Capitolo 3

Rappresentazione geometrica di vettori e matrici

3.1 Vettori

3.1.1 Interpretazione geometrica

Nel caso generale un vettore è definito come $x \in \mathbb{R}^n$; nel caso di due o tre dimensioni i vettori sono rappresentabili come segmenti orientati, caratterizzati dalla lunghezza ($\|x\|_2$), dalla direzione e dal verso (rappresentati dal versore $\hat{x} = \frac{x}{\|x\|_2}$ per cui $\|\hat{x}\|_2 = 1$).

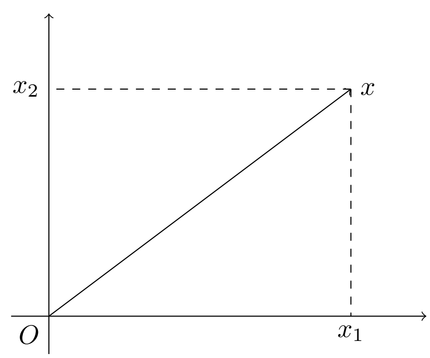

Figura 3.1: Vettore con due componenti

La somma di due vettori $x + y$ è ottenibile per via grafica o con la *regola del parallelogramma* o con il *metodo punta coda*. La sottrazione $x - y$ invece col secondo metodo (è il segmento che unisce le punte di x e y).

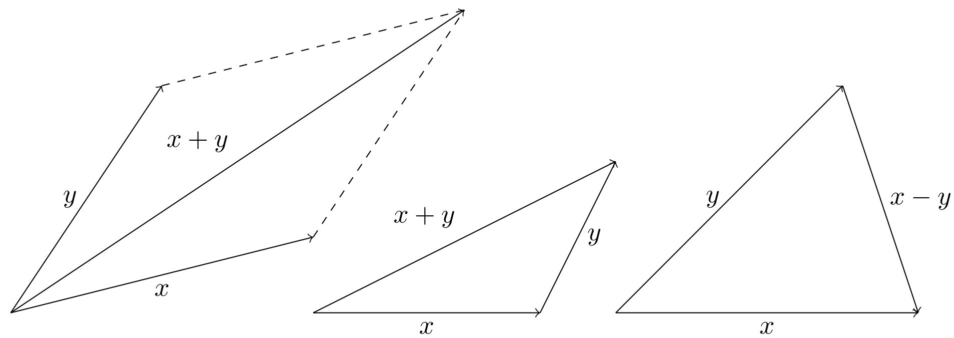

Figura 3.2: Somma e sottrazione

E' definita anche la moltiplicazione scalare per vettore αx, che produce uno 'scalamento' del vettore lungo la direzione.

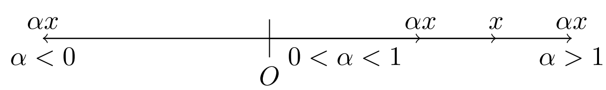

Figura 3.3: Moltiplicazione vettore per scalare

$< x_1, ..., x_k >= \{\alpha_1 x_1 + ... + \alpha_k x_k | \alpha_1 ... \alpha_k \in \mathbb{R}\}$ è detto *sottospazio vettoriale* generato da $x_1, ..., x_k$. Ad esempio per $k = 1$ si ha la retta $\langle x_1 \rangle$ su cui giace x_1 (se non nullo), per $k = 2$ si ha il piano $< x_1, x_2 >$ su cui giacciono x_1 e x_2 (se non allineati); k è la dimensione del sottospazio (salvo degenerazioni, che avvengono quando $x_1, ..., x_k$ sono *linearmente dipendenti*).

3.1.2 Prodotto scalare

Dati $x, y \in \mathbb{R}^n$ il *prodotto scalare* $x \cdot y$ o $< x, y >$ è dato da $\sum_{i=1}^{n} x_i y_i$. Con la convenzione che i vettori siano colonne si puo' definire anche come $x^T y$.
$x^T y = \|x\|_2 \|y\|_2 \cos\theta$

Esempio Presi tre vettori $x = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ $y = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ $z = \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}$
$x^T y = -2 + 0 + 0 = -2$
$\|x\|_2 = \sqrt{1+4} = \sqrt{5}$, $\|y\|_2 = \sqrt{4+1} = \sqrt{5}$
$\cos\theta = \frac{x^T y}{\|x\|_2 \|y\|_2} = \frac{-2}{\sqrt{5}\sqrt{5}} = -\frac{2}{5}$
$x^T z = 3 + 12 = 15$
$\|x\|_2 = \sqrt{5}$, $\|z\|_2 = 3\sqrt{5}$
$\cos\theta = \frac{15}{15} = 1$
Si possono dedurre informazioni geometriche:

- $x^T y > 0 \Rightarrow \theta < \frac{\pi}{2}$

- $x^T y < 0 \Rightarrow \theta > \frac{\pi}{2}$
- $x^T y = 0 \Rightarrow \theta = \frac{\pi}{2}$ (x e y sono ortogonali)

Proiezione Dati due vettori x e y, trovare la *proiezione* p di y su x.
$\cos\theta = \frac{\|p\|_2}{\|y\|_2} \Rightarrow \|p\|_2 = \frac{x^T y}{\|x\|_2}$
$\hat{p} = \hat{x} = \frac{x}{\|x\|_2} \Rightarrow p = \|p\|_2\, \hat{p} = \frac{x^T y}{\|x\|_2} \cdot \frac{x}{\|x\|_2} = \frac{x^T y}{x^T x} x$ ($x^T x$ è la lunghezza al quadrato)
$x = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ $y = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$
$x^T y = -2$, $x^T x = 5$
$p = -\frac{2}{5} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{2}{5} \\ -\frac{4}{5} \\ 0 \end{pmatrix}$

3.1.3 Basi

$x_1 ... x_k \in \mathbb{R}^n$ si dicono *base ortogonale* se $x_i^T x_j = 0\ \forall i \neq j$, si dicono *base ortonormale* se $x_i^T x_j = \delta_{ij}$ (delta di Kronecker).

La *base canonica* è costituita da vettori del tipo $e_k = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ con l'unico 1 in $k-$esima posizione.

Teorema Se $x_1 ... x_k$ è una base ortogonale di vettori non nulli, allora $x_1 ... x_k$ sono vettori linearmente indipendenti.
Dimostrazione: Considero una combinazione lineare $v = \lambda_1 x_1 + ... + \lambda_k x_k = 0$, la tesi è che $\lambda_i = 0$.
Calcolo $x_i^T v = x_i (\lambda_1 x_1 + ... + \lambda_k x_k) = \lambda_1 x_i^T x_1 + \lambda_2 x_i^T x_2 + ... + \lambda_i x_i^T x_i + ... + \lambda_k x_i^T x_k =$
$= \lambda_i x_i^T x_i = \lambda_i \|x_i\|_2^2$
$x_i^T v = \lambda_i \|x_i\|_2^2$ Il primo membro è nullo perchè $v = 0$ per ipotesi, quindi o $\|x_i\|_2^2 = 0$ (il che è falso, perchè $x_i = 0$ contraddice l'ipotesi) oppure $\lambda_i = 0$, questo è vero, il che chiude la dimostrazione.

Teorema $x_1 ... x_k$ base ortogonale, $u \in < x_1, ..., x_k > \Rightarrow u = \sum_{i=1}^{k} \frac{u^T x_i}{x_i^T x_i} x_i$ (u è la somma delle sue proiezioni sugli x_i)

3.2 Matrici

Consideriamo una matrice $A \in \mathbb{R}^{m \times n}$ come una trasformazione del vettore in input x, $A : \mathbb{R}^n \to \mathbb{R}^m$, $x \mapsto Ax$

Esempio 1 $A = 0$

$x \mapsto 0 \cdot x = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$

E' la trasformazione che prende ogni vettore e lo collassa nell'origine.

Esempio 2 $A = I \in \mathbb{R}^{n \times n}$

$x \in \mathbb{R}^n \to Ix = x$

Trasformazione identica, lascia invariati i valori.

Esempio 3 $A = 2I = \begin{pmatrix} 2 & & 0 \\ & \ddots & \\ 0 & & 2 \end{pmatrix}$

$x \mapsto 2Ix = 2x$

Omotetia, dilatazione del vettore.

Esempio 4 $A = -I$

$x \mapsto -Ix = -x$

Simmetria rispetto all'origine

Esempio 5 $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 3}$

$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \to Ax = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Ha l'effetto di schiacciare il vettore solo sul piano orizzontale $< e_1, e_2 >$ levando la terza dimensione (foto dall'alto).

Esempio 6 $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 3}$

$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to Ax = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$

Il vettore viene immerso in uno spazio tridimensionale sul piano $x_3 = 0$.

3.3 Applicazioni lineari

Le trasformazioni associate a matrici vengono dette *trasformazioni* (o *applicazioni*) *lineari*, e vengono caratterizzate da due proprietà algebriche:

- $A(x + y) = Ax + Ay$ (la somma degli input corrisponde alla somma degli output)
- $A(\alpha x) = \alpha Ax$ (ad input scalato corrisponde output scalato)

3.3.1 Significato geometrico delle operazioni fra matrici

Somma $A, B \Longrightarrow A + B$
$x \mapsto (A + B)\, x = Ax + Bx$
Gli effetti di A e B si sommano.

Prodotto per scalare $A \Longrightarrow \alpha A$
$x \mapsto (\alpha A)\, x = \alpha\, (Ax)$
Gli effetti vengono scalati.

Prodotto matriciale $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times p} \Longrightarrow A \cdot B \in \mathbb{R}^{m\times p}$
$x \in \mathbb{R}^p \to (A \cdot B)\, x = A\,(Bx)$
Geometricamente è come comporre le trasformazioni (composizione di A dopo B).
$x \in \mathbb{R}^p \longrightarrow Bx \in \mathbb{R}^n \longrightarrow (A \cdot B)\, x \in \mathbb{R}^m$
Esempio
$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$
$A \cdot B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (immergere e poi fotografare equivale a riprodurre la scena di partenza)
$B \cdot A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ (fotografare e poi immergere equivale ad azzerare l'altezza):
$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto (BA)\, x = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$

Matrice inversa $A \in \mathbb{R}^{n\times n}$, $\det A \neq 0 \Rightarrow A^{-1}$corrisponde alla trasformazione inversa (dato l'output ottiene l'input)

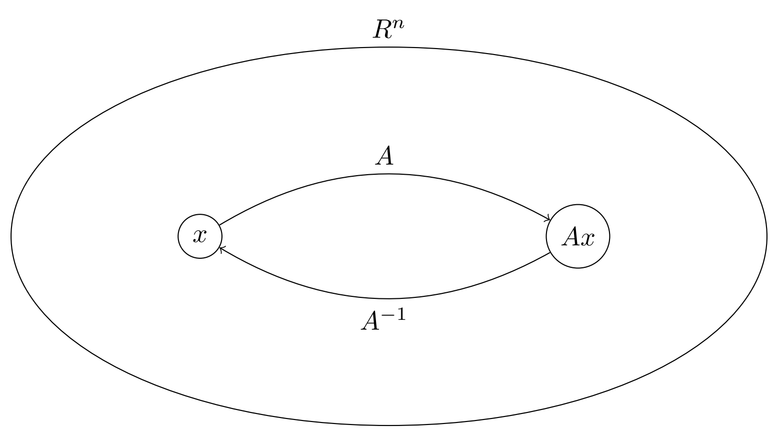

Figura 3.4: Rappresentazione dell'inversa

3.3.2 Nucleo e Immagine di una matrice

Si definisce *nucleo* di A ($\ker A$)

$$\mathfrak{N}(A) := \{x \in \mathbb{R}^n : \ Ax = 0\}$$

Si definisce *immagine* di A ($\mathrm{Im}A$)

$$\mathfrak{R}(A) := \{y \in \mathbb{R}^m : \ \exists x \in \mathbb{R}^n : \ y = Ax\}$$

$$A \text{ surgettiva } \iff \mathfrak{R}(A) = \mathbb{R}^m$$

Teorema Una matrice è iniettiva se e solo se il nucleo è l'origine (*nucleo banale*)

$$A \text{ iniettiva } \iff \mathfrak{N}(A) = \{0\}$$

Dimostrazione dell'implicazione a destra: Per assurdo supponiamo che $\exists x \neq 0 : x \in \mathfrak{N}(A) \Rightarrow Ax = 0 = A \cdot 0$ che contraddice l'ipotesi di iniettività.
Dimostrazione dell'implicazione a sinistra: Considero $v \neq w \in \mathbb{R}^n$, suppongo per assurdo $Av = Aw$ quindi $A(v - w) = 0 \Rightarrow v - w \in \mathfrak{N}(A) = \{0\} \Rightarrow v - w = 0 \Rightarrow v = w$

Proprietà $\mathfrak{N}(A)$ e $\mathfrak{R}(A)$ sono sottospazi vettoriali (cioè presi due vettori all'interno di uno dei due, la somma tra loro e il prodotto per uno scalare rimangono al loro interno).

3.3.2.1 Dimensioni e base di Nucleo e Immagine

Prendiamo $\mathfrak{R}(A)$, si definisce la sua dimensione *rango* o *caratteristica* della matrice: $\mathrm{rk}(A) := \dim \mathfrak{R}(A)$.
Consideriamo la base canonica di $\mathbb{R}^n$ $\{e_1, ..., e_n\}$ e un vettore appartenente ad $\mathbb{R}^n$, questo è scrivibile come $x = x_1e_1 + x_2e_2 + ... + x_ne_n$.
$y = Ax = A(x_1e_1 + ... + x_ne_n) = x_1(Ae_1) + x_2(Ae_2) + ... + x_n(Ae_n)$
Il vettore canonico e_j è formato da tanti zeri ed un uno nella posizione dell'indice, pertanto, moltiplicandolo per la matrice, restituisce la corrispondente colonna della matrice (analogamente $e_j^T A$ restituisce le righe): pertanto $\mathfrak{R}(A) = \langle \text{colonne di } A \rangle$.
Si deduce che $\mathrm{rk}(A) \leqslant \min(m, n)$.

$$A \text{ iniettiva } \iff \dim \mathfrak{N}(A) = 0$$

$$A \text{ surgettiva } \iff \dim \mathfrak{R}(A) = \mathrm{rk}(A) = m$$

Teorema Per ogni matrice, la somma della dimensione del nucleo e del rango è uguale al numero delle colonne.

$$\forall A \in \mathbb{R}^{m \times n} \ \dim \mathfrak{N}(A) + \mathrm{rk}(A) = n$$

Corollario 1 $\mathfrak{N}(A) = 0 \iff \mathrm{rk}(A) = n$

Corollario 2 Se $m = n$, $\mathfrak{N}(A) = 0 \Rightarrow A$ è iniettiva e surgettiva $\Rightarrow \exists A^{-1} \Rightarrow \det A \neq 0$

Corollario 3 Se $m < n$ $\mathrm{rk}(A) < n \Rightarrow \dim \mathfrak{N}(A) > 0 \Rightarrow \mathfrak{N}(A) \neq \{0\}$

Esempi Calcolo di nucleo e immagine
1) $A = 0$ $\mathfrak{N}(A) = \mathbb{R}^n$, $\mathfrak{R}(A) = \{0\}$, $\mathrm{rk}(A) = 0$
2) $A = I$ $\mathfrak{N}(A) = \{0\}$, $\mathfrak{R}(A) = \mathbb{R}^n$, $\mathrm{rk}(I) = n$
3) $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ è surgettiva ma non iniettiva, infatti $\mathfrak{N}(A) = \left\{ x = \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} \mid x_3 \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$ $(\dim \mathfrak{N}(A) = 1)$, $\mathfrak{R}(A) = \mathbb{R}^2$
4) $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ $\mathfrak{N}(A) = \{0\}$, $\mathfrak{R}(A) = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_3 = 0 \right\} \Rightarrow \mathrm{rk}(A) = 2$ quindi $\mathfrak{R}(A) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle$

3.4 Isometrie

Le isometrie sono applicazioni lineari che conservano le distanze.
Considero due punti x e y, li vediamo come punte dei vettori partenti da un punto O, pertanto per ottenere la distanza, possiamo valutare $\|x-y\|_2$. Data una matrice A, essa corrisponde ad una isometria se $\|x-y\|_2 = \|Ax - Ay\|_2 = \|A(x-y)\|_2$, ovvero se la lunghezza di un vettore è uguale alla lunghezza del vettore moltiplicato a sinistra per la matrice:

$$A \text{ isometria } \Leftrightarrow \forall v \in \mathbb{R}^n \ \|Av\|_2 = \|v\|_2$$

Lemma Considerato un vettore canonico e_j, sappiamo che Ae_j fornisce la j-esima colonna di A e che $e_j^T A$ fornisce la j-esima riga: un elemento a_{ij} è quindi individuato dal prodotto $e_i^T A e_j$.

3.4.1 Matrici ortogonali

Una matrice quadrata si dice *ortogonale* se $A^T A = I$.

Osservazione A ortogonale $\Longleftrightarrow A^{-1} = A^T \Longleftrightarrow AA^T = I \Longleftrightarrow A$ ha colonne ortonormali
Le matrici ortogonali sono l'alter ego algebrico delle isometrie.

Teorema A ortogonale $\Longleftrightarrow A$ isometria

Dimostrazione Implicazione a destra: Essendo A ortogonale, applico la definizione e l'osservazione ottenendo la proprietà geometrica.
$\forall v \in \mathbb{R}^n : \|Av\|_2^2 = (Av)^T (Av) = v^T \left(A^T A\right) v = v^T I v = v^T v = \|v\|_2^2$
Implicazione a sinistra: Considero la base ortonormale $e_1..e_n$ $(e_i^T e_j = \delta_{ij})$, se A è una isometria allora anche $Ae_1...Ae_n$ è una base ortonormale, quindi $(Ae_i)^T (Ae_j) = \delta_{ij}$ ovvero, per il lemma, $e_i^T \left(A^T A\right) e_j = \left(A^T A\right)_{ij} = \delta_{ij} \Rightarrow A^T A = I$.

Esempi

- I è ortogonale ($I^{-1} = I^T I$), ha colonne ortonormali (il prodotto scalare fra le colonne è zero).
- $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ non è ortogonale, il prodotto scalare fra le colonne è 3.
- $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ non è ortogonale; se vado a normalizzare (la lunghezza è $\sqrt{2}$) le colonne pero' diventa ortogonale (in quanto per costruzione le lunghezze valgono 1), ovvero $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$.
- Considero una matrice di permutazione (ovvero la matrice che ha l'effetto di permutare il vettore a cui è applicata), che è ortogonale (essendo composta da vettori della base ortonormale permutati), $\Pi \in \mathbb{R}^{5\times 5}$ e un vettore $v \in \mathbb{R}^5$: $(e_1|e_4|e_5|e_3|e_2) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$ e $\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix}$, il loro prodotto è $\Pi v = \begin{pmatrix} v_1 \\ v_5 \\ v_4 \\ v_2 \\ v_3 \end{pmatrix}$.
- La matrice ottenuta tramite il metodo di Gauss senza scambi è $A = LU$, se ci sono scambi (pivoting) si considera la matrice di permutazione Π che riassume gli scambi effettuati e si ha $\Pi A = LU$

Corollario del teorema Se A, B sono matrici ortogonali allora $A \cdot B$ e A^T sono anch'esse ortogonali.
Il determinante di una matrice ortogonale può solo valere ± 1 infatti $1 = \det\left(A^T A\right) = \det\left(A^T\right)\det A = (\det A)^2 \Rightarrow \det A = \pm 1$

3.4.2 Rotazioni in $\mathbb{R}^2$

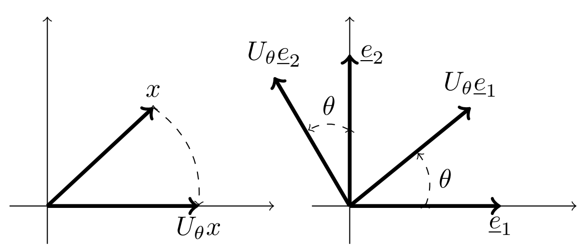

Figura 3.5: Rotazioni

Dato un vettore in input, vorrei ruotarlo in senso antiorario. Considero i vettori canonici e_1 ed e_2 e i corrispondenti ruotati di θ $U_\theta e_1$ e $U_\theta e_2$:

$$U_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Verifichiamo l'ortogonalità:
$U_\theta^T U_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos^2\theta + \sin^2\theta & \cos\theta(-\sin\theta) + \sin\theta\cos\theta \\ 0 & 1 \end{pmatrix} = I$
La rotazione inversa è quindi l'inversa della matrice (essendo ortogonale è la trasposta):
$U^{-1} = U^T = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = U_{-\theta}$

Dato $x \in \mathbb{R}^2$, cerco U_θ tale che $U_\theta x = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$ con $\alpha > 0$ opportuno.

$U_\theta x = \begin{pmatrix} x_1\cos\theta - x_2\sin\theta \\ x_1\sin\theta + x_2\cos\theta \end{pmatrix}$, pongo $x_1\sin\theta + x_2\cos\theta = 0 \Rightarrow \frac{\sin\theta}{\cos\theta} = \tan\theta = -\frac{x_2}{x_1}$, quindi si ottiene $\cos\theta = \frac{x_1}{\|x\|_2} = \frac{x_1}{\sqrt{x_1^2+x_2^2}}$ e $\sin\theta = -\frac{x_2}{\|x\|_2} = -\frac{x_2}{\sqrt{x_1^2+x_2^2}}$. Quindi $U_\theta x = \begin{pmatrix} \|x\|_2 \\ 0 \end{pmatrix}$.

3.4.3 Rotazioni di Givens $G(i,j,\theta)$

Sono caratterizzate dagli indici delle direzioni che individuano il piano di rotazione, e dall'angolo θ di rotazione. La matrice che descrive la rotazione è una matrice identità eccetto nelle posizioni di indice i e j. Ha l'effetto di una rotazione nelle direzioni e_i e e_j, lasciando invariate le altre
$G_{ii} = c$, $G_{ij} = -s$, $G_{ji} = s$, $G_{jj} = c$

$$G(i,j,\theta)\,x = \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ cx_i - sx_j \\ x_{i+1} \\ \vdots \\ x_{j-1} \\ sx_j + cx_j \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix}$$

Chiamando $y := G(i,j,\theta)\,x$,

posso ottenere $y = \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ \sqrt{x_i^2 + x_j^2} \\ x_{i+1} \\ \vdots \\ x_{j-1} \\ 0 \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix}$ scegliendo $c = \frac{x_i}{\sqrt{x_i^2+x_j^2}}$, $s = -\frac{x_j}{\sqrt{x_i^2+x_j^2}}$

Esempio $x = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 3 \end{pmatrix}$

Come primo passo voglio azzerare x_5 facendo perno su x_1(nel senso che solo $x_1$cambierà):

$G(1,5,\theta) = \begin{pmatrix} c & 0 & 0 & 0 & -s \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ s & 0 & 0 & 0 & c \end{pmatrix}$ Calcolo $c = \frac{1}{\sqrt{1+9}} = \frac{1}{\sqrt{10}}$ e $s = -\frac{3}{10}$ ottenendo $x' = G(1,5,\theta)\, x = \begin{pmatrix} \sqrt{10} \\ -1 \\ 2 \\ 0 \\ 0 \end{pmatrix}$.

Si può procedere per successive rotazioni per azzerare tutti gli elementi tranne uno mantenendo sempre lo stesso perno.

$G(1,2,\theta') = \begin{pmatrix} c & -s & 0 & 0 & 0 \\ s & c & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ $c = \frac{\sqrt{10}}{\sqrt{10+1}} = \frac{\sqrt{10}}{\sqrt{11}}$, $s = \frac{1}{\sqrt{11}}$, $x'' = G(1,2,\theta')\, x' = \begin{pmatrix} \sqrt{11} \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$

$G(1,3,\theta'') = \begin{pmatrix} c & 0 & -s & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ s & 0 & c & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ $c = \frac{\sqrt{11}}{\sqrt{15}}$, $s = -\frac{2}{\sqrt{15}}$, $y = G(1,3,\theta'')\, x'' = \begin{pmatrix} \sqrt{15} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

Abbiamo quindi in realtà composto varie rotazioni: $G(1,3,\theta'')\, G(1,2,\theta')\, G(1,5,\theta)\, x = y$

Essendo tutte isometrie, il risultato finale dovrebbe corrispondere alla lunghezza di x: $\|x\|_2 = \sqrt{1+1+4+9} = \sqrt{15}$

Abbiamo considerato casi per cui $i < j$. Cosa cambia se $j < i$? Cambia la visualizzazione della matrice, vengono invertite le posizioni di s (che va in alto a destra) e $-s$ (che va in basso a sinistra).

Esercizi Calcolare una sequenza di rotazioni di Givens che portano il primo vettore nel secondo:

1) $\begin{pmatrix} -2 \\ 0 \\ 3 \\ 1 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} \alpha \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

2) $\begin{pmatrix} 0 \\ 1 \\ 4 \\ -1 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 \\ 0 \\ \beta \\ 0 \end{pmatrix}$

3.4.3.1 Fattorizzazione QR tramite rotazioni

E' definibile un analogo del metodo di Gauss?
Consideriamo una matrice 6×4 e scegliamo al primo passo l'elemento a_{11} come perno e l'elemento a_{21} sarà quello da azzerare: $G_{12} := G(1, 2, \theta)$, $G_{12}A$ è il prodotto che tocca le righe in cui giacciono perno ed elemento da azzerare; si prosegue sempre mantenendo lo stesso perno, alla fine delle rotazioni sulla colonna scelta, siamo nella situazione $G_{16}G_{15}G_{14}G_{13}G_{12}A$ in cui la prima colonna è tutta azzerata sotto la diagonale. Passo alla seconda colonna usando come perno a_{22} pertanto verranno applicate le rotazioni per $i = 2$ e $j = 3 \to 6$: siccome la rotazione applicata agli zeri produce zeri, la prima colonna non viene 'sporcata'. Analogamente si applica anche alla terza e alla quarta colonna, ottenendo: $G_{46}G_{45}G_{36}G_{35}G_{34}G_{26}G_{25}G_{24}G_{23}G_{16}G_{15}G_{14}G_{13}G_{12}A$ avrà una forma triangolare superiore nella parte quadrata (definita R), più altre due righe di zeri.
Il costo dell'algoritmo di Gauss per una matrice $m \times n$ è $\frac{1}{2}n^2\left(m - \frac{n}{3}\right)$, mentre per Givens è $2n^2\left(m - \frac{n}{3}\right)$ ovvero quattro volte l'eliminazione gaussiana.
Analogamente al metodo di Gauss che costruisce due matrici tali che $A = LU$, il risultato col metodo di Givens è $A = Q\begin{pmatrix} R \\ 0 \end{pmatrix}$. Chiamata infatti G la matrice prodotto delle matrici di rotazione, essa sarà una matrice ortogonale, da cui discende che $G^{-1} = G^T =: Q$, $A = Q\begin{pmatrix} R \\ 0 \end{pmatrix}$. Il procedimento è noto come *fattorizzazione QR* (analogamente a quello di Gauss chiamato fattorizzazione LU).
Il metodo di Gauss è stabile (a patto di fare il pivoting) tranne in qualche caso, Givens è invece sempre stabile.

Proprietà Si enunciano due teoremi.
Teorema 1: Sia $A \in \mathbb{R}^{m \times n}$ con $m \geqslant n$, se $\mathrm{rk}(A) = n \Rightarrow \det R \neq 0$ (la matrice R è invertibile).
Teorema 2: Se $\mathrm{rk}(A) = n \Rightarrow \mathfrak{R}(A) = \mathfrak{R}(Q_0)$ dove Q_0 è la parte della matrice $Q \in \mathbb{R}^{m \times m}$ che comprende solo le prime n colonne $q_1, ..., q_n$.
Corollario: $q_1, ..., q_n$ è una base ortonormale di $\mathfrak{R}(A)$

3.4.4 Riflessioni di Householder

La matrice che definisce una *riflessione di Householder* è $P = I - 2ww^T$ con tutte matrici quadrate e w è un vettore colonna di lunghezza 1 ($\|w\|_2 = 1$)[1].
Per ricavare l'intepretazione geometrica calcoliamo $Px \in \mathbb{R}^m$ dato $x \in \mathbb{R}^m$.
$Px = \left(I - 2ww^T\right)x = Ix - 2w\left(w^Tx\right) = x - 2\left(w^Tx\right)w$, quest'ultimo termine, essendo $w^Tw = 1$ lo si può vedere come $2\frac{w^Tx}{w^Tw}w$ che è il doppio della proiezione di x su w: Px si trova togliendo ad x due volte la sua proiezione su w (quindi è x riflesso rispetto all'iperpiano perpendicolare a w)

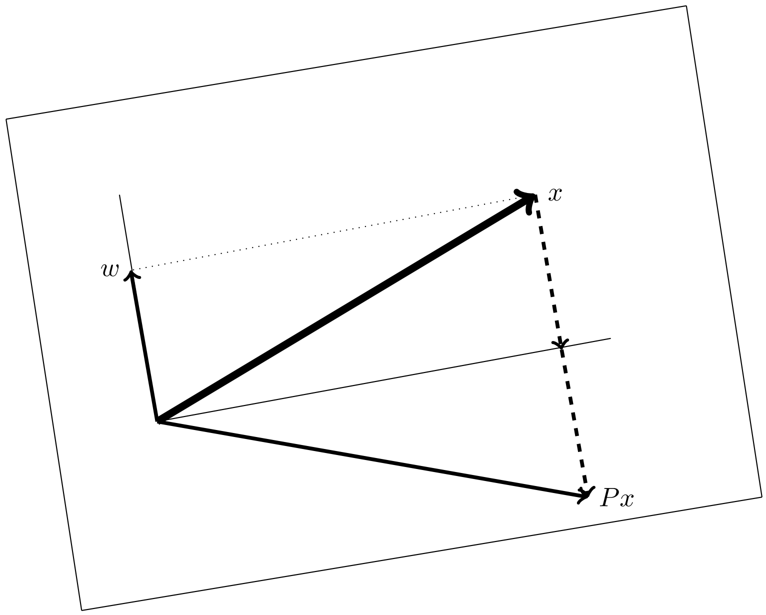

Figura 3.6: Riflessione di Householder

Proposizione $P = P^T$, $P^2 = I$ ($P^TP = I \Rightarrow P$ è ortogonale)
Dimostrazione:
$P^T = \left(I - 2ww^T\right)^T = I^T - \left(2ww^T\right)^T = I - 2\left(w^T\right)^Tw^T = P$
$P^2 = \left(I - 2ww^T\right)\left(I - 2ww^T\right) = I - 2ww^T - 2ww^T + 4w\left(w^Tw\right)w^T = I - 4ww^T + 4ww^T = I$

3.4.4.1 Fattorizzazione QR tramite riflessioni

Usando le riflessioni di Householder si può azzerare un vettore in un colpo solo?
Problema: Dato $x \in \mathbb{R}^n$ si cerca w (quindi P) tale che $Px = \alpha e_1$. Essendo P ortogonale, $\|x\|_2 = \|Px\|_2 = \|\alpha e_1\|_2 = |\alpha|$, scegliamo ad esempio $\alpha = +\|x\|_2$: $Px = x - 2\left(w^Tx\right)w$ deve essere uguale ad $\alpha e_1 = \|x\|_2 e_1$.
$x - 2\left(w^Tx\right)w = \|x\|_2 e_1 \Rightarrow 2\left(w^Tx\right)w = x - \|x\|_2 e_1 =: u$
Confrontando i versori (w lo è già), $w = \hat{u} = \frac{u}{\|u\|_2}$

[1] Il prodotto riga per colonna restituisce uno scalare, mentre il prodotto colonna per riga restituisce una matrice.

Algoritmo 3.1 Algoritmo di fattorizzazione

$u := x - \|x\|_2 \, e_1$

$w := \frac{u}{\|u\|_2}$

$P := I - 2ww^T$

Gli ultimi due passi si riassumono come $P = I - 2\frac{uu^T}{u^T u}$ (nella frazione la parte di sopra è una matrice, quella di sotto uno scalare)

Esempio $x = \begin{pmatrix} 3 \\ 1 \\ 5 \\ 1 \end{pmatrix}$, $\|x\|_2 = \alpha = \sqrt{9+1+25+1} = 6 \Rightarrow Px = \begin{pmatrix} 6 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ (come verifica)

$u = x - \|x\|_2 \, e_1 = \begin{pmatrix} -3 \\ 1 \\ 5 \\ 1 \end{pmatrix} \Rightarrow P = I - \frac{2}{u^T u} uu^T$

$u^T u = \begin{pmatrix} -3 & 1 & 5 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ 1 \\ 5 \\ 1 \end{pmatrix} = 36$

$uu^T = \begin{pmatrix} -3 \\ 1 \\ 5 \\ 1 \end{pmatrix} \begin{pmatrix} -3 & 1 & 5 & 1 \end{pmatrix} = \begin{pmatrix} 9 & -3 & -15 & -3 \\ -3 & 1 & 5 & 1 \\ -15 & 5 & 25 & 5 \\ -3 & 1 & 5 & 1 \end{pmatrix}$ Si vede che anche questa matrice è simmetrica, come tutta la riflessione di Householder.

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \frac{1}{18} \begin{pmatrix} 9 & -3 & -15 & -3 \\ -3 & 1 & 5 & 1 \\ -15 & 5 & 25 & 5 \\ -3 & 1 & 5 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{6} & \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{17}{18} & -\frac{5}{18} & -\frac{1}{18} \\ \frac{5}{6} & -\frac{5}{18} & -\frac{7}{18} & -\frac{5}{18} \\ \frac{1}{6} & -\frac{1}{18} & -\frac{5}{18} & \frac{17}{18} \end{pmatrix}$$

Fattorizzazione QR Consideriamo la matrice 6×4 da triangolarizzare.
Al primo passo, troviamo P_1 in modo che dia la prima colonna con α in prima posizione e cinque zeri.
Al secondo passo, consideriamo la seconda colonna ridotta a cinque elementi (dall'elemento a_{22} in giù), otteniamo una matrice P_2 con dimensione 5×5, pertanto la inseriamo in una matrice $\tilde{P}_2$ che la contiene in questo modo: $\tilde{P}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & & & & & \\ 0 & & & & & \\ 0 & & & P_2 & & \\ 0 & & & & & \\ 0 & & & & & \end{pmatrix}$. Si dimostra che questa matrice così costruita azzera gli elementi della seconda colonna come richiesto: $\tilde{P}_2 P_1 A$. Analogamente si procede per le altre colonne, ottenendo $\tilde{P}_4 \tilde{P}_3 \tilde{P}_2 P_1 A = \begin{pmatrix} R \\ 0 \end{pmatrix}$.

3.4.4.2 Differenze con gli altri due metodi

Metodo	Costo	Risultato
Gauss	$\frac{1}{2}n^2\left(m-\frac{n}{3}\right)$	$A = LU$
Givens	$2n^2\left(m-\frac{n}{3}\right)$	$A = Q\begin{pmatrix} R \\ 0 \end{pmatrix}$
Householder	$n^2\left(m-\frac{n}{3}\right)$	$A = Q\begin{pmatrix} R \\ 0 \end{pmatrix}$

Tabella 3.1: Confronto fra metodi di triangolarizzazione

C'è una differenza fra Givens ed Householder: Givens azzera un elemento alla volta, quindi vede se ci sono degli zeri in qualche elemento (*matrici sparse*), mentre Householder non se ne accorge. Si può programmare il metodo di Givens in modo tale che salti operazioni nel caso trovi uno zero.

3.4.5 Esercizi

Applicare Givens ed Householder per azzerare i seguenti vettori in modo che $x \Rightarrow \begin{pmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{pmatrix}$:

$x = \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}$

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

$x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$