7. THE JORDAN CANONICAL FORM

7.1 Introduction

Not every matrix has enough linearly independent eigenvectors to be diagonalizable. However by using similarity transformations every square matrix can be transformed to the Jordan canonical form, which is almost diagonal.

Example 7.1.1 Consider the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}.$$

It has eigenvalues $\lambda_{1,2} = 2$, but only one linearly independent eigenvector.

$$\begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Leftrightarrow \quad 3x_2 = 0,$$

We choose for example

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Example 7.1.2. The $r_i \times r_i$ matrix

$$J_i = \begin{bmatrix} \lambda_1 & 1 & 0 & \cdots & 0 \\ 0 & \lambda_1 & 1 & \cdots & 0 \\ \vdots & & & \vdots & \vdots \\ 0 & \cdots & \cdots & \lambda_1 & 1 \\ 0 & & & 0 & \lambda_1 \end{bmatrix}$$

has only one lin.ind. eigenvector $[x_1, 0, \ldots, 0]^T = x_1 \boldsymbol{e}_1$.

7.2 The Jordan canonical form

Theorem 7.2.1. The Jordan canonical form. Every nxn–matrix A is similar to a matrix

$$J = diag\left[J_1, J_2, J_3, \ldots, J_p\right],$$

where every J_i is an $r_i \times r_i$ matrix of the form presented above. The representation is unique up to permutations of the diagonal blocks.

Proof: (See Gantmacher: Matrizenrechnung, 1970 or Ortega: Matrix Theory 1987, maybe the simplest one is due Väliahon and can be found in the Prasolov's book (1994)).

Exercise 7.2.1. Find a matrix P so that in the preceding theorem

$$PJP^{-1} = diag\left[J_2, J_1, J_3, \dots, J_p\right].$$

Definition. The matrix J presented in theorem 8.2.1 is called ***the Jordan canonical form*** of A. The blocks J_i are called ***Jordan blocks.*** If $r_i = 1\ \forall i$, then $p = n$ and A is similar to a diagonal matrix $(J_1 = \lambda_i)$. If $r_i = n$, then $J = J_1$.

NB. By using exercise 7.2.1 we can find similarity transformations that permutes the Jordan blocks into the order wanted. So every matrix is similar to a unique Jordan canonical form and so the **matrices *A* and *B* are similar if and only if they have the same Jordan canonical form.**

Example 7.2.1. We now consider all the possible Jordan canonical forms of the 4x4 matrix.

a) $$J = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \qquad r_i = 1, p = 4$$

b) $$J = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \qquad r_1 = r_2 = 2, p = 2$$

c) $$J = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \qquad r_1 = r_2 = 1, r_3 = 2, p = 3$$

d) $$J = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \qquad r_1 = 1, r_2 = 3,\ p = 2$$

e) $$J=\begin{bmatrix}2&1&0&0\\0&2&1&0\\0&0&2&1\\0&0&0&2\end{bmatrix}\qquad r_1=4,\quad p=1$$

Where every matrix here has two as *its* eigenvalues and characteristic polynomial $(\lambda-2)^4$. We observe that matrices can have the same eigenvalues but different Jordan forms. We now concentrate on eigenvectors.

There is one linearly independent eigenvector corresponding to every Jordan block. The lin.ind. eigenvectors of the matrix J = *diag* $[J_1, J_2, \ldots, J_p]$ are

$$\boldsymbol{e}_1, \boldsymbol{e}_{r_1+1}, \boldsymbol{e}_{r_1+r_2+1}, \ldots, \boldsymbol{e}_{r_1+r_2+\ldots+r_{p-1}+1}.$$

The eigenvecotrs of the matrices in the previous exercise are:

(a) $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$, (b) $\boldsymbol{e}_1, \boldsymbol{e}_3$, (c) $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, (d) $\boldsymbol{e}_1, \boldsymbol{e}_2$, (e) $\boldsymbol{e}_1$

We use the amount of the eigenvectors to define:

Definition. Algebraic multiplicity of the eigenvalue λ_i is *its* multiplicity as a root of the characteristic polynomial. *Its* ***geometric multiplicity*** is the amount of linearly independent eigenvectors corresponding to λ_i. If the geometric multiplicity < algebraic multiplicity then the matrix is ***defective.***

Example 7.2.2 In the preceding example the algebraic multiplicity of the eigenvalue $\lambda = 2$ of the 4x4 matrices (a), (b), (c), (d), (e) is 4, because $|\lambda I-A| = (\lambda-2)^4$. The geometric multiplicity is however a) 4, b) 2, c) 3, d) 2, e) 1.

If the matrix A does not have n linearly independent eigenvectors it is not diagonalizable by similar transformations. However it can be transformed to Jordan canonical form by similar transformations. We now acquaint ourselves with the structure of this similarity transformation matrix.

Let A be similar to J:

$$A = SJS^{-1} \Leftrightarrow AS = SJ \Leftrightarrow$$

$$A\left[\boldsymbol{s}_1,\boldsymbol{s}_2,\ldots,\boldsymbol{s}_n\right]=\left[\boldsymbol{s}_1,\boldsymbol{s}_2,\ldots,\boldsymbol{s}_n\right]\begin{bmatrix}J_1&0&\cdots&0\\0&J_2&0&0\\0&&\cdots&J_p\end{bmatrix}$$

Direct computations shows that for the first block

$$\begin{cases} A\boldsymbol{s}_1 = \lambda_1 \boldsymbol{s}_1 \\ A\boldsymbol{s}_2 = \lambda_1 \boldsymbol{s}_2 + \boldsymbol{s}_1 \Leftrightarrow (A - \lambda_1 I)\boldsymbol{s}_2 = \boldsymbol{s}_1 \\ A\boldsymbol{s}_i = \lambda_1 \boldsymbol{s}_i + \boldsymbol{s}_{i-1} \Leftrightarrow (A - \lambda_1 I)\boldsymbol{s}_i = \boldsymbol{s}_{i-1}, i = 2, \dots, r_1 \end{cases}$$

The vector set $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \dots, \boldsymbol{s}_{r_1}\}$ is called a ***Jordan chain.*** The vector $\boldsymbol{s}_1$ is an ***eigenvector***. The vectors $\boldsymbol{s}_2, \dots, \boldsymbol{s}_{r_1}$ are called ***generalized eigenvectors.*** The same thing happens for other blocks also, so if $A = SJS^{-1}$ then the column vectors of S form p Jordan chains.

$$\{\ \boldsymbol{s}_1, \boldsymbol{s}_2, \dots, \boldsymbol{s}_{r_1}\ \}, \{\ \boldsymbol{s}_{r_1+1}, \dots, \boldsymbol{s}_{r_1+r_2}\ \}, \dots, \{\ \boldsymbol{s}_{r_1+r_2+\dots+r_{p-1}+1}, \dots, \boldsymbol{s}_n\ \}.$$

Example 7.2.3 We consider the matrix

$$A = \frac{1}{2}\begin{bmatrix} 5 & -1 \\ 1 & 3 \end{bmatrix}.$$

It has eigenvalues

$$|A - \lambda I| = \begin{vmatrix} \frac{5}{2} - \lambda & -\frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} - \lambda \end{vmatrix} = 0 \Leftrightarrow \lambda_{1,2} = 2,$$

and the corresponding eigenvectors can be solved from the equation

$$\begin{bmatrix} 5 & -1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad .$$

One solution is

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and there are no more linearly independent solutions. The Jordan chain is

$$\boldsymbol{s}_1 = \boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

$$(A - 2I)\boldsymbol{s}_2 = \boldsymbol{s}_1 \Leftrightarrow \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where $x_2 = 0, x_1 = 2$, so

$$s_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

So we get the chain

$$\{s_1, s_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right\},$$

hence

$$S = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}, \text{ and } \quad S^{-1} = \begin{bmatrix} 0 & 1 \\ 1/2 & -1/2 \end{bmatrix}.$$

and

$$A[s_1, s_2] = A \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix},$$

Finally we compute the Jordan canonical form of the matrix A

$$J = S^{-1}AS = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 5/2 & -1/2 \\ 1/2 & 3/2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

Exercise 7.2.2 Compute the algebraic and geometric multiplicities of matrices

a) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ b) $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ d) $\begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}$

Which of these matrices are diagonalizable? Find the Jordan canonical forms and show that the matrices truly are similar to their canonical forms.

7.3 Constructing Jordan canonical form

Usually the Jordan canonical form is used to prove theoretical results, so we seldom need to find it. Whatever algorithm is used to form the canonical form it is extremely sensitive to numerical errors. That is why symbolic computation is needed. It is not easy to do either because eigenvalues should be calculated accurately. Unfortunately, a general method for computing roots an equation of the n^{th} order does not excist. The following algorithm finds out how many Jordan blocks correspond to every distinct eigenvalue, respectively and what their dimensions are. After we know this it is easy to construct the Jordan canonical form.

1. Compute the distinct eigenvalues $\beta_1, \beta_2, \ldots, \beta_s$ of the nxn matrix A. Preferably accurately.

2. Let β_i be an eigenvalue. For every eigenvalue compute the rank

$$r_j(\beta_i) = rank\left[(A-\beta_i I)^j\right] \quad 1 \le j \le n$$

If $r_k(\beta_i) = r_{k+1}(\beta_i)$, then $r_j(\beta_i) = r_k(\beta_i)$, for every $j \ge k$.

3. Compute the numbers

$$b_1(\beta_i) = n - 2r_1(\beta_i) + r_2(\beta_i)$$

$$b_m(\beta_i) = r_{m+1}(\beta_i) - 2r_m(\beta_i) + r_{m-1}(\beta_i), \qquad m \ge 2$$

In Jordan canonical form there occurs precisely $b_m(\beta_i)$ mxm Jordan blocks corresponding to eigenvalue β_i.

Example 7.3.1

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$$

and has spectrum $\sigma(A) = \{3,3\}$. Now $\beta_1 = 3$.

$$r_1(3) = rank\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} = 1, \quad r_2(3) = rank\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0,$$

hence $r_j(3) = 0$, when $j \ge 2$.

$$b_1(3) = 0, \quad b_2(3) = 1.$$

There is one 2x2 block corresponding to the eigenvalue 3 in the Jordan canonical form,

$$J = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}.$$

Exercise 7.3.1 What is the Jordan canonical form of the matrix

$$\mathrm{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 2 & 1 & -1 & -1 & 0 & -1 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ -1 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}?$$

7.4 Cayley-Hamilton-theorem

We introduce an interesting result in this chapter. The Cayley-Hamilton-theorem says that every matrix satisfies *its* own characteristic polynomial, in other words replacing every variable λ in the characteristic polynomial with the matrix A gives us a zero matrix as the result. We first prove a lemma, which shows how the spectrum of a matrix polynomial is dependent on the matrix.

Theorem 7.4.1. Let $A \in \mathbf{C}^{nxn}$, $\sigma(A) = \{\lambda_1,\ldots,\lambda_n\}$ and

$$p(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \ldots + a_0$$

be some polynomial. The corresponding ***matrix polynomial***

$$p(A) = A^m + a_{m-1}A^{m-1} + \ldots + a_0 I$$

has spectrum

$$\sigma(p(A)) = \{p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_n)\}.$$

Proof.

$$A = SJS^{-1}, J = diag\left[J_1, J_2, \ldots, J_p\right].$$

$$A^i = (SJS^{-1})^i = (SJS^{-1})(SJS^{-1})(SJS^{-1})\ldots(SJS^{-1}) = SJ^iS^{-1},$$

hence

$$p(A) = SJ^mS^{-1} + a_{m-1}SJ^{m-1}S^{-1} + \ldots + a_0 I$$

$$= S(J^m + a_{m-1}J^{m-1} + \ldots + a_0 I)S^{-1} = Sp(J)S^{-1}.$$

Because $J = diag[J_1, J_2, \ldots, J_p]$, so then $J^i = diag\left[J_1^i, \; J_2^i, \; \ldots, \; J_p^i\right]$, and

$$p(J) = diag\left[\; p(J_1), p(J_2), \ldots, p(J_p) \;\right].$$

Furthermore, because

$$J_1 = \begin{bmatrix} \lambda_1 & 1 & & \cdots & 0 \\ 0 & \lambda_1 & 1 & \cdots & 0 \\ .. & .. & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \lambda_1 & 1 \\ 0 & \ldots & & 0 & \lambda_1 \end{bmatrix},$$

$$J_1^i = \begin{bmatrix} \lambda_1^i & x & x & \cdots & x \\ 0 & \lambda_1^i & x & \cdots & x \\ .. & .. & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \lambda_1^i & x \\ 0 & \ldots & & 0 & \lambda_1^i \end{bmatrix},$$

where x can be $\neq 0$, so

$$p(J_1) = \begin{bmatrix} p(\lambda_1) & x & x & \cdots & x \\ 0 & p(\lambda_1) & x & \cdots & x \\ .. & .. & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & p(\lambda_1) & x \\ 0 & \ldots & & 0 & p(\lambda_1) \end{bmatrix},$$

and the block $p(J_1)$ has eigenvalue $p(\lambda_1)$. The same thing happens in the other blocks as well, so the claim holds. The matrix polynomial $p(A)$ has at least as many eigenvectors as A because the Jordan blocks are transformed into upper triangular blocks when forming the polynomial. There might be more eigenvectors because the Jordan blocks can fall apart. ❑

Example 7.4.1 Let

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

and $p(\lambda) = \lambda^2 + 1$,

$$p(A) = A^2 + I = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 4 \\ 0 & 5 \end{bmatrix},$$

so the spectrum of $p(A)$ is

$$\sigma(p(A)) = \{5,5\}.$$

On the other hand $p(2) = 5$, so the theorem says that

$$\sigma(p(A)) = \{p(2), p(2)\} = \{5,5\}.$$

Example 7.4.2 The matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

has only one linearly independent eigenvector. If $p(\lambda) = \lambda^2$, then

$$p(A) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

and $p(A)$has two linearly independent eigenvectors.

Theorem 7.4.2(Cayley-Hamilton) If $p(\lambda)$ is characteristic polynomial of A, then $p(A) = O$.

Proof.

$$p(\lambda) = |\lambda I - A| = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n),$$

so

$$p(A) = (A - \lambda_1 I)(A - \lambda_2 I)\ldots(A - \lambda_n I).$$

We first assume, that A is diagonalizable, that is $A = SDS^{-1}$, $D = diag[\lambda_1, \lambda_2,\ldots,\lambda_n]$. Now

$$p(A) = S(D - \lambda_1 I)S^{-1}S(D - \lambda_2 I)S^{-1}\ldots S(D - \lambda_n I)S^{-1}$$

$$= S(D - \lambda_1 I)(D - \lambda_2 I)\ldots(D - \lambda_n I)S^{-1}$$

$$= S\ \ diag[0, \lambda_2 - \lambda_1,\ldots,\lambda_n - \lambda_1]\ \ diag[\lambda_1 - \lambda_2, 0,\ldots,\lambda_n - \lambda_2)]..$$

$$diag[\lambda_1 - \lambda_n,\ldots,0]S^{-1} = S \cdot O \cdot S^{-1} = O.$$

More generally if $A = SJS^{-1}$ and J is Jordan canonical form of A, then

$$p(A) = S(J - \lambda_1 I)(J - \lambda_2 I)\ldots(J - \lambda_n I)S^{-1}.$$

Let

$$J = diag\left[J_1, J_2,\ldots, J_p\right].$$

Eigenvalues occur in blocks r_i times and we get

$$p(A) = S\left[\ (J-\lambda_1 I)^{r_1}\ (J-\lambda_1 I)^{r_2} \ldots (J-\lambda_n I)^{r_p}\ \right]S^{-1}.$$

We have

$$(J-\lambda_1 I)^{r_1} = diag\left[\ (J_1-\lambda_1 I)^{r_1}, (J_2-\lambda_1 I)^{r_1}, \ldots, (J_p-\lambda_1 I)^{r_1}\ \right]$$

Block

$$(J_1-\lambda_1 I)^{r_1} = \left(\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \vdots & & & & \\ 0 & 0 & & \cdots & 1 \\ 0 & 0 & & \cdots & 0 \end{bmatrix}\right)^{r_1} = O,$$

so the first block of the first factor in the product is zero matrix. The second block of the second factor is zero vector and so on, so finally $p(A) = O$. ❑

Exercise 7.4.1 Cayley-Hamilton theorem can be proved using Cramers rule. It says

$$det(\lambda I\text{-}A)I = (\lambda I\text{-}A)\ adj(\lambda I\text{-}A).$$

$adj(\lambda I\text{-}A)$ is a matrix polynomial of order n-1 (it is formed by using the subdeterminants of an $(n\text{-}1)x(n\text{-}1)$ matrix). We denote

$$adj(\lambda I - A) = B_{n-1}\lambda^{n-1} + B_{n-2}\lambda^{n-2} + \ldots + B_0,$$

where B_i are nxn matrices. The previous equation now has the form

$$det(\lambda I - A)I = (\lambda I - A)adj(\lambda I - A)$$

$$= \lambda^n B_{n-1} + \lambda^{n-1}\left[B_{n-2} - AB_{n-1}\right] + \ldots + \lambda\left[B_0 - AB_1\right] - AB_0.$$

Write $det(\lambda I\text{-}A)$ as a polynomial of λ and substitute A into the right and left hand sides of the equation, and you will get matrices B_i as a result. Substitute them into the right hand side equation to conclude the arguement.

7.5 Minimal polynomial

Definition 7.5.1 The ***minimal polynomial*** of a matrix A is the polynomial $\pi(\lambda)$ that has the lowest order leading coefficient one such that

$\pi(A) = 0.$

Theorem 7.5.1 Let A be an nxn matrix and $\beta_1, \beta_2, \ldots, \beta_q$ *its* distinct eigenvalues. Let s_i be the size of the largest Jordan block corresponding to β_i. Now the order of the minimal polynomial of A is $s_1 + s_2 + \ldots + s_q$ and the minimal polynomial is

$$\pi(\lambda) = (\lambda - \beta_1)^{s_1} (\lambda - \beta_2)^{s_2} \cdots (\lambda - \beta_q)^{s_q}.$$

Proof. (Similar to proof of the Cayley-Hamilton theorem). It follows from the fact that the block $(J_i - \beta_i I)$ needs to be raised to a power s_i to get a zero matrix, where s_i is the size of the largest block corresponding to the eigenvalue β_i. ❑

Example 7.5.1

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 2.$$

The minimal polynomial is $(\lambda-2)^3$. The characteristic polynomial is $(\lambda-2)^4$.

Theorem 7.5.1 can be used to create an easy method for finding the minimal - polynomial, when the matrix is small. First we find the characteristic polynomial of A and separate it into factors. We assume that characteristic polynomial is

$$\left|\lambda I - A\right| = \prod_{k=1}^{s} (\lambda - \beta_k)^{m_k}, \quad \sum_{k=1}^{s} m_k = n.$$

The minimal polynomial of the matrix A is

$$\psi(\lambda) = \prod_{k=1}^{s} (\lambda - \beta_k)^{n_k}, \quad 1 \le n_k \le m_k$$

The order n_k can be found by trial and error. We reduce the order of the factors in order and check if A is root of the polynomial. If it is a root, then we continue reducing the order, if it is not, then we move to the next factor.

Definition. Elementary Divisors. Elementary divisors of matrix A are the characteristic polynomials of its Jordan blocks.

Exercise 7.5.1 Prove that similar matrices have the same minimal polynomial.

Exercise 7.5.2 Find the Jordan canonical form of the matrix

$$A = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 4 & 0 \\ -1 & 3 & 1 \end{bmatrix}$$

Find the eigenvectors and generalized eigenvectors, characteristic polynomial, minimal polynomial and the characteristic polynomials of each Jordan block.

Exercise 7.5.3 Prove that the minimal polynomial is the characteristic polynomial if and only if only each block corresponds to a distinct eigenvalue.