lyryx with Open Texts

LINEAR ALGEBRA with Applications

Open Edition

Adapted for

Emory University

Math 221

Linear Algebra

Sections 1 & 2

Lectured and adapted by

Le Chen

April 15, 2021

le.chen@emory.edu

Course page

http://math.emory.edu/~lchen41/teaching/2021_Spring_Math221

ADAPTABLE | ACCESSIBLE | AFFORDABLE

by W. Keith Nicholson

Contents

3. Determinants and Diagonalization

Contents

With each square matrix we can calculate a number, called the determinant of the matrix, which tells us whether or not the matrix is invertible. In fact, determinants can be used to give a formula for the inverse of a matrix. They also arise in calculating certain numbers (called eigenvalues) associated with the matrix. These eigenvalues are essential to a technique called diagonalization that is used in many applications where it is desired to predict the future behaviour of a system. For example, we use it to predict whether a species will become extinct.

Determinants were first studied by Leibnitz in 1696, and the term "determinant" was first used in 1801 by Gauss is his *Disquisitiones Arithmeticae*. Determinants are much older than matrices (which were introduced by Cayley in 1878) and were used extensively in the eighteenth and nineteenth centuries, primarily because of their significance in geometry (see Section 4.4). Although they are somewhat less important today, determinants still play a role in the theory and application of matrix algebra.

3.1 The Cofactor Expansion

In Section 2.4 we defined the determinant of a 2×2 matrix $A = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]$ as follows:[1]

$$\det A = \left| \begin{array}{cc} a & b \\ c & d \end{array} \right| = ad - bc$$

and showed (in Example 2.4.4) that A has an inverse if and only if $\det A \neq 0$. One objective of this chapter is to do this for *any* square matrix A. There is no difficulty for 1×1 matrices: If $A = [a]$, we define $\det A = \det [a] = a$ and note that A is invertible if and only if $a \neq 0$.

If A is 3×3 and invertible, we look for a suitable definition of $\det A$ by trying to carry A to the identity matrix by row operations. The first column is not zero (A is invertible); suppose the (1, 1)-entry a is not zero. Then row operations give

$$A = \left[\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array} \right] \rightarrow \left[\begin{array}{ccc} a & b & c \\ ad & ae & af \\ ag & ah & ai \end{array} \right] \rightarrow \left[\begin{array}{ccc} a & b & c \\ 0 & ae-bd & af-cd \\ 0 & ah-bg & ai-cg \end{array} \right] = \left[\begin{array}{ccc} a & b & c \\ 0 & u & af-cd \\ 0 & v & ai-cg \end{array} \right]$$

where $u = ae - bd$ and $v = ah - bg$. Since A is invertible, one of u and v is nonzero (by Example 2.4.11); suppose that $u \neq 0$. Then the reduction proceeds

$$A \rightarrow \left[\begin{array}{ccc} a & b & c \\ 0 & u & af-cd \\ 0 & v & ai-cg \end{array} \right] \rightarrow \left[\begin{array}{ccc} a & b & c \\ 0 & u & af-cd \\ 0 & uv & u(ai-cg) \end{array} \right] \rightarrow \left[\begin{array}{ccc} a & b & c \\ 0 & u & af-cd \\ 0 & 0 & w \end{array} \right]$$

where $w = u(ai - cg) - v(af - cd) = a(aei + bfg + cdh - ceg - afh - bdi)$. We define

$$\det A = aei + bfg + cdh - ceg - afh - bdi \tag{3.1}$$

and observe that $\det A \neq 0$ because $a \det A = w \neq 0$ (is invertible).

To motivate the definition below, collect the terms in Equation 3.1 involving the entries a, b, and c in row 1 of A:

$$\begin{aligned} \det A = \left| \begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array} \right| &= aei + bfg + cdh - ceg - afh - bdi \\ &= a(ei - fh) - b(di - fg) + c(dh - eg) \\ &= a \left| \begin{array}{cc} e & f \\ h & i \end{array} \right| - b \left| \begin{array}{cc} d & f \\ g & i \end{array} \right| + c \left| \begin{array}{cc} d & e \\ g & h \end{array} \right| \end{aligned}$$

This last expression can be described as follows: To compute the determinant of a 3×3 matrix A, multiply each entry in row 1 by a sign times the determinant of the 2×2 matrix obtained by deleting the row and column of that entry, and add the results. The signs alternate down row 1, starting with $+$. It is this observation that we generalize below.

[1]Determinants are commonly written $|A| = \det A$ using vertical bars. We will use both notations.

Example 3.1.1

$$\det\begin{bmatrix} 2 & 3 & 7 \\ -4 & 0 & 6 \\ 1 & 5 & 0 \end{bmatrix} = 2\begin{vmatrix} 0 & 6 \\ 5 & 0 \end{vmatrix} - 3\begin{vmatrix} -4 & 6 \\ 1 & 0 \end{vmatrix} + 7\begin{vmatrix} -4 & 0 \\ 1 & 5 \end{vmatrix}$$
$$= 2(-30) - 3(-6) + 7(-20)$$
$$= -182$$

This suggests an inductive method of defining the determinant of any square matrix in terms of determinants of matrices one size smaller. The idea is to define determinants of 3×3 matrices in terms of determinants of 2×2 matrices, then we do 4×4 matrices in terms of 3×3 matrices, and so on.

To describe this, we need some terminology.

Definition 3.1 Cofactors of a Matrix

Assume that determinants of $(n-1)\times(n-1)$ matrices have been defined. Given the $n\times n$ matrix A, let

A_{ij} denote the $(n-1)\times(n-1)$ matrix obtained from A by deleting row i and column j.

*Then the $(i,\ j)$-**cofactor** $c_{ij}(A)$ is the scalar defined by*

$$c_{ij}(A) = (-1)^{i+j}\det(A_{ij})$$

*Here $(-1)^{i+j}$ is called the **sign** of the $(i,\ j)$-position.*

The sign of a position is clearly 1 or -1, and the following diagram is useful for remembering it:

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}$$

Note that the signs alternate along each row and column with $+$ in the upper left corner.

Example 3.1.2

Find the cofactors of positions $(1,\ 2)$, $(3,\ 1)$, and $(2,\ 3)$ in the following matrix.

$$A = \begin{bmatrix} 3 & -1 & 6 \\ 5 & 2 & 7 \\ 8 & 9 & 4 \end{bmatrix}$$

Solution. Here A_{12} is the matrix $\begin{bmatrix} 5 & 7 \\ 8 & 4 \end{bmatrix}$ that remains when row 1 and column 2 are deleted. The sign of position $(1,\ 2)$ is $(-1)^{1+2} = -1$ (this is also the $(1,\ 2)$-entry in the sign diagram), so the $(1,\ 2)$-cofactor is

$$c_{12}(A) = (-1)^{1+2} \begin{vmatrix} 5 & 7 \\ 8 & 4 \end{vmatrix} = (-1)(5\cdot 4 - 7\cdot 8) = (-1)(-36) = 36$$

Turning to position $(3,\ 1)$, we find

$$c_{31}(A) = (-1)^{3+1} A_{31} = (-1)^{3+1} \begin{vmatrix} -1 & 6 \\ 2 & 7 \end{vmatrix} = (+1)(-7-12) = -19$$

Finally, the $(2,\ 3)$-cofactor is

$$c_{23}(A) = (-1)^{2+3} A_{23} = (-1)^{2+3} \begin{vmatrix} 3 & -1 \\ 8 & 9 \end{vmatrix} = (-1)(27+8) = -35$$

Clearly other cofactors can be found—there are nine in all, one for each position in the matrix.

We can now define $\det A$ for any square matrix A

Definition 3.2 Cofactor expansion of a Matrix

Assume that determinants of $(n-1)\times(n-1)$ matrices have been defined. If $A = \left[a_{ij}\right]$ is $n\times n$ define

$$\det A = a_{11}c_{11}(A) + a_{12}c_{12}(A) + \cdots + a_{1n}c_{1n}(A)$$

*This is called the **cofactor expansion** of $\det A$ along row 1.*

It asserts that $\det A$ can be computed by multiplying the entries of row 1 by the corresponding cofactors, and adding the results. The astonishing thing is that $\det A$ can be computed by taking the cofactor expansion along *any row or column*: Simply multiply each entry of that row or column by the corresponding cofactor and add.

Theorem 3.1.1: Cofactor Expansion Theorem[2]

The determinant of an $n\times n$ matrix A can be computed by using the cofactor expansion along any row or column of A. That is $\det A$ can be computed by multiplying each entry of the row or column by the corresponding cofactor and adding the results.

The proof will be given in Section **??**.

[2]The cofactor expansion is due to Pierre Simon de Laplace (1749–1827), who discovered it in 1772 as part of a study of linear differential equations. Laplace is primarily remembered for his work in astronomy and applied mathematics.

Example 3.1.3

Compute the determinant of $A = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 7 & 2 \\ 9 & 8 & -6 \end{bmatrix}$.

Solution. The cofactor expansion along the first row is as follows:

$$\begin{aligned} \det A &= 3c_{11}(A) + 4c_{12}(A) + 5c_{13}(A) \\ &= 3\begin{vmatrix} 7 & 2 \\ 8 & -6 \end{vmatrix} - 4\begin{vmatrix} 1 & 2 \\ 9 & -6 \end{vmatrix} + 3\begin{vmatrix} 1 & 7 \\ 9 & 8 \end{vmatrix} \\ &= 3(-58) - 4(-24) + 5(-55) \\ &= -353 \end{aligned}$$

Note that the signs alternate along the row (indeed along *any* row or column). Now we compute $\det A$ by expanding along the first column.

$$\begin{aligned} \det A &= 3c_{11}(A) + 1c_{21}(A) + 9c_{31}(A) \\ &= 3\begin{vmatrix} 7 & 2 \\ 8 & -6 \end{vmatrix} - \begin{vmatrix} 4 & 5 \\ 8 & -6 \end{vmatrix} + 9\begin{vmatrix} 4 & 5 \\ 7 & 2 \end{vmatrix} \\ &= 3(-58) - (-64) + 9(-27) \\ &= -353 \end{aligned}$$

The reader is invited to verify that $\det A$ can be computed by expanding along any other row or column.

The fact that the cofactor expansion along *any row or column* of a matrix A always gives the same result (the determinant of A) is remarkable, to say the least. The choice of a particular row or column can simplify the calculation.

Example 3.1.4

Compute $\det A$ where $A = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 5 & 1 & 2 & 0 \\ 2 & 6 & 0 & -1 \\ -6 & 3 & 1 & 0 \end{bmatrix}$.

Solution. The first choice we must make is which row or column to use in the cofactor expansion. The expansion involves multiplying entries by cofactors, so the work is minimized when the row or column contains as many zero entries as possible. Row 1 is a best choice in this matrix (column 4 would do as well), and the expansion is

$$\begin{aligned} \det A &= 3c_{11}(A) + 0c_{12}(A) + 0c_{13}(A) + 0c_{14}(A) \\ &= 3\begin{vmatrix} 1 & 2 & 0 \\ 6 & 0 & -1 \\ 3 & 1 & 0 \end{vmatrix} \end{aligned}$$

This is the first stage of the calculation, and we have succeeded in expressing the determinant of the 4×4 matrix A in terms of the determinant of a 3×3 matrix. The next stage involves this 3×3 matrix. Again, we can use any row or column for the cofactor expansion. The third column is preferred (with two zeros), so

$$\begin{aligned} \det A &= 3\left(0\begin{vmatrix} 6 & 0 \\ 3 & 1 \end{vmatrix} - (-1)\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} + 0\begin{vmatrix} 1 & 2 \\ 6 & 0 \end{vmatrix}\right) \\ &= 3[0 + 1(-5) + 0] \\ &= -15 \end{aligned}$$

This completes the calculation.

Computing the determinant of a matrix A can be tedious. For example, if A is a 4×4 matrix, the cofactor expansion along any row or column involves calculating four cofactors, each of which involves the determinant of a 3×3 matrix. And if A is 5×5, the expansion involves five determinants of 4×4 matrices! There is a clear need for some techniques to cut down the work.[3]

The motivation for the method is the observation (see Example 3.1.4) that calculating a determinant is simplified a great deal when a row or column consists mostly of zeros. (In fact, when a row or column consists *entirely* of zeros, the determinant is zero—simply expand along that row or column.)

Recall next that one method of *creating* zeros in a matrix is to apply elementary row operations to it. Hence, a natural question to ask is what effect such a row operation has on the determinant of the matrix. It turns out that the effect is easy to determine and that elementary *column* operations can be used in the same way. These observations lead to a technique for evaluating determinants that greatly reduces the labour involved. The necessary information is given in Theorem 3.1.2.

Theorem 3.1.2

Let A denote an $n \times n$ matrix.

1. *If A has a row or column of zeros,* $\det A = 0$.
2. *If two distinct rows (or columns) of A are interchanged, the determinant of the resulting matrix is* $-\det A$.
3. *If a row (or column) of A is multiplied by a constant u, the determinant of the resulting matrix is* $u(\det A)$.
4. *If two distinct rows (or columns) of A are identical,* $\det A = 0$.

[3]If $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ we can calculate $\det A$ by considering $\begin{bmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{bmatrix}$ obtained from A by adjoining columns 1 and 2 on the right. Then $\det A = aei + bfg + cdh - ceg - afh - bdi$, where the positive terms aei, bfg, and cdh are the products down and to the right starting at a, b, and c, and the negative terms ceg, afh, and bdi are the products down and to the left starting at c, a, and b. **Warning**: This rule does **not** apply to $n \times n$ matrices where $n > 3$ or $n = 2$.

5. *If a multiple of one row of A is added to a different row (or if a multiple of a column is added to a different column), the determinant of the resulting matrix is* $\det A$.

Proof. We prove properties 2, 4, and 5 and leave the rest as exercises.

Property 2. If A is $n \times n$, this follows by induction on n. If $n = 2$, the verification is left to the reader. If $n > 2$ and two rows are interchanged, let B denote the resulting matrix. Expand $\det A$ and $\det B$ along a row *other than* the two that were interchanged. The entries in this row are the same for both A and B, but the cofactors in B are the negatives of those in A (by induction) because the corresponding $(n-1) \times (n-1)$ matrices have two rows interchanged. Hence, $\det B = -\det A$, as required. A similar argument works if two columns are interchanged.

Property 4. If two rows of A are equal, let B be the matrix obtained by interchanging them. Then $B = A$, so $\det B = detA$. But $\det B = -\det A$ by property 2, so $\det A = \det B = 0$. Again, the same argument works for columns.

Property 5. Let B be obtained from $A = \left[a_{ij}\right]$ by adding u times row p to row q. Then row q of B is

$$(a_{q1} + ua_{p1},\ a_{q2} + ua_{p2},\ \dots,\ a_{qn} + ua_{pn})$$

The cofactors of these elements in B are the same as in A (they do not involve row q): in symbols, $c_{qj}(B) = c_{qj}(A)$ for each j. Hence, expanding B along row q gives

$$\begin{aligned}
\det A &= (a_{q1} + ua_{p1})c_{q1}(A) + (a_{q2} + ua_{p2})c_{q2}(A) + \cdots + (a_{qn} + ua_{pn})c_{qn}(A) \\
&= [a_{q1}c_{q1}(A) + a_{q2}c_{q2}(A) + \cdots + a_{qn}c_{qn}(A)] + u[a_{p1}c_{q1}(A) + a_{p2}c_{q2}(A) + \cdots + a_{pn}c_{qn}(A)] \\
&= \det A + u \det C
\end{aligned}$$

where C is the matrix obtained from A by replacing row q by row p (and both expansions are along row q). Because rows p and q of C are equal, $\det C = 0$ by property 4. Hence, $\det B = detA$, as required. As before, a similar proof holds for columns. □

To illustrate Theorem 3.1.2, consider the following determinants.

$$\begin{vmatrix} 3 & -1 & 2 \\ 2 & 5 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 0$$ (because the last row consists of zeros)

$$\begin{vmatrix} 3 & -1 & 5 \\ 2 & 8 & 7 \\ 1 & 2 & -1 \end{vmatrix} = -\begin{vmatrix} 5 & -1 & 3 \\ 7 & 8 & 2 \\ -1 & 2 & 1 \end{vmatrix}$$ (because two columns are interchanged)

$$\begin{vmatrix} 8 & 1 & 2 \\ 3 & 0 & 9 \\ 1 & 2 & -1 \end{vmatrix} = 3\begin{vmatrix} 8 & 1 & 2 \\ 1 & 0 & 3 \\ 1 & 2 & -1 \end{vmatrix}$$ (because the second row of the matrix on the left is 3 times the second row of the matrix on the right)

$$\begin{vmatrix} 2 & 1 & 2 \\ 4 & 0 & 4 \\ 1 & 3 & 1 \end{vmatrix} = 0$$ (because two columns are identical)

$$\begin{vmatrix} 2 & 5 & 2 \\ -1 & 2 & 9 \\ 3 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 9 & 20 \\ -1 & 2 & 9 \\ 3 & 1 & 1 \end{vmatrix}$$ (because twice the second row of the matrix on the left was added to the first row)

The following four examples illustrate how Theorem 3.1.2 is used to evaluate determinants.

Example 3.1.5

Evaluate $\det A$ when $A = \begin{bmatrix} 1 & -1 & 3 \\ 1 & 0 & -1 \\ 2 & 1 & 6 \end{bmatrix}$.

Solution. The matrix does have zero entries, so expansion along (say) the second row would involve somewhat less work. However, a column operation can be used to get a zero in position (2, 3)—namely, add column 1 to column 3. Because this does not change the value of the determinant, we obtain

$$\det A = \begin{vmatrix} 1 & -1 & 3 \\ 1 & 0 & -1 \\ 2 & 1 & 6 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 4 \\ 1 & 0 & 0 \\ 2 & 1 & 8 \end{vmatrix} = -\begin{vmatrix} -1 & 4 \\ 1 & 8 \end{vmatrix} = 12$$

where we expanded the second 3×3 matrix along row 2.

Example 3.1.6

If $\det \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix} = 6$, evaluate $\det A$ where $A = \begin{bmatrix} a+x & b+y & c+z \\ 3x & 3y & 3z \\ -p & -q & -r \end{bmatrix}$.

Solution. First take common factors out of rows 2 and 3.

$$\det A = 3(-1) \det \begin{bmatrix} a+x & b+y & c+z \\ x & y & z \\ p & q & r \end{bmatrix}$$

Now subtract the second row from the first and interchange the last two rows.

$$\det A = -3 \det \begin{bmatrix} a & b & c \\ x & y & z \\ p & q & r \end{bmatrix} = 3 \det \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix} = 3 \cdot 6 = 18$$

The determinant of a matrix is a sum of products of its entries. In particular, if these entries are polynomials in x, then the determinant itself is a polynomial in x. It is often of interest to determine which values of x make the determinant zero, so it is very useful if the determinant is given in factored form. Theorem 3.1.2 can help.

Example 3.1.7

Find the values of x for which $\det A = 0$, where $A = \begin{bmatrix} 1 & x & x \\ x & 1 & x \\ x & x & 1 \end{bmatrix}$.

Solution. To evaluate $\det A$, first subtract x times row 1 from rows 2 and 3.

$$\det A = \begin{vmatrix} 1 & x & x \\ x & 1 & x \\ x & x & 1 \end{vmatrix} = \begin{vmatrix} 1 & x & x \\ 0 & 1-x^2 & x-x^2 \\ 0 & x-x^2 & 1-x^2 \end{vmatrix} = \begin{vmatrix} 1-x^2 & x-x^2 \\ x-x^2 & 1-x^2 \end{vmatrix}$$

At this stage we could simply evaluate the determinant (the result is $2x^3-3x^2+1$). But then we would have to factor this polynomial to find the values of x that make it zero. However, this factorization can be obtained directly by first factoring each entry in the determinant and taking a common factor of $(1-x)$ from each row.

$$\begin{aligned} \det A = \begin{vmatrix} (1-x)(1+x) & x(1-x) \\ x(1-x) & (1-x)(1+x) \end{vmatrix} &= (1-x)^2 \begin{vmatrix} 1+x & x \\ x & 1+x \end{vmatrix} \\ &= (1-x)^2(2x+1) \end{aligned}$$

Hence, $\det A = 0$ means $(1-x)^2(2x+1) = 0$, that is $x = 1$ or $x = -\frac{1}{2}$.

Example 3.1.8

If a_1, a_2, and a_3 are given show that

$$\det \begin{bmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{bmatrix} = (a_3 - a_1)(a_3 - a_2)(a_2 - a_1)$$

Solution. Begin by subtracting row 1 from rows 2 and 3, and then expand along column 1:

$$\det \begin{bmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{bmatrix} = \det \begin{bmatrix} 1 & a_1 & a_1^2 \\ 0 & a_2 - a_1 & a_2^2 - a_1^2 \\ 0 & a_3 - a_1 & a_3^2 - a_1^2 \end{bmatrix} = \begin{bmatrix} a_2 - a_1 & a_2^2 - a_1^2 \\ a_3 - a_1 & a_3^2 - a_1^2 \end{bmatrix}$$

Now $(a_2 - a_1)$ and $(a_3 - a_1)$ are common factors in rows 1 and 2, respectively, so

$$\begin{aligned} \det \begin{bmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{bmatrix} &= (a_2 - a_1)(a_3 - a_1) \det \begin{bmatrix} 1 & a_2 + a_1 \\ 1 & a_3 + a_1 \end{bmatrix} \\ &= (a_2 - a_1)(a_3 - a_1)(a_3 - a_2) \end{aligned}$$

The matrix in Example 3.1.8 is called a Vandermonde matrix, and the formula for its determinant can be generalized to the $n \times n$ case (see Theorem 3.2.7).

If A is an $n \times n$ matrix, forming uA means multiplying *every* row of A by u. Applying property 3 of Theorem 3.1.2, we can take the common factor u out of each row and so obtain the following useful result.

Theorem 3.1.3

If A is an $n \times n$ matrix, then $\det(uA) = u^n \det A$ for any number u.

The next example displays a type of matrix whose determinant is easy to compute.

Example 3.1.9

Evaluate $\det A$ if $A = \begin{bmatrix} a & 0 & 0 & 0 \\ u & b & 0 & 0 \\ v & w & c & 0 \\ x & y & z & d \end{bmatrix}$.

Solution. Expand along row 1 to get $\det A = a \begin{vmatrix} b & 0 & 0 \\ w & c & 0 \\ y & z & d \end{vmatrix}$. Now expand this along the top row to get $\det A = ab \begin{vmatrix} c & 0 \\ z & d \end{vmatrix} = abcd$, the product of the main diagonal entries.

A square matrix is called a **lower triangular matrix** if all entries above the main diagonal are zero (as in Example 3.1.9). Similarly, an **upper triangular matrix** is one for which all entries below the main diagonal are zero. A **triangular matrix** is one that is either upper or lower triangular. Theorem 3.1.4 gives an easy rule for calculating the determinant of any triangular matrix. The proof is like the solution to Example 3.1.9.

Theorem 3.1.4

If A is a square triangular matrix, then det A is the product of the entries on the main diagonal.

Theorem 3.1.4 is useful in computer calculations because it is a routine matter to carry a matrix to triangular form using row operations.

Block matrices such as those in the next theorem arise frequently in practice, and the theorem gives an easy method for computing their determinants. This dovetails with Example 2.4.11.

Theorem 3.1.5

Consider matrices $\left[\begin{array}{cc} A & X \\ 0 & B \end{array}\right]$ *and* $\left[\begin{array}{cc} A & 0 \\ Y & B \end{array}\right]$ *in block form, where A and B are square matrices. Then*

$$\det \left[\begin{array}{cc} A & X \\ 0 & B \end{array}\right] = \det A \det B \text{ and } \det \left[\begin{array}{cc} A & 0 \\ Y & B \end{array}\right] = \det A \det B$$

Proof. Write $T = \det \left[\begin{array}{cc} A & X \\ 0 & B \end{array}\right]$ and proceed by induction on k where A is $k \times k$. If $k=1$, it is the cofactor expansion along column 1. In general let $S_i(T)$ denote the matrix obtained from T by deleting row i and column 1. Then the cofactor expansion of $\det T$ along the first column is

$$\det T = a_{11} \det(S_1(T)) - a_{21} \det(S_2(T)) + \cdots \pm a_{k1} \det(S_k(T)) \tag{3.2}$$

where $a_{11}, a_{21}, \cdots, a_{k1}$ are the entries in the first column of A. But $S_i(T) = \left[\begin{array}{cc} S_i(A) & X_i \\ 0 & B \end{array}\right]$ for each $i = 1, 2, \cdots, k$, so $\det(S_i(T)) = \det(S_i(A)) \cdot \det B$ by induction. Hence, Equation 3.2 becomes

$$\begin{aligned} \det T &= \{a_{11} \det(S_1(T)) - a_{21} \det(S_2(T)) + \cdots \pm a_{k1} \det(S_k(T))\} \det B \\ &= \{\det A\} \det B \end{aligned}$$

as required. The lower triangular case is similar. □

Example 3.1.10

$$\det\begin{bmatrix} 2 & 3 & 1 & 3 \\ 1 & -2 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 4 & 0 & 1 \end{bmatrix} = -\begin{vmatrix} 2 & 1 & 3 & 3 \\ 1 & -1 & -2 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 4 & 1 \end{vmatrix} = -\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix}\begin{vmatrix} 1 & 1 \\ 4 & 1 \end{vmatrix} = -(-3)(-3) = -9$$

The next result shows that $\det A$ is a linear transformation when regarded as a function of a fixed column of A. The proof is Exercise 3.1.21.

Theorem 3.1.6

Given columns $\mathbf{c}_1, \cdots, \mathbf{c}_{j-1}, \mathbf{c}_{j+1}, \cdots, \mathbf{c}_n$ *in* $\mathbb{R}^n$, *define* $T : \mathbb{R}^n \to \mathbb{R}$ *by*

$$T(\mathbf{x}) = \det\begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_{j-1} & \mathbf{x} & \mathbf{c}_{j+1} & \cdots & \mathbf{c}_n \end{bmatrix} \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

Then, for all $\mathbf{x}$ *and* $\mathbf{y}$ *in* $\mathbb{R}^n$ *and all* a *in* $\mathbb{R}$,

$$T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) \quad \text{and} \quad T(a\mathbf{x}) = aT(\mathbf{x})$$

Exercises for 3.1

Exercise 3.1.1 Compute the determinants of the following matrices.

a) $\begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}$

b) $\begin{bmatrix} 6 & 9 \\ 8 & 12 \end{bmatrix}$

c) $\begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}$

d) $\begin{bmatrix} a+1 & a \\ a & a-1 \end{bmatrix}$

e) $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

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

g) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

h) $\begin{bmatrix} 0 & a & 0 \\ b & c & d \\ 0 & e & 0 \end{bmatrix}$

i) $\begin{bmatrix} 1 & b & c \\ b & c & 1 \\ c & 1 & b \end{bmatrix}$

j) $\begin{bmatrix} 0 & a & b \\ a & 0 & c \\ b & c & 0 \end{bmatrix}$

k) $\begin{bmatrix} 0 & 1 & -1 & 0 \\ 3 & 0 & 0 & 2 \\ 0 & 1 & 2 & 1 \\ 5 & 0 & 0 & 7 \end{bmatrix}$

l) $\begin{bmatrix} 1 & 0 & 3 & 1 \\ 2 & 2 & 6 & 0 \\ -1 & 0 & -3 & 1 \\ 4 & 1 & 12 & 0 \end{bmatrix}$

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

n) $\begin{bmatrix} 4 & -1 & 3 & -1 \\ 3 & 1 & 0 & 2 \\ 0 & 1 & 2 & 2 \\ 1 & 2 & -1 & 1 \end{bmatrix}$

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

p) $\begin{bmatrix} 0 & 0 & 0 & a \\ 0 & 0 & b & p \\ 0 & c & q & k \\ d & s & t & u \end{bmatrix}$

---

b. 0

d. -1

f. -39

h. 0

j. $2abc$

l. 0

n. -56

p. $abcd$

Exercise 3.1.2 Show that $\det A = 0$ if A has a row or column consisting of zeros.

Exercise 3.1.3 Show that the sign of the position in the last row and the last column of A is always $+1$.

Exercise 3.1.4 Show that $\det I = 1$ for any identity matrix I.

Exercise 3.1.5 Evaluate the determinant of each matrix by reducing it to upper triangular form.

a) $\left[\begin{array}{rrr} 1 & -1 & 2 \\ 3 & 1 & 1 \\ 2 & -1 & 3 \end{array}\right]$ b) $\left[\begin{array}{rrr} -1 & 3 & 1 \\ 2 & 5 & 3 \\ 1 & -2 & 1 \end{array}\right]$

c) $\left[\begin{array}{rrrr} -1 & -1 & 1 & 0 \\ 2 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 1 & 3 & -1 & 2 \end{array}\right]$ d) $\left[\begin{array}{rrrr} 2 & 3 & 1 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 5 & 1 & 1 \\ 1 & 1 & 2 & 5 \end{array}\right]$

---

b. -17

d. 106

Exercise 3.1.6 Evaluate by cursory inspection:

a. $\det \left[\begin{array}{ccc} a & b & c \\ a+1 & b+1 & c+1 \\ a-1 & b-1 & c-1 \end{array}\right]$

b. $\det \left[\begin{array}{ccc} a & b & c \\ a+b & 2b & c+b \\ 2 & 2 & 2 \end{array}\right]$

---

b. 0

Exercise 3.1.7 If $\det \left[\begin{array}{ccc} a & b & c \\ p & q & r \\ x & y & z \end{array}\right] = -1$ compute:

a. $\det \left[\begin{array}{ccc} -x & -y & -z \\ 3p+a & 3q+b & 3r+c \\ 2p & 2q & 2r \end{array}\right]$

b. $\det \left[\begin{array}{ccc} -2a & -2b & -2c \\ 2p+x & 2q+y & 2r+z \\ 3x & 3y & 3z \end{array}\right]$

---

b. 12

Exercise 3.1.8 Show that:

a. $\det \left[\begin{array}{ccc} p+x & q+y & r+z \\ a+x & b+y & c+z \\ a+p & b+q & c+r \end{array}\right] = 2 \det \left[\begin{array}{ccc} a & b & c \\ p & q & r \\ x & y & z \end{array}\right]$

b. $\det \left[\begin{array}{ccc} 2a+p & 2b+q & 2c+r \\ 2p+x & 2q+y & 2r+z \\ 2x+a & 2y+b & 2z+c \end{array}\right] = 9 \det \left[\begin{array}{ccc} a & b & c \\ p & q & r \\ x & y & z \end{array}\right]$

---

b.
$$\begin{aligned}
&\det \left[\begin{array}{ccc} 2a+p & 2b+q & 2c+r \\ 2p+x & 2q+y & 2r+z \\ 2x+a & 2y+b & 2z+c \end{array}\right] \\
&= 3 \det \left[\begin{array}{ccc} a+p+x & b+q+y & c+r+z \\ 2p+x & 2q+y & 2r+z \\ 2x+a & 2y+b & 2z+c \end{array}\right] \\
&= 3 \det \left[\begin{array}{ccc} a+p+x & b+q+y & c+r+z \\ p-a & q-b & r-c \\ x-p & y-q & z-r \end{array}\right] \\
&= 3 \det \left[\begin{array}{ccc} 3x & 3y & 3z \\ p-a & q-b & r-c \\ x-p & y-q & z-r \end{array}\right] \cdots
\end{aligned}$$

Exercise 3.1.9 In each case either prove the statement or give an example showing that it is false:

a. $\det(A+B) = \det A + \det B$.

b. If $\det A = 0$, then A has two equal rows.

c. If A is 2×2, then $\det(A^T) = \det A$.

d. If R is the reduced row-echelon form of A, then $\det A = \det R$.

e. If A is 2×2, then $\det(7A) = 49 \det A$.

f. $\det(A^T) = -\det A$.

g. $\det(-A) = -\det A$.

h. If $\det A = \det B$ where A and B are the same size, then $A = B$.

---

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

d. False. $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \to R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

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

h. False. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

Exercise 3.1.10 Compute the determinant of each matrix, using Theorem 3.1.5.

a. $\begin{bmatrix} 1 & -1 & 2 & 0 & -2 \\ 0 & 1 & 0 & 4 & 1 \\ 1 & 1 & 5 & 0 & 0 \\ 0 & 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$

b. $\begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ -1 & 3 & 1 & 4 & 0 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & -1 & 0 & 2 \\ 0 & 0 & 3 & 0 & 1 \end{bmatrix}$

---

b. 35

Exercise 3.1.11 If $\det A = 2$, $\det B = -1$, and $\det C = 3$, find:

a) $\det \begin{bmatrix} A & X & Y \\ 0 & B & Z \\ 0 & 0 & C \end{bmatrix}$

b) $\det \begin{bmatrix} A & 0 & 0 \\ X & B & 0 \\ Y & Z & C \end{bmatrix}$

c) $\det \begin{bmatrix} A & X & Y \\ 0 & B & 0 \\ 0 & Z & C \end{bmatrix}$

d) $\det \begin{bmatrix} A & X & 0 \\ 0 & B & 0 \\ Y & Z & C \end{bmatrix}$

---

b. -6

d. -6

Exercise 3.1.12 If A has three columns with only the top two entries nonzero, show that $\det A = 0$.

Exercise 3.1.13

a. Find $\det A$ if A is 3×3 and $\det(2A) = 6$.

b. Under what conditions is $\det(-A) = \det A$?

Exercise 3.1.14 Evaluate by first adding all other rows to the first row.

a. $\det \begin{bmatrix} x-1 & 2 & 3 \\ 2 & -3 & x-2 \\ -2 & x & -2 \end{bmatrix}$

b. $\det \begin{bmatrix} x-1 & -3 & 1 \\ 2 & -1 & x-1 \\ -3 & x+2 & -2 \end{bmatrix}$

---

b. $-(x-2)(x^2+2x-12)$

Exercise 3.1.15

a. Find b if $\det \begin{bmatrix} 5 & -1 & x \\ 2 & 6 & y \\ -5 & 4 & z \end{bmatrix} = ax + by + cz$.

b. Find c if $\det \begin{bmatrix} 2 & x & -1 \\ 1 & y & 3 \\ -3 & z & 4 \end{bmatrix} = ax + by + cz$.

---

b. -7

Exercise 3.1.16 Find the real numbers x and y such that $\det A = 0$ if:

a) $A = \begin{bmatrix} 0 & x & y \\ y & 0 & x \\ x & y & 0 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & x & x \\ -x & -2 & x \\ -x & -x & -3 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & x & x^2 & x^3 \\ x & x^2 & x^3 & 1 \\ x^2 & x^3 & 1 & x \\ x^3 & 1 & x & x^2 \end{bmatrix}$

d) $A = \begin{bmatrix} x & y & 0 & 0 \\ 0 & x & y & 0 \\ 0 & 0 & x & y \\ y & 0 & 0 & x \end{bmatrix}$

---

b. $\pm\frac{\sqrt{6}}{2}$

d. $x = \pm y$

Exercise 3.1.17 Show that

$$\det \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & x & x \\ 1 & x & 0 & x \\ 1 & x & x & 0 \end{bmatrix} = -3x^2$$

Exercise 3.1.18 Show that

$$\det \begin{bmatrix} 1 & x & x^2 & x^3 \\ a & 1 & x & x^2 \\ p & b & 1 & x \\ q & r & c & 1 \end{bmatrix} = (1-ax)(1-bx)(1-cx).$$

Exercise 3.1.19
Given the polynomial $p(x) = a + bx + cx^2 + dx^3 + x^4$, the matrix $C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a & -b & -c & -d \end{bmatrix}$ is called the **companion matrix** of $p(x)$. Show that $\det(xI - C) = p(x)$.

Exercise 3.1.20 Show that

$$\det \begin{bmatrix} a+x & b+x & c+x \\ b+x & c+x & a+x \\ c+x & a+x & b+x \end{bmatrix}$$
$$= (a+b+c+3x)[(ab+ac+bc) - (a^2+b^2+c^2)]$$

Exercise 3.1.21 . Prove Theorem 3.1.6. [*Hint*: Expand the determinant along column j.]

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ and $A = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{x}+\mathbf{y} & \cdots & \mathbf{c}_n \end{bmatrix}$ where $\mathbf{x}+\mathbf{y}$ is in column j. Expanding $\det A$ along column j (the one containing $\mathbf{x}+\mathbf{y}$):

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) = \det A &= \sum_{i=1}^{n}(x_i+y_i)c_{ij}(A) \\ &= \sum_{i=1}^{n} x_i c_{ij}(A) + \sum_{i=1}^{n} y_i c_{ij}(A) \\ &= T(\mathbf{x}) + T(\mathbf{y}) \end{aligned}$$

Similarly for $T(a\mathbf{x}) = aT(\mathbf{x})$.

Exercise 3.1.22 Show that

$$\det \begin{bmatrix} 0 & 0 & \cdots & 0 & a_1 \\ 0 & 0 & \cdots & a_2 & * \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{n-1} & \cdots & * & * \\ a_n & * & \cdots & * & * \end{bmatrix} = (-1)^k a_1 a_2 \cdots a_n$$

where either $n = 2k$ or $n = 2k+1$, and $*$-entries are arbitrary.

Exercise 3.1.23 By expanding along the first column, show that:

$$\det \begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} = 1 + (-1)^{n+1}$$

if the matrix is $n \times n$, $n \geq 2$.

Exercise 3.1.24 Form matrix B from a matrix A by writing the columns of A in reverse order. Express $\det B$ in terms of $\det A$.

If A is $n \times n$, then $\det B = (-1)^k \det A$ where $n = 2k$ or $n = 2k+1$.

Exercise 3.1.25 Prove property 3 of Theorem 3.1.2 by expanding along the row (or column) in question.

Exercise 3.1.26 Show that the line through two distinct points (x_1, y_1) and (x_2, y_2) in the plane has equation

$$\det \begin{bmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix} = 0$$

Exercise 3.1.27 Let A be an $n \times n$ matrix. Given a polynomial $p(x) = a_0 + a_1 x + \cdots + a_m x^m$, we write

$p(A) = a_0 I + a_1 A + \cdots + a_m A^m$. For example, if $p(x) = 2 - 3x + 5x^2$, then
$p(A) = 2I - 3A + 5A^2$. The *characteristic polynomial* of A is defined to be $c_A(x) = \det[xI - A]$, and the Cayley-Hamilton theorem asserts that $c_A(A) = 0$ for any matrix A.

a. Verify the theorem for

i. $A = \left[\begin{array}{rr} 3 & 2 \\ 1 & -1 \end{array} \right]$ ii. $A = \left[\begin{array}{rrr} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 8 & 2 & 2 \end{array} \right]$

b. Prove the theorem for $A = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]$

3.2 Determinants and Matrix Inverses

In this section, several theorems about determinants are derived. One consequence of these theorems is that a square matrix A is invertible if and only if $\det A \neq 0$. Moreover, determinants are used to give a formula for A^{-1} which, in turn, yields a formula (called Cramer's rule) for the solution of any system of linear equations with an invertible coefficient matrix.

We begin with a remarkable theorem (due to Cauchy in 1812) about the determinant of a product of matrices. The proof is given at the end of this section.

Theorem 3.2.1: Product Theorem

If A and B are $n \times n$ matrices, then $\det(AB) = \det A \det B$.

The complexity of matrix multiplication makes the product theorem quite unexpected. Here is an example where it reveals an important numerical identity.

Example 3.2.1

If $A = \left[\begin{array}{rr} a & b \\ -b & a \end{array}\right]$ and $B = \left[\begin{array}{rr} c & d \\ -d & c \end{array}\right]$ then $AB = \left[\begin{array}{cc} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{array}\right]$.

Hence $\det A \det B = \det(AB)$ gives the identity

$$(a^2+b^2)(c^2+d^2) = (ac-bd)^2 + (ad+bc)^2$$

Theorem 3.2.1 extends easily to $\det(ABC) = \det A \det B \det C$. In fact, induction gives

$$\det(A_1A_2\cdots A_{k-1}A_k) = \det A_1 \det A_2 \cdots \det A_{k-1} \det A_k$$

for any square matrices $A_1, \dots, A_k$ of the same size. In particular, if each $A_i = A$, we obtain

$$\det(A^k) = (detA)^k, \quad \text{for any } k \geq 1$$

We can now give the invertibility condition.

Theorem 3.2.2

An $n \times n$ matrix A is invertible if and only if $\det A \neq 0$. When this is the case, $\det(A^{-1}) = \frac{1}{\det A}$

<u>Proof.</u> If A is invertible, then $AA^{-1} = I$; so the product theorem gives

$$1 = \det I = \det(AA^{-1}) = \det A \det A^{-1}$$

Hence, $\det A \neq 0$ and also $\det A^{-1} = \frac{1}{\det A}$.

Conversely, if $\det A \neq 0$, we show that A can be carried to I by elementary row operations (and invoke Theorem 2.4.5). Certainly, A can be carried to its reduced row-echelon form R, so $R = E_k \cdots E_2 E_1 A$ where the E_i are elementary matrices (Theorem 2.5.1). Hence the product theorem gives

$$\det R = \det E_k \cdots \det E_2 \det E_1 \det A$$

Since $\det E \neq 0$ for all elementary matrices E, this shows $\det R \neq 0$. In particular, R has no row of zeros, so $R = I$ because R is square and reduced row-echelon. This is what we wanted. □

Example 3.2.2

For which values of c does $A = \begin{bmatrix} 1 & 0 & -c \\ -1 & 3 & 1 \\ 0 & 2c & -4 \end{bmatrix}$ have an inverse?

Solution. Compute $\det A$ by first adding c times column 1 to column 3 and then expanding along row 1.

$$\det A = \det \begin{bmatrix} 1 & 0 & -c \\ -1 & 3 & 1 \\ 0 & 2c & -4 \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 0 \\ -1 & 3 & 1-c \\ 0 & 2c & -4 \end{bmatrix} = 2(c+2)(c-3)$$

Hence, $\det A = 0$ if $c = -2$ or $c = 3$, and A has an inverse if $c \neq -2$ and $c \neq 3$.

Example 3.2.3

If a product $A_1 A_2 \cdots A_k$ of square matrices is invertible, show that each A_i is invertible.

Solution. We have $\det A_1 \det A_2 \cdots \det A_k = \det (A_1 A_2 \cdots A_k)$ by the product theorem, and $\det (A_1 A_2 \cdots A_k) \neq 0$ by Theorem 3.2.2 because $A_1 A_2 \cdots A_k$ is invertible. Hence

$$\det A_1 \det A_2 \cdots \det A_k \neq 0$$

so $\det A_i \neq 0$ for each i. This shows that each A_i is invertible, again by Theorem 3.2.2.

Theorem 3.2.3

If A is any square matrix, $\det A^T = \det A$.

Proof. Consider first the case of an elementary matrix E. If E is of type I or II, then $E^T = E$; so certainly $\det E^T = \det E$. If E is of type III, then E^T is also of type III; so $\det E^T = 1 = \det E$ by Theorem 3.1.2. Hence, $\det E^T = \det E$ for every elementary matrix E.

Now let A be any square matrix. If A is not invertible, then neither is A^T; so $\det A^T = 0 = \det A$ by Theorem 3.2.2. On the other hand, if A is invertible, then $A = E_k \cdots E_2 E_1$, where the E_i are elementary matrices (Theorem 2.5.2). Hence, $A^T = E_1^T E_2^T \cdots E_k^T$ so the product theorem gives

$$\begin{aligned}\det A^T = \det E_1^T \det E_2^T \cdots \det E_k^T &= \det E_1 \det E_2 \cdots \det E_k \\ &= \det E_k \cdots \det E_2 \det E_1 \\ &= \det A\end{aligned}$$

This completes the proof. □

Example 3.2.4

If $\det A = 2$ and $\det B = 5$, calculate $\det(A^3B^{-1}A^TB^2)$.

Solution. We use several of the facts just derived.

$$\begin{aligned}\det(A^3B^{-1}A^TB^2) &= \det(A^3)\det(B^{-1})\det(A^T)\det(B^2) \\ &= (\det A)^3 \tfrac{1}{\det B} \det A (\det B)^2 \\ &= 2^3 \cdot \tfrac{1}{5} \cdot 2 \cdot 5^2 \\ &= 80\end{aligned}$$

Example 3.2.5

A square matrix is called **orthogonal** if $A^{-1} = A^T$. What are the possible values of $\det A$ if A is orthogonal?

Solution. If A is orthogonal, we have $I = AA^T$. Take determinants to obtain

$$1 = \det I = \det(AA^T) = \det A \det A^T = (\det A)^2$$

Since $\det A$ is a number, this means $\det A = \pm 1$.

Hence Theorems 2.6.4 and 2.6.5 imply that rotation about the origin and reflection about a line through the origin in $\mathbb{R}^2$ have orthogonal matrices with determinants 1 and -1 respectively. In fact they are the *only* such transformations of $\mathbb{R}^2$. We have more to say about this in Section 8.2.

Adjugates

In Section 2.4 we defined the adjugate of a 2×2 matrix $A = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$ to be $\operatorname{adj}(A) = \left[\begin{array}{rr} d & -b \\ -c & a \end{array}\right]$. Then we verified that $A(\operatorname{adj} A) = (\det A)I = (\operatorname{adj} A)A$ and hence that, if $\det A \neq 0$, $A^{-1} = \frac{1}{\det A} \operatorname{adj} A$. We are now able to define the adjugate of an arbitrary square matrix and to show that this formula for the inverse remains valid (when the inverse exists).

Recall that the (i, j)-cofactor $c_{ij}(A)$ of a square matrix A is a number defined for each position (i, j) in the matrix. If A is a square matrix, the **cofactor matrix of** A is defined to be the matrix $\left[c_{ij}(A)\right]$ whose (i, j)-entry is the (i, j)-cofactor of A.

Definition 3.3 Adjugate of a Matrix

The **adjugate**[4] *of* A, *denoted* $\text{adj}(A)$, *is the transpose of this cofactor matrix; in symbols,*

$$\text{adj}(A) = \left[c_{ij}(A)\right]^T$$

This agrees with the earlier definition for a 2×2 matrix A as the reader can verify.

Example 3.2.6

Compute the adjugate of $A = \left[\begin{array}{rrr} 1 & 3 & -2 \\ 0 & 1 & 5 \\ -2 & -6 & 7 \end{array}\right]$ and calculate $A(\text{adj}\, A)$ and $(\text{adj}\, A)A$.

Solution. We first find the cofactor matrix.

$$\left[\begin{array}{ccc} c_{11}(A) & c_{12}(A) & c_{13}(A) \\ c_{21}(A) & c_{22}(A) & c_{23}(A) \\ c_{31}(A) & c_{32}(A) & c_{33}(A) \end{array}\right] = \left[\begin{array}{ccc} \left|\begin{array}{rr} 1 & 5 \\ -6 & 7 \end{array}\right| & -\left|\begin{array}{rr} 0 & 5 \\ -2 & 7 \end{array}\right| & \left|\begin{array}{rr} 0 & 1 \\ -2 & -6 \end{array}\right| \\ -\left|\begin{array}{rr} 3 & -2 \\ -6 & 7 \end{array}\right| & \left|\begin{array}{rr} 1 & -2 \\ -2 & 7 \end{array}\right| & -\left|\begin{array}{rr} 1 & 3 \\ -2 & -6 \end{array}\right| \\ \left|\begin{array}{rr} 3 & -2 \\ 1 & 5 \end{array}\right| & -\left|\begin{array}{rr} 1 & -2 \\ 0 & 5 \end{array}\right| & \left|\begin{array}{rr} 1 & 3 \\ 0 & 1 \end{array}\right| \end{array}\right]$$

$$= \left[\begin{array}{rrr} 37 & -10 & 2 \\ -9 & 3 & 0 \\ 17 & -5 & 1 \end{array}\right]$$

Then the adjugate of A is the transpose of this cofactor matrix.

$$\text{adj}\, A = \left[\begin{array}{rrr} 37 & -10 & 2 \\ -9 & 3 & 0 \\ 17 & -5 & 1 \end{array}\right]^T = \left[\begin{array}{rrr} 37 & -9 & 17 \\ -10 & 3 & -5 \\ 2 & 0 & 1 \end{array}\right]$$

The computation of $A(\text{adj}\, A)$ gives

$$A(\text{adj}\, A) = \left[\begin{array}{rrr} 1 & 3 & -2 \\ 0 & 1 & 5 \\ -2 & -6 & 7 \end{array}\right] \left[\begin{array}{rrr} 37 & -9 & 17 \\ -10 & 3 & -5 \\ 2 & 0 & 1 \end{array}\right] = \left[\begin{array}{rrr} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{array}\right] = 3I$$

and the reader can verify that also $(\text{adj}\, A)A = 3I$. Hence, analogy with the 2×2 case would indicate that $\det A = 3$; this is, in fact, the case.

The relationship $A(\text{adj}\, A) = (\det A)I$ holds for any square matrix A. To see why this is so,

[4] This is also called the classical adjoint of A, but the term "adjoint" has another meaning.

consider the general 3×3 case. Writing $c_{ij}(A) = c_{ij}$ for short, we have

$$\text{adj } A = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}^T = \begin{bmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{bmatrix}$$

If $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ in the usual notation, we are to verify that $A(\text{adj } A) = (\det A)I$. That is,

$$A(\text{adj } A) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{bmatrix} = \begin{bmatrix} \det A & 0 & 0 \\ 0 & \det A & 0 \\ 0 & 0 & \det A \end{bmatrix}$$

Consider the (1, 1)-entry in the product. It is given by $a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13}$, and this is just the cofactor expansion of $\det A$ along the first row of A. Similarly, the (2, 2)-entry and the (3, 3)-entry are the cofactor expansions of $\det A$ along rows 2 and 3, respectively.

So it remains to be seen why the off-diagonal elements in the matrix product $A(\text{adj } A)$ are all zero. Consider the (1, 2)-entry of the product. It is given by $a_{11}c_{21} + a_{12}c_{22} + a_{13}c_{23}$. This *looks* like the cofactor expansion of the determinant of *some* matrix. To see which, observe that c_{21}, c_{22}, and c_{23} are all computed by *deleting* row 2 of A (and one of the columns), so they remain the same if row 2 of A is changed. In particular, if row 2 of A is replaced by row 1, we obtain

$$a_{11}c_{21} + a_{12}c_{22} + a_{13}c_{23} = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = 0$$

where the expansion is along row 2 and where the determinant is zero because two rows are identical. A similar argument shows that the other off-diagonal entries are zero.

This argument works in general and yields the first part of Theorem 3.2.4. The second assertion follows from the first by multiplying through by the scalar $\frac{1}{\det A}$.

Theorem 3.2.4: Adjugate Formula

If A is any square matrix, then

$$A(\text{adj } A) = (\det A)I = (\text{adj } A)A$$

In particular, if det A $\neq$ 0, the inverse of A is given by

$$A^{-1} = \frac{1}{\det A} \text{adj } A$$

It is important to note that this theorem is *not* an efficient way to find the inverse of the matrix A. For example, if A were 10×10, the calculation of $\text{adj } A$ would require computing $10^2 = 100$ determinants of 9×9 matrices! On the other hand, the matrix inversion algorithm would find A^{-1} with about the same effort as finding $\det A$. Clearly, Theorem 3.2.4 is not a *practical* result: its virtue is that it gives a formula for A^{-1} that is useful for *theoretical* purposes.

Example 3.2.7

Find the $(2,\ 3)$-entry of A^{-1} if $A = \begin{bmatrix} 2 & 1 & 3 \\ 5 & -7 & 1 \\ 3 & 0 & -6 \end{bmatrix}$.

Solution. First compute

$$\det A = \begin{vmatrix} 2 & 1 & 3 \\ 5 & -7 & 1 \\ 3 & 0 & -6 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 7 \\ 5 & -7 & 11 \\ 3 & 0 & 0 \end{vmatrix} = 3 \begin{vmatrix} 1 & 7 \\ -7 & 11 \end{vmatrix} = 180$$

Since $A^{-1} = \frac{1}{\det A} \operatorname{adj} A = \frac{1}{180} \left[c_{ij}(A) \right]^T$, the $(2,\ 3)$-entry of A^{-1} is the $(3,\ 2)$-entry of the matrix $\frac{1}{180} \left[c_{ij}(A) \right]$; that is, it equals $\frac{1}{180} c_{32}(A) = \frac{1}{180} \left(- \begin{vmatrix} 2 & 3 \\ 5 & 1 \end{vmatrix} \right) = \frac{13}{180}$.

Example 3.2.8

If A is $n \times n$, $n \geq 2$, show that $\det(\operatorname{adj} A) = (\det A)^{n-1}$.

Solution. Write $d = \det A$; we must show that $\det(\operatorname{adj} A) = d^{n-1}$. We have $A(\operatorname{adj} A) = dI$ by Theorem 3.2.4, so taking determinants gives $d \det(\operatorname{adj} A) = d^n$. Hence we are done if $d \neq 0$. Assume $d = 0$; we must show that $\det(\operatorname{adj} A) = 0$, that is, $\operatorname{adj} A$ is not invertible. If $A \neq 0$, this follows from $A(\operatorname{adj} A) = dI = 0$; if $A = 0$, it follows because then $\operatorname{adj} A = 0$.

Cramer's Rule

Theorem 3.2.4 has a nice application to linear equations. Suppose

$$A\mathbf{x} = \mathbf{b}$$

is a system of n equations in n variables $x_1,\ x_2,\ \dots,\ x_n$. Here A is the $n \times n$ coefficient matrix, and $\mathbf{x}$ and $\mathbf{b}$ are the columns

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

of variables and constants, respectively. If $\det A \neq 0$, we left multiply by A^{-1} to obtain the solution $\mathbf{x} = A^{-1}\mathbf{b}$. When we use the adjugate formula, this becomes

$$\begin{aligned}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} &= \frac{1}{\det A}(\operatorname{adj} A)\mathbf{b} \\ &= \frac{1}{\det A}\begin{bmatrix} c_{11}(A) & c_{21}(A) & \cdots & c_{n1}(A) \\ c_{12}(A) & c_{22}(A) & \cdots & c_{n2}(A) \\ \vdots & \vdots & & \vdots \\ c_{1n}(A) & c_{2n}(A) & \cdots & c_{nn}(A) \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}\end{aligned}$$

Hence, the variables $x_1, x_2, \dots, x_n$ are given by

$$\begin{aligned} x_1 &= \tfrac{1}{\det A}\left[b_1c_{11}(A) + b_2c_{21}(A) + \cdots + b_nc_{n1}(A)\right] \\ x_2 &= \tfrac{1}{\det A}\left[b_1c_{12}(A) + b_2c_{22}(A) + \cdots + b_nc_{n2}(A)\right] \\ &\vdots \qquad\qquad \vdots \\ x_n &= \tfrac{1}{\det A}\left[b_1c_{1n}(A) + b_2c_{2n}(A) + \cdots + b_nc_{nn}(A)\right] \end{aligned}$$

Now the quantity $b_1c_{11}(A) + b_2c_{21}(A) + \cdots + b_nc_{n1}(A)$ occurring in the formula for x_1 looks like the cofactor expansion of the determinant of a matrix. The cofactors involved are $c_{11}(A)$, $c_{21}(A)$, $\dots$, $c_{n1}(A)$, corresponding to the first column of A. If A_1 is obtained from A by replacing the first column of A by $\mathbf{b}$, then $c_{i1}(A_1) = c_{i1}(A)$ for each i because column 1 is deleted when computing them. Hence, expanding $\det(A_1)$ by the first column gives

$$\begin{aligned} \det A_1 &= b_1c_{11}(A_1) + b_2c_{21}(A_1) + \cdots + b_nc_{n1}(A_1) \\ &= b_1c_{11}(A) + b_2c_{21}(A) + \cdots + b_nc_{n1}(A) \\ &= (\det A)x_1 \end{aligned}$$

Hence, $x_1 = \frac{\det A_1}{\det A}$ and similar results hold for the other variables.

Theorem 3.2.5: Cramer's Rule[5]

If A is an invertible $n \times n$ matrix, the solution to the system

$$A\mathbf{x} = \mathbf{b}$$

of n equations in the variables $x_1, x_2, \dots, x_n$ is given by

$$x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A}, \quad \cdots, \quad x_n = \frac{\det A_n}{\det A}$$

where, for each k, A_k is the matrix obtained from A by replacing column k by $\mathbf{b}$.

[5]Gabriel Cramer (1704–1752) was a Swiss mathematician who wrote an introductory work on algebraic curves. He popularized the rule that bears his name, but the idea was known earlier.

Example 3.2.9

Find x_1, given the following system of equations.

$$\begin{aligned} 5x_1 + x_2 - x_3 &= 4 \\ 9x_1 + x_2 - x_3 &= 1 \\ x_1 - x_2 + 5x_3 &= 2 \end{aligned}$$

Solution. Compute the determinants of the coefficient matrix A and the matrix A_1 obtained from it by replacing the first column by the column of constants.

$$\det A = \det \begin{bmatrix} 5 & 1 & -1 \\ 9 & 1 & -1 \\ 1 & -1 & 5 \end{bmatrix} = -16$$

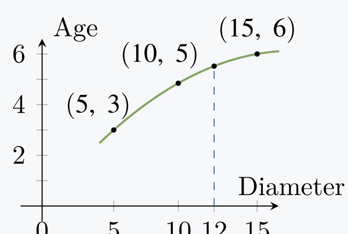
$$\det A_1 = \det \begin{bmatrix} 4 & 1 & -1 \\ 1 & 1 & -1 \\ 2 & -1 & 5 \end{bmatrix} = 12$$

Hence, $x_1 = \frac{\det A_1}{\det A} = -\frac{3}{4}$ by Cramer's rule.

Cramer's rule is *not* an efficient way to solve linear systems or invert matrices. True, it enabled us to calculate x_1 here without computing x_2 or x_3. Although this might seem an advantage, the truth of the matter is that, for large systems of equations, the number of computations needed to find *all* the variables by the gaussian algorithm is comparable to the number required to find *one* of the determinants involved in Cramer's rule. Furthermore, the algorithm works when the matrix of the system is not invertible and even when the coefficient matrix is not square. Like the adjugate formula, then, Cramer's rule is *not* a practical numerical technique; its virtue is theoretical.

Polynomial Interpolation

Example 3.2.10



A forester wants to estimate the age (in years) of a tree by measuring the diameter of the trunk (in cm). She obtains the following data:

	Tree 1	Tree 2	Tree 3
Trunk Diameter	5	10	15
Age	3	5	6

Estimate the age of a tree with a trunk diameter of 12 cm.

Solution.
The forester decides to "fit" a quadratic polynomial

$$p(x) = r_0 + r_1 x + r_2 x^2$$

to the data, that is choose the coefficients r_0, r_1, and r_2 so that $p(5) = 3$, $p(10) = 5$, and $p(15) = 6$, and then use $p(12)$ as the estimate. These conditions give three linear equations:

$$\begin{aligned} r_0 + 5r_1 + 25r_2 &= 3 \\ r_0 + 10r_1 + 100r_2 &= 5 \\ r_0 + 15r_1 + 225r_2 &= 6 \end{aligned}$$

The (unique) solution is $r_0 = 0$, $r_1 = \frac{7}{10}$, and $r_2 = -\frac{1}{50}$, so

$$p(x) = \tfrac{7}{10}x - \tfrac{1}{50}x^2 = \tfrac{1}{50}x(35 - x)$$

Hence the estimate is $p(12) = 5.52$.

As in Example 3.2.10, it often happens that two variables x and y are related but the actual functional form $y = f(x)$ of the relationship is unknown. Suppose that for certain values $x_1, x_2, \dots, x_n$ of x the corresponding values $y_1, y_2, \dots, y_n$ are known (say from experimental measurements). One way to estimate the value of y corresponding to some other value a of x is to find a polynomial[6]

$$p(x) = r_0 + r_1x + r_2x^2 + \cdots + r_{n-1}x^{n-1}$$

that "fits" the data, that is $p(x_i) = y_i$ holds for each $i = 1, 2, \dots, n$. Then the estimate for y is $p(a)$. As we will see, such a polynomial always exists if the x_i are distinct.

The conditions that $p(x_i) = y_i$ are

$$\begin{array}{ccccccccc} r_0 &+& r_1x_1 &+& r_2x_1^2 &+ \cdots +& r_{n-1}x_1^{n-1} &=& y_1 \\ r_0 &+& r_1x_2 &+& r_2x_2^2 &+ \cdots +& r_{n-1}x_2^{n-1} &=& y_2 \\ \vdots && \vdots && \vdots && \vdots && \vdots \\ r_0 &+& r_1x_n &+& r_2x_n^2 &+ \cdots +& r_{n-1}x_n^{n-1} &=& y_n \end{array}$$

In matrix form, this is

$$\left[\begin{array}{ccccc} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{array} \right] \left[\begin{array}{c} r_0 \\ r_1 \\ \vdots \\ r_{n-1} \end{array} \right] = \left[\begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_n \end{array} \right] \tag{3.3}$$

It can be shown (see Theorem 3.2.7) that the determinant of the coefficient matrix equals the product of all terms $(x_i - x_j)$ with $i > j$ and so is nonzero (because the x_i are distinct). Hence the equations have a unique solution $r_0, r_1, \dots, r_{n-1}$. This proves

[6]A **polynomial** is an expression of the form $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ where the a_i are numbers and x is a variable. If $a_n \neq 0$, the integer n is called the degree of the polynomial, and a_n is called the leading coefficient. See Appendix **??**.

Theorem 3.2.6

Let n data pairs $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ be given, and assume that the x_i are distinct. Then there exists a unique polynomial

$$p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_{n-1}x^{n-1}$$

such that $p(x_i) = y_i$ for each $i = 1, 2, \dots, n$.

The polynomial in Theorem 3.2.6 is called the **interpolating polynomial** for the data.

We conclude by evaluating the determinant of the coefficient matrix in Equation 3.3. If $a_1, a_2, \dots, a_n$ are numbers, the determinant

$$\det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}$$

is called a **Vandermonde determinant**.[7] There is a simple formula for this determinant. If $n = 2$, it equals $(a_2 - a_1)$; if $n = 3$, it is $(a_3 - a_2)(a_3 - a_1)(a_2 - a_1)$ by Example 3.1.8. The general result is the product

$$\prod_{1 \le j < i \le n} (a_i - a_j)$$

of all factors $(a_i - a_j)$ where $1 \le j < i \le n$. For example, if $n = 4$, it is

$$(a_4 - a_3)(a_4 - a_2)(a_4 - a_1)(a_3 - a_2)(a_3 - a_1)(a_2 - a_1)$$

Theorem 3.2.7

Let $a_1, a_2, \dots, a_n$ be numbers where $n \ge 2$. Then the corresponding Vandermonde determinant is given by

$$\det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix} = \prod_{1 \le j < i \le n} (a_i - a_j)$$

Proof. We may assume that the a_i are distinct; otherwise both sides are zero. We proceed by induction on $n \ge 2$; we have it for $n = 2, 3$. So assume it holds for $n - 1$. The trick is to replace a_n

[7] Alexandre Théophile Vandermonde (1735–1796) was a French mathematician who made contributions to the theory of equations.

by a variable x, and consider the determinant

$$p(x) = \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^{n-1} \\ 1 & x & x^2 & \cdots & x^{n-1} \end{bmatrix}$$

Then $p(x)$ is a polynomial of degree at most $n-1$ (expand along the last row), and $p(a_i) = 0$ for each $i = 1, 2, \dots, n-1$ because in each case there are two identical rows in the determinant. In particular, $p(a_1) = 0$, so we have $p(x) = (x-a_1)p_1(x)$ by the factor theorem (see Appendix **??**). Since $a_2 \neq a_1$, we obtain $p_1(a_2) = 0$, and so $p_1(x) = (x-a_2)p_2(x)$. Thus $p(x) = (x-a_1)(x-a_2)p_2(x)$. As the a_i are distinct, this process continues to obtain

$$p(x) = (x-a_1)(x-a_2)\cdots(x-a_{n-1})d \tag{3.4}$$

where d is the coefficient of x^{n-1} in $p(x)$. By the cofactor expansion of $p(x)$ along the last row we get

$$d = (-1)^{n+n} \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-2} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^{n-2} \end{bmatrix}$$

Because $(-1)^{n+n} = 1$ the induction hypothesis shows that d is the product of all factors $(a_i - a_j)$ where $1 \leq j < i \leq n-1$. The result now follows from Equation 3.4 by substituting a_n for x in $p(x)$. □

Proof of Theorem 3.2.1. If A and B are $n \times n$ matrices we must show that

$$\det(AB) = \det A \det B \tag{3.5}$$

Recall that if E is an elementary matrix obtained by doing one row operation to I_n, then doing that operation to a matrix C (Lemma 2.5.1) results in EC. By looking at the three types of elementary matrices separately, Theorem 3.1.2 shows that

$$\det(EC) = \det E \det C \quad \text{for any matrix } C \tag{3.6}$$

Thus if $E_1, E_2, \dots, E_k$ are all elementary matrices, it follows by induction that

$$\det(E_k \cdots E_2 E_1 C) = \det E_k \cdots \det E_2 \det E_1 \det C \text{ for any matrix } C \tag{3.7}$$

Lemma. If A has no inverse, then $\det A = 0$.

Proof. Let $A \to R$ where R is reduced row-echelon, say $E_n \cdots E_2 E_1 A = R$. Then R has a row of zeros by Part (4) of Theorem 2.4.5, and hence $\det R = 0$. But then Equation 3.7 gives $\det A = 0$ because $\det E \neq 0$ for any elementary matrix E. This proves the Lemma.

Now we can prove Equation 3.5 by considering two cases.

Case 1. A has no inverse. Then AB also has no inverse (otherwise $A[B(AB)^{-1}] = I$) so A is invertible by Corollary 2.4.2 to Theorem 2.4.5. Hence the above Lemma (twice) gives

$$\det(AB) = 0 = 0 \det B = \det A \det B$$

proving Equation 3.5 in this case.

Case 2. A has an inverse. Then A is a product of elementary matrices by Theorem 2.5.2, say $A = E_1E_2\cdots E_k$. Then Equation 3.7 with $C = I$ gives

$$\det A = \det(E_1E_2\cdots E_k) = \det E_1 \det E_2 \cdots \det E_k$$

But then Equation 3.7 with $C = B$ gives

$$\det(AB) = \det[(E_1E_2\cdots E_k)B] = \det E_1 \det E_2 \cdots \det E_k \det B = \det A \det B$$

and Equation 3.5 holds in this case too. □

Exercises for 3.2

Exercise 3.2.1 Find the adjugate of each of the following matrices.

a) $\begin{bmatrix} 5 & 1 & 3 \\ -1 & 2 & 3 \\ 1 & 4 & 8 \end{bmatrix}$ b) $\begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ d) $\frac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$

---

b. $\begin{bmatrix} 1 & -1 & -2 \\ -3 & 1 & 6 \\ -3 & 1 & 4 \end{bmatrix}$

d. $\frac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix} = A$

Exercise 3.2.2 Use determinants to find which real values of c make each of the following matrices invertible.

a) $\begin{bmatrix} 1 & 0 & 3 \\ 3 & -4 & c \\ 2 & 5 & 8 \end{bmatrix}$ b) $\begin{bmatrix} 0 & c & -c \\ -1 & 2 & 1 \\ c & -c & c \end{bmatrix}$

c) $\begin{bmatrix} c & 1 & 0 \\ 0 & 2 & c \\ -1 & c & 5 \end{bmatrix}$ d) $\begin{bmatrix} 4 & c & 3 \\ c & 2 & c \\ 5 & c & 4 \end{bmatrix}$

e) $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & c \\ 2 & c & 1 \end{bmatrix}$ f) $\begin{bmatrix} 1 & c & -1 \\ c & 1 & 1 \\ 0 & 1 & c \end{bmatrix}$

---

b. $c \neq 0$

d. any c

f. $c \neq -1$

Exercise 3.2.3 Let A, B, and C denote $n \times n$ matrices and assume that $\det A = -1$, $\det B = 2$, and $\det C = 3$. Evaluate:

a) $\det(A^3BC^TB^{-1})$ b) $\det(B^2C^{-1}AB^{-1}C^T)$

---

b. -2

Exercise 3.2.4 Let A and B be invertible $n \times n$ matrices. Evaluate:

a) $\det(B^{-1}AB)$ b) $\det(A^{-1}B^{-1}AB)$

---

b. 1

Exercise 3.2.5 If A is 3×3 and $\det(2A^{-1}) = -4$ and $\det(A^3(B^{-1})^T) = -4$, find $\det A$ and $\det B$.

Exercise 3.2.6 Let $A = \begin{bmatrix} a & b & c \\ p & q & r \\ u & v & w \end{bmatrix}$ and assume that $\det A = 3$. Compute:

a. $\det(2B^{-1})$ where $B = \begin{bmatrix} 4u & 2a & -p \\ 4v & 2b & -q \\ 4w & 2c & -r \end{bmatrix}$

b. $\det(2C^{-1})$ where $C = \begin{bmatrix} 2p & -a+u & 3u \\ 2q & -b+v & 3v \\ 2r & -c+w & 3w \end{bmatrix}$

---

b. $\frac{4}{9}$

Exercise 3.2.7 If $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = -2$ calculate:

a. $\det \begin{bmatrix} 2 & -2 & 0 \\ c+1 & -1 & 2a \\ d-2 & 2 & 2b \end{bmatrix}$

b. $\det \begin{bmatrix} 2b & 0 & 4d \\ 1 & 2 & -2 \\ a+1 & 2 & 2(c-1) \end{bmatrix}$

c. $\det(3A^{-1})$ where $A = \begin{bmatrix} 3c & a+c \\ 3d & b+d \end{bmatrix}$

---

b. 16

Exercise 3.2.8 Solve each of the following by Cramer's rule:

a) $\begin{array}{rl} 2x + y = & 1 \\ 3x + 7y = & -2 \end{array}$

b) $\begin{array}{rl} 3x + 4y = & 9 \\ 2x - y = & -1 \end{array}$

c) $\begin{array}{rl} 5x + y - z = & -7 \\ 2x - y - 2z = & 6 \\ 3x + 2z = & -7 \end{array}$

d) $\begin{array}{rl} 4x - y + 3z = & 1 \\ 6x + 2y - z = & 0 \\ 3x + 3y + 2z = & -1 \end{array}$

---

b. $\frac{1}{11}\begin{bmatrix} 5 \\ 21 \end{bmatrix}$

d. $\frac{1}{79}\begin{bmatrix} 12 \\ -37 \\ -2 \end{bmatrix}$

Exercise 3.2.9 Use Theorem 3.2.4 to find the (2, 3)-entry of A^{-1} if:

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

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

---

b. $\frac{4}{51}$

Exercise 3.2.10 Explain what can be said about $\det A$ if:

a) $A^2 = A$

b) $A^2 = I$

c) $A^3 = A$

d) $PA = P$ and P is invertible

e) $A^2 = uA$ and A is $n \times n$

f) $A = -A^T$ and A is $n \times n$

g) $A^2 + I = 0$ and A is $n \times n$

---

b. $\det A = 1, \ -1$

d. $\det A = 1$

f. $\det A = 0$ if n is odd; nothing can be said if n is even

Exercise 3.2.11 Let A be $n \times n$. Show that $uA = (uI)A$, and use this with Theorem 3.2.1 to deduce the result in Theorem 3.1.3: $\det(uA) = u^n \det A$.

Exercise 3.2.12 If A and B are $n \times n$ matrices, if $AB = -BA$, and if n is odd, show that either A or B has no inverse.

Exercise 3.2.13 Show that $\det AB = \det BA$ holds for any two $n \times n$ matrices A and B.

Exercise 3.2.14 If $A^k = 0$ for some $k \geq 1$, show that A is not invertible.

Exercise 3.2.15 If $A^{-1} = A^T$, describe the cofactor matrix of A in terms of A.

---

dA where $d = \det A$

Exercise 3.2.16 Show that no 3×3 matrix A exists such that $A^2 + I = 0$. Find a 2×2 matrix A with this property.

Exercise 3.2.17 Show that $\det(A + B^T) = \det(A^T + B)$ for any $n \times n$ matrices A and B.

Exercise 3.2.18 Let A and B be invertible $n \times n$ matrices. Show that $\det A = \det B$ if and only if $A = UB$ where U is a matrix with $\det U = 1$.

Exercise 3.2.19 For each of the matrices in Exercise 2, find the inverse for those values of c for which it exists.

---

b. $\frac{1}{c}\begin{bmatrix} 1 & 0 & 1 \\ 0 & c & 1 \\ -1 & c & 1 \end{bmatrix}$, $c \neq 0$

d. $\frac{1}{2}\begin{bmatrix} 8-c^2 & -c & c^2-6 \\ c & 1 & -c \\ c^2-10 & c & 8-c^2 \end{bmatrix}$

f. $\frac{1}{c^3+1}\begin{bmatrix} 1-c & c^2+1 & -c-1 \\ c^2 & -c & c+1 \\ -c & 1 & c^2-1 \end{bmatrix}$, $c \neq -1$

Exercise 3.2.20 In each case either prove the statement or give an example showing that it is false:

a. If $\operatorname{adj} A$ exists, then A is invertible.

b. If A is invertible and $\operatorname{adj} A = A^{-1}$, then $\det A = 1$.

c. $\det(AB) = \det(B^T A)$.

d. If $\det A \neq 0$ and $AB = AC$, then $B = C$.

e. If $A^T = -A$, then $\det A = -1$.

f. If $\operatorname{adj} A = 0$, then $A = 0$.

g. If A is invertible, then $\operatorname{adj} A$ is invertible.

h. If A has a row of zeros, so also does $\operatorname{adj} A$.

i. $\det(A^T A) > 0$ for all square matrices A.

j. $\det(I + A) = 1 + \det A$.

k. If AB is invertible, then A and B are invertible.

l. If $\det A = 1$, then $\operatorname{adj} A = A$.

m. If A is invertible and $\det A = d$, then $\operatorname{adj} A = dA^{-1}$.

---

b. T. $\det AB = \det A \det B = \det B \det A = \det BA$.

d. T. $\det A \neq 0$ means A^{-1} exists, so $AB = AC$ implies that $B = C$.

f. F. If $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ then $\operatorname{adj} A = 0$.

h. F. If $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ then $\operatorname{adj} A = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}$

j. F. If $A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ then $\det(I + A) = -1$ but $1 + \det A = 1$.

l. F. If $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ then $\det A = 1$ but $\operatorname{adj} A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \neq A$

Exercise 3.2.21 If A is 2×2 and $\det A = 0$, show that one column of A is a scalar multiple of the other. [*Hint*: Definition 2.5 and Part (2) of Theorem 2.4.5.]

Exercise 3.2.22 Find a polynomial $p(x)$ of degree 2 such that:

a. $p(0) = 2$, $p(1) = 3$, $p(3) = 8$

b. $p(0) = 5$, $p(1) = 3$, $p(2) = 5$

---

b. $5 - 4x + 2x^2$.

Exercise 3.2.23 Find a polynomial $p(x)$ of degree 3 such that:

a. $p(0) = p(1) = 1$, $p(-1) = 4$, $p(2) = -5$

b. $p(0) = p(1) = 1$, $p(-1) = 2$, $p(-2) = -3$

---

b. $1 - \frac{5}{3}x + \frac{1}{2}x^2 + \frac{7}{6}x^3$

Exercise 3.2.24 Given the following data pairs, find the interpolating polynomial of degree 3 and estimate the value of y corresponding to $x = 1.5$.

a. (0, 1), (1, 2), (2, 5), (3, 10)

b. (0, 1), (1, 1.49), (2, −0.42), (3, −11.33)

c. (0, 2), (1, 2.03), (2, −0.40), (−1, 0.89)

---

b. $1 - 0.51x + 2.1x^2 - 1.1x^3; 1.25$, so $y = 1.25$

Exercise 3.2.25 If $A = \begin{bmatrix} 1 & a & b \\ -a & 1 & c \\ -b & -c & 1 \end{bmatrix}$ show that $\det A = 1 + a^2 + b^2 + c^2$. Hence, find A^{-1} for any a, b, and c.

Exercise 3.2.26

a. Show that $A = \begin{bmatrix} a & p & q \\ 0 & b & r \\ 0 & 0 & c \end{bmatrix}$ has an inverse if and only if $abc \neq 0$, and find A^{-1} in that case.

b. Show that if an upper triangular matrix is invertible, the inverse is also upper triangular.

---

b. Use induction on n where A is $n \times n$. It is clear if $n = 1$. If $n > 1$, write $A = \begin{bmatrix} a & X \\ 0 & B \end{bmatrix}$ in block form where B is $(n-1) \times (n-1)$. Then $A^{-1} = \begin{bmatrix} a^{-1} & -a^{-1}XB^{-1} \\ 0 & B^{-1} \end{bmatrix}$, and this is upper triangular because B is upper triangular by induction.

Exercise 3.2.27 Let A be a matrix each of whose entries are integers. Show that each of the following conditions implies the other.

1. A is invertible and A^{-1} has integer entries.

2. $\det A = 1$ or -1.

Exercise 3.2.28 If $A^{-1} = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 3 \\ 3 & 1 & -1 \end{bmatrix}$ find $\operatorname{adj} A$.

$-\frac{1}{21}\begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 3 \\ 3 & 1 & -1 \end{bmatrix}$

Exercise 3.2.29 If A is 3×3 and $\det A = 2$, find $\det(A^{-1} + 4 \operatorname{adj} A)$.

Exercise 3.2.30 Show that $\det \begin{bmatrix} 0 & A \\ B & X \end{bmatrix} = \det A \det B$ when A and B are 2×2. What if A and B are 3×3? [*Hint*: Block multiply by $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$.]

Exercise 3.2.31 Let A be $n \times n$, $n \geq 2$, and assume one column of A consists of zeros. Find the possible values of $\operatorname{rank}(\operatorname{adj} A)$.

Exercise 3.2.32 If A is 3×3 and invertible, compute $\det(-A^2(\operatorname{adj} A)^{-1})$.

Exercise 3.2.33 Show that $\operatorname{adj}(uA) = u^{n-1} \operatorname{adj} A$ for all $n \times n$ matrices A.

Exercise 3.2.34 Let A and B denote invertible $n \times n$ matrices. Show that:

a. $\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-2} A$ (here $n \geq 2$) [*Hint*: See Example 3.2.8.]

b. $\operatorname{adj}(A^{-1}) = (\operatorname{adj} A)^{-1}$

c. $\operatorname{adj}(A^T) = (\operatorname{adj} A)^T$

d. $\operatorname{adj}(AB) = (\operatorname{adj} B)(\operatorname{adj} A)$ [*Hint*: Show that $AB \operatorname{adj}(AB) = AB \operatorname{adj} B \operatorname{adj} A$.]

---

b. Have $(\operatorname{adj} A)A = (\det A)I$; so taking inverses, $A^{-1} \cdot (\operatorname{adj} A)^{-1} = \frac{1}{\det A} I$. On the other hand, $A^{-1} \operatorname{adj}(A^{-1}) = \det(A^{-1})I = \frac{1}{\det A} I$. Comparison yields $A^{-1}(\operatorname{adj} A)^{-1} = A^{-1} \operatorname{adj}(A^{-1})$, and part **(b)** follows.

d. Write $\det A = d$, $\det B = e$. By the adjugate formula $AB \operatorname{adj}(AB) = deI$, and $AB \operatorname{adj} B \operatorname{adj} A = A[eI] \operatorname{adj} A = (eI)(dI) = deI$. Done as AB is invertible.

3.3 Diagonalization and Eigenvalues

The world is filled with examples of systems that evolve in time—the weather in a region, the economy of a nation, the diversity of an ecosystem, etc. Describing such systems is difficult in general and various methods have been developed in special cases. In this section we describe one such method, called *diagonalization,* which is one of the most important techniques in linear algebra. A very fertile example of this procedure is in modelling the growth of the population of an animal species. This has attracted more attention in recent years with the ever increasing awareness that many species are endangered. To motivate the technique, we begin by setting up a simple model of a bird population in which we make assumptions about survival and reproduction rates.

Example 3.3.1

Consider the evolution of the population of a species of birds. Because the number of males and females are nearly equal, we count only females. We assume that each female remains a juvenile for one year and then becomes an adult, and that only adults have offspring. We make three assumptions about reproduction and survival rates:

1. The number of juvenile females hatched in any year is twice the number of adult females alive the year before (we say the **reproduction rate** is 2).
2. Half of the adult females in any year survive to the next year (the **adult survival rate** is $\frac{1}{2}$).
3. One quarter of the juvenile females in any year survive into adulthood (the **juvenile survival rate** is $\frac{1}{4}$).

If there were 100 adult females and 40 juvenile females alive initially, compute the population of females k years later.

Solution. Let a_k and j_k denote, respectively, the number of adult and juvenile females after k years, so that the total female population is the sum $a_k + j_k$. Assumption 1 shows that $j_{k+1} = 2a_k$, while assumptions 2 and 3 show that $a_{k+1} = \frac{1}{2}a_k + \frac{1}{4}j_k$. Hence the numbers a_k and j_k in successive years are related by the following equations:

$$\begin{aligned} a_{k+1} &= \tfrac{1}{2}a_k + \tfrac{1}{4}j_k \\ j_{k+1} &= 2a_k \end{aligned}$$

If we write $\mathbf{v}_k = \left[\begin{array}{c} a_k \\ j_k \end{array} \right]$ and $A = \left[\begin{array}{cc} \frac{1}{2} & \frac{1}{4} \\ 2 & 0 \end{array} \right]$ these equations take the matrix form

$$\mathbf{v}_{k+1} = A\mathbf{v}_k, \quad \text{for each } k = 0,\ 1,\ 2,\ \dots$$

Taking $k = 0$ gives $\mathbf{v}_1 = A\mathbf{v}_0$, then taking $k = 1$ gives $\mathbf{v}_2 = A\mathbf{v}_1 = A^2\mathbf{v}_0$, and taking $k = 2$ gives $\mathbf{v}_3 = A\mathbf{v}_2 = A^3\mathbf{v}_0$. Continuing in this way, we get

$$\mathbf{v}_k = A^k\mathbf{v}_0, \quad \text{for each } k = 0,\ 1,\ 2,\ \dots$$

Since $\mathbf{v}_0 = \begin{bmatrix} a_0 \\ j_0 \end{bmatrix} = \begin{bmatrix} 100 \\ 40 \end{bmatrix}$ is known, finding the population profile $\mathbf{v}_k$ amounts to computing A^k for all $k \geq 0$. We will complete this calculation in Example 3.3.12 after some new techniques have been developed.

Let A be a fixed $n \times n$ matrix. A sequence $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots$ of column vectors in $\mathbb{R}^n$ is called a **linear dynamical system**[8] if $\mathbf{v}_0$ is known and the other $\mathbf{v}_k$ are determined (as in Example 3.3.1) by the conditions

$$\mathbf{v}_{k+1} = A\mathbf{v}_k \text{ for each } k = 0, 1, 2, \dots$$

These conditions are called a **matrix recurrence** for the vectors $\mathbf{v}_k$. As in Example 3.3.1, they imply that

$$\mathbf{v}_k = A^k \mathbf{v}_0 \text{ for all } k \geq 0$$

so finding the columns $\mathbf{v}_k$ amounts to calculating A^k for $k \geq 0$.

Direct computation of the powers A^k of a square matrix A can be time-consuming, so we adopt an indirect method that is commonly used. The idea is to first **diagonalize** the matrix A, that is, to find an invertible matrix P such that

$$P^{-1}AP = D \text{ is a diagonal matrix} \tag{3.8}$$

This works because the powers D^k of the diagonal matrix D are easy to compute, and Equation 3.8 enables us to compute powers A^k of the matrix A in terms of powers D^k of D. Indeed, we can solve Equation 3.8 for A to get $A = PDP^{-1}$. Squaring this gives

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD^2P^{-1}$$

Using this we can compute A^3 as follows:

$$A^3 = AA^2 = (PDP^{-1})(PD^2P^{-1}) = PD^3P^{-1}$$

Continuing in this way we obtain Theorem 3.3.1 (even if D is not diagonal).

Theorem 3.3.1

If $A = PDP^{-1}$ then $A^k = PD^kP^{-1}$ for each $k = 1, 2, \dots$.

Hence computing A^k comes down to finding an invertible matrix P as in equation Equation 3.8. To do this it is necessary to first compute certain numbers (called eigenvalues) associated with the matrix A.

[8] More precisely, this is *a linear discrete* dynamical system. Many models regard $\mathbf{v}_t$ as a continuous function of the time t, and replace our condition between $\mathbf{b}_{k+1}$ and $A\mathbf{v}_k$ with a differential relationship viewed as functions of time.

Eigenvalues and Eigenvectors

Definition 3.4 Eigenvalues and Eigenvectors of a Matrix

If A is an $n \times n$ matrix, a number λ is called an ***eigenvalue*** *of A if*

$$A\mathbf{x} = \lambda \mathbf{x} \text{ for some column } \mathbf{x} \neq \mathbf{0} \text{ in } \mathbb{R}^n$$

In this case, $\mathbf{x}$ is called an ***eigenvector*** *of A corresponding to the eigenvalue λ, or a* ***λ-eigenvector*** *for short.*

Example 3.3.2

If $A = \left[\begin{array}{rr} 3 & 5 \\ 1 & -1 \end{array} \right]$ and $\mathbf{x} = \left[\begin{array}{r} 5 \\ 1 \end{array} \right]$ then $A\mathbf{x} = 4\mathbf{x}$ so $\lambda = 4$ is an eigenvalue of A with corresponding eigenvector $\mathbf{x}$.

The matrix A in Example 3.3.2 has another eigenvalue in addition to $\lambda = 4$. To find it, we develop a general procedure for *any* $n \times n$ matrix A.

By definition a number λ is an eigenvalue of the $n \times n$ matrix A if and only if $A\mathbf{x} = \lambda\mathbf{x}$ for some column $\mathbf{x} \neq \mathbf{0}$. This is equivalent to asking that the homogeneous system

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

of linear equations has a nontrivial solution $\mathbf{x} \neq \mathbf{0}$. By Theorem 2.4.5 this happens if and only if the matrix $\lambda I - A$ is not invertible and this, in turn, holds if and only if the determinant of the coefficient matrix is zero:

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

This last condition prompts the following definition:

Definition 3.5 Characteristic Polynomial of a Matrix

If A is an $n \times n$ matrix, the ***characteristic polynomial*** *$c_A(x)$ of A is defined by*

$$c_A(x) = \det(xI - A)$$

Note that $c_A(x)$ is indeed a polynomial in the variable x, and it has degree n when A is an $n \times n$ matrix (this is illustrated in the examples below). The above discussion shows that a number λ is an eigenvalue of A if and only if $c_A(\lambda) = 0$, that is if and only if λ is a **root** of the characteristic polynomial $c_A(x)$. We record these observations in

Theorem 3.3.2

Let A be an $n \times n$ matrix.

1. *The eigenvalues λ of A are the roots of the characteristic polynomial $c_A(x)$ of A.*
2. *The λ-eigenvectors $\mathbf{x}$ are the nonzero solutions to the homogeneous system*
$$(\lambda I - A)\mathbf{x} = \mathbf{0}$$
of linear equations with $\lambda I - A$ as coefficient matrix.

In practice, solving the equations in part 2 of Theorem 3.3.2 is a routine application of gaussian elimination, but finding the eigenvalues can be difficult, often requiring computers (see Section 8.5). For now, the examples and exercises will be constructed so that the roots of the characteristic polynomials are relatively easy to find (usually integers). However, the reader should not be misled by this into thinking that eigenvalues are so easily obtained for the matrices that occur in practical applications!

Example 3.3.3

Find the characteristic polynomial of the matrix $A = \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix}$ discussed in Example 3.3.2, and then find all the eigenvalues and their eigenvectors.

Solution. Since $xI - A = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} - \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} x-3 & -5 \\ -1 & x+1 \end{bmatrix}$ we get

$$c_A(x) = \det \begin{bmatrix} x-3 & -5 \\ -1 & x+1 \end{bmatrix} = x^2 - 2x - 8 = (x-4)(x+2)$$

Hence, the roots of $c_A(x)$ are $\lambda_1 = 4$ and $\lambda_2 = -2$, so these are the eigenvalues of A. Note that $\lambda_1 = 4$ was the eigenvalue mentioned in Example 3.3.2, but we have found a new one: $\lambda_2 = -2$.

To find the eigenvectors corresponding to $\lambda_2 = -2$, observe that in this case

$$(\lambda_2 I - A)\mathbf{x} = \begin{bmatrix} \lambda_2 - 3 & -5 \\ -1 & \lambda_2 + 1 \end{bmatrix} = \begin{bmatrix} -5 & -5 \\ -1 & -1 \end{bmatrix}$$

so the general solution to $(\lambda_2 I - A)\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ where t is an arbitrary real number.

Hence, the eigenvectors $\mathbf{x}$ corresponding to λ_2 are $\mathbf{x} = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ where $t \neq 0$ is arbitrary.

Similarly, $\lambda_1 = 4$ gives rise to the eigenvectors $\mathbf{x} = t \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, $t \neq 0$ which includes the observation in Example 3.3.2.

Note that a square matrix A has *many* eigenvectors associated with any given eigenvalue λ.

In fact *every* nonzero solution $\mathbf{x}$ of $(\lambda I - A)\mathbf{x} = \mathbf{0}$ is an eigenvector. Recall that these solutions are all linear combinations of certain basic solutions determined by the gaussian algorithm (see Theorem 1.3.2). Observe that any nonzero multiple of an eigenvector is again an eigenvector,[9] and such multiples are often more convenient.[10] Any set of nonzero multiples of the basic solutions of $(\lambda I - A)\mathbf{x} = \mathbf{0}$ will be called a set of **basic eigenvectors** corresponding to λ.

Example 3.3.4

Find the characteristic polynomial, eigenvalues, and basic eigenvectors for

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

Solution. Here the characteristic polynomial is given by

$$c_A(x) = \det \begin{bmatrix} x-2 & 0 & 0 \\ -1 & x-2 & 1 \\ -1 & -3 & x+2 \end{bmatrix} = (x-2)(x-1)(x+1)$$

so the eigenvalues are $\lambda_1 = 2$, $\lambda_2 = 1$, and $\lambda_3 = -1$. To find all eigenvectors for $\lambda_1 = 2$, compute

$$\lambda_1 I - A = \begin{bmatrix} \lambda_1 - 2 & 0 & 0 \\ -1 & \lambda_1 - 2 & 1 \\ -1 & -3 & \lambda_1 + 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & -3 & 4 \end{bmatrix}$$

We want the (nonzero) solutions to $(\lambda_1 I - A)\mathbf{x} = \mathbf{0}$. The augmented matrix becomes

$$\left[\begin{array}{rrr|r} 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & -3 & 4 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

using row operations. Hence, the general solution $\mathbf{x}$ to $(\lambda_1 I - A)\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ where t is arbitrary, so we can use $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as the basic eigenvector corresponding to $\lambda_1 = 2$. As the reader can verify, the gaussian algorithm gives basic eigenvectors $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ \frac{1}{3} \\ 1 \end{bmatrix}$ corresponding to $\lambda_2 = 1$ and $\lambda_3 = -1$, respectively. Note that to eliminate fractions, we could instead use $3\mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$ as the basic λ_3-eigenvector.

[9] In fact, any nonzero linear combination of λ-eigenvectors is again a λ-eigenvector.

[10] Allowing nonzero multiples helps eliminate round-off error when the eigenvectors involve fractions.

Example 3.3.5

If A is a square matrix, show that A and A^T have the same characteristic polynomial, and hence the same eigenvalues.

Solution. We use the fact that $xI - A^T = (xI - A)^T$. Then

$$c_{A^T}(x) = \det\left(xI - A^T\right) = \det\left[(xI - A)^T\right] = \det(xI - A) = c_A(x)$$

by Theorem 3.2.3. Hence $c_{A^T}(x)$ and $c_A(x)$ have the same roots, and so A^T and A have the same eigenvalues (by Theorem 3.3.2).

The eigenvalues of a matrix need not be distinct. For example, if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ the characteristic polynomial is $(x-1)^2$ so the eigenvalue 1 occurs twice. Furthermore, eigenvalues are usually not computed as the roots of the characteristic polynomial. There are iterative, numerical methods (for example the QR-algorithm in Section 8.5) that are much more efficient for large matrices.

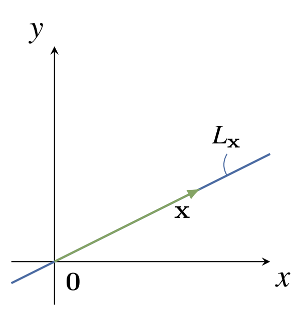
A-Invariance

If A is a 2×2 matrix, we can describe the eigenvectors of A geometrically using the following concept. A line L through the origin in $\mathbb{R}^2$ is called **A-invariant** if $A\mathbf{x}$ is in L whenever $\mathbf{x}$ is in L. If we think of A as a linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$, this asks that A carries L into itself, that is the image $A\mathbf{x}$ of each vector $\mathbf{x}$ in L is again in L.

Example 3.3.6

The x axis $L = \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \mid x \text{ in } \mathbb{R} \right\}$ is A-invariant for any matrix of the form

$$A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \text{ because } \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} ax \\ 0 \end{bmatrix} \text{ is } L \text{ for all } \mathbf{x} = \begin{bmatrix} x \\ 0 \end{bmatrix} \text{ in } L$$



To see the connection with eigenvectors, let $\mathbf{x} \neq \mathbf{0}$ be any nonzero vector in $\mathbb{R}^2$ and let $L_\mathbf{x}$ denote the unique line through the origin containing $\mathbf{x}$ (see the diagram). By the definition of scalar multiplication in Section 2.6, we see that $L_\mathbf{x}$ consists of all scalar multiples of $\mathbf{x}$, that is

$$L_\mathbf{x} = \mathbb{R}\mathbf{x} = \{t\mathbf{x} \mid t \text{ in } \mathbb{R}\}$$

Now suppose that $\mathbf{x}$ is an eigenvector of A, say $A\mathbf{x} = \lambda\mathbf{x}$ for some λ in $\mathbb{R}$. Then if $t\mathbf{x}$ is in $L_\mathbf{x}$ then

$$A(t\mathbf{x}) = t(A\mathbf{x}) = t(\lambda\mathbf{x}) = (t\lambda)\mathbf{x} \text{ is again in } L_\mathbf{x}$$

That is, $L_\mathbf{x}$ is A-invariant. On the other hand, if $L_\mathbf{x}$ is A-invariant then $A\mathbf{x}$ is in $L_\mathbf{x}$ (since $\mathbf{x}$ is in $L_\mathbf{x}$). Hence $A\mathbf{x} = t\mathbf{x}$ for some t in $\mathbb{R}$, so $\mathbf{x}$ is an eigenvector for A (with eigenvalue t). This proves:

Theorem 3.3.3

Let A be a 2×2 matrix, let $\mathbf{x} \neq \mathbf{0}$ be a vector in $\mathbb{R}^2$, and let $L_{\mathbf{x}}$ be the line through the origin in $\mathbb{R}^2$ containing $\mathbf{x}$. Then

$$\mathbf{x} \text{ is an eigenvector of } A \quad \text{if and only if} \quad L_{\mathbf{x}} \text{ is } A\text{-invariant}$$

Example 3.3.7

1. If θ is not a multiple of π, show that $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ has no real eigenvalue.

2. If m is real show that $B = \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}$ has a 1 as an eigenvalue.

Solution.

1. A induces rotation about the origin through the angle θ (Theorem 2.6.4). Since θ is not a multiple of π, this shows that no line through the origin is A-invariant. Hence A has no eigenvector by Theorem 3.3.3, and so has no eigenvalue.

2. B induces reflection Q_m in the line through the origin with slope m by Theorem 2.6.5. If $\mathbf{x}$ is any nonzero point on this line then it is clear that $Q_m\mathbf{x} = \mathbf{x}$, that is $Q_m\mathbf{x} = 1\mathbf{x}$. Hence 1 is an eigenvalue (with eigenvector $\mathbf{x}$).

If $\theta = \frac{\pi}{2}$ in Example 3.3.7, then $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ so $c_A(x) = x^2 + 1$. This polynomial has no root in $\mathbb{R}$, so A has no (real) eigenvalue, and hence no eigenvector. In fact its eigenvalues are the complex numbers i and $-i$, with corresponding eigenvectors $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$ In other words, A *has* eigenvalues and eigenvectors, just not real ones.

Note that *every* polynomial has complex roots,[11] so every matrix has complex eigenvalues. While these eigenvalues may very well be real, this suggests that we really should be doing linear algebra over the complex numbers. Indeed, everything we have done (gaussian elimination, matrix algebra, determinants, etc.) works if all the scalars are complex.

[11] This is called the *Fundamental Theorem of Algebra* and was first proved by Gauss in his doctoral dissertation.

Diagonalization

An $n \times n$ matrix D is called a **diagonal matrix** if all its entries off the main diagonal are zero, that is if D has the form

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \text{diag}\,(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n)$$

where $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n$ are numbers. Calculations with diagonal matrices are very easy. Indeed, if $D = \text{diag}\,(\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n)$ and $E = \text{diag}\,(\mu_1,\ \mu_2,\ \ldots,\ \mu_n)$ are two diagonal matrices, their product DE and sum $D+E$ are again diagonal, and are obtained by doing the same operations to corresponding diagonal elements:

$$\begin{aligned} DE &= \text{diag}\,(\lambda_1\mu_1,\ \lambda_2\mu_2,\ \ldots,\ \lambda_n\mu_n) \\ D+E &= \text{diag}\,(\lambda_1+\mu_1,\ \lambda_2+\mu_2,\ \ldots,\ \lambda_n+\mu_n) \end{aligned}$$

Because of the simplicity of these formulas, and with an eye on Theorem 3.3.1 and the discussion preceding it, we make another definition:

Definition 3.6 Diagonalizable Matrices

*An $n \times n$ matrix A is called **diagonalizable** if*

$$P^{-1}AP \text{ is diagonal for some invertible } n \times n \text{ matrix } P$$

*Here the invertible matrix P is called a **diagonalizing matrix** for A.*

To discover when such a matrix P exists, we let $\mathbf{x}_1,\ \mathbf{x}_2,\ \ldots,\ \mathbf{x}_n$ denote the columns of P and look for ways to determine when such $\mathbf{x}_i$ exist and how to compute them. To this end, write P in terms of its columns as follows:

$$P = [\mathbf{x}_1,\ \mathbf{x}_2,\ \cdots,\ \mathbf{x}_n]$$

Observe that $P^{-1}AP = D$ for some diagonal matrix D holds if and only if

$$AP = PD$$

If we write $D = \text{diag}\,(\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n)$, where the λ_i are numbers to be determined, the equation $AP = PD$ becomes

$$A\,[\mathbf{x}_1,\ \mathbf{x}_2,\ \cdots,\ \mathbf{x}_n] = [\mathbf{x}_1,\ \mathbf{x}_2,\ \cdots,\ \mathbf{x}_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

By the definition of matrix multiplication, each side simplifies as follows

$$\begin{bmatrix} A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \cdots & \lambda_n\mathbf{x}_n \end{bmatrix}$$

Comparing columns shows that $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$ for each i, so

$$P^{-1}AP = D \quad \text{ if and only if } A\mathbf{x}_i = \lambda_i \mathbf{x}_i \text{ for each } i$$

In other words, $P^{-1}AP = D$ holds if and only if the diagonal entries of D are eigenvalues of A and the columns of P are corresponding eigenvectors. This proves the following fundamental result.

Theorem 3.3.4

Let A be an $n \times n$ matrix.

1. *A is diagonalizable if and only if it has eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ such that the matrix $P = \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{array} \right]$ is invertible.*

2. *When this is the case, $P^{-1}AP = \mathbf{diag}\,(\lambda_1, \lambda_2, \dots, \lambda_n)$ where, for each i, λ_i is the eigenvalue of A corresponding to $\mathbf{x}_i$.*

Example 3.3.8

Diagonalize the matrix $A = \left[\begin{array}{rrr} 2 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 3 & -2 \end{array} \right]$ in Example 3.3.4.

Solution. By Example 3.3.4, the eigenvalues of A are $\lambda_1 = 2$, $\lambda_2 = 1$, and $\lambda_3 = -1$, with corresponding basic eigenvectors $\mathbf{x}_1 = \left[\begin{array}{r} 1 \\ 1 \\ 1 \end{array} \right]$, $\mathbf{x}_2 = \left[\begin{array}{r} 0 \\ 1 \\ 1 \end{array} \right]$, and $\mathbf{x}_3 = \left[\begin{array}{r} 0 \\ 1 \\ 3 \end{array} \right]$ respectively. Since the matrix $P = \left[\begin{array}{ccc} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{array} \right] = \left[\begin{array}{rrr} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 3 \end{array} \right]$ is invertible, Theorem 3.3.4 guarantees that

$$P^{-1}AP = \left[\begin{array}{ccc} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{array} \right] = \left[\begin{array}{rrr} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{array} \right] = D$$

The reader can verify this directly—easier to check $AP = PD$.

In Example 3.3.8, suppose we let $Q = \left[\begin{array}{ccc} \mathbf{x}_2 & \mathbf{x}_1 & \mathbf{x}_3 \end{array} \right]$ be the matrix formed from the eigenvectors $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ of A, but in a *different order* than that used to form P. Then $Q^{-1}AQ = \mathbf{diag}\,(\lambda_2, \lambda_1, \lambda_3)$ is diagonal by Theorem 3.3.4, but the eigenvalues are in the *new* order. Hence we can choose the diagonalizing matrix P so that the eigenvalues λ_i appear in any order we want along the main diagonal of D.

In every example above each eigenvalue has had only one basic eigenvector. Here is a diagonalizable matrix where this is not the case.

Example 3.3.9

Diagonalize the matrix $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

Solution. To compute the characteristic polynomial of A first add rows 2 and 3 of $xI - A$ to row 1:

$$\begin{aligned} c_A(x) &= \det \begin{bmatrix} x & -1 & -1 \\ -1 & x & -1 \\ -1 & -1 & x \end{bmatrix} = \det \begin{bmatrix} x-2 & x-2 & x-2 \\ -1 & x & -1 \\ -1 & -1 & x \end{bmatrix} \\ &= \det \begin{bmatrix} x-2 & 0 & 0 \\ -1 & x+1 & 0 \\ -1 & 0 & x+1 \end{bmatrix} = (x-2)(x+1)^2 \end{aligned}$$

Hence the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = -1$, with λ_2 repeated twice (we say that λ_2 has *multiplicity* two). However, A *is* diagonalizable. For $\lambda_1 = 2$, the system of equations $(\lambda_1 I - A)\mathbf{x} = \mathbf{0}$ has general solution $\mathbf{x} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as the reader can verify, so a basic λ_1-eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Turning to the repeated eigenvalue $\lambda_2 = -1$, we must solve $(\lambda_2 I - A)\mathbf{x} = \mathbf{0}$. By gaussian elimination, the general solution is $\mathbf{x} = s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ where s and t are arbitrary. Hence the gaussian algorithm produces *two* basic λ_2-eigenvectors $\mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{y}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ If we take $P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ we find that P is invertible. Hence $P^{-1}AP = \mathbf{diag}\,(2, \ -1, \ -1)$ by Theorem 3.3.4.

Example 3.3.9 typifies every diagonalizable matrix. To describe the general case, we need some terminology.

Definition 3.7 Multiplicity of an Eigenvalue

An eigenvalue λ of a square matrix A is said to have ***multiplicity*** *m if it occurs m times as a root of the characteristic polynomial $c_A(x)$.*

For example, the eigenvalue $\lambda_2 = -1$ in Example 3.3.9 has multiplicity 2. In that example the gaussian algorithm yields two basic λ_2-eigenvectors, the same number as the multiplicity. This works in general.

Theorem 3.3.5

A square matrix A is diagonalizable if and only if every eigenvalue λ of multiplicity m yields exactly m basic eigenvectors; that is, if and only if the general solution of the system $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has exactly m parameters.

One case of Theorem 3.3.5 deserves mention.

Theorem 3.3.6

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

The proofs of Theorem 3.3.5 and Theorem 3.3.6 require more advanced techniques and are given in Chapter 5. The following procedure summarizes the method.

Diagonalization Algorithm

To diagonalize an $n \times n$ matrix A:

Step 1. Find the distinct eigenvalues λ of A.

Step 2. Compute a set of basic eigenvectors corresponding to each of these eigenvalues λ as basic solutions of the homogeneous system $(\lambda I - A)\mathbf{x} = \mathbf{0}$.

Step 3. The matrix A is diagonalizable if and only if there are n basic eigenvectors in all.

Step 4. If A is diagonalizable, the $n \times n$ matrix P with these basic eigenvectors as its columns is a diagonalizing matrix for A, that is, P is invertible and $P^{-1}AP$ is diagonal.

The diagonalization algorithm is valid even if the eigenvalues are nonreal complex numbers. In this case the eigenvectors will also have complex entries, but we will not pursue this here.

Example 3.3.10

Show that $A = \left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right]$ is not diagonalizable.

Solution 1. The characteristic polynomial is $c_A(x) = (x-1)^2$, so A has only one eigenvalue $\lambda_1 = 1$ of multiplicity 2. But the system of equations $(\lambda_1 I - A)\mathbf{x} = \mathbf{0}$ has general solution $t\left[\begin{array}{c} 1 \\ 0 \end{array}\right]$, so there is only one parameter, and so only one basic eigenvector $\left[\begin{array}{c} 1 \\ 2 \end{array}\right]$. Hence A is not diagonalizable.

Solution 2. We have $c_A(x) = (x-1)^2$ so the only eigenvalue of A is $\lambda = 1$. Hence, if A were diagonalizable, Theorem 3.3.4 would give $P^{-1}AP = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] = I$ for some invertible matrix

P. But then $A = PIP^{-1} = I$, which is not the case. So A cannot be diagonalizable.

Diagonalizable matrices share many properties of their eigenvalues. The following example illustrates why.

Example 3.3.11

If $\lambda^3 = 5\lambda$ for every eigenvalue of the diagonalizable matrix A, show that $A^3 = 5A$.

Solution. Let $P^{-1}AP = D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Because $\lambda_i^3 = 5\lambda_i$ for each i, we obtain

$$D^3 = \text{diag}(\lambda_1^3, \dots, \lambda_n^3) = \text{diag}(5\lambda_1, \dots, 5\lambda_n) = 5D$$

Hence $A^3 = (PDP^{-1})^3 = PD^3P^{-1} = P(5D)P^{-1} = 5(PDP^{-1}) = 5A$ using Theorem 3.3.1. This is what we wanted.

If $p(x)$ is any polynomial and $p(\lambda) = 0$ for every eigenvalue of the diagonalizable matrix A, an argument similar to that in Example 3.3.11 shows that $p(A) = 0$. Thus Example 3.3.11 deals with the case $p(x) = x^3 - 5x$. In general, $p(A)$ is called the *evaluation* of the polynomial $p(x)$ at the matrix A. For example, if $p(x) = 2x^3 - 3x + 5$, then $p(A) = 2A^3 - 3A + 5I$—note the use of the identity matrix.

In particular, if $c_A(x)$ denotes the characteristic polynomial of A, we certainly have $c_A(\lambda) = 0$ for each eigenvalue λ of A (Theorem 3.3.2). Hence $c_A(A) = 0$ for every diagonalizable matrix A. This is, in fact, true for *any* square matrix, diagonalizable or not, and the general result is called the Cayley-Hamilton theorem. It is proved in Section **??** and again in Section **??**.

Linear Dynamical Systems

We began Section 3.3 with an example from ecology which models the evolution of the population of a species of birds as time goes on. As promised, we now complete the example—Example 3.3.12 below.

The bird population was described by computing the female population profile $\mathbf{v}_k = \left[\begin{array}{c} a_k \\ j_k \end{array} \right]$ of the species, where a_k and j_k represent the number of adult and juvenile females present k years after the initial values a_0 and j_0 were observed. The model assumes that these numbers are related by the following equations:

$$\begin{aligned} a_{k+1} &= \tfrac{1}{2}a_k + \tfrac{1}{4}j_k \\ j_{k+1} &= 2a_k \end{aligned}$$

If we write $A = \left[\begin{array}{cc} \frac{1}{2} & \frac{1}{4} \\ 2 & 0 \end{array} \right]$ the columns $\mathbf{v}_k$ satisfy $\mathbf{v}_{k+1} = A\mathbf{v}_k$ for each $k = 0, 1, 2, \dots$.

Hence $\mathbf{v}_k = A^k\mathbf{v}_0$ for each $k = 1, 2, \dots$. We can now use our diagonalization techniques to determine the population profile $\mathbf{v}_k$ for all values of k in terms of the initial values.

Example 3.3.12

Assuming that the initial values were $a_0 = 100$ adult females and $j_0 = 40$ juvenile females, compute a_k and j_k for $k = 1, 2, \dots$.

Solution. The characteristic polynomial of the matrix $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ 2 & 0 \end{bmatrix}$ is $c_A(x) = x^2 - \frac{1}{2}x - \frac{1}{2} = (x-1)(x+\frac{1}{2})$, so the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -\frac{1}{2}$ and gaussian elimination gives corresponding basic eigenvectors $\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -\frac{1}{4} \\ 1 \end{bmatrix}$. For convenience, we can use multiples $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$ respectively. Hence a diagonalizing matrix is $P = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$ and we obtain

$$P^{-1}AP = D \text{ where } D = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}$$

This gives $A = PDP^{-1}$ so, for each $k \geq 0$, we can compute A^k explicitly:

$$\begin{aligned} A^k = PD^kP^{-1} &= \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & (-\frac{1}{2})^k \end{bmatrix} \frac{1}{6} \begin{bmatrix} 4 & 1 \\ -2 & 4 \end{bmatrix} \\ &= \frac{1}{6} \begin{bmatrix} 4+2(-\frac{1}{2})^k & 1-(-\frac{1}{2})^k \\ 8-8(-\frac{1}{2})^k & 2+4(-\frac{1}{2})^k \end{bmatrix} \end{aligned}$$

Hence we obtain

$$\begin{aligned} \begin{bmatrix} a_k \\ j_k \end{bmatrix} = \mathbf{v}_k = A^k\mathbf{v}_0 &= \frac{1}{6} \begin{bmatrix} 4+2(-\frac{1}{2})^k & 1-(-\frac{1}{2})^k \\ 8-8(-\frac{1}{2})^k & 2+4(-\frac{1}{2})^k \end{bmatrix} \begin{bmatrix} 100 \\ 40 \end{bmatrix} \\ &= \frac{1}{6} \begin{bmatrix} 440+160(-\frac{1}{2})^k \\ 880-640(-\frac{1}{2})^k \end{bmatrix} \end{aligned}$$

Equating top and bottom entries, we obtain exact formulas for a_k and j_k:

$$a_k = \frac{220}{3} + \frac{80}{3}\left(-\frac{1}{2}\right)^k \text{ and } j_k = \frac{440}{3} + \frac{320}{3}\left(-\frac{1}{2}\right)^k \text{ for } k = 1, 2, \cdots$$

In practice, the exact values of a_k and j_k are not usually required. What is needed is a measure of how these numbers behave for large values of k. This is easy to obtain here. Since $(-\frac{1}{2})^k$ is nearly zero for large k, we have the following approximate values

$$a_k \approx \frac{220}{3} \text{ and } j_k \approx \frac{440}{3} \text{ if } k \text{ is large}$$

Hence, in the long term, the female population stabilizes with approximately twice as many juveniles as adults.

Definition 3.8 Linear Dynamical System

If A is an $n \times n$ matrix, a sequence $\mathbf{v}_0,\ \mathbf{v}_1,\ \mathbf{v}_2,\ \dots$ of columns in $\mathbb{R}^n$ is called a ***linear dynamical system*** *if $\mathbf{v}_0$ is specified and $\mathbf{v}_1,\ \mathbf{v}_2,\ \dots$ are given by the matrix recurrence $\mathbf{v}_{k+1} = A\mathbf{v}_k$ for each $k \geq 0$. We call A the* ***migration*** *matrix of the system.*

We have $\mathbf{v}_1 = A\mathbf{v}_0$, then $\mathbf{v}_2 = A\mathbf{v}_1 = A^2\mathbf{v}_0$, and continuing we find

$$\mathbf{v}_k = A^k\mathbf{v}_0 \text{ for each } k = 1,\ 2,\ \cdots \tag{3.9}$$

Hence the columns $\mathbf{v}_k$ are determined by the powers A^k of the matrix A and, as we have seen, these powers can be efficiently computed if A is diagonalizable. In fact Equation 3.9 can be used to give a nice "formula" for the columns $\mathbf{v}_k$ in this case.

Assume that A is diagonalizable with eigenvalues $\lambda_1,\ \lambda_2,\ \dots,\ \lambda_n$ and corresponding basic eigenvectors $\mathbf{x}_1,\ \mathbf{x}_2,\ \dots,\ \mathbf{x}_n$. If $P = \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{array}\right]$ is a diagonalizing matrix with the $\mathbf{x}_i$ as columns, then P is invertible and

$$P^{-1}AP = D = \text{diag}\,(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n)$$

by Theorem 3.3.4. Hence $A = PDP^{-1}$ so Equation 3.9 and Theorem 3.3.1 give

$$\mathbf{v}_k = A^k\mathbf{v}_0 = (PDP^{-1})^k\mathbf{v}_0 = (PD^kP^{-1})\mathbf{v}_0 = PD^k(P^{-1}\mathbf{v}_0)$$

for each $k = 1,\ 2,\ \dots$. For convenience, we denote the column $P^{-1}\mathbf{v}_0$ arising here as follows:

$$\mathbf{b} = P^{-1}\mathbf{v}_0 = \left[\begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_n \end{array}\right]$$

Then matrix multiplication gives

$$\begin{aligned}
\mathbf{v}_k &= PD^k(P^{-1}\mathbf{v}_0) \\
&= \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{array}\right] \left[\begin{array}{cccc} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{array}\right] \left[\begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_n \end{array}\right] \\
&= \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{array}\right] \left[\begin{array}{c} b_1\lambda_1^k \\ b_2\lambda_2^k \\ \vdots \\ b_3\lambda_n^k \end{array}\right] \\
&= b_1\lambda_1^k\mathbf{x}_1 + b_2\lambda_2^k\mathbf{x}_2 + \cdots + b_n\lambda_n^k\mathbf{x}_n
\end{aligned} \tag{3.10}$$

for each $k \geq 0$. This is a useful **exact formula** for the columns $\mathbf{v}_k$. Note that, in particular,

$$\mathbf{v}_0 = b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \cdots + b_n\mathbf{x}_n$$

However, such an exact formula for $\mathbf{v}_k$ is often not required in practice; all that is needed is to *estimate* $\mathbf{v}_k$ for large values of k (as was done in Example 3.3.12). This can be easily done if A has a largest eigenvalue. An eigenvalue λ of a matrix A is called a **dominant eigenvalue** of A if it has multiplicity 1 and

$$|\lambda| > |\mu| \text{ for all eigenvalues } \mu \neq \lambda$$

where $|\lambda|$ denotes the absolute value of the number λ. For example, $\lambda_1 = 1$ is dominant in Example 3.3.12.

Returning to the above discussion, suppose that A has a dominant eigenvalue. By choosing the order in which the columns $\mathbf{x}_i$ are placed in P, we may assume that λ_1 is dominant among the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of A (see the discussion following Example 3.3.8). Now recall the exact expression for $\mathbf{v}_k$ in Equation 3.10 above:

$$\mathbf{v}_k = b_1\lambda_1^k\mathbf{x}_1 + b_2\lambda_2^k\mathbf{x}_2 + \cdots + b_n\lambda_n^k\mathbf{x}_n$$

Take λ_1^k out as a common factor in this equation to get

$$\mathbf{v}_k = \lambda_1^k\left[b_1\mathbf{x}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{x}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{x}_n\right]$$

for each $k \geq 0$. Since λ_1 is dominant, we have $|\lambda_i| < |\lambda_1|$ for each $i \geq 2$, so each of the numbers $(\lambda_i/\lambda_1)^k$ become small in absolute value as k increases. Hence $\mathbf{v}_k$ is approximately equal to the first term $\lambda_1^k b_1\mathbf{x}_1$, and we write this as $\mathbf{v}_k \approx \lambda_1^k b_1\mathbf{x}_1$. These observations are summarized in the following theorem (together with the above exact formula for $\mathbf{v}_k$).

Theorem 3.3.7

Consider the dynamical system $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots$ with matrix recurrence

$$\mathbf{v}_{k+1} = A\mathbf{v}_k \text{ for } k \geq 0$$

where A and $\mathbf{v}_0$ are given. Assume that A is a diagonalizable $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ and corresponding basic eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, and let $P = \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{array}\right]$ be the diagonalizing matrix. Then an exact formula for $\mathbf{v}_k$ is

$$\mathbf{v}_k = b_1\lambda_1^k\mathbf{x}_1 + b_2\lambda_2^k\mathbf{x}_2 + \cdots + b_n\lambda_n^k\mathbf{x}_n \text{ for each } k \geq 0$$

where the coefficients b_i come from

$$\mathbf{b} = P^{-1}\mathbf{v}_0 = \left[\begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_n \end{array}\right]$$

Moreover, if A has dominant[12] eigenvalue λ_1, then $\mathbf{v}_k$ is approximated by

$$\mathbf{v}_k = b_1\lambda_1^k\mathbf{x}_1 \text{ for sufficiently large } k.$$

Example 3.3.13

Returning to Example 3.3.12, we see that $\lambda_1 = 1$ is the dominant eigenvalue, with eigenvector $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Here $P = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$ and $\mathbf{v}_0 = \begin{bmatrix} 100 \\ 40 \end{bmatrix}$ so $P^{-1}\mathbf{v}_0 = \frac{1}{3}\begin{bmatrix} 220 \\ -80 \end{bmatrix}$. Hence $b_1 = \frac{220}{3}$ in the notation of Theorem 3.3.7, so

$$\begin{bmatrix} a_k \\ j_k \end{bmatrix} = \mathbf{v}_k \approx b_1\lambda_1^k\mathbf{x}_1 = \tfrac{220}{3}1^k\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

where k is large. Hence $a_k \approx \frac{220}{3}$ and $j_k \approx \frac{440}{3}$ as in Example 3.3.12.

This next example uses Theorem 3.3.7 to solve a "linear recurrence." See also Section ??.

Example 3.3.14

Suppose a sequence $x_0,\ x_1,\ x_2,\ \dots$ is determined by insisting that

$$x_0 = 1,\ x_1 = -1,\ \text{ and } x_{k+2} = 2x_k - x_{k+1} \text{ for every } k \geq 0$$

Find a formula for x_k in terms of k.

Solution. Using the linear recurrence $x_{k+2} = 2x_k - x_{k+1}$ repeatedly gives

$$x_2 = 2x_0 - x_1 = 3, \quad x_3 = 2x_1 - x_2 = -5, \quad x_4 = 11, \quad x_5 = -21,\ \dots$$

so the x_i are determined but no pattern is apparent. The idea is to find $\mathbf{v}_k = \begin{bmatrix} x_k \\ x_{k+1} \end{bmatrix}$ for each k instead, and then retrieve x_k as the top component of $\mathbf{v}_k$. The reason this works is that the linear recurrence guarantees that these $\mathbf{v}_k$ are a dynamical system:

$$\mathbf{v}_{k+1} = \begin{bmatrix} x_{k+1} \\ x_{k+2} \end{bmatrix} = \begin{bmatrix} x_{k+1} \\ 2x_k - x_{k+1} \end{bmatrix} = A\mathbf{v}_k \text{ where } A = \begin{bmatrix} 0 & 1 \\ 2 & -1 \end{bmatrix}$$

The eigenvalues of A are $\lambda_1 = -2$ and $\lambda_2 = 1$ with eigenvectors $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, so the diagonalizing matrix is $P = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$.

Moreover, $\mathbf{b} = P_0^{-1}\mathbf{v}_0 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ so the exact formula for $\mathbf{v}_k$ is

$$\begin{bmatrix} x_k \\ x_{k+1} \end{bmatrix} = \mathbf{v}_k = b_1\lambda_1^k\mathbf{x}_1 + b_2\lambda_2^k\mathbf{x}_2 = \tfrac{2}{3}(-2)^k\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \tfrac{1}{3}1^k\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Equating top entries gives the desired formula for x_k:

$$x_k = \tfrac{1}{3}\left[2(-2)^k + 1\right] \text{ for all } k = 0,\ 1,\ 2,\ \dots$$

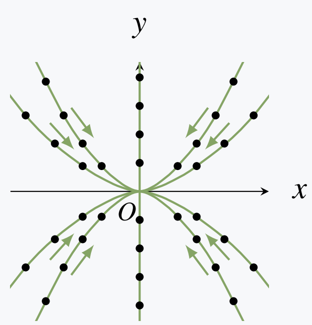
The reader should check this for the first few values of k.

[12]Similar results can be found in other situations. If for example, eigenvalues λ_1 and λ_2 (possibly equal) satisfy $|\lambda_1| = |\lambda_2| > |\lambda_i|$ for all $i > 2$, then we obtain $\mathbf{v}_k \approx b_1\lambda_1^k x_1 + b_2\lambda_2^k x_2$ for large k.

Graphical Description of Dynamical Systems

If a dynamical system $\mathbf{v}_{k+1} = A\mathbf{v}_k$ is given, the sequence $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots$ is called the **trajectory** of the system starting at $\mathbf{v}_0$. It is instructive to obtain a graphical plot of the system by writing $\mathbf{v}_k = \left[\begin{array}{c} x_k \\ y_k \end{array}\right]$ and plotting the successive values as points in the plane, identifying $\mathbf{v}_k$ with the point (x_k, y_k) in the plane. We give several examples which illustrate properties of dynamical systems. For ease of calculation we assume that the matrix A is simple, usually diagonal.

Example 3.3.15

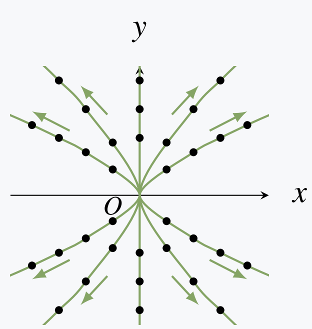


Let $A = \left[\begin{array}{cc} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{array}\right]$ Then the eigenvalues are $\frac{1}{2}$ and $\frac{1}{3}$, with corresponding eigenvectors $\mathbf{x}_1 = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$ and $\mathbf{x}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$. The exact formula is

$$\mathbf{v}_k = b_1 \left(\tfrac{1}{2}\right)^k \left[\begin{array}{c} 1 \\ 0 \end{array}\right] + b_2 \left(\tfrac{1}{3}\right)^k \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$$

for $k = 0, 1, 2, \dots$ by Theorem 3.3.7, where the coefficients b_1 and b_2 depend on the initial point $\mathbf{v}_0$. Several trajectories are plotted in the diagram and, for each choice of $\mathbf{v}_0$, the trajectories converge toward the origin because both eigenvalues are less than 1 in absolute value. For this reason, the origin is called an **attractor** for the system.

Example 3.3.16

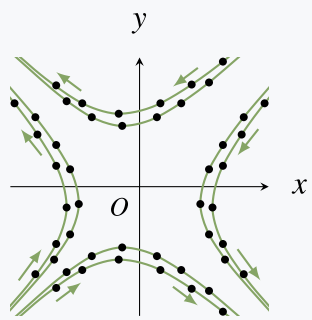


Let $A = \left[\begin{array}{cc} \frac{3}{2} & 0 \\ 0 & \frac{4}{3} \end{array}\right]$. Here the eigenvalues are $\frac{3}{2}$ and $\frac{4}{3}$, with corresponding eigenvectors $\mathbf{x}_1 = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$ and $\mathbf{x}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$ as before. The exact formula is

$$\mathbf{v}_k = b_1 \left(\tfrac{3}{2}\right)^k \left[\begin{array}{c} 1 \\ 0 \end{array}\right] + b_2 \left(\tfrac{4}{3}\right)^k \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$$

for $k = 0, 1, 2, \dots$. Since both eigenvalues are greater than 1 in absolute value, the trajectories diverge away from the origin for every choice of initial point V_0. For this reason, the origin is called a **repellor** for the system.

Example 3.3.17



Let $A = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$. Now the eigenvalues are $\frac{3}{2}$ and $\frac{1}{2}$, with corresponding eigenvectors $\mathbf{x}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The exact formula is

$$\mathbf{v}_k = b_1 \left(\tfrac{3}{2}\right)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix} + b_2 \left(\tfrac{1}{2}\right)^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

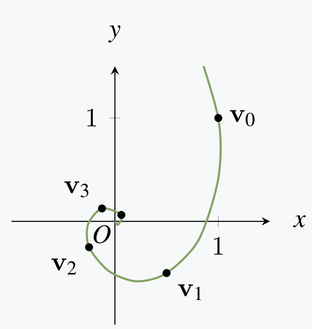
for $k = 0, 1, 2, \dots$. In this case $\frac{3}{2}$ is the dominant eigenvalue so, if $b_1 \neq 0$, we have $\mathbf{v}_k \approx b_1 \left(\frac{3}{2}\right)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ for large k and $\mathbf{v}_k$ is approaching the line $y = -x$.

However, if $b_1 = 0$, then $\mathbf{v}_k = b_2 \left(\frac{1}{2}\right)^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and so approaches the origin along the line $y = x$. In general the trajectories appear as in the diagram, and the origin is called a **saddle point** for the dynamical system in this case.

Example 3.3.18

Let $A = \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}$. Now the characteristic polynomial is $c_A(x) = x^2 + \frac{1}{4}$, so the eigenvalues are the complex numbers $\frac{i}{2}$ and $-\frac{i}{2}$ where $i^2 = -1$. Hence A is not diagonalizable as a real matrix. However, the trajectories are not difficult to describe. If we start with $\mathbf{v}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ then the trajectory begins as

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\frac{1}{4} \\ -\frac{1}{4} \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -\frac{1}{8} \\ \frac{1}{8} \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} \frac{1}{16} \\ \frac{1}{16} \end{bmatrix}, \mathbf{v}_5 = \begin{bmatrix} \frac{1}{32} \\ -\frac{1}{32} \end{bmatrix}, \mathbf{v}_6 = \begin{bmatrix} -\frac{1}{64} \\ -\frac{1}{64} \end{bmatrix}, \dots$$



The first five of these points are plotted in the diagram. Here each trajectory spirals in toward the origin, so the origin is an attractor. Note that the two (complex) eigenvalues have absolute value less than 1 here. If they had absolute value greater than 1, the trajectories would spiral out from the origin.

Google PageRank

Dominant eigenvalues are useful to the Google search engine for finding information on the Web. If an information query comes in from a client, Google has a sophisticated method of establishing the "relevance" of each site to that query. When the relevant sites have been determined, they are placed in order of importance using a ranking of *all* sites called the PageRank. The relevant sites with the highest PageRank are the ones presented to the client. It is the construction of the PageRank that is our interest here.

The Web contains many links from one site to another. Google interprets a link from site j to site i as a "vote" for the importance of site i. Hence if site i has more links to it than does site j, then i is regarded as more "important" and assigned a higher PageRank. One way to look at this is to view the sites as vertices in a huge directed graph (see Section 2.2). Then if site j links to site i there is an edge from j to i, and hence the (i, j)-entry is a 1 in the associated adjacency matrix (called the *connectivity* matrix in this context). Thus a large number of 1s in row i of this matrix is a measure of the PageRank of site i.[13]

However this does not take into account the PageRank of the sites that link to i. Intuitively, the higher the rank of these sites, the higher the rank of site i. One approach is to compute a dominant eigenvector $\mathbf{x}$ for the connectivity matrix. In most cases the entries of $\mathbf{x}$ can be chosen to be positive with sum 1. Each site corresponds to an entry of $\mathbf{x}$, so the sum of the entries of sites linking to a given site i is a measure of the rank of site i. In fact, Google chooses the PageRank of a site so that it is proportional to this sum.[14]

Exercises for 3.3

Exercise 3.3.1 In each case find the characteristic polynomial, eigenvalues, eigenvectors, and (if possible) an invertible matrix P such that $P^{-1}AP$ is diagonal.

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

c) $A = \begin{bmatrix} 7 & 0 & -4 \\ 0 & 5 & 0 \\ 5 & 0 & -2 \end{bmatrix}$ d) $A = \begin{bmatrix} 1 & 1 & -3 \\ 2 & 0 & 6 \\ 1 & -1 & 5 \end{bmatrix}$

e) $A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 6 & -6 \\ 1 & 2 & -1 \end{bmatrix}$ f) $A = \begin{bmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}$

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

i) $A = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix}$, $\lambda \neq \mu$

---

b. $(x-3)(x+2); 3; -2; \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix};$

$P = \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix}; P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}.$

[13] For more on PageRank, visit https://en.wikipedia.org/wiki/PageRank.

[14] See the articles "Searching the web with eigenvectors" by Herbert S. Wilf, UMAP Journal 23(2), 2002, pages 101–103, and "The worlds largest matrix computation: Google's PageRank is an eigenvector of a matrix of order 2.7 billion" by Cleve Moler, Matlab News and Notes, October 2002, pages 12–13.

d. $(x-2)^3; 2; \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$; No such P; Not diagonalizable.

f. $(x+1)^2(x-2); -1,\ -2; \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$; No such P; Not diagonalizable. Note that this matrix and the matrix in Example 3.3.9 have the same characteristic polynomial, but that matrix is diagonalizable.

h. $(x-1)^2(x-3); 1,\ 3; \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ No such P; Not diagonalizable.

Exercise 3.3.2 Consider a linear dynamical system $\mathbf{v}_{k+1} = A\mathbf{v}_k$ for $k \geq 0$. In each case approximate $\mathbf{v}_k$ using Theorem 3.3.7.

a. $A = \begin{bmatrix} 2 & 1 \\ 4 & -1 \end{bmatrix}$, $\mathbf{v}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

b. $A = \begin{bmatrix} 3 & -2 \\ 2 & -2 \end{bmatrix}$, $\mathbf{v}_0 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$

c. $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 3 \\ 1 & 4 & 1 \end{bmatrix}$, $\mathbf{v}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

d. $A = \begin{bmatrix} 1 & 3 & 2 \\ -1 & 2 & 1 \\ 4 & -1 & -1 \end{bmatrix}$, $\mathbf{v}_0 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$

---

b. $V_k = \frac{7}{3}2^k \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

d. $V_k = \frac{3}{2}3^k \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

Exercise 3.3.3 Show that A has $\lambda = 0$ as an eigenvalue if and only if A is not invertible.

Exercise 3.3.4 Let A denote an $n \times n$ matrix and put $A_1 = A - \alpha I$, α in $\mathbb{R}$. Show that λ is an eigenvalue of A if and only if $\lambda - \alpha$ is an eigenvalue of A_1. (Hence, the eigenvalues of A_1 are just those of A "shifted" by α.) How do the eigenvectors compare?

$A\mathbf{x} = \lambda\mathbf{x}$ if and only if $(A - \alpha I)\mathbf{x} = (\lambda - \alpha)\mathbf{x}$. Same eigenvectors.

Exercise 3.3.5 Show that the eigenvalues of $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ are $e^{i\theta}$ and $e^{-i\theta}$. (See Appendix ??)

Exercise 3.3.6 Find the characteristic polynomial of the $n \times n$ identity matrix I. Show that I has exactly one eigenvalue and find the eigenvectors.

Exercise 3.3.7 Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ show that:

a. $c_A(x) = x^2 - \operatorname{tr} Ax + \det A$, where $\operatorname{tr} A = a + d$ is called the **trace** of A.

b. The eigenvalues are $\frac{1}{2}\left[(a+d) \pm \sqrt{(a-d)^2 + 4bc}\right]$.

Exercise 3.3.8 In each case, find $P^{-1}AP$ and then compute A^n.

a. $A = \begin{bmatrix} 6 & -5 \\ 2 & -1 \end{bmatrix}$, $P = \begin{bmatrix} 1 & 5 \\ 1 & 2 \end{bmatrix}$

b. $A = \begin{bmatrix} -7 & -12 \\ 6 & -10 \end{bmatrix}$, $P = \begin{bmatrix} -3 & 4 \\ 2 & -3 \end{bmatrix}$ [*Hint*: $(PDP^{-1})^n = PD^nP^{-1}$ for each $n = 1,\ 2,\ \dots$.]

---

b. $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, so $A^n = P\begin{bmatrix} 1 & 0 \\ 0 & 2^n \end{bmatrix}P^{-1} = \begin{bmatrix} 9 - 8 \cdot 2^n & 12(1-2^n) \\ 6(2^n - 1) & 9 \cdot 2^n - 8 \end{bmatrix}$

Exercise 3.3.9

a. If $A = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ verify that A and B are diagonalizable, but AB is not.

b. If $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ find a diagonalizable matrix A such that $D + A$ is not diagonalizable.

---

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

Exercise 3.3.10 If A is an $n \times n$ matrix, show that A is diagonalizable if and only if A^T is diagonalizable.

Exercise 3.3.11 If A is diagonalizable, show that each of the following is also diagonalizable.

a. A^n, $n \geq 1$

b. kA, k any scalar.

c. $p(A)$, $p(x)$ any polynomial (Theorem 3.3.1)

d. $U^{-1}AU$ for any invertible matrix U.

e. $kI + A$ for any scalar k.

---

b. and d. $PAP^{-1} = D$ is diagonal, then b. $P^{-1}(kA)P = kD$ is diagonal, and d. $Q(U^{-1}AU)Q = D$ where $Q = PU$.

Exercise 3.3.12 Give an example of two diagonalizable matrices A and B whose sum $A + B$ is not diagonalizable.

$\left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right]$ is not diagonalizable by Example 3.3.8. But $\left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right] = \left[\begin{array}{cc} 2 & 1 \\ 0 & -1 \end{array}\right] + \left[\begin{array}{cc} -1 & 0 \\ 0 & 2 \end{array}\right]$ where $\left[\begin{array}{cc} 2 & 1 \\ 0 & -1 \end{array}\right]$ has diagonalizing matrix $P = \left[\begin{array}{cc} 1 & -1 \\ 0 & 3 \end{array}\right]$ and $\left[\begin{array}{cc} -1 & 0 \\ 0 & 2 \end{array}\right]$ is already diagonal.

Exercise 3.3.13 If A is diagonalizable and 1 and -1 are the only eigenvalues, show that $A^{-1} = A$.

Exercise 3.3.14 If A is diagonalizable and 0 and 1 are the only eigenvalues, show that $A^2 = A$.

We have $\lambda^2 = \lambda$ for every eigenvalue λ (as $\lambda = 0, 1$) so $D^2 = D$, and so $A^2 = A$ as in Example 3.3.9.

Exercise 3.3.15 If A is diagonalizable and $\lambda \geq 0$ for each eigenvalue of A, show that $A = B^2$ for some matrix B.

Exercise 3.3.16 If $P^{-1}AP$ and $P^{-1}BP$ are both diagonal, show that $AB = BA$. [*Hint*: Diagonal matrices commute.]

Exercise 3.3.17 A square matrix A is called **nilpotent** if $A^n = 0$ for some $n \geq 1$. Find all nilpotent diagonalizable matrices. [*Hint*: Theorem 3.3.1.]

Exercise 3.3.18 Let A be any $n \times n$ matrix and $r \neq 0$ a real number.

a. Show that the eigenvalues of rA are precisely the numbers $r\lambda$, where λ is an eigenvalue of A.

b. Show that $c_{rA}(x) = r^n c_A\left(\frac{x}{r}\right)$.

---

b. $c_{rA}(x) = \det[xI - rA]$
$= r^n \det\left[\frac{x}{r}I - A\right] = r^n c_A\left[\frac{x}{r}\right]$

Exercise 3.3.19

a. If all rows of A have the same sum s, show that s is an eigenvalue.

b. If all columns of A have the same sum s, show that s is an eigenvalue.

Exercise 3.3.20 Let A be an invertible $n \times n$ matrix.

a. Show that the eigenvalues of A are nonzero.

b. Show that the eigenvalues of A^{-1} are precisely the numbers $1/\lambda$, where λ is an eigenvalue of A.

c. Show that $c_{A^{-1}}(x) = \frac{(-x)^n}{\det A} c_A\left(\frac{1}{x}\right)$.

---

b. If $\lambda \neq 0$, $A\mathbf{x} = \lambda\mathbf{x}$ if and only if $A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}$. The result follows.

Exercise 3.3.21 Suppose λ is an eigenvalue of a square matrix A with eigenvector $\mathbf{x} \neq \mathbf{0}$.

a. Show that λ^2 is an eigenvalue of A^2 (with the same $\mathbf{x}$).

b. Show that $\lambda^3 - 2\lambda + 3$ is an eigenvalue of $A^3 - 2A + 3I$.

c. Show that $p(\lambda)$ is an eigenvalue of $p(A)$ for any nonzero polynomial $p(x)$.

---

b. $(A^3-2A-3I)\mathbf{x} = A^3\mathbf{x} - 2A\mathbf{x} + 3\mathbf{x} = \lambda^3\mathbf{x} - 2\lambda\mathbf{x} + 3\mathbf{x} = (\lambda^3 - 2\lambda - 3)\mathbf{x}$.

Exercise 3.3.22 If A is an $n \times n$ matrix, show that $c_{A^2}(x^2) = (-1)^n c_A(x) c_A(-x)$.

Exercise 3.3.23 An $n \times n$ matrix A is called nilpotent if $A^m = 0$ for some $m \geq 1$.

a. Show that every triangular matrix with zeros on the main diagonal is nilpotent.

b. If A is nilpotent, show that $\lambda = 0$ is the only eigenvalue (even complex) of A.

c. Deduce that $c_A(x) = x^n$, if A is $n \times n$ and nilpotent.

---

b. If $A^m = 0$ and $A\mathbf{x} = \lambda\mathbf{x}$, $\mathbf{x} \neq \mathbf{0}$, then $A^2\mathbf{x} = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$. In general, $A^k\mathbf{x} = \lambda^k\mathbf{x}$ for all $k \geq 1$. Hence, $\lambda^m\mathbf{x} = A^m\mathbf{x} = 0\mathbf{x} = \mathbf{0}$, so $\lambda = 0$ (because $\mathbf{x} \neq \mathbf{0}$).

Exercise 3.3.24 Let A be diagonalizable with real eigenvalues and assume that $A^m = I$ for some $m \geq 1$.

a. Show that $A^2 = I$.

b. If m is odd, show that $A = I$. [*Hint*: Theorem **??**]

---

a. If $A\mathbf{x} = \lambda\mathbf{x}$, then $A^k\mathbf{x} = \lambda^k\mathbf{x}$ for each k. Hence $\lambda^m\mathbf{x} = A^m\mathbf{x} = \mathbf{x}$, so $\lambda^m = 1$. As λ is real, $\lambda = \pm 1$ by the Hint. So if $P^{-1}AP = D$ is diagonal, then $D^2 = I$ by Theorem 3.3.4. Hence $A^2 = PD^2P = I$.

Exercise 3.3.25 Let $A^2 = I$, and assume that $A \neq I$ and $A \neq -I$.

a. Show that the only eigenvalues of A are $\lambda = 1$ and $\lambda = -1$.

b. Show that A is diagonalizable. [*Hint*: Verify that $A(A+I) = A+I$ and $A(A-I) = -(A-I)$, and then look at nonzero columns of $A+I$ and of $A-I$.]

c. If $Q_m : \mathbb{R}^2 \to \mathbb{R}^2$ is reflection in the line $y = mx$ where $m \neq 0$, use (b) to show that the matrix of Q_m is diagonalizable for each m.

d. Now prove (c) geometrically using Theorem 3.3.3.

Exercise 3.3.26 Let $A = \left[\begin{array}{rrr} 2 & 3 & -3 \\ 1 & 0 & -1 \\ 1 & 1 & -2 \end{array}\right]$ and $B = \left[\begin{array}{rrr} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 2 & 0 & 0 \end{array}\right]$. Show that $c_A(x) = c_B(x) = (x+1)^2(x-2)$, but A is diagonalizable and B is not.

Exercise 3.3.27

a. Show that the only diagonalizable matrix A that has only one eigenvalue λ is the scalar matrix $A = \lambda I$.

b. Is $\left[\begin{array}{rr} 3 & -2 \\ 2 & -1 \end{array}\right]$ diagonalizable?

---

a. We have $P^{-1}AP = \lambda I$ by the diagonalization algorithm, so $A = P(\lambda I)P^{-1} = \lambda PP^{-1} = \lambda I$.

b. No. $\lambda = 1$ is the only eigenvalue.

Exercise 3.3.28 Characterize the diagonalizable $n \times n$ matrices A such that $A^2 - 3A + 2I = 0$ in terms of their eigenvalues. [*Hint*: Theorem 3.3.1.]

Exercise 3.3.29 Let $A = \left[\begin{array}{cc} B & 0 \\ 0 & C \end{array}\right]$ where B and C are square matrices.

a. If B and C are diagonalizable via Q and R (that is, $Q^{-1}BQ$ and $R^{-1}CR$ are diagonal), show that A is diagonalizable via $\left[\begin{array}{cc} Q & 0 \\ 0 & R \end{array}\right]$

b. Use (a) to diagonalize A if $B = \left[\begin{array}{cc} 5 & 3 \\ 3 & 5 \end{array}\right]$ and $C = \left[\begin{array}{rr} 7 & -1 \\ -1 & 7 \end{array}\right]$.

Exercise 3.3.30 Let $A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$ where B and C are square matrices.

a. Show that $c_A(x) = c_B(x)c_C(x)$.

b. If $\mathbf{x}$ and $\mathbf{y}$ are eigenvectors of B and C, respectively, show that $\begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}$ are eigenvectors of A, and show how every eigenvector of A arises from such eigenvectors.

Exercise 3.3.31 Referring to the model in Example 3.3.1, determine if the population stabilizes, becomes extinct, or becomes large in each case. Denote the adult and juvenile survival rates as A and J, and the reproduction rate as R.

	R	A	J
$a.$	2	$\frac{1}{2}$	$\frac{1}{2}$
$b.$	3	$\frac{1}{4}$	$\frac{1}{4}$
$c.$	2	$\frac{1}{4}$	$\frac{1}{3}$
$d.$	3	$\frac{3}{5}$	$\frac{1}{5}$

---

b. $\lambda_1 = 1$, stabilizes.

d. $\lambda_1 = \frac{1}{24}(3+\sqrt{69}) = 1.13$, diverges.

Exercise 3.3.32 In the model of Example 3.3.1, does the final outcome depend on the initial population of adult and juvenile females? Support your answer.

Exercise 3.3.33 In Example 3.3.1, keep the same reproduction rate of 2 and the same adult survival rate of $\frac{1}{2}$, but suppose that the juvenile survival rate is ρ. Determine which values of ρ cause the population to become extinct or to become large.

Exercise 3.3.34 In Example 3.3.1, let the juvenile survival rate be $\frac{2}{5}$ and let the reproduction rate be 2. What values of the adult survival rate α will ensure that the population stabilizes?

Extinct if $\alpha < \frac{1}{5}$, stable if $\alpha = \frac{1}{5}$, diverges if $\alpha > \frac{1}{5}$.

Supplementary Exercises for Chapter 3

Exercise 3.1 Show that

$$\det \begin{bmatrix} a+px & b+qx & c+rx \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{bmatrix} = (1+x^3)\det \begin{bmatrix} a & b & c \\ p & q & r \\ u & v & w \end{bmatrix}$$

Exercise 3.2

a. Show that $(A_{ij})^T = (A^T)_{ji}$ for all i, j, and all square matrices A.

b. Use (a) to prove that $\det A^T = \det A$. [*Hint*: Induction on n where A is $n \times n$.]

---

b. If A is 1×1, then $A^T = A$. In general, $\det[A_{ij}] = \det\left[(A_{ij})^T\right] = \det\left[(A^T)_{ji}\right]$ by (a) and induction. Write $A^T = \left[a'_{ij}\right]$ where $a'_{ij} = a_{ji}$, and expand $\det A^T$ along column 1.

$$\begin{aligned} \det A^T &= \sum_{j=1}^{n} a'_{j1}(-1)^{j+1}\det[(A^T)_{j1}] \\ &= \sum_{j=1}^{n} a_{1j}(-1)^{1+j}\det[A_{1j}] = \det A \end{aligned}$$

where the last equality is the expansion of $\det A$ along row 1.

Exercise 3.3 Show that $\det \begin{bmatrix} 0 & I_n \\ I_m & 0 \end{bmatrix} = (-1)^{nm}$ for all $n \geq 1$ and $m \geq 1$.

Exercise 3.4 Show that

$$\det \begin{bmatrix} 1 & a & a^3 \\ 1 & b & b^3 \\ 1 & c & c^3 \end{bmatrix} = (b-a)(c-a)(c-b)(a+b+c)$$

Exercise 3.5 Let $A = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$ be a 2×2 matrix with rows R_1 and R_2. If $\det A = 5$, find $\det B$ where

$$B = \begin{bmatrix} 3R_1 + 2R_3 \\ 2R_1 + 5R_2 \end{bmatrix}$$

Exercise 3.6 Let $A = \begin{bmatrix} 3 & -4 \\ 2 & -3 \end{bmatrix}$ and let $\mathbf{v}_k = A^k\mathbf{v}_0$ for each $k \geq 0$.

a. Show that A has no dominant eigenvalue.

b. Find $\mathbf{v}_k$ if $\mathbf{v}_0$ equals:

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

ii. $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

iii. $\begin{bmatrix} x \\ y \end{bmatrix} \neq \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ or $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$