CAPITOLO 6

Determinanti, autovalori e autospazi

1. Proprietà del determinante

Sia
$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
una matrice 2×2. Se $ad - bc \neq 0$ allora A è invertibile e
$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$
Invece, se $ad - bc = 0$ allora A ha rango al massimo uno e non è invertibile.

Vorremo adesso cercare di determinare un polinomio che abbia una simile proprietà per matrici $n \times n$, cioè matrici quadrate più grandi.

Definizione 6.1. Sia n un intero positivo. Il *determinante* det è una funzione $M_{n\times n}(K) \to K$ tale che

(1) $\det(I_n) = 1$.

(2) det è lineare sulle righe. Siano $v_1, \ldots, v_n, w \in K^n$ e $\lambda \in K$ allora
$$\det \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{i-1} \\ v_i + \lambda w \\ v_{i+1} \\ \vdots \\ v_n \end{pmatrix} = \det \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{i-1} \\ v_i \\ v_{i+1} \\ \vdots \\ v_n \end{pmatrix} + \lambda \det \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{i-1} \\ w \\ v_{i+1} \\ \vdots \\ v_n \end{pmatrix}$$

(3) det è alternante, cioè se A è una matrice $n \times n$ con due righe identiche allora $\det(A) = 0$.

Questa definizione è una forma matematica di wishful thinking, cioè definiamo una funzione semplicemente richiedendo le proprietà che vorremmo avere e speriamo (ottimisticamente) che la funzione esista e sia ben definita e unica. Nel nostro caso l'esistenza è abbastanza complicata, non è invece troppo difficile mostrare che il determinante, quando esiste, è unico.

Iniziamo col mostrare che $ad - bc$ è un determinante per le matrici 2×2:

Lemma 6.2. *La funzione* $\det : M_{2\times 2}(K) \to K$ *tale che*
$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$
è un determinante.

Dimostrazione. Se $A = I_2$, cioè $a = d = 1, b = c = 0$, troviamo che $ad - bc = 1$, quindi $\det(I_2) = 1$.

Se A ha due righe identiche, cioè $c = a, d = b$, troviamo che $ad - bc = ab - ba = 0$, quindi $\det(A) = 0$.

Rimane da mostrare la linearià. Iniziamo con la linearità nella seconda riga. Abbiamo che

$$\begin{aligned}
\det\begin{pmatrix} a & b \\ c_1+\lambda c_2 & d_1+\lambda d_2 \end{pmatrix} &= a(d_1+\lambda d_2)-b(c_1+\lambda c_2) \\
&= ad_1-bc_1+\lambda(ad_2-bc_2) \\
&= \det\begin{pmatrix} a & b \\ c_1 & d_1 \end{pmatrix}+\lambda\det\begin{pmatrix} a & b \\ c_2 & d_2 \end{pmatrix}
\end{aligned}$$

La linearità sulle prima righe si mostra in modo simile. □

Adesso mostriamo che se $\det : M_{n\times n}(K) \to K$ esiste allora è unico. Per mostrare l'unicità iniziamo studiando che cosa succede al determinante quando applichiamo operazioni elementari sulle righe.

Lemma 6.3. *Sia A una matrice $n\times n$ e sia A' ottenuta da A sommando λ volte la i-esima riga alla j-esima riga, allora* $\det(A)=\det(A')$.

Dimostrazione. Usando la linearità in riga j troviamo

$$\det\begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_{j-1} \\ v_j+\lambda v_i \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix} = \det\begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_{j-1} \\ v_j \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix} + \lambda\det\begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_{j-1} \\ v_i \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix}$$

La seconda matrice a destra ha due righe identiche. Quindi il suo determinante è zero e troviamo che $\det(A')=\det(A)+0$. □

Dalla linearità sulle righe segue direttamente che se moltiplichiamo una riga con λ allora anche il determinante viene moltiplicato con λ.

Rimane quindi da studiare che cosa succede quando si scambiano due righe.

Lemma 6.4. *Sia A una matrice $n\times n$ e A' la matrice ottenuta da A scambiando due righe, allora*

$$\det(A') = -\det(A).$$

Dimostrazione. Usando il fatto che il determinante è alternante (per la prima equazione) e lineare sulle righe (per le altre equazioni) troviamo

$$
\begin{aligned}
0 = \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_i + v_j \\ v_{i+1} \\ \vdots \\ v_{j-1} \\ v_j + v_i \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix} &= \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_i \\ v_{i+1} \\ \vdots \\ v_{j-1} \\ v_j + v_i \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix} + \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_j \\ v_{i+1} \\ \vdots \\ v_{j-1} \\ v_j + v_i \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix} \\
&= \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_i \\ v_{i+1} \\ \vdots \\ v_{j-1} \\ v_i \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix} + \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_i \\ v_{i+1} \\ \vdots \\ v_{j-1} \\ v_j \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix} + \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_j \\ v_{i+1} \\ \vdots \\ v_{j-1} \\ v_i \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix} + \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_j \\ v_{i+1} \\ \vdots \\ v_{j-1} \\ v_j \\ v_{j+1} \\ \vdots \\ v_n \end{pmatrix}
\end{aligned}
$$

Delle quattro matrici alla fine, due hanno due righe identiche e quindi il loro determinante è zero. Le altre due matrice sono A e A'. Dunque troviamo

$$0 = \det(A) + \det(A').$$

□

Lemma 6.5. *Se A ha una riga composta esclusivamente da zeri, allora* $\det(A) = 0$.

Dimostrazione. Scambiando le righe di A cambia solo il segno di $\det(A)$, quindi possiamo assumere che sia l'ultima riga ad essere zero. Sia $w \in K^n$, allora usando la linearità troviamo

$$\det(A) = \det \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ \vec{0} \end{pmatrix} = \det \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ 0w \end{pmatrix} = 0 \det \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ w \end{pmatrix} = 0.$$

□

Adesso possiamo mostrare che esiste al massimo una funzione det.

Proposizione 6.6. *Se* $\det : M_{n\times n}(K) \to K$ *esiste, allora è unico. Inoltre, abbiamo che* $\det(A) = 0$ *se e solo se* $\mathrm{rg}(A) < n$.

Dimostrazione. Sia A una matrice e B una forma a scala di A, allora

$$\det(B) = c\det(A)$$

con $c = (-1)^e \lambda_1 \dots \lambda_t$, dove e è il numero di scambi di righe e i λ_i sono i fattori con cui abbiamo moltiplicato le righe. Questi fattori sono nonzero, quindi c è diverso da zero e $\det(A) = \frac{1}{c}\det(B)$.

Se il rango di A non è n, allora B ha una riga costituita da zeri. Dal Lemma precedente segue che $\det(B) = 0$ e quindi $\det(A) = \frac{1}{c}\det(B) = 0$.

Se il rango di A è n possiamo applicare operazioni elementari sulle righe fino ad ottenere $B = I_n$ e abbiamo che $\det(A) = \frac{1}{c}\det(I_n) = \frac{1}{c} \neq 0$. In particolare $\det(A)$ è determinata dalle operazioni necessarie per ottenere la forma a scala I_n. □

La proposizione illustra come si può calcolare il determinante usando le operazioni sulle righe per la matrice A. In molti casi questo è anche il modo più veloce per calcolare il determinante.

Corollario 6.7. *Sia A una matrice $n \times n$.*

(1) Se le righe di A sono linearmente dipendenti allora $\det(A) = 0$.
(2) Abbiamo $\det(\lambda A) = \lambda^n \det(A)$.
(3) Se A è una matrice triangolare superiore ($A_{ij} = 0$ se $i > j$) allora $\det(A)$ *è il prodotto degli elementi sulla diagonale:*

$$\det(A) = a_{11}a_{22}\dots a_{nn}.$$

Similmente, se A è una matrice triangolare inferiore ($A_{ij} = 0$ se $i < j$) allora $\det(A)$ *è il prodotto degli elementi sulla diagonale.*

Dimostrazione. (1) Se le righe di A sono linearmente dipendenti allora $\mathrm{rg}(A) < n$ e $\det(A) = 0$.
(2) Moltiplicando una riga di A con λ cambiamo il determinante di un fattore λ. Nella matrice λA abbiamo moltiplicato tutte le n righe con λ, quindi il determinante sarà cambiato di un fattore λ^n.
(3) Sia A una matrice triangolare superiore. Consideriamo prima il caso in cui tutti gli a_{ii} siano diversi da zero. Allora tutti gli elementi sulla diagonale sono pivot, $\mathrm{rg}(A) = n$ e A è in forma a scala. Possiamo adesso trovare una forma a scala ridotta semplicemente sommando multipli di righe su altre righe. La matrice ottenuta soddisfa $a_{ij} = 0$ per ogni $i \neq j$, e non abbiamo cambiato il determinante.

L'ultima matrice è ottenuta da I_n moltiplicando i-esima riga con a_{ii}, quindi $\det(A) = a_{11}a_{22}\dots a_{nn}\det(I_n) = a_{11}a_{22}\dots a_{nn}$.

Nel caso in cui uno degli a_{ii} sia zero, allora $\mathrm{rg}(A) < n$ e $\det(A) = 0$ e abbiamo finito.

Per le matrici triangolari inferiori si procede similmente.

□

Esempio 6.8. Applicando operazioni sulle righe troviamo

$$\det\begin{pmatrix}1 & 2 & 3\\ 1 & 3 & 4\\ 1 & 3 & 5\end{pmatrix} = \det\begin{pmatrix}1 & 2 & 3\\ 0 & 1 & 1\\ 0 & 1 & 2\end{pmatrix} = \det\begin{pmatrix}1 & 2 & 3\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}$$

L'ultima matrice è una matrice triangolare superiore. Quindi il suo determinante è il prodotto degli elementi sulla diagonale e quindi il determinante è 1.

2. Teorema di Binet

In questa sezione mostriamo il rapporto tra il determinante ed il prodotto tra matrici.

Lemma 6.9. *Per le matrici elementari abbiamo che*

$$\det(P(i,j)) = -1;\ \det(S(i,j;\lambda)) = 1;\ \det(M(i,\lambda)) = \lambda$$

Dimostrazione. La matrice $P(i,j)$ è ottenuta dalla matrice I_n scambiando due righe, quindi $\det(P(i,j)) = -\det(I_n) = -1$.

Le matrici $S(i,j;\lambda)$ sono matrici triangolari, quindi il determinate è il prodotto degli elementi sulla diagonale (Corollario 6.7) e $\det(S(i,j;\lambda)) = 1$. Similmente troviamo che $\det M(i,\lambda) = \lambda$. □

Teorema 6.10 (Binet). *Siano $A, B \in M_{n\times n}(K)$. Allora*

$$\det(AB) = \det(A)\det(B).$$

Dimostrazione. Cominciamo con il caso in cui A sia una matrice elementare.

Se $A = P(i,j)$ abbiamo che $\det(A) = -1$ e che la matrice AB è la matrice ottenuta da B scambiano due righe (Proposizione 2.12), quindi $\det(AB) = -\det(B)$. In particolare $\det(AB) = \det(A)\det(B)$.

Se $A = M(i,\lambda)$, allora $\det(A) = \lambda$. La matrice AB è ottenuta da B moltiplicando la i-esima riga per λ (Proposizione 2.12). Quindi $\det(AB) = \lambda\det(B)$, e $\det(AB) = \det(A)\det(B)$.

Se $A = S(i,j;\lambda)$, allora $\det(A) = 1$. La matrice AB è ottenuta da B sommando un multiplo di una riga ad un'altra. In particolare $\det(AB) = \det(B)$ e quindi $\det(AB) = \det(A)\det(B)$.

Assumiamo adesso che A non sia una matrice elementare ma che $\mathrm{rg}(A) = n$. Allora A è il prodotto di matrici elementari (Dimostrazione della Proposizione 2.14). Tramite induzione sul numero delle matrici elementari nel prodotto si trova che $\det(AB) = \det(A)\det(B)$.

Se $\mathrm{rg}(A) < n$, allora la moltiplicazione per A non è suriettiva. L'immagine della moltiplicazione per AB è contenuta nell'immagine della moltiplicazione per A, quindi non è suriettiva e $\mathrm{rg}(AB) < n$. In particolare $\det(A) = 0$ e $\det(AB) = 0$ e quindi $\det(A)\det(B) = 0 = \det(AB)$. □

Corollario 6.11. *Se A è una matrice invertibile, allora* $\det(A) \neq 0$ *e*

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

.

Dimostrazione. Da $AA^{-1} = I_n$ segue che

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1}).$$

□

Corollario 6.12. *Abbiamo*

$$\det(A) = \det(A^T)$$

Dimostrazione. Se $\mathrm{rg}(A) < n$ allora anche $\mathrm{rg}(A^T) < n$ (Corollario 4.58) e quindi $\det(A) = 0$ e $\det(A^T) = 0$.

Se A è una matrice elementare abbiamo che $\det(A) = \det(A^T)$. Questo segue dal fatto che $P(i,j)^T = P(i,j)$, $M(i,\lambda)^T = M(i,\lambda)$ e $S(i,j;\lambda)^T = S(j,i;\lambda)$.

Se $\mathrm{rg}(A) = n$ allora possiamo scrivere $A = E_1 \dots E_n$, con E_i matrici elementari. Allora $A^T = E_n^T E_{n-1}^T \dots E_1^T$. Utilizzando due volte il Teorema di Binet troviamo che

$$\det(A^T) = \det(E_n^T)\det(E_{n-1}^T)\dots\det(E_1^T) = \det(E_n)\dots\det(E_1) = \det(A)$$

□

Osservazione 6.13. Le operazioni sulle righe di A corrispondono ad operazioni sulle le colonne di A^T. Essendo $\det(A) = \det(A^T)$, segue che possiamo anche applicare operazioni sulle colonne di A per calcolare il determinante. Quindi se scambiano due colonne allora il determinante cambia segno, se sommiamo un multiplo di una colonna ad un'altra, allora il determinante non cambia, e se moltiplichiamo una colonna per un numero λ allora il determinante viene moltiplicato per λ.

Esempio 6.14. Abbiamo

$$\begin{aligned}
\det\begin{pmatrix} 2 & 1 & -1 & 3 \\ -1 & 3 & 1 & 2 \\ 2 & -1 & 4 & 1 \\ 3 & -2 & 1 & 4 \end{pmatrix} &= -\det\begin{pmatrix} -1 & 3 & 1 & 2 \\ 2 & 1 & -1 & 3 \\ 2 & -1 & 4 & 1 \\ 3 & -2 & 1 & 4 \end{pmatrix} \\
&= -\det\begin{pmatrix} -1 & 3 & 1 & 2 \\ 0 & 7 & 1 & 7 \\ 0 & 5 & 6 & 5 \\ 0 & 7 & 4 & 10 \end{pmatrix} \\
&= \det\begin{pmatrix} -1 & 1 & 3 & 2 \\ 0 & 1 & 7 & 7 \\ 0 & 6 & 5 & 5 \\ 0 & 4 & 7 & 10 \end{pmatrix} \\
&= \det\begin{pmatrix} -1 & 1 & 3 & 2 \\ 0 & 1 & 7 & 7 \\ 0 & 0 & -37 & -37 \\ 0 & 0 & -21 & -18 \end{pmatrix} \\
&= -37\det\begin{pmatrix} -1 & 1 & 3 & 2 \\ 0 & 1 & 7 & 7 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -21 & -18 \end{pmatrix} \\
&= -37\det\begin{pmatrix} -1 & 1 & 3 & 2 \\ 0 & 1 & 7 & 7 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \\
&= (-37)\cdot(-1)\cdot(1)\cdot(1)\cdot(3) = 111
\end{aligned}$$

3. Esistenza del determinante

L'esistenza è più complicata. Daremo una formula per $\det(A)$ e verificheremo che soddisfi le proprietà della Definizione 6.1. A questo punto diamo una formula senza motivazione.

Per una matrice $n \times n$ A e interi i, j tali che $1 \leq i, j \leq n$ denotiamo con $A^{i,j}$ la matrice $(n-1) \times (n-1)$ ottenuta da A togliendo la riga i-esima e la colonna j-esima.

Sia adesso $d_n : M_{n\times n}(K) \to K$ la funzione che per $n = 1$ manda la matrice (α) nel numero $\alpha \in K$ e, per $n \geq 2$, definiamo ricorsivamente

$$d_n(A) := \sum_{i=1}^{n} (-1)^{i+1} a_{i1} d_{n-1}(A^{i,1})$$

Teorema 6.15. *La funzione d_n è il determinante.*

Dimostrazione. La dimostrazione è per induzione su n. Per $n = 1$ è chiaro che la funzione $(\alpha) \mapsto \alpha$ è lineare sull'unica riga, e manda $\det(I_1) = \det(1) = 1$. La condizione sull'alternanza è da controllare per $n \geq 2$.

Assumiamo che d_{n-1} sia il determinante e dimostriamo che d_n è il determinante. La prima cosa da mostrare è $d_n(I_n) = 1$.

$$d_n(I_n) = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} d_{n-1}(I_n^i).$$

Per $i \neq 1$ abbiamo che $a_{i,1} = 0$ quindi la sommatoria si riduce al solo indice $i = 1$ per cui $a_{1,1} = 1$. Perciò

$$d_n(I_n) = d_{n-1}(I_n^{1,1}).$$

Togliendo la prima riga e la prima colonna da I_n troviamo I_{n-1} quindi $d_{n-1}(I_n^{1,1}) = d_{n-1}(I_{n-1}) = 1$ (usando l'ipotesi induttiva).

Dimostriamo ora la linearità sulle righe. Supponiamo che la riga k-esima di A si possa scrivere come $v_k + \lambda w_k$. Sia B la matrice ottenuta da A mettendo v_k come riga k e sia C la matrice ottenuta da A mettendo w_k come riga k. Verifichiamo che $d_n(A) = d_n(B) + \lambda d_n(C)$. Per $i \neq k$ troviamo $a_{ij} = b_{ij} = c_{ij}$ e per ipotesi induttiva

$$d_{n-1}(A^{i,1}) = d_{n-1}(B^{i,1}) + \lambda d_{n-1}(C^{i,1}).$$

Per $i = k$ troviamo invece $A^{k,1} = B^{k,1} = C^{k,1}$ e $a_{kj} = b_{kj} + \lambda c_{kj}$. Sostituendo otteniamo:

$$\begin{aligned}
d_n(A) &= \sum_{i=1}^{n} (-1)^{i+1} a_{i1} d_{n-1}(A^{i,1}) \\
&= (-1)^{k+1} a_{k1} d_{n-1}(A^{k,1}) + \sum_{i=1, i\neq k}^{n} (-1)^{i+1} a_{i1} d_{n-1}(A^{i,1}) \\
&= (-1)^{k+1} (b_{k1} + \lambda c_{k1}) d_{n-1}(A^{k,1}) + \\
&\qquad + \sum_{i=1, i\neq k}^{n} (-1)^{i+1} a_{i1} (d_{n-1}(B^{i,1}) + \lambda d_{n-1}(C^{i,1})) \\
&= \sum_{i=1}^{n} (-1)^{i+1} b_{i1} d_{n-1}(B^{i,1}) + \lambda \sum_{i=1}^{n} (-1)^{i+1} c_{i1} d_{n-1}(C^{i,1}). \\
&= d_n(B) + \lambda d_n(C)
\end{aligned}$$

Dimostriamo ora l'alternanza. Assumiamo che la riga k-esima e la riga m-esima di A siano uguali, con $k < m$. Allora se $i \neq k, m$ abbiamo che la matrice $A^{i,1}$ ha due righe identiche e per ipotesi induttiva troviamo che $d_{n-1}(A^{i,1}) = 0$ per $i \neq k, m$. Dal fatto che la riga k e la riga m coincidono troviamo che le matrici $A^{k,1}$ e $A^{m,1}$ hanno le stesse righe, ma in una ordine diverso. La riga k di $A^{m,1}$ è la riga $m-1$ di $A^{k,1}$, e le righe $k+1, \ldots, m-1$ di $A^{m,1}$ sono le righe $k, \ldots, m-2$ di $A^{k,1}$. Quindi dobbiamo usare $m - k + 1$ scambi di righe per ottenere $A^{k,1}$ da $A^{m,1}$. Essendo d_{n-1} il determinante esso è alternante quindi scambiare due righe cambia il segno di d_{n-1}, quindi $d_{n-1}(A^{k,1}) = (-1)^{m-k+1} d_{n-1}(A^{m,1})$. (Proposizione 6.4)

$$\begin{aligned}
d_n(A) &= \sum_{i=1}^{n} (-1)^{i+1} a_{i1} d_{n-1}(A^{i,1}) \\
&= (-1)^{k+1} a_{k1} d_{n-1}(A^{k,1}) + (-1)^{m+1} a_{m1} d_{n-1}(A^{m,1}) \\
&= (-1)^{k+1} a_{m1} (-1)^{m-k+1} d_{n-1}(A^{m,1}) + (-1)^{m+1} a_{m1} d_{n-1}(A^{m,1}) \\
&= 0.
\end{aligned}$$

□

Teorema 6.16. *Il determinante esiste ed è unico.*

Dimostrazione. Abbiamo visto che d_n è il determinante in $M_{n,n}(K)$ per ogni $n \geq 1$. Nella sezione precedente avevamo dimostrato l'unicità. □

Corollario 6.17. *(Sviluppo di Laplace) Per ogni $j \in \{1, \dots n\}$ abbiamo*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A^{i,j})$$

(sviluppo lunga la colonna j).

Per ogni $i \in \{1, \dots n\}$ abbiamo che

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A^{i,j})$$

(sviluppo lunga la riga i.)

Dimostrazione. La formula per $d_n(A)$ è lo sviluppo di $\det(A)$ lungo la prima colonna.

Sia B la matrice tale che la prima colonna di B sia la colonna j-esima di A e la seconda fino alla j-esima colonna di B siano la prima fino alla $j-1$-esima colonna di A. Allora B è ottenuta da A facendo $j-1$ scambi di colonna. Inoltre abbiamo che $B^{i,1} = A^{i,j}$ e $b_{i1} = a_{ij}$. Quindi

$$\begin{aligned} \det(A) &= (-1)^{j-1} \det(B) \\ &= (-1)^{j-1} \sum_{i=1}^{n} (-1)^{i-1} b_{i1} \det(B^{i,1}) \\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A^{i,j}) \end{aligned}$$

e troviamo la formula per lo sviluppo lunga la colonna j.

Se adesso sia $B = A^T$ allora $B^{j,i} = A^{i,j}$ e $b_{ji} = a_{ij}$. Allora sviluppando $\det(B)$ lunga la j-esima colonna dà

$$\begin{aligned} \det(A) &= \det(B) \\ &= \sum_{j=1}^{n} (-1)^{i+j} b_{ji} B^{j,i} \\ &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} A^{i,j} \end{aligned}$$

e troviamo lo sviluppo lungo la riga i. □

Esempio 6.18. Facciamo ora un primo calcolo utilizzando lo sviluppo lungo la seconda colonna.

$$\begin{aligned} &\det \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 3 & 5 \end{pmatrix} \\ = \quad & (-1)^{1+2}\, 2 \det \begin{pmatrix} 1 & 4 \\ 1 & 5 \end{pmatrix} + (-1)^{2+2}\, 3 \det \begin{pmatrix} 1 & 3 \\ 1 & 5 \end{pmatrix} + (-1)^{3+2}\, 3 \det \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} \\ = \quad & -2(1 \cdot 5 - 1 \cdot 4) + 3(1 \cdot 5 - 1 \cdot 3) - 3(1 \cdot 4 - 1 \cdot 3) = -2 + 6 - 3 = 1 \end{aligned}$$

Osservazione 6.19. Per calcolare il determinante di una matrice $n \times n$ utilizzando lo sviluppo di Laplace si devono calcolare n determinanti di matrici $(n-1) \times (n-1)$. Continuando con questo ragionamento possiamo concludere di dover calcolare $n!$ determinanti di matrici 1×1. Quindi dobbiamo fare circa $n!$ operazioni aritmetiche (somme e moltiplicazioni) per calcolare il determinante.

Se, invece, utilizziamo il procedimento di Gauß per calcolare il determinante dobbiamo arrivare alla forma a scala di A. Per questo dobbiamo ottenere $\frac{1}{2}n(n-1)$ entrate uguali a zero. Per avere un'entrata uguale a zero dobbiamo fare circa n operazioni aritmetiche ed in totale dobbiamo quindi fare circa $\frac{1}{2}n^2(n-1)$ operazioni aritmetiche.

Per n grande abbiamo che $\frac{1}{2}n^3$ è molto più piccolo di $n!$, e in pratica si nota già con $n = 4$ che è molto più efficiente applicare il procedimento di Gauß che lo sviluppo di Laplace.

4. Esistenza del determinante (approccio alternativo)*

In questa sezione diamo un approccio alternativo per la dimostrazione dell'esistenza del determinante. Questo approccio è più elegante, ma richiede più teoria.

DEFINIZIONE 6.20. Una permutazione di $\{1, \dots, n\}$ è una mappa biettiva $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$.

Denotiamo con S_n tutte le permutazioni di $\{1, \dots, n\}$.

ESEMPIO 6.21. Ci sono due permutazioni di $\{1, 2\}$, abbiamo $\sigma(1) = 1, \sigma(2) = 2$ e abbiamo $\sigma(1) = 2$, $\sigma(2) = 1$.

Ci sono sei permutazioni di $\{1, 2, 3\}$. Per $\sigma(1)$ abbiamo tre scelte, una volta scelto $\sigma(1)$ abbiamo ancora due possibilità per $\sigma(2)$.

In generale abbiamo $n!$ permutazioni di $\{1, \dots, n\}$.

Ogni permutazione di $\{1, \dots, n\}$ si può identificare con uno schema di n caselle marcate in un quadro di dimensione $n \times n$. Si consideri nella i-esima riga la casella nella $\sigma(i)$-esima colonna. In quel modo si determina uno schema tale che in ogni riga e in ogni colonna c'è un'unica casella marcata.

ESEMPIO 6.22. Per $\sigma \in S_3$ con $\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 1$ troviamo

	X	
		X
X		

DEFINIZIONE 6.23. Sia σ una permutazione di $\{1, \dots, n\}$. Allora il *segno* di σ, denotato con $\operatorname{sign}(\sigma)$, è $(-1)^e$ dove $e = \#\{(i, j) \mid \sigma(i) > \sigma(j); i < j\}$.

LEMMA 6.24. *Per $\sigma \in S_n$ abbiamo*

$$\operatorname{sign}(\sigma) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}$$

DIMOSTRAZIONE. I prodotti $\prod_{i<j} i - j$ e $\prod_{i<j} \sigma(i) - \sigma(j)$ contengono gli stessi fattori, ma potrebbero avere segno diverso. In particolare

$$\frac{\prod_{i<j} \sigma(j) - \sigma(i)}{\prod_{i<j} j - i} \in \{-1, 1\}$$

Per ogni coppia (i, j) con $j > i$ e $\sigma(i) > \sigma(j)$ troviamo un -1. Se abbiamo invece $\sigma(j) > \sigma(i)$ troviamo un 1. Quindi

$$\operatorname{sign}(\sigma) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}$$

□

LEMMA 6.25. *Se* $\sigma, \tau \in S_n$ *allora*

$$\operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau).$$

DIMOSTRAZIONE. Abbiamo

$$\operatorname{sign}(\sigma\tau) = \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j-i} = \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \prod_{i<j} \frac{\tau(j) - \tau(i)}{j-i}$$

Il secondo fattore a destra è $\operatorname{sign}(\tau)$. Rimane da mostrare che il primo fattore è $\operatorname{sign}(\sigma)$.

Abbiamo che il primo fattore

$$\prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)}$$

è uguale a

$$\begin{aligned}
&= \prod_{\substack{i<j \\ \tau(i)<\tau(j)}} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \prod_{\substack{i<j \\ \tau(i)>\tau(j)}} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \\
&= \prod_{\substack{i<j \\ \tau(i)<\tau(j)}} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \prod_{\substack{i>j \\ \tau(i)<\tau(j)}} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \\
&= \prod_{\tau(i)<\tau(j)} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \\
&= \prod_{i<j} \frac{\sigma(i) - \sigma(j)}{j-i} = \operatorname{sign}(\sigma).
\end{aligned}$$

Nella penultima equazione abbiamo usato che τ è biettiva. Quindi troviamo

$$\operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau).$$

□

DEFINIZIONE 6.26. Una *trasposizione* è una permutazione che permuta soltanto due elementi k, m, quindi $\sigma(i) = i$ per $i \neq k, m$ e $\sigma(k) = m, \sigma(m) = k$.

ESEMPIO 6.27. Una trasposizione soddisfa $\operatorname{sign}(\tau) = -1$. Quindi se σ è un prodotto di un numero pari di trasposizioni allora $\operatorname{sign}(\sigma) = 1$, se è un prodotto di un numero dispari di trasposizioni allora $\operatorname{sign}(\sigma) = -1$.

Adesso definiamo

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

e mostriamo che soddisfa le tre condizioni della Definizione 6.1

PROPOSIZIONE 6.28. *Abbiamo che* $\det(I_n) = 1$.

DIMOSTRAZIONE. Se $\sigma \neq \mathrm{id}$ allora per almeno un i abbiamo $\sigma(i) \neq i$ e quindi $a_{i\sigma(i)} = 0$. Nella somma che definisce il determinante abbiamo che tutti i prodotti sono zero, tranne il prodotto per $\sigma = \mathrm{id}$. Per $\sigma = \mathrm{id}$ troviamo $\operatorname{sign}(\sigma) = 1$ e

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = (+1) \prod_{i=1}^{n} a_{ii} = 1$$

□

PROPOSIZIONE 6.29. *Abbiamo che* det *è lineare sulle righe.*

DIMOSTRAZIONE. Siano B, C due matrici con tutte le righe identiche, meno la riga k. Sia A la matrice con le stesse righe di B e C, tranne la riga k, che è la riga k di A più λ volte riga k di B. Quindi scriviamo $A_{ij} = B_{ij} = C_{ij} = a_{ij}$ se $i \neq k$. Invece scriviamo $A_{kj} = b_{kj} + \lambda c_{kj}$, $B_{kj} = b_{kj}$ e $C_{kj} = c_{kj}$. Allora

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{k-1\sigma(k-1)} (b_{k\sigma(k)} + \lambda c_{k\sigma(k)}) \dots a_{n\sigma(n)} \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{k-1\sigma(k-1)} b_{k\sigma(k)} \dots a_{n\sigma(n)} + \dots \\
&\quad \lambda \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{k-1\sigma(k-1)} c_{k\sigma(k)} \dots a_{n\sigma(n)} \\
&= \det(A) + \lambda \det(B)
\end{aligned}$$

□

Per mostrare che det è alternante usiamo il seguente fatto

LEMMA 6.30. *Sia τ una trasposizione e sia A_n il sottoinsieme di tutte le permutazioni pari. Allora $D := \{\sigma\tau \mid \sigma \in A_n\}$ è l'insieme delle permutazioni dispari. Inoltre abbiamo che $\sigma_1\tau = \sigma_2\tau$ se e solo se $\sigma_1 = \sigma_2$.*

DIMOSTRAZIONE. La moltiplicazione di una permutazione con τ cambia il segno, e quindi definisce mappe $f : A_n \to D$ a $g : D \to A_n$. Dal fatto che $\sigma^2 = \mathrm{id}$ segue che $f \circ g = \mathrm{id}_D$ e $g \circ f : \mathrm{id}_{A_n}$. Quindi f è biiettiva. Dal fatto che f è suriettiva segue che ogni permutazione è della forma $\sigma\tau$, con $\sigma \in A_k$ e dal fatto che è iniettiva segue che questa descrizione è unica. □

PROPOSIZIONE 6.31. *La funzione* det *è alternante.*

DIMOSTRAZIONE. Sia A una matrice con due righe identiche, diciamo la riga k e la riga m. Sia τ la trasposizione di k ed m.

Sia ora σ una permutazione pari. Siano a, b tali che $\sigma(a) = k$ e $\sigma(b) = m$. Sia $\sigma_1 = \sigma\tau$. Allora $\tau(i) = \sigma(i)$ per $i \neq k, m$, ma $\sigma(k) = \sigma_1(m)$ e $\sigma(m) = \sigma_1(k)$. Dal fatto che la riga k-esima e la riga m-esima sono uguali segue che $a_{k\sigma(k)}a_{m\sigma(m)} = a_{k\sigma(m)}a_{m\sigma(k)} = a_{m\sigma_1(m)}a_{k\sigma_1(k)}$. Quindi

$$\prod_{i=1}^{n} a_{i\sigma(i)} = \prod_{i=1}^{n} a_{i\sigma_1(i)}.$$

Ma $\operatorname{sign}(\sigma_1) = -\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau)$. Quindi il contributo di σ e σ_1 alla somma si cancella. Ripetendo questa operazione troviamo che tutti i termini si cancellano a coppie e

$$\det(A) = 0.$$

□

TEOREMA 6.32. *Il determinante esiste ed è unico.*

DIMOSTRAZIONE. Nei tre lemmi precedenti abbiamo mostrato che esiste una funzione che soddisfa le tre condizione della Definizione 6.1. Nella sezione precedente abbiamo già mostrato che se esiste, allora è unica. □

Adesso possiamo ritrovare la formula della sezione precedente:

Sia A una matrice $n \times n$. Sia $j \in \{1, \ldots, n\}$. Riscriviamo la formula per il determinante in un modo conveniente per i calcoli.

$$\begin{aligned}
\det(A) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{k=1}^{n} a_{k\sigma(k)} \\
&= \sum_{i=1}^{n} \left(\sum_{\sigma \in S_n | \sigma(i)=j} \operatorname{sign}(\sigma) \prod_{k=1}^{n} a_{k\sigma(k)} \right) \\
&= \sum_{i=1}^{n} \left(\sum_{\sigma \in S_n | \sigma(i)=j} \operatorname{sign}(\sigma) a_{ij} \prod_{k=1;k\neq i}^{n} a_{i\sigma(i)} \right) \\
&= \sum_{i=1}^{n} a_{ij} \left(\sum_{\sigma \in S_n | \sigma(i)=j} \operatorname{sign}(\sigma) \prod_{k=1;k\neq i}^{n} a_{k\sigma(k)} \right)
\end{aligned}$$

Sia $A^{i,j}$ la matrice ottenuta da A togliendo la i-esima riga e la k-esima colonna. Un calcolo diretto ci mostra che

$$\det(A^{i,j}) = (-1)^{i+j} \left(\sum_{\sigma \in S_n | \sigma(i)=j} \operatorname{sign}(\sigma) \prod_{k=1;k\neq i}^{n} a_{k\sigma(k)} \right)$$

Combinando le due formule otteniamo i due seguenti teoremi di sviluppo di Laplace.

Teorema 6.33 (Sviluppo lungo la j-esima colonna).

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A^{i,j})$$

Usando $\det(A) = \det(A^T)$ troviamo anche che

Teorema 6.34 (Sviluppo lungo la i-esima riga).

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A^{i,j})$$

5. La matrice dei cofattori

Sia A una matrice $n \times n$. Definiamo la matrice dei cofattori $\operatorname{cof}(A)$ come

$$\operatorname{cof}(A)_{i,j} = (-1)^{i+j} \det(A^{i,j})$$

Proposizione 6.35. *Abbiamo che*

$$A \operatorname{cof}(A)^T = \operatorname{cof}(A)^T A = \det(A) I_n$$

Dimostrazione. Dalla formula del prodotto di matrici troviamo

$$\begin{aligned}
(A \operatorname{cof}(A)^T)_{ij} &= \sum_{k=1}^{n} A_{ik} (\operatorname{cof}(A)^T)_{k,j} \\
&= \sum_{k=1}^{n} (-1)^{k+j} A_{ik} \det A^{j,k}.
\end{aligned}$$

Se $i = j$ allora troviamo la formula del sviluppo di Laplace lungo la riga i, quindi $(A \operatorname{cof}(A)^T)_{ii} = \det(A)$.

Se $i \neq j$ allora troviamo lo sviluppo lungo riga j della matrice $\tilde{A}$, dove $\tilde{A}$ è la matrice A dove la j-esima riga è sostituta con la i-esima riga di A. Quindi $\tilde{A}$ ha due righe identiche e $\det(\tilde{A}) = 0$. In particolare $(A \operatorname{cof}(A))_{ij} = 0$ per $j \neq i$.

Per l'altro prodotto si procede in modo simile. □

COROLLARIO 6.36. *Se A è una matrice $n \times n$ tale che* $\det(A) \neq 0$. *Allora*

$$A^{-1} = \frac{1}{\det(A)} \operatorname{cof}(A)^T.$$

OSSERVAZIONE 6.37. Questo risultato rimostra che se $\det(A) \neq 0$ allora A è invertibile. Questa formula è una generalizzazione della formula per le matrici 2×2 che abbiamo dato all'inizio del capitolo.

Questa formula per la matrice inversa ha delle applicazione teoriche, ma non è molto utile per determinare A^{-1} se $n \geq 3$.

ESEMPIO 6.38. Se

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

allora

$$\operatorname{cof}(A)^T = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

ESEMPIO 6.39. Se

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

allora

$$\operatorname{cof}(A)^T = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

TEOREMA 6.40 (Regola di Cramer). *Sia $A \in M_{n\times n}(K)$ e $b \in K^n$. Se* $\det(A) \neq 0$ *allora l'unica soluzione di $Ax = b$ soddisfa*

$$x_i = \frac{\det A_i}{\det(A)} \text{ per } 1 \leq i \leq n$$

dove A_i è la matrice ottenuta da A sostituendo la i-esima colonna con il vettore b.

DIMOSTRAZIONE. Dal fatto che $\det(A) \neq 0$ segue che A è invertibile. Allora l'unica soluzione di $Ax = b$ è $x = A^{-1}b$. Tramite la formula per il calcolo di A^{-1} troviamo che

$$\begin{aligned} x_i &= (A^{-1}b)_i \\ &= \frac{1}{\det(A)} (\operatorname{cof}(A)^T b)_i \\ &= \frac{1}{\det(A)} \sum_{j=1}^{n} \operatorname{cof}(A)^T_{i,j} b_j \\ &= \frac{1}{\det(A)} \sum_{j=1}^{n} (-1)^{i+j} \det(A^{j,i}) b_j \end{aligned}$$

L'ultima somma è lo sviluppo di Laplace lungo la i-esima colonna di A_i. □

ESEMPIO 6.41. Consideriamo

$$\begin{cases} 2x_1 - x_2 = 4 \\ -3x_1 + 2x_2 = -3 \end{cases}$$

Allora con la regola di Cramer troviamo che

$$x_1 = \frac{\det\begin{pmatrix} 4 & -1 \\ -3 & 2 \end{pmatrix}}{\det\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}} = 5, \quad x_2 = \frac{\det\begin{pmatrix} 2 & 4 \\ -3 & -3 \end{pmatrix}}{\det\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}} = 6$$

6. Determinante di un endomorfismo

Se V, W sono spazi vettoriali della stessa dimensione e B, B' basi per V e W, allora $M_B^{B'}(f)$ è una matrice quadrata. In generale il determinante di tale matrice dipende dalla scelta delle basi B, B'. Nel caso particolare in cui $V = W$ e $B = B'$ questo non vale più.

PROPOSIZIONE 6.42. *Sia $f : V \to V$ un endomorfismo e siano B, B' basi di V. Allora*

$$\det(M_B^B(f)) = \det(M_{B'}^{B'}(f))$$

DIMOSTRAZIONE. Sappiamo che

$$M_{B'}^{B'}(f) = T_B^{B'} M_B^B(f) T_{B'}^B = (T_{B'}^B)^{-1} M_B^B(f) T_{B'}^B$$

Quindi con la formula di Binet troviamo

$$\begin{aligned} \det(M_{B'}^{B'}(f)) &= \det((T_{B'}^B)^{-1} M_B^B(f) T_{B'}^B) \\ &= \det((T_{B'}^B)^{-1}) \det(M_B^B(f)) \det(T_{B'}^B) \\ &= \frac{1}{\det(T_{B'}^B)} \det M_B^B(f) \det(T_{B'}^B) \\ &= \det M_B^B(f) \end{aligned}$$

□

Quindi il determinante è indipendente della scelta della base e possiamo dare la seguente definizione.

DEFINIZIONE 6.43. Sia V uno spazio vettoriale finitamente generato. Sia $f : V \to V$ un endomorfismo. Allora il *determinante* di f è $\det(M_B^B(f))$ per una scelta qualsiasi di una base B di V.

7. Autovalori e autovettori

Sia $f : V \to V$ un endomorfismo di uno spazio vettoriale V di dimensione finita.

DEFINIZIONE 6.44. Un *autovettore* v di f è un vettore $v \in V, v \neq \overrightarrow{0}$, tale che $f(v) = \lambda v$ per un $\lambda \in K$. Un numero λ tale che esiste un $v \in V, v \neq \overrightarrow{0}$, $f(v) = \lambda v$ è chiamato *autovalore*.

Per un numero λ definiamo $E_\lambda = \ker(f - \lambda \operatorname{id}_V)$ l'autospazio di λ.

Sia $A \in M_{n\times n}(K)$ una matrice quadrata. Allora un autovettore di A è un vettore $v \in K^n \setminus \{\overrightarrow{0}\}$ tale che $Av = \lambda v$ per un $\lambda \in K$. In quel caso λ è chiamato un autovalore di A.

In particolare ogni $v \in E_\lambda$ soddisfa $(f - \lambda \operatorname{id}_V)(v) = \overrightarrow{0}$, quindi $f(v) = \lambda v$. Cioè ogni $v \in E_\lambda$ con $v \neq \overrightarrow{0}$ è un autovettore di autovalore λ e tutti gli autovettori con autovalore λ sono in E_λ.

Per trovare tutti i possibili autovalori utilizziamo il *polinomio caratteristico.*

DEFINIZIONE 6.45. Sia A una matrice $n \times n$ allora $p_A(t) = \det(A - tI_n)$ è il *polinomio caratteristico* di A.

Se $f : V \to V$ è un endomorfismo e B è una base di V allora il *polinomio caratteristico di f* è quello di $A = M_B^B(f)$.

Lemma 6.46. *Il polinomio caratteristico di A è della forma*

$$(-1)^n t^n + (-1)^{n-1}\operatorname{tr}(A)t^{n-1} + a_{n-2}t^{n-2} + a_{n-3}t^{n-3} + \cdots + a_1 t + \det(A)$$

con $a_i \in K$ per $i = 1, 2, \ldots, n-2$.

Dimostrazione. Mostriamo questo per induzione su n. Se $n = 2$ abbiamo che $p_A(t) = t^2 - (a_{11} + a_{22})t + a_{11}a_{22} - a_{21}a_{21}$ e abbiamo concluso.

Se $n > 2$ allora lo sviluppo lungo la prima colonna dà

$$\det(A - tI_n) = (a_{11} - t)p_{A^{1,1}}(t) + \sum_{i=2}^{n}(-1)^{i+1}a_{1i}\det((A - tI_n)^{1,i})$$

Si può facilmente dimostrare che il grado in t di $\det((A - tI_n)^{1,i})$ è al massimo il numero di t che compaiono nella sottomatrice. Quel numero è al massimo $n-2$ se $i \neq 1$. Quindi il coefficienti di t^n e t^{n-1} sono quelli di $(a_{11}-t)p_{A^{1,1}}(t)$, per induzione troviamo che quel polinomio è

$$(a_{11} - t)((-1)^{n-1}t^{n-1} + (-1)^{n-2}\operatorname{tr} A^{1,1}t^{n-2} + \text{ termini di grado } \leq n-3)$$

Questo è

$$(-1)^{n-1+1}t^n + (-1)^{n-1}a_{11} + (-1)^{n-2+1}\operatorname{tr}(A^{1,1}) + \text{ termini di grado } \leq n-2$$

Quindi il coefficiente di t^n è $(-1)^n$ ed il coefficiente di t^{n-1} è $\operatorname{tr}(A)$.

Il coefficiente di t^0 è semplicemente $\det(A - 0I_n) = \det(A)$. $\square$

Lemma 6.47. *Due matrici simili hanno lo stesso polinomio caratteristico. In particolare, il polinomio caratteristico di un endomorfismo non dipende dalla scelta della base.*

Dimostrazione. Sia Q una matrice $n\times n$ invertibile. Allora, usando $\det(Q^{-1}) = 1/\det(Q)$ e il teorema di Binet, troviamo

$$\begin{aligned}
p_A(t) = \det(A - tI_n) &= \det(Q)\det(A - tI_n)\det(Q^{-1}) \\
&= \det(Q(A - tI_n)Q^{-1}) \\
&= \det(QAQ^{-1} - Q(tI_n)Q^{-1}) \\
&= \det(QAQ^{-1} - tQQ^{-1}) \\
&= \det(QAQ^{-1} - tI_n) = p_{QAQ^{-1}}(t)
\end{aligned}$$

Quindi A e QAQ^{-1} hanno lo stesso polinomio caratteristico.

Se $f : V \to V$ è un endomorfismo e B e B' sono basi per V allora $M_B^B(f) = QM_{B'}^B(f)Q^{-1}$, con $Q = T_{B'}^B$, quindi le due matrici sono simili e hanno lo stesso polinomio caratteristico. $\square$

Esempio 6.48. Sia A una matrice 2×2

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Allora il polinomio caratteristico di A è

$$\det(A - tI_2) = \det\begin{pmatrix} a-t & b \\ c & d-t \end{pmatrix} = (a-t)(d-t) - bc = t^2 - (a+d)t + ad - bc$$

Lemma 6.49. *Abbiamo che $\dim E_\lambda > 0$ se solo se $p_A(\lambda) = 0$. In particolare gli autovalori di A sono precisamente gli zeri di $p_A(t)$,*

Dimostrazione. Per il teorema del rango abbiamo che $\dim E_\lambda = n - \operatorname{rg}(A - \lambda I_n)$. Quindi $\dim E_\lambda > 0$ se e solo se $\operatorname{rg}(A - \lambda I) < n$. L'ultima affermazione è equivalente a $\det(A - \lambda I_n) = 0$. $\square$

Esempio 6.50. Sia
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$
Allora il polinomio caratteristico è $t^2 - 1$. Quindi gli autovalori di A sono $-1, +1$. Sia adesso
$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$
Allora il polinomio caratteristico è $t^2 + 1$. Quindi A non ha autovalori reali, ma ha autovalori complessi $+i, -i$.

Definizione 6.51. Sia $p(t)$ il polinomio caratteristico di una matrice $A \in M_{n\times n}(K)$. Per un autovalore λ scriviamo $p(t) = (t-\lambda)^{m_a} q(t)$, con $q(\lambda) \neq 0$. Allora chiamiamo $m_a = m_a(\lambda)$ la *molteplicità algebrica* di λ.

Chiamiamo $m_g(\lambda) = \dim E_\lambda$ la *molteplicità geometrica* di λ.

Lemma 6.52. *Sia λ un autovalore. Allora abbiamo che*
$$1 \leq m_g(\lambda) \leq m_a(\lambda)$$

Dimostrazione. Sia $\{v_1, \ldots, v_r\}$ una base di E_λ. Possiamo quindi estenderla ad una base B di V, trovando che $M_B^B(f)$ è
$$\left(\begin{array}{c|c} \lambda I_r & * \\ \hline 0 & A' \end{array}\right).$$
Quindi $\det(A - tI_n) = (\lambda - t)^r \det(A' - \lambda I_{n-r})$, dove la formula si ottiene facilmente iterando lo sviluppo per le prime r colonne. In particolare otteniamo che $m_g(\lambda) = r \leq m_a(\lambda)$. □

Esempio 6.53. Negli esempi precedenti avevamo ottenuto polinomi caratteristici $t^2 - 1 = (t-1)(t+1)$ e $(t^2+1) = (t-i)(t+i)$. In quei casi la molteplicità algebrica è 1 per ogni autovalore. Consideriamo ora
$$A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \text{ e } B = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix},$$
dove $p_A(t) = p_B(t) = (t-3)^2$. Troviamo quindi che la molteplicità algebrica di $\lambda = 3$ è $m_a(3) = 2$.

In questo caso abbiamo due possibili situazioni. Infatti
$$A - 3I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ e } B - 3I_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$
Quindi, per la matrice A, otteniamo che la moltiplicità geometrica $m_g(3) = 1$ mentre, per la matrice B, $m_g(3) = 2$.

Vedremo che sarà fondamentale distinguere questi due casi.

Lemma 6.54. *Sia A la matrice associata ad un endomorfismo f di uno spazio vettoriale V. Siano $\lambda_1, \ldots, \lambda_s$ autovalori distinti di A e siano $B_i := \{v_1^{(i)}, \ldots, v_{r_i}^{(i)}\}$ basi per E_{λ_i}. (Quindi r_i è la molteplicità geometrica di λ_i, $m_g(\lambda_i) = r_i$)*

Allora $B_1 \cup \cdots \cup B_s$ è un insieme linearmente indipendente.

Dimostrazione. Sia
$$\sum_{i=1}^{s} \sum_{j=1}^{r_i} \mu_j^{(i)} v_j^{(i)} = \overrightarrow{0} \tag{3}$$
Se applichiamo f a destra e a sinistra troviamo
$$\sum_{i=1}^{s} \lambda_i \sum_{j=1}^{r_i} \mu_j^{(i)} v_j^{(i)} = \overrightarrow{0}$$

Prendiamo ora λ_s volte la prima equazione e sottraiamola alla seconda, trovando che

$$\sum_{i=1}^{s-1}(\lambda_i - \lambda_s)\sum_{j=1}^{r_i}\mu_j^{(i)}v_j^{(i)} = \overrightarrow{0}$$

Procediamo ora per induzione sul valore s. Il caso $s = 1$ è banale. Facciamo anche il caso $s = 2$. Otteniamo che

$$(\lambda_1 - \lambda_2)(\sum_j \mu_j^{(1)}v_j^{(1)}) = \overrightarrow{0}$$

Dal fatto che i vettori $v_j^{(1)}$ sono una base di E_{λ_1} segue che sono linearmente indipendenti. Inoltre gli autovalori λ_1 e λ_2 sono distinti, quindi $(\lambda_1 - \lambda_2) \neq 0$. Dall'indipendenza lineare otteniamo quindi che tutti i coefficienti $\mu_j^{(1)}$ devono essere zero. Scambiando i ruoli di λ_1 e λ_2 troviamo che anche $\mu_j^{(2)} = 0$ per tutti gli indici j.

Similmente, per $s > 2$, l'ipotesi induttiva ci dice che per $i < s$ e $1 \leq j \leq r_i$ allora

$$(\lambda_i - \lambda_s)\mu_j^{(i)} = 0$$

per $i < s$. Quindi $\mu_j^{(i)} = 0$ per $i < s$. Nell'equazione (3) rimane solamente

$$\sum \mu_j^{(s)}v_j^{(s)} = \overrightarrow{0}.$$

Dal fatto che $v_1^{(s)}, \ldots, v_{i_s}^{(s)}$ è una base di E_{λ_s} segue che $\mu_j^{(s)} = 0$ per ogni $j = 1, \ldots, i_s$. □

Definizione 6.55. Un endomorfismo $f : V \to V$ è *diagonalizzabile* se esiste una base B per V tale che $M_B^B(f)$ è una matrice diagonale. Una matrice A è diagonalizzabile se esiste una matrice Q tale che $Q^{-1}AQ$ è una matrice diagonale.

Proposizione 6.56. *Sia V uno spazio vettoriale finitamente generato di dimensione n. Sia $f : V \to V$ un endomorfismo associato ad una matrice A. Allora f è diagonalizzabile se e solo se $p_A(t)$ è un prodotto di fattori lineari e per ogni autovalore abbiamo che la molteplicità algebra e geometrica coincidono.*

Dimostrazione. Siano E_{λ_i} gli autospazi e siano B_i basi per gli autospazi. Abbiamo che f è diagonalizzabile se e solo se $B = B_1 \cup \cdots \cup B_s$ sono un insieme di generatori di V. Gli elementi di B sono linearmente indipendenti. Quindi generano un sottospazio di dimensione $\sum m_g(\lambda_i) \leq \sum m_a(\lambda_i) \leq n = \deg(p_A(t))$. Quindi f è diagonalizzabile se e solo se

(1) $\sum m_a(\lambda_i) = n$
(2) Per ogni i abbiamo $m_g(\lambda_i) = m_a(\lambda_i)$.

La prima condizione equivale al fatto che p sia un prodotto di fattori lineari, mentre la seconda condizione implica che ogni autovalore ha molteplicità algebrica uguale a quella geometrica. □

Corollario 6.57. *Se $p(t)$ è il prodotto di n fattori lineari distinti allora A è diagonalizzabile.*

Dimostrazione. Abbiamo che $p(t)$ è prodotto di fattori lineari. Inoltre abbiamo per ogni autovalore che

$$1 \leq m_a(\lambda_i) \leq m_g(\lambda_i) = 1$$

Quindi $m_a(\lambda_i) = m_g(\lambda_i) = 1$. □

Osservazione 6.58. Un autovalore λ con $m_a(\lambda) = 1$ viene solitamente chiamato *autovalore semplice*.

Esempio 6.59. Consideriamo le seguenti matrici 2×2

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$$

Per la prima matrice abbiamo che $p(t) = t^2 + 1$. Quindi p non ha zeri in $\mathbf{R}$, quindi non è diagonalizzabile su $\mathbf{R}$. Invece possiamo scrivere $p(t) = (t - i)(t + i)$ in $\mathbf{C}$ e quindi è diagonalizzabile su $\mathbf{C}$. La seconda matrice ha $p(t) = (t^2-1) = (t-1)(t+1)$. Quindi il polinomio caratteristico è un prodotto di fattori lineari distinti. La terza matrice ha $p(t) = (t-3)^2$. Quindi ha un unico autovalore con molteplicità algebrica due. La molteplicità geometrica è $\dim \mathrm{Sol}(A - 3I_2, \overrightarrow{0})$ che è generato da $(1, 0)$. Quindi la molteplicità geometrica è 1 e A non è diagonalizzabile.

Esempio 6.60. Sia

$$A = \begin{pmatrix} 3 & 4 & 4 \\ -3 & -6 & -8 \\ 3 & 7 & 9 \end{pmatrix}$$

Allora

$$\begin{aligned} p_A(t) &= \det \begin{pmatrix} 3-t & 4 & 4 \\ -3 & -6-t & -8 \\ 3 & 7 & 9-t \end{pmatrix} \\ &= \det \begin{pmatrix} 3-t & 4 & 4 \\ -3 & -6-t & -8 \\ 0 & 1-t & 1-t \end{pmatrix} \\ &= \det \begin{pmatrix} 3-t & 4 & 0 \\ -3 & -6-t & -2+t \\ 0 & 1-t & 0 \end{pmatrix} \\ &= -(-2+t)((3-t)(1-t)+0) \end{aligned}$$

Quindi gli autovalori sono $\{1, 2, 3\}$.

Per trovare gli autovettori dobbiamo determinare $\mathrm{Sol}(A - \lambda I_3, \overrightarrow{0})$ per ogni autovalore λ. Sostituendo $\lambda = 1, 2, 3$ troviamo autovettori $v_1 = (2, -2, 1)$ per $\lambda = 1$, $v_2 = (0, -1, 1)$ per $\lambda = 2$ e $v_3 = (1, -3, 3)$ per $\lambda = 3$.

Quindi per

$$Q = \begin{pmatrix} 2 & 1 & 0 \\ -2 & -3 & -1 \\ 1 & 3 & 1 \end{pmatrix}$$

troviamo

$$Q^{-1}AQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

8. Forma di Jordan

Abbiamo visto due ostruzioni alla diagonalizzabilità di una matrice. La prima è che il polinomio caratteristico non sia un prodotto di fattori lineari. Questo problema si può risolvere passando ad un campo più grande (un campo algebricamente chiuso). L'altro problema è che la molteplicità geometrica e algebrica possono essere diverse.

Queste molteplicità non cambiano passando ad un campo più grande. La soluzione è quella di utilizzare la forma di Jordan.

Esempio 6.61. Sia

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}.$$

Allora la molteplicità algebrica di 3 è 2, ma $\mathrm{Sol}((A-3I_2), \overrightarrow{0})$ è uno-dimensionale. Quindi A non è diagonalizzabile. Invece $(A-3I_2)^2$ è la matrice 0 e quindi $\mathrm{Sol}((A-\lambda)^2, \overrightarrow{0})$ ha dimensione uguale alla molteplicità algebrica.

Definizione 6.62. Un *blocco di Jordan* $J_{n,\lambda}$ di lunghezza n con autovalore λ è una matrice $n \times n$

$$J_{n,\lambda} = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

Osservazione 6.63. Il polinomio caratteristico di $J_{n,\lambda}$ è $(\lambda-t)^n$. Quindi $J_{n,\lambda}$ ha un unico autovalore λ con molteplicità algebrica n. La matrice $J_{n,\lambda}-\lambda I_n$ è già in forma a scala ed ha rango $n-1$. Quindi la molteplicità geometrica è 1.

Vediamo che la matrice $(J_{n,\lambda}-\lambda I_n)^k \neq 0$ per $k < n$ e $(J_{n,\lambda}-\lambda I_n)^k = 0$ per $k \geq n$.

Definizione 6.64. Una matrice $J \in M_{n\times n}(K)$ è in forma di Jordan se

(1) $J_{i,i+1} \in \{0, 1\}$ per $i = 1, \dots, n-1$.
(2) Se $J_{i,i+1} = 1$ allora $J_{i,i} = J_{i+1,i+1}$.
(3) $J_{i,k} = 0$ per $k \neq i, i+1$.

Osservazione 6.65. Una matrice di Jordan ha entrate nonzero solamente sulla diagonale e sulla prima “sopradiagonale”. In particolare possiamo suddividere la nostra matrice in blocchi, costituiti da matrici di Jordan sulla diagonale e matrici composte di soli zeri altrove.

Osservazione 6.66. Gli autovalori dei blocchi di Jordan non devono essere distinti. In particolare una matrice diagonale è una matrice di Jordan costituita da n blocchi 1×1.

Possiamo ora mostrare che per ogni matrice $A \in M_{n\times n}(K)$ tale che $p_A(t)$ è un prodotto di fattori lineari allora esiste una matrice Q per cui $Q^{-1}AQ$ è in forma di Jordan.

Definizione 6.67. Un vettore $v \neq \overrightarrow{0}$ è chiamato *autovettore generalizzato* di *autovalore* λ se per qualche intero i abbiamo $(A-\lambda I_n)^i v = \overrightarrow{0}$.

Lemma 6.68. *Sia $A \in M_{n\times n}(K)$. Sia λ uno zero di $p_A(t)$ e sia $m = m_a(\lambda)$ la sua molteplicità algebrica. Allora esistono due sottospazi V_1, V_2 di K^n tali che*

(1) $\dim V_1 = m$, $V_1 \oplus V_2 = V$,
(2) la moltiplicazione con A manda V_1 in V_1 e V_2 in V_2.
(3) sia A_i una matrice dell'endomorfismo A ristretto a V_i. Allora $p_{A_1}(t) = (\lambda-t)^m$ e $p_{A_2}(\lambda) \neq 0$.

Dimostrazione. Consideriamo i sottospazi $W_i := \mathrm{Sol}((A-\lambda I_n)^i, \overrightarrow{0}) \subset K^n$. Dal fatto che $p_A(\lambda) = 0$ segue che $\dim W_1 \geq 1$. Per ogni i abbiamo $W_i \subset W_{i+1}$. Inoltre, per ogni i tale che $W_i \neq W_{i+1}$ abbiamo che $n \geq \dim W_{i+1} \geq \dim W_i + 1$. Quindi esiste un j tale che $W_{j-1} \neq W_j$ ma $W_j = W_i$ per ogni $i > j$. Prendiamo $V_1 = W_j$ e $V_2 = \mathrm{im}((A-\lambda I_n)^j)$.

Dalla formula della dimensioni (teorema del rango applicato a $(A-\lambda I_n)^j$) segue che $\dim V_1 + \dim V_2 = n$. Mostriamo ora che

$$K^n = V_1 \oplus V_2$$

Se un vettore $v \in V_1 \cap V_2$ allora $(A - \lambda I_n)^j v = \overrightarrow{0}$ $(v \in V_1)$ ed esiste $w \in K^n$ con $v = (A - \lambda I_n)^j w$ $(v \in V_2)$. Quindi $w \in W_{2j}$. Ma $W_j = W_{2j}$, e $w \in W_i$. Allora $v = (A - \lambda I_n)^j w = 0$. Quindi $V_1 \cap V_2 = \{\overrightarrow{0}\}$. Dal fatto che $\dim V_1 + \dim V_2 = n$ adesso segue $K^n = V_1 \oplus V_2$ e abbiamo mostrato il primo punto.

Per il secondo punto notiamo che se $v \in W_i$ allora $w := (A - \lambda I_n)v \in W_{i-1}$. Quindi $Av = w + \lambda v \in W_i$ e A manda V_1 in se stesso.

Sia adesso $v \in V_2$. Quindi esiste un vettore w tale che $v = (A-\lambda I_n)^i w$. Notiamo inoltre che vale la seguente proprietà, $A(A - \lambda I_n)) = (A^2 - \lambda A) = (A - \lambda I_n)A$. Quindi $A(A - \lambda I_n)^i = (A - \lambda I_n)^i A$ e

$$Av = A(A - \lambda I_n)^i w = (A - \lambda I_n)^i (Aw).$$

Ottenendo quindi Av è nell'immagine di $(A - \lambda I_n)^i$. Quindi A mappa V_2 su se stesso e abbiamo mostrato il secondo punto.

Per il terzo punto studiamo i polinomi caratteristici. Siano A_1 ed A_2 le matrici ottenute riscrivendo la matrice A rispetto (rispettivamente) a V_1 e V_2. Se $p_{A_2}(\lambda) = 0$, allora esiste un $v \in V_2$, con $v \neq \overrightarrow{0}$ tale che $Av = \lambda v$ e $v \in V_1$. Quindi $v \in V_1 \cap V_2$, che contraddice $V_1 \cap V_2 = \{\overrightarrow{0}\}$, quindi $p_{A_2}(\lambda) \neq 0$.

Costruiamo ora induttivamente una base per V_1. Sia $\{v_1, \dots, v_{m_1}\}$ una base di $W_1 = \mathrm{Sol}((A - \lambda I), \overrightarrow{0})$. Estendiamola poi ad una base $\{v_1, \dots, v_{m_2}\}$ di $W_2 = \mathrm{Sol}((A - \lambda I)^2, \overrightarrow{0})$ iterando il procedimento fino a trovare una base

$$B = \{v_1, \dots, v_{m_j}\}$$

di $W_j = V_1$.

Adesso $(A-\lambda I_n)$ manda W_i in W_{i-1}, quindi se $v_k \in W_i \backslash W_{i-1}$, allora $k > m_j - 1$ e $(A - \lambda)v_k = w$, con $w \in W_{i-1}$. La nostra scelta della base ci permette di scrivere $w = \sum_{j=1}^{m_{j-1}} c_j v_j$ con $c_j \in K$ e otteniamo quindi che $Av_k = \lambda v_k + w = \lambda v_k + \sum_{j=1}^{m_{j-1}} c_j v_j$.

Dal fatto che $k > m_{j-1}$ segue adesso che la sotto-matrice A_1 della matrice A rispetto alla base B ha soltanto λ sulla diagonale ed è una matrice triangolare superiore. In particolare $\det(A_1 - tI_n) = (\lambda - t)^m$. □

Proposizione 6.69. *Se $p_A(t)$ è un prodotto di fattori lineari, allora A ammette una forma di Jordan. In particolare se $K = \mathbf{C}$ allora A ha una forma di Jordan.*

Dimostrazione. Lo dimostreremo per induzione sulla dimensione, n. Il caso di $n = 1$ è banale, perché ogni matrice 1×1 è una matrice diagonale, in particolare è in forma di Jordan.

Se $p_A(t)$ ha due zeri distinti allora il lemma precedente implica che possiamo trovare matrici A_1, A_2, Q tali che

$$Q^{-1}AQ = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right)$$

Usando l'ipotesi induttiva troviamo delle matrici R_1, R_2 tali che $R_1^{-1}A_1R_1$ è in forma di Jordan J_1, e $R_2^{-1}A_2R_2$ è in forma di Jordan J_2. Se

$$R = \left(\begin{array}{c|c} R_1 & 0 \\ \hline 0 & R_2 \end{array}\right)$$

allora un calcolo diretto mostra che

$$R^{-1}Q^{-1}AQR = \left(\begin{array}{c|c} J_1 & 0 \\ \hline 0 & J_2 \end{array}\right).$$

Quindi abbiamo trovato che A ammette una forma di Jordan.

Rimane il caso che $P_A(t) = (\lambda - t)^n$. Definiamo adesso $W_i = \ker(A - \lambda I_n)^i$. Sappiamo che per $i = n$ abbiamo $W_i = K^n$.

Sia j il valore più grande tale che $W_{j-1} \neq K^n$. Allora possiamo scegliere $v_j^{(1)}, \dots, v_j^{(t_1)}$ linearmente indipendenti tali che

$$\operatorname{span}(v_j^{(1)}, \dots, v_j^{(t)}) \oplus W_{j-1} = K^n.$$

Adesso, per $i = 1, \dots, j-1, k = 1, \dots t_1$ prendiamo $v_i^{(k)} = (A - \lambda I_n)v_{i+1}^{(k)}$. Scegliamo inoltre $v_{j-1}^{t_1+1}, \dots, v_{j-1}^{(t_2)}$ tali che $\operatorname{span}(v_j^{(1)}, \dots, v_j^{(t)}, v_{j-1}^{(1)}, \dots, v_{j-1}^{t_2}) \oplus W_{j-2} = K^n$.

Ripetendo questo procedimento, alla fine troviamo che la matrice rispetto a $\{v_1^{1)}, \dots, v_1^{(j)}, \dots\}$ è in forma di Jordan se tutti i vettori sono linearmente indipendenti. □

Il modo per trovare la forma di Jordan è basato su questa dimostrazione. Si inizia determinando $V_i = \ker(A - \lambda I_n)^i$. Si prende una base per W_i, e si applica $A - \lambda I_n$ a questa base $i-1$ volte, e si stabilisce la parte di V_i corrispondente ad una potenza più piccola. La parte più complicata sarà quella di capite a quali autovettori sono associati gli autovettori generalizzati determinati al passo successivo.

Esempio 6.70. Sia

$$A = \begin{pmatrix} 1 & 0 & -2 & 0 & 0 \\ -2 & 3 & -2 & 0 & 0 \\ 2 & 0 & 5 & 0 & 0 \\ 0 & 2 & 2 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

Allora il polinomio caratteristico è

$$(t-3)^5$$

Abbiamo inoltre che $(A-3)^2$ è

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 \end{pmatrix}$$

e $(A-3)^3 = 0$.

Abbiamo quindi che $\dim W_3 = 5, \dim W_2 = 4$, $\dim W_1 = 2$, e possiamo concludere che dovremo ottenere un blocco di lunghezza 3 ed uno di lunghezza 2.

Il vettore $(0,1,0,0,0) \in V_3 \setminus V_2$. Quindi possiamo usarlo per trovare un blocco di Jordan. Allora $(A - 3I_5)(0,1,0,0,0) = (0,0,0,2,0)$ e $(A-3)^2(0,1,0,0,0) = (0,0,0,-2,2)$.

Come secondo generatore di un blocco dobbiamo scegliere un vettore che sia in V_2, ma non in $V_1 \oplus \operatorname{span}(0,0,0,2,0)$. Possiamo per esempio prendere $(1,0,0,0,0)$, trovando $(A - 3I_3)(1,0,0,0,0) = (-2,2,2,0,0)$.

Rispetto alla base $\{v_1, 2e_4, e_2, v_2, e_1\}$, $v_1 = (0,0,0,-2,2); v_2 = (-2,-2,2,0,0)$ troviamo che

$$\left(\begin{array}{ccc|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{array}\right)$$

Osservazione 6.71. Siano A una matrice $n \times n$ e λ un suo autovalore. Denotiamo, come sopra, $W_i = \operatorname{Sol}((A - \lambda I_n)^i, \overrightarrow{0})$. Allora W_1 è formato unicamente da autovettori, ed è quindi l'autospazio di λ. Inoltre $\dim W_1$ è il numero di blocchi di Jordan. Lo spazio W_2 contiene anche autovettori generalizzati, associati a blocchi di lunghezza almeno 2. Il numero di tali blocchi sarà $\dim W_2 - \dim W_1$. Similmente troviamo che $\dim W_i - \dim W_{i-1}$ è il numero di blocchi di lunghezza almeno i. Quindi

$$\dim W_i - \dim W_{i-1} - (\dim W_{i+1} - \dim W_i) = 2\dim W_i - \dim W_{i+1} - \dim W_{i-1}$$

è il numero di blocchi di lunghezza i.