Jordan normal form

1 Review the concepts of coordinate changes for linear maps, characteristic polynomials, eigenvalues, eigenvectors, diagonalization, Jordan normal form.

Let V be a vector space over $\mathbb{C}$ with $\dim V = n$, and $A: V \to V$ be a linear map. We start with matrix representation of A. Let $\mathcal{B} = (e_1, e_2, \dots, e_n)$ be a basis of V. Each Ae_i is a linear combination of $e_1, e_2, \dots, e_n$. Write

$$Ae_i = \sum_{j=1}^{n} \alpha_{ji} e_j .$$

The matrix $(\alpha_{ij})_{1 \le i,j \le n}$ is called the matrix representation or representing matrix of A in the basis $\mathcal{B}$. We denote $[A]_{\mathcal{B}} = (\alpha_{ij})_{1 \le i,j \le n}$. Each vector $x \in V$ is a linear combination of $e_1, \dots, e_n$. Write $x = \sum_{i=1}^{n} x_i e_i$. The column vector

$$[x]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

is called the coordinate representation of x in the basis $\mathcal{B}$. Thus,

$$[A]_{\mathcal{B}} = \left([Ae_1]_{\mathcal{B}} \ [Ae_2]_{\mathcal{B}} \ \cdots \ [Ae_n]_{\mathcal{B}} \right) \qquad (1)$$

and $[Ax]_{\mathcal{B}} = [A]_{\mathcal{B}} [x]_{\mathcal{B}}$.

Now we consider the coordinate changes. Let $\mathcal{B}' = (e_1', \dots, e_n')$ be another basis of V. The matrix representing the change of basis from $\mathcal{B}$ to $\mathcal{B}'$ is defined as

$$[I]_{\mathcal{B} \to \mathcal{B}'} = \left([e_1']_{\mathcal{B}} \ [e_2']_{\mathcal{B}} \ \cdots \ [e_n']_{\mathcal{B}} \right). \qquad (2)$$

Note that $[I]_{\mathcal{B} \to \mathcal{B}'}$ is an invertible matrix. We have

$$[x]_{\mathcal{B}} = [I]_{\mathcal{B} \to \mathcal{B}'} [x]_{\mathcal{B}'}. \qquad (3)$$

Since $[x]_{\mathcal{B}'} = [P]^{-1}_{\mathcal{B}\to\mathcal{B}'}[x]_{\mathcal{B}}$, we get the identity $[P]_{\mathcal{B}'\to\mathcal{B}} = [P]^{-1}_{\mathcal{B}\to\mathcal{B}'}$.

We have

$$[Ae_i]_{\mathcal{B}} = [P]_{\mathcal{B}\to\mathcal{B}'}[Ae_i]_{\mathcal{B}'} = [P]_{\mathcal{B}\to\mathcal{B}'}[A]_{\mathcal{B}'}[e_i]_{\mathcal{B}'} = [P]_{\mathcal{B}\to\mathcal{B}'}[A]_{\mathcal{B}'}[P]_{\mathcal{B}'\to\mathcal{B}}[e_i]_{\mathcal{B}}.$$

Then

$$([Ae_1]_{\mathcal{B}} \cdots [Ae_n]_{\mathcal{B}}) = [P]_{\mathcal{B}\to\mathcal{B}'}[A]_{\mathcal{B}'}[P]_{\mathcal{B}'\to\mathcal{B}}\underbrace{([e_1]_{\mathcal{B}} \cdots [e_n]_{\mathcal{B}})}_{=I_n}.$$

Thus,

$$[A]_{\mathcal{B}} = [P]_{\mathcal{B}\to\mathcal{B}'}[A]_{\mathcal{B}'}[P]^{-1}_{\mathcal{B}\to\mathcal{B}'}. \qquad (4)$$

We see that the representing matrices of A in different bases are conjugate to one another. The concepts of characteristic polynomials, eigenvalues, eigenspaces, diagonalizability which will be discussed do not depend on the choice of basis for V. However, choosing a basis is needed when we want to do calculation.

The characteristic polynomial of A is defined as $p_A(z) = \det(A - z\,Id_V)$. Each root of this polynomial is called an eigenvalue of A. If λ is an eigenvalue then the space

$$E_\lambda = \{x \in V : (A - \lambda Id_V)x = 0\} \qquad (5)$$

is nontrivial and called the eigenspace associate with λ.

In many circumstances, we want to find a basis of V in which the representing matrix of A is simple. If there exists a basis $\mathcal{B}$ of V such that $[A]_{\mathcal{B}}$ is diagonal then A is said to be diagonalizable, or semisimple. Not all linear transformations are diagonalizable. However, for every linear transformation A there exists a basis $\mathcal{B}$ such that $[A]_{\mathcal{B}}$ is of Jordan normal form, i.e.

$$[A]_{\mathcal{B}} = \begin{pmatrix} \boxed{\begin{matrix} \alpha_1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \alpha_1 \end{matrix}} & & & 0 \\ & \boxed{\begin{matrix} \alpha_2 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \alpha_2 \end{matrix}} & & \\ & & \ddots & \\ 0 & & & \boxed{\begin{matrix} \alpha_m & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \alpha_m \end{matrix}} \end{pmatrix}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_m$ are not necessarily distinct. Each block is called a <u>Jordan block</u>. If $\mathcal{B}$ diagonalizes A then $[A]_{\mathcal{B}}$ is a diagonal matrix, which is also of Jordan normal form where every Jordan block is of size 1.

To examine the diagonalizability and Jordan normal form of A, we view V as a module over the principal ring $\mathbb{C}[z]$ via the ring morphism $\mathbb{C}[z] \to \mathrm{End}(V)$, $f \mapsto f(A)$. The linear transformation $f(A)$ is defined as

$$f(A) = c_0 \, Id_V + \sum_{j=1}^{m} c_j \underbrace{A^j}_{\substack{= A \circ A \circ \cdots \circ A \\ (j \text{ times})}} \qquad \text{if} \quad f(z) = c_0 + \sum_{j=1}^{m} c_j z^j.$$

Because V is a finitely generated module over $\mathbb{C}$, it is also finitely generated over $\mathbb{C}[z]$. The Cayley-Hamilton theorem says that $p_A(A) = 0$. Thus, $p_A(z)$ is an exponent of V. Write $p_A(z) = (z-\lambda_1)^{r_1} \cdots (z-\lambda_m)^{r_m}$ where $\lambda_1, \ldots, \lambda_m$ are pairwise distinct complex numbers. Each polynomial $z - \lambda_j$ is a prime in $\mathbb{C}[z]$. By the structure theorem of finitely generated modules over a principal ring (Theorem 7.5, Lang "Algebra" p. 149),

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_m) \qquad \text{as} \quad \mathbb{C}[z]\text{-modules},$$

where $V(\lambda_j) = \ker(A - \lambda_j Id_V)^{r_j}$. This is an invariant submodule of V. The structure theorem further states that for each $\lambda \in \{\lambda_1, \dots, \lambda_m\}$,

$$V(\lambda) \simeq \mathbb{C}[z]/(z-\lambda)^{\nu_1} \oplus \dots \oplus \mathbb{C}[z]/(z-\lambda)^{\nu_s} \quad \text{as } \mathbb{C}[z]\text{-modules},$$

where $1 \le \nu_1 \le \nu_2 \le \dots \le \nu_s$ and the sequence $\nu_1, \dots, \nu_s$ is uniquely determined. Write $V(\lambda) = V(\lambda)_1 \oplus \dots \oplus V(\lambda)_s$ where $V(\lambda)_j \simeq \mathbb{C}[z]/(z-\lambda)^{\nu_j}$. Then

$$\begin{aligned} E_\lambda = \{v \in V : (z-\lambda)v = 0\} &= \{v \in V(\lambda) : (z-\lambda)v = 0\} \\ &= \{v = v_1 + \dots + v_s : v_j \in V(\lambda)_j,\ (z-\lambda)v_j = 0\} \\ &= \text{linear span}\{v_1, v_2, \dots, v_s\}, \end{aligned}$$

where $v_j \neq 0$ is an element of $V(\lambda)_j$ such that $(z-\lambda)v_j = 0$. Thus, $\dim E_\lambda = s$ called the <u>geometric multiplicity</u> of λ. ~~It is also~~ In other words, the geometric multiplicity of λ is the dimension of the eigenspace associate with λ. It is also the number of cyclic modules whose exponent is a power of $(z-\lambda)$ in the decomposition of V.

Let $(z-\lambda)^r$ be the power of $(z-\lambda)$ in $p_A(z)$. Because each $V(\lambda_j)$ is an invariant subspace of V, there is a basis of V in which the representing matrix of A is of block form.

$$\begin{pmatrix} \boxed{V(\lambda_1)} & & & 0 \\ & \boxed{V(\lambda_2)} & & \\ & & \ddots & \\ 0 & & & \boxed{V(\lambda_m)} \end{pmatrix}$$

Thus, $p_A(z) = p_{A|_{V(\lambda_1)}}(z) \cdots p_{A|_{V(\lambda_m)}}(z)$. Because $(z-\lambda_j)^{r_j}$ is an exponent of $V(\lambda_j)$,

λ_j is the only eigenvalue of $A|_{V(\lambda_j)}$. Thus, $p_{A|_{V(\lambda_j)}}(z)$ is a power of $(z-\lambda_j)$. Thus, $p_{A|_{V(\lambda_j)}}(z)=(z-\lambda_j)^{r_j}$. Then $\dim V(\lambda_j) = \deg p_{A|_{V(\lambda_j)}} = r_j$. We have showed that $\dim V(\lambda) = r$. Let $p_j(z)$ be the characteristic polynomial of $A|_{V(\lambda)_j}$. Because each $V(\lambda)_j$ is an invariant subspace of $V(\lambda)$, there is a basis of $V(\lambda)$ in which the representing matrix of $A|_{V(\lambda)}$ is of block form.

$$\begin{pmatrix} \boxed{V(\lambda)_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{V(\lambda)_s} \end{pmatrix}$$

Thus, $(z-\lambda)^r = p_{A|_{V(\lambda)}}(z) = p_1(z)\cdots p_s(z)$. Hence, each $p_j(z)$ is a power of $(z-\lambda)$. Because $V(\lambda)_j \simeq \mathbb{C}[z]/(z-\lambda)^{\nu_j}$ as $\mathbb{C}[z]$-modules, they are isomorphic as $\mathbb{C}$-modules. Thus, $\dim V(\lambda)_j = \nu_j$. Then $\deg p_j = \nu_j$ and

$$r = \deg p_{A|_{V(\lambda)}} = \deg p_1 + \cdots + \deg p_s = \nu_1 + \cdots + \nu_s .$$

As a consequence, $r \geq s$. The number r is called the <u>algebraic multiplicity</u> of λ. It is the exponent of $(z-\lambda)$ in the characteristic polynomial of A. It is also the sum of the dimensions of cyclic modules whose exponents are powers of $(z-\lambda)$ in the decomposition of V. We see that the algebraic multiplicity is always greater than or equal to the geometric multiplicity. They are equal if and only if $\nu_1 = \cdots = \nu_s = 1$, i.e. $E(\lambda) = V(\lambda)$. Note that we always have $E(\lambda) \subset V(\lambda)$. The following statements are equivalent.

(i) A is diagonalizable.

(ii) The algebraic multiplicity is equal to the geometric multiplicity for every eigenvalue.

(iii) $\nu_1 = \cdots = \nu_s = 1$ for every eigenvalue.

(iv) $\dim E(\lambda_j) = r_j$ for all $1 \leq j \leq m$.

(v) $\left((A - \lambda \mathrm{Id}_V)^2 v = 0 \Rightarrow (A - \lambda \mathrm{Id}_V) v = 0 \right)$ for every eigenvalue λ and vector $v \in V$.

(vi) $V = E(\lambda_1) \oplus \cdots \oplus E(\lambda_m)$.

<u>Diagonalization algorithm</u>

Let $A: V \to V$ be a linear transformation and $[A]_{\mathcal{B}_0}$ be its representing matrix in basis $\mathcal{B}_0$. Our goal is to find a basis $\mathcal{B}$ such that $[A]_{\mathcal{B}}$ is diagonal.

1) Calculate the characteristic polynomial $p_A(z) = \det(A - z \mathrm{Id}_V) = \det([A]_{\mathcal{B}_0} - z I_n)$.

2) Write $p_A(z) = (-1)^n (z - \lambda_1)^{r_1} \cdots (z - \lambda_m)^{r_m}$.

3) Find a basis for
$$E(\lambda) = \{ v \in V : (A - \lambda \mathrm{Id}_V) v = 0 \} \simeq \{ [v]_{\mathcal{B}_0} : ([A]_{\mathcal{B}_0} - \lambda I_n) [v]_{\mathcal{B}_0} = 0 \}$$

for each $\lambda = \lambda_1, \lambda_2, \dots, \lambda_m$.

- If $\dim E(\lambda_j) < r_j$ for some j then A is not diagonalizable. The algorithm stops.
- If $\dim E(\lambda_j) = r_j$ for all j then A is diagonalizable.

4) Let $\mathcal{B}_j$ be the basis of $E(\lambda_j)$ and $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_m)$ be the basis of V obtained by concatenating $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_m$ in that order. (the order of vectors within each $\mathcal{B}_j$ does not matter). Write $\mathcal{B} = (v_1, v_2, \dots, v_n)$. This is a basis that diagonalizes A.

$$[A]_{\mathcal{B}} = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_1 \end{matrix}} & & & 0 \\ & \boxed{\begin{matrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_2 \end{matrix}} & & \\ & & \ddots & \\ 0 & & & \boxed{\begin{matrix} \lambda_m & & \\ & \ddots & \\ & & \lambda_m \end{matrix}} \end{pmatrix} = [I]^{-1}_{\mathcal{B}_0 \to \mathcal{B}} [A]_{\mathcal{B}_0} [I]_{\mathcal{B}_0 \to \mathcal{B}}.$$

Recall $[P]_{\mathcal{B}_0 \to \mathcal{B}} = ([v_1]_{\mathcal{B}_0} \cdots [v_n]_{\mathcal{B}_0})$.

In case A is not diagonalizable, we want to find a basis of V in which the representing matrix of A is of Jordan normal form. By the analysis following the structure theorem, the algebraic multiplicity of λ is equal to the sum of the size of Jordan blocks whose diagonal entries are λ whereas the geometric multiplicity is equal to the number of Jordan blocks whose diagonal entries are λ.

$$V = \underbrace{\ker(A - \lambda_1 Id_V)^{r_1}}_{V(\lambda_1)} \oplus \cdots \oplus \underbrace{\ker(A - \lambda_m Id_V)^{r_m}}_{V(\lambda_m)}.$$

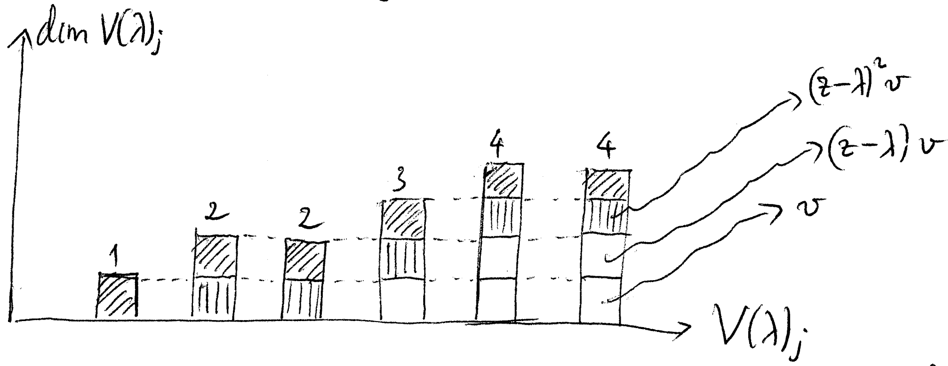
For each $\lambda = \lambda_1, \lambda_2, \dots, \lambda_m$ we write

$$V(\lambda) = \underbrace{V(\lambda)_1}_{\simeq \mathbb{C}[z]/(z-\lambda)^{\nu_1}} \oplus \cdots \oplus \underbrace{V(\lambda)_s}_{\simeq \mathbb{C}[z]/(z-\lambda)^{\nu_s}},$$

where $1 \leq \nu_1 \leq \nu_2 \leq \dots \leq \nu_s$. We will see that each $V(\lambda)_j$ corresponds to a Jordan block of A.

As a $\mathbb{C}$-module, $\mathbb{C}[z]/(z-\lambda)^{\nu}$ has a basis $1 + (z-\lambda)^{\nu}\mathbb{C}[z]$, $(z-\lambda) + (z-\lambda)^{\nu}\mathbb{C}[z], \dots, (z-\lambda)^{\nu-1} + (z-\lambda)^{\nu}\mathbb{C}[z]$. Because $V(\lambda)_j \simeq \mathbb{C}[z]/(z-\lambda)^{\nu_j}$ as $\mathbb{C}$-modules, $V(\lambda)_j$ has a basis $v, (z-\lambda).v, \dots, (z-\lambda)^{\nu_j-1}.v$ where v corresponds to $1 + (z-\lambda)^{\nu_j}\mathbb{C}[z]$ in the isomorphism. We can characterize v by the fact that it is an element in V such that $(z-\lambda)^{\nu_j}v = 0$ but $(z-\lambda)^{\nu_j-1}v \neq 0$. In order to find a basis for $V(\lambda)$, we need to find "v" for $V(\lambda)_s, V(\lambda)_{s-1}, \dots, V(\lambda)_1$ in that order to avoid collecting linearly dependent vectors as we move from $V(\lambda)_j$ to $V(\lambda)_i$.

First, we determine the numbers $\nu_1, \nu_2, \dots, \nu_s$. Suppose the solution spaces of $(A-\lambda Id_V)v=0$, $(A-\lambda Id_V)^2 v = 0, \dots, (A-\lambda Id_V)^N v = 0$ have dimensions $\alpha_1 < \alpha_2 < \dots < \alpha_N = r$ respectively. Note that $N = \nu_s$.



The squares ▨ denote basis vectors for $(A-\lambda Id_V)v = 0$. The square ▨ and ▥ denote basis vectors for $(A-\lambda Id_V)^2 v = 0$... We have

$$\# ▨ = \alpha_1$$

$$\# ▨ + \# ▥ = \alpha_2$$

....

Put $\beta_1 = \alpha_1$, $\beta_2 = \alpha_2 - \alpha_1, \dots, \beta_N = \alpha_N - \alpha_{N-1}$. Then β_j is equal to the number of $\nu_1, \nu_2, \dots, \nu_s$ that are $\geq j$. Put $\gamma_1 = \beta_1 - \beta_2, \dots, \gamma_{N-1} = \beta_{N-1} - \beta_N$, $\gamma_N = \beta_N$. Then γ_j is the number of $\nu_1, \dots, \nu_s$ that are equal to j. Once we get $\gamma_1, \dots, \gamma_N$ we can obtain $\nu_1, \dots, \nu_s$. For example, if $(\gamma_1, \dots, \gamma_N) = (0, 1, 0, 3, 1)$ then $(\nu_1, \dots, \nu_s) =$
$= (2, 4, 4, 4, 5)$.

Next, we find a basis of $V(\lambda)$ in which $A|_{V(\lambda)}$ is of Jordan normal form. We know that this form has s Jordan blocks whose sizes are $\nu_1, \dots, \nu_s$. Take $v_{1,1} \in V$ such that $(A-\lambda Id_V)^{\nu_s} v_{1,1} = 0$ and $(A-\lambda Id_V)^{\nu_s - 1} v_{1,1} \neq 0$.
Put $v_{1,2} = (A-\lambda Id_V)v_{1,1}$, $v_{1,3} = (A-\lambda Id_V)^2 v_{1,1}, \dots, v_{1,\nu_s} = (A-\lambda Id_V)^{\nu_s - 1} v_{1,1}$.
Then $v_{1,1}, v_{1,2}, \dots, v_{1,\nu_s}$ form a basis for $V(\lambda)_s$. We have

$$Av_{1,k-1} = (A-\lambda Id_V)v_{1,k-1} + \lambda v_{1,k-1} = \lambda v_{1,k-1} + v_{1,k} \quad \forall 2\le k\le \nu_s - 1.$$

$$Av_{1,\nu_s} = \lambda v_{1,\nu_s}.$$

In the basis $\mathcal{B}_1 = (v_{1,\nu_s}, v_{1,\nu_s-1}, \dots, v_{1,1})$, $A|_{V(\lambda)_s}$ is represented by the matrix

$$\left([Av_{1,\nu_s}] \cdots [Av_{1,1}]\right) = \underbrace{\begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}}_{\text{size } \nu_s}.$$

We have found the first piece of the desired basis for $V(\lambda)$. There are $s-1$ other pieces to be found. Take $v_{2,1} \in V$ such that $(A-\lambda Id_V)^{\nu_{s-1}} v_{2,1} = 0$, $(A-\lambda Id_V)^{\nu_{s-1}-1} v_{2,1} \neq 0$ and that $v_{2,1}$ is linearly independent of $\mathcal{B}_1$. Put $v_{2,2} = (A-\lambda Id_V)v_{2,1}, \dots, v_{2,\nu_{s-1}} = (A-\lambda Id_V)^{\nu_{s-1}-1} v_{2,1}$. Then $v_{2,1}, v_{2,2}, \dots, v_{2,\nu_{s-1}}$ form a basis for $V(\lambda)_{s-1}$. In the basis $\mathcal{B}_2 = (v_{2,\nu_{s-1}}, v_{2,\nu_{s-1}-1}, \dots, v_{2,1})$, $A|_{V(\lambda)_{s-1}}$ is represented by the matrix

$$\left([Av_{2,\nu_{s-1}}] \cdots [Av_{2,1}]\right) = \underbrace{\begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}}_{\text{size } \nu_{s-1}}.$$

This is the second piece of our desired basis for $V(\lambda)$. There are $s-2$ other pieces to be found. Take $v_{3,1} \in V$ such that $(A-\lambda Id_V)^{\nu_{s-2}} v_{3,1} = 0$, $(A-\lambda Id_V)^{\nu_{s-2}-1} v_{3,1} \neq 0$ and that $v_{3,1}$ is linearly independent of $\mathcal{B}_1 \cup \mathcal{B}_2$. We keep doing this procedure until all s pieces of the desired basis for $V(\lambda)$ is found. We concatenate these pieces to get $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_s)$. In this basis, $A|_{V(\lambda)}$ is represented by the matrix

$$[A|_{V(\lambda)}]_{\mathcal{B}} = \begin{pmatrix} \left.\boxed{\begin{matrix}\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{matrix}}\right\}\nu_s & & & \\ & \left.\boxed{\begin{matrix}\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{matrix}}\right\}\nu_{s-1} & & \\ & & \ddots & \\ & & & \left.\boxed{\begin{matrix}\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{matrix}}\right\}\nu_1 \end{pmatrix} = [I]^{-1}_{\mathcal{B}_0\to\mathcal{B}}\,[A|_{V(\lambda)}]_{\mathcal{B}_0}\,[I]_{\mathcal{B}_0\to\mathcal{B}}.$$

Recall $[I]_{\mathcal{B}_0\to\mathcal{B}} = \Big(\underbrace{[v_{1,\nu_s}]_{\mathcal{B}_0}\cdots [v_{1,1}]_{\mathcal{B}_0}}_{\text{from } \mathcal{B}_1}\ \underbrace{[v_{2,\nu_{s-1}}]_{\mathcal{B}_0},\dots,[v_{2,1}]_{\mathcal{B}_0}}_{\text{from } \mathcal{B}_2}\cdots\underbrace{[v_{s,\nu_1}]_{\mathcal{B}_0}\cdots[v_{s,1}]_{\mathcal{B}_0}}_{\text{from } \mathcal{B}_s}\Big)$.

After getting the desired basis for each $V(\lambda_i)$, we concatenate them to get a basis for V in which A is of Jordan normal form. The matrix representing the transformation is obtained by appending the matrices $[I]_{\mathcal{B}_0\to\mathcal{B}}$ together. By the above analysis, we can write an algorithm.

Jordan normal form algorithm

Let $A: V\to V$ be a linear transformation and $[A]_{\mathcal{B}_0}$ be its representing matrix in basis $\mathcal{B}_0$. Our goal is to find a basis $\mathcal{B}$ such that $[A]_{\mathcal{B}}$ is of Jordan normal form.

1) Calculate the characteristic polynomial $p_A(z) = \det(A - z\,Id_V) = \det([A]_{\mathcal{B}_0} - zI_n)$. Write $p_A(z) = (-1)^n (z-\lambda_1)^{r_1}\cdots(z-\lambda_m)^{r_m}$ where $\lambda_1,\dots,\lambda_m$ are pairwise distinct complex numbers and $r_1,\dots,r_m \geq 1$.

2) For each λ_j,

- Calculate the spaces $\ker(A-\lambda_j Id_V)$, $\ker(A-\lambda_j Id_V)^2, \dots$ until we first reach the number N with $\ker(A-\lambda_j Id_V)^N = V$. Denote the dimensions of those spaces as $\alpha_1 < \alpha_2 < \dots < \alpha_N = r_j$ respectively.

- Calculate $\beta_1 = \alpha_1,\ \beta_2 = \alpha_2 - \alpha_1, \dots, \beta_N = \alpha_N - \alpha_{N-1}$.
- Calculate $\gamma_1 = \beta_1 - \beta_2, \dots, \gamma_{N-1} = \beta_{N-1} - \beta_{N-2},\ \gamma_N = \beta_N$.
- Calculate $\nu_1 \leq \nu_2 \leq \dots \leq \nu_s$ such that γ_k is number of $\nu_1, \dots, \nu_s$ that are equal to k.
- Find $v_{1,1} \in \ker(A - \lambda_j Id_V)^{\nu_s} \setminus \ker(A - \lambda_j Id_V)^{\nu_s - 1}$. Calculate $v_{1,2} = (A - \lambda_j Id_V) v_{1,1}$, $v_{1,3} = (A - \lambda_j Id_V) v_{1,2}, \dots, v_{1,\nu_s} = (A - \lambda_j Id_V) v_{1,\nu_s - 1}$.
- Find $v_{2,1} \in \ker(A - \lambda_j Id_V)^{\nu_{s-1}} \setminus \ker(A - \lambda_j Id_V)^{\nu_{s-1} - 1}$ such that $v_{2,1}$ is linearly independent of $v_{1,1}, \dots, v_{1,\nu_s}$. Calculate $v_{2,2} = (A - \lambda_j Id_V) v_{2,1}, \dots, v_{2,\nu_{s-1}} =$ $= (A - \lambda_j Id_V) v_{2,\nu_{s-1} - 1}$.
- Find $v_{3,1} \in \ker(A - \lambda_j Id_V)^{\nu_{s-2}} \setminus \ker(A - \lambda_j Id_V)^{\nu_{s-2} - 1}$ such that $v_{3,1}$ is linearly independent of $v_{1,1}, \dots, v_{1,\nu_s}, v_{2,1}, \dots, v_{2,\nu_{s-1}}$. Calculate $v_{3,2} = (A - \lambda_j Id_V) v_{3,1}$, $\dots, v_{3,\nu_{s-2}} = (A - \lambda_j Id_V) v_{3,\nu_{s-2} - 1}$.

$\dots\dots$

- Find $v_{s,1}$ $\dots$
- Put $B_j = (\underbrace{v_{1,\nu_s}, \dots, v_{1,1}}, \underbrace{v_{2,\nu_{s-1}}, \dots, v_{2,1}}, \dots, \underbrace{v_{s,\nu_1}, \dots, v_{s,1}})$,

and

$$J_j = \begin{pmatrix} \left.\boxed{\begin{matrix} \lambda_j & 1 & & \\ & \lambda_j & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{matrix}}\right\}\nu_s & & 0 \\ & \ddots & \\ 0 & & \left.\boxed{\begin{matrix} \lambda_j & 1 & \\ & \ddots & 1 \\ & & \lambda_j \end{matrix}}\right\}\nu_1 \end{pmatrix}$$

3) Put $B = (B_1, B_2, \dots, B_m)$ and

$$J = \begin{pmatrix} J_1 & & 0 \\ & J_2 & \\ 0 & & \ddots \\ & & & J_m \end{pmatrix}$$

and $[I]_{\mathcal{B}_0\to\mathcal{B}} = \left([v_{1,\nu_1}]_{\mathcal{B}_0} \cdots [v_{1,1}]_{\mathcal{B}_0}\ [v_{2,\nu_2-1}]_{\mathcal{B}_0}\cdots[v_{2,1}]_{\mathcal{B}_0}\cdots [v_{s,\nu_s}]_{\mathcal{B}_0}\cdots[v_{s,1}]_{\mathcal{B}_0}\right).$

Then
$$[A]_{\mathcal{B}} = J = [I]^{-1}_{\mathcal{B}_0\to\mathcal{B}}\,[A]_{\mathcal{B}_0}\,[I]_{\mathcal{B}_0\to\mathcal{B}}.$$

$\boxed{2}$ Examples of diagonalizing a matrix

Example 1 Consider matrix $A=\begin{pmatrix}1&2&0\\0&3&0\\2&-4&2\end{pmatrix}$.

It can be viewed as a linear map from $\mathbb{C}^3$ to $\mathbb{C}^3$ whose representing matrix in the standard basis $\mathcal{B}_0$ is the given matrix.

The characteristic polynomial of A is $p_A(z)=\det(A-zI_3)=-(z-1)(z-2)(z-3)$.
The eigenvalues are $(\lambda_1,\lambda_2,\lambda_3)=(1,2,3)$. Now we compute the eigenspaces

$$E(\lambda_1)=\{x\in\mathbb{C}^3 : (A-I_3)x=0\}.$$

$$A-I_3=\begin{pmatrix}0&2&0\\0&2&0\\2&-4&1\end{pmatrix}\xrightarrow[r_1\to r_1/2]{r_2\to r_2-r_1}\begin{pmatrix}0&1&0\\0&0&0\\2&4&1\end{pmatrix}\xrightarrow[r_2\leftrightarrow r_3]{r_1\leftrightarrow r_3}\begin{array}{c}\begin{array}{ccc}x_1\downarrow & x_2\downarrow & x_3\downarrow\end{array}\\ \begin{pmatrix}2&-4&1\\0&1&0\\0&0&0\end{pmatrix}\end{array}.$$

Thus, $E(\lambda_1)=\{(x_1,x_2,x_3): x_2=0,\ x_3=-2a,\ x_1=a\}$.

$\dim E(\lambda_1)=1=$ multiplicity of root λ_1 of $p_A(z)$.

$E(\lambda_1)$ has a basis consisting of $v_1=\begin{pmatrix}1\\0\\-2\end{pmatrix}$.

$$E(\lambda_2)=\{x\in\mathbb{C}^3 : (A-2I_3)x=0\}$$

$$A-2I_3=\begin{pmatrix}-1&2&0\\0&1&0\\2&-4&0\end{pmatrix}\xrightarrow[r_1\to -r_1]{r_3\to r_3+2r_1}\begin{pmatrix}1&-2&0\\0&1&0\\0&0&0\end{pmatrix}$$

Thus, $E(\lambda_2)=\{(x_1,x_2,x_3): x_2=0,\ x_1=2x_2=0,\ x_3=a\}$

$\dim E(\lambda_2) = 1 =$ multiplicity of root λ_2 of $p_A(z)$.

$E(\lambda_2)$ has a basis consisting of $v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

$$E(\lambda_3) = \{x \in \mathbb{C}^3 : (A - 3I_3)x = 0\}.$$

$$A - 3I_3 = \begin{pmatrix} -2 & 2 & 0 \\ 0 & 0 & 0 \\ 2 & -4 & -1 \end{pmatrix} \xrightarrow[\substack{r_1 \to -r_1/2 \\ r_2 \leftrightarrow r_3}]{r_3 \to r_3 + r_1} \begin{pmatrix} 1 & -1 & 0 \\ 0 & -2 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus, $E(\lambda_3) = \{(x_1, x_2, x_3) : x_3 = -2a,\ x_2 = a,\ x_1 = a\}$.

$\dim E(\lambda_3) = 1 =$ multiplicity of root λ_3 of $p_A(z)$.

$E(\lambda_3)$ has a basis consisting of $v_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$.

Therefore, A is diagonalizable. In the basis $\mathcal{B} = (v_1, v_2, v_3)$, A is of diagonal form. The matrix representing the change of basis is $[P]_{\mathcal{B}_0 \to \mathcal{B}} = (v_1\ v_2\ v_3)$.

$$\underbrace{\begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}}_{[A]_{\mathcal{B}}} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ -2 & 1 & -2 \end{pmatrix}^{-1} \underbrace{\begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 2 & -4 & 2 \end{pmatrix}}_{[A]_{\mathcal{B}_0}} \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ -2 & 1 & -2 \end{pmatrix}}_{[P]_{\mathcal{B}_0 \to \mathcal{B}}}.$$

Example 2

Consider matrix $A = \begin{pmatrix} 2 & 4 & 6 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{pmatrix}$

It can be viewed as a linear map from $\mathbb{C}^3$ to $\mathbb{C}^3$ whose representing matrix in the standard basis $\mathcal{B}_0$ is the given matrix.

The characteristic polynomial of A is $p_A(z) = \det(A - zI_3) = -(z-2)^2(z-4)$. The eigenvalues are $(\lambda_1, \lambda_2) = (2, 4)$ where λ_1 is of multiplicity 2. Now we compute the eigenspaces.

$$E(\lambda_1) = \{ z \in \mathbb{C}^3 : (A - 2I_3) z = 0 \}.$$

$$A - 2I_3 = \begin{pmatrix} 0 & 4 & 6 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix} \xrightarrow[\substack{r_2 \to r_2/2 \\ r_1 \to \frac{r_1}{2} - 3r_2}]{r_3 \to r_3 - r_2} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus, $E(\lambda_1) = \{ (z_1, z_2, z_3) : z_1 = a, z_2 = 0, z_3 = 0 \}$. Because $\dim E(\lambda_1) = 1$ which is less than the multiplicity of λ_1, A is not diagonalizable.

3 Examples of making Jordan normal form

Example 1 Consider matrix $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$.

It can be viewed as a linear map from $\mathbb{C}^3$ to $\mathbb{C}^3$ whose representing matrix in the standard basis $\mathcal{B}_0$ is the given matrix.

The characteristic polynomial of A is $p_A(z) = -(z-1)^2(z-2)$. The eigenvalues are $(\lambda_1, \lambda_2) = (1, 2)$ where λ_1 is of multiplicity 2.

We find a desired basis for $\mathbb{C}^3(\lambda_1)$.

$$A - I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[r_2 \to \frac{r_2}{2}]{r_3 \to r_3 - \frac{r_2}{2}} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\ker(A - I_3) = \text{linear span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

$$(A - I_3)^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[\substack{r_2 \to r_2 - r_1 \\ r_1 \to r_1/2}]{r_3 \to r_3 - \frac{r_2}{2}} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\ker((A - I_3)^2) = \text{linear span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

We stop because $\dim\ker((A-I_3)^2) = 2 =$ multiplicity of λ_1. We get the sequence $(\alpha_1, \alpha_2) = (1,2)$. Then $\beta_1 = \alpha_1 = 1$, $\beta_2 = \alpha_2 - \alpha_1 = 1$. Then $\gamma_1 = \beta_1 - \beta_2 = 0$, $\gamma_2 = \beta_2 = 1$. This implies $\nu_1 = 2$. There is one Jordan block of size 2, which is $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. A vector in $\ker((A-I_3)^2) \setminus \ker(A-I_3)$ is $v_{1,1} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

$$v_{1,2} = (A - I_3)v_{1,1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

The desired basis for $\mathbb{C}^3(\lambda_1)$ is thus $(v_{1,2}, v_{1,1})$.

Now we find a desired basis for $\mathbb{C}^3(\lambda_2)$.

$$A - 2I_3 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[r_2 \to -r_2]{r_1 \to -r_1 - r_2} \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\ker(A-2I_3) = \text{linear span}\left\{\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}\right\}$$

We stop because $\dim\ker(A-2I_3) = 1 =$ multiplicity of λ_2. The desired basis of $\mathbb{C}^3(\lambda_2)$ consists of $v'_{1,1} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$.

Therefore, in the basis $\mathcal{B} = (v_{1,2}, v_{1,1}, v'_{1,1})$, A is of Jordan normal form. The matrix representing the change of basis is $[I]_{\mathcal{B}_0 \to \mathcal{B}} = (v_{1,2} \; v_{1,1} \; v'_{1,1})$.

$$\underbrace{\begin{pmatrix} \boxed{\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}} & \\ & \boxed{2} \end{pmatrix}}_{[A]_{\mathcal{B}}} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}}_{[A]_{\mathcal{B}_0}} \underbrace{\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}}_{[I]_{\mathcal{B}_0 \to \mathcal{B}}}.$$

Example 2

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

It can be viewed as a linear map from $\mathbb{C}^4$ to $\mathbb{C}^4$ whose representing matrix in the standard basis $\mathcal{B}_0$ is the given matrix.

The characteristic polynomial of A is $p_A(z) = (z-1)^3(z-3)$. The eigenvalues are $(\lambda_1, \lambda_2) = (1, 3)$ where λ_1 is of multiplicity 3.

We find a desired basis for $\mathbb{C}^4(\lambda_1)$.

$$A - I_4 = \begin{pmatrix} 0 & 0 & 3 & 0 \\ 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} \xrightarrow[\substack{r_1 \to r_1/3 \\ r_1 \leftrightarrow r_2}]{r_4 \to r_4 - r_1} \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{aligned} \ker(A - I_4) &= \{(z_1, z_2, z_3, z_4) \in \mathbb{C}^4 : z_2 = a,\ z_4 = b,\ z_3 = 0,\ z_1 = -2a - 3b\} \\ &= \{(-2a-3b, a, 0, b) : a, b \in \mathbb{C}\} \\ &= \text{linear span} \left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

$$(A - I_4)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & 4 & 12 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[r_1 \leftrightarrow r_2]{r_2 \to r_2/2} \begin{pmatrix} 1 & 2 & 6 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{aligned} \ker((A - I_4)^2) &= \{(-2a - 6b - 3c, a, b, c) : a, b, c \in \mathbb{C}\} \\ &= \text{linear span} \left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -6 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\} \end{aligned}$$

We stop because $\dim \ker((A-I_4)^2) = 3 =$ multiplicity of λ_1. We get the sequence $(\alpha_1, \alpha_2) = (2, 3)$. Thus, $\beta_1 = \alpha_1 = 2$, $\beta_2 = \alpha_2 - \alpha_1 = 1$. Then $\gamma_1 = \beta_1 - \beta_2 = 1$, $\gamma_2 = \beta_2 = 1$. This implies $(\nu_1, \nu_2) = (1, 2)$. There is one Jordan block of size 1 and one of size 2, which are

$$\begin{pmatrix} \boxed{\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}} & \\ & \boxed{1} \end{pmatrix}.$$

A vector in $\ker((A-I_4)^2) \setminus \ker(A-I_4)$ is $v_{1,1} = \begin{pmatrix} -6 \\ 0 \\ 1 \\ 0 \end{pmatrix}$.

$$v_{1,2} = (A - I_4) v_{1,1} = \begin{pmatrix} 0 & 0 & 3 & 0 \\ 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} -6 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -6 \\ 0 \\ 3 \end{pmatrix}.$$

A vector in $\ker(A - I_4)$ that is independent of $v_{1,1}$ and $v_{1,2}$ is $v_{2,1} = \begin{pmatrix} -3 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

The desired basis for $\mathbb{C}^4(\lambda_1)$ is $\mathcal{B}_1 = (v_{1,2}, v_{1,1}, v_{2,1})$.

Now we find a desired basis for $\mathbb{C}^4(\lambda_2)$.

$$A - 3I_4 = \begin{pmatrix} -2 & 0 & 3 & 0 \\ 1 & 0 & 0 & 3 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 3 & -2 \end{pmatrix} \xrightarrow[\substack{r_3 \to (-r_3 + r_1)/8 \\ r_2 \to -r_2/2}]{r_1 \to r_1 + 2r_2} \begin{pmatrix} 0 & 0 & 3 & 6 \\ 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[\substack{r_2 \leftrightarrow r_3 \\ r_3 \leftrightarrow r_4 \\ r_1 \to r_1 - 3r_3}]{\substack{r_1 \to r_1 \cancel{\times 8} - 3r_3 - 6r_4 \\ r_1 \leftrightarrow r_2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\ker(A - 3I_4) = \{(0, a, 0, 0) : a \in \mathbb{C}\}$$

$$= \text{linear span} \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

We stop because $\dim \ker(A - 3I_4) = 1 =$ multiplicity of λ_2. The desired basis $\mathcal{B}_2$ of $\mathbb{C}^4(\lambda_2)$ consists of $v'_{1,1} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

Therefore, in the basis $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2) = (v_{1,2}, v_{1,1}, v_{2,1}, v'_{1,1})$, A is of Jordan normal form. The matrix representing the change of basis is $[P]_{\mathcal{B}_0 \to \mathcal{B}} = (v_{1,2} \; v_{1,1} \; v_{2,1} \; v'_{1,1})$.

$$\underbrace{\begin{pmatrix} \boxed{\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}} & & \\ & 1 & \\ & & 3 \end{pmatrix}}_{[A]_{\mathcal{B}}} = \begin{pmatrix} 3 & -6 & -3 & 0 \\ -6 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \end{pmatrix}^{-1} \underbrace{\begin{pmatrix} 1 & 0 & 3 & 0 \\ 1 & 3 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 \end{pmatrix}}_{[A]_{\mathcal{B}_0}} \underbrace{\begin{pmatrix} 3 & -6 & -3 & 0 \\ -6 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \end{pmatrix}}_{[P]_{\mathcal{B}_0 \to \mathcal{B}}}.$$