Chapter 8
Matrices

Most students by now have been exposed to the language of matrices. They arise naturally in many subject areas but mainly in the context of solving a simultaneous system of linear equations. In this chapter, we will give a review of matrices, systems of linear equations, inverses, determinants, and eigenvectors and eigenvalues. The next chapter will apply what is learned here to linear systems of differential equations.

W.A. Adkins and M.G. Davidson, *Ordinary Differential Equations*, Undergraduate Texts in Mathematics, DOI 10.1007/978-1-4614-3618-8_8,

8.1 Matrix Operations

A ***matrix*** is a rectangular array of entities and is generally written in the following way:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}.$$

We let $\mathcal{R}$ denote the set of entities that will be in use at any particular time. Each x_{ij} is in $\mathcal{R}$, and in this text, $\mathcal{R}$ can be one of the following sets:

$\mathbb{R}$ or $\mathbb{C}$	The scalars
$\mathbb{R}[t]$ or $\mathbb{C}[t]$	Polynomials with real or complex entries
$\mathbb{R}(s)$ or $\mathbb{C}(s)$	The real or complex rational functions
$C^k(I,\mathbb{R})$ or $C^k(I,\mathbb{C})$	Real- or complex-valued functions with k continuous derivatives

Notice that addition and multiplication are defined on $\mathcal{R}$. Below we will extend these operations to matrices.

The following are examples of matrices.

Example 1.

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 9 \end{bmatrix}, \quad C = \begin{bmatrix} i & 2-i \\ 1 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} \sin t \\ \cos t \\ \tan t \end{bmatrix}, \quad E = \begin{bmatrix} \dfrac{s}{s^2-1} & \dfrac{1}{s^2-1} \\ \dfrac{-1}{s^2-1} & \dfrac{s+2}{s^2-1} \end{bmatrix}.$$

It is a common practice to use capital letters, like A, B, C, D, and E, to denote matrices. The ***size*** of a matrix is determined by the number of rows m and the number of columns n and written $m \times n$. In Example 1, A is a 2×3 matrix, B is a 1×3 matrix, C and E are 2×2 matrices, and D is a 3×1 matrix. A matrix is ***square*** if the number of rows is the same as the number of columns. Thus, C and E are square matrices. An entry in a matrix is determined by its position. If X is a matrix, the (i,j) ***entry*** is the entry that appears in the ith row and jth column. We denote it in two ways: $\mathrm{ent}_{ij}(X)$ or more simply X_{ij}. Thus, in Example 1, $A_{13} = 3$, $B_{12} = -1$, and $C_{22} = 0$. We say that two matrices X and Y are ***equal*** if the corresponding entries are equal, that is, $X_{ij} = Y_{ij}$, for all indices i and j. Necessarily, X and Y must be the same size. The ***main diagonal*** of a square $n \times n$ matrix X is the vector formed from the entries $X_{i\,i}$, for $i = 1, \ldots, n$. The main diagonal of C is $(i, 0)$ and the main diagonal of E is $(\frac{s}{s^2-1}, \frac{s+2}{s^2-1})$. A matrix is said to be a ***real matrix*** if each entry is real and a ***complex matrix*** if each entry is complex. Since every real number is also complex, every real matrix is also a complex matrix. Thus, A and B are real (and complex) matrices while C is a complex matrix.

Even though a matrix is a structured array of entities in $\mathcal{R}$, it should be viewed as a single object just as a word is a single object though made up of many letters. We let $M_{m,n}(\mathcal{R})$ denote the set of all $m \times n$ matrices with entries in $\mathcal{R}$. If the focus is on matrices of a certain size and not on the entries, we will sometimes write $M_{m,n}$. The following definitions highlight various kinds of matrices that commonly arise:

1. A ***diagonal*** matrix D is a square matrix in which all entries off the main diagonal are 0. We can say this in another way:

$$D_{i\,j} = 0 \text{ if } i \neq j.$$

   Examples of diagonal matrices are:

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \quad \begin{bmatrix} e^t & 0 & 0 \\ 0 & e^{4t} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} \frac{1}{s} & 0 & 0 & 0 \\ 0 & \frac{2}{s-1} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{-1}{s-2} \end{bmatrix}.$$

   It is convenient to write $\text{diag}(d_1, \ldots, d_n)$ to represent the diagonal $n \times n$ matrix with $(d_1, \ldots, d_n)$ on the diagonal. Thus, the diagonal matrices listed above are $\text{diag}(1, 4)$, $\text{diag}(e^t, e^{4t}, 1)$, and $\text{diag}(\frac{1}{s}, \frac{2}{s-1}, 0, -\frac{1}{s-2})$, respectively.
2. The ***zero*** matrix $\mathbf{0}$ is the matrix with each entry 0. The size is usually determined by the context. If we need to be specific, we will write $\mathbf{0}_{m,n}$ to mean the $m \times n$ zero matrix. Note that the square zero matrix, $\mathbf{0}_{n,n}$ is diagonal and is $\text{diag}(0, \ldots, 0)$.
3. The ***identity matrix***, I, is the square matrix with ones on the main diagonal and zeros elsewhere. The size is usually determined by the context, but if we want to be specific, we write I_n to denote the $n \times n$ identity matrix. The 2×2 and the 3×3 identity matrices are

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

4. We say a square matrix is ***upper triangular*** if each entry below the main diagonal is zero. We say a square matrix is ***lower triangular*** if each entry above the main diagonal is zero. The matrices

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 3 & 5 \\ 0 & 0 & 3 \\ 0 & 0 & -4 \end{bmatrix}$$

are upper triangular, and

$$\begin{bmatrix} 4 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & -7 \end{bmatrix}$$

are lower triangular.

5. Suppose A is an $m \times n$ matrix. The ***transpose*** of A, denoted A^t, is the $n \times m$ matrix obtained by turning the rows of A into columns. In terms of the entries we have, more explicitly,

$$(A^t)_{i\,j} = A_{j\,i}.$$

This expression reverses the indices of A. Simple examples are

$$\begin{bmatrix} 2 & 3 \\ 9 & 0 \\ 1 & 4 \end{bmatrix}^t = \begin{bmatrix} 2 & 9 & 1 \\ 3 & 0 & 4 \end{bmatrix} \quad \begin{bmatrix} e^t \\ e^{-t} \end{bmatrix}^t = \begin{bmatrix} e^t & e^{-t} \end{bmatrix} \quad \begin{bmatrix} \frac{1}{s} & \frac{2}{s^3} \\ \frac{2}{s^2} & \frac{3}{s} \end{bmatrix}^t = \begin{bmatrix} \frac{1}{s} & \frac{2}{s^2} \\ \frac{2}{s^3} & \frac{3}{s} \end{bmatrix}.$$

Matrix Algebra

There are three matrix operations that make up the algebraic structure of matrices: addition, scalar multiplication, and matrix multiplication.

Addition

Suppose A and B are two matrices of the same size. We define ***matrix addition***, $A + B$, entrywise by the following formula:

$$(A + B)_{i\,j} = A_{i\,j} + B_{i\,j}.$$

Thus, if

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

then

$$A + B = \begin{bmatrix} 1+4 & -2-1 & 0+0 \\ 4-3 & 5+8 & -3+1 \end{bmatrix} = \begin{bmatrix} 5 & -3 & 0 \\ 1 & 13 & -2 \end{bmatrix}.$$

Corresponding entries are added. Addition preserves the size of matrices. We can symbolize this in the following way:

$$+ : M_{m,n}(\mathcal{R}) \times M_{m,n}(\mathcal{R}) \to M_{m,n}(\mathcal{R}).$$

Addition satisfies the following properties:

Proposition 2. *Suppose A, B, and C are $m \times n$ matrices. Then*

$$A + B = B + A, \qquad \text{(commutative)}$$

$$(A + B) + C = A + (B + C), \qquad \text{(associative)}$$

$$A + \mathbf{0} = A, \qquad \text{(additive identity)}$$

$$A + (-A) = \mathbf{0}. \qquad \text{(additive inverse)}$$

Scalar Multiplication

Suppose A is an matrix and $c \in \mathcal{R}$. We define ***scalar multiplication***, $c \cdot A$, (but usually we will just write cA), entrywise by the following formula

$$(cA)_{ij} = cA_{ij}.$$

For example,

$$-2 \begin{bmatrix} 1 & 9 \\ -3 & 0 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} -2 & -18 \\ 6 & 0 \\ -4 & -10 \end{bmatrix}.$$

Scalar multiplication preserves the size of matrices. Thus,

$$\cdot : \mathcal{R} \times M_{m,n}(\mathcal{R}) \to M_{m,n}(\mathcal{R}).$$

Scalar multiplication satisfies the following properties:

Proposition 3. *Suppose A and B are matrices of the same size. Suppose $c_1, c_2 \in \mathcal{R}$. Then*

$$c_1(A + B) = c_1 A + c_1 B, \qquad \text{(distributive)}$$

$$(c_1 + c_2)A = c_1 A + c_2 A, \qquad \text{(distributive)}$$

$$c_1(c_2A) = (c_1c_2)A, \qquad \text{(associative)}$$

$$1A = A, \qquad \text{(1 is a multiplicative identity)}$$

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

Matrix Multiplication

Matrix multiplication is more complicated than addition and scalar multiplication. We will define it in two stages: first on row and column matrices and then on general matrices.

A ***row matrix*** or ***row vector*** is a matrix which has only one row. Thus, row vectors are in $M_{1,n}$. Similarly, a ***column matrix*** or ***column vector*** is a matrix which has only one column. Thus, column vectors are in $M_{m,1}$. We frequently will denote column and row vectors by lower case boldface letters like $\mathbf{v}$ or $\mathbf{x}$ instead of capital letters. It is unnecessary to use double subscripts to indicate the entries of a row or column matrix: if $\mathbf{v}$ is a row vector, then we write $\mathbf{v}_i$ for the ith entry instead of $\mathbf{v}_{1\,i}$. Similarly for column vectors. Suppose $\mathbf{v} \in M_{1,n}$ and $\mathbf{w} \in M_{n,1}$. We define the product $\mathbf{v} \cdot \mathbf{w}$ (or preferably $\mathbf{vw}$) to be the scalar given by

$$\mathbf{vw} = v_1w_1 + \cdots + v_nw_n.$$

Even though this formula looks like the scalar product or dot product that you likely have seen before, keep in mind that $\mathbf{v}$ is a row vector while $\mathbf{w}$ is a column vector. For example, if

$$\mathbf{v} = \begin{bmatrix} 1 & 3 & -2 & 0 \end{bmatrix} \text{ and } \mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 9 \end{bmatrix},$$

then

$$\mathbf{vw} = 1 \cdot 1 + 3 \cdot 3 + (-2) \cdot 0 + 0 \cdot 9 = 10.$$

Now suppose that A is any matrix. It is often convenient to distinguish the rows of A in the following way. Let $\mathrm{Row}_i(A)$ denotes the ith row of A. Then

$$A = \begin{bmatrix} \mathrm{Row}_1(A) \\ \mathrm{Row}_2(A) \\ \vdots \\ \mathrm{Row}_m(A) \end{bmatrix}.$$

In a similar way, if B is another matrix, we can distinguish the columns of B. Let $\mathrm{Col}_j(B)$ denote the jth column of B, then

$$B = \begin{bmatrix} \mathrm{Col}_1(B) & \mathrm{Col}_2(B) & \cdots & \mathrm{Col}_p(B) \end{bmatrix}.$$

Now let $A \in M_{mn}$ and $B \in M_{np}$. We define the ***matrix product*** of A and B to be the $m \times p$ matrix given entrywise by $\mathrm{ent}_{i\,j}(AB) = \mathrm{Row}_i(A)\,\mathrm{Col}_j(B)$. In other words, the (i, j)-entry of the product of A and B is the ith row of A times the jth column of B. We thus have

$$AB = \begin{bmatrix} \mathrm{Row}_1(A)\,\mathrm{Col}_1(B) & \mathrm{Row}_1(A)\,\mathrm{Col}_2(B) & \cdots & \mathrm{Row}_1(A)\,\mathrm{Col}_p(B) \\ \mathrm{Row}_2(A)\,\mathrm{Col}_1(B) & \mathrm{Row}_2(A)\,\mathrm{Col}_2(B) & \cdots & \mathrm{Row}_2(A)\,\mathrm{Col}_p(B) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Row}_m(A)\,\mathrm{Col}_1(B) & \mathrm{Row}_m(A)\,\mathrm{Col}_2(B) & \cdots & \mathrm{Row}_m(A)\,\mathrm{Col}_p(B) \end{bmatrix}.$$

Notice that each entry of AB is given as a product of a row vector and a column vector. Thus, it is necessary that the number of columns of A (the first matrix) match the number of rows of B (the second matrix). This common number is n. The resulting product is an $m \times p$ matrix. We thus write

$$\cdot : M_{m,n}(\mathcal{R}) \times M_{n,p}(\mathcal{R}) \to M_{m,p}(\mathcal{R}).$$

In terms of the entries of A and B, we have

$$\mathrm{ent}_{i\,j}(AB) = \mathrm{Row}_i(A)\,\mathrm{Col}_j(B) = \sum_{k=1}^{n} \mathrm{ent}_{i\,k}(A)\mathrm{ent}_{k\,j}(B) = \sum_{k=1}^{n} A_{i,k} B_{k,j}.$$

Example 4.

1. If $A = \begin{bmatrix} 2 & 1 \\ -1 & 3 \\ 4 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ 2 & -2 \end{bmatrix}$, then AB is defined because the number of columns of A is the number of rows of B. Further, AB is a 3×2 matrix and

$$AB = \begin{bmatrix} \begin{bmatrix} 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} & \begin{bmatrix} 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} \\ \begin{bmatrix} -1 & 3 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} & \begin{bmatrix} -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} \\ \begin{bmatrix} 4 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} & \begin{bmatrix} 4 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 4 & -7 \\ 4 & 8 \end{bmatrix}.$$

2. If $A = \begin{bmatrix} e^t & 2e^t \\ e^{2t} & 3e^{2t} \end{bmatrix}$ and $B = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, then

$$AB = \begin{bmatrix} -2e^t + 2e^t \\ -2e^{2t} + 3e^{2t} \end{bmatrix} = \begin{bmatrix} 0 \\ e^{2t} \end{bmatrix}.$$

Notice in the definition (and the example) that in a given column of AB, the corresponding column of B appears as the second factor. Thus,

$$\mathrm{Col}_j(AB) = A\,\mathrm{Col}_j(B). \tag{1}$$

Similarly, in each row of AB, the corresponding row of A appears and we get

$$\mathrm{Row}_i(A)B = \mathrm{Row}_i(AB). \tag{2}$$

Notice too that even though the product AB is defined, it is not necessarily true that BA is defined. This is the case in part 1 of the above example due to the fact that the number of columns of B (2) does not match the number of rows of A (3). Even when AB and BA are defined, it is not necessarily true that they are equal. Consider the following example:

Example 5. Suppose

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 1 \\ 4 & -1 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & -1 \end{bmatrix} = \begin{bmatrix} 10 & -1 \\ 12 & -3 \end{bmatrix}$$

while

$$BA = \begin{bmatrix} 2 & 1 \\ 4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 7 \\ 4 & 5 \end{bmatrix}.$$

These products are not the same. This example shows that matrix multiplication is *not* commutative. On the other hand, there can be two special matrices A and B for which $AB = BA$. In this case, we say A and B ***commute***. For example, if

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

then

$$AB = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = BA.$$

Thus A and B commute. However, such occurrences are special. The other properties that we are used to in an algebra are valid. We summarize them in the following proposition.

Proposition 6. *Suppose A, B, and C are matrices whose sizes are such that each line below is defined. Suppose $c_1, c_2 \in \mathcal{R}$. Then*

$$\begin{aligned}
&A(BC) = (AB)C, && \text{(associatvie)} \\
&A(cB) = (cA)B = c(AB), && \text{(commutes with scalar multiplication)} \\
&(A + B)C = AC + BC, && \text{(distributive)} \\
&A(B + C) = AB + AC, && \text{(distributive)} \\
&IA = AI = A. && \text{(I is a multiplicative identity)}
\end{aligned}$$

We highlight two useful formulas that follow from these algebraic properties. If A is an $m \times n$ matrix, then

$$A\mathbf{x} = x_1 \operatorname{Col}_1(A) + \cdots x_n \operatorname{Col}_n(A), \qquad \text{where } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{3}$$

and

$$\mathbf{y}A = y_1 \mathrm{Row}_1(A) + \cdots y_m \mathrm{Row}_m(A), \quad \text{where} \mathbf{y} = [y_1 \cdots y_m]. \tag{4}$$

Henceforth, we will use these algebraic properties without explicit reference. The following result expresses the relationship between multiplication and transposition of matrices

Theorem 7. *Let A and B be matrices such that AB is defined. Then $B^t A^t$ is defined and*

$$B^t A^t = (AB)^t.$$

Proof. The number of columns of B^t is the same as the number of rows of B while the number of rows of A^t is the number of columns of A. These numbers agree since AB is defined so $B^t A^t$ is defined. If n denotes these common numbers, then

$$(B^t A^t)_{i\,j} = \sum_{k=1}^{n} (B^t)_{i\,k} (A^t)_{k\,j} = \sum_{k=1}^{n} A_{j\,k} B_{k\,i} = (AB)_{j\,i} = ((AB)^t)_{i\,j}.$$

All entries of $A^t B^t$ and $(AB)^t$ agree so they are equal. $\square$

Exercises

1–3. Let $A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ -1 & 2 \end{bmatrix}$, and $C = \begin{bmatrix} 0 & 2 \\ -3 & 4 \\ 1 & 1 \end{bmatrix}$. Compute the following matrices.

1. $B + C$, $B - C$, $2B - 3C$
2. AB, AC, BA, CA
3. $A(B + C)$, $AB + AC$, $(B + C)A$
4. Let $A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ -1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 0 \end{bmatrix}$. Find C so that $3A + C = 4B$.

5–9. Let $A = \begin{bmatrix} 3 & -1 \\ 0 & -2 \\ 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 1 & 1 & -3 \\ 0 & -1 & 4 & -1 \end{bmatrix}$, and $C = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 3 & 1 \\ 0 & 1 & 8 \\ 1 & 1 & 7 \end{bmatrix}$. Find the following products.

5. AB
6. BC
7. CA
8. $B^t A^t$
9. ABC
10. Let $A = \begin{bmatrix} 1 & 4 & 3 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 \\ 0 \\ -1 \\ -2 \end{bmatrix}$. Find AB and BA.

11–13. Let $A = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \\ -1 & -2 & -5 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 4 & -1 \end{bmatrix}$, $C = \begin{bmatrix} 3 & -2 \\ 3 & -2 \end{bmatrix}$. Verify the following facts:

11. $A^2 = 0$
12. $B^2 = I_2$
13. $C^2 = C$

14–15. Compute $AB - BA$ in each of the following cases.

14. $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

15. $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 & -2 \\ 3 & -2 & 4 \\ -3 & 5 & -1 \end{bmatrix}$

16. Let $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}$. Show that there are no numbers a and b so that $AB - BA = I$, where I is the 2×2 identity matrix.

17. Suppose that A and B are 2×2 matrices.
    1. Show by example that it need not be true that $(A + B)^2 = A^2 + 2AB + B^2$.
    2. Find conditions on A and B to insure that the equation in Part (a) is valid.

18. If $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, compute A^2 and A^3.

19. If $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, compute B^n for all n.

20. If $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, compute A^2, A^3, and more generally, A^n for all n.

21. Let $A = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be a matrix with two rows v_1 and v_2. (The number of columns of A is not relevant for this problem.) Describe the effect of multiplying A on the left by the following matrices:

$$(a) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \; (b) \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} \; (c) \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \; (d) \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \; (e) \begin{bmatrix} 1 & 0 \\ 0 & a \end{bmatrix}$$

22. Let $E(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$. Show that $E(\theta_1 + \theta_2) = E(\theta_1)E(\theta_2)$.

23. Let $F(\theta) = \begin{bmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{bmatrix}$. Show that $F(\theta_1 + \theta_2) = F(\theta_1)F(\theta_2)$.

24. Let $D = \operatorname{diag}(d_1, \ldots, d_n)$ and $E = \operatorname{diag}(e_1, \ldots, e_n)$. Show that

$$DE = \operatorname{diag}(d_1e_1, \ldots, d_ne_n).$$

8.2 Systems of Linear Equations

Most students have learned various techniques for finding the solution of a system of linear equations. They usually include various forms of elimination and substitutions. In this section, we will learn the Gauss-Jordan elimination method. It is essentially a highly organized method involving elimination and substitution that always leads to the solution set. This general method has become the standard for solving systems. At first reading, it may seem to be a bit complicated because of its description for general systems. However, with a little practice on a few examples, it is quite easy to master. We will as usual begin with our definitions and proceed with examples to illustrate the needed concepts. To make matters a bit cleaner, we will stick to the case where $\mathcal{R} = \mathbb{R}$. Everything we do here will work for $\mathcal{R} = \mathbb{C}$, $\mathbb{R}(s)$, or $\mathbb{C}(s)$ as well. (A technical difficulty for general $\mathcal{R}$ is the lack of inverses.)

If $x_1, \ldots, x_n$ are variables, then the equation

$$a_1 x_1 + \cdots + a_n x_n = b$$

is called a ***linear equation*** in the unknowns $x_1, \ldots, x_n$. A ***system of linear equations*** is a set of m linear equations in the unknowns $x_1, \ldots, x_n$ and is written in the form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array} \tag{1}$$

The entries a_{ij} are in $\mathbb{R}$ and are called ***coefficients***. Likewise, each b_j is in $\mathbb{R}$. A key observation is that (1) can be rewritten in matrix form as:

$$A\mathbf{x} = \mathbf{b}, \tag{2}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

We call A the ***coefficient matrix***, $\mathbf{x}$ the ***variable matrix***, and $\mathbf{b}$ the ***output matrix***. Any column vector $\mathbf{x}$ with entries in $\mathbb{R}$ that satisfies equation (1) (or, equivalently, equation (2)) is called a ***solution***. If a system has a solution, we say it is ***consistent***; otherwise, it is ***inconsistent***. The ***solution set*** is the set of all solutions. The system is said to be ***homogeneous*** if $\mathbf{b} = \mathbf{0}$, otherwise it is called ***nonhomogeneous***. The homogeneous case is especially important. We call the solution set to

$$A\mathbf{x} = \mathbf{0}$$

the ***null space*** of A and denote it by $\mathrm{NS}(A)$. Another important matrix associated with (2) is the ***augmented matrix***:

$$[A \mid b\,] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right],$$

where the vertical line only serves to separate A from $\mathbf{b}$.

Example 1. Write the coefficient, variable, output, and augmented matrices for the following system:

$$\begin{aligned} -2x_1 + 3x_2 - x_3 &= 4 \\ x_1 - 2x_2 + 4x_3 &= 5. \end{aligned}$$

Determine whether the following vectors are solutions:

$$\mathbf{v_1} = \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{v_2} = \begin{bmatrix} 7 \\ 7 \\ 3 \end{bmatrix}, \quad \mathbf{v_3} = \begin{bmatrix} 10 \\ 7 \\ 1 \end{bmatrix}, \quad \mathbf{v_4} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

► **Solution.** The coefficient matrix is $A = \begin{bmatrix} -2 & 3 & -1 \\ 1 & -2 & 4 \end{bmatrix}$, the variable matrix is $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, the output matrix is $\mathbf{b} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, and the augmented matrix is $\left[\begin{array}{ccc|c} -2 & 3 & -1 & 4 \\ 1 & -2 & 4 & 5 \end{array}\right]$. The system is nonhomogeneous. Notice that

$$A\mathbf{v_1} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \quad \text{and } A\mathbf{v_2} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \text{ while } A\mathbf{v_3} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Therefore, $\mathbf{v_1}$ and $\mathbf{v_2}$ are solutions, $\mathbf{v_3}$ is not a solution but since $A\mathbf{v_3} = \mathbf{0}$, we have $\mathbf{v_3}$ is in the null space of A. Finally, $\mathbf{v_4}$ is not the right size and thus cannot be a solution. ◄

Remark 2. When only 2 or 3 variables are involved in an example, we will frequently use the variables x, y, and z instead of the subscripted variables x_1, x_2, and x_3.

Linearity

It is convenient to think of $\mathbb{R}^n$ as the set of column vectors $M_{n,1}(\mathbb{R})$. If A is an $m \times n$ real matrix, then for each column vector $\mathbf{x} \in \mathbb{R}^n$, the product, $A\mathbf{x}$, is a column vector in $\mathbb{R}^m$. Thus, the matrix A induces a map which we also denote just by $A : \mathbb{R}^n \to \mathbb{R}^m$ given by matrix multiplication. It satisfies the following important property.

Proposition 3. *The map $A : \mathbb{R}^n \to \mathbb{R}^m$ is linear. In other words,*

1. $A(\mathbf{x} + \mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y})$
2. $A(c\mathbf{x}) = cA(\mathbf{x})$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Proof. This follows directly from Propositions 3 and 6 of Sect. 8.1. □

Linearity is an extremely important property for it allows us to describe the structure of the solution set to $A\mathbf{x} = \mathbf{b}$ in a particularly nice way.

Proposition 4. *With A, $\mathbf{x}$, and $\mathbf{b}$ as above, we have two possibilities. Either the solution set to $A\mathbf{x} = \mathbf{b}$ is the empty set or we can write all solutions in the following form:*

$$\mathbf{x}_\mathrm{p} + \mathbf{x}_\mathrm{h},$$

where $\mathbf{x}_\mathrm{p}$ is a fixed particular solution and $\mathbf{x}_\mathrm{h}$ is any vector in $\mathrm{NS}(A)$.

Remark 5. We will write the solution set to $A\mathbf{x} = \mathbf{b}$, when it is nonempty, as

$$x_\mathrm{p} + \mathrm{NS}(A).$$

Proof. Suppose $\mathbf{x}_\mathrm{p}$ is a fixed particular solution and $\mathbf{x}_\mathrm{h} \in \mathrm{NS}(A)$. Then $A(\mathbf{x}_\mathrm{p}+\mathbf{x}_\mathrm{h}) = A\mathbf{x}_\mathrm{p} + A\mathbf{x}_\mathrm{h} = \mathbf{b} + \mathbf{0} = \mathbf{b}$. This implies that $\mathbf{x}_\mathrm{p} + \mathbf{x}_\mathrm{h}$ is a solution. On the other hand, suppose $\mathbf{x}$ is a solution to $A\mathbf{x} = \mathbf{b}$. Let $\mathbf{x}_\mathrm{h} = \mathbf{x} - \mathbf{x}_\mathrm{p}$. Then $A\mathbf{x}_\mathrm{h} = A(\mathbf{x} - \mathbf{x}_\mathrm{p}) = A\mathbf{x} - A\mathbf{x}_\mathrm{p} = \mathbf{b} - \mathbf{b} = 0$. This means that $\mathbf{x}_\mathrm{h}$ is in the null space of A and we get $\mathbf{x} = \mathbf{x}_\mathrm{p} + \mathbf{x}_\mathrm{h}$. □

Remark 6. The solution set being empty is a legitimate possibility. For example, the simple equation $0x = 1$ has empty solution set. The system of equations $A\mathbf{x}=\mathbf{0}$ is called the ***associated homogeneous system***. It should be mentioned that the particular solution $\mathbf{x}_\mathrm{p}$ is not necessarily unique. In Chap. 5, we saw a similar theorem for a second order differential equation $Ly = f$. That theorem provided a strategy for solving such differential equations: First we solved the homogeneous equation $Ly = 0$ and second found a particular solution (using variation of parameters or undetermined coefficients). For a linear system of equations, the matter is much simpler; the Gauss-Jordan method will give the whole solution set at one time. We will see that it has the above form.

Homogeneous Systems

The homogeneous case, $A\mathbf{x} = \mathbf{0}$, is of particular interest. Observe that $\mathbf{x} = \mathbf{0}$ is always a solution so $\mathrm{NS}(A)$ is never the empty set. But much more is true.

Proposition 7. *The solution set,* $\mathrm{NS}(A)$*, to a homogeneous system* $A\mathbf{x} = \mathbf{0}$ *is a linear space. In other words, if* $\mathbf{x}$ *and* $\mathbf{y}$ *are solutions to the homogeneous system and* c *is a scalar, then* $\mathbf{x} + \mathbf{y}$ *and* $c\mathbf{x}$ *are also solutions.*

Proof. Suppose $\mathbf{x}$ and $\mathbf{y}$ are in $\mathrm{NS}(A)$. Then $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x}+A\mathbf{y} = \mathbf{0}+\mathbf{0} = \mathbf{0}$. This shows that $\mathbf{x} + \mathbf{y} \in \mathrm{NS}(A)$. Now suppose $c \in \mathbb{R}$. Then $A(c\mathbf{x}) = cA\mathbf{x} = c\mathbf{0} = \mathbf{0}$. Hence $c\mathbf{x} \in \mathrm{NS}(A)$. Thus $\mathrm{NS}(A)$ is a linear space. ☐

Corollary 8. *The solution set to a general system of linear equations,* $A\mathbf{x} = \mathbf{b}$*, is either*

1. *Empty*
2. *Unique*
3. *Infinite*

Proof. The associated homogeneous system $A\mathbf{x} = \mathbf{0}$ has solution set, $\mathrm{NS}(A)$, that is either equal to the trivial set $\{\mathbf{0}\}$ or an infinite set. To see this suppose that $\mathbf{x}$ is a nonzero solution to $A\mathbf{x} = \mathbf{0}$, then by Proposition 7, all multiples, $c\mathbf{x}$, are in $\mathrm{NS}(A)$ as well. Therefore, by Proposition 4, if there is a solution to $A\mathbf{x} = \mathbf{b}$, it is unique or there are infinitely many. ☐

The Elementary Equation and Row Operations

We say that two systems of equations are ***equivalent*** if their solution sets are the same. This definition implies that the variable matrix is the same for each system.

Example 9. Consider the following systems of equations:

$$\begin{aligned} 2x + 3y &= 5 \\ x - y &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x &= 1 \\ y &= 1. \end{aligned}$$

The solution set to the second system is transparent. For the first system, there are some simple operations that easily lead to the solution: First, switch the two equations around. Next, multiply the equation $x - y = 1$ by -2 and add the result to the second equation. We then obtain

$$\begin{aligned} x - \ y &= 0 \\ 5y &= 5 \end{aligned}$$

Next, multiply the second equation by $\frac{1}{5}$ to get $y = 1$. Then add this equation to the first. We get $x = 1$ and $y = 1$. Thus, they both have the same solution set, namely, the single vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. They are thus equivalent. When used in the right way, these kinds of operations can transform a complicated system into a simpler one. We formalize these operations in the following definition:

Suppose $A\mathbf{x} = \mathbf{b}$ is a given system of linear equations. The following three operations are called ***elementary equation operations***:

1. Switch the order in which two equations are listed.
2. Multiply an equation by a nonzero scalar.
3. Add a multiple of one equation to another.

Notice that each operation produces a new system of linear equations but leaves the size of the system unchanged. Furthermore, we have the following proposition.

Proposition 10. *An elementary equation operation applied to a system of linear equations is an equivalent system of equations.*

Proof. Suppose S is a system of equations and S' is a system obtained from S by switching two equations. A vector $\mathbf{x}$ is a solution to S if and only if it satisfies each equation in the system. If we switch the order of two of the equations, then $\mathbf{x}$ still satisfies each equation, and hence is a solution to S'. Notice that applying the same elementary equation operation to S' produces S. Hence, a solution to S' is a solution to S. It follows that S and S' are equivalent. The proof for the second and third elementary equation operations is similar. □

The main idea in solving a system of linear equations is to perform a finite sequence of elementary equation operations to transform a system into simpler system where the solution set is transparent. Proposition 10 implies that the solution set of the simpler system is the same as original system. Let us consider our example above.

Example 11. Use elementary equation operations to transform

$$\begin{aligned} 2x + 3y &= 5, \\ x - \ y &= 0 \end{aligned}$$

into

$$\begin{aligned} x &= 1, \\ y &= 1. \end{aligned}$$

► **Solution.**

$$\begin{aligned} 2x + 3y &= 5 \\ x - y &= 0 \end{aligned}$$ Switch the order of the two equations.

$$\begin{aligned} x - y &= 0 \\ 2x + 3y &= 5 \end{aligned}$$ Add −2 times the first equation to the second equation.

$$\begin{aligned} x - y &= 0 \\ 5y &= 5 \end{aligned}$$ Multiply the second equation by $\frac{1}{5}$.

$$\begin{aligned} x - y &= 0 \\ y &= 1 \end{aligned}$$ Add the second equation to the first.

$$\begin{aligned} x \quad &= 1 \\ y &= 1 \end{aligned}$$ ◄

Each operation produces a new system equivalent to the first by Proposition 10. The end result is a system where the solution is transparent. Since $y = 1$ is apparent in the fourth system, we could have stopped and used the method of ***back substitution***, that is, substitute $y = 1$ into the first equation and solve for x. However, it is in accord with the Gauss–Jordan elimination method to continue as we did to eliminate the variable y in the first equation of the fourth system.

You will notice that the variables x and y play no prominent role in any of the calculations. They merely serve as placeholders for the coefficients, some of which change with each operation. We thus simplify the notation by performing the elementary operations on just the augmented matrix. The elementary equation operations become the ***elementary row operations*** which act on the augmented matrix of the system.

The elementary row operations on a matrix are:

1. Switch two rows
2. Multiply a row by a nonzero constant
3. Add a multiple of one row to another

The following notations for these operations will be useful:

1. p_{ij} - switch rows i and j
2. $m_i(a)$ - multiply row i by $a \neq 0$
3. $t_{ij}(a)$ - add to row j the value of a times row i

The effect of p_{ij} on a matrix A is denoted by $p_{ij}(A)$. Similarly for the other elementary row operations.

The corresponding operations when applied to the augmented matrix for the system in Example 11 becomes:

$$\left[\begin{array}{rr|r} 2 & 3 & 5 \\ 1 & -1 & 0 \end{array}\right] \xrightarrow{p_{12}} \left[\begin{array}{rr|r} 1 & -1 & 0 \\ 2 & 3 & 5 \end{array}\right] \xrightarrow{t_{12}(-2)} \left[\begin{array}{rr|r} 1 & -1 & 0 \\ 0 & 5 & 5 \end{array}\right]$$

$$\xrightarrow{m_2(1/5)} \left[\begin{array}{rr|r} 1 & -1 & 0 \\ 0 & 1 & 1 \end{array}\right] \xrightarrow{t_{21}(1)} \left[\begin{array}{rr|r} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right].$$

Above each arrow is the notation for the elementary row operation performed to produce the next augmented matrix. The sequence of elementary row operations chosen follows a certain strategy: Starting from left to right and top down, one tries to isolate a 1 in a given column and produce 0's above and below it. This corresponds to isolating and eliminating variables.

Let us consider three illustrative examples. The sequence of elementary row operation we perform is in accord with the Gauss–Jordan method which we will discuss in detail later on in this section. For now, verify each step. The end result will be an equivalent system for which the solution set will be transparent.

Example 12. Consider the following system of linear equations:

$$\begin{array}{rcrcrcl} 2x & + & 3y & + & 4z & = & 9 \\ x & + & 2y & - & z & = & 2 \end{array}.$$

Find the solution set and write it in the form $\mathbf{x}_\mathrm{p} + \mathrm{NS}(A)$.

► **Solution.** We first will write the augmented matrix and perform a sequence of elementary row operations:

$$\left[\begin{array}{rrr|r} 2 & 3 & 4 & 9 \\ 1 & 2 & -1 & 2 \end{array}\right] \xrightarrow{p_{12}} \left[\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 2 & 3 & 4 & 9 \end{array}\right] \xrightarrow{t_{12}(-2)} \left[\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 0 & -1 & 6 & 5 \end{array}\right]$$

$$\xrightarrow{m_2(-1)} \left[\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 0 & 1 & -6 & -5 \end{array}\right] \xrightarrow{t_{21}(-2)} \left[\begin{array}{rrr|r} 1 & 0 & 11 & 12 \\ 0 & 1 & -6 & -5 \end{array}\right].$$

The last augmented matrix corresponds to the system

$$\begin{array}{rcrcrcl} x & + & & & 11z & = & 12 \\ & & y & - & 6z & = & -5. \end{array}$$

In the first equation, we can solve for x in terms of z, and in the second equation, we can solve for y in terms of z. We refer to z as a ***free variable*** and let $z = \alpha$ be a parameter in $\mathbb{R}$. Then we obtain

$$\begin{aligned} x &= 12 - 11\alpha \\ y &= -5 + 6\alpha \\ z &= \quad\ \alpha. \end{aligned}$$

In vector form, we write

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 12 - 11\alpha \\ -5 + 6\alpha \\ \alpha \end{bmatrix} = \begin{bmatrix} 12 \\ -5 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -11 \\ 6 \\ 1 \end{bmatrix}.$$

The vector, $\mathbf{x}_{\mathrm{p}} = \begin{bmatrix} 12 \\ -5 \\ 0 \end{bmatrix}$ is a particular solution (corresponding to $\alpha = 0$) while all multiples of the vector $\begin{bmatrix} -11 \\ 6 \\ 1 \end{bmatrix}$ gives the null space of A. We have thus written the solution in the form $\mathbf{x}_{\mathrm{p}} + \mathrm{NS}(A)$. In this case, there are infinitely many solutions. ◀

Example 13. Find the solution set for the system

$$\begin{aligned} 3x + 2y + z &= 4 \\ 2x + 2y + z &= 3 \\ x + y + z &= 0. \end{aligned}$$

► **Solution.** Again we start with the augmented matrix and apply elementary row operations. Occasionally, we will apply more than one operation at a time. When this is so, we stack the operations above the arrow with the topmost operation performed first followed in order by the ones below it:

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 4 \\ 2 & 2 & 1 & 3 \\ 1 & 1 & 1 & 0 \end{array}\right] \xrightarrow{p_{13}} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 2 & 1 & 3 \\ 3 & 2 & 1 & 4 \end{array}\right] \xrightarrow{\substack{t_{12}(-2) \\ t_{13}(-3)}} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 3 \\ 0 & -1 & -2 & 4 \end{array}\right]$$

$$\xrightarrow{p_{23}} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & 4 \\ 0 & 0 & -1 & 3 \end{array}\right] \xrightarrow{\substack{m_2(-1) \\ m_3(-1)}} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & -4 \\ 0 & 0 & 1 & -3 \end{array}\right]$$

$$\xrightarrow{\substack{t_{32}(-2) \\ t_{31}(-1)}} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{array}\right] \xrightarrow{t_{21}(-1)} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{array}\right].$$

The last augmented matrix corresponds to the system

$$\begin{aligned} x &= 1 \\ y &= 2 \\ z &= -3. \end{aligned}$$

The solution set is transparent: $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$. In this example, we note that $\mathrm{NS}(A) = \{\mathbf{0}\}$, and the system has a unique solution. ◄

Example 14. Solve the following system of linear equations:

$$\begin{aligned} x + 2y + 4z &= -2 \\ x + y + 3z &= 1 \\ 2x + y + 5z &= 2. \end{aligned}$$

► **Solution.** Again we begin with the augmented matrix and perform elementary row operations.

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & -2 \\ 1 & 1 & 3 & 1 \\ 2 & 1 & 5 & 2 \end{array}\right] \xrightarrow[t_{13}(-2)]{t_{12}(-1)} \left[\begin{array}{ccc|c} 1 & 2 & 4 & -2 \\ 0 & -1 & -1 & 3 \\ 0 & -3 & -3 & 6 \end{array}\right]$$

$$\xrightarrow{m_2(-1)} \left[\begin{array}{ccc|c} 1 & 2 & 4 & -2 \\ 0 & 1 & 1 & -3 \\ 0 & -3 & -3 & 6 \end{array}\right] \xrightarrow{t_{23}(3)} \left[\begin{array}{ccc|c} 1 & 2 & 4 & -2 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & -3 \end{array}\right]$$

$$\xrightarrow{m_3(-1/3)} \left[\begin{array}{ccc|c} 1 & 2 & 4 & -2 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[t_{21}(-2)]{t_{31}(2) \atop t_{32}(3)} \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

The system that corresponds to the last augmented matrix is

$$\begin{aligned} x + 2z &= 0 \\ y + z &= 0 \\ 0 &= 1. \end{aligned}$$

The last equation, which is shorthand for $0x + 0y + 0z = 1$, clearly has no solution. Thus, the system has no solution. ◄

Reduced Matrices

These last three examples typify what happens in general and illustrate the three possible outcomes discussed in Corollary 8: infinitely many solutions, a unique solution, or no solution at all. The most involved case is when the solution set has infinitely many solutions. In Example 12, a single parameter α was needed to parameterize the set of solutions. However, in general, there may be many parameters needed. We will always want to use the least number of parameters possible, without dependencies among them. In each of the three preceding examples, it was transparent what the solution was by considering the system determined by the last listed augmented matrix. The last matrix was in a certain sense reduced as simple as possible.

We say that a matrix is in ***row echelon form (REF)*** if the following three conditions are satisfied:

1. The nonzero rows lie above the zero rows.
2. The first nonzero entry in a nonzero row is 1. (We call such a 1 a ***leading one***.)
3. For any two adjacent nonzero rows, the leading one of the upper row is to the left of the leading one of the lower row. (We say the leading ones are in echelon form.)

We say a matrix is in ***row reduced echelon form (RREF)*** if it also satisfies

4. The entries above each leading one are zero.

Example 15. Determine which of the following matrices are row echelon form, row reduced echelon form, or neither. For the matrices in row echelon form, determine the columns (**C**) of the leading ones. If a matrix is not in row reduced echelon form, explain which conditions are violated.

$$(1)\ \begin{bmatrix} 1 & 0 & -3 & 11 & 2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad (2)\ \begin{bmatrix} 0 & 1 & 0 & 1 & 4 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (3)\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(4)\ \begin{bmatrix} 1 & 0 & 0 & 4 & 3 & 0 \\ 0 & 2 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (5)\ \begin{bmatrix} 1 & 1 & 2 & 4 & -7 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)\ \begin{bmatrix} 0 & 1 & 0 & 2 \\ 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

► **Solution.**

1. (REF): Leading ones are in the first, third, and fourth columns. It is not reduced because there is a nonzero entry above the leading one in the third column.
2. (RREF): The leading ones are in the second and third columns.
3. Neither: The zero row is not at the bottom.
4. Neither: The first nonzero entry in the second row is not 1.
5. (REF): Leading ones are in the first and fifth columns. It is not reduced because there is a nonzero entry above the leading one in the fifth column.
6. Neither: The leading ones are not in echelon form. ◄

Suppose a matrix A transforms by elementary row operations into a matrix A' which is in row reduced echelon form. We will sometimes say that A ***row reduces*** to A'. The ***rank*** of A, denoted Rank$A$, is the number of nonzero rows in A'. The definition of row reduced echelon form is valid for arbitrary matrices and not just matrices that come from a system of linear equations, that is, the augmented matrix. Suppose though that we consider a system $A\mathbf{x} = \mathbf{b}$, where A is an $m \times n$ matrix. Suppose the augmented matrix $[A|\mathbf{b}]$ is row reduced to a matrix $[A'|\mathbf{b}']$. Let $r =$ Rank$A$ and $r_{\mathbf{b}} =$ Rank$[A|\mathbf{b}]$. The elementary row operations that row reduce $[A|\mathbf{b}]$ to $[A'|\mathbf{b}']$ are the same elementary row operations the row reduce A to A'. Further, A' is row reduced echelon form. Hence, r is the number of nonzero rows in A' and $r \leq r_{\mathbf{b}}$. Consider the following possibilities:

1. If $r < r_{\mathbf{b}}$, then there is a row of the form $[0 \cdots 0|1]$ in $[A'|\mathbf{b}']$ in which case $r_{\mathbf{b}} = r + 1$. Such a row translates into the equation

$$0x_1 + \cdots + 0x_n = 1,$$

which means the system is inconsistent. This is what occurs in Example 14. Recall that the augmented matrix row reduces as follows:

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & -2 \\ 1 & 1 & 3 & 1 \\ 2 & 1 & 5 & 2 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right],$$

where $\to \cdots \to$ denotes the sequence of elementary row operations used. Notice that Rank$(A) = 2$ while the presence of $[000|1]$ in the last row of $[A'|\mathbf{b}']$ gives Rank$[A|\mathbf{b}] = 3$. There are no solutions.
2. Suppose $r = r_{\mathbf{b}}$. The variables that correspond to the columns where the leading ones occur are called the ***leading variables*** or ***dependent variables***. Since each nonzero row has a leading one, there are r leading variables. All of the other variables are called ***free variables***, and we are able to solve each leading variable in terms of the free variables and so there are solutions. The system $A\mathbf{x} = \mathbf{b}$ is consistent. Consider two subcases:

a. Suppose $r < n$. Since there are a total of n (the number of columns of A) variables, there are $n - r$ free variables and hence infinitely many solutions. This is what occurs in Example 12. Recall

$$\left[\begin{array}{ccc|c} 2 & 3 & 4 & 9 \\ 1 & 2 & -1 & 2 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 11 & 12 \\ 0 & 1 & -6 & -5 \end{array}\right].$$

Here, we have $r = \mathrm{Rank}A = \mathrm{Rank}[A|\mathbf{b}]$=2 and $n = 3$. There is exactly $n - r = 3 - 2 = 1$ free variable.

b. Now suppose $r = n$. Then every variable is a leading variable. There are no free variables so there is a unique solution. This is what occurs in Example 13. Recall

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 4 \\ 2 & 2 & 1 & 3 \\ 1 & 1 & 1 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{array}\right].$$

Here we have $r = \mathrm{Rank}A = \mathrm{Rank}[A|\mathbf{b}] = 3$ and $n = 3$. There are no free variables. The solution is unique.

We summarize our discussion in the following proposition.

Proposition 16. *Let A be an $m \times n$ matrix and $\mathbf{b}$ an $n \times 1$ column vector. Let $r = \mathrm{Rank}A$ and $r_{\mathbf{b}} = \mathrm{Rank}[A|\mathbf{b}]$:*

1. *If $r < r_{\mathbf{b}}$, then $A\mathbf{x} = \mathbf{b}$ is inconsistent.*
2. *If $r = r_{\mathbf{b}}$, then $A\mathbf{x} = \mathbf{b}$ is consistent. Further,*

   a. *if $r < n$, there are $n - r > 0$ free variables and hence infinitely many solutions.*
   b. *if $r = n$, there is exactly one solution.*

Example 17. Suppose the following matrices are obtained by row reducing the augmented matrix of a system of linear equations. Identify the leading and free variables and write down the solution set. Assume the variables are $x_1, x_2, \ldots$.

$$(1) \left[\begin{array}{cccc|c} 1 & 1 & 4 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \quad (2) \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$(3) \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 5 \end{array}\right] \quad (4) \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

► **Solution.** 1. The zero row provides no information and can be ignored. The variables are x_1, x_2, x_3, and x_4. The leading ones occur in the first and fourth column. Therefore, x_1 and x_4 are the leading variables. The free variables are x_2 and x_3. Let $\alpha = x_2$ and $\beta = x_3$. The first row implies the equation $x_1 + x_2 + 4x_3 = 2$. We solve for x_1 and obtain $x_1 = 2 - x_2 - 4x_3 = 2 - \alpha - 4\beta$. The second row implies the equation $x_4 = 3$. Thus,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 - \alpha - 4\beta \\ \alpha \\ \beta \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 3 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -4 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

where α and β are arbitrary parameters in $\mathbb{R}$.

2. x_1 is the leading variable. $\alpha = x_2$ and $\beta = x_3$ are free variables. The first row implies $x_1 = 1 - \alpha$. The solution is

$$\mathbf{x} = \begin{bmatrix} 1 - \alpha \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

where α and β are in $\mathcal{R}$.

3. The leading variables are x_1, x_2, and x_3. There are no free variables. The solution set is

$$\mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}.$$

4. The row $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$ implies the solution set is empty. ◄

The Gauss–Jordan Elimination Method

Now that you have seen several examples, we present the Gauss-Jordan elimination method for any matrix. It is an algorithm to transform any matrix to row reduced echelon form using a finite number of elementary row operations. When applied to an augmented matrix of a system of linear equations, the solution set can be readily discerned. It has other uses as well so our description will be for an arbitrary matrix.

Algorithm 18. The Gauss–Jordan Elimination Method Let A be a matrix. There is a finite sequence of elementary row operations that transform A to a matrix in row reduced echelon form. There are two stages of the process: (1) The first stage is called ***Gaussian elimination*** and transforms a given matrix to row echelon form, and (2) the second stage is called ***Gauss–Jordan elimination*** and transforms a matrix in row echelon form to row reduced echelon form.

From A to REF: Gaussian Elimination

1. Let $A_1 = A$. If $A_1 = \mathbf{0}$, then A is in row echelon form.
2. If $A_1 \neq 0$, then in the first nonzero column from the left (say the jth column), locate a nonzero entry in one of the rows (say the ith row with entry a):

   a. Multiply that row by the reciprocal of that nonzero entry: $m_i(1/a)$.
   b. Permute that row with the top row: $p_{1\,i}$. There is now a 1 in the $(1, j)$ entry.
   c. If b is a nonzero entry in the (i, j) position for $i \neq 1$, add $-b$ times the first row to the ith row: $t_{1\,j}(-b)$. Do this for each row below the first.

   The transformed matrix will have the following form:

$$\begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & & & \\ \vdots & \ddots & \vdots & & & A_2 & \\ 0 & \cdots & 0 & 0 & & & \end{bmatrix}.$$

   The $*$'s in the first row are unknown entries, and A_2 is a matrix with fewer rows and columns than A_1.

3. If $A_2 = \mathbf{0}$, we are done. The above matrix is in row echelon form.
4. If $A_2 \neq \mathbf{0}$, apply step (2) to A_2. Since there are zeros to the left of A_2 and the only elementary row operations we apply affect the rows of A_2 (and not all of A), there will continue to be zeros to the left of A_2. The result will be a matrix of the form

$$\begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & * & * & \cdots & * \\ & & & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\ \vdots & \ddots & \vdots & 0 & 0 & \cdots & 0 & 0 & & & \\ & & & \vdots & \vdots & & \vdots & \vdots & & A_3 & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & & & \end{bmatrix}.$$

5. If $A_3 = 0$, we are done. Otherwise, continue repeating step (2) until a matrix $A_k = \mathbf{0}$ is obtained.

From REF to RREF: Gauss–Jordan Elimination

1. The leading ones now become apparent in the previous process. We begin with the rightmost leading one. Suppose it is in the kth row and lth column. If there

is a nonzero entry (b say) above that leading one, we add $-b$ times the kth row to it: $t_{k\,j}(-b)$. We do this for each nonzero entry in the lth column. The result is zeros above the rightmost leading one. (The entries to the left of a leading one are zeros. This process preserves that property.)

2. Now repeat the process described above to each leading one moving right to left. The result will be a matrix in row reduced echelon form.

Example 19. Use the Gauss–Jordan method to row reduce the following matrix to row reduced echelon form:

$$\begin{bmatrix} 2 & 3 & 8 & 0 & 4 \\ 3 & 4 & 11 & 1 & 8 \\ 1 & 2 & 5 & 1 & 6 \\ -1 & 0 & -1 & 0 & 1 \end{bmatrix}.$$

► **Solution.** The Gauss–Jordan algorithm produces

$$\begin{bmatrix} 2 & 3 & 8 & 0 & 4 \\ 3 & 4 & 11 & 1 & 8 \\ 1 & 2 & 5 & 1 & 6 \\ -1 & 0 & -1 & 0 & 1 \end{bmatrix} \xrightarrow{p_{13}} \begin{bmatrix} 1 & 2 & 5 & 1 & 6 \\ 3 & 4 & 11 & 1 & 8 \\ 2 & 3 & 8 & 0 & 4 \\ -1 & 0 & -1 & 0 & 1 \end{bmatrix} \xrightarrow[t_{14}(1)]{t_{12}(-3)\; t_{13}(-2)}$$

$$\begin{bmatrix} 1 & 2 & 5 & 1 & 6 \\ 0 & -2 & -4 & -2 & -10 \\ 0 & -1 & -2 & -2 & -8 \\ 0 & 2 & 4 & 1 & 7 \end{bmatrix} \xrightarrow{m_2(-1/2)} \begin{bmatrix} 1 & 2 & 5 & 1 & 6 \\ 0 & 1 & 2 & 1 & 5 \\ 0 & -1 & -2 & -2 & -8 \\ 0 & 2 & 4 & 1 & 7 \end{bmatrix} \xrightarrow{t_{23}(1)\; t_{24}(-2)}$$

$$\begin{bmatrix} 1 & 2 & 5 & 1 & 6 \\ 0 & 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & -1 & -3 \\ 0 & 0 & 0 & -1 & -3 \end{bmatrix} \xrightarrow{m_3(-1)\; t_{34}(1)} \begin{bmatrix} 1 & 2 & 5 & 1 & 6 \\ 0 & 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{t_{32}(-1)\; t_{31}(-1)}$$

$$\begin{bmatrix} 1 & 2 & 5 & 0 & 3 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{t_{21}(-2)} \begin{bmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In the first step, we observe that the first column is nonzero so it is possible to produce a 1 in the upper left-hand corner. This is most easily accomplished by $p_{1,3}$. The next set of operations produces 0's below this leading one. We repeat this procedure on the submatrix to the right of the zeros. We produce a one in the 2, 2 position by $m_2(-\frac{1}{2})$, and the next set of operations produce zeros below this second leading one. Now notice that the third column below the second leading one is zero. There are no elementary row operations that can produce a leading one in the (3, 3) position that involve just the third and fourth row. We move over to the fourth column and observe that the entries below the second leading one are not both zero. The elementary row operation $m_3(-1)$ produces a leading one in the (3, 4) position and the subsequent operation produces a zero below it. At this point, A has been transformed to row echelon form. Now starting at the rightmost leading one, the 1 in the (3, 4) position, we use operations of the form $t_{3\,i}(a)$ to produce zeros above that leading one. This is applied to each column that contains a leading one. The result is in row reduced echelon form. ◄

The student is encouraged to go carefully through Examples 12–14. In each of those examples, the Gauss–Jordan Elimination method was used to transform the augmented matrix to the matrix in row reduced echelon form.

A Basis of the Null Space

When the Gauss-Jordan elimination method is used to compute the null space of A, the solution space takes the form

$$\{\alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k : \alpha_1, \ldots, \alpha_k \in \mathbb{R}\}. \tag{3}$$

We simplify the notation and write $\operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ given by (3). The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ turn out to also be linearly independent. The notion of linear independence for linear spaces of functions was introduced earlier in the text and that same notion extends to vectors in $\mathbb{R}^n$.

Before we give the definition of linear independence, consider the following example.

Example 20. Find the null space of

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

► **Solution.** We augment A with the zero vector and row reduce:

$$\left[\begin{array}{rrrrr|r} 2 & 3 & 1 & 4 & -2 & 0 \\ 1 & 1 & -1 & 2 & 3 & 0 \\ 3 & 5 & 3 & 6 & -7 & 0 \\ 4 & 5 & -1 & 8 & 4 & 0 \end{array}\right] \xrightarrow{p_{12}} \left[\begin{array}{rrrrr|r} 1 & 1 & -1 & 2 & 3 & 0 \\ 2 & 3 & 1 & 4 & -2 & 0 \\ 3 & 5 & 3 & 6 & -7 & 0 \\ 4 & 5 & -1 & 8 & 4 & 0 \end{array}\right]$$

$$\xrightarrow{\substack{t_{12}(-2) \\ t_{13}(-3) \\ t_{14}(-4)}} \left[\begin{array}{rrrrr|r} 1 & 1 & -1 & 2 & 3 & 0 \\ 0 & 1 & 3 & 0 & -8 & 0 \\ 0 & 2 & 6 & 0 & -16 & 0 \\ 0 & 1 & 3 & 0 & -8 & 0 \end{array}\right] \xrightarrow{\substack{t_{23}(-2) \\ t_{24}(-1)}} \left[\begin{array}{rrrrr|r} 1 & 1 & -1 & 2 & 3 & 0 \\ 0 & 1 & 3 & 0 & -8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{t_{21}(-1)} \left[\begin{array}{rrrrr|r} 1 & 0 & -4 & 2 & 11 & 0 \\ 0 & 1 & 3 & 0 & -8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

If the variables are $x_1, \ldots, x_5$, then x_1 and x_2 are the leading variables and x_3, x_4, and x_5 are the free variables. Let $\alpha = x_3$, $\beta = x_4$, and $\gamma = x_5$. Then

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 4\alpha - 2\beta - 11\gamma \\ -3\alpha + 8\gamma \\ \alpha \\ \beta \\ \gamma \end{bmatrix} = \alpha \begin{bmatrix} 4 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} -11 \\ 8 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

It follows that

$$\mathrm{NS}(A) = \mathrm{Span}\left\{ \begin{bmatrix} 4 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -11 \\ 8 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$ ◄

We say that vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$ are ***linearly independent*** if the equation

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$$

implies that the coefficients $a_1, \ldots, a_k$ are all zero. Otherwise, they are said to be ***linearly dependent***.

Example 21. Show that the vectors

$$\mathbf{v_1} = \begin{bmatrix} 4 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \ \text{and} \ \mathbf{v}_3 = \begin{bmatrix} -11 \\ 8 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

that span the null space in Example 20 are linearly independent.

► **Solution.** The equation $a_1\mathbf{v_1} + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 = \mathbf{0}$ means

$$\begin{bmatrix} 4a_1 - 2a_3 - 11a_3 \\ -3a_1 + 8a_3 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

From the last three rows, it is immediate that $a_1 = a_2 = a_3 = 0$. This implies the linear independence. ◄

Suppose V is a linear subspace of $\mathbb{R}^n$. By this we mean that V is in $\mathbb{R}^n$ and is closed under addition and scalar multiplication. That is, if

1. $\mathbf{v}_1, \mathbf{v}_2$ are in V, then so is $\mathbf{v}_1 + \mathbf{v}_2$
2. $c \in \mathbb{R}$ and $\mathbf{v} \in V$, then $c\mathbf{v} \in V$.

We say that the set of vectors $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ forms a ***basis*** of V if $\mathcal{B}$ is linearly independent and spans V. Thus, in Example 20, the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathrm{NS}(A)$. We observe that the number of vectors is three: one for each free variable.

Theorem 22. *Suppose A is an $m \times n$ matrix of rank r then there are $n - r$ vectors that form a basis of the null space of A.*

Proof. Suppose $[A|\mathbf{0}]$ is row reduced to $[R|\mathbf{0}]$, in row reduced echelon form. By Theorem 16, there are $n - r$ free variables. Let $k = n - r$. Suppose $f_1, \ldots, f_k$ are the columns of R that do not contain leading ones. Then $x_{f_1}, \ldots, x_{f_k}$ are the free variables. Let $\alpha_j = x_{f_j}$, $j = 1, \ldots, k$ be parameters. Solving for $\mathbf{x}$ in $R\mathbf{x} = \mathbf{0}$ in terms of the free variables, we get vectors $v_1, \ldots, v_k$ such that

$$\mathbf{x} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k.$$

It follows that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ span $\mathrm{NS}(A)$. Since R is in row reduced echelon form, the f_j^{th} entry of $\mathbf{v}_j$ is one while the f_j^{th} entry of $\mathbf{v}_i$ is zero, when $i \neq j$. If $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$, then the f_j^{th} entry of the left-hand side is a_j, and hence, $a_j = 0$, $j = 1, \ldots k$. It follows that $v_1, \ldots, v_k$ are linearly independent and hence form a basis of $\mathrm{NS}(A)$. □

Exercises

1–2. For each system of linear equations, identify the coefficient matrix A, the variable matrix $\mathbf{x}$, the output matrix $\mathbf{b}$, and the augmented matrix $[A|\mathbf{b}]$.

1.

$$\begin{aligned} x + 4y + 3z &= 2 \\ x + y - z &= 4 \\ 2x + \quad\quad z &= 1 \\ y - z &= 6 \end{aligned}$$

2.

$$\begin{aligned} 2x_1 - 3x_2 + 4x_3 + x_4 &= 0 \\ 3x_1 + 8x_2 - 3x_3 - 6x_4 &= 1 \end{aligned}$$

3. Suppose $A = \begin{bmatrix} 1 & 0 & -1 & 4 & 3 \\ 5 & 3 & -3 & -1 & -3 \\ 3 & -2 & 8 & 4 & -3 \\ -8 & 2 & 0 & 2 & 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ 3 \\ -4 \end{bmatrix}$. Write out the system of linear equations that corresponds to $A\mathbf{x} = \mathbf{b}$.

4–9. In the following, matrices identify those that are in row reduced echelon form. If a matrix is not in row reduced echelon form, find a single elementary row operation that will transform it to row reduced echelon form and write the new matrix.

4. $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & -4 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix}$

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

7. $A = \begin{bmatrix} 0 & 1 & 0 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

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

9. $A = \begin{bmatrix} 1 & 0 & 1 & 0 & 3 \\ 0 & 1 & 3 & 4 & 1 \\ 3 & 0 & 3 & 0 & 9 \end{bmatrix}$

10 –18. Use the Gauss-Jordan elimination method to row reduce each matrix.

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

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

12. $\begin{bmatrix} 0 & -2 & 3 & 2 & 1 \\ 0 & 2 & -1 & 4 & 0 \\ 0 & 6 & -7 & 0 & -2 \\ 0 & 4 & -6 & -4 & -2 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 2 & 1 & 1 & 5 \\ 2 & 4 & 0 & 0 & 6 \\ 1 & 2 & 0 & 1 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{bmatrix}$

14. $\begin{bmatrix} -1 & 0 & 1 & 1 & 0 & 0 \\ -3 & 1 & 3 & 0 & 1 & 0 \\ 7 & -1 & -4 & 0 & 0 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ -1 & 2 & 0 \\ 1 & 6 & 8 \\ 0 & 4 & 4 \end{bmatrix}$

16. $\begin{bmatrix} 5 & 1 & 8 & 1 \\ 1 & 1 & 4 & 0 \\ 2 & 0 & 2 & 1 \\ 4 & 1 & 7 & 1 \end{bmatrix}$

17. $\begin{bmatrix} 2 & 8 & 0 & 0 & 6 \\ 1 & 4 & 1 & 1 & 7 \\ -1 & -4 & 0 & 1 & 0 \end{bmatrix}$

18. $\begin{bmatrix} 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 \end{bmatrix}$

19–25. Solve the following systems of linear equations:

19.
$$\begin{aligned} x + 3y \phantom{{}+2z} &= 2 \\ 5x \phantom{{}-y} + 3z &= -5 \\ 3x - y + 2z &= -4 \end{aligned}$$

20.
$$\begin{aligned} 3x_1 + 2x_2 + 9x_3 + 8x_4 &= 10 \\ x_1 + \phantom{x_2 +{}} x_3 + 2x_4 &= 4 \\ -2x_1 + x_2 + x_3 - 3x_4 &= -9 \\ x_1 + x_2 + 4x_3 + 3x_4 &= 3 \end{aligned}$$

21.
$$\begin{aligned} -x + 4y &= -3x \\ x - y &= -3y \end{aligned}$$

22.
$$\begin{aligned} -2x_1 - 8x_2 - x_3 - x_4 &= -9 \\ -x_1 - 4x_2 \phantom{{}- x_3} - x_4 &= -8 \\ x_1 + 4x_2 + x_3 + x_4 &= 6 \end{aligned}$$

23.
$$\begin{aligned} 2x + 3y + 8z &= 5 \\ 2x + y + 10z &= 3 \\ 2x \phantom{{}+y} + 8z &= 4 \end{aligned}$$

24.
$$\begin{aligned} x_1 + x_2 + x_3 + 5x_4 &= 3 \\ x_2 + x_3 + 4x_4 &= 1 \\ x_1 \qquad + x_3 + 2x_4 &= 2 \\ 2x_1 + 2x_2 + 3x_3 + 11x_4 &= 8 \\ 2x_1 + x_2 + 2x_3 + 7x_4 &= 7 \end{aligned}$$

25.
$$\begin{aligned} x_1 + x_2 \qquad &= 3 + x_1 \\ x_2 + 2x_3 &= 4 \qquad + x_2 + x_3 \\ x_1 + 3x_2 + 4x_3 &= 11 + x_1 + 2x_2 + 2x_3 \end{aligned}$$

26–32. For each of the following matrices A, find a basis of the null space.

26. $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

27. $A = \begin{bmatrix} 1 & 3 \\ 3 & 5 \end{bmatrix}$

28. $A = \begin{bmatrix} 2 & 6 \\ 1 & 3 \\ -1 & -3 \end{bmatrix}$

29. $A = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 4 & 1 \end{bmatrix}$

30. $A = \begin{bmatrix} 1 & 1 & 3 & -2 \\ 1 & 4 & 1 & 1 \\ 4 & 7 & 10 & -5 \end{bmatrix}$

31. $A = \begin{bmatrix} -1 & 2 & 1 & 1 \\ 6 & -2 & 3 & 1 \\ 2 & -1 & 0 & 4 \\ 5 & 1 & 7 & -9 \end{bmatrix}$

32. $A = \begin{bmatrix} 2 & 3 & 1 & 9 & 3 \\ 0 & 1 & -3 & 4 & 0 \\ 2 & 1 & 7 & 1 & 3 \\ 4 & 4 & 8 & 10 & 6 \end{bmatrix}$

33. Suppose the homogeneous system $A\mathbf{x} = \mathbf{0}$ has the following two solutions: $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$. Is $\begin{bmatrix} 5 \\ -1 \\ 4 \end{bmatrix}$ a solution? Why or why not?

34. For what value of k will the following system have a solution:

$$\begin{aligned} x_1 + x_2 - x_3 &= 2 \\ 2x_1 + 3x_2 + x_3 &= 4 \\ x_1 - 2x_2 - 10x_3 &= k \end{aligned}$$

35–36. Let $A = \begin{bmatrix} 1 & 3 & 4 \\ -2 & 1 & 7 \\ 1 & 1 & 0 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{b}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

35. Solve $A\mathbf{x} = \mathbf{b}_i$, for each $i = 1, 2, 3$.
36. Solve the above systems simultaneously by row reducing

$$[A|b_1|b_2|b_3] = \left[\begin{array}{ccc|ccc} 1 & 3 & 4 & 1 & 1 & 1 \\ -2 & 1 & 7 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

8.3 Invertible Matrices

Let A be a square matrix. A matrix B is said to be an ***inverse*** of A if $BA = AB = I$. In this case, we say A is ***invertible*** or ***nonsingular***. If A is not invertible, we say A is ***singular***.

Example 1. Suppose

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

Show that A is invertible and an inverse is

$$B = \begin{bmatrix} -1 & -1 \\ 4 & 3 \end{bmatrix}.$$

► **Solution.** Observe that

$$AB = \begin{bmatrix} 3 & 1 \\ -4 & -1 \end{bmatrix}\begin{bmatrix} -1 & -1 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$BA = \begin{bmatrix} -1 & -1 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ -4 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$ ◄

The following proposition says that when A has an inverse, there can only be one.

Proposition 2. *Let A be an invertible matrix. Then the inverse is unique.*

Proof. Suppose B and C are inverses of A. Then

$$B = BI = B(AC) = (BA)C = IC = C.$$ □

Because of uniqueness, we can properly say *the inverse* of A when A is invertible. In Example 1, the matrix $B = \begin{pmatrix} -1 & -1 \\ 4 & 3 \end{pmatrix}$ is the inverse of A; there are no others. It is standard convention to denote the inverse of A by A^{-1}.

We say that B is a ***left inverse*** of A if $BA = I$ and a ***right inverse*** if $AB = I$. For square matrices, we have the following proposition which we will not prove. This proposition tells us that it is enough to check that either $AB = I$ or $BA = I$. It is not necessary to check both products.

Proposition 3. *Suppose A is a square matrix and B is a left or right inverse of A. Then A is invertible and $A^{-1} = B$.*

For many matrices, it is possible to determine their inverse by inspection. For example, the identity matrix I_n is invertible and its inverse is I_n: $I_n I_n = I_n$. A diagonal matrix $\mathrm{diag}(a_1, \ldots, a_n)$ is invertible if each $a_i \neq 0$, $i = 1, \ldots, n$. The inverse then is simply $\mathrm{diag}\left(\frac{1}{a_1}, \ldots, \frac{1}{a_n}\right)$. However, if one of the a_i is zero, then the matrix in not invertible. Even more is true. If A has a zero row, say the ith row, then A is not invertible. To see this, we get from (2) of Sect. 8.1 that $\mathrm{Row}_i(AB) = \mathrm{Row}_i(A)B = \mathbf{0}$. Hence, there is no matrix B for which $AB = I$. Similarly, a matrix with a zero column cannot be invertible.

Proposition 4. *Let A and B be invertible matrices. Then*

1. *A^{-1} is invertible and $(A^{-1})^{-1} = A$.*
2. *AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*

Proof. Suppose A and B are invertible. The symmetry of the equation $A^{-1}A = AA^{-1} = I$ says that A^{-1} is invertible and $(A^{-1})^{-1} = A$. We also have

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I.$$

This shows $(AB)^{-1} = B^{-1}A^{-1}$. □

The following corollary easily follows by induction:

Corollary 5. *If $A = A_1 \cdots A_k$ is the product of invertible matrices, then A is invertible and $A^{-1} = A_k^{-1} \cdots A_1^{-1}$.*

The Elementary Matrices

When an elementary row operation is applied to the identity matrix I, the resulting matrix is called an ***elementary matrix***.

Example 6. Show that each of the following matrices are elementary matrices:

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

► **Solution.** We have

$$m_2(3)I = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} = E_1,$$

$$t_{23}(2)I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = E_2,$$

$$p_{12}I = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = E_3,$$

where the size of the identity matrix matches the size of the given matrix. ◄

The following example shows a useful relationship between an elementary row operation and left multiplication by the corresponding elementary matrix.

Example 7. Use the elementary matrices in Example 6 to show that multiplying a matrix A on the left by E_i produces the same effect as applying the corresponding elementary row operation to A.

► **Solution.** 1. Let $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$. Then $E_1 A = \begin{bmatrix} A_1 \\ 3A_2 \end{bmatrix} = m_2(3)A$.

2. Let $A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix}$. Then $E_2 A = \begin{bmatrix} A_1 \\ A_2 \\ 2A_2 + A_3 \end{bmatrix} = t_{23}(2)A$.

3. Let $A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix}$. Then $E_3 A = \begin{bmatrix} A_2 \\ A_1 \\ A_3 \\ A_4 \end{bmatrix} = p_{12}A$. ◄

This important relationship is summarized in general as follows:

Proposition 8. *Let e be an elementary row operation. That is, let e be one of $p_{i\,j}$, $t_{i\,j}(a)$, or $m_i(a)$. Let $E = eI$ be the elementary matrix obtained by applying e to the identity matrix. Then*

$$e(A) = EA.$$

In other words, when A is multiplied on the left by the elementary matrix E, it is the same as applying the elementary row operation e to A.

If $E = eI$ is an elementary matrix, then we can apply a second elementary row operation to E to get I back. For example, consider the elementary matrices in Example 6. It is easy to see that

$$m_2(1/3)E_1 = I, \quad t_{23}(-2)E_2 = I, \quad \text{and} \quad p_{12}E_3 = I.$$

In general, each elementary row operation is reversible by the another elementary row operation. Switching two rows is reversed by switching those two rows again, thus $p_{ij}^{-1} = p_{ij}$. Multiplying a row by a nonzero constant is reversed by multiplying that same row by the reciprocal of that nonzero constant, thus $(m_i(a))^{-1} = m_i(1/a)$. Finally, adding a multiple of a row to another is reversed by adding the negative multiple of first row to the second, thus $(t_{ij}(a))^{-1} = t_{ij}(-a)$. If e is an elementary row operation, we let e^{-1} denote the inverse row operation. Thus,

$$ee^{-1}A = A,$$

for any matrix A. These statements imply

Proposition 9. *An elementary matrix is invertible and the inverse is an elementary matrix.*

Proof. Let e be an elementary row operation and e^{-1} its inverse elementary row operation. Let $E = e(I)$ and $B = e^{-1}I$. By Proposition 8, we have

$$EB = E(e^{-1}I) = ee^{-1}I = I.$$

It follows the E is invertible and $E^{-1} = B = e^{-1}(I)$ is an elementary matrix. □

We now have a useful result that will be used later.

Theorem 10. *Let A by an $n \times n$ matrix. Then the following are equivalent:*

1. *A is invertible.*
2. *The null space of A, $\mathrm{NS}(A)$, consists only of the zero vector.*
3. *A row reduces (by Gauss–Jordan) to the identity matrix.*

Proof. Suppose A is invertible and $\mathbf{c} \in \mathrm{NS}(A)$. Then $A\mathbf{c} = \mathbf{0}$. Multiply both sides by A^{-1} to get

$$\mathbf{c} = A^{-1}A\mathbf{c} = A^{-1}\mathbf{0} = \mathbf{0}.$$

Thus, the null space of A consists only of the zero vector.

Suppose now that the null space of A is trivial. Then the system $A\mathbf{x} = \mathbf{0}$ is equivalent to

$$\begin{aligned} x_1 \qquad\qquad &= 0 \\ x_2 \qquad &= 0 \\ \ddots \quad & \\ x_n &= 0 \end{aligned}$$

Thus, the augmented matrix $[A \mid \mathbf{0}]$ reduces to $[I \mid \mathbf{0}]$. The same elementary row operations row reduces A to the identity.

Now suppose A row reduces to the identity. Then there is a sequence of elementary matrices, $E_1, \ldots, E_k$, (corresponding to the elementary row operations) such that $E_1 \cdots E_k A = I$. Let $B = E_1 \cdots E_k$. Then $BA = I$ and this implies A is invertible. □

Corollary 11. *Suppose A is an $n \times n$ matrix. If the null space of A is not the zero vector, then A row reduces to a matrix that has a zero row.*

Proof. Suppose A transforms by elementary row operations to R which is in row reduced echelon form. The system $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions, and this only occurs if there are one or more free variables in the system $R\mathbf{x} = \mathbf{0}$. The number of leading variables which is the same as the number of nonzero rows is thus less than n. Hence, there are some zero rows. □

Inversion Computations

Let $\mathbf{e}_i$ be the column vector with 1 in the ith position and 0's elsewhere. By Equation 8.1.(1), the equation $AB = I$ implies that $A\,\mathrm{Col}_i(B) = \mathrm{Col}_i(I) = \mathbf{e}_i$. This means that the solution to $A\mathbf{x} = \mathbf{e}_i$ is the ith column of the inverse of A, when A is invertible. We can thus compute the inverse of A one column at a time using the Gauss–Jordan elimination method on the augmented matrix $[A|\mathbf{e}_i]$. Better yet, though, is to perform the Gauss–Jordan elimination method on the matrix $[A|I]$, that is, the matrix A augmented with I. If A is invertible, it will reduce to a matrix of the form $[I|B]$ and B will be A^{-1}. If A is not invertible, it will not be possible to produce the identity in the first slot.

We illustrate this in the following two examples.

Example 12. Determine whether the matrix

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

is invertible. If it is, compute the inverse.

► **Solution.** We will augment A with I and follow the procedure outlined above:

$$\left[\begin{array}{rrr|rrr} 2 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 3 & -1 & 4 & 0 & 0 & 1 \end{array}\right] \xrightarrow[p_{13}]{t_{13}(-1)} \left[\begin{array}{rrr|rrr} 1 & -1 & 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 2 & 0 & 3 & 1 & 0 & 0 \end{array}\right] \xrightarrow[t_{23}(-2)]{t_{13}(-2)}$$

$$\left[\begin{array}{rrr|rrr} 1 & -1 & 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 3 & -2 & -2 \end{array}\right] \xrightarrow[t_{31}(-1)]{m_3(-1)\; t_{32}(-1)} \left[\begin{array}{rrr|rrr} 1 & -1 & 0 & 2 & -2 & -1 \\ 0 & 1 & 0 & 3 & -1 & -2 \\ 0 & 0 & 1 & -3 & 2 & 2 \end{array}\right] \xrightarrow{t_{21}(1)}$$

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 5 & -3 & -3 \\ 0 & 1 & 0 & 3 & -1 & -2 \\ 0 & 0 & 1 & -3 & 2 & 2 \end{array}\right].$$

It follows that A is invertible and $A^{-1} = \begin{bmatrix} 5 & -3 & -3 \\ 3 & -1 & -2 \\ -3 & 2 & 2 \end{bmatrix}$. ◄

Example 13. Let $A = \begin{bmatrix} 1 & -4 & 0 \\ 2 & 1 & 3 \\ 0 & -7 & 3 \end{bmatrix}$. Determine whether A is invertible. If it is, find its inverse.

► **Solution.** Again, we augment A with I and row reduce:

$$\left[\begin{array}{rrr|rrr} 1 & -4 & 0 & 1 & 0 & 0 \\ 2 & 1 & 3 & 0 & 1 & 0 \\ 0 & 9 & 3 & 0 & 0 & 1 \end{array}\right] \xrightarrow[t_{23}(-1)]{t_{12}(-2)} \left[\begin{array}{rrr|rrr} 1 & -4 & 0 & 1 & 0 & 0 \\ 0 & 9 & 3 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 & 1 \end{array}\right].$$

We can stop at this point. Notice that the row operations produced a $\mathbf{0}$ row in the reduction of A. This implies A cannot be invertible. ◄

Solving a System of Equations

Suppose A is a square matrix with a known inverse. Then the equation $A\mathbf{x} = \mathbf{b}$ implies $\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$ and thus gives the solution.

Example 14. Solve the following system:

$$\begin{aligned} 2x + \quad & + 3z = 1 \\ y + & \; z = 2 \\ 3x - y + & 4z = 3. \end{aligned}$$

► **Solution.** The coefficient matrix is

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

whose inverse we computed in the example above:

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

The solution to the system is thus

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} 5 & -3 & -3 \\ 3 & -1 & -2 \\ -3 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -10 \\ -5 \\ 7 \end{bmatrix}.$$ ◄

Exercises

1–12. Determine whether the following matrices are invertible. If so, find the inverse:

1. $\begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}$

2. $\begin{bmatrix} 3 & 2 \\ 4 & 3 \end{bmatrix}$

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

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

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

6. $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

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

8. $\begin{bmatrix} 1 & 0 & -2 \\ 2 & -2 & 0 \\ 1 & 2 & -1 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 3 & 0 & 1 \\ 2 & 2 & -2 & 0 \\ 1 & -1 & 0 & 4 \\ 1 & 2 & 3 & 9 \end{bmatrix}$

10. $\begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$

11. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

12. $\begin{bmatrix} -3 & 2 & -8 & 2 \\ 0 & 2 & -3 & 5 \\ 1 & 2 & 3 & 5 \\ 1 & -1 & 1 & -1 \end{bmatrix}$

13–18. Solve each system $A\mathbf{x} = \mathbf{b}$, where A and $\mathbf{b}$ are given below, by first computing A^{-1} and applying it to $A\mathbf{x} = \mathbf{b}$ to get $\mathbf{x} = A^{-1}\mathbf{b}$.

13. $A = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$

15. $A = \begin{bmatrix} 1 & 0 & -2 \\ 2 & -2 & 0 \\ 1 & 2 & -1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$

16. $A = \begin{bmatrix} 1 & -1 & 1 \\ -2 & 5 & -2 \\ 0 & 2 & -1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

17. $A = \begin{bmatrix} 1 & 3 & 0 & 1 \\ 2 & 2 & -2 & 0 \\ 1 & -1 & 0 & 4 \\ 1 & 2 & 3 & 9 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}$

18. $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ -2 \\ 1 \end{bmatrix}$

19. Suppose A is an invertible matrix. Show that A^t is invertible and give a formula for the inverse.

20. Let $E(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$. Show $E(\theta)$ is invertible and find its inverse.

21. Let $F(\theta) = \begin{bmatrix} \sinh\theta & \cosh\theta \\ \cosh\theta & \sinh\theta \end{bmatrix}$. Show $F(\theta)$ is invertible and find its inverse.

22. Suppose A is invertible and $AB = AC$. Show that $B = C$. Give an example of a nonzero matrix A (not invertible) with $AB = AC$, for some B and C, but $B \neq C$.

8.4 Determinants

In this section, we will discuss the definition of the determinant and some of its properties. For our purposes, the determinant is a very useful number that we can associate to a square matrix. The determinant has an wide range of applications. It can be used to determine whether a matrix is invertible. Cramer's rule gives the unique solution to a system of linear equations as the quotient of determinants. In multidimensional calculus, the Jacobian is given by a determinant and expresses how area or volume changes under a transformation. Most students by now are familiar with the definition of the determinant for a 2×2 matrix: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The ***determinant*** of A is given by

$$\det(A) = ad - bc.$$

It is the product of the diagonal entries minus the product of the off diagonal entries. For example, $\det \begin{bmatrix} 1 & 3 \\ 5 & -2 \end{bmatrix} = 1 \cdot (-2) - 5 \cdot 3 = -17$.

The definition of the determinant for an $n \times n$ matrix is decidedly more complicated. We will present an inductive definition. Let A be an $n \times n$ matrix and let $A(i, j)$ be the matrix obtained from A by deleting the ith row and jth column. Since $A(i, j)$ is an $(n-1) \times (n-1)$ matrix, we can inductively define the (i, j) ***minor***, $\mathrm{Minor}_{i\,j}(A)$, to be the determinant of $A(i, j)$:

$$\mathrm{Minor}_{i\,j}(A) = \det(A(i, j)).$$

The following theorem, whose proof we omit, is the basis for the definition of the determinant.

Theorem 1 (Laplace Expansion Formulas). *Suppose A is an $n \times n$ matrix. Then the following numbers are all equal and we call this number the* ***determinant of A****:*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \mathrm{Minor}_{i\,j}(A) \quad \textit{for each } i$$

and

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \mathrm{Minor}_{i\,j}(A) \quad \textit{for each } j.$$

Any of these formulas can thus be taken as the definition of the determinant. In the first formula, the index i is fixed and the sum is taken over all j. The entries $a_{i,j}$ thus fill out the ith row. We therefore call this formula the ***Laplace expansion of***

the determinant along the ith row or simply a *row expansion*. Since the index i can range from 1 to n, there are n row expansions. In a similar way, the second formula is called the ***Laplace expansion of the determinant along the jth column*** or simply a ***column expansion*** and there are n column expansions. The presence of the factor $(-1)^{i+j}$ alternates the signs along the row or column according as $i + j$ is even or odd. The ***sign matrix***

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

is a useful tool to organize the signs in an expansion.

It is common to use the absolute value sign $|A|$ to denote the determinant of A. This should not cause confusion unless A is a 1×1 matrix, in which case we will not use this notation.

Example 2. Find the determinant of the matrix

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

► **Solution.** For purposes of illustration, we compute the determinant in two ways. First, we expand along the first row:

$$\det A = 1 \cdot \begin{vmatrix} -2 & 4 \\ 0 & 5 \end{vmatrix} - 2 \begin{vmatrix} 3 & 4 \\ 1 & 5 \end{vmatrix} + (-2) \begin{vmatrix} 3 & -2 \\ 1 & 0 \end{vmatrix} = 1 \cdot (-10) - 2 \cdot (11) - 2(2) = -36.$$

Second, we expand along the second column:

$$\det A = (-)2 \begin{vmatrix} 3 & 4 \\ 1 & 5 \end{vmatrix} + (-2) \begin{vmatrix} 1 & -2 \\ 1 & 5 \end{vmatrix} (-)0 \begin{vmatrix} 1 & -2 \\ 3 & 4 \end{vmatrix} = (-2) \cdot 11 - 2 \cdot (7) = -36.$$

Of course, we get the same answer; that is what the theorem guarantees. Observe though that the second column has a zero entry which means that we really only needed to compute two minors. In practice, we usually try to use an expansion along a row or column that has a lot of zeros. Also note that we use the sign matrix to adjust the signs on the appropriate terms. ◄

Properties of the Determinant

The determinant has many important properties. The three listed below show how the elementary row operations affect the determinant. They are used extensively to simplify many calculations.

Proposition 3. *Let A be an $n \times n$ matrix. Then*

1. $\det p_{i,j}(A) = -\det A$.
2. $\det m_i(a)(A) = a \det A$.
3. $\det t_{i,j}(A) = \det A$.

Proof. We illustrate the proof for the 2×2 case. Let $A = \begin{bmatrix} r & s \\ t & u \end{bmatrix}$. We then have

1. $|p_{1,2}(A)| = \begin{vmatrix} t & u \\ r & s \end{vmatrix} = ts - ru = -|A|$.
2. $|m_1(a)(A)| = \begin{vmatrix} ar & as \\ t & u \end{vmatrix} = aru - ast = a|A|$.
3. $|t_{1,2}(a)(A)| = \begin{vmatrix} r & s \\ t + ar & u + as \end{vmatrix} = r(u + as) - s(t + ar) = ru - st = |A|$. □

Another way to express 2. is

$2'$. If $A' = m_i(a)A$ then $\det A = \frac{1}{a} \det A'$.

Corollary 4. *Let E be an elementary matrix. Consider the three cases: If*

1. $E = p_{i\,j} I$, *then* $\det E = -1$.
2. $E = m_i(a) I$, *then* $\det E = a$.
3. $E = t_{i\,j} I$, *then* $\det E = 1$.

Furthermore,

$$\det EA = \det E \det A.$$

Proof. Let $A = I$ in Proposition 3 to get the stated formulas for $\det E$. Now let A be an arbitrary square matrix. The statement $\det EA = \det E \det A$ is now just a restatement of Proposition 3 and the fact that $eA = EA$ for an elementary row operation e and its associated elementary matrix E. □

Further important properties include:

Proposition 5. *1.* $\det A = \det A^t$.
2. If A has a zero row (or column), then $\det A = 0$.
3. If A has two equal rows (or columns), then $\det A = 0$.
4. If A is upper or lower triangular, then the determinant of A is the product of the diagonal entries.

Proof. 1. The transpose changes row expansions to column expansions and column expansions to row expansions. By Theorem 1, they are all the same.
2. All coefficients of the minors in an expansion along a zero row (column) are zero so the determinant is zero.
3. If the ith and jth rows are equal, then $\det A = \det(p_{i,j}A) = -\det A$ and this implies $\det A = 0$. If A has two equal columns, then A^t has two equal rows. Thus, $\det A = \det A^t = 0$.
4. Suppose A is upper triangular. Expansion along the first column gives $a_{11}Minor_{11}(A)$. But $Minor_{11}(A)$ is an upper triangular matrix of size one less than A. By induction, $\det A$ is the product of the diagonal entries. Since the transpose changes lower triangular matrices to upper triangular matrices, we get that the determinant of a lower triangular matrix is likewise the product of the diagonal entries. □

Example 6. Use elementary row operations to find $\det A$ if

$$(1)\ A = \begin{bmatrix} 2 & 4 & 2 \\ -1 & 3 & 5 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad (2)\ A = \begin{bmatrix} 1 & 0 & 5 & 1 \\ -1 & 2 & 1 & 3 \\ 2 & 2 & 16 & 6 \\ 3 & 1 & 0 & 1 \end{bmatrix}.$$

► **Solution.** Again we will write the elementary row operation that we have used above the equal sign.

$$(1)\quad \begin{vmatrix} 2 & 4 & 2 \\ -1 & 3 & 5 \\ 0 & 1 & 1 \end{vmatrix} \overset{m_1(\frac{1}{2})}{=} 2\begin{vmatrix} 1 & 2 & 1 \\ -1 & 3 & 5 \\ 0 & 1 & 1 \end{vmatrix} \overset{t_{12}(1)}{=} 2\begin{vmatrix} 1 & 2 & 1 \\ 0 & 5 & 6 \\ 0 & 1 & 1 \end{vmatrix}$$

$$\overset{p_{23}}{=} -2\begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 5 & 6 \end{vmatrix} \overset{t_{23}(-5)}{=} -2\begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = -2$$

For the first equality, we have used (3′) above, and in the last equality, we have used the fact that the last matrix is upper triangular and its determinant is the product of the diagonal entries:

$$(2) \qquad \begin{vmatrix} 1 & 0 & 5 & 1 \\ -1 & 2 & 1 & 3 \\ 2 & 2 & 16 & 6 \\ 3 & 1 & 0 & 1 \end{vmatrix} \overset{\substack{t_{12}(1) \\ t_{13}(-2) \\ t_{14}(-3)}}{=} \begin{vmatrix} 1 & 0 & 5 & 1 \\ 0 & 2 & 6 & 4 \\ 0 & 2 & 6 & 4 \\ 0 & 1 & -15 & -2 \end{vmatrix} = 0,$$

with the last equality because two rows are equal. ◀

In the following example, we use elementary row operations to zero out entries in a column and then use a Laplace expansion formula.

Example 7. Find the determinant of

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

► **Solution.**

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 4 & 2 & -1 \\ 2 & 2 & 3 & 0 \\ -1 & 1 & 2 & 4 \\ 0 & 1 & 3 & 2 \end{vmatrix} \overset{\substack{t_{1,2}(-2) \\ t_{1,3}(1)}}{=} \begin{vmatrix} 1 & 4 & 2 & -1 \\ 0 & -6 & -1 & 2 \\ 0 & 5 & 4 & 3 \\ 0 & 1 & 3 & 2 \end{vmatrix} \\ &= \begin{vmatrix} -6 & -1 & 2 \\ 5 & 4 & 3 \\ 1 & 3 & 2 \end{vmatrix} \overset{\substack{t_{3,1}(6) \\ t_{3,2}(-5)}}{=} \begin{vmatrix} 0 & 17 & 14 \\ 0 & -11 & -7 \\ 1 & 3 & 2 \end{vmatrix} \\ &= \begin{vmatrix} 17 & 14 \\ -11 & -7 \end{vmatrix} = -119 + 154 = 35, \end{aligned}$$

◀

The following two theorems state very important properties about the determinant.

Theorem 8. *A square matrix A is invertible if and only if* $\det A \neq 0$.

Proof. Apply Gauss–Jordan to A to get R, a matrix in row reduce echelon form. There is a sequence of elementary row operations $e_1, \ldots, e_k$ such that

$$e_1 \cdots e_k A = R.$$

Let $E_i = e_i I$, $i = 1, \ldots, k$ be the corresponding elementary matrices. Then

$$E_1 \cdots E_k A = R.$$

By repeatedly using Corollary 4, we get

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

Now suppose A is invertible. By Theorem 10 of Sect. 8.3, $R = I$. Since each factor $\det E_i \neq 0$ by Corollary 4, it follows that $\det A \neq 0$. On the other hand, if A is not invertible, then R has a zero row by Theorem 10 of Sect. 8.3 and Corollary 11 of Sect. 8.3. By Proposition 5, we have $\det R = 0$. Since each factor E_i has a nonzero determinant, it follows that $\det A = 0$. □

Theorem 9. *If A and B are square matrices of the same size, then*

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

Proof. We consider two cases.

1. Suppose A is invertible. Then there is a sequence of elementary matrices such that $A = E_1 \cdots E_k$. Now repeatedly use Corollary 4 to get

$$\det AB = \det E_1 \cdots E_k B = \det E_1 \cdots \det E_k \det B = \det A \det B.$$

2. Now suppose A is not invertible. Then AB is not invertible for otherwise there would be a C such that $(AB)C = I$. But by associativity of the product, we have $A(BC) = I$, and this implies A is invertible (with inverse BC). Now by Theorem 8, we have $\det AB = 0$ and $\det A = 0$ and the result follows. □

The Cofactor and Adjoint Matrices

Again, let A be a square matrix. We define the ***cofactor*** matrix, $\mathrm{Cof}(A)$, of A to be the matrix whose (i, j)-entry is $(-1)^{i+j}\mathrm{Minor}_{i,j}$. We define the ***adjoint*** matrix, $\mathrm{Adj}(A)$, of A by the formula $\mathrm{Adj}(A) = (\mathrm{Cof}(A))^t$. The important role of the adjoint matrix is seen in the following theorem and its corollary.

Theorem 10. *For A a square matrix, we have*

$$A\ \mathrm{Adj}(A) = \mathrm{Adj}(A)\ A = \det(A)I.$$

Proof. The (i, j) entry of $A\ \mathrm{Adj}(A)$ is

$$\sum_{k=0}^{n} A_{i\,k}(\mathrm{Adj}(A))_{k\,j} = \sum_{k=0}^{n}(-1)^{k+j} A_{i\,k}\mathrm{Minor}_{j\,k}(A).$$

When $i = j$, this is a Laplace expansion formula and is hence $\det A$ by Theorem 1. When $i \neq j$, this is the expansion of a determinant for a matrix with two equal rows and hence is zero by Proposition 5. □

The following corollary immediately follows.

Corollary 11 (The Adjoint Inversion Formula). *If* $\det A \neq 0$, *then*

$$A^{-1} = \frac{1}{\det A}\mathrm{Adj}(A).$$

The inverse of a 2×2 matrix is a simple matter: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then $\mathrm{Adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ and if $\det(A) = ad - bd \neq 0$, then

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{1}$$

For an example, suppose $A = \begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix}$. Then $\det(A) = 1 - (6) = -5 \neq 0$ so A is invertible and $A^{-1} = \frac{-1}{5}\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} \frac{-1}{5} & \frac{-3}{5} \\ \frac{-2}{5} & \frac{-1}{5} \end{bmatrix}$.

The general formula for the inverse of a 3×3 is substantially more complicated and difficult to remember. Consider though an example.

Example 12. Let

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

Find its inverse if it is invertible.

► **Solution.** We expand along the first row to compute the determinant and get $\det(A) = 1\det\begin{bmatrix} 4 & 1 \\ 0 & 3 \end{bmatrix} - 2\det\begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix} = 1(12) - 2(4) = 4$. Thus, A is invertible.

The cofactor of A is $\mathrm{Cof}(A) = \begin{bmatrix} 12 & -4 & 4 \\ -6 & 3 & -2 \\ 2 & -1 & 2 \end{bmatrix}$ and $\mathrm{Adj}(A) = \mathrm{Cof}(A)^t = \begin{bmatrix} 12 & -6 & 2 \\ -4 & 3 & -1 \\ 4 & -2 & 2 \end{bmatrix}$. The inverse of A is thus

$$A^{-1} = \frac{1}{4}\begin{bmatrix} 12 & -6 & 2 \\ -4 & 3 & -1 \\ 4 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 3 & \frac{-3}{2} & \frac{1}{2} \\ -1 & \frac{3}{4} & \frac{-1}{4} \\ 1 & \frac{-1}{2} & \frac{1}{2} \end{bmatrix}.$$ ◀

In our next example, we will consider a matrix with entries in $\mathcal{R} = \mathbb{R}[s]$. Such matrices will arise naturally in Chap. 9.

Example 13. Let

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

Find the inverse of the matrix

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

► **Solution.** A straightforward computation gives

$$\det(sI - A) = (s-4)(s^2+1).$$

The matrix of minors for $sI - A$ is

$$\begin{bmatrix} (s-1)(s-2)-3 & -3 & s-1 \\ -2(s-2)-1 & (s-1)(s-2)-1 & -(s-1)-2 \\ 6+(s-1) & -3(s-1) & (s-1)^2 \end{bmatrix}.$$

After simplifying, we obtain the cofactor matrix

$$\begin{bmatrix} s^2-3s-1 & 3 & s-1 \\ 2s-3 & s^2-3s+1 & s+1 \\ s+5 & 3s-3 & (s-1)^2 \end{bmatrix}.$$

The adjoint matrix is

$$\begin{bmatrix} s^2-3s-1 & 2s-3 & s+5 \\ 3 & s^2-3s+1 & 3s-3 \\ s-1 & s+1 & (s-1)^2 \end{bmatrix}.$$

Finally, we obtain the inverse:

$$(sI-A)^{-1} = \begin{bmatrix} \frac{s^2-3s-1}{(s-4)(s^2+1)} & \frac{2s-3}{(s-4)(s^2+1)} & \frac{s+5}{(s-4)(s^2+1)} \\ \frac{3}{(s-4)(s^2+1)} & \frac{s^2-3s+1}{(s-4)(s^2+1)} & \frac{3s-3}{(s-4)(s^2+1)} \\ \frac{s-1}{(s-4)(s^2+1)} & \frac{s+1}{(s-4)(s^2+1)} & \frac{(s-1)^2}{(s-4)(s^2+1)} \end{bmatrix}. \qquad \blacktriangleleft$$

Cramer's Rule

We finally consider a well-known theoretical tool used to solve a system $A\mathbf{x} = \mathbf{b}$ when A is invertible. Let $A(i, \mathbf{b})$ denote the matrix obtained by replacing the ith column of A with the column vector $\mathbf{b}$. We then have the following theorem:

Theorem 14. *Suppose* $\det A \neq 0$. *Then the solution to* $A\mathbf{x} = \mathbf{b}$ *is given coordinate-wise by the formula*

$$\mathbf{x}_i = \frac{\det A(i, \mathbf{b})}{\det A}.$$

Proof. Since A is invertible, we have

$$\begin{aligned} \mathbf{x}_i &= (A^{-1}\mathbf{b})_i = \sum_{k=1}^{n} (A^{-1})_{i\,k} \mathbf{b}_k \\ &= \frac{1}{\det A} \sum_{k=1}^{n} (-1)^{i+k} \mathrm{Minor}_{k\,i}(A) \mathbf{b}_k \\ &= \frac{1}{\det(A)} \sum_{k=1}^{n} (-1)^{i+k} \mathbf{b}_k \mathrm{Minor}_{k\,i}(A) = \frac{\det A(i, \mathbf{b})}{\det A}. \end{aligned} \qquad \square$$

The following example should convince you that Cramer's rule is mainly a theoretical tool and not a practical one for solving a system of linear equations. The Gauss–Jordan elimination method is usually far more efficient than computing $n+1$ determinants for a system $A\mathbf{x} = \mathbf{b}$, where A is $n \times n$.

Example 15. Solve the following system of linear equations using Cramer's rule:

$$\begin{aligned} x + y + z &= 0 \\ 2x + 3y - z &= 11 \\ x \quad\;\; + z &= -2 \end{aligned}$$

► **Solution.** We have

$$\det A = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & -1 \\ 1 & 0 & 1 \end{vmatrix} = -3,$$

$$\det A(1,\mathbf{b}) = \begin{vmatrix} 0 & 1 & 1 \\ 11 & 3 & -1 \\ -2 & 0 & 1 \end{vmatrix} = -3,$$

$$\det A(2,\mathbf{b}) = \begin{vmatrix} 1 & 0 & 1 \\ 2 & 11 & -1 \\ 1 & -2 & 1 \end{vmatrix} = -6,$$

$$\text{and } \det A(3,\mathbf{b}) = \begin{vmatrix} 1 & 1 & 0 \\ 2 & 3 & 11 \\ 1 & 0 & -2 \end{vmatrix} = 9,$$

where $\mathbf{b} = \begin{bmatrix} 0 \\ 11 \\ -2 \end{bmatrix}$. Since $\det A \neq 0$, Cramer's rule gives

$$x_1 = \frac{\det A(1,\mathbf{b})}{\det A} = \frac{-3}{-3} = 1,$$

$$x_2 = \frac{\det A(2,\mathbf{b})}{\det A} = \frac{-6}{-3} = 2,$$

and

$$x_3 = \frac{\det A(3,\mathbf{b})}{\det A} = \frac{9}{-3} = -3.$$

◄

Exercises

1–9. Find the determinant of each matrix given below in three ways: a row expansion, a column expansion, and using row operations to reduce to a triangular matrix.

1. $\begin{bmatrix} 1 & 4 \\ 2 & 9 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 1 \\ 4 & 4 \end{bmatrix}$

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

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

5. $\begin{bmatrix} 4 & 0 & 3 \\ 8 & 1 & 7 \\ 3 & 4 & 1 \end{bmatrix}$

6. $\begin{bmatrix} 3 & 98 & 100 \\ 0 & 2 & 99 \\ 0 & 0 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 0 & 1 & -2 & 4 \\ 2 & 3 & 9 & 2 \\ 1 & 4 & 8 & 3 \\ -2 & 3 & -2 & 4 \end{bmatrix}$

8. $\begin{bmatrix} -4 & 9 & -4 & 1 \\ 2 & 3 & 0 & -4 \\ -2 & 3 & 5 & -6 \\ -3 & 2 & 0 & 1 \end{bmatrix}$

9. $\begin{bmatrix} 2 & 4 & 2 & 3 \\ 1 & 2 & 1 & 4 \\ 4 & 8 & 4 & 6 \\ 1 & 9 & 11 & 13 \end{bmatrix}$

10–15. Find the inverse of $(sI - A)$ and determine for which values of s $\det(sI - A) = 0$.

10. $\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$

11. $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}$

14. $\begin{bmatrix} 1 & -3 & 3 \\ -3 & 1 & 3 \\ 3 & -3 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 0 & 4 & 0 \\ -1 & 0 & 0 \\ 1 & 4 & -1 \end{bmatrix}$

16–24. Use the adjoint formula for the inverse for the matrices given below.

16. $\begin{bmatrix} 1 & 4 \\ 2 & 9 \end{bmatrix}$

17. $\begin{bmatrix} 1 & 1 \\ 4 & 4 \end{bmatrix}$

18. $\begin{bmatrix} 3 & 4 \\ 2 & 6 \end{bmatrix}$

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

20. $\begin{bmatrix} 4 & 0 & 3 \\ 8 & 1 & 7 \\ 3 & 4 & 1 \end{bmatrix}$

21. $\begin{bmatrix} 3 & 98 & 100 \\ 0 & 2 & 99 \\ 0 & 0 & 1 \end{bmatrix}$

22. $\begin{bmatrix} 0 & 1 & -2 & 4 \\ 2 & 3 & 9 & 2 \\ 1 & 4 & 8 & 3 \\ -2 & 3 & -2 & 4 \end{bmatrix}$

23. $\begin{bmatrix} -4 & 9 & -4 & 1 \\ 2 & 3 & 0 & -4 \\ -2 & 3 & 5 & -6 \\ -3 & 2 & 0 & 1 \end{bmatrix}$

24. $\begin{bmatrix} 2 & 4 & 2 & 3 \\ 1 & 2 & 1 & 4 \\ 4 & 8 & 4 & 6 \\ 1 & 9 & 11 & 13 \end{bmatrix}$

25–28. Use Cramer's rule to solve the system $A\mathbf{x} = \mathbf{b}$ for the given matrices A and $\mathbf{b}$.

25. $A = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

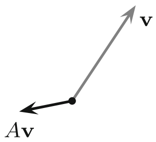
26. $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$

27. $A = \begin{bmatrix} 1 & 0 & -2 \\ 2 & -2 & 0 \\ 1 & 2 & -1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$

28. $A = \begin{bmatrix} 1 & 3 & 0 & 1 \\ 2 & 2 & -2 & 0 \\ 1 & -1 & 0 & 4 \\ 1 & 2 & 3 & 9 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}$

8.5 Eigenvectors and Eigenvalues

Suppose A is a square $n \times n$ matrix. Again, it is convenient to think of $\mathbb{R}^n$ as the set of column vectors $M_{n,1}(\mathbb{R})$. If $\mathbf{v} \in \mathbb{R}^n$, then A transforms $\mathbf{v}$ to a new vector $A\mathbf{v}$ as seen below.

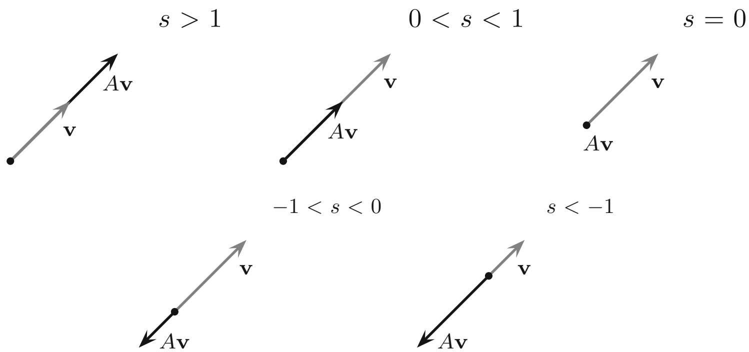


As illustrated, A rotates and compresses (or stretches) a given vector $\mathbf{v}$. However, if a vector $\mathbf{v}$ points in the right direction, then A acts in a much simpler manner.

We say that s is an ***eigenvalue*** of A if there is a ***nonzero*** vector $\mathbf{v}$ in $\mathbb{R}^n$ such that

$$A\mathbf{v} = s\mathbf{v}. \tag{1}$$

The vector $\mathbf{v}$ is called an ***eigenvector***[1] associated to s. The pair $(s, \mathbf{v})$ is called an ***eigenpair***. One should think of eigenvectors as the directions for which A acts by stretching or compressing vectors by the length determined by the eigenvalue s: if $s > 1$, the eigenvector is stretched; if $0 < s < 1$, then the eigenvector is compressed; if $s < 0$, the direction of the eigenvector is reversed; and if $s = 0$, the eigenvector is in the null space of A. See the illustration below.



This notion is very important and has broad applications in mathematics, computer science, physics, engineering, and economics. For example, the Google page rank algorithm is based on this concept.

[1]Eigenvectors and eigenvalues are also called characteristic vectors and characteristic values, respectively.

In this section, we discuss how to find the eigenpairs for a given matrix A. We begin with a simple example.

Example 1. Suppose $A = \begin{bmatrix} -3 & 1 \\ -4 & 2 \end{bmatrix}$. Let

$$\mathbf{v_1} = \begin{bmatrix} 2 \\ 8 \end{bmatrix}, \ \mathbf{v_2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \mathbf{v_3} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \text{and} \ \mathbf{v_4} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Show that $\mathbf{v_1}$ and $\mathbf{v_2}$ are eigenvectors for A. What are the associated eigenvalues? Show that $\mathbf{v_3}$ and $\mathbf{v_4}$ are not eigenvectors.

► **Solution.** We simply observe that

1. $$A\mathbf{v_1} = \begin{bmatrix} -3 & 1 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} = 1 \begin{bmatrix} 2 \\ 8 \end{bmatrix}.$$

   Thus, $\mathbf{v_1}$ an eigenvector with eigenvalue 1.

2. $$A\mathbf{v_2} = \begin{bmatrix} -3 & 1 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

   Thus, $\mathbf{v_2}$ an eigenvector with eigenvalue -2.

3. $$A\mathbf{v_3} = \begin{bmatrix} -3 & 1 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ -6 \end{bmatrix}.$$

   We see that $A\mathbf{v_3}$ is not a multiple of $\mathbf{v_3}$. It is not an eigenvector.
4. Eigenvectors must be nonzero so $\mathbf{v_4}$ is not an eigenvector. ◄

To find the eigenvectors and eigenvalues for A, we analyze (1) a little closer. Let us rewrite it as $sv - Av = 0$. By inserting the identity matrix, we get $sIv - Av = 0$, and by the distributive property, we see (1) is equivalent to

$$(sI - A)v = 0. \tag{2}$$

Said another way, a nonzero vector $\mathbf{v}$ is an eigenvector for A if and only if it is in the null space of $sI - A$ for an eigenvalue s. Let E_s be the null space of $sI - A$; it is called the ***eigenspace*** for A with eigenvalue s. So once an eigenvalue is known, the corresponding eigenspaces are easily computed. How does one determine the eigenvalues? By Theorems 10 of Sect. 8.3 and 8 of Sect. 8.4, s is an eigenvalue if and only if

$$\det(sI - A) = 0. \tag{3}$$

As a function of s, we let $c_A(s) = \det(sI - A)$; it is called the ***characteristic polynomial of*** A, for it is a polynomial of degree n (assuming A is an $n \times n$ matrix). The matrix $sI - A$ is called the ***characteristic matrix***, and (3) is called the ***characteristic equation***. By solving the characteristic equation, we determine the eigenvalues.

Example 2. Determine the characteristic polynomial, $c_A(s)$, for

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

as given in Example 1. Find the eigenvalues and corresponding eigenspaces.

► **Solution.** The characteristic matrix is

$$sI - A = \begin{bmatrix} s+3 & -1 \\ 4 & s-2 \end{bmatrix},$$

and the characteristic polynomial is given as follows:

$$\begin{aligned} c_A(s) &= \det(sI - A) = \det \begin{bmatrix} s+3 & -1 \\ 4 & s-2 \end{bmatrix} \\ &= (s+3)(s-2) + 4 = s^2 + s - 2 = (s+2)(s-1). \end{aligned}$$

The characteristic equation is $(s+2)(s-1) = 0$, and hence, the eigenvalues are -2 and 1. The eigenspaces, E_s, are computed as follows:

E_{-2}: We let $s = -2$ in the characteristic matrix and row reduce the corresponding augmented matrix $[-2I - A|\mathbf{0}]$ to get the null space of $-2I - A$. We get

$$\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 4 & -4 & 0 \end{array}\right] \xrightarrow{t_{1\,2}(-4)} \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

If x and y are the variables, then y is a free variable. Let $y = \alpha$. Then $x = y = \alpha$. From this, we see that

$$E_{-2} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

E_1: We let $s = 1$ in the characteristic matrix and row reduce the corresponding augmented matrix $[1I - A|\mathbf{0}]$ to get the null space of $1I - A$. We get

$$\left[\begin{array}{cc|c} 4 & -1 & 0 \\ 4 & -1 & 0 \end{array}\right] \xrightarrow{m_1(1/4)} \left[\begin{array}{cc|c} 1 & -1/4 & 0 \\ 4 & -1 & 0 \end{array}\right] \xrightarrow{t_{12}(-4)} \left[\begin{array}{cc|c} 1 & -1/4 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Again $y = \alpha$ is the free variable and $x = \frac{1}{4}\alpha$. It follows that the null space of $1I - A$ is all multiples of the vector $\begin{bmatrix} 1/4 \\ 1 \end{bmatrix}$. Since we are considering all multiples of a vector, we can clear the fraction (by multiplying by 4) and write

$$E_1 = \mathrm{Span}\left\{\begin{bmatrix} 1 \\ 4 \end{bmatrix}\right\}.$$

We will routinely do this. ◀

Remark 3. In Example 1, we found that the vector $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 8 \end{bmatrix}$ was an eigenvector with eigenvalue 1. We observe that $\mathbf{v}_1 = 2\begin{bmatrix} 1 \\ 4 \end{bmatrix} \in E_1$. In like manner, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in E_{-2}$.

Example 4. Determine the characteristic polynomial for

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

Find the eigenvalues and corresponding eigenspaces.

► **Solution.** The characteristic matrix is

$$sI - A = \begin{bmatrix} -3 & 1 & -1 \\ -3 & s+1 & -3 \\ -1 & 1 & s-3 \end{bmatrix},$$

and the characteristic polynomial is calculated by expanding along the first row as follows:

$$\begin{aligned} c_A(s) &= \det(sI - A) = \det\begin{bmatrix} s-3 & 1 & -1 \\ -3 & s+1 & -3 \\ -1 & 1 & s-3 \end{bmatrix} \\ &= (s-3)((s+1)(s-3)+3) - (-3(s-3)-3) - (-3+s+1) \end{aligned}$$

$$= s^3 - 5s^2 + 8s - 4 = (s-1)(s-2)^2.$$

It follows that the eigenvalues are 1 and 2. The eigenspaces, E_s, are computed as follows:

E_1: We let $s = 1$ in the characteristic matrix and row reduce the corresponding augmented matrix $[I - A|\mathbf{0}]$. We forego the details but we get

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

If x, y, and z are the variables, then z is the free variable. Let $z = \alpha$. Then $x = \alpha$, $y = 3\alpha$, and $z = \alpha$. From this, we see that

$$E_1 = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}\right\}.$$

E_2: We let $s = 2$ in the characteristic matrix and row reduce the corresponding augmented matrix $[2I - A|\mathbf{0}]$ to get

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Here we see that $y = \alpha$ and $z = \beta$ are free variable and $x = \alpha - \beta$. It follows that the null space of $2I - A$ is

$$E_2 = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}.$$ ◀

For a diagonal matrix, the eigenpairs are simple to find. Let

$$A = \operatorname{diag}(a_1, \ldots, a_n) = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & a_n \end{bmatrix}.$$

Let $\mathbf{e}_i$ be the column vector in $\mathbb{R}^n$ with 1 in the ith position and zeros elsewhere; $\mathbf{e}_i$ is the ith column of the identity matrix I_n. A simple calculation gives

$$A\mathbf{e}_i = a_i \mathbf{e}_i.$$

Thus, for a diagonal matrix, the eigenvalues are the diagonal entries and the eigenvectors are the coordinate axes in $\mathbb{R}^n$. In other words, the coordinate axes point in the directions for which A scales vectors.

If A is a real matrix, then it can happen that the characteristic polynomial have complex roots. In this case, we view A a complex matrix and make all computation over the complex numbers. If $\mathbb{C}^n$ denotes the $n \times 1$ column vectors with entries in $\mathbb{C}$, then we view A as transforming vectors $\mathbf{v} \in \mathbb{C}^n$ to vectors $A\mathbf{v} \in \mathbb{C}^n$.

Example 5. Determine the characteristic polynomial for

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

Find the eigenvalues and corresponding eigenspaces.

► **Solution.** The characteristic matrix is

$$sI - A = \begin{bmatrix} s-2 & 5 \\ -1 & s+2 \end{bmatrix},$$

and the characteristic polynomial is given as follows:

$$\begin{aligned} c_A(s) &= \det(sI - A) = \det \begin{bmatrix} s-2 & 5 \\ -1 & s+2 \end{bmatrix} \\ &= (s-2)(s+2) + 5 = s^2 + 1. \end{aligned}$$

The eigenvalues are $\pm i$. The eigenspaces, E_s, are computed as follows:

E_i: We let $s = i$ in the characteristic matrix and row reduce the corresponding augmented matrix $[iI - A|\mathbf{0}]$. We get

$$\left[\begin{array}{cc|c} i-2 & 5 & 0 \\ -1 & i+2 & 0 \end{array}\right] \xrightarrow[m_1(-1)]{p_{12}} \left[\begin{array}{cc|c} 1 & -i-2 & 0 \\ i-2 & 5 & 0 \end{array}\right]$$
$$\xrightarrow{t_{12}(-(i-2))} \left[\begin{array}{cc|c} 1 & -i-2 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

If x and y are the variables, then y is a free variable. Let $y = \alpha$. Then $x = \alpha(i+2)$. From this, we see that

$$E_i = \text{Span}\left\{\begin{bmatrix} i+2 \\ 1 \end{bmatrix}\right\}.$$

E_{-i}: We let $s = -i$ in the characteristic matrix and row reduce the corresponding augmented matrix $[-iI - A|\mathbf{0}]$. We get

$$\left[\begin{array}{cc|c} -i-2 & 5 & 0 \\ -1 & -i+2 & 0 \end{array}\right] \xrightarrow[m_1(-1)]{p_{12}} \left[\begin{array}{cc|c} 1 & i-2 & 0 \\ -i-2 & 5 & 0 \end{array}\right]$$
$$\xrightarrow{t_{12}(i+2)} \left[\begin{array}{cc|c} 1 & i-2 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Again $y = \alpha$ is the free variable and $x = (2-i)\alpha$. It follows that

$$E_{-i} = \text{Span}\left\{\begin{bmatrix} 2-i \\ 1 \end{bmatrix}\right\}.$$ ◀

Remark 6. It should be noted in this example that the eigenvalues and eigenvectors are complex conjugate of one another. In other words, if $s = i$ is an eigenvalue, then $\overline{s} = -i$ is another eigenvalue. Further, if $\mathbf{v} = \alpha \begin{bmatrix} i+2 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue i, then $\overline{\mathbf{v}} = \alpha \begin{bmatrix} -i+2 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue $\overline{i} = -i$ and vice versa. The following theorem shows that this happens in general as long as A is a real matrix. Thus, E_{-i} may be computed by simply taking the complex conjugate of E_i.

Theorem 7. *Suppose A is a real $n \times n$ matrix with a complex eigenvalue s. Then $\overline{s}$ is an eigenvalue and*

$$E_{\overline{s}} = \overline{E_s}.$$

Proof. Suppose s is a complex eigenvalue and $\mathbf{v} \in E_s$ is a corresponding eigenvector. Then $A\mathbf{v} = s\mathbf{v}$. Taking complex conjugates and keeping in mind that $\overline{A} = A$ since A is real, we get $A\overline{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{s\mathbf{v}} = \overline{s}\,\overline{\mathbf{v}}$. It follows that $\overline{\mathbf{v}} \in E_{\overline{s}}$ and $\overline{E_s} \subset E_{\overline{s}}$. This argument is symmetric so $\overline{E_{\overline{s}}} \subset E_s$. These two statements imply $\overline{E_s} = E_{\overline{s}}$. □

Exercises

1–7. For each of the following matrices A determine the characteristic polynomial, the eigenvalues, and the eigenspaces.

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

2. $A = \begin{bmatrix} -6 & -5 \\ 7 & 6 \end{bmatrix}$

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

4. $A = \begin{bmatrix} 7 & 4 \\ -16 & -9 \end{bmatrix}$

5. $A = \begin{bmatrix} 9 & 12 \\ -8 & -11 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$

7. $A = \begin{bmatrix} 20 & -15 \\ 30 & -22 \end{bmatrix}$

8–13. For the following problems, A and its characteristic polynomial are given. Find the eigenspaces for each eigenvalue.

8. $A = \begin{bmatrix} 3 & -2 & 2 \\ 9 & -7 & 9 \\ 5 & -4 & 6 \end{bmatrix}$, $c_A(s) = (s+1)(s-1)(s-2)$

9. $A = \begin{bmatrix} 8 & -5 & 5 \\ 0 & -2 & 0 \\ -10 & 5 & -7 \end{bmatrix}$, $c_A(s) = (s+2)^2(s-3)$

10. $A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & -1 & 4 \\ 0 & -2 & 5 \end{bmatrix}$, $c_A(s) = (s-1)^2(s-3)$

11. $A = \begin{bmatrix} 2 & 0 & 0 \\ 2 & -3 & 6 \\ 1 & -3 & 6 \end{bmatrix}, c_A(s) = s(s-2)(s-3)$

12. $A = \begin{bmatrix} -6 & 11 & -16 \\ 4 & -4 & 8 \\ 7 & -10 & 16 \end{bmatrix}, c_A(s) = (s-2)^3$

13. $A = \begin{bmatrix} -8 & 13 & -19 \\ 8 & -8 & 14 \\ 11 & -14 & 22 \end{bmatrix}, c_A(s) = (s-2)(s^2-4s+5)$