4

Eigenvalues and Eigenvectors

Because in this book we are interested in state spaces, many of the linear operations we consider will be performed on vectors within a single space, resulting in transformed vectors within that same space. Thus, we will consider operators of the form $A: \boldsymbol{X} \to \boldsymbol{X}$. A special property of these operators is that they lead to special vectors and scalars known as *eigenvectors* and *eigenvalues*. These quantities are of particular importance in the stability and control of linear systems. In this chapter, we will discuss eigenvalues and eigenvectors and related concepts such as singular values.

4.1 A-Invariant Subspaces

In any space and for any operator *A*, there are certain subspaces in which, if we take a vector and operate on it using *A*, the result remains in that subspace. Formally,

> ***A-Invariant Subspace:*** Let $\mathbf{X}_1$ be a subspace of linear vector space $\mathbf{X}$. This subspace is said to be *A-invariant* if for every vector $\boldsymbol{x} \in \mathbf{X}_1$, $A\boldsymbol{x} \in \mathbf{X}_1$. When the operator *A* is understood from the context, then $\mathbf{X}_1$ is sometimes said to be simply "invariant." (4.1)

Finite-dimensional subspaces can always be thought of as lines, planes, or hyperplanes that pass through the origin of their parent space. In the next section, we will consider *A*-invariant subspaces consisting of lines through the origin, i.e., the *eigenvectors*.

4.2 Definitions of Eigenvectors and Eigenvalues

Recall that a linear operator A is simply a rule that assigns a new vector $A\boldsymbol{x}$ to an old vector $\boldsymbol{x}$. In general, operator A can take arbitrary actions on vector $\boldsymbol{x}$, scaling it and "moving" it throughout the space. However, there are special situations in which the action of A is simply to scale the vector $\boldsymbol{x}$ for some particular vectors $\boldsymbol{x}$ as pictured in Figure 4.1.

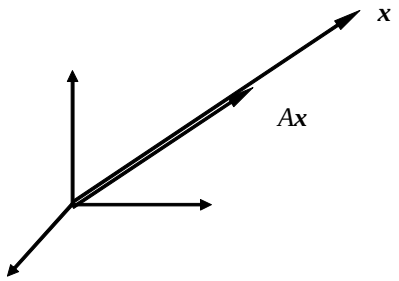

Figure 4.1 Scaling action of an operator acting on vector **x**.

If we denote the scaling factor in this situation by λ, then we have the relationship

$$A\boldsymbol{x} = \lambda\boldsymbol{x} \tag{4.2}$$

Note that this relationship will not hold for all vectors $\boldsymbol{x}$ and all scalars λ, but only for special specific instances. These are the *eigenvalues*[M] and *eigenvectors*.[M]

`eig(A)`

Notice that the eigenvectors clearly define one-dimensional A-invariant subspaces. In fact, the span of any collection of different eigenvectors will be an A-invariant subspace as well.

> ***Eigenvalues and Eigenvectors:*** In the relationship $A\boldsymbol{x} = \lambda\boldsymbol{x}$, the nonzero values of $\boldsymbol{x}$ are *eigenvectors*, and the corresponding values for λ (which *may* be zero) are the *eigenvalues*. (4.3)

Note that Equation (4.2) above also implies that

$$\begin{aligned} 0 &= \lambda\boldsymbol{x} - A\boldsymbol{x} \\ &= (\lambda I - A)\boldsymbol{x} \end{aligned} \tag{4.4}$$

We can therefore interpret an eigenvector as being a vector from the null space of $\lambda I - A$ corresponding to an eigenvalue λ.

A term that we will use in future chapters when collectively referring to an operator's eigenvalues is *spectrum*. The spectrum of an operator is simply the set of all its eigenvalues.

Example 4.1: Electric Fields

In an *isotropic* dielectric medium, the relationship between the electric field vector E and the electric flux density (also known as the *displacement vector*) is known to be

$$D = \varepsilon E \tag{4.5}$$

where the quantity ε is the *dielectric constant*. However, there are dielectrics that are *anisotropic*, which means that their dielectric constants depend on direction. For these materials,

$$D_i = \sum_{j=1}^{3} \varepsilon_{ij} E_j \tag{4.6}$$

or, in matrix form,

$$\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}$$

or

$$D = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix} E \tag{4.7}$$

For a particular medium with a known set of dielectric constants $\varepsilon_{ij}, i, j = 1, 2, 3,$ find the directions, if any, in which the electric field and the flux density are collinear.

Solution:

If the electric field and the flux density are collinear, then they are proportional to one another, so we could say that $D = \lambda E$ giving

$$\lambda E = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix} E \tag{4.8}$$

Recognizing that this is an eigenvalue-eigenvector problem, we can say that the electric field and flux density vectors are collinear only in the directions given by the eigenvectors of the matrix of dielectric constants (and the constants of proportionality λ will be the corresponding eigenvalues).

4.3 Finding Eigenvalues and Eigenvectors

The first step in determining eigenvalues and eigenvectors is to force $\lambda I - A$ to have a nontrivial null space. The easiest way to do this is to determine the matrix representation for A and set $|\lambda I - A| = 0$. The result for an $n \times n$ A-matrix will be an n^{th} order polynomial equation in the variable λ. By the fundamental theorem of algebra, this polynomial will have exactly n roots, and therefore each matrix of dimensions $n \times n$, or equivalently, each n-dimensional operator, will have exactly n eigenvalues (although we will find later that this does *not* imply that there will also be n eigenvectors). These, however, may be repeated and/or complex. Complex eigenvalues, of course, will appear in conjugate pairs.

Remark: The reader may have noticed that Equation (1.49) implies that, at least for a SISO system for which the division operation is defined, the denominator of the transfer function of a system is the n^{th}-order polynomial $|sI - A|$. This is, of course, the same polynomial whose roots give us the eigenvalues. The temptation is therefore to guess that the poles and eigenvalues are the same, but this is *not* the case. Although all the roots of $|sI - A|$ are eigenvalues, not all eigenvalues are poles of the transfer function. We will see the reason for this in Chapter 8.

Example 4.2: Simple Eigenvalue Problem

Find the eigenvalues for the operator represented by matrix

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

Solution:

Using the determinant rule,

$$\begin{aligned}|\lambda I - A| &= \begin{vmatrix} \lambda - 3 & 2 \\ 1 & \lambda - 4 \end{vmatrix} \\ &= (\lambda - 3)(\lambda - 4) - 2 \\ &= \lambda^2 - 7\lambda + 10 \\ &= (\lambda - 2)(\lambda - 5)\end{aligned}$$

Therefore, $|\lambda I - A| = 0$ gives $\lambda_1 = 2$ and $\lambda_2 = 5$ as the two eigenvalues.

To determine the eigenvectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ corresponding, respectively, to these eigenvalues, we realize that $\boldsymbol{x}_i \in \mathrm{N}(\lambda_i I - A)$ by construction, so we simply find a vector in this null space[M] to act as eigenvector $\boldsymbol{x}_i$.

`null(A)`

Example 4.3: Simple Eigenvector Problem

Determine the eigenvectors for the matrix operator given in the previous example.

Solution:

Treating each eigenvalue separately, consider λ_1 and seek solutions to:

$$(\lambda_1 I - A)\boldsymbol{x} = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix} \boldsymbol{x} = 0$$

It is obvious that the matrix $\lambda_1 I - A$ has a single linearly independent column and therefore has rank one. The dimension of its null space is therefore one and there will be a single vector solution, which is $\boldsymbol{x}_1 = \alpha \begin{bmatrix} 2 & 1 \end{bmatrix}^{\mathrm{T}}$ for any scalar α. For the other eigenvalue,

$$(\lambda_2 I - A)\boldsymbol{x} = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \boldsymbol{x} = 0$$

yields $\boldsymbol{x}_2 = \beta \begin{bmatrix} 1 & -1 \end{bmatrix}^{\mathrm{T}}$ for any scalar β.

By convention, we often eliminate the arbitrary constant from the eigenvectors by normalizing them. For example, the two eigenvectors for the example above would be given as

$$\boldsymbol{x}_1 = \begin{bmatrix} 2/\sqrt{5} \\ \\ 1/\sqrt{5} \end{bmatrix} \qquad \boldsymbol{x}_2 = \begin{bmatrix} 1/\sqrt{2} \\ \\ -1/\sqrt{2} \end{bmatrix}$$

It is understood that any multiple of such normalized eigenvectors will also be an eigenvector.

Example 4.4: Complex Eigenvalues and Eigenvectors

Suppose we have an operator that rotates vectors in $\Re^3$ clockwise about some other given vector $\boldsymbol{v}$ by an angle θ, as shown in Figure 4.2. This is a linear operator. If the axis of rotation is the direction of the vector $\boldsymbol{r} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$ and the angle of rotation is $\theta = 60°$ it can be shown that the matrix representation is ([4]):

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

Find the eigenvalues and eigenvectors for this operator.

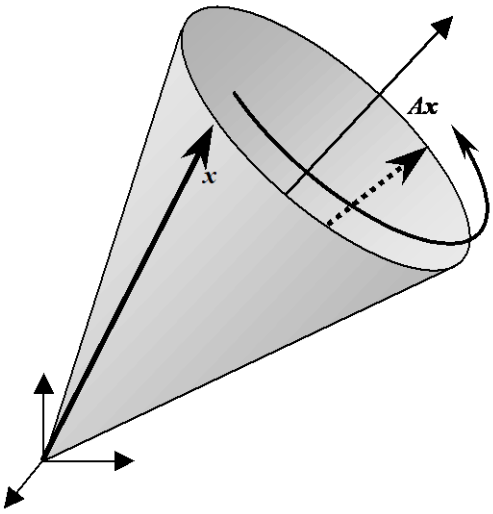

Figure 4.2 A vector rotating about a fixed, arbitrary axis. The rotating vector sweeps out the surface of a cone centered on the axis of rotation.

Solution:

Before proceeding with the computation of the eigenvalues and eigenvectors, consider the original definition. An eigenvector is a vector that is merely scaled by the action of the operator. In this situation, it is geometrically obvious that the only vector that will not "move" under this three-dimensional rotation operator will be a vector along the axis of rotation itself. That is, any vector of the form $\alpha \boldsymbol{v}$ will only rotate about itself, and in fact it will be scaled only by a factor of unity, i.e., its length will not change. We would expect any other vector in the space to rotate on the surface of the cone defined by the description of the operator. Thus, we can logically reason the existence of an eigenvalue of 1 and a corresponding eigenvector of $\alpha \boldsymbol{v}$.

By computing $|\lambda I - A|$ and finding the roots of the resulting polynomial, we find that the eigenvalues are $\lambda_1 = 1$, $\lambda_2 = \frac{1}{2} + j\frac{\sqrt{3}}{2}$, and $\lambda_3 = \frac{1}{2} - j\frac{\sqrt{3}}{2}$, where $j = \sqrt{-1}$. Furthermore, by finding the null spaces of the (complex) matrices $\lambda_i I - A$, we can compute the corresponding eigenvectors as

$$\boldsymbol{x}_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \qquad \boldsymbol{x}_2 = \begin{bmatrix} \frac{1}{2} + j\frac{\sqrt{3}}{2} \\ \frac{1}{2} - j\frac{1}{2\sqrt{3}} \\ j\frac{1}{\sqrt{3}} \end{bmatrix} \qquad \boldsymbol{x}_3 = \begin{bmatrix} \frac{1}{2} - j\frac{\sqrt{3}}{2} \\ \frac{1}{2} + j\frac{1}{2\sqrt{3}} \\ -j\frac{1}{\sqrt{3}} \end{bmatrix}$$

The first point to notice here is that the complex eigenvectors corresponding to the complex pair of eigenvalues is itself a complex conjugate pair. Second, we see that the observation about the real eigenvalue and its corresponding eigenvector is indeed true. However we did not predict the existence of the complex pair of eigenvalues and eigenvectors. In general, it can be said that complex eigenvalues and eigenvectors do not conform to the same geometric interpretation as real-valued eigenvalues and eigenvectors. Nevertheless, they are just as important for most purposes, including stability theory and control systems that we study in later chapters.

Example 4.5: Eigenvalues and Eigenvectors of Operators on Function Spaces

Let **V** denote the linear vector space of polynomials in x of degree ≤ 2. Also consider a linear operator T that takes an arbitrary vector $p(x)$ and transforms it into the vector

$$(Tp)(x) = \frac{d}{dx}[(1-x)p(x)] + 2p(0)$$

Find the operator's eigenvalues and eigenvectors if it is expressed in the basis $\{\boldsymbol{e}_i\} = \{1, x, x^2\}$.

Solution:

It is possible to find the eigenvalues and eigenvectors of such an operator directly, by manipulating the polynomials in the equation $(Tp)(x) = \lambda p(x)$. However, it is considerably easier to first express the operator as a matrix in the basis given, and then use existing numerical tools to compute the eigenvalues and eigenvectors.

To find the linear operator's matrix representation, we determine its effects on the basis vectors;

$$T(1) = \frac{d}{dx}\left[(1-x)\cdot 1\right] + 2\cdot(1) = 1 \qquad (= \begin{bmatrix}1 & 0 & 0\end{bmatrix}^{\mathrm{T}})$$

$$T(x) = \frac{d}{dx}\left[(1-x)\cdot x\right] + 2\cdot(0) = 1-2x \qquad (= \begin{bmatrix}1 & -2 & 0\end{bmatrix}^{\mathrm{T}})$$

$$T(x^2) = \frac{d}{dx}\left[(1-x)\cdot x^2\right] + 2\cdot(0) = 2x-3x^2 \qquad (= \begin{bmatrix}0 & 2 & -3\end{bmatrix}^{\mathrm{T}})$$

Therefore,

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

In such a case, where the matrix is upper- or lower-triangular, it can be easily shown that the elements on the diagonal will be the eigenvalues. We therefore have by inspection, $\lambda_1 = 1$, $\lambda_2 = -2$, and $\lambda_3 = -3$. For the unnormalized eigenvectors,

$$\boldsymbol{v}_1 = \mathrm{N}(1\cdot I - A) = \begin{bmatrix}1\\0\\0\end{bmatrix} \quad \boldsymbol{v}_2 = \mathrm{N}(-2\cdot I - A) = \begin{bmatrix}1\\-3\\0\end{bmatrix} \quad \boldsymbol{v}_3 = \mathrm{N}(-3\cdot I - A) = \begin{bmatrix}1\\-4\\2\end{bmatrix} \tag{4.9}$$

Note that written as polynomials in their original space **V**, these vectors are $\boldsymbol{v}_1 = 1$, $\boldsymbol{v}_2 = 1-3x$, and $\boldsymbol{v}_3 = 1-4x+2x^2$. We can use these expressions to

verify that the concept of eigenvalues and eigenvectors is not restricted to matrices by performing operator T on them directly as polynomials:

$$T(1) = 1 \qquad (= \lambda_1 \boldsymbol{v}_1)$$
$$T(1-3x) = 6x - 2 = -2(1-3x) \qquad (= \lambda_2 \boldsymbol{v}_2)$$
$$T(1-4x+2x^2) = -6x^2 + 12x - 3 = -3(1-4x+2x^2) \qquad (= \lambda_3 \boldsymbol{v}_3)$$

4.4 The Basis of Eigenvectors

In this section, we will show the result of changing the basis of an operator to the basis formed by a set of n eigenvectors. Such bases will be shown to have simplifying effects on the matrix representations of operators. These simplified forms of operators are known as *canonical*[M] *forms*. However, to form such a basis of eigenvectors, we must first know whether or not a complete set of n linearly independent eigenvectors exists, and if they do not, what we do about it.

`canon(sys,type)`

4.4.1 Changing to the Basis of Eigenvectors

For the time being, we will assume that our matrix of interest has a complete set of n linearly independent eigenvectors. This will allow us to change the basis of the operator to the basis of eigenvectors without any trouble. To do so by following the procedure of Section 2.2.4, we will need the matrix that relates the old basis vectors, which we will assume are the standard basis vectors, to the new basis vectors, which are the eigenvectors. As we found in Example 2.7, it is considerably easier to write expressions for the new basis vectors in terms of the old basis vectors. That is, if $\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n\}$ represents the old (standard) basis and $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\}$ represents the set of n eigenvectors, then it is obvious that

$$\boldsymbol{x}_i = \left[\boldsymbol{e}_1 \,\middle|\, \boldsymbol{e}_2 \,\middle|\, \cdots \,\middle|\, \boldsymbol{e}_n\right] \begin{bmatrix} x_{1i} \\ \hdashline x_{2i} \\ \hdashline \vdots \\ \hdashline x_{ni} \end{bmatrix}$$

or

$$\left[\boldsymbol{x}_1 \,\middle|\, \boldsymbol{x}_2 \,\middle|\, \cdots \,\middle|\, \boldsymbol{x}_n\right] = \left[\boldsymbol{e}_1 \,\middle|\, \boldsymbol{e}_2 \,\middle|\, \cdots \,\middle|\, \boldsymbol{e}_n\right] \left[\begin{array}{c|c|c|c} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{array}\right]$$

where x_{ji} is the j^{th} component of vector $\boldsymbol{x}_i$. Thus, as we did in Example 2.7, we construct the matrix

$$M = B^{-1} = \left[\begin{array}{c|c|c|c} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{array}\right] = \left[\boldsymbol{x}_1 \mid \boldsymbol{x}_2 \mid \cdots \mid \boldsymbol{x}_n\right] \tag{4.10}$$

The matrix *M* is known as the *modal matrix*, for reasons that will become clear in Chapter 6.

Then, from Equation (3.25), we find that

$$\hat{A} = M^{-1}AM \tag{4.11}$$

Before presenting an example, it is useful to predict what form this "new" operator matrix $\hat{A}$ will take. It is known that the i^{th} column of an operator matrix consists of the representation of the effect of that operator acting on the i^{th} basis vector. It is also known that when an operator operates on an eigenvector, the result is a scaled version of that eigenvector, i.e., $A\boldsymbol{x} = \lambda\boldsymbol{x}$. Therefore, when the basis is the set of eigenvectors, then the operator acting on the basis vectors will give vectors with components only along those same basis vectors. The i^{th} column of *A* should have nonzero entries in only the i^{th} position, i.e., the matrix $\hat{A}$ will be *diagonal*.

Example 4.6: Diagonalization of Operators

Change the matrix operator

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

which is expressed in the standard basis, into the basis of its own eigenvectors.

Solution:

We first find the eigenvalues:

$$0 = |\lambda I - A| = \begin{vmatrix} \lambda - 3 & 0 & -2 \\ 0 & \lambda - 3 & 2 \\ -2 & 2 & \lambda - 1 \end{vmatrix}$$
$$= \lambda^3 - 7\lambda^2 + 7\lambda + 15$$
$$= (\lambda - 5)(\lambda - 3)(\lambda + 1)$$

so $\lambda_1 = 5$, $\lambda_2 = 3$, and $\lambda_3 = -1$. Finding the (normalized) eigenvectors,

$$\boldsymbol{x}_1 = \mathrm{N}(\lambda_1 I - A) = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \qquad \boldsymbol{x}_2 = \mathrm{N}(\lambda_2 I - A) = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$\boldsymbol{x}_3 = \mathrm{N}(\lambda_3 I - A) = \begin{bmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

We therefore have the modal matrix

$$M = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix}$$

and the transformed operator matrix

$$\hat{A} = M^{-1} A M = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

which is, of course, diagonal.

4.4.2 Repeated Eigenvalues

In the examples considered so far, each eigenvalue has been distinct. The first case in which we will encounter matrices that do *not* have a complete set of n independent eigenvectors is when one or more of the eigenvalues are repeated. When an eigenvalue is repeated, i.e., the same λ is a multiple root of $|\lambda I - A|$,

some new situations can arise. If, for example, eigenvalue λ_j is repeated m_j times, we say that m_j is the *algebraic multiplicity* of λ_j. Two cases can arise:

1. If we compute the nullity of $\lambda_j I - A$ and find it to be $q(\lambda_j I - A) = n - r(\lambda_j I - A) = m_j$, then by definition, the dimension of $\mathrm{N}(\lambda_j I - A)$ is m_j. It would therefore be possible to find m_j linearly independent eigenvectors associated with the same eigenvalue λ_j. This would be done in the usual way, generating eigenvectors $\boldsymbol{x}_i, \ldots, \boldsymbol{x}_{m_j} \in \mathrm{N}(\lambda_j I - A)$ as in the previous chapter, probably with the help of a computer.[M]

`null(A)`

2. If $q(\lambda_j I - A) < m_j$, then we cannot find m_j eigenvectors because $\mathrm{N}(\lambda_j I - A)$ is not big enough. We therefore conclude that although every $n \times n$ matrix (or *n*-dimensional operator) has *n* eigenvalues, there are *not* always a full set of *n* eigenvectors.

Case 1 represents an eigenvector problem that is no more difficult than if all the eigenvalues were distinct. Such a situation defines an *A*-invariant subspace of dimension greater than one associated with a single eigenvalue as follows:

> ***Eigenspace:*** The set of all eigenvectors corresponding to an eigenvalue λ_i forms a basis for a subspace of **X**, called the *eigenspace* of λ_i. This eigenspace also happens to be the null space of a transformation defined as $\lambda_i I - A$. (4.12)

It is clear that the eigenspace of λ_i is *A*-invariant and has dimension equal to $q(\lambda_i I - A)$. The above definition is sometimes restricted to the real eigenvectors of *A* when **X** itself is real. This allows us to avoid defining complex subspaces of real vector spaces.

However, in case 2, we are faced with certain problems related to the use of eigenvectors. Most importantly, we clearly cannot diagonalize the *A*-matrix by changing to the basis of eigenvectors because we will not have a sufficient set of *n* eigenvectors that we need to construct an $n \times n$ transformation matrix. To resolve this difficulty, we introduce the "next best thing" to an eigenvector, called *a generalized eigenvector*.*

---

* Our use of the term "generalized eigenvector" is the classical one. In other texts (including MATLAB's manual entry under `EIG`), a generalized eigenvector is a vector $\boldsymbol{x}$ that solves the so-called "generalized eigenvalue problem," i.e., $(A - \lambda B)\boldsymbol{x} = 0$. The two usages are unrelated.

4.4.3 Generalized Eigenvectors

When an eigenvalue λ_i has an algebraic multiplicity $m_i > 1$, it may have any number of eigenvectors g_i, where $g_i \le m_i$, equal to the nullity of $\lambda_i I - A$:

$$g_i = q(\lambda_i I - A) \tag{4.13}$$

This number will be referred to as the geometric multiplicity of λ_i because it represents the dimension of the eigenspace of λ_i. Considering all possible eigenvalues, we will have a total of $\sum_i g_i \le n$ eigenvectors for an *n*-dimensional operator *A*.

When $\sum_i g_i < n$, we will have an insufficient number of eigenvectors to construct a modal matrix as in (4.10) and thereby diagonalize the operator *A*. However, we can define so-called "generalized eigenvectors" that serve a similar purpose. The modal matrix thus constructed will not diagonalize the operator, but it will produce a similar form known as the *Jordan canonical form*. Canonical forms, such as the diagonal form, which is a special case of a Jordan form, are simply standardized matrix structures that have particularly convenient forms for different purposes. We will see the usefulness of canonical forms in future chapters.

We begin discussing generalized eigenvectors by recalling that if $\boldsymbol{x}_1$ is a *regular* eigenvector corresponding to eigenvalue λ_1, then $A\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$. If for this eigenvalue $g_1 < m_1$ and if we can find a nontrivial solution $\boldsymbol{x}_2$ to the equation

$$A\boldsymbol{x}_2 = \lambda_1 \boldsymbol{x}_2 + \boldsymbol{x}_1 \tag{4.14}$$

that is linearly independent of $\boldsymbol{x}_1$, then $\boldsymbol{x}_2$ will be a generalized eigenvector. If $m_1 - g_1 = 1$, then $\boldsymbol{x}_2$ is the only generalized eigenvector necessary. If $m_1 - g_1 = p_1$, where $p_1 > 1$, then further generalized eigenvectors can be found from the "chain"

$$\begin{aligned} A\boldsymbol{x}_3 &= \lambda_1 \boldsymbol{x}_3 + \boldsymbol{x}_2, \\ A\boldsymbol{x}_4 &= \lambda_1 \boldsymbol{x}_4 + \boldsymbol{x}_3, \\ &\vdots \end{aligned}$$

until p_1 such vectors are found.

This definition for generalized eigenvectors therefore suggests one method for computing them, which is commonly called the *bottom-up* method:

1. For repeated eigenvalue λ_i with algebraic multiplicity m_i and geometric multiplicity g_i given by (4.13), first find g_i regular eigenvectors $\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{ig_i}$.
2. For each of these regular eigenvectors, compute nontrivial solutions to

$$(A - \lambda_i I)\boldsymbol{x}_{j+g_i} = \boldsymbol{x}_j \tag{4.15}$$

   for $j = i1, \ldots, ig_i$, which can be recognized as a rearranged version of Equation (4.14). If the computed $\boldsymbol{x}_{j+g_i}$ is linearly independent of all previously computed regular and generalized eigenvectors, then it is a generalized eigenvector.
3. To complete the set of $p_i = m_i - g_i$ generalized eigenvectors, continue to compute linearly independent solutions of

$$(A - \lambda_i I)\boldsymbol{x}_{j+(k+1)g_i} = \boldsymbol{x}_{j+kg_i},$$
$$\vdots$$

   for all $j = i1, \ldots, ig_1$ and $k = 1, 2, \ldots$ until all generalized eigenvectors are found.

This technique is illustrated in the following example.

Example 4.7: A Small Number of Generalized Eigenvectors

Determine all eigenvalues, regular eigenvectors, and generalized eigenvectors for the operator represented by the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ -3 & 3 & 5 \end{bmatrix}$$

Solution:

First computing the eigenvalues:

$$0 = |\lambda I - A| = \begin{vmatrix} \lambda - 1 & -2 & 0 \\ 0 & \lambda - 1 & 0 \\ 3 & -3 & \lambda - 5 \end{vmatrix}$$
$$= (\lambda - 5)(\lambda - 1)^2$$

so $\lambda_1 = 5$ and $\lambda_2 = 1\,(m_2 = 2)$. For eigenvalue $\lambda_1 = 5$, we will have only a single regular eigenvector, computed as $\boldsymbol{x}_1 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}}$ from the solution of the equation

$$(5 \cdot I - A)\boldsymbol{x}_1 = \begin{bmatrix} 4 & -2 & 0 \\ 0 & 4 & 0 \\ 3 & -3 & 0 \end{bmatrix} \boldsymbol{x}_1 = 0$$

For $\lambda_2 = 1\,(m_2 = 2)$, we find that

$$g_2 = q(1 \cdot I - A) = q\left(\begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 0 \\ 3 & -3 & -4 \end{bmatrix} \right) = 1$$

so that there will be a single regular eigenvector, calculated as $\boldsymbol{x}_2 = \begin{bmatrix} 4/5 & 0 & 3/5 \end{bmatrix}^{\mathrm{T}}$ from the equation

$$(1 \cdot I - A)\boldsymbol{x}_2 = 0$$

We therefore require $p_2 = m_2 - g_2 = 1$ generalized eigenvectors for λ_2. To compute it, we construct the linear algebraic equation

$$(A - 1 \cdot I)\boldsymbol{x}_3 = \boldsymbol{x}_2$$

or

$$\begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ -3 & 3 & 4 \end{bmatrix} \boldsymbol{x}_3 = \begin{bmatrix} 4/5 \\ 0 \\ 3/5 \end{bmatrix} \tag{4.16}$$

Equation (4.16) will of course have a homogeneous solution, which we may

discard, because it will be exactly the regular eigenvector $\boldsymbol{x}_2$. The particular solution, as obtained using the methods of Section 3.3.1, is $\boldsymbol{x}_3 = \begin{bmatrix} -3/5 & 2/5 & -3/5 \end{bmatrix}^{\mathrm{T}}$. This is linearly independent of $\boldsymbol{x}_2$ and is therefore the generalized eigenvector (the *only* one). Note that unlike regular eigenvectors, generalized eigenvectors are particular solutions to linear equations and therefore cannot be multiplied by an arbitrary constant. Nor are they unique.

Chains of Generalized Eigenvectors

It may be noticed that the procedure listed above has no stopping condition. That is, the search simply continues until all generalized eigenvectors are found, but no guidance is provided as to which indices *j* and *k* will provide these linearly independent vectors. A more procedural method can be devised if one considers the "chaining" of eigenvectors in more detail. To generate this algorithmic procedure, we first introduce the concept of the *index* of an eigenvalue.

> ***Index of Eigenvalue:*** The *index* of a repeated eigenvalue λ_i, denoted η_i, is the smallest integer η such that
>
> $$r(A-\lambda_i I)^{\eta} = n - m_i \tag{4.17}$$
>
> where n is the dimension of the space (size of the matrix), and m_i is the algebraic multiplicity of λ_i.

The rest of the algorithm is then calculated as follows:

1. For eigenvalue λ_i with index η_i, find all linearly independent solutions to the simultaneous set of matrix equations

$$\begin{aligned} (A-\lambda_i I)^{\eta_i}\boldsymbol{x} &= 0 \\ (A-\lambda_i I)^{\eta_i - 1}\boldsymbol{x} &\neq 0 \end{aligned} \tag{4.18}$$

   Each such solution will start a different "chain" of generalized eigenvectors. Because $r(A-\lambda_i I)^{\eta_i} = n - m_i$, we know there will be no more than $n-(n-m_i) = m_i$ such solutions, each of which is a generalized eigenvector. Denote these solutions $\boldsymbol{v}_1^1, \ldots, \boldsymbol{v}_{m_i}^1$. [Note that there might be fewer than m_i such vectors, because although there are exactly m_i solutions to $(A-\lambda_i I)^{\eta_i}\boldsymbol{x} = 0$, some of them might

not satisfy $(A-\lambda_i I)^{\eta_i-1}\boldsymbol{x}\neq 0$. For generality, we will assume there are m_i solutions to (4.18)].

2. Begin generating further generalized eigenvectors by computing the chain for each $j=1,\ldots,m_i$:

$$\begin{aligned}(A-\lambda_i I)\boldsymbol{v}_j^1 &= \boldsymbol{v}_j^2\\ (A-\lambda_i I)\boldsymbol{v}_j^2 &= \boldsymbol{v}_j^3\\ (A-\lambda_i I)\boldsymbol{v}_j^3 &= \boldsymbol{v}_j^4\\ &\vdots\end{aligned}$$

until we get to the point that

$$(A-\lambda_i I)\boldsymbol{v}_j^{\eta_i}=0$$

which of course indicates that $\boldsymbol{v}_j^{\eta_i}$ is a *regular* eigenvector. The chains thus end when regular eigenvectors are reached.

3. The *length* of the chains thus found will be η_i. There also may be chains of shorter length. If the chains of length η_i do not produce the full set of p_i generalized eigenvectors, begin the procedure again by finding all solutions to the set of equations

$$\begin{aligned}(A-\lambda_i I)^{\eta_i-1}\boldsymbol{x}&=0\\ (A-\lambda_i I)^{\eta_i-2}\boldsymbol{x}&\neq 0\end{aligned}\qquad(4.19)$$

and repeating the procedure. This will produce chains of length η_i-1. Continue until all generalized eigenvectors have been found.

4. Repeat for any other repeated eigenvalues.

The method outlined in these steps is usually referred to as the *top-down* method for finding eigenvectors. Although it is more computationally intensive and difficult to perform "by hand," it is easily programmed on a computer and has the added advantage that it can simultaneously find the regular eigenvectors as well. The previous method is known as the bottom-up method because it starts with known regular eigenvectors. The following examples illustrate the top-down method for simple matrices and more complex matrices.

Example 4.8: Top-Down Method for a Small Matrix

Use the top-down method to find the generalized eigenvector for the same 3×3 matrix as in Example 4.7.

Solution:

Recall that for this example, the eigenvalues were $\lambda_1 = 5$ and $\lambda_2 = 1\,(m_2 = 2)$, and the (regular) eigenvector corresponding to $\lambda_1 = 5$ was $\boldsymbol{x}_1 = \begin{bmatrix}0 & 0 & 1\end{bmatrix}^{\mathrm{T}}$. Now consider the repeated eigenvalue $\lambda_2 = 1\,(m_2 = 2)$. To find the index η_2, first note that $n - m_2 = 3 - 2 = 1$:

$$r\left((A - 1\cdot I)^1\right) = r\left(\begin{bmatrix}0 & 2 & 0\\ 0 & 0 & 0\\ -3 & 3 & 4\end{bmatrix}\right) = 2 \qquad (\neq n - m_2)$$

$$r\left((A - 1\cdot I)^2\right) = r\left(\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ -12 & 6 & 16\end{bmatrix}\right) = 1 \qquad (= n - m_2)$$

Therefore $\eta_2 = 2$.

So now we seek all solutions to the equations

$$\left(A - 1\cdot I\right)^2 \boldsymbol{x} = \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ -12 & 6 & 16\end{bmatrix}\boldsymbol{x} = 0$$

$$\left(A - 1\cdot I\right)^1 \boldsymbol{x} = \begin{bmatrix}0 & 2 & 0\\ 0 & 0 & 0\\ -3 & 3 & 4\end{bmatrix}\boldsymbol{x} \neq 0$$

To solve such a problem, we simply seek a vector $\boldsymbol{x} \in \mathrm{N}\left(\left(A - 1\cdot I\right)^2\right)$ such that $\boldsymbol{x} \notin \mathrm{N}\left(\left(A - 1\cdot I\right)^1\right)$. This is most easily done by computer. In this case, it is easily observed that the vector $\boldsymbol{v}_1^1 = \begin{bmatrix}1 & 2 & 0\end{bmatrix}^{\mathrm{T}}$ is one such solution. Now to generate the chain:

$$(A-1\cdot I)\boldsymbol{v}_1^1=\boldsymbol{v}_1^2=\begin{bmatrix}0&2&0\\0&0&0\\-3&3&4\end{bmatrix}\begin{bmatrix}1\\2\\0\end{bmatrix}=\begin{bmatrix}4\\0\\3\end{bmatrix}$$

$$(A-1\cdot I)\boldsymbol{v}_1^2=\boldsymbol{v}_1^3=\begin{bmatrix}0&2&0\\0&0&0\\-3&3&4\end{bmatrix}\begin{bmatrix}4\\0\\3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

So clearly the chain ends with the vector $\boldsymbol{v}_1^2=\begin{bmatrix}4&0&3\end{bmatrix}^{\mathrm{T}}$, which will be the *regular* eigenvector corresponding to $\lambda_2=1$. This eigenvector, which we prefer to denote as $\boldsymbol{x}_2$ after normalization, will appear exactly the same as eigenvector $\boldsymbol{x}_2$ in Equation (4.16) of Example 4.7. The generalized eigenvector chained to this regular eigenvector is $\boldsymbol{v}_1^1=\begin{bmatrix}1&2&0\end{bmatrix}^{\mathrm{T}}$. Note that this generalized eigenvector is *not* the same as the generalized eigenvector computed in the bottom-up example. Generalized eigenvectors are not unique.

Example 4.9: Multiple Chains of Generalized Eigenvectors

Find all eigenvalues, eigenvectors, and generalized eigenvectors for the operator represented by the matrix:

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

Solution:

The first thing one should notice about this example is that the last two rows are identical. The matrix, therefore, will be rank-deficient and must have at least one zero eigenvalue, because $|A|=|0\cdot I-A|=0$. (In general, the A matrix itself will have rank deficiency equal to the geometric multiplicity of any zero eigenvalues.) However, to determine all the eigenvalues, we find all the zeros to the polynomial

$$|\lambda I-A|=(\lambda-2)^5\lambda$$

which provide eigenvalues of $\lambda_1 = 0$ and $\lambda_2 = 2$ $(m_2 = 5)$. For $\lambda_1 = 0$, we can compute the corresponding eigenvector by finding a vector in the (one-dimensional) null space of $A - 0 \cdot I$ $(= A)$. This will give $\boldsymbol{x}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}^{\mathrm{T}}$ (unnormalized).

For $\lambda_2 = 2$ $(m_2 = 5)$, we begin by computing the value $n - m_2 = 6 - 5 = 1$. The index η_2 must be found from the sequence of calculations

$$\begin{aligned} r(A - 2 \cdot I) &= 4 \qquad (\neq n - m_2) \\ r\left((A - 2 \cdot I)^2\right) &= 2 \qquad (\neq n - m_2) \\ r\left((A - 2 \cdot I)^3\right) &= 1 \qquad (= n - m_2) \end{aligned}$$

so $\eta_2 = 3$. This implies that there will be a chain of generalized eigenvectors, of length three, ending with a regular eigenvector. Because $m_2 = 5$, we also know there must also be either a second chain of length two, or two more chains each of length one. To resolve this question, we note that because $r(A - 2 \cdot I) = 4$, the eigenspace of λ_2 will be of dimension $n - 4 = 2$, so there can only be two regular eigenvectors and, hence, two chains. We will therefore expect to find a second chain of length two.

Returning to the computation of the first chain, we must find a solution to the pair of equations

$$\begin{aligned} (A - 2 \cdot I)^3 \boldsymbol{x} &= 0 \\ (A - 2 \cdot I)^2 \boldsymbol{x} &\neq 0 \end{aligned} \tag{4.20}$$

where

$$(A - 2 \cdot I)^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -4 & 4 \\ 0 & 0 & 0 & 0 & 4 & -4 \end{bmatrix} \qquad (A - 2 \cdot I)^2 = \begin{bmatrix} 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 0 & -2 & 2 \end{bmatrix}$$

One such solution is the vector $\boldsymbol{v}_1^1 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$. Beginning the chain, we find

$$\begin{aligned} \boldsymbol{v}_1^2 &= (A - 2 \cdot I)\boldsymbol{v}_1^1 = \begin{bmatrix} 1 & -1 & 2 & -2 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \\ \boldsymbol{v}_1^3 &= (A - 2 \cdot I)\boldsymbol{v}_1^2 = \begin{bmatrix} 2 & 2 & 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \\ \boldsymbol{v}_1^4 &= (A - 2 \cdot I)\boldsymbol{v}_1^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \end{aligned}$$

Therefore, $\boldsymbol{v}_1^3$ is a regular eigenvector chained, in order, to generalized eigenvectors $\boldsymbol{v}_1^2$ and $\boldsymbol{v}_1^1$. The reader can verify that any other solutions to the system of equations (4.20) produce only vectors that are linearly dependent on the vectors that form this chain. However, one need not perform this check, because it is known that there can only be one chain of length three.

To find the chain of length two, we compute the solution to the equations

$$\begin{aligned} (A - 2 \cdot I)^2 \boldsymbol{x} &= 0 \\ (A - 2 \cdot I)^1 \boldsymbol{x} &\neq 0 \end{aligned} \tag{4.21}$$

one of which is $\boldsymbol{v}_2^1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$. This produces the chain

$$\begin{aligned} \boldsymbol{v}_2^2 &= (A - 2 \cdot I)\boldsymbol{v}_2^1 = \begin{bmatrix} 0 & 0 & 2 & -2 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \\ \boldsymbol{v}_2^3 &= (A - 2 \cdot I)\boldsymbol{v}_2^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \end{aligned}$$

so the second chain consists of regular eigenvector $\boldsymbol{v}_2^2$ and generalized eigenvector $\boldsymbol{v}_2^1$. It is very important to note that there are other solutions to Equation (4.21), but that all of them result in chains that consist of vectors that are linearly dependent on those previously found. It is necessary that the modal matrix we construct be nonsingular, so all regular and generalized eigenvectors must by linearly independent. We discard all the dependent solutions.

4.4.4 When n Independent Eigenvectors Exist

It is clear by now that the case of *n* independent eigenvectors makes construction of the modal matrix much simpler. There are certain matrices for which this will always be true. As we have already noted, the first is when all the eigenvalues of the matrix are distinct. By the construction of the null space of $\lambda I - A$, as in Equation (4.4), we know that each distinct eigenvalue will have associated with

it at least one eigenvector. When all *n* eigenvalues of an operator are distinct, we can also rely on the following result:

THEOREM: The eigenvectors of an operator *A*, all of whose eigenvalues are distinct, are linearly independent. (4.22)

PROOF: We will construct this proof by contradiction. Let matrix *A* have distinct eigenvalues, i.e., $\lambda_i \neq \lambda_j$ for all $i \neq j$. Suppose further that there are only *k* linearly independent eigenvectors among the complete set of *n*. Without loss of generality, let these *k* linearly independent eigenvectors be the first *k*. Then for all $k < j \leq n$, eigenvector $\boldsymbol{x}_j$ can be written as a linear combination of vectors in $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$. Write this linear combination as $\boldsymbol{x}_j = \sum_{i=1}^{k} a_{ij} \boldsymbol{x}_i$, where not all the coefficients a_{ij} are zero. Therefore,

$$A\boldsymbol{x}_j = \lambda_j \boldsymbol{x}_j = \sum_{i=1}^{k} a_{ij} \lambda_j \boldsymbol{x}_i \tag{4.23}$$

In addition,

$$A\boldsymbol{x}_j = A\left(\sum_{i=1}^{k} a_{ij} \boldsymbol{x}_i\right) = \sum_{i=1}^{k} a_{ij} A\boldsymbol{x}_i = \sum_{i=1}^{k} a_{ij} \lambda_i \boldsymbol{x}_i \tag{4.24}$$

Comparing (4.23) to (4.24), we have

$$\sum_{i=1}^{k} a_{ij} \lambda_j \boldsymbol{x}_i = \sum_{i=1}^{k} a_{ij} \lambda_i \boldsymbol{x}_i$$

giving,

$$\sum_{i=1}^{k} a_{ij} \left(\lambda_i - \lambda_j\right) \boldsymbol{x}_i = 0$$

Because this summation only runs up to *k*, i.e., through the independent $\boldsymbol{x}_i$'s, equality to zero implies that all coefficients $a_{ij}\left(\lambda_i - \lambda_j\right)$ are zero. However, because it is known that not

all a_{ij}'s are zero, then for some $i \neq j$, $\lambda_i = \lambda_j$. Because this contradicts our original assumption, we conclude that all the vectors $\boldsymbol{x}_i$, $i = 1, \ldots, n$ are linearly independent.

There is a second situation in which we can guarantee the existence of *n* independent eigenvectors, and that is when the matrix is hermitian (refer to Section 3.2.2). In this case, we first show, using the following theorem, that the eigenvectors corresponding to different eigenvalues are orthogonal, and that all eigenvalues are real.

THEOREM: If matrix *A* is hermitian, all its eigenvalues will be real, and it has a full set of *n* orthogonal eigenvectors, regardless of the existence of any repeated eigenvalues. (4.25)

PROOF: We first demonstrate that the eigenvalues for *A* are all real. We already know that each eigenvalue will have at least one eigenvector, so we can write $A\boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i$, or $A\boldsymbol{e}_i - \lambda_i \boldsymbol{e}_i = 0$. Then taking the inner product of each side of this equation with $\boldsymbol{e}_i$, we get the equality

$$\langle \boldsymbol{e}_i, A\boldsymbol{e}_i - \lambda_i \boldsymbol{e}_i \rangle = 0 \tag{4.26}$$

Then with the following manipulations:

$$\begin{aligned} 0 &= \langle \boldsymbol{e}_i, A\boldsymbol{e}_i \rangle - \langle \boldsymbol{e}_i, \lambda_i \boldsymbol{e}_i \rangle \\ &= \langle A\boldsymbol{e}_i, \boldsymbol{e}_i \rangle - \lambda_i \langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle \\ &= \langle \lambda_i \boldsymbol{e}_i, \boldsymbol{e}_i \rangle - \lambda_i \langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle \\ &= \overline{\lambda}_i \langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle - \lambda_i \langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle \\ &= \left(\overline{\lambda}_i - \lambda_i \right) \langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle \end{aligned} \tag{4.27}$$

where the transition from the first to the second line in this sequence is possible because of the self-adjointness of the operator *A*. Now because $\langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle > 0$, we must have $\overline{\lambda}_i - \lambda_i = 0$, which is only possible if λ_i is real.

Now take two different eigenvalues λ_i and λ_j. Consider the inner product $\langle \boldsymbol{e}_i, A\boldsymbol{e}_j \rangle$:

$$\begin{aligned}
\langle \boldsymbol{e}_i, A\boldsymbol{e}_j \rangle &= \langle \boldsymbol{e}_i, \lambda_j \boldsymbol{e}_j \rangle \\
\langle A\boldsymbol{e}_i, \boldsymbol{e}_j \rangle &= \lambda_j \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle \\
\langle \lambda_i \boldsymbol{e}_i, \boldsymbol{e}_j \rangle &= \lambda_j \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle \\
\overline{\lambda}_i \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle &= \lambda_j \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle \\
\lambda_i \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle &= \lambda_j \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle
\end{aligned}$$

implying $(\lambda_i - \lambda_j)\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = 0$. But we have stipulated that $\lambda_i \neq \lambda_j$, so $\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = 0$ and the eigenvectors are thus orthogonal.

Now we are prepared to show that hermitian matrices have *n* linearly independent eigenvectors.

THEOREM: Hermitian matrices have a complete set of *n* regular (i.e., not generalized) eigenvectors. (4.28)

PROOF: Suppose a hermitian (self-adjoint) matrix *A* has a (real) eigenvalue λ_i of arbitrary algebraic multiplicity. Suppose further that it has a generalized eigenvector, that is, a vector $\boldsymbol{x}_{i+1}$ chained to a regular eigenvector $\boldsymbol{x}_i$ such that $A\boldsymbol{x}_{i+1} - \lambda_i \boldsymbol{x}_{i+1} = \boldsymbol{x}_i$. Taking the inner product of both sides of this equality with $\boldsymbol{x}_i$,

$$\begin{aligned}
\langle A\boldsymbol{x}_{i+1} - \lambda_i \boldsymbol{x}_{i+1}, \boldsymbol{x}_i \rangle &= \langle \boldsymbol{x}_i, \boldsymbol{x}_i \rangle \\
&= \langle A\boldsymbol{x}_{i+1}, \boldsymbol{x}_i \rangle - \langle \lambda_i \boldsymbol{x}_{i+1}, \boldsymbol{x}_i \rangle \\
&= \langle \boldsymbol{x}_{i+1}, A\boldsymbol{x}_i \rangle - \lambda_i \langle \boldsymbol{x}_{i+1}, \boldsymbol{x}_i \rangle \\
&= \langle \boldsymbol{x}_{i+1}, \lambda_i \boldsymbol{x}_i \rangle - \lambda_i \langle \boldsymbol{x}_{i+1}, \boldsymbol{x}_i \rangle \\
&= \lambda_i \langle \boldsymbol{x}_{i+1}, \boldsymbol{x}_i \rangle - \lambda_i \langle \boldsymbol{x}_{i+1}, \boldsymbol{x}_i \rangle \\
&= 0
\end{aligned}$$

But if $\boldsymbol{x}_i$ is an eigenvector, we cannot have $\langle \boldsymbol{x}_i, \boldsymbol{x}_i \rangle = 0$. Therefore, we have shown by contradiction that there are no generalized eigenvectors. If there are no generalized eigenvectors, then there must be *n* regular eigenvectors.

Example 4.10: Moments of Inertia

Consider a system of point masses m_i, $i = 1, \cdots, n$, each located by a vector $r_i = \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}^{\mathrm{T}}$, that are fixed relative to one another. Such a system has an *inertia matrix* defined as

$$J = \begin{bmatrix} J_{xx} & J_{xy} & J_{xy} \\ J_{yx} & J_{yy} & J_{yz} \\ J_{zx} & J_{zy} & J_{zz} \end{bmatrix}$$

$$\triangleq \begin{bmatrix} \sum_{i=1}^{n} m_i \left(y_i^2 + z_i^2 \right) & -\sum_{i=1}^{n} m_i x_i y_i & -\sum_{i=1}^{n} m_i x_i z_i \\ -\sum_{i=1}^{n} m_i y_i x_i & \sum_{i=1}^{n} m_i \left(x_i^2 + z_i^2 \right) & -\sum_{i=1}^{n} m_i y_i z_i \\ -\sum_{i=1}^{n} m_i z_i x_i & -\sum_{i=1}^{n} m_i z_i y_i & \sum_{i=1}^{n} m_i \left(x_i^2 + y_i^2 \right) \end{bmatrix} \tag{4.29}$$

(A similar set of definitions applies to rigid bodies, with incremental masses and integrals used instead of point masses and summations.) With this definition and the angular velocity vector ω, the angular momentum of the system Ω can be written as $\Omega = J\omega$ and the kinetic energy can be written as $K = \frac{1}{2} \omega^{\mathrm{T}} J \omega$. Find a set of three coordinate axes for the system such that if the angular velocity is along one of those coordinate axes, the angular momentum will be along one of them as well.

Solution:

The problem statement implies that the angular velocity is to be proportional to the angular momentum, i.e., $\Omega = \lambda \omega$, where λ is a constant of proportionality. This in turn implies that $J\omega = \lambda \omega$, or again, the eigenvalue-eigenvector problem $(J - \lambda I)\omega = 0$. If we solve this problem to find the eigenvectors υ_i, $i = 1, 2, 3$, and use them as new basis vectors to diagonalize the system, we will obtain a new inertia matrix

$$\bar{J} = \begin{bmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{bmatrix} \tag{4.30}$$

where the quantities J_i, $i = 1, 2, 3,$ are known as the *principal moments of inertia*. The eigenvectors are called *principal axes of inertia*. (How can we be sure three such axes exist?) If we denote by ω_i, $i = 1, 2, 3,$ the components of the angular velocity about the three principal axes, then we will have

$$\Omega = \bar{J}\omega = \begin{bmatrix} J_1\omega_1 \\ J_2\omega_2 \\ J_3\omega_3 \end{bmatrix} \tag{4.31}$$

and

$$K = \tfrac{1}{2}\omega^{\mathrm{T}}\bar{J}\omega = \tfrac{1}{2}\left(J_1\omega_1^2 + J_1\omega_2^2 + J_1\omega_3^2\right) \tag{4.32}$$

The principal axes of an object will always be equal to any axes of symmetry of the body.

4.4.5 Jordan Canonical Forms

We have mentioned that the motivation for computing generalized eigenvectors is that we can use them in the modal matrix when we have an insufficient number of regular eigenvectors. We have seen that when the modal matrix, whose columns constitute a new basis of eigenvectors, contains only regular eigenvectors, then the matrix operator becomes a diagonal matrix of eigenvalues. This is because $A\boldsymbol{x} = \lambda\boldsymbol{x}$ for regular eigenvectors $\boldsymbol{x}$.

When generalized eigenvectors are included in the modal matrix, the resulting matrix operator in the new basis will not be diagonal because for generalized eigenvectors $\boldsymbol{x}_i$ chained to regular eigenvector $\boldsymbol{x}_\eta$, we have

$$\begin{aligned} A\boldsymbol{x}_\eta &= \lambda\boldsymbol{x}_\eta \\ A\boldsymbol{x}_{\eta-1} &= \lambda\boldsymbol{x}_{\eta-1} + \boldsymbol{x}_\eta \\ &\vdots \\ A\boldsymbol{x}_2 &= \lambda\boldsymbol{x}_2 + \boldsymbol{x}_3 \\ A\boldsymbol{x}_1 &= \lambda\boldsymbol{x}_1 + \boldsymbol{x}_2 \end{aligned}$$

From these relationships we define a *Jordan block* as a submatrix with the following structure:

$$\begin{matrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & & 0 \\ 0 & 0 & \lambda & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 & 0 \\ 0 & 0 & 0 & \ddots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{matrix}$$

The size of this block will be the length of the chain that defines it. This implies that the size of the largest Jordan block will be the size of the longest chain of eigenvectors, which has been defined as the index η of the eigenvalue. In a transformed *A*-matrix, there will be as many Jordan blocks corresponding to a particular eigenvalue as there are chains of eigenvectors corresponding to that eigenvalue, and their size will vary according to the lengths of the respective chains. Thus, there will be as many Jordan blocks for an eigenvalue as there are regular eigenvectors associated with that eigenvalue, which we know is the geometric multiplicity g_i. There also may be one or more Jordan blocks corresponding to different eigenvalues, and they will, of course, be calculated independently of the Jordan blocks for other eigenvalues. In the most trivial case, a matrix may have a complete set of *n* regular eigenvectors, so all its Jordan blocks, whether they belong to distinct eigenvalues or not, will be of size one, which gives the diagonal matrix in the new basis. This is why we say that the diagonal form is a special case of a Jordan form.

When transforming a matrix into its Jordan form[M], it is important to group regular eigenvectors and their chained generalized eigenvectors into the modal matrix in the reverse order in which the chain was produced using the top-down method. The regular eigenvector is first, followed sequentially by chained generalized eigenvectors. This will produce a coherent Jordan block in the transformed matrix. The blocks themselves can be arranged in the transformed matrix by rearranging the order in which the chains are placed as columns in the modal matrix. We will see in Chapters 6 and 8 the usefulness of Jordan forms. The following example illustrates these properties on the previously presented examples.

`jordan(A)`

Example 4.11: Jordan Forms for Some Previous Examples

Transform the *A*-matrices used in the previous two examples into their Jordan forms by constructing modal matrices with the regular and generalized eigenvectors.

$$A_1 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ -3 & 3 & 5 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 3 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Solution:

All of the eigenvalues, eigenvectors, and generalized eigenvectors have already been computed for these two examples. For matrix A_1, we construct the modal matrix with these vectors:

$$M_1 = \left[\boldsymbol{x}_1 \;\vdots\; \boldsymbol{v}_1^2 \;\vdots\; \boldsymbol{v}_1^1 \right] = \begin{bmatrix} 0 & 4 & 1 \\ 0 & 0 & 2 \\ 1 & 3 & 0 \end{bmatrix}$$

$$= \left[\begin{array}{c|c|c} \text{regular} & \text{regular} & \text{generalized} \\ \text{eigenvector} & \text{eigenvector} & \text{eigenvector} \end{array} \right]$$

then

$$\hat{A}_1 = M_1^{-1} A_1 M_1 = \left[\begin{array}{c|cc} 5 & 0 & 0 \\ \hline 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right]$$

For matrix A_2, the following (sequentially ordered) chains of eigenvectors and generalized eigenvectors were found:

$$\underset{(\lambda_1=0)}{\boldsymbol{x}_1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \qquad \{\boldsymbol{v}_1^3, \boldsymbol{v}_1^2, \boldsymbol{v}_1^1\} = \left[\begin{array}{rr|r} 2 & 1 & 0 \\ 2 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{array} \right] \qquad \{\boldsymbol{v}_2^2, \boldsymbol{v}_2^1\} = \left[\begin{array}{r|r} 0 & 0 \\ 0 & 0 \\ 2 & 0 \\ -2 & 0 \\ 0 & 1 \\ 0 & 1 \end{array} \right]$$

therefore if we form the modal matrix

$$M_2 = \begin{bmatrix} 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 2 & 0 \\ 0 & 0 & -2 & 1 & -2 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

we get

$$\hat{A}_2 = M_2^{-1} A_2 M_2 = \left[\begin{array}{c:ccc:cc} 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

which is seen to be a Jordan canonical form with a single 1×1 Jordan block corresponding to $\lambda_1 = 0$, and two Jordan blocks, of sizes 3×3 and 2×2, which is the order in which the eigenvector chains were arranged in modal matrix M_2.

4.5 Singular Values

Eigenvalues and eigenvectors of matrices have important application in the stability and control of systems of differential equations, as we will see in Chapters 7 and 10. For many of our purposes, the eigenvalues and eigenvectors sufficiently characterize the numerical properties of our operators. However, they are defined only for square matrices that represent operators $\Re^n \to \Re^n$. They cannot be computed for nonsquare matrices.

In this section we will introduce the concept of *singular values* and *singular value decompositions* (SVD). These tools will apply with equal validity for square and nonsquare matrices alike. They have many purposes, both directly, as in modern methods for robust controller design, and indirectly, as the underlying numerical algorithm in such everyday computations as rank determination.

> ***Singular Value Decomposition:*** Let A be an arbitrary $m \times n$ matrix. There exist two orthonormal matrices, U and V, such that

$$A = U\Sigma V^{\mathrm{T}} \tag{4.33}$$

where

$$\Sigma = \left[\begin{array}{c|c} S & 0 \\ \hline 0 & 0 \end{array}\right] \tag{4.34}$$

and S is the diagonal matrix, $S = diag\{\sigma_1, \sigma_2, \ldots, \sigma_r\}$. The r values $\sigma_1, \sigma_2, \ldots, \sigma_r$ are arranged such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. Together with the values $\sigma_{r+1} = \cdots = \sigma_n = 0$ they are known as *singular values* for matrix A. The factorization (4.33) is the *singular value decomposition*[M] of A.

`svd(A)`

4.5.1 Properties of Singular Value Decompositions

The first thing to note about an SVD is that the value r, i.e., the number of nonzero singular values, is equal to the rank of matrix A. A matrix therefore has as many zero singular values as its rank deficiency. These singular values are also the positive square roots of the eigenvalues of $A^{\mathrm{T}}A$. The columns of matrix U are the *left singular vectors*, are orthonormal, and are the eigenvectors of AA^{T}. The columns of V are the *right singular vectors*, are also orthonormal, and are the eigenvectors of $A^{\mathrm{T}}A$. However, while all these properties might suggest an easy way to compute an SVD, this is *not* a numerically accurate way to do it. In practice, there are several numerically robust methods for iteratively computing the singular value decomposition (see [6], [9], and [12]). We will not be further concerned with the actual computation of SVDs, but rather with their application in linear systems problems.

Example 4.12: Singular Value Decomposition

For the rank-2 matrix, give an SVD.

$$A = \begin{bmatrix} 4 & 8 & -2 & 0 \\ 3 & 3 & -3 & 1 \\ 1 & 5 & 1 & -1 \end{bmatrix}$$

Solution:

As mentioned, SVDs are never actually computed by hand, but rather with some numerically accurate computer programs. Therefore, we will give here, for purposes of illustrating such a decomposition, only the result:

$$
\begin{aligned}
A &= U\Sigma V^{\mathrm{T}} \\
&= \begin{bmatrix} 0.8165 & 0 & -0.5774 \\ 0.4082 & -0.7071 & 0.5774 \\ 0.4082 & 0.7071 & 0.5774 \end{bmatrix} \begin{bmatrix} 11.225 & 0 & 0 & 0 \\ 0 & 3.7417 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \\
&\qquad \begin{bmatrix} 0.4362 & -0.378 & 0.8138 & -0.066 \\ 0.8729 & 0.378 & -0.2825 & 0.1243 \\ -0.2182 & 0.7559 & 0.4977 & 0.365 \\ 0 & -0.378 & -0.1009 & 0.9203 \end{bmatrix}^{\mathrm{T}}
\end{aligned}
$$

Notice that there are indeed two nonzero singular values, according to the rank of *A*.

4.5.2 Some Applications of Singular Values

Listed in this section are some common uses for singular values and SVDs [6]. It is assumed that a good computer program is available for computing the SVD.

Norm Computations

In Chapter 3 it was noted that although the euclidean, or 2-norm, of a matrix operator was relatively easy to define, it was not at that time clear how such a norm was actually computed. Generally, 2-norms are computed with the aid of the singular values:

$$\|A\|_2 = \sigma_1 \tag{4.35}$$

which is the largest singular value. This can actually be readily shown by considering Equation (3.36):

$$
\begin{aligned}
\|A\|_2 &= \sup_{\boldsymbol{x}\neq 0} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \\
&= \sup_{\boldsymbol{x}\neq 0} \frac{\langle A\boldsymbol{x}, A\boldsymbol{x}\rangle^{1/2}}{\langle \boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}} = \sup_{\boldsymbol{x}\neq 0} \frac{\langle \boldsymbol{x}, A^T A\boldsymbol{x}\rangle^{1/2}}{\langle \boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}}
\end{aligned} \tag{4.36}
$$

for real matrix A. Now because $A^{\mathrm{T}}A$ is obviously symmetric, we can consider using an orthonormal set $\{\boldsymbol{e}_i\}$ of its eigenvectors as a basis in which to expand vector $\boldsymbol{x}$, $\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i$, where each eigenvector $\boldsymbol{e}_i$ corresponds to an eigenvalue σ_i^2 of $A^{\mathrm{T}}A$ (notice the notation for these eigenvalues). Doing this gives

$$
\begin{aligned}
\|A\|_2 &= \sup_{\boldsymbol{x}\neq 0} \frac{\left\langle \sum_{j=1}^{n} x_j \boldsymbol{e}_j, A^{\mathrm{T}}A \sum_{i=1}^{n} x_i \boldsymbol{e}_i \right\rangle^{1/2}}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2}} \\
&= \sup_{\boldsymbol{x}\neq 0} \frac{\left\langle \sum_{j=1}^{n} x_j \boldsymbol{e}_j, \sum_{i=1}^{n} x_i A^{\mathrm{T}} A \boldsymbol{e}_i \right\rangle^{1/2}}{\left[\sum_{i=1}^{n} x_i^2 \right]^{1/2}} \\
&= \sup_{\boldsymbol{x}\neq 0} \frac{\left\langle \sum_{j=1}^{n} x_j \boldsymbol{e}_j, \sum_{i=1}^{n} x_i \sigma_i^2 \boldsymbol{e}_i \right\rangle^{1/2}}{\left[\sum_{i=1}^{n} x_i^2 \right]^{1/2}} \\
&= \sup_{\boldsymbol{x}\neq 0} \frac{\left[\sum_{i=1}^{n} x_i \sigma_i^2 \left\langle \sum_{j=1}^{n} x_j \boldsymbol{e}_j, \boldsymbol{e}_i \right\rangle \right]^{1/2}}{\left[\sum_{i=1}^{n} x_i^2 \right]^{1/2}} \\
&= \sup_{\boldsymbol{x}\neq 0} \frac{\left[\sum_{i=1}^{n} x_i^2 \sigma_i^2 \right]^{1/2}}{\left[\sum_{i=1}^{n} x_i^2 \right]^{1/2}} \\
&= \sigma_1
\end{aligned}
$$

where σ_1 is the largest singular value of A. One can verify that the supremum (i.e., the least upper bound) achieved in the final line of this development occurs

when $\boldsymbol{x} = \boldsymbol{e}_1$, i.e., the right singular vector of A (the eigenvector of $A^{\mathrm{T}}A$) corresponding to singular value σ_1^2.

If A is a square invertible matrix (full rank), then we also have

$$\left\|A^{-1}\right\|_2 = \frac{1}{\sigma_n} \tag{4.37}$$

The Frobenius norm can also be computed using singular values:

$$\begin{aligned} \|A\|_F &= \left[\sigma_1^2 + \cdots + \sigma_n^2\right]^{1/2} \\ &= \left[\sigma_1^2 + \cdots + \sigma_r^2\right]^{1/2} \end{aligned} \tag{4.38}$$

Rank Determination and Condition Number

The SVD is by far the most common method for determining the rank of a matrix. As previously stated, the rank of a matrix is the number of nonzero singular values for that matrix.

Determinating the condition[M] number of a matrix is a related function. The *condition number* is a relative measure of how close a matrix is to rank deficiency and can be interpreted as a measure of how much numerical error is likely to be introduced by computations involving that matrix. For a square matrix, the condition number is a measure of how close the matrix is to being singular. The condition number is defined as $cond(A) = \sigma_1/\sigma_n$, i.e., the largest singular value over the smallest. By this definition, it is clear that a singular matrix, which has at least one zero singular value, will have an infinite or undefined condition number. A matrix with a large condition number is said to be *ill-conditioned*, and a matrix with a small condition number (which is desirable) is said to be *well-conditioned*.

`cond(A)`

When matrix A is used in a system of linear simultaneous equations, $A\boldsymbol{x} = \boldsymbol{y}$, the condition number is an approximation for the amount by which errors in A or $\boldsymbol{y}$ might be amplified in the solution for $\boldsymbol{x}$.

Matrix Inverses

If A is a square matrix, then of course U and V will also be square matrices. In such a case, SVDs are useful in computing the inverse of nonsingular matrix A. Because the inverse of an orthonormal matrix is equal to its transpose, we have

$$\begin{aligned} A^{-1} &= \left(U\Sigma V^{\mathrm{T}}\right)^{-1} = V\Sigma^{-1}U^{\mathrm{T}} \\ &= V \cdot diag(1/\sigma_1, \ldots, 1/\sigma_n) \cdot U^{\mathrm{T}} \end{aligned} \tag{4.39}$$

From this expression it can be seen how the "size" of the inverse depends on the singular values and why (4.37) is true.

Computing Subspaces

By partitioning the SVD into blocks corresponding to the location of the zero and nonzero singular values, we can exploit some special properties of the matrices *U* and *V*:

$$A = U\Sigma V^{\mathrm{T}} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \end{bmatrix} \tag{4.40}$$

where the matrices are partitioned compatibly. It can be shown that the columns of matrix U_1 form an orthonormal basis of the range space of *A*, and the columns of V_2 form an orthonormal basis for the null space of *A*.

Projection Operators

Using the same partitions for the *U* and *V* matrices as above, the SVD is an easy way to compute projection operators. For the projection operator that projects vectors onto the range space of *A*, we may use:

$$P_{\mathrm{R}(A)} = U_1 U_1^{\mathrm{T}} = AA^{+} \tag{4.41}$$

and for the projection onto the null space of *A*, we may use

$$P_{\mathrm{N}(A)} = V_2 V_2^{\mathrm{T}} = I - A^{+}A \tag{4.42}$$

Solving Simultaneous Equations

Consider the equation $A\boldsymbol{x} = \boldsymbol{y}$, where $\boldsymbol{x} \in \Re^n$, $\boldsymbol{y} \in \Re^m$, and $A: \Re^n \rightarrow \Re^m$. As we have noted above, if *A* is nonsingular, then a solution to this system in terms of the SVD can be given as

$$\boldsymbol{x} = V \cdot \Sigma^{+} \cdot U^{\mathrm{T}} \cdot \boldsymbol{y} \tag{4.43}$$

where

$$\Sigma^{+} = \left[\begin{array}{c|c} S^{-1} & 0 \\ \hline 0 & 0 \end{array}\right]$$

and the zero blocks are sized appropriately according to the dimensions of U and V.

This formula also applies in the two cases considered in the previous chapter, i.e., the overdetermined and underdetermined cases, regardless of whether a unique solution exists, an infinite number of solutions exist, or no exact solutions exist. If there are zero singular values, we replace the corresponding

$$\frac{1}{\sigma_i} = \frac{1}{0}$$

terms by *zero*, and use exactly the same formula, i.e., (4.43). Equation (4.43) therefore gives the pseudoinverse solution to such systems, automatically implementing the left- or right-pseudoinverse as appropriate. In the event that no exact solution exists, (4.43) will give the solution of least-squared error. In the event that multiple solutions exist, (4.43) returns the solution with minimum norm. Adding to $\boldsymbol{x}$ any linear combination of the columns of V_2 will then be adding homogeneous terms from the null space of A.

Example 4.13: Singular Value Decompositions in Robotics

In this extended example, we will illustrate the use of SVDs in two aspects of a simple robotics problem, both involving the inverse kinematics of a simple robot. Another example will be given in Section 5.3.3. The problem concerns the determination of joint angles for a three-link robot. The reader need not have any previous exposure to robotics.

The robotic *forward* kinematics problem is simply the problem of determining the position and orientation of the tip of the robot, given the joint angles. The *inverse* kinematics problem is the problem of determining the joint angles for a given tip position. In our problem, we wish to determine the vector of angular joint velocities

$$\dot{\Theta} = \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix}$$

such that the tip of the third link travels with a prescribed velocity in a two-dimensional coordinate system *XY* (see Figure 4.3).

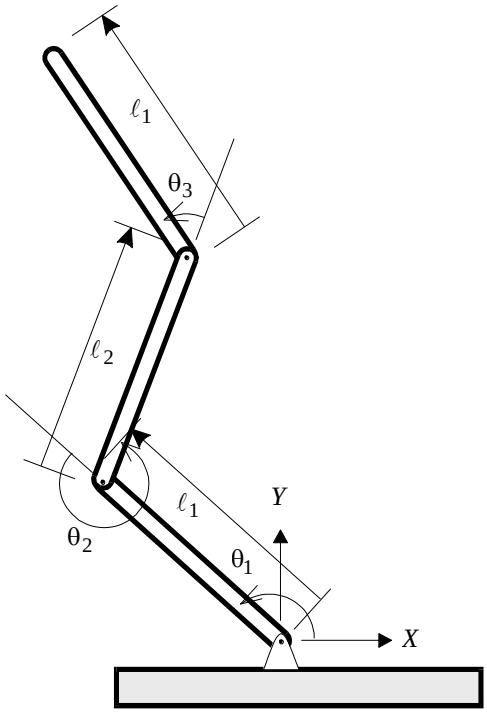

Figure 4.3 A three-link robot arm in a plane. Coordinates of the tip of the third link are measured in the Cartesian system XY, while individual joint angles are measured positively counterclockwise from the axis of the preceding link, except for the first link, which is measured from the horizontal.

The forward kinematic equations for this robot arm can be written from basic trigonometry as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} l_1 \cos(\theta_1) + l_2 \cos(\theta_1 + \theta_2) + l_3 \cos(\theta_1 + \theta_2 + \theta_3) \\ l_1 \sin(\theta_1) + l_2 \sin(\theta_1 + \theta_2) + l_3 \sin(\theta_1 + \theta_2 + \theta_3) \end{bmatrix} \tag{4.44}$$
$$\triangleq \boldsymbol{f}(\Theta) = \begin{bmatrix} f_1(\Theta) \\ f_2(\Theta) \end{bmatrix}$$

This vector function $\boldsymbol{f} \in \Re^2$ of the vector of joint angles $\Theta \in \Re^3$ specifies the two coordinates of the tip of link three in terms of the three joint angles.

Often it is the roboticist's job to specify the three motor velocities

$$\dot{\Theta} = \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix}$$

necessary to produce a particular desired velocity

$$\dot{\mathrm{X}}_{desired} = \begin{bmatrix} \dot{x}_{desired} \\ \dot{y}_{desired} \end{bmatrix}$$

of the tip of the robot. This represents the inverse velocity kinematics problem, which can be simplified with some new notation:

$$\begin{aligned} \dot{\mathrm{X}}_{desired} &= \frac{d}{dt}\boldsymbol{f}(\Theta) = \frac{\partial \boldsymbol{f}(\Theta)}{\partial \Theta}\frac{d\Theta}{dt} \\ &\triangleq J(\Theta)\dot{\Theta} \end{aligned} \tag{4.45}$$

where $J(\Theta) \triangleq \partial \boldsymbol{f}(\Theta)/\partial\Theta$ is known as the *jacobian* of the nonlinear transformation $\boldsymbol{f}$. It is, in this case, a 2×3 matrix that depends nonlinearly on the joint angles Θ. Using the formula given in Appendix A for the derivative of a vector with respect to another vector, we can obtain for this example:

$$J(\Theta) = \begin{bmatrix} -l_1 S_1 - l_2 S_{12} - l_3 S_{123} & -l_2 S_{12} - l_3 S_{123} & -l_3 S_{123} \\ l_1 C_1 + l_2 C_{12} + l_3 C_{123} & l_2 C_{12} + l_3 C_{123} & l_3 C_{123} \end{bmatrix} \tag{4.46}$$

where $C_1 \triangleq \cos(\theta_1)$, $S_{12} \triangleq \sin(\theta_1 + \theta_2)$, etc. If at any given time we can accurately measure the angles Θ and thus compute $J(\Theta)$, then (4.45) can be considered a linear system of two equations and three unknowns.

Of course, if the jacobian $J(\Theta)$ were square and nonsingular, it could be inverted, and hence, (4.45) could be solved to determine the necessary joint velocities $\dot{\Theta}$. However, in this case the jacobian is nonsquare. Hence, a pseudoinverse solution is possible.

$$\dot{\Theta} = J^{+}(\Theta)\dot{\mathrm{X}}_{desired} \tag{4.47}$$

or, as computed using the singular value decomposition,

$$\dot{\Theta} = V \cdot \Sigma^{+} \cdot U^{\mathrm{T}} \cdot \dot{\mathrm{X}}_{desired} \tag{4.48}$$

where $J = U\Sigma V^{\mathrm{T}}$.

Note that, according to the discussion of this pseudoinverse in Chapter 3, this solution, in the case of multiple possible solutions to underdetermined systems, minimizes a secondary criterion, namely the squared norm of the vector $\dot{\Theta}$, or, equivalently, $\frac{1}{2}\dot{\Theta}^{\mathrm{T}}\dot{\Theta}$. This has the effect of minimizing the vector of joint velocities for a given endpoint velocity.

When the jacobian of such a system is rank deficient, a particular physical phenomenon occurs. This is known as kinematic *singularity*[*] and results in a robot that, in that joint angle configuration Θ, is unable to produce endpoint velocities in one or more directions of its workspace. For example, as seen in Figure 4.4, such a robot whose joints are all aligned is unable to instantaneously produce velocities along the axis of its aligned links.

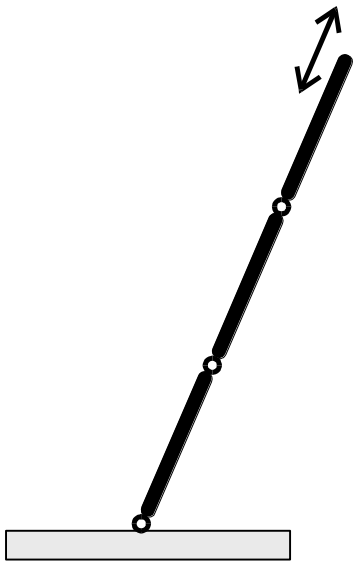

Figure 4.4 A three-link planar robot in a singular configuration. It is unable to instantaneously produce velocity in the direction of the arrow.

However, the concept of singularity can also be considered a matter of degree, i.e., in some configurations a robot may be closer to singularity than in other configurations. This can be displayed graphically with what are known as

[*] Note that the usage of this term is consistent with our notion of matrix singularity. When the nonsquare jacobian is rank-deficient, the pseudoinverse will not exist because the matrix JJ^{T} will be square and rank-deficient and, thus, singular.

manipulability ellipses in the robotics field. A manipulability ellipse is an ellipse, usually drawn with its center at the tip of the robot, whose axes depict the ability of the robot to produce velocity in the directions of those axes, with a unit joint angle velocity. For example, the three configurations in Figure 4.5 are shown with approximate manipulability ellipses attached.

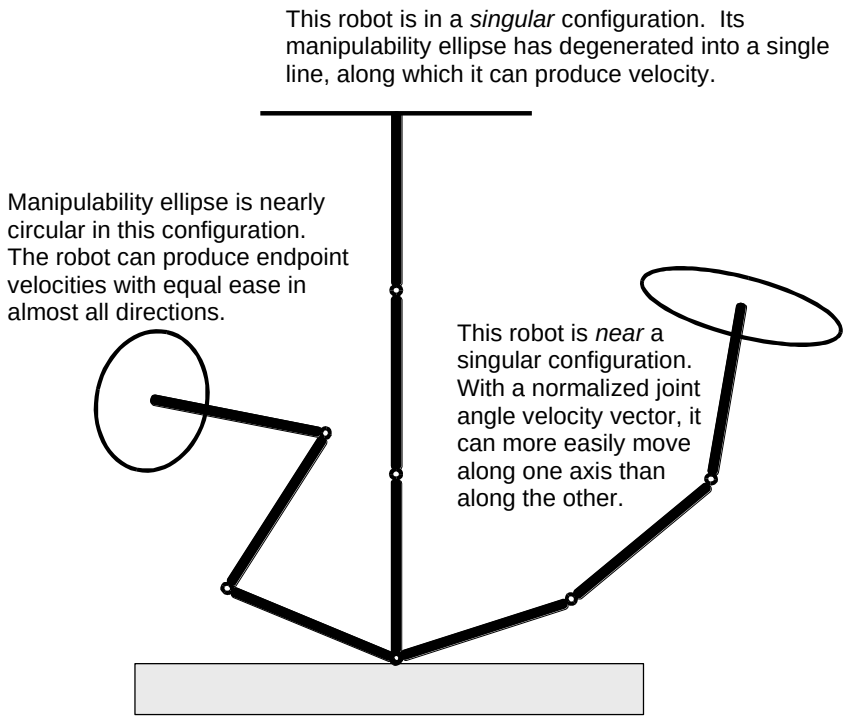

Figure 4.5 Three configurations of a robot with different degrees of manipulability.

The manipulability ellipses may be drawn by computing the SVD of the jacobian. The axis lengths of the ellipses are equal to the singular values of J, and the directions of those axes are equal to the left singular vectors (columns of U). These dimensions will be further illustrated in Example 5.1. In practice, the pseudoinverse solution itself may be augmented by a term that is proportional to the direction of the gradient of a manipulability measure, $M(\Theta)$, e.g.,

$$M(\Theta) = \sqrt{\left| JJ^{\mathrm{T}} \right|} = \prod_{i=1}^{n} \sigma_i \tag{4.49}$$

where the σ_i's are the singular values of *J*. This has the effect of maximizing the manipulator's distance from the nearest singularity. However, such a term must be first projected onto the null space of the jacobian, lest it prevent the robot from reaching the desired velocity $\dot{\mathrm{X}}_{desired}$. That is, one can use

$$\begin{aligned}\dot{\Theta} &= J^{+}(\Theta)\dot{\mathrm{X}}_{desired} + P_{\mathrm{N}(J)}\nabla M(\Theta) \\ &= J^{+}(\Theta)\dot{\mathrm{X}}_{desired} + \left[I - J^{+}(\Theta)J(\Theta)\right]\nabla M(\Theta)\end{aligned} \tag{4.50}$$

which, when multiplied by the jacobian, returns

$$\begin{aligned}J(\Theta)\dot{\Theta} &= J(\Theta)J^{+}(\Theta)\dot{\mathrm{X}}_{desired} + \left[J(\Theta) - J(\Theta)J^{+}(\Theta)J(\Theta)\right]\nabla M(\Theta) \\ &= \dot{\mathrm{X}}_{desired}\end{aligned}$$

demonstrating that the homogeneous term does not affect $\dot{\mathrm{X}}_{desired}$. Referring to (4.43) and (4.42), Equation (4.50) can be computed using the SVD:

$$\dot{\Theta} = V \cdot \Sigma^{+} \cdot U^{\mathrm{T}} \cdot \dot{\mathrm{X}}_{desired} + V_2 V_2^{\mathrm{T}} \nabla M(\Theta) \tag{4.51}$$

with the SVD of *J* partitioned as in (4.40).

4.6 Summary

In this chapter, we have presented new material concerning eigenvalues, eigenvectors, and singular values. Again we stress the geometric interpretation of these concepts: that eigenvectors are *A*-invariant subspaces. The convenience of knowing these invariant subspaces will become apparent in Chapter 6. Furthermore, the notion of a singular value is introduced, partly as a counterpart to eigenvalues and eigenvectors for nonsquare matrices, but mostly because they are finding increasing use in modern control theory, robust and adaptive control, and signal processing. The fact that the largest singular value provides the norm of a matrix is equally valid in time- and frequency-domain, and this has recently led to significant developments in control theory. In particular, in the field of optimal control, wherein a cost criterion is to be minimized by a stabilizing control input, the use of the singular value has provided a means to minimize frequency-domain criteria, whereas previously, optimal control had been primarily performed in time-domain.

Other important developments in this chapter are:

- Eigenvectors and eigenvalues, like other concepts introduced in previous chapters, are not to be considered properties of *matrices* but rather of

operators. This brings us back to the notion of invariance, i.e., a subspace on which an operator has only a scaling effect.

- We investigated the basis of eigenvectors as a change-of-basis matrix. This results in Jordan forms in general, and diagonal forms in special cases. The importance of these forms is partly illustrated in Exercise 4.14, where it is shown how diagonalization aids the solution of differential systems. This will be further emphasized in Chapter 6.
- The concept of generalized eigenvectors was discussed in some detail. The use of generalized eigenvectors is limited except for computing Jordan forms, because they are not invariant subspaces, as are regular eigenvectors.
- As we have mentioned, the singular value, briefly introduced here, is indispensible for many matrix computations and numerical methods. Algorithms for computing them are numerically stable and are the basis for many of the other computations performed by computer-aided engineering software packages. Quantities such as rank, condition, and the solutions to simultaneous equations are often computed via an underlying SVD.

Chapter 5 is the final chapter in Part I. It presents a mathematical treatment of other kinds of functions on matrices, i.e., not necessarily linear operators. These functions will be required when we solve state space systems, analyze their stability, and develop control and estimation methods for them.

4.7 Problems

4.1 Let T be an operator that rotates vectors in $\Re^3$ by an angle 45° counterclockwise about the vector $\boldsymbol{x} = [1 \quad 0 \quad 1]^{\mathrm{T}}$ and magnifies them by a factor of four. Determine all real-valued eigenvectors and their corresponding eigenvalues.

4.2 Find all eigenvalues and eigenvectors for the matrix

$$A = \begin{bmatrix} 7 & 2 & 2 & 0 \\ -2 & 2 & -1 & 0 \\ -7 & -4 & 0 & 1 \\ 2 & 1 & 1 & 3 \end{bmatrix}$$

4.3 Find the eigenvalues and corresponding eigenvectors of the following matrices:

a) $\begin{bmatrix} 1 & -2 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ -3 & -3 & -5 \end{bmatrix}$ c) $\begin{bmatrix} 0 & 1 \\ -\omega_n^2 & 0 \end{bmatrix}$

4.4 Prove that for matrix A with eigenvalues λ_i, $|A| = \prod_{i=1}^{n} \lambda_i$.

4.5 Find the eigenvalues, eigenvectors (real and/or generalized), and Jordan form for the matrices:

a) $\begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 0 & -4 \\ 0 & 3 & 0 \\ -2 & 0 & -1 \end{bmatrix}$ c) $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & -1 & 1 \end{bmatrix}$

4.6 Let **V** be the vector space of polynomials of degree less than or equal to two, with inner product

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \bar{\boldsymbol{f}}(0)\boldsymbol{g}(0) + \bar{\boldsymbol{f}}(1)\boldsymbol{g}(1) + \bar{\boldsymbol{f}}(2)\boldsymbol{g}(2)$$

Also let T be a linear operator defined by

$$(T\boldsymbol{p})(x) = \frac{d}{dx}\left(x\boldsymbol{p}(\boldsymbol{x})\right)$$

a) Find the matrix for the operator T in the basis $\{1, \boldsymbol{x}, \boldsymbol{x}^2\}$.

b) Find the eigenvalues and eigenvectors of T.

c) Determine whether or not T is self-adjoint, i.e., its adjoint is the same as itself. Why or why not?

4.7 Find the eigenvalues and eigenvectors of the operator

$$A = \begin{bmatrix} 24 & 18 & 9 \\ -28 & 26 & 4 \\ 56 & -28 & 4 \end{bmatrix}$$

Then determine coefficients a_i such that

$$\begin{bmatrix} 4 \\ 4 \\ -2 \end{bmatrix} = \sum_{i=1}^{3} a_i \boldsymbol{e}_i$$

where the $\boldsymbol{e}_i$'s are eigenvectors of A.

4.8 Find a similarity transformation that reduces the following matrices to their Jordan forms. Identify the Jordan blocks.

$$A = \begin{bmatrix} 7 & -11 & 30 \\ -14 & 13 & -43 \\ -8 & 10 & 29 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}$$

4.9 Find the eigenvalues of the matrix

$$\begin{bmatrix} \alpha & \omega \\ -\omega & \alpha \end{bmatrix}$$

4.10 Let **V** be the linear space of polynomials of degree less than or equal to two, with the inner product defined as

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \int_{-1}^{1} \bar{\boldsymbol{f}}(x) \boldsymbol{g}(x)\, dx$$

Consider the linear operator

$$(T\boldsymbol{p})(x) = \tfrac{1}{2} x(1-x)\boldsymbol{p}''(x) + x\boldsymbol{p}'(x) + \boldsymbol{p}(x) + x^2 \boldsymbol{p}(0)$$

a) Find the matrix of T with respect to the basis $\{1, x, x^2\}$.

b) Find the eigenvalues and eigenvectors of T.

c) Determine whether T is hermitian.

4.11 Let **W** be the linear space of complex-valued functions, over the complex scalars, that are integrable on the interval $x \in [0,1]$. Let operator T be defined by

$$(T\boldsymbol{f})(x) = \int_0^1 (x - 2t)\,\boldsymbol{f}(t)\,dt$$

a) Show that T is a linear operator.

b) Find the range space $\mathrm{R}(T)$.

c) Find all nonzero eigenvalues of T and their corresponding eigenvectors.

d) Find a polynomial $\boldsymbol{p}(t) \in \mathbf{W}$ of degree two such that $T\boldsymbol{p} = 0$.

4.12 Let **V** be the space of functions spanned by $\cos t$ and $\sin t$. An operator T is defined as

$$(T\boldsymbol{f})(t) = \int_0^{2\pi} \sin(t - x)\,\boldsymbol{f}(x)\,dx$$

a) Find the matrix of the operator T in the basis $\{\cos t, \sin t\}$.

b) Find the eigenvalues and eigenvectors of T.

4.13 Suppose two square matrices A and $\hat{A}$ are similar to one another; i.e., $\hat{A} = M^{-1}AM$ for some orthonormal matrix M. Show explicitly that the eigenvalues of $\hat{A}$ are the same as those of A. Do the two matrices have the same eigenvectors?

4.14 For the electrical circuit in Example 1.2, find a set of coordinates that transforms the describing equations into a pair of *decoupled* first-order differential equations, i.e., two first-order equations that can be solved independently of one another. Are these coordinates physically meaningful? (If necessary, use some nominal component values to simplify the algebra.)

4.15 Write MATLAB (or code in another language) to detect the presence of and find the generalized eigenvectors for an arbitrary matrix. Use it to find the Jordan form for the matrices given in Example 4.11.

4.8 References and Further Reading

Much of the material in this chapter is standard and can be found in any text on matrix computations, such as [5], including the control systems texts [1] and [2]. Explicit discussion of the computation of generalized eigenvectors can be difficult to find; ours is closely related to [1]. For further discussion of eigenvalues and eigenvectors for function spaces, see [3] and [8]. For more examples systems with *eigenfunctions*, see [8].

Our treatment of SVDs should be regarded as merely an overview. Good introductions to the theory of SVDs can be found in [6] and [12]. The use of SVDs for the solution of least-squares problems is found in [7]. Computation of the SVDs is discussed in some detail in [6], [12], and [13]. Their application in robotics, as we introduced by example, can be found in [9], [10], and [11] (for more information on robot kinematics itself, see [4] and [10]).

[1] Brogan, William L., *Modern Control Theory*, 3rd edition, Prentice-Hall, 1991.

[2] Chen, Tsi-Chong, *Linear System Theory and Design*, Holt, Rinehart, and Winston, 1984.

[3] Coddington, Earl A., and Norman Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, 1955.

[4] Fu, K. S., R. C. Gonzalez, and C. S. G. Lee, *Robotics: Control, Sensing, Vision, and Intelligence*, McGraw-Hill, 1987.

[5] Golub, Gene H., and Charles F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 1989.

[6] Klema, Virginia, and Alan J. Laub, "The Singular Value Decomposition: Its Computation and Some Applications", *IEEE Transactions on Automatic Control*, vol. AC-25, no. 2, April 1980, pp. 63-79.

[7] Lawson, Charles L., and Richard J. Hanson, *Solving Least Squares Problems*, Prentice-Hall, 1974.

[8] Lebedev, N. N., I. P. Skalskaya, and Y. S. Uflyand, *Worked Problems in Applied Mathematics*, Dover, 1965.

[9] Maciejewski, Anthony A., and Charles A. Klein, "The Singular Value Decomposition: Computation and Applications to Robotics," *International Journal of Robotics Research*, vol. 8, no. 6, December 1989, pp. 164-176.

[10] Nakamura, Yoshihiko, *Advanced Robotics: Redundancy and Optimization*, Addison-Wesley, 1991.

[11] Nakamura, Yoshihiko, and Yingti Xu, "Geometrical Fusion Method for Multi-Sensor Robotic Systems," *IEEE International Conference on Robotics and Automation*, 1989, pp. 668-673.

[12] Press, William H. Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, 1988.

[13] Ulrey, Renard R., Anthony A. Maciejewski, and Howard J. Siegel, "Parallel Algorithms for Singular Value Decomposition," *Proc. 8th International Parallel Processing Symposium*, April 1994, pp. 524-533.