PROGRAMMA DI
ISTITUZIONI DI MATEMATICHE I

Corso di Laurea in ARCHITETTURA
A.A. 2006/2007 (Alberto PARMEGGIANI)

Indice

1 Algebra lineare

1.1 Vettori nel piano $\mathbb{R}^2$, nello spazio tridimensionale $\mathbb{R}^3$, e nello spazio n-dimensionale $\mathbb{R}^n$

Si consideri l'insieme dei **vettori** del piano

$$\mathbb{R}^2 = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} ; x_1, x_2 \in \mathbb{R} \right\}.$$

Un punto (x_1, x_2) del piano cartesiano è identificato con un elemento $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$. Dati $\alpha \in \mathbb{R}$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ e $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, si hanno (per definizione) le operazioni di "prodotto per uno scalare" e di "somma tra vettori"

$$\alpha x = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \end{bmatrix} \qquad \text{e} \qquad x + y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}.$$

Il vettore nullo è $0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (ed è identificato con il punto $(0, 0)$).
Due vettori $v, w \neq 0$ si dicono *paralleli* se sono uno un multiplo dell'altro, cioè se esiste $\alpha \in \mathbb{R}$ con $\alpha \neq 0$ tale che $v = \alpha w$. Si può inoltre definire un'operazione, detta *prodotto scalare*, che associa a due vettori x, y un numero come segue

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 \,.$$

Talvolta si usa il simbolo $x \cdot y$ al posto di $\langle x, y \rangle$. Due vettori v e w si dicono *ortogonali* (in simboli $v \perp w$) se $\langle v, w \rangle = 0$.

La *norma* (o lunghezza) del vettore x è per definizione

$$\|x\| = \sqrt{x_1^2 + x_2^2} = \sqrt{\langle x, x \rangle}\,.$$

Tutte le precedenti operazioni si estendono naturalmente allo spazio dei vettori tridimensionali

$$\mathbb{R}^3 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} ; x_1, x_2, x_3 \in \mathbb{R} \right\},$$

ed anche allo spazio dei vettori n-dimensionali

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} ;\; x_1, \ldots, x_n \in \mathbb{R} \right\}.$$

In questo caso dati $\alpha \in \mathbb{R}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$ e $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$

$$\alpha x = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \alpha x_3 \\ \vdots \\ \alpha x_n \end{bmatrix}, \quad x + y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

e

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3 + \ldots + x_n y_n \,.$$

Inoltre,

$$\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2 + \ldots + x_n^2}\,.$$

- Si ha che $\mathbb{R}^n$, munito delle operazioni di "prodotto per uno scalare" e di "somma tra vettori", risulta essere uno *spazio vettoriale.*

Sia k un numero naturale, per *combinazione lineare* di k vettori $v_1, v_2, \cdots v_k$ di $\mathbb{R}^n$ si intende una qualsiasi espressione della forma $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$, dove $\alpha_1, \alpha_2, \cdots, \alpha_k$ sono k numeri reali.

Dati k vettori del piano $v_1, v_2, \cdots v_k$, essi si dicono *linearmente indipendenti* se dal fatto che $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$ segue che $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$, cioè quando vale

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0 \Longrightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0.$$

Qualora non siano linearmente indipendenti i vettori si dicono *linearmente dipendenti.* Si ha che

- due vettori di $\mathbb{R}^n$ sono linearmente dipendenti se e solo se sono uno un multiplo dell'altro;
- **tre** vettori di $\mathbb{R}^2$ sono **sempre** linearmente dipendenti;
- **quattro** vettori di $\mathbb{R}^3$ sono **sempre** linearmente dipendenti,
- e, più in generale, $n+1$ vettori di $\mathbb{R}^n$ sono **sempre** linearmente dipendenti.
- L'importanza di avere una famiglia $\{v_1, \ldots, v_k\}$ di vettori linearmente indipendenti di $\mathbb{R}^n$ risiede nel fatto che se un vettore è combinazione lineare dei $v_1, \ldots, v_k$, allora i coefficienti della combinazione lineare sono **unici**.

È importante osservare che i vettori possono essere pensati tanto come *vettori colonna* $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ quanto come *vettori riga* $(x_1, x_2, \ldots, x_n)$. In generale si utilizzano indifferentemente una o l'altra scrittura.

1.2 Matrici ed operazioni su matrici

Dati due numeri naturali m ed n (si pensi ad esempio ad $m, n \in \{1, 2, 3\}$) una matrice $m \times n$ è una tabella con m righe ed n colonne. Ad esempio,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

è una generica matrice 3×2.

Se $m = n$ la matrice si dice *quadrata.*

Date due matrici $m \times n$, A e B, la somma $A + B$ è (per definizione) la matrice $m \times n$ ottenuta sommando A con B "componente per componente". Ad esempio, se

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \qquad \text{e} \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

allora

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \\ a_{31} + b_{31} & a_{32} + b_{32} \end{bmatrix}.$$

La moltiplicazione di una matrice A per uno scalare $\alpha \in \mathbb{R}$ è per definizione la matrice avente per componenti quelle di A moltiplicate per lo scalare α. Ad esempio, se A è la matrice scritta sopra,

$$\alpha A = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} \\ \alpha a_{21} & \alpha a_{22} \\ \alpha a_{31} & \alpha a_{32} \end{bmatrix}.$$

Date una matrice $m \times n$, A, ed una matrice $n \times k$, B, il prodotto AB è la matrice $m \times k$ che ha nel posto determinato dall'incrocio tra la i-esima riga e la j-esima colonna il prodotto scalare tra la i-esima riga della matrice A e la j-esima colonna della matrice B (si opera cioè un prodotto "righe-per-colonne"). In particolare se

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad \text{e} \qquad B = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$$

allora

$$AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} \end{bmatrix}.$$

Quindi tramite il prodotto di una matrice $m \times n$ per un matrice (vettore colonna) $n \times 1$, il prodotto di una matrice $m \times n$, A, per una matrice $n \times k$, $B = [b_1|b_2|\ldots|b_k]$ dove $b_1, b_2, \ldots, b_k \in \mathbb{R}^n$ sono le colonne di B, si ottiene operando k volte il prodotto righe-per-colonne della matrice A per ciascuna colonna b_j $(j = 1, \ldots, k)$ di B :

$$AB = [Ab_1|Ab_2|\ldots|Ab_k]$$

che risulta perciò essere una matrice $m \times k$.

- Si ha la seguente proprietà di *linearità*: data la matrice $m \times n$, A, per ogni scalare $\alpha, \beta \in \mathbb{R}$ e vettori colonna (matrici $n \times 1$) $x, y \in \mathbb{R}^n$ si ha

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay.$$

- La matrice *identità* I è la matrice con tutti 1 sulla diagonale principale e 0 altrove.

Osservazione. *Ricordiamo ancora che una matrice $n \times 1$, x, si identifica con un vettore $x \in \mathbb{R}^n$.*

1.3 Metodi di risoluzione di sistemi di equazioni lineari

Un sistema lineare di m equazioni in n incognite si rappresenta con una matrice $m \times n$ (la matrice dei coefficienti), A, un vettore b di $\mathbb{R}^m$ (il vettore dei "termini noti"), ed un vettore x di $\mathbb{R}^n$ (il vettore delle incognite). Cioè il sistema $[A|b]$ si può rappresentare nella forma

$$Ax = b\,.$$

Per risolvere il sistema si può ricorrere a sostituzioni successive oppure alla *riduzione di Gauss-Jordan* che può essere descritta come segue.

- Si costruisce la matrice $[A|b]$;
- si opera sulla matrice $[A|b]$ agendo con *operazioni elementari* **sulle righe** della matrice fino a ridurre la matrice alla forma $[S|c]$, dove S è una forma *ridotta a scala* di A;
- si risolve il sistema partendo dall'ultima equazione in $[S|c]$ fino ad arrivare alla prima (con sostituzioni successive).

È importante osservare che procedendo con la riduzione di Gauss-Jordan ci si trova davanti ad una serie di **possibili alternative**:

- il sistema non ammette soluzione,
- il sistema ammette un'unica soluzione,
- il sistema ammette ∞^k soluzioni per un certo $k = n - r$, dove r è il numero di **pivot** della matrice a scala S. Il numero r si chiama *rango della matrice* A, e si indica con $\mathrm{rg}(A)$. Si noti che $r \leq \min\{m, n\}$.
- Il caso di non esistenza della soluzione si ha quando, dopo aver effettuato la riduzione di Gauss-Jordan, si trova un'equazione impossibile nel sistema $[S|c]$ (ad esempio, $0 = c_m$ con $c_m \neq 0$).

Si ha l'importante teorema di Rouché-Capelli.

Teorema. (Rouché-Capelli) *Data A, matrice $m \times n$, e dato il vettore $b \in \mathbb{R}^m$, il sistema $Ax = b$ è risolubile (cioè "ammette soluzione", o anche "è compatibile") per $x \in \mathbb{R}^n$ se e solo se*

$$\mathrm{rg}(A) = \mathrm{rg}([A|b]).$$

In tal caso, posto $r = \mathrm{rg}(A)$, il sistema ammette ∞^{n-r} soluzioni.

1.4 Proprietà delle soluzioni di sistemi di equazioni lineari.

Consideriamo il sistema di m equazioni in n incognite

$$Ax = b, \quad b \in \mathbb{R}^m.$$

Quando $b = 0$ il sistema si dice *omogeneo*. Consideriamo l'insieme delle soluzioni

$$S_b = \{x \in \mathbb{R}^n;\ Ax = b\}.$$

Valgono le seguenti proprietà:

- se x ed y sono due soluzioni del sistema omogeneo (cioè $x, y \in S_0$) allora per ogni $\alpha, \beta \in \mathbb{R}$ anche $\alpha x + \beta y$ è soluzione del sistema omogeneo (cioè $\alpha x + \beta y \in S_0$);
- se $\xi_0 \in S_b$ allora per ogni soluzione $x \in S_0$ si ha che $\xi_0 + x \in S_b$. In simboli, si ha che

  $$S_b = \{\xi_0 + x\,;\ x \in S_0\}$$

  (ξ_0 si dice essere una soluzione *particolare* del sistema non omogeneo). Si usa la notazione $S_b = \xi_0 + S_0$.

1.5 Sottospazi vettoriali e sottovarietà lineari affini

Un insieme $V \subset \mathbb{R}^n$ è un *sottospazio vettoriale* se

$$\alpha v + \beta v' \in V \qquad \forall \alpha,\, \beta \in \mathbb{R} \quad \forall v, v' \in V.$$

Quindi, ad esempio, è un **sottospazio vettoriale** *l'insieme S_0 delle soluzioni di un sistema lineare* **omogeneo**.

Dati k vettori, $v_1, \cdots, v_k$, lo *spazio generato* da tali vettori è l'insieme

$$V = \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k\,;\, \alpha_1, \cdots, \alpha_k \in \mathbb{R}\}\,.$$

Tale insieme si denota con il simbolo $\mathrm{Span}\,\{v_1, v_2, \cdots, v_k\}$ ed i vettori $v_1, \cdots, v_k$ sono detti i *generatori* del sottospazio V. Si ha che *V è un* **sottospazio vettoriale** *di $\mathbb{R}^n$.*

- Si dice che la famiglia **ordinata** di vettori $\{v_1, v_2, \cdots, v_k\}$ forma una *base* di un sottospazio vettoriale V se i $v_1, \ldots, v_k$ sono **linearmente indipendenti** e

  $$\mathrm{Span}\,\{v_1, v_2, \cdots, v_k\} = V\,.$$

  La *dimensione* del sottospazio vettoriale V è il numero di elementi che formano una base.

- Dati i sottospazi V e W di $\mathbb{R}^n$ allora:
  - l'insieme $V \cap W$ è un sottospazio vettoriale di $\mathbb{R}^n$;
  - l'insieme $V + W := \{v + w;\ v \in V,\, w \in W\}$ è un sottospazio vettoriale di $\mathbb{R}^n$;
  - si dice che V e W *si sommano direttamente* se ogni vettore $x \in V + W$ si scrive in maniera **univoca** come $x = v + w$ dove $v \in V$ e $w \in W$. In tal caso si scrive $V \oplus W$. Vale che V e W *si sommano direttamente se e solo se* $V \cap W = \{0\}$.

- Se A è una matrice $m \times n$ allora
  - l'*immagine* di A è l'insieme

    $$\mathrm{Im}\,A = \{Ax\,;\, x \in \mathbb{R}^n\};$$

  - il *nucleo* di A è l'insieme

    $$\mathrm{Ker}\,A = \{x \in \mathbb{R}^n;\ Ax = 0\}.$$

  - È importante osservare che

    $$\mathrm{Ker}\,A \neq \{0\} \iff \textit{i vettori colonna di } A \textit{ sono } \textbf{linearmente dipendenti};$$

  - $\mathrm{Im}\,A$ e $\mathrm{Ker}\,A$ sono **sottospazi vettoriali** di $\mathbb{R}^m$ ed $\mathbb{R}^n$, **rispettivamente**;
  - per l'insieme S_0 delle soluzioni del sistema omogeneo $Ax = 0$ vale $S_0 = \mathrm{Ker}\,A$.

- Osserviamo che data la matrice $m \times n$, $A = [\vec{a}_1|\vec{a}_2|\ldots|\vec{a}_n]$, dove i vettori $\vec{a}_j \in \mathbb{R}^m$ sono le colonne di A, e dati i vettori $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ e $b \in \mathbb{R}^m$, il sistema $Ax = b$ è equivalente a scrivere

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \ldots + x_n\vec{a}_n = b,$$

e quindi *il sistema è risolubile se e solo se* $b \in \operatorname{Im} A$. Se ne deduce anche che

$$\operatorname{Im} A = \operatorname{Span}\{\text{colonne di } A\}.$$

È anche importante osservare che dal Teorema di Rouché-Capelli segue che l'appartenenza del vettore b allo spazio delle colonne di A (che dà la risolubilità del sistema) si legge nel sistema equivalente $[S|c]$, dove S è una riduzione a scala di A, ed esattamente nelle condizioni $c_{r+1} = \ldots = c_m = 0$, dove $r = \operatorname{rg}(A)$.

- Si ha che $\operatorname{rg}(A) = \dim \operatorname{Im} A$.

- La matrice A si dice *singolare* se $\operatorname{Ker} A \neq \{0\}$, il che è equivalente a $\operatorname{rg}(A) < n$.

Data la matrice A, $m \times n$, si definisce *trasposta di A* la matrice A^T *le cui righe sono le colonne di A*. Essa è quindi una matrice $n \times m$. Per esempio

$$\text{se} \quad A = \begin{bmatrix} 1 & 2 \\ -1 & 7 \\ 3 & 0 \end{bmatrix} \quad \text{allora} \quad A^T = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 7 & 0 \end{bmatrix}.$$

La matrice trasposta di A è anche denotata con il simbolo ${}^tA$.

- Se A è una matrice $m \times n$ e B è una matrice $n \times k$, allora vale $(AB)^T = B^TA^T$, che quindi risulta essere una matrice $k \times m$.

- Vale la (importantissima) relazione: $\operatorname{rg}(A) = \operatorname{rg}(A^T)$. Si ha quindi che

$$\operatorname{rg}(A) = \dim \operatorname{Span}\{\text{colonne di } A\} = \dim \operatorname{Span}\{\text{righe di } A\}.$$

Un esempio molto importante di base di $\mathbb{R}^3$ è l'insieme formato dai vettori

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Infatti, e_1, e_2 ed e_3 sono linearmente indipendenti ed $\mathbb{R}^3 = \operatorname{Span}\{e_1, e_2, e_3\}$.

In generale dato un insieme di generatori di uno spazio vettoriale si può estrarre da essi una base procedendo come segue:

- si costruisce una matrice avente per **colonne** tali vettori (per fissare le idee consideriamo il caso di quattro vettori in $\mathbb{R}^3$);

- si utilizzano le operazioni elementari **sulle righe** della matrice fino ad ottenere una forma scala S,

- le colonne della matrice di partenza **corrispondenti** alle colonne di S che contengono i pivot di S formano una base.

- **Equivalentemente**, si costruisce una matrice avente per **righe** i vettori dati e si utilizzano le operazioni elementari (sempre sulle righe) per ridurre la matrice ad una forma a scala. I vettori riga che contengono i pivot della matrice a scala, riletti come vettori colonna, formano una base.

Possiamo inoltre fare le seguenti considerazioni:

- i vettori $v_1, \ldots, v_k$ di $\mathbb{R}^n$ sono linearmente indipendenti se e solo se, definendo $A = [v_1|\ldots|v_k]$ (matrice $n \times k$), il sistema $Ax = 0$ ha come **unica** soluzione la soluzione nulla 0 (e quindi se e solo se la matrice A è non-singolare, e cioè, ancora, se e solo se $\mathrm{rg}(A) = k$);

- in generale, per ottenere l'equazione cartesiana di $\mathrm{Span}\{v_1, \ldots, v_k\}$ basta usare il Teorema di Rouché-Capelli sul sistema $[A|x]$ e leggere le equazioni dalle condizioni di compatibilità del sistema, inoltre le colonne di A corrispondenti ai pivot di una forma a scala di A sono una **base** per $\mathrm{Span}\{v_1, \ldots, v_k\}$;

- avendo definito *sottovarietà lineare affine* l'insieme $\xi_0 + V = \{\xi_0 + v;\ v \in V\}$, dove V è un sottospazio vettoriale di $\mathbb{R}^n$, si ha quindi che

- per ottenere l'equazione cartesiana della sottovarietà lineare affine $\xi_0 + V$, dove $V = \mathrm{Span}\{v_1, \ldots, v_k\}$, basta operare come nel caso vettoriale sul sistema $[A|x - \xi_0]$.

- Si pone $\dim(\xi_0 + V) = \dim V$.

- Supponiamo che $V = \mathrm{Span}\{v_1, \ldots, v_k\}$. Consideriamo ancora la matrice $n \times k$, $A = [v_1|v_2|\ldots|v_k]$. Si può allora pensare a $\xi_0 + V$ come l'insieme S_b delle soluzioni del sistema $Ax = b$, dove $b = A\xi_0$, cioè $\xi_0 + V = S_b$.

È chiaro che in un sottospazio vettoriale possono essere scelte diverse basi. Quindi i vettori appartenenti a tale sottospazio possono essere rappresentati usando diversi sistemi di coordinate. Un'operazione fondamentale è il cambiamento di sistema di coordinate, ossia passare dalle coordinate di un vettore espresse rispetto ad una base a quelle scritte rispetto ad una base diversa (si veda il paragrafo **1.11**).

1.6 Determinante per matrici $n \times n$ e proprietà

Il determinante di un numero (matrice 1×1) è il numero stesso.
Data una matrice 2×2

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

il determinate di A è per definizione il numero

$$\det A = ad - bc\,.$$

Data una matrice 3×3

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

il determinante di A è definito nel modo seguente

$$\det A = a_{11}\det A_{11} - a_{12}\det A_{12} + a_{13}\det A_{13}$$

dove

$$A_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \quad A_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \quad A_{13} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

Definendo

$$\mathcal{A}_{ij} = (-1)^{i+j}\det A_{ij}, \quad 1 \leq i, j \leq n,$$

il **complemento algebrico dell'elemento** a_{ij}, si ha che la formula sopra si scrive anche

$$\det A = a_{11}\mathcal{A}_{11} + a_{12}\mathcal{A}_{12} + a_{13}\mathcal{A}_{13}.$$

In generale, nel caso $n \times n$, si ha che il determinante di una matrice A può essere calcolato nel modo seguente (sviluppo di Laplace). Dati $i, j \in \{1, 2, \ldots, n\}$ sia A_{ij} la matrice ottenuta dalla matrice A sopprimendo la i–esima riga e la j–esima colonna, e sia $\mathcal{A}_{ij} = (-1)^{i+j}\det A_{ij}$, $1 \leq i, j \leq n$, il complemento algebrico di a_{ij}. Fissiamo un numero $k \in \{1, 2, \ldots, n\}$. Allora vale il seguente sviluppo del determinante rispetto alla k–esima riga:

$$\det A = (-1)^{k+1}a_{k1}\det A_{k1} + (-1)^{k+2}a_{k2}\det A_{k2} + \ldots + (-1)^{k+n}a_{kn}\det A_{kn} =$$

$$= a_{k1}\mathcal{A}_{k1} + a_{k2}\mathcal{A}_{k2} + \ldots + a_{kn}\mathcal{A}_{kn}.$$

Vale una formula analoga rispetto alle colonne. Infatti, fissato $k \in \{1, 2, \ldots, n\}$, si ha il seguente sviluppo del determinante rispetto alla k-esima colonna:

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

$$= a_{1k}\mathcal{A}_{1k} + a_{2k}\mathcal{A}_{2k} + \ldots + a_{nk}\mathcal{A}_{nk}.$$

Il determinante soddisfa le seguenti proprietà:

- $\det I = 1$;
- $\det A = \det(A^T)$;
- data la matrice A sia B la matrice ottenuta scambiando nella matrice A due colonne (o due righe), allora $\det B = -\det A$; perciò, in particolare, se due colonne (o due righe) di A sono uguali, allora $\det A = 0$;
- siano $v_1, v_3, \ldots, v_n$ e w vettori colonna ed α e β due numeri reali allora

  $$\det\,[\alpha v_1 + \beta w|v_2|\ldots|v_n] = \alpha\det\,[v_1|v_2|\ldots|v_n] + \beta\det[w|v_2|\ldots|v_n],$$

  e vale anche sempre che

  $$\det\,[v_1|\ldots|\alpha v_i + \beta v_j|\ldots|v_j|\ldots|v_n] = \alpha\det\,[v_1|\ldots|v_i|\ldots|v_j|\ldots|v_n].$$

  La formula precedente si applica anche ad ogni altra coppia di colonne (e continua a valere ragionando sulle righe invece che sulle colonne);
- vale la formula di Binét: $\det(AB) = (\det A)(\det B) = \det(BA)$.

1.7 Matrici invertibili e calcolo della matrice inversa

Una matrice **quadrata** A si dice *invertibile* se esiste una matrice B tale che $AB = BA = I$. Si ha che *la matrice inversa se esiste è unica.* Si ha inoltre che *una matrice è invertibile se e solo se essa ammette una unica inversa destra*, e che *una matrice è invertibile se e solo se* $\mathrm{rg}(A) = n$. La matrice inversa (o brevemente *l'inversa*) della matrice A si denota con il simbolo A^{-1}. Vale il seguente teorema.

Teorema. *La matrice A è invertibile se e solo se $\det A \neq 0$. In particolare si ha che se A è invertibile, allora* $\det(A^{-1}) = \dfrac{1}{\det A}$.

Per calcolare l'inversa di una matrice A si procede come segue:

- si costruisce la matrice $[A|I]$
- si procede con operazioni elementari sulle righe della matrice $[A|I]$ fino a che non ci si riduce ad una matrice della forma $[I|C]$ per una qualche matrice C (quello che conta è avere nel primo "blocco" la matrice identità!)
- si conclude che $A^{-1} = C$.

L'uso del determinante permette anche di dare le seguenti formule per l'inversa di A.

- Nel caso 2×2, data la matrice $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ con $\det A \neq 0$, allora si ha la formula

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

- Nel caso $n \times n$, se $\mathcal{A} = (\mathcal{A}_{ij})_{1 \leq i,j \leq n}$ è la matrice la cui entrata al posto ij è $\mathcal{A}_{ij}$, $1 \leq i,j \leq n$, si definisce la **matrice aggiunta** di A tramite la formula

$$\mathrm{Agg}(A) = \mathcal{A}^T,$$

e si ha che

$$A^{-1} = \frac{1}{\det A}\mathrm{Agg}(A).$$

- In particolare, considerando il sistema $Ax = b$, A matrice $n \times n$ *invertibile*, allora la soluzione è univocamente data da $\xi = A^{-1}b$. Di più, si ha la *formula di Cramer* per la soluzione $\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$,

$$\xi_i = \frac{\det B_i}{\det A}, \quad 1 \leq i \leq n,$$

dove B_i è la matrice che si ottiene *sostituendo alla colonna i-esima di A il vettore termine noto b*, $i = 1, 2, \ldots, n$.

1.8 Prodotto scalare di vettori in $\mathbb{R}^n$ e proiezione di un vettore lungo un altro

Ricordiamo che, dati i vettori $x, y \in \mathbb{R}^n$, si è definito *prodotto scalare* il numero reale

$$\langle x, y \rangle = \left\langle \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = \sum_{j=1}^{n} x_j y_j.$$

È importante notare che $\langle x, y \rangle = x^T y$ (prodotto del vettore riga x^T col vettore colonna y). Quindi è facile vedere che si ha la seguente formula: *per ogni matrice $n \times n$, A, si ha*

$$\langle Ax, y \rangle = \langle x, A^T y \rangle.$$

Proprietà del prodotto scalare:

- simmetria: $\langle v, w \rangle = \langle w, v \rangle$, per ogni $v, w \in \mathbb{R}^n$;
- bilinearità: $\begin{cases} \langle \alpha v + \beta v', w \rangle = \alpha \langle v, w \rangle + \beta \langle v', w \rangle \\ \\ \langle v, \alpha w + \beta w' \rangle = \alpha \langle v, w \rangle + \beta \langle v, w' \rangle \end{cases}$, per tutti gli $\alpha, \beta \in \mathbb{R}$ e tutti i $v, v', w, w' \in \mathbb{R}^n$;
- $\langle v, v \rangle \geq 0$, per tutti i $v \in \mathbb{R}^n$, e $\langle v, v \rangle = 0$ se e solo se $v = 0$.

La **lunghezza** (o **norma**) di un vettore $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ è stata definita da

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{j=1}^{n} x_j^2}.$$

Si osservi che per ogni $x \in \mathbb{R}^n$ ed ogni $\alpha \in \mathbb{R}$

$$\|\alpha x\| = |\alpha| \, \|x\|, \quad \text{e} \quad \|x\| = 0 \iff x = 0.$$

Dati poi due punti P_0, P_1 di $\mathbb{R}^n$, si definisce la loro *distanza* (euclidea) tramite

$$\mathsf{dist}(P_0, P_1) := \|P_0 - P_1\|.$$

Si noti che $\mathsf{dist}(P_0, P_1) = \mathsf{dist}(P_1, P_0)$.

Vale la *disuguaglianza di Cauchy-Schwarz: per ogni coppia di vettori $x, y \in \mathbb{R}^n$ si ha*

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

e l'uguaglianza sussiste se e solo se x è multiplo di y.

Proprietà della norma:

- $\|x\| \geq 0$ per tutti gli $x \in \mathbb{R}^n$, e $\|x\| = 0$ se e solo se $x = 0$;

- $\|\alpha x\| = |\alpha|\|x\|$ per tutti gli $\alpha \in \mathbb{R}$ e tutti gli $x \in \mathbb{R}^n$;
- dalla disuguaglianza di Cauchy-Schwarz segue la **disuguaglianza triangolare**
$$\|x+y\| \leq \|x\| + \|y\|, \quad \forall x, y \in \mathbb{R}^n.$$

Ora, poiché dalla disuguaglianza di Cauchy-Schwarz risulta quindi essere
$$-1 \leq \frac{\langle x, y\rangle}{\|x\|\|y\|} \leq 1, \ \forall x, y \neq 0,$$
si ha che **esiste un unico angolo** $\alpha \in [0, \pi]$ tale che
$$\frac{\langle x, y\rangle}{\|x\|\|y\|} = \cos\alpha.$$
Tale angolo si chiama *l'angolo (convesso) formato dai vettori x e y.*

- Si dice che due vettori v e w di $\mathbb{R}^n$ sono *ortogonali* se $\langle v, w\rangle = 0$.
- Dato $v \in \mathbb{R}^n$, si ha che $\langle v, w\rangle = 0$ per ogni $w \in \mathbb{R}^n$ se e solo se $v = 0$.
- Dato un sottospazio V di $\mathbb{R}^n$, si definisce il *sottospazio ortogonale di V* l'insieme
$$V^\perp := \{w \in \mathbb{R}^n;\ \langle v, w\rangle = 0,\ \forall v \in V\}.$$
Si ha che $V^\perp$ è un sottospazio vettoriale di $\mathbb{R}^n$, che $\mathbb{R}^n = V \oplus V^\perp$, e che $\dim V^\perp = n - \dim V$.
- Due sottospazi U, V si dicono essere *ortogonali* quando $\langle u, v\rangle = 0$ per tutti gli $u \in U$ e $v \in V$. Una condizione necessaria e sufficiente affinché U, con base $\{u_1, \ldots, u_\ell\}$, e V, con base $\{v_1, \ldots, v_k\}$, siano ortogonali è che valga $\langle u_i, v_j\rangle = 0$ per ogni $i = 1, \ldots, \ell$ e $j = 1 \ldots, k$. In particolare deve essere $U \cap V = \{0\}$.
- Quando $A = A^T$, allora $(\operatorname{Ker} A)^\perp = \operatorname{Im} A$. Infatti, se $v \in \operatorname{Im} A$, allora $v = Av'$ per un certo $v' \in \mathbb{R}^n$, e quindi per ogni $w \in \operatorname{Ker} A$ si ha
$$\langle v, w\rangle = \langle Av', w\rangle = \langle v', A^T w\rangle = \langle v', Aw\rangle = 0,$$
poiché $A = A^T$ e $Aw = 0$. Ma allora $\langle v, w\rangle = 0$ per ogni $w \in \operatorname{Ker} A$, da cui $v \in (\operatorname{Ker} A)^\perp$. Quindi $\operatorname{Im} A \subset (\operatorname{Ker} A)^\perp$ ed essendo
$$\dim \operatorname{Im} A = \operatorname{rg}(A) = n - \dim \operatorname{Ker} A = \dim(\operatorname{Ker} A)^\perp,$$
ne segue che $\operatorname{Im} A = (\operatorname{Ker} A)^\perp$.
- Dal punto precedente segue che $\mathbb{R}^n = (\operatorname{Ker} A) \oplus (\operatorname{Im} A)$, con somma diretta **ortogonale**.

1.9 Basi ortonormali, processo di ortonormalizzazione, matrici ortogonali

Una base $\{u_1, \ldots, u_n\}$ di $\mathbb{R}^n$ si dice **ortonormale** se valgono le relazioni

$$\|u_1\| = \|u_2\| = \ldots = \|u_n\| = 1, \quad \langle u_j, u_r\rangle = 0 \text{ se } j \neq r.$$

Le basi ortonormali di $\mathbb{R}^n$ sono di fondamentale importanza, in quanto esse permettono di scrivere facilmente le coordinate di un vettore qualsiasi v rispetto ad esse. Infatti: *se $\{u_1, \ldots, u_n\}$ è una base ortonormale di $\mathbb{R}^n$ allora ogni vettore $v \in \mathbb{R}^n$ può essere scritto* **in maniera univoca** *nella forma*

$$v = \langle v, u_1\rangle u_1 + \langle v, u_2\rangle u_2 + \ldots + \langle v, u_n\rangle u_n.$$

Quindi **le coordinate di** v **rispetto alla base** $\{u_1, \ldots, u_n\}$ sono i numeri $\langle v, u_1\rangle, \ldots, \langle v, u_n\rangle$.

- **Osservazione.** I vettori di una famiglia $\{u_1, \ldots, u_k\} \subset \mathbb{R}^n$ ortonormale sono **linearmente indipendenti**.

- Dati i vettori non nulli $v, w \in \mathbb{R}^n$, si definisce *proiezione di v lungo w* il vettore

$$\mathsf{pr}_w(v) := \frac{\langle v, w\rangle}{\|w\|^2} w.$$

  Si noti il seguente fatto importantissimo (perché fondamentale per il processo di ortonormalizzazione): *il vettore $v - \mathsf{pr}_w(v)$ è ortogonale a w*, cioè

$$\langle v - \mathsf{pr}_w(v), w\rangle = 0.$$

Vale il seguente importante teorema.

Teorema. (Processo di ortonormalizzazione di Gram-Schmidt) *Data la base $\{v_1, \ldots, v_n\}$ di $\mathbb{R}^n$ è sempre possibile costruire a partire da essa una base ortonormale $\{u_1, \ldots, u_n\}$.*

La dimostrazione del teorema è costruttiva. Facciamo vedere la costruzione nei casi $n = 2, 3$.

Caso $n = 2$. Sia $\{v_1, v_2\}$ una base di $\mathbb{R}^2$. Si pone

$$u_1 := \frac{v_1}{\|v_1\|}.$$

Quindi $\|u_1\| = 1$. Sia poi $\tilde{u}_2 := v_2 - \langle v_2, u_1\rangle u_1$. Allora $\langle \tilde{u}_2, u_1\rangle = 0$. Per finire basta prendere

$$u_2 = \frac{\tilde{u}_2}{\|\tilde{u}_2\|} = \frac{v_2 - \langle v_2, u_1\rangle u_1}{\|v_2 - \langle v_2, u_1\rangle u_1\|}.$$

Caso $n = 3$. Sia $\{v_1, v_2, v_3\}$ una base di $\mathbb{R}^3$. Si pone

$$u_1 := \frac{v_1}{\|v_1\|}.$$

Quindi $\|u_1\| = 1$. Sia poi $\tilde{u}_2 := v_2 - \langle v_2, u_2\rangle u_2$. Allora $\langle \tilde{u}_2, u_1\rangle = 0$, e quindi prendiamo

$$u_2 = \frac{\tilde{u}_2}{\|\tilde{u}_2\|} = \frac{v_2 - \langle v_2, u_1\rangle u_1}{\|v_2 - \langle v_2, u_1\rangle u_1\|}.$$

Ora prendiamo $\tilde{u}_3 := v_3 - \langle v_3, u_2\rangle u_2 - \langle v_3, u_1\rangle u_1$, e, per finire, essendo $\langle \tilde{u}_3, u_1\rangle = \langle \tilde{u}_3, u_2\rangle = 0$,

$$u_3 = \frac{v_3 - \langle v_3, u_2\rangle u_2 - \langle v_3, u_1\rangle u_1}{\|v_3 - \langle v_3, u_2\rangle u_2 - \langle v_3, u_1\rangle u_1\|}.$$

Il caso generale n-dimensionale procede per induzione esattamente secondo le linee sopra esposte.

Osservazione. Una formulazione apparentemente più generale del processo di ortonormalizzazione di Gram-Schmidt è la seguente:

Teorema. *A partire da una base $\{v_1, \ldots, v_k\}$ del sottospazio vettoriale $V \subset \mathbb{R}^n$ si può sempre costruire un'altra base di V che sia ortonormale.*

Data una base ortonormale $\{u_1, \ldots, u_n\}$ di $\mathbb{R}^n$, costruiamo la matrice R la cui j-esima colonna è il vettore u_j, $j = 1, \ldots, n$:

$$R = [u_1|u_2|\ldots|u_n].$$

Allora la matrice trasposta di R si scrive

$$R^T = \begin{bmatrix} u_1^T \\ u_2^T \\ \ldots \\ u_n^T \end{bmatrix},$$

le cui righe sono i vettori **riga** u_j^T. Dalla regola di moltiplicazione righe-per-colonne tra matrici si ha perciò

$$R^T R = \begin{bmatrix} u_1^T u_1 & u_1^T u_2 & \ldots & u_1^T u_n \\ u_2^T u_1 & u_2^T u_2 & \ldots & u_2^T u_n \\ \vdots & \vdots & \ddots & \vdots \\ u_n^T u_1 & u_n^T u_2 & \ldots & u_n^T u_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} = I.$$

Poiché R^T è invertibile, essa possiede un'unica inversa destra che quindi è R, da cui segue $R^T = R^{-1}$. Tali matrici, cioè quelle matrici *le cui colonne (equivalentemente le cui righe) formano una base ortonormale di* $\mathbb{R}^n$ si dicono **ortogonali**, e sono di fondamentale importanza perché conservano il prodotto scalare tra vettori, e quindi le loro lunghezze, e quindi gli angoli tra i vettori. Infatti se R è una matrice ortogonale $n \times n$ allora

$$\langle Rv, Rw\rangle = \langle v, R^T Rw\rangle = \langle v, w\rangle,\ \forall v, w \in \mathbb{R}^n.$$

Quindi, prendendo $v = w$ si ottiene $\|Rv\| = \|v\|$, cioè Rv ha la stessa lunghezza di v. Allora, indicato con α l'angolo tra Rv e Rw e β l'angolo tra v e w, si ha che

$$\cos\alpha = \frac{\langle Rv, Rw\rangle}{\|Rv\|\|Rw\|} = \frac{\langle v, w\rangle}{\|v\|\|w\|} = \cos\beta,$$

da cui segue che $\alpha = \beta$, essendo per ipotesi $\alpha, \beta \in [0, \pi]$.

- Dato $u \in \mathbb{R}^n$, la funzione $f_u\colon \mathbb{R}^n \ni v \longmapsto \langle v, u\rangle \in \mathbb{R}$ determina univocamente u e, reciprocamente, è determinata univocamente da u. Ciò si vede facilmente prendendo una base ortonormale $\{u_1, \ldots, u_n\}$ di $\mathbb{R}^n$ e notando che

$$u = \langle u, u_1\rangle u_1 + \langle u, u_2\rangle u_2 + \ldots + \langle u, u_n\rangle u_n = f_u(u_1)u_1 + f_u(u_2)u_2 + \ldots + f_u(u_n)u_n.$$

1.10 Prodotto vettoriale di vettori in $\mathbb{R}^3$ e sua interpretazione geometrica

Dati due vettori $v = (v_1, v_2, v_3)^T$ e $w = (w_1, w_2, w_3)^T$ dello spazio $\mathbb{R}^3$, il loro *prodotto vettoriale* è il vettore $v \times w$ determinato dalla funzione, detta *prodotto misto*,

$$f_{v\times w}(u) = \langle u, v \times w\rangle \overset{\text{def}}{=} \det \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix}, \text{ ove } u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$

In altre parole, il vettore $v \times w$ è determinato dal conoscere tutti i possibili prodotti scalari $\langle u, v \times w\rangle$, u vettore qualsiasi di $\mathbb{R}^3$, i quali, a loro volta, *per definizione* valgono $\det[u|v|w]$. Geometricamente, il prodotto misto $\langle u, v \times w\rangle$ rappresenta il volume (orientato, cioè con segno!) del parallelepipedo avente per spigoli u, v e w.

- È quindi chiaro che $\langle u, v \times w\rangle = 0$ *se e solo se* u, v *e* w *sono linearmente dipendenti.*
- In particolare, $v \times w$ *è ortogonale sia a* v *che a* w.

Poiché la funzione $f_{v\times w}$ determina univocamente il vettore $v \times w$, ne segue che $v \times w$ si può anche scrivere come il seguente determinante "formale" (in quanto $\mathbf{e}_1, \mathbf{e}_2$ ed $\mathbf{e}_3$, vettori ortonormali della base canonica di $\mathbb{R}^3$, non sono numeri ma vettori!)

$$v \times w = \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix},$$

in altre parole

$$v \times w = \begin{bmatrix} v_2w_3 - v_3w_2 \\ -(v_1w_3 - v_3w_1) \\ v_1w_2 - v_2w_1 \end{bmatrix}.$$

Geometricamente, il prodotto vettoriale dei vettori v e w rappresenta l'area "orientata" (cioè con segno!) del parallelogramma avente per lati v e w.

- È quindi chiaro che $v \times w = 0$ *se e solo se* v *e* w *sono uno multiplo dell'altro.*

Valgono le seguenti proprietà: per ogni $u, v, w \in \mathbb{R}^3$,

- $v \times v = 0$;
- $v \times w = -(w \times v)$;
- per ogni $\lambda \in \mathbb{R}$, $v \times (\lambda w) = \lambda(v \times w) = (\lambda v) \times w$;
- $(v + w) \times u = v \times u + w \times u$;
- $\|v \times w\| = \sqrt{\|v\|^2\|w\|^2 - \langle v, w\rangle^2} = \|v\|\,\|w\| \sin\alpha$, dove $\alpha \in [0, \pi]$ è l'angolo tra v e w.

1.11 Applicazioni lineari e loro rappresentazione matriciale, matrici del cambiamento di base

Siano V e W due sottospazi vettoriali di $\mathbb{R}^k$ ed $\mathbb{R}^s$, rispettivamente, di dimensioni rispettive n ed m .

Una *applicazione lineare* è una funzione

$$L : V \longrightarrow W$$

soddisfacente la seguente proprietà: per ogni $v, w \in V$ e per ogni $\alpha, \beta \in \mathbb{R}$,

$$L(\alpha v + \beta w) = \alpha L(v) + \beta L(w)\,.$$

Si ha che, fissata una base $\{f_1, f_2, \cdots, f_n\}$ di V ed una $\{g_1, g_2, \cdots, g_m\}$ di W, alla trasformazione lineare L è associata una matrice A_L la cui prima colonna è formata dalle componenti del vettore $L(f_1)$ rispetto alla base $g_1, g_2, \cdots, g_m$, la seconda colonna sono le componenti di $L(f_2)$ rispetto alla stessa base e così via. Si ha quindi che A_L è una matrice $m \times n$.

Una tecnica molto importante è quella di cambiare sistema di riferimento per descrivere più semplicemente, e quindi risolvere in modo più efficiente, un dato problema. Se si impone che il cambiamento di sistema di riferimento mantenga fissa l'origine da noi scelta inizialmente, si arriva al concetto di *cambiamento di base.* Una base $\mathcal{V} = \{v_1, \ldots, v_n\}$ di $\mathbb{R}^n$ è infatti la scelta di un sistema di riferimento rispetto al quale i vettori $v \in \mathbb{R}^n$ vengono rappresentati come combinazione lineare dei v_j, cioè

$$v = \sum_{j=1}^{n} \alpha_j v_j, \quad \alpha_1, \ldots, \alpha_n \in \mathbb{R}.$$

Si ha così che **le coordinate di v rispetto alla base data sono i numeri** $\alpha_1, \ldots, \alpha_n$. Il problema è allora calcolare le coordinate di v rispetto ad un'altra base $\mathcal{W} = \{w_1, \ldots, w_n\}$ di $\mathbb{R}^n$. La soluzione è fornita dalla *matrice $M_{\mathcal{V},\mathcal{W}}$ associata alla* **mappa identità***, cioè la matrice associata all'applicazione lineare identità* $\mathrm{Id}\colon \mathbb{R}^n \ni v \longmapsto v \in \mathbb{R}^n$*, rispetto alle basi $\mathcal{V}$ (di partenza) e $\mathcal{W}$ (di arrivo).* Infatti si parte con l'osservare che ogni elemento della base $\mathcal{V}$ si scrive in termini della base $\mathcal{W}$ nella maniera seguente:

$$v_j = \sum_{k=1}^{n} a_{kj} w_k, \quad j = 1, \ldots, n.$$

È importante notare l'ordine degli indici nelle coordinate di v_j rispetto alla base $\mathcal{W}$: il vettore $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$ è la j-esima colonna della matrice $M_{\mathcal{V},\mathcal{W}}$. Ora, avendo $M_{\mathcal{V},\mathcal{W}}$ e dato $v \in \mathbb{R}^n$ qualsiasi, si ha

$$v = \sum_{j=1}^{n} \alpha_j v_j, \quad \text{ma anche} \quad v = \sum_{k=1}^{n} \beta_k w_k,$$

per cui

$$v = \sum_{j=1}^{n} \alpha_j \left(\sum_{k=1}^{n} a_{kj} w_k \right) = \sum_{j,k=1}^{n} a_{kj} \alpha_j w_k = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{kj} \alpha_j \right) w_k = \sum_{k=1}^{n} \beta_k w_k,$$

da cui segue (per l'unicità della scrittura di v come combinazione lineare degli elementi della base $\mathcal{W}$) che

$$\sum_{j=1}^{n} a_{kj}\alpha_j = \beta_k, \quad k = 1, \dots, n,$$

e cioè, in notazione matriciale (righe-per-colonne), i $(\beta_1, \dots, \beta_n)$ si ottengono dalla relazione

$$M_{\mathcal{V},\mathcal{W}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

A questo punto è anche chiaro che *la matrice $M_{\mathcal{W},\mathcal{V}}$ del cambiamento inverso, cioè del passaggio dalla base $\mathcal{W}$ alla base $\mathcal{V}$ è data dalla matrice $M^{-1}_{\mathcal{V},\mathcal{W}}$. Essa risolve il problema di passare dalle coordinate di v rispetto alla base $\mathcal{W}$ alle coordinate di v rispetto alla base $\mathcal{V}$.*

Esempio. Per fissare le idee, diamo un esempio del calcolo di $M_{\mathcal{V},\mathcal{W}}$ in dimensione 2. Siano allora

$$\mathcal{V} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix} \right\}, \quad \text{e} \quad \mathcal{W} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

Calcoliamo allora le coordinate di $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ rispetto a $\mathcal{W}$. Per fare ciò, risolviamo il sistema lineare

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} = a_{11} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + a_{21} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

La matrice associata al sistema è dunque

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ -1 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 3/2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -1/2 \\ 0 & 1 & 3/2 \end{array}\right],$$

cioè $a_{11} = -1/2$ e $a_{21} = 3/2$. Alla stessa maniera, da

$$\begin{bmatrix} 2 \\ -3 \end{bmatrix} = a_{12} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + a_{22} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

segue

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ -1 & 1 & -3 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & -1/2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 5/2 \\ 0 & 1 & -1/2 \end{array}\right],$$

cioè $a_{12} = 5/2$ e $a_{22} = -1/2$. Quindi la matrice cercata è

$$M_{\mathcal{V},\mathcal{W}} = \begin{bmatrix} -1/2 & 5/2 \\ 3/2 & -1/2 \end{bmatrix}.$$

È **importante** ora notare le seguenti cose:

- i sistemi che forniscono le colonne di $M_{\mathcal{V},\mathcal{W}}$ potevano essere *contemporaneamente* risolti operando sulle righe della matrice

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 2 \\ -1 & 1 & 2 & -3 \end{array}\right]$$

  in modo da ottenere la matrice $[I|C]$. Si ha che C è la matrice cercata $M_{\mathcal{V},\mathcal{W}}$;

- il metodo fornisce anche la seguente procedura (completamente generale) per calcolare la matrice $M_{\mathcal{V},\mathcal{W}}$ del cambiamento di base dalla $\mathcal{V} = \{v_1, \dots, v_n\}$ alla $\mathcal{W} = \{w_1, \dots, w_n\}$:
$$M_{\mathcal{V},\mathcal{W}} = [w_1|\dots|w_n]^{-1}[v_1|\dots|v_n].$$

1.12 Forme quadratiche e diagonalizzazione

Sia A una matrice quadrata $n \times n$. Il *polinomio caratteristico* di A è

$$p(\lambda) = \det(A - \lambda I) \qquad (\lambda \in \mathbb{R}).$$

Il polinomio caratteristico **ha grado** n, e le sue radici si chiamano *autovalori* (della matrice A). Se λ_0 è un autovalore l'insieme Ker $(A - \lambda_0 I)$ è *l'autospazio* associato all'autovalore λ_0. Gli elementi di Ker $(A - \lambda_0 I)$ sono gli *autovettori* di A (associati all'autovalore λ_0).

- Si dice che la *molteplicità algebrica* dell'autovalore λ_0 è m se
$$(\lambda - \lambda_0)^m \text{ divide } p(\lambda), \quad \text{ma } (\lambda - \lambda_0)^{m+1} \text{ non divide } p(\lambda).$$

- Si dice che la *molteplicità geometrica* dell'autovalore λ_0 è r se
$$\dim \operatorname{Ker}(A - \lambda_0 I) = r.$$

- Si ha sempre che
$$\text{molteplicità geometrica}(\lambda_0) \leq \text{molteplicità algebrica}(\lambda_0).$$

Si definisce *traccia* della matrice A, quadrata $n \times n$, la somma degli elementi sulla diagonale principale di A, cioè il numero
$$\operatorname{Tr}(A) = \sum_{j=1}^{n} a_{jj}.$$

Data la matrice A quadrata $n \times n$, valgono le seguenti relazioni tra gli autovalori $\lambda_1, \dots, \lambda_n$ di A, $\operatorname{Tr}(A)$ e $\det(A)$:

- $\operatorname{Tr}(A) = \sum_{j=1}^{n} \lambda_j$;
- $\det(A) = \lambda_1 \lambda_2 \dots \lambda_n$;
- quando A è 2×2 si ha
$$\det(A - \lambda I) = \lambda^2 - \operatorname{Tr}(A)\lambda + \det(A).$$

Una matrice quadrata A si dice *simmetrica* se coincide con la sua trasposta, cioè se $A = A^T$. Vale il seguente risultato.

Teorema. *Sia A una matrice simmetrica $n \times n$. Allora, gli autovalori di A, $\lambda_1, \cdots, \lambda_n$, ripetuti secondo la rispettiva molteplicità algebrica, sono reali ed autovettori associati ad* *autovalori distinti sono tra loro ortogonali. Di più, si può trovare una* **base** *di* $\mathbb{R}^n$ *fatta di autovettori di A. Tale base, eventualmente tramite un processo di ortonormalizzazione di Gram-Schmidt, può essere scelta* **ortonormale**.

È importante notare il fatto seguente.
Sia A una matrice quadrata $n \times n$ **qualsiasi** (cioè non necessariamente simmetrica) la quale *abbia tutti gli autovalori* $\lambda_1, \dots, \lambda_n$ *in* $\mathbb{R}$ *ed i cui relativi autovettori formino una* **base** *di* $\mathbb{R}^n$. Sia allora Λ la matrice diagonale i cui elementi sulla diagonale principale sono gli autovalori $\lambda_1, \dots, \lambda_n$ di A, e siano $v_1, \dots, v_n$ i corrispondenti autovettori (relativi a $\lambda_1, \dots, \lambda_n$, rispettivamente). Definiamo $S = [v_1|v_2|\dots|v_n]$. Allora S è invertibile perché $v_1, \dots, v_n$ formano una base di $\mathbb{R}^n$. Dall'equazione agli autovalori $Av_j = \lambda_j v_j$, $j = 1, \dots, n$, segue che (righe-per-colonne)

$$AS = [Av_1|Av_2|\dots|Av_n] = [\lambda_1 v_1|\lambda_2 v_2|\dots|\lambda_n v_n] = S\Lambda,$$

e quindi si ha la relazione

$$A = S\Lambda S^{-1}.$$

Una tale matrice si dice allora **diagonalizzabile**. Quindi, il teorema precedente assicura che *tutte le matrici simmetriche sono diagonalizzabili.*
Osserviamo anche che quando A *è simmetrica e gli autovettori* $v_1, \dots, v_n$ *si scelgono in modo da formare una base* **ortonormale** *di* $\mathbb{R}^n$, allora la matrice S associata agli autovettori di A è **ortogonale**, per cui vale $S^{-1} = S^T$ e quindi anche la relazione

$$A = S\Lambda S^T.$$

Data A, matrice simmetrica $n \times n$, una *forma quadratica* Q è una funzione che associa ad ogni vettore x con n componenti il numero $Q(x) = \langle Ax, x\rangle$. Sappiamo quindi che esiste una matrice ortogonale S tale che $A = S\Lambda S^T$, con

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix},$$

e con $\lambda_1, \lambda_2, \cdots, \lambda_n$ gli autovalori (eventualmente ripetuti) della matrice A. Allora, se si pone $y = S^T x$, nelle "nuove" coordinate y si ha

$$Q(x) = Q(Sy) = \langle \Lambda y, y\rangle = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

2 Geometria del piano e dello spazio

2.1 Piani nello spazio tridimensionale

Sia P_0 un punto di $\mathbb{R}^3$ ed n un vettore di $\mathbb{R}^3$ diverso da zero. Allora il piano passante per il punto P_0 e perpendicolare al vettore n è l'insieme

$$\pi = \{P \in \mathbb{R}^3 \, : \, \langle P - P_0, n\rangle = 0\}.$$

- Il vettore n si chiama *(una) direzione normale al piano* π.

Osserviamo che posto $P = (x, y, z)$, $P_0 = (x_0, y_0, z_0)$ e $n = (n_1, n_2, n_3)^T$ la formula precedente si scrive come
$$n_1 x + n_2 y + n_3 z - (n_1 x_0 + n_2 y_0 + n_3 z_0) = 0\,.$$
In generale quindi un piano scritto secondo la *rappresentazione cartesiana* è dato da
$$ax + by + cz + d = 0$$
dove a, b, c, d sono numeri assegnati. È allora chiaro che data la precedente rappresentazione cartesiana il vettore $(a, b, c)^T$ è un vettore normale al piano. Siano dati tre punti $P_0 = (x_0, y_0, z_0), P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_2, y_2, z_2)$ tali che i vettori $P_1 - P_0$ e $P_2 - P_0$ siano linearmente indipendenti (tali punti si dicono *non allineati*), cioè
$$(P_1 - P_0) \times (P_2 - P_0) \neq 0\,.$$
Allora **il** piano π passante per P_0, P_1 e P_2 è dato dall'equazione
$$\det \begin{bmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ x_2 - x_0 & y_2 - y_0 & z_2 - z_0 \end{bmatrix} = 0\,,$$
cioè, posto $P = (x, y, z)$, la formula precedente può essere riscritta come segue
$$\langle (P - P_0), (P_1 - P_0) \times (P_2 - P_0) \rangle = 0\,.$$
Il piano π contenente i punti P_0, P_1 e P_2, con $P_1 - P_0$ e $P_2 - P_0$ *linearmente indipendenti*, ha anche la seguente *rappresentazione parametrica*
$$\pi = \{P_0 + \alpha(P_1 - P_0) + \beta(P_2 - P_0)\,;\, \alpha, \beta \in \mathbb{R}\}\,.$$

- I vettori $P_1 - P_0$ e $P_2 - P_0$ si chiamano *direzioni del piano* π.

Dalla rappresentazione parametrica è immediato passare a quella cartesiana infatti basta scegliere, ad esempio,
$$P_0 = P_0 \qquad \text{e} \qquad n = (P_1 - P_0) \times (P_2 - P_0)$$
Due piani π_1 e π_2 si dicono

- *perpendicolari* se sono ortogonali le loro direzioni normali;
- *paralleli* se le loro direzioni normali sono una multiplo dell'altra.

Osservazione. Notiamo quindi che l'equazione cartesiana di una retta di $\mathbb{R}^3$ rappresenta l'intersezione di due piani distinti e non paralleli.

2.2 Rette nello spazio tridimensionale

Sia $v = (v_1, v_2, v_3)^T$ un vettore (non nullo!) di $\mathbb{R}^3$ e $P_0 = (x_0, y_0, z_0)$ un punto (fissato). La retta r passante per P_0 con *direzione* v è l'insieme
$$r = \{P_0 + tv\,;\; t \in \mathbb{R}\} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 + tv_1 \\ y_0 + tv_2 \\ z_0 + tv_3 \end{bmatrix}\,;\; t \in \mathbb{R} \right\}.$$

La precedente è detta rappresentazione della retta r in *forma parametrica.* Si può passare dalla rappresentazione parametrica alla *rappresentazione cartesiana* ponendo

$$\begin{cases} x = x_0 + tv_1 \\ y = y_0 + tv_2 \\ z = z_0 + tv_3 \end{cases}$$

"eliminando" il parametro t dalle precedenti espressioni ci si riduce ad espressioni della forma

$$\begin{cases} ax + by + cz + d = 0 \\ a_1x + b_1y + c_1z + d_1 = 0 \end{cases}$$

dove a, b, c, a_1, b_1, c_1 sono numeri opportuni che dipendono da P_0 e v. Ci si può porre il problema inverso ossia come passare dalla rappresentazione cartesiana a quella parametrica.

In questo caso sono assegnate le espressioni (in modo che non siano una un multiplo dell'altra!)

$$\begin{cases} ax + by + cz + d = 0 \\ a_1x + b_1y + c_1z + d_1 = 0\,. \end{cases}$$

Basta scegliere un punto P_0 tra le soluzioni del precedente sistema (essendo formato da due equazioni in due incognite ammette sicuramente, almeno, ∞^1 soluzioni) e prendere come direzione v il prodotto vettoriale $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \times \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$.

Siano r ed s due rette con direzione v_r e v_s rispettivamente. Si dice che

- r è *parallela* ad s se v_r è un multiplo di v_s;
- r è *perpendicolare* ad s se r ed s si **intersecano** e v_r è perpendicolare a v_s, cioè $r \cap s \neq \emptyset$ e $\langle v_r, v_s \rangle = 0$;
- r ed s sono *sghembe* se v_r **non** è un multiplo di v_s ma $r \cap s = \emptyset$.

2.3 Distanza punto–retta e punto–piano

Sia $P_0 = (x_0, y_0, z_0)$ un punto di $\mathbb{R}^3$ e sia r una retta assegnata, ad esempio, in forma cartesiana:

$$\begin{cases} ax + by + cz + d = 0 \\ a_1x + b_1y + c_1z + d_1 = 0\,. \end{cases}$$

Sia v il vettore

$$v = \frac{(a, b, c)^T \times (a_1, b_1, c_1)^T}{\|(a, b, c)^T \times (a_1, b_1, c_1)^T\|}\,.$$

Allora la distanza del punto P_0 dalla retta r è data dalla formula

$$d(P_0; r) = \|(P_1 - P_0) \times v\|\,,$$

dove P_1 è un punto della retta r fissato in modo arbitrario (la distanza del punto P_0 dalla retta r è indipendente dalla scelta del punto P_1).

Sia $P_0 = (x_0, y_0, z_0)$ un punto di $\mathbb{R}^3$ e sia π una piano assegnato, ad esempio, in forma cartesiana:

$$ax + by + cz + d = 0\,.$$

Allora la distanza del punto P_0 dal piano π è data dalla formula

$$d(P_0;\pi) = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}$$

Osserviamo che se $P_0 = (x_0, y_0)$ è un punto di $\mathbb{R}^2$ ed r è la retta di equazione cartesiana

$$ax + by + c = 0\,,$$

allora la distanza del punto P_0 dalla retta r è data dalla formula

$$d(P_0;r) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}\,.$$

2.4 Coniche, quadriche e forme normali

Per la classificazione delle quadriche si faccia riferimento al Capitolo 6 di [2]. Quella delle coniche segue da quella delle quadriche considerando i casi in cui $\mathrm{rg}(A)$ è uguale a 2 oppure 1. Basta usare le relative tabelle per la classificazione delle quadriche sostituendo

- *parabola* al posto di cilindro parabolico;
- *ellisse* al posto di cilindro ellittico;
- *iperbole* al posto di cilindro iperbolico;
- *punto* al posto di retta;
- *retta* al posto di piano.

Testo seguito:

[1] S.Abeasis, Elementi di algebra lineare e geometria, Zanichelli (Bologna), 2000.

[2] P.Albano-A.Parmeggiani, Elementi Introduttivi di Matematica, 2002.

Altri testi consigliati:

[3] M.Bramanti-C.D.Pagani-S.Salsa, Matematica, Zanichelli (Bologna), 2000.