Linear Algebra Primer

Daniel S. Stutts, Ph.D.

Original Edition: 12/1991 Revised: 4/1/04 Current Edition: 8-3-2014 (fixed typos)

1 Introduction

This primer was written to provide a brief overview of the main concepts and methods in elementary linear algebra. It was not intended to take the place of any of the many elementary linear algebra texts in the market. It contains relatively few examples and no exercises. The interested reader will find more in depth coverage of these topics in introductory text books. Much of the material including the order in which it is presented comes from Howard Anton's "Elementary Linear Algebra" 2^{nd} Ed., John Wiley, 1977. Another excellent basic text is "Linear Algebra and Its Applications," by Charles G. Cullen. A more advanced text is "Linear Algebra and its Applications" by Gilbert Strang.

The author hopes that this primer will answer some of your questions as they arise, and provide some motivation (prime the pump, so to speak) for you to explore the subject in more depth. At the very least, you now have a list (albeit a short one) of references from which to obtain more in depth explanation.

It should be noted that the examples given here have been motivated by the solution of consistent systems of equations which have an equal number of unknowns and equations. Therefore, only the analysis of square (n by n) matrices have been presented. Furthermore, only the properties of real matrices (those with real elements) have been included.

1.1 Explanation of Notation Used

For clarity of notation, **bold** symbols are used to denote vectors and matrices. For matrices, upper case bold letters are used, and for vectors, which are $n \times 1$ matrices, bold lower case letters are used. Non-bold symbols are used to denote scalar quantities.

Subscripts are used to denote elements of matricies or vectors. Superscripts (when not referring to exponentiation) are used to identify eigenvectors and their respective components.

1.1.1 Indicial Notation

A matrix $\mathbf{A}$ may be described by *indicial* notation. The term located at the i^{th} row and j^{th} column is denoted be the scalar a_{ij}.

Thus, ij^{th} component of the sum of two matrices, $\mathbf{A}$ and $\mathbf{B}$, may be written: $[\mathbf{A}+\mathbf{B}]_{ij} = a_{ij} + b_{ij}$

Example 1

$$\mathbf{C} = \mathbf{A}_{(2\times2)} + \mathbf{B}_{(2\times2)} = \begin{bmatrix} (a_{11}+b_{11}) & (a_{12}+b_{12}) \\ (a_{21}+b_{21}) & (a_{22}+b_{22}) \end{bmatrix}$$

Hence, for example, $c_{12} = a_{12} + b_{12}$.

2 Linear Systems of Equations

The following systems of equations

$$\begin{gathered} a_{11}x_{11} + a_{12}x_{12} + \cdots + a_{ln}x_{ln} = b_1 \\ a_{21}x_{21} + a_{22}x_{22} + \cdots + a_{2n}x_{2n} = b_2 \\ \vdots \\ a_{n1}x_{nl} + a_{n2}x_{n2} + \cdots + a_{nn}x_{nn} = b_n \end{gathered} \tag{1}$$

may be written in the matrix form as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{ln} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{2}$$

or

$$\mathbf{Ax} = \mathbf{b} \tag{3}$$

where $\mathbf{A}$ is an n by n matrix and $\mathbf{x}$ and $\mathbf{b}$ are n by 1 matrices or vectors.

While the solution of systems of linear equations provides one significant motivation to study matrices and their properties, there are numerous other applications for matrices. All applications of matrices require a reasonable degree of understanding of matrix and vector properties.

3 Matrix Properties and Definitions

For any matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, the following hold:

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
3. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$

   Provided that $\mathbf{AB}$ and $\mathbf{AC}$ are defined

   Definition 1 $\mathbf{AB}$ *is defined if* $\mathbf{B}$ *has the same number of rows as* $\mathbf{A}$ *has columns.*

   $\mathbf{A}_{(m\times r)}\mathbf{B}_{(r\times n)} = \mathbf{AB}_{(m\times n)}$

4. *Identity Matrix:* $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$ where

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

5. *Zero Matrix:* $\mathbf{0A} = \mathbf{A0} = \mathbf{0}$

$$\mathbf{0} = \begin{bmatrix} 0 & 0 & 0 & . & . \\ 0 & 0 & 0 & . & . \\ 0 & 0 & 0 & . & . \\ . & . & . & . & . \\ . & . & . & . & 0 \end{bmatrix}$$

6. $\mathbf{A} + \mathbf{0} = \mathbf{A}$

3.1 The transpose operation and its properties

The transpose of a matrix $\mathbf{A}$, written as $\mathbf{A}^T$, is the matrix $\mathbf{A}$ with its off-diagonal components reflected across the main diagonal. Hence, defining the components of the matrix $\mathbf{A}$ as a_{ij} in indicial notation, the components of $\mathbf{A}^T$ are given by a_{ji}.

3.1.1 Properties of the transpose operation

1. $(\mathbf{A}^T)^T = \mathbf{A}$
2. $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
3. $(k\mathbf{A})^T = \mathbf{k}\mathbf{A}^T$
4. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

A matrix $\mathbf{A}$ is termed symmetric if $\mathbf{A}^T = \mathbf{A}$ and *skew-symmetric* if $\mathbf{A}^T = -\mathbf{A}$. Clearly, a skew-symmetric matrix can only have zero diagonal terms.

3.2 Multiplication of a Matrix

There are a few rules for matrix and vector multiplication which must be considered in addition to the rules for the more familiar scalar algebra. These are enumerated with examples below.

1. *Multiplication by a scalar* The product of a scalar and the matrix with each of its elements multiplied by the scalar. For example:

   $\alpha\mathbf{A} = [\alpha a_{ij}] = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} \\ \alpha a_{21} & \alpha a_{22} \end{bmatrix}$

2. *Multiplication by a matrix* The product of two matrices, $\mathbf{A}$ and $\mathbf{B}$, when it is defined is another matrix, $\mathbf{C}$

   $$\mathbf{C} = \mathbf{AB} \tag{1}$$

   where the components of $\mathbf{C}$ may be computed as follows:

   $$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = c_{ij} \tag{2}$$

   Example 2

   $$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

Note that in general, matrices are not commutative over multiplication:

$$\mathbf{AB} \neq \mathbf{BA} \tag{3}$$

This fact leads to the definition of *pre-multiplication* and *post-multiplication:*

Definition 1 *A matrix* $\mathbf{A}$ *is said to be pre-multiplied by a matrix* $\mathbf{B}$ *when* $\mathbf{B}$ *multiplies* $\mathbf{A}$ *from the left, and post-multiplied when* $\mathbf{B}$ *is multiplied by* $\mathbf{A}$ *from the left.*

Other terminology for the direction of multiplication in common use is *left multiplication* and *right multiplication.* In Equation (3), on the r.h.s., $\mathbf{A}$ is being pre-multiplied by $\mathbf{B}$, and on the l.h.s. $\mathbf{A}$ is being post-multiplied by $\mathbf{B}$.

3. *Multiplication by a vector:*

$$\mathbf{Ax} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} \tag{4}$$

Pre-multiplication of a matrix by a vector requires taking the transpose of the vector first in order to comply with the rules of matrix multiplication.

3.3 Matrix Inverse

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \quad \text{if} \quad \mathbf{A}^{-1} \quad \text{exists,}$$

3.4 The Determinant Operation

Recall from vector mechanics that in three space (dimensions) the vector product of any two vectors $\mathbf{A} = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})$ and $\mathbf{B} = (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})$ is defined as

$$\mathbf{A} \times \mathbf{B} = det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

$$= (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \tag{5}$$

This method of computing the determinant is called cofactor expansion. A cofactor is the signed minor of a given element in a matrix. A minor M_{ij} is the determinant of the sub matrix which remains after the i^{th} row and the j^{th} column of the matrix are deleted. In this case, we have

$$M_{11} = (a_2b_3 - a_3b_2) \tag{6}$$

$$M_{12} = (a_1b_3 - a_3b_1) \tag{7}$$

$$M_{13} = (a_1b_2 - a_2b_1) \tag{8}$$

The cofactors are given by

$$C_{ij} = (-1)^{i+j} M_{ij} \tag{9}$$

Hence,

$$C_{11} = (-1)^{1+1} M_{11} = M_{11} \tag{10}$$

$$C_{12} = (-1)^{1+2} M_{12} = -M_{12} \tag{11}$$

etc.

In the above example,

$$det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = c_{11}\mathbf{i} + c_{12}\mathbf{j} + c_{13}\mathbf{k} \tag{12}$$

Example 3 *Let* $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 3 & -2 \end{bmatrix}$ *then*

$$det\mathbf{A} = 2(-2) - 1(3) = -7$$

Example 4 *Let* $\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 3 & -2 & 1 \\ 1 & -1 & 2 \end{bmatrix}$

then

$$det\mathbf{A} = 2(-2(2) - (-1)(1)) - 1(3(2) - 1(1)) + 0(3(-1) - (-2)(1)) = 2(-3) - 5 = -11$$

Expansion by cofactors is mostly useful for small matrices (less than 4×4). For larger matrices, the number of operations becomes prohibitively large. For example:

A 2 by 2 matrix requires 2 multiplications

A 3 by 3 matrix requires 9 multiplications

However, a 4 by 4 matrix requires the computation of 4+4! = 28 signed elementary products.

A 10 by 10 matrix would require 10 + 10! = 3,628,810 signed elementary products!

This trend suggests that soon even the largest and fastest computers would choke on such a computation.

For large matrices, the determinant is best computed using *row reduction.*

Row reduction consists of using elementary row and column operations to reduce a matrix down to a simpler form, usually upper or lower triangular form.

This is accomplished by multiplying one row by a constant and adding it another row to produce a zero at the desired position.

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

Reduce $\boldsymbol{A}$ *to upper triangular form, i.e., all zeros under the main diagonal (2 -2 2).*

Multiplying row 1 by -1/2 and adding it to row 3 yields

$$\begin{bmatrix} 2 & 1 & 0 \\ 3 & -2 & 1 \\ 0 & -1.5 & 2 \end{bmatrix}$$

Similarly, multiplying row 1 by -3/2 and adding it to row 2 yields

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & -3.5 & 1 \\ 0 & -1.5 & 2 \end{bmatrix}$$

multiplying row 2 by $-\frac{1.5}{3.5} = -\frac{3}{7}$ and adding it to row 3

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & -\frac{7}{2} & 1 \\ 0 & 0 & \frac{11}{7} \end{bmatrix}$$

The determinant is now easily computed by multiplying the elements of the main diagonal.

$$det\mathbf{A} = 2(-\tfrac{7}{2})(\tfrac{11}{7}) = -11$$

This type of row reduction is called Gaussian elimination and is much more efficient than the cofactor expansion technique for large matrices.

3.5 Properties of Determinant Operations

1. If A is a square matrix then $det(\mathbf{A}^T) = det(\mathbf{A})$
2. $det(k\mathbf{A}) = k^n det(\mathbf{A})$ where $\mathbf{A}$ is an $n \times n$ matrix and k is a scalar
3. $det(\mathbf{AB}) = det(\mathbf{A})det(\mathbf{B})$ where $\mathbf{A}$ and $\mathbf{B}$ are square matrices of the same size.
4. A square matrix $\mathbf{A}$ is invertible if and only if $det(\mathbf{A}) \neq 0$

Proof 1 *If* A *is invertible then* $\mathbf{AA}^{-1} = \mathbf{I} \Rightarrow det\left(\mathbf{AA}^{-1}\right) = (det\mathbf{A})\left(det\mathbf{A}^{-1}\right) = det\mathbf{I} = 1$ *Thus,* $det\mathbf{A}^{-1} = \frac{1}{det\mathbf{A}} \Rightarrow det\mathbf{A} \neq 0.$ □

An important implication of this result is the following.

5. For a homogeneous system of linear equations $\mathbf{Ax} = \mathbf{0}$

   There exists a nontrivial $\mathbf{x}$ ($\mathbf{x} \neq \mathbf{0}$) if and only if $det(\mathbf{A}) = 0$

   Proof 2 *Assume* $det(\mathbf{A}) \neq \mathbf{0}$. *This implies that* $\mathbf{A}^{-1}$ *exists.*
   $\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{Ix} = 0$. *This implies that* $\mathbf{x} = 0$, *(a contradiction), so the only way that* $\boldsymbol{x}$ *can be other than zero is for* $det(\mathbf{A}) = 0$, *and hence, for non-trivial* $\mathbf{x}$, $\mathbf{A}^{-1}$ *not to exist.* □

   This result is used *very* often in applied mathematics, physics and engineering.

6. If $\mathbf{A}$ is invertible then
   $\mathbf{A}^{-1} = \frac{1}{det(\mathbf{A})} adj(\mathbf{A})$ where $adj\mathbf{A}$ is the *adjoint* of A.

   Definition 2 *The adjoint of a matrix* $\boldsymbol{A}$ *is defined as the transpose of the cofactor matrix of* $\boldsymbol{A}$.

Another way to calculate the inverse of a matrix is by Gaussian elimination. This method is easier to apply on larger matrices.

Since $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, we start with the matrix $\mathbf{A}$ which we want to invert on the left and the identity matrix on the right. We then do elementary row operations (Gaussian Elimination) on the matrix while simultaneously doing the same operations on $\mathbf{I}$. This can be accomplished by adjoining the two matrices to form a matrix of the form [$\mathbf{A}$ $\mathbf{I}$].

Example 6

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix} \tag{13}$$

Adjoining $\boldsymbol{A}$ *with* $\boldsymbol{I}$ *yields*

$$\begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{bmatrix} \tag{14}$$

Adding -2 times the first row to the second row and -1 times the first row to the third yields

$$\begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{bmatrix} \tag{15}$$

Adding 2 times the 2^{nd} *row to the third yields*

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

Multiplying the third row by -1 yields

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

Adding 3 times the third row to the second and -3 times the third row to the first yields

$$\begin{bmatrix} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{bmatrix}$$

Finally, adding -2 times the second row to the first yields:

$$\begin{bmatrix} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{bmatrix}$$

Thus,

$$\mathbf{A}^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

3.5.1 Cramer's Rule

If $\mathbf{Ax} = \mathbf{b}$ is a system of n linear equations in n unknowns such that $det(\mathbf{A}) \neq 0$ then, the system has a unique solution which is given by

$$x_1 = \frac{det(A_1)}{det(A)},\; x_2 = \frac{det(A_2)}{det(A)},\; x_3 = \frac{det(A_3)}{det(A)} \tag{16}$$

where $\mathbf{A}_j$ is the matrix obtained by re placing the entries of the j^{th} column $\mathbf{A}$ by the entries in the vector $\mathbf{b}$, where

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{pmatrix} \tag{17}$$

Example 7 *Consider the following system:*

$$\begin{array}{rcrcrl} 2x_1 & + & 4x_2 & - & 2x_3 & = 18 \\ & & 2x_2 & + & 3x_3 & = -2 \\ x_1 & & & + & 5x_3 & = -7 \end{array}$$

solve for x_1, x_2 and x_3.

Solution: *Recasting the system in matrix-vector form, we have*

$$\begin{bmatrix} 2 & 4 & -2 \\ 0 & 2 & 3 \\ 1 & 0 & 5 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 18 \\ -2 \\ -7 \end{pmatrix}$$

Next, we define the three matrices formed by replacing in turn each of the columns of $\boldsymbol{A}$ *with* $\boldsymbol{b}$*:*

$$\mathbf{A_1} = \begin{bmatrix} 18 & 4 & -2 \\ -2 & 2 & 3 \\ -7 & 0 & 5 \end{bmatrix} \mathbf{A_2} = \begin{bmatrix} 2 & 18 & -2 \\ 0 & -2 & 3 \\ 1 & -7 & 5 \end{bmatrix}$$

$$\mathbf{A_3} = \begin{bmatrix} 2 & 4 & 18 \\ 0 & 2 & -2 \\ 1 & 0 & -7 \end{bmatrix}$$

Next, we compute the individual determinants:

$$\begin{aligned} det(\boldsymbol{A}) &= 2[2(5) - (-2(1))] - 3[2(0) - 4(1)] = 36 \\ det(\mathbf{A}_1) &= (-7)[4(3) - (-2)(2)] + 5[18(2) - 4(-2)] = 108 \\ det(\mathbf{A}_2) &= 2[-2(5) - (3)(-7)] + 1[18(3) - (-2)(-2)] = 72 \\ det(\mathbf{A}_3) &= 2[2(-7) - (-2)(0)] + 1[4(-2) - 18(2)] = -72 \end{aligned}$$

Thus,

$$x_1 = \frac{det(\mathbf{A}_1)}{det(\mathbf{A})} = \frac{108}{36} = 3, \; x_2 = \frac{det(\mathbf{A}_2)}{det(\mathbf{A})} = \frac{72}{36} = 2, \; x_3 = \frac{det(\mathbf{A}_3)}{det\mathbf{A})} = -\frac{72}{36} = -2 \tag{18}$$

Note that we took advantage of any zero elements by expanding the cofactors along the rows or columns that contained them.

Cramer's rule is particularly efficient to use on 2×2 system.

Consider a general 2nd-order example

Example 8

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{19}$$

$$det(A) = a_{11}a_{22} - a_{12}a_{21}, \; det(A_1) = b_1a_{22} - a_{12}b_2, \; det(A_2) = a_{11}b_2 - b_1a_{21} \tag{20}$$

Thus,

$$x_1 = \frac{det(A_1)}{det(A)} = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \; x_2 = \frac{det(A_2)}{det(A)} = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}} \tag{21}$$

3.6 The Characteristic Polynomial and Eigenvalues and Vectors

We have shown that the system

$$\mathbf{Ax} = \lambda\mathbf{x} \tag{22}$$

where λ is a scalar value. Equivalently, we may write

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \tag{23}$$

Hence, from Theorem (5) on page 7, there exists a nontrivial $\mathbf{x}$ if and only if

$$det(\mathbf{A} - \lambda\mathbf{I}) = 0 \tag{24}$$

Evaluation of the above results in a polynomial in λ. This is the so called *characteristic polynomial* and its roots λ_i are the *characteristic values* or *eigenvalues*. Evaluation of (24) yields,

$$\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2}\cdots + c_n = 0 \tag{25}$$

Furthermore, the solution of (23), $\mathbf{x}^i$, corresponding to the i^{th} eigenvalue, is the i^{th} eigenvector of the matrix $\mathbf{A}$. It can be shown that the matrix th$\mathbf{A}$ itself satisfies the characteristic polynomial.

$$\mathbf{A^n} + c_1\mathbf{A^{n-1}} + c_2\mathbf{A^{n-2}} + \cdots + c_n\mathbf{I} = 0$$

This result is known as *Cayley-Hamilton Theorem*. It may be shown that the matrix $\mathbf{A}$ is also anihilated by a minimum polynomial of degree less than or equal to that of the characteristic polynomial.

Example 9 *Find the eigenvalues and eigenvectors of*

$$\mathbf{A} = \begin{bmatrix} 4 & -5 \\ 1 & -2 \end{bmatrix}$$

Solution:

$$det\left(\begin{bmatrix} 4 & -5 \\ 1 & -2 \end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = det\begin{bmatrix} 4-\lambda & -5 \\ 1 & -2-\lambda \end{bmatrix}$$

$$= (4-\lambda)(-2-\lambda) + 5 = 0$$

$$\lambda^2 - 2\lambda - 3 = 0$$

or

$$(\lambda - 3)(\lambda + 1) = 0$$

Thus, $\lambda_1 = -1$*, and* $\lambda_2 = 3$ *The eigenvector of* $\boldsymbol{A}$ *corresponding to* $\lambda = -1$ *may be found as follows:*

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x}^1 = \begin{bmatrix} 4-(-1) & -5 \\ 1 & -2-(-1) \end{bmatrix}\begin{pmatrix} x_1^1 \\ x_2^1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

or

$$\begin{bmatrix} 5 & -5 \\ 1 & -1 \end{bmatrix}\begin{pmatrix} x_1^1 \\ x_2^1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which has an obvious solution of

$$\mathbf{x}^1 = c_1\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

where c_1 *is any scalar.*

Similarly, substituting $\lambda_2 = 3$ yields

$$\left[\begin{array}{cc} 1 & -5 \\ 1 & -5 \end{array}\right]\left(\begin{array}{c} x_1^2 \\ x_2^2 \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \end{array}\right)$$

Thus,

$$\mathbf{x}^2 = c_2\left(\begin{array}{c} 5 \\ 1 \end{array}\right)$$

The calculation of eigenvalues and eigenvectors has many applications. One important application is the similarity transformation. Under certain conditions, a general system of equations may be transformed into a diagonal system. The most important case is that of symmetric matrices which will be discussed later. In other words, the system $\mathbf{Ax} = \mathbf{b}$ may be transformed into an equivalent system $\mathbf{Dy} = \mathbf{c}$ where $\mathbf{D}$ is the diagonal matrix – making the solution of $\mathbf{Dy} = \mathbf{c}$ especially easy. Two matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be similar if $det(\mathbf{A}) = det(\mathbf{B})$. Another way of saying the above is if $\mathbf{A}$ and $\mathbf{B}$ are square matrices $\mathbf{B}$ is similar to $\mathbf{A}$ if and only if there is an invertible matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{PBP}^{-1}$. It turns out that an $n \times n$ matrix $\mathbf{A}$ is diagonalizable if $\mathbf{A}$ has n linearly independent eigenvectors.

Example 10 *Since the previous example had two independent eigenvectors (i.e., $\mathbf{x}^1 \neq K\mathbf{x}^2$ for any scalar K)*

$$\mathbf{A} = \left[\begin{array}{cc} 4 & -5 \\ 1 & -2 \end{array}\right] \tag{26}$$

should be diagonalizable. A matrix composed of $\mathbf{x}^1$ and $\mathbf{x}^2$ as its two columns will diagonalize $\boldsymbol{A}$. We show that this is so by trying it!

$$\mathbf{P}^{-1}\mathbf{AP} = \left[\begin{array}{cc} -\frac{1}{4} & \frac{5}{4} \\ \frac{1}{4} & -\frac{1}{4} \end{array}\right]\left[\begin{array}{cc} -4 & 5 \\ 1 & -2 \end{array}\right]\left[\begin{array}{cc} 1 & 5 \\ 1 & 1 \end{array}\right]$$

$$= \frac{1}{4}\left[\begin{array}{cc} -1 & 5 \\ 1 & -1 \end{array}\right]\left[\begin{array}{cc} -1 & 15 \\ 1 & -3 \end{array}\right] = \left[\begin{array}{cc} -1 & 0 \\ 0 & 3 \end{array}\right]$$

We see that $\boldsymbol{D}$ is composed of λ_1 and λ_2 on its main diagonal and zeros elsewhere. Hence, the matrix, $\boldsymbol{A}$, was indeed similar to a diagonal matrix.

3.7 Special Properties of Symmetric Matrices

Symmetric matrices have several special properties. The principal ones for an n by n symmetric matrix are enumerated below:

1. Symmetric real matrices (those with real elements) have n real eigenvalues.
2. Eigenvectors corresponding to distinct real roots are orthogonal.

   Proof 3 *If* $\mathbf{A} = \mathbf{A}^T$*, and*
   $$\mathbf{A}\mathbf{x}^1 = \lambda_1 \mathbf{x}^1 \tag{27}$$
   and
   $$\mathbf{A}\mathbf{x}^2 = \lambda_2 \mathbf{x}^2 \tag{28}$$
   Then,
   $$\lambda_1 \mathbf{x}^{2T}\mathbf{A}\mathbf{x}^1 = \lambda_1 \mathbf{x}^{2T}\mathbf{x}^1 \tag{29}$$
   and
   $$\lambda_2 \mathbf{x}^{1T}\mathbf{A}\mathbf{x}^2 = \lambda_2 \mathbf{x}^{1T}\mathbf{x}^2 \tag{30}$$
   Since $\mathbf{x}^{2T}\mathbf{x}^1 = \mathbf{x}^{1T}\mathbf{x}^2$ *and*
   $$\mathbf{x}^{2T}\mathbf{A}\mathbf{x}^1 = \mathbf{x}^{1T}\mathbf{A}\mathbf{x}^2 \tag{31}$$
   subtraction of (29) from (30) yields
   $$(\lambda_2 - \lambda_1)\,\mathbf{x}^{1T}\mathbf{x}^2 = 0 \tag{32}$$
   Hence, if $\lambda_1 \neq \lambda_2$*,* $\mathbf{x}^1$ *and* $\mathbf{x}^2$ *are orthogonal as claimed.* □

3. If λ_r is a root of the characteristic polynomial of algebraic multiplicity m, there exist m independent eigenvectors corresponding to λ_r.

One of the most important consequences of the above for symmetric matrices is that all symmetric matrices are similar to a diagonal matrix. This fact has powerful consequences in the solution of systems of linear ordinary differential equations with constant coefficients which result from the application of Newton's 2^{nd} law, or Hamilton's principle. Essentially, such systems, which usually result from symmetric operators, may be uncoupled by a similarity transformation, and hence, each ordinary differential equation solved individually. Exceptions to this rule include systems modeled with general viscous damping, and those with gyroscopic inertial terms.

The following examples were contributed by Dr. Geroid P. MacSithigh.

Example 11 *(symmetric)*

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Characteristic polynomial: $(\lambda - 4)^2(\lambda - 5) = 0$

$$\lambda_1 = 5,\ \lambda_2 = \lambda_3 = 4$$

$$\lambda_1 = 5$$

$$\mathbf{x}^1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

$$\lambda_2 = 4,\ \mathbf{x}^2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},\ \mathbf{x}^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{33}$$

Example 12 *(non-symmetric)*

$$\mathbf{A} = \begin{bmatrix} 5 & 0 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Characteristic polynomial: $(\lambda - 5)^2(\lambda - 3) = 0$

$$\lambda_1 = 3,\ \lambda_2 = \lambda_3 = 5$$

$$\lambda_1 = 3, \quad \mathbf{x}^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\lambda_2 = 5; \quad \mathbf{x}^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Thus, only one eigenvector corresponding to $\lambda = 5$.

4 Bibliography

1. Anton, Howard, “Elementary Linear Algebra,” 1977, *John Wiley&Sons*

2. Cullen, Charles G., “Matrices and Linear Transformations,” 1972, 2^{nd} Ed., *Addison-Wesley, Reprinted by Dover 1990.*

3. Strang, Gilbert, “Linear Algebra and Its Applications,” 1988, 3^{rd} Ed., *International Thomson Publishing.*