Corso di Geometria

Docente: Guido Pezzini

a.a. 2019/2020

Appunti vari

versione del 20.11.2019

1. Verifica della formula della matrice inversa

Sia A una matrice $n \times n$ tale che $\det(A) \neq 0$. La formula per la matrice inversa propone un'espressione per l'entrata al posto (i,j) di A^{-1}. Se chiamiamo $b_{i,j}$ quest'entrata, abbiamo

$$b_{i,j} = \frac{(-1)^{i+j}\det(A_{j,i})}{\det(A)}$$

dove come al solito $A_{j,i}$ è la matrice ottenuta da A rimuovendo la j-esima riga e la i-esima colonna. Verifichiamo che effettivamente

$$A \cdot A^{-1} = I_n$$

Calcoliamo l'entrata $c_{1,1}$ al posto $(1,1)$ del prodotto $A \cdot A^{-1}$. Quest'entrata vale

$$c_{1,1} = a_{1,1}b_{1,1} + a_{1,2}b_{2,1} + \ldots + a_{1,n}b_{n,1} = \sum_{k=1}^{n} a_{1,k}b_{k,1} = \ldots$$

e usando la formula per $c_{i,j}$, otteniamo

$$\ldots = \sum_{k=1}^{n} a_{1,k}\left(\frac{(-1)^{1+k}\det(A_{1,k})}{\det(A)}\right) = \frac{\sum_{k=1}^{n}(-1)^{k+1}a_{1,k}\det(A_{1,k})}{\det(A)} = \ldots$$

Osserviamo che il numeratore della frazione ottenuta è proprio lo sviluppo del determinante di A secondo la prima riga. Per cui

$$\ldots = 1$$

il che corrisponde all'entrata al posto $(1,1)$ di I_n. Lo stesso ragionamento si applica per tutte le entrate sulla diagonale, cioè per $c_{2,2}, c_{3,3}\ldots, c_{n,n}$.

Verifichiamo ora che fuori dalla diagonale la matrice $A \cdot A^{-1}$ ha entrate nulle. Calcoliamo $c_{1,2}$: in modo simile a prima otteniamo

$$c_{1,2} = \sum_{k=1}^{n} a_{1,k}b_{k,2} = \frac{\sum_{k=1}^{n}(-1)^{k+2}a_{1,k}\det(A_{2,k})}{\det(A)}$$

Qui il numeratore non è lo sviluppo secondo una qualche riga del determinante di A. Però osserviamo che è lo sviluppo del determinante di un'altra matrice $\widetilde{A}$, ottenuta copiando tutte le righe di A tranne la seconda, e mettendo al posto della seconda riga la prima riga di A:

$$\widetilde{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n-1} & a_{1,n} \\ a_{\mathbf{1},1} & a_{\mathbf{1},2} & \ldots & a_{\mathbf{1},n-1} & a_{\mathbf{1},n} \\ a_{3,1} & a_{3,2} & \ldots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,n-1} & a_{n,n} \end{pmatrix}$$

Infatti, lo sviluppo del determinante di $\widetilde{A}$ per la seconda riga vale

$$\det(\widetilde{A}) = \sum_{k=1}^{n}(-1)^{k+2}a_{1,k}\det(\widetilde{A}_{2,k}) = \ldots$$

Ricordiamo che $\widetilde{A}_{2,k}$ è ottenuta da $\widetilde{A}$ rimuovendo la seconda riga e la k-esima colonna. Per cui abbiamo $\widetilde{A}_{2,k} = A_{2,k}$ per ogni k. Otteniamo

$$\ldots = \sum_{k=1}^{n} (-1)^{k+2} a_{1,k} \det(A_{2,k})$$

Però sappiamo che $\det(\widetilde{A}) = 0$, perché la matrice ha due righe uguali. Concludiamo che

$$c_{1,2} = 0$$

Allo stesso modo si verifica che $c_{i,j} = 0$ per ogni $i \neq j$. Quindi otteniamo l'uguaglianza voluta

$$A \cdot A^{-1} = I_n$$

2. Dimostrazione dell'invarianza del rango per le operazioni elementari di riga

Dimostriamo che le operazioni elementari di riga non cambiano il rango di una matrice. Vediamo l'argomentazione per una matrice 2×2; il caso generale è più difficile da illustrare (perché compaiono molti più minori, ovviamente), ma il ragionamento è esattamente lo stesso. Per brevità, scriviamo qui sinteticamente "determinante $k \times k$ di una matrice" per intendere il *determinante di un minore $k \times k$ della matrice.*

Sia dunque

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matrice qualsiasi 2×2. Esaminiamo le tre possibili operazioni elementari.

La prima operazione scambia le due righe di A, ottenendo quindi la matrice

$$B = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$

Il rango è il massimo r tale che la matrice ha un minore $r \times r$ invertibile. Confrontiamo i *determinanti* di tutti i minori di A e di B, iniziando dagli 1×1 (scritti da sinistra a destra e dall'alto in basso); osserviamo anche che $\det(B) = -\det(A)$. Otteniamo:

grandezza	A	B
1×1	a, b, c, d	c, d, b, a
2×2	$\det(A)$	$-\det(A)$

Il rango corrisponde all'ultima riga della tabella in cui vedo almeno un numero non nullo. È evidente che questo succede alla stessa riga per A e per B, perchè (a meno di un segno cambiato) i numeri che vedo per B sono gli stessi di quelli di A, è solo cambiato l'ordine in cui li ho scritti. Il rango quindi non cambia se scambiamo due righe della matrice.

La seconda operazione moltiplica una riga fissata per un numero reale non nullo α. Applichiamola alla prima riga, il ragionamento con la seconda riga è identico. Otteniamo stavolta

$$B = \begin{pmatrix} \alpha \cdot a & \alpha \cdot b \\ c & d \end{pmatrix}$$

Sappiamo che stavolta $\det(B) = \alpha \cdot \det(A)$, per cui, in questo caso, la tabella dei determinanti dei minori è la seguente:

grandezza	A	B
1×1	a, b, c, d	$\alpha \cdot a, \alpha \cdot b, c, d$
2×2	$\det(A)$	$\alpha \cdot \det(A)$

Qui osserviamo che i numeri in tabella sono rimasti gli stessi, oppure sono stati moltiplicati per α. Ciascuno di essi, se era diverso da zero prima, lo è ancora, e se era uguale a zero prima, lo è ancora.

Deduciamo che anche in questo caso $\mathrm{rg}(B) = \mathrm{rg}(A)$, cioè il rango non è cambiato dopo l'operazione di riga.

La terza operazione somma ad una riga fissata R_i un multiplo di un'altra R_j, multiplo per un numero reale β. Applichiamola facendo $R_1 \to R_1 + \beta R_2$, il caso in cui cambio R_2 invece di R_1 è uguale. Otteniamo stavolta

$$B = \begin{pmatrix} a + \beta c & b + \beta d \\ c & d \end{pmatrix}$$

Stavolta $\det(B) = \det(A)$, per cui, in questo caso, la tabella dei determinanti dei minori è la seguente:

grandezza	A	B
1×1	a, b, c, d	$a + \beta c, b + \beta d, c, d$
2×2	$\det(A)$	$\det(A)$

La riga dei determinanti 2×2 non è cambiata, per cui non resta che occuparci della riga dei determinanti 1×1. Qui i numeri sono cambiati, e potrebbero essere persino "comparsi" degli zeri dove non ce n'erano! (Ad esempio se $a \neq 0 \neq c$ e abbiamo scelto $\beta = -a/c$.)

Dobbiamo analizzare tutti i casi possibili. Ricordiamo che conta solo *se almeno uno dei determinanti* 1×1 *è diverso da zero*. Controlliamo che, riguardo a questa condizione, quello che succede in A succede sempre anche in B.

(1) Se tutti i determinanti dei minori 1×1 di A sono nulli, allora A è la matrice nulla, e lo è anche B, per cui anche i determinanti dei suoi minori 1×1 sono nulli.
(2) Supponiamo che almeno uno fra i determinanti dei minori 1×1 di A sia non nullo. Abbiamo i seguenti casi possibili:
   (a) vale $c \neq 0$, oppure $d \neq 0$. Osserviamo che c e d compaiono anche fra i determinanti dei minori 1×1 di B, per cui anche uno dei determinanti dei minori 1×1 di B è non nullo. In questo caso siamo a posto. L'altra possibilità è:
   (b) invece, vale $c = d = 0$. In questo caso uno fra a e b deve essere non nullo. Inoltre, visto che $c = d = 0$, i numeri a e b compaiono inalterati fra i determinanti dei minori 1×1 di B. Anche in questo caso siamo a posto.

Deduciamo che anche in questo caso $\mathrm{rg}(B) = \mathrm{rg}(A)$, cioè il rango non è cambiato dopo l'operazione di riga.

3. Dimostrazione del teorema degli orlati

Sia A matrice $m \times n$, sia M minore $k \times k$ di A, e supponiamo che valgano le ipotesi del teorema degli orlati. Cioè $\det(M) \neq 0$, e ogni orlato di M ha determinante nullo. Dimostriamo che $\mathrm{rg}(A) = k$.

Sappiamo che il rango di A non cambia riordinando le righe, e per lo stesso motivo anche riordinando le colonne. Per cui possiamo riordinare prima di tutto le righe e le colonne di A in modo che M coinvolga proprio le prime k righe e le prime k colonne di A. Otteniamo una matrice A' "a blocchi":

$$A' = \begin{pmatrix} M & B \\ C & D \end{pmatrix}$$

dove chiamiamo B, C e D i "blocchi" delle altre entrate di A'. Ricordiamo: $\mathrm{rg}(A') = \mathrm{rg}(A)$.

Eseguiamo ora l'algoritmo di Gauß alla matrice M. Visto che M è invertibile, otteniamo una matrice triangolare superiore con tutti i pivot $p_1, \ldots, p_k$ sulla diagonale:

$$M' = \begin{pmatrix} p_1 & m_{1,2} & m_{1,3} & \ldots & m_{1,k} \\ 0 & p_2 & m_{2,3} & \ldots & m_{2,k} \\ 0 & 0 & p_3 & \ldots & m_{3,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & p_k \end{pmatrix}$$

Ora consideriamo di nuovo A', e applichiamo *le stesse operazioni di riga che avevamo fatto per ridurre a scalini* M, applicate alle righe "complete" di A' invece che alle righe "troncate" di M. Otteniamo una matrice a blocchi

$$A'' = \begin{pmatrix} M' & B' \\ C & D \end{pmatrix}$$

Osserviamo che B è probabilmente cambiata in un'altra matrice B', perchè abbiamo applicato a tutta A' le operazioni di riga che facevamo su M, ma C e D non sono cambiate, perchè queste operazioni di riga coinvolgono solo le prime k righe.

Usando i pivot di M', possiamo usare ancora le operazioni di riga su A'' e mettere a zero tutte le entrate di C. Otteniamo una matrice

$$A''' = \begin{pmatrix} M' & B' \\ 0 & D' \end{pmatrix}$$

dove per "0" intendiamo un "blocco", delle stesse dimensioni di C, dove tutte le entrate sono nulle.

Anche qui probabilmente D è cambiata, diventando un'altra matrice D'.

Attenzione: quello che vogliamo far vedere è che D' è *già la matrice nulla*. Supponiamo che non lo sia, e consideriamo un'entrata a di D' con $a \neq 0$, in modo che a sia il più a sinistra possibile fra le entrate non nulle di D' (come se stessimo continuando l'algoritmo di Gauß).

Consideriamo N, orlato $(k+1) \times (k+1)$ di M' in A''' ottenuto considerando le prime k righe, le prime k colonne, più la riga e la colonna dove compare a.

Visto che a è il più a sinistra possibile fra le entrate non nulle di D', ci sono tutti zeri alla sua sinistra. Per cui il minore N ha la forma seguente:

$$N = \begin{pmatrix} p_1 & m_{1,2} & m_{1,3} & \dots & m_{1,k} & * \\ 0 & p_2 & m_{2,3} & \dots & m_{2,k} & * \\ 0 & 0 & p_3 & \dots & m_{3,k} & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & p_k & * \\ 0 & 0 & 0 & \dots & 0 & a \end{pmatrix}$$

e allora $\det(N) = p_1 \cdot p_2 \cdot \ldots \cdot p_k \cdot a$ è diverso da zero.

Ora, A''' era ottenuta da A prima riordinando delle righe e delle colonne, e poi facendo certe operazioni di riga. Le righe e le colonne di A''' che coinvolgono il minore N corrispondono perciò a certe righe e certe colonne di A (quelle prima del riordinamento). Esse determinano un minore $\widetilde{N}$ che è un orlato di M in A (così come N è un orlato di M' in A''').

Considerato come abbiamo usato l'algoritmo di Gauß, il minore N di A''' è ottenuto applicando le stesse operazioni di riga sul minore $\widetilde{N}$ di A. Quindi $\det(\widetilde{N})$ è anch'esso diverso da zero. Questo è impossibile, perché avrei trovato un orlato invertibile di M in A, cosa che è esclusa dalle ipotesi del teorema.

Quindi un'entrata come a non può esistere, e concludiamo che D' è la matrice tutta nulla. In altre parole, A''' ha le prime k righe non tutte nulle, ma quelle dopo sono tutte nulle. Quindi A''' è già a scalini, e il suo rango è k.

Visto che A''' ha lo stesso rango di A, segue la tesi del teorema: il rango di A è proprio k.

4. Risulati sulle basi di uno spazio vettoriale

Questa sezione riporta i risultati e le dimostrazioni viste a lezione, e può essere letta *al posto* delle sezioni 7, 8 e 8.1 della Parte 4 degli appunti del Prof. Savo.

Ricordiamo le definizioni seguenti.

(1) Dei vettori $v_1, \dots, v_k$ (eventualmente ripetuti) di uno spazio vettoriale V si dicono *linearmente indipendenti* se l'unico modo di scrivere il vettore nullo come combinazione lineare di questi vettori
$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k = O$$
è scegliendo tutti i coefficienti uguali a zero:
$$\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$$

(2) Dei vettori $v_1, \dots, v_k$ (eventualmente ripetuti) di uno spazio vettoriale V si dicono *generatori* di V se ogni vettore di V si può scrivere come combinazione lineare di $v_1, \dots, v_k$. In altre parole, sono generatori di V se per ogni $v \in V$ esistono coefficienti $\beta_1, \dots, \beta_k \in \mathbb{R}$ tali che
$$v = \beta_1 v_1 + \ldots + \beta_k v_k$$

(3) Una k-upla ordinata di vettori $(v_1, \dots, v_k)$ si dice *base* di V se i vettori $v_1, \dots, v_k$ sono linearmente indipendenti e sono generatori di V.

(4) Come caso limite, nelle definizioni precedenti si considera anche la lista "vuota" di vettori, cioè "nessun vettore". Si dice allora che l'*insieme vuoto* $\varnothing$ per definizione è linearmente indipendente.

(5) Sempre come caso limite, l'insieme vuoto per definizione è anche un insieme di generatori di V, se però V contiene solo il vettore nullo, cioè $V = \{O\}$. Allora $\varnothing$ è per definizione anche una base di[1] $V = \{O\}$.

[1]Ma $\varnothing$ non è base di alcun altro spazio vettoriale!

(6) Uno spazio vettoriale V si dice *finitamente generato* se esiste un numero finito di vettori che sono generatori di V.

Ricordiamo che nel corso assumiamo *implicitamente* che **tutti gli spazi vettoriali considerati siano finitamente generati**.

Osservazione 4.1. Siano $v_1, \ldots, v_k$ linearmente indipendenti.

(1) Ricordiamo che allora nessuno di questi vettori è combinazione lineare degli altri.
(2) Ricordiamo anche che fra questi vettori non può esserci il vettore nullo, perché se ad esempio $v_1 = O$, allora avremmo
$$1 \cdot v_1 + 0 \cdot v_2 + \ldots + 0 \cdot v_k = O$$
e avremmo scritto il vettore nullo come combinazione lineare di questi vettori, avendo non tutti i coefficienti nulli.

Osservazione 4.2. Siano ora $v_1, \ldots, v_k$ generatori di V, e sia $u \in V$ un vettore qualunque. Allora possiamo sicuramente aggiungerlo all'elenco, cioè anche $v_1, \ldots, v_k, u$ sono generatori di V. Infatti, se già ogni vettore si può scrivere come combinazione lineare di $v_1, \ldots, v_k$, a maggior ragione si può scrivere come combinazione lineare di $v_1, \ldots, v_k, u$.

In certi casi (ma non sempre!), si può anche *togliere* un vettore da un elenco di generatori, come afferma il lemma seguente.

Lemma 4.3. *Siano $v_1, \ldots, v_k$ generatori di V. Se uno di essi è combinazione lineare degli altri, mettiamo v_k, allora possiamo "scartarlo" dall'elenco. Cioè anche $v_1, \ldots, v_{k-1}$ sono generatori di V.*

Dimostrazione. Stiamo supponendo che esistano coefficienti $\gamma_1, \ldots, \gamma_{k-1} \in \mathbb{R}$ tali che
$$v_k = \gamma_1 v_1 + \ldots + \gamma_{k-1} v_{k-1}$$
Verifichiamo che $v_1, \ldots, v_{k-1}$ sono generatori di V. Sia $v \in V$ vettore qualsiasi. Sappiamo che $v_1, \ldots, v_k$ sono generatori, quindi esistono coefficienti $\beta_1, \ldots, \beta_k$ tali che
$$v = \beta_1 v_1 + \ldots + \beta_k v_k = \ldots$$
Sostituiamo l'espressione che abbiamo per v_k, ottenendo
$$\ldots = \beta_1 v_1 + \ldots + \beta_{k-1} v_{k-1} + \beta_k(\gamma_1 v_1 + \ldots + \gamma_{k-1} v_{k-1}) =$$
$$= (\beta_1 + \beta_k \gamma_1) v_1 + (\beta_2 + \beta_k \gamma_2) v_2 + \ldots + (\beta_{k-1} + \beta_k \gamma_{k-1}) v_{k-1}$$
Abbiamo cioè scritto v (qualsiasi) anche come combinazione lineare di $v_1, \ldots, v_{k-1}$. Quindi essi sono generatori di V. □

Corollario 4.4. *Dato un insieme di generatori di uno spazio vettoriale V, si possono scartare alcuni vettori se necessario, fino ad ottenere una base. In particolare, ogni spazio vettoriale ha almeno una base.*

Dimostrazione. Se i generatori non formano già una base, vuol dire che sono linearmente dipendenti, quindi almeno uno è combinazione lineare degli altri. Grazie al Lemma 4.3, possiamo scartarlo. Proseguiamo, finché non otteniamo un insieme finito di generatori linearmente indipendenti, cioè (messi in una n-upla) una base. □

Discutiamo ora il procedimento "opposto", cioè la possibilità di *aggiungere* vettori ad un elenco, fino ad ottenere una base.

Lemma 4.5. *Siano $v_1, \ldots, v_k$ vettori linearmente indipendenti di uno spazio vettoriale V. Allora si possono aggiungere all'elenco altri vettori, se necessario, fino ad ottenere una base. (Si dice anche "completarli ad una base".)*

Dimostrazione. Se $v_1, \ldots, v_k$ sono generatori di V allora abbiamo già finito. Altrimenti, consideriamo una base $\mathcal{B} = (w_1, \ldots, w_n)$ di V, e in particolare w_1. Se è combinazione lineare di $v_1, \ldots, v_k$, allora lo ignoriamo. Invece, se non è combinazione lineare di $v_1, \ldots, v_k$, lo aggiungiamo alla lista, ponendo $v_{k+1} = w_1$.

In questo secondo caso, i vettori $v_1, \ldots, v_k, v_{k+1}$ sono ancora linearmente indipendenti. Infatti, se una loro combinazione lineare è uguale al vettore nullo:

$$\alpha_1 v_1 + \ldots + \alpha_k v_k + \alpha_{k+1} v_{k+1} = O$$

allora sappiamo che α_{k+1} è zero, perché altrimenti il vettore $v_{k+1} = w_1$ sarebbe combinazione lineare degli altri vettori (col solito trucco di portare gli altri dall'altro lato e dividere tutto per α_{k+1}), e siamo supponendo che non sia così. Quindi in effetti $\alpha_{k+1} = 0$, allora la combinazione lineare che abbiamo in realtà è

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = O$$

e concludiamo che tutti i coefficienti sono nulli, perché $v_1, \ldots, v_k$ sono linearmente indipendenti. Quindi effettivamente $v_1, \ldots, v_{k+1}$ sono linearmente indipendenti.

Proseguiamo con lo stesso procedimento applicato a w_2, w_3, eccetera. Alla fine arriviamo ad un elenco "esteso" di vettori $v_1, \ldots, v_k, \ldots, v_m$, ancora linearmente indipendenti. Ogni vettore della base $\mathcal{B}$ o è in quest'elenco perché lo abbiamo aggiunto, oppure non è in elenco perché lo abbiamo ignorato, e questo perché era combinazione lineare di vettori che già avevamo.

In ogni caso, ogni vettore di $\mathcal{B}$ si scrive come combinazione lineare di $v_1, \ldots, v_m$. Visto che ogni vettore di V si scrive come combinazione lineare di vettori di $\mathcal{B}$, concludiamo che ogni vettore di V si scrive come combinazione lineare di $v_1, \ldots, v_m$. In altre parole, $(v_1, \ldots, v_m)$ è una base di V. $\square$

Siamo pronti a dare il prossimo teorema, che è il risultato principale di questa sezione.

Teorema 4.6. *Siano $(v_1, \ldots, v_k)$ e $(w_1, \ldots, w_m)$ basi di uno stesso spazio vettoriale V. Allora $k = m$, cioè esse sono formate dallo stesso numero di vettori.*

Dimostrazione. Supponiamo per assurdo che $k > m$. Cioè la prima delle due basi è in effetti

$$(v_1, \ldots, v_m, v_{m+1}, \ldots, v_k)$$

Osserviamo che $w_1, \ldots, w_m$ sono generatori di V, quindi in particolare anche il vettore v_1 si può scrivere come loro combinazione lineare. Cioè esistono coefficienti $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ tali che

$$v_1 = \alpha_1 w_1 + \ldots + \alpha_m w_m$$

I coefficienti $\alpha_1, \ldots, \alpha_m$ non possono essere tutti nulli, perché se lo fossero avremmo $v_1 = O$, e questo sarebbe in contraddizione col fatto che $(v_1, \ldots, v_k)$ è una base (i suoi vettori devono essere linearmente indipendenti).

Quindi almeno uno fra $\alpha_1, \ldots, \alpha_m$ è diverso da zero. Sia ad esempio $\alpha_1 \neq 0$ (se è un altro, basta riordinare i vettori). Segue

$$w_1 = \frac{1}{\alpha_1} v_1 - \frac{\alpha_2}{\alpha_1} w_2 - \ldots - \frac{\alpha_m}{\alpha_1} w_m$$

Ora, i vettori $v_1, w_1, \ldots, w_m$ sono generatori di V (vedi Osservazione 4.2). D'altronde, grazie al Lemma 4.3, il vettore w_1 si può scartare da $v_1, w_1, \ldots, w_m$. Cioè anche $v_1, w_2, \ldots, w_m$ sono generatori di V.

Proseguiamo usando v_2 invece di v_1: visto che $v_1, w_2, \ldots, w_m$ sono generatori di V, possiamo scrivere v_2 come loro combinazione lineare:

$$v_2 = \beta_1 v_1 + \beta_2 w_2 + \beta_3 w_3 + \ldots + \beta_m w_m$$

per certi coefficienti $\beta_1, \ldots, \beta_m \in \mathbb{R}$. Come prima, non tutti questi coefficienti possono essere nulli, altrimenti avremmo $v_2 = O$ (impossibile).

In più, affermiamo che β_1 non può essere l'*unico* non nullo. Infatti, se invece β_1 fosse proprio l'unico non nullo, allora avremmo

$$v_2 = \beta_1 v_1$$

e questo sarebbe in contraddizione col fatto che $(v_1, \ldots, v_k)$ è una base di V (i suoi vettori devono essere linearmente indipendenti). Per cui, almeno uno fra $\beta_2, \beta_3, \ldots, \beta_m$ è non nullo. Sia ad esempio $\beta_2 \neq 0$ (come prima, se è un altro, basta riordinare i vettori). Allora

$$w_2 = -\frac{\beta_1}{\beta_2} v_1 + \frac{1}{\beta_2} v_2 - \frac{\beta_3}{\beta_2} w_3 - \ldots - \frac{\beta_m}{\beta_2} w_m$$

Come prima, aggiungiamo v_2 alla lista, ottenendo che $v_1, v_2, w_2, w_3, \ldots, w_m$ sono generatori di V, e usiamo il Lemma 4.3 per scartare w_2. Concludiamo che anche i vettori $v_1, v_2, w_3, w_4, \ldots, w_m$ sono generatori di V.

Possiamo proseguire nello stesso modo, rimpiazzando vettori nell'elenco "dei w" con vettori dall'elenco "dei v". Il procedimento finisce quando "non abbiamo più w", cioè fino ad ottenere $v_1, v_2, \dots, v_m$.

Come abbiamo visto, ad ogni passo otteniamo sempre un elenco di generatori, per cui $v_1, v_2, \dots, v_m$ sono generatori di V. Allora l'elemento successivo nella prima base, cioè v_{k+1}, deve essere combinazione lineare di $v_1, v_2, \dots, v_m$. Questo però contraddice il fatto che $(v_1, \dots, v_k)$ è una base (i suoi vettori devono essere linearmente indipendenti).

Per cui $k > m$ è impossibile. Con lo stesso argomento, invertendo i ruoli delle due basi, si dimostra che anche $k < m$ è impossibile, per cui $k = m$. □

Grazie al Teorema 4.6, possiamo dare la seguente definizione.

Definizione 4.7. Sia V spazio vettoriale. Il numero di vettori di una base qualsiasi di V si chiama *dimensione* di V, e si scrive $\dim(V)$.

Mettendo insieme quanto ottenuto qui sopra, concludiamo con una proposizione molto utile.

Proposizione 4.8. *Sia V spazio vettoriale, e siano $v_1, \dots, v_k$ vettori di V.*

(1) *Se $k > \dim(V)$, allora $v_1, \dots, v_k$ sono linearmente dipendenti.*
(2) *Se $k < \dim(V)$, allora $v_1, \dots, v_k$ non sono generatori di V.*
(3) *Se $k = \dim(V)$ e $v_1, \dots, v_k$ sono linearmente indipendenti, allora $(v_1, \dots, v_k)$ è una base di V.*
(4) *Se $k = \dim(V)$ e $v_1, \dots, v_k$ sono generatori di V, allora $(v_1, \dots, v_k)$ è una base di V.*

Dimostrazione. (1) Se fossero linearmente indipendenti, grazie al Lemma 4.5 si potrebbero completare se necessario ad una base, che però avrebbe più elementi della dimensione di V, il che è impossibile.

(2) Se fossero generatori di V, grazie al Corollario 4.4 si potrebbe scartarne qualcuno, se necessario, fino ad arrivare a una base. Questa base però avrebbe meno elementi della dimensione di V, il che è impossibile.

(3) Visto che sono linearmente indipendenti, grazie al Lemma 4.5 si possono completare se necessario ad una base. Ma questa non può avere più elementi della dimensione di V. Visto che già avevamo proprio quel numero di elementi, essi formavano già una base.

(4) Visto che sono generatori di V, grazie al Corollario 4.4 se ne possono scartare alcuni, se necessario, fino ad arrivare a una base. Ma questa non può avere meno elementi della dimensione di V. Visto che già avevamo proprio quel numero di elementi, essi formavano già una base. □

5. Dimostrazione della formula di Grassmann

Siano U e W sottospazi vettoriali di uno spazio vettoriale V. La formula afferma che allora

$$\dim(U + W) + \dim(U \cap W) = \dim(U) + \dim(W)$$

Vediamone una dimostrazione. Scriviamo $k = \dim(U \cap W)$, $n = \dim(U)$ e $m = \dim(W)$.

Fissiamo una base

$$(v_1, \dots, v_k)$$

di $U \cap W$. Visto che si tratta di vettori (in particolare) contenuti in U, e sono linearmente indipendenti, li possiamo completare se necessario ad una base di U

$$(v_1, \dots, v_k, u_{k+1}, \dots, u_n)$$

Osserviamo che $(v_1, \dots, v_k)$ potrebbe già essere una base di U, questo vorrebbe dire che non c'è stato bisogno di aggiungere alcun vettore "u". Questo non compromette il resto della dimostrazione.

Completiamo $(v_1, \dots, v_k)$ *anche* ad una base di W

$$(v_1, \dots, v_k, w_{k+1}, \dots, w_m)$$

Consideriamo ora

$$(v_1, \dots, v_k, u_{k+1}, \dots, u_n, w_{k+1}, \dots, w_m)$$

Si tratta di $k + (n - k) + (m - k)$ vettori, cioè di $n + m - k$ vettori. Se dimostriamo che è una base di $U + W$, avremo dimostrato la formula di Graßmann.

(1) Dimostriamo che $v_1, \dots, v_k, u_{k+1}, \dots, u_n, w_{k+1}, \dots, w_m$ sono generatori di $U+W$. Questo è ovvio, perché sono stati ottenuti mettendo insieme dei generatori di U con dei generatori di W, semplicemente evitando di scrivere due volte i vettori "ripetuti" $v_1, \dots, v_k$.

(2) Dimostriamo che $v_1, \dots, v_k, u_{k+1}, \dots, u_n, w_{k+1}, \dots, w_m$ sono linearmente indipendenti. Supponiamo allora di avere dei coefficienti in modo che

$$\alpha_1 v_1 + \ldots + \alpha_k v_k + \beta_{k+1} u_{k+1} + \ldots + \beta_n u_n + \gamma_{k+1} w_{k+1} + \ldots + \gamma_m w_m = O \tag{5.1}$$

e dimostriamo che l'unica possibilità è avere tutti i coefficienti nulli. Riscriviamo l'uguaglianza come

$$\alpha_1 v_1 + \ldots + \alpha_k v_k + \beta_{k+1} u_{k+1} + \ldots + \beta_n u_n = -\gamma_{k+1} w_{k+1} - \ldots - \gamma_m w_m$$

Quest'uguaglianza vuol dire che abbiamo scritto lo <u>stesso</u> vettore (visto che c'è l'uguaglianza!) in due modi: a sinistra è espresso come combinazione lineare di vettori di U, e a destra è espresso come combinazione lineare di vettori di W. Quindi questo vettore è contemporaneamente in U e in W, cioè è un elemento di $U \cap W$.

Lo possiamo riscrivere come combinazione lineare di $v_1, \dots, v_k$, cioè esistono dei coefficienti δ_i tali che

$$-\gamma_{k+1} w_{k+1} - \ldots - \gamma_m w_m = \delta_1 v_1 + \ldots + \delta_k v_k$$

Portiamo tutto da una parte, ottenendo

$$\delta_1 v_1 + \ldots + \delta_k v_k + \gamma_{k+1} w_{k+1} + \ldots + \gamma_m w_m = O$$

Ora ricordiamo che $v_1, \dots, v_k, w_{k+1}, \dots, w_m$ sono linearmente indipendenti, per cui l'unica possibilità è avere i coefficienti nulli. In altre parole otteniamo $\delta_1 = \ldots = \delta_k = 0$ e soprattutto

$$\gamma_{k+1} = \ldots = \gamma_m = 0$$

cosa che già va nella direzione che volevamo. Allora riscriviamo l'uguaglianza (5.1), tenendo conto che gli ultimi coefficienti sono tutti nulli:

$$\alpha_1 v_1 + \ldots + \alpha_k v_k + \beta_{k+1} u_{k+1} + \ldots + \beta_n u_n = O$$

Ora, ricordiamo che $v_1, \dots, v_k, u_{k+1}, \dots, u_n$ sono linearmente indipendenti, per cui l'unica possibilità è che

$$\alpha_1 = \ldots = \alpha_k = 0, \quad \beta_{k+1} = \ldots = \beta_n = 0$$

Cioè abbiamo dimostrato che l'unico caso in cui l'uguaglianza (5.1) può essere vera è quella in cui tutti i coefficienti sono nulli. In altre parole, i vettori $v_1, \dots, v_k, u_{k+1}, \dots, u_n, w_{k+1}, \dots, w_m$ sono effettivamente linearmente indipendenti.

Questo conclude la dimostrazione.

6. Sulle coordinate di un vettore di $\mathbb{R}^n$ rispetto a una base

A lezione abbiamo fatto delle osservazioni semplici sulle coordinate di un vettore qualsiasi $v \in \mathbb{R}^n$ rispetto a una base data $\mathcal{B} = (v_1, \dots, v_n)$, non necessariamente uguale alla base canonica. Queste osservazioni non sono presenti sugli appunti di Savo (mi pare).

Sia $M = \mathrm{Mat}(v_1, \dots, v_n)$ la matrice che ha per colonne i vettori $v_1, \dots, v_n$. Scriviamo

$$v = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

e cerchiamo le coordinate di v rispetto a $\mathcal{B}$, cioè calcoliamo $\mathcal{F}_{\mathcal{B}}(v)$. Poniamo queste coordinate come incognite:

$$\mathcal{F}_{\mathcal{B}}(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Allora stiamo cercando una soluzione (in realtà l'unica soluzione) del sistema

$$x_1 v_1 + \ldots + x_n v_n = v$$

Abbiamo già osservato a lezione che questo è semplicemente un altro modo di scrivere il sistema seguente:

$$M \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \tag{6.1}$$

Ricordiamo che $v_1, \dots, v_n$ sono linearmente indipendenti, per cui M ha rango n, cioè è invertibile. Per cui il teorema di Cramer ci dice che

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = M^{-1} \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

In altre parole abbiamo

$$\mathcal{F}_{\mathcal{B}}(v) = M^{-1} \cdot v$$

Questa formula sarà utile in seguito. In linea di principio, si può usarla per calcolare le coordinate di v in pratica, solo che è necessario invertire la matrice M per farlo. Spesso è più veloce semplicemente risolvere il sistema (6.1) usando l'algoritmo di Gauß.

Però, notiamo anche che M non dipende da v, dipende solo da $\mathcal{B}$. Per cui, se si devono calcolare le coordinate di *molti* vettori rispetto alla stessa base $\mathcal{B}$, può essere più veloce calcolare M^{-1} una volta per tutte, invece di risolvere tanti sistemi diversi.

7. Sulle basi degli autospazi

Sia $f \colon V \to V$ endomorfismo di uno spazio vettoriale. Supponiamo f abbia almeno un autovalore, e siano $\lambda_1, \dots, \lambda_h$ i suoi autovalori (distinti). Siano $\mathcal{B}_1, \dots, \mathcal{B}_h$ basi rispettivamente degli autospazi $E(\lambda_1), \dots, E(\lambda_h)$.

In questa sezione dimostriamo che mettendo insieme tutti i vettori delle basi $\mathcal{B}_1, \dots, \mathcal{B}_h$ si ottiene una famiglia B di vettori linearmente indipendenti.

Procediamo per assurdo, cioè supponiamo che siano linearmente dipendenti, cioè supponiamo che esista una combinazione lineare di essi uguale al vettore nullo e tale che non tutti i coefficienti siano nulli, e cerchiamo di ottenere una contraddizione.

Scriviamo una combinazione lineare siffatta. Per semplicità, possiamo omettere gli addendi in cui il coefficiente è zero. Inoltre possiamo dare dei nomi semplici ai vettori coinvolti: anche se sono presi da basi diverse, diamo loro semplicemente lo stesso nome "v", li scriviamo in un ordine qualsiasi (indipendentemente dalla base da cui provengono), e mettiamo un indice che segua l'ordine in cui li scriviamo. Cioè la nostra combinazione lineare è

$$a_1 v_1 + \ldots + a_t v_t = O \tag{7.1}$$

dove $a_1, \dots, a_t$ sono tutti non nulli, e ciascun vettore v_i proviene da una qualsiasi delle basi $\mathcal{B}_1, \dots, \mathcal{B}_h$. Chiaramente sui v_i non ci sono restrizioni "di provenienza": vettori diversi possono provenire da basi diverse, oppure dalla stessa base.

Questo modo di assegnare nomi agli autovettori non tiene conto di qual è l'autovalore relativo a ciascun autovettore. Possiamo reintrodurre questa informazione, semplicemente rinominando *anche* gli autovalori. Diciamo cioè, per esempio, che α_i è l'autovalore dell'autovettore v_i. Chiaramente ciascun α_i sarà quindi uno dei numeri reali $\lambda_1, \dots, \lambda_h$, solo con un nome diverso. Inoltre osserviamo che fra i numeri $\alpha_1, \dots, \alpha_t$ potrebbero esserci anche ripetizioni, perché nulla vieta a più vettori fra i $v_1, \dots, v_t$ di provenire dalla stessa base, cioè di avere lo stesso autovalore.

Ora facciamo un'ipotesi cruciale. Consideriamo **tutte le possibili combinazioni lineari** di vettori di B che diano come risultato il vettore nullo e in cui non tutti i coefficienti siano nulli. La combinazione lineare (7.1) è una di esse, ma potrebbero essercene delle altre. Consideriamo il *numero di addendi*, cioè il numero di coefficienti non nulli che compaiono in ciascuna di queste combinazioni lineari. È un numero ≥ 1, che può variare a seconda della combinazione lineare considerata. Anche se sono possibili molte combinazioni lineari diverse (in linea di principio anche infinite), ci sarà un numero **minimo** possibile di coefficienti non nulli.

Ecco: supponiamo che t sia proprio questo minimo, cioè che la nostra combinazione lineare (7.1) sia una di quelle che ha il minimo numero possibile di coefficienti non nulli.

Facciamo ora l'immagine tramite f di entrambi i membri della (7.1), e ricordiamo che $f(O) = O$. Otteniamo

$$f(a_1 v_1 + \ldots + a_t v_t) = O$$

cioè

$$a_1 f(v_1) + \ldots + a_t f(v_t) = O$$

Ricordiamo anche che $f(v_i) = \alpha_i v_i$ per le notazioni che abbiamo scelto, per cui otteniamo

$$a_1 \alpha_1 v_1 + \ldots + a_t \alpha_t v_t = O \tag{7.2}$$

Ora moltiplichiamo la (7.1) per α_t e sottraiamola dalla (7.2). È un'operazione che ricorda le operazioni elementari di riga in una matrice. Viene

$$(a_1 \alpha_1 v_1 + \ldots + a_t \alpha_t v_t) - \alpha_t (a_1 v_1 + \ldots + a_t v_t) = O - \alpha_t O = O$$

e raccogliendo

$$(\alpha_1 - \alpha_t) a_1 v_1 + (\alpha_2 - \alpha_t) a_2 v_2 + \ldots + (\alpha_{t-1} - \alpha_t) a_{t-1} v_{t-1} + (\alpha_t - \alpha_t) a_t v_t = O$$

cioè

$$(\alpha_1 - \alpha_t) a_1 v_1 + (\alpha_2 - \alpha_t) a_2 v_2 + \ldots + (\alpha_{t-1} - \alpha_t) a_{t-1} v_{t-1} = O$$

Questa è una combinazione lineare di elementi di B, che fa il vettore nullo. Non sappiamo se e quali suoi coefficienti siano non nulli, ma sicuramente il numero di coefficienti non nulli è *inferiore* a t. Visto che t era il minimo possibile, questa non può essere una delle combinazioni lineari che stavamo considerando. L'unica possibilità è che *tutti* i suoi coefficienti siano nulli.

Ricordiamo che $a_1, \ldots, a_t$ sono tutti non nulli, per cui deve valere

$$\alpha_1 - \alpha_t = \alpha_2 - \alpha_t = \ldots = \alpha_{t-1} - \alpha_t = 0$$

e cioè

$$\alpha_1 = \ldots = \alpha_t$$

In altre parole $\alpha_1, \ldots, \alpha_t$ sono tutti uguali. D'altronde, come avevamo già osservato, ripetizioni fra gli α_i non erano affatto escluse.

Però allora $v_1, \ldots, v_t$ hanno tutti lo stesso autovalore, cioè provengono dalla stessa base. Quindi sono linearmente indipendenti, e questo contraddice la (7.1): <u>assurdo</u>.

Abbiamo ottenuto un assurdo, ipotizzando che B fosse fatta di vettori linearmente dipendenti. Per cui è una famiglia di vettori linearmente indipendenti, che è quello che volevamo dimostrare.

8. Forme bilineari: introduzione

Nella parte del corso su algebra lineare abbiamo visto tecniche di risoluzione di sistemi lineari omogenei. Esse possono essere interpretate come dei cambiamenti (lineari) di coordinate, fatti in modo che le equazioni assumano una forma particolarmente semplice.

Ad esempio, l'equazione lineare omogenea in tre variabili

$$x - y + z = 0$$

ha spazio delle soluzioni di dimensione 2, e una base è data dai vettori

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Completiamoli a una base di $\mathbb{R}^3$, aggiungendo

$$u_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Le coordinate di un vettore

$$v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

rispetto alla base $\mathcal{B} = (u_1, u_2, u_3)$ sono i numeri x', y', z' tali che

$$v = x' u_1 + y' u_2 + z' u_3$$

Usando la teoria vista finora, si può scrivere la relazione fra le coordinate x, y, z di v nella base canonica, e le coordinate x', y', z' di v rispetto a $\mathcal{B}$. In questo caso è facile farlo anche esplicitamente senza usare la teoria: basta sostituire i vettori v, u_1, u_2, u_3 nell'espressione qui sopra. Si ottiene

$$\begin{cases} x &= x' + z' \\ y &= x' + y' \\ z &= y' \end{cases}$$

Ora, le soluzioni dell'equazione originaria $x - y + z = 0$ sono le combinazioni lineari di u_1 e u_2, quindi sono caratterizzate dalla proprietà $z' = 0$. Infatti, sostituendo x, y, z nell'equazione otteniamo

$$0 = (x' + z') - (x' + y') + y' = z'$$

Abbiamo trasformato quindi l'equazione da cui siamo partiti. Considerata nelle nuove coordinate, quelle rispetto alla base $\mathcal{B}$ invece che alla base canonica, diventa l'equazione molto più semplice $z' = 0$.

Il nostro obiettivo è riprodurre un procedimento simile, ma partendo da equazioni omogenee di secondo grado. Potremo usare l'algebra lineare, perché considereremo equazioni esprimibili mediante il prodotto fra matrici. Partiremo dal fatto che il prodotto scalare stesso è dato da una formula che è un polinomio di grado 2, ad esempio in $\mathbb{R}^2$:

$$\left\langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} z \\ t \end{pmatrix} \right\rangle = \begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} z \\ t \end{pmatrix} = xz + yt$$

Vorremo considerare "varianti" del prodotto scalare, date mettendo una matrice fissata "in mezzo" ai due vettori, cioè con una formula del tipo

$$\begin{pmatrix} x & y \end{pmatrix} \cdot A \cdot \begin{pmatrix} z \\ t \end{pmatrix}$$

dove in questo caso A è una matrice 2×2, e in generale sarà $n \times n$.

Studiando la matrice A e le sue proprietà, riusciremo a mostrare che i polinomi che considereremo (di secondo grado, omogenei, in più variabili) possono essere scritti in una forma semplice, in seguito a un cambio (lineare) di variabili.

Questo riduce lo studio degli insiemi delle soluzioni allo studio di una lista di poche equazioni "standard", le cui soluzioni sono abbastanza semplici da descrivere geometricamente, almeno in $\mathbb{R}^2$ ed $\mathbb{R}^3$.

Infine, lo studio di queste "varianti" del prodotto scalare si rivela di importanza fondamentale in fisica (es. in teoria della relatività e in meccanica Hamiltoniana).

9. Forme bilineari: definizione

Definizione 9.1. Un'applicazione $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ si dice *forma bilineare* se esiste una matrice $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ tale che

$$b(u, v) = u^t \cdot A \cdot v$$

(dove consideriamo u, v come vettori colonna, per cui u^t è semplicemente u scritto come vettore riga.)

Esempio 9.2. Sia $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definita come il prodotto scalare:

$$b(u, v) = \langle u, v \rangle$$

Allora b è una forma bilineare, dove la matrice A della definizione è la matrice identità.

Esempio 9.3. Sia

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

La matrice A definisce la forma bilineare $b\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ data da

$$b\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} z \\ t \end{pmatrix}\right) = \begin{pmatrix} x & y \end{pmatrix} \cdot A \cdot \begin{pmatrix} z \\ t \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} z \\ t \end{pmatrix} =$$

$$\begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} 2z - t \\ 3z + t \end{pmatrix} = 2xz - xt + 3yz + yt$$

Sia $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ una forma bilineare, e sia A come nella definizione. Scriviamo

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

Sia come al solito $(e_1, \dots, e_n)$ la base canonica di $\mathbb{R}^n$. Calcoliamo $b(e_1, e_1)$:

$$e_1^t \cdot A \cdot e_1 = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \cdot A \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix} = a_{11}$$

Possiamo calcolare allo stesso modo $b(e_i, e_j)$ per qualsiasi i e j, ad esempio

$$b(e_1, e_2) = e_1^t \cdot A \cdot e_2 = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} = a_{12}$$

Si ottiene facilmente

$$b(e_i, e_j) = a_{ij}$$

In altre parole, vale la formula

$$A = \begin{pmatrix} b(e_1, e_1) & \dots & b(e_1, e_n) \\ \vdots & \ddots & \vdots \\ b(e_n, e_1) & \dots & b(e_n, e_n) \end{pmatrix}$$

Deduciamo anche che *la forma bilineare b determina <u>univocamente</u> la matrice A*. Possiamo dare allora la seguente definizione.

Definizione 9.4. Sia b forma bilineare e A matrice come sopra. La matrice A si dice *matrice canonica* di b.

Esempio 9.5. Sia $b\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ l'applicazione data da

$$b\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} z \\ t \end{pmatrix}\right) = xz + 2xt - 4yz - yt$$

Questa b è una forma bilineare. Infatti, calcoliamo quella che dovrebbe essere la matrice canonica di b, se fossimo sicuri che b è una forma bilineare:

$$b(e_1, e_1) = b\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 1, \quad b(e_1, e_2) = b\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 2,$$

$$b(e_2, e_1) = b\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = -4, \quad b(e_2, e_2) = b\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = -1$$

e poniamo

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

Si verifica facilmente che in effetti

$$b\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} z \\ t \end{pmatrix}\right) = \begin{pmatrix} x & y \end{pmatrix} \cdot A \cdot \begin{pmatrix} z \\ t \end{pmatrix}$$

quindi b è una forma bilineare, e A è la sua matrice canonica.

Come per il prodotto scalare, si verifica facilmente che una forma bilineare b soddisfa le proprietà seguenti:

(1) per ogni $u_1, u_2, v \in \mathbb{R}^n$ si ha

$$b(u_1 + u_2, v) = b(u_1, v) + b(u_2, v)$$

(2) per ogni $u, v \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$ si ha

$$b(\lambda u, v) = \lambda b(u, v)$$

(3) per ogni $u, v_1, v_2 \in \mathbb{R}^n$ si ha

$$b(u, v_1 + v_2) = b(u, v_1) + b(u, v_2)$$

(4) per ogni $u, v \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$ si ha

$$b(u, \lambda v) = \lambda b(u, v)$$

cioè b è "lineare rispetto a ciascuna delle due variabili u, v". Non è difficile dimostrare che se un'applicazione $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ soddisfa queste quattro proprietà, allora è una forma bilineare. Il procedimento è lo stesso che si usa per dimostrare che un'applicazione lineare $\mathbb{R}^n \to \mathbb{R}^m$ è data dalla moltiplicazione a sinistra per una matrice $m \times n$.

10. Forme bilineari e cambiamento di base

Definizione 10.1. Sia $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ una forma bilineare, e sia $\mathcal{B} = (v_1, \ldots, v_n)$ una base di $\mathbb{R}^n$. Si dice *matrice di b rispetto alla base $\mathcal{B}$* la matrice

$$A = \begin{pmatrix} b(v_1, v_1) & \ldots & b(v_1, v_n) \\ \vdots & \ddots & \vdots \\ b(v_n, v_1) & \ldots & b(v_n, v_n) \end{pmatrix}$$

La matrice canonica di b quindi è semplicemente la matrice di b rispetto alla base canonica.

Proposizione 10.2. *Siano $b, \mathcal{B}$ e A come nella Definizione 10.1, siano $u, v \in \mathbb{R}^n$ e siano*

$$\mathcal{F}_{\mathcal{B}}(u) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}, \quad \mathcal{F}_{\mathcal{B}}(v) = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$$

le colonne delle coordinate di u e v nella base $\mathcal{B}$. Allora

$$b(u, v) = \mathcal{F}_{\mathcal{B}}(u)^t \cdot A \cdot \mathcal{F}_{\mathcal{B}}(v)$$

cioè, più esplicitamente, vale

$$b(u, v) = \begin{pmatrix} c_1 & \ldots & c_n \end{pmatrix} \cdot \begin{pmatrix} b(v_1, v_1) & \ldots & b(v_1, v_n) \\ \vdots & \ddots & \vdots \\ b(v_n, v_1) & \ldots & b(v_n, v_n) \end{pmatrix} \cdot \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$$

Dimostrazione. Abbiamo

$$u = c_1 v_1 + \ldots + c_n v_n, \quad v = d_1 v_1 + \ldots + d_n v_n$$

quindi

$$\begin{aligned} b(u, v) &= b(c_1 v_1 + \ldots + c_n v_n, v) = c_1 b(v_1, v) + \ldots + c_n b(v_1, v) = \\ &= \begin{pmatrix} c_1 & \ldots & c_n \end{pmatrix} \cdot \begin{pmatrix} b(v_1, v) \\ \vdots \\ b(v_n, v) \end{pmatrix} \end{aligned}$$

Calcoliamo il vettore colonna che compare come secondo fattore. Abbiamo

$$b(v_1, v) = b(v_1, d_1 v_1 + \ldots + d_n v_n) = d_1 b(v_1, v_1) + \ldots + d_n b(v_1, v_n),$$

e così fino a

$$b(v_n, v) = b(v_n, d_1 v_1 + \ldots + d_n v_n) = d_1 b(v_n, v_1) + \ldots + d_n b(v_n, v_n)$$

Ma allora

$$\begin{pmatrix} b(v_1, v) \\ \vdots \\ b(v_n, v) \end{pmatrix} = \begin{pmatrix} b(v_1, v_1) & \ldots & b(v_1, v_n) \\ \vdots & \ddots & \vdots \\ b(v_n, v_1) & \ldots & b(v_n, v_n) \end{pmatrix} \cdot \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$$

e quindi

$$b(u, v) = \begin{pmatrix} c_1 & \ldots & c_n \end{pmatrix} \cdot \begin{pmatrix} b(v_1, v_1) & \ldots & b(v_1, v_n) \\ \vdots & \ddots & \vdots \\ b(v_n, v_1) & \ldots & b(v_n, v_n) \end{pmatrix} \cdot \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$$

che è quello che volevamo dimostrare. □

Si calcola facilmente la matrice di una forma bilineare b rispetto ad una nuova base, conoscendo la matrice di passaggio da una base all'altra. Supponiamo di avere due basi, $\mathcal{B}$ e $\mathcal{B}'$, di $\mathbb{R}^n$, sia A la matrice di b rispetto alla base $\mathcal{B}$. Calcoliamo la matrice A' di b rispetto alla base $\mathcal{B}'$.

Dobbiamo trovare una matrice A' tale che

$$b(u, v) = \mathcal{F}_{\mathcal{B}'}(u)^t \cdot A' \cdot \mathcal{F}_{\mathcal{B}'}(v)$$

Sia M la matrice di passaggio da $\mathcal{B}$ a $\mathcal{B}'$, come nella sezione sui cambi di base per le applicazioni lineari. In altre parole, M è la matrice che ha per colonne le coordinate dei vettori di $\mathcal{B}'$ rispetto alla base $\mathcal{B}$:

$$M = \mathrm{Mat}(\mathcal{F}_{\mathcal{B}}(v_1'), \ldots, \mathcal{F}_{\mathcal{B}}(v_n'))$$

dove $\mathcal{B}' = (v_1', \ldots, v_n')$. Sappiamo allora che M permette di passare facilmente dalle coordinate di un vettore qualsiasi $u \in \mathbb{R}^n$ rispetto alla base $\mathcal{B}'$ a quelle rispetto a $\mathcal{B}$:

$$\mathcal{F}_{\mathcal{B}}(u) = M \cdot \mathcal{F}_{\mathcal{B}'}(u)$$

Inoltre sappiamo che

$$b(u, v) = \mathcal{F}_{\mathcal{B}}(u)^t \cdot A \cdot \mathcal{F}_{\mathcal{B}}(v)$$

e quindi

$$b(u, v) = (M \cdot \mathcal{F}_{\mathcal{B}'}(u))^t \cdot A \cdot M \cdot \mathcal{F}_{\mathcal{B}'}(v) = \mathcal{F}_{\mathcal{B}'}(u)^t \cdot M^t \cdot A \cdot M \cdot \mathcal{F}_{\mathcal{B}'}(v)$$

Per cui

$$A' = M^t \cdot AM$$

Osserviamo che questa ricorda la formula che lega le matrici di un'applicazione lineare rispetto a due basi diverse. La differenza è che qui si usa M^t invece di M^{-1}.

Definizione 10.3. Due matrici A e A', entrambe $n \times n$, si dicono *congruenti* se esiste una matrice invertibile M tale che

$$A' = M^t \cdot A \cdot M$$

11. Forme bilineari non degeneri

Definizione 11.1. Sia $b \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ forma bilineare, con matrice canonica A. La forma bilineare b si dice *degenere* se $\det(A) = 0$, si dice *non degenere* se invece $\det(A) \neq 0$.

Osservazione 11.2. La matrice di b rispetto ad un'altra base qualsiasi $\mathcal{B}'$ è della forma $M^t AM$, dove M è la matrice del cambiamento di base dalla base canonica a $\mathcal{B}'$. Visto che $\det(M) \neq 0$ e che $\det(M^t) = \det(M)$, abbiamo

$$\det(A') = \det(M)^2 \cdot \det(A)$$

per cui A ha determinante nullo se e solo se A' ha determinante nullo. Quindi una forma bilineare è degenere se la sua matrice rispetto a una base <u>qualsiasi</u> ha determinante nullo.

Notiamo che, data una matrice A, quadrata $n \times n$, possiamo considerare **contemporaneamente** la forma bilineare $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ di matrice canonica A, e l'endomorfismo $f\colon \mathbb{R}^n \to \mathbb{R}^n$ di matrice canonica A.

Notiamo che b è non degenere se e solo se A è invertibile (per definizione), e questo è anche equivalente al fatto che f sia un isomorfismo.

Al contrario, b è degenere se e solo se f non è un isomorfismo. Ricordiamo che allora vale $\mathrm{Ker}(f) \neq \{O\}$ (e anche $\mathrm{Im}(f) \neq \mathbb{R}^n$).

È utile considerare anche un secondo endomorfismo $g\colon \mathbb{R}^n \to \mathbb{R}^n$, quello che ha matrice canonica A^t.

Allora osserviamo che, dati $u, v \in \mathbb{R}^n$, vale

$$b(u,v) = u^t \cdot \underbrace{A \cdot v}_{=f(v)} = u^t \cdot f(v)$$

e

$$b(u,v) = \underbrace{u^t \cdot A}_{=g(u)^t} \cdot v = g(u)^t \cdot v$$

Proposizione 11.3. *Consideriamo una forma bilineare $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Le seguenti affermazioni sono equivalenti:*

(1) *b è degenere,*
(2) *esiste un vettore $u_0 \in \mathbb{R}^n$, diverso dal vettore nullo, tale che $b(u_0, v) = 0$ per ogni $v \in \mathbb{R}^n$,*
(3) *esiste $v_0 \in \mathbb{R}^n$, diverso dal vettore nullo, tale che $b(u, v_0) = 0$ per ogni $u \in \mathbb{R}^n$.*

Dimostrazione. Dimostriamo che (1) implica (3), usando le formule viste qui sopra. Visto che b è degenere, allora $\mathrm{Ker}(f)$ contiene almeno un vettore non nullo v_0. Per ogni $u \in \mathbb{R}^n$ vale

$$b(u, v_0) = u^t \cdot f(v_0) = u^t \cdot O = 0$$

Allo stesso modo si dimostra che (1) implica (2): basta osservare che se b è degenere anche $\mathrm{Ker}(g)$ contiene almeno un vettore non nullo u_0, e per ogni $v \in \mathbb{R}^n$ vale

$$b(u_0, v) = g(u)^t \cdot v = O \cdot v = 0$$

Viceversa, dimostriamo che (3) implica (1). Supponiamo che esista v_0 non nullo tale che $b(u, v_0) = 0$ per ogni $u \in \mathbb{R}^n$. Scriviamo

$$v_0 = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

e

$$Av_0 = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$$

Sapendo che $b(u, v_0) = 0$ qualsiasi sia u, calcoliamo Av_0 usando i vettori della base canonica $e_1, \ldots, e_n$ al posto di u:

$$0 = b(e_1, v_0) = \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} \cdot \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = d_1$$

Allo stesso modo si dimostra $d_2 = \ldots = d_n = 0$. Quindi $Av_0 = O$, cioè $v_0 \in \mathrm{Ker}(f)$. Ma allora f non è un isomorfismo, e deduciamo che $\det(A) = 0$.

Ugualmente si dimostra che (2) implica (1): se esiste u_0 tale che $b(u_0, v) = 0$ per qualsiasi v, si calcola $u^t A$ usando $e_1, \ldots, e_n$ al posto di v, e si ottiene $u^t A = O$, cioè $g(u)^t = O$. Quindi g non è un isomorfismo, e $\det(A^t) = \det(A)$ deve essere 0. □

Esempio 11.4. La forma bilineare b di matrice canonica

$$A = \begin{pmatrix} 1 & -3 \\ -2 & 6 \end{pmatrix}$$

è degenere. Troviamo $v_0 \neq O$ tale che $b(u, v_0) = 0$ per ogni u. Per la dimostrazione qui sopra, basta trovare v_0 tale che $Av_0 = O$. Si tratta cioè di risolvere un sistema omogeneo di due equazioni in due incognite (le entrate di v_0), e prendere una soluzione non tutta nulla. Ad esempio

$$v_0 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Allora, dato un vettore qualunque

$$u = \begin{pmatrix} x \\ y \end{pmatrix}$$

abbiamo

$$b(u, v_0) = \begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ -2 & 6 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Troviamo invece u_0 non nullo tale che $b(u_0, v) = 0$ per ogni v. Dobbiamo trovare u_0 non nullo tale che $u_0^t A = O$, ad esempio

$$u_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Allora, dato un vettore qualunque

$$v = \begin{pmatrix} z \\ t \end{pmatrix}$$

abbiamo

$$b(u_0, v) = \begin{pmatrix} 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ -2 & 6 \end{pmatrix} \cdot \begin{pmatrix} z \\ t \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Osserviamo che vale $b(u_0, u_0) = 0$ e $b(v_0, v_0) = 0$.

Esempio 11.5. Sia b una forma bilineare. Anche se b è non degenere, è possibile che esistano vettori $u \in \mathbb{R}^n$ non nulli tali che $b(u, u) = 0$. Ad esempio, se b ha matrice canonica (con determinante non nullo)

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

allora sappiamo subito che $b(e_1, e_1) = 0$ (l'angolo in alto a sinistra di A è sempre uguale a $b(e_1, e_1)$).

12. Forme bilineari simmetriche e antisimmetriche

Definizione 12.1. Una forma bilineare $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ si dice *simmetrica* se la sua matrice canonica è una matrice simmetrica, e si dice *antisimmetrica* se la sua matrice canonica è antisimmetrica.

Proposizione 12.2. *Sia b una forma bilineare. Allora b è simmetrica se e solo se, per ogni $u, v \in \mathbb{R}^n$, vale*

$$b(u, v) = b(v, u)$$

Inoltre b è antisimmetrica se e solo se, per ogni $u, v \in \mathbb{R}^n$, vale

$$b(u, v) = -b(v, u)$$

Dimostrazione. Se $b(u, v) = b(v, u)$ qualsiasi siano u e v, è evidente dalla definizione che la sua matrice canonica A è simmetrica.

Viceversa, supponiamo che A sia simmetrica, cioè $A = A^t$. Prendiamo due vettori qualsiasi $u, v \in \mathbb{R}^n$:

$$b(u, v) = u^t \cdot A \cdot v = u^t \cdot A^t \cdot v = \left(v^t \cdot A \cdot u\right)^t = v^t \cdot A \cdot u = b(v, u)$$

dove la penultima uguaglianza è vera perché $v^t \cdot A \cdot u$ è un numero reale, cioè una matrice 1×1, che perciò è uguale alla sua trasposta.

L'affermazione riguardo a b antisimmetrica si dimostra in modo simile. $\square$

13. Il teorema di Sylvester

Sia $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ forma bilineare simmetrica, con matrice canonica A.

Come già fatto prima, possiamo interpretare A anche come matrice canonica di un endomorfismo $f\colon \mathbb{R}^n \to \mathbb{R}^n$. Ricordiamo il teorema spettrale: essendo A una matrice simmetrica, l'endomorfismo f è diagonalizzabile, ed esiste una base **ortonormale** $\mathcal{B}'$ di $\mathbb{R}^n$ fatta di autovettori di f. Cioè esiste una matrice invertibile M tale che $M^t = M^{-1}$ e la matrice

$$D = M^{-1} \cdot A \cdot M$$

è diagonale.

A questo punto l'osservazione importante è che vale anche

$$D = M^t \cdot A \cdot M$$

per cui D è anche la matrice di b rispetto alla base $\mathcal{B}'$. Cioè

$$D = \begin{pmatrix} b(v_1', v_1') & \dots & b(v_1', v_n') \\ \vdots & \ddots & \vdots \\ b(v_n', v_1') & \dots & b(v_n', v_n') \end{pmatrix}$$

Per ogni v_i' sia λ_i il suo autovalore, cioè $f(v_i') = \lambda_i v_i'$. Se λ_i è diverso da 0, poniamo

$$w_i = \frac{v_i'}{\sqrt{|\lambda_i|}}$$

Se invece $\lambda_i = 0$, allora poniamo

$$w_i = v_i'$$

Visto che abbiamo solo riscalato alcuni vettori di $\mathcal{B}'$ per dei fattori non nulli, sappiamo che anche $\mathcal{C} = (w_1, \dots, w_n)$ è una base di $\mathbb{R}^n$.

Non è difficile vedere che la matrice di f rispetto alla base $\mathcal{C}$ è sempre la matrice D. Calcoliamo invece la matrice di b rispetto a $\mathcal{C}$.

Dato che D è la matrice diagonale che ha sulla diagonale principale i numeri $\lambda_1, \dots, \lambda_n$, sappiamo che $b(v_i', v_i') = \lambda_i$, e $b(v_i', v_j') = 0$ se $i \neq j$. Allora la matrice di b nella base $\mathcal{C}$ è la matrice diagonale che ha sulla diagonale principale i numeri $b(w_1, w_1), \dots, b(w_n, w_n)$.

Ora, se λ_i è positivo, abbiamo

$$b(w_i, w_i) = b\left(\frac{v_i'}{\sqrt{|\lambda_i|}}, \frac{v_i'}{\sqrt{|\lambda_i|}}\right) = \frac{b(v_i', v_i')}{|\lambda_i|} = \frac{\lambda_i}{|\lambda_i|} = 1$$

Se invece λ_i è negativo, abbiamo

$$b(w_i, w_i) = b\left(\frac{v_i'}{\sqrt{|\lambda_i|}}, \frac{v_i'}{\sqrt{|\lambda_i|}}\right) = \frac{b(v_i', v_i')}{|\lambda_i|} = \frac{\lambda_i}{|\lambda_i|} = -1$$

Se infine $\lambda_i = 0$, abbiamo

$$b(w_i, w_i) = b(v_i', v_i') = 0$$

Quindi la matrice di b rispetto alla base $\mathcal{C}$ è una matrice diagonale, con sulla diagonale numeri (eventualmente diversi fra loro) che sono 1, oppure -1, oppure 0. Abbiamo tanti 1 quanti sono i λ_i positivi, tanti -1 quanti sono i λ_i negativi, e tanti 0 quanti sono i λ_i nulli. Osserviamo che ci sono tanti zeri sulla diagonale quant'è la dimensione dell'autospazio $E(0)$, cioè $\mathrm{Ker}(f)$.

Possiamo scrivere la matrice a "blocchi":

$$\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

dove p è il numero di 1 sulla diagonale, q è il numero di -1 sulla diagonale, la matrice I_p è come al solito la matrice identità $p \times p$, e il blocco in basso a destra è una sottomatrice nulla $r \times r$.

Nella prossima definizione riformuliamo la stessa cosa, contando gli autovalori distinti, ma ciascuno con la sua molteplicità geometrica.

Definizione 13.1. Sia b forma bilineare simmetrica come sopra, e A la sua matrice canonica. Si dice *segnatura* di A (e di b) la coppia (p, q) dove p è il numero di autovalori positivi distinti della matrice canonica di b (ciascuno contato tante volte quant'è la sua molteplicità geometrica), q è il numero di autovalori distinti negativi (ciascuno contato tante volte quant'è la sua molteplicità geometrica).

In conclusione, dal nostro ragionamento otteniamo il teorema seguente.

Teorema 13.2 (di Sylvester)**.** *Sia $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ forma bilineare simmetrica, e sia (p, q) la sua segnatura. Allora esiste una base ortogonale di $\mathbb{R}^n$ tale che la matrice di b rispetto a questa base è diagonale, della forma*

$$\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Osservazione 13.3. Attenzione: è possibile che non ci siano λ_i uguali a zero. Se ci sono, allora b è degenere, altrimenti b è non degenere. Se non ci sono, allora $p + q = n$ e la matrice di b è

$$\begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$$

È anche possibile che non ci siano autovalori positivi, o non ci siano autovalori negativi, ecc...

Concludiamo la sezione con una precisazione sui numeri p e q. In linea di principio è possibile che esista un'altra base in cui b ha matrice della stessa forma del teorema di Sylvester, dove però compaiano un numero diverso di entrate uguali a 1, oppure un numero diverso di entrate uguali a -1.

La definizione che abbiamo dato della segnatura (p, q) di b non è sufficiente a escludere questa possibilità. Vediamo allora nel seguente corollario il motivo per cui questo non può succedere.

Corollario 13.4. *Sia $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ forma bilineare simmetrica, di segnatura (p, q), e sia $\mathcal{B}$ base di $\mathbb{R}^n$ in cui b abbia matrice*

$$\begin{pmatrix} I_m & 0 & 0 \\ 0 & -I_l & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Allora m è la dimensione massima di un sottospazio di $\mathbb{R}^n$ tale che $b(u, u) > 0$ per ogni vettore u non nullo nel sottospazio. Inoltre l è la dimensione massima di un sottospazio di $\mathbb{R}^n$ tale che $b(u, u) < 0$ per ogni vettore u non nullo nel sottospazio. In particolare abbiamo $p = m$ e $q = l$.

Dimostrazione. Denotiamo $\mathcal{B} = (v_1, \ldots, v_n)$, e osserviamo che $U = L[v_1, \ldots, v_m]$ è un sottospazio di $\mathbb{R}^n$ tale che $b(u, u) > 0$ per ogni vettore non nullo $u \in U$ (il conto è identico al caso del prodotto scalare). Quindi la "dimensione massima" citata nel corollario è almeno m.

Sia U' un sottospazio per cui $b(u, u) > 0$ per ogni $u \in U'$ non nullo. Supponiamo per assurdo che $\dim(U') > m$, e consideriamo $W = L[v_{m+1}, v_{m+2}, \ldots, v_n]$. Scriviamo un vettore qualsiasi $w \in W$ usando i vettori della base data di W, cioè $w = c_{m+1}v_{m+1} + \ldots + c_n v_n$, e calcoliamo $b(w, w)$:

$$b(w, w) = c_{m+1}^2 b(v_{m+1}, v_{m+1}) + \ldots + c_n^2 b(v_n, v_n) \leq 0$$

Osserviamo che $\dim(W) = n - m$. Dalla formula di Graßmann abbiamo

$$\dim(U' \cap W) = \dim(U') + \dim(W) - \dim(U' + W) = \dim(U') + (n - m) - \dim(U' + W) > 0$$

dove la disuguaglianza alla fine vale perché $\dim(U') > m$ e $\dim(U' + W) \leq n$. Allora esiste un vettore non nullo u in $U' \cap W$, però $b(u, u)$ dovrebbe essere contemporaneamente > 0 (perché $u \in U'$) e ≤ 0 (perché $u \in W$). Assurdo. Quindi l'affermazione riguardo a m nel corollario è vera.

L'affermazione riguardo a l si dimostra in modo analogo, e l'ultima affermazione del corollario segue dal fatto che abbiamo espresso m e l in modo indipendente dalla base scelta. □

Esempio 13.5. Consideriamo la forma bilineare $b\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ di matrice canonica

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

Si tratta di una matrice simmetrica, per cui b è simmetrica. Gli autovalori di A sono 5, -1 e 0, le molteplicità geometriche sono tutte uguali a 1, e tre autovettori per i rispettivi autovalori sono

$$\begin{pmatrix}1\\2\\0\end{pmatrix}, \qquad \begin{pmatrix}0\\0\\1\end{pmatrix}, \qquad \begin{pmatrix}-2\\1\\0\end{pmatrix}$$

Sono già tutti ortogonali, perché sono autovettori di autovalori tutti diversi. Una base ortonormale di autovalori $\mathcal{B}' = (v_1', v_2', v_3')$ è ottenuta riscalando per la norma:

$$v_1' = \frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\\0\end{pmatrix}, \qquad v_2' = \begin{pmatrix}0\\0\\1\end{pmatrix}, \qquad v_3' = \frac{1}{\sqrt{5}}\begin{pmatrix}-2\\1\\0\end{pmatrix}$$

La matrice di b in questa base è diagonale, con gli autovalori sulla diagonale:

$$D = \begin{pmatrix}5 & 0 & 0\\0 & -1 & 0\\0 & 0 & 0\end{pmatrix}$$

Facciamo la verifica per v_1':

$$b(v_1', v_1') = \begin{pmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0\end{pmatrix} \cdot \begin{pmatrix}1 & 2 & 0\\2 & 4 & 0\\0 & 0 & -1\end{pmatrix} \cdot \begin{pmatrix}\frac{1}{\sqrt{5}}\\\frac{2}{\sqrt{5}}\\0\end{pmatrix} = \begin{pmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0\end{pmatrix} \cdot \begin{pmatrix}\frac{5}{\sqrt{5}}\\\frac{10}{\sqrt{5}}\\0\end{pmatrix} = 5$$

Visto che c'è un autovalore positivo e un autovalore negativo, la segnatura di b è $(1,1)$.

Per ottenere una matrice come nel teorema di Sylvester, dobbiamo riscalare opportunamente i vettori v_1', v_2', v_3'. Il fattore che usiamo è $1/\sqrt{|\lambda_i|}$ se l'autovalore λ_i è non nullo, invece lasciamo il vettore così com'è se l'autovalore λ_i è nullo. Otteniamo:

$$w_1 = \frac{1}{\sqrt{|5|}}v_1' = \frac{1}{\sqrt{5}}\begin{pmatrix}\frac{1}{\sqrt{5}}\\\frac{2}{\sqrt{5}}\\0\end{pmatrix} = \begin{pmatrix}\frac{1}{5}\\\frac{2}{5}\\0\end{pmatrix}$$

$$w_2 = \frac{1}{\sqrt{|-1|}}v_2' = \frac{1}{\sqrt{1}}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}0\\0\\1\end{pmatrix}$$

$$w_3 = v_3' = \begin{pmatrix}\frac{-2}{\sqrt{5}}\\\frac{1}{\sqrt{5}}\\0\end{pmatrix}$$

In questa base, si verifica facilmente che la matrice di b è

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

Definizione 13.6. Sia $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ forma bilineare simmetrica, di segnatura (p,q). Essa si dice

(1) *definita positiva* se $p = n$,
(2) *definita negativa* se $q = n$,
(3) *semidefinita positiva* se $q = 0$,
(4) *semidefinita negativa* se $p = 0$.

Notiamo che se b è definita positiva allora è per forza non degenere, e lo stesso vale se b è definita negativa. Dal Corollario 13.4 segue immediatamente il seguente.

Corollario 13.7. *Una forma bilineare $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ è definita positiva se e solo se $b(u,u) > 0$ per ogni $u \in \mathbb{R}^n$ non nullo. Inoltre b è definita negativa se e solo se $b(u,u) < 0$ per ogni $u \in \mathbb{R}^n$ non nullo.*

14. Forme quadratiche

Definizione 14.1. Sia $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ una forma bilineare simmetrica. Si dice *forma quadratica* associata a b l'applicazione $q\colon \mathbb{R}^n \to \mathbb{R}$ definita come

$$q(v) = b(v, v)$$

per ogni $v \in \mathbb{R}^n$.

Esempio 14.2. L'applicazione $q\colon \mathbb{R}^3 \to \mathbb{R}$ data da

$$q\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2$$

è la forma quadratica associata al prodotto scalare. La stessa cosa vale per ogni n: se $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ è il prodotto scalare, allora la forma quadratica associata è data da

$$q(v) = \|v\|^2$$

Esempio 14.3. L'applicazione $q\colon \mathbb{R}^2 \to \mathbb{R}$ data da

$$q\begin{pmatrix} x \\ y \end{pmatrix} = x^2 - y^2$$

è la forma quadratica associata alla forma bilineare simmetrica che ha matrice canonica

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

Esempio 14.4. L'applicazione $q\colon \mathbb{R}^2 \to \mathbb{R}$ data da

$$q\begin{pmatrix} x \\ y \end{pmatrix} = xy$$

è la forma quadratica associata alla forma bilineare simmetrica che ha matrice canonica

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

Osservazione 14.5. Consideriamo un polinomio di secondo grado omogeneo qualsiasi $p(x_1, x_2, \ldots, x_n)$, in n variabili. Allora esiste una forma bilineare b per cui p sia uguale alla forma quadratica q associata a b. Infatti p si può scrivere come

$$p(x_1, \ldots, x_n) = a_{11}x_1^2 + a_{12}x_1x_2 + \ldots + a_{nn}x_n^2$$

cioè chiamiamo a_{ij} il coefficiente davanti a x_ix_j. Facendo questo possiamo assumere che $i \leq j$ in tutti i monomi, visto che $x_ix_j = x_jx_i$.

Allora si vede immediatamente che basta prendere la forma bilineare simmetrica b che ha matrice canonica

$$A = \begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{12}}{2} & \cdots & \frac{a_{1n}}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{23}}{2} & \cdots & \frac{a_{2n}}{2} \\ \frac{a_{13}}{2} & \frac{a_{23}}{2} & a_{33} & \cdots & \frac{a_{3n}}{2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{a_{1n}}{2} & \frac{a_{2n}}{2} & \frac{a_{3n}}{2} & \cdots & a_{nn} \end{pmatrix}$$

Cioè, al posto di riga i e colonna j mettiamo $\frac{a_{ij}}{2}$ se $i < j$, mettiamo $\frac{a_{ji}}{2}$ se $i > j$, e mettiamo a_{ii} se $i = j$.

Mostriamo ora che, usando la forma quadratica q, si può risalire alla forma bilineare simmetrica b da cui proviene. Infatti, per la bilinearità di b, per ogni coppia di vettori u, v abbiamo

$$q(u+v) = b(u+v, u+v) = q(u) + q(v) + b(u, v) + b(v, u) = q(u) + q(v) + 2b(u, v)$$

cioè

$$b(u, v) = \frac{1}{2}\left(q(u+v) - q(u) - q(v)\right)$$

e questo esprime b in funzione di q (si chiama *formula di polarizzazione*). Ne deduciamo che q determina univocamente b.

Infine, applichiamo il teorema di Sylvester alle forme quadratiche. Esso mostra che data una forma quadratica su $\mathbb{R}^n$ se ne può trovare una "espressione canonica", in termini delle coordinate dei vettori rispetto ad una certa base di $\mathbb{R}^n$.

Infatti, applicando il teorema di Sylvester alla forma bilineare b di forma quadratica associata q, otteniamo:

Teorema 14.6. *Sia $q\colon \mathbb{R}^n \to \mathbb{R}$ una forma quadratica, e sia (p, q) la segnatura della forma bilineare corrispondente. Allora esiste una base ortonormale $\mathcal{B}$ di $\mathbb{R}^n$ tale che*

$$q(v) = X_1^2 + \ldots + X_p^2 - X_{p+1}^2 - \ldots - X_{p+q}^2$$

dove $X_1, \ldots, X_n$ sono le coordinate di v rispetto alla base $\mathcal{B}$.

Esempio 14.7. Sia $q\colon \mathbb{R}^3 \to \mathbb{R}$ la forma quadratica

$$q(x, y, z) = x^2 + 4xy + 4y^2 - z^2$$

La forma bilineare $b\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ corrispondente a q ha matrice canonica

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

cioè si tratta della stessa forma bilineare dell'Esempio 13.5. Applichiamo anche in questo caso il teorema di Sylvester: sappiamo che esiste una base $\mathcal{C}$ in cui q è data dalla formula $X_1^2 - X_2^2$, dove X_1, X_2, X_3 sono le coordinate rispetto $\mathcal{C}$.

Rendiamo esplicito il *cambio di coordinate* e verifichiamo che viene effettivamente questa formula semplice per q. Dobbiamo passare dalle coordinate rispetto alla base canonica $\mathcal{E}$ a quelle rispetto alla base nuova $\mathcal{C} = (w_1, w_2, w_3)$. Ricordiamo che

$$w_1 = \begin{pmatrix} \frac{1}{5} \\ \frac{2}{5} \\ 0 \end{pmatrix}, \qquad w_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad w_3 = \begin{pmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix}$$

Costruiamo la matrice M di passaggio da $\mathcal{E}$ a $\mathcal{C}$:

$$M = \begin{pmatrix} \frac{1}{5} & 0 & \frac{-2}{\sqrt{5}} \\ \frac{2}{5} & 0 & \frac{1}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix}$$

Scriviamo un vettore qualsiasi $v \in \mathbb{R}^3$ come un vettore colonna le cui entrate sono le incognite x, y, z:

$$v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Come al solito, denotiamo con $\mathcal{F}_\mathcal{C}(v)$ le coordinate di v rispetto a $\mathcal{C}$. Sappiamo dalla teoria che

$$v = M \cdot \mathcal{F}_\mathcal{C}(v)$$

Vogliamo considerare le coordinate di v rispetto a $\mathcal{C}$ come *nuove coordinate*, denotiamole per semplicità X, Y, Z invece di X_1, X_2, X_3. Allora

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & 0 & \frac{-2}{\sqrt{5}} \\ \frac{2}{5} & 0 & \frac{1}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

Otteniamo le seguenti espressioni di x, y, z in termini di X, Y, Z:

$$\begin{cases} x = \frac{1}{5}X - \frac{2}{\sqrt{5}}Z \\ y = \frac{2}{5}X + \frac{1}{\sqrt{5}}Z \\ z = Y \end{cases}$$

Ora possiamo scrivere $q(v)$ nelle nuove coordinate. Partiamo dall'espressione nelle vecchie coordinate

$$q(v) = x^2 + 4xy + 4y^2 - z^2 = \dots$$

e sostituiamo x, y, z con le loro espressioni nelle nuove coordinate

$$\dots = \left(\frac{1}{5}X - \frac{2}{\sqrt{5}}Z\right)^2 + 4\left(\frac{1}{5}X - \frac{2}{\sqrt{5}}Z\right)\left(\frac{2}{5}X + \frac{1}{\sqrt{5}}Z\right) + 4\left(\frac{2}{5}X + \frac{1}{\sqrt{5}}Z\right)^2 - Y^2 =$$
$$= X^2 - Y^2$$

Osservazione: abbiamo espresso qui le vecchie coordinate rispetto alle nuove. Se avessimo bisogno, per qualche motivo, di esprimere le nuove in funzione delle vecchie, partiremmo dall'espressione

$$v = M \cdot \mathcal{F}_{\mathcal{C}}(v)$$

da cui segue

$$M^{-1} \cdot v = \mathcal{F}_{\mathcal{C}}(v)$$

Ricordiamo che $\mathcal{C}$ non è necessariamente ortonormale (nel teorema di Sylvester, era $\mathcal{B}'$ la base sicuramente ortonormale, non $\mathcal{C}$). Quindi M^{-1} non è necessariamente uguale a M^t. Comunque vale

$$\begin{pmatrix} \frac{1}{5} & 0 & \frac{-2}{\sqrt{5}} \\ \frac{2}{5} & 0 & \frac{1}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

e visto che

$$\begin{pmatrix} \frac{1}{5} & 0 & \frac{-2}{\sqrt{5}} \\ \frac{2}{5} & 0 & \frac{1}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \end{pmatrix}$$

deduciamo le formule

$$\begin{cases} X &= x + 2y \\ Y &= z \\ Z &= -\frac{2}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y \end{cases}$$

15. Equazioni di secondo grado in $\mathbb{R}^2$

Come applicazione delle sezioni precedenti, vediamo come si possono studiare le soluzioni di un'equazione del tipo

$$p(x, y) = d$$

dove $p(x, y)$ è un polinomio omogeneo di grado 2, e $d \in \mathbb{R}$ è una costante.

Osservamo che $p(x, y)$ è una forma quadratica. Infatti, essendo omogeneo di grado 2, si può scrivere come

$$p(x, y) = ax^2 + bxy + cy^2$$

Questa è una forma quadratica. La forma bilineare simmetrica b corrispondente si trova come nell'Osservazione 14.5. In questo caso si trova la matrice canonica

$$A = \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}$$

Dal Teorema 14.6, sappiamo che esiste una base *ortonormale* $\mathcal{B}$ tale che, nelle coordinate x', y' rispetto alla base $\mathcal{B}$, la forma quadratica diventa come nel teorema.

Discutiamo le varie possibilità, a seconda della segnatura (p, q) di b. Ricordiamo che p, q sono interi non negativi, e $p + q \leq 2$ visto che siamo in $\mathbb{R}^2$.

(1) Se la segnatura è $(0, 0)$, allora A è la matrice nulla, e il polinomio p diventa il polinomio costantemente nullo

$$p(x', y') = 0$$

(Attenzione: il polinomio diventa il polinomio nullo, non l'equazione $p(x, y) = d$.)

Quindi: se $d = 0$ allora l'equazione diventa $0 = 0$, soddisfatta da tutti gli elementi di $\mathbb{R}^2$. Se invece $d \neq 0$, l'equazione è $0 = d$, non soddisfatta da alcun elemento di $\mathbb{R}^2$.

In conclusione, se la segnatura è $(0,0)$ allora l'insieme delle soluzioni è tutto $\mathbb{R}^2$, oppure è l'insieme vuoto.

(2) Supponiamo ora che la segnatura sia $(2,0)$. Allora
$$p(x',y') = (x')^2 + (y')^2$$
Se $d < 0$ l'equazione $p(x',y') = d$ non ha soluzioni. Se $d = 0$, ha come unica soluzione $x' = 0, y' = 0$, cioè un punto. Se $d > 0$, allora l'insieme delle soluzioni è la circonferenza di raggio $\sqrt{d}$. Questo ovviamente nelle variabili x', y'.

Nelle variabili x, y, sarà la circonferenza eventualmente dilatata, o "deformata", a seconda di come è fatto il passaggio di coordinate da x, y a x', y'.

(3) In modo simile si tratta il caso della segnatura $(0,2)$, perché la forma quadratica diventa
$$p(x',y') = -(x')^2 - (y')^2$$
e l'equazione diventa $-(x')^2 - (y')^2 = d$, cioè $(x')^2 + (y')^2 = -d$. Si ricade quindi nel caso precedente, con $-d$ al posto di d.

(4) Se la segnatura è $(1,0)$, allora
$$p(x',y') = (x')^2$$
e l'equazione diventa $(x')^2 = d$. Questo impone condizioni su x', ma non su y' che quindi può essere qualsiasi. Se $d < 0$ allora non abbiamo soluzioni. Se $d = 0$ allora $x' = 0$, perciò l'insieme delle soluzioni è la retta dei punti che hanno coordinata $x' = 0$ e coordinata y' qualsiasi.

Se $d > 0$ allora x' può essere $\sqrt{d}$ oppure $-\sqrt{d}$. L'insieme delle soluzioni dunque è formato da due rette: una retta è quella dei punti con coordinata x' uguale a $\sqrt{d}$ e coordinata y' qualsiasi. La seconda retta è quella dei punti con coordinata x' uguale a $-\sqrt{d}$, e coordinata y' qualsiasi.

(5) Se la segnatura è $(0,1)$, allora
$$p(x',y') = -(y')^2$$
e l'equazione diventa $(y')^2 = -d$. L'insieme delle soluzioni si può descrivere come nel caso precedente, solo che qui saranno scambiate le coordinate x' e y', e d andrà sostituito con $-d$.

(6) Rimane il caso in cui la segnatura è $(1,1)$. In questo caso la forma quadratica diventa
$$p(x',y') = (x')^2 - (y')^2$$
e l'equazione diventa $(x')^2 - (y')^2 = d$. Se $d = 0$, quest'equazione diventa
$$(x' + y')(x' - y') = 0$$
che è soddisfatta se $x' = y'$ oppure se $x' = -y'$. L'insieme delle soluzioni quindi è formato dall'unione di due rette: quella dei punti tali che $x' = y'$, e quella dei punti tali che $x' = -y'$.

Se invece $d \neq 0$, allora l'insieme delle soluzioni è un'iperbole.

16. Cenni di geometria affine in $\mathbb{R}^n$

Definizione 16.1. Un sottoinsieme $U \subseteq \mathbb{R}^n$ si dice *sottospazio affine* se esiste un elemento $\overline{u} \in U$ tale che l'insieme
$$U_0 = \{u - \overline{u} \mid u \in U\}$$
è un sottospazio vettoriale di $\mathbb{R}^n$. Si usa anche la notazione $U_0 = U - \overline{u}$.

Osservazione 16.2. Sia U un sottospazio affine di $\mathbb{R}^n$, e sia $\overline{\overline{u}}$ un elemento qualsiasi di U. Osserviamo che $w = \overline{\overline{u}} - \overline{u}$ è in $U_0 = U - \overline{u}$. Visto che U_0 è un sottospazio vettoriale, l'insieme
$$\{v - w \mid v \in U_0\}$$
è uguale a U_0 stesso. Quindi abbiamo
$$U_0 = \{u - \overline{u} \mid u \in U\} = \{u - \overline{u} - w \mid u \in U\} = \{u - \overline{\overline{u}} \mid u \in U\}.$$
In altre parole U_0 <u>non dipende</u> dall'elemento $\overline{u}$ di U.

Osservazione 16.3. Dall'osservazione precedente si deducono immediatamente le seguenti proprietà:

(1) se $u, u' \in U$ allora $u - u' \in U_0$;
(2) se $u \in U$ e $v \in U_0$ allora $u + v \in U$;
(3) U è esso stesso un sottospazio vettoriale di $\mathbb{R}^n$ se e solo se $O \in U$.

Definizione 16.4. Siano $U \subseteq \mathbb{R}^n$ ed U_0 come nella definizione precedente. Il sottospazio U_0 si dice *sottospazio vettoriale soggiacente* ad U. La *dimensione* di U è definita come $\dim(U_0)$. Inoltre

(1) U si dice *retta (affine)* se $\dim(U) = 1$,
(2) U si dice *piano (affine)* se $\dim(U) = 2$,
(3) U si dice *iperpiano (affine)* se $\dim(U) = n - 1$.

Esempio 16.5. Sia S un sistema di equazioni lineari in n variabili, eventualmente non omogeneo. Allora $U = \mathrm{Sol}(S)$ è un sottospazio affine di $\mathbb{R}^n$, **se S è compatibile**. Infatti, dato S_0 il sistema omogeneo associato ad S, e data una soluzione $\overline{u}$ di S, si verifica facilmente che $U - \overline{u}$ è l'insieme delle soluzioni di U_0.

Viceversa, dato un sottospazio affine non vuoto U di $\mathbb{R}^n$, esiste un sistema lineare S (eventualmente non omogeneo) che ha per soluzioni proprio U. Infatti, basta considerare un sistema omogeneo S_0 che abbia per soluzioni U_0. Fissato ora $\overline{u}$ in U, basta prendere come termini noti per S le equazioni di S_0 calcolate nelle entrate di $\overline{u}$.