale.
Consideriamo ad esempio il piano reale: per fare questo avete usato gli elementi di $\mathbb{R}^2$ chiamandoli sia "punti" che "vettori" e osservando che la differenza di due punti P, Q dà un vettore, $\vec{PQ} = Q - P$ che è un vettore applicato in P e con punto finale in Q.

Bisogna inoltre osservare che quanto avete visto nel primo modulo di questo corso fa uso della geometria euclidea del piano e dello spazio 3-dimensionale studiata a scuola, che è fondata sugli assiomi di Euclide; siete così riusciti a tradurre problemi geometrici in problemi algebrici, facendo della geometria analitica, usando cioé coordinate ed equazioni anziché enti geometrici astratti.

In questo secondo modulo di Geometria 1 introduciamo la struttura affine su un K-spazio vettoriale finitamente generato: questo ci permette tre cose:

- studiare la geometria affine ed euclidea in dimensione n;

- rivisitare quanto visto nel primo modulo ma senza far uso della geometria fondata sugli assiomi euclidei, tutto quello che ci serve sono gli spazi vettoriali, che ci permettono di dimostrare, facendo uso solo dell'algebra, quanto già visto: un esempio è 5.9; se facciamo geometria euclidea partendo dagli assiomi di Euclide, dobbiamo assumere come assioma il quinto postulato (per un punto passa una ed una sola parallela ad una retta data), se invece facciamo geometria partendo dagli assiomi della teoria degli insiemi, e cioé usando gli spazi vettoriali, il quinto postulato diventa una proposizione dimostrabile (vedi 5.9)

- fare geometria affine non necessariamente sul campo reale, ma più in generale su un campo K, che può ad esempio essere il campo dei razionali, o dei complessi, o un campo finito...

Qui sotto trovate una definizione molto astratta di spazio affine: la definizione generale viene data solo per mettere bene in evidenza che una cosa sono i punti ed un'altra i vettori, e che la struttura affine dipende appunto dal fatto che due punti individuano un vettore, ma nel seguito consideriamo un caso particolare di spazio affine, cioé la struttura affine di un K-spazio vettoriale V finitamente generato (in realtà però si può dimostrare che ci si può sempre ricondurre a questo caso!)

Definizione 5.1 Sia V un K-spazio vettoriale; *uno spazio affine su* V è una coppia (A, ϕ) dove A è un insieme non vuoto, i cui elementi sono detti punti di A, e ϕ è una applicazione:

$$\begin{array}{rccc} \phi: & A \times A & \rightarrow & V \\ & (P, Q) & \mapsto & \phi(P, Q) \end{array}$$

che verifichi le seguenti condizioni:

(1) $\forall P \in A, \forall v \in V, \exists! Q \in A$ tale che $\phi(P, Q) = v$

(2) $\forall P, Q, R \in A$ si ha $\phi(P, Q) + \phi(Q, R) = \phi(P, R)$

Osservazione 5.2 Nella definizione precedente, (1) dice che fissata una origine P, ritrovo V come l'insieme dei $\phi(P, Q)$; cioé ad ogni punto P "attacco" lo spazio vettoriale V.

Proposizione-Definizione 5.3 Sia V un K-spazio vettoriale, $dimV = n$; l'applicazione:

$$\begin{array}{rccc} \phi: & V \times V & \rightarrow & V \\ & (a,b) & \mapsto & b-a \end{array}$$

rende V uno spazio affine su se stesso, che verrà denotato nel seguito con V_a. Si pone:

$$dimV_a := dimV$$

Se $V = K^n$, V_a si indica con $\mathbb{A}^n(K)$ o solo con $\mathbb{A}^n$ se è chiaro su quale campo stiamo lavorando.

Dimostrazione Le condizioni (1) e (2) valgono, infatti:

(1): $\forall a \in V, \forall v \in V, \exists! b \in V$ tale che $\phi(a,b) = v$; basta prendere $b = a + v$

(2): $\forall p, q, r \in V$ si ha $(q-p) + (r-q) = r - p$.

Notazione 5.4 Noi studiamo solo spazi affini di tipo V_a; chiamiamo gli elementi di V_a, di solito indicati con lettere maiuscole, punti, mentre continuiamo a chiamare vettori gli elementi di V.

Nel seguito quindi V è un K-spazio vettoriale fissato di $dimV = n > 0$.

Definizione 5.5 Sia W un sottospazio vettoriale di V e sia Q un punto di V_a. *Il sottospazio affine di V_a passante per il punto Q e di giacitura W, o parallelo a W*, si definisce come:

$$S := \{P \in V_a \mid P - Q \in W\} = \{P \in V_a \mid P = Q + w,\ w \in W\}$$

Scriveremo quindi

$$S = Q + W$$

Si osservi che per definizione un sottospazione affine non è mai vuoto: se $S = Q + W$, $S \neq \emptyset$ perché $0 \in W \Rightarrow Q \in S$.

Il sottospazio affine di V_a passante per il punto Q e di giacitura W si può pensare come il laterale del gruppo $(V, +)$ rispetto al sottogruppo W individuato da Q. .

Proposizione-Definizione 5.6 Siano W un sottospazio vettoriale di V, Q un punto di V_a ed $S = Q + W$. Allora:

a) $\forall Q' \in S$ risulta $S = Q' + W$;

b) W è l'unico sottospazio vettoriale di V tale che la giacitura di S sia W; possiamo quindi scrivere

$$W = giacS$$

In altre parole, un sottospazio affine è individuato dalla sua giacitura e da un suo qualunque punto. Si pone:

$$dimS := dim(giac(S))$$

Dimostrazione a): Consideriamo lo spazio vettoriale quoziente V/W; dire che $S = Q+W$ equivale a dire che $S = [Q]$ in V/W; quindi $S = [Q']$ per ogni $Q' \in S$.

b) La giacitura di S è unica, infatti se $S = Q + W = Q' + U$ per quanto appena visto sulla scelta del punto si ha $S = Q + W = Q + U$; allora

$$w \in W \Rightarrow Q + w \in S = Q + U \Rightarrow \exists u \in U, Q + w = Q + u \Rightarrow w \in U$$

e viceversa, cioé $W = U$.

Definizione 5.7 Un sottospazio affine di dimensione 1 è detto *retta*;

un sottospazio affine di dimensione 2 è detto *piano*;

un sottospazio affine di dimensione $n - 1$ è detto *iperpiano*.

Se S è un sottospazio affine di V_a, diremo che un vettore $v \in V$ è un vettore parallelo ad S, e scriveremo $v//S$, se $v \in giac(S)$.
La giacitura di una retta R si chiama anche la *direzione di* R, e un suo vettore $\neq 0$ si chiama vettore di direzione di R.

Definizione 5.8 Siano S, T sottospazi affini di V_a, con $0 < dimS \leq dimT$; S e T si dicono

paralleli , e scriviamo $S//T$, se $giacS \subseteq giacT$;

sghembi se non sono paralleli e se $S \cap T = \emptyset$;

incidenti se non sono paralleli e se $S \cap T \neq \emptyset$.

Proposizione 5.9 (Quinto Postulato) *Sia P un punto di V_a e sia S un suo sottospazio affine di dim > 0. Allora esiste ed è unico il sottospazio affine per P e $//S$, di dimensione $= dimS$.*

Dimostrazione $P + giacS$ soddisfa i requisiti richiesti, ed è l'unico a farlo.

Definizione 5.10 *Un sistema di coordinate affini, o riferimento affine*, nello spazio affine V_a è dato da un punto $A \in V_a$ e da una base $\mathcal{B} = (v_1, \ldots, v_n)$ per V; lo denoteremo con $\mathcal{R} = A\mathcal{B}$.

Sia $\mathcal{R} = A\mathcal{B}$ un riferimento fissato; se $P \in V_a$, e se $P - A = (x_1, \ldots, x_n)_{\mathcal{B}}$, diciamo che $(x_1, \ldots, x_n)$ sono le coordinate affini di P rispetto ad $\mathcal{R}$, e scriviamo

$$P = (x_1, \ldots, x_n)_{\mathcal{R}}$$

Risulta $A = (0, \ldots, 0)_{\mathcal{R}}$; A è detto *origine del sistema di riferimento* $\mathcal{R}$.

Si chiama *riferimento affine standard* su $\mathbb{A}^n(K)$ il sistema di riferimento affine $\mathcal{S} = 0\mathcal{E}$ dove 0 è il vettore nullo di K^n ed $\mathcal{E}$ è la base canonica; se $P = (x_1, \ldots, x_n) \in \mathbb{A}^n$, si ha $P = (x_1, \ldots, x_n)_{\mathcal{S}}$.

Più in generale, se l'origine del riferimento $\mathcal{R}$ coincide con il vettore nullo 0 di V, si ha

$$P = (x_1, \ldots, x_n)_{\mathcal{R}} = (x_1, \ldots, x_n)_{\mathcal{B}}$$

Osserviamo che, se $\mathcal{R} = A\mathcal{B}$ è un riferimento affine su V_a, e $P = (x_1, \ldots, x_n)_{\mathcal{R}}$, $Q = (y_1, \ldots, y_n)_{\mathcal{R}}$, il vettore $Q - P$ ha come coordinate rispetto a $\mathcal{B}$ la differenza delle coordinate:

$$Q - P = (Q - A) - (P - A) = (y_1, \ldots, y_n)_{\mathcal{B}} - (x_1, \ldots, x_n)_{\mathcal{B}} = (y_1 - x_1, \ldots, y_n - x_n)_{\mathcal{B}}$$

Osservazione 5.11 Fissato un riferimento affine su V_a, è possibile dare equazioni parametriche e cartesiane per un sottospazio affine rispetto a quel riferimento, in analogia con quanto si fa per i sottospazi vettoriali rispetto ad una base: si veda l'esercizio 1 del foglio di esercizi 6.

Proposizione 5.12 *Sia $\{S_i\}_{i\in I}$ una famiglia di sottospazi affini di V_a tali che $\cap_{i\in I} S_i \neq \emptyset$. Allora $\cap_{i\in I} S_i$ è un sottospazio affine, e se $Q \in \cap_{i\in I} S_i$, risulta*

$$\bigcap_{i\in I} S_i = Q + \bigcap_{i\in I} giac(S_i)$$

Dimostrazione Sia $Q \in \cap_{i\in I} S_i$ e $W_i := giac(S_i)$; allora $S_i = Q + W_i \ \forall i \in I$. Poniamo

$$S := Q + \bigcap_{i\in I} W_i$$

Allora S è un sottospazio affine perché l'intersezione di sottospazi vettoriali è un sottospazio vettoriale, e si ha

$$P \in S \iff P - Q \in \bigcap_{i\in I} W_i \iff P - Q \in W_i \ \forall i \in I \iff P \in S_i \ \forall i \in I \iff P \in \bigcap_{i\in I} S_i$$

Proposizione-Definizione 5.13 Se $P_0, \dots, P_s$ sono punti di V_a, il sottospazio affine generato da $\{P_0, \dots, P_s\}$, indicato con $\overline{P_0 \dots P_s}$, è il più piccolo sottospazio affine contenente $P_0, \dots P_s$, o equivalentemente l'intersezione di tutti i sottospazi affini contenenti $P_0, \dots P_s$.

Risulta:

$$\overline{P_0 \dots P_s} = P_0 + < P_1 - P_0, \dots, P_s - P_0 >$$

Si osservi che per definizione $\overline{P_0 \dots P_s}$ non dipende dall'ordine dei punti, quindi anziché P_0 si può scegliere uno qualsiasi dei punti.

Dimostrazione Poniamo $S := P_0 + < P_1 - P_0, \dots, P_s - P_0 >$; allora S è un sottospazio affine che contiene $P_0 \dots P_s$, quindi $\overline{P_0 \dots P_s} \subseteq S$.
D'altra parte se T è un sottospazio affine che contiene $P_0, \dots, P_s$, si può scrivere $T = P_0 + giacT$ e risulta $P_1 - P_0, \dots, P_s - P_0 \in giacT$, quindi $giacT \supseteq giacS$ da cui $T \supseteq S$; ma allora

$$\overline{P_0 \dots P_s} = \bigcap_{T \ sott. \ affine \ \supseteq P_0, \dots, P_s} T \ \supseteq \ S$$

■

Definizione 5.14 Se $P_0, \dots, P_s$ sono punti di V_a, si ha sempre $dim\overline{P_0 \dots P_s} \leq s$.
Se $dim\overline{P_0 \dots P_s} = s$ i punti $P_0, \dots, P_s$ si dicono *affinemente indipendenti*, se $dim\overline{P_0 \dots P_s} < s$ si dicono *affinemente dipendenti*; i punti $P_0, \dots, P_s$ sono affinemente dipendenti se e solo se i vettori $P_1 - P_0, \dots, P_s - P_0$ sono l.d.

I punti $P_0, \dots, P_s$ si dicono *allineati*, risp. *complanari*, se esiste una retta, risp. un piano, che li contiene, cioé se $dim\overline{P_0 \dots P_s} \leq 1$, risp. ≤ 2.

Esempi 5.15 *a*) Due punti P_0, P_1 sono aff. ind. $\iff$ $P_0 \neq P_1$, nel qual caso $\overline{P_0 P_1}$ è la retta $P_0 + < P_1 - P_0 >$.

b) Tre punti P_0, P_1, P_2 sono aff. ind. $\iff$ non stanno su una retta $\iff$ $P_1 - P_0, P_2 - P_0$ sono l.i., nel qual caso $\overline{P_0 P_1 P_2}$ è il piano $P_0 + < P_1 - P_0, P_2 - P_0 >$.

D'altra parte, tre punti distinti P_0, P_1, P_2 sono aff. dip. $\iff$ $P_1 - P_0 = \rho(P_2 - P_0)$ per un qualche $\rho \in K$, e in questo caso la retta $\overline{P_0\,P_1\,P_2}$ che li contiene è $P_0 + < P_1 - P_0 >$.

c) Quattro punti P_0, P_1, P_2, P_3 sono aff. ind. $\iff$ non stanno su un piano $\iff$ $P_1 - P_0, P_2 - P_0, P_3 - P_0$ sono l.i., nel qual caso $\overline{P_0\,P_1\,P_2\,P_3}$ è il sottospazio 3-dimensionale $P_0 + < P_1 - P_0, P_2 - P_0, P_3 - P_0 >$; quindi possiamo trovare quattro punti aff. ind. solo se la dimensione dello spazio ambiente V_a è almeno 3.

Per esempio, sono aff. ind. in $\mathbb{A}^3(\mathbb{R})$ i punti:

$$(0,0,0), (1,0,0), (0,1,0), (0,0,1)$$

5.2 Affinità

Definizione 5.16 Sia V un K-spazio vettoriale, $dimV = n$, e sia $v \in V$. La *traslazione definita dal vettore* v è l'applicazione così definita:

$$\begin{array}{rccc} t_v : & V_a & \to & V_a \\ & P & \mapsto & P+v \end{array}$$

Si osservi che una traslazione t_v è biettiva. Infatti, è su, perché $\forall Q \in V_a$, $Q = t_v(Q-v)$. La t_v è 1-1 perché $t_v(P) = t_v(Q) \Rightarrow P+v = Q+v \Rightarrow P = Q$.

Proposizione-Definizione 5.17 L'insieme delle traslazioni di V_a con l'operazione di composizione costituisce un gruppo, detto *il gruppo delle traslazioni* e denotato con T_V; tale gruppo è isomorfo al gruppo additivo di V tramite l'isomorfismo:

$$\begin{array}{rccc} \tau : & (V,+) & \to & (T_V, \circ) \\ & v & \mapsto & t_v \end{array}$$

Dimostrazione T_V è un gruppo: se $v, w \in V$, si ha

$$(t_v \circ t_w)(P) = (P+w)+v = P+(v+w) = t_{v+w}(P)$$

quindi $t_v \circ t_w = t_{v+w} \in T_V$; l'elemento neutro è $id_V = t_0$ e ogni elemento ha inverso: $t_v^{-1} = t_{-v}$.

La τ è su e anche 1-1 perché $t_v = t_w \Rightarrow t_v(0) = t_w(0) \Rightarrow v = w$; τ è un morfismo perché $t_v \circ t_w = t_{v+w}$.

Esempio 5.18 Consideriamo $\mathbb{A}^2 = \mathbb{A}^2(\mathbb{R})$ con il riferimento standard, e sia $v = (1,2) \in \mathbb{R}^2$; allora

$$\begin{array}{rccc} t_v : & \mathbb{A}^2 & \to & \mathbb{A}^2 \\ & (x,y) & \mapsto & (x+1, y+2) \end{array}$$

applicare la traslazione al piano significa che tutto il piano scorre su se stesso parallelamente al vettore v.

Definizione 5.19 Siano V e W K-spazi vettoriali di dimensione finita, e sia $\varphi : V \to W$ una applicazione lineare; chiameremo ancora applicazione lineare φ l'applicazione di spazi affini che agisce come φ:

$$\begin{array}{rccc} \varphi : & V_a & \to & W_a \\ & P & \mapsto & \varphi(P) \end{array}$$

Una *applicazione affine* $f : V_a \to W_a$ è la composizione di una applicazione lineare $\varphi : V_a \to W_a$ con una traslazione $t_w : W_a \to W_a$:

$$f = t_w \circ \varphi$$

Risulta quindi

$$f(P) = w + \varphi(P)$$

Osservazione 5.20 Una traslazione t_v non è mai lineare a meno che $v = 0$; infatti, se $v \neq 0$, $t_v(0) \neq 0$.

Proposizione 5.21 *Siano V e W K-spazi vettoriali di dimensione finita, e sia $f : V_a \to W_a$ una applicazione affine. Allora f si scrive in modo unico come composizione di una applicazione lineare con una traslazione, cioé se $f = t_w \circ \varphi$ e anche $f = t_v \circ \psi$, si ha $t_v = t_w$ e $\varphi = \psi$. Chiameremo φ* la parte lineare di f *e* t_w la traslazione associata ad f.

La f è biettiva se e solo se la sua parte lineare φ è biettiva.

Dimostrazione Sia $f = t_w \circ \varphi = t_v \circ \psi$ con $\varphi, \psi : V \to W$ lineari e $v, w \in W$; allora $f(0) = w + \varphi(0) = v + \psi(0) \Rightarrow w + 0 = v + 0 \Rightarrow w = v$ da cui $\forall x \in V_a$, $f(x) = w + \varphi(x) = w + \psi(x) \Rightarrow \varphi(x) = \psi(x) \Rightarrow \varphi = \psi$.

Per finire, $f = t_w \circ \varphi$ è composizione di una biezione, t_w, con la sua parte lineare, quindi f è biettiva se e solo se φ è biettiva.

Osservazione 5.22 $f = t_w \circ \varphi$ è nota se è nota φ e l'immagine $f(P)$ di un punto P; infatti in tal caso

$$f(P) = w + \varphi(P)$$

cioé

$$w = f(P) - \varphi(P)$$

In particolare

$$w = f(0)$$

Definizione 5.23 Siano V e W K-spazi vettoriali di dimensione finita n, e sia $f : V_a \to W_a$ una applicazione affine, $f = t_w \circ \varphi$; f è detta *isomorfismo affine di V_a su W_a* se è biettiva, o equivalentemente se φ è un isomorfismo vettoriale.

Un isomorfismo affine di V_a su se stesso è detto *affinità di V_a*; dunque una affinità di V_a è della forma $t_v \circ \varphi$ con $v \in V$ e $\varphi \in GL(V)$.

Proposizione-Definizione 5.24 Sia V un K-spazio vettoriale, $dimV = n$. L'insieme:

$$Aff(V) := \{f : V_a \to V_a, \;\; f \text{ affinita}'\}$$

con l'operazione di composizione è un gruppo, detto *il gruppo affine di V_a*; i suoi sottogruppi sono detti *gruppi di trasformazioni affini di V_a*.

Se $V = K^n$, $Aff(V)$ si chiama *il gruppo affine di ordine n su K*.

Dimostrazione L'operazione è interna: se $f = t_v \circ \varphi$, $g = t_w \circ \psi$, allora $(f \circ g)(x) = f(g(x)) = f(w + \psi(x)) = v + \varphi(w + \psi(x)) = v + \varphi(w) + (\varphi \circ \psi)(x) = (t_{v+\varphi(w)} \circ (\varphi \circ \psi))(x)$, quindi

$$f \circ g = (t_v \circ \varphi) \circ (t_w \circ \psi) = t_{v+\varphi(w)} \circ (\varphi \circ \psi) \in Aff(V_a) \qquad (\star)$$

$id_{V_a} = t_0 \circ id_V \in Aff(V)$;

se $f = t_v \circ \varphi \in Aff(V)$, f^{-1} esiste perché f è biettiva, ed esiste anche φ^{-1} perché $\varphi \in GL(V)$. Ponendo $v' := \varphi^{-1}(-v)$, da $(\star)$ si ha:

$$(t_v \circ \varphi) \circ (t_{v'} \circ \varphi^{-1}) = t_{v+\varphi(v')} \circ (\varphi \circ \varphi^{-1}) = t_{v-v} \circ id_V = id_{V_a}$$

quindi è un'affinità.

Proposizione 5.25 *Sia V un K-spazio vettoriale, $dimV = n$. Allora*

a) Le traslazioni di V_a sono tutte e sole le affinità della forma $t_v \circ id_V$; T_V è un gruppo di trasformazioni affini.

b) Se P è un punto fissato di V_a, l'insieme:

$$Aff(V)_P := \{f \in Aff(V), \;\; f(P) = P\}$$

con l'operazione di composizione è un gruppo di trasformazioni affini, detto il gruppo delle affinità che fissano il punto P.

Tale gruppo è isomorfo al gruppo $GL(V)$ tramite l'isomorfismo:

$$\begin{array}{cccc} \mu : & (Aff(V)_P, \circ) & \rightarrow & (GL(V), \circ) \\ & t_v \circ \varphi & \mapsto & \varphi \end{array}$$

Dimostrazione *a)*: basta osservare che $T_V \subset Aff(V)$, e che T_V è un gruppo per 5.17 rispetto alla stessa operazione $\circ$ che abbiamo in $Aff(V)$.

b): $Aff(V)_P$ è un sottogruppo di $Aff(V)$, infatti:

$f(P) = P$, $g(P) = P \Rightarrow (f \circ g)(P) = f(g(P)) = P$;

$id_{V_a}(P) = P$;

se $f(P) = P$, f^{-1}, che esiste perché $f \in Aff(V)$, è un'affinità tale che $f^{-1}(P) = P$.

μ è un morfismo di gruppi: da $(\star)$ in dimostrazione di 5.24 si ha che la parte lineare di $(t_v \circ \varphi) \circ (t_w \circ \psi)$ è $\varphi \circ \psi$, cioé $\mu(f \circ g) = \mu(\varphi) \circ \mu(\psi)$;

μ è 1-1 per 5.21 e 5.22;

μ è su, infatti sia $\varphi \in GL(V)$; allora $f := t_{P-\varphi(P)} \circ \varphi$ fissa P e $\mu(f) = \varphi$. ∎

Proposizione 5.26 *Sia V un K-spazio vettoriale, $dimV = n$, $f = t_v \circ \varphi$ una affinità di V_a, e S un sottospazio affine; allora $f(S)$ è un sottospazio affine della stessa dimensione di S e di giacitura $\varphi\,(giac\,S)$. Più precisamente, se Q è un punto di S, si ha:*

$$f(S) = f(Q) + \varphi\,(giac\,S) \qquad (\dagger)$$

Dimostrazione Sia $giacS = W$, e $Q \in S$, quindi $S = Q + W$. Allora

$$f(S) = f(Q + W) = \{f(Q + w), \;\; w \in W\} = \{v + \varphi(Q + w), \;\; w \in W\} =$$
$$= \{(v + \varphi(Q)) + \varphi(w), \;\; w \in W\} = (v + \varphi(Q)) + \varphi(W) = f(Q) + \varphi(W)$$

Così abbiamo provato $(\dagger)$; essendo ϕ un automorfismo, $dim(\phi(W)) = dimW$, da cui $dimf(S) = dimS$.

∎

Proposizione 5.27 *Sia V un K-spazio vettoriale, $dimV = n$, $f = t_v \circ \varphi$ una affinità di V_a, e S, T sottospazi affini. Allora S e T sono sghembi, risp. incidenti, risp. paralleli se e solo se $f(S)$ e $f(T)$ lo sono.*

Questo si esprime dicendo che essere sghembi, essere incidenti, ed essere paralleli sono *proprietà affini*, cioé si mantengono per affinità.

Dimostrazione Supponiamo per esempio $dimS \leq dimT$. Si ha:

$$S//T \iff giac\,S \subseteq giac\,T \iff \underbrace{\varphi(giac\,S)}_{giac\,(f(S))} \subseteq \underbrace{\varphi(giac\,T)}_{giac\,(f(T))} \iff f(S)\,//\,f(T)$$

Ora supponiamo S e T non paralleli, e dunque anche $f(S)$, $f(T)$ non paralleli. Si ha:
$S \cap T \neq \emptyset \iff f(S) \cap f(T) \neq \emptyset$ perché f è un'applicazione, ed è biettiva, quindi ha una inversa; quindi S, T sono incidenti se e solo se $f(S), f(T)$ lo sono, e S e T sono sghembi se e solo se $f(S)$ e $f(T)$ lo sono. ■

Definizione 5.28 Sia V un K-spazio vettoriale, $dimV = n$, $f = t_v \circ \varphi$ una affinità di V_a, e S un sottospazio affine; diciamo che S *è un sottospazio fisso per* f se $f(S) = S$; diciamo invece che S *è un sottospazio di punti fissi per* f se $f(P) = P\ \forall P \in S$.

5.3 Equazioni di una affinità

Proposizione 5.29 *Siano V un K-spazio vettoriale, $\dim V = n$, $\mathcal{R} = A\mathcal{B}$ un riferimento affine fissato , e sia $f = t_v \circ \varphi$ una affinità.*
Allora se $P = (x_1, \ldots, x_n)_{\mathcal{R}}$, $f(P) = (y_1, \ldots, y_n)_{\mathcal{R}}$, e $f(A) = (c_1, \ldots, c_n)_{\mathcal{R}}$, risulta:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = M_{\mathcal{B}}(\varphi) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \tag{$\diamondsuit$}$$

Dimostrazione Si ha

$$f(P) - f(A) = (v + \varphi(P)) - (v + \varphi(A)) = \varphi(P) - \varphi(A) = \varphi(P - A) \tag{$\circ$}$$

da cui

$$f(P) - A = \varphi(P - A) + f(A) - A \tag{$*$}$$

Ricordiamo che $P - A = (x_1, \ldots, x_n)_{\mathcal{B}}$, $\quad f(P) - A = (y_1, \ldots, y_n)_{\mathcal{B}}$, $\quad f(A) - A = (c_1, \ldots, c_n)_{\mathcal{B}}$.

Scrivendo $(*)$ in coordinate rispetto alla base $\mathcal{B}$ si ha quindi $(\diamondsuit)$.

Osservazione 5.30 Ponendo $X := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $Y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $C := \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, le $(\diamondsuit)$ si scrivono in forma compatta:

$$Y = M_{\mathcal{B}}(\varphi)\, X + C$$

Se vogliamo le equazioni per esteso, sia $M_{\mathcal{B}}(\varphi) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$; allora $(\diamondsuit)$ è

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

cioé

$$\begin{cases} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n + c_1 \\ \vdots \\ y_n = a_{n1}x_1 + \ldots + a_{nn}x_n + c_n \end{cases} \tag{$**$}$$

Definizione 5.31 Le $(**)$ o equivalentemente le $(\diamondsuit)$ vengono dette *le equazioni della f rispetto al riferimento $\mathcal{R}$.*

Osservazione 5.32 Da 5.22 si ha $v = f(0)$. Nelle notazioni di 5.29 si ha $f(A) - A = (c_1, \ldots, c_n)_{\mathcal{B}}$, quindi se $A = 0$ allora $(c_1, \ldots, c_n)_{\mathcal{B}} = f(0) - 0 = v$. Quindi se scriviamo le equazioni di $f = t_v \circ \varphi$ in un sistema di riferimento $\mathcal{R} = 0\mathcal{B}$ dove $0 = 0_V$ denota come sempre il vettore nullo, troviamo le $(\diamondsuit)$ dove la matrice colonna C è data dalle coordinate del vettore di traslazione v rispetto alla base $\mathcal{B}$.

Osservazione 5.33 *a)* Se $\mathcal{R} = 0\mathcal{B}$ le equazioni di una traslazione t_v sono

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}, \qquad \text{dove } v = (c_1, \dots, c_n)_{\mathcal{B}}$$

b) Se $\mathcal{R} = A\mathcal{B}$ e $f \in Aff(V)_A$ le equazioni di $f = t_v \circ \varphi$ sono:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = M_{\mathcal{B}}(\varphi) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Esempio 5.34 Consideriamo $\mathbb{A}^2(\mathbb{R})$ con il riferimento standard $\mathcal{S} = 0\mathcal{E}$; determiniamo una affinità che porti la retta $R : x + y + 1 = 0$ nella retta $S : x - y + 2 = 0$.

Ci sono vari modi per trovarla, uno è questo: cerchiamo $\varphi \in GL(\mathbb{R}^2)$ tale che $\varphi(giac\, R) = giac\, S$:

$giac\, R : x + y = 0$ quindi $giac\, R =< (1, -1) >$,

$giac\, S : x - y = 0$ quindi $giac\, S =< (1, 1) >$.

Consideriamo la base di $\mathbb{R}^2$ $\mathcal{B} = (\underbrace{(1, -1)}_{=v_1}, \underbrace{(1, 1)}_{=v_2})$, e l'automorfismo di $\mathbb{R}^2$:

$$\begin{array}{rccc} \varphi : & \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ & (1, -1) & \mapsto & (1, 1) \\ & (1, 1) & \mapsto & (1, -1) \end{array}$$

quindi $M_{\mathcal{B}}(\varphi) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Cerchiamo la matrice $M_{\mathcal{E}}(\varphi)$: si ha

$$e_1 = \frac{1}{2}v_1 + \frac{1}{2}v_2, \quad e_2 = -\frac{1}{2}v_1 + \frac{1}{2}v_2 \Rightarrow$$

$$\varphi(e_1) = \frac{1}{2}\varphi(v_1) + \frac{1}{2}\varphi(v_2) = e_1, \quad \varphi(e_2) = -\frac{1}{2}\varphi(v_1) + \frac{1}{2}\varphi(v_2) = -e_2 \Rightarrow$$

$$M_{\mathcal{E}}(\varphi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Si ha $R = P+ < (1, -1) >$ con $P = (-2, 1)$ e $S = Q+ < (1, 1) >$ con $Q = (1, 3)$; quindi se $f = t_v \circ \varphi$ è tale che $f(P) = Q$, allora f soddisfa la richiesta $f(R) = S$ (si veda per esempio la dimostrazione di 5.26 , (†)).

Da 5.22

$$v = f(P) - \varphi(P) = Q - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = (1, 3) - (-2, -1) = (3, 4)$$

Quindi equazioni per f nel riferimento standard $\mathcal{S}$ sono

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

Se invece avessimo voluto fare i conti direttamente rispetto al riferimento $\mathcal{R} = 0\mathcal{B}$:

$$P - 0 = (-2, 1) = (-\frac{3}{2}, -\frac{1}{2})_{\mathcal{B}} \Rightarrow P = (-\frac{3}{2}, -\frac{1}{2})_{\mathcal{R}}, Q - 0 = (1, 3) = (-1, 2)_{\mathcal{B}} \Rightarrow Q = (-1, 2)_{\mathcal{R}}$$

quindi se imponiamo $Q = f(P)$:

$$v = Q - \varphi(P) = ((-1, 2) - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \end{pmatrix})_{\mathcal{B}} = ((-1, 2) - (-\frac{1}{2}, -\frac{3}{2}))_{\mathcal{B}} = (-\frac{1}{2}, \frac{7}{2})_{\mathcal{B}}$$

da cui equazioni per f nel riferimento $\mathcal{R}$ sono

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -\frac{1}{2} \\ \frac{7}{2} \end{pmatrix}$$

Un altro modo per risolvere l'esercizio è, visto che R ed S sono incidenti in un punto, trovare il punto $A = R \cap S$; si deve avere $f(A) = A$, quindi equazioni per f nel riferimento $A\mathcal{B}$ sono

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

6 Spazi euclidei

6.1 La struttura di spazio euclideo su uno spazio vettoriale reale

Definizione 6.1 Se V è uno spazio v.e., allora lo spazio affine V_a è detto spazio euclideo, e verrà denotato con V_e.

Quindi uno spazio euclideo è un $\mathbb{R}$-spazio vettoriale dotato di un prodotto scalare su cui consideriamo la struttura affine.

Se $V = \mathbb{R}^n$ con il prodotto scalare standard lo spazio affine $\mathbb{A}^n(\mathbb{R})$ è uno spazio euclideo, che denoteremo con E^n.

In uno spazio euclideo un riferimento $A\mathcal{N}$ dove $\mathcal{N}$ è una base ortonormale è detto *riferimento cartesiano* o *sistema di coordinate cartesiane.*

Esempio 6.2 In E^n il riferimento standard $0\mathcal{E}$ è un riferimento cartesiano.

In E^2 il riferimento $A\mathcal{B}$ dove $A = (1, 2)$ e

$$\mathcal{B} = \left((\frac{2\sqrt{2}}{3}, \frac{1}{3}), (-\frac{1}{3}, \frac{2\sqrt{2}}{3}) \right)$$

è un riferimento cartesiano.

Definizione 6.3 Sia V_e uno spazio euclideo e siano P, Q due suoi punti. La *distanza tra P e Q* è il numero reale ≥ 0 così definito:

$$d(P, Q) := ||Q - P|| = \sqrt{(Q - P) \bullet (Q - P)}$$

Osservazione 6.4 Sia V_e uno spazio euclideo e $\mathcal{R}$ un riferimento cartesiano; allora

$$d((x_1, \ldots, x_n)_{\mathcal{R}}, (y_1, \ldots, y_n)_{\mathcal{R}}) = \sqrt{\sum_{i=1}^{n}(y_i - x_i)^2}$$

Definizione 6.5 *a*) Sia V uno spazio vettoriale euclideo, e siano $v, v' \in V$, $v \neq 0, v' \neq 0$; allora la disuguaglianza di Schwarz dice che

$$|v \bullet v'| \leq ||v|| \cdot ||v'||$$

quindi

$$-1 \leq \frac{v \bullet v'}{||v|| \cdot ||v'||} \leq 1$$

Ne segue che esiste un unico θ tale che

$$cos\,\theta = \frac{v \bullet v'}{||v|| \cdot ||v'||}, \qquad 0 \leq \theta \leq \pi \tag{$\star$}$$

Possiamo quindi definire *l'angolo convesso tra i vettori v e v'* come l'unico θ che verifica ($\star$), ossia come l'unico θ tale che

$$v \bullet v' = cos\,\theta\, ||v|| \cdot ||v'||, \qquad 0 \leq \theta \leq \pi$$

cioé

$$\theta = arccos\frac{v \bullet v'}{||v|| \cdot ||v'||}$$

$b)$Sia V_e uno spazio euclideo, $dimV_e = n$, e sia $R = Q + < v >$ una retta; esistono esattamente due versori su $giacR$, cioé $\pm\frac{v}{||v||}$.

Una retta orientata è una retta R insieme alla scelta di uno dei due versori di direzione.

Siano R ed S due rette orientate, con versori fissati rispettivamente u e w. L'*angolo convesso delle due rette orientate* è definito come l'angolo convesso θ tra u e w, cioé è il θ tale che $0 \leq \theta \leq \pi$, e

$$\cos\theta = u \bullet w$$

Le rette R ed S sono ortogonali se $\theta = \frac{\pi}{2}$, cioé se $u \bullet w = 0$.

Proposizione 6.6 *L'applicazione distanza:*

$$\begin{array}{cccc} d: & V_e \times V_e & \rightarrow & \mathbb{R} \\ & (P,Q) & \mapsto & d(P,Q) \end{array}$$

gode delle seguenti proprietà:

a) $d(P,Q) \geq 0 \quad \forall P,Q \in V_e, \quad d(P,Q) = 0 \iff P = Q$

b) $d(P,Q) = d(Q,P) \quad \forall P,Q \in V_e$

c) $d(P,Q) + d(Q,R) \geq d(P,R) \quad \forall P,Q,R \in V_e$

Dimostrazione Segue dalle proprietà della norma (4.15).

6.2 Isometrie

Definizione 6.7 Sia V_e uno spazio euclideo, $dimV_e = n$; una *isometria di* V_e è una affinità $f = t_v \circ \varphi$ tale che $\varphi \in O(V)$.

La f è *una isometria diretta* se $\varphi \in SO(V)$, *una isometria inversa* se $\varphi \in O(V) \setminus SO(V)$; quindi f è diretta $\iff$ $det\varphi = 1$, f è inversa $\iff$ $det\varphi = -1$.

Una isometria diretta che lascia fisso un punto P è detta *rotazione di centro* P.

Proposizione-Definizione 6.8

$$Isom(V) := \{f \in Aff(V), \quad f \text{ isometria}\}$$

è un sottogruppo di $Aff(V)$, detto *il gruppo delle isometrie di* V.

$$Isom^+(V) := \{f \in Isom(V), \quad f \text{ isometria diretta}\}$$

è un sottogruppo di $Isom(V)$, detto *il gruppo delle isometrie dirette di* V.

Se P è un punto, l'insieme delle rotazioni di centro P:

$$Isom^+(V)_P = Isom^+(V) \cap Aff(V)_P$$

è un sottogruppo di $Isom^+(V)$.

Dimostrazione Esercizio

Osservazione 6.9 Una isometria $f = t_v \circ \varphi$ conserva le distanze e gli angoli. Infatti, φ conserva la norma, quindi se P e Q sono due punti,

$$d(f(P), f(Q)) = ||f(Q) - f(P)|| = ||v + \varphi(Q) - (v + \varphi(P))|| = ||\varphi(Q - P)|| = ||Q - P|| = d(P, Q)$$

Inoltre, se R e S sono due rette orientate, con versori fissati rispettivamente u e w, e θ è l'angolo convesso di R ed S, le immagini delle due rette sono le rette orientate $f(R)$ ed $f(S)$ con versori fissati rispettivamente $\varphi(u)$ e $\varphi(w)$; si ha

$$giacf(R) = \varphi(< u >) =< \varphi(u) >, \quad giacf(S) = \varphi(< w >) =< \varphi(w) >$$

Se η è l'angolo convesso di $f(R)$ ed $f(S)$, si ha quindi:

$$\cos\theta = u \bullet w = \varphi(u) \bullet \varphi(w) = cos\eta$$

e $0 \leq \theta \leq \pi$, $0 \leq \eta \leq \pi$, da cui $\theta = \eta$.

6.3 Le isometrie di E^2

Nel seguito consideriamo E^2 dove è fissato il riferimento standard $0\mathcal{E}$; $O = (0,0)$ denota l'origine dei vettori.

Definizione 6.10 Una $\varphi \in O(\mathbb{R}^2)$, $\varphi \neq id$, che fissa tutti i vettori di una retta vettoriale r è detta *riflessione*, ed r si chiama asse di riflessione.

Analogamente, una f isometria di E^2, $f \neq id$, che fissa tutti i punti di una retta r è detta *riflessione*, ed r si chiama asse di riflessione.

Definizione 6.11 Una isometria $f = t_v \circ \varphi$ di E^2 con φ riflessione di asse r, e $v \neq 0$, $v//r$, è detta *glissoriflessione*, ed r si chiama *l'asse di glissoriflessione.*

Notazione 6.12 Nel seguito poniamo

$$R_\theta = \begin{pmatrix} cos\theta & -sen\theta \\ sen\theta & cos\theta \end{pmatrix}$$

$$A_\theta = \begin{pmatrix} cos\theta & sen\theta \\ sen\theta & -cos\theta \end{pmatrix} = R_\theta \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Osservazione 6.13 Ricordiamo che (foglio di esercizi 5)

$$SO(2) = \{R_\theta, \quad \theta \in \mathbb{R}\}, \qquad O(2) \setminus SO(2) = \{A_\theta, \quad \theta \in \mathbb{R}\}.$$

Sia $f = t_v \circ \varphi$ una isometria. Poiché $\varphi \in O(\mathbb{R}^2)$ e $\mathcal{E}$ è base ortonormale, si ha $M_\mathcal{E}(\varphi) \in O(2) = SO(2) \cup (\mathcal{O}(2) \setminus \mathcal{SO}(2))$, quindi $M_\mathcal{E}(\varphi)$ è della forma R_θ oppure A_θ.

Le equazioni di f sono dunque, se $P = (x,y) = (x,y)_\mathcal{S}$, $f(P) = (x', y')$, $v = f(0) = (a,b)$,

$$\begin{cases} x' = xcos\theta - ysen\theta + a \\ y' = xsen\theta + ycos\theta + b \end{cases} \qquad \textit{se } f \textit{ diretta}$$

$$\begin{cases} x' = xcos\theta + ysen\theta + a \\ y' = xsen\theta - ycos\theta + b \end{cases} \qquad \textit{se } f \textit{ inversa}$$

Cominciamo a studiare φ.

Lemma 6.14 Nelle notazioni precedenti

a) se φ è diretta è una rotazione vettoriale, o rotazione di centro O, di equazione

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R_\theta \begin{pmatrix} x \\ y \end{pmatrix}$$

che non ha autovalori reali (quindi non ci sono rette fisse) per $\theta \neq 0, \pi + 2k\pi$; un versore $(cos\alpha, sen\alpha) \in S^1$ viene mandato da f in $(cos(\alpha+\theta), sen(\alpha+\theta))$.

Se $\theta = 0$ c'è un unico autovalore, 1, con $\mu(1) = 2$, e $\varphi = id$, se $\theta = \pi$ c'è un unico autovalore, -1, con $\mu(-1) = 2$, e $\varphi = -id$.

$b)$ se φ è inversa ha equazione

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A_\theta \begin{pmatrix} x \\ y \end{pmatrix}$$

osserviamo che A_θ è simmetrica, quindi φ ha una base ortonormale di autovettori; precisamente, ha due autovalori, 1 e -1, e i relativi autospazi sono

$$V_1 =< (cos\theta + 1, sen\theta) >, \qquad V_{-1} =< (cos\theta - 1, sen\theta) >$$

e si ha che $(cos\theta + 1, sen\theta)$ e $(cos\theta - 1, sen\theta)$ sono ortogonali.

Ne segue che φ è la riflessione rispetto alla retta V_1.

Dimostrazione $a)$ Cerchiamo gli autovalori:

$$p_\varphi = det\begin{pmatrix} cos\theta - t & -sen\theta \\ sen\theta & cos\theta - t \end{pmatrix} = cos^2\theta - 2tcos\theta + t^2 + sen^2\theta = t^2 - 2cos\theta t + 1$$

Quindi le radici $cos\theta \pm \sqrt{cos^2\theta - 1}$ sono reali $\iff cos^2\theta = 1 \iff cos\theta = \pm 1 \iff \theta = 0, \pi + 2k\pi$

Se $\theta = 0$, $R_\theta = I$, cioé $\varphi = id$; c'è un unico autovalore doppio, 1, e $V_1 = \mathbb{R}^2$.

Se $\theta = \pi$, $R_\theta = -I$, cioé $\varphi = -id$; c'è un unico autovalore doppio, -1, e $V_{-1} = \mathbb{R}^2$. In questo caso $\varphi(x, y) = -(x, y)$, cioé φ porta una retta per O in se stessa invertendone il senso.

Consideriamo un versore $(cos\alpha, sen\alpha) \in S^1$; si ha:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R_\theta \begin{pmatrix} cos\alpha \\ sen\alpha \end{pmatrix} = \begin{pmatrix} cos\theta cos\alpha - sen\theta sen\alpha \\ sen\theta cos\alpha + cos\theta sen\alpha \end{pmatrix} = \begin{pmatrix} cos(\alpha + \theta) \\ sen(\alpha + \theta) \end{pmatrix}$$

quindi φ ruota di θ ogni versore della circonferenza unitaria, e quindi anche ogni retta per O.

$b)$ Cerchiamo gli autovalori:

$$p_\varphi = det\begin{pmatrix} cos\theta - t & sen\theta \\ sen\theta & -cos\theta - t \end{pmatrix} = -cos^2\theta - tcos\theta + tcos\theta + t^2 - sen^2\theta = t^2 - 1$$

quindi ci sono due autovalori, 1 e -1, di molteplicità 1, e si ha:

$$V_1 : \quad (cos\theta - 1)x + sen\theta y = 0 \quad \Rightarrow \quad V_1 =< \underbrace{(cos\theta + 1, sen\theta)}_{=:u} >$$

$$V_{-1} : \quad (cos\theta + 1)x + sen\theta y = 0 \quad \Rightarrow \quad V_{-1} =< \underbrace{(cos\theta - 1, sen\theta)}_{=:w} >$$

Le due rette V_1 e V_{-1} sono ortogonali. Consideriamo la base ortonormale (di autovettori):

$$\mathcal{B} = (\frac{u}{||u||}, \frac{w}{||w||})$$

Poiché $\varphi(\frac{u}{||u||}) = \frac{u}{||u||}$, e $\varphi(\frac{w}{||w||}) = -\frac{w}{||w||}$, la matrice di φ rispetto a $\mathcal{B}$ è

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

e le equazioni di φ sono, se $v = (X, Y)_\mathcal{B}$, $\varphi(v) = (X', Y')_\mathcal{B}$:

$$\begin{cases} X' = X \\ Y' = -Y \end{cases}$$

Quindi φ è una riflessione di asse V_1 (tutti i vettori di V_1 sono fissi); geometricamente φ ribalta il piano su se stesso lasciando ferma appunto la retta V_1, mentre V_{-1} si ribalta su se stessa: $\varphi((0, Y)) = (0, -Y)$.

Osservazione 6.15 Sia $f = t_v \circ \varphi$ una isometria che fissa un punto A.

Mettiamoci nel riferimento $\mathcal{R} = A\mathcal{E}$; da 5.33 sappiamo che le equazioni di $f = t_v \circ \varphi$ sono, se $M := M_\mathcal{E}(\varphi) \in O(2)$:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = M \begin{pmatrix} x \\ y \end{pmatrix}$$

Osserviamo che $\{P - A, P \in E^2\}$ è identificabile allo spazio vettoriale $\mathbb{R}^2$, e che, inoltre, $P = (x, y)_\mathcal{R} \iff P - A = (x, y)_\mathcal{E} = (x, y)$, $f(P) = (x', y')_\mathcal{R} \iff f(P) - A = (x', y')_\mathcal{E} = (x', y')$; quindi f è un endomorfismo unitario dello spazio vettoriale $\{P - A, P \in E^2\}$, e quanto visto in 6.14 è ancora valido, pensando ai vettori con origine in A anziché in O, cioé traslando del vettore $A - O$ i risultati di 6.14.

Quindi, se f è una isometria con un punto fisso A, è una rotazione di centro A se è diretta, ed è una riflessione se è inversa. Inoltre, nel secondo caso, ha una retta R di punti fissi (che è l'asse di riflessione) e una retta fissa S ortogonale ad R, entrambe passanti per A; ma allora ogni retta ortogonale ad R è fissa.

Lemma 6.16 *Sia $g = t_u \circ \varphi$ una isometria inversa, con φ riflessione di asse R e $u \perp R$. Allora g è una riflessione di asse una retta S, con $S // R$.*

Dimostrazione Sia $\mathcal{B} = (v_1, v_2)$ la base ortonormale di autovettori per φ dove $\varphi(v_1) = v_1$, $\varphi(v_2) = -v_2$; allora $R =< v_1 >$, $u = (0, c)_\mathcal{B}$ e le equazioni di φ rispetto a $\mathcal{B}$ sono:

$$\begin{cases} X' = X \\ Y' = -Y \end{cases}$$

Sia ora $\mathcal{R} = 0\mathcal{B}$; allora le equazioni di g rispetto a $\mathcal{R}$ sono:

$$\begin{cases} X' = X \\ Y' = -Y + c \end{cases}$$

quindi g fissa tutti i punti tali che $Y = -Y + c$, cioé tutti i punti della retta $S : Y = \frac{c}{2}$, che è $//R$.
■

Teorema 6.17 (Teorema di Chasles, 1831)
Una isometria di E^2 che fissa un punto è una rotazione se è diretta, una riflessione se è inversa. Una isometria di E^2 che non fissa alcun punto è una traslazione se è diretta, una glissoriflessione se è inversa.

Dimostrazione Sia $f = t_v \circ \varphi$. Se f ha un punto fisso, l'enunciato segue da 6.15.

A) Sia f diretta e senza punti fissi.

Da (6.13) le equazioni di f sono:

$$\begin{cases} x' = xcos\theta - ysen\theta + a \\ y' = xsen\theta + ycos\theta + b \end{cases}$$

I punti fissi di f sono i punti $P = (x, y)$ tali che $f(P) = P$; questi sono le soluzioni del sistema:

$$\begin{cases} x = xcos\theta - ysen\theta + a \\ y = xsen\theta + ycos\theta + b \end{cases}$$

cioé le soluzioni di

$$\begin{cases} (cos\theta - 1)x - (sen\theta)y + a = 0 \\ (sen\theta)x + (cos\theta - 1)y + b = 0 \end{cases}$$

che non ha soluzioni se e solo se

$$rg\begin{pmatrix} cos\theta - 1 & -sen\theta \\ sen\theta & cos\theta - 1 \end{pmatrix} < rg\begin{pmatrix} cos\theta - 1 & -sen\theta & a \\ sen\theta & cos\theta - 1 & b \end{pmatrix}$$

cioé se la prima matrice ha rango 1 e la seconda ha rango 2, oppure se la prima ha rango 0 e la seconda ha rango 1. Ma se il rango della prima matrice è < 2, allora è necessariamente 0, infatti:

$$det\begin{pmatrix} cos\theta - 1 & -sen\theta \\ sen\theta & cos\theta - 1 \end{pmatrix} = 0 \iff cos^2\theta - 2cos\theta + 1 + sen^2\theta = 0 \iff$$

$$\iff 2(1 - cos\theta) = 0 \iff cos\theta = 1 \iff \theta = 0 \iff \begin{pmatrix} cos\theta - 1 & -sen\theta \\ sen\theta & cos\theta - 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Quindi f non ha punti fissi se e solo se $\theta = 0$ e inoltre

$$rg\begin{pmatrix} cos\theta - 1 & -sen\theta & a \\ sen\theta & cos\theta - 1 & b \end{pmatrix} = 1 \iff (a, b) \neq (0, 0).$$

Quindi f non ha punti fissi $\iff M_{\mathcal{E}}(\varphi) = I$ e $(a, b) \neq 0 \iff f = t_v$ con $v \neq 0$.

B) Sia f inversa e senza punti fissi.

Da (6.13) le equazioni di f sono:

$$\begin{cases} x' = xcos\theta + ysen\theta + a \\ y' = xsen\theta - ycos\theta + b \end{cases}$$

I punti fissi di f sono le soluzioni del sistema:

$$\begin{cases} (cos\theta - 1)x + (sen\theta)y + a = 0 \\ (sen\theta)x + (-cos\theta - 1)y + b = 0 \end{cases}$$

e dato che

$$det\begin{pmatrix} cos\theta - 1 & sen\theta \\ sen\theta & -cos\theta - 1 \end{pmatrix} = 1 - cos^2\theta - sen^2\theta = 0 \quad \forall\theta$$

e d'altra parte $\forall\theta$ la matrice $\begin{pmatrix} cos\theta - 1 & sen\theta \\ sen\theta & -cos\theta - 1 \end{pmatrix}$ non è la matrice nulla, per ogni θ la non esistenza di punti fissi equivale alla condizione:

$$rg\begin{pmatrix} cos\theta - 1 & sen\theta & a \\ sen\theta & -cos\theta - 1 & b \end{pmatrix} = 2$$

e dato che le prime due colonne sono sempre l.d., questa equivale a

$$(a, b) \text{ non proporzionale a } (cos\theta - 1, sen\theta).$$

Ricordiamo che l'asse di riflessione V_1 di φ è generato da $u = (cos\theta + 1, sen\theta)$ e che $w = (cos\theta - 1, sen\theta)$ è ortogonale a u (6.14), per cui f non ha punti fissi se e solo se $v = (a, b)$ non è ortogonale a V_1.
Ma allora la proiezione ortogonale v' di v su V_1 è $\neq 0$, cioé $v = v' + v$" con $v' \in V_1$, $v' \neq 0$, e $v" \perp v'$.
Ne segue:

$$f = t_v \circ \varphi = t_{v'+v"} \circ \varphi = t_{v'} \circ (t_{v"} \circ \varphi) \qquad \text{con } v' \neq 0, v' \in V_1$$

Il lemma 6.16 ci dice che $t_{v"} \circ \varphi$ è una riflessione di asse una retta parallela a V_1, per cui f è una glissoriflessione.

Lemma 6.18 *Se f è una riflessione, allora $f^2 = id$.*

Basta osservare che $f = t_v \circ \varphi$ con φ che ha due autovalori, 1 e -1, e se v_1 e v_2 sono autovettori relativi rispettivamente a 1 e a -1, allora $\varphi^2(v_1) = v_1$ e $\varphi^2(v_2) = -(-v_2) = v_2$ cioé $\varphi^2 = id$. Quindi

$$f^2 = t_{v+\varphi(v)} \circ \varphi^2 = t_{v+\varphi(v)}$$

ma f^2 fissa i punti dell' asse di riflessione di f, e l'unica traslazione che fissa dei punti è id. ■

References

[A] Michael Artin: *Algebra*. Bollati Boringhieri 1997.

[C] Maurizio Cornalba: http://mate.unipv.it/cornalba/alglin.html

[C2] Charles W. Curtis *Linear Algebra - An Introductory Approach*. UTM Springer-Verlag 1984.

[S] Edoardo Sernesi: *Geometria 1*. Bollati Boringhieri 1989.