# E

# Linear Algebra: Eigenproblems

# TABLE OF CONTENTS

## §E.1. Introduction

This Appendix focuses on the conventional algebraic eigenvalue problem. It summarily covers fundamental material of use in the main body of the book. It does not dwell on eigensolution methods for two reasons. First, those require specialized treatment beyond our intended scope. Second, efficient methods are implemented as black boxes in high level programming languages such as *Matlab*. Most properties and results will be stated without proof. Suggested textbooks and monographs devoted to the topic are listed under **Notes and Bibliography**.

## §E.2. The Standard Algebraic Eigenproblem

Consider the linear equation system (D.16), namely $\mathbf{A}\mathbf{x} = \mathbf{y}$, in which $\mathbf{A}$ is a square $n \times n$ matrix, while $\mathbf{x}$ and $\mathbf{y}$ are $n$-vectors. (Entries of $\mathbf{A}$, $\mathbf{x}$ and $\mathbf{y}$ may be real or complex numbers.) Suppose that the right-hand side vector $\mathbf{y}$ is required to be a multiple $\lambda$ of the solution vector $\mathbf{x}$:

$$\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x}, \tag{E.1}$$

or, written in full,

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & \lambda x_1 \\ a_{21}x_2 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & \lambda x_2 \\ \cdots & & \cdots & & \cdots & & \cdots & & \cdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & \lambda x_n. \end{array} \tag{E.2}$$

These equations are trivially verified for $\mathbf{x} = \mathbf{0}$. The interesting solutions, called *nontrivial*, are those for which $\mathbf{x} \neq \mathbf{0}$. Pairing a nontrivial solution with a corresponding $\lambda$ that satisfies (E.1) or (E.2) gives an *eigensolution pair*, or *eigenpair* for short.

Equation (E.1) or (E.2) is called an *algebraic eigensystem*. The determination of nontrivial eigenpairs $\{\lambda, \mathbf{x}\}$ is known as the standard (classical, conventional) *algebraic eigenproblem*. In the sequel, the qualifier *algebraic* will be usually omitted. Properties linked to an eigenproblem are collectively called *spectral properties* of $\mathbf{A}$, a term that comes from original applications to optics. The set of eigenvalues of $\mathbf{A}$ is called the *spectrum* of that matrix.

### §E.2.1. Characteristic Equation

The eigenystem (E.1) can be rearranged into the homogeneous form

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}. \tag{E.3}$$

A nontrivial solution $\mathbf{x} \neq \mathbf{0}$ of (E.3) is possible if and only if the coefficient matrix $\mathbf{A} - \lambda\mathbf{I}$ is singular. Such a condition can be expressed as the vanishing of the determinant

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{E.4}$$

When this determinant is expanded, we obtain an algebraic polynomial equation in $\lambda$ of degree $n$:

$$P(\lambda) = \lambda^n + \alpha_1\lambda^{n-1} + \cdots + \alpha_n = 0. \tag{E.5}$$

This is known as the *characteristic equation* of the matrix $\mathbf{A}$. The left-hand side is called the *characteristic polynomial*. From the Fundamental Theorem of Algebra we know that a $n^{th}$ degree polynomial has $n$ (possibly complex) roots $\lambda_1, \lambda_2, \ldots, \lambda_n$. Those $n$ solutions, generically denoted by $\lambda_i$ in which $i = 1, \ldots n$, are called the *eigenvalues*, *eigenroots* or *characteristic values*[1] of matrix $\mathbf{A}$. The characteristic polynomial of the transposed matrix is $\det(\mathbf{A}^T - \lambda\mathbf{I}) = \det(\mathbf{A} - \lambda\mathbf{I}) = P(\lambda)$. Consequently the eigenvalues of $\mathbf{A}$ and $\mathbf{A}^T$ are the same.

The coefficients of the characteristic polynomial of a real $\mathbf{A}$ are real. However, its roots (the eigenvalues) may be complex. For real $\mathbf{A}$ these will occur in complex conjugate pairs. If the matrix is symmetric real, however, the occurrence of real roots is guaranteed, as discussed in §E.3.

We will occasionally denote by $\lambda_1, \lambda_2, \ldots \lambda_p$, $p \leq n$, all the *distinct* eigenvalues of $\mathbf{A}$.

Eigenvalues may be ordered according to different conventions, which are stated in conjunction with specific examples. For future use, the diagonal matrix of eigenvalues will be denoted by $\mathbf{\Lambda} = \mathbf{diag}[\lambda_i]$. This will be called the *eigenvalue matrix*.

### §E.2.2. Eigenvectors

Associated to each eigenvalue $\lambda_i$ there is a nonzero vector $\mathbf{x}_i$ that satisfies (E.1) instantiated for $\lambda \to \lambda_i$:

$$\mathbf{A}\,\mathbf{x}_i = \lambda_i\,\mathbf{x}_i. \qquad \mathbf{x}_i \neq 0. \tag{E.6}$$

This $\mathbf{x}_i$ is called a *right eigenvector* or *right characteristic vector*. This is often abbreviated to just *eigenvector*, in which case the "right" qualifier is tacitly understood. If $\mathbf{A}$ is real and so is $\lambda_i$, $\mathbf{x}_i$ has real entries. But it $\lambda_i$ is complex, entries of $\mathbf{x}_i$ will generally be complex.

The *left eigenvectors* of $\mathbf{A}$ are the right eigenvectors of its transpose, and are denoted by $\mathbf{y}_i$:

$$\mathbf{A}^T\,\mathbf{y}_i = \lambda_i\,\mathbf{y}_i, \qquad \mathbf{y}_i \neq 0. \tag{E.7}$$

Note that the same index $i$ can be used to pair $\mathbf{x}_i$ and $\mathbf{y}_i$ since the eigenvalues of $\mathbf{A}$ and $\mathbf{A}^T$ coalesce. Transposing both sides of (E.7) the definition may be restated as $\mathbf{y}_i^T\,\mathbf{A} = \lambda_i\,\mathbf{y}_i^T$.

There is an inherent magnitude indeterminacy in (E.6) since $\beta\,\mathbf{x}_i$, in which $\beta$ is an arbitrary nonzero factor, is also a right eigenvector. Likewise for $\mathbf{y}_i$. Therefore eigenvectors are often *normalized* according to some convention. A popular one is unit dot product of left and right eigenvectors:

$$\mathbf{y}_i^T\,\mathbf{x}_i = \mathbf{x}_i^T\,\mathbf{y}_i = 1. \tag{E.8}$$

A set of left and right eigenvectors $\{\mathbf{y}_i\}$ and $\{\mathbf{x}_j\}$, where $i$, $j$ range over $1 \leq k \leq n$ integers, is said to be *biorthonormal* with respect to $\mathbf{A}$ if they verify, with $\delta_{ij}$ denoting the Kronecker delta:

$$\mathbf{y}_i^T\,\mathbf{x}_j = \delta_{ij}, \qquad \mathbf{y}_i^T\,\mathbf{A}\,\mathbf{x}_i = \lambda_i\,\delta_{ij}. \tag{E.9}$$

[1] The terminology *characteristic value* is that found in publications before 1950; it derives from the works of Cauchy in French. *Eigenvalue* is a German-English hybrid. In German (and Dutch) "eigen" is used in the sense of special, inherent, or peculiar. It has become a popular qualifier attached (as prefix) to several items pertaining to the topic. For example: eigenvector, eigenproblem, eigensystem, eigenroot, eigenspace, eigenfunction, eigenfrequency, etc. *Characteristic equation* and *characteristic polynomial* are exceptions. "Eigen" hits the ear better than the long-winded "characteristic."

### §E.2.3. Eigenpairs of Matrix Powers

Premultiplying (E.1) by $\mathbf{A}$ yields

$$\mathbf{A}^2\,\mathbf{x}_i = \lambda_i\,\mathbf{A}\,\mathbf{x}_i = \lambda_i^2\,\mathbf{x}_i. \tag{E.10}$$

Thus the eigenvalues of $\mathbf{A}^2$ are $\lambda_i^2$, whereas the eigenvectors do not change. Continuing the process, it is easily shown that the eigenvalues of $\mathbf{A}^k$, where $k$ is a positive integer, are $\lambda_i^k$.

To investigate negative exponents, assume that $\mathbf{A}$ is nonsingular so that $\mathbf{A}^{-1}$ exists. Premultiplying (E.1) by $\mathbf{A}^{-1}$ yields

$$\mathbf{I}\,\mathbf{x}_i = \lambda_i\,\mathbf{A}^{-1}\,\mathbf{x}_i, \quad \text{or} \quad \mathbf{A}^{-1}\,\mathbf{x}_i = \frac{1}{\lambda_i}\,\mathbf{x}_i. \tag{E.11}$$

Consequently the eigenvalues of the inverse are obtained by taking the reciprocals of the original eigenvalues, whereas eigenvectors remain unchanged. The generalization to arbitrary negative integer exponents is immediate: the eigenvalues of $\mathbf{A}^{-k}$, where $k$ is a positive integer, are $\lambda_i^{-k}$.

Extension of these results to non-integer powers (for example, the matrix square root) as well as to general matrix functions, requires the use of the spectral decomposition studied in §E.5.

### §E.2.4. Multiplicities

The following terminology is introduced to succintly state conditions for matrix reduction to diagonal form. They are largely taken from [659].

An eigenvalue $\lambda$ is said to have *algebraic multiplicity* $m_a$ if it is a root of multiplicity $m_a$ of the characteristic polynomial. If $m_a = 1$, the eigenvalue is said to be *simple*; in which case it corresponds to an isolated root of $P(\lambda) = 0$. If $m_a > 1$ the eigenvalue is said to be *nonsimple* or *multiple*. Since $P(\lambda) = 0$ has $n$ roots, we can state that the sum of algebraic multiplicities of all *distinct* eigenvalues must equal $n$. For example, the $n \times n$ identity matrix $\mathbf{I}$ has one eigenvalue of value 1 that has algebraic multiplicity $m_a = n$.

An eigenvalue $\lambda$ is said to have *geometric multiplicity* $m_g$ if the maximum number of linearly independent eigenvectors associated with it is $m_g$. This is equal to the kernel dimension of $\mathbf{A} - \lambda\,\mathbf{I}$. Note that $1 \le m_g \le m_a$. If an eigenvalue is simple, that is $m_a = 1$, obviously $m_g = 1$.

A matrix is *derogatory* if the geometric multiplicity of at least one of its eigenvalues is greater than one. Qbviously a matrix with only simple eigenvalues cannot be derogatory.

A multiple eigenvalue is *semi-simple* (a confusing name) if its algebraic and geometric multiplicities agree: $m_a = m_g$. A multiple eigenvalue that is not semi-simple is called *defective*.

### §E.2.5. Similarity Transformation

Two square matrices, $\mathbf{A}$ and $\mathbf{B}$, are said to be *similar* if there is a nonsingular matrix $\mathbf{T}$ such that

$$\mathbf{B} = \mathbf{T}^{-1}\,\mathbf{A}\,\mathbf{T}. \tag{E.12}$$

The mapping $\mathbf{A} \to \mathbf{B}$ is called a *similarity transformation*. The characteristic polynomials of the two matrices are identical. Consequently the two matrices have the same eigenvalues and each eigenvalue has the same algebraic multiplicity. An eigenvector $\mathbf{x}_A$ of $\mathbf{A}$ is transformed into the eigenvector $\mathbf{x}_B = \mathbf{T}\,\mathbf{x}_A$ of $\mathbf{B}$. It can be shown that geometric multiplicities are also preserved.

A key objective behind solving the eigenproblem of $\mathbf{A}$ is to transform $\mathbf{A}$ to a simpler form via (E.12). This process is called *reduction*. The simplest reduced form is the diagonal matrix for eigenvalues $\mathbf{\Lambda} = \mathbf{diag}(\lambda_i)$. This reduction, called *diagonalization*, is desirable but not always possible.

### §E.2.6. Diagonalization

Diagonalization via (E.12) can be done if $\mathbf{A}$ falls into one of two cases: (1) all eigenvalues are single, and (2) some eigenvalues are multiple but their geometric and algebraic multiplicities agree (i.e., they are not defective). Such matrices are called *diagonalizable*.

A diagonalizable matrix $\mathbf{A}$ has a *complete* set of $n$ linearly independent right eigenvectors $\{\mathbf{x}_i\}$ as well as left eigenvectors $\mathbf{y}_i$. Either set spans the Euclidean $n$-space and can be biorthogonalized so they verify (E.9) for $k = n$. To express this in compact matrix form, it is convenient to collect those sets into two $n \times n$ *eigenvector matrices* built by stacking eigenvectors as columns:

$$\mathbf{Y}^T = [\,\mathbf{y}_1^T \quad \mathbf{y}_2^T \quad \ldots \quad \mathbf{y}_n^T\,], \qquad \mathbf{X} = [\,\mathbf{x}_1 \quad \mathbf{x}_2 \quad \ldots \quad \mathbf{x}_n\,]. \tag{E.13}$$

These matrices verify

$$\mathbf{Y}^T\,\mathbf{A}\,\mathbf{X} = \mathbf{X}^T\mathbf{A}^T\mathbf{Y} = \mathbf{\Lambda} = \mathbf{diag}[\lambda_i], \qquad \mathbf{Y}^T\mathbf{X} = \mathbf{X}^T\mathbf{Y} = \mathbf{I}. \tag{E.14}$$

From the last relation, $\mathbf{X}^{-1} = \mathbf{Y}^T$ and $\mathbf{Y}^{-1} = \mathbf{X}^T$ whence

$$\mathbf{X}^{-1}\,\mathbf{A}\,\mathbf{X} = \mathbf{Y}^T\mathbf{A}^T\mathbf{Y}^{-T} = \mathbf{\Lambda} = \mathbf{diag}[\lambda_i]. \tag{E.15}$$

The similarity transformations shown above are a particular case of (E.12). The spectral decompositions studied in §E.5 follow immediately from these relations.

On premultiplying the first and last terms of (E.15) by $\mathbf{X}$, the original eigenproblem can be recast as

$$\mathbf{A}\,\mathbf{X} = \mathbf{X}\,\mathbf{\Lambda}, \tag{E.16}$$

This all-matrix form is often found in the literature.

### §E.2.7. Other Reduced Forms

If the matrix $\mathbf{A}$ is derogatory, diagonalization is not possible. Other reduced forms are used in such a case. They have the common goal of trying to simplify the original eigenproblem. Some possible choices are:

*Jordan Block Form*. An upper bidiagonal matrix with eigenvalues in its diagonal and ones or zeros in its superdiagonal. This reduction is always possible but may be numerically unstable.

*Upper Triangular Form*. This reduction is always possible and can be made stable. Eigenvalues occupy the diagonal of this form.

*Companion Form*. Closely related to the characteristic polynomial coefficients. Very compact but hardly used in numerical computation as it is prone to disastrous instability.

Intermediate reduced forms between the original matrix and the final reduced form include Hessemberg and tridiagonal matrices.

The sequence of operations to pass from the original to intermediate and final reduced forms may be studied in the specialized literature. See **Notes and Bibliography**.

**Example E.1.**

Consider the real unsymmetric matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 3 & 2 \\ -1 & 2 & 3 \end{bmatrix}. \tag{E.17}$$

The characteristic polynomial is $P(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = 8 - 14\lambda + 7\lambda^2 - \lambda^3$. This has the roots $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 4$, which are sorted in ascending order. The associated (unnormalized) right eigenvectors are $\mathbf{x}_1 = [\,-2\;\;-2\;\;1\,]^T$, $\mathbf{x}_2 = [\,-2\;\;1\;\;1\,]^T$ and $\mathbf{x}_3 = [\,1\;\;1\;\;1\,]^T$. The associated (unnormalized) left eigenvectors are $\mathbf{y}_1 = [\,0\;\;-1\;\;1\,]^T$, $\mathbf{y}_2 = [\,-1\;\;1\;\;0\,]^T$ and $\mathbf{y}_3 = [\,-3\;\;5\;\;4\,]^T$. These results can be used to construct the (diagonal) eigenvalue matrix and the (unnormalized) eigenvector matrices:

$$\boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}, \quad \mathbf{X} = [\,\mathbf{x}_1\;\;\mathbf{x}_2\;\;\mathbf{x}_3\,] = \begin{bmatrix} -2 & 3 & -1 \\ -2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{Y} = [\,\mathbf{y}_1\;\;\mathbf{y}_2\;\;\mathbf{y}_3\,] = \begin{bmatrix} 0 & -1 & -3 \\ -1 & 1 & 5 \\ 1 & 0 & 4 \end{bmatrix}. \tag{E.18}$$

We can now form the matrix products in (E.15):

$$\mathbf{Y}^T\,\mathbf{X} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{bmatrix}, \qquad \mathbf{Y}^T\,\mathbf{A}\,\mathbf{X} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 24 \end{bmatrix}. \tag{E.19}$$

These products are not yet $\mathbf{I}$ and $\boldsymbol{\Lambda}$ because the eigenvectors need to be appropriately normalized. To accomplish that, divide $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$ by $+\sqrt{3}$, $+\sqrt{2}$, $+\sqrt{6}$, $+\sqrt{3}$, $-\sqrt{2}$, and $+\sqrt{6}$, respectively. Note that the scale factors for $\mathbf{x}_2$ and $\mathbf{y}_2$ must have opposite signs to make the normalization work; that is, one of those vectors must be "flipped." Then

$$\mathbf{X} = \begin{bmatrix} -2/\sqrt{3} & 3/\sqrt{2} & 1/\sqrt{6} \\ -2/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} 0 & 1/\sqrt{2} & -3/\sqrt{6} \\ -1/\sqrt{3} & -1/\sqrt{2} & 5/\sqrt{6} \\ 1/\sqrt{3} & 0 & 4/\sqrt{6} \end{bmatrix}. \tag{E.20}$$

These do verify $\mathbf{X}^T\,\mathbf{X} = \mathbf{I}$ and $\mathbf{X}^T\,\mathbf{A}\,\mathbf{X} = \boldsymbol{\Lambda}$. It is easy to check that the biorthonormalized matrices (E.20) verify $\mathbf{X}^{-1} = \mathbf{Y}^T$ and $\mathbf{Y}^{-1} = \mathbf{X}^T$.

## §E.3. Real Symmetric Real Matrices

Real symmetric matrices satisfying $\mathbf{A}^T = \mathbf{A}$ are of special importance in the finite element method, and thus deserve VIM treatment[2] in this Appendix.

### §E.3.1. Spectral Properties

In linear algebra books dealing with the algebraic eigenproblem it is shown that all eigenvalues of a real symmetric matrix are real and semi-simple. Consequently symmetric matrices are diagonalizable and thus possess a complete set of linearly independent eigenvectors. Right and left eigenvectors coalesce: $\mathbf{y}_i = \mathbf{x}_i$, and such qualifiers may be omitted.

The eigenvectors $\mathbf{x}_i$ can be *orthonormalized* so that

$$\mathbf{x}_i^T\,\mathbf{x}_j = \delta_{ij}, \qquad \mathbf{x}_i^T\,\mathbf{A}\,\mathbf{x}_j = \lambda_i\,\delta_{ij}. \tag{E.21}$$

[2] VIM stands for Very Important Matrix.

This is (E.9) specialized to $\mathbf{y}_i \rightarrow \mathbf{x}_i$. If eigenvectors are stacked as columns of a matrix $\mathbf{X}$, the foregoing orthonormality conditions can be compactly stated as

$$\mathbf{X}^T\,\mathbf{X} = \mathbf{I}, \qquad \mathbf{X}^T\,\mathbf{A}\,\mathbf{X} = \mathbf{\Lambda} = \mathbf{diag}[\lambda_i]. \tag{E.22}$$

Properties (a–c) also hold for the more general class of *Hermitian matrices*.[3] A Hermitian matrix is equal to its own conjugate transpose. This extension has limited uses for structural analysis, but is important in other application areas.

### §E.3.2. Positivity

Let $\mathbf{A}$ be an $n \times n$ real symmetric matrix and $\mathbf{v}$ an arbitrary real $n$-vector. $\mathbf{A}$ is said to be *positive definite* (p.d.) if

$$\mathbf{v}^T\mathbf{A}\mathbf{v} > 0, \quad \forall \quad \mathbf{v} \neq \mathbf{0}. \tag{E.23}$$

A positive definite matrix has rank $n$. This property can be quickly checked by computing the $n$ eigenvalues $\lambda_i$ of $\mathbf{A}$. If all $\lambda_i > 0$, $\mathbf{A}$ is p.d. (This property follows on setting $\mathbf{x}$ to the orthonormalized eigenvectors of $\mathbf{A}$, and using $\mathbf{v} \rightarrow \mathbf{x}_i$ to conclude that $\lambda_i = \mathbf{x}_i^T\,\mathbf{A}\,\mathbf{x}_i > 0$.)

$\mathbf{A}$ is said to be *nonnegative*[4] if zero equality is allowed in (E.23)

$$\mathbf{v}^T\,\mathbf{A}\,\mathbf{v} \geq 0, \quad \forall \quad \mathbf{v} \neq 0. \tag{E.24}$$

A p.d. matrix is also nonnegative but the converse is not necessarily true. This property can be quickly checked by computing the $n$ eigenvalues $\lambda_i$ of $\mathbf{A}$. If $n - r$ eigenvalues are zero and the rest are positive, $\mathbf{A}$ is nonnegative with rank $r$.

An $n \times n$ symmetric real matrix $\mathbf{A}$ that has at least one negative eigenvalue is called *indefinite*.

The extension of these definitions to general matrices is discussed in specialized texts.

For the concept of *inertia* of Hermitian matrices, which include symmetric matrices as special case, see Appendix P. This concept is important in dynamics.

**Example E.2.**

Consider the real symmetric matrix

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 4 \end{bmatrix}. \tag{E.25}$$

The characteristic polynomial is $P(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = 10 - 17\lambda + 8\lambda^2 - \lambda^3$. This has the roots $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 5$, which are sorted in ascending order. The associated (unnormalized) eigenvectors are $\mathbf{x}_1 = [\,1\;\;1\;\;0\,]^T$, $\mathbf{x}_2 = [\,1\;\;-1\;\;1\,]^T$ and $\mathbf{x}_3 = [\,-1\;\;1\;\;2\,]^T$. These eigenpairs can be used to construct the (diagonal) eigenvalue matrix and the (unnormalized) eigenvector matrix:

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \qquad \mathbf{X} = [\,\mathbf{x}_1\;\;\mathbf{x}_2\;\;\mathbf{x}_3\,] = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 1 & 2 \end{bmatrix}. \tag{E.26}$$

[3] Also called *self-adjoint matrices*.

[4] A property called *positive semidefinite* by some authors.

We can now form the matrix combinations in (E.22):

$$\mathbf{X}^T\,\mathbf{X} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}, \qquad \mathbf{X}^T\,\mathbf{A}\,\mathbf{X} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 30 \end{bmatrix}. \tag{E.27}$$

These are not yet $\mathbf{I}$ and $\mathbf{\Lambda}$ because eigenvectors need to be normalized to unit length. To accomplish that, divide $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ by $+\sqrt{2}$, $+\sqrt{3}$ and $+\sqrt{6}$, respectively. Then

$$\mathbf{X} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} \end{bmatrix}, \quad \mathbf{X}^T\,\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}, \quad \mathbf{X}^T\,\mathbf{A}\,\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \mathbf{\Lambda}. \tag{E.28}$$

## §E.4. Normal and Orthogonal Matrices

Let $\mathbf{A}$ be an $n \times n$ real square matrix. This matrix is called *normal* if

$$\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T \tag{E.29}$$

A normal matrix is called *orthogonal* if

$$\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T = \mathbf{I} \qquad \text{or} \quad \mathbf{A}^T = \mathbf{A}^{-1} \tag{E.30}$$

All eigenvalues of an orthogonal matrix have modulus one, and the matrix has rank $n$.

The generalization of the orthogonality property to complex matrices, for which transposition is replaced by conjugation, leads to *unitary* matrices. These are not required, however, for the material covered in the text.

## §E.5. Spectral Decomposition and Matrix Functions

Let $\mathbf{A}$ be a $n \times n$ symmetric real matrix. As noted in §E.3, it posseses a complete set of real eigenvalues $\lambda_i$ and corresponding eigenvectors $\mathbf{x}_i$, for $i = 1, \ldots n$. The latter are assumed to be orthonormalized so that the conditions (E.21) hold. The *spectral decomposition* of $\mathbf{A}$ is the following expansion in terms of rank-one matrices:

$$\mathbf{A} = \sum_i^n \lambda_i\,\mathbf{x}_i\,\mathbf{x}_i^T. \tag{E.31}$$

This can be easily proven by postmultiplying both sides of (E.31) by $\mathbf{x}_j$ and using (E.21):

$$\mathbf{A}\,\mathbf{x}_j = \sum_i^n \lambda_i\,\mathbf{x}_i\,\mathbf{x}_i^T\,\mathbf{x}_j = \sum_i^n \lambda_i\,\mathbf{x}_i\,\delta_{ij} = \lambda_j\,\mathbf{x}_j. \tag{E.32}$$

Assume next that $\mathbf{A}$ is nonsingular so all $\lambda_i \neq 0$. Since the eigenvalues of the inverse are the reciprocal of the original eigenvalues as per (E.11), the spectral decomposition of the inverse is

$$\mathbf{A}^{-1} = \sum_i^n \frac{1}{\lambda_i}\,\mathbf{x}_i\,\mathbf{x}_i^T. \tag{E.33}$$

This is valid for more general matrix powers:

$$\mathbf{A}^m \stackrel{\text{def}}{=} \sum_i^n \lambda_i^m \, \mathbf{x}_i \, \mathbf{x}_i^T. \tag{E.34}$$

in which the exponent $m$ may be arbitrary as long as all $\lambda_i^m$ exist. This represents a generalization of the integer powers studied in §E.2.3. The expression (E.34) can be in fact extended to other matrix functions as a definition. For example, the matrix logarithm may be defined as

$$\log(\mathbf{A}) \stackrel{\text{def}}{=} \sum_i^n \log(\lambda_i)\, \mathbf{x}_i \, \mathbf{x}_i^T, \quad \text{iff} \quad \lambda_i \neq 0 \quad \forall\, i, \tag{E.35}$$

whence $\log(\mathbf{A})$ is also symmetric real if all $\lambda_i > 0$, that is, $\mathbf{A}$ is p.d. More generally, if $f(.)$ is a scalar function,

$$f(\mathbf{A}) \stackrel{\text{def}}{=} \sum_i^n f(\lambda_i)\, \mathbf{x}_i \, \mathbf{x}_i^T, \tag{E.36}$$

in which function applications to each $\lambda_i$ must make sense. For the matrix exponential function $\exp(\mathbf{A})$, no restrictions apply since the exponential of any real number is well defined.

The concept can be extended to a diagonalizable $n \times n$ matrix $\mathbf{A}$ that has eigenvalues $\lambda_i$, with corresponding right eigenvectors $\mathbf{x}_i$ and left eigenvectors $\mathbf{y}_i$. The latter are assumed to be biorthonormalized as per (E.9). For notational simplicity all eigenvalues and eigenvectors are assumed to be real. The sp[ectral decomposition is

$$\mathbf{A} = \sum_i^n \lambda_i \, \mathbf{x}_i \mathbf{y}_i^T. \tag{E.37}$$

The proof is identical to that of (E.32). If $\mathbf{A}$ and/or its eigenvectors contain complex entries, the transpose operator is replaced by conjugate transpose.

These decompositions play a fundamental role in the investigation of spectral properties, as well as the definition and study of arbitrary matrix functions.

**Example E.3.**

Consider the real symmetric matrix (E.27), with eigenvalues (in ascending order) $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 5$, and the orthonormalized eigenvectors listed in (E.28). Its spectral decomposition is $\lambda_1 \mathbf{x}_1 \mathbf{x}_1^T + \lambda_2 \mathbf{x}_2 \mathbf{x}_2^T + \lambda_3 \mathbf{x}_3 \mathbf{x}_3^T$, which numerically works out to be

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 4 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{2}{3}\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} + \frac{1}{6}\begin{bmatrix} 1 & -1 & -2 \\ -1 & 1 & 2 \\ -2 & 2 & 4 \end{bmatrix}. \tag{E.38}$$

(The number in front of the matrices is just a scale factor, not the eigenvalue.)

**Example E.4.**

Consider the real unsymmetric matrix (E.17), with eigenvalues (in ascending order) $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 4$, and the orthonormalized eigenvectors listed in (E.20). Its spectral decomposition is $\lambda_1 \mathbf{x}_1 \mathbf{y}_1^T + \lambda_2 \mathbf{x}_2 \mathbf{y}_2^T + \lambda_3 \mathbf{x}_3 \mathbf{y}_3^T$, which numerically works out to be

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 3 & 2 \\ -1 & 2 & 3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 0 & 2 & -2 \\ 0 & 2 & -2 \\ 0 & -1 & 1 \end{bmatrix} + \begin{bmatrix} 3 & -3 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \end{bmatrix} + \frac{2}{3}\begin{bmatrix} -3 & 5 & 4 \\ -3 & 5 & 4 \\ -3 & 5 & 4 \end{bmatrix}. \tag{E.39}$$

## §E.6. Matrix Deflation

In eigenproblem solution processes it often happens that one or more eigenpairs are either known in advance, or are obtained before the others.[5] A procedure by which those eigenpairs are "deactivated" in some manner is called *matrix deflation* or simply *deflation* Two deflation schemes are commonly used in practice. They are associated with the names Hotelling and Wielandt, respectively.

### §E.6.1. Hotelling Deflation

Let $\mathbf{A}$ be a diagonalizable $n \times n$ real matrix with the spectral decomposition (E.37), in which the eigenvectors satisfy the biorthonormality condition (E.9). Suppose that eigenvalue $\lambda_i$, with associated right eigenvector $\mathbf{x}_i$ and left eigenvector $\mathbf{y}_i$, are known. They are assumed to be real for notational simplicity. A *deflated matrix* $\tilde{\mathbf{A}}_i$ is obtained by subtracting a rank one correction:

$$\tilde{\mathbf{A}}_i = \mathbf{A} - \lambda_i\, \mathbf{x}_i \mathbf{y}_i^T. \tag{E.40}$$

Obviously this has a spectral decomposition identical to (E.37), except that $\lambda_i$ is replaced by zero. For example if $i = 1$, the deflated matrix $\tilde{\mathbf{A}}_1$ has eigenvalues $0, \lambda_2, \ldots \lambda_n$, and the same right and left eigenvectors. If $\mathbf{A}$ is symmetric, $\mathbf{y}_i = \mathbf{x}_i$ and $\tilde{\mathbf{A}}_i$ is also symmetric. This deflation scheme produces nothing new if $\lambda_i$ is zero.

If $\mathbf{A}$ and/or its eigenvectors contain complex entries, the transpose operator is replaced by the conjugate transpose.

**Example E.5.**

Consider the real symmetric matrix (E.27), with eigenvalues (in ascending order) $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 5$, and the orthonormalized eigenvectors listed in (E.28). Deflating $\lambda_1$, $\lambda_2$ and $\lambda_3$ in turn through (E.40) gives

$$\tilde{\mathbf{A}}_1 = \frac{1}{6}\begin{bmatrix} 7 & -1 & 4 \\ -1 & 7 & -4 \\ 4 & -4 & 4 \end{bmatrix}, \quad \tilde{\mathbf{A}}_2 = \frac{1}{3}\begin{bmatrix} 4 & -1 & -5 \\ -1 & 4 & 5 \\ -5 & 5 & 10 \end{bmatrix}, \quad \tilde{\mathbf{A}}_3 = \frac{1}{2}\begin{bmatrix} 3 & -3 & -2 \\ -3 & 3 & 2 \\ -2 & 2 & 8 \end{bmatrix}. \tag{E.41}$$

The eigenvalues of these matrices are $\{0, 1, 2\}$, $\{0, 1, 5\}$ and $\{0, 2, 5\}$, respectively. The eigenvectors remain the same.

**Example E.6.**

Consider the real unsymmetric matrix (E.17), with eigenvalues (in ascending order) $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 4$, and the biorthonormalized eigenvectors listed in (E.20). Deflating $\lambda_1$, $\lambda_2$ and $\lambda_3$ in turn through (E.40) gives

$$\tilde{\mathbf{A}}_1 = \frac{1}{3}\begin{bmatrix} 3 & 1 & 8 \\ -3 & 7 & 8 \\ -3 & 7 & 8 \end{bmatrix}, \quad \tilde{\mathbf{A}}_2 = \begin{bmatrix} -2 & 4 & 2 \\ -2 & 4 & 2 \\ -2 & 3 & 3 \end{bmatrix}, \quad \tilde{\mathbf{A}}_3 = \frac{1}{2}\begin{bmatrix} 9 & -7 & -2 \\ 3 & -1 & -2 \\ 3 & -4 & 1 \end{bmatrix}. \tag{E.42}$$

The eigenvalues of these matrices are $\{0, 2, 4\}$, $\{0, 1, 4\}$ and $\{0, 1, 2\}$, respectively. The right and left eigenvectors remain the same.

---

[5] The second scenario is the case in the classical power iteration method, which was important for "large" eigenproblems before 1960. The largest eigenvalue in modulus, also known as the *dominant* one, is obtained along with the associated eigenvector(s). The matrix is then deflated so that the next-in-modulus (subdominant) eigenvalue becomes the dominant one.

### §E.6.2. Wielandt Deflation

This deflation scheme is more general than the previous one, which is included as special case. Let $\mathbf{A}$ be a diagonalizable $n \times n$ real matrix, with eigenvectors that satisfy the biorthonormality condition (E.9). Again suppose that eigenvalue $\lambda_i$, with associated right eigenvector $\mathbf{x}_i$ and left eigenvector $\mathbf{y}_i$, are known. All are assumed to be real for notational simplicity. Let $\mathbf{w}_i$ be a column $n$-vector satisfying $\mathbf{w}_i^T\mathbf{x}_i = 1$, and $\sigma$ a real scalar called the *spectral shift* or simply *shift*. The Wielandt-deflated matrix is

$$\tilde{\mathbf{A}}_i = \mathbf{A} - \sigma \mathbf{x}_i \mathbf{w}_i^T. \tag{E.43}$$

Premultiplying both sides by $\mathbf{y}_j^T$ yields

$$\mathbf{y}_j^T\, \tilde{\mathbf{A}}_i = \mathbf{y}_j^T\, \mathbf{A} - \sigma\delta_{ij}\mathbf{w}_i^T \quad \text{or} \quad \tilde{\mathbf{A}}_i^T\, \mathbf{y}_j = \mathbf{A}^T\, \mathbf{y}_j = \lambda_j\, \mathbf{y}_j, \quad \text{if} \quad j \neq i. \tag{E.44}$$

Consequently the eigenvalues $\lambda_j$ for $j \neq i$, as well as the corresponding left eigenvectors $\mathbf{y}_j$, are preserved by the deflation process. What happens for $i = j$? Postmultiplying by $\mathbf{x}_i$ yields

$$\tilde{\mathbf{A}}_i\, \mathbf{x}_i = \mathbf{A}\, \mathbf{x}_i - \sigma\, \mathbf{x}_i, \quad \text{or} \quad \tilde{\mathbf{A}}_i\, \mathbf{x}_i - (\lambda_i - \sigma)\, \mathbf{x}_i = 0. \tag{E.45}$$

This shows that eigenvector $\mathbf{x}_i$ is preserved whereas eigenvalue $\lambda_i$ is shifted by $-\sigma$. The right eigenvectors $\mathbf{x}_j$ for $i \neq j$ are not generally preserved; they become linear combinations of $\mathbf{x}_i$ and $\mathbf{x}_j$. Following [659, p. 118], assume $\tilde{\mathbf{x}}_j = \mathbf{x}_j - \gamma_j\, \mathbf{x}_i$. Then

$$\tilde{\mathbf{A}}_i\, \tilde{\mathbf{x}}_j = \lambda_j\, \mathbf{x}_j - (\gamma_j\, \lambda_i + \sigma\, \mathbf{w}^T\, \mathbf{x}_j - \sigma\, \gamma_j)\, \mathbf{x}_i. \tag{E.46}$$

This shows that $\tilde{\mathbf{x}}_j$ is an eigenvector of $\tilde{\mathbf{A}}_i$ if

$$\gamma_i = \frac{\mathbf{w}^T\mathbf{x}_j}{1 - (\lambda_i - \lambda_j)/\sigma}, \quad \text{for} \quad j \neq i. \tag{E.47}$$

This is undefined if the denominator vanishes: $\lambda_i - \lambda_j = \sigma$, which happens if $\lambda_i - \sigma$ becomes a multiple eigenvalue of the deflated matrix.

The process shown so far only requires knowledge of the right eigenvector $\mathbf{x}_i$. This *need not be normalized*, since that is taken care of by the unit dot product with $\mathbf{w}_i$. If $\sigma = \lambda_i$ and $\mathbf{w}_i = c\mathbf{y}_i$, in which $c$ is determined by $\mathbf{w}_i^T\mathbf{x}_i = 1$, the Hotelling deflation (E.40) is recovered. This choice preserves *all* eigenvectors while shifting $\lambda_i$ to zero, and retains symmetry if $\mathbf{A}$ is real symmetric.

### §E.6.3. Matrix Order Reduction

The freedom to choose $\mathbf{w}$ in the Wielandt deflation (E.43) can be exploited to combine deflation with matrix order reduction. As in §E.6.2, assume that eigenvalue $\lambda_i$ as well as the right eigenvector $\mathbf{x}_i$ are known, and are both assumed real. Denote the $j^{th}$ entry of $\mathbf{x}_i$ by $x_{ij}$ and assume $x_{ij} \neq 0$. Pick $\mathbf{w}_i = \mathbf{e}_j^T\mathbf{A}$, in which $\mathbf{e}_j$ is the $n$-vector with all zero entries except the $j^{th}$ one, which is one. Consequently the chosen $\mathbf{w}_i$ is $\mathbf{a}_j$, the $j^{th}$ row of $\mathbf{A}$. To identify that row, the deflated matrix will be denoted by $\mathbf{A}_{ij}$:

$$\tilde{\mathbf{A}}_{ij} = \mathbf{A} - S\, \mathbf{x}_i\, \mathbf{a}_j = (\mathbf{I} - S\, \mathbf{x}_i\, \mathbf{e}_j^T)\, \mathbf{A}, \tag{E.48}$$

in which $S = 1/x_{ij}$. By inspection the $j^{th}$ row of $\tilde{\mathbf{A}}_{ij}$ is zero. All eigenvectors $\tilde{\mathbf{x}}_j$ of $\tilde{\mathbf{A}}_{ij}$ with $i \neq j$ will have their $j^{th}$ entry equal to zero. Consequently that row and the corresponding column can be deleted, which reduces its order by one. The order-reduced $(n-1) \times (n-1)$ matrix is denoted by $\hat{\mathbf{A}}_{ij}$. The operation is called *deflate-and-order-reduce*. It is better explained by an example.

**Example E.7**.

Consider the real symmetric matrix (E.27), with eigenvalues (in ascending order) $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 5$. The matrix and (unnormalized) eigenvector matrix are reproduced for convenience:

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 4 \end{bmatrix}, \qquad \mathbf{X} = [\, \mathbf{x}_1 \;\; \mathbf{x}_2 \;\; \mathbf{x}_3 \,] = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \mathbf{Y}. \tag{E.49}$$

To deflate-and-reduce $\lambda_1$, we pick eigenvector $\mathbf{x}_1$, the first entry of which is $x_{11} = 1$, and the first row of $\mathbf{A}$: $\mathbf{a}_1 = [\, 1 \;\; -1 \;\; -1 \,]$, to get

$$\tilde{\mathbf{A}}_{11} = \mathbf{A} - \frac{1}{x_{11}} \mathbf{x}_1 \mathbf{a}_1 = \begin{bmatrix} 0 & 0 & 0 \\ -3 & 3 & 2 \\ -1 & 1 & 4 \end{bmatrix} \Rightarrow \hat{\mathbf{A}}_{11} = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}. \tag{E.50}$$

The eigenvalues of $\tilde{\mathbf{A}}_{11}$ and $\hat{\mathbf{A}}_{11}$ are $\{0, 2, 5\}$ and $\{2, 5\}$, respectively. The (unnormalized) eigenvector matrices are

$$\tilde{\mathbf{X}}_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \tilde{\mathbf{Y}}_{11} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \quad \hat{\mathbf{X}}_{11} = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}, \quad \hat{\mathbf{Y}}_{11} = \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}. \tag{E.51}$$

Since the deflated matrices (E.50) are unsymmetric, the left and right eigenvalues no longer coalesce. Note that left eigenvectors except $\mathbf{y}_1$ are preserved, whereas right eigenvectors except $\mathbf{x}_1$ change.

What happens if $\lambda_1$ is deflated via the second row? We get

$$\tilde{\mathbf{A}}_{12} = \mathbf{A} - \frac{1}{x_{12}} \mathbf{x}_1 \mathbf{a}_2 = \begin{bmatrix} 3 & -3 & -2 \\ 0 & 0 & 0 \\ -1 & 1 & 4 \end{bmatrix} \Rightarrow \hat{\mathbf{A}}_{12} = \begin{bmatrix} 3 & -2 \\ -1 & 4 \end{bmatrix}. \tag{E.52}$$

Deflated and reduced matrices are different from those in (E.50) although they have the same eigenvalues.

Deflation via the third row is not possible since the third entry of $\mathbf{x}_1$ is zero.

The deflate-and-order-reduce process can be obviously continued by renaming $\hat{\mathbf{A}}_{11}$ or $\hat{\mathbf{A}}_{12}$ to $\mathbf{A}$, and shifting one of its eigenvalues, say $\lambda_1 = 2$, to zero. The reduced matrix will be $1 \times 1$ with eigenvalue 5, so the resulting $\hat{\mathbf{A}}_{1j}$ ($j = 1$ or 2) will be $[\, 5 \,]$.

## §E.6.4. Nullspace Reduction

If several eigenpairs are known, they can be simultaneously deactivated instead of one by one. The procedure is known as *block deflation*.

The Hotelling deflation (E.40) is trivially extended to the block case: just subtract several rank-one terms. Eigenvectors are not altered. On the other hand, extending the general Wielandt deflation (E.43) is not trivial because each single eigenpair deflation changes the right eigenvectors, and generally no particular computational advantage accrues from doing blocks. But there is a special case where it deserves attention: deflate-and-order-reduce a null space.

Suppose the $n \times n$ diagonalizable real matrix $\mathbf{A}$ is $m \geq 1$ times singular. That is, its kernel has dimension $m$. Reordering the characteristic roots as necessary, we can state that $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$ is a null eigenvalue of algebraic multiplicity $m$. The associated eigenspace matrix of right eigenvectors is denoted by

$$\mathbf{X}_m = [\, \mathbf{x}_1 \quad \mathbf{x}_2 \quad \ldots \quad \mathbf{x}_m \,]. \tag{E.53}$$

(These eigenvectors are not necessarily orthonormal, but must be linearly independent.) For brevity $\mathbf{X}_m$ will be referred to as the *nullspace basis*, or simply *nullspace*, of $\mathbf{A}$. From a computational standpoint it is important to realize that $\mathbf{X}_m$ can be obtained without solving the eigenproblem (E.1). Gaussian row reduction is sufficient.[6] For example, the built-in function `NullSpace` of *Mathematica* returns that matrix.

Suppose that we want to get rid of both multiple singularity and nullspace by block deflate-and-reduce to end up with a $(n-m) \times (n-m)$ matrix with the same nonzero eigenvalues. To do this, pick $m$ rows to zero out, and stack the corresponding rows of $\mathbf{A}$ to form the $m \times n$ matrix $\mathbf{A}_m$. Then

$$\tilde{\mathbf{A}}_m = \mathbf{A} - \mathbf{X}_m\,\mathbf{S}\,\mathbf{U}\,\mathbf{A}_m. \tag{E.54}$$

Here $\mathbf{S}$ is a $m \times m$ diagonal scaling matrix that contains the reciprocals of the eigenvector entries corresponding to the rows to be cleared, while $\mathbf{U}$ is a $m \times m$ unit upper triangular matrix, the unknown entries of which are chosen so that targetted rows of the deflated matrix vanish.[7] It is of course possible to combine $\mathbf{S}$ and $\mathbf{U}$ into a single upper triangular matrix $\mathbf{U}_S = \mathbf{S}\,\mathbf{U}$. This contains eigenvector entry reciprocals in its diagonal and thus it is not necessarity unit upper triangular.

If $m = 1$, $\mathbf{S}$ and $\mathbf{U}$ collapse to $1 \times 1$ matrices containing $1/x_{ij}$ and 1, respectively, and (E.54) reduces to the single vector case (E.48).

The procedure is illustrated by a symbolic example.

**Example E.8**. The following $3 \times 3$ symbolic symmetric matrix of rank 1 was discovered through *Mathematica*. It depends on three real scalar parameters: $b$, $c$, and $d$, which must be nonzero:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} \frac{b\,c}{d} & b & c \\ b & \frac{b\,d}{c} & d \\ c & d & \frac{c\,d}{b} \end{bmatrix}, \tag{E.55}$$

in which $\mathbf{a}_i$ denotes the $i^{th}$ row of $\mathbf{A}$. The eigenvalues and (unnormalized) eigenvectors are

$$\lambda_1 = \lambda_2 = 0, \quad \lambda_3 = \frac{b\,c}{d} + \frac{b\,d}{c} + \frac{c\,d}{b}, \qquad \mathbf{X} = [\,\mathbf{x}_1 \;\; \mathbf{x}_2 \;\; \mathbf{x}_3\,] = \begin{bmatrix} -d/b & -d/c & b/d \\ 0 & 1 & b/c \\ 1 & 0 & 1 \end{bmatrix}. \tag{E.56}$$

The nullspace is spanned by $[\,\mathbf{x}_1 \quad \mathbf{x}_2\,]$. To block deflate it so that the first two rows of the deflated matrix vanish we apply (E.54):

$$\begin{aligned} \tilde{\mathbf{A}}_2 &= \mathbf{A} - [\,\mathbf{x}_1 \quad \mathbf{x}_2\,]\,\mathbf{S}\,\mathbf{U}\,[\,\mathbf{a}_1 \quad \mathbf{a}_2\,] \\ &= \mathbf{A} - \begin{bmatrix} -d/b & -d/c \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1/1 & 0 \\ 0 & 1/1 \end{bmatrix} \begin{bmatrix} 1 & u_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{b\,c}{d} & b & c \\ b & \frac{b\,d}{c} & d \end{bmatrix} \\ &= \begin{bmatrix} c + \frac{b\,c}{d} + \frac{b\,d}{c} + d\,u_{12} & b + d + \frac{b\,d^2}{c^2} + \frac{d^2\,u_{12}}{c} & c + \frac{c\,d}{b} + \frac{d^2}{c} + \frac{d^2\,u_{12}}{b} \\ 0 & 0 & 0 \\ c - \frac{b\,c}{d} - b\,u_{12} & d - \frac{b\,(c + d\,u_{12})}{c} & c\,(\frac{d}{b} - 1) - d\,u_{12} \end{bmatrix} \end{aligned} \tag{E.57}$$

[6] At least in exact arithmetic; in floating-point work the detection of rank is an ill-posed problem.

[7] The effect of $\mathbf{U}$ is essentially to appropriately scale and combine the original eigenvectors stacked in $\mathbf{X}_m$ so they become equivalent to the modified right eigenvectors in the one-vector-at-a-time deflation.

A simple calculation shows that taking $u_{12} = -c/d - b\,c/d^2 - b/(c\,d)$ makes the first row of $\tilde{\mathbf{A}}_2$ null, reducing it to

$$\tilde{\mathbf{A}}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ c + b^2\left(\frac{1}{c} + \frac{c}{d^2}\right) & d + b^2\left(\frac{1}{d} + \frac{d}{c^2}\right) & \frac{b\,c}{d} + \frac{b\,d}{c} + \frac{c\,d}{b} \end{bmatrix} \tag{E.58}$$

Removing the first two rows and columns reduces this to the $1 \times 1$ matrix $[\, b\,c/d + b\,d/c + c\,d/b \,] = [\, \lambda_3 \,]$.

## Notes and Bibliography

The best overall coverage of the algebraic eigenproblem is still the monograph of Wilkinson [803], although some of the computational methods described therein are dated; especially for sparse matrices. For symmetric matrices the most elegant exposition is Parlett's textbook [548], reprinted by SIAM.

Hotelling deflation was proposed in 1933 while its generalization: Wielandt deflation, appeared in 1944; see [374,803] for historical references. Block Wielandt deflation was introduced by Brauer [105] in 1952. Strangely, the nullspace deflation procedure described in §E.6.4 is not found in the literature despite its usefulness.

In FEM models the solution of very large eigensystems (thousands or millions of DOF) occurs frequently in applications such as structural analysis for vibration and stability. The matrices involved: mass, material stiffness, and geometric stiffness, are typically very sparse, meaning that the proportion of nonzero entries is very small. Specialized methods for efficient handling of such problems have been steadily developed since the 1960s. A comprehensive coverage from the standpoint of underlying theory is provided in the textbook by Saad [659], which has been republished by SIAM in a revised and updated 2011 edition. The book by Stewart [702] focuses more on theory, especially perturbation results.

The most exhaustive study of matrix functions, combining theory and computer implementation, is the textbook by Higham [363], which has a comprehensive reference list until 2007.

## Exercises for Appendix E
## Linear Algebra: Eigenproblems

**EXERCISE E.1** What are the eigenvalues and eigenvectors of a diagonal matrix?

**EXERCISE E.2** Matrix $\mathbf{A}$ is changed to $\mathbf{A} - \sigma\,\mathbf{I}$, in which $\sigma$ is a scalar called the *spectral shift* and $\mathbf{I}$ is the identity matrix. Explain what happens to the eigenvalues and eigenvectors of $\mathbf{A}$.

**EXERCISE E.3** Show that the eigenvalues of a real symmetric square matrix are real, and that all eigenvector entries are real.

**EXERCISE E.4** Let the $n$ real eigenvalues $\lambda_i$ of a real $n \times n$ symmetric matrix $\mathbf{A}$ be classified into two subsets: $r$ eigenvalues are nonzero whereas $n - r$ are zero. Show that $\mathbf{A}$ has rank $r$.

**EXERCISE E.5** Show that the characteristic polynomials of $\mathbf{A}$ and $\mathbf{T}^{-1}\,\mathbf{A}\,\mathbf{T}$, in which $\mathbf{T}$ is a nonsingular transformation matrix, are identical. Conclude that the similarity transformation (E.12) preserves eigenvalues and their algebraic multiplicities.

**EXERCISE E.6** Let $\mathbf{Q}$ be a real orthogonal matrix: $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$. Show that all of its eigenvalues $\lambda_i$, which are generally complex, have unit modulus.

**EXERCISE E.7** Let $\mathbf{A}$ be real skew-symmetric, that is, $\mathbf{A} = -\mathbf{A}^T$. Show that all eigenvalues of $\mathbf{A}$ are purely imaginary or zero.

**EXERCISE E.8** Let $\mathbf{P}$ be a real square matrix that satisfies

$$\mathbf{P}^2 = \mathbf{P}. \tag{EE.1}$$

Such matrices are called *idempotent*, and also *orthogonal projectors*. Show that all eigenvalues of $\mathbf{P}$ are either zero or one. *Hint*: (E.10).

**EXERCISE E.9** Two conforming diagonalizable square matrices $\mathbf{A}$ and $\mathbf{B}$ are said to commute if $\mathbf{AB} = \mathbf{BA}$. Show that necessary and sufficient conditions for this to happen is that $\mathbf{A}$ and $\mathbf{B}$ have identical eigenvectors. *Hint*: for sufficiency use $\mathbf{A} = \mathbf{X}\,\mathbf{\Lambda}_A\,\mathbf{X}^{-1}$ and $\mathbf{B} = \mathbf{X}\,\mathbf{\Lambda}_B\,\mathbf{X}^{-1}$.

**EXERCISE E.10**

(Advanced) A matrix whose elements are equal on any line parallel to the main diagonal is called a Toeplitz matrix. (They arise in finite difference or finite element discretizations of regular one-dimensional grids.) Show that if $\mathbf{T}_1$ and $\mathbf{T}_2$ are any two Toeplitz matrices, they commute: $\mathbf{T}_1\mathbf{T}_2 = \mathbf{T}_2\mathbf{T}_1$. Hint: do a Fourier transform to show that the eigenvectors of any Toeplitz matrix are of the form $\{e^{i\omega nh}\}$; then apply the results of the previous Exercise.

**EXERCISE E.11** Let $\mathbf{A}$ and $\mathbf{B}$ be two diagonalizable square matrices of equal order that commute. Show that the eigenvalues of $\mathbf{AB}$ are the product of the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$ (Frobenius).