Sui determinanti e l'indipendenza lineare di vettori

1. Si dice che m vettori $\mathbf{v}_1$, $\mathbf{v}_2$, ... , $\mathbf{v}_m$ di $\mathbb{R}^n$ sono **linearmente indipendenti**, se una loro combinazione lineare può dare il vettore nullo *solo* se i coefficienti della combinazione sono *tutti* nulli:
$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + ... + a_m\mathbf{v}_m = \mathbf{0} \qquad \Longrightarrow \qquad a_1 = a_2 = ... = a_m = 0.$$

2. Sia A una matrice quadrata di ordine n. Definiamo il **determinante** di A (in simboli $\det A$ o $|A|$) *per ricorsione* come
$$\begin{cases} \det A = \det(a_{11}) = a_{11} \text{ se } n = 1 \\ \det A = a_{11}A_{11} + a_{12}A_{12} + ... + a_{1n}A_{1n} \end{cases}$$
ove per ogni $k \in \{1, ..., n\}$ con A_{1k} rappresentiamo il *complemento algebrico di* a_{1k}, cioè il prodotto di $(-1)^{1+k}$ per il *minore complementare di* a_{1k} (cioè il determinante della matrice che si ottiene da A togliendo la prima riga e la k-esima colonna).

3. **Teorema di Laplace** (con ovvia simbologia): fissata comunque la riga $h \in \{1, ..., n\}$ (o la colonna $k \in \{1, ..., n\}$) si ha
$$\det A = a_{h1}A_{h1} + a_{h2}A_{h2} + ... + a_{hn}A_{hn} = a_{1k}A_{1k} + a_{2k}A_{2k} + ... + a_{nk}A_{nk}$$

4. **Proprietà**

   a) Se A ha *una riga* (oppure una colonna) che *è il vettore nullo* si ha $\det(A) = 0$: utilizzare il teorema di Laplace per sviluppare lungo la riga (o colonna) nulla!

   b) Se *si scambiano due righe* (o due colonne) di A il determinante cambia segno, ma non valore assoluto:
$$\begin{vmatrix} \vdots \\ \mathbf{a} \\ \vdots \\ \mathbf{b} \\ \vdots \end{vmatrix} = - \begin{vmatrix} \vdots \\ \mathbf{b} \\ \vdots \\ \mathbf{a} \\ \vdots \end{vmatrix}$$
   Per vederlo basta pensare di sviluppare i due determinanti successivamente secondo tutte le righe che non sono state toccate dallo scambio, fino ad avere una combinazione lineare dei determinanti di matrici di ordine due che provengono dalle due righe scambiate: in questi sviluppi il coefficiente del determinante della matrice formata con le colonne di posto i e j estratte dalla matrice $\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$ coincide con il coefficiente del determinante della matrice formata con le colonne di posto i e j estratte dalla matrice $\begin{pmatrix} \mathbf{b} \\ \mathbf{a} \end{pmatrix}$ poiché non ci sono stati cambiamenti nel resto della matrice. La tesi segue dal fatto che per ogni i e j si ha $\begin{vmatrix} b_i & b_j \\ a_i & a_j \end{vmatrix} = - \begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix}$

   c) Se A contiene *due righe* (o due colonne) *uguali* si ha $\det A = 0$

   Infatti pensando che nella matrice al punto (**b**) sia $\mathbf{a} = \mathbf{b}$ si ricava $\det A = -\det A$ e quindi $\det A = 0$

   d) Se A' è la matrice ottenuta da A *moltiplicando una* sua *riga* (o una sua colonna) per un numero λ si ha $\det A' = \lambda \det A$

   Infatti basta sviluppare secondo la riga (o la colonna) moltiplicata per λ per mettere in evidenza il coefficiente λ comune a tutti i prodotti $a_{hk}A_{hk}$.

e) Se si pensa una riga (o colonna) di A come la *somma di due righe* (o colonne): $\mathbf{a} = \mathbf{b} + \mathbf{c}$, $\det A$ *coincide con la somma dei determinanti delle due matrici* B e C ottenute da A sostituendo alla riga (o colonna) $\mathbf{a}$ una volta la $\mathbf{b}$ e l'altra la $\mathbf{c}$.

Infatti, sviluppiamo i determinanti di B e C rispettivamente secondo la riga (o la colonna) $\mathbf{b}$ e secondo la riga (o la colonna) $\mathbf{c}$: indicando con A_k il complemento algebrico dell'elemento a_k di posto k di $\mathbf{a}$

$$\det B + \det C = \sum_{k=1}^{n} b_k A_k + \sum_{k=1}^{n} c_k A_k = \sum_{k=1}^{n} (b_k + c_k) A_k = \sum_{k=1}^{n} a_k A_k = \det A$$

f) Se A' è la matrice ottenuta da A *sommando a una* sua *riga* (o a una sua colonna) $\mathbf{a}$ *un'altra* $\mathbf{b}$ *moltiplicata per un numero* λ si ha $\det A' = \det A$

Infatti la riga (o la colonna) $\mathbf{a}+\lambda\mathbf{b}$ si presenta proprio come una somma di righe (o colonne). Applichiamo il punto (**e**) e osserviamo che la matrice che contiene $\lambda\mathbf{b}$ ha due righe (o colonne) proporzionali: ricordando (**c**) si ha

$$\det A' = \sum_{k=1}^{n} (a_k + \lambda b_k) A_k = \sum_{k=1}^{n} a_k A_k + \lambda \sum_{k=1}^{n} b_k A_k = \sum_{k=1}^{n} a_k A_k = \det A$$

g) Teorema di Binet: se A e B sono due matrici quadrate di ordine n, $\det AB = \det A \det B$

5. Una matrice A quadrata di ordine n è *invertibile se e solo se* $\det A$ *è non nullo.*
   Infatti, $\boxed{\text{se } A \text{ è invertibile}}$ esiste una matrice B tale che $AB = I$ (matrice identica) e quindi $\det A \det B = \det AB = \det I = 1$: perché questo sia possibile $\det A$ non deve essere nullo; anzi deve essere $\det B = \dfrac{1}{\det A}$.
   Viceversa, usando il teorema di Laplace si vede che per ogni scelta della riga h si ha $\det A = a_{h1}A_{h1} + a_{h2}A_{h2} + ... + a_{hn}A_{hn}$ e ricordando 4(**c**) si ha che se si sceglie una riga $k \neq h$ risulta $a_{k1}A_{h1} + a_{k2}A_{h2} + ... + a_{kn}A_{hn} = 0$ poiché questa somma corrisponde ad aver sostituito in A la riga h con una delle altre e quindi aver calcolato il determinante di una matrice con due righe uguali. Dunque

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} = \begin{pmatrix} \det A & 0 & \dots & 0 \\ 0 & \det A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \det A \end{pmatrix} = (\det A) \cdot I$$

   cioè $\boxed{\text{se } \det A \text{ è non nullo}}$ esiste l'inversa ed è $B = \dfrac{1}{\det A} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$

6. Sia A una matrice quadrata di ordine n: essa ha *determinante diverso da* 0 se e solo se i suoi n *vettori colonna* (o equivalentemente i suoi n vettori riga) sono *linearmente indipendenti.*

   Dimostrazione. $\boxed{\text{Se } \det A \neq 0}$ la matrice è invertibile cioè esiste una matrice B tale che $AB = I$. Allora si vede che la combinazione lineare dei vettori colonna di A mediante gli elementi della colonna j-esima di B dà la j-esima colonna di I cioè

$$\begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} = \mathbf{e}_j$$

   ove $\mathbf{a}_k$ è la k-esima colonna di A ed $\mathbf{e}_j$ è il vettore di $\mathbb{R}^n$ che ha tutti gli elementi nulli tranne quello di posto j, che vale 1.

I vettori $\mathbf{e}_1$, $\mathbf{e}_2$, ... , $\mathbf{e}_n$ sono ovviamente linearmente indipendenti in $\mathbb{R}^n$, poiché
$$a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + ... + a_n\mathbf{e}_n = (a_1, a_2, ..., a_n) = \mathbf{0}$$
se e solo se tutti gli a_j sono nulli: quindi non possono essere generati per combinazione lineare da n vettori tra di loro dipendenti (e quindi non in grado di generare tutto $\mathbb{R}^n$). Dunque gli n vettori colonna di A che generano tutti i vettori $\mathbf{e}_1$, $\mathbf{e}_2$, ... , $\mathbf{e}_n$ devono essere linearmente indipendenti.

Viceversa $\boxed{\text{se gli } n \text{ vettori } \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \text{ sono linearmente indipendenti}}$ essi sono una base di $\mathbb{R}^n$; in particolare ogni $\mathbf{e}_j$ è combinazione lineare di $\mathbf{a}_1$, $\mathbf{a}_2$, ... , $\mathbf{a}_n$

$$\mathbf{e}_j = b_{1j}\mathbf{a}_1 + b_{2j}\mathbf{a}_2 + ... + b_{nj}\mathbf{a}_n$$

quindi la matrice B ottenuta accostando i vettori colonna $(b_{1j}, b_{2j}, ..., b_{nj})^T$ è l'inversa della A e se la matrice A è invertibile il suo determinante è non nullo.

Quanto all'enunciato per le righe basta ricordare che le righe di A sono le colonne di A^T e che $\det A = \det A^T$.

7. Sia ora A una matrice di tipo (m, n) *non necessariamente quadrata.* Denotiamo con A' una sottomatrice quadrata di ordine r di A. Se $\det A' \neq 0$ gli r vettori riga e gli r vettori colonna di A dalla cui "intersezione" si ottiene la matrice A' sono linearmente indipendenti.

   Infatti consideriamo ad esempio le r colonne $\mathbf{a}_{i_1}$, $\mathbf{a}_{i_2}$,..., $\mathbf{a}_{i_r}$ con r righe delle quali si forma A': se una loro combinazione lineare $\lambda_1\mathbf{a}_{i_1} + \lambda_2\mathbf{a}_{i_2} + ... + \lambda_r\mathbf{a}_{i_r}$ è nulla, deve essere nulla in particolare la combinazione lineare dei sottovettori colonna di $\mathbf{a}_{i_1}$, $\mathbf{a}_{i_2}$,..., $\mathbf{a}_{i_r}$ che si ottengono prendendo esattamente le righe che compongono la matrice A', che però sono linearmente indipendenti, poiché $\det A' \neq 0$: quindi $\lambda_1 = \lambda_2 = ... = \lambda_r = 0$ e di conseguenza sono linearmente indipendenti (in base alla definizione di indipendenza) anche i vettori colonna di A [(1)].

8. Sia ancora A una matrice di tipo (m, n) *non necessariamente quadrata.* Se r vettori colonna (o riga) di A sono linearmente indipendenti, la matrice che si ottiene accostandoli contiene sicuramente una sottomatrice quadrata A' di ordine r tale che $\det A' \neq 0$.

   Svolgiamo la dimostrazione per i vettori colonna. Supponiamo (per non usare indici come i_j) che gli r vettori linearmente indipendenti siano le prime r colonne di A.
   Osserviamo che scambi di riga nella matrice B ottenuta accostandoli producono una matrice formata ancora da r vettori indipendenti [(2)] che chiameremo di nuovo B.
   In particolare, per rendere più economico il percorso che faremo per la dimostrazione potremmo pensare di aver spostato in fondo alla matrice le eventuali righe formate esclusivamente da zeri, che non possono comunque essere più di $n - r$, altrimenti le colonne di B sarebbero di fatto r vettori di uno spazio con meno di r componenti (le altre sono nulle!) e quindi non potrebbero essere linearmente indipendenti.

[1)] Si consiglia di provare a fare il ragionamento scrivendo il dettaglio su una matrice a coefficienti generici di tipo (3,4), supponendo $r = 2$.

[2)] Infatti ciò equivale a elencare le componenti in un ordine uguale per tutti i vettori colonna, ma diverso da quello naturale. Ciò significa solo cambiare l'ordine in cui si elencano i vettori della base standard. Ad esempio scrivere $\begin{pmatrix} a_{1i} \\ a_{3i} \\ a_{2i} \end{pmatrix}$ invece di $\begin{pmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \end{pmatrix}$ significa che si sono considerati i vettori della base standard nell'ordine **i**, **k**, **j** e si sono scritte le componenti di ogni vettore secondo i vettori della base presi in quest'ordine. Osserviamo però che in questa differente scrittura il vettore **k** ha componenti $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ e il vettore **j** ha componenti $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$!

Calcoliamo il determinante della sottomatrice $\begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_r \end{pmatrix}$ di B ottenuta accostando le sue prime r righe: se è diverso da zero il teorema è dimostrato; altrimenti (vedi punto 6) almeno una riga di tale sottomatrice dipende linearmente dalle altre e quindi può essere ottenuta come loro combinazione lineare. A meno di scambi di righe possiamo supporre che la riga combinazione lineare delle altre sia l'ultima delle r, cioè che sia $\mathbf{b}_r = \sum_{h=1}^{r-1} (\mu_{rh}\mathbf{b}_h)$, spostare tale riga in fondo alla matrice B e ricominciare dal calcolo del determinante della matrice formata dalle prime r righe della nuova matrice.
Proseguendo in questo modo o in qualche passo si individua una matrice quadrata di ordine r con determinante diverso da zero oppure, dopo aver accostato alle prime $r-1$ righe via via ciascuna delle altre, si ha una matrice in cui le ultime $n-r+1$ righe sono combinazioni lineari delle precedenti:

$$\begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_{r-1} \\ \sum_{h=1}^{r-1} (\mu_{rh}\mathbf{b}_h) \\ \vdots \\ \sum_{h=1}^{r-1} (\mu_{mh}\mathbf{b}_h) \end{pmatrix}$$

Notiamo che fin qui abbiamo sempre supposto che nella matrice di cui si calcola il determinante la prima riga sia sempre $\mathbf{b}_1$ e quindi non abbiamo ancora esaminato i determinanti di tutte le sottomatrici di B di ordine r; ma l'annullarsi del determinante delle $n-r+1$ matrici esaminate garantisce che anche le altre sottomatrici di B di ordine r hanno determinante nullo.
Infatti (proprietà (4**f**)) aggiungendo ad una riga di una matrice una combinazione lineare delle altre il determinante non cambia e quindi ad esempio

$$\begin{vmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_{r-3} \\ \sum_{h=1}^{r-1} (\mu_{rh}\mathbf{b}_h) \\ \sum_{h=1}^{r-1} \left(\mu_{r+1\ h}\mathbf{b}_h\right) \\ \sum_{h=1}^{r-1} \left(\mu_{r+2\ h}\mathbf{b}_h\right) \end{vmatrix} = \begin{vmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_{r-3} \\ \mu_{r\ r-2}\mathbf{b}_{r-2} + \mu_{r\ r-1}\mathbf{b}_{r-1} \\ \mu_{r+1\ r-2}\mathbf{b}_{r-2} + \mu_{r+1\ r-1}\mathbf{b}_{r-1} \\ \mu_{r+2\ r-2}\mathbf{b}_{r-2} + \mu_{r+2\ r-1}\mathbf{b}_{r-1} \end{vmatrix} = \text{ per la proprietà } 4(\mathbf{e})$$

$$= \begin{vmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_{r-3} \\ \mu_{r\ r-2}\mathbf{b}_{r-2} \\ \mu_{r+1\ r-2}\mathbf{b}_{r-2} + \mu_{r+1\ r-1}\mathbf{b}_{r-1} \\ \mu_{r+2\ r-2}\mathbf{b}_{r-2} + \mu_{r+2\ r-1}\mathbf{b}_{r-1} \end{vmatrix} + \begin{vmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_{r-3} \\ \mu_{r\ r-1}\mathbf{b}_{r-1} \\ \mu_{r+1\ r-2}\mathbf{b}_{r-2} + \mu_{r+1\ r-1}\mathbf{b}_{r-1} \\ \mu_{r+2\ r-2}\mathbf{b}_{r-2} + \mu_{r+2\ r-1}\mathbf{b}_{r-1} \end{vmatrix}$$

Iterando quest'ultimo passaggio si vede che il determinante iniziale è una combinazione con coefficienti che sono prodotti di coefficienti $\mu_{h\ r-2}$ e $\mu_{h\ r-1}$ (con $h = r, r+1, r+2$) di determinanti in cui due (o anche tre) righe coincidono, dato che nelle tre righe finali si trovano solo i vettori $\mathbf{b}_{r-2}$ o $\mathbf{b}_{r-1}$, e quindi sono nulli!

Ora supponiamo che tutti i determinanti delle matrici quadrate di ordine r estratti da B siano nulli e riprendiamo in esame la matrice formata dai primi r vettori riga: se il determinante è nullo, anche i corrispondenti vettori colonna $\begin{pmatrix} a_{11} \\ \vdots \\ a_{r1} \end{pmatrix}, ..., \begin{pmatrix} a_{1r} \\ \vdots \\ a_{rr} \end{pmatrix}$ sono linearmente dipendenti, cioè esistono dei coefficienti non tutti nulli $\lambda_1, ..., \lambda_r$ tali che

$$\lambda_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{r1} \end{pmatrix} + ... + \lambda_r \begin{pmatrix} a_{1r} \\ \vdots \\ a_{rr} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Allora considerando la combinazione lineare delle colonne della matrice B tramite gli stessi coefficienti si trova

$$\lambda_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{r1} \\ \sum_{h=1}^{r-1} \left(\mu_{r+1\,h}\, a_{h1}\right) \\ \vdots \\ \sum_{h=1}^{r-1} \left(\mu_{mh} a_{h1}\right) \end{pmatrix} + ... + \lambda_r \begin{pmatrix} a_{1r} \\ \vdots \\ a_{rr} \\ \sum_{h=1}^{r-1} \left(\mu_{r+1\,h} a_{hr}\right) \\ \vdots \\ \sum_{h=1}^{r-1} \left(\mu_{mh} a_{hr}\right) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sum_{h=1}^{r-1} \mu_{r+1\,h} \left(\sum_{k=1}^{r} \lambda_k a_{hk}\right) \\ \vdots \\ \sum_{h=1}^{r-1} \mu_{mh} \left(\sum_{h=1}^{r} \lambda_k a_{hk}\right) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

e quindi i vettori colonna della matrice B non sono indipendenti, contro l'ipotesi iniziale [(3)].

9. Mettendo insieme i risultati dimostrati nei punti 7 e 8 si vede che in una matrice A di tipo (m, n) non necessariamente quadrata, r *vettori* colonna (o riga) *sono linearmente indipendenti se e solo se la matrice che si ottiene accostandoli contiene una sottomatrice quadrata* A' di ordine r *tale che* $\det A' \neq 0$.

   Ne consegue che è equivalente definire il **rango di una matrice** A come il più grande numero di vettori riga (o colonna) linearmente indipendenti contenuti in A oppure come il più grande intero r tale che in A sia contenuta una sottomatrice quadrata A' di ordine r avente determinante diverso da 0.

10. Notiamo che la parte della dimostrazione che evidenzia che se i determinanti delle $n - r + 1$ matrici ottenute considerando le prime $r - 1$ righe e via via ciascuna delle altre sono nulli allora sono nulli anche tutti i determinanti delle altre matrici di ordine r estraibili da B è quanto sta dietro la regola di Kronecker per il calcolo del rango di una matrice.

[3)] Anche in questa dimostrazione la varietà di indici e parametri può risultare ostica: si consiglia di provare a fare il ragionamento scrivendo i passaggi in dettaglio su una matrice B a coefficienti generici formata da 2 o al massimo 3 colonne aventi 5 componenti (quindi $m = 5$ e $r = 2$ o 3).