8. Forme canoniche razionali

Sia $f : V \to V$ un endomorfismo di uno spazio vettoriale di dimensione finita su un campo K. Come si è osservato, non è in generale detto che si possa trovare una base di V rispetto alla quale la matrice di f sia in forma canonica di Jordan. Si può però trovare una base di V rispetto alla quale la matrice di f è di una forma particolarmente semplice, detta *forma canonica razionale*, o *forma canonica di Frobenius*. Sia $P(X)$ il polinomio caratteristico di f, e sia $P = \prod_{i=1}^m P_i^{k_i}$ la sua decomposizione in fattori irriducibili, dove i P_i sono a due a due primi fra loro. Indichiamo con W_i il nucleo di $P_i^{k_i}(f)$. Sappiamo che i W_i sono sottospazi invarianti e che $V = \bigoplus_i W_i$. Rispetto a una base di V costruita mettendo insieme una base di W_1, una di W_2, e così via, la matrice di f è dunque diagonale a blocchi della forma

$$\begin{pmatrix} A_1 & 0 & \dots & & \\ 0 & A_2 & 0 & \dots & \\ \dots & 0 & A_3 & 0 & \dots \\ & & \dots & & \\ & & \dots & 0 & A_m \end{pmatrix},$$

dove A_i è la matrice della applicazione lineare f_i da W_i in sè indotta da f. Poiché il polinomio caratteristico di f_i è $P_i^{k_i}$, questo ci permette di limitarci a studiare il caso in cui $P(X)$ è una potenza di un polinomio irriducibile $Q(X)$ di grado h. Sia k l'intero positivo tale che $Q^k(f) = 0$ ma $Q^{k-1}(f) \neq 0$, e poniamo $V_i = \ker(Q^i(f))$ per ogni intero i tale che $0 \leq i \leq k$. È chiaro che i V_i sono invarianti e che

$$\{0\} = V_0 \subset \cdots \subset V_{i-1} \subset V_i \subset \cdots \subset V_k = V\,.$$

Osserviamo che, se $i > 1$ e $v_1, \dots, v_s$ sono elementi di V_i che sono indipendenti modulo V_{i-1}, allora $Q(f)(v_1), \dots, Q(f)(v_s)$ sono indipendenti modulo V_{i-2}. Supponiamo infatti che $\sum a_i Q(f)(v_i) \in V_{i-2}$. Allora $Q(f)^{i-1}(\sum a_i v_i) = 0$, e quindi $\sum a_i v_i$ appartiene a V_{i-1}. Dato che i v_i sono indipendenti modulo V_{i-1} se ne deduce che tutti gli a_i sono nulli.

Costruiamo ora una base V costruendo una opportuna base di V_i modulo V_{i-1} per ogni i. Iniziamo con $i = k$. Sia $v_{k,1}$ un elemento di V_k non appartenente a V_{k-1}, e notiamo che $v_{k,1}, f(v_{k,1}), f^2(v_{k,1}), \dots, f^{h-1}(v_{k,1})$ sono indipendenti modulo V_{k-1}. Per dimostrarlo ci baseremo sul seguente semplice risultato.

Lemma (8.1). *Sia W un sottospazio invariante di V, e sia v un elemento di V. Allora*

$$\mathcal{I} = \{Q \in K[X] : Q(f)(v) \in W\}$$

è un ideale in $K[X]$.

La dimostrazione è immediata. Se P e Q appartengono a $\mathcal{I}$ allora

$$(P + Q)(f)(v) = P(f)(v) + Q(f)(v) \in W\,.$$

D'altra parte, se $R \in K[X]$, allora

$$(RP)(f) = R(f)(P(f)(v)) \in W$$

poiché $P(f)(v) \in W$ e W è invariante. Questo dimostra il lemma.

Applichiamo il lemma con $W = V_{k-1}$ e $v = v_{k,1}$. Siano gli a_i scalari tali che

$$\sum_{i=0}^{h-1} a_i f^i(v_{k,1}) \in V_{k-1}\,.$$

Se poniamo $R(X) = \sum_{i=0}^{h-1} a_i X^i$, questo equivale a dire che $R(f)(v_{k,1}) \in V_{k-1}$, cioè che $R(X) \in \mathcal{I}$. Dato che $Q^{k-1}(f)(Q(f)(v_{k,1})) = 0$, anche $Q(X)$ appartiene a $\mathcal{I}$. Per il lemma che si è appena dimostrato, $\mathcal{I}$ è un ideale in $K[X]$, e quindi ha un generatore monico che divide sia Q che R. Dato che Q è irriducibile e il grado di R è strettamente minore del grado di Q, se $R \neq 0$ questo generatore deve essere 1. Dato che $V_k \neq V_{k-1}$, questo è assurdo. L'unica via di uscita è che sia $R = 0$, cioè che tutti gli a_i siano nulli. Questo mostra che $v_{k,1}, f(v_{k,1}), f^2(v_{k,1}), \ldots, f^{h-1}(v_{k,1})$ sono indipendenti modulo V_{k-1}. Notiamo inoltre che il sottospazio W' di V generato da V_{k-1} e da $v_{k,1}, f(v_{k,1}), f^2(v_{k,1}), \ldots, f^{h-1}(v_{k,1})$ è invariante. In effetti ogni elemento v di W' si scrive come somma di un elemento di V_{k-1} e di una combinazione lineare di $v_{k,1}, f(v_{k,1}), f^2(v_{k,1}), \ldots, f^{h-1}(v_{k,1})$. Dunque $f(v)$ è somma di un elemento di V_{k-1} e di una combinazione lineare di $v_{k,1}, f(v_{k,1}), f^2(v_{k,1}), \ldots, f^h(v_{k,1})$. Ma d'altra parte il fatto che $Q(f)(v_{k,1})$ appartenga a V_{k-1} ci dice che $f^h(v_{k,1})$ è somma di un elemento di V_{k-1} e di una combinazione lineare di $v_{k,1}, f(v_{k,1}), f^2(v_{k,1}), \ldots, f^{h-1}(v_{k,1})$. Quindi $f(v) \in W'$. Se $v_{k,1}, f(v_{k,1}), f^2(v_{k,1}), \ldots, f^{h-1}(v_{k,1})$ non è una base di V_k modulo V_{k-1}, scegliamo un vettore $v_{k,2}$ appartenente a V_k ma non a W'. Ragionando esattamente come sopra si mostra che $v_{k,2}, f(v_{k,2}), f^2(v_{k,2}), \ldots, f^{h-1}(v_{k,2})$ sono indipendenti modulo W' e che il sottospazio generato da W' e da $v_{k,2}, f(v_{k,2}), f^2(v_{k,2}), \ldots, f^{h-1}(v_{k,2})$ è invariante. Iterando questo procedimento si giunge, in un numero finito di passi, a costruire una base di V_k modulo V_{k-1} della forma

$$v_{k,1}, f(v_{k,1}), \ldots, f^{h-1}(v_{k,1}), \ldots, v_{k,n_k}, f(v_{k,n_k}), \ldots, f^{h-1}(v_{k,n_k})\,.$$

Come si è osservato,

$$Q(f)(v_{k,1}), \ldots, Q(f)(f^{h-1}(v_{k,1})), \ldots, Q(f)(v_{k,n_k}), \ldots, Q(f)(f^{h-1}(v_{k,n_k}))$$

sono indipendenti modulo V_{k-2}. Ragionando come sopra si possono trovare elementi $v_{k-1,1}, \ldots, v_{k-1,n_{k-1}}$ di V_{k-1} tali che

$$\begin{aligned}&Q(f)(v_{k,1}), \ldots, Q(f)(f^{h-1}(v_{k,1})), \ldots, Q(f)(v_{k,n_k}), \ldots, Q(f)(f^{h-1}(v_{k,n_k}))\\ &\qquad v_{k-1,1}, f(v_{k-1,1}), \ldots, f^{h-1}(v_{k-1,1}), \ldots, v_{k-1,n_{k-1}}, f(v_{k-1,n_{k-1}}), \ldots, f^{h-1}(v_{k-1,n_{k-1}})\end{aligned}$$

siano una base di V_{k-1} modulo V_{k-2}. Iterando questa costruzione si giunge a trovare una base di V della forma

$$
\begin{array}{l}
v_{k,1},\dots,f^{h-1}(v_{k,1}),Q(f)(v_{k,1}),\dots,Q(f)(f^{h-1}(v_{k,1})),\dots,Q^{k-1}(v_{k,1}),\dots,Q^{k-1}(f^{h-1}(v_{k,1})),\\
\qquad\qquad\dots\qquad\qquad\dots\qquad\qquad\dots\\
v_{k,n_k},\dots,f^{h-1}(v_{k,n_k}),\dots,Q^{k-1}(v_{k,n_k}),\dots,Q^{k-1}(f^{h-1}(v_{k,n_k})),\\
v_{k-1,1},\dots,f^{h-1}(v_{k-1,1}),\dots,Q^{k-2}(v_{k-1,1}),\dots,Q^{k-2}(f^{h-1}(v_{k-1,1})),\\
\qquad\qquad\dots\qquad\qquad\dots\qquad\qquad\dots\\
v_{k-1,n_{k-1}},\dots,f^{h-1}(v_{k-1,n_{k-1}}),\dots,Q^{k-2}(v_{k-1,n_{k-1}}),\dots,Q^{k-2}(f^{h-1}(v_{k-1,n_{k-1}})),\\
\qquad\qquad\dots\qquad\qquad\dots\\
\qquad\qquad\dots\\
v_{1,1},\dots,f^{h-1}(v_{1,1}),\\
\qquad\dots\\
v_{1,n_1},\dots,f^{h-1}(v_{1,n_1}).
\end{array}
$$

Dato che, per ogni scelta di i,j,s,t, il vettore $Q^i(f)(f^j(v_{s,t}))$ è combinazione lineare di $v_{s,t},f(v_{s,t}),\dots,f^{hi+j}(v_{s,t})$, un altro sistema di generatori per V è

$$
\begin{array}{l}
v_{k,1},f(v_{k,1})\dots,f^{hk-1}(v_{k,1}),\\
\qquad\dots\\
v_{k,n_k},f(v_{k,n_k})\dots,f^{hk-1}(v_{k,n_k}),\\
v_{k-1,1},f(v_{k-1,1})\dots,f^{h(k-1)-1}(v_{k-1,1}),\\
\qquad\dots\\
v_{k-1,n_{k-1}},f(v_{k-1,n_{k-1}})\dots,f^{h(k-1)-1}(v_{k-1,n_{k-1}}),\\
\qquad\dots\\
\qquad\dots\\
v_{1,1},\dots,f^{h-1}(v_{1,1}),\\
\qquad\dots\\
v_{1,n_1},\dots,f^{h-1}(v_{1,n_1}).
\end{array}
$$

Dato che questo sistema di generatori consta di $\dim(V)$ elementi, è anch'esso una base di V. È questa la base cercata. Per vedere come è fatta la matrice di f rispetto a questa base notiamo che, per ogni scelta di s e t, si ha che $Q^s(f)(v_{s,t})=0$, e quindi $f^{hs}(v_{s,t})$ è combinazione lineare di $v_{s,t},f(v_{s,t}),\dots,f^{hs-1}(v_{s,t})$. Più esattamente, se $Q^s(X)=\sum a_iX^i$, allora

$$
f(f^{hs-1}(v_{s,t}))=f^{hs}(v_{s,t})=-\sum_{i=0}^{hs-1}a_if^i(v_{s,t})\,.
$$

La matrice di f rispetto alla nostra base è dunque una matrice diagonale a blocchi

$$
\begin{pmatrix}
C_1 & 0 & \dots & & \\
0 & C_2 & 0 & \dots & \\
\dots & 0 & C_3 & 0 & \dots\\
 & & \dots & & \\
 & & \dots & 0 & C_n
\end{pmatrix}
$$

dove $n = n_1 + \cdots + n_k$ e, per ogni j tale che $n_k + \cdots + n_i < j \leq n_k + \cdots + n_{i-1}$,

$$C_j = \begin{pmatrix} 0 & 0 & & \ldots & 0 & -c_0 \\ 1 & 0 & & \ldots & 0 & -c_1 \\ 0 & 1 & 0 & \ldots & 0 & -c_2 \\ & & \ldots & & \ldots & \\ & & \ldots & & \ldots & \\ & & \ldots & 1 & 0 & -c_{hi-2} \\ & & & \ldots & 1 & -c_{hi-1} \end{pmatrix},$$

dove i è il più grande intero tale che $j \leq n_i + \cdots + n_k$ e

$$Q^i(X) = X^{hi} + c_{hi-1}X^{hi-1} + c_{hi-2}X^{hi-2} + \cdots + c_1 X + c_0 \,.$$

Abbiamo dunque dimostrato in generale il seguente risultato.

Teorema (8.2) (Forma normale razionale o di Frobenius). *Sia V uno spazio vettoriale di dimensione finita sul campo K, e sia f un endomorfismo di V. Vi è una base di V rispetto alla quale la matrice di f è una matrice diagonale a blocchi*

$$\begin{pmatrix} B_1 & 0 & \ldots & & \\ 0 & B_2 & 0 & \ldots & \\ \ldots & 0 & B_3 & 0 & \ldots \\ & & \ldots & & \\ & & \ldots & 0 & B_n \end{pmatrix},$$

dove

$$B_i = \begin{pmatrix} 0 & 0 & & \ldots & 0 & -b_0 \\ 1 & 0 & & \ldots & 0 & -b_1 \\ 0 & 1 & 0 & \ldots & 0 & -b_2 \\ & & \ldots & & \ldots & \\ & & \ldots & & \ldots & \\ & & \ldots & 1 & 0 & -b_{s-2} \\ & & & \ldots & 1 & -b_{s-1} \end{pmatrix}$$

e

$$X^s + b_{s-1}X^{s-1} + b_{s-2}X^{s-2} + \cdots + b_1 X + b_0$$

è una potenza di un polinomio irriducibile in $K[X]$.