

Chapter 8

Eigenvalues

So far, our applications have concentrated on statics: unchanging equilibrium configurations of physical systems, including mass/spring chains, circuits, and structures, that are modeled by linear systems of algebraic equations. It is now time to set our universe in motion. In general, a *dynamical system* refers to the (differential) equations governing the temporal behavior of some physical system: mechanical, electrical, chemical, fluid, etc. Our immediate goal is to understand the behavior of the simplest class of linear dynamical systems — first order autonomous linear systems of ordinary differential equations. As always, complete analysis of the linear situation is an essential prerequisite to making progress in the vastly more complicated nonlinear realm.

We begin with a very quick review of the scalar case, where solutions are simple exponential functions. We use the exponential form of the scalar solution as a template for a possible solution in the vector case. Substituting the resulting formula into the system leads us immediately to the equations defining the so-called eigenvalues and eigenvectors of the coefficient matrix. Further developments will require that we become familiar with the basic theory of eigenvalues and eigenvectors, which prove to be of absolutely fundamental importance in both the mathematical theory and its applications, including numerical algorithms.

The present chapter develops the most important properties of eigenvalues and eigenvectors. The applications to dynamical systems will appear in Chapter 9, while applications to iterative systems and numerical methods is the topic of Chapter 10. Extensions of the eigenvalue concept to differential operators acting on infinite-dimensional function space, of essential importance for solving linear partial differential equations modelling continuous dynamical systems, will be covered in later chapters. Each square matrix has a collection of one or more complex scalars called eigenvalues and associated vectors, called eigenvectors. Viewing the matrix as a linear transformation, the eigenvectors indicate directions of pure stretch and the eigenvalues the degree of stretching. Most matrices are complete, meaning that their (complex) eigenvectors form a basis of the underlying vector space. When expressed in terms of its eigenvector basis, the matrix assumes a very simple diagonal form, and the analysis of its properties becomes extremely simple. A particularly important class are the symmetric matrices, whose eigenvectors form an orthogonal basis of \mathbb{R}^n ; in fact, this is by far the most common way for orthogonal bases to appear. Incomplete matrices are trickier, and we relegate them and their associated non-diagonal Jordan canonical form to the final section. The numerical computation of eigenvalues and eigenvectors is a challenging issue, and must be deferred until Section 10.6. *Note:* Unless you are prepared to consult that section now, solving the computer-based problems in this chapter will require

access to a computer package that can accurately compute eigenvalues and eigenvectors.

A non-square matrix A does not have eigenvalues. As an alternative, the square roots of the eigenvalues of associated square Gram matrix $K = A^T A$ serve to define its singular values. Singular values and principal component analysis appear in an increasingly broad range of modern applications, including statistics, data mining, image processing, semantics, language and speech recognition, and learning theory. The singular values are used to define the condition number of a matrix, quantifying the difficulty of accurately solving an associated linear system.

8.1. Simple Dynamical Systems.

The purpose of this section is to motivate the concepts of eigenvalue and eigenvector by attempting to solve the simplest class of dynamical systems — first order homogeneous linear systems of ordinary differential equations. We begin with a review of the scalar case, introducing basic notions of stability in preparation for the general version, to be treated in depth in Chapter 9. Readers who are uninterested in such motivational material may skip ahead to Section 8.2 without penalty.

Scalar Ordinary Differential Equations

Consider the elementary scalar ordinary differential equation

$$\frac{du}{dt} = a u. \quad (8.1)$$

Here $a \in \mathbb{R}$ is a real constant, while the unknown $u(t)$ is a scalar function. As you learned in first year calculus, the general solution to (8.1) is an exponential function

$$u(t) = c e^{at}. \quad (8.2)$$

The integration constant c is uniquely determined by a single *initial condition*

$$u(t_0) = b \quad (8.3)$$

imposed at an initial time t_0 . Substituting $t = t_0$ into the solution formula (8.2),

$$u(t_0) = c e^{at_0} = b, \quad \text{and so} \quad c = b e^{-at_0}.$$

We conclude that

$$u(t) = b e^{a(t-t_0)}. \quad (8.4)$$

is the unique solution to the scalar initial value problem (8.1), (8.3).

Example 8.1. The radioactive decay of an isotope, say Uranium 238, is governed by the differential equation

$$\frac{du}{dt} = -\gamma u. \quad (8.5)$$

Here $u(t)$ denotes the amount of the isotope remaining at time t , and the coefficient $\gamma > 0$ governs the decay rate. The solution is given by an exponentially decaying function $u(t) = c e^{-\gamma t}$, where $c = u(0)$ is the initial amount of radioactive material.

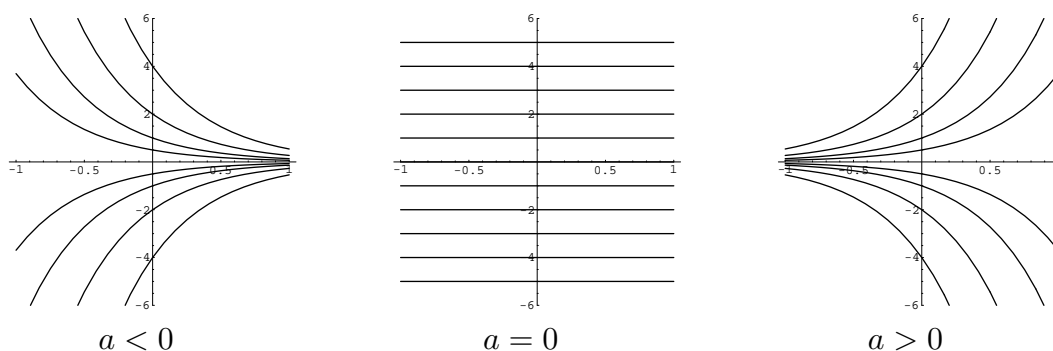


Figure 8.1. Solutions to $\dot{u} = a u$.

The isotope's *half-life* t^* is the time it takes for half of a sample to decay, that is when $u(t^*) = \frac{1}{2}u(0)$. To determine t^* , we solve the algebraic equation

$$e^{-\gamma t^*} = \frac{1}{2}, \quad \text{so that} \quad t^* = \frac{\log 2}{\gamma}. \quad (8.6)$$

Let us make some elementary, but pertinent observations about this simple linear dynamical system. First of all, since the equation is homogeneous, the zero function $u(t) \equiv 0$ (corresponding to $c = 0$) is a constant solution, known as an *equilibrium solution* or *fixed point*, since it does not depend on t . If the coefficient $a > 0$ is positive, then the solutions (8.2) are exponentially growing (in absolute value) as $t \rightarrow +\infty$. This implies that the zero equilibrium solution is *unstable*. The initial condition $u(t_0) = 0$ produces the zero solution, but if we make a tiny error (either physical, numerical, or mathematical) in the initial data, say $u(t_0) = \varepsilon$, then the solution $u(t) = \varepsilon e^{a(t-t_0)}$ will eventually get very far away from equilibrium. More generally, any two solutions with very close, but not equal, initial data, will eventually become arbitrarily far apart: $|u_1(t) - u_2(t)| \rightarrow \infty$ as $t \rightarrow \infty$. One consequence is an inherent difficulty in accurately computing the long time behavior of the solution, since small numerical errors will eventually have very large effects.

On the other hand, if $a < 0$, the solutions are exponentially decaying in time. In this case, the zero solution is *stable*, since a small error in the initial data will have a negligible effect on the solution. In fact, the zero solution is *globally asymptotically stable*. The phrase “asymptotically stable” implies that solutions that start out near equilibrium eventually return; more specifically, if $u(t_0) = \varepsilon$ is small, then $u(t) \rightarrow 0$ as $t \rightarrow \infty$. “Globally” implies that all solutions, no matter how large the initial data is, return to equilibrium. In fact, for a linear system, the stability of an equilibrium solution is inevitably a global phenomenon.

The borderline case is when $a = 0$. Then all the solutions to (8.1) are constant. In this case, the zero solution is *stable* — indeed, *globally stable* — but not asymptotically stable. The solution to the initial value problem $u(t_0) = \varepsilon$ is $u(t) \equiv \varepsilon$. Therefore, a solution that starts out near equilibrium will remain near, but will not asymptotically return. The three qualitatively different possibilities are illustrated in Figure 8.1.

First Order Dynamical Systems

The simplest class of *dynamical systems* consist of n first order ordinary differential

equations for n unknown functions

$$\frac{du_1}{dt} = f_1(t, u_1, \dots, u_n), \quad \dots \quad \frac{du_n}{dt} = f_n(t, u_1, \dots, u_n),$$

which depend on a scalar variable $t \in \mathbb{R}$, which we usually view as time. We will often write the system in the equivalent vector form

$$\frac{d\mathbf{u}}{dt} = \mathbf{f}(t, \mathbf{u}). \quad (8.7)$$

A vector-valued solution $\mathbf{u}(t) = (u_1(t), \dots, u_n(t))^T$ serves to parametrize a curve in \mathbb{R}^n , called a *solution trajectory*. A dynamical system is called *autonomous* if the time variable t does not appear explicitly on the right hand side, and so has the system has the form

$$\frac{d\mathbf{u}}{dt} = \mathbf{f}(\mathbf{u}). \quad (8.8)$$

Dynamical systems of ordinary differential equations appear in an astonishing variety of applications, and have been the focus of intense research activity since the early days of calculus.

We shall concentrate most of our attention on the very simplest case: a homogeneous, linear, autonomous dynamical system

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u}, \quad (8.9)$$

in which A is a constant $n \times n$ matrix. In full detail, the system consists of n linear ordinary differential equations

$$\begin{aligned} \frac{du_1}{dt} &= a_{11}u_1 + a_{12}u_2 + \dots + a_{1n}u_n, \\ \frac{du_2}{dt} &= a_{21}u_1 + a_{22}u_2 + \dots + a_{2n}u_n, \\ &\vdots \\ \frac{du_n}{dt} &= a_{n1}u_1 + a_{n2}u_2 + \dots + a_{nn}u_n, \end{aligned} \quad (8.10)$$

involving n unknown functions $u_1(t), u_2(t), \dots, u_n(t)$. In the autonomous case, the coefficients a_{ij} are assumed to be (real) constants. We seek not only to develop basic solution techniques for such dynamical systems, but to also understand their behavior from both a qualitative and quantitative standpoint.

Drawing our inspiration from the exponential solution formula (8.2) in the scalar case, let us investigate whether the vector system admits any solutions of a similar exponential form

$$\mathbf{u}(t) = e^{\lambda t} \mathbf{v}. \quad (8.11)$$

Here λ is a constant scalar, so $e^{\lambda t}$ is a scalar function of t , while $\mathbf{v} \in \mathbb{R}^n$ is a constant vector. In other words, the components $u_i(t) = v_i e^{\lambda t}$ of our desired solution are assumed to be

constant multiples of the *same* exponential function. Since \mathbf{v} is assumed to be constant, the derivative of $\mathbf{u}(t)$ is easily found:

$$\frac{d\mathbf{u}}{dt} = \frac{d}{dt} (e^{\lambda t} \mathbf{v}) = \lambda e^{\lambda t} \mathbf{v}.$$

On the other hand, since $e^{\lambda t}$ is a scalar, it commutes with matrix multiplication, and so

$$A\mathbf{u} = A e^{\lambda t} \mathbf{v} = e^{\lambda t} A\mathbf{v}.$$

Therefore, $\mathbf{u}(t)$ will solve the system (8.9) if and only if

$$\lambda e^{\lambda t} \mathbf{v} = e^{\lambda t} A\mathbf{v},$$

or, canceling the common scalar factor $e^{\lambda t}$,

$$\lambda \mathbf{v} = A\mathbf{v}.$$

The result is a system of algebraic equations relating the vector \mathbf{v} and the scalar λ . Analysis of this system and its ramifications will be the topic of the remainder of this chapter. After gaining a complete understanding, we will return, suitably armed, to the main issue, solution of linear dynamical systems, in Chapter 9.

8.2. Eigenvalues and Eigenvectors.

We inaugurate our discussion of eigenvalues and eigenvectors with the fundamental definition.

Definition 8.2. Let A be an $n \times n$ matrix. A scalar λ is called an *eigenvalue* of A if there is a *non-zero* vector $\mathbf{v} \neq \mathbf{0}$, called an *eigenvector*, such that

$$A\mathbf{v} = \lambda \mathbf{v}. \tag{8.12}$$

Thus, the matrix A effectively stretches the eigenvector \mathbf{v} by an amount specified by the eigenvalue λ . In this manner, the eigenvectors specify the directions of pure stretch for the linear transformation defined by the matrix A .

Remark: The odd-looking terms “eigenvalue” and “eigenvector” are hybrid German–English words. In the original German, they are *Eigenwert* and *Eigenvektor*, which can be fully translated as “proper value” and “proper vector”. For some reason, the half-translated terms have acquired a certain charm, and are now standard. The alternative English terms *characteristic value* and *characteristic vector* can be found in some (mostly older) texts. Oddly, the term *characteristic equation*, to be defined below, is still used.

The requirement that the eigenvector \mathbf{v} be nonzero is important, since $\mathbf{v} = \mathbf{0}$ is a trivial solution to the eigenvalue equation (8.12) for *any* scalar λ . Moreover, as far as solving linear ordinary differential equations goes, the zero vector $\mathbf{v} = \mathbf{0}$ only gives the trivial zero solution $\mathbf{u}(t) \equiv \mathbf{0}$.

The eigenvalue equation (8.12) is a system of linear equations for the entries of the eigenvector \mathbf{v} — provided the eigenvalue λ is specified in advance — but is “mildly”

nonlinear as a combined system for λ and \mathbf{v} . Gaussian elimination per se will not solve the problem, and we are in need of a new idea. Let us begin by rewriting the equation in the form

$$(A - \lambda I)\mathbf{v} = \mathbf{0}, \quad (8.13)$$

where I is the identity matrix of the correct size[†]. Now, for given λ , equation (8.13) is a homogeneous linear system for \mathbf{v} , and always has the trivial zero solution $\mathbf{v} = \mathbf{0}$. But we are specifically seeking a nonzero solution! According to Theorem 1.45, a homogeneous linear system has a nonzero solution $\mathbf{v} \neq \mathbf{0}$ if and only if its coefficient matrix, which in this case is $A - \lambda I$, is singular. This observation is the key to resolving the eigenvector equation.

Theorem 8.3. *A scalar λ is an eigenvalue of the $n \times n$ matrix A if and only if the matrix $A - \lambda I$ is singular, i.e., of rank $< n$. The corresponding eigenvectors are the nonzero solutions to the eigenvalue equation $(A - \lambda I)\mathbf{v} = \mathbf{0}$.*

We know a number of ways to characterize singular matrices, including the vanishing determinant criterion given in (det0■). Therefore, the following result is an immediate corollary.

Proposition 8.4. *A scalar λ is an eigenvalue of the matrix A if and only if λ is a solution to the characteristic equation*

$$\det(A - \lambda I) = 0. \quad (8.14)$$

In practice, when finding eigenvalues and eigenvectors by hand, one first solves the characteristic equation (8.14). Then, for each eigenvalue λ one uses standard linear algebra methods, i.e., Gaussian elimination, to solve the corresponding linear system (8.13) for the eigenvector \mathbf{v} .

Example 8.5. Consider the 2×2 matrix

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

We compute the determinant in the characteristic equation using (1.34):

$$\det(A - \lambda I) = \det \begin{pmatrix} 3 - \lambda & 1 \\ 1 & 3 - \lambda \end{pmatrix} = (3 - \lambda)^2 - 1 = \lambda^2 - 6\lambda + 8.$$

Thus, the characteristic equation is a quadratic polynomial equation, and can be solved by factorization:

$$\lambda^2 - 6\lambda + 8 = (\lambda - 4)(\lambda - 2) = 0.$$

We conclude that A has two eigenvalues: $\lambda_1 = 4$ and $\lambda_2 = 2$.

[†] Note that it is not legal to write (8.13) in the form $(A - \lambda)\mathbf{v} = \mathbf{0}$ since we do not know how to subtract a scalar λ from a matrix A . Worse, if you type $A - \lambda$ in MATLAB, it will subtract λ from *all* the entries of A , which is *not* what we are after!

For each eigenvalue, the corresponding eigenvectors are found by solving the associated homogeneous linear system (8.13). For the first eigenvalue, the eigenvector equation is

$$(A - 4I)\mathbf{v} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \text{or} \quad \begin{aligned} -x + y &= 0, \\ x - y &= 0. \end{aligned}$$

The general solution is

$$x = y = a, \quad \text{so} \quad \mathbf{v} = \begin{pmatrix} a \\ a \end{pmatrix} = a \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

where a is an arbitrary scalar. Only the nonzero solutions[†] count as eigenvectors, and so the eigenvectors for the eigenvalue $\lambda_1 = 4$ must have $a \neq 0$, i.e., they are all nonzero scalar multiples of the basic eigenvector $\mathbf{v}_1 = (1, 1)^T$.

Remark: In general, if \mathbf{v} is an eigenvector of A for the eigenvalue λ , then so is any nonzero scalar multiple of \mathbf{v} . (Why?) In practice, we only distinguish linearly independent eigenvectors. Thus, in this example, we shall say “ $\mathbf{v}_1 = (1, 1)^T$ is *the* eigenvector corresponding to the eigenvalue $\lambda_1 = 4$ ”, when we really mean that the eigenvectors for $\lambda_1 = 4$ consist of all nonzero scalar multiples of \mathbf{v}_1 .

Similarly, for the second eigenvalue $\lambda_2 = 2$, the eigenvector equation is

$$(A - 2I)\mathbf{v} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The solution $(-a, a)^T = a(-1, 1)^T$ is the set of scalar multiples of the eigenvector $\mathbf{v}_2 = (-1, 1)^T$. Therefore, the complete list of eigenvalues and eigenvectors (up to scalar multiple) is

$$\lambda_1 = 4, \quad \mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda_2 = 2, \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Example 8.6. Consider the 3×3 matrix

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

[†] If, at this stage, you end up with a linear system with only the trivial zero solution, you’ve done something wrong! Either you don’t have a correct eigenvalue — maybe you made a mistake setting up and/or solving the characteristic equation — or you’ve made an error solving the homogeneous eigenvector system.

Using the formula (1.82) for a 3×3 determinant, we compute the characteristic equation

$$\begin{aligned} 0 = \det(A - \lambda I) &= \det \begin{pmatrix} -\lambda & -1 & -1 \\ 1 & 2 - \lambda & 1 \\ 1 & 1 & 2 - \lambda \end{pmatrix} \\ &= (-\lambda)(2 - \lambda)^2 + (-1) \cdot 1 \cdot 1 + (-1) \cdot 1 \cdot 1 - \\ &\quad - 1 \cdot (2 - \lambda)(-1) - 1 \cdot 1 \cdot (-\lambda) - (2 - \lambda) \cdot 1 \cdot (-1) \\ &= -\lambda^3 + 4\lambda^2 - 5\lambda + 2. \end{aligned}$$

The resulting cubic polynomial can be factorized:

$$-\lambda^3 + 4\lambda^2 - 5\lambda + 2 = -(\lambda - 1)^2(\lambda - 2) = 0.$$

Most 3×3 matrices have three different eigenvalues, but this particular one has only two: $\lambda_1 = 1$, which is called a double eigenvalue since it is a double root of the characteristic equation, along with a simple eigenvalue $\lambda_2 = 2$.

The eigenvector equation (8.13) for the double eigenvalue $\lambda_1 = 1$ is

$$(A - I)\mathbf{v} = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The general solution to this homogeneous linear system

$$\mathbf{v} = \begin{pmatrix} -a - b \\ a \\ b \end{pmatrix} = a \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

depends upon two free variables: $y = a$, $z = b$. Any nonzero solution forms a valid eigenvector for the eigenvalue $\lambda_1 = 1$, and so the general eigenvector is any non-zero linear combination of the two “basis eigenvectors” $\mathbf{v}_1 = (-1, 1, 0)^T$, $\widehat{\mathbf{v}}_1 = (-1, 0, 1)^T$.

On the other hand, the eigenvector equation for the simple eigenvalue $\lambda_2 = 2$ is

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

The general solution

$$\mathbf{v} = \begin{pmatrix} -a \\ a \\ a \end{pmatrix} = a \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

consists of all scalar multiple of the eigenvector $\mathbf{v}_2 = (-1, 1, 1)^T$.

In summary, the eigenvalues and (basis) eigenvectors for this matrix are

$$\begin{aligned}\lambda_1 &= 1, & \mathbf{v}_1 &= \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, & \hat{\mathbf{v}}_1 &= \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \\ \lambda_2 &= 2, & \mathbf{v}_2 &= \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.\end{aligned}\tag{8.15}$$

In general, given an eigenvalue λ , the corresponding *eigenspace* $V_\lambda \subset \mathbb{R}^n$ is the subspace spanned by all its eigenvectors. Equivalently, the eigenspace is the kernel

$$V_\lambda = \ker(A - \lambda I).\tag{8.16}$$

In particular, λ is an eigenvalue if and only if $V_\lambda \neq \{\mathbf{0}\}$ is a nontrivial subspace, and then every nonzero element of V_λ is a corresponding eigenvector. The most economical way to indicate each eigenspace is by writing out a basis, as in (8.15).

Example 8.7. The characteristic equation of the matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 2 & 0 & 1 \end{pmatrix}$ is

$$0 = \det(A - \lambda I) = -\lambda^3 + \lambda^2 + 5\lambda + 3 = -(\lambda + 1)^2(\lambda - 3).$$

Again, there is a double eigenvalue $\lambda_1 = -1$ and a simple eigenvalue $\lambda_2 = 3$. However, in this case the matrix

$$A - \lambda_1 I = A + I = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 0 & 1 \\ 2 & 0 & 2 \end{pmatrix}$$

has only a one-dimensional kernel, spanned by $(2, -1, -2)^T$. Thus, even though λ_1 is a double eigenvalue, it only admits a one-dimensional eigenspace. The list of eigenvalues and eigenvectors is, in a sense, incomplete:

$$\lambda_1 = -1, \quad \mathbf{v}_1 = \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}, \quad \lambda_2 = 3, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

Example 8.8. Finally, consider the matrix $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -2 \\ 2 & 2 & -1 \end{pmatrix}$. The characteristic equation is

$$0 = \det(A - \lambda I) = -\lambda^3 + \lambda^2 - 3\lambda - 5 = -(\lambda + 1)(\lambda^2 - 2\lambda + 5).$$

The linear factor yields the eigenvalue -1 . The quadratic factor leads to two complex roots, $1 + 2i$ and $1 - 2i$, which can be obtained via the quadratic formula. Hence A has one real and two complex eigenvalues:

$$\lambda_1 = -1, \quad \lambda_2 = 1 + 2i, \quad \lambda_3 = 1 - 2i.$$

Complex eigenvalues are as important as real eigenvalues, and we need to be able to handle them too. To find the corresponding eigenvectors, which will also be complex, we need to solve the usual eigenvalue equation (8.13), which is now a complex homogeneous linear system. For example, the eigenvector(s) for $\lambda_2 = 1 + 2i$ are found by solving

$$(A - (1 + 2i I))\mathbf{v} = \begin{pmatrix} -2i & 2 & 0 \\ 0 & -2i & -2 \\ 2 & 2 & -2 - 2i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This linear system can be solved by Gaussian elimination (with complex pivots). A simpler approach is to work directly: the first equation $-2ix + 2y = 0$ tells us that $y = ix$, while the second equation $-2iy - 2z = 0$ says $z = -iy = x$. If we trust our calculations so far, we do not need to solve the final equation $2x + 2y + (-2 - 2i)z = 0$, since we know that the coefficient matrix is singular and hence it must be a consequence of the first two equations. (However, it does serve as a useful check on our work.) So, the general solution $\mathbf{v} = (x, ix, x)^T$ is an arbitrary constant multiple of the complex eigenvector $\mathbf{v}_2 = (1, i, 1)^T$. The eigenvector equation for $\lambda_3 = 1 - 2i$ is similarly solved for the third eigenvector $\mathbf{v}_3 = (1, -i, 1)^T$.

Summarizing, the matrix under consideration has three complex eigenvalues and three corresponding eigenvectors, each unique up to (complex) scalar multiple:

$$\begin{aligned} \lambda_1 &= -1, & \lambda_2 &= 1 + 2i, & \lambda_3 &= 1 - 2i, \\ \mathbf{v}_1 &= \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, & \mathbf{v}_2 &= \begin{pmatrix} 1 \\ i \\ 1 \end{pmatrix}, & \mathbf{v}_3 &= \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix}. \end{aligned}$$

Note that the third complex eigenvalue is the complex conjugate of the second, and the eigenvectors are similarly related. This is indicative of a general fact for real matrices:

Proposition 8.9. *If A is a real matrix with a complex eigenvalue $\lambda = \mu + i\nu$ and corresponding complex eigenvector $\mathbf{v} = \mathbf{x} + i\mathbf{y}$, then the complex conjugate $\bar{\lambda} = \mu - i\nu$ is also an eigenvalue with complex conjugate eigenvector $\bar{\mathbf{v}} = \mathbf{x} - i\mathbf{y}$.*

Proof: First take complex conjugates of the eigenvalue equation (8.12):

$$\bar{A} \bar{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \bar{\lambda} \bar{\mathbf{v}}.$$

Using the fact that a real matrix is unaffected by conjugation, so $\bar{A} = A$, we conclude $A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$, which is the eigenvalue equation for the eigenvalue $\bar{\lambda}$ and eigenvector $\bar{\mathbf{v}}$. *Q.E.D.*

As a consequence, when dealing with real matrices, one only needs to compute the eigenvectors for *one* of each complex conjugate pair of eigenvalues. This observation effectively halves the amount of work in the unfortunate event that we are confronted with complex eigenvalues.

Remark: The reader may recall that we said one should never use determinants in practical computations. So why have we reverted to using determinants to find eigenvalues? The truthful answer is that the practical computation of eigenvalues and eigenvectors *never*

resorts to the characteristic equation! The method is fraught with numerical traps and inefficiencies when (a) computing the determinant leading to the characteristic equation, then (b) solving the resulting polynomial equation, which is itself a nontrivial numerical problem, [30], and, finally, (c) solving each of the resulting linear eigenvector systems. Indeed, if we only know an approximation $\tilde{\lambda}$ to the true eigenvalue λ , the approximate eigenvector system $(A - \tilde{\lambda}I)\mathbf{v} = \mathbf{0}$ has a nonsingular coefficient matrix, and hence only admits the trivial solution — which does not even qualify as an eigenvector! Nevertheless, the characteristic equation does give us important theoretical insight into the structure of the eigenvalues of a matrix, and can be used on small, e.g., 2×2 and 3×3 , matrices, when exact arithmetic is employed. Numerical algorithms for computing eigenvalues and eigenvectors are based on completely different ideas, and will be discussed in Section 10.6.

Basic Properties of Eigenvalues

If A is an $n \times n$ matrix, then its *characteristic polynomial* is

$$p_A(\lambda) = \det(A - \lambda I) = c_n \lambda^n + c_{n-1} \lambda^{n-1} + \cdots + c_1 \lambda + c_0. \quad (8.17)$$

The fact that $p_A(\lambda)$ is a polynomial of degree n is a consequence of the general determinantal formula (1.81). Indeed, every term is plus or minus a product of n matrix entries containing one from each row and one from each column. The term corresponding to the identity permutation is obtained by multiplying the diagonal entries together, which, in this case, is

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) = (-1)^n \lambda^n + (-1)^{n-1} (a_{11} + a_{22} + \cdots + a_{nn}) \lambda^{n-1} + \cdots. \quad (8.18)$$

All of the other terms have at most $n - 2$ diagonal factors $a_{ii} - \lambda$, and so are polynomials of degree $\leq n - 2$ in λ . Thus, (8.18) is the only summand containing the monomials λ^n and λ^{n-1} , and so their respective coefficients are

$$c_n = (-1)^n, \quad c_{n-1} = (-1)^{n-1} (a_{11} + a_{22} + \cdots + a_{nn}) = (-1)^{n-1} \operatorname{tr} A, \quad (8.19)$$

where $\operatorname{tr} A$, the sum of its diagonal entries, is called the *trace* of the matrix A . The other coefficients c_{n-2}, \dots, c_1, c_0 in (8.17) are more complicated combinations of the entries of A . However, setting $\lambda = 0$ implies $p_A(0) = \det A = c_0$, and hence the constant term in the characteristic polynomial equals the determinant of the matrix. In particular, if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a 2×2 matrix, its characteristic polynomial has the form

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \\ &= \lambda^2 - (a + d)\lambda + (ad - bc) = \lambda^2 - (\operatorname{tr} A)\lambda + (\det A). \end{aligned} \quad (8.20)$$

As a result of these considerations, the characteristic equation of an $n \times n$ matrix A is a polynomial equation of degree n . According to the Fundamental Theorem of Algebra (see Corollary 16.63) every (complex) polynomial of degree n can be completely factored:

$$p_A(\lambda) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n). \quad (8.21)$$

The complex numbers $\lambda_1, \dots, \lambda_n$, some of which may be repeated, are the *roots* of the characteristic equation $p_A(\lambda) = 0$, and hence the eigenvalues of the matrix A . Therefore, we immediately conclude:

Theorem 8.10. *An $n \times n$ matrix A has at least one and at most n distinct complex eigenvalues.*

Most $n \times n$ matrices — meaning those for which the characteristic polynomial factors into n *distinct* factors — have *exactly* n complex eigenvalues. More generally, an eigenvalue λ_j is said to have *multiplicity* m if the factor $(\lambda - \lambda_j)$ appears exactly m times in the factorization (8.21) of the characteristic polynomial. An eigenvalue is *simple* if it has multiplicity 1. In particular, A has n distinct eigenvalues if and only if all its eigenvalues are simple. In all cases, when the eigenvalues are counted in accordance with their multiplicity, every $n \times n$ matrix has a total of n possibly repeated eigenvalues.

An example of a matrix with just one eigenvalue, of multiplicity n , is the $n \times n$ identity matrix I , whose only eigenvalue is $\lambda = 1$. In this case, *every* nonzero vector in \mathbb{R}^n is an eigenvector of the identity matrix, and so the eigenspace is all of \mathbb{R}^n . At the other extreme, the “bidiagonal” *Jordan block matrix*

$$J_\lambda = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \ddots & \ddots \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}, \quad (8.22)$$

also has only one eigenvalue, λ , again of multiplicity n . But in this case, J_λ has only one eigenvector (up to scalar multiple), which is the first standard basis vector \mathbf{e}_1 , and so its eigenspace is one-dimensional. (You are asked to prove this in Exercise ■.)

Remark: If λ is a complex eigenvalue of multiplicity k for the real matrix A , then its complex conjugate $\bar{\lambda}$ also has multiplicity k . This is because complex conjugate roots of a real polynomial necessarily appear with identical multiplicities.

Remark: If $n \leq 4$, then one can, in fact, write down an explicit formula for the solution to a polynomial equation of degree n , and hence explicit (but not particularly helpful) formulae for the eigenvalues of general 2×2 , 3×3 and 4×4 matrices. As soon as $n \geq 5$, there is no explicit formula (at least in terms of radicals), and so one must usually resort to numerical approximations. This remarkable and deep algebraic result was proved by the young Norwegian mathematician Nils Hendrik Abel in the early part of the nineteenth century, [59].

If we explicitly multiply out the factored product (8.21) and equate the result to the characteristic polynomial (8.17), we find that its coefficients c_0, c_1, \dots, c_{n-1} can be written as certain polynomials of the roots, known as the *elementary symmetric polynomials*. The first and last are of particular importance:

$$c_0 = \lambda_1 \lambda_2 \cdots \lambda_n, \quad c_{n-1} = (-1)^{n-1} (\lambda_1 + \lambda_2 + \cdots + \lambda_n). \quad (8.23)$$

Comparison with our previous formulae for the coefficients c_0 and c_{n-1} leads to the following useful result.

Proposition 8.11. *The sum of the eigenvalues of a matrix equals its trace:*

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}. \quad (8.24)$$

The product of the eigenvalues equals its determinant:

$$\lambda_1 \lambda_2 \cdots \lambda_n = \det A. \quad (8.25)$$

Remark: For repeated eigenvalues, one must add or multiply them in the formulae (8.24), (8.25) according to their multiplicity.

Example 8.12. The matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 2 & 0 & 1 \end{pmatrix}$ considered in Example 8.7 has trace and determinant

$$\operatorname{tr} A = 1, \quad \det A = 3.$$

These fix, respectively, the coefficient of λ^2 and the constant term in the characteristic equation. This matrix has two distinct eigenvalues: -1 , which is a double eigenvalue, and 3 , which is simple. For this particular matrix, formulae (8.24), (8.25) become

$$1 = \operatorname{tr} A = (-1) + (-1) + 3, \quad 3 = \det A = (-1)(-1)3.$$

Note that the double eigenvalue contributes twice to the sum and product.

8.3. Eigenvector Bases and Diagonalization.

Most of the vector space bases that play a distinguished role in applications consist of eigenvectors of a particular matrix. In this section, we show that the eigenvectors for any “complete” matrix automatically form a basis for \mathbb{R}^n or, in the complex case, \mathbb{C}^n . In the following subsection, we use the eigenvector basis to rewrite the linear transformation determined by the matrix in a simple diagonal form. The most important case — symmetric and positive definite matrices — will be treated in the next section.

The first task is to show that eigenvectors corresponding to distinct eigenvalues are automatically linearly independent.

Lemma 8.13. *If $\lambda_1, \dots, \lambda_k$ are distinct eigenvalues of the same matrix A , then the corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent.*

Proof: The result is proved by induction on the number of eigenvalues. The case $k = 1$ is immediate since an eigenvector cannot be zero. Assume that we know the result is valid for $k - 1$ eigenvalues. Suppose we have a vanishing linear combination:

$$c_1 \mathbf{v}_1 + \cdots + c_{k-1} \mathbf{v}_{k-1} + c_k \mathbf{v}_k = \mathbf{0}. \quad (8.26)$$

Let us multiply this equation by the matrix A :

$$\begin{aligned} A(c_1 \mathbf{v}_1 + \cdots + c_{k-1} \mathbf{v}_{k-1} + c_k \mathbf{v}_k) &= c_1 A \mathbf{v}_1 + \cdots + c_{k-1} A \mathbf{v}_{k-1} + c_k A \mathbf{v}_k \\ &= c_1 \lambda_1 \mathbf{v}_1 + \cdots + c_{k-1} \lambda_{k-1} \mathbf{v}_{k-1} + c_k \lambda_k \mathbf{v}_k = \mathbf{0}. \end{aligned}$$

On the other hand, if we multiply the original equation by λ_k , we also have

$$c_1 \lambda_k \mathbf{v}_1 + \cdots + c_{k-1} \lambda_k \mathbf{v}_{k-1} + c_k \lambda_k \mathbf{v}_k = \mathbf{0}.$$

Subtracting this from the previous equation, the final terms cancel and we are left with the equation

$$c_1(\lambda_1 - \lambda_k) \mathbf{v}_1 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k) \mathbf{v}_{k-1} = \mathbf{0}.$$

This is a vanishing linear combination of the first $k - 1$ eigenvectors, and so, by our induction hypothesis, can only happen if all the coefficients are zero:

$$c_1(\lambda_1 - \lambda_k) = 0, \quad \dots \quad c_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

The eigenvalues were assumed to be distinct, so $\lambda_j \neq \lambda_k$ when $j \neq k$. Consequently, $c_1 = \cdots = c_{k-1} = 0$. Substituting these values back into (8.26), we find $c_k \mathbf{v}_k = \mathbf{0}$, and so $c_k = 0$ also, since the eigenvector $\mathbf{v}_k \neq \mathbf{0}$. Thus we have proved that (8.26) holds if and only if $c_1 = \cdots = c_k = 0$, which implies the linear independence of the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$. This completes the induction step. *Q.E.D.*

The most important consequence of this result is when a matrix has the maximum allotment of eigenvalues.

Theorem 8.14. *If the $n \times n$ real matrix A has n distinct real eigenvalues $\lambda_1, \dots, \lambda_n$, then the corresponding real eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ form a basis of \mathbb{R}^n . If A (which may now be either a real or a complex matrix) has n distinct complex eigenvalues, then the corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ form a basis of \mathbb{C}^n .*

If a matrix has multiple eigenvalues, then there may or may not be an eigenvector basis of \mathbb{R}^n (or \mathbb{C}^n). The matrix in Example 8.6 admits an eigenvector basis, whereas the matrix in Example 8.7 does not. In general, it can be proved[†] that the dimension of the eigenspace is less than or equal to the eigenvalue's multiplicity. In particular, every simple eigenvalue has a one-dimensional eigenspace, and hence, up to scalar multiple, only one associated eigenvector.

Definition 8.15. An eigenvalue λ of a matrix A is called *complete* if its eigenspace $V_\lambda = \ker(A - \lambda I)$ has the same dimension as its multiplicity. The matrix A is *complete* if all its eigenvalues are.

Note that a simple eigenvalue is automatically complete, since its eigenspace is the one-dimensional subspace spanned by the corresponding eigenvector. Thus, only multiple eigenvalues can cause a matrix to be incomplete.

[†] This is a consequence of Theorem 8.43.

Remark: The multiplicity of an eigenvalue λ_i is sometimes referred to as its *algebraic multiplicity*. The dimension of the eigenspace V_λ is its *geometric multiplicity*, and so completeness requires that the two multiplicities are equal. The word “complete” is not completely standard; other common terms for such matrices are *perfect*, *semi-simple* and, as discussed below, *diagonalizable*.

Theorem 8.16. *An $n \times n$ real or complex matrix A is complete if and only if its eigenvectors span \mathbb{C}^n . In particular, any $n \times n$ matrix that has n distinct eigenvalues is complete.*

Or, stated another way, a matrix is complete if and only if it admits an eigenvector basis of \mathbb{C}^n . Most matrices are complete. Incomplete $n \times n$ matrices, which have fewer than n linearly independent complex eigenvectors, are more tricky to deal with, and we relegate most of the messy details to Section 8.6.

Remark: We already noted that complex eigenvectors of a real matrix always appear in conjugate pairs: $\mathbf{v} = \mathbf{x} \pm i\mathbf{y}$. It can be shown that the real and imaginary parts of these vectors form a real basis for \mathbb{R}^n . (See Exercise ■ for the underlying principle.)

For instance, in Example 8.8, the complex eigenvectors are $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \pm i \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. The vectors $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, consisting of the real eigenvector and the real and imaginary parts of the complex eigenvectors, form a basis for \mathbb{R}^3 .

Diagonalization

Let $L: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation on n -dimensional Euclidean space. As we know, cf. Theorem 7.5, L is determined by multiplication by an $n \times n$ matrix. However, the matrix representing a given linear transformation will depend upon the choice of basis for the underlying vector space \mathbb{R}^n . Linear transformations having a complicated matrix representation in terms of the standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ may be considerably simplified by choosing a suitably adapted basis $\mathbf{v}_1, \dots, \mathbf{v}_n$. We are now in a position to understand how to effect such a simplification.

For example, the linear transformation $L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ 2x + 4y \end{pmatrix}$ studied in Example 7.19 is represented by the matrix $A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$ — when expressed in terms of the standard basis of \mathbb{R}^2 . In terms of the alternative basis $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, it is represented by the diagonal matrix $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, indicating it has a simple stretching action on the new basis vectors: $A\mathbf{v}_1 = 2\mathbf{v}_1$, $A\mathbf{v}_2 = 3\mathbf{v}_2$. Now we can understand the reason for this simplification. *The new basis consists of the two eigenvectors of the matrix A .* This observation is indicative of a general fact: representing a linear transformation in terms

of an eigenvector basis has the effect of replacing its matrix representative by a simple diagonal form. The effect is to *diagonalize* the original coefficient matrix.

According to (7.26), if $\mathbf{v}_1, \dots, \mathbf{v}_n$ form a basis of \mathbb{R}^n , then the corresponding matrix representative of the linear transformation $L[\mathbf{v}] = A\mathbf{v}$ is given by the similar matrix $B = S^{-1}AS$, where $S = (\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n)^T$ is the matrix whose columns are the basis vectors. In the preceding example, $S = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$, and we find that $S^{-1}AS = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ is a diagonal matrix.

Definition 8.17. A square matrix A is called *diagonalizable* if there exists a nonsingular matrix S and a diagonal matrix $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ such that

$$S^{-1}AS = \Lambda. \quad (8.27)$$

A diagonal matrix represents a linear transformation that simultaneously stretches[†] in the direction of the basis vectors. Thus, every diagonalizable matrix represents a elementary combination of (complex) stretching transformations.

To understand the diagonalization equation (8.27), we rewrite it in the equivalent form

$$AS = S\Lambda. \quad (8.28)$$

Using the basic property (1.11) of matrix multiplication, one easily sees that the k^{th} column of this $n \times n$ matrix equation is given by

$$A\mathbf{v}_k = \lambda_k \mathbf{v}_k.$$

Therefore, the columns of S are necessarily eigenvectors, and the entries of the diagonal matrix Λ are the corresponding eigenvalues! And, as a result, a diagonalizable matrix A must have n linearly independent eigenvectors, i.e., an eigenvector basis, to form the columns of the nonsingular diagonalizing matrix S . Since the diagonal form Λ contains the eigenvalues along its diagonal, it is uniquely determined up to a permutation of its entries.

Now, as we know, not every matrix has an eigenvector basis. Moreover, even when it exists, the eigenvector basis may be complex, in which case S is a complex matrix, and the entries of the diagonal matrix Λ are the complex eigenvalues. Thus, we should distinguish between complete matrices that are diagonalizable over the complex numbers and the more restrictive class of matrices which can be diagonalized by a real matrix S .

Theorem 8.18. *A matrix is complex diagonalizable if and only if it is complete. A matrix is real diagonalizable if and only if it is complete and has all real eigenvalues.*

[†] A negative diagonal entry represents the combination of a reflection and stretch. Complex entries correspond to a complex stretching transformation. See Section 7.2 for details.

Example 8.19. The 3×3 matrix $A = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ considered in Example 8.5 has eigenvector basis

$$\mathbf{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

We assemble these to form the eigenvector matrix

$$S = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{and so} \quad S^{-1} = \begin{pmatrix} -1 & 0 & -1 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

The diagonalization equation (8.27) becomes

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

with the eigenvalues of A appearing on the diagonal of Λ , in the same order as the eigenvectors.

Remark: If a matrix is not complete, then it cannot be diagonalized. A simple example is a matrix of the form $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$ with $c \neq 0$, which represents a shear in the direction of the x axis. Incomplete matrices will be the subject of the Section 8.6.

8.4. Eigenvalues of Symmetric Matrices.

Fortunately, the matrices that arise in most applications are complete and, in fact, possess some additional structure that ameliorates the calculation of their eigenvalues and eigenvectors. The most important class are the symmetric, including positive definite, matrices. In fact, not only are the eigenvalues of a symmetric matrix necessarily real, the eigenvectors always form an *orthogonal basis* of the underlying Euclidean space. In such situations, we can tap into the dramatic power of orthogonal bases that we developed in Chapter 5. In fact, this is by far the most common way for orthogonal bases to appear — as the eigenvector bases of symmetric matrices.

Theorem 8.20. *Let $A = A^T$ be a real symmetric $n \times n$ matrix, Then*

- (a) *All the eigenvalues of A are real.*
- (b) *Eigenvectors corresponding to distinct eigenvalues are orthogonal.*
- (c) *There is an orthonormal basis of \mathbb{R}^n consisting of n eigenvectors of A .*

In particular, all symmetric matrices are complete.

We defer the proof of Theorem 8.20 until the end of this section.

Remark: Orthogonality is with respect to the standard dot product on \mathbb{R}^n . As we noted in Section 7.5, the transpose or adjoint operation is intimately connected with the dot product. An analogous result holds for self-adjoint linear transformations on \mathbb{R}^n ; see Exercise ■ for details.

Example 8.21. The 2×2 matrix $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ considered in Example 8.5 is symmetric, and so has real eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 2$. You can easily check that the corresponding eigenvectors $\mathbf{v}_1 = (1, 1)^T$ and $\mathbf{v}_2 = (-1, 1)^T$ are orthogonal: $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, and hence form an orthogonal basis of \mathbb{R}^2 . An orthonormal basis is provided by the unit eigenvectors

$$\mathbf{u}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix},$$

obtained by dividing each eigenvector by its length: $\mathbf{u}_k = \mathbf{v}_k / \|\mathbf{v}_k\|$.

Example 8.22. Consider the symmetric matrix $A = \begin{pmatrix} 5 & -4 & 2 \\ -4 & 5 & 2 \\ 2 & 2 & -1 \end{pmatrix}$. A straightforward computation produces its eigenvalues and eigenvectors:

$$\begin{aligned} \lambda_1 &= 9, & \lambda_2 &= 3, & \lambda_3 &= -3, \\ \mathbf{v}_1 &= \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, & \mathbf{v}_2 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, & \mathbf{v}_3 &= \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}. \end{aligned}$$

As the reader can check, the eigenvectors form an orthogonal basis of \mathbb{R}^3 . The orthonormal eigenvector basis promised by Theorem 8.20 is obtained by dividing each eigenvector by its norm:

$$\mathbf{u}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{pmatrix}.$$

We can characterize positive definite matrices by their eigenvalues.

Theorem 8.23. A symmetric matrix $K = K^T$ is positive definite if and only if all of its eigenvalues are strictly positive.

Proof: First, if $K > 0$, then, by definition, $\mathbf{x}^T K \mathbf{x} > 0$ for all nonzero vectors $\mathbf{x} \in \mathbb{R}^n$. In particular, if $\mathbf{x} = \mathbf{v}$ is an eigenvector with (necessarily real) eigenvalue λ , then

$$0 < \mathbf{v}^T K \mathbf{v} = \mathbf{v}^T (\lambda \mathbf{v}) = \lambda \|\mathbf{v}\|^2, \quad (8.29)$$

which immediately proves that $\lambda > 0$. Conversely, suppose K has all positive eigenvalues. Let $\mathbf{u}_1, \dots, \mathbf{u}_n$ be the orthonormal eigenvector basis of \mathbb{R}^n guaranteed by Theorem 8.20, with $K \mathbf{u}_j = \lambda_j \mathbf{u}_j$. Then, writing

$$\mathbf{x} = c_1 \mathbf{u}_1 + \dots + c_n \mathbf{u}_n, \quad \text{we have} \quad K \mathbf{x} = c_1 \lambda_1 \mathbf{u}_1 + \dots + c_n \lambda_n \mathbf{u}_n.$$

Therefore, using the orthonormality of the eigenvectors,

$$\mathbf{x}^T K \mathbf{x} = (c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n) \cdot (c_1 \lambda_1 \mathbf{u}_1 + \cdots + c_n \lambda_n \mathbf{u}_n) = \lambda_1 c_1^2 + \cdots + \lambda_n c_n^2 > 0$$

for any $\mathbf{x} \neq \mathbf{0}$, since only $\mathbf{x} = \mathbf{0}$ has coordinates $c_1 = \cdots = c_n = 0$. This proves that K is positive definite. *Q.E.D.*

Remark: The same proof shows that K is positive semi-definite if and only if all its eigenvalues satisfy $\lambda \geq 0$. A positive semi-definite matrix that is not positive definite admits a zero eigenvalue and one or more *null eigenvectors*, i.e., solutions to $K\mathbf{v} = \mathbf{0}$. Every nonzero element $\mathbf{0} \neq \mathbf{v} \in \ker K$ of the kernel is a null eigenvector.

Example 8.24. Consider the symmetric matrix $K = \begin{pmatrix} 8 & 0 & 1 \\ 0 & 8 & 1 \\ 1 & 1 & 7 \end{pmatrix}$. Its characteristic equation is

$$\det(K - \lambda I) = -\lambda^3 + 23\lambda^2 - 174\lambda + 432 = -(\lambda - 9)(\lambda - 8)(\lambda - 6),$$

and so its eigenvalues are 9, 8 and 6. Since they are all positive, K is a positive definite matrix. The associated eigenvectors are

$$\lambda_1 = 9, \quad \mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \lambda_2 = 8, \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \lambda_3 = 6, \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

Note that the eigenvectors form an orthogonal basis of \mathbb{R}^3 , as guaranteed by Theorem 8.20. As usual, we can construct an corresponding orthonormal eigenvector basis

$$\mathbf{u}_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix},$$

by dividing each eigenvector by its norm.

Proof of Theorem 8.20: First recall that if $A = A^T$ is real, symmetric, then

$$(A\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot (A\mathbf{w}) \quad \text{for all} \quad \mathbf{v}, \mathbf{w} \in \mathbb{C}^n, \quad (8.30)$$

where we use the Euclidean dot product for real vectors and, more generally, the Hermitian dot product $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \overline{\mathbf{w}}$ when they are complex. (See Exercise ■.)

To prove property (a), suppose λ is a complex eigenvalue with complex eigenvector $\mathbf{v} \in \mathbb{C}^n$. Consider the Hermitian dot product of the complex vectors $A\mathbf{v}$ and \mathbf{v} :

$$(A\mathbf{v}) \cdot \mathbf{v} = (\lambda \mathbf{v}) \cdot \mathbf{v} = \lambda \|\mathbf{v}\|^2.$$

On the other hand, by (8.30),

$$(A\mathbf{v}) \cdot \mathbf{v} = \mathbf{v} \cdot (A\mathbf{v}) = \mathbf{v} \cdot (\lambda \mathbf{v}) = \mathbf{v}^T \overline{\lambda \mathbf{v}} = \overline{\lambda} \|\mathbf{v}\|^2.$$

Equating these two expressions, we deduce

$$\bar{\lambda} \|\mathbf{v}\|^2 = \lambda \|\mathbf{v}\|^2.$$

Since \mathbf{v} is an eigenvector, it is nonzero, $\mathbf{v} \neq \mathbf{0}$, and so $\bar{\lambda} = \lambda$. This proves that the eigenvalue λ is real.

To prove (b), suppose

$$A\mathbf{v} = \lambda\mathbf{v}, \quad A\mathbf{w} = \mu\mathbf{w},$$

where $\lambda \neq \mu$ are distinct real eigenvalues. Then, again by (8.30),

$$\lambda \mathbf{v} \cdot \mathbf{w} = (A\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot (A\mathbf{w}) = \mathbf{v} \cdot (\mu\mathbf{w}) = \mu \mathbf{v} \cdot \mathbf{w},$$

and hence

$$(\lambda - \mu) \mathbf{v} \cdot \mathbf{w} = 0.$$

Since $\lambda \neq \mu$, this implies that $\mathbf{v} \cdot \mathbf{w} = 0$ and hence the eigenvectors \mathbf{v}, \mathbf{w} are orthogonal.

Finally, the proof of (c) is easy if all the eigenvalues of A are distinct. Theorem 8.14 implies that the eigenvectors form a basis of \mathbb{R}^n , and part (b) proves they are orthogonal. (An alternative proof starts with orthogonality, and then applies Proposition 5.4 to prove that the eigenvectors form a basis.) To obtain an orthonormal basis, we merely divide the eigenvectors by their lengths: $\mathbf{u}_k = \mathbf{v}_k / \|\mathbf{v}_k\|$, as in Lemma 5.2.

To prove (c) in general, we proceed by induction on the size n of the matrix A . The case of a 1×1 matrix is trivial. (Why?) Let A have size $n \times n$. We know that A has at least one eigenvalue, λ_1 , which is necessarily real. Let \mathbf{v}_1 be the associated eigenvector. Let

$$V^\perp = \{ \mathbf{w} \in \mathbb{R}^n \mid \mathbf{v}_1 \cdot \mathbf{w} = 0 \}$$

denote the orthogonal complement to the eigenspace V_{λ_1} — the set of all vectors orthogonal to the first eigenvector. Proposition 5.48 implies that $\dim V^\perp = n - 1$, and we choose an orthonormal basis $\mathbf{y}_1, \dots, \mathbf{y}_{n-1}$. Now, if \mathbf{w} is any vector in V^\perp , so is $A\mathbf{w}$ since, by symmetry,

$$\mathbf{v}_1 \cdot (A\mathbf{w}) = \mathbf{v}_1^T A\mathbf{w} = (A\mathbf{v}_1)^T \mathbf{w} = \lambda_1 \mathbf{v}_1^T \mathbf{w} = \lambda_1 (\mathbf{v}_1 \cdot \mathbf{w}) = 0.$$

Thus, A defines a linear transformation on V^\perp , and is represented by a symmetric $(n-1) \times (n-1)$ matrix with respect to the chosen orthonormal basis $\mathbf{y}_1, \dots, \mathbf{y}_{n-1}$. Our induction hypothesis implies that there is an orthonormal basis of V^\perp consisting of eigenvectors $\mathbf{u}_2, \dots, \mathbf{u}_n$ of A . Appending the unit eigenvector $\mathbf{u}_1 = \mathbf{v}_1 / \|\mathbf{v}_1\|$ to this collection will complete the orthonormal basis of \mathbb{R}^n . *Q.E.D.*

The Spectral Theorem

Since a real, symmetric matrix admits an eigenvector basis, it is diagonalizable. Moreover, since we can arrange that the eigenvectors form an orthonormal basis, the diagonalizing matrix takes a particularly simple form. Recall that an $n \times n$ matrix Q is called *orthogonal* if and only if its columns form an orthonormal basis of \mathbb{R}^n . Alternatively, one characterizes orthogonal matrices by the condition $Q^{-1} = Q^T$, as in Definition 5.18.

Therefore, when we use the orthonormal eigenvector basis in the diagonalization formula (8.27), the result is the *Spectral Theorem* that governs the diagonalization of symmetric matrices.

Theorem 8.25. *Let A be a real, symmetric matrix. Then there exists an orthogonal matrix Q such that*

$$A = Q \Lambda Q^{-1} = Q \Lambda Q^T, \quad (8.31)$$

where Λ is a real diagonal matrix. The eigenvalues of A appear on the diagonal of Λ , while the eigenvectors are the corresponding columns of Q .

Remark: The term “spectrum” refers to the eigenvalues of a matrix or, more generally, a linear operator. The terminology is motivated by physics. The spectral energy lines of atoms, molecules and nuclei are characterized as the eigenvalues of the governing quantum mechanical linear operators, [102, 106].

Warning: The spectral factorization $A = Q \Lambda Q^T$ and the Gaussian factorization $A = L D L^T$ of a regular symmetric matrix, cf. (1.52), are completely different. In particular, the eigenvalues are *not* the pivots, so $\Lambda \neq D$.

The *spectral decomposition* (8.31) provides us with an alternative means of diagonalizing the associated quadratic form $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$, i.e., of completing the square. We write

$$q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = \mathbf{x}^T Q \Lambda Q^T \mathbf{x} = \mathbf{y}^T \Lambda \mathbf{y} = \sum_{i=1}^n \lambda_i y_i^2, \quad (8.32)$$

where $\mathbf{y} = Q^T \mathbf{x} = Q^{-1} \mathbf{x}$ are the coordinates of \mathbf{x} with respect to the orthonormal eigenvalue basis of A . In particular, $q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$ and so A is positive definite if and only if each eigenvalue $\lambda_i > 0$ is strictly positive, reconfirming Theorem 8.23.

Example 8.26. For the 2×2 matrix $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ considered in Example 8.21, the orthonormal eigenvectors produce the diagonalizing orthogonal rotation matrix $Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$. The reader can check the spectral factorization

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} = A = Q \Lambda Q^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

According to (8.32), the associated quadratic form is diagonalized as

$$q(\mathbf{x}) = 3x_1^2 + 2x_1x_2 + 3x_2^2 = 4y_1^2 + 2y_2^2,$$

where $\mathbf{y} = Q^T \mathbf{x}$, i.e., $y_1 = \frac{x_1 + x_2}{\sqrt{2}}$, $y_2 = \frac{-x_1 + x_2}{\sqrt{2}}$.

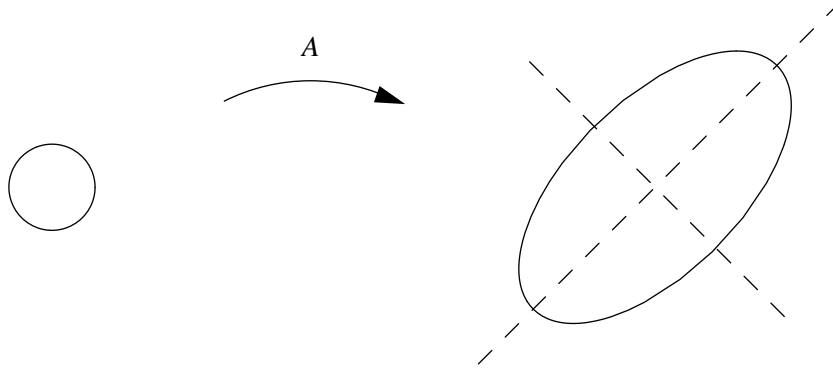


Figure 8.2. Stretching a Circle into an Ellipse.

We note that you can always choose Q to be a proper orthogonal matrix, so $\det Q = 1$, since an improper orthogonal matrix can be made proper by multiplying one of its columns by -1 , which does not affect its status as an eigenvector matrix. Since a proper orthogonal matrix Q represents a rigid rotation of \mathbb{R}^n , the diagonalization of a symmetric matrix can be interpreted as a rotation of the coordinate system in which the orthogonal eigenvectors line up along the coordinate axes. Therefore, a linear transformation $L(\mathbf{x}) = A\mathbf{x}$ represented by a positive definite matrix $A > 0$ can be regarded as a combination of stretches along a mutually orthogonal set of directions. In elasticity, the stress tensor of a deformed body is represented by a positive definite matrix. Its eigenvalues are known as the *principal stretches* and its eigenvectors the principal directions of the elastic deformation.

A good way to visualize this is to consider the effect of the linear transformation on the unit (Euclidean) sphere

$$S_1 = \{ \|\mathbf{x}\| = 1 \}.$$

Stretching the sphere in orthogonal directions will map it to an ellipsoid $E = L[S_1]$ whose axes are aligned with the directions of stretch. Explicitly, if the linear transformation is prescribed by $\mathbf{y} = A\mathbf{x}$, then

$$E = L[S_1] = \{ A\mathbf{x} \mid \|\mathbf{x}\| = 1 \} = \{ \mathbf{y} \mid \|A^{-1}\mathbf{y}\| = 1 \}. \quad (8.33)$$

For example, the matrix $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ considered above represents the linear transformation

$$\tilde{x} = 3x + y, \quad \tilde{y} = x + 3y.$$

In this instance, the unit circle $x^2 + y^2 = 1$ is mapped to an ellipse

$$\left(\frac{3\tilde{x} - \tilde{y}}{8} \right)^2 + \left(\frac{-\tilde{x} + 3\tilde{y}}{8} \right)^2 = \frac{5}{32} \tilde{x}^2 - \frac{3}{16} \tilde{x}\tilde{y} + \frac{5}{32} \tilde{y}^2 = 1,$$

whose principal axes line up with the eigenvectors $\mathbf{u}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, $\mathbf{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$; see Figure 8.2. The eigenvalues, 4, 2, prescribe the ellipse's *semi-axes*.

Optimization Principles for Eigenvalues

As we learned in Chapter 4, the solution to a linear system with positive definite coefficient matrix can be characterized by a minimization principle. Thus, it should come as no surprise that eigenvalues of positive definite, and even more general symmetric matrices, can also be characterized by some sort of optimization procedure. A number of basic numerical algorithms for computing eigenvalues, of both matrices and, later on, differential operators are based on such optimization principles.

First consider the relatively simple case of a diagonal matrix $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. We assume that the diagonal entries, which are the *same* as the eigenvalues, appear in decreasing order,

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n, \quad (8.34)$$

so λ_1 is the largest eigenvalue, while λ_n is the smallest. The effect of Λ on a vector $\mathbf{y} = (y_1, y_2, \dots, y_n)^T \in \mathbb{R}^n$ is to multiply its entries by the diagonal eigenvalues: $\Lambda \mathbf{y} = (\lambda_1 y_1, \lambda_2 y_2, \dots, \lambda_n y_n)^T$. In other words, the linear transformation represented by the coefficient matrix Λ has the effect of stretching[†] the i^{th} coordinate direction by the factor λ_i . In particular, the maximal stretch occurs in the \mathbf{e}_1 direction, with factor λ_1 , while the minimal (or largest negative) stretch occurs in the \mathbf{e}_n direction, with factor λ_n . The germ of the optimization principles for characterizing the extreme eigenvalues is contained in this geometrical observation.

Let us turn our attention to the associated quadratic form

$$q(\mathbf{y}) = \mathbf{y}^T \Lambda \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2. \quad (8.35)$$

Note that $q(t\mathbf{e}_1) = \lambda_1 t^2$, and hence if $\lambda_1 > 0$, then $q(\mathbf{y})$ has no maximum; on the other hand, if $\lambda_1 \leq 0$, so all eigenvalues are non-positive, then $q(\mathbf{y}) \leq 0$ for all \mathbf{y} , and its maximal value is $q(\mathbf{0}) = 0$. Thus, in either case, a strict maximization of $q(\mathbf{y})$ is of no help.

Suppose, however, that we try to maximize $q(\mathbf{y})$ but restrict \mathbf{y} to be a unit vector (in the Euclidean norm):

$$\|\mathbf{y}\|^2 = y_1^2 + \dots + y_n^2 = 1.$$

In view of (8.34),

$$q(\mathbf{y}) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2 \leq \lambda_1 y_1^2 + \lambda_1 y_2^2 + \dots + \lambda_1 y_n^2 = \lambda_1 (y_1^2 + \dots + y_n^2) = \lambda_1.$$

Moreover, $q(\mathbf{e}_1) = \lambda_1$. We conclude that the maximal value of $q(\mathbf{y})$ over all unit vectors *is* the largest eigenvalue of Λ :

$$\lambda_1 = \max \{ q(\mathbf{y}) \mid \|\mathbf{y}\| = 1 \}.$$

By the same reasoning, its minimal value equals the smallest eigenvalue:

$$\lambda_n = \min \{ q(\mathbf{y}) \mid \|\mathbf{y}\| = 1 \}$$

[†] If $\lambda_i < 0$, then the effect is to stretch and reflect.

Thus, we can characterize the two extreme eigenvalues by optimization principles, albeit of a slightly different character than we treated in Chapter 4.

Now suppose A is any symmetric matrix. We use the spectral decomposition (8.31) to diagonalize the associated quadratic form

$$q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = \mathbf{x}^T Q \Lambda Q^T \mathbf{x} = \mathbf{y}^T \Lambda \mathbf{y}, \quad \text{where} \quad \mathbf{y} = Q^T \mathbf{x} = Q^{-1} \mathbf{x},$$

as in (8.32). According to the preceding discussion, the maximum of $\mathbf{y}^T \Lambda \mathbf{y}$ over all unit vectors $\|\mathbf{y}\| = 1$ is the largest eigenvalue λ_1 of Λ , which is the *same* as the largest eigenvalue of A . Moreover, since Q is an orthogonal matrix, Proposition 7.24 tell us that it maps unit vectors to unit vectors:

$$1 = \|\mathbf{y}\| = \|Q^T \mathbf{x}\| = \|\mathbf{x}\|,$$

and so the maximum of $q(\mathbf{x})$ over all unit vectors $\|\mathbf{x}\| = 1$ is the same maximum eigenvalue λ_1 . Similar reasoning applies to the smallest eigenvalue λ_n . In this fashion, we have established the basic optimization principles for the extreme eigenvalues of a symmetric matrix.

Theorem 8.27. *If A is a symmetric matrix, then*

$$\lambda_1 = \max \{ \mathbf{x}^T A \mathbf{x} \mid \|\mathbf{x}\| = 1 \}, \quad \lambda_n = \min \{ \mathbf{x}^T A \mathbf{x} \mid \|\mathbf{x}\| = 1 \}, \quad (8.36)$$

are, respectively its largest and smallest eigenvalues.

The maximal value is achieved when $\mathbf{x} = \pm \mathbf{u}_1$ is one of the unit eigenvectors corresponding to the largest eigenvalue; similarly, the minimal value is at $\mathbf{x} = \pm \mathbf{u}_n$.

Remark: In multivariable calculus, the eigenvalue λ plays the role of a *Lagrange multiplier* for the constrained optimization problem. See [9] for details.

Example 8.28. The problem is to maximize the value of the quadratic form

$$q(x, y) = 3x^2 + 2xy + 3y^2$$

for all x, y lying on the unit circle $x^2 + y^2 = 1$. This maximization problem is precisely of form (8.36). The coefficient matrix is $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, whose eigenvalues are, according to Example 8.5, $\lambda_1 = 4$ and $\lambda_2 = 2$. Theorem 8.27 implies that the maximum is the largest eigenvalue, and hence equal to 4, while its minimum is the smallest eigenvalue, and hence equal to 2. Thus, evaluating $q(x, y)$ on the unit eigenvectors, we conclude that

$$q\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = 2 \leq q(x, y) \leq 4 = q\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \quad \text{for all} \quad x^2 + y^2 = 1.$$

In practical applications, the restriction of the quadratic form to unit vectors may not be particularly convenient. One can, however, rephrase the eigenvalue optimization principles in a form that utilizes general vectors. If $\mathbf{v} \neq \mathbf{0}$ is any nonzero vector, then $\mathbf{x} = \mathbf{v}/\|\mathbf{v}\|$ is a unit vector. Substituting this expression for \mathbf{x} in the quadratic form

$q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ leads to the following optimization principles for the extreme eigenvalues of a symmetric matrix:

$$\lambda_1 = \max \left\{ \frac{\mathbf{v}^T A \mathbf{v}}{\|\mathbf{v}\|^2} \mid \mathbf{v} \neq \mathbf{0} \right\}, \quad \lambda_n = \min \left\{ \frac{\mathbf{v}^T A \mathbf{v}}{\|\mathbf{v}\|^2} \mid \mathbf{v} \neq \mathbf{0} \right\}. \quad (8.37)$$

Thus, we replace optimization of a quadratic polynomial over the unit sphere by optimization of a rational function over all of $\mathbb{R}^n \setminus \{\mathbf{0}\}$. Referring back to Example 8.28, the maximum value of

$$r(x, y) = \frac{3x^2 + 2xy + 3y^2}{x^2 + y^2} \quad \text{for all} \quad \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is equal to 4, the same maximal eigenvalue of the corresponding coefficient matrix.

What about characterizing one of the intermediate eigenvalues? Then we need to be a little more sophisticated in designing the optimization principle. To motivate the construction, look first at the diagonal case. If we restrict the quadratic form (8.35) to vectors $\tilde{\mathbf{y}} = (0, y_2, \dots, y_n)^T$ whose first component is zero, we obtain

$$q(\tilde{\mathbf{y}}) = q(0, y_2, \dots, y_n) = \lambda_2 y_2^2 + \dots + \lambda_n y_n^2.$$

The maximum value of $q(\tilde{\mathbf{y}})$ over all such $\tilde{\mathbf{y}}$ of norm 1 is, by the same reasoning, the second largest eigenvalue λ_2 . Moreover, we can characterize such vectors geometrically by noting that $\tilde{\mathbf{y}} \cdot \mathbf{e}_1 = 0$, and so they are orthogonal to the first standard basis vector, which also happens to be the eigenvector of Λ corresponding to the eigenvalue λ_1 . Similarly, if we want to find the j^{th} largest eigenvalue λ_j , we maximize $q(\mathbf{y})$ over all unit vectors $\hat{\mathbf{y}}$ whose first $j-1$ components vanish, $y_1 = \dots = y_{j-1} = 0$, or, stated geometrically, over all vectors $\hat{\mathbf{y}}$ such that $\|\hat{\mathbf{y}}\| = 1$ and $\hat{\mathbf{y}} \cdot \mathbf{e}_1 = \dots = \hat{\mathbf{y}} \cdot \mathbf{e}_{j-1} = 0$, i.e., over all vectors orthogonal to the first $j-1$ eigenvectors of Λ .

A similar reasoning based on the Spectral Theorem 8.25 and the orthogonality of eigenvectors of symmetric matrices, leads to the general result.

Theorem 8.29. *Let A be a symmetric matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ and corresponding orthogonal eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. Then the maximal value of the quadratic form $\mathbf{x}^T A \mathbf{x}$ over all unit vectors which are orthogonal to the first $j-1$ eigenvectors is its j^{th} eigenvalue:*

$$\lambda_j = \max \left\{ \mathbf{x}^T A \mathbf{x} \mid \|\mathbf{x}\| = 1, \mathbf{x} \cdot \mathbf{v}_1 = \dots = \mathbf{x} \cdot \mathbf{v}_{j-1} = 0 \right\}. \quad (8.38)$$

Thus, at least in principle, one can compute the eigenvalues and eigenvectors of a symmetric matrix by the following recursive procedure. First, find the largest eigenvalue λ_1 by the basic maximization principle (8.36) and its associated eigenvector \mathbf{v}_1 by solving the eigenvector system (8.13). The next largest eigenvalue λ_2 is then characterized by the constrained minimization principle (8.38), and so on. Although of theoretical interest, this algorithm is not effective in practical numerical computations.

8.5. Singular Values.

We have already indicated the central role played by the eigenvalues and eigenvectors of a square matrix in both theory and applications. Much more evidence to this effect will appear in the ensuing chapters. Alas, non-square matrices do not have eigenvalues (why?), and so, at first glance, do not appear to possess any quantities of comparable significance. However, our earlier treatment of least squares minimization problems as well as the equilibrium equations for structures and circuits made essential use of the symmetric, positive semi-definite *square* Gram matrix $K = A^T A$ — which can be naturally formed even when A is rectangular. Perhaps the eigenvalues of K might play a comparably important role for general matrices. Since they are not easily related to the eigenvalues of A — which, in the truly rectangular case, don't even exist — we shall endow them with a new name.

Definition 8.30. The *singular values* $\sigma_1, \dots, \sigma_n$ of an $m \times n$ matrix A are the square roots, $\sigma_i = \sqrt{\lambda_i}$, of the eigenvalues of the associated Gram matrix $K = A^T A$. The corresponding eigenvectors of K are known as the *singular vectors* of A .

Since K is necessarily positive semi-definite, its eigenvalues are always non-negative, $\lambda_i \geq 0$, and hence the singular values of A are also all non-negative[†], $\sigma_i \geq 0$ — no matter whether A itself has positive, negative, or even complex eigenvalues, or is rectangular and has no eigenvalues at all. However, for symmetric matrices, there is a direct connection between the two quantities:

Proposition 8.31. If $A = A^T$ is a symmetric matrix, its singular values are the absolute values of its eigenvalues: $\sigma_i = |\lambda_i|$; its singular vectors coincide with the associated eigenvectors.

Proof: When A is symmetric, $K = A^T A = A^2$. So, if $A\mathbf{v} = \lambda\mathbf{v}$, then $K\mathbf{v} = A^2\mathbf{v} = \lambda^2\mathbf{v}$. Thus, every eigenvector \mathbf{v} of A is also an eigenvector of K with eigenvalue λ^2 . Therefore, the eigenvector basis of A is an eigenvector basis for K , and hence forms a complete system of singular vectors for A also. *Q.E.D.*

The standard convention is to label the singular values in *decreasing* order, so that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$. Thus, σ_1 will always denote the largest or *dominant* singular value. If $A^T A$ has repeated eigenvalues, the singular values of A are repeated with the same multiplicities.

Example 8.32. Let $A = \begin{pmatrix} 3 & 5 \\ 4 & 0 \end{pmatrix}$. The associated Gram matrix is $K = A^T A = \begin{pmatrix} 25 & 15 \\ 15 & 25 \end{pmatrix}$, with eigenvalues $\lambda_1 = 40$ and $\lambda_2 = 10$. Thus, the singular values of A are $\sigma_1 = \sqrt{40} \approx 6.3246\dots$ and $\sigma_2 = \sqrt{10} \approx 3.1623\dots$. Note that these are *not* the same as its eigenvalues, namely $\lambda_1 = \frac{1}{2}(3 + \sqrt{89}) \approx 6.2170\dots$, $\lambda_2 = \frac{1}{2}(3 - \sqrt{89}) \approx -3.2170$.

[†] *Warning:* Some authors, e.g., [123], only designate the nonzero σ_i 's as singular values.

A rectangular matrix Σ will be called *diagonal* if its only nonzero entries are on the main diagonal starting in the upper left hand corner, and so $\sigma_{ij} = 0$ for $i \neq j$. An example is the matrix

$$\Sigma = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

whose only nonzero entries are in the diagonal $(1, 1)$ and $(2, 2)$ positions. (Its last diagonal entry happens to be 0.)

The generalization of the spectral factorization to non-symmetric matrices is known as the *singular value decomposition*, commonly abbreviated as SVD. Unlike the spectral decomposition, the singular value decomposition applies to arbitrary real rectangular matrices.

Theorem 8.33. *Any real $m \times n$ matrix A can be factorized*

$$A = P \Sigma Q^T \quad (8.39)$$

into the product of an $m \times m$ orthogonal matrix P , the $m \times n$ diagonal matrix Σ that has the first $l = \min\{m, n\}$ singular values of A as its diagonal entries, and an $n \times n$ orthogonal matrix Q^T .

Proof: Writing the factorization (8.39) as $AQ = P\Sigma$, and looking at the columns of the resulting matrix equation, we are led to the vector equations

$$A\mathbf{u}_i = \sigma_i \mathbf{v}_i, \quad i = 1, \dots, n, \quad (8.40)$$

relating the orthonormal columns of $Q = (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_n)$ to the orthonormal columns of $P = (\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_m)$. The scalars σ_i in (8.40) are the diagonal entries of Σ or, if $m < i \leq n$, equal to 0. The fact that P and Q are both orthogonal matrices means that their column vectors form orthonormal bases for, respectively, \mathbb{R}^m and \mathbb{R}^n under the Euclidean dot product. In this manner, the singular values indicate how far the linear transformation represented by the matrix A stretches a distinguished set of orthonormal basis vectors.

To construct the required bases, we prescribe $\mathbf{u}_1, \dots, \mathbf{u}_n$ to be the orthonormal eigenvector basis of the Gram matrix $K = A^T A$; thus

$$A^T A \mathbf{u}_j = K \mathbf{u}_j = \lambda_j \mathbf{u}_j = \sigma_j^2 \mathbf{u}_j.$$

We claim that the image vectors $\mathbf{w}_i = A\mathbf{u}_i$ are automatically orthogonal. Indeed, in view of the orthonormality of the \mathbf{u}_i ,

$$\mathbf{w}_i \cdot \mathbf{w}_j = \mathbf{w}_i^T \mathbf{w}_j = (A\mathbf{u}_i)^T A\mathbf{u}_j = \mathbf{u}_i^T A^T A \mathbf{u}_j = \sigma_j^2 \mathbf{u}_i^T \mathbf{u}_j = \sigma_j^2 \mathbf{u}_i \cdot \mathbf{u}_j = \begin{cases} 0, & i \neq j, \\ \sigma_i^2, & i = j. \end{cases} \quad (8.41)$$

Consequently, $\mathbf{w}_1, \dots, \mathbf{w}_n$ form an orthogonal system of vectors having respective norms

$$\|\mathbf{w}_i\| = \sqrt{\mathbf{w}_i \cdot \mathbf{w}_i} = \sigma_i.$$

Since $\mathbf{u}_1, \dots, \mathbf{u}_n$ form a basis of \mathbb{R}^n , their images $\mathbf{w}_1 = A\mathbf{u}_1, \dots, \mathbf{w}_n = A\mathbf{u}_n$ span $\text{rng } A$. Suppose that A has r non-zero singular values, so $\sigma_{r+1} = \dots = \sigma_n = 0$. Then

the corresponding image vectors $\mathbf{w}_1, \dots, \mathbf{w}_r$ are non-zero, mutually orthogonal vectors, and hence form an orthogonal basis for $\text{rng } A$. Since the dimension of $\text{rng } A$ is equal to its rank, this implies that the number of non-zero singular values is $r = \text{rank } A$. The corresponding unit vectors

$$\mathbf{v}_i = \frac{\mathbf{w}_i}{\sigma_i} = \frac{A\mathbf{u}_i}{\sigma_i}, \quad i = 1, \dots, r, \quad (8.42)$$

are an orthonormal basis for $\text{rng } A$. Let us further select an orthonormal basis $\mathbf{v}_{r+1}, \dots, \mathbf{v}_m$ for its orthogonal complement $\text{coker } A = (\text{rng } A)^\perp$. The combined set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_m$ clearly forms an orthonormal basis of \mathbb{R}^m , and satisfies (8.40) as required. In this manner, the resulting orthonormal bases $\mathbf{u}_1, \dots, \mathbf{u}_n$ and $\mathbf{v}_1, \dots, \mathbf{v}_m$ form the respective columns of the orthogonal matrices Q, P in the singular value decomposition (8.39). *Q.E.D.*

Warning: If $m < n$, then only the first m singular values appear along the diagonal of Σ . It follows from the proof that the remaining $n - m$ singular values are all zero.

Example 8.34. For the matrix $A = \begin{pmatrix} 3 & 5 \\ 4 & 0 \end{pmatrix}$ considered in Example 8.32, the orthonormal eigenvector basis of $K = A^T A = \begin{pmatrix} 25 & 15 \\ 15 & 25 \end{pmatrix}$ is given by the unit singular vectors $\mathbf{u}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ and $\mathbf{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$. Thus, $Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$. On the other hand, according to (8.42),

$$\mathbf{v}_1 = \frac{A\mathbf{u}_1}{\sigma_1} = \frac{1}{\sqrt{40}} \begin{pmatrix} 4\sqrt{2} \\ 2\sqrt{2} \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}, \quad \mathbf{v}_2 = \frac{A\mathbf{u}_2}{\sigma_2} = \frac{1}{\sqrt{10}} \begin{pmatrix} \sqrt{2} \\ -2\sqrt{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix},$$

and thus $P = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix}$. You may wish to validate the resulting singular value decomposition

$$A = \begin{pmatrix} 3 & 5 \\ 4 & 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{40} & 0 \\ 0 & \sqrt{10} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = P\Sigma Q^T.$$

As their name suggests, the singular values can be used to detect singular matrices. Indeed, the singular value decomposition tells us some interesting new geometrical information about matrix multiplication, supplying additional details to the discussion begun in Section 2.5 and continued in Section 5.6. The next result follows directly from the proof of Theorem 8.33.

Theorem 8.35. Let $\sigma_1, \dots, \sigma_r > 0$ be the non-zero singular values of the $m \times n$ matrix A . Let $\mathbf{v}_1, \dots, \mathbf{v}_m$ and $\mathbf{u}_1, \dots, \mathbf{u}_n$ be the orthonormal bases of, respectively, \mathbb{R}^m and \mathbb{R}^n provided by the columns of P and Q in its singular value decomposition $A = P\Sigma Q^T$. Then (i) $r = \text{rank } A$,

- (ii) $\mathbf{u}_1, \dots, \mathbf{u}_r$ form an orthonormal basis for $\text{corng } A$,
- (iii) $\mathbf{u}_{r+1}, \dots, \mathbf{u}_n$ form an orthonormal basis for $\ker A$,
- (iv) $\mathbf{v}_1, \dots, \mathbf{v}_r$ form an orthonormal basis for $\text{rng } A$,
- (v) $\mathbf{v}_{r+1}, \dots, \mathbf{v}_m$ form an orthonormal basis for $\text{coker } A$.

We already noted in Section 5.6 that the linear transformation $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by matrix multiplication, $L[\mathbf{x}] = A\mathbf{x}$, can be interpreted as a projection from \mathbb{R}^n to $\text{corng } A$ followed by an invertible map from $\text{corng } A$ to $\text{rng } A$. The singular value decomposition tells us that not only is the latter map invertible, it is simply a combination of stretches in the r mutually orthogonal singular directions $\mathbf{u}_1, \dots, \mathbf{u}_r$, whose magnitudes equal the nonzero singular values. In this way, we have at last reached a complete understanding of the subtle geometry underlying the simple operation of matrix multiplication.

An alternative interpretation of the singular value decomposition is to view the two orthogonal matrices in (8.39) as defining rigid rotations/reflections. Therefore, in all cases, a linear transformation from \mathbb{R}^n to \mathbb{R}^m is composed of three ingredients:

- (i) A rotation/reflection of the domain space \mathbb{R}^n , as prescribed by Q^T , followed by
- (ii) a simple stretching map of the coordinate vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$ of domain space, mapping \mathbf{e}_i to $\sigma_i \mathbf{e}_i$ in the target space \mathbb{R}^m , followed by
- (iii) a rotation/reflection of the target space, as prescribed by P .

In fact, in most cases we can choose both P and Q to be proper orthogonal matrices representing rotations; see Exercise ■.

Condition Number, Rank, and Principal Component Analysis

The singular values not only provide a nice geometric interpretation of the action of the matrix, they also play a key role in modern computational algorithms. The relative magnitudes of the singular values can be used to distinguish well-behaved linear systems from ill-conditioned systems which are much trickier to solve accurately. Since the number of nonzero singular values equals its rank, a square matrix that has one or more zero singular values is singular. A matrix with one or more very small singular values should be considered to be close to singular. Such ill-conditioning is traditionally quantified as follows.

Definition 8.36. The *condition number* of an $n \times n$ matrix is the ratio between its largest and smallest singular value: $\kappa(A) = \sigma_1/\sigma_n$.

A matrix with a very large condition number is said to be *ill-conditioned*; in practice, this occurs when the condition number is larger than the reciprocal of the machine's precision, e.g., 10^6 for most single precision arithmetic. As the name implies, it is much harder to solve a linear system $A\mathbf{x} = \mathbf{b}$ when its coefficient matrix is ill-conditioned. In the extreme case when A has one or more zero singular values, so $\sigma_n = 0$, its condition number is infinite, and the linear system is singular, with either no solution or infinitely many solutions.

The accurate computation of the rank of a matrix can be a numerical challenge. Small numerical errors in the entries can have a dramatic effect. For example, $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ 3 & 3 & -3 \end{pmatrix}$

has rank $r = 1$, but a tiny change, say to $\tilde{A} = \begin{pmatrix} 1.00001 & 1 & -1 \\ 2 & 2.00001 & -2 \\ 3 & 3 & -3.00001 \end{pmatrix}$, will produce a nonsingular matrix with rank $r = 3$. The latter matrix, however, is very close to singular, and this is highlighted by their respective singular values. For the first matrix, they are $\sigma_1 = \sqrt{42} \approx 6.48$ and $\sigma_2 = \sigma_3 = 0$, reconfirming that A has rank 1, whereas for \tilde{A} we find $\sigma_1 \approx 6.48075$ while $\sigma_2 \approx \sigma_3 \approx .000001$. The fact that the second and third singular values are very small indicates that \tilde{A} is very close to a matrix of rank 1 and should be viewed as a numerical or experimental perturbation of such a matrix. Thus, the most effective practical method for computing the rank of a matrix is to first assign a threshold, e.g., 10^{-5} , for singular values, and then treat any small singular value lying below the threshold as if it were zero.

This idea underlies the method of *Principal Component Analysis* that is playing an increasingly visible role in modern statistics, data analysis and data mining, imaging, and a variety of other fields, [89]. The singular vectors associated with the larger singular values indicate the *principal components* of the matrix, while small singular values indicate unimportant directions. In applications, the columns of the matrix A represent the data vectors, which are normalized to have mean $\mathbf{0}$. The corresponding Gram matrix $K = A^T A$ can be identified as the associated *covariance matrix*. Its eigenvectors are the principal components that serve indicate directions of correlation and clustering to be found in the data. Classification of patterns in images, sounds, semantics, and so on are being successfully handled by this powerful new approach.

The Pseudoinverse

With the singular value decomposition in hand, we are able to introduce a generalization of the inverse that applies to cases when the matrix in question is singular or even rectangular. We begin with the diagonal case. Let Σ be an $m \times n$ diagonal matrix with r nonzero diagonal entries $\sigma_1, \dots, \sigma_r$. We define the *pseudoinverse* of Σ to be the $n \times m$ diagonal matrix Σ^+ whose nonzero diagonal entries are the reciprocals $1/\sigma_1, \dots, 1/\sigma_r$. For example, if

$$\Sigma = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{then} \quad \Sigma^+ = \begin{pmatrix} \frac{1}{5} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The zero diagonal entries are *not* inverted. In particular, if Σ is a nonsingular square diagonal matrix, then its pseudoinverse and ordinary inverse are the same: $\Sigma^+ = \Sigma^{-1}$.

Definition 8.37. The *pseudoinverse* of an $m \times n$ matrix A with singular value decomposition $A = P \Sigma Q^T$ is the $n \times m$ matrix $A^+ = Q \Sigma^+ P^T$.

Note that the latter equation is the singular value decomposition of the pseudoinverse A^+ , and hence its nonzero singular values are the reciprocals of the nonzero singular values of A . If A is a non-singular square matrix, then its pseudoinverse agrees with its ordinary

inverse, since

$$A^{-1} = (P \Sigma Q^T)^{-1} = (Q^{-1})^T \Sigma^{-1} P^{-1} = Q \Sigma^+ P^T = A^+,$$

where we used the fact that the inverse of an orthogonal matrix is equal to its transpose.

If A is square and nonsingular, then, as we know, the solution to the linear system $A \mathbf{x} = \mathbf{b}$ is given by $\mathbf{x}^* = A^{-1} \mathbf{b}$. For a general coefficient matrix, the vector $\mathbf{x}^* = A^+ \mathbf{b}$ obtained by applying the pseudoinverse to the right hand side plays a distinguished role — it is the *least squares solution* to the system! In this manner, the pseudoinverse provides us with a direct route to least squares solutions to systems of linear equations.

Theorem 8.38. *Consider the linear system $A \mathbf{x} = \mathbf{b}$. Let $\mathbf{x}^* = A^+ \mathbf{b}$, where A^+ is the pseudoinverse of A . If $\ker A = \{\mathbf{0}\}$, then $\mathbf{x}^* \in \text{corng } A$ is the least squares solution to the system. If, more generally, $\ker A \neq \{\mathbf{0}\}$, then \mathbf{x}^* is the least squares solution of minimal Euclidean norm among all vectors that minimize the least squares error $\|A \mathbf{x} - \mathbf{b}\|$.*

Proof: To show that $\mathbf{x}^* = A^+ \mathbf{b}$ is the least squares solution to the system, we must check that it satisfies the normal equations $A^T A \mathbf{x}^* = A^T \mathbf{b}$. Using the definition of the pseudoinverse and the singular value decomposition (8.39), we find

$$\begin{aligned} A^T A \mathbf{x}^* &= A^T A A^+ \mathbf{b} = (P \Sigma Q^T)^T (P \Sigma Q^T) (Q \Sigma^+ P^T) \mathbf{b} \\ &= Q \Sigma^T \Sigma \Sigma^+ P^T \mathbf{b} = Q \Sigma^T P^T \mathbf{b} = A^T \mathbf{b}, \end{aligned}$$

where the next to last equality is left as Exercise ■ for the reader. This proves that \mathbf{x}^* solves the normal equations, and hence minimizes the least squares error[†].

Thus, when $\text{rank } A = n$, the vector \mathbf{x}^* is the unique least squares solution to the system. More generally, if A has rank $r < n$, so, $\ker A \neq \{\mathbf{0}\}$, only the first r singular values are nonzero, and therefore the last $n - r$ rows of Σ^+ are all zero. This implies that the last $n - r$ entries of the vector $\mathbf{c} = \Sigma^+ P^T \mathbf{b}$ are also all zero, so $\mathbf{c} = (c_1, \dots, c_r, 0, \dots, 0)^T$. We conclude that

$$\mathbf{x}^* = A^+ \mathbf{b} = Q \Sigma^+ P^T \mathbf{b} = Q \mathbf{c} = c_1 \mathbf{u}_1 + \dots + c_r \mathbf{u}_r$$

is a linear combination of the first r singular vectors, and hence, by Theorem 8.35, $\mathbf{x}^* \in \text{corng } A$. The most general least squares solution has the form $\mathbf{x} = \mathbf{x}^* + \mathbf{z}$ where $\mathbf{z} \in \ker A$, and the fact that $\|\mathbf{x}^*\|$ is minimized follows as in Theorem 5.57. Q.E.D.

When forming the pseudoinverse, we see that very small singular values lead to very large entries in Σ^+ , which will cause numerical difficulties when computing the least squares solution $\mathbf{x}^* = A^+ \mathbf{b}$ to the linear system. A common and effective computational strategy to avoid the effects of small singular values is to replace the corresponding diagonal entries of the pseudoinverse Σ^+ by 0. This has the effect of regularizing ill-conditioned matrices that are very close to singular — rather than solve the system directly for $\mathbf{x} = A^{-1} \mathbf{b}$, one would employ the suitably regularized pseudoinverse.

Finally, we note that practical numerical algorithms for computing singular values and the singular value decomposition can be found in [68, 123]

[†] In Chapter 4, this was proved under the assumption that $\ker A = \{\mathbf{0}\}$. You are asked to establish the general case in Exercise ■.

8.6. Incomplete Matrices and the Jordan Canonical Form.

Unfortunately, not all matrices are complete. Matrices that do not have enough (complex) eigenvectors to form a basis are considerably less pleasant to work with. However, as they occasionally appear in applications, it is worth learning how to handle them. We shall show how to supplement the eigenvectors in order to obtain a basis in which the matrix assumes a simple, but now non-diagonal form. The resulting construction is named after the nineteenth century French mathematician Camille Jordan[‡].

Throughout this section, A will be an $n \times n$ matrix, with either real or complex entries. We let $\lambda_1, \dots, \lambda_k$ denote the *distinct* eigenvalues of A . We recall that Theorem 8.10 guarantees that every matrix has at least one (complex) eigenvalue, so $k \geq 1$. Moreover, we are assuming that $k < n$, as otherwise A would be complete.

Definition 8.39. A *Jordan chain* of length j is a sequence of non-zero vectors $\mathbf{w}_1, \dots, \mathbf{w}_j \in \mathbb{C}^n$ that satisfies

$$A\mathbf{w}_1 = \lambda\mathbf{w}_1, \quad A\mathbf{w}_i = \lambda\mathbf{w}_i + \mathbf{w}_{i-1}, \quad i = 2, \dots, j, \quad (8.43)$$

where λ is an eigenvalue of A .

Note that the initial vector \mathbf{w}_1 in a Jordan chain is a genuine eigenvector, and so Jordan chains only exist when λ is an eigenvalue. The other vectors, $\mathbf{w}_2, \dots, \mathbf{w}_j$, are *generalized eigenvectors*, in accordance with the following definition.

Definition 8.40. A nonzero vector $\mathbf{w} \neq \mathbf{0}$ that satisfies

$$(A - \lambda I)^k \mathbf{w} = \mathbf{0} \quad (8.44)$$

for some $k > 0$ and $\lambda \in \mathbb{C}$ is called a *generalized eigenvector* of the matrix A .

Note that every ordinary eigenvector is automatically a generalized eigenvector, since we can just take $k = 1$ in (8.44); the converse is not necessarily valid. We shall call the minimal value of k for which (8.44) holds the *index* of the generalized eigenvector. Thus, an ordinary eigenvector is a generalized eigenvector of index 1. Since $A - \lambda I$ is nonsingular whenever λ is not an eigenvalue of A , its k^{th} power $(A - \lambda I)^k$ is also nonsingular. Therefore, generalized eigenvectors can only exist when λ is an ordinary eigenvalue of A — there are no additional “generalized eigenvalues”.

Lemma 8.41. The i^{th} vector \mathbf{w}_i in a Jordan chain (8.43) is a generalized eigenvector of index i .

Proof: By definition, $(A - \lambda I)\mathbf{w}_1 = \mathbf{0}$, and so \mathbf{w}_1 is an eigenvector. Next, we have $(A - \lambda I)\mathbf{w}_2 = \mathbf{w}_1$, and so $(A - \lambda I)^2\mathbf{w}_2 = (A - \lambda I)\mathbf{w}_1 = \mathbf{0}$. Thus, \mathbf{w}_2 a generalized eigenvector of index 2. A simple induction proves that $(A - \lambda I)^i\mathbf{w}_i = \mathbf{0}$. *Q.E.D.*

[‡] No relation to Wilhelm Jordan of Gauss–Jordan fame.

Example 8.42. Consider the 3×3 Jordan block $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. The only eigenvalue is $\lambda = 2$, and $A - 2I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. We claim that the standard basis vectors \mathbf{e}_1 , \mathbf{e}_2 and \mathbf{e}_3 form a Jordan chain. Indeed, $A\mathbf{e}_1 = 2\mathbf{e}_1$, and hence $\mathbf{e}_1 \in \ker(A - 2I)$ is a genuine eigenvector. Furthermore, $A\mathbf{e}_2 = 2\mathbf{e}_2 + \mathbf{e}_1$, and $A\mathbf{e}_3 = 2\mathbf{e}_3 + \mathbf{e}_2$, as you can easily check. Thus, \mathbf{e}_1 , \mathbf{e}_2 and \mathbf{e}_3 satisfy the Jordan chain equations for the eigenvalue $\lambda = 2$. Note that \mathbf{e}_2 lies in the kernel of $(A - 2I)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, and so is a generalized eigenvector of index 2. Indeed, every vector of the form $\mathbf{w} = a\mathbf{e}_1 + b\mathbf{e}_2$ with $b \neq 0$ is a generalized eigenvector of index 2. (When $b = 0$, $a \neq 0$, the vector $\mathbf{w} = a\mathbf{e}_1$ is an ordinary eigenvector of index 1.) Finally, $(A - 2I)^3 = O$, and so every vector $\mathbf{v} \in \mathbb{R}^3$, including \mathbf{e}_3 , is a generalized eigenvector of index 3 (or less).

A basis of \mathbb{R}^n or \mathbb{C}^n is called a *Jordan basis* for the matrix A if it consists of one or more Jordan chains that have no elements in common. Thus, for the Jordan matrix in Example 8.42, the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is, in fact, a Jordan basis. An eigenvector basis qualifies as a Jordan basis, since each eigenvector belongs to a Jordan chain of length 1. Jordan bases are the desired extension of eigenvector bases, and every square matrix has one.

Theorem 8.43. Every $n \times n$ matrix admits a Jordan basis of \mathbb{C}^n . The first elements of the Jordan chains form a maximal system of linearly independent eigenvectors. Moreover, the number of generalized eigenvectors in the Jordan basis that belong to the Jordan chains associated with the eigenvalue λ is the same as the eigenvalue's multiplicity.

Example 8.44. Consider the matrix

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

With some work, its characteristic equation is found to be

$$p_A(\lambda) = \det(A - \lambda I) = \lambda^5 + \lambda^4 - 5\lambda^3 - \lambda^2 + 8\lambda - 4 = (\lambda - 1)^3(\lambda + 2)^2 = 0,$$

and hence A has two eigenvalues: $\lambda_1 = 1$, which is a triple eigenvalue, and $\lambda_2 = -2$, which is double. Solving the associated homogeneous systems $(A - \lambda_j I)\mathbf{v} = \mathbf{0}$, we discover that, up to constant multiple, there are only two eigenvectors: $\mathbf{v}_1 = (0, 0, 0, -1, 1)^T$ for $\lambda_1 = 1$ and, anticipating our final numbering, $\mathbf{v}_4 = (-1, 1, 1, -2, 0)^T$ for $\lambda_2 = -2$. Thus, A is far from complete.

To construct a Jordan basis, we first note that since A has 2 linearly independent eigenvectors, the Jordan basis will contain two Jordan chains; the one associated with the

triple eigenvalue $\lambda_1 = 1$ has length 3, while $\lambda_2 = -2$ admits a Jordan chain of length 2. To construct the former, we need to first solve the system $(A - I)\mathbf{w} = \mathbf{v}_1$. Note that the coefficient matrix is singular — it must be since 1 is an eigenvalue — and the general solution is $\mathbf{w} = \mathbf{v}_2 + t\mathbf{v}_1$ where $\mathbf{v}_2 = (0, 1, 0, 0, -1)^T$, and t is the free variable. The appearance of an arbitrary multiple of the eigenvector \mathbf{v}_1 in the solution is not unexpected; indeed, the kernel of $A - I$ is the eigenspace for $\lambda_1 = 1$. We can choose any solution, e.g., \mathbf{v}_2 as the second element in the Jordan chain. To find the last element of the chain, we solve $(A - I)\mathbf{w} = \mathbf{v}_2$ to find $\mathbf{w} = \mathbf{v}_3 + t\mathbf{v}_1$ where $\mathbf{v}_3 = (0, 0, 0, 1, 0)^T$ can be used as the last element of this Jordan chain. Similarly, to construct the Jordan chain for the second eigenvalue, we solve $(A + 2I)\mathbf{w} = \mathbf{v}_4$ and find $\mathbf{w} = \mathbf{v}_5 + t\mathbf{v}_4$ where $\mathbf{v}_5 = (-1, 0, 0, -2, 1)^T$. Thus, the desired Jordan basis is

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -2 \\ 0 \end{pmatrix}, \quad \mathbf{v}_5 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ -2 \\ 1 \end{pmatrix},$$

with $A\mathbf{v}_1 = \mathbf{v}_1$, $A\mathbf{v}_2 = \mathbf{v}_1 + \mathbf{v}_2$, $A\mathbf{v}_3 = \mathbf{v}_2 + \mathbf{v}_3$, $A\mathbf{v}_4 = -2\mathbf{v}_4$, $A\mathbf{v}_5 = \mathbf{v}_4 - 2\mathbf{v}_5$.

To prove Theorem 8.43, we begin with a simple lemma.

Lemma 8.45. *If $\mathbf{v}_1, \dots, \mathbf{v}_n$ forms a Jordan basis for the matrix A , it also forms a Jordan basis for $B = A - cI$, for any scalar c .*

Proof: We note that the eigenvalues of B are of the form $\lambda - c$, where λ is an eigenvalue of A . Moreover, given a Jordan chain $\mathbf{w}_1, \dots, \mathbf{w}_j$ of A , we have

$$B\mathbf{w}_1 = (\lambda - c)\mathbf{w}_1, \quad B\mathbf{w}_i = (\lambda - c)\mathbf{w}_i + \mathbf{w}_{i-1}, \quad i = 2, \dots, j,$$

so $\mathbf{w}_1, \dots, \mathbf{w}_j$ is also a Jordan chain for B corresponding to the eigenvalue $\lambda - c$. *Q.E.D.*

The proof of Theorem 8.43 will be done by induction on the size n of the matrix. The case $n = 1$ is trivial, since any nonzero element of \mathbb{C} is a Jordan basis for a 1×1 matrix $A = (a)$. To perform the induction step, we assume that the result is valid for all matrices of size $\leq n - 1$. Let A be an $n \times n$ matrix. According to Theorem 8.10, A has at least one complex eigenvalue λ . Let $B = A - \lambda I$. Since λ is an eigenvalue of A , we know that 0 is an eigenvalue of B . This means that $\ker B \neq \{\mathbf{0}\}$, and so $r = \text{rank } B < n$. Moreover, by Lemma 8.45, any Jordan basis of B is also a Jordan basis for A , and so we can concentrate all our attention on the singular matrix B from now on.

We note that $W = \text{rng } B \subset \mathbb{C}^n$ is an invariant subspace, i.e., $B\mathbf{w} \in W$ whenever $\mathbf{w} \in W$, cf. Exercise ■. Moreover, since B is singular, $\dim W = r = \text{rank } B < n$. Thus, by fixing a basis of W , we can realize the restriction $B: W \rightarrow W$ as multiplication by an $r \times r$ matrix. The fact that $r < n$ allows us to invoke the induction hypothesis, and deduce the existence of a Jordan basis $\mathbf{w}_1, \dots, \mathbf{w}_r \in W \subset \mathbb{C}^n$ for the action of B on the subspace W . Our goal is to complete this collection to a full Jordan basis on \mathbb{C}^n .

To this end, we append two additional kinds of vectors. Suppose that the Jordan basis of W contains k null Jordan chains associated with its zero eigenvalue. Each null Jordan

chain consists of vectors $\mathbf{w}_1, \dots, \mathbf{w}_j \in W$ satisfying $B\mathbf{w}_1 = \mathbf{0}$, $B\mathbf{w}_2 = \mathbf{w}_1$, \dots , $B\mathbf{w}_j = \mathbf{w}_{j-1}$. The number of null Jordan chains is equal to the number of linearly independent null eigenvectors of B in $W = \text{rng } B$, that is $k = \dim(\ker B \cap \text{rng } B)$. To each null Jordan chain, we append a vector $\mathbf{w}_{j+1} \in \mathbb{C}^n$ such that $B\mathbf{w}_{j+1} = \mathbf{w}_j$, noting that \mathbf{w}_{j+1} exists because $\mathbf{w}_j \in \text{rng } B$. Appending \mathbf{w}_{j+1} , we deduce that $\mathbf{w}_1, \dots, \mathbf{w}_{j+1} \in \mathbb{C}^n$ forms a null Jordan chain, of length $j + 1$, for B . Having extended all the null Jordan chains in W , the resulting collection contains $r + k$ vectors in \mathbb{C}^n arranged in nonoverlapping Jordan chains. To complete to a basis, we include $n - r - k$ additional linearly independent null vectors $\mathbf{z}_1, \dots, \mathbf{z}_{n-r-k} \in \ker B \setminus \text{rng } B$ that lie outside its range. Since $B\mathbf{z}_j = \mathbf{0}$, each \mathbf{z}_j forms a null Jordan chain of length 1. We claim that the complete collection consisting of the non-null Jordan chains in W , the k extended null chains, and the additional null vectors $\mathbf{z}_1, \dots, \mathbf{z}_{n-r-k}$, forms the desired Jordan basis. By construction, it consists of nonoverlapping Jordan chains. The only remaining issue is proving that the vectors are linear independent, which is left as a challenge for the reader in Exercise ■. Q.E.D.

Just as an eigenvector basis diagonalizes a complete matrix, a Jordan basis provides a particularly simple form for an incomplete matrix, known as the *Jordan canonical form*.

Definition 8.46. An $n \times n$ matrix of the form[†]

$$J_{\lambda,n} = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \ddots & \ddots \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}, \quad (8.45)$$

in which λ is a real or complex number, is known as a *Jordan block*.

In particular, a 1×1 Jordan block is merely a scalar $J_{\lambda,1} = \lambda$. Since every matrix has at least one (complex) eigenvector — see Theorem 8.10 — the Jordan block matrices have the least possible number of eigenvectors.

Lemma 8.47. The $n \times n$ Jordan block matrix $J_{\lambda,n}$ has a single eigenvalue, λ , and a single independent eigenvector, \mathbf{e}_1 . The standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$ form a Jordan chain for $J_{\lambda,n}$.

Definition 8.48. A *Jordan matrix* is a square matrix of block diagonal form

$$J = \text{diag}(J_{\lambda_1,n_1}, J_{\lambda_2,n_2}, \dots, J_{\lambda_k,n_k}) = \begin{pmatrix} J_{\lambda_1,n_1} & & & \\ & J_{\lambda_2,n_2} & & \\ & & \ddots & \\ & & & J_{\lambda_k,n_k} \end{pmatrix}, \quad (8.46)$$

in which one or more Jordan blocks, not necessarily of the same size, lie along the diagonal, while all off-diagonal blocks are zero.

[†] All non-displayed entries are zero.

Note that the only non-zero entries in a Jordan matrix are those on the diagonal, which can have any complex value, and those on the superdiagonal, which are either 1 or 0. The positions of the superdiagonal 1's uniquely prescribes the Jordan blocks.

For example, the 6×6 matrices

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix},$$

are all Jordan matrices; the first is a diagonal matrix, consisting of 6 distinct 1×1 Jordan blocks; the second has a 4×4 Jordan block followed by a 2×2 block that happen to have the same diagonal entries; the last has three 2×2 Jordan blocks.

As a direct corollary of Lemma 8.47 combined with the matrix's block structure, cf. Exercise ■, we obtain a complete classification of the eigenvectors and eigenvalues of a Jordan matrix.

Lemma 8.49. *The Jordan matrix (8.46) has eigenvalues $\lambda_1, \dots, \lambda_k$. The standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$ form a Jordan basis; the Jordan chains are labeled by the Jordan blocks.*

Thus, in the preceding examples of Jordan matrices, the first has three double eigenvalues, 1, 2, 3, and corresponding linearly independent eigenvectors $\mathbf{e}_1, \mathbf{e}_6; \mathbf{e}_2, \mathbf{e}_5; \mathbf{e}_3, \mathbf{e}_4$, each of which belongs to a Jordan chain of length 1. The second matrix has only one eigenvalue, -1 , but two Jordan chains, namely $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ and $\mathbf{e}_5, \mathbf{e}_6$. The last has eigenvalues 0, 1, 2 and three Jordan chains, namely $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3, \mathbf{e}_4$, and $\mathbf{e}_5, \mathbf{e}_6$. In particular, the only complete Jordan matrices are the diagonal matrices, all of whose Jordan blocks are of size 1×1 .

Theorem 8.50. *Let A be an $n \times n$ real or complex matrix. Let $S = (\mathbf{w}_1 \ \mathbf{w}_2 \ \dots \ \mathbf{w}_n)$ be the matrix whose columns are a Jordan basis of A . Then S places A in Jordan canonical form*

$$S^{-1}AS = J = \text{diag}(J_{\lambda_1, n_1}, J_{\lambda_2, n_2}, \dots, J_{\lambda_k, n_k}). \quad (8.47)$$

The diagonal entries of the similar Jordan matrix J are the eigenvalues of A . In particular, A is complete (diagonalizable) if and only if every Jordan block is of size 1×1 or, equivalently, all Jordan chains are of length 1. The Jordan canonical form of A is uniquely determined up to a permutation of the diagonal Jordan blocks.

For instance, the matrix $A = \begin{pmatrix} -1 & 0 & 1 & 0 & 0 \\ -2 & 2 & -4 & 1 & 1 \\ -1 & 0 & -3 & 0 & 0 \\ -4 & -1 & 3 & 1 & 0 \\ 4 & 0 & 2 & -1 & 0 \end{pmatrix}$ considered in Example 8.44

has the following Jordan basis matrix and Jordan canonical form

$$S = \begin{pmatrix} 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & -2 & -2 \\ 1 & -1 & 0 & 0 & 1 \end{pmatrix}, \quad J = S^{-1}AS = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}.$$