Jordan normal form

Example: The matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable in any field.

Proof: It has eigenvalue 1 of multiplicity two, hence could only be similar to $\boldsymbol{I}_2$. But for any regular $\boldsymbol{R}$: $\boldsymbol{R}^{-1}\boldsymbol{I}_2\boldsymbol{R} = \boldsymbol{I}_2 \neq \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Jordan normal form

Example: The matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable in any field.

Definition: A *Jordan block* is a square matrix of the form

$$\boldsymbol{J}_\lambda = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

Theorem: Every square complex matrix $\boldsymbol{A}$ is similar to a matrix $\boldsymbol{J}$ in the so called *Jordan normal form*

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_{\lambda_1} & & \\ & \ddots & \\ & & \boldsymbol{J}_{\lambda_k} \end{pmatrix}$$

Each Jordan block $\boldsymbol{J}_{\lambda_i}$ corresponds to an eigenvalue λ_i of $\boldsymbol{A}$.
A λ_i may yield several Jordan blocks, indeed of various sizes.

Fact: For each λ, the number of blocks and their sizes are uniquely determined by $\boldsymbol{A}$. Hence the Jordan normal form of $\boldsymbol{A}$ is unique upto a permutation of the Jordan blocks on the diagonal.

Observation: A diagonalizable matrix has Jordan blocks of size one.

Generalized eigenvectors

When $\boldsymbol{A}$ is diagonalizable, i.e. $\boldsymbol{AR} = \boldsymbol{RD}$,

then the columns of $\boldsymbol{R}$ are eigenvectors of $\boldsymbol{A}$.

What can we say about matrices that are not diagonalizable?

Proposition: Let $\boldsymbol{AR} = \boldsymbol{RJ}_\lambda$.

If $\boldsymbol{x}_i$ is the i-th column of $\boldsymbol{R}$, then it satisfies $(\boldsymbol{A} - \lambda \boldsymbol{I})^i \boldsymbol{x}_i = \boldsymbol{0}$.

Proof:

$$
\begin{array}{cccc|cccc}
 & & & & \lambda & 1 & & \\
 & \boldsymbol{RJ}_\lambda & & & & \lambda & \dots & 1 \\
 & & & & & & & \lambda \\
\hline
\boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n & \lambda \boldsymbol{x}_1 & \boldsymbol{x}_1 + \lambda \boldsymbol{x}_2 & \dots & \boldsymbol{x}_{n-1} + \lambda \boldsymbol{x}_n
\end{array}
$$

$$
\begin{array}{lclcl}
\boldsymbol{A}\boldsymbol{x}_1 = \lambda \boldsymbol{x}_1 & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}_1 = \boldsymbol{0} & & \\
\boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{x}_1 + \lambda \boldsymbol{x}_2 & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}_2 = \boldsymbol{x}_1 & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})^2 \boldsymbol{x}_2 = \boldsymbol{0} \\
 & \vdots & & & \\
\boldsymbol{A}\boldsymbol{x}_n = \boldsymbol{x}_{n-1} + \lambda \boldsymbol{x}_n & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}_n = \boldsymbol{x}_{n-1} & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})^n \boldsymbol{x}_n = \boldsymbol{0}
\end{array}
$$

Generalized eigenvectors

When $\boldsymbol{A}$ is diagonalizable, i.e. $\boldsymbol{AR} = \boldsymbol{RD}$,

then the columns of $\boldsymbol{R}$ are eigenvectors of $\boldsymbol{A}$.

What can we say about matrices that are not diagonalizable?

Proposition: Let $\boldsymbol{AR} = \boldsymbol{RJ}_\lambda$.

If $\boldsymbol{x}_i$ is the i-th column of $\boldsymbol{R}$, then it satisfies $(\boldsymbol{A} - \lambda \boldsymbol{I})^i \boldsymbol{x}_i = \mathbf{0}$.

Definition: *Generalized eigenvector* of a matrix $\boldsymbol{A}$ for an eigenvalue λ is any vector $\boldsymbol{x}$ satisfying $(\boldsymbol{A} - \lambda \boldsymbol{I})^i \boldsymbol{x} = \mathbf{0}$ for some $i \in \mathbb{N}$.

They form *chains* $\boldsymbol{x}_k, \dots, \boldsymbol{x}_2, \boldsymbol{x}_1, \mathbf{0}$, where $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}_i = \boldsymbol{x}_{i-1}$.

Analogously, for a linear map f we get $f(\boldsymbol{x}_i) - \lambda \boldsymbol{x}_i = \boldsymbol{x}_{i-1}$.

In another notation: $\boldsymbol{x} \in \ker((\boldsymbol{A} - \lambda \boldsymbol{I})^i)$, or $\boldsymbol{x} \in \ker((f - \lambda id)^i)$.

Theorem: (equivalent version of Jordan's normal form theorem)
Each finitely generated space V over $\mathbb{C}$ and linear $f : V \to V$ has a basis from chains of generalized eigenvectors of the map f.

Note: Also holds for any $\mathbb{K}$, when eigenvalues have algebraic multiplicity $\dim(V)$, i.e. if $p_{[f]_{X,X}}(t)$ decomposes into linear terms.

Example

The matrix $\boldsymbol{A} = \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix}$ is similar to a matrix in the Jordan normal form with two blocks $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, because

$$\boldsymbol{AR} = \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{RJ}$$

$(3, 2, 1)^T$ is an eigenvector for 2, i.e. $(\boldsymbol{A} - 2\boldsymbol{I}_3)(3, 2, 1)^T = \mathbf{0}$ and
$(1, 1, 1)^T$ is an eigenvector for 1, i.e. $(\boldsymbol{A} - 1\boldsymbol{I}_3)(1, 1, 1)^T = \mathbf{0}$.

The middle column of the matrix $\boldsymbol{R}$ however satisfies
$\boldsymbol{A} \cdot (2, 2, 1)^T = (3, 2, 1)^T + 2 \cdot (2, 2, 1)^T \quad \Longrightarrow$
$(\boldsymbol{A} - 2\boldsymbol{I}_3)\ (2, 2, 1)^T = (3, 2, 1)^T \quad \Longrightarrow$
$(\boldsymbol{A} - 2\boldsymbol{I}_3)^2 (2, 2, 1)^T = (\boldsymbol{A} - 2\boldsymbol{I}_3)(3, 2, 1)^T = \mathbf{0}$.

Proof of the theorem — Part 1

By induction on $\dim(V)$. For each eigenvalue λ we introduce the map $g_\lambda(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda\boldsymbol{x}$. We fix some eigenvalue λ arbitrarily.

Since both f and id are linear maps, $g_\lambda = f - \lambda id$ is also linear.

Denote $W = g_\lambda(V)$, the range of the map g_λ.

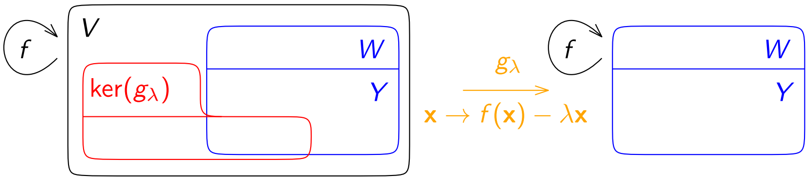
Since g_λ is a linear map, W is a vector space. Indeed W is a *subspace* of V, because $\forall \boldsymbol{x} \in V : g_\lambda(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda\boldsymbol{x} \in V$.

Next, $\dim(W) < \dim(V)$ because the eigenvector $\boldsymbol{u}$ for λ satisfies $g_\lambda(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda\boldsymbol{x} = \mathbf{0}$, i.e. $\dim(\ker(g_\lambda)) \geq 1$ and thus $\dim(V) = \dim(g_\lambda(V)) + \dim(\ker(g_\lambda)) = \dim(W) + \dim(\ker(g_\lambda))$.

The map f can be restricted to W, since for $g_\lambda(\boldsymbol{x}) \in W$ we have $f(g_\lambda(\boldsymbol{x})) = f(f(\boldsymbol{x}) - \lambda\boldsymbol{x}) = f(f(\boldsymbol{x})) - \lambda f(\boldsymbol{x}) = g_\lambda(f(\boldsymbol{x})) \in W$.

According to the inductive hypothesis for f and W, the subspace W has a basis Y from chains of generalized eigenvectors of f.

Example for the first part of the proof



For $[f]_{K,K} = \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix}$ a $\lambda = 2$ is $[g_2]_{K,K} = \begin{pmatrix} -3 & 7 & -5 \\ -2 & 5 & -4 \\ -1 & 3 & -3 \end{pmatrix} \sim\sim \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$

$Z = \{(3,2,1)^T\}$ is a basis of $\ker(g_2)$ so $\dim(W) = 3 - 1 = 2$.

When we extend Z by $\boldsymbol{e}^1, \boldsymbol{e}^2$ to a basis of V, we get $\{g_2(\boldsymbol{e}^1), g_2(\boldsymbol{e}^2)\} = \{(-3,-2,-1)^T, (7,5,3)^T\}$ as a basis of W.

Note that $W \cap \ker(g_2) \neq \emptyset$. This intersection has dimension 1.

There are two chains that form the basis Y of the subspace W: the first is $(3,2,1)^T$ for $\lambda = 2$ and the next is $(1,1,1)^T$ for $\lambda = 1$. (Both have length one, so they contain "ordinary" eigenvectors.)

Proof of theorem — Part 2

Denote $d = \dim(\ker(g_\lambda))$ and $d' = \dim(\ker(g_\lambda) \cap W)$.

Arrange the basis Y into r strings so that the first d' corresponds to λ and others correspond to the other eigenvalues $\lambda', \ldots, \lambda'^{\cdots'}$:

$$
\begin{array}{ccccccccc}
\boldsymbol{y}^1_{k_1} & \xrightarrow{g_\lambda} & \cdots\cdots\cdots & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_2 & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_1 & \xrightarrow{g_\lambda} & \mathbf{0} \\
& \boldsymbol{y}^2_{k_2} \xrightarrow{g_\lambda} & \cdots & \xrightarrow{g_\lambda} & \boldsymbol{y}^2_2 & \xrightarrow{g_\lambda} & \boldsymbol{y}^2_1 & \xrightarrow{g_\lambda} & \mathbf{0} \\
& & & & & & & & \vdots \\
& & \boldsymbol{y}^{d'}_{k_{d'}} & \xrightarrow{g_\lambda} & \cdots & \xrightarrow{g_\lambda} & \boldsymbol{y}^{d'}_1 & \xrightarrow{g_\lambda} & \mathbf{0} \\
& & \boldsymbol{y}^{d'+1}_{k_{d'+1}} & \xrightarrow{g_{\lambda'}} & \cdots & \xrightarrow{g_{\lambda'}} & \boldsymbol{y}^{d'+1}_1 & \xrightarrow{g_{\lambda'}} & \mathbf{0} \\
& & & & & & & & \vdots \\
& & & & & \cdots & \boldsymbol{y}^r_1 & \xrightarrow{g_{\lambda'^{\cdots'}}} & \mathbf{0}
\end{array}
$$

As chains of Y are in W, we can extend each of the first d' chains by some $\boldsymbol{x}^i \in V$ so that $g_\lambda(\boldsymbol{x}^i) = \boldsymbol{y}^i_{k_i}$ for $i \in \{1, \ldots, d'\}$.

The vectors $\boldsymbol{y}^1_1, \ldots, \boldsymbol{y}^{d'}_1$ form the basis of the space $\ker(g_\lambda) \cap W$.

Complete them by $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^{d-d'}$ to a basis of $\ker(g_\lambda)$ (other than Z) and get $d - d'$ new chains of length 1 formed by $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^{d-d'}$.

That yields chains

$$
\begin{array}{ccccccccccccc}
\boldsymbol{x}^1 & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_{k_1} & \xrightarrow{g_\lambda} & \cdots & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_2 & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_1 & \xrightarrow{g_\lambda} & \mathbf{0} \\
 & & & & & & & & & & \vdots \\
 & & \boldsymbol{x}^{d'} & \xrightarrow{g_\lambda} & \boldsymbol{y}^{d'}_{k_{d'}} & \xrightarrow{g_\lambda} & \cdots & \xrightarrow{g_\lambda} & \boldsymbol{y}^{d'}_1 & \xrightarrow{g_\lambda} & \mathbf{0} \\
 & & & & \boldsymbol{y}^{d'+1}_{k_{d'+1}} & \xrightarrow{g_{\lambda'}} & \cdots & \xrightarrow{g_{\lambda'}} & \boldsymbol{y}^{d'+1}_1 & \xrightarrow{g_{\lambda'}} & \mathbf{0} \\
 & & & & & & & & & & \vdots \\
 & & & & & & & \cdots & \boldsymbol{y}^r_1 & \xrightarrow{g_{\lambda'\cdots'}} & \mathbf{0} \\
 & & & & & & & & \boldsymbol{z}^1 & \xrightarrow{g_\lambda} & \mathbf{0} \\
 & & & & & & & & & & \vdots \\
 & & & & & & & & \boldsymbol{z}^{d-d'} & \xrightarrow{g_\lambda} & \mathbf{0}
\end{array}
$$

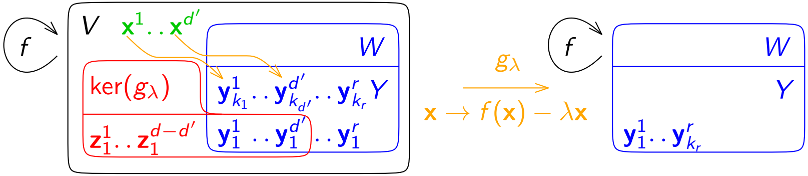
In our example:

$$
\begin{array}{ccccc}
(2,2,1)^T & \xrightarrow{g_2} & (3,2,1)^T & \xrightarrow{g_2} & \mathbf{0} \\
 & & (1,1,1)^T & \xrightarrow{g_1} & \mathbf{0}
\end{array}
$$

We have no $\boldsymbol{z}^i$ because $d = d' = 1$.

We added $d = \dim(\ker(g_\lambda))$ vectors to the basis of W,
so in total we have as many as is the dimension of the space V.

We show that they are linearly independent and therefore they form a basis of the space V.



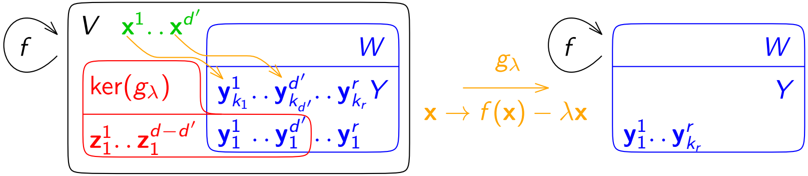
Consider a linear combination $\sum_i a_i \boldsymbol{x}^i + \sum_i b_{i,j} \boldsymbol{y}^i_j + \sum_i c_i \boldsymbol{z}^i = \mathbf{0}$.

Since $\mathbf{0} = g_\lambda(\mathbf{0}) = g_\lambda \left(\sum_i a_i \boldsymbol{x}^i + \sum_{i,j} b_{i,j} \boldsymbol{y}^i_j + \sum_i c_i \boldsymbol{z}^i \right) = \sum_{i,j} b'_{i,j} \boldsymbol{y}^i_j$,
where the vectors $\boldsymbol{y}^i_j$ are linearly independent, we must have $0 = b'_{i,j} =$

$$\begin{cases} a_i & \text{for } i \leq d', j = k_i \\ b_{i,j+1} & \text{for } i \leq d', j < k_i \\ (\lambda^* - \lambda) b_{i,j} & \text{for } i > d', j = k_i \\ (\lambda^* - \lambda) b_{i,j} + b_{i,j+1} & \text{for } i > d', j < k_i \end{cases}$$

where $\lambda^* \neq \lambda$ matches the i-th chain.

It follows from $g_\lambda(\boldsymbol{x}^i) = \boldsymbol{y}^i_{k_i}$ and $g_\lambda(\boldsymbol{y}^i_j) = \boldsymbol{y}^i_{j-1}$ for $i \leq d'$; while for $i > d' : g_\lambda(\boldsymbol{y}^i_1) = f(\boldsymbol{y}^i_1) - \lambda \boldsymbol{y}^i_1 = \lambda^* \boldsymbol{y}^i_1 - \lambda \boldsymbol{y}^i_1 = (\lambda^* - \lambda)\boldsymbol{y}^i_1$ and for $j > 1$ also: $g_\lambda(\boldsymbol{y}^i_j) = f(\boldsymbol{y}^i_j) - \lambda \boldsymbol{y}^i_j = f(\boldsymbol{y}^i_j) - \lambda^* \boldsymbol{y}^i_j + (\lambda^* - \lambda)\boldsymbol{y}^i_j = g_{\lambda^*}(\boldsymbol{y}^i_j) + (\lambda^* - \lambda)\boldsymbol{y}^i_j = \boldsymbol{y}^i_{j-1} + (\lambda^* - \lambda)\boldsymbol{y}^i_j$.



Consider a linear combination $\sum_i a_i \boldsymbol{x}^i + \sum_i b_{i,j} \boldsymbol{y}^i_j + \sum_i c_i \boldsymbol{z}^i = \mathbf{0}$.

Since $\mathbf{0} = g_\lambda(\mathbf{0}) = g_\lambda\left(\sum_i a_i \boldsymbol{x}^i + \sum_{i,j} b_{i,j} \boldsymbol{y}^i_j + \sum_i c_i \boldsymbol{z}^i\right) = \sum_{i,j} b'_{i,j} \boldsymbol{y}^i_j$,

where the vectors $\boldsymbol{y}^i_j$ are linearly independent, we must have $0 = b'_{i,j} = \begin{cases} a_i & \text{for } i \leq d', j = k_i \\ b_{i,j+1} & \text{for } i \leq d', j < k_i \\ (\lambda^* - \lambda) b_{i,j} & \text{for } i > d', j = k_i \\ (\lambda^* - \lambda) b_{i,j} + b_{i,j+1} & \text{for } i > d', j < k_i \end{cases}$

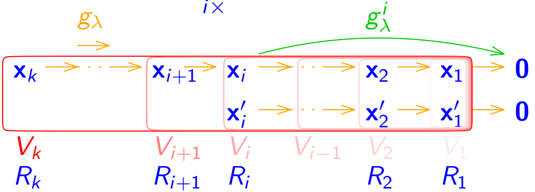
where $\lambda^* \neq \lambda$ matches the i-th chain.

The first case gives: $\forall i : a_i = 0$, the next: $\forall i \leq d', \forall j > 1 : b_{i,j} = 0$ and the other two: $\forall i > d', \forall j : b_{i,j} = 0$. In the combination, only the coefficients $b_{i,1}$ for$i \leq d'$ and c_i remain, but they are also zero, since the vectors $\boldsymbol{y}^1_1, \ldots, \boldsymbol{y}^{d'}_1, \boldsymbol{z}^1, \ldots, \boldsymbol{z}^{d-d'}$ form a basis of $\ker(g_\lambda)$.

Calculation of chains corresponding to λ

Notation: Map $g_\lambda^i = \underbrace{g_\lambda \circ g_\lambda \circ \cdots \circ g_\lambda}_{i\times}$ ... corresponds to $(\boldsymbol{A} - \lambda \boldsymbol{I})^i$

Procedure:



- We determine the sequence of spaces $V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_k$, where $V_i = \ker(g_\lambda^i)$ and $k = \min\{i : \ker(g_\lambda^i) = \ker(g_\lambda^{i+1})\}$.
- We set $R_{k+1} = \emptyset$ and for i from k to 1:
  - calculate the set $g_\lambda(R_{i+1})$

    ... we extend the already establlished chains
  - and extend it by vectors from $V_i \setminus V_{i-1}$ to a linearly independent set R_i of size $\dim(V_i) - \dim(V_{i-1})$

    ... we add to R_i the beginnings of new chains

A Jordan cell of size i corersponds to a chain that begins some $\boldsymbol{x}_i \in R_i \setminus g_\lambda(R_{i+1})$ followed by its images $\boldsymbol{x}_{i-j} = g_\lambda^j(\boldsymbol{x}_i) \in R_{i-j}$.

Example

$$\boldsymbol{A} = \begin{pmatrix} -2 & -3 & 6 & 2 & -3 & -2 & -8 \\ -2 & 0 & 4 & 0 & -1 & -1 & -4 \\ 0 & 1 & 1 & 0 & 3 & -1 & -1 \\ 2 & 2 & -4 & 0 & 4 & 1 & 4 \\ 1 & 0 & -2 & 0 & 1 & 1 & 2 \\ -2 & -3 & 4 & 1 & -4 & 1 & -5 \\ 2 & 3 & -4 & -1 & 5 & 0 & 5 \end{pmatrix}$$

$p_{\boldsymbol{A}}(t) =$

$= t^7 - 6t^6 + 15t^5 - 20t^4 + 15t^3 - 6t^2 + t$

$= t \cdot (t-1)^6$

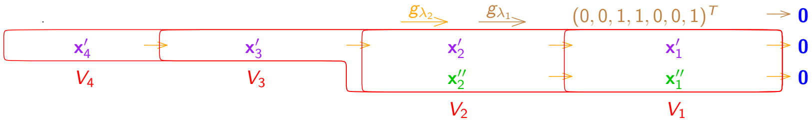
Eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = 1$.

Since the algebraic multiplicity of λ_1 is 1, it has geometric multiplicity 1 as well and it corersponds to a Jordan cell of size 1.

$$\xrightarrow{g_{\lambda_1}} \quad (0,0,1,1,0,0,1)^T \quad \longrightarrow \mathbf{0}$$

We choose an eigenvector $\boldsymbol{x}_1 = (0,0,1,1,0,0,1)^T$ for λ_1.

Example



The matrix

$$\boldsymbol{B} = \boldsymbol{A} - \lambda_2 \boldsymbol{I}_7 = \begin{pmatrix} -3 & -3 & 6 & 2 & -3 & -2 & -8 \\ -2 & -1 & 4 & 0 & -1 & -1 & -4 \\ 0 & 1 & 0 & 0 & 3 & -1 & -1 \\ 2 & 2 & -4 & -1 & 4 & 1 & 4 \\ 1 & 0 & -2 & 0 & 0 & 1 & 2 \\ -2 & -3 & 4 & 1 & -4 & 0 & -5 \\ 2 & 3 & -4 & -1 & 5 & 0 & 4 \end{pmatrix}$$

has rank 5.

$\dim(V_1) = 7 - 5 = 2$.

The eigenvalue $\lambda_2 = 1$ thus corresponds to two Jordan cels, i.e. to two chains.

The chain lengths can be derived from dimensions of $V_2, V_3, \ldots$

$\text{rank}(\boldsymbol{B}^2) = 3 \Rightarrow \dim(V_2) = 4 \Rightarrow$ both chains have length at least 2

$\text{rank}(\boldsymbol{B}^3) = 2 \Rightarrow \dim(V_3) = 5 \Rightarrow$ one lenght is 2 and the other 4.

Jordan normal form is

$$\boldsymbol{J} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Example — calculation of generalized eigenvectors

$$\xrightarrow{g_{\lambda_2}} \quad \xrightarrow{g_{\lambda_1}} \quad (0,0,1,1,0,0,1)^T \rightarrow \mathbf{0}$$

$$\underbrace{(1,0,0,0,0,0,0)^T}_{V_4} \rightarrow \underbrace{(-3,-2,0,2,1,-2,2)^T}_{V_3} \rightarrow (4,1,1,-2,-1,0,-1)^T \rightarrow (-2,0,-1,0,0,0,0)^T \rightarrow \mathbf{0}$$

$$\underbrace{(0,0,0,3,-1,-4,2)^T}_{V_2} \rightarrow \underbrace{(1,-3,-1,-3,0,-3,0)^T}_{V_1} \rightarrow \mathbf{0}$$

Choose e.g. $\boldsymbol{x}_4' = (1,0,0,0,0,0,0)^T \in V_4$, then
$\boldsymbol{x}_3' = g_{\lambda_2}(\boldsymbol{x}_4') = \boldsymbol{B}\boldsymbol{x}_4' = (-3,-2,0,2,1,-2,2)^T \in V_3$ and
$\boldsymbol{x}_2' = g_{\lambda_2}(\boldsymbol{x}_3') = \boldsymbol{B}\boldsymbol{x}_3' = (4,1,1,-2,-1,0,-1)^T \in V_2$.

Choose vector $\boldsymbol{x}_2'' \in V_2 \setminus V_1$ linearly independent on $\boldsymbol{x}_2'$ (we show later how), e.g. $\boldsymbol{x}_2'' = (0,0,0,3,-1,-4,2)^T$.

Now $\boldsymbol{x}_1' = g_{\lambda_2}(\boldsymbol{x}_2') = \boldsymbol{B}\boldsymbol{x}_2' = (-2,0,-1,0,0,0,0)^T \in V_1$
and $\boldsymbol{x}_1'' = g_{\lambda_2}(\boldsymbol{x}_2'') = \boldsymbol{B}\boldsymbol{x}_2'' = (1,-3,-1,-3,0,-3,0)^T \in V_1$.

The desired regular matrix $\boldsymbol{R}$ for $\boldsymbol{AR} = \boldsymbol{RJ}$ is

$$\boldsymbol{R} = \begin{pmatrix} | & | & | & | & | & | & | \\ \boldsymbol{x}_1 & \boldsymbol{x}_1' & \boldsymbol{x}_2' & \boldsymbol{x}_3' & \boldsymbol{x}_4' & \boldsymbol{x}_1'' & \boldsymbol{x}_2'' \\ | & | & | & | & | & | & | \end{pmatrix} = \begin{pmatrix} 0 & -2 & 4 & -3 & 1 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 & -3 & 0 \\ 1 & -1 & 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & -2 & 2 & 0 & -3 & 3 \\ 0 & 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -2 & 0 & -3 & -4 \\ 1 & 0 & -1 & 2 & 0 & 0 & 2 \end{pmatrix}$$

Example — choice of $\boldsymbol{x}_2''$

Calculate the basis of V_2, i.e. of the space $\ker(\boldsymbol{B}^2)$.

$$\boldsymbol{B}^2 = \begin{pmatrix} 4 & 4 & -8 & -2 & 6 & 2 & 10 \\ 1 & 2 & -2 & -1 & 3 & 0 & 3 \\ 1 & -1 & -2 & 0 & -2 & 2 & 3 \\ -2 & -5 & 4 & 2 & -8 & 1 & -5 \\ -1 & -2 & 2 & 1 & -3 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -3 & 2 & 1 & -5 & 1 & -2 \end{pmatrix} \sim\sim \begin{pmatrix} 1 & 0 & -2 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 2 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & -1 & -3 \end{pmatrix} \quad \Rightarrow \quad \ker(\boldsymbol{B}^2) =$$

$$= \mathcal{L}((-2,0,-1,0,0,0,0)^T, (0,2,0,1,-1,0,0)^T, (1,-1,0,-1,0,-1,0)^T, (2,-1,0,-3,0,0,-1)^T)$$

Put the basis row-wise into a matrix and transform it to a echelon form.

$$\begin{pmatrix} -2 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & -1 & 0 & -1 & 0 \\ 2 & -1 & 0 & -3 & 0 & 0 & -1 \end{pmatrix} \sim\sim \begin{pmatrix} 3 & 0 & 0 & 0 & -2 & -5 & 1 \\ 0 & 3 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 3 & 0 & 4 & 10 & -2 \\ 0 & 0 & 0 & 3 & -1 & -4 & 2 \end{pmatrix} = \boldsymbol{M}_1$$

Do the same for the space V_1, where we add $\boldsymbol{x}_2'$ to the basis.

$$\boldsymbol{B} = \begin{pmatrix} -3 & -3 & 6 & 2 & -3 & -2 & -8 \\ -2 & -1 & 4 & 0 & -1 & -1 & -4 \\ 0 & 1 & 0 & 0 & 3 & -1 & -1 \\ 2 & 2 & -4 & -1 & 4 & 1 & 4 \\ 1 & 0 & -2 & 0 & 0 & 1 & 2 \\ -2 & -3 & 4 & 1 & -4 & 0 & -5 \\ 2 & 3 & -4 & -1 & 5 & 0 & 4 \end{pmatrix} \sim\sim \begin{pmatrix} 1 & 0 & -2 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \Rightarrow$$

$$\ker(\boldsymbol{B}) = \mathcal{L}((2,0,-1,0,0,0,0)^T, (1,-1,0,-1,0,-1,0)^T$$

$$\begin{pmatrix} -2 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & -1 & 0 & 0 \\ 4 & 1 & 1 & -2 & -1 & 0 & -1 \end{pmatrix} \sim\sim \begin{pmatrix} 3 & 0 & 0 & -3 & -1 & -1 & -1 \\ 0 & 3 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 3 & 6 & 2 & 2 & 2 \end{pmatrix} = \boldsymbol{M}_2$$

The row of $\boldsymbol{M}_1$ with pivot in another column, that are pivots of $\boldsymbol{M}_2$, is $\boldsymbol{x}_2''$.