6. Spazi euclidei ed hermitiani

6.1 In [GA] 5.4 abbiamo definito il prodotto scalare fra vettori di $\mathbb{R}^n$ (che d'ora in poi chiameremo **prodotto scalare standard** su $\mathbb{R}^n$) e abbiamo considerato le seguenti proprietà:

bilinearità	$\langle v, w+u\rangle = \langle v,w\rangle + \langle v,u\rangle, \langle v+u, w\rangle = \langle v,w\rangle + \langle u,w\rangle$ e $\langle cv, w\rangle = c\langle v,w\rangle = \langle v, cw\rangle, \quad \forall\ v,w,u \in \mathbb{R}^n, \forall\ c \in \mathbb{R}$
simmetria	$\langle v,w\rangle = \langle w,v\rangle, \quad \forall\ v,w \in \mathbb{R}^n$
definito positivo	$\langle v,v\rangle \geq 0\ \forall\ v \in \mathbb{R}^n$, e $\langle v,v\rangle = 0$ se e solo se $v = O$.

Ora vogliamo esaminare la situazione su uno spazio vettoriale qualsiasi: la terza proprietà ci imporrà di distinguere, ad un certo punto, fra spazi vettoriali reali e complessi.

6.2 Definizione. Sia V uno spazio vettoriale sul campo K. Una **forma bilineare** su V è una funzione $f : V \times V \to K$ che è lineare su ognuna delle due componenti separatamente, ovvero tale che, $\forall\ v,w,u \in V, \forall\ a,b \in K$

i) $f(av+bu, w) = af(v,w) + bf(u,w)$

ii) $f(v, aw+bu) = af(v,w) + bf(v,u)$.

f si dice **simmetrica** se $f(v,w) = f(w,v)\ \forall\ v,w \in V$.

6.3 Esempi. (1) La formula $f((x_1,x_2),(y_1,y_2)) = 3x_1y_1 + x_2y_2 - 2x_1y_2$ definisce una forma bilineare su $\mathbb{R}^2$; e $f(p,q) = \int_0^1 p(t)q(t)dt$ dà una forma bilineare simmetrica su $\mathbb{R}[t]$.

(2) Ogni matrice $A \in M(n \times n, K)$ dà luogo a una forma bilineare f_A su K^n in questo modo: se $X = {}^t(x_1,\ldots,x_n)$ e $Y = {}^t(y_1,\ldots,y_n)$, $f_A(X,Y) := {}^tXAY$.

In particolare, se $K = \mathbb{R}$ e $A = I_n$, f_A risulta essere il prodotto scalare standard su $\mathbb{R}^n$.

(3) Sia V uno spazio vettoriale di dimensione n sul campo K, e sia $\mathcal{B} = (v_1,\ldots,v_n)$ una sua base; ogni matrice $A \in M(n \times n, K)$ dà luogo a una forma bilineare f_A in questo modo: se $v,w \in V$, e $[v]_\mathcal{B} = X = {}^t(x_1,\ldots,x_n), [w]_\mathcal{B} = Y = {}^t(y_1,\ldots,y_n)$, allora $f_A(X,Y) := {}^tXAY$.

6.4 Proposizione. f_A è simmetrica se e solo se la matrice A è simmetrica.

Dimostrazione. Se ${}^tA = A$, si ha

$$f_A(X,Y) = {}^tXAY = {}^t({}^tXAY) = {}^tY({}^tA)X = {}^tYAX = f_A(Y,X)$$

(il secondo passaggio è dovuto al fatto che ${}^tXAY$ è una matrice 1×1). L'altra implicazione si ottiene calcolando f_A sui vettori della base canonica: $f_A(e_i,e_j) =$

${}^t e_i A e_j = a_{ij}$, ma anche $f_A(e_i, e_j) = f_A(e_j, e_i) = {}^t e_j A e_i = a_{ji}$, quindi A è simmetrica.

6.5 Proposizione. Sia V uno spazio vettoriale di dimensione n sul campo K, e sia $\mathcal{B} = (v_1, \ldots, v_n)$ una sua base; ogni forma bilineare simmetrica f è del tipo f_A se $A \in M(n \times n, K)$ è definita come $a_{ij} := f(v_i, v_j)$.

Scriveremo $A = M_{\mathcal{B}}(f)$, e diremo che A **rappresenta la forma bilineare f in base $\mathcal{B}$**, o che A è associata a f rispetto a $\mathcal{B}$.

Dimostrazione. Sia A la matrice definita da $a_{ij} := f(v_i, v_j)$, e siano $v, w \in V$ con $[v]_{\mathcal{B}} = X, [w]_{\mathcal{B}} = Y$. Allora per definizione si ha:

$f(v, w) = f(x_1 v_1 + \cdots + x_n v_n, y_1 v_1 + \cdots + y_n v_n) = \sum_1^n x_i f(v_i, y_1 v_1 + \cdots + y_n v_n) = \sum_{i,j=1}^n x_i y_j f(v_i, v_j) = \sum_{i,j=1}^n x_i y_j a_{ij} = {}^t X A Y = f_A(X, Y)$.

6.6 Come avviene per gli operatori, si pone il problema di confrontare le matrici associate a una forma bilineare f rispetto a due basi diverse $\mathcal{B}, \mathcal{B}'$ di V. Il legame è il seguente:

$$M_{\mathcal{B}'}(f) = {}^t P M_{\mathcal{B}}(f) P$$

dove $P = M(\mathcal{B}, \mathcal{B}')$ è la matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$; infatti siano $v, w \in V$: se poniamo $[v]_{\mathcal{B}} = X, [w]_{\mathcal{B}} = Y, [v]_{\mathcal{B}'} = X', [w]_{\mathcal{B}'} = Y', M_{\mathcal{B}}(f) = A, M_{\mathcal{B}'}(f) = A'$, risulta $X = PX', Y = PY'$ e dunque si ha

$$f(v, w) = {}^t X A Y = {}^t(PX') A (PY') = ({}^t X')({}^t P A P) Y' = {}^t X' A' Y',$$

da cui si ottiene $A' = {}^t PAP$. In questo caso le matrici A e A' sono dette congruenti, e si può dimostrare che matrici per cui vale $A' = {}^t PAP$ rappresentano la stessa forma bilineare.

6.7 Definizione. Due matrici $A, A' \in M(n \times n, K)$ si dicono **congruenti** se esse rappresentano la stessa forma bilineare, ovvero se esiste una matrice $P \in GL(n, K)$ tale che $A' = {}^t PAP$.

Esempio. Le matrici congruenti alla matrice identità, cioè del tipo ${}^t PP$ con $P \in GL(n, K)$, rappresentano il prodotto scalare standard su $\mathbb{R}^n$ (vedi es. 6.3 (2)).

6.8 Definizione. Sia V uno spazio vettoriale **reale**. Chiameremo **prodotto scalare** una forma bilineare simmetrica f che sia anche definita positiva, cioè tale che $\forall v \neq O, f(v, v) > 0$. In genere, per un prodotto scalare useremo la notazione $\langle v, w \rangle$ invece di $f(v, w)$.

La coppia $(V, \langle\ ,\ \rangle)$ viene detta **spazio euclideo** reale.

6.9 Dunque ogni matrice associata a un prodotto scalare deve avere questa proprietà (oltre a essere simmetrica): $\forall\ X \in \mathbb{R}^n, X \neq O,\ {}^tXAX > 0$. E viceversa, la forma bilineare rappresentata da una matrice simmetrica A con questa proprietà è un prodotto scalare. C'è un modo per vedere quando ciò accade: la matrice A, oltre che essere simmetrica, deve avere il determinante positivo (e dunque deve essere invertibile), e positivi anche tutti i determinanti dei minori che si ottengono cancellando, via via, l'ultima riga e l'ultima colonna, fino a giungere a (a_{11}). Questa condizione è necessaria e sufficiente.

6.10 Esempi. 1. $A = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 4 & 2 \\ 1 & 2 & 9 \end{pmatrix}$ rappresenta un prodotto scalare poiché $detA = 79$, $det \begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix} = 11$, $det(a_{11}) = 3$.

2. Consideriamo $(\mathbb{R}^4, l)$, dove l è la forma di Lorentz, così definita:

$$l(X,Y) = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4.$$

Notiamo che esistono dei vettori non nulli X per cui vale $l(X,X) = 0$: questa è la principale differenza rispetto al prodotto scalare standard su $\mathbb{R}^4$, che si esprime dicendo che l non è definita positiva. La forma di Lorentz dunque non è un prodotto scalare, infatti $M_{\mathcal{C}}(l)$ ha determinante negativo.

3. L'integrale descritto in 6.3(1) è un prodotto scalare su $\mathbb{R}[t]$ e su $\mathbb{R}_n[t]$; infatti se $p(t)$ è un polinomio non nullo, per continuità $\int_0^1 p^2(t)dt > 0$.

6.11 Proposizione. Sia $(V, \langle , \rangle)$ uno spazio euclideo di dimensione finita e siano $v_1, \ldots, v_r$ vettori non nulli a due a due ortogonali: allora essi sono linearmente indipendenti.

Dimostrazione. Sia $a_1v_1 + \cdots + a_rv_r = O$: ne segue che per ogni j,

$$\langle a_1v_1 + \cdots + a_rv_r, v_j\rangle = \langle O, v_j\rangle = 0$$

ma anche per ipotesi

$$\langle a_1v_1 + \cdots + a_rv_r, v_j\rangle = \langle a_jv_j, v_j\rangle = a_j||v_j||^2$$

da cui $a_j = 0$, essendo $v_j \neq O$.

6.12 Teorema. Sia $(V, \langle , \rangle)$ uno spazio euclideo di dimensione finita: ogni sua base può essere ortogonalizzata, ovvero trasformata in una base ortogonale, mediante il procedimento di Gram-Schmidt.

Dimostrazione. Ricordiamo la nozione di proiezione di un vettore su un vettore non nullo (cfr. [GA] capitolo 5): $pr_w v := \frac{\langle v,w\rangle}{\langle w,w\rangle} w$ è un multiplo di w, e vale $(v - pr_w v) \perp w$, poiché $(v - pr_w v) \perp pr_w v$.

Sia $\mathcal{B} = (v_1, \ldots, v_n)$. Chiamiamo $v_1 := w_1$, e descriviamo il primo passo del procedimento di Gram-Schmidt.

Sia $w_2 := v_2 - pr_{w_1} v_2$, allora $w_2 \neq O$, altrimenti v_1, v_2 sarebbero linearmente dipendenti; $w_2 \perp w_1$, per cui dalla proposizione 6.11 w_1, w_2 sono linearmente indipendenti, e $\mathcal{L}(v_1, v_2) = \mathcal{L}(w_1, w_2)$.

Il passo successivo è il seguente: definiamo $w_3 := v_3 - pr_{w_1} v_3 - pr_{w_2} v_3$. Allora $w_3 \neq O$, altrimenti v_1, v_2, v_3 sarebbero linearmente dipendenti (dato che $\mathcal{L}(v_1, v_2) = \mathcal{L}(w_1, w_2)$); $w_3 \perp w_1$ e $w_3 \perp w_2$: infatti

$$\langle w_3, w_1\rangle = \langle v_3 - \frac{\langle w_1, v_3\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle w_2, v_3\rangle}{\langle w_2, w_2\rangle} w_2, w_1\rangle = \langle v_3, w_1\rangle - \frac{\langle w_1, v_3\rangle}{\langle w_1, w_1\rangle}\langle w_1, w_1\rangle = 0$$

$$\langle w_3, w_2\rangle = \langle v_3 - \frac{\langle w_1, v_3\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle w_2, v_3\rangle}{\langle w_2, w_2\rangle} w_2, w_2\rangle = \langle v_3, w_2\rangle - \frac{\langle w_2, v_3\rangle}{\langle w_2, w_2\rangle}\langle w_2, w_2\rangle = 0$$

per cui w_1, w_2, w_3 sono linearmente indipendenti dalla Proposizione 6.10, e infine $\mathcal{L}(v_1, v_2, v_3) = \mathcal{L}(w_1, w_2, w_3)$.

Si procede in questo modo fino all'ultimo vettore $w_n := v_n - pr_{w_1} v_n \cdots - pr_{w_{n-1}} v_n$, e si ha $V = \mathcal{L}(v_1, \ldots, v_n) = \mathcal{L}(w_1, \ldots, w_n)$, dunque $(w_1, \ldots, w_n)$ è la base cercata.

6.13 Corollario. Ogni spazio euclideo di dimensione finita ha basi ortonormali. Se $\mathcal{B} = (v_1, \ldots, v_n)$ è una base ortonormale, vale $\langle v_i, v_j\rangle = \delta_{ij}$, e $M_{\mathcal{B}}(<,>) = I$.

Dimostrazione. Segue dal teorema, dividendo i vettori trovati per la loro norma. La seconda parte è ovvia.

6.14 Definizione. Sia $(V, \langle,\rangle)$ uno spazio euclideo, e W un suo sottospazio. Il complemento ortogonale di W è $W^\perp := \{v \in V / \langle v, w\rangle = 0\ \forall w \in W\}$. La proiezione ortogonale π di V su W è così definita: se $v = w + u$, con $w \in W$ e $u \in W^\perp$, $\pi(v) = w$.

6.15 Proposizione. Sia $(V, \langle,\rangle)$ uno spazio euclideo di dimensione n, e W un suo sottospazio. Allora $W^\perp$ è un sottospazio vettoriale di V e vale $V = W \oplus W^\perp$; inoltre la proiezione ortogonale π di V su W è una applicazione lineare.

Dimostrazione. E' immediato verificare che $W^\perp$ è un sottospazio vettoriale di V.

Dimostriamo ora che $V = W \oplus W^\perp$. (1) $W \cap W^\perp = O$: infatti se $w \in W$ sta anche in $W^\perp$, vale $\langle w, w\rangle = 0$, ma questo è impossibile se $w \neq O$.

(2) Per mostrare che $V = W + W^\perp$, fisso una base ortonormale $(w_1, \ldots, w_k)$ di W e scrivo un generico $v \in V$ come $v = \sum_1^k a_j w_j + (v - \sum_1^k a_j w_j)$: devo determinare

i numeri a_j di modo che $v - \sum_1^k a_j w_j \in W^\perp$, cioè in modo che $\langle v - \sum_1^k a_j w_j, w_i \rangle = 0$ per ogni indice i. Per linearità, mi basta scegliere $a_j = \langle v, w_j \rangle$.

(3) Siano $a, b \in \mathbb{R}$ e $v_1, v_2 \in V$: dunque esistono $w_1, w_2 \in W$ e $u_1, u_2 \in W^\perp$ tali che $v_1 = w_1 + u_1, v_2 = w_2 + u_2$ da cui

$$av_1 + bv_2 = (aw_1 + bw_2) + (au_1 + bu_2).$$

Perciò

$$\pi(av_1 + bv_2) = aw_1 + bw_2 = a\pi(v_1) + b\pi(v_2)$$

che dice la linearità di $\pi : V \to W$.

6.16 Proposizione. Sia V uno spazio euclideo e $\mathcal{B} = (v_1, \ldots, v_n)$ una sua base ortonormale. Allora per ogni $v \in V$, $[v]_\mathcal{B} = (\langle v, v_1 \rangle, \ldots, \langle v, v_n \rangle)$, e se v e w hanno come coordinate rispetto a $\mathcal{B}$ i vettori X e Y, allora $\langle v, w \rangle = {}^tX \cdot Y = \sum_{i=1}^n x_i y_i$.

Inoltre la matrice del cambiamento di base fra due basi ortonormali è una matrice ortogonale.

Dimostrazione. Se $[v]_\mathcal{B} = (x_1, \ldots, x_n)$, risulta

$$\langle v, v_j \rangle = \langle x_1 v_1 + \cdots + x_n v_n, v_j \rangle = x_j \langle v_j, v_j \rangle = x_j;$$

nello stesso modo si dimostra l'altra uguaglianza.

Sia $\mathcal{B}' = (w_1, \ldots, w_n)$ un'altra base ortonormale di V. Le colonne di $P := M(\mathcal{B}, \mathcal{B}')$ sono $X_j = [w_j]_\mathcal{B}$, ma vale ${}^tX_j \cdot X_k = \langle w_j, w_k \rangle = \delta_{j,k}$, perciò P ha colonne di norma uno e a due a due ortogonali, dunque è una matrice ortogonale.

Su uno spazio euclideo si possono considerare i concetti di lunghezza, distanza, angolo, come su $\mathbb{R}^n$ dotato del prodotto scalare standard. Le proprietà sono le stesse (a motivo dell'esistenza di basi ortonormali e della Proposizione 6.16), e si possono vedere in [GA] capitolo 5.

6.17 Passiamo ora a considerare **spazi vettoriali complessi**, in particolare consideriamo $\mathbb{C}^n$, su cui vorremmo definire il concetto, per esempio, di lunghezza. Se considero $\mathbb{C}^n$ come spazio vettoriale reale, ovvero se lo identifico con $\mathbb{R}^{2n}$, di modo che $X = (x_1, \ldots, x_n) = (a_1 + ib_1, \ldots, a_n + ib_n) = (a_1, b_1, \ldots, a_n, b_n)$, allora risulta

$$||X|| = \sqrt{a_1^2 + b_1^2 + \cdots + a_n^2 + b_n^2} = \sqrt{x_1\overline{x_1} + \cdots + x_n\overline{x_n}} = \sqrt{|x_1|^2 + \cdots + |x_n|^2}.$$

Ma estendendo in modo naturale il prodotto scalare a $\mathbb{C}^n$, avrei invece $||X||^2 = x_1^2 + \cdots + x_n^2 \in \mathbb{C}$, e quindi esisterebbero vettori non nulli, come per esempio $X = (1, i) \in \mathbb{C}^2$, che hanno lunghezza nulla, il che contrasta con l'idea di lunghezza.

Dunque il prodotto scalare va definito su $\mathbb{C}^n$ di modo che la lunghezza di un vettore continui a essere un numero reale non negativo, e non succeda l'anomalia descritta sopra: si pone perciò (**prodotto hermitiano standard** in $\mathbb{C}^n$)

$$X \cdot Y = \Sigma x_i \overline{y_i}.$$

E' facile dimostrare la seguente proposizione:

6.18 Proposizione. Il prodotto hermitiano standard gode delle seguenti proprietà:

(1) $(aX + bZ) \cdot Y = a(X \cdot Y) + b(Z \cdot Y)$ e $Y \cdot (aX + bZ) = \overline{a}Y \cdot X + \overline{b}Y \cdot Z$ (linearità complessa o hermitiana)
(2) $X \cdot Y = \overline{Y \cdot X}$ (simmetria complessa o hermitiana)
(3) Se $X \neq O$, $||X||^2 = X \cdot X = \Sigma x_i \overline{x_i} = \Sigma |x_i|^2 > 0$ (definito positivo).

Notiamo in particolare che la proprietà (2) non ci permette di applicare al prodotto hermitiano standard la teoria delle forme bilineari simmetriche. Definiamo perciò un analogo complesso, le forme hermitiane.

6.19 Definizione. Sia V uno spazio vettoriale complesso. Una **forma hermitiana** su V è una funzione $f : V \times V \to \mathbb{C}$ lineare e simmetrica in senso complesso, cioè che soddisfa, per $v, w, u \in V$, e $a, b \in \mathbb{C}$:

(1) $f(v, w) = \overline{f(w, v)}$
(2) $f(av + bu, w) = af(v, w) + bf(u, w)$
Essa è definita positiva se soddisfa anche
(3) Se $v \neq O$, $||v||^2 = f(v, v) > 0$.

6.20 Proposizione. Sia V uno spazio vettoriale complesso di dimensione n, e sia $\mathcal{B} = (v_1, \ldots, v_n)$ una sua base; ogni matrice $A \in M(n \times n, \mathbb{C})$ tale che $A = \overline{{}^tA}$, ovvero ogni **matrice hermitiana**, dà luogo a una forma hermitiana f_A in questo modo: se

$$v, w \in V,\ [v]_{\mathcal{B}} = X = {}^t(x_1, \ldots, x_n), [w]_{\mathcal{B}} = Y = {}^t(y_1, \ldots, y_n),$$

allora $f_A(v, w) := {}^tXA\overline{Y}$; viceversa, ogni forma hermitiana f è del tipo f_A se si sceglie $a_{ij} := f(v_i, v_j)$ (risulta ovviamente che $A = \overline{{}^tA}$). Scriveremo $A = M_{\mathcal{B}}(f)$.

Dimostrazione. Sia A una matrice hermitiana, ovvero tale che $A = \overline{{}^tA}$: verifichiamo che f_A soddisfa la definizione 6.21. Se $[u]_{\mathcal{B}} = Z$, allora $f_A(av + bu, w) =$

$${}^t(aX + bZ)A\overline{Y} = (a{}^tX + b{}^tZ)A\overline{Y} = a{}^tXA\overline{Y} + b{}^tZA\overline{Y} = af_A(v, w) + bf_A(u, w).$$

Inoltre $f_A(Y, X) := {}^tYA\overline{X} = {}^t({}^tYA\overline{X}) = {}^t({}^tY\overline{{}^tA}\overline{X}) = \overline{{}^tXA\overline{Y}} = \overline{f_A(X, Y)}$.

Sia poi A data da $a_{ij} = f(v_i, v_j)$: come nel caso reale si dimostra che $f(v, w) = f_A(v, w)$; inoltre dalle proprietà di f si ha $a_{ij} = f(v_i, v_j) = \overline{f(v_j, v_i)} = \overline{a_{ji}}$, perciò $A = \overline{{}^t A}$.

Nota. Se f è il prodotto hermitiano standard, $I = M_{\mathcal{C}}(f)$.

6.21 Consideriamo ora uno spazio vettoriale complesso V di dimensione n e una forma hermitiana su V definita positiva (o **prodotto hermitiano**), che indichiamo con $\langle , \rangle$: dunque, se $v \neq O, \langle v, v \rangle$ è un numero reale positivo. La coppia $(V, \langle , \rangle)$ è detta uno **spazio hermitiano**. Su V possiamo considerare basi ortogonali e basi ortonormali; per esempio, la base canonica è una base ortonormale di $\mathbb{C}^n$ con il prodotto hermitiano standard.

6.22 Proposizione. Gli autovalori di una matrice hermitiana sono numeri reali. Le radici del polinomio caratteristico di una matrice simmetrica reale sono reali, e dunque sono autovalori.

Dimostrazione. Denotiamo con $\langle , \rangle$ il prodotto hermitiano standard su $\mathbb{C}^n$. Sia M una matrice hermitiana e sia $X \neq O$ tale che $MX = \lambda X$. Allora $\langle MX, X \rangle = \langle \lambda X, X \rangle = \lambda ||X||^2$, ma anche

$$\langle MX, X \rangle = {}^t(MX)\overline{X} = {}^tX^tM\overline{X} = {}^tX(\overline{MX}) = \langle X, MX \rangle = \langle X, \lambda X \rangle = \overline{\lambda}||X||^2,$$

quindi $\overline{\lambda} = \lambda$.

Nel secondo caso, basta osservare che una matrice simmetrica reale è una particolare matrice hermitiana.

6.23 Teorema Spettrale nel caso simmetrico. (1) Se $A \in M(n \times n, \mathbb{R})$ è simmetrica, esiste una matrice ortogonale P tale che ${}^tPAP = P^{-1}AP$ è diagonale (e viceversa).

(2) Sia V uno spazio euclideo: operatori simmetrici si diagonalizzano in base ortonormale, ovvero esiste una base ortonormale di V formata da autovettori dell'operatore, e gli elementi della matrice diagonale sono i corrispondenti autovalori.

Dimostrazione. (1) Sia λ un autovalore di A (esiste dalla proposizione 6.22) e v_1 un suo autovettore di norma 1. Procedo per induzione su n: se $n = 1$, ho concluso, perché ogni matrice 1×1 è diagonale.

Altrimenti, estendo v_1 a una base ortonormale $\mathcal{B} = (v_1, \ldots, v_n)$ di $\mathbb{R}^n$. Sia $\tilde{P} = M(\mathcal{C}, \mathcal{B})$: le colonne di P sono i vettori v_j, e $\tilde{P}$ è ortogonale (proposizione 6.16). Sia $M = {}^t\tilde{P}AP$: è facile verificare che M è simmetrica, e che la prima colonna (e anche riga) di M è ${}^t(\lambda, 0, \ldots, 0)$; infatti vale $m_{i,1} = {}^tv_iAv_1 = {}^tv_i\lambda_1v_1$.

Questo dice che M è della forma $\begin{pmatrix} \lambda & O \\ O & N \end{pmatrix}$, dove $N \in M((n-1)\times(n-1),\mathbb{R})$ è una matrice simmetrica, perciò per induzione esiste $Q \in O(n-1)$ con ${}^tQNQ = D$ diagonale. Se considero $\tilde{Q} \in O(n)$, $\tilde{Q} = \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix}$, è facile verificare che la matrice

$$B := {}^t(\tilde{P}\tilde{Q})A(\tilde{P}\tilde{Q}) = {}^t\tilde{Q}({}^t\tilde{P}A\tilde{P})\tilde{Q} = {}^t\tilde{Q}M\tilde{Q}$$

è diagonale: dunque la matrice cercata è $P := \tilde{P}\tilde{Q}$.

Viceversa, sia $D = {}^tPAP = P^{-1}AP$: allora $A = PDP^{-1} = PD({}^tP)$, quindi A è simmetrica.

(2) Sia T un operatore sullo spazio euclideo $(V, \langle,\rangle)$ dotato di una base ortonormale $\mathcal{B}$, e sia A una matrice simmetrica tale che $A = M_{\mathcal{B}}(T)$. Allora esiste una matrice ortogonale P tale che ${}^tPAP = P^{-1}AP = D$ è diagonale, dunque D è simile ad A, perciò rappresenta l'operatore T rispetto ad un'altra base $\mathcal{B}'$, e P è la matrice del cambiamento di base. Poiché P è ortogonale, anche $\mathcal{B}'$ è una base ortonormale.

6.24 Esempio. Diagonalizzare in base ortonormale l'operatore rappresentato in base canonica dalla matrice $D = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

D ha autovalori $\pm\sqrt{5}, 0$; posso formare la matrice P che diagonalizza prendendo come colonne un vettore di norma uno per ogni autospazio. Per esempio, posso scegliere $P = \frac{1}{\sqrt{10}} \begin{pmatrix} 2 & 2 & \sqrt{2} \\ \sqrt{5} & -\sqrt{5} & 0 \\ 1 & 1 & -2\sqrt{2} \end{pmatrix}$.

Esercizi.

(6.1) Scrivere la matrice che rappresenta il prodotto scalare standard di $\mathbb{R}^3$ rispetto alla base $\mathcal{B} = ((1,1,1),(1,1,0),(-1,0,0))$.

Soluzione. $M = \begin{pmatrix} 3 & 2 & -1 \\ 2 & 2 & -1 \\ -1 & -1 & 1 \end{pmatrix}$.

(6.2) Scrivere una base ortonormale per il sottospazio di $\mathbb{R}^3$ generato dai vettori $(1,2,3), (-1,-1,0)$.

Soluzione. Uso il procedimento di Gram Schmidt per ottenere una base ortogonale. Sia $w_1 = (1,2,3)$, e $w_2 = (-1,-1,0) - pr_{(1,2,3)}(-1,-1,0) = (-1,-1,0) -$

$(-3/14)(1,2,3) = (-11/14, -8/14, 9/14)$. Ora devo dividere i due vettori per la loro norma: $||w_1|| = \sqrt{14}$ e $||w_2|| = \sqrt{266}/14$, e ottengo la base ortonormale cercata: $(1/\sqrt{14}, 2/\sqrt{14}, 3/\sqrt{14}), (-11/\sqrt{266}, -8/\sqrt{266}, 9/\sqrt{266})$.

(6.3) Dimostrare che se f è un prodotto scalare su V, l'unico vettore v con la proprietà: $f(w,v) = 0 \ \forall\ w \in V$, è il vettore nullo (detto in altri termini, $V^\perp = O$).

Dimostrazione. Sia f una forma bilineare simmetrica e A una matrice che la rappresenta. Se f è un prodotto scalare, la matrice A è non singolare, e quindi l'unica soluzione del sistema omogeneo $Av = O$ è $v = O$. Ma $f(w,v) = {}^t w(Av)$ deve essere zero per ogni w, e questo è possibile solo se $Av = O$, cosa che implica, come abbiamo detto, che $v = O$.

(6.4) Ortonormalizzare con il procedimento di Gram-Schmidt la seguente base di $\mathbb{R}^3$: $\mathcal{B} = ((1,1,1),(1,1,0),(4,0,0))$.

Dimostrazione. Risulta $w_1 = (1,1,1)$,
$w_2 = (1,1,0) - pr_{(1,1,1)}(1,1,0) = (1,1,0) - (2/3)(1,1,1) = (1/3, 1/3, -2/3)$,
$w_3 = (4,0,0) - pr_{(1,1,1)}(4,0,0) - pr_{(1/3,1/3,-2/3)}(4,0,0) = (4,0,0) - (4/3)(1,1,1) - (4/3)(9/6)(1/3,1/3,-2/3) = (4,0,0) - (4/3,4/3,4/3) - (2/3,2/3,-4/3) = (2,-2,4/3)$.
Dunque ottengo la base ortonormale dividendo per le norme: $||w_1|| = \sqrt{3}$, $||w_2|| = \sqrt{6}/3$, $||w_3|| = 2\sqrt{11}/3$, da cui la base:

$$((1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}), (1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6}), (3/\sqrt{11}, -3/\sqrt{11}, 2/\sqrt{11}).$$

(6.5) Considerare $\mathbb{R}_2[t]$ con il prodotto scalare dato dall'integrazione fra 0 e 1, $W = \mathcal{L}(t^2)$; calcolare $W^\perp$. Calcolare anche la lunghezza di $t^2 - 1$.

Soluzione. $W^\perp = \{p(t) = a_0 + a_1 t + a_2 t^2 / \int_0^1 t^2 p(t)dt = 0\}$. La condizione diventa:

$$0 = \int_0^1 t^2 p(t)dt = a_0 \int_0^1 t^2 dt + a_1 \int_0^1 t^3 dt + a_2 \int_0^1 t^4 dt = \frac{a_0}{3} + \frac{a_1}{4} + \frac{a_2}{5},$$

dunque i polinomi che compongono $W^\perp$ sono quelli i cui coefficienti soddisfano $\frac{a_0}{3} + \frac{a_1}{4} + \frac{a_2}{5} = 0$.

Osserviamo che $W^\perp$ ha dimensione due, e $W \cap W^\perp = O$.

Inoltre risulta $||t^2 - 1||^2 = \int_0^1 (t^2-1)^2 dt = \int_0^1 (t^4 - 2t^2 + 1)dt = \frac{1}{5} - \frac{2}{3} + 1 = \frac{8}{15}$, ovvero la lunghezza di $t^2 - 1$ è $\sqrt{8/15}$.

(6.6) Se $V = \mathbb{R}^3$, W è la retta passante per l'origine di direzione $(1, 1, -1)$ e π è la proiezione di V su W, calcolare $\pi(3, 4, 5)$ e $\pi(-1, -1, 1)$.

Dimostrazione. Per definizione, $W = \mathcal{L}((1, 1, -1))$ e dunque si vede subito che, essendo $(-1, -1, 1) \in W$, vale $\pi(-1, -1, 1) = (-1, -1, 1)$.

Calcoliamo $W^\perp$: si ha $(x, y, z) \in W^\perp$ se $x + y - z = 0$, ovvero se $z = x + y$.

Perciò $(3, 4, 5) = (k, k, -k) + (x, y, x + y)$ da cui segue

$$k + x = 3, k + y = 4, -k + x + y = 5$$

che dà $k = \frac{2}{3}$ perciò $\pi(3, 4, 5) = (\frac{2}{3}, \frac{2}{3}, -\frac{2}{3})$.

(6.7) Dimostrare che $f : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$, $f((z_1, z_2), (w_1, w_2)) = iz_1\overline{w_2} - iz_2\overline{w_1}$ è una forma hermitiana. E' un prodotto hermitiano?

Dimostrazione. La eventuale matrice associata sarebbe $A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$, che è una matrice hermitiana; siccome vale

$$(z_1, z_2) \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} \overline{w_1} \\ \overline{w_2} \end{pmatrix} = (-iz_2, iz_1) \begin{pmatrix} \overline{w_1} \\ \overline{w_2} \end{pmatrix} =$$

$$= -iz_2\overline{w_1} + iz_1\overline{w_2} = f((z_1, z_2), (w_1, w_2)),$$

f è una forma hermitiana.

Non è un prodotto hermitiano poiché, per esempio, $f((1, 0), (1, 0)) = 0$.

(6.8) Dire quali fra le seguenti sono forme bilineari simmetriche e quali forme hermitiane:

$f((x, y, z), (a, b, c)) = 3ax + 2by + 4xc - 3xb$

$g((z_1, z_2), (w_1, w_2)) = z_1w_1 + \overline{z_1}w_2$

$h(v, w) = v \times w$, dove $\times$ designa il prodotto vettoriale in $\mathbb{R}^3$.

Soluzione. h non è né una forma bilineare né una forma hermitiana poiché i suoi valori sono in $\mathbb{R}^3$.

Di f e g calcoliamo le eventuali matrici, che sono $A = \begin{pmatrix} 3 & -3 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. Si verifica facilmente che A riproduce f, dunque f è una forma bilineare, non simmetrica, mentre B non riproduce g.

(6.9) Dire quali fra le seguenti matrici sono simmetriche, definite positive, ortogonali, hermitiane.

$$A = \begin{pmatrix} i & -i \\ i & 0 \end{pmatrix} \; B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \; C = \begin{pmatrix} 3 & 1-3i \\ 1+3i & 0 \end{pmatrix} \; D = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Soluzione. Simmetriche: B, D. Hermitiane: B, C, D. Fra queste, le definite positive (si controlla con il metodo dei determinanti) sono: C. Ortogonali: B.

(6.10) Diagonalizzare in base ortonormale la matrice $A = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Soluzione. Gli autovalori di A sono 4 e 0 con molteplicità due. Gli autospazi sono:
$V_4 = \{(x, y, z)/y = x, z = 0\}, V_0 = \{(x, y, z)/x = -y\}$.

In essi scelgo

$$v_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), v_2 = (\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0), v_3 = (0, 0, 1).$$

Si ottiene $P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ e $A' = ({}^tP)AP = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.