Forme bilineari e prodotti scalari

Il prodotto scalare standard di $\mathbb{R}^n$ può anche essere scritto come un prodotto riga per colonna

$$\langle \underline{u}, \underline{v} \rangle_{\mathbb{R}^n} = \underline{u}^t I \underline{v}$$

dove I è la matrice identità. Possiamo generalizzare questa nozione definendo per una qualsiasi matrice $A \in \mathrm{Mat}(n)$

$$\phi_A(\underline{u}, \underline{v}) = \underline{u}^t A \underline{v}.$$

In questo modo otteniamo una funzione di due variabili $\phi_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ che, per le proprietà del prodotto riga per colonna, è bilineare come il prodotto scalare standard.

Osservazione 1. *Se $A = (a_{ij})$ allora $\phi_A(\underline{e}_i, \underline{e}_j) = a_{ij}$*

La situazione è simile a quella che abbiamo visto per le applicazioni lineari: scelta una base (qui abbiamo preso la base canonica) possiamo associare ad ogni matrice un'applicazione bilineare, e ad ogni applicazione bilineare una matrice poichè *ogni applicazione bilineare è determinata in modo unico dai suoi valori su tutte le possibili coppie ordinate di elementi di una base di $\mathbb{R}^n$.*

Proposizione 2. *Sia $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ un'applicazione bilineare. Definiamo la matrice A con coefficienti $a_{ij} = \psi(\underline{e}_i, \underline{e}_j)$ allora $\psi = \phi_A$, ossia per ogni coppia di vettori $\underline{u}, \underline{v} \in \mathbb{R}^n$ abbiamo $\psi(\underline{u}, \underline{v}) = \phi_A(\underline{u}, \underline{v})$. A si dice la matrice associata all'applicazione bilineare ψ rispetto alla base canonica di $\mathbb{R}^n$.*

dim: Scriviamo $\underline{u} = x_1\underline{e}_1 + \dots x_n\underline{e}_n$, $\underline{v} = y_1\underline{e}_1 + \dots + y_n\underline{e}_n$, allora, applicando la linearità prima in una variabile, poi nell'altra

$$\begin{aligned}
\psi(\underline{u}, \underline{v}) &= \psi(x_1\underline{e}_1 + \dots x_n\underline{e}_n, \underline{v}) = x_1\psi(\underline{e}_1, \underline{v}) + \dots + x_n\psi(\underline{e}_n, \underline{v}) = \\
=& x_1\psi(\underline{e}_1, y_1\underline{e}_1 + \dots + y_n\underline{e}_n) + \dots + x_n\psi(\underline{e}_n, y_1\underline{e}_1 + \dots + y_n\underline{e}_n) = \\
=& x_1(y_1\psi(\underline{e}_1, \underline{e}_1) + \dots + y_n\psi(\underline{e}_1, \underline{e}_n)) + \dots + x_n(y_1\psi(\underline{e}_n, \underline{e}_1) + \dots + y_n\psi(\underline{e}_n, \underline{e}_n)) = \\
=& x_1y_1a_{11} + x_1y_2a_{12} + \dots + x_1y_na_{1n} + \dots x_ny_1a_{n1} + \dots x_ny_na_{nn} = \\
=& \sum_{i,j=1}^{n} x_iy_ja_{ij} = \underline{x}^t A \underline{y} = \phi_A(\underline{u}, \underline{v})
\end{aligned}$$

Sempre come nel caso delle applicazioni lineari, anche per le forme bilineari possiamo definire la matrice associata ad una forma bilineare rispetto ad una base qualsiasi.

Definizione 3. *Sia $\psi : \mathbb{R}^n \times \mathbb{R}^n \to R$ una forma bilineare, $\mathcal{B} = \{\underline{u}_1, \dots, \underline{u}_n\}$ una base di $\mathbb{R}^n$. Allora la matrice di ψ rispetto alla base $\mathcal{B}$ è la matrice*

$$A = (a_{ij}) = \psi(\underline{u}_i, \underline{u}_j).$$

Se $\underline{\alpha}, \underline{\beta}$ sono le coordinate di $\underline{v}, \underline{w}$ rispetto alla base $\mathcal{B}$ allora si avrà

$$\psi(\underline{v}, \underline{w}) = \underline{\alpha}^t A \underline{\beta}.$$

Con la stessa dimostrazione della proposizione 2 possiamo dimostrare che ψ è determinata in maniera univoca dalla sua matrice associata rispetto ad una base qualsiasi.

Esempio 1. Consideriamo la base di $\mathbb{R}^3$

$$\mathcal{B} = \left\{ \underline{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \underline{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \underline{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

La matrice con coefficienti $a_{ij} = \langle \underline{v}_i, \underline{v}_j \rangle_{\mathbb{R}^3}$ è la matrice

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

che è la matrice del prodotto scalare standard di $\mathbb{R}^3$ rispetto alla base $\mathcal{B}$ e ha questa proprietà: consideriamo i vettori $\underline{x}_i = (a_i, b_i, c_i)^t \; i = 1, 2$ di coordinate $\underline{\xi}_i = (a_i - b_i, b_i - c_i, c_i)^t$ rispetto alla base $\mathcal{B}$. Il prodotto

$$\xi_1^t A \xi_2 = a_1 a_2 + b_1 b_2 + c_1 c_2 = \langle \underline{x}_1, \underline{x}_2 \rangle_{\mathbb{R}^3}$$

da come risultato il loro prodotto scalare.

Tutto il materiale di queste note si generalizza a spazi vettoriali qualsiasi. Usando le coordinate rispetto ad una base ci riportiamo al caso di $\mathbb{R}^n$. Nell'esempio successivo vediamo una forma bilineare molto importante in analisi. Possiede tutte le proprietà del prodotto scalare standard di $\mathbb{R}^n$

Esempio 2. Consideriamo i polinomi di $V = \mathbb{R}^4[x]$ come funzioni continue dall'intervallo $(0, 1)$ in $\mathbb{R}$, definiamo un'applicazione lineare sullo spazio dei polinomi di grado qualsiasi $\int_0^1 dx = I : \mathbb{R}[x] \to \mathbb{R}$ ponendo

$$I(1) = 1, \quad I(x) = \frac{1}{2}, \quad I(x^2) = \frac{1}{3}, \dots I(x^n) = \frac{1}{n+1} \dots$$

(per chi conosce gli integrali, I è l'integrale sull intervallo $(0, 1)$ del polinomio). Definiamo una forma bilineare $B : V \times V \to \mathbb{R}$ ponendo $B(p, q) = I(pq)$ La forma è bilineare perchè se $p, q, r \in V, k \in \mathbb{R}$ abbiamo $I((p+kq)r) = I(r(p + kq)) = I(pr + kqr) = I(pr) + kI(qr) = B(p, r) + kB(q, r)$ visto che I è lineare. Calcolando B su tutte le possibili coppie di elementi della base canonica (tenendo conto che $B(p, q) = B(q, p)$) troviamo la matrice

$$A = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{pmatrix}$$

Proposizione 4. *Data una forma bilineare $\psi : V \times V \to \mathbb{R}$, sia A la sua matrice associata rispetto a una base $\mathcal{B}$ e sia $\mathcal{B}'$ un'altra base di V. Allora la matrice associata a ψ rispetto a $\mathcal{B}'$ è la matrice $A' = ({}_{\mathcal{B}}M_{\mathcal{B}'})^t A \, {}_{\mathcal{B}}M_{\mathcal{B}'}$, dove ${}_{\mathcal{B}}M_{\mathcal{B}'}$, qui per brevità M, è la matrice di passaggio da $\mathcal{B}$ a $\mathcal{B}'$*

dim: La dimostrazione è simile a quella vista per le applicazioni lineari: se $\underline{v}'_i, \underline{v}'_j$ sono due vettori della base $\mathcal{B}'$, le loro coordinate rispetto a $\mathcal{B}'$ sono i vettori $\underline{e}_i, \underline{e}_j$ della base canonica di $\mathbb{R}^n$, mentre le loro cordinate rispetto a $\mathcal{B}$ sono i vettori $M\underline{e}_i, M\underline{e}_j$, le colonne i e j di M. Allora calcolando $\psi(\underline{v}'_i, \underline{v}'_j)$ nelle coordinate rispetto alla base $\mathcal{B}$ usando la matrice A, detti a'_{ij} i coefficienti di A', abbiamo

$$a'_{ij} = \psi(\underline{v}'_i, \underline{v}'_j) = (M\underline{e}_j)^t AM\underline{e}_i = \underline{e}_j^t M^t AM\underline{e}_i = \underline{e}_j^t(M^t AM)\underline{e}_i = (M^t AM)_{ij}.$$

Quindi il coefficiente i, j della matrice A' è il coefficiente i, j della matrice $M^t AM$.

La relazione tra matrici che rappresentano la stessa forma bilineare rispetto a basi diverse si chiama congruenza.

Definizione 5. *Siano $A, B \in \mathrm{Mat}(n)$ due matrici quadrate. B si dice congruente ad A se esiste una matrice invertibile M tale che*

$$B = M^t AM$$

Esercizio. Dimostrare che

a) La congruenza di matrici è una relazione di equivalenza.

b) Matrici congruenti hanno lo stesso rango

Osservazione 6. *Congruenza e similitudine sono relazioni diverse sulle matrici: ad esempio*

$$\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} = I$$

ma la matrice $4I$ non è simile alla matrice identità. Però se la matrice M è ortogonale, la congruenza è anche una similitudine.

Esempio 3. Riprendendo l'Esempio 1, detta M la matrice di passaggio dalla base canonica alla base $\mathcal{B}$, abbiamo che la matrice A associata al prodotto standard di $\mathbb{R}^3$ rispetto a $\mathcal{B}$ è proprio $M^t IM = M^t M$, visto che I è la matrice del prodotto standard rispetto alla base canonica.

Forme bilineari simmetriche

Una proprietà importante del prodotto scalare standard è la simmetria nelle due variabili. Una forma bilineare che abbia questa proprietà si dice simmetrica. Segue dall'osservazione 1 e dalla dimostrazione della proposizione 2 che la matrice associata ad una forma bilineare è una matrice simmetrica. Viceversa se A è una matrice simmetrica, $\alpha(\underline{u},\underline{v}) = \underline{u}^t A\underline{v}$ è una forma bilineare simmetrica. Una forma simmetrica si dice anche *prodotto scalare*

Definizione 7. *Sia $\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ un prodotto scalare su $\mathbb{R}^n$ (o più in generale su uno spazio V), la forma quadratica associata ad α è la funzione $Q_\alpha : \mathbb{R}^n \to \mathbb{R}$, $Q_\alpha(\underline{v}) = \alpha(\underline{v},\underline{v})$*

Ad esempio, la forma quadratica associata al prodotto scalare standard di $\mathbb{R}^n$ è la norma al quadrato di un vettore.

Il nome forma quadratica deriva dal fatto che una volta scelta una base $\mathcal{B}$ di V, possiamo rappresentare una forma bilineare α con una matrice A e allora $q_\alpha(\underline{v})$ è un polinomio omogeneo di grado 2 nelle coordinate $(x_1,\dots,x_n)^t$ di $\underline{v}$ rispetto a $\mathcal{B}$:

$$p_\alpha(\underline{v}) = \alpha(\underline{v},\underline{v}) = \underline{x}^t A\underline{x} = \sum_{i=1}^n x_i(\sum_{j=1}^n a_{ij}x_j) = \sum_{ij=1}^n a_{ij}x_ix_j$$

Esempio 4. Sia $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, la forma quadratica associata alla forma bilineare definita dalla matrice è

$$(x,y)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2xy$$

Con il cambiamento di variabili $X = \frac{1}{\sqrt{2}}(x+y)$, $Y = \frac{1}{\sqrt{2}}(x-y)$ abbiamo che $2xy = X^2 - Y^2$. Questo cambiamento corrisponde a cambiare base dalla base canonica alla base $\mathcal{B} = \{(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})^t,(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}})^t\}$. Detta M la matrice di cambiamento abbiamo $M^tAM = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ che ha come forma quadratica associata $X^2 - Y^2$.

Non tutti i prodotti scalari sono definiti positivi come il prodotto standard di $\mathbb{R}^n$, si hanno diversi casi:

Definizione 8. *Sia $\alpha : V \times V \to \mathbb{R}$ un forma bilineare simmetrica. Diciamo che B è*

- *Degenere se $\exists \underline{u} \neq \underline{0} \in V \mid \forall v \in V\ \alpha(\underline{u},\underline{v}) = 0$*
- *Definita positiva (negativa) se $\forall \underline{u} \in V\ \underline{u} \neq 0 \Rightarrow \alpha(\underline{u},\underline{u}) > 0 \quad (< 0)$*
- *Semidefinita positiva (negativa) se $\forall \underline{u} \in V\ \Rightarrow \alpha(\underline{u},\underline{u}) \geq 0 \quad (\leq 0)$*

- *Indefinita se esistono* $\underline{v}, \underline{w} \in V$ *tali che* $\alpha(\underline{v}, \underline{v}) > 0$, $\alpha(\underline{w}, \underline{w}) < 0$, *ossia* B *non è definita positiva o negativa.*

La stessa terminologia si usa anche per le forme quadratiche associate.

Teorema 9. *Sia* $\alpha : V \times V \to \mathbb{R}$ *un prodotto scalare. Allora* α *è degenere se e solo se la matrice di* B *rispetto a una base qualsiasi di* V *non ha rango massimo.*

dim: Sia A la matrice associata rispetto a una base $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ di V. Supponiamo che il rango di A non sia massimo. Allora esiste un vettore di $\mathbb{R}^n$, $\underline{c} = (c_1, \dots, c_n)^t \neq 0$ tale che $A\underline{c} = \underline{0}$, quindi per qualsiasi $\underline{v} \in V$, posto $\underline{u} = c_1\underline{v}_1 + \cdots + c_n\underline{v}_n$, detto $\underline{b} \in \mathbb{R}^n$ il vettore delle coordinate di $\underline{v}$ rispetto a $\mathcal{B}$, abbiamo $\alpha(\underline{u}, \underline{v}) = \underline{b}^t A\underline{c} = \underline{b}^t\underline{0} = \underline{0}$.

Viceversa se $\underline{0} \neq \underline{v} \in V$ tale che $\alpha(\underline{v}, \underline{u}) = 0\,\forall\underline{u} \in V$, $\underline{c}$ sono le coordinate di $\underline{v}$ rispetto a una base $\mathcal{B} = \{\underline{u}_1, \dots, \underline{u}_n\}$ e A la matrice di α rispetto a questa base, allora si avrà $\underline{x}^t A\underline{c} = \underline{x}^t(A\underline{c}) = \langle \underline{x}, A\underline{c}\rangle_{\mathbb{R}^n} = 0$ per ogni $\underline{x} \in \mathbb{R}^n$. Poichè il prodotto standard di $\mathbb{R}^n$ è definito positivo, questo implica $A\underline{c} = 0$ con $\underline{c} \neq \underline{0}$ visto che sono le coordinate di un vettore non nullo rispetto ad una base di V. Dunque A non può avere rango massimo.

Definizione 10. *Sia* $\alpha : V \times V \to \mathbb{R}$ *un prodotto scalare. Il nucleo o radicale di* α *è*

$$V^\perp = \{\underline{u} \in V \,|\, \forall\underline{v} \in V,\ \alpha(\underline{u}, \underline{v}) = 0\}$$

Inoltre (anche se queste definizioni si usano soprattutto nel caso di prodotti scalari definiti positivi) diciamo che due vettori $\underline{u}, \underline{v} \in V$ *sono ortogonali se* $\alpha(\underline{u}, \underline{v}) = 0$, *e se* $U \subset V$ *è un sottospazio l' ortogonale di* U *(rispetto ad* α*) è*

$$U^\perp = \{\underline{v} \in V \,|\, \forall\underline{u} \in U,\ \alpha(\underline{u}, \underline{v}) = 0\}$$

Esempio 5. La forma bilineare definita nell'Esempio 2 è un prodotto scalare definito positivo, viene detta metrica L^2 ed è importante in analisi. In $\mathbb{R}^3$, ponendo $\alpha(\underline{u}, \underline{v}) = \underline{u}^t A_i \underline{v}$, dove A_i è una delle matrici

$$A_1 = -I_3, \qquad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

otteniamo forme rispettivamente definita negativa, semidefinita positiva, indefinita su $\mathbb{R}^3$.

La forma A_2 è degenere e per questa forma abbiamo $(\mathbb{R}^3)^\perp = L[\underline{e}_3]$ e $L[\underline{e}_1, \underline{e}_3]^\perp = L[\underline{e}_2, \underline{e}_3]$, quindi l'intersezione tra un sottospazio E e il sottospazio dei vettori ortogonali a tutti i vettori di E può essere non vuota.

Anche quando la forma α è indefinita ci sono vettori non nulli tali che $\alpha(\underline{u}, \underline{u}) = 0$. Ad esempio se $\underline{u} = (1, 0, 1)^t$, $\underline{u}^t A_3 \underline{u} = 0$, e dunque $L[\underline{u}] \cap L[\underline{u}]^\perp$ contiene vettori non nulli (in particolare il vettore $\underline{u}$ stesso).Però, quando la forma è non degenere indefinita, possiamo sempre trovare un vettore $\underline{v}$ tale che $\alpha(\underline{u}, \underline{v}) \neq 0$.

Questi due esempi mostrano come il teorema che dice che possiamo sempre scomporre $\mathbb{R}^n$ comme somma diretta di un sottospazio E e del suo complemento ortogonale si generalizzi solo alle forme definite positive o negative. In effetti, essenzialmente tutto quello che abbiamo visto per il prodotto standard di $\mathbb{R}^n$ resta valido (eventualmente con segni diversi) per forme che siano definite (positive o negative).

Abbiamo il seguente teorema

Teorema 11 (Matrice in forma di Sylvester)**.** *Data una forma bilineare simmetrica $\alpha : V \times V \to \mathbb{R}$ esiste una base $\mathcal{B} = \{\underline{u}_1, \ldots, \underline{u}_n\}$ di V tale che la matrice associata ad α rispetto a $\mathcal{B}$ sia diagonale con i primi p elementi della diagonale uguali a 1, i successivi q elementi della diagonale uguali a -1 e gli ultimi $n - p - q$ elementi nulli.*

Il teorema può essere dimostrato essenzialmente adattando l'algoritmo di Gram-Schmidt al caso di forme che non siano definite positive. Qui però lo dimostriamo usando il teorema spettrale. Per semplicità ci limitiamo al caso di forme definite in $\mathbb{R}^n$ (il caso generale si dimostra nello stesso modo usando le coordinate rispetto a una base qualsiasi).

Sia quindi A la matrice della forma α rispetto alla base canonica. La matrice A è simmetrica, quindi esiste una base ortonormale $\mathcal{B}$ =$\{\underline{u}_1, \ldots, \underline{u}_n\}$ di autovettori. Consideriamo gli autovalori (non necessariamente distinti)

$$\lambda_1, \ldots, \lambda_p, \lambda_{p+1}, \ldots \lambda_{p+q}, \lambda_{p+q+1}, \ldots \lambda_n$$

ordinati in modo che i primi p siano quelli positivi, i successivi q quelli negativi e gli ultimi nulli (ossia 0 ripetuto $n - \mathrm{rk}(A)$ volte). La matrice di α rispetto a $\mathcal{B}$ è la matrice diagonale D con gli autovalori λ_i sulla diagonale: infatti se M è la matrice di passaggio dalla base canonica a $\mathcal{B}$, M è ortogonale quindi $M^t AM = M^{-1}AM = D$. In maniera equivalente possiamo osservare che se $i \neq j$

$$\underline{u}_i^t A\underline{u}_j = \underline{u}_i^t(A\underline{u}_j) = \underline{u}_i^t(\lambda_j \underline{u}_j) = 0$$

poichè gli elementi della base $\mathcal{B}$ sono autovettori ortonormali a due a due, mentre

$$\underline{u}_i^t A\underline{u}_i = \underline{u}_i^t(\lambda_i \underline{u}_i) = \lambda_i \|\underline{u}_i\|^2 = \lambda_i.$$

Consideriamo ora la base

$$\mathcal{D} = \left\{ \frac{\underline{u}_1}{\sqrt{\lambda_1}}, \ldots, \frac{\underline{u}_p}{\sqrt{\lambda_p}}, \frac{\underline{u}_{p+1}}{\sqrt{-\lambda_{p+1}}}, \ldots \frac{\underline{u}_{p+q}}{\sqrt{-\lambda_{p+q}}}, \underline{u}_{p+q+1}, \ldots \underline{u}_n \right\}$$

e denotiamo con $\underline{v}_i$ l'i-esimo vettore. Come prima $\underline{v}_i^t A\underline{v}_j = 0$ perchè i $\underline{v}_i$ sono autovettori ortogonali a due a due mentre il prodotto $\underline{v}_i^t A\underline{v}_i$ è uguale a $\frac{\lambda_i}{|\lambda_i|}$ per $i \leq p+q$, uguale a zero se $i > p+q$. Rispetto a questa base la matrice della forma α è in forma di Sylvester.

I numeri $p, q, n-p-p$ non dipendono dalla base scelta, ma hanno un significato intrinseco: $n-p-q$ è n meno il rango della matrice (o la molteplicità di 0 come autovalore di A), p, q sono le dimensioni dei sottospazi più grandi su cui α è definita positiva o negativa.

Proposizione 12. *Sia $\alpha : V \times V \to \mathbb{R}$ una forma bilineare simmetrica, $\mathcal{B}$ una base rispetto alla quale la matrice di α è in forma di Sylvester. Allora p è la dimensione massima di un sottospazio di V sul quale la restrizione di α è definita positiva. Di conseguenza p, q non dipendono dalla base $\mathcal{B}$.*

dim: Siano $\{\underline{v}_i\}$, $i = 1, \dots, n$ i vettori di $\mathcal{B}$. Allora per $i \leq p$ abbiamo che $\alpha(\underline{v}_i, \underline{v}_i) = 1$, per $p+1 \leq i \leq p+q$, $\alpha(\underline{v}_i, \underline{v}_i) = -1$ e per $p+q+1 \leq i \leq n$, $\alpha(\underline{v}_i, \underline{v}_i) = 0$. Sia U un sottospazio sul quale la restrizione di α è definita positiva e supponiamo per assurdo che $\dim U > p$. Posto $W = L[\underline{v}_{p+1}, \dots, \underline{v}_n]$, abbiamo che la restrizione di α su $\underline{W}$ è semidefinita negativa. Per la formula di Grassmann

$$n \geq \dim(U+W) = \dim U + \dim W - \dim(U \cap W) > p + n - p - \dim(U \cap W).$$

Poichè la seconda diseuguaglianza è stretta (abbiamo supposto $\dim U > p$) dovremmo avere $\dim(U \cap W) > 0$ e nell'intersezione dovrebbe esserci un vettore non nullo. Però detto $\underline{u}$ questo vettore dovremmo avere $\alpha(\underline{u}, \underline{u}) > 0$, visto che $\underline{u} \in U$, ma anche $\alpha(\underline{u}, \underline{u}) \leq 0$ poichè $\underline{u} \in W$. Questo è un assurdo, dunque p è la dimensione massima che può avere un sottospazio sul quale α sia definita positiva. Poichè p e $n-p-q$ non dipendono dalla base, neanche q dipende dalla base (o si dimostra allo stesso modo la proposizione analoga per q).

Possiamo quindi definire

Definizione 13. *La coppia di numeri (p, q) si dice segnatura della forma bilineare α e della forma quadratica associata.*

Quindi la segnatura è *un invariante della classe di congruenza* della matrice, ossia calcolandola per matrici che rappresentano la stessa forma α (e sono dunque congruenti) troviamo lo sstesso risultato. Viceversa, si può vedere (esercizio) che due matrici in forma di Sylvester con diversi $p, q, n-p-q$ non sono congruenti dunque la segnatura è un *invariante completo* della classe di congruenza:

Corollario 14. *Due matrici sono congruenti se e solo se hanno la stessa segnatura.*

La segnatura di una matrice si può anche calcolare calcolando gli autovalori e vedendo quanti sono positivi, quanti negativi e quanti nulli, però questo non è sempre il modo più semplice.

Esempio 6. Riprendiamo la forma B sullo spazio dei polinomi e la sua matrice A rispetto alla base canonica viste nell'Esempio 2. Siccome la norma al quadrato di un polinomio è l'integrale su un intervallo del polinomio al quadrato che è sempre maggiore o uguale a zero e zero solo per il polinomio nullo, questa è una forma definita positiva. Per metterla in forma di Sylvester basta quindi ortogonalizzare con l'algoritmo di Gram-Schmidt: prendiamo il polinomio costante $p_1 = 1$ come primo elemento. Abbiamo $B(1,1) = 1$, $B(1,x) = 1/2$ poniamo quindi $p_2 = x - 1/2$ e abbiamo $B(p_1,p_2) = 0$. Usando la matrice A (o chi conosce gli integrali può usarli per fare questi conti) $B(p_2,p_2) = 1/12$, $B(p_1,x^2) = 1/3$, $B(p_2,x^2) = 1/12$, quindi poniamo

$$p_3 = x^2 - \frac{1}{3}p_1 - p_2 = x^2 - x + \frac{1}{6}$$

Sempre usando la matrice abbiamo $B(p_3,p_3) = 1/180$, $B(x^4,p_1) = 1/5$, $B(x^4,p_2) = 3/40$, $B(x^4,p_3) = 1/120$ dunque

$$p_4 = x^3 - \frac{1}{5}p_1 - \frac{9}{10}p_2 - \frac{3}{2}p_3 = x^4 - \frac{3}{2}x^2 + \frac{3}{5}x$$

I polinomi p_1, p_2, p_3, p_4 formano una base ortogonale di $\mathbb{R}^4[x]$ rispetto a questo prodotto scalare. La norma al quadrato di p_4 è $B(p_4,p_4) = 1/350$, dividendo questi polinomi per le loro norme otteniamo una base ortonormale con matrice associata I_4, la forma di Sylvester di una forma bilineare definita positiva su uno spazio di dimensione 4.

D'altra parte, diagonalizzare la matrice A è molto più complicato.

Il criterio di Sylvester

Per stabilire se una matrice sia definita positiva o negativa, si può usare un semplice criterio. Chiamiamo minori principali di una matrice A di ordine n i minori $\mu_{1,1}, \mu_{12,12}, \ldots, \mu_{1\ldots n-1,1\ldots n-1}, A$ ossia i minori ottenuti prendendo le prime i righe e i colonne di A per ogni $i = 1, \ldots, n$.

Teorema 15 (Criterio di Sylvester)**.** *Sia A una matrice simmetrica con determinante diverso da zero, d_i i determinanti dei suoi minori principali, allora A è definita positiva se e solo se $d_i > 0\, \forall i = 1, \ldots, n$, definita negativa se e solo se $(-1)^i d_i > 0 \forall i = 1, \ldots, n$.*

La matrice deve avere rango massimo, ossia la forma deve essere non degenere, ma ci si può sempre ridurre a questo caso considerando la restrizione della forma a un complementare del suo radicale.