To study systems of first-order equations such as (4.40) systematically it is convenient to introduce vectors. We define $\mathbf{y}$ to be a point in n-dimensional Euclidean space, E_n, with coordinates $(y_1, \dots, y_n)$. Unless otherwise indicated, E_n will represent *real* n-dimensional Euclidean space, that is, the coordinates $(y_1, \dots, y_n)$ of the vector $\mathbf{y}$ are real numbers. However, the entire theory developed here carries over to the complex case with only minor changes, which will be indicated where necessary. We next define functions

$$\hat{f}_j(t, \mathbf{y}) = f_j(t, y_1, \dots, y_n), \quad j = 1, \dots, n$$

and thus the system (4.40) can be written in the form

$$\begin{aligned} y_1' &= \hat{f}_1(t, \mathbf{y}) \\ y_2' &= \hat{f}_2(t, \mathbf{y}) \\ &\vdots \\ y_n' &= \hat{f}_n(t, \mathbf{y}). \end{aligned} \tag{4.45}$$

Proceeding heuristically (we will be more precise below), we next observe that $\hat{f}_1, \dots, \hat{f}_n$ can be regarded as n components of the vector-valued function $\mathbf{f}$ defined by

$$\mathbf{f}(t, \mathbf{y}) = \operatorname{col}(\hat{f}_1(t, \mathbf{y}), \dots, \hat{f}_n(t, \mathbf{y})),$$

where col means column vector. We also define

$$\mathbf{y}' = \operatorname{col}(y_1', \dots, y_n').$$

Thus the system of n first-order equations (4.40) (and all the systems which arose earlier in this section (see also (4.45))) can be written in the very compact form

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}) \tag{4.46}$$

Equation (4.46) resembles the familiar single first-order equation $y' = f(t, y)$, with y, f replaced by the vectors $\mathbf{y}$, $\mathbf{f}$, respectively.

Example 3. We may write the system (4.42) above

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= g(t, y_1, y_2) \end{aligned}$$

as $\mathbf{y}' = \mathbf{f}(t, \mathbf{y})$ with $\mathbf{y} = (y_1, y_2)$ and

$$\begin{aligned} \hat{f}_1(t, \mathbf{y}) &= f_1(t, y_1, y_2) = y_2 \\ \hat{f}_2(t, \mathbf{y}) &= f_2(t, y_1, y_2) = g(t, y_1, y_2) \end{aligned}$$

so that

$$\mathbf{f}(t, \mathbf{y}) = \operatorname{col}(y_2, g(t, y_1, y_2)).$$

The Euclidean length of the vector $\mathbf{y}$ is defined by the relation

$$\|\mathbf{y}\| = (|y_1|^2 + \cdots + |y_n|^2)^{1/2} = \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2}.$$

Notice that $|y_i|$ is well defined for y_i complex and thus $\|\mathbf{y}\|$ is also defined for a complex vector $\mathbf{y}$. We need the notion of length in order to measure distances between solutions of systems. However, for the purpose of dealing with systems such as (4.46) it turns out to be more convenient to define a different quantity for the length (or norm) of a vector $\mathbf{y}$ than the familiar Euclidean length, namely,

$$|\mathbf{y}| = |y_1| + |y_2| + \cdots + |y_n| = \sum_{i=1}^{n} |y_i|.$$

Again, $|\mathbf{y}|$ is well defined for either real or complex vectors $\mathbf{y}$. No confusion need arise from using the absolute value sign for different purposes; on the left-hand side $|\mathbf{y}|$ is the notation for length of the vector $\mathbf{y}$; on the right-hand side we sum the absolute values of the components of $\mathbf{y}$. Observe, for example, if $\mathbf{y} = (3+i, 3-i)$, then $\|\mathbf{y}\| = (|3+i|^2 + |3-i|^2)^{1/2} = (10+10)^{1/2} = (20)^{1/2}$ and $|\mathbf{y}| = |3+i| + |3-i| = (10)^{1/2} + (10)^{1/2} = 2(10)^{1/2}$; clearly $|\mathbf{y}| > \|\mathbf{y}\|$ in this case and in fact, $|\mathbf{y}| = \sqrt{2}\|\mathbf{y}\|$. In general, the quantities $\|\mathbf{y}\|$ and $|\mathbf{y}|$ are related, as follows.

Exercise

4. If $\mathbf{y}$ is an n-dimensional vector, E_n, show that

$$\|\mathbf{y}\| \le |\mathbf{y}| \le \sqrt{n}\,\|\mathbf{y}\|$$

[*Hint*: Use the inequality $2|uv| \le |u|^2 + |v|^2$ and show $\|\mathbf{y}\|^2 \le |\mathbf{y}|^2 \le n\|\mathbf{y}\|^2$.]

The important point about this inequality is that $|\mathbf{y}|$ is small if and only if $\|\mathbf{y}\|$ is small.

The length function $|\mathbf{y}|$ has the following important properties:

i) $|\mathbf{y}| \ge 0$ and $|\mathbf{y}| = 0$ if and only if $\mathbf{y} = \mathbf{0}$.
ii) if c is any complex number, $|c\mathbf{y}| = |c|\,|\mathbf{y}|$.
iii) for all $\mathbf{y}$ and $\mathbf{z}$, $|\mathbf{y}+\mathbf{z}| \le |\mathbf{y}| + |\mathbf{z}|$.

The proofs are immediate from well-known properties of complex numbers. For example, to prove (ii) we have

$$|c\mathbf{y}| = \sum_{j=1}^{n} |cy_j| = \sum_{j=1}^{n} |c|\,|y_j| = |c| \sum_{j=1}^{n} |y_j| = |c|\,|\mathbf{y}|.$$

Similarly for (iii) we use the inequality $|u+v| \le |u| + |v|$ valid for any complex numbers u and v.

Exercise

5. Show that the Euclidean length $\|\mathbf{y}\|$ of a vector $\mathbf{y}$ also satisfies the properties (i), (ii), (iii) above. [*Hint*: To prove (iii) you will need to apply the Schwarz inequality for sums, that is,

$$\left|\sum_{i=1}^{n} a_i b_i\right|^2 \le \sum_{i=1}^{n} |a_i|^2 \sum_{i=1}^{n} |b_i|^2$$

Eigenvalues, Eigenvectors, and Linear Systems with Constant Coefficients

We have seen in Chapter 1 how to solve the scalar equation $y' = ay$, and we know that every solution is of the form $e^{at}c$, where c is a constant. In this chapter we will learn how to find a fundamental matrix of the system $\mathbf{y}' = A\mathbf{y}$, where A is a *constant* $n \times n$ matrix. The explicit calculation of a fundamental matrix will lead us naturally to the study of eigenvalues and eigenvectors of matrices. As in Chapter 4, some knowledge of linear algebra is essential, but for students with this knowledge, this chapter, which contains the results of Sections 3.4 and 3.5 as very special cases, can be studied instead of those sections. We emphasize that the techniques discussed in this chapter are not applicable to systems for which the coefficient matrix is not constant.

5.1 THE EXPONENTIAL OF A MATRIX

In order to find a fundamental matrix of the system

$$\mathbf{y}' = A\mathbf{y}, \tag{5.1}$$

we first need to define *the exponential of a matrix*. If M is a $n \times n$ matrix, we define the matrix $\exp M$ (or e^M) to be the sum of the series

$$\exp M = I + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \cdots + \frac{M^k}{k!} + \cdots = \sum_{k=0}^{\infty} \frac{M^k}{k!}, \tag{5.2}$$

where I is the $n \times n$ identity matrix. (Note that $M^0 = I$ and $0! = 1$.) To justify this definition, we must show that the right-hand side of (5.2) makes sense. It is not difficult to define a suitable notion of convergence of a series of matrices and to show, using this definition, that $\exp M$ is well defined for every matrix $M \in \mathscr{F}_{nn}$. This is done in Appendix 4.

An important property of the exponential matrix is that if $P, M \in \mathscr{F}_{nn}$ and if P and M commute ($MP = PM$), then

$$\exp(M+P) = \exp M \cdot \exp P. \tag{5.3}$$

To prove this, we apply the definition (5.2) to the left-hand side of (5.3). We obtain

$$\exp(M+P) = \sum_{k=0} \frac{(M+P)^k}{k!}. \tag{5.4}$$

By the binomial theorem and $MP = PM$,

$$(M+P)^k = \sum_{l=0}^{k} \frac{k!}{l!(k-l)!} M^l P^{k-l}.$$

(If x, y are real or complex numbers and $k > 0$ is an integer, the binomial theorem states that $(x+y)^k = \sum_{l=0}^{k} k!/[l!(k-l)!]\, x^l y^{k-l}$. If x and y are matrices which commute, the same result holds.) Therefore, canceling $k!$, we obtain

$$\exp(M+P) = \sum_{k=0}^{\infty} \left[\sum_{l=0}^{k} \frac{M^l}{l!} \frac{P^{k-l}}{(k-l)!} \right]. \tag{5.5}$$

On the other hand,

$$\exp M \cdot \exp P = \sum_{i=0}^{\infty} \frac{M^i}{i!} \cdot \sum_{j=0}^{\infty} \frac{P^j}{j!}.$$

By multiplication of absolutely convergent series, we have

$$\exp M \cdot \exp P = \sum_{k=0}^{\infty} C_k$$

$$C_k = \sum_{l=0}^{k} \frac{M^l}{l!} \frac{P^{k-1}}{(k-l)!}. \tag{5.6}$$

Comparison of (5.6) with (5.5) proves (5.3).

A useful property is that if T is a nonsingular $n \times n$ matrix,

$$T^{-1}(\exp M)\, T = \exp(T^{-1} M T). \tag{5.7}$$

Exercises

1. Verify (5.7). [*Hint*: Use (5.2).]
2. If $M \in \mathscr{F}_{nn}$ show that:

   a) $\exp(c_1 M + c_2 M) = \exp c_1 M \exp c_2 M$ for any $c_1, c_2 \in \mathscr{F}$.

b) $(e^{M})^{-1}=e^{-M}$.

c) $(e^{M})^{k}=e^{kM}$, where k is any integer.

d) $e^{0}=I$, where 0 is the $n\times n$ zero matrix.

We are now ready to establish the basic result for linear systems with constant coefficients

$$y'=Ay. \tag{5.1}$$

Theorem 1. The matrix

$$\Phi(t)=\exp At \tag{5.8}$$

is the fundamental matrix of (5.1) *with* $\Phi(0)=I$ *on* $-\infty<t<\infty$.

Proof. That $\Phi(0)=I$ is obvious from (5.2). Using (5.2) with $M=At$ (well defined for $-\infty<t<\infty$ and every $n\times n$ matrix A), we have by differentiation[†]

$$(\exp At)'=A+\frac{A^2t}{1!}+\frac{A^3t^2}{2!}+\cdots+\frac{A^kt^{k-1}}{(k-1!)}+\cdots=A\exp At,$$

$-\infty<t<\infty$. Therefore, exp At is a solution matrix of (5.1) (its columns are solutions of (5.1)). Since $\det\Phi(0)=\det I=1$, Theorem 2, Section 4.3, shows that $\Phi(t)$ is a fundamental matrix of (5.1). ∎

It follows from Theorem 1 and Eq. 4.30 (Section 4.3) that every solution ϕ of the system (5.1) has the form

$$\boldsymbol{\phi}(t)=(\exp At)\,\mathbf{c}\qquad(-\infty<t<\infty) \tag{5.9}$$

for a suitably chosen constant vector **c**.

Exercises

3. Show that if $\boldsymbol{\phi}$ is that solution of (5.1) satisfying $\boldsymbol{\phi}(t_0)=\boldsymbol{\eta}$, then $\boldsymbol{\phi}(t)=[\exp A(t-t_0)]\boldsymbol{\eta}$ $-\infty<t<\infty$.

4. Show that if $\Phi(t)=e^{At}$, then $\Phi^{-1}(t)=e^{-At}$.

We now proceed to find some fundamental matrices in certain special cases; that is, we evaluate $\exp At$ for certain matrices A.

Example 1. Find a fundamental matrix of the system $\mathbf{y}'=A\mathbf{y}$ if A is a diagonal matrix,

$$A=\begin{bmatrix} d_1 & & & 0\\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n\end{bmatrix}.$$

[†] It is easy to prove that the familiar theorems on differentiation of power series (Section 6.2) with real or complex coefficients hold essentially without change for power series having $n\times n$ matrices as coefficients.

From (5.2),

$$\exp At = I + \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix} \frac{t}{1!} + \begin{bmatrix} d_1^2 & & \\ & \ddots & \\ & & d_n^2 \end{bmatrix} \frac{t^2}{2!} + \cdots$$

$$+ \begin{bmatrix} d_1^k & & 0 \\ & \ddots & \\ 0 & & d_n^k \end{bmatrix} \frac{t^k}{k!} + \cdots$$

$$= \begin{bmatrix} \exp d_1 t & & & 0 \\ & \exp d_2 t & & \\ & & \ddots & \\ 0 & & & \exp d_n t \end{bmatrix}$$

and by Theorem 1 this is a fundamental matrix. This result is, of course, obvious, since in the present case each equation of the system is $y_k' = d_k y_k$ ($k = 1, \ldots, n$) and can be integrated separately.

Example 2. Find a fundamental matrix of $\mathbf{y}' = A\mathbf{y}$ if

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

Since

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

and since these two matrices commute, we have

$$\exp At = \exp \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} t \cdot \exp \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} t$$

$$= \begin{bmatrix} e^{3t} & 0 \\ 0 & e^{3t} \end{bmatrix} \left\{ I + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} t + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 \frac{t^2}{2!} + \cdots \right\}.$$

But

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

and the infinite series terminates after two terms. Therefore,

$$\exp At = e^{3t} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

and by Theorem 1 this is a fundamental matrix.

Exercises

5. Find a fundamental matrix of the system $\mathbf{y}'=A\mathbf{y}$ if

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

and check your answer by direct integration of the given system.

6. Find a fundamental matrix of the system $\mathbf{y}'=A\mathbf{y}$ if

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

where A is an $n\times n$ matrix.

7. Find a fundamental matrix of the system $\mathbf{y}'=A\mathbf{y}$, where A is the $n\times n$ matrix

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

8. What is wrong with the following calculation for an arbitrary continuous matrix $A(t)$?

$$\frac{d}{dt}\left[\exp\left(\int_{t_0}^{t} A(s)\,ds\right)\right]=A(t)\exp\left(\int_{t_0}^{t} A(s)\,ds\right),$$

so that $\exp\left(\int_{t_0}^{t} A(s)\,ds\right)$ is a fundamental matrix of $\mathbf{y}'=A(t)\mathbf{y}$ for any continuous matrix $A(t)$.

9. Consider the system

$$t\mathbf{y}'=A\mathbf{y},$$

where A is a constant matrix. Show that $|t|^A=e^{A\log|t|}$ is a fundamental matrix for $t\neq 0$ in two ways: (i) by direct substitution, (ii) by making the change of variable $|t|=e^s$.

5.2 EIGENVALUES AND EIGENVECTORS OF MATRICES

You will have noticed that the examples and exercises presented so far, all of which involve the calculation of e^{At}, are of a rather special form. In order to be able to handle more complicated problems and in order to obtain a general representation of solutions of (5.1) (that is, if we want to evaluate explicitly the entries of the matrix $\exp(At)$, we will need to introduce the notions of eigenvalue and eigenvector of a matrix.

To motivate these concepts, consider the system $\mathbf{y}'=A\mathbf{y}$, and look for a solution of the form

$$\boldsymbol{\phi}(t)=e^{\lambda t}\mathbf{c},\quad \mathbf{c}\neq\mathbf{0},$$

where the constant λ and the vector $\mathbf{c}$ are to be determined. Such a form is suggested by the above examples. Substitution shows that $e^{\lambda t}\mathbf{c}$ is a solution if and only if

$$\lambda e^{\lambda t}\mathbf{c} = Ae^{\lambda t}\mathbf{c}.$$

Since $e^{\lambda t} \neq 0$, this condition becomes

$$(\lambda I - A)\,\mathbf{c} = \mathbf{0}$$

which can be regarded as a linear homogeneous algebraic system for the vector $\mathbf{c}$. By elementary linear algebra (see for example [3], Theorem 1, Section 3.8 and Theorem 2, Section 4.5) this system has a nontrivial solution if and only if λ is chosen in such a way that

$$\det(\lambda I - A) = 0.$$

This suggests the following definitions. Unless otherwise stated we shall assume that our field of scalars is the complex numbers.

Definition 1. *Let A be a real or complex $n \times n$ matrix. An eigenvalue of A is a scalar λ such that the algebraic system*

$$(\lambda I - A)\,\mathbf{x} = \mathbf{0} \tag{5.10}$$

has a nontrivial solution. Any such nontrivial solution of (5.10) *is called an eigenvector of A corresponding to the eigenvalue λ.**

Definition 2. *The polynomial of degree n,*

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

*is called the characteristic polynomial of A.***

Therefore, the calculation preceding Definition 1 shows that $e^{\lambda t}\mathbf{c}$ is a solution of the linear system $\mathbf{y}' = A\mathbf{y}$ if and only if λ is an eigenvalue of A and $\mathbf{c}$ is a corresponding eigenvector. We will return to a discussion of the system $\mathbf{y}' = A\mathbf{y}$ in Section 5.3 after we have become familiar with properties of eigenvalues and eigenvectors.

In view of the remarks immediately preceding Definition 1, the eigenvalues of A are the roots of the polynomial equation $p(\lambda) = 0$. As $p(\lambda)$ is a polynomial of degree n, there are exactly n eigenvalues, not necessarily distinct. In particular, there is at least one eigenvalue and one eigenvector for every matrix A. If $\lambda = \lambda_0$ is a simple root of the equation $p(\lambda) = 0$, then λ_0 is called a *simple eigenvalue*. If $\lambda = \lambda_0$ is a k-fold root of the equation $p(\lambda) = 0$

* Even though the entries of A are real, the scalar λ may be complex (see Example 1 following).

** The function p defined by the expression $p(\lambda) = \det(\lambda I - A)$ is a polynominal of degree n. We shall tacitly assume that such determinantal polynomials obey the rules of determinants.

(that is, $p(\lambda)$ has $(\lambda-\lambda_0)^k$, but not $(\lambda-\lambda_0)^{k+1}$, as a factor), then λ_0 is an *eigenvalue of multiplicity k*. Since the constant term in $p(\lambda)$ is $p(0)=\det(-A)$, if $\lambda=0$ is not an eigenvalue of A, then $p(0)\neq 0$, and in this case A is nonsingular.

Example 1. Find the eigenvalues and corresponding eigenvectors of the matrix

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

The eigenvalues of A are roots of the equation

$$\det(A-\lambda I)=\det\begin{bmatrix} 3-\lambda & 5 \\ -5 & 3-\lambda \end{bmatrix}=\lambda^2-6\lambda+34=0.$$

Thus, $\lambda_{1,2}=3\pm 5i$. The eigenvector

$$\mathbf{u}=\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

corresponding to the eigenvalue $\lambda_1=3+5i$ must satisfy the linear homogeneous algebraic system

$$(A-\lambda_1 I)\,\mathbf{u}=\begin{bmatrix} -5i & 5 \\ -5 & -5i \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}=0.$$

Thus, u_1, u_2 satisfy the system of equations

$$-iu_1+u_2=0$$
$$-u_1-iu_2=0$$

and, therefore,

$$\mathbf{u}=\alpha\begin{bmatrix} 1 \\ i \end{bmatrix}$$

is an eigenvector for any constant α. Similarly, the eigenvector

$$\mathbf{v}=\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

corresponding to the eigenvalue $\lambda_2=3-5i$ is found to be

$$\mathbf{v}=\beta\begin{bmatrix} i \\ 1 \end{bmatrix}$$

for any constant β.

Example 2. Find the eigenvalues of the matrix

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

Consider the equation $\det(\lambda I-A)=0$.

$$\det\begin{bmatrix} \lambda-2 & -1 \\ 1 & \lambda-4 \end{bmatrix}=(\lambda-2)(\lambda-4)+1=\lambda^2-6\lambda+9=0.$$

Thus, $\lambda=3$ is an eigenvalue of A of multiplicity two. To find a corresponding eigenvector we consider the system

$$(3I-A)\mathbf{c}=\mathbf{0}$$

or

$$\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{matrix} c_1-c_2=0 \\ c_1-c_2=0. \end{matrix}$$

Any vector $\mathbf{c}$ with components $c_1=c_2$ is an eigenvector. Thus, the vector

$$\mathbf{c}=\alpha\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where α is any scalar, is an eigenvector corresponding to the eigenvalue $\lambda=3$.

Exercises

1. Compute the eigenvalues and corresponding eigenvectors of each of the following matrices.

a) $\begin{bmatrix} -1 & 2 \\ -2 & -1 \end{bmatrix}$ b) $\begin{bmatrix} -3 & 1 & 7 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$

d) $\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$ e) $\begin{bmatrix} 1 & 0 & 3 \\ 8 & 1 & -1 \\ 5 & 1 & -1 \end{bmatrix}$ f) $\begin{bmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}$

g) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ h) $\begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

i) $\begin{bmatrix} 2 & -3 & 3 \\ 4 & -5 & 3 \\ 4 & -4 & 2 \end{bmatrix}$ (eigenvalues are $-1, 2, -2$)

j) $\begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 2 & 0 & 1 \end{bmatrix}$ (eigenvalues are $-1, -1, 3$)

k) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}$ (eigenvalues are $-1, -2, -3$)

l) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix}$ (eigenvalues are $-1, -1, -2$)

m) $\begin{bmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

n) $\begin{bmatrix} 3 & -1 & -4 & 2 \\ 2 & 3 & -2 & -4 \\ 2 & -1 & -3 & 2 \\ 1 & 2 & -1 & -3 \end{bmatrix}$ [*Hint*: Characteristic polynomial is $(\lambda-1)^2(\lambda+1)^2$.]

2. Show that if A is a triangular matrix of the form

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

the eigenvalues of A are $\lambda=a_{ii}$, where $i=1,\ldots,n$.

You will note that in Example 1 preceding, the two eigenvectors $\mathbf{u}$ and $\mathbf{v}$ are linearly independent if $\alpha\neq 0$ and $\beta\neq 0$, since

$$\det[\mathbf{u},\mathbf{v}]=\det\begin{bmatrix} \alpha & \beta i \\ \alpha i & \beta \end{bmatrix}=2\alpha\beta\neq 0.$$

Therefore, the vectors $\mathbf{u}$ and $\mathbf{v}$ form a basis of (complex) two-dimensional Euclidean space. However, in Example 2, the eigenvectors form only a one-dimensional subspace. In applications to differential equations as well as in matrix theory it is important to know whether the set of all eigenvectors (corresponding to the various eigenvalues) of a given matrix A form a basis. As Example 1 shows, even if the matrix A is real, the eigenvectors may have complex components. Thus, we consider the eigenvectors as vectors with complex components. If the $n\times n$ matrix A has n distinct eigenvalues, the corresponding eigenvectors form a basis for complex n-dimensional Euclidean space.

Theorem 1. *A set of k eigenvectors corresponding to any k distinct eigenvalues is linearly independent.*

Proof. We shall prove the theorem by induction on the number k of eigenvectors. For $k=1$, the result is trivial. Now, assume that every set of $(p-1)$ eigenvectors corresponding to $(p-1)$ distinct eigenvalues of a given matrix A is linearly independent. Let $\mathbf{v}_1,\ldots,\mathbf{v}_p$ be eigenvectors of A corresponding to the eigenvalues $\lambda_1,\ldots,\lambda_p$, respectively, with $\lambda_i\neq\lambda_j$ for $i\neq j$. Suppose that there exist constants $c_1, c_2,\ldots,c_p$, not all zero, such that

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_p\mathbf{v}_p=\mathbf{0}. \tag{5.11}$$

We may assume $c_1\neq 0$. Applying $A-\lambda_1 I$ to both sides of this equation, and using $(A-\lambda_1 I)\,\mathbf{v}_j=(\lambda_j-\lambda_1)\,\mathbf{v}_j (j=1,\ldots,n)$, we obtain

$$c_2(\lambda_2-\lambda_1)\,\mathbf{v}_2+c_3(\lambda_3-\lambda_1)\,\mathbf{v}_3+\cdots+c_p(\lambda_p-\lambda_1)\,\mathbf{v}_p=\mathbf{0}. \tag{5.12}$$

But $\mathbf{v}_2, \mathbf{v}_3,\ldots,\mathbf{v}_p$ are linearly independent by the inductive hypothesis, and therefore $c_j(\lambda_j-\lambda_1)=0$, where $j=2,3,\ldots,p$. Since $\lambda_j\neq\lambda_1$, where $j=2$,

$3, \ldots, p$, we have $c_j = 0$, where $j = 2, 3, \ldots, p$, and (5.11) becomes $c_1\mathbf{v}_1 = 0$. Since $\mathbf{v}_1 \neq 0$, $c_1 = 0$ as well, which shows that $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are linearly independent. This proves the theorem by induction. ∎

We remind you that since the characteristic polynomial of an $n \times n$ matrix is a polynomial of degree n there are at most n distinct eigenvalues. Since the eigenvectors span a subspace of n-dimensional space, there are at most n linearly independent eigenvectors. Of course, in any case, there exists at least one eigenvector, since there is at least one (distinct) eigenvalue.

Example 3. Determine the subspace spanned by the eigenvectors of the matrix A in Example 2.

As we saw in Example 2

$$\mathbf{u} = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \alpha \neq 0$$

is an eigenvector corresponding to the eigenvalue $\lambda = 3$ of A of multiplicity 2 for any value $\alpha \neq 0$. Since $\lambda = 3$ is the only eigenvalue of A, every eigenvector of A is of this form for some $\alpha \neq 0$. Thus, the set of all eigenvectors of A is the subspace of two-dimensional space spanned by the vector

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Clearly, this subspace is the line passing through the point (1, 1) and the origin.

Exercises

3. Determine the subspace, and its dimension, spanned by the eigenvectors of each matrix in Exercise 1.
4. In the matrix A of Exercise 2 assume the diagonal elements a_{ii}, where $i = 1, \ldots, n$, are all distinct. Find the dimension of the subspace spanned by the eigenvectors of A.

5.3 CALCULATION OF A FUNDAMENTAL MATRIX

We have seen in Theorem 1, Section 5.1, that $\exp tA$ is a fundamental matrix of the linear system with constant coefficients, $\mathbf{y}' = A\mathbf{y}$. We have also seen in Examples 1 and 2, Section 5.1, how to compute $\exp tA$ in certain special cases; in particular, we have seen how to compute $\exp tA$ when A is diagonal. We will now show how to compute a fundamental matrix Φ of the system $\mathbf{y}' = A\mathbf{y}$ when A has n linearly independent eigenvectors. This is, in particular, true if all the eigenvalues of A are distinct. We postpone to Section 5.5 consideration of the completely general case of an arbitrary matrix A.

Suppose the matrix A has n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ corresponding to the (not necessarily distinct) eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Motivated by the discussion at the beginning of Section 5.2, we claim that each vector function

$$\boldsymbol{\phi}_j(t)=\exp(\lambda_j t)\,\mathbf{v}_j \qquad j=1,\dots,n$$

is a solution of $\mathbf{y}'=A\mathbf{y}$, on $-\infty<t<\infty$. For,

$$\begin{aligned}\boldsymbol{\phi}_j'(t)&=\exp(\lambda_j t)\,\lambda_j\mathbf{v}_j\\&=\exp(\lambda_j t)\,A\mathbf{v}_j\\&=A\exp(\lambda_j t)\,\mathbf{v}_j\\&=A\boldsymbol{\phi}_j(t)\qquad j=1,\dots,n,\end{aligned}$$

where we have used the fact that $A\mathbf{v}_j=\lambda_j\mathbf{v}_j, j=1,\dots,n$. Define

$$\Phi(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\dots,\boldsymbol{\phi}_n(t)].$$

Since each column of Φ is a solution of $\mathbf{y}'=A\mathbf{y}$, Φ is a solution matrix of $\mathbf{y}'=A\mathbf{y}$ on $-\infty<t<\infty$. We have

$$\det\Phi(0)=\det[\mathbf{v}_1,\dots,\mathbf{v}_n]\neq 0,$$

because the vectors $\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n$ are linearly independent. It now follows from Theorem 2, Section 4.3, that det $\Phi(t)\neq 0$ for $-\infty<t<\infty$ and that $\Phi(t)$ is a fundamental matrix of $\mathbf{y}'=A\mathbf{y}$ on $-\infty<t<\infty$. We have therefore proved the following result.

Theorem 1. *Let A be a constant matrix (real or complex). Suppose $\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n$ are n linearly independent eigenvectors corresponding respectively to the eigenvalues $\lambda_1,\lambda_2,\dots,\lambda_n$. Then*

$$\Phi(t)=[\exp(\lambda_1 t)\,\mathbf{v}_1,\exp(\lambda_2 t)\,\mathbf{v}_2,\dots,\exp(\lambda_n t)\,\mathbf{v}_n]$$

is a fundamental matrix of the linear system with constant coefficients $\mathbf{y}'=A\mathbf{y}$ on $-\infty<t<\infty$. In particular this is the case if the eigenvalues $\lambda_1,\lambda_2,\dots,\lambda_n$ are distinct.

Example 1. Find a fundamental matrix of the system $\mathbf{y}'=A\mathbf{y}$ if

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

By Example 1, Section 5.2, $\lambda_1=3+5i$ and $\lambda_2=3-5i$ are eigenvalues of A and

$$\mathbf{v}_1=\begin{bmatrix}1\\i\end{bmatrix},\qquad \mathbf{v}_2=\begin{bmatrix}i\\1\end{bmatrix}$$

are (linearly independent) eigenvectors corresponding to λ_1, λ_2, respectively. By Theorem 1

$$\Phi(t)=\begin{bmatrix}e^{(3+5i)t}&ie^{(3-5i)t}\\ie^{(3+5i)t}&e^{(3-5i)t}\end{bmatrix}$$

is a fundamental matrix on $-\infty<t<\infty$.

In general, Theorem 1 does not yield exp tA, even though it does yield a fundamental matrix $\Phi(t)$ of $y'=Ay$. By Corollary 2 to Theorem 2, Section 4.3, since exp tA and $\Phi(t)$ are both fundamental matrices of $y'=Ay$ on $-\infty<t<\infty$, there exists a nonsingular matrix C such that

$$\exp tA=\Phi(t)\,C. \tag{5.13}$$

Setting $t=0$ in (5.13), we obtain $C=\Phi^{-1}(0)$. Thus,

$$\exp tA=\Phi(t)\,\Phi^{-1}(0). \tag{5.14}$$

Example 2. Find exp tA if A is the matrix in Example 1.

By (5.14), Example 1 and Example 3, Section 5.2, we have successively

$$\begin{aligned}\exp tA&=\begin{bmatrix} e^{(3+5i)t} & ie^{(3-5i)t} \\ ie^{(3+5i)t} & e^{(3-5i)t}\end{bmatrix}\begin{bmatrix} 1 & i \\ i & 1\end{bmatrix}^{-1}\\
&=\frac{1}{2}\begin{bmatrix} e^{(3+5i)t} & ie^{(3-5i)t} \\ ie^{(3+5i)t} & e^{(3-5i)t}\end{bmatrix}\begin{bmatrix} 1 & -i \\ -i & 1\end{bmatrix}\\
&=\frac{1}{2}\begin{bmatrix} e^{(3+5i)t}+e^{(3-5i)t} & -i(e^{(3+5i)t}-e^{(3-5i)t}) \\ i(e^{(3+5i)t}-e^{(3-5i)t}) & e^{(3+5i)t}+e^{(3-5i)t}\end{bmatrix}\\
&=e^{3t}\begin{bmatrix} \cos 5t & \sin 5t \\ -\sin 5t & \cos 5t\end{bmatrix}.\end{aligned}$$

If A is real, exp tA is real from the definition (5.2). Thus, Eq. 5.14 gives at the same time a way of constructing a real fundamental matrix, whenever A is real. Example 2 is a special case of this remark.

Exercises

1. Find a fundamental matrix of the system $y'=Ay$; also find exp tA for each of the following coefficient matrices.

a) $A=\begin{bmatrix} 1 & 2 \\ 4 & 3\end{bmatrix}$ (see Exercise 1(c), Section 5.2)

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

c) $A=\begin{bmatrix} 1 & 0 & 3 \\ 8 & 1 & -1 \\ 5 & 1 & -1\end{bmatrix}$

d) $A=\begin{bmatrix} 2 & -3 & 3 \\ 4 & -5 & 3 \\ 4 & -4 & 2\end{bmatrix}$ (see Exercise 1(i), Section 5.2)

e) $A=\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6\end{bmatrix}$ (see Exercise 1(k), Section 5.2)

2. Show that the scalar second-order differential equation $u'' + pu' + qu = 0$ is equivalent to the system $\mathbf{y}' = A\mathbf{y}$ with

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

and compute the eigenvalues λ_1, λ_2 of A.

3. Compute a fundamental matrix for the system in Exercise 2 if $\lambda_1 \neq \lambda_2$, that is, if $p^2 \neq 4q$, and construct the general solution of the scalar second-order equation.

An alternative way of producing a real fundamental matrix if A is a real 2×2 matrix is contained in the following exercises.

Exercises

4. Given the matrix

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

show that $A^2 = -I$, $A^3 = -A$, $A^4 = I$ and compute A^m, where m is an arbitrary positive integer.

5. Use the result of Exercise 4 and the definition (5.2) to show that

$$e^{tA} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}.$$

$$\left[\textit{Hint}: \cos t = 1 - \frac{t^2}{2!} + \frac{t^4}{4!} + \cdots, \qquad \sin t = t - \frac{t^3}{3!} + \frac{t^5}{5!} - \cdots.\right]$$

6. Compute e^{tA}, if

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

[*Hint*: Use Exercises 4 and 5.]

We close this section with the solution of the nonhomogeneous system

$$\mathbf{y}' = A\mathbf{y} + \mathbf{g}(t), \tag{5.15}$$

where A is a constant matrix and $\mathbf{g}$ is a given continuous function on $-\infty < t < \infty$. The variation of constants formula (Theorem 1, Section 4.4) with $\Phi(t) = \exp tA$ as a fundamental matrix of the homogeneous system now becomes particularly simple in appearance. We have $\Phi^{-1}(s) = \exp(-sA)$, $\Phi(t)\,\Phi^{-1}(s) = \exp[(t-s)A]$; if the initial condition is $\boldsymbol{\phi}(t_0) = \boldsymbol{\eta}$, $\boldsymbol{\phi}_h(t) = \exp[(t-t_0)A]\,\boldsymbol{\eta}$ and the solution of (5.15) is

$$\boldsymbol{\phi}(t) = \exp[(t-t_0)A]\,\boldsymbol{\eta} + \int_{t_0}^{t} \exp[(t-s)A]\,\mathbf{g}(s)\,ds \qquad -\infty < t < \infty, \tag{5.16}$$

where e^{tA} is the fundamental matrix of the homogeneous system that we can construct by the method shown in this section. Note how easy it is to compute the inverse of Φ and also $\Phi(t)\,\Phi^{-1}(s)$ in this case. However, it may not be possible to evaluate the integral in (5.16) explicitly except in special cases.

Example 3. Find the solution ϕ of the system $\mathbf{y}' = A\mathbf{y} + \mathbf{g}(t)$ satisfying the initial condition

$$\boldsymbol{\phi}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{if} \quad A = \begin{bmatrix} 3 & 5 \\ -5 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{g}(t) = \begin{bmatrix} e^{-t} \\ 0 \end{bmatrix}.$$

From Example 2 preceding, we have

$$\exp tA = e^{3t} \begin{bmatrix} \cos 5t & \sin 5t \\ -\sin 5t & \cos 5t \end{bmatrix}.$$

Substituting in (5.16), we obtain (using $t_0 = 0$)

$$\boldsymbol{\phi}(t) = e^{3t} \begin{bmatrix} \cos 5t & \sin 5t \\ -\sin 5t & \cos 5t \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \int_0^t e^{3(t-s)} \begin{bmatrix} \cos 5(t-s) & \sin 5(t-s) \\ -\sin 5t(t-s) & \cos 5(t-s) \end{bmatrix} \begin{bmatrix} e^{-s} \\ 0 \end{bmatrix} ds$$

$$= e^{3t} \begin{bmatrix} \sin 5t \\ \cos 5t \end{bmatrix} + \int_0^t e^{3(t-s)} e^{-s} \begin{bmatrix} \cos 5(t-s) \\ -\sin 5(t-s) \end{bmatrix} ds.$$

In this case, we can evaluate the integrals as follows

$$\boldsymbol{\phi}(t) = e^{3t} \begin{bmatrix} \sin 5t \\ \cos 5t \end{bmatrix} + e^{3t} \int_0^t e^{-4s} \begin{bmatrix} \cos 5t \cos 5s + \sin 5t \sin 5s \\ -\sin 5t \cos 5s + \cos 5t \sin 5s \end{bmatrix} ds.$$

Using the formulas (these can be found by integration by parts)

$$\int_0^t e^{-4s} \cos 5s \, ds = \frac{e^{-4s}}{16+25} (-4 \cos 5s + 5 \sin 5s) \Big|_{s=0}^{s=t}$$

$$\int_0^t e^{-4s} \sin 5s \, ds = \frac{e^{-4s}}{16+25} (-4 \sin 5s - 5 \cos 5s) \Big|_{s=0}^{s=t}$$

we obtain

$$\boldsymbol{\phi}(t) = e^{3t} \begin{bmatrix} \sin 5t \\ \cos 5t \end{bmatrix} + e^{3t} \begin{bmatrix} \cos 5t \left\{ \frac{e^{-4t}}{41} (-4 \cos 5t + 5 \sin 5t) + \frac{4}{41} \right\} \\ \quad + \sin 5t \left\{ \frac{e^{-4t}}{41} (-4 \sin 5t - 5 \cos 5t) + \frac{5}{41} \right\} \\ -\sin 5t \left\{ \frac{e^{-4t}}{41} (-4 \cos 5t + 5 \sin 5t) + \frac{4}{41} \right\} \\ \quad + \cos 5t \left\{ \frac{e^{-4t}}{41} (-4 \sin 5t - 5 \cos 5t) + \frac{5}{41} \right\} \end{bmatrix}.$$

Further simplification seems pointless. You will note that even such a simple example as above leads to a rather complicated answer.

Exercises

7. Find the solution $\boldsymbol{\phi}$ of the system

$$\mathbf{y}' = A\mathbf{y} + \mathbf{g}(t)$$

in each of the following cases:

a) $\boldsymbol{\phi}(0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$, $\mathbf{g}(t) = \begin{bmatrix} e^t \\ 1 \end{bmatrix}$ (see Exercise 1a).

b) $\boldsymbol{\phi}(0) = 0$, $A = \begin{bmatrix} 0 & 1 & -0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}$, $\mathbf{g}(t) = \begin{bmatrix} 0 \\ 0 \\ e^{-t} \end{bmatrix}$

(see Exercise 1e).

c) $\boldsymbol{\phi}(1) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $A = \begin{bmatrix} 2 & -3 & 3 \\ 4 & -5 & 3 \\ 4 & -4 & 2 \end{bmatrix}$, $\mathbf{g}(t)$ arbitrary

(see Exercise 1d).

8. By converting to an equivalent system, find the general solution of the scalar equation

$$y'' - y = f(t),$$

where f is continuous, by using the theory of this section.

9. Use the results of this section and Exercise 5 to find the general solution of the scalar equation

$$y'' + y = f(t),$$

where f is continuous.

10. Suppose m is not an eigenvalue of the matrix A. Show that the nonhomogeneous system

$$\mathbf{y}' = A\mathbf{y} + \mathbf{c}e^{mt}$$

has a solution of the form

$$\boldsymbol{\phi}(t) = \mathbf{p}e^{mt}$$

and calculate the vector $\mathbf{p}$ in terms of A and $\mathbf{c}$.

11. Suppose m is not an eigenvalue of the matrix A. Show that the nonhomogeneous system

$$\mathbf{y}' = A\mathbf{y} + \sum_{j=0}^{k} \mathbf{c}_j t^j e^{mt}$$

has a solution of the form

$$\boldsymbol{\phi}(t) = \sum_{j=0}^{k} \mathbf{p}_j t^j e^{mt}.$$