Complementi di geometria: Sottospazi invarianti, disuguaglianze sul rango.

Def: Sia $f: X \to Y$ e sia $Z \in X$, l'immagine di Z rispetto ad f è data da
$f(Z) = \{ f(x) ; x \in Z \}$

Proposizione: Se $f: X \to Y$ è una trasformazione lineare e Z è un sottospazio di X, allora $f(Z)$ è sottospazio di Y

Def: Sia $A \in \mathbb{C}^{m \times n}$ e sia Z un sottospazio di $\mathbb{C}^n$, sia $a: \mathbb{C}^n \to \mathbb{C}^m$, $a(x) = Ax$

allora l'immagine di Z rispetto alla matrice A è data da $A(Z)=\mathcal{A}(Z)=\{Ax, x\in Z\}$.

Proprietà: Se $A\in\mathbb{C}^{m\times n}$, $B\in\mathbb{C}^{n\times p}$, allora $A(\operatorname{Im}B)=\operatorname{Im}(AB)$

Dim: 1. $A(\operatorname{Im}B)\subseteq \operatorname{Im}(AB)$.

2. $\operatorname{Im}(AB)\subseteq A(\operatorname{Im}B)$.

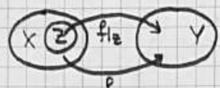
1. Se $x\in A(\operatorname{Im}B)$, $\exists z\in \operatorname{Im}B: x=Az$, $\exists w: z=Bw$, $x=ABw \Rightarrow x\in\operatorname{Im}(AB)$

2. Se $x\in\operatorname{Im}(AB)$, $\exists w$. $x=ABw$, $z:=Bw \Rightarrow x=Az \Rightarrow z\in\operatorname{Im}B$, $x\in A(\operatorname{Im}B)$ $\square$

Esempio: $A=\begin{bmatrix}1&0&1&0\\2&1&1&1\\1&0&-1&0\end{bmatrix}$, $B=\begin{bmatrix}1&0\\0&0\\-1&1\\0&1\end{bmatrix}$ $A(\operatorname{Im}B)=\operatorname{Im}(AB)=$

$=\operatorname{Im}\begin{bmatrix}0&1\\1&2\\2&-1\end{bmatrix}$

Def: Sia $f:X\to Y$, $Z\subseteq X$, la restrizione di f a Z è data da $f_{|Z}:Z\to Y$ tale che $(\forall x\in Z)\ f_{|Z}(x)=f(x)$.



Proposizione: Se $f:X\to Y$ è una trasformazione lineare e $Z\subseteq X$ è un sottospazio di X, allora $f_{|Z}:Z\to Y$ è a sua volta una trasformazione lineare.

Def: Sia $f:X\to X$ una trasformazione lineare e sia Z un sottospazio di X. Diciamo che Z è invariante rispetto ad f se $f(Z)\subset Z$

Se $A\in\mathbb{C}^{m\times m}$ e Z è un sottospazio di $\mathbb{C}^m$, Z è invariante rispetto ad A se $A(Z)\subseteq Z$.

Esempio: $A=\begin{bmatrix}1&-1&-1\\-1&4&5\\1&-2&-3\end{bmatrix}$, $Z=\operatorname{Im}M$, $M=\begin{bmatrix}0&1\\1&-1\\-1&1\end{bmatrix}$ quindi $\dim Z=2$, vogliamo far vedere che $A(Z)\subseteq Z$

$A(Z)=A(\operatorname{Im}M)=\operatorname{Im}(AM)\overset{?}{\subseteq} Z=\operatorname{Im}M$. Procediamo in questo modo: $[M\,|\,AM]$ e poi facciamo la riduzione di Gauss e verifichiamo che tutte le colonne di AM siano combinazione lineare delle colonne di M.

$AM=\begin{bmatrix}0&1\\-1&0\\1&0\end{bmatrix}$ $[M\,|\,AM]=\left[\begin{array}{cc|cc}0&1&0&1\\1&-1&-1&0\\-1&1&1&0\end{array}\right]\to\left[\begin{array}{cc|cc}1&-1&-1&0\\0&1&0&1\\0&0&0&0\end{array}\right]$ Questa è la forma di Gauss, le prime due colonne sono indipendenti mentre le ultime due dipendono linearmente dalle prime due e quindi $\operatorname{Im}(AM)\subseteq M$ ossia $A(Z)\subseteq Z$.

Se Q è una trasformazione lineare $Q: V \to W$, Z sottospazio di V invariante rispetto ad Q ossia $Q(Z) \subseteq Z$. La restrizione $Q|_Z : Z \to W$, $(\forall x \in Z)\ Q|_Z(x) = Q(x)$.
In questo caso possiamo dire che $Q(x) \subseteq Z$ quindi in realtà è ben definita anche la trasformazione lineare $Q|_Z : Z \to Z$, $(\forall x \in Z)\ Q|_Z(x) = Q(x)$. Per i sottospazi invarianti useremo questa seconda definizione.

Proprietà: $Q: V \to V$ trasformazione lineare, Z sottospazio di V, $Q(Z) \subseteq Z$, $m = \dim V$, $n = \dim Z$, $B = \{b_1, \dots, b_n, b_{n+1}, \dots, b_m\}$, scegliamo questa base in modo tale che B è base di V e che $\bar{B} = \{b_1, \dots, b_n\}$ è base di Z. Allora:

$$[Q]_{B,B} = \begin{array}{cc} \begin{array}{cc} \scriptstyle n & \scriptstyle m-n \end{array} & \\ \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] & \begin{array}{l} \}\, n \\ \}\, m-n \end{array} \end{array} \quad \text{Inoltre } A_{11} = [Q|_Z]_{\bar{B},\bar{B}}.$$

Dim.: $[Q]_{B,B} = [\underbrace{[Q(b_1)]_B, \dots, [Q(b_n)]_B}_{n}, \underbrace{[Q(b_{n+1})]_B, \dots, [Q(b_m)]_B}_{m-n}]$, a noi interessano le prime n colonne

Osservazione chiave $Q(b_1) \in Z$ poiché $Q(Z) \subseteq Z$ e $b_1 \in Z$. $Z = \operatorname{Im}[b_1, \dots, b_n]$, quindi $Q(b_1) = \alpha_1 b_1 + \dots + \alpha_n b_n$. $[Q(b_1)]_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \}\, n \\ \\ \}\, m-n \end{array}$ la presenza di questi $m-n$ zeri strutturali ci dà la struttura del teorema.

$[Q|_Z]_{\bar{B},\bar{B}} = [[Q|_Z(b_1)]_{\bar{B}}, \dots, [Q|_Z(b_n)]_{\bar{B}}] = [[Q(b_1)]_{\bar{B}}, \dots, [Q(b_n)]_{\bar{B}}]$, ricordando la scrittura precedente $Q(b_1) = \alpha_1 b_1 + \dots + \alpha_n b_n$, quindi $[Q(b_1)]_{\bar{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$ ma questo coincide con le prime n caselle del caso precedente e quindi $[Q|_Z]_{\bar{B},\bar{B}} = A_{11}$.

<u>Osservazione</u>: $\sigma(Q) = \sigma([Q]_{B,B}) = \sigma\left(\underbrace{\left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right]}_{A}\right) = \sigma(A_{11}) \cup \sigma(A_{22})$ poiché $\chi_A(\lambda) = \det(\lambda I - A) = \det \left[\begin{array}{c|c} \lambda I - A_{11} & -A_{12} \\ \hline 0 & \lambda I - A_{22} \end{array}\right] =$
$= \det(\lambda I - A_{11}) \cdot \det(\lambda I - A_{22}) = \chi_{A_{11}}(\lambda) \cdot \chi_{A_{22}}(\lambda)$.
Quindi $\sigma(Q) = \sigma(Q|_Z) \cup \sigma(A_{22})$.

La possibilità di poter restringere gli autovalori sugli spazi invarianti ci darà delle informazioni importanti sulle proprietà strutturali dei sistemi di controllo che vedremo più avanti nello specifico.

Esempio: $A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 4 & 5 \\ 1 & -2 & -3 \end{bmatrix}$, $Z = \operatorname{Im} M$, $M = \begin{bmatrix} 0 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}$, $Q: \mathbb{C}^3 \to \mathbb{C}^3$, $Q(x) = Ax$

(*) questo è più vantaggioso poiché ha più zeri

→ Basi di Z, il primo è la prima col + la seconda

→ basta che sia L.I. dai precedenti

Vogliamo calcolare $[Q]_{B,B}$ con $B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^{(*)}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} = \{b_1, b_2, b_3\}$

Dato che B poi costituirà una matrice che dovrà essere invertita, tanto più le colonne ci danno una matrice simile all'identità, meno conti dovremo fare

$$[Q]_{B,B}=\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}^{-1}A\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix},\quad \begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}^{-1}=\begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix}$$

facciamo il conto per righe perché così è più semplice

$$\begin{matrix}1&0&0\\0&1&1\\0&0&1\end{matrix}$$

Quindi
$$\begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}1&0&-1\\-1&-1&5\\1&1&-3\end{bmatrix}=\left[\begin{array}{cc|c}1&0&-1\\-1&-1&5\\\hline \boxed{0\;\;0}&&2\end{array}\right]$$

Verifichiamo se la struttura prevista torna. I due zeri strutturali ci sono, se non ci fossero stati avrebbe significato un errore di calcolo, $A_{11}=\begin{bmatrix}1&0\\-1&-1\end{bmatrix}=[Q|_Z]_{\bar{B}}$, inoltre $\sigma(A)=\sigma(A_{11})\cup\sigma(A_{22})=$
$=\sigma\left(\begin{bmatrix}1&0\\-1&-1\end{bmatrix}\right)\cup\sigma([2])=\{1,-1\}\cup\{2\}=\{1,-1,2\}$, inoltre $\sigma(Q|_Z)=\{1,-1\}$

<u>Proprietà</u>: Sia $Q: V\to V$ una trasformazione lineare e siano $Z_1, Z_2, \dots, Z_\ell$ sottospazi di V tali che:

1. $Z_1\oplus Z_2\oplus\dots\oplus Z_\ell=V$
2. $Q(Z_i)\subseteq Z_i$, per $i=1,\dots,\ell$

Sia $r_i=\dim Z_i$ e sia B_i una base di Z_i, per $i=1,\dots,\ell$.
Sia $B=\{B_1,B_2,\dots,B_\ell\}$ una base di V, allora

$$[Q]_{B,B}=\begin{array}{c}\begin{matrix}r_1&r_2&\dots&r_\ell\end{matrix}\\\left[\begin{array}{c|c|c|c}A_1&0&0&0\\\hline 0&A_2&0&0\\\hline 0&0&\ddots&0\\\hline 0&0&0&A_\ell\end{array}\right]\end{array}\begin{matrix}r_1\\r_2\\\vdots\\r_\ell\end{matrix},\quad A_i=[Q|_{Z_i}]_{B_i,B_i}.$$

Dim: $B_1=\{b_1,b_2,\dots,b_{r_1}\}$, e mostriamo che il primo blocco di colonne di $[Q]_{B,B}$ ha questa struttura. In modo analogo si procede per le altre colonne. $[Q(b_1)]_B$ è la prima colonna rispetto a $[Q]_{B,B}$. $Q(b_1)\in Z_1$ perché $b_1\in Z_1$ e $Q(Z_1)\subseteq Z_1$, ma $Z_1=\mathrm{Im}[b_1,b_2,\dots,b_{r_1}]$ e quindi $Q(b_1)=\alpha_1b_1+\alpha_2b_2+\dots+\alpha_{r_1}b_{r_1}$, quindi conosciamo le coordinate di $Q(b_1)$ rispetto alla base B: $[Q(b_1)]_B=[\underbrace{\alpha_1,\alpha_2,\dots,\alpha_{r_1}}_{r_1}\ \underbrace{0\dots 0}_{m-r_1}]^T$. Rimane da dimostrare che $A_i=[Q|_{Z_i}]_{B_i,B_i}$.
$[Q|_{Z_1}]_{B_1,B_1}=[[Q|_{Z_1}(b_1)]_{B_1},\dots[Q|_{Z_1}(b_{r_1})]_{B_1}]=[[Q(b_1)]_{B_1},\dots,[Q(b_{r_1})]_{B_1}]$, considerando la prima colonna $Q(b_1)=\alpha_1b_1+\alpha_2b_2+\dots+\alpha_{r_1}b_{r_1}$ ma quindi $[Q(b_1)]_{B_1}=[\alpha_1,\alpha_2,\dots,\alpha_{r_1}]^T$ che coincide con la prima colonna di A_1, cioè le prime r_1 componenti della prima colonna della matrice complessiva. □

<u>Proprietà</u>: $\sigma(Q)=\sigma(A_1)\cup\sigma(A_2)\cup\dots\cup\sigma(A_\ell)$

Dim: $\sigma(Q) = \sigma(\underbrace{[Q]_{B,B}}_{A})$, $\chi_A(\lambda) = \det(\lambda I - A) = \det \begin{bmatrix} \lambda I - A_1 & & & \\ & \lambda I - A_2 & & \\ & & \ddots & \\ & & & \lambda I - A_\ell \end{bmatrix} =$

$= \det(\lambda I - A_1)\det(\lambda I - A_2)\cdots\det(\lambda I - A_\ell) =$

$= \chi_{A_1}(\lambda)\chi_{A_2}(\lambda)\cdots\chi_{A_\ell}(\lambda)$.

Quindi $\sigma(A) = \sigma(A_1) \cup \sigma(A_2) \cup \ldots \cup \sigma(A_\ell)$. □

Scritto in modo più astratto $\sigma(Q) = \sigma(Q|_{Z_1}) \cup \sigma(Q|_{Z_2}) \cup \ldots \cup \sigma(Q|_{Z_\ell})$.

Esempio: $A = \begin{bmatrix} 1 & -1 & -1 & 0 \\ 2 & 0 & -1 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ $Z_1 = \mathrm{Im}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$, $Z_2 = \mathrm{Im}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 0 \\ 0 & 0 \end{bmatrix}$, $Z_3 = \mathrm{Im}\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

Vogliamo far vedere che questi tre spazi sono invarianti rispetto alla matrice A e calcolare la matrice diagonale che rappresenta la trasformazione lineare come nella proposizione precedente. Inoltre, per prima cosa dobbiamo mostrare che $Z_1 \oplus Z_2 \oplus Z_3 = \mathbb{C}^4$

$Z_1 + Z_2 + Z_3 = \mathrm{Im} \underbrace{\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ -1 & -1 & 0 & 0 \\ -1 & 0 & 0 & -1 \end{bmatrix}}_{M}$ Si può far vedere che il rango di questa matrice è 4, ad esempio facendo vedere che il determinante è non nullo e che quindi: $\mathrm{Im}\, M = \mathbb{C}^4$. Inoltre $Z_2 + Z_3 = Z_2 \oplus Z_3$ poiché i vettori sono indipendenti, allo stesso modo $Z_1 + (Z_2 \oplus Z_3) =$

$= Z_1 \oplus Z_2 \oplus Z_3 = \mathbb{C}^4$. Ora verifichiamo che Z_i sono invarianti.

$Q(Z_1) \subseteq Z_1 \iff A(Z_1) = A(\mathrm{Im}[1,1,-1,-1]^T) = \mathrm{Im}\, A[1,1,-1,-1]^T = \mathrm{Im}[1,1,-1,1] \in Z_1$

$A(Z_2) = \mathrm{Im}\, A \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}^T = \mathrm{Im}\begin{bmatrix} 1 & -1 \\ 3 & 0 \\ -1 & 1 \\ 0 & 0 \end{bmatrix} \subseteq Z_2$. che sono combinazioni lineari di vettori di Z_2

$A(Z_3) = \mathrm{Im}\, A[1\ 0\ 0\ -1]^T = \mathrm{Im}[1\ 0\ 0\ -1] \subseteq Z_3$.

$$B = \Bigg\{ \underbrace{\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}}_{Z_1}, \underbrace{\overset{2^\circ \text{col.}}{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}}, \overset{1^\circ - 2^\circ}{\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}}}_{Z_2}, \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}}_{Z_3} \Bigg\} \quad [Q]_{B,B} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 \end{bmatrix}^{-1} A \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 \end{bmatrix} =$$

cerchiamo una combinazione delle colonne per trovare la base canonica

$$= \begin{bmatrix} -1 & 0 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 2 & 1 \\ 1 & 0 & 3 & 0 \\ -1 & 1 & -2 & 0 \\ -1 & 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} \overset{A_1}{1} & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}_{A_3}^{A_2}$$

la struttura ottenuta è quella desiderata. $A_1 = [Q|_{Z_1}]_{\{b_1\}}$ e così via,

$\sigma(A) = \sigma(A_1) \cup \sigma(A_2) \cup \sigma(A_3) = \sigma([1]) \cup \sigma\left(\begin{bmatrix} 0 & 3 \\ -1 & 2 \end{bmatrix}\right) \cup \sigma([1])$

$\chi_{A_2}(\lambda) = \det(\lambda I - A_2) = \det\begin{pmatrix}\lambda & -3\\ 1 & \lambda-2\end{pmatrix} = \lambda(\lambda-2)+3 = \lambda^2-2\lambda+3$, $\lambda = 1\pm\sqrt{1-3} = 1\pm\sqrt{2}j$

$\sigma(A) = \{1, 1+\sqrt{2}j, 1-\sqrt{2}j\}$.

<u>Proprietà</u> (<u>Disuguaglianza del rango di Silvester</u>): Sia $A\in\mathbb{C}^{n\times m}$, $B\in\mathbb{C}^{m\times p}$, allora
$\operatorname{rank} AB \geq \operatorname{rank} A + \operatorname{rank} B - m$.

Dim: $a:\mathbb{C}^m\to\mathbb{C}^n$, $a(x)=Ax$, $a|_{\operatorname{Im}B}: \operatorname{Im}B\to\mathbb{C}^n$, $\operatorname{Ker} A = \operatorname{Ker} a \supset \operatorname{Ker} a|_{\operatorname{Im}B}$

Applichiamo il teorema nullità + rango: $\operatorname{rank} A + \dim\operatorname{Ker} A = m$, inoltre
$\dim\operatorname{Ker}A = \dim\operatorname{Ker} a \geq \dim\operatorname{Ker} a|_{\operatorname{Im}B}$. $\dim\operatorname{Ker} a|_{\operatorname{Im}B} + \operatorname{rank} a|_{\operatorname{Im}B} = \operatorname{rank} B$
$\operatorname{rank} a|_{\operatorname{Im}B} \overset{\text{x def.}}{=} \dim\operatorname{Im} a|_{\operatorname{Im}B} = \dim A(\operatorname{Im}B) = \dim\operatorname{Im}(AB) = \operatorname{rank} AB$
Quindi $\dim\operatorname{Ker} a|_{\operatorname{Im}B} = \operatorname{rank}B - \operatorname{rank}AB$.
$\operatorname{rank} A = m - \dim\operatorname{Ker}A \leq m - \dim\operatorname{Ker} a|_{\operatorname{Im}B} = m-(\operatorname{rank}B + \operatorname{rank}AB)$
ovvero $\operatorname{rank}AB \geq \operatorname{rank}A + \operatorname{rank}B - m$. □

<u>Proprietà</u>: Se $A\in\mathbb{C}^{n\times m}$, $B\in\mathbb{C}^{m\times p}$ allora $\operatorname{rank}AB \leq \min\{\operatorname{rank}A, \operatorname{rank}B\}$

Dim: $\operatorname{Im}(AB) = A(\operatorname{Im}B) \subset A(\mathbb{C}^m) = \operatorname{Im}A \Rightarrow \operatorname{rank}AB \leq \operatorname{rank}A$. Ora basta notare che
$\operatorname{rank}(AB) = \operatorname{rank}(B^TA^T)$ e quindi $\operatorname{rank}B^TA^T \leq \operatorname{rank}B^T = \operatorname{rank}B$. □

Quindi $\operatorname{rank}A + \operatorname{rank}B - m \leq \operatorname{rank}AB \leq \min\{\operatorname{rank}A, \operatorname{rank}B\}$

Esempio: $A = \begin{bmatrix}1&0&1&1\\1&2&1&1\\1&0&0&0\\0&1&0&0\end{bmatrix}$, $B = \begin{bmatrix}1&0&1\\-1&1&0\\0&1&1\\1&0&1\end{bmatrix}$ $\operatorname{rank}A = 3$ $\operatorname{rank}B = 2$ $\operatorname{rank}AB \geq 3+2-4 = 1$ $\operatorname{rank}AB \leq \min\{3,2\} = 2$

$1^a+2^a=3^a$

Complementi di geometria: Polinomio minimo, autospazi generalizzati

Def: Sia $p(x) = a_0 + a_1x + \ldots + a_nx^n$, $a_0, a_1, \ldots, a_n \in \mathbb{C}$ e definiamo $p: \mathbb{C}^{m\times m}\to\mathbb{C}^{m\times m}$, tale che se $A\in\mathbb{C}^{m\times m}$, allora $p(A) = a_0I + a_1A + a_2A^2 + \ldots + a_nA^n$

Esempio: $p(x) = 1+2x+x^2$, $A = \begin{bmatrix}1&0&1\\0&1&1\\1&0&2\end{bmatrix}$, $p(A) = 1\cdot I + 2A + A^2 =$

$$= \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} + \begin{bmatrix}2&0&2\\0&2&2\\2&0&4\end{bmatrix} + \begin{bmatrix}2&0&3\\1&1&3\\3&0&5\end{bmatrix} = \begin{bmatrix}5&0&5\\1&4&5\\5&0&10\end{bmatrix}$$

Ricordiamo che $\mathbb{C}^{m\times m}$ è uno spazio vettoriale su $\mathbb{C}$ e $\dim\mathbb{C}^{m\times m} = m^2$

e chiama base di $\mathbb{C}^{2\times 2} = \left\{\begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}0&1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}\right\}$ nello stesso modo si può costruire una base per $\mathbb{C}^{m\times n}$

<u>Proprietà</u>: Se $A\in\mathbb{C}^{n\times n}$ esiste un polinomio non nullo p tale che $p(A)=0$

Dim: $\mathbb{C}^{n\times n}$ sono uno spazio vettoriale con $\dim \mathbb{C}^{n\times n} = n^2$. $A^0=I, A^1, A^2, \dots, A^{(n^2)}$ sono n^2+1 matrici quindi queste matrici non possono essere tutte linearmente indipendenti, quindi esisteranno coefficienti tali che $\alpha_0 I+\alpha_1 A^1+\dots+\alpha_{n^2}A^{(n^2)}=0$ con $\alpha_0,\alpha_1,\dots,\alpha_{n^2}$ non tutti nulli. Quindi se consideriamo $p(x)=\alpha_0+\alpha_1 x+\dots+\alpha_{n^2}x^{(n^2)}$, p è tale che $p(A)=0$ □

<u>Def</u>: Il polinomio p è monico se $p(x)=a_0+a_1x+\dots+a_{n-1}x^{n-1}+x^n$.

<u>Def</u>: Se $A\in\mathbb{C}^{n\times n}$ il polinomio minimo di A è il polinomio monico μ_A di grado più basso tale che $\mu_A(A)=0$.

Esempio: $A=\begin{bmatrix}1&0&0\\1&1&1\\1&0&2\end{bmatrix}$, $A^0=I$, A è l.i. da I, $A^2=\begin{bmatrix}1&0&0\\3&1&3\\3&0&4\end{bmatrix}$ verifichiamo se A^2 è l.i. da I e A. Facciamo la riduzione di Gauss e scriviamo nelle coordinate della base canonica le matrici.

$$B=\left\{\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}, \begin{bmatrix}0&0&0\\1&0&0\\0&0&0\end{bmatrix}, \dots\right\} \quad [A]_B=\begin{bmatrix}1\\1\\1\\0\\1\\0\\0\\1\\2\end{bmatrix}$$

questo procedimento si chiama vettorizzazione e abbiamo incolonnato le colonne della matrice.
Quindi ora ci chiediamo se A^2 è l.i. rispetto ad I e A.

$$M=[[A^0]_B,[A^1]_B,[A^2]_B]=\begin{bmatrix}1&1&1\\0&1&3\\0&1&3\\0&0&0\\1&1&1\\0&0&0\\0&0&0\\0&1&3\\1&2&4\end{bmatrix}\to\begin{bmatrix}1&1&1\\0&1&3\\0&1&3\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&1&3\\0&1&3\end{bmatrix}\to\begin{bmatrix}1&1&1\\0&1&3\\ & 0 & \end{bmatrix}$$

Quindi A^2 dipende linearmente da I e A. Il polinomio minimo avrà grado 2.

$p(A)=\alpha_0 I+\alpha_1 A+\alpha_2 A^2=0 \quad \alpha_0[I]_B+\alpha_1[A]_B+\alpha_2[A^2]_B=0$

$\underbrace{[[I]_B,[A]_B,[A^2]_B]}_{M}[\alpha_0,\alpha_1,\alpha_2]^T=0$ ma quindi stiamo cercando un elemento del Kernel della matrice M, sappiamo che il $\operatorname{Ker} M$ coincide con quello della sua riduzione di Gauss: $\operatorname{Ker} M=\operatorname{Span}[2,-3,1]^T$, quindi $\alpha_0=2$, $\alpha_1=-3$, $\alpha_2=1$

$2A^0-3A+A^2=0$ quindi $\mu_A(x)=2-3x+x^2=(x-2)(x-1)$.

Calcoliamo anche il polinomio caratteristico $\chi_A(\lambda)=\det(\lambda I-A)=\det\begin{bmatrix}\lambda-1&0&0\\-1&\lambda-1&-1\\-1&0&\lambda-2\end{bmatrix}=$ $=(\lambda-1)^2(\lambda-2)$, $\chi_A(\lambda)$ e $\mu_A(x)$ hanno le stesse radici, la differenza è che la radice 1 ha molteplicità 2, il polinomio minimo è un fattore del polinomio caratteristico.

Proprietà: Il polinomio minimo è unico

Dim: $A \in \mathbb{C}^{m \times m}$ assumiamo per assurdo $\exists p_1, p_2$ polinomi monici tali che $p_1 \neq p_2$, $gr(p_1) = gr(p_2)$ e $p_1(A) = p_2(A) = 0$. Consideriamo $(p_1 - p_2)(A) = 0 = p_1(A) - p_2(A)$ quindi abbiamo un terzo polinomio che annulla A ma questo è tale che $gr(p_1 - p_2) = gr(p_1) - 1$ ma questo è assurdo perché contraddice il fatto che p_1 e p_2 sono polinomi minimi. $\lightning$ $\square$

Proprietà: Le radici del polinomio minimo μ_A sono autovalori di A.

Dim: Per assurdo $\exists \bar{\lambda} \in \mathbb{C}$ tale che $\mu_A(\bar{\lambda}) = 0$ e $\bar{\lambda} \notin \sigma(A)$.

$\mu_A(\lambda) = (\lambda - \bar{\lambda}) p(\lambda)$ perché $\bar{\lambda}$ è una radice di μ_A. Inoltre $\mu_A(A) = (A - \bar{\lambda} I) p(A) = 0$ poiché μ_A è il polinomio minimo. Poiché $\bar{\lambda}$ non è autovalore $\det(A - \bar{\lambda} I) \neq 0$ e quindi è invertibile quindi moltiplicando ambo i membri per $(A - \bar{\lambda} I)^{-1}$ otteniamo $0 = p(A)$ quindi $p(A)$ si annulla su A ma ha grado inferiore a μ_A e questo è assurdo perché μ_A è il polinomio minimo. $\lightning$ $\square$

Se $\chi_A(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_\ell)^{m_\ell}$ allora $\mu_A(\lambda) = (\lambda - \lambda_1)^{\mu_1} (\lambda - \lambda_2)^{\mu_2} \cdots (\lambda - \lambda_\ell)^{\mu_\ell}$ e $1 \leq \mu_i \leq m_i$, $i = 1, \dots, \ell$. Se valutiamo il polinomio minimo su A otteniamo:
$\mu_A(A) = (A - \lambda_1 I)^{\mu_1} (A - \lambda_2 I)^{\mu_2} \cdots (A - \lambda_\ell I)^{\mu_\ell} = 0$ ci interesserà studiare il kernel delle matrici $A - \lambda_i I$: $Ker(A - \lambda_i I)^{\mu_i}$ che è l'autospazio generalizzato associato a λ_i.

Def: Sia $A \in \mathbb{C}^{m \times m}$, $v \in \mathbb{C}^n$, $v \neq 0$ è un autovettore generalizzato di A associato all'autovalore λ se $\exists i \in \mathbb{N}$ tale che $(A - \lambda I)^i v = 0$. L'insieme degli autovettori generalizzati si chiama autospazio generalizzato. V_λ: autospazio generalizzato associato a λ.

Chiaramente se $Av = \lambda v \Rightarrow (A - \lambda I)v = 0 \Rightarrow v \in V_\lambda$, quindi gli autovettori sono anche autovettori generalizzati ma non è vero il contrario, in generale l'autospazio generalizzato è più grande dell'autospazio.

Proprietà: $Ker(A - \lambda I)^i \subseteq Ker(A - \lambda I)^{i+1}$

Dim: Se $v \in Ker(A - \lambda I)^i \Rightarrow (A - \lambda I)^i v = 0 \Rightarrow (A - \lambda I)(A - \lambda I)^i v = 0 \Rightarrow (A - \lambda I)^{i+1} v = 0 \Rightarrow$ $v \in Ker(A - \lambda I)^{i+1}$. $\square$

Proprietà: $Ker(A - \lambda I)^{i+1} = Ker(A - \lambda I)^i \Rightarrow Ker(A - \lambda I)^{i+2} = Ker(A - \lambda I)^{i+1}$.

Dim: Sappiamo già che $Ker(A - \lambda I)^{i+1} \subseteq Ker(A - \lambda I)^{i+2}$, dimostriamo il contrario.
Sia $v \in Ker(A - \lambda I)^{i+2} \Rightarrow (A - \lambda I)^{i+2} v = 0 \Rightarrow (A - \lambda I)^{i+1} (A - \lambda I) v = 0$, quindi $(A - \lambda I) v \in Ker(A - \lambda I)^{i+1} = Ker(A - \lambda I)^i \Rightarrow (A - \lambda I) v \in Ker(A - \lambda I)^i \Rightarrow (A - \lambda I)^{i+1} v = 0$ $\square$

$\bar{I} = \min\{i \in \mathbb{N} : Ker(A - \lambda I)^i = Ker(A - \lambda I)^{i+1}\}$ possiamo scrivere la seguente catena di inclusioni

$Ker(A-\lambda I) \subset Ker(A-\lambda I)^2 \subset \ldots \subset Ker(A-\lambda I)^r = Ker(A-\lambda I)^{r+1} = \ldots$

$V_\lambda = Ker(A-\lambda I)^r$.

Esempio: $A = \begin{bmatrix} -2 & 0 & 0 \\ -5 & 1 & -1 \\ -3 & 0 & 1 \end{bmatrix}$ $\chi_A(\lambda) = det(\lambda I - A) = det \left[\begin{array}{cc|c} \lambda+2 & 0 & 0 \\ 5 & \lambda-1 & 1 \\ \hline 3 & 0 & \lambda-1 \end{array}\right] = (\lambda+2)(\lambda-1)^2$

$\sigma(A) = \{1, -2\}$

$\lambda = 1$ $Ker(A-\lambda I) = Ker \begin{bmatrix} -3 & 0 & 0 \\ -5 & 0 & -1 \\ -3 & 0 & 0 \end{bmatrix} = Im \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $Ker(A-\lambda I)^2 = Ker \begin{bmatrix} 9 & 0 & 0 \\ 18 & 0 & 0 \\ 9 & 0 & 0 \end{bmatrix} = Im \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

↳ ha rango 2 quindi ha dim 1

la 1ª colonna è nulla

↳ ha rango 1, quindi ha dim 2

$Ker(A-\lambda I)^3 = Ker \begin{bmatrix} 81 & 0 & 0 \\ 162 & 0 & 0 \\ 81 & 0 & 0 \end{bmatrix} = Im \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = Ker(A-\lambda I)^2 = V_1$

$\lambda = -2$ $Ker(A+2I) = Ker \begin{bmatrix} 0 & 0 & 0 \\ -5 & 3 & -1 \\ -3 & 0 & 3 \end{bmatrix} = Im \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $Ker(A+2I)^2 = Ker \begin{bmatrix} 0 & 0 & 0 \\ -12 & 9 & -6 \\ -9 & 0 & 9 \end{bmatrix} = Im \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = V_2$

Abbiamo trovato una base di vettori generalizzati di $\mathbb{C}^3$, $B = \{V_1, V_2\}$

<u>Lemma di Bezout</u>: Se f, g sono polinomi coprimi, cioè non hanno radici in comune, allora esistono polinomi a, b tali che $af + bg = 1$.

Esempio: $f(\lambda) = \lambda - 1$, $g(\lambda) = (\lambda-2)\lambda$, f e g sono coprimi $a(\lambda) = a_1\lambda + a_0$, $b(\lambda) = b_0$

$(a_1\lambda + a_0)(\lambda - 1) + b_0(\lambda-2)\lambda = 1 \iff \lambda^2(a_1 + b_0) + \lambda(a_0 - a_1 - 2b_0) + (-a_0 - 1) = 0$

$\begin{cases} a_1 + b_0 = 0 \\ a_0 - a_1 - 2b_0 = 0 \\ -a_0 - 1 = 0 \end{cases} \iff \begin{cases} a_0 = 1 \\ b_0 = -a_1 \\ -1 - a_1 + 2a_1 = 0 \end{cases} \iff \begin{cases} a_1 = 1 \\ a_0 = -1 \\ b_0 = -1 \end{cases}$ $a(\lambda) = \lambda - 1$, $b(\lambda) = -1$

[Vedi procedimento analitico dei regolatori di FCA]

<u>Proprietà</u>: Se $A \in \mathbb{C}^{m \times m}$, f, g sono polinomi coprimi allora $Ker\, f(A) \cap Ker\, g(A) = \{0\}$.

Dim: Dal lemma di Bezout $\exists a, b$ polinomi tali che $af + bg = 1$, se applichiamo questi polinomi ad A otteniamo $a(A)f(A) + b(A)g(A) = I$. Per assurdo $\exists v \neq 0 : v \in Ker\, f(A) \wedge v \in Ker\, g(A) \Rightarrow f(A)v = g(A)v = 0$ da cui $a(A)\underbrace{f(A)v}_{0} + b(A)\underbrace{g(A)v}_{0} = v$ da cui $v = 0$ ma questo è in contraddizione con l'ipotesi $v \neq 0$. □

<u>Proprietà</u>: Se $A \in \mathbb{C}^{m \times m}$, f, g polinomi coprimi, allora $Ker\, f(A)g(A) = Ker\, f(A) \oplus Ker\, g(A)$.

Dim: $Ker\, f(A) + Ker\, g(A) = Ker\, f(A) \oplus Ker\, g(A)$ perché nella proprietà precedente abbiamo fatto vedere che la loro intersezione è $\{0\}$. Dobbiamo dimostrare due cose:

1. $Ker\, f(A) \oplus Ker\, g(A) \subseteq Ker\, f(A)g(A)$
2. $dim\, Ker\, f(A) \oplus Ker\, g(A) = dim\, Ker\, f(A)g(A)$

1. Sia $v \in \mathrm{Ker}\, f(A) \oplus \mathrm{Ker}\, g(A) \Rightarrow \exists\, v_f, v_g : v = v_f + v_g,\ f(A)v_f = 0,\ g(A)v_g = 0$
Consideriamo $f(A)g(A)v = f(A)g(A)(v_f + v_g) = f(A)g(A)v_f + f(A)g(A)v_g$ il prodotto di polinomi è commutativo quindi $g(A)\underbrace{f(A)v_f}_{=0} + f(A)\underbrace{g(A)v_g}_{=0} = 0$ ma questo significa che $v \in \mathrm{Ker}\, f(A)g(A)$.
2. $\dim \mathrm{Ker}\, f(A)g(A) = n - \mathrm{rank}\, f(A)g(A)$ per il teorema nullità + rango, poi per la disuguaglianza di Silvester abbiamo che $\mathrm{rank}\, f(A)g(A) \geq \mathrm{rank}\, f(A) + \mathrm{rank}\, g(A) - n$
Quindi $n - \mathrm{rank}\, f(A)g(A) \leq n - \mathrm{rank}\, f(A) + n - \mathrm{rank}\, g(A) = \dim \mathrm{Ker}\, f(A) + \dim \mathrm{Ker}\, g(A)$ ma per il teorema di Grassmann $= \dim \mathrm{Ker}\, f(A) \oplus \mathrm{Ker}\, g(A)$. □

<u>Teorema di decomposizione primaria</u>: Sia $A \in \mathbb{C}^{n\times n}$ con polinomio minimo $\mu_A(\lambda) = (\lambda-\lambda_1)^{\mu_1}(\lambda-\lambda_2)^{\mu_2} \cdots (\lambda-\lambda_\ell)^{\mu_\ell}$ e siano $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_\ell}$ gli autospazi generalizzati di A. Allora:
1. $V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_\ell} = \mathbb{C}^n$
2. $V_{\lambda_i} = \mathrm{Ker}\,(A-\lambda_i I)^{\mu_i}$, $i = 1, \dots, \ell$.

Dim: $\mu_A(A) = (A-\lambda_1 I)^{\mu_1} \cdots (A-\lambda_\ell I)^{\mu_\ell}$ che sono tutti polinomi coprimi, quindi $\mathrm{Ker}\, \mu_A(A) = \mathrm{Ker}\,(A-\lambda_1 I)^{\mu_1} \oplus \dots \oplus \mathrm{Ker}\,(A-\lambda_\ell I)^{\mu_\ell}$, però sappiamo che $\mu_A(A) = 0$ quindi $\mathrm{Ker}\, \mu_A(A) = \mathbb{C}^n$, inoltre se chiamiamo $d_i = \dim \mathrm{Ker}\,(A-\lambda_i I)^{\mu_i}$, $i = 1, \dots, \ell$
$n = d_1 + d_2 + \dots + d_\ell$. Rimane da mostrare che per $i = 1, \dots, \ell$ $V_{\lambda_i} = \mathrm{Ker}\,(A-\lambda_i I)^{\mu_i}$
$i = 1$, $V_{\lambda_1} = \mathrm{Ker}\,(A-\lambda_1 I)^{\mu_1}$, ricordiamo che $V_{\lambda_1} = \mathrm{Ker}\,(A-\lambda_1 I)^{\bar{i}}$ dove $\bar{i} = \min \{ i \in \mathbb{N} : \mathrm{Ker}\,(A-\lambda_1 I)^i = \mathrm{Ker}\,(A-\lambda_1 I)^{i+1} \}$. Dimostriamo che $\bar{i} = \mu_1$.
Per assurdo $\bar{i} > \mu_1$ oppure $\bar{i} < \mu_1$.
1. $\bar{i} > \mu_1$, $\dim \mathrm{Ker}\,(A-\lambda_1 I)^{\mu_1} < \dim \mathrm{Ker}\,(A-\lambda_1 I)^{\mu_1+1}$, consideriamo quindi $\dim \mathrm{Ker}\,(A-\lambda_1 I)^{\mu_1+1}(A-\lambda_2)^{\mu_2} \cdots (A-\lambda_\ell)^{\mu_\ell} = \dim \mathrm{Ker}\,(A-\lambda_1 I)^{\mu_1+1} + \dots + \dim \mathrm{Ker}\,(A-\lambda_\ell I)^{\mu_\ell}$
$> d_1 + d_2 + \dots + d_\ell = n$ ma questo è assurdo perché la dimensione del kernel non può essere maggiore della dimensione dello spazio $\lightning$
2. $\bar{i} < \mu_1$, $\mathrm{Ker}\,(A-\lambda_1 I)^{\bar{i}} = \mathrm{Ker}\,(A-\lambda_1 I)^{\mu_1}$, $\dim \mathrm{Ker}\,(A-\lambda_1 I)^{\bar{i}}(A-\lambda_2 I)^{\mu_2} \cdots (A-\lambda_\ell I)^{\mu_\ell} =$
$= \dim \mathrm{Ker}\,(A-\lambda_1 I)^{\bar{i}} + \dim \mathrm{Ker}\,(A-\lambda_2 I)^{\mu_2} + \dots + \dim \mathrm{Ker}\,(A-\lambda_\ell I)^{\mu_\ell} = d_1 + d_2 + \dots + d_\ell = n$
Ma se definiamo $p(\lambda) = (\lambda-\lambda_1)^{\bar{i}}(\lambda-\lambda_2)^{\mu_2} \cdots (\lambda-\lambda_\ell)^{\mu_\ell}$, $\dim \mathrm{Ker}\, p(A) = n$ ma per il teorema nullità più rango $p(A) = 0$ ma $gr(p) < g(\mu_A)$ con μ_A pol. minimo. $\lightning$ □

Complementi di geometria: Autospazi generalizzati, teorema di Hamilton-Cayley

Ricordiamo la definizione di sottospazio invariante $Z \in \mathbb{C}^n$, $A \in \mathbb{C}^{n\times n}$, $A(Z) \subseteq Z$, è possibile trovare una famiglia di spazi invarianti che sono dati dal kernel di un generico polinomio nella matrice A.

Proprietà: Se $A \in \mathbb{C}^{n\times n}$, p è un polinomio, allora $A(\mathrm{Ker}\, p(A)) \subset \mathrm{Ker}\, p(A)$, cioè il $\mathrm{Ker}\, p(A)$ è invariante rispetto ad A.

Dim: Se $v \in Ker\, p(A) \Rightarrow p(A)v = 0 \Rightarrow A p(A) v = 0 \Leftrightarrow p(A) A v = 0$ perché $A p(A) = \lambda p(\lambda)|_{\lambda=A}$ e $p(A)A = p(\lambda)\lambda|_{\lambda=A}$ ma $\lambda p(\lambda) = p(\lambda)\lambda$, $\Leftrightarrow Av \in Ker\, p(A)$, quindi il $Ker\, p(A)$ è invariante rispetto ad A. $\square$

Conseguenza $A(Ker(A-\lambda I)) \subseteq Ker(A-\lambda I)$, stessa cosa per gli autospazi generalizzati.
$A(Ker(A-\lambda I)^i) \subseteq Ker(A-\lambda I)^i$

Ricordiamo il teorema di decomposizione primaria per cui se $A \in \mathbb{C}^{m\times m}$, $\chi_A(\lambda) = (\lambda-\lambda_1)^{m_1} \cdots (\lambda-\lambda_\ell)^{m_\ell}$ $V_{\lambda_1} \oplus \dots \oplus V_{\lambda_\ell} = \mathbb{C}^m$, inoltre ora sappiamo che $A(V_{\lambda_i}) \subseteq V_{\lambda_i}$, $i=1,\dots,\ell$. Per un risultato visto in qualche lezione fa $B = B_1 \cup B_2 \cup \dots \cup B_\ell$ con B_i base di V_{λ_i}, $i=1,\dots,\ell$. Se consideriamo $\mathcal{a}: \mathbb{C}^m \to \mathbb{C}^m$ con $\mathcal{a}(x) = Ax$

$$[\mathcal{a}]_{B,B} \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_\ell \end{bmatrix}, \quad [\mathcal{a}|_{V_{\lambda_i}}]_{B_i,B_i} = A_i, \quad [\mathcal{a}]_{B,B} = [b_1,\dots b_m]^{-1} A [b_1,\dots,b_m]$$

Questo procedimento può essere visto come una generalizzazione della diagonalizzazione e questo procedimento è sempre possibile.

<u>Proprietà</u>: $\sigma(\mathcal{a}|_{V_\lambda}) = \{\lambda\}$.

Dim: $\mathcal{a}|_{V_\lambda}: V_\lambda \to V_\lambda$ ben definito perché V_λ è invariante su A. Per assurdo $\exists \mu \neq \lambda : \mu \in \sigma(\mathcal{a}|_{V_\lambda})$ $\Rightarrow \exists w \neq 0: \mathcal{a}|_{V_\lambda}(w) = \mu w \Rightarrow Aw = \mu w \Rightarrow w \in V_\mu$, ma $w \in V_\lambda \Rightarrow w \in V_\mu \cap V_\lambda$ ma $V_\mu \cap V_\lambda = \{0\}$ con $\mu \neq \lambda \Rightarrow w = 0$ $\lightning$ $\square$

<u>Proprietà</u>: Se $A \in \mathbb{C}^{m\times m}$, $\chi_A(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2} \cdots (\lambda-\lambda_\ell)^{m_\ell}$ allora $\dim V_{\lambda_i} = m_i$, $i = 1,\dots,\ell$ (cioè la dimensione dell'autospazio generalizzato associato a λ_i coincide con la molteplicità algebrica dell'autovalore).

Dim: $d_i := \dim V_{\lambda_i}$, $i=1,\dots,\ell$, ricordandoci che il polinomio caratteristico di A è uguale al polinomio caratteristico di $[\mathcal{a}]_{B,B}$ diagonale a blocchi perché sono collegate da un cambio di coordinate, inoltre il polinomio caratteristico di $[\mathcal{a}]_{B,B}$ è il prodotto dei polinomi caratteristici di A_i.

$$\chi_A(\lambda) = \det \begin{bmatrix} \lambda I - A_1 & & 0 \\ & \lambda I - A_2 & \\ 0 & & \ddots \ \lambda I - A_\ell \end{bmatrix} = \det(\lambda I - A_1) \cdots \det(\lambda I - A_\ell) = \chi_{A_1}(\lambda)\chi_{A_2}(\lambda) \cdots \chi_{A_\ell}(\lambda)$$
$$\underbrace{\quad}_{d_1}\ \underbrace{\quad}_{d_2} \cdots \underbrace{\quad}_{d_\ell}$$
$$= (\lambda-\lambda_1)^{d_1}(\lambda-\lambda_2)^{d_2} \cdots (\lambda-\lambda_\ell)^{d_\ell} \text{ quindi}$$
$m_i = d_i$, $i = 1,\dots,\ell$ $\square$

Conseguenza

<u>Proprietà</u>: Se $A \in \mathbb{C}^{m\times m}$ con $\chi_A(\lambda) = (\lambda-\lambda_1)^{m_1} \cdots (\lambda-\lambda_\ell)^{m_\ell}$, $\mu_A(\lambda) = (\lambda-\lambda_1)^{\mu_1} \cdots (\lambda-\lambda_\ell)^{\mu_\ell}$ allora $\mu_i \leq m_i$, $i=1,\dots,\ell$.

Dim: basta farlo per $i=1$. $d_1 = \dim V_{\lambda_1}$, $\mu_1 = \min\{i \in \mathbb{N} : Ker(A-\lambda_1 I)^i = Ker(A-\lambda_1 I)^{i+1}\}$.
il valore massimo che potremo avere di i può al massimo essere d_1 e quindi $\mu_1 \leq d_1 = m_1$

Teorema di Hamilton - Cayley: Se $A \in \mathbb{C}^{m \times m}$ allora $\chi_A(A) = 0$.

Dim: $\exists$ polinomio $p : \chi_A(\lambda) = p(\lambda)\mu_A(\lambda)$ che è una conseguenza della proprietà precedente.
$\chi_A(A) = p(A)\mu_A(A)$ ma per definizione di polinomio minimo $\mu_A(A) = 0$. $\square$

Esempio: $A = \begin{bmatrix} 3 & 1 \\ -4 & -1 \end{bmatrix}$ $\chi_A(\lambda) = \det\begin{bmatrix} \lambda-3 & -1 \\ +4 & \lambda+1 \end{bmatrix} = (\lambda-3)(\lambda+1)+4 = \lambda^2-2\lambda-3+4 = \lambda^2-2\lambda+1 = (\lambda-1)^2$

$$\chi_A(A) = A^2-2A+I = \begin{bmatrix} 5 & 2 \\ -8 & -3 \end{bmatrix} - 2\begin{bmatrix} 3 & 1 \\ -4 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Esempio: $A = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 2 & 0 \\ -2 & 2 & 1 \end{bmatrix}$ $\chi_A(\lambda) = \det\begin{bmatrix} \lambda-1 & -2 & 1 \\ 1 & \lambda-2 & 0 \\ 2 & -2 & \lambda-1 \end{bmatrix} = -1(-2(\lambda-1)+2) + (\lambda-2)((\lambda-1)^2-2) =$

$= -(-2\lambda+4) + (\lambda-2)(\lambda^2-2\lambda-1) =$

$= 2(\lambda-2) + (\lambda-2)(\lambda^2-2\lambda-1) =$

$= (\lambda-2)(2+\lambda^2-2\lambda-1) = (\lambda-2)(\lambda^2-2\lambda+1) = (\lambda-2)(\lambda-1)^2$. $\sigma(A) = \{2, 1\}$

$\lambda = 1$ $Ker(A-I) = Ker\begin{bmatrix} 0 & 2 & -1 \\ -1 & 1 & 0 \\ -2 & 2 & 0 \end{bmatrix} = Im\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ non abbiamo raggiunto la dimensione dell'autospazio che è pari alla molteplicità algebrica che in questo caso è 2, quindi

rango 2 quindi dimensione 1

(2) (1) (2)

abbiamo fare ancora un passo e qui siamo sicuri che sarà l'ultimo passo perché la dimensione dovrà per forza crescere ma poi raggiungiamo 2 e quindi ci fermiamo.

$$Ker(A-I)^2 = Ker\begin{bmatrix} 0 & 0 & 0 \\ -1 & -1 & 1 \\ -2 & -2 & 2 \end{bmatrix} = Im\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \end{bmatrix} = Im \begin{array}{c} \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} \\ \begin{array}{cc} b_1 & b_2 \end{array} \end{array}$$

rank = 1
dim = 2

$$\lambda = 2 \quad Ker(A-2I) = Ker\begin{bmatrix} -1 & 2 & -1 \\ -1 & 0 & 0 \\ -2 & 2 & -1 \end{bmatrix} = Im \begin{array}{c} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \\ b_3 \end{array} = V_2 \quad B = \{b_1, b_2, b_3\}$$

rango 2
dim 1

(0) (1) (2)

per colonne | per riga

Calcoliamo ora $[a]_{B,B} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 & 2 & -1 \\ -1 & 2 & 0 \\ -2 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & -1 \\ -1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & 0 \\ -3 & 2 & 2 \\ -4 & 3 & 4 \end{bmatrix} =$

$= \left[\begin{array}{cc|c} -1 & 1 & 0 \\ -4 & 3 & 0 \\ \hline 0 & 0 & 2 \end{array}\right]$ Abbiamo ottenuto la struttura a blocchi desiderata e $A_1 = \begin{bmatrix} -1 & 1 \\ -4 & 3 \end{bmatrix} = [a|_{V_1}]_{B_1,B_1}$

e $A_2 = [2] = [a|_{V_2}]_{B_2,B_2}$. Chiaramente $\sigma(A_2) = \{2\}$.

Calcoliamo per verifica $\sigma(A_1)$: $\chi_{A_1}(\lambda) = \lambda^2-2\lambda+1 = (\lambda-1)^2$ e quindi $\sigma(A_1) = \{1\}$

Esercizio: $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -2 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 0 & 3 \end{bmatrix}$ vogliamo trovare polinomio caratteristico, polinomio minimo autospazi generalizzati e mettere la matrice in forma diagonale a blocchi come abbiamo fatto prima.

$$\chi_A(\lambda) = \det\begin{bmatrix} \lambda-1 & 0 & 0 & 0 \\ 1 & \lambda & 0 & 2 \\ -1 & -1 & \lambda-2 & -2 \\ -1 & -1 & 0 & \lambda-3 \end{bmatrix} = (\lambda-1)(\lambda(\lambda-2)(\lambda-3)+2(\lambda-2)) = (\lambda-2)(\lambda-1)(\lambda^2-3\lambda+2) =$$
$$= (\lambda-2)(\lambda-1)(\lambda-2)(\lambda-1) = (\lambda-2)^2(\lambda-1)^2 \qquad \sigma(A)=\{2,1\}$$

$$\lambda=1,\ \mathrm{Ker}(A-I) = \underset{\substack{\text{rango } 2\\ \text{dim}=2}}{\mathrm{Ker}} \begin{array}{c} \begin{array}{cccc} (-1) & (1) & & \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & -2 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 0 & 2 \end{bmatrix} \\ \begin{array}{cccc} (3) & & & (4) \end{array} \end{array} = \mathrm{Im}\begin{array}{c}\begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 0 & 0 \\ 0 & -1 \end{bmatrix} \\ \begin{array}{cc} b_1 & b_2 \end{array}\end{array} = V_1$$

e ci fermiamo perché abbiamo già raggiunto la dimensione corretta

$$\lambda=2,\ \mathrm{Ker}(A-2I) = \underset{\substack{\text{rango } 3\\ \text{dim } 1}}{\mathrm{Ker}} \begin{bmatrix} -1 & 0 & 0 & 0 \\ -1 & -2 & 0 & -2 \\ 1 & 1 & 0 & 2 \\ 1 & 1 & 0 & 1 \end{bmatrix} = \mathrm{Im}\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

non abbiamo raggiunto la molteplicità e quindi dobbiamo proseguire

$$\mathrm{Ker}(A-2I)^2 = \underset{\substack{\text{rango } 2\\ \text{dim } 2}}{\mathrm{Ker}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & -1 \end{bmatrix} = \mathrm{Im}\begin{array}{c}\begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & -1 \end{bmatrix} \\ \begin{array}{cc} b_3 & b_4 \end{array}\end{array} = V_2. \quad \mu_A(\lambda) = \underbrace{(\lambda-1)}_{\text{raggiunto al 1° step}}\underbrace{(\lambda-2)^2}_{\text{raggiunto al 2° step}}$$

$$[a]_{B,B} = [b_1 b_2 b_3 b_4]^{-1} A [b_1 b_2 b_3 b_4]$$ oppure possiamo sfruttare la struttura di $[a]_{B,B} = \left[\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right]$

$$A_1 = [a|_{V_1}]_{B_1} = [[a(b_1)]_{B_1}, [a(b_2)]_{B_1}] \quad B_1 = \left\{\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ -1 \end{bmatrix}\right\} \quad a(b_1) = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} = 1\cdot b_1$$

$$a(b_2) = Ab_2 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ -1 \end{bmatrix} = 1\cdot b_2 \qquad \text{Quindi } A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A_2 = [[a(b_3)]_{B_2}, [a(b_4)]_{B_2}], \quad B_2 = \left\{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}\right\} \quad a(b_3) = Ab_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \end{bmatrix} = 2\cdot b_3$$

$$a(b_4) = Ab_4 = \begin{bmatrix} 0 \\ 2 \\ -1 \\ -2 \end{bmatrix} = -1\cdot b_3 + 2\cdot b_4 \qquad \text{Quindi } A_2 = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \qquad [a]_{B,B} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

Rivediamo ora cosa sono le norme

Def: Dato uno spazio vettoriale V, una norma su V è una funzione $\|\cdot\|: V \to \mathbb{R}$ tale che

1. $(\forall x \in V)\ \|x\| \geq 0.$
2. $\|x\| = 0 \iff x = 0.$
3. $(\forall x \in V)(\forall \alpha \in \mathbb{C})\ \|\alpha x\| = |\alpha|\,\|x\|.$
4. $(\forall x, y \in V)\ \|x+y\| \leq \|x\| + \|y\|.$

Esempi: $V = \mathbb{R}^n$, $x \in \mathbb{R}^n$, $x = [x_1, x_2, \dots, x_n]^T$, la norma più comune è la norma euclidea

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{x^T x}$$

$\|x\|_1 = |x_1| + |x_2| + \dots + |x_n|$ è la somma dei valori assoluti delle componenti

$\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}$

$V = \mathbb{C}^n$, $x \in \mathbb{C}^n$, $x = [x_1, x_2, \dots, x_n]^T$

$\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2} = \sqrt{x^* \cdot x}$ con x^* che rappresenta x trasposto coniugato

$\|\cdot\|_1$ e $\|\cdot\|_\infty$ sono definite in modo del tutto analogo.

Definiamo ora le norme matriciali indotte dalle norme vettoriali

Def: Sia $\underset{\text{norma vettoriale}}{\|\cdot\|_v}$ una norma su $\mathbb{C}^n$ e $\mathbb{C}^n$, definiamo $\underset{\text{norma matriciale}}{\|\cdot\|_M} : \mathbb{C}^{n \times n} \to \mathbb{R}$ tale che

$$(\forall A \in \mathbb{C}^{n \times n}) \ \|A\|_M = \sup_{\substack{x \in \mathbb{C}^n \\ x \neq 0}} \frac{\|Ax\|_v}{\|x\|_v}$$

Proprietà: $\|\cdot\|_M$ è una norma su $\mathbb{C}^{n \times n}$

Questa norma può essere interpretata come il massimo rapporto che c'è sulla norma del vettore Ax e il vettore x, quindi se interpretiamo A come un operatore lineare che agisce su x questa è una sorta di massimo guadagno di ampiezza che può essere ottenuto dall'operatore lineare associato ad A.

Proprietà: Se $A \in \mathbb{C}^{n \times n}$, $z \in \mathbb{C}^n$, $\|Az\|_v \leq \|A\|_M \cdot \|z\|_v$ (le norme sono compatibili)

Dim: $\|A\|_M = \sup_{x \neq 0} \frac{\|Ax\|_v}{\|x\|_v} \geq \frac{\|Az\|_v}{\|z\|_v}$ quindi $\|A\|_M \|z\|_v \geq \|Az\|_v$

Vediamo come si calcolano le norme indotte più comuni: $A \in \mathbb{C}^{n \times n}$, $A = (a_{ij})$

$\|A\|_2 = \max\{\sqrt{|\lambda|}, \lambda \in \sigma(A^*A)\}$, A^*A è hermitiana e ha solo autovalori reali, la norma 2 è infatti anche detta norma spettrale.

$$\|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|, \qquad \|A\|_\infty = \max_{i=1,\dots,n} \sum_{j=1}^{n} |a_{ij}|$$

Esempio: $A = \begin{array}{cc} \begin{bmatrix} 1 & 2 & -3 \\ 0 & -1 & 0 \\ -4 & 1 & 0 \end{bmatrix} & \begin{array}{c} 6 \\ 1 \\ 5 \end{array} \\ \begin{array}{ccc} 5 & 4 & 3 \end{array} & \end{array}$ $\quad \|A\|_1 = 5$; $\|A\|_\infty = 6$

$A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$ $\quad \|A\|_2 = \max\{\sqrt{4}, \sqrt{9}\} = 3$.

$A^*A = A^TA = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$