Chapter 6

CHAPTER SIX THE JORDAN CANONICAL FORM AND APPLICATIONS

6.1 Introduction

In this chapter we elaborate upon the investigations into similarity which were begun in Chapter 5 and bring our understanding of the matter to a satisfactory and elegant conclusion in the presentation of the "Jordan[1] canonical form." This term refers to a special form that a matrix may be transformed into under similarity.

We saw in Chapter 5 that the similarity transformation of a matrix into a special form is of interest from the point of view of applications and that problems of transforming a matrix under similarity are quite interesting in themselves. The diagonalization of symmetric matrices was applied to quadratic forms in Section 5.6 and to the inertia tensor in Section 5.7. We will see in Section 6.3 that the Jordan canonical form is of use in solving systems of differential equations.

It would be convenient if every real matrix were orthogonally similar to a diagonal matrix, but unfortunately, it is only the symmetric matrices that have this property. In problems involving similarity, say similarity to an upper triangular matrix, factorization of the characteristic polynomial is always a stumbling block and so any result must carry along the necessary assumptions regarding it. It has been proved that there is no "quadratic formula" type method for solving polynomial equations of degree five and larger, and so we can feel sure that this factorization must be assumed separately. Is there a best result that can be stated, with reasonable assumptions, regarding similarity? An answer will soon appear.

In this section, we will review the theory and methods developed in Chapter 5 regarding similarity, diagonalization, eigenvalues and eigenvectors, and the characteristic and minimum polynomials. In addition, we will consider several examples and present the definition of the Jordan block, a fundamental unit in the discussion that follows.

[1] It is named for the French mathematician Camille Jordan (1838-1922).

REVIEW

Let A be an $n \times n$ matrix. We will use the notation of Chapter 5 for the characteristic and minimum polynomials of A, and we will rely on the definitions of eigenvalue and eigenvector from that same chapter.

6.1.1 Summary of Previous Results:

(a) The eigenvalues of A are the roots of the characteristic polynomial, $p_A(\lambda)$, of A. To find an eigenvector of A corresponding to the eigenvalue λ_0, one finds a solution of the homogeneous system $(A - \lambda_0 I)X = 0$. (Theorem 5.2.1)

(b) A is similar to a diagonal matrix if and only if A has n linearly independent eigenvectors. If $X_1, \ldots, X_n$ are independent eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$ and $S = [X_1 \ldots X_n]$, then S is nonsingular and $S^{-1}AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. (Theorem 5.3.1)

(c) If A is a real matrix and $p_A(\lambda)$ factors completely or A is a complex matrix, then A is orthogonally similar (resp., unitarily similar) to an upper triangular matrix. The eigenvalues of an upper triangular matrix are the entries on the diagonal. (Schur's Theorem, Theorem 5.4.2)

(d) A is a real symmetric matrix if and only if A is orthogonally similar to a diagonal matrix. In this case, the eigenvalues of A are real. (Corollary 5.4.1 (the Principal Axes Theorem) and Theorem 5.2.4)

(e) A is a hermitian matrix over the complex numbers if and only if A is unitarily similar to a diagonal matrix and the eigenvalues of A are real. (Corollary 5.4.2 and Theorem 5.2.4)

(f) Eigenvectors corresponding to distinct eigenvalues of A are linearly independent. If A is hermitian or real symmetric, then eigenvectors corresponding to distinct eigenvalues are orthogonal. (Theorem 5.2.3 and Theorem 5.4.3)

(g) A matrix satisfies its characteristic polynomial; that is, $p_A(A) = 0$. (Theorem 5.5.1 (the Cayley-Hamilton Theorem))

(h) The monic polynomial $m_A(\lambda)$ of least degree satisfying $m_A(A) = 0$ is the minimal polynomial of A. If $p_A(\lambda) = (a_1 - \lambda)^{m_1} \cdots (a_s - \lambda)^{m_s}$, then $m_A(\lambda) = (a_1 - \lambda)^{n_1} \cdots (a_s - \lambda)^{n_s}$, where $1 \leq n_i \leq m_i$ for $i = 1, \ldots, s$. That is, $m_A(\lambda)$ is a factor of $p_A(\lambda)$ and $m_A(\lambda)$ contains each of the linear factors of $p_A(\lambda)$. (Definition, Theorem 5.5.2, and Theorem 5.5.3)

(i) Similar matrices have the same minimum and characteristic polynomials. In particular, similar matrices have the same eigenvalues. (Theorem 5.2.2 and Exercise 7 of section 5.5)

Having reviewed these facts from Chapter 5, let us consider some easy examples to gain some experience with and appreciation for the theory we have just reviewed.

6.1.2 Examples

Example 6.1.1. (a) Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. Then A is upper triangular and so the eigenvalues of A are 1 and 2 (See part (c) of the Summary). Since the eigenvalues of A are distinct, the corresponding eigenvectors are linearly independent (See part (f) of the Summary), and so A is similar to the diagonal matrix $\operatorname{diag}(1, 2)$ by part (b) of the Summary.

(b) Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Then the only eigenvalue of A is 1 and it has multiplicity 2. Now $A - 1I$ has rank 1. It follows that there are at most $2 - 1 = 1$ independent eigenvectors, and so, A is not similar to a diagonal matrix (See part (b) of the Summary).

(c) Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then $p_A(\lambda) = \lambda^2 + 1$ and so the characteristic polynomial does not factor completely over the real numbers. It follows that A is not similar to a diagonal matrix over the real numbers. However, considering A as a matrix over the complex numbers, $p_A(\lambda)$ factors as $p_A(\lambda) = (\lambda - i)(\lambda + i)$. Thus A has two distinct complex eigenvalues and so there is a complex matrix S with:

$$S^{-1}AS = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

(d) Let $A = \begin{bmatrix} 3 & 1 \\ 1 & 7 \end{bmatrix}$. Then A is symmetric and so A is similar to a real diagonal matrix (See Summary part (d)).

(e) Let $A = \begin{bmatrix} 1 & 1 - i \\ 1 + i & 7 \end{bmatrix}$. Then A is a hermitian matrix and so A is similar to a real diagonal matrix (See Summary part (e)).

□

We will consider now the fundamental elements that make up the Jordan canonical form of a matrix.

JORDAN BLOCKS

The reader might recall that in both the "diagonalization" process and the "upper triangularization" process, the order in which the eigenvalues occurred on the diagonal of the resulting matrix was arbitrary in that any order desired could be obtained. The order could be controlled by choosing the eigenvectors in the proper order. So, for example, if a is an eigenvalue of A of multiplicity m, one could arrange to have a appear in the first m entries of the resulting similar upper triangular matrix.

We consider now a special type of matrix that has a single eigenvalue. We will see in Section 6.2 that any matrix (with $p_A(\lambda)$ factoring completely) is similar to a matrix with these special matrices on the diagonal.

A Jordan block is an $m \times m$ matrix J of the form

$$J = \begin{bmatrix} a & 1 & 0 & \dots & 0 & 0 \\ 0 & a & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & a \end{bmatrix}.$$

We say that a is the eigenvalue associated with J, and we see that in the matrix J, each entry on the diagonal is an a and each entry on the "superdiagonal" (the entries above the diagonal) is a 1. All other entries are 0. For example,

$$[3], \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \text{ and } \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

are Jordan blocks, but

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

are not Jordan blocks. It is not difficult to calculate the minimum and characteristic polynomials for a Jordan block.

Theorem 6.1.1. *Let J be an $m \times m$ Jordan block with eigenvalue a. Then $p_J(\lambda) = (-1)^m(\lambda - a)^m$ and $m_J(\lambda) = (\lambda - a)^m$.* (For a Jordan block the characteristic and minimum polynomials are equal, except possibly for sign.)

Proof. Since J is upper triangular, we see that

$$p_J(\lambda) = |J - \lambda I| = (a - \lambda)^m = (-1)^m(\lambda - a)^m.$$

By previous results, we know that $m_J(\lambda)$ is a factor of $p_J(\lambda)$ and so $m_J(\lambda) = (\lambda - a)^k$, where $1 \leq k \leq m$ and k is the least integer satisfying

$$m_J(J) = (J - aI)^k = 0.$$

Now

$$J - aI = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

and we see that

$$(J-aI)^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

$$\vdots$$

$$(J-AI)^{m-1} = \begin{bmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} (J-aI)^m &= (J-aI)(J-aI)^{m-1} \\ &= 0. \end{aligned}$$

From this we see that $m_J(\lambda) = (\lambda - a)^m$. □

Now let us consider the eigenvectors associated with an $m \times m$ Jordan block J with eigenvalue a. Since

$$J - aI = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix},$$

it is not hard to see that the first $m-1$ row vectors are linearly independent and that $J-aI$ has rank $m-1$. From this we see that J has only $m-(m-1)=1$ linearly independent eigenvectors. Let us find conditions under which an $m \times m$ matrix A is similar to a Jordan block.

Let A be an $m \times m$ matrix and assume A is similar to the Jordan block J with a on the diagonal. Then since similar matrices have the same characteristic polynomials, $p_A(\lambda) = (a-\lambda)^m$. Let S be the nonsingular matrix with $S^{-1}AS = J$ and assume $S = [X_1 \dots X_m]$, where X_j is the j-th column vector of S. Then we get $AS = SJ$ and so

$$A[X_1 \dots X_m] = [X_1 \dots X_m] \begin{bmatrix} a & 1 & 0 & \dots & 0 \\ 0 & a & 1 & \dots & 0 \\ 0 & 0 & a & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a \end{bmatrix}.$$

It follows that

$$\left[\begin{array}{cccc} AX_1 & AX_2 & \dots & AX_m \end{array}\right] = \left[\begin{array}{ccccc} aX_1 & X_1 + aX_2 & X_2 + aX_3 & \dots & X_{m-1} + aX_m \end{array}\right]$$

and so

$$\begin{aligned} AX_1 &= aX_1 \\ AX_2 &= X_1 + aX_2 \\ &\vdots \\ AX_m &= X_{m-1} + aX_m. \end{aligned}$$

Rewriting this we obtain

$$\begin{aligned} (A - aI)X_1 &= 0 \\ (A - aI)X_2 &= X_1 \\ (A - aI)X_3 &= X_2 \\ &\vdots \\ (A - aI)X_m &= X_{m-1}. \end{aligned}$$

Notice that X_1 is an eigenvector. The other vectors $X_2, \dots, X_m$ are called **generalized eigenvectors**, and $X_1, \dots, X_m$ is called a **Jordan basis**. This proves one part of the following theorem.

Theorem 6.1.2. *An $m \times m$ matrix A is similar to an $m \times m$ Jordan block J with eigenvalue a if and only if there are independent $m \times 1$ column vectors $X_1, \dots, X_m$ satisfying*

$$\begin{aligned} (A - aI)X_1 &= 0 \\ (A - aI)X_2 &= X_1 \\ (A - aI)X_3 &= X_2 \\ &\vdots \\ (A - aI)X_m &= X_{m-1}. \end{aligned}$$

(A Jordan block corresponds to a string of generalized eigenvectors.)

Proof. See Exercise 11. □

Example 6.1.2. Let $A = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}$. Then $p_A(\lambda) = (\lambda - 2)^2$, so $\lambda = 2$ is an eigenvalue of multiplicity 2. The rank of $A - 2I = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ is 1 and so there is only one independent eigenvector. It follows that A is not similar to a diagonal matrix. Let $X_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Then X_1

is an eigenvector of A associated with the eigenvalue $\lambda = 2$, and there is no other eigenvector independent from X_1. Let us attempt to find a vector X_2 so that X_1, X_2 forms a Jordan basis. We need to solve the equation $(A - 2I)X_2 = X_1$ or

$$\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

We see that $X_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is a solution. Now let $S = [X_1 X_2] = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$. Then

$$S^{-1}AS = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix},$$

which is a Jordan block.

□

Exercises

For each of the matrices in Exercises 1 - 5 determine which are similar to diagonal matrices. Give reasons for your conclusion.

1. $\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$

2. $\begin{bmatrix} 3 & 7 \\ 7 & -2 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$

4. $\begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$

5. $\begin{bmatrix} 1 & i & 1-i \\ -i & 3 & 1 \\ 1+i & 1 & 2 \end{bmatrix}$

6. Which of the following matrices are Jordan blocks? Give reasons.

   (a) $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$

   (b) $[2]$

   (c) $\begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$

   (d) $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

   (e) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

7. Find a Jordan block J that is similar to the matrix $A = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}$.

8. Find a Jordan block J that is similar to the matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

9. For the matrices A and J in Exercise 7, find a matrix S such that $S^{-1}AS = J$.

10. For the matrices A and J in Exercise 8, find a matrix S such that $S^{-1}AS = J$.

11. Prove the remaining part of Theorem 6.1.2.

12. Let A and $S = [X_1 X_2 X_3]$ be 3×3 matrices and assume that

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

Write the relationships satisfied by the matrix A and the column vectors of S.

13. Follow the instructions in Exercises 12 assuming that

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

6.2 The Jordan Canonical Form

As we have observed before, not every matrix is similar to a diagonal matrix. By Theorem 5.4.2, we know that if the characteristic polynomial of a matrix A factors completely, then A is similar to an upper triangular matrix. One wonders if this is the best result that can be obtained. The answer is "no" and in this chapter we investigate this "closest-to-diagonal" matrix that can be obtained by similarity transformations.

Let A be an $n \times n$ matrix and let S be the set of all matrices that are similar to A. If the characteristic polynomial of $A, p_A(\lambda)$, factors completely, then we know that A is similar to an upper triangular matrix U. But this upper triangular matrix U is not unique. For example, suppose $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. Then A itself is upper triangular, and, of course, A is similar to itself so that $A \in S$. But A has two distinct eigenvalues (1 and 2) so that A is similar to the diagonal matrix $D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. It follows that $D \in S$ and so D is a second upper triangular matrix in S. Since A is also similar to the diagonal matrix $D' = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, one can see that "absolute" uniqueness is probably impossible to achieve.

The following theorem identifies a "closest-to-diagonal" matrix J in the class S of matrices that are similar to a given matrix A, and states that this matrix J is unique (more or less). Because of this, J is called a **canonical form**, and being named after its founder, it is called the **Jordan canonical form**. The matrix J is called the Jordan canonical form "of A," and J is said to be "in" Jordan canonical form. In this context the word "canonical" has nothing to do with church law, but rather carries the implication of "simplest" and "unique." The proof is omitted here, but outlined in Appendix F.

Theorem 6.2.1. *Let A be an $n \times n$ matrix and assume that $p_A(\lambda)$ factors completely. Then A is similar to a matrix J of the form*

$$J = \begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_k \end{bmatrix},$$

where $J_1, \ldots, J_k$ are Jordan blocks. The matrix J is unique except for the order of the blocks $J_1, \ldots, J_k$, which can occur in any order. (If the characteristic polynomial factors completely, the matrix is similar to a matrix in Jordan form.)

The above theorem is an "existence" theorem in that it states the existence of a quantity, J in this case, but offers no help in finding it. There are not many parameters in the make-up of the matrix J. We need to know how many blocks, the size of the blocks, and the eigenvalue associated with the blocks. These parameters may often be determined by investigating properties of the original matrix A.

6.2.1 PROPERTIES OF THE JORDAN FORM

It is, in general, difficult to find the Jordan canonical form of a matrix, but knowledge of certain elementary facts simplifies the task. In the following discussion we will assume that A is an $n \times n$ matrix and the characteristic polynomial of A factors completely, say $p_A(\lambda) = (a_1 - \lambda)^{m_1} \ldots (a_s - \lambda)^{m_s}$, where $a_1, \ldots, a_s$ are distinct. Further, let the minimum polynomial of A be $m_A(\lambda) = (\lambda - a_1)^{n_1} \ldots (\lambda - a_s)^{n_s}$. Let J be the Jordan canonical form of A, and assume $J_1, \ldots, J_k$ are the Jordan blocks of J.

Since J and A are similar they have the same characteristic polynomial, and since J is upper triangular, the eigenvalues of J lie on the diagonal. From this it is easy to see that the following theorem is true.

Theorem 6.2.2. *The sum of the orders of the blocks in which a_i occurs on the diagonal is m_i; that is, a_i occurs on the diagonal of J m_i times.* (An eigenvalue of multiplicity m occurs m times on the diagonal of the Jordan form.)

Now let S be a nonsingular matrix such that $S^{-1}AS = J$, or $AS = SJ$. If $S = [X_1 \ldots X_n]$, where X_j is the j-th column of S, then $X_1, \ldots, X_n$ are linearly independent and we have

$$\begin{aligned}
AS &= \begin{bmatrix} AX_1 & \ldots & AX_n \end{bmatrix} \\
&= SJ \\
&= \begin{bmatrix} X_1 & \ldots & X_n \end{bmatrix} \begin{bmatrix} J_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & J_k \end{bmatrix} \\
&= \begin{bmatrix} a_1X_1 & X_1 + a_1X_2 & \ldots & X_r - 1 + a_1X_r & a_2X_r + 1 & X_r + 1 + a_2X_r + 2 & \ldots \end{bmatrix},
\end{aligned}$$

where a_i is the eigenvalue associated with J_i and J_1 is $r \times r$. If we let $A_i = A - a_iI$ and if we equate the columns of AS and SJ, we have

$$\begin{array}{rclcrclcrcl}
AX_1 & = & a_1X_1 & \Longrightarrow & A_1X_1 & = & 0 & & A_2X_r + 1 & = & 0 \\
AX_2 & = & X_1 + a_1X_2 & \Longrightarrow & A_1X_2 & = & X_1 & & A_2X_r + 2 & = & X_{r+1} \\
 & \vdots & & \vdots & & \vdots & & & & \vdots & \\
AX_r & = & X_{r-1} + a_1X_r & \Longrightarrow & A_1X_r & = & X_{r-1} & & & \text{etc.} &
\end{array}$$

A basis of the above form is called a **Jordan basis**. From the above computation one sees that $X_1, X_{r+1}, \ldots$ are linearly independent eigenvectors and there is one of them for each Jordan block of J. We have shown the following theorem:

Theorem 6.2.3. *The number of blocks associated with the eigenvalue a_i is equal to the number of linearly independent eigenvectors associated with a_i.* (There is a block in J for each independent eigenvector.)

Although it is harder to see, the following also holds:

Theorem 6.2.4. *The order of the largest block associated with a_i is n_i, the exponent of $(\lambda - a_i)$ in $m_A(\lambda)$.* (The largest block with a given eigenvalue is multiplicity of the eigenvalue in the minimum polynomial.)

Proof. By assumption, $m_A(\lambda) = (\lambda - a_1)^{n_1} \ldots (\lambda - a_s)^{n_s}$ is the minimum polynomial of A and since similar matrices have the same minimum polynomial (See Exercise 7, Section 5.5), $m_J(\lambda) = m_A(\lambda)$. Notice what happens when one multiplies matrices (assuming the products are defined) that are in "block-diagonal form":

$$\begin{bmatrix} A_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & A_k \end{bmatrix} \begin{bmatrix} B_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & B_k \end{bmatrix} = \begin{bmatrix} A_1B_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & A_kB_k \end{bmatrix}.$$

Assume now that J_1 is the largest block associated with the eigenvalue a_1 and that J_1 is $r \times r$. Recall that by Theorem 6.1.1, the minimum polynomial of J_1 is $(\lambda - a_1)^r$. That is, r is the least power of $(J_1 - a_1 I)$ that is the zero matrix. Now substitute J into $m_J(\lambda)$ and apply these observations:

$$\begin{aligned}
m_J(J) &= (J - a_1 I)^{n_1} \ldots (J - a_1 I)^{n_s} \\
&= \begin{bmatrix} J_1 - a_1 I & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & J_k - a_1 I \end{bmatrix}^{n_1} \cdots \begin{bmatrix} J_1 - a_s I & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & J_k - a_s I \end{bmatrix}^{n_s} \\
&= \begin{bmatrix} (J_1 - a_1 I)^{n_1} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & (J_k - a_1 I)^{n_1} \end{bmatrix} \cdots \begin{bmatrix} (J_1 - a_s I)^{n_s} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & (J_k - a_s I)^{n_s} \end{bmatrix} \\
&= \begin{bmatrix} (J_1 - a_1 I)^{n_1} \ldots (J_1 - a_s I)^{n_s} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & (J_k - a_1 I)^{n_1} \ldots (J_k - a_s I)^{n_s} \end{bmatrix}.
\end{aligned}$$

Now since $m_J(J) = 0$, all blocks on the diagonal of this last matrix must be zero. It follows that

$$(J_1 - a_1 I)^{n_1} \ldots (J_1 - a_s I)^{n_s} = 0.$$

But $a_1 \neq a_2, \ldots, a_s$, and so each of the terms $(J_1 - a_2 I), \ldots, (J_1 - a_s I)$ have nonzero entries on their diagonals. It follows that $(J_1 - a_1 I) n_1 = 0$, and so $n_1 = r$, the order of J_1. Now J_1 was assumed to be the largest block associated with a_1 and so if J_2 also has the eigenvalue a_1, $(J_2 - a_1 I)^{n_1} = 0$. But n_1 is the least integer satisfying this condition, so $n_1 = r$ and the theorem is proved since the order of the blocks may be arbitrarily chosen. $\square$

6.2.2 FINDING THE JORDAN FORM

While it is not in general easy to find the Jordan canonical form of a matrix, the above results provide enough information that it is possible to determine the Jordan form in certain cases. Finding the characteristic and minimum polynomials is the first step. Next, knowing the nature of the associated eigenvectors provides further clues. Often these two steps prove sufficient, but in other cases, one must try to find a Jordan basis. The following examples illustrate the various possibilities.

Example 6.2.1. (a) Assume that A is a matrix such that:

$$p_A(\lambda) = (1 - \lambda)^3 (2 - \lambda)^2$$
$$m_A(\lambda) = (\lambda - 1)^2 (\lambda - 2).$$

Then if J is the Jordan canonical form of A, we know by Theorem 6.2.2 that 1 appears three times on the diagonal of J and 2 appears twice. By Theorem 6.2.4, the order of the largest block associated with the eigenvalue 2 is 1. From this we can see that

$$J = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} & & & 0 \\ & [1] & & \\ & & [2] & \\ 0 & & & [2] \end{bmatrix}.$$

Here we have shown the Jordan blocks that lie on the diagonal and represent the 0's that lie outside the blocks by a single 0.

(b) If A is such that

$$p_A(\lambda) = (1 - \lambda)^3 (2 - \lambda)^2 = -m_A(\lambda),$$

then the largest block with a 2 on the diagonal has order 2 and the largest block with a 1 on the diagonal has order 3. From this we see that the Jordan canonical form is:

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

(c) Assume that A is a 4×4 matrix with

$$p_A(\lambda) = (1-\lambda)^4 \text{ and } m_A(\lambda) = (\lambda - 1)^2.$$

By Theorem 6.2.4, the largest Jordan block associated with the eigenvalue 1 is 2×2. This leaves two possibilities. If A has three independent eigenvectors, then there are three blocks in the Jordan canonical form J and

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

On the other hand, if A has only two independent eigenvectors, then J has two Jordan blocks on its diagonal and so

$$J = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} & 0 \\ 9 & \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \end{bmatrix}.$$

(d) Knowing the characteristic and minimum polynomials and the number of linearly independent eigenvectors may not be sufficient to determine the Jordan canonical form of a matrix. Suppose that A is a 7×7 matrix with

$$p_A(\lambda) = (1-\lambda)^7 \text{ and } m_A(\lambda) = (\lambda - 1)^3.$$

If A has three linearly independent eigenvectors, then the Jordan canonical form J of A must have three Jordan blocks. Because of the minimum polynomial of A, the largest block must be 3×3. These conditions don't determine J. J could consist of a 3×3 Jordan block and two 2×2 blocks or J could have two 3×3 Jordan blocks and one 1×1 block.

(e) Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$. Then $p_A(\lambda) = (1-\lambda)^3$ and so $\lambda = 1$ is the only eigenvalue of A. Let us examine the eigenvectors of A in order to determine the Jordan canonical form of A. Consider

$$(A - 1I)X = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0.$$

We see that $x_2 = 0$ and $x_1 = -x_3$, and so $X_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ is a solution and there is only one linearly independent solution. It follows from Theorem 6.2.3 that there is only one Jordan block in the Jordan canonical form J of A and so

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

To make the example interesting, let us find a matrix S such that $S^{-1}AS = J$. From the previous discussion we know that we want to take $S = [X_1 X_2 X_3]$, where X_1, X_2, X_3 is a Jordan basis; that is, $(A-I)X_1 = 0, (A-I)X_2 = X_1$ and $(A-I)X_3 = X_2$. We compute solutions to these equations:

$$(A-I)X_1 = 0; \text{ take } X_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$(A-I)X_2 = X_1; \text{ take } X_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \text{ or } \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$(A-I)X_3 = X_2; \text{ take } X_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Then X_1, X_2, X_3 are linearly independent and so they form a Jordan basis. If we let

$$S = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 0 \end{bmatrix},$$

then

$$S^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

and $S^{-1}AS = J$.

(f) Finding a Jordan basis is not always straightforward. Suppose that

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

The characteristic polynomial is $(2-\lambda)^3$ and we see that

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

and so $A-2I$ has rank 1. This means that there are two independent eigenvectors and therefore two Jordan blocks, a 2×2 block and a 1×1 block. To find a Jordan basis we must find X_1, X_2, X_3 satisfying

$$\begin{aligned}(A-2I)X_1 &= 0\\ (A-2I)X_2 &= X_1\\ (A-2I)X_3 &= 0\end{aligned}$$

Solving for eigenvectors X_1 and X_3, consider

$$(A-2I)X = \begin{bmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We see that $X_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and $X_3 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ are obvious choices. Now let's find X_2.

We must solve

$$(A-2I)X_2 = X_1$$

or

$$(A-2I)X_2 = \begin{bmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Considering the first and last rows, we see that there's no solution. Perhaps we should switch eigenvectors and try to solve

$$(A-2I)X_2 = X_3$$

or

$$(A-2I)X_2 = \begin{bmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Again, no solution. The theorem above states that our matrix A is similar to a matrix in Jordan form and it implies the existence of a Jordan basis. The theorem doesn't guarantee that the Jordan basis will be easy to find! We need to replace the eigenvector X_1 by another eigenvector so that $(A-2I)X_2 = X_1$ does have a solution. Let's let

$$X_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Then X_1 is and eigenvector and X_2 can be chosen to be $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. The matrix $S = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ will properly transform the matrix A to its Jordan form.

□

Exercises

In Exercises 1 - 6, find the Jordan canonical form of the matrix satisfying the given conditions.

1. A is a 3×3 matrix with $p_A(\lambda) = (-2-\lambda)^3$ and $m_A(\lambda) = (\lambda+2)^2$.
2. A is a 3×3 matrix with $p_A(\lambda) = (2-\lambda)^3$ and $m_A(\lambda) = (\lambda-2)^3$.
3. A is a 3×3 matrix with $p_A(\lambda) = (2-\lambda)^3$ and $(A-2I)^2 = 0$, but $A \neq 2I$.
4. A is a 5×5 matrix with $p_A(\lambda) = (2-\lambda)^3(3-\lambda)^2$ and $m_A(\lambda) = (\lambda-2)^2(\lambda-3)$.
5. A is a 6×6 matrix with $p_A(\lambda) = (2-\lambda)^4(3-\lambda)^2$ and $m_A(\lambda) = (\lambda-2)^2(\lambda-3)^2$. The matrix $A-2I$ has rank 4.
6. A is a 6×6 matrix with $p_A(\lambda) = (1-\lambda)^4(-2-\lambda)^2$ and $m_A(\lambda) = (\lambda+2)^2(\lambda-1)^2$. The matrix $A-I$ has rank 3.

   Find the Jordan canonical form of the matrices in Exercises 7 - 12.

7. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
8. $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$
9. $\begin{bmatrix} 2 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$
10. $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$
11. $\begin{bmatrix} 0 & 1 & 0 \\ -8 & 6 & 1 \\ -1 & -1 & -1 \end{bmatrix}$
12. $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$
13. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix}$. Find a matrix S such that $S^{-1}AS$ is in Jordan canonical form.
14. Find the Jordan canonical form J of $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ and find a nonsingular matrix S with $S^{-1}AS = J$.
15. Give an example of a 3×3 matrix A satisfying: (a) $\lambda = 2$ is the only eigenvalue of A, and (b) there are two linearly independent eigenvectors of A associated with this eigenvalue.

16. Give an example of a 3×3 matrix A satisfying: (a) $\lambda = 1$ is the only eigenvalue of A, and (b) there is only one linearly independent eigenvector of A associated with this eigenvalue.

17. Let A be a 3×3 matrix satisfying: (a) $\lambda = 1$ is the only eigenvalue of A, and (b) there are three linearly independent eigenvectors of A associated with the eigenvalue of $\lambda = 1$. Show that $A = I$.

6.3 Systems of Constant Coefficient Differential Equations (optional)

The Jordan canonical form of a matrix is of use in solving differential equations. In most elementary courses in differential equations, a general discussion of systems of first order, constant coefficient linear differential equations is not included. This omission is mainly due to the unavailability of the Jordan canonical form which is necessary in solving such systems. It seems appropriate to include a discussion of these systems here. In what follows, the reader will find some background in differential equations helpful, but not absolutely necessary.

By a **system of first order constant coefficient linear differential equations** we mean a system of the form

$$\begin{array}{rcccccccc} x_1' & = & a_{11}x_1 & + & \ldots & + & a_{1n}x_n & + & f_1(t) \\ \vdots & & & & & & & & \vdots \\ x_n' & = & a_{n1}x_1 & + & \ldots & + & a_{nn}x_n & + & f_n(t) \end{array} \tag{6.3.1}$$

Of course a **solution** of System 6.3.1 is a collection of functions $x_1(t), \ldots, x_n(t)$ that satisfy the equations on some interval.

These systems of equations arise naturally: If we consider the n-th order constant coefficient linear differential equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1y' + a_0y = g(t), \tag{6.3.2}$$

then this equation can be "reduced" to a system of the form 6.3.1 by making the substitutions

$$\begin{array}{rcl} x_1 & = & y \\ x_2 & = & y' \\ & \vdots & \\ x_n & = & y^{(n-1)} \end{array} \tag{6.3.3}$$

Using the substitutions 6.3.3 in 6.3.2 we obtain

$$\begin{array}{rcl} x_1' & = & x_2 \\ x_2' & = & x_3 \\ & \vdots & \\ x_{n-1}' & = & x_n \\ x_n' & = & -a_{n-1}x_n - \ldots - a_1x_2 - a_0x_1 + g(t), \end{array} \tag{6.3.4}$$

which is a system of equations of the form 6.3.1.

A further natural occurrence of linear systems of differential equations arises in applying Kirchhoff's laws (see Section 1.7) to electrical networks involving inductors and capacitors, applying laws of motion to coupled spring-mass systems, and in other physical situations. To solve a first order system of linear differential equations we will need to be able to solve a single first order equation. We start with a brief review.

6.3.1 SOLVING A FIRST ORDER EQUATION

The general first order linear differential equation is an equation of the form

$$y' + f(t)y = g(t), \tag{6.3.5}$$

where f and g are assumed to be continuous. This equation can be solved by using the integrating factor

$$p(t) = e^{\ \ f(t)dt} \tag{6.3.6}$$

If both sides of Equation 6.3.5 are multiplied by $p(t)$, we obtain

$$y'p(t) + f(t)yp(t) = g(t)p(t). \tag{6.3.7}$$

But notice that

$$\left.\begin{aligned} \frac{d}{dt}(yp(t)) &= y'p(t) + yp'(t) \\ &= y'p(t) + yf(t)p(t) \end{aligned}\right] \tag{6.3.8}$$

since

$$p'(t) = e^{\ \ f(t)dt} = p(t)f(t).$$

Now combining 6.3.7 and 6.3.8, we get

$$\frac{d}{dt}(yp(t)) = g(t)p(t);$$

integrating, we obtain

$$yp(t) = \int g(t)p(t)dt + C$$

or

$$y = \frac{1}{p(t)} \int g(t)p(t)dt + \frac{C}{p(t)}. \tag{6.3.9}$$

The assumption of continuity of the functions $f(t)$ and $g(t)$ guarantees that the integrals in 6.3.6 and 6.3.9 exist. The solution 6.3.9 is called the **general solution** of 6.3.5 and it involves the unknown constant C. An initial condition of the form $y(a) = b$ determines the value of C and gives the unique solution that satisfies this condition.

Example 6.3.1. Consider the equation

$$y' + 2ty = t.$$

Applying the above method we multiply by the integrating factor

$$p(t) = e^{\quad 2tdt} = e^{t^2}$$

and obtain

$$y'e^{t^2} + 2te^{t^2}y = te^{t^2}.$$

Rewriting this equation we get

$$\frac{d}{dt}ye^{t^2} = te^{t^2}$$

and integrating both sides with respect to t gives

$$ye^{t^2} = \int te^{t^2}dt = \frac{e^{t^2}}{2} + C.$$

It follows that

$$y = \frac{1}{2} + Ce^{-t^2}.$$

If a solution is desired that satisfies $y(0) = 2$, then we get

$$2 = \frac{1}{2} + Ce^0$$

so that $C = \dfrac{3}{2}$ and we obtain the specific solution

$$y = \frac{1}{2} + \frac{3}{2}e^{-t^2}.$$

$\square$

Just as matrix notation simplified the expression and solution of systems of linear equations, the same holds true for systems of differential equations.

6.3.2 MATRIX NOTATION

Let us now return to the problem of finding a solution to the system of equations 6.3.1. We introduce matrix notation to write Equations 6.3.1 in a more compact "matrix" differential equation. Let

$$X = X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix},$$

$$A = [a_{ij}],$$

and

$$F = F(t) = \begin{bmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{bmatrix}.$$

For a “matrix” function $X(t)$ we define the derivative X' of X by

$$X'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}.$$

Using this notation, the system 6.3.1 can then be expressed as a matrix differential equation of the form

$$X' = AX + F(t) \tag{6.3.10}$$

6.3.3 USING THE JORDAN FORM

The first step in solving the matrix equation 6.3.10 is as follows: A is similar to a matrix J in Jordan canonical form (J may have complex numbers on its diagonal), say $S^{-1}AS = J$ or $A = SJS^{-1}$. Then

$$X' = SJS^{-1}X + F(t)$$

or

$$S^{-1}X' = JS^{-1}X + S^{-1}F(t).$$

Let $Y = S^{-1}X$ and $G(t) = S^{-1}F(t)$. Then Equation 6.3.10 becomes

$$Y' = JY + G(t). \tag{6.3.11}$$

Now let $J = \begin{bmatrix} J_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_k \end{bmatrix}$, where each J_i is Jordan block. Assume that J_i is $n_i \times n_i$ and let $Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_k \end{bmatrix}$ and $G(t) = \begin{bmatrix} G_1(t) \\ \vdots \\ G_k(t) \end{bmatrix}$, where Y_i and $G_i(t)$ are $n_i \times 1$. Then we get

$$
\begin{aligned}
Y' = \begin{bmatrix} Y_1' \\ \vdots \\ Y_k' \end{bmatrix}
&= \begin{bmatrix} J_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_k \end{bmatrix} \begin{bmatrix} Y_1 \\ \vdots \\ Y_k \end{bmatrix} + \begin{bmatrix} G_1(t) \\ \vdots \\ G_k(t) \end{bmatrix} \\
&= \begin{bmatrix} J_1 Y_1 + G_1(t) \\ \vdots \\ J_k Y_k + G_k(t) \end{bmatrix}.
\end{aligned}
$$

From this we see that we need only solve the systems

$$Y_i' = J_i Y_i + G_i(t) \tag{6.3.12}$$

for $i = 1, \dots, k$, where each J_i is a Jordan block. Given the solutions Y_i, let $Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_k \end{bmatrix}$ and take $X = SY$. X is then the solution of 6.3.10 and so we need only consider a system of equations of the form

$$Z' = JZ + H(t), \tag{6.3.13}$$

where J is a Jordan block. Let $Z = \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}$, $H(t) = \begin{bmatrix} h_1(t) \\ \vdots \\ h_m(t) \end{bmatrix}$, and $J = \begin{bmatrix} a & 1 & 0 & \dots & 0 & 0 \\ 0 & a & 1 & \dots & 0 & 0 \\ 0 & 0 & a & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a & 1 \\ 0 & 0 & 0 & \dots & 0 & a \end{bmatrix}$.

The system 6.3.12 is then equivalent to

$$
\begin{aligned}
z_1' &= a z_1 + z_2 + h_1(t) \\
&\vdots \\
z_{m-1}' &= a z_{m-1} + z_m + h_{m-1}(t) \\
z_m' &= a z_m + h(t).
\end{aligned} \tag{6.3.14}
$$

The problem of solving a system of differential equations has now been reduced to solving a system associated with a single Jordan block. We tackle that problem next.

6.3.4 SOLVING WHEN JORDAN BLOCKS OCCUR

Solving systems determined by Jordan blocks as in 6.3.14 is relatively straightforward. We start from the bottom and work up: solve the last equation for z_m and substitute z_m into the next to the last equation and solve for z_{m-1}, and so forth.. At each step we need to solve a constant coefficient first order linear differential equation of the form $y' - ay = f(t)$.

To solve $y' - ay = f(t)$, we proceed as before: Find the integrating factor $e^{-adt} = e^{-at}$. Multiply the equation by the integrating factor and get

$$y'e^{-at} - ae^{-at}y = e^{-at}f(t).$$

Notice that the left hand side of the equation is the derivative of ye^{-at}, and so we have

$$\frac{d}{dt}(ye^{-at}) = e^{-at}f(t).$$

Now we can integrate both sides to obtain

$$ye^{-at} = \int e^{-at}f(t)dt$$

and solve for y, obtaining $y = e^{at}\int e^{-at}f(t)dt$.

Example 6.3.2. Solve $Z' = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} Z + \begin{bmatrix} e^t \\ e^{2t} \end{bmatrix}$. We must solve the system:

(a) $z_1' = 2z_1 + z_2 + e^t$

(b) $z_2' = 2z_2 + e^{2t}$.

1. Solve (b) first: The integrating factor is e^{-2t}; multiply and get

$$z_2'e^{-2t}z_2 = \frac{d}{dt}z_2e^{-2t}z_2 = \frac{d}{dt}z_2e^{-2t} = e^{-2t}e^{2t} = 1.$$

   Integrating, we get $z_1e^{-2t} = \int 1dt - t + C$, so that $z_2 = te^{2t} + Ce^{2t}$.

2. Substitution into (a) gives

$$z'1 = 2z_1 + te^{2t} + Ce^{2t} + e^t;$$

   now solve for z_1. The integrating factor is e^{-2t}. Multiply, manipulate and get

$$z_1'e^{-2t} - 2e^{-2t}z_1 = \frac{d}{dt}z_1e^{-2t} = t + C + e^{-t}.$$

   Integrating, we get

$$z_1e^{-2t} = \frac{t^2}{2} + Ct - e^{-t} + C'$$

   or

$$z_1 = \frac{t^2}{2}e^{2t} + Cte^{2t} - e^t + C'e^{2t}.$$

It follows that the general solution of the system of equations is:

$$z_1 = \frac{t^2}{2}e^{2t} + Cte^{2t} - e^t + C'e^{2t}$$
$$z_2 = te^{2t} + Ce^{2t}.$$

If a solution satisfying the initial conditions $z_1(0) = 2, z_2(0) = 3$ is desired, we can solve for C and $C' : z_2(0) = 0e^{2\cdot 0} + Ce^{2\cdot 0} = 3$, so $C = 3$. $z_1(0) = \frac{0}{2}e^{2\cdot 0} + C\cdot 0\cdot e^0 - e^0 + C'e^0 = 2$, so $C' = 1$ and we get the following solution of the initial value problem:

$$z_1 = \frac{t^2}{2}e^{2t} + 3te^{2t} - e^t + e^{2t}$$
$$z_2 = te^{2t} + 3e^{2t}.$$

□

The above example illustrates only one step in a rather long and complicated process. To illustrate the entire process, we will go through the details of the following example:

Example 6.3.3. Find the general solution of

$$\begin{aligned} x_1' &= 3x_1 + x_2 + x_3 + 1 \\ x_2' &= 2x_1 + 2x_2 + x_3 + e^t \\ x_3' &= -6x_1 - 3x_2 - 2x_3 + e^{2t}. \end{aligned}$$

We first formulate the problem in matrix form:

$$X' = \begin{bmatrix} 3 & 1 & 1 \\ 2 & 2 & 1 \\ -6 & -3 & -2 \end{bmatrix} X + \begin{bmatrix} 1 \\ e^t \\ e^{2t} \end{bmatrix}$$

and calculate the Jordan canonical form J of the matrix

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 2 & 2 & 1 \\ -6 & -3 & -2 \end{bmatrix}.$$

Now $p_A(\lambda) = (1-\lambda)^3$ and $m_A(\lambda) = (\lambda - 1)^2$ since $(A - 1I)^2 = 0$. We see that

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

We will need to know the matrix S that satisfies $S^{-1}AS = J$. Recall that $S = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}$, where X_1, X_2, X_3 is a Jordan basis. Because of the form of J, these basis vectors must satisfy $(A-I)X_1 = 0, (A-I)X_2 = X_1, (A-I)X_3 = 0$. After some experimentation, we get

$$X_1 = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}, X_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \text{ and } X_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

Thus $S = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ -3 & 1 & 3 \end{bmatrix}$. Also, $S^{-1} = -\frac{1}{2}\begin{bmatrix} -1 & -1 & 0 \\ -4 & -2 & -2 \\ 1 & -1 & 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 \\ 4 & 2 & 2 \\ -1 & 1 & 0 \end{bmatrix}$.

By the previous section, we must solve

$$Y' = JY + G(t),$$

where $Y = S^{-1}X$ and $G(t) = S^{-1}F(t) = \begin{bmatrix} \frac{1}{2}(1+e^t) \\ 2+e^t+e^{2t} \\ \frac{1}{2}(e^t-1) \end{bmatrix}$. We let $Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ and solve the system, which you will recall involves solving the systems determined by the blocks of J. The 1×1 block gives the equation

$$y_3' = y_3 + \frac{1}{2}(e^t - 1).$$

The solution is $y_3 = \frac{1}{2}te^t + Ce^t + \frac{1}{2}$.

The 2×2 block gives the system

$$\begin{aligned} y_1' &= y_1 + y_2 + \frac{1}{2}(1+e^t) \\ y_2' &= y_2 + 2 + e^t + e^{2t}. \end{aligned}$$

Solving the second equation, we get

$$y_2 = -2 + te^t + e^{2t} + C'e^t.$$

Substituting into the first equation gives the equation

$$y_1' - y_1 = y_2 + \frac{1}{2}(1+e^t) = -\frac{3}{2} + te^t + e^{2t} + \left(C' + \frac{1}{2}\right)e^t.$$

Solving this equation we get:

$$\begin{aligned} y_1 &= \frac{3}{2} + \frac{t^2}{2}e^t + e^{2t} + \left(C' + \frac{1}{2}\right)te^t + C''e^t \\ &= \frac{3}{2} + e^t\left(\frac{t^2}{2} + \left(C' + \frac{1}{2}\right)t + C''\right) + e^{2t}. \end{aligned}$$

We have found y_1, y_2 and y_3 above and so we have determined Y. Now $Y = S^{-1}X$, so $X = SY$ and so

$$X = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ -3 & 1 & 3 \end{bmatrix} \begin{bmatrix} \frac{3}{2} + e^t \left(\frac{t^2}{2} + \left(C' + \frac{1}{2} \right) t + C'' \right) + e^{2t} \\ -2 + te^t + e^{2t} + C'e^t \\ \frac{1}{2}te^t + Ce^t + \frac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} 1 + e^t \left(\frac{t^2}{2} + C't + C'' - C \right) + e^{2t} \\ 2 + e^t \left(\frac{t^2}{2} + (C'+1)t + C'' + C \right) + e^{2t} \\ -6 + e^t \left(-\frac{3t^2}{2} - 3C't - 3C'' + C' + C \right) - 2e^{2t} \end{bmatrix}$$

is the general solution of the original system of equations.

To find the specific solution that satisfies the initial conditions $x_1(0) = 0, x_2(0) = 1, x_3(0) = 1$, substitute $t = 0$ and solve for C, C' and C''. We obtain the following system of equations:

$$\begin{array}{rclcl} x_1(0) & = & 0 & = & 1 + C'' - C + 1 \quad \text{or} \quad C'' - C = -2 \\ x_2(0) & = & 1 & = & 2 + C'' + C + 1 \quad \text{or} \quad C'' + C = -2 \\ x_3(0) & = & 1 & = & -6 - 3C'' + C' + C - 2 \quad \text{or} \quad -3C'' + C' + C = 9. \end{array}$$

Solving the system of linear equations, we get $C = 0, C' = 3$, and $C'' = -2$. The specific solution is then

$$X = \begin{bmatrix} 1 + e^t \left(\frac{t^2}{2} + 3t + -2 \right) + e^{2t} \\ 2 + e^t \left(\frac{t^2}{2} + 4t - 2 \right) + e^{2t} \\ -6 + e^t \left(-\frac{3t^2}{2} - 9t + 9 \right) - 2e^{2t} \end{bmatrix}.$$

$\square$

Exercises

In Exercises 1 - 4, find an integrating factor and use it to find the general solution of the given first order linear equation. Then find a specific solution which satisfies the given initial condition.

1. $y' - 2y = e^t, y(0) = 1$
2. $y' + 3y = t, y(0) = -1$
3. $y' - ty = t, y(0) = 0$
4. $y' - t^{-1}y = t^2, y(1) = -1$

In Exercises 5 and 6, express the systems in the form of a matrix equation, that is, find matrices X, A, and $F(t)$ with $X' = AX + F(t)$.

5.

$$\begin{aligned} x_1' &= 2x_1 + x_2 + e^t \\ x'2 &= 3x_2 + e^{2t} \end{aligned}$$

6.

$$\begin{aligned} x_1' &= 3x_1 + x_2 + \sin t \\ x_2' &= 2x_1 - 2x_2 + t \end{aligned}$$

7. Find the general solution of

$$X' = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} X + \begin{bmatrix} t \\ e^t \end{bmatrix}$$

and then find the specific solution that satisfies the following initial conditions: $x_1(0) = 1, x_2(0) = -2$.

8. Find the general solution of

$$X' = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} X + \begin{bmatrix} t^2 \\ t \end{bmatrix}$$

and then find the specific solution that satisfies the following initial conditions: $x_1(0) = -1, x_2(0) = 0$.

9. Solve the following initial value problem:

$$X' = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} X + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$x_1(0) = 0, x_2(0) = 1, x_3(0) = 0$.

10. Find the general solution of the following system of equations:

$$\begin{aligned} x_1' &= 2x_1 - x_2 + x_3 + 1 \\ x_2' &= x_1 + x_3 + e^t \\ x_3' &= x_1 - 2x_2 + 3x_3 + e^{-t} \end{aligned}$$

11. Find the general solution of

$$X' = \begin{bmatrix} -1 & 1 \\ 4 & -1 \end{bmatrix} X.$$