LA FORMA CANONICA DI JORDAN

R. BENEDETTI

Inquadriamo la questione e richiamiamo alcuni fatti che saranno considerati noti al lettore.

Sia V uno spazio vettoriale su $\mathbb{K}$, $\dim V = n$. Sullo spazio End (V) degli endomorfismi di V, consideriamo la relazione di equivalenza definita da:

$f, g \in$ End (V) *si dicono* coniugati *e scriveremo* $f \sim g$, *se è solo se esiste* $h \in GL(V)$ *tale che* $g = h^{-1} \circ f \circ h$.

Nel caso in cui $V = \mathbb{K}^n$, End $(\mathbb{K}^n)$ è canonicamente identificato con lo spazio delle matrici $n \times n$ $M(n, \mathbb{K})$, $GL(V)$ con $GL(n, \mathbb{K})$ e si usa dire che due matrici coniugate sono *simili.* Vogliamo studiare il quoziente End $(V)/\sim$ ed in particolare $M(n, \mathbb{K})/\sim$. Per ogni endomorfismo f di V ed ogni base $\mathcal{B}$ di V, indicheremo con $M_{\mathcal{B}}(f)$ la matrice che esprime f in termini del sistema di coordinate associato a $\mathcal{B}$.

- Sono fatti tra loro equivalenti:
  (1) $f \sim g$
  (2) Per qualsiasi base $\mathcal{B}$, le matrici $M_{\mathcal{B}}(f)$ e $M_{\mathcal{B}}(g)$ sono simili.
  (3) Esistono basi $\mathcal{B}$ e $\mathcal{B}'$ di V tali che $M_{\mathcal{B}}(f) = M_{\mathcal{B}'}(g)$.
- **Alcuni invarianti.** Se $g \sim f$ allora:
  (1) f e g hanno lo stesso polinomio caratteristico, $p_f(t) = p_g(t) \in \mathbb{K}[t]$.
  (2) f e g hanno lo stesso spettro di autovalori, Spettro(f) = Spettro(g)
  (3) Per ogni autovalore λ che hanno necessariamente in comune, le molteplicità algebriche in quanto radici del polinomio caratteristico coincidono, $m_f(\lambda) = m_g(\lambda)$, e i rispettivi autospazi hanno la stessa dimensione, $\dim V_\lambda(f) = \dim V_\lambda(g)$.
- **Endomorfismi triangolabili.** Un endomorfismo f è triangolabile se esiste una base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}(f)$ è triangolare superiore. Sono fatti equivalenti
  (1) f è triangolabile.
  (2) Il polinomio caratteristico di f è completamente fattorizzabile in $\mathbb{K}[t]$:

$$p_f(t) = \pm \prod_{j=1}^{k} (t - \lambda_j)^{m_j}$$

$$\lambda_j \neq \lambda_i, \text{ se } i \neq j; \ \sum_j m_j = n \ .$$

  (3) Esiste una base $\mathcal{B} = \{v_1, \dots, v_n\}$, con bandiera Span $\{v_1\} \subset$ Span $\{v_1, v_2\} \subset$ Span $\{v_1, v_2, v_3\} \subset \cdots \subset V$ f–invariante.

  Essere o no triangolabile è una proprietà invariante rispetto alla coniugazione, per cui possiamo restringere la relazione di equivalenza all'insieme $\mathcal{T}(V)$ degli endomorfismi triangolabili. Poiché $\mathbb{C}$ è un campo algebricamente chiuso, allora $\mathcal{T}(V) =$ End (V) per ogni $\mathbb{C}$-spazio vettoriale V.

La teoria della "forma canonica di Jordan" fornirà un insieme completo di invarianti ed una descrizione esaustiva del quoziente $\mathcal{T}(V)/\sim$.

1. Ideale e polinomio caratteristico di un endomorfismo

Dati $f \in \text{End }(V)$ e un polinomio $p(t) = a_0 + a_1 t + \cdots + a_k t^k \in \mathbb{K}[t]$, poniamo

$$p(f) := a_0 f^0 + a_1 f^1 + \ldots a_k f^k \in \text{End }(V)$$

dove $f^0 = Id$, $f^1 = f$, $f^n = f^{n-1} \circ f$.

Il seguente semplice Lemma è cruciale in molti degli argomenti che seguono.

Lemma 1.1. *Se $p(t) = a(t)b(t)$ allora $p(f) = a(f) \circ b(f) = b(f) \circ a(f)$.*

Dimostrazione. Ci si riduce al caso evidente di un monomio per cui cui $t^m = t^r t^s$, $f^m = f^r \circ f^s = f^s \circ f^r$.

□

Definiamo

$$I(f) = \{p(t) \in \mathbb{K}[t];\ p(f) = 0\}$$

Lemma 1.2.
(1) *$I(f)$ contiene polinomi di grado ≥ 1.*
(2) *Per ogni $p(t), q(t) \in I(f)$, allora $p(t) + q(t) \in I(f)$; per ogni $p(t) \in I(f)$ e ogni $q(t) \in \mathbb{K}[t]$, allora $q(t)p(t) \in I(f)$. Un sottoinsieme non vuoto di $\mathbb{K}[t]$ che verifichi queste proprietà è detto un ideale, pertanto $I(f)$ è detto* l'ideale dell'endomorfismo.
(3) *Se $f \sim g$, allora $I(f) = I(g)$.*

Dimostrazione. Se $f = 0$, allora $t^n \in I(f)$ per ogni $n \geq 1$. Sia $f \neq 0$ Se $m > n^2$, allora gli endomorfismi $f^0, f^1, \ldots, f^m$ sono linearmente dipendenti per cui esiste $p(t)$ di grado k $m \geq k \geq 1$ tale che $p(f) = 0$. Le proprietà che fanno di $I(f)$ un ideale di $\mathbb{K}[t]$ sono di verifica immediata, utilizzando anche il Lemma 1.1. Se $g = h \circ f \circ h^{-1}$, allora per ogni polinomio $p(t)$, $p(g) = h \circ p(f) \circ h^{-1}$. Poiché h è invertibile, $p(g) = 0$ se e solo se $p(f) = 0$.

□

Un polinomio minimo per f è un polinomio $q(t) \in I(f)$ di grado minimo tra i polinomi di grado ≥ 1 che stanno in $I(f)$.

Lemma 1.3. *$I(f)$ è generato da ogni suo polinomio minimo $q(t)$, cioè $I(f) = (q(t)) := \{p(t)q(t);\ p(t) \in \mathbb{K}[t]\}$.*

Dimostrazione. Sia $a(t) \in I(f)$. Eseguendo la divisione Euclidea ("con il resto") sapiamo che esistono unici polinomi $Q(t)$ e $R(t)$ tali che

$$a(t) = Q(t)q(t) + R(t), \quad \text{grado } R(t) < \text{grado } q(t)\ .$$

Poiché $I(f)$ è un ideale, $R(t) = a(t) - Q(t)q(t) \in I(f)$. Poiché $q(t)$ è un polinomio minimo, allora $R(t) = R$ è uno scalare. Infine $RId = a(f) - Q(f) \circ q(f) = 0 - Q(f) \circ 0 = 0$, implica necessariamente che $R = 0$. Dunque $a(t) = Q(t)q(t)$ come voluto.

□

Una conseguenza immediata del precedente Lemma è che esiste un unico polinomio minimo *monico* $q_f(t)$ (cioè con coefficiente del monomio di grado massimo uguale a 1). Questo $q_f(t)$ è detto "il"*polinomio minimo di f*. Segue immediatamente da quanto visto che

Corollario 1.4. *Se $f \sim g$, allora $q_f(t) = q_g(t)$ (cioè disponiamo di un altro invariante polinomiale oltre il già noto polinomio caratteristico $p_f(t)$).*

In effetti i due polinomi invarianti sono strettamente imparentati.

Proposizione 1.5. *$p_f(t) \in I(f)$ e quindi $q_f(t)$ divide $p_f(t)$.*

Dimostrazione. Lo dimostreremo nel caso in cui $f \in \mathcal{T}(V)$. Daremo poi alla fine una indicazione di come dimostrarlo in generale. Sia allora $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V con bandiera f–invariante, per cui $f(v_1) = \mu_1 v_1$, $f(v_2) = *v_1 + \mu_2 v_2$, $f(v_3) = *v_1 + *v_2 + \mu_3 v_3$, ecc., dove gli asterischi stanno a indicare qualche scalare che non ha importanza specificare, i μ_j sono gli autovalori di f e sono possibili ripetizioni. Sappiamo che

$$p_f(f) = \pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_1 Id)$$

e che i fattori possono essere permutati a piacimento senza alterare il risultato della composizione. Basta dimostrare che $p_f(f)(v_j) = 0$ per ogni vettore di $\mathcal{B}$. Adesso $\pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_2 Id) \circ (f - \mu_1 Id)(v_1) = \pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_2 Id)(0) = 0$; $\pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_3 Id) \circ (f - \mu_1 Id) \circ (f - \mu_2 Id)(v_2) = \pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_3 Id) \circ (f - \mu_1 Id)(*v_1) = \pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_3 Id)(0) = 0$. Iterando in modo induttivo la procedura si conclude appunto che $p_f(f)(v_j) = 0$ per ogni j, come voluto.

Vediamo come si può operare quando f non è necessariamente triangolabile. Consideriamo intanto il caso reale $\mathbb{K} = \mathbb{R}$. In questo caso possiamo usare la procedura della "complessificazione" descritta per esempio nella nota [3] sul TSH reperibile in nella sezione "Didattica" di

http://www.dm.unipi.it/ benedett

Allora $p_f(t)$ coincide con il polinomio caratteristico $p_{f_\mathbb{C}}(t)$ dell'endomorfismo complessificato $f_\mathbb{C} \in \text{End}\,(V_\mathbb{C})$. Poiché f è la restrizione di $f_\mathbb{C}$ alla parte reale V di $V_\mathbb{C}$, e $p_{f_\mathbb{C}}(f_\mathbb{C}) = 0$ (perchè su $\mathbb{C}$ tutti gli endomorfismi sono triangolabili), possiamo concludere che anche $p_f(f) = 0$. In generale, dato f, supponiamo esista un campo $\mathbb{F}$ che estende $\mathbb{K}$ (per cui $\mathbb{K}[t] \subset \mathbb{F}[t]$) tale che $p_f(t) \in \mathbb{K}[t]$ sia completamente fattorizzabile in $\mathbb{F}[t]$. Fissiamo una base $\mathcal{B}$ di V e sia A la matrice che rappresenta f in quel sistema di coordinate. Allora $p_f(f) = 0$ se e solo se $p_A(A) = 0$. Adesso $X \to AX$ definisce anche un endomorfismo di $\mathbb{F}^n$, che sappiamo essere triangolabile. Allora $p_A(A) = 0$ che comunque è l' identità tra matrici su $\mathbb{K} \subset \mathbb{F}$ che volevamo. L'esistenza di estensioni $\mathbb{K} \subset \mathbb{F}$ con questa proprietà è garantita da una teoria elementare che almeno gli studenti di Matematica dovrebbero incontrare nel corso di Algebra del primo anno della laurea triennale.

□

Lemma 1.6. *Se $p(t) \in I(f)$ e $\lambda \in \mathbb{K}$ è un autovalore di f allora $p(\lambda) = 0$.*

Dimostrazione. Sia $v \neq 0$, un autovettore tale che $f(v) = \lambda v$. Allora $0 = p(f)(v) = p(\lambda)v$. Poichè $v \neq 0$, necessariamente $p(\lambda) = 0$.

□

Tiriamo allora alcune conseguenze di quanto visto quando $f \in \mathcal{T}(V)$.

Corollario 1.7. *Se f è un endomorfismo triangolabile allora*

$$p_f(t) = \pm \prod_{j=1}^{k} (t - \lambda_j)^{m_j}$$

$$q_f(t) = \pm \prod_{j=1}^{k} (t - \lambda_j)^{r_j}$$

$$\lambda_j \neq \lambda_i, \text{ se } i \neq j; \ \sum_j m_j = n; \ 1 \leq r_j \leq m_j, \ j = 1, \dots k \ .$$

Dimostrazione. Il polinomio caratteristico è completamente fattorizzabile in $\mathbb{K}[t]$ perché f è triangolabile. Il polinomio minimo $q_f(t)$ è di quella forma per qualche $0 \leq r_j \leq m_j$, perché il polinomio minimo divide quello caratteristico. Infine ogni $r_j \geq 1$ perché ogni autovalore è radice di $q_f(t) \in I(f)$.

$\square$

2. Decomposizione primaria

Si ricordi per esempio che f è diagonalizzabile se e solo se $V = \oplus V_{\lambda_j}(f)$ è la somma diretta degli autospazi di f, cioè è la somma diretta di sottospazi f–invarianti sui quali la restrizione di f è particolarmente semplice, trattandosi di un multiplo dell'identità. Analogamente, il programma che ci porterà infine alla "forma normale di Jordan" per un arbitrario endomorfismo f che sia solo triangolabile, si propone di decomporre in modo "canonico" V in somma diretta di sottospazi f–invarianti sui quali la restrizione di f sia la "più semplice possibile". Il "Teorema di decomposizione primaria" che andiamo a discutere è il passo fondamentale in quella direzione. Richiamiamo prima alcune nozioni riguardanti i polinomi (almeno gli studenti di Matematica potranno ancora una volta riferirsi al corso di Algebra). Dati due polinomi non nulli $a(t), b(t) \in \mathbb{K}[t]$, poniamo

$$I(a(t), b(t)) = \{p(t) = m(t)a(t) + n(t)b(t);\ m(t), n(t) \in \mathbb{K}[t]\}\ .$$

Non è difficile mostrare usando come prima la divisione Euclidea tra polinomi che

(1) $I = I(a(t), b(t))$ è un ideale in $\mathbb{K}[t]$ che contiene polinomi non nulli.
(2) Esiste un unico polinomio monico $d(t)$ di grado minimo tra i polinomi non nulli contenuti nell'ideale, e questo $d(t)$ genera $I = (d(t))$.
(3) Il polinomio $d(t)$ verifica le seguenti proprietà:
  - E' un divisore comune di $a(t)$ e $b(t)$, cioè $d(t)|a(t)$ e $d(t)|b(t)$;
  - Se $p(t)$ è un altro divisore comune allora $p(t)|d(t)$. Per questo motivo $d(t)$ è detto il "massimo comun divisore" $\mathrm{MCD}(a(t), b(t))$

Segue in particolare dalla precedente discussione che il MCD (come tutti gli elementi di $I(a(t), b(t))$) può essere scritto nella forma

(Identità di Bezout) $d(t) = m(t)a(t) + n(t)b(t)$ per qualche $m(t), n(t) \in \mathbb{K}[t]$.

Possiamo infine enunciare

Proposizione 2.1. (Decomposizione primaria.) *Sia* $f \in \mathrm{End}\,(V)$ *e sia* $p(t) \in I(f)$ *non nullo. Supponiamo che* $p(t) = a(t)b(t)$ *e che* $MCD(a(t), b(t)) = 1$ *(cioè* $p(t)$ *è prodotto di due fattori "coprimi"). Allora:*

(1) Si ha la decomposizione in somma diretta

$$V = \mathrm{Ker}(a(f)) \oplus \mathrm{Ker}(b(f))$$

e i due addendi diretti sono f*–invarianti.*

(2) Se $g = h \circ f \circ h^{-1}$*, così che* $V = \mathrm{Ker}(a(f)) \oplus \mathrm{Ker}(b(f)) = \mathrm{Ker}(a(g)) \oplus \mathrm{Ker}(b(g))$*, allora* $h(\mathrm{Ker}(a(f))) = \mathrm{Ker}(a(g))$ *e* $h(\mathrm{Ker}(b(f))) = \mathrm{Ker}(b(g))$*.*

Dimostrazione. Usando l'identità di Bezout abbiamo che

$$id = m(f) \circ a(f) + n(f) \circ b(f)$$

da cui per ogni $v \in V$

$$v = m(f) \circ a(f)(v) + n(f) \circ b(f)(v)\ .$$

Osserviamo che $b(f) \circ m(f) \circ a(f)(v) = m(f) \circ p(f)(v) = 0$ perché $p(t) \in I(f)$. Dunque $m(f) \circ a(f)(v) \in \mathrm{Ker}(b(f))$. Analogamente $n(f) \circ b(f)(v) \in \mathrm{Ker}(a(f))$. Abbiamo quindi dimostrato che $V = \mathrm{Ker}(a(f)) + \mathrm{Ker}(b(f))$. D'altra parte se $v \in \mathrm{Ker}(a(f)) \cap \mathrm{Ker}(b(f))$ $v = m(f) \circ a(f)(v) + n(f) \circ b(f)(v) = 0$, per cui è una somma diretta. Il fatto che gli addendi diretti siano f-invarianti è di semplice verifica usando ancora una volta il Lemma 1.1.

Venendo al punto (2), si ha per esempio $a(g) = h \circ a(f) \circ h^{-1}$ da cui per ogni $v \in V$, $a(g) \circ h(v) = h \circ a(f)(v)$ da cui ricaviamo l'inclusione $h(\mathrm{Ker}(a(f))) \subset \mathrm{Ker}(a(g))$. Per simmetria ricaviamo anche l'inclusione opposta. Si ragiona analogamente con $b(t)$. La Proposizione è dimostrata. $\square$

Tiriamo le conseguenze quando $f \in \mathcal{T}(V)$. Poiché $p_f(t) \in I(f)$ è completamente fattorizzabile, e i fattori corrispondenti ad autovalori distinti sono chiaramente coprimi, applicando ripetutamente la Proposizione precedente abbiamo

Corollario 2.2. *Se f è triangolabile allora*

$$V = \oplus_{j=1}^{k} \mathrm{Ker}(f - \lambda_j Id)^{m_j}$$

ed ogni addendo diretto è f-invariante.

Vogliamo adesso cominciare ad analizzare le proprietà della restrizione di f ad ogni addendo diretto. Premettiamo alcune osservazioni che valgono per ogni endomorfismo.

Lemma 2.3. *Sia $h \in \mathrm{End}\,(V)$. Allora:*

(1) Per ogni $n \geq 1$, $\mathrm{Ker}(h^n) \subset \mathrm{Ker}(h^{n+1})$.

(2) Se $\mathrm{Ker}(h^n) = \mathrm{Ker}(h^{n+1})$ allora $\mathrm{Ker}(h^n) = \mathrm{Ker}(h^m)$ per ogni $m \geq n$.

(3) Se $p = k \circ h \circ k^{-1}$ è coniugato ad h, allora per ogni $n \geq 1$, $k(\mathrm{Ker}(h^n)) = \mathrm{Ker}(p^n)$. In particolare ogni dimensione $d_n = \dim \mathrm{Ker}(h^n)$ è invariante per coniugazione.

Dimostrazione. La prima affermazione è evidente. Per mostrare la seconda basta far vedere che $\mathrm{Ker}(h^{n+1}) = \mathrm{Ker}(h^{n+2})$, e concludere poi per induzione. In effetti basta dimostrare che vale l'inclusione $\mathrm{Ker}(h^{n+2}) \subset \mathrm{Ker}(h^{n+1})$. Ma $h^{n+2}(v) = h^{n+1}(h(v)) = 0$ se e solo se $h(v) \in \mathrm{Ker}(h^{n+1}) = \mathrm{Ker}(h^n)$. Dunque $h^n(h(v)) = h^{n+1}(v) = 0$ come voluto. Per dimostrare il punto (3) si ragiona come nella dimostrazione di (2) in Proposizione 2.1.

$\square$

Proposizione 2.4. *Sia f un endomorfismo triangolabile con polinomio caratteristico e polinimio minimo come stabilito nel Corollario 1.7. Sia λ un autovalore di f di molteplicità m rispetto al polinomio caratteristico, di molteplicità r rispetto al polinomio minimo. Indichiamo con $W_\lambda := \mathrm{Ker}(f - \lambda Id)^m$ l'addendo diretto corrispondente a λ nella decomposizione primaria associata al polinomio caratteristico $p_f(t)$. Indichiamo con $g_\lambda = f|W_\lambda$. Allora:*

(1) $\dim W_\lambda = m$, $p_{g_\lambda}(t) = \pm(t-\lambda)^m$, $q_{g_\lambda}(t) = \pm(t-\lambda)^r$.

(2) Ogni dimensione $d_n = \dim \mathrm{Ker}(f - \lambda Id)^n$ è invariante per coniugazione; $\dim V_\lambda(f) = d_1 < d_2 \cdots < d_r = \dim W_\lambda = m$.

(3) La decomposizione primaria associata al polinomio caratteristico di f coincide con la decomposizione primaria associata al polinomio minimo.

Dimostrazione. Applicando il Lemma precedente a $h := (f - \lambda Id)$ vediamo che la stringa dei successivi nuclei è contenuta in W_λ; inizia con l'autospazio $V_\lambda(f)$ e termina con tutto W_λ. Ne segue che ogni g_λ ammette solo l'autovalore λ. D'altra parte, usando un sistema di coordinate adattato alla decomposizione primaria, si vede immediatamente che $p_f(t) = \prod_\lambda p_{g_\lambda}(t)$ da cui necessariamente $p_{g_\lambda}(t) = \pm(t-\lambda)^m$, e quindi $\dim W_\lambda = m$. Se uno dei g_λ avesse polinomio minimo di grado strettamente minore di r, allora $\prod_\lambda q_{g_\lambda}(t)$ sarebbe un polinomio appartenente a $I(f)$ di grado ≥ 1 e strettamente minore del grado di $q_f(t)$, che è impossibile. Il punto (1) è così dimostrato. La prima affermazione del punto (2) segue dal punto (3) del Lemma 2.3 osservando che $f \sim f'$ se e solo se $(f - \alpha Id) \sim (f' - \alpha Id)$ per ogni scalare α. La seconda segue dal precedente punto (1) e dal punto (2) del Lemma 2.3. Il punto (3) è ora evidente. $\square$

Riassumenndo quanto fatto, lo studio degli endomorfismi triangolabili a meno di coniugazione è ridotto allo studio degli endomorfismi "speciali" $g : W \to W$ tali che $p_g(t) = \pm(t-\lambda)^m$ (per cui $\dim W = m$), e $p_g(t) = \pm(t-\lambda)^r$, per qualche $1 \leq r \leq m$. Inoltre posto $h := g - \lambda Id$, la successione di dimensioni strettamente crescenti $\dim V_\lambda(g) = d_1 < d_2 \cdots < d_r = \dim W = m$ è invariante per coniugazione. Tutte queste informazioni sono incorporate nella stringa di invarianti numerici

$$[\lambda, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)] \ .$$

Mostreremo che questo è l'invariante completo che cercavamo.

Come già osservato, se g e g' sono endomorfismi speciali su W con lo stesso autovalore λ, allora $g \sim g'$ se e solo se $h \sim h'$. La dimostrazione del seguente Lemma è immediata.

Lemma 2.5. *Se g è speciale su W con con stringa di invarianti $[\lambda, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)]$, allora $h = g - \lambda Id$ è speciale su W con stringa di invarianti $[0, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)]$.*

Un endomorfismo speciale h con autovalore $\lambda = 0$ è detto *nilpotente*. Se g è speciale di autovalore λ, allora $h = g - \lambda Id$ è detto la parte nilpotente di g. Possiamo allora riassumere quanto osservato qui sopra dicendo che non è restrittivo assumere che $\lambda = 0$, cioè che tutto lo studio è ridotto al caso degli endomorfismi nilpotenti. Osserviamo anche che se W viene decomposto in somma diretta di sottospazi h-invarianti, allora la restizione di h ad ogni addendo diretto è nilpotente.

Cominciamo con l' analizzare i due casi estremi: $r = 1$, $r = m$.

$r = 1$: questo equivale a dire che la stringa invariante è $[0, 1, (m)]$, cioè che $W = V_0(h)$ e siamo nel caso diagonalizzabile.

$r = m$:

Definizione 2.6. Una base di W è detta *ciclica* per h nilpotente se è della forma $\mathcal{B} = \{h^{m-1}(v), h^{m-2}(v), \ldots, h(v), v\}$ per qualche $v \in W$.

La dimostrazione del seguente lemma è immediata.

Lemma 2.7. *Indicate con E^j, $j = 1, \ldots, m$, le colonne della matrice identità $m \times m$. Allora $\mathcal{B}$ è ciclica per h se e solo se le colonne della matrice $M_{\mathcal{B}}(h)$ sono $C^1 = 0$, $C^2 = E^1$, $\ldots$, $C^m = E^{m-1}$.*

La matrice $m \times m$ della forma descritta nel Lemma è indicata con $J(0, m)$ ed è detta *blocco di Jordan di taglia m e autovalore* 0. Analogamente, per ogni $\lambda \in \mathbb{K}$, poniamo $J(\lambda, m) = \lambda I + J(0, m)$ il corrispondente blocco di Jordan di autovalore λ.

Possiamo infine enunciare

Lemma 2.8. *Sia h nilpotente. Allora sono fatti tra loro equivalenti:*

(1) $p_h(t) = q_h(t) = t^m$;

(2) h ha stringa invariante $[0, m, (1, 2, 3, \ldots, m)]$.

(3) Esiste una base ciclica per h.

Dimostrazione. (1) è equivalente a (2) perché le dimensioni sono forzate a crescere ad ogni passo di 1. E' evidente che (3) implica (1). Dimostriamo l'altra implicazione. Poiché $r = m$, esiste sicuramente v tale che $h^{m-1}(v) \neq 0$. Basta allora dimostrare che gli m vettori di $\mathcal{B} = \{h^{m-1}(v), h^{m-2}(v), \ldots, h(v), v\}$ sono linearmente indipendenti. Sia allora $\sum_{j=0}^{m-1} a_j h^j(v) = 0$ e mostriamo che necessariamente tutti gli $a_j = 0$. Applicando h, ricaviamo la relazione $\sum_{j=1}^{m-1} a_{j-1} h^j(v) = 0$. Iterando l'applicazione di h otteniamo un sistema di relazioni che termina con $a_0 h^{m-1}(v) = 0$. Poiché $h^{m-1}(v) \neq 0$, necessariamente $a_0 = 0$, e risalendo in modo induttivo le relazioni trovate si conclude che tutti i coefficienti sono nulli.

$\square$

Abbiamo così trovato i primi due casi di forma normale di Jordan per gli endomorfismi speciali e nilpotenti:

Il blocco $J(0, m)$ corrispondente alla stringa $[0, m, (1, 2, 3, \dots, m)]$;

La matrice diagonale ("a blocchi") formata da m blocchi di Jordan $J(0, 1)$.

Possiamo infine enunciare il teorema della forma normale di Jordan per gli endomorfismi nilpotenti in generale.

Theorem 2.9. *Sia h un endomorfismo nilpotente su W con stringa invariante $[0, r, (d_1 < d_2 < \dots < d_{r-1} < d_r = m)]$. Allora:*

(1) Esiste una decomposizione di $W = \oplus Z_j$ in somma diretta di sottospazi h-invarianti tale che la restrizione di h ad ogni Z_j ammette una base ciclica. Indicata con $\mathcal{B}$ la base di W ottenuta unendo tali basi cicliche, allora $M_{\mathcal{B}}(h)$ è una matrice diagonale a blocchi dove ogni blocco è di Jordan del tipo $J(0, s_j)$, $\sum_j s_j = m$. Questa $M_{\mathcal{B}}(h)$ è detta una forma di Jordan per h, mentre $\mathcal{B}$ è detta una base di Jordan per h.

(2) La forma di Jordan di h è unica a meno di permutazione dei blocchi ed è completamente ed esplicitamente determinabile in funzione soltanto della stringa invariante $[0, r, (d_1 < d_2 < \dots < d_{r-1} < d_r = m)]$, che risulta quindi essere un invariante completo.

Dimostrazione. Consideriamo un frammento a tre termini del tipo $\mathrm{Ker}(h^{j-2}) \subset \mathrm{Ker}(h^{j-1}) \subset \mathrm{Ker}(h^j)$, $1 \leq j-2$, $j \leq r$, della successione strettamente crescente di nuclei. Illustriamo qui la costruzione fondamentale che porterà alla dimostrazione del Teorema. Fissiamo una decomposizione in somma diretta

$$\mathrm{Ker}(h^j) = \mathrm{Ker}(h^{j-1}) \oplus U$$

dove $t = \dim U = d_j - d_{j-1}$, e fissiamo una base $\{u_1, \dots u_t\}$ di U. Applichiamo h ottenendo $\{h(u_1), \dots h(u_t)\}$. Affermiamo che valgono i seguenti fatti:

(1) $\{h(u_1), \dots h(u_t)\}$ è contenuto in $\mathrm{Ker}(h^{j-1})$.
(2) I vettori in $\{h(u_1), \dots h(u_t)\}$ sono lineramente indipendenti.
(3) $\mathrm{Ker}(h^{j-2}) \cap \mathrm{Span}\{h(u_1), \dots h(u_t)\} = \{0\}$.

Dimostriamo queste affermazioni. La prima è evidente. Sia $\sum_i a_i h(u_i) = 0$. Allora $\sum_i a_i u_i \in \mathrm{Ker}(h) \subset \mathrm{Ker}(h^{j-1})$ ma allora tutti i coefficienti sono nulli perché gli u_i sono linearmente indipendenti e $\mathrm{Ker}(h^{j-1}) \cap U = \{0\}$. Questo prova il secondo punto. La prova del terzo è del tutto simile.

Procediamo adesso alla costruzione passo per passo della base $\mathcal{B}$ di W con le proprietà volute. Il lettore deve immaginare che i vettori di questa base saranno arrangiati in una tabella a gradini, che scendono verso destra di altezza e lunghezza variabili. Daremo una dopo l'altra le righe di questa tabella di vettori. Queste righe saranno allineate sul bordo sinistro della tabella. La lunghezza e l'altezza di ciascun gradino sarà esplicitamente calcolabile solamente in funzione della stringa invariante di h.

Partiamo dal frammento a tre termini all'estremità destra: $\mathrm{Ker}(h^{r-2}) \subset \mathrm{Ker}(h^{r-1}) \subset \mathrm{Ker}(h^r)$. Applichiamo la costruzione fondamentale. La prima riga della tabella di vettori sarà data da $(u_1, \dots, u_t)$ dove $t = m - d_{r-1}$. Applichiamo h e otteniamo un primo segmento $(h(u_1), \dots, h(u_t))$ della seconda riga. Consideriamo adesso il frammento a tre termini ottenuto scivolando di una posizione verso sinistra: $\mathrm{Ker}(h^{r-3}) \subset \mathrm{Ker}(h^{r-2}) \subset \mathrm{Ker}(h^{r-1})$ e applichiamo ancora una volta la costruzione fondamentale *raffinandola però come segue*: imponiamo che la base di U sia della forma $h(u_1), h(u_2), \dots, h(u_t), u_{t+1}, \dots, u_{t'}$, e questi vettori formeranno la seconda riga della tabella, mentre $h^2(u_1), h^2(u_2), \dots, h^2(u_t), h(u_{t+1}), \dots, h(u_{t'})$ sarà il segmento iniziale della terza riga. Questo è possibile perché grazie al fatto (3) osservato prima, $\mathrm{Span}\{h(u_1), h(u_2), \dots, h(u_t)\}$ può far parte di un complementare U di $\mathrm{Ker}(h^{r-2})$ in $\mathrm{Ker}(h^{r-1})$. Notiamo che $t' = t + t" = d_{r-1} - d_{r-2}$, $t = d_r - d_{r-1}$, per cui $0 \leq t" = t' - t = 2d_{r-1} - (d_{r-2} + d_r)$. Se $t" = 0$, allora le prime due righe fanno parte dello stesso gradino di altezza maggiore o uguale a 2 e lunghezza uguale a t. Se $t" > O$, allora il primo gradino ha lunghezza uguale a t e altezza uguale a 1, mentre il secondo gradino ha lunghezza maggiore o uguale a $t"$. Consideriamo adesso il frammento a tre termini ottenuto scivolando di un' altra posizione verso sinistra. E applichiamo ancora una volta la versione raffinata della costruzione fondamentale. Iteriamo la procedura fino a quando è possibile. Ne risulta una tabella a gradini di vettori che verifica le proprietà dette in precedenza. Per costruzione l'unione di questi vettori forma una base $\mathcal{B}$ di W. D'altra parte, leggendo dal basso verso l'alto ciascuna colonna della tabella vediamo vettori indipendenti del tipo $\{h^j(u), h^{j-1}(u), \dots, u\}$ tali che, ancora per costruzione, $h^{j+1}(v) = 0$. Dunque lo spazio generato da ogni colonna è h-invariante e munito per costruzione di una base ciclica per la restrizione di h. Dunque l'esistenza di una forma di Jordan per h è dimostrata. Riguardo all'unicità. A meno di permutazione dei blocchi (cioè a meno di permutazioni dei vettori di una "base di Jordan" per h) possiamo assumere che i blocchi siano disposti lungo la diagonale in modo che la taglia decresca. Possiamo allora organizzare i vettori della base in una tabella a gradini con le proprietà formali della costruzione precedente, dove l'altezza e la lunghezza di ogni gradino è esplicitamente esprimibile in funzione delle taglie dei blocchi. A questo punto è facile convincersi che questa tabella è il risultato di una implementazione della costruzione che ha portato alla dimostrazione dell'esistenza, per cui la lunghezza e l'altezza di ogni gradino (quindi la taglia di ogni blocco) è esplicitamente esprimibile in funzione solamente della stringa invariante di h. Il teorema è così dimostrato.

□

Osservazioni 2.10. A parte quelle relative al primo gradino, volutamente non abbiamo esplicitato l' espressione delle lunghezze e delle altezze dei gradini in funzione di $[0, r, (d_1 < d_2 < \dots < d_{r-1} < d_r = m)]$. Può essere un utile esercizio per il lettore farlo, seguendo la costruzione induttiva di una base di Jordan. Può essere anche istruttivo ricavare da $[0, r, (d_1 < d_2 < \dots < d_{r-1} < d_r = m)]$ informazioni parziali relative alla forma normale di Jordan. Per esempio il lettore può dimostrare per esercizio che il numero totale dei blocchi è uguale a d_1, mentre la taglia massima al variare dei blocchi è uguale a r.

Per completezza enunciamo la versione generale del Teorema sulla forma canonica di Jordan per endomorfismi triangolabili arbitrari (non necessariamente nilpotenti). La dimostrazione è una conseguenza immediata della discussione precedente.

Teorema 2.1. *Sia f un endomorfismo triangolabile su V. Possiamo associare a f l'invariante per coniugazione $\mathcal{J}(h)$ dato dall'insieme delle stringhe invarianti $[\lambda, r_\lambda, (d_1^\lambda < \dots < d_{r_\lambda}^\lambda = m_\lambda)]$, al variare di λ nello spettro di f. Allora:*

(1) Esiste una base $\mathcal{B}$ (detta di Jordan per f) tale che $M_{\mathcal{B}}(f)$ è una matrice diagonale a blocchi tale che ogni blocco è della forma $J(\lambda, s)$ dove λ varia nello spettro di f. Tale matrice è detta una forma di Jordan per f.

(2) La forma di Jordan è unica a meno di permutazione dei blocchi ed è completamente e esplicitamente determinata in funzione dell'invariante $\mathcal{J}$ che risulta quindi un invariante completo.

3. La forma normale di Jordan reale

E' un'altra applicazione della procedura della "complessificazione" degli endomorfismi reali descritta con i dettagli per esempio nella nota [3] sul TSH reperibile in

http://www.dm.unipi.it/ benedett.

Quindi ci limitiamo qui a descrivere rapidamente il risultato. Dato un endomorfismo reale $f : V \to V$ (non necessariamente triangolabile) il complessificato $f_{\mathbb{C}} : V_C \to V_{\mathbb{C}}$ è tale che $f_{\mathbb{C}}|V = f$ e $p_f(t) = p_{f_{\mathbb{C}}}(t) \in \mathbb{R}[t]$. Allora la decomposizione primaria di $V_{\mathbb{C}}$ relativa a $p_{f_{\mathbb{C}}}$ può presentare addendi diretti $f_{\mathbb{C}}$-invarianti di due tipi:

(1) $\mathrm{Ker}(f_{\mathbb{C}} - \lambda Id)^m$, quando λ è reale di molteplicità m. In questo caso $\mathrm{Ker}(f_{\mathbb{C}} - \lambda Id)^m$ è il complessificato di $\mathrm{Ker}(f - \lambda Id)^m$.

(2) $\mathrm{Ker}(f_{\mathbb{C}} - \alpha Id)^k \oplus \mathrm{Ker}(f_{\mathbb{C}} - \overline{\alpha} Id)^k$, quando $\alpha \neq \overline{\alpha}$ ed entrambi hanno la stessa molteplicità k. Questo addendo è invariante per la coniugazione su $V_{\mathbb{C}}$, $z \to \overline{z}$ ed è il complessificato della sua parte reale cioè del luogo dei suoi punti tali che $z\overline{z}$.

Nel primo caso si applica a $\mathrm{Ker}(f - \lambda Id)^m$ la discussione precedente (su $\mathbb{R}$) e ne risulta una base di Jordan per la restrizione di f che è anche una base di Jordan reale per la restrizione di $f_{\mathbb{C}}$ a $\mathrm{Ker}(f_{\mathbb{C}} - \lambda Id)^m$.

Nel secondo caso, poiché $f_{\mathbb{C}}$ è il complessificato di f reale, risulta che se $\mathcal{B} = \{v_1, \dots, v_k\}$ è una base di Jordan per la restrizione di $f_{\mathbb{C}}$ a $\mathrm{Ker}(f_{\mathbb{C}} - \alpha Id)^k$ allora $\overline{\mathcal{B}} = \{\overline{v}_1, \dots, \overline{v}_k\}$ è una base di Jordan per la restrizione di $f_{\mathbb{C}}$ a $\mathrm{Ker}(f_{\mathbb{C}} - \overline{\alpha} Id)^k$. Prendendo le parti reali e immaginarie dei vettori di $\mathcal{B}$

$$(v_j + \overline{v}_j)/2, \quad -i(v_j - \overline{v}_j)/2$$

si effettua un cambiamento di base da $\mathcal{B}$ ad una base reale $\mathcal{B}_{\mathbb{R}}$ che è per definizione una *base di Jordan reale* della restrizione di f alla parte reale di $\mathrm{Ker}(f_{\mathbb{C}} - \alpha Id)^k \oplus \mathrm{Ker}(f_{\mathbb{C}} - \overline{\alpha} Id)^k$. La corrispondente matrice rappresentativa è detta in *forma di Jordan reale*. Questa è diagonale a blocchi dove ogni blocco reale relativo alla coppia di autovalori $(\alpha, \overline{\alpha})$ è indicato con $J_{\mathbb{R}}(\alpha, 2s)$. Questo presenta lungo la diagonale s blocchi 2×2 della forma

$$\begin{bmatrix} |\alpha|\cos(\theta) & -|\alpha|\sin(\theta) \\ |\alpha|\sin(\theta) & |\alpha|\cos(\theta) \end{bmatrix}$$

dove $\alpha = |\alpha|e^{i\theta}$, mentre al di sopra della diagonale presenta blocchi uguali alla matrice identità 2×2. Una volta precisata la forma di questi blocchi reali relativi ad autovalori non reali, lasciamo al lettore la cura di formulare la versione reale del teorema di esistenza e unicità della forma canonica di Jordan (reale).