Linear Algebra Notes

Shreas

September 28, 2025

Contents

Chapter 1

Linear Equations

1.1 Fields

We let F denote either the set of real numbers or the the set of complex numbers.

1. Addition is commutative,
$$x + y = y + x$$
for all x and y in F.

2. Addition is associative,
$$x + (y + z) = (x + y) + z$$
for all x, y and z in F.

3. There is a unique element 0 (zero) in F such that $x + 0 = x$, for every $x \in F$.

4. To each $x \in F$ there corresponds a unique element $(-x) \in F$ such that $x + (-x) = 0$.

5. Multiplication is commutative,
$$xy = yx$$
for all $x, y \in F$.

6. Multiplication is associative,
$$x(yz) = (xy)z$$
for all $x, y, z \in F$.

7. There is a unique non-zero element 1 (one) in F such that $x1 = x \ \forall x \in F$.

8. To each non-zero $x \in F$ there corresponds a unique element $x^{-1} \in F$ such that $xx^{-1} = 1$.

9. Multiplication distributes over addition; that is, $x(y + z) = xy + xz, \ \forall x, y, z \in F$.

Suppose one has a set F of objects $x, y, z, \ldots$ and two operations on the elements of F as follows. The first operation, called addition, associates with each pair of elements $x, y \in F$ an element $(x + y) \in F$; the second operation, called multiplication, associates with each pair x, y an element $xy \in F$; and these two operations satisfy conditions (1)-(9) above. The set F, together with these two operations, is then called a **field**.

1.2 Systems of Linear Equations

Suppose F is a field. We consider the problem of finding n scalars $x_1, \ldots, x_n$ which satisfy the conditions

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= y_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= y_2 \\ \vdots \quad + \quad \vdots \quad + \cdots + \quad \vdots \quad &= \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= y_m \end{aligned} \tag{1.1}$$

where $y_1, \cdots, y_m$ and A_{ij}, $1 \le i \le m$, $1 \le j \le n$, are given elements of F. We call 1.1 a **system of** m **linear equations in** n **unknowns**. Any n-tuple $(x_1, \cdots, x_n)$ of elements of F which satisfies each of the equations in 1.1 is called a solution of the system. If $y_1 = y_2 = \cdots = y_m = 0$, we say that the system is **homogeneous**, or that each of the equations is homogeneous.

For the general equation 1.1, suppose we select m scalars $c_1, \ldots, c_m$, multiply the jth equation by c_j and then add. We obtain the equation

$$(c_1A_{11} + \cdots + c_mA_{m1})x_1 + \cdots + (c_1A_{1n} + \cdots + c_mA_{mn})x_n = c_1y_1 + \cdots + c_my_m.$$

Such an equation we shall call a **linear combination** of the equations in 1.1. Any solution of the entire system of equations 1.1 will also be a solution of this new equation. This is the fundamental idea of the elimination process.

Theorem 1.2.1. *Equivalent systems of linear equations have exactly the same solutions.*

1.3 Matrices and Elementary Row Operations

In forming linear combinations of linear equations there is no need to continue writing the unknowns $x_1, \ldots, x_n$, since one actually computes only with the coefficients A_{ij} and the scalars y_i. We shall now abbreviate the system 1.1 by

$$AX = Y$$

where

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}.$$

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

We call A the **matrix coefficients** of the system. The rectangular array displayed above is not a matrix, but it is a representation of a matrix. An $m \times n$ **matrix over the field** F is a function A from the set of pairs of integers $(i, j), 1 \leq i \leq m, 1 \leq j \leq n$ into the field F. The **entries** of the matrix A are the scalars $A(i, j) = A_{ij}$. Thus X is, or defines, an $n \times 1$ matrix and Y is an $m \times 1$ matrix.

We restrict our attention to three **elementary row operations** on an $m \times n$ matrix A over the field F:

1. multiplication of one row of A by a non-zero scalar c;

2. replacement of the rth row of A by row r plus c times row s, c any scalar and $r \neq s$;

3. interchange of two rows of A.

An elementary row operation is a special type of function e which associated with each $m \times n$ matrix A an $m \times n$ matrix $e(A)$. One can precisely describe e in the three cases as follows:

1. $e(A)_{ij} = A_{ij}$ if $i \neq r$, $e(A)_{rj} = cA_{rj}$.

2. $e(A)_{ij} = A_{ij}$ if $i \neq r$, $e(A)_{rj} = A_{rj} + cA_{sj}$.

3. $e(A)_{ij} = A_{ij}$ if i is different from both r and s, $e(A)_{rj} = A_{sj}, e(A)_{sj} = A_{rj}$.

In defining $e(A)$, it is not important how many columns A has, but the number of rows of A is crucial. We shall agree that an elementary row operation e is defined on the class of all $m \times n$ matrices over F, for some fixed m but any n. In other words, a particular e is defined on the class of all m-rowed matrices over F.

Theorem 1.3.1. *To each elementary row operation e there corresponds an elementary row operation e_1, of the same type as e, such that $e_1(e(A)) = e(e_1(A))$ for each A. In other words, the inverse operation of an elementary row operation exists and is an elementary row operation of the same type.*

Proof. (1) Suppose e is the operation which multiplies the rth row of a matrix by the non-zero scalar c. Let e_1 be the operation which multiplies row r by c^{-1}. (2) Suppose e is the operation which replaces row r by row r plus c times row s, $r \neq s$. Let e_1 be the operation which replaces row r by row r plus $(-c)$ times row s. (3) If e interchanges rows r and s, let $e_1 = e$. In each of these three cases we clearly have $e_1(e(A)) = e(e_1(A)) = A$ for each A.

□

Definition 1.3.2. If A and B are $m \times n$ matrices over the field F, we say that B is **row-equivalent to** A if B can be obtained from A by a finite sequence of elementary row operations.

Theorem 1.3.3. *If A and B are row-equivalent $m \times n$ matrices, the homogeneous systems of linear equations $AX = 0$ and $BX = 0$ have exactly the same solutions.*

Proof. Suppose we pass from A to B by a finite sequence of elementary row operations:

$$A = A_0 \to A_1 \to \cdots \to A_k = B.$$

It is enough to prove that the systems $A_jX = 0$ and $A_{j+1}X = 0$ have the same solutions, i.e., that one elementary row operation does not disturb the set of solutions.

So suppose that B is obtained from A by a single elementary row operation. No matter which of the three types the operation is, (1), (2), or (3), each equation in the system $BX = 0$

will be a linear combination of the equations in the system $AX = 0$. Since the inverse of an elementary row operation is an elementary row operation, each equation in $AX = 0$ will also be a linear combination of the equations in $BX = 0$. Hence these two systems are equivalent, and by Theorem 1.2.1 they have the same solutions.

□

Example 1.3.4. Suppose F is the field of complex numbers and

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

In performing row operations it is often convenient to combine several operations of type (2). With this in mind

$$\begin{bmatrix} -1 & i \\ -i & 3 \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 2+i \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} 0 & 1 \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Thus the system of equations

$$\begin{aligned} -x_1 + ix_2 &= 0 \\ -ix_1 + 3x_2 &= 0 \\ x_1 + 2x_2 &= 0 \end{aligned}$$

has only the trivial solutions $x_1 = x_2 = 0$.

Definition 1.3.5. An $m \times n$ matrix R is called **row-reduced** if:

1. the first non-zero entry in ech non-zero row of R is 1;
2. each column of R which contains the leading non-zero entry of some row has all its other entries 0.

Example 1.3.6. One example of a row-reduced matrix is the $n \times n$ identity matrix I. This is the

$n \times n$ matrix defined by

$$I_{ij} = \delta_{ij} = \begin{cases} 1, & \text{if} \quad i = j \\ 0, & \text{if} \quad i \neq j. \end{cases}$$

This is the first of many occasions on which we shall use the **Kronecker delta** (δ).

Two examples of matrices which are not row-reduced are:

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

The second matrix fails to satisfy condition (a), because the leading non-zero entry of the first row is not 1. The first matrix does satisfy condition (a), but fails to satisfy condition (b) in column 3.

Theorem 1.3.7. *Every $m \times n$ matrix over the field F is row-equivalent to a row-reduced matrix.*

Proof. Let A be an $m \times n$ matrix over F. If every entry in the first row of A is 0, then condition (a) is satisfied in so far as row 1 is concerned. If row 1 has a non-zero entry, let k be the smallest positive integer j for which $A_{1j} \neq 0$. Multiply row 1 by A_{1k}^{-1}, and then condition (a) is satisfied with regard to row 1. Now for each $i \geq 2$, add $(-A_{ik})$ times row 1 to row i. Now the leading non-zero entry of row 1 occurs in column k, that entry is 1, and every other entry in column k is 0.

Now consider the matrix which has resulted from above. If every entry in row 2 is 0, we do nothing to row 2. If some entry in row 2 is different from 0, we multiply row 2 by a scalar so that the leading non-zero entry is 1. In the event that row 1 had a leading non-zero entry in column k, this leading non-zero entry of row 2 cannot occur in column k; say it occurs in column $k_r \neq k$. By adding suitable multiples of row 2 to the various rows, we can arrange that all entries in column k' are 0, except the 1 in row 2. The important thing to notice is this: In carrying out these last operations, we will not change the entries of row 1 in column $1, \ldots, k$; nor will we change any entry of column k. Of course, if row 1 was idenically 0, the operations with row 2 will not affect row 1.

Working with one row at a time in the above manner, it is clear that in a finite number of steps we will arrive at a row-reduced matrix.

□

1.4 Row-Reduced Echelon Matrices

Definition 1.4.1. An $m \times n$ matrix R is called a **row-reduced echelon matrix** if:

1. R is row-reduced;
2. every row of R which has all its entries 0 occurs below every row which has a non-zero entry;
3. if rows $1, \ldots, r$ are the non-zero rows of R, and if the leading non-zero entry of row i occurs in column k_i, $i = 1, \ldots, r$, then $k_1 < k_2 < \cdots < k_r$.

One can also describe an $m \times n$ row-reduced echelon matrix R as follows. Either every entry in R is 0, or there exists a positive integer r, $1 \leq r \leq m$, and r positive integers $k_1, \cdots, k_r$ with $1 \leq k_i \leq n$ and

1. $R_{ij} = 0$ for $i > r$, and $R_{ij} = 0$ if $j < k_i$.
2. $R_{ik_j} = \delta_{ij}, 1 \leq i \leq r, 1 \leq j \leq r$.
3. $k_1 < \cdots < k_r$.

Example 1.4.2. Two examples of row-reduced echelon matrices are the $n \times n$ identity matrix, and the $m \times n$ **zero matrix** $0^{m,n}$, in which all entries are 0.

Theorem 1.4.3. *Every $m \times n$ matrix A is row-equivalent to a row-reduced echelon matrix.*

Proof. We know that A is row-equivalent to a row-reduced matrix. All that we need observe is that by performing a finite number of row interchanges on a row-reduced matrix we can bring it to a row-reduced echelon form.

□

Theorem 1.4.4. *If A is an $m \times n$ matrix and $m < n$, then the homogeneous system of linear equations $AX = 0$ has a non-trivial solution.*

Proof. Let R be a row-reduced echelon matrix which is row-equivalent to A. Then the system $AX = 0$ and $RX = 0$ have the same solutions by Theorem 1.3.3. If r is the number of non-zero rows in R, then certainly $r \leq m$, and since $m < n$, we have $r < n$. It follows immediately that $AX = 0$ has a non-trivial solution.

□

Theorem 1.4.5. *If A is an $n \times n$ (square) matrix, then A is row-equivalent to the $n \times n$ identity matrix if and only if the system of equations $AX = 0$ has only the trivial solutions.*

Proof. If A is row-equivalent to I, then $AX = 0$ and $IX = 0$ have the same solutions. Conversely, suppose $AX = 0$ has only the trivial solution $X = 0$. Let R be an $n \times n$ row-reduced echelon matrix which is row-equivalent to A, and let r be the number of non-zero rows of R. Then $RX = 0$ has no trivial solution. Thus $r \geq n$. But since R has n rows, certainly $r \leq n$, and we have $r = n$. Since this means that R actually has a leading non-zero entry of 1 in each of its n rows, and since these 1 's occur each in a different one of the n columns, R must be the $n \times n$ identity matrix.

□

We form the **augmented matrix** A' of the system $AX = Y$. This is the $m \times (n + 1)$ matrix whose first n columns are the columns of A and whose last column is Y. More precisely,

$$\begin{aligned} A'_{ij} &= A_{ij}, \text{ if } j \leq n \\ A'_{(n+1)} &= y_i. \end{aligned}$$

Suppose we perform a sequence of elementary row operations on A, arriving at a row-reduced echelon matrix R. If we perform this same sequence of row operations on the augmented matrix A', we will arrive at a matrix R' whose first n columns are the columns of R and whose last column contains certain scalars $z_1, \ldots, z_m$. The scalars z_i are the entries of the $m \times 1$ matrix

$$Z = \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}$$

which results from applying the sequence of row operations to the matrix Y. The systems $AX = Y$ and $RX = Z$ are equivalent and hence have the same solutions.

Exercises

Exercise 1.4.6. Find all the solutions to the following system of equations by row reducing the coefficient matrix:

$$\begin{aligned} \frac{1}{3}x_1 + 2x_2 - 6x_3 &= 0 \\ -4x_1 + 5x_3 &= 0 \\ -3x_1 + 6x_2 - 13x_3 &= 0 \\ -\frac{7}{3}x_1 + 2x_2 - \frac{8}{3}x_3 &= 0 \end{aligned}$$

Solution 1.4.7. The given homogeneous system is:

$$\begin{aligned} \frac{1}{3}x_1 + 2x_2 - 6x_3 &= 0 \\ -4x_1 + 5x_3 &= 0 \\ -3x_1 + 6x_2 - 13x_3 &= 0 \\ -\frac{7}{3}x_1 + 2x_2 - \frac{8}{3}x_3 &= 0 \end{aligned}$$

The coefficient matrix is:

$$A = \begin{bmatrix} \frac{1}{3} & 2 & -6 \\ -4 & 0 & 5 \\ -3 & 6 & -13 \\ -\frac{7}{3} & 2 & -\frac{8}{3} \end{bmatrix}$$

We seek all vectors $x = (x_1, x_2, x_3)^T$ such that $Ax = 0$. To simplify, multiply rows containing fractions by appropriate non-zero constants:

- Multiply row 1 by 3 : $R_1 \longrightarrow 3R_1$.
- Multiply row 4 by 3 : $R_3 \longrightarrow 3R_4$.

This yields an equivalent system with coefficient matrix:

$$B = \begin{bmatrix} 1 & 6 & -18 \\ -4 & 0 & 5 \\ -3 & 6 & -13 \\ -7 & 6 & -8 \end{bmatrix}$$

Let R_1, R_2, R_3, R_4 denote the rows of B. We can eliminate entries below the pivot in column 1:

- $R_2 \longrightarrow R_2 + 4R_1$:
$$(-4, 0, 5) + 4(1, 6, -18) = (0, 24, -67).$$
- $R_3 \longrightarrow R_3 + 3R_1$:
$$(-3, 6, -13) + 3(1, 6, -18) = (0, 24, -67).$$
- $R_4 \longrightarrow R_4 + 7R_1$:
$$(-7, 6, -8) + 7(1, 6, -18) = (0, 48, -134).$$

The matrix now becomes

$$\begin{bmatrix} 1 & 6 & -18 \\ 0 & 24 & -67 \\ 0 & 24 & -67 \\ 0 & 48 & -134 \end{bmatrix}$$

We can now eliminate redundant rows.

- $R_3 \to R_3 - R_2$:
$$(0, 24, -67) - (0, 24, -67) = (0, 0, 0).$$
- $R_4 \to R_4 - 2R_2$:
$$(0, 48, -134) - 2(0, 24, -67) = (0, 0, 0)$$

The reduced echelon row form is:

$$\begin{bmatrix} 1 & 6 & -18 \\ 0 & 24 & -67 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The non-zero rows correspond to the equations:

$$x_1 + 6x_2 - 18x_3 = 0 \tag{1.2}$$
$$24x_2 - 67x_3 = 0 \tag{1.3}$$

From equation (1.3)

$$24x_2 = 67x_3 \implies x_2 = \frac{67}{24}x_3.$$

We can substitute this into equation (1.2)

$$x_1 + 6\left(\frac{67}{24}x_3\right) - 18x_3 = 0 \implies x_1 + \frac{67}{4}x_3 - 18x_3 = 0$$

$$x_1 = 18x_3 - \frac{67}{4}x_3 = \frac{5}{4}x_3.$$

Let $x_3 = t$, where $t \in \mathbb{R}$. Then:

$$x_1 = \frac{5}{4}t, x_2 = \frac{67}{24}t, x_3 = t.$$

To eliminate fractions, let $t = 24s$, where $s \in \mathbb{R}$. Then:

$$x_1 = 30s, x_2 = 67s, x_3 = 24s.$$

Hence the general solution is

$$\boxed{(x_1, x_2, x_3) = s(30, 67, 24), s \in \mathbb{R}}$$

Exercise 1.4.8. Find a row-reduced echelon matrix which is row equivalent to

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

What are the solutions of $AX = 0$?

Solution 1.4.9. We begin with the matrix

$$A = \begin{bmatrix} 1 & -i \\ 2 & 2 \\ i & 1+i \end{bmatrix}$$

and aim to find a row-reduced echelon matrix row-equivalent to A. Then, we solve the homogeneous system $AX = 0$, where $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{C}^2$. We apply elementary row operations to transform A into row-reduced echelon form. The first row has a pivot in column 1. To eliminate entries below it:

- Row 2: $R_2 \to R_2 - 2R_1$:

$$(2, 2) - 2(1, -i) = (0, 2 + 2i) = (0, 2(i + 1)).$$

- Row 3: $R_3 \to R_3 - iR_1$:

$$(i, 1+i) - i(1, -i) = (0, i).$$

Now the matrix becomes

$$\begin{bmatrix} 1 & -i \\ 0 & 2(i+1) \\ 0 & i \end{bmatrix}.$$

We now focus on column 2. The first non-zero entry below row 1 is in row 2. To eliminate the entry in row 3:

- Row 3: $R_3 \to R_3 - \left(\frac{i}{2(i+1)}\right) R_2$:

$$i - \left(\frac{i}{2(i+1)}\right) \cdot 2(i+1) = i - i = 0.$$

The matrix becomes

$$\begin{bmatrix} 1 & -i \\ 0 & 2(i+1) \\ 0 & 0 \end{bmatrix}.$$

We now scale row 2 to make the pivot 1:

- Row 2: $R_2 \to \frac{1}{2(1+i)} R_2$:

$$(0, 2(i+1)) \to (0, 1).$$

Now the matrix is

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

Then we use row 2 to eliminate the entry above it in row 1:

- Row 1: $R_1 \to R_1 + iR_2$:

$$(1, -i) + i(0, 1) = (1, 0).$$

The final row-reduced echelon form is:

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

The system $AX = 0$ is equivalent to $RX = 0$, which gives:

- $1 \cdot x_1 + 0 \cdot x_2 = 0 \implies x_1 = 0$.
- $0 \cdot x_1 + 1 \cdot x_2 = 0 \implies x_2 = 0$.
- The third row is $0 - 0 = 0$, which is always true.

Thus the only solution is the trivial one: $x_1 = 0, x_2 = 0$.

Exercise 1.4.10. Describe explicitly all 2×2 row-reduced echelon matrices.

Solution 1.4.11.

1. Zero matrix
$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$
This trivially satisfies all conditions.

2. One non-zero row (top row). The second row must be 0. The top row must have a leading 1.

   (a) Leading 1 in column 1
   $$\begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix}.$$
   Here, a is any scalar. The leading 1 is in column 1, and the entry above it and the entry below it are zero.

(b) Leading 1 in column 2

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

The first entry must be 0 to ensure the leading 1 is the first non-zero entry.

3. Two non-zero rows. Both rows must have leading 1's, with the second row's pivot to the right of the first. The only possibility is the 2×2 identity matrix

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

Exercise 1.4.12. Consider the system of equations

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ 2x_1 + 2x_3 &= 1 \\ x_1 - 3x_2 + 4x_2 &= 2. \end{aligned}$$

Does this system have a solution? If so, describe explicitly all solutions.

Solution 1.4.13. We consider the system of equations

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ 2x_1 + 2x_3 &= 1 \\ x_1 - 3x_2 + 4x_2 &= 2. \end{aligned}$$

Our goal is to determine whether this system has a solution and, if so, describe all solutions explicitly. The augmented matrix representing the system is

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

We apply elementary row operations to transform the matrix into row-echelon form.

- $R_2 \longrightarrow R_2 - 2R_1$:
$$(2, 0, 2, 1) - 2(1, -1, 2, 1) = (0, 2, -2, -1).$$
- $R_3 \longrightarrow R_3 - R_1$:
$$(1, -3, 4, 2) - (1, -1, 2, 1) = (0, -2, 2, 1).$$

The matrix becomes:
$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & 2 & -2 & -1 \\ 0 & -2 & 2 & 1 \end{array}\right].$$

- $R_3 \longrightarrow R_3 + R_2$:
$$(0, -2, 2, 1) + (0, 2, -2, -1) = (0, 0, 0, 0).$$

The row-reduced form is:
$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & 2 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The third row $0 = 0$ is always true, so the system is consistent. The remaining equations are:

$$x_1 - x_2 + 2x_3 = 1 \tag{1.4}$$
$$2x_2 - 2x_3 = -1 \tag{1.5}$$

From equation 1.5:
$$2x_2 - 2x_3 = 1 \implies x_2 - x_3 = -\frac{1}{2}.$$

Let $x_3 = t$, where t is a free parameter. Then: $x_2 = t - \frac{1}{2}$. Substitute into equation 1.5:

$$x_1 - \left(t - \frac{1}{2}\right) + 2t = 1 \implies x_1 + t + \frac{1}{2} = 1 \implies x_1 = \frac{1}{2} - t.$$

The system has infinitely many solutions. All solutions are given by:

$$x_1 = \frac{1}{2} - t, \quad x_2 = -\frac{1}{2} + t, \quad x_3 = t$$

where t is any scalar (real or complex).

Exercise 1.4.14. Give an example of a system of two linear equations in two unknowns which has no solution.

Solution 1.4.15. Consider the following two equations:

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

These equations represent two parallel lines with the same slope (-1) but different y-intercepts. Since parallel lines never intersect, the system has no solution.

Exercise 1.4.16. Show that the system

$$\begin{aligned} x_1 - 2x_2 + x_3 + 2x_4 &= 1 \\ x_1 + x_2 - x_3 + x_4 &= 2 \\ x_1 + 7x_2 - 5x_3 - x_4 &= 3 \end{aligned}$$

has no solution.

Solution 1.4.17. The system of equations is

$$\begin{aligned} x_1 - 2x_2 + x_3 + 2x_4 &= 1 \\ x_1 + x_2 - x_3 + x_4 &= 2 \\ x_1 + 7x_2 - 5x_3 - x_4 &= 3. \end{aligned}$$

The augmented matrix is:

$$\left[\begin{array}{cccc|c} 1 & -2 & 1 & 2 & 1 \\ 1 & 1 & -1 & 1 & 2 \\ 1 & 7 & -5 & -1 & 3 \end{array}\right].$$

We eliminate the first column below the pivot:

- $R_2 \longrightarrow R_2 - R_1$:

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

- $R_3 \longrightarrow R_3 - R_1$:

$$(1, 7, -5, -1, 3) - (1, -2, 1, 2, 1) = (0, 9, -6, -3, 2).$$

The matrix becomes

$$\left[\begin{array}{cccc|c} 1 & -2 & 1 & 2 & 1 \\ 0 & 3 & -2 & -1 & 1 \\ 0 & 9 & -6 & -3 & 2 \end{array}\right].$$

We then eliminate the second column below the pivot:

- $R_3 \longrightarrow R_3 - 3R_2$:

$$(0, 9, -6, -3, 2) - 3(0, 3, -2, -1, 1) = (0, 0, 0, 0, -1).$$

The reduced matrix is

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

The third row corresponds to the equation:

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -1$$

which simplifies to

$$0 = -1.$$

This is a contradiction, meaning the system has no solution.

Exercise 1.4.18. Find all solutions of

$$\begin{aligned} 2x_1 - 3x_2 - 7x_3 + 5x_4 + 2x_5 &= -2 \\ x_1 - 2x_2 - 4x_3 + 3x_4 + x_5 &= -2 \\ 2x_1 - 4x_3 + 2x_4 + x_5 &= 3 \\ x_1 - 5x_2 - 7x_3 + 6x_4 + 2x_5 &= -7. \end{aligned}$$

Solution 1.4.19. The system of equations is

$$\begin{aligned} 2x_1 - 3x_2 - 7x_3 + 5x_4 + 2x_5 &= -2 \\ x_1 - 2x_2 - 4x_3 + 3x_4 + x_5 &= -2 \\ 2x_1 - 4x_3 + 2x_4 + x_5 &= 3 \\ x_1 - 5x_2 - 7x_3 + 6x_4 + 2x_5 &= -7. \end{aligned}$$

The augmented matrix is,

$$\left[\begin{array}{ccccc|c} 2 & -3 & -7 & 5 & 2 & -2 \\ 1 & -2 & -4 & 3 & 1 & -2 \\ 2 & 0 & -4 & 2 & 1 & 3 \\ 1 & -5 & -7 & 6 & 2 & -7 \end{array}\right].$$

To simplify, we can swap rows to have a leading 1 in the first row:

$$R_1 \iff R_2$$

$$\left[\begin{array}{ccccc|c} 1 & -2 & -4 & 3 & 1 & -2 \\ 2 & -3 & -7 & 5 & 2 & -2 \\ 2 & 0 & -4 & 2 & 1 & 3 \\ 1 & -5 & -7 & 6 & 2 & -7 \end{array}\right].$$

We then eliminate below pivot in column 1:

- $R_2 \to R_2 - 2R_1$.
- $R_3 \to R_3 - 2R_1$.

- $R_4 \longrightarrow R_4 - R_1$.

$$\left[\begin{array}{ccccc|c} 1 & -2 & -4 & 3 & 1 & -2 \\ 0 & 1 & 1 & -1 & 0 & 2 \\ 0 & 4 & 4 & -4 & -1 & 7 \\ 0 & -3 & -3 & 3 & 1 & -5 \end{array}\right].$$

We then eliminate below pivot in column 2 :

- $R_3 \longrightarrow R_3 - 4R_2$.
- $R_4 \longrightarrow R_4 + 3R_2$.

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

Simplify rows 3 and 4 :

- $R_4 \longrightarrow R_4 + R_3$.
- $R_3 \longrightarrow -R_3$.

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

Eliminate above pivot in column 5 :

- $R_1 \longrightarrow R_1 - R_3$.

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

The reduced system is,

$$\begin{aligned} x_1 - 2x_2 - 4x_3 + 3x_4 &= -3 \\ x_2 + x_3 - x_4 &= 2 \\ x_5 &= 1. \end{aligned}$$

Let $x_3 = s$ and $x_4 = t$ be free variables. From the second equation:

$$x_2 = 2 - s + t.$$

From the first equation:

$$x_1 = -3 + 2x_2 + 4x_3 - 3x_4 = -3 + 2(2 - s + t) + 4s - 3t = 1 + 2s - t.$$

Therefore, the general solution is

$$(1 + 2s - t, \quad 2 - s + t, \quad s, \quad t, \quad 1).$$

Exercise 1.4.20. Let

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

For which triples (y_1, y_2, y_3) does the system $AX = Y$ have a solution?

Solution 1.4.21. Given the matrix

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

and a vector $Y = (y_1, y_2, y_3)^T$, we want to determine for which triples (y_1, y_2, y_3) the system

$$AX = Y$$

has a solution.
A square matrix A is invertible if and only if its determinant is non-zero. If A is invertible, then for every Y, the system $AX = Y$ has a unique solution. Let us compute the determinant of A:

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

Using cofactor expansion along the first row:

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

$$= 3 \cdot (1 \cdot 0 - 1 \cdot (-3)) + 1 \cdot (2 \cdot 0 - 1 \cdot 1) + 2 \cdot (2 \cdot (-3) - 1 \cdot 1)$$

$$= 3 \cdot 3 + 1 \cdot (-1) + 2 \cdot (-7) = 9 - 1 - 14 = -6.$$

Since $\det(A) = -6 \neq 0$, the matrix A is invertible. Because A is invertible, the equation $AX = Y$ has a solution for every vector $Y \in \mathbb{R}^3$.

Exercise 1.4.22. Let

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

For which (y_1, y_2, y_3, y_4) does the system of equations $AX = Y$ have a solution?

Solution 1.4.23. Let

$$A = \begin{bmatrix} 3 & -6 & 2 & -1 \\ -2 & 4 & 1 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & -2 & 1 & 0 \end{bmatrix}$$

and perform row operations to reduce A to row-echelon form.

- Swap rows 1 and 4 :

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

- Add 2 times row 1 to row 2:

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

- Subtract 3 times row 1 from row 4:

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

- Divide row 2 by 3:

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

- Subtract row 2 from row 3, and add row 2 to row 4:

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

The system is $AX = Y$, where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}.$$

Now after applying all the row operations, the system is:

1. $x_1 - 2x_2 + x_3 = y_4$.
2. $3x_3 + 3x_4 = y_2 + 2y_4$.
3. $x_3 + x_4 = y_3$.
4. $-x_3 - x_4 = y_1 - 3y_4$.

From equations (2), (3) and (4):

- From (2): $x_3 + x_4 = \frac{y_2+2y_4}{3}$.
- From (3): $x_3 + x_4 = y_3$.
- From (4): $x_3 + x_4 = -y_1 + 3y_4$.

Equation these equations:

$$y_3 = \frac{y_2 + 2y_4}{3} \text{ and } y_3 = -y_1 + 3y_4.$$

Rewriting,

$$3y_3 = y_2 + 2y_4 \text{ and } y_1 + y_3 = 3y_4.$$

Therefore the system $AX = Y$ has a solution if and only if:

$$\boxed{(y_1, y_2, y_3, y_4) \text{ such that } y_1 + y_3 = 3y_4 \text{ and } y_2 + 2y_4 = 3y_3.}$$

Exercise 1.4.24. Suppose R and R' are 2×3 row-reduced echelon matrices and that the systems $RX = 0$ and $R'X = 0$ have exactly the same solutions. Prove that $R = R'$.

Solution 1.4.25. Let R and R' be 2×3 row-reduced echelon matrices. Assume that the homogeneous systems

$$RX = 0 \text{ and } R'X = 0$$

have exactly the same solution sets. We aim to prove that $R = R'$. For any matrix A, the null space is the set of all vectors X such that $AX = 0$. The row space of A is the span of its rows. The row space of a matrix is the orthogonal complement of its null space. That is,

$$\text{row}(A) = (\text{null}(A))^{\perp}.$$

In our case, since

$$\text{null}(R) = \text{null}(R'),$$

it follows that

$$\text{row}(R) = (\text{null}(R))^{\perp} = (\text{null}(R'))^{\perp} = \text{row}(R').$$

Thus the row spaces of R and R' are identical. A key property of row-reduced echelon form is that for any given subspace of $\mathbb{R}^3$, there is exactly one matrix in row-reduced echelon form whose rows form a basis for that subspace. Since both R and R' are in row-reduced echelon form and have the same row space, their rows must be the same. Therefore,

$$\boxed{R = R'.}$$

1.5 Matrix Multiplication

Suppose B is an $n \times p$ matrix over a field F with rows $\beta_1, \ldots, \beta_n$ and the from B we construct a matrix C with rows $\gamma_1, \ldots, \gamma_m$ by forming certain linear combinations

$$\gamma_i = A_{i1}\beta_1 + A_{i2}\beta_2 + \cdots + A_{in}\beta_n. \tag{1.6}$$

The rows of C are determined by the mn scalars A_{ij} which are themselves the entries of an $m \times n$ matrix A. If 1.6 is expanded to

$$(C_{i1} \cdots C_{ip}) = \sum_{r=1}^{n} (A_{ir}B_{r1} \cdots A_{ir}B_{rp})$$

we see that the entries of C are given by

$$C_{ij} = \sum_{r=1}^{n} A_{ir} B_{rj}.$$

Definition 1.5.1. Let A be an $m \times n$ matrix over the field F and let B be an $n \times p$ matrix over F. The product AB is the $m \times p$ matrix C whose i, j entry is

$$C_{ij} = \sum_{r=1}^{n} A_{ir} B_{rj}.$$

The product of two matrices need not be defined; the product is defined if and only if the number of columns in the first matrix coincides with the number of rows in the second matrix. Even when the products AB and BA are both defined it need not be true that $AB = BA$; in other words, matrix multiplication is not commutative.

Theorem 1.5.2. *If A, B, C are matrices over the field F such that the products BC and $A(BC)$ are defined, then so are the products AB, $(AB)C$ and*

$$A(BC) = (AB)C.$$

Proof. Suppose B is an $n \times p$ matrix. Since BC is defined, C is a matrix with p rows, and BC has n rows. Because $A(BC)$ is defined we may assume A is an $m \times n$ matrix. Thus the product AB exists and is an $m \times p$ matrix, from which it follows that the product $(AB)C$ exists. To show that $A(BC) = (AB)C$ means to show that

$$[A(BC)]_{ij} = [(AB)C]_{ij}$$

for each i, j. By definition

$$\begin{aligned}
[A(BC)]_{ij} &= \sum_r A_{ir}(BC)_{rj} \\
&= \sum_r A_{ir} \sum_s B_{rs}C_{sj} \\
&= \sum_r \sum_s A_{ir}B_{rs}C_{sj} \\
&= \sum_s \sum_r A_{ir}B_{rs}C_{sj} \\
&= \sum_s (\sum_r A_{ir}B_{rs})C_{sj} \\
&= \sum_s (AB)_{is}C_{sj} \\
&= [(AB)C)]_{ij}.
\end{aligned}$$

□

When A is an $n \times n$ matrix, the product AA is defined. We shall denote this matrix by A^2. In general, the product $AA \cdots A$ (k times) is unambiguously defined, and we shall denote this product by A^k. The relation $A(BC) = (AB)C$ implies among other things that linear combinations of linear combinations of the rows of C are again linear combinations of the rows of C.

Definition 1.5.3. An $m \times n$ matrix is said to be an elementary matrix if it can be obtained from the $m \times m$ identity matrix by means of a single elementary row operation.

Theorem 1.5.4. *Let e be an elementary row operation and let E be the $m \times m$ elementary matrix $E = e(I)$. Then, for every $m \times n$ matrix A,*

$$e(A) = EA.$$

Proof. The point of the proof is that the entry in the i th row and jth column of the product matrix EA is obtained from the ith row of E and the j th column of A. The three types of elementary row operations should be taken up separately. We shall give a detailed proof for an operation of type (ii). Suppose $r \neq s$ and e is the operation 'replacement of row r by row r plus c times row s.' Then

$$E_{ik} = \begin{cases} \delta_{ik}, & i \neq r \\ \delta_{rk} + c\delta_{sk}, & i = r. \end{cases}$$

Therefore,

$$(EA)_{ij} = \sum_{k=1}^{m} E_{ik} A_{kj} = \begin{cases} A_{ik}, & i \neq r \\ A_{rj} + cA_{sj}, & i = r. \end{cases}$$

In other words $EA = e(A)$.

□

Corollary 1.5.5. *Let A and B be $m \times n$ matrices over the field F. Then B is row-equivalent to A if and only if $B = PA$, where P is a product of $m \times m$ elementary matrices.*

Proof. Suppose $B = PA$ where $P = E_s \cdots E_2 E_1$ and the E_i are $m \times m$ elementary matrices. Then $E_1 A$ is row-equivalent to A, and $E_2(E_1 A)$ is row-equivalent to $E_1 A$. So $E_2 E_1 A$ is row-equivalent to A; and continuing in this way we see that $(E_s \cdots E_1)A$ is row-equivalent to A.

Now suppose that B is row-equivalent to A. Let $E_1, E_2, \ldots, E_s$ be the elementary matrices corresponding to some sequence of elementary row operations which carries A into B. Then $B = (E_s \cdots E_1)A$.

□

Exercises

Exercise 1.5.6. Let

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -1 \end{bmatrix}.$$

Compute ABC and CAB.

Solution 1.5.7.

$$AB = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 3 + (-1) \cdot 1 + 1 \cdot (-1) \\ 1 \cdot 3 + 2 \cdot 1 + 1 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 6 - 1 - 1 \\ 3 + 2 - 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}.$$

$$ABC = (AB)C = \begin{bmatrix} 4 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} = \begin{bmatrix} 4 \cdot 1 & 4 \cdot (-1) \\ 4 \cdot 1 & 4 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 4 & -4 \\ 4 & -4 \end{bmatrix}.$$

$$CA = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + (-1) \cdot 1 & 1 \cdot (-1) + (-1) \cdot 2 & 1 \cdot 1 + (-1) \cdot 1 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 0 \end{bmatrix}.$$

$$CAB = (CA)B = \begin{bmatrix} 1 & -3 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 3 + (-3) \cdot 1 + 0 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}.$$

Exercise 1.5.8. Let

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

Verify directly that $A(AB) = A^2B$.

Solution 1.5.9.

$$AB = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 1 & 3 \\ 4 & 4 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + (-1) \cdot 1 + 1 \cdot 4 & 1 \cdot (-2) + (-1) \cdot 3 + 1 \cdot 4 \\ 2 \cdot 2 + 0 \cdot 1 + 1 \cdot 4 & 2 \cdot (-2) + 0 \cdot 3 + 1 \cdot 4 \\ 3 \cdot 2 + 0 \cdot 1 + 1 \cdot 4 & 3 \cdot (-2) + 0 \cdot 3 + 1 \cdot 4 \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ 8 & 0 \\ 10 & -2 \end{bmatrix}.$$

$$A(AB) = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & -1 \\ 8 & 0 \\ 10 & -2 \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 + (-1) \cdot 8 + 1 \cdot 10 & 1 \cdot (-1) + (-1) \cdot 0 + 1 \cdot (-2) \\ 2 \cdot 5 + 0 \cdot 8 + 1 \cdot 10 & 2 \cdot (-1) + 0 \cdot 0 + 1 \cdot (-2) \\ 3 \cdot 5 + 0 \cdot 8 + 1 \cdot 10 & 3 \cdot (-1) + 0 \cdot 0 + 1 \cdot (-2) \end{bmatrix}.$$

$$A(AB) = \begin{bmatrix} 7 & -3 \\ 20 & -4 \\ 25 & -5 \end{bmatrix}.$$

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

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

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

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

$$= \begin{bmatrix} 2 \cdot 2 + (-1) \cdot 1 + 1 \cdot 4 & 2 \cdot (-2) + (-1) \cdot 3 + 1 \cdot 4 \\ 5 \cdot 2 + (-2) \cdot 1 + 3 \cdot 4 & 5 \cdot (-2) + (-2) \cdot 3 + 3 \cdot 4 \\ 6 \cdot 2 + (-3) \cdot 1 + 4 \cdot 4 & 6 \cdot (-2) + (-3) \cdot (-2) 4 \cdot 4 \end{bmatrix}$$

$$A^2B = \begin{bmatrix} 7 & -3 \\ 20 & -4 \\ 25 & -5 \end{bmatrix}.$$

Therefore,

$$\boxed{A^2B = A(AB)}$$

Exercise 1.5.10. Find two different 2×2 matrices A such that $A^2 = 0$ but $A \neq 0$.

Solution 1.5.11. We want non-zero 2×2 matrices A such that $A^2 = 0$. Such matrices are called **nilpotent matrices**. Two examples of nilpotent matrices are

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Exercise 1.5.12. For the matrix

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

find elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k \cdots E_2 E_1 A = I.$$

Solution 1.5.13. Let

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

We perform row operations to reduce A to the identity matrix I.

1. Eliminate below pivot in column 1.

   - $R_2 \to R_2 - 2R_1$.
   - $R_3 \to R_3 - 3R_1$.

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

2. Normalize pivot in row 2.

   - $R_2 \to \frac{1}{2}R_2$.

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

3. Eliminate above and below pivot in column 2.

   - $R_1 \to R_1 + R_2$.
   - $R_3 \to R_3 - 3R_2$.

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

4. Normalize pivot in row 3.

   - $R_3 \to -2R_3$.

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

5. Eliminate above pivot in column 3.

   - $R_1 \to R_1 - \frac{1}{2}R_3$.
   - $R_2 \to R_2 + \frac{1}{2}R_3$.

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

Each row operation corresponds to an elementary matrix.

Exercise 1.5.14. Let

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

Is there a matrix C such that $CA = B$?

Solution 1.5.15. The matrix A is 3×2 and the matrix B is 2×2. For the product CA to be defined, the number of columns in C must equal the number of rows in A, so C must have

3 columns. Since B is 2×2, the product CA must also be 2×2, so C must have 2 rows. Therefore, C is a 2×3 matrix.
Let
$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}.$$
Then compute CA :
$$\begin{bmatrix} c_{11} + 2c_{12} + c_{13} & -c_{11} + 2c_{12} \\ c_{21} + 2c_{22} + c_{23} & -c_{21} + 2c_{22} \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ -4 & 4 \end{bmatrix}.$$
This gives the system of equations:
From the first row:

1. $c_{11} + 2c_{12} + c_{13} = 3.$

2. $-c_{11} + 2c_{12} = 1.$

From the second row:

1. $c_{21} + 2c_{22} + c_{23} = -4.$

2. $-c_{21} + 2c_{22} = 4.$

For the first row, from (2):
$$c_{11} = 2c_{12} - 1.$$
Substitute into (1):
$$\begin{aligned} (2c_{12} - 1) + 2c_{12} + c_13 &= 3 \\ 4c_{12} - 1 + c_{13} &= 3 \\ c_{13} &= 4 - 4c_{12} \end{aligned}$$
So for any choice of c_{22}, we can find c_{21} and c_{23}. Since we can choose values for c_{12} and c_{22} freely and solve for the remaining variables, there exist infinitely many matrices C such that $CA = B$.

Exercise 1.5.16. Let A be an $m \times n$ matrix and B an $n \times k$ matrix. Show that the columns of $C = AB$ are linear combinations of A. If $\alpha_1, \ldots, \alpha_n$ are the columns of A and $\gamma_1, \ldots, \gamma_k$ are the columns of C, then

$$\gamma_j = \sum_{r=1}^{n} B_{rj} \alpha_r.$$

Solution 1.5.17. Let A be an $m \times n$ matrix with columns

$$\alpha_1, \alpha_2, \ldots, \alpha_n$$

so that

$$A = [\alpha_1 \mid \alpha_2 \mid \cdots \mid \alpha_n].$$

Let B be an $n \times k$ matrix with entries B_{rj}, where $r = 1, \ldots, n$ and $j = 1, \ldots, k$. Let $C = AB$ be an $m \times k$ matrix with columns

$$\gamma_1, \gamma_2, \ldots, \gamma_k.$$

The (i, j)-entry of C is given by:

$$C_{ij} = (AB)_{ij} = \sum_{r=1}^{n} A_{ir} B_{rj}.$$

The j th column of C, denoted by γ_j, has entries:

$$\gamma_j = \begin{bmatrix} C_{1j} \\ C_{2j} \\ \vdots \\ C_{mj} \end{bmatrix} = \begin{bmatrix} \sum_{r=1}^{n} A_{1r} B_{rj} \\ \sum_{r=1}^{n} A_{2r} B_{rj} \\ \vdots \\ \sum_{r=1}^{n} A_{mr} B_{rj} \end{bmatrix}.$$

This can be rewritten as:

$$\gamma_j = \sum_{r=1}^{n} B_{rj} \begin{bmatrix} A_{1r} \\ A_{2r} \\ \vdots \\ A_{mr} \end{bmatrix} = \sum_{r=1}^{n} B_{rj}\alpha_r.$$

Each column of γ_j of $C = AB$ is a linear combination of the columns of A, with the coefficients given by the entries of the j-th column of B :

$$\gamma_j = \sum_{r=1}^{n} B_{rj}\alpha_r.$$

Exercise 1.5.18. Let A and B be 2×2 matrices such that $AB = I$. Prove that $BA = I$.

Solution 1.5.19. Since A and B are 2×2 matrice and $AB = I$, we take determinants on both sides:

$$\det(AB) = \det(I).$$

Using the multiplicative property of determinants:

$$\det(A) \cdot \det(B) = 1.$$

This implies $\det(A) \neq 0$. Hence, A is invertible. Since A is invertible, there exists a matrix A^{-1} such that

$$A^{-1}A = AA^{-1} = I.$$

Starting from $AB = I$, multiply both sides by A^{-1} :

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

Simplifying:

$$(A^{-1}A)B = A^{-1} \implies IB = A^{-1} \implies B = A^{-1}.$$

Since $B = A^{-1}$, we have:

$$BA = A^{-1}A = I$$

as desired.

Exercise 1.5.20. Let

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

be a 2×2 matrix. We inquire when it is possible to find 2×2 matrices A and B such that $C = AB - BA$. Prove that such matrices can be found if and only if $C_{11} + C_{22} = 0$.

Solution 1.5.21. Let A and B be 2×2 matrices. Then

$$\operatorname{trace}(AB) = \operatorname{trace}(BA).$$

So,

$$\operatorname{trace}(C) = \operatorname{trace}(AB - BA) = \operatorname{trace}(AB) - \operatorname{trace}(BA) = 0.$$

In terms of entries, $C_{11} + C_{22} = 0$. We show that every traceless matrix is a commutator.

1. Case 1: C is invertible.

   Since $\operatorname{trace}(C) = 0$, the eigenvalues are λ and $-\lambda$ for some $\lambda \neq 0$. Then C is diagonizable: there exists an invertible matrix P such that

   $$P^{-1}CP = \begin{bmatrix} \lambda & 0 \\ 0 & -\lambda \end{bmatrix}.$$

   Now, define:

   $$A_0 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 0 & 0 \\ \lambda & 0 \end{bmatrix}.$$

   Then:

   $$A_0B_0 = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix}, \quad B_0A_0 = \begin{bmatrix} 0 & 0 \\ 0 & \lambda \end{bmatrix}.$$

So,

$$A_0B_0 - B_0A_0 = \begin{bmatrix} \lambda & 0 \\ 0 & -\lambda \end{bmatrix}.$$

Thus $P^{-1}CP$ is a commutator. Let:

$$A = PA_0P^{-1}, \quad B = PB_0P^{-1}.$$

Then:

$$AB - BA = P(A_0B_0 - B_0A_0)P^{-1} = P(P^{-1}CP)P^{-1} = C.$$

So, C is a commutator.

2. Case 2: C is not invertible.

   Since $\operatorname{trace}(C) = 0$, both eigenvalues are 0. Then C is either the zero matrix or nilpotent.

   - If $C = 0$, take $A = B \implies AB - BA = 0$.
   - If $C \neq 0$, then C is similar to a Jordan block:

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

   We can show that J is a commutator. For example, take:

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

   Then:

   $$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \implies AB - BA = J.$$

   So, J is a commutator. By similarity, C is also a commutator.

Thus, in all cases C is a commutator.

1.6 Invertible Matrices

Definition 1.6.1. Let A be an $n \times n$ (square) matrix over the field F. An $n \times n$ matrix B such that $BA = I$ is called a left inverse of A; an $n \times n$ matrix B such that $AB = I$ is called a right inverse of A. If $AB = BA = I$, then B is called a two-sided inverse of A and A is said to be invertible.

Lemma 1.6.2. *If A has a left inverse B and a right inverse C, then $B = C$.*

Proof. Suppose $BA = I$ and $AC = I$. Then

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

□

Thus if A has a left and a right inverse, A is invertible and has a unique two-sided inverse, which we shall denote by A^{-1} and simply call the inverse of A.

Theorem 1.6.3. *Let A and B be $n \times n$ matrices over F.*

1. *If A is invertible, so is A^{-1} and $(A^{-1})^{-1} = A$.*
2. *If both A and B are invertible, so is AB, and $(AB)^{-1} = B^{-1}A^{-1}$.*

Proof. The first statement is evident from the symmetry of the definition. The second follows upon verification of the relations

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

□

Corollary 1.6.4. *A product of invertible matrices is invertible.*

Theorem 1.6.5. *An elementary matrix is invertible.*

Proof. Let E be an elementary matrix corresponding to the elementary row operation e. If e_1 is the inverse operation of e (Theorem 1.3.1) and $E_1 = e_1(I)$, then

$$EE_1 = e(E_1) = e(e_1(I)) = I$$

and

$$E_1E = e_1(E) = e_1(e(I)) = I$$

so that E is invertible and $E_1 = E^{-1}$.

□

Theorem 1.6.6. *If A is an $n \times n$ matrix, the following are equivalent.*

1. *A is invertible.*
2. *A is row-equivalent to the $n \times n$ identity matrix.*
3. *A is a product of elementary matrices.*

Proof. Let R be a row-reduced echelon matrix which is row-equivalent to A. By Theorem 1.5.4,

$$R = E_k \cdots E_2E_1A$$

where $E_1, \ldots, E_k$ are elementary matrices. Each E_j is invertible, and so

$$A = E_1^{-1} \cdots E_k^{-1}R.$$

Since products of invertible matrices are invertible, we see that A is invertible if and only if R is invertible. Since R is a square row-reduced echelon matrix, R is invertible if and only if each row of R contains a non-zero entry, that is, if and only if $R = I$. We have now shown that A is invertible if and only if $R = I$, and if $R = I$ then $A = E_k^{-1} \cdots E_1^{-1}$. It should now be apparent that (1), (2) and (3) are equivalent statements about A.

□

Corollary 1.6.7. *If A is an invertible $n \times n$ matrix and if a sequence of elementary row operations reduces A to the identity, then that same sequence of operations when applied to I yields A^{-1}.*

Theorem 1.6.8. *For an $n \times n$ matrix A, the following are equivalent.*

1. *A is invertible.*
2. *The homogeneous system $AX = 0$ has only the trivial solution $X = 0$.*
3. *The system of equations $AX = Y$ has a solution X for each $n \times 1$ matrix Y.*

Proof. According to Theorem 1.4.5, condition (2) is equivalent to the fact that A is row-equivalent to the identity matrix. By Theorem 1.6.6, (1) and (2) are therefore equivalent. If A is invertible, the solution of $AX = Y$ is $X = A^{-1}Y$. Conversely, suppose $AX = Y$ has a solution for each given Y. Let R be a row-reduced echelon matrix which is row-equivalent to A. We wish to show that $R = I$. That amounts to showing that the last row of R is not (identically) 0. Let

$$E = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

If the system $RX = E$ can be solved for X, the last row of R cannot be 0. We know that $R = PA$, where P is invertible. Thus $RX = E$ if and only if $AX = P^{-1}E$. According to (3), the latter system has a solution.

□

Corollary 1.6.9. *A square matrix with either a left or right inverse is invertible.*

Proof. Let A be an $n \times n$ matrix. Suppose A has a left inverse, i.e., a matrix B such that $BA = I$. Then $AX = 0$ has only the trivial solution, because $X = IX = B(AX)$. Therefore

A is invertible. On the other hand, suppose A has a right inverse, i.e., a matrix C such that $AC = I$. Then C has a left inverse and is therefore invertible. It then follows that $A = C^{-1}$ and so A is invertible with inverse C.

□

Corollary 1.6.10. *Let $A = A_1 A_2 \cdots A_k$, where $A_1, \ldots, A_k$ are $n \times n$ (square) matrices. Then A is invertible if and only if each A_j is invertible.*

Proof. We have already shown that the product of two invertible matrices is invertible. From this one sees easily that if each A_j is invertible then A is invertible.
Suppose now that A is invertible. We first prove that A_k is invertible. Suppose X is an $n \times 1$ matrix and $A_k X = 0$. Then $AX = (A_1 \cdots A_{k-1})A_k X = 0$. Since A is invertible we must have $X = 0$. The system of equations $A_k X = 0$ thus has no non-trivial solution, so A_k is invertible. But now $A_1 \cdots A_{k-1} = AA_k^{-1}$ is invertible. By the preceeding argument, A_{k-1} is invertible. Continuing this way, we conclude that each A_j is invertible.

□

Suppose A is an $m \times n$ matrix and we wish to solve the system of equations $AX = Y$. If R is a row-reduced echelon matrix which is row-equivalent to A, then $R = PA$ where P is an $m \times m$ invertible matrix. The solutions of the system $AX = Y$ are exactly the same as the solutions of the system $RX = PY(= Z)$.

Exercises

Exercise 1.6.11. Let

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

Find a row-reduced echelon matrix R which is row-equivalent to A and an invertible 3×3 matrix P such that $R = PA$.

Solution 1.6.12. Let

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

We perform elementary row operations to reduce A to row-reduced echelon form R.

1. Eliminate entries below the first pivot.
   - $R_2 \longrightarrow R_2 + R_1$.
   - $R_3 \longrightarrow R_3 - R_1$.

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

2. Normalize the second row.
   - $R_2 \longrightarrow \frac{1}{2}R_2$.

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

3. Eliminate entries above and below the second pivot.
   - $R_1 \longrightarrow R_1 - 2R_2$
   - $R_3 \longrightarrow R_3 + 4R_2$.

$$A \longrightarrow \begin{bmatrix} 1 & 0 & -3 & -5 \\ 0 & 1 & 2 & \frac{5}{2} \\ 0 & 0 & 8 & 11 \end{bmatrix}.$$

4. Normalize the third row.
   - $R_3 \longrightarrow \frac{1}{8}R_3$.

$$A \longrightarrow \begin{bmatrix} 1 & 0 & -3 & -5 \\ 0 & 1 & 2 & \frac{5}{2} \\ 0 & 0 & 1 & \frac{11}{8} \end{bmatrix}.$$

5. Eliminate entries above the third pivot.
   - $R_1 \to R_1 + 3R_3$.
   - $R_2 \to R_2 - 2R_3$.

$$A \to \begin{bmatrix} 1 & 0 & 0 & -\frac{7}{8} \\ 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{11}{8} \end{bmatrix} = R.$$

We want an invertible 3×3 matrix P such that $R = PA$. We perform the same row operations on the identity matrix I_3. We start with:

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

1. $R_2 \to R_2 + R_1$.
$$\to \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2. $R_3 \to R_3 - R_1$.
$$\to \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

3. $R_2 \to \frac{1}{2}R_2$.
$$\to \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

4. $R_1 \to R_1 - 2R_2$.
$$\to \begin{bmatrix} 0 & -1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

5. $R_3 \longrightarrow R_3 + 4R_2$.
$$\longrightarrow \begin{bmatrix} 0 & -1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

6. $R_3 \longrightarrow \frac{1}{8}R_3$.
$$\longrightarrow \begin{bmatrix} 0 & -1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \end{bmatrix}.$$

7. $R_1 \longrightarrow R_1 + 3R_3$.
$$\longrightarrow \begin{bmatrix} \frac{3}{8} & -\frac{1}{4} & \frac{3}{8} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \end{bmatrix}.$$

8. $R_2 \longrightarrow R_2 - 2R_3$.
$$\longrightarrow \begin{bmatrix} \frac{3}{8} & -\frac{1}{4} & \frac{3}{8} \\ \frac{1}{4} & 0 & -\frac{1}{4} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \end{bmatrix}.$$

9. Multiply by 8 to clear denominators.
$$P = \frac{1}{8}\begin{bmatrix} 3 & -2 & 3 \\ 2 & 0 & -2 \\ 1 & 2 & 1 \end{bmatrix}.$$

Therefore,
$$\boxed{R = \begin{bmatrix} 1 & 0 & 0 & -\frac{7}{8} \\ 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{11}{8} \end{bmatrix}, \quad P = \frac{1}{8}\begin{bmatrix} 3 & -2 & 3 \\ 2 & 0 & -2 \\ 1 & 2 & 1 \end{bmatrix}.}$$

Exercise 1.6.13. Let
$$A = \begin{bmatrix} 2 & 0 & i \\ 1 & -3 & -i \\ i & 1 & 1 \end{bmatrix}.$$

Find a row-reduced echelon matrix R which is row-equivalent to A and and invertible 3×3 matrix P such that $R = PA$.

Solution 1.6.14. We begin with the matrix

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

and seek an invertible matrix P such that $PA = R$, where R is the row-reduced echelon form of A. To find P, we augment the identity matrix I_3 with A.

$$[I \mid A] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 0 & i \\ 0 & 1 & 0 & 1 & -3 & -i \\ 0 & 0 & 1 & i & 1 & 1 \end{array}\right].$$

Performing row operations on this augmented matrix will transform the left block into P and the right block into R.

1. Normalize the first pivot.

   - $R_1 \longrightarrow \frac{1}{2}R_1$

$$\longrightarrow \left[\begin{array}{ccc|ccc} \frac{1}{2} & 0 & 0 & 1 & 0 & \frac{i}{2} \\ 0 & 1 & 0 & 1 & -3 & -i \\ 0 & 0 & 1 & i & 1 & 1 \end{array}\right].$$

2. Eliminate below the first pivot.

   - $R_2 \longrightarrow R_2 - R_1$.
   - $R_3 \longrightarrow R_3 - iR_1$.

$$\longrightarrow \left[\begin{array}{ccc|ccc} \frac{1}{2} & 0 & 0 & 1 & 0 & \frac{i}{2} \\ -\frac{1}{2} & 1 & 0 & 0 & -3 & -\frac{3i}{2} \\ -\frac{i}{2} & 0 & 1 & 0 & 1 & \frac{3}{2} \end{array}\right].$$

3. Normalize the second pivot.

- $R_2 \longrightarrow -\frac{1}{3}R_2$.

$$\longrightarrow \left[\begin{array}{ccc|ccc} \frac{1}{2} & 0 & 0 & 1 & 0 & \frac{i}{2} \\ \frac{1}{6} & -\frac{1}{3} & 0 & 0 & 1 & \frac{i}{2} \\ -\frac{i}{2} & 0 & 1 & 0 & 1 & \frac{3}{2} \end{array}\right].$$

4. Eliminate below the second pivot.

- $R_3 \longrightarrow R_3 - R_2$.

$$\longrightarrow \left[\begin{array}{ccc|ccc} \frac{1}{2} & 0 & 0 & 1 & 0 & \frac{i}{2} \\ \frac{1}{6} & -\frac{1}{3} & 0 & 0 & 1 & \frac{i}{2} \\ -\frac{3i+1}{6} & \frac{1}{3} & 1 & 0 & 0 & \frac{3-i}{2} \end{array}\right]$$

5. Normalize the third pivot.

- $R_3 \longrightarrow \frac{2}{3-i}R_3 = \frac{3+i}{5}R_3$.

$$\longrightarrow \left[\begin{array}{ccc|ccc} \frac{1}{2} & 0 & 0 & 1 & 0 & \frac{i}{2} \\ \frac{1}{6} & -\frac{1}{3} & 0 & 0 & 1 & \frac{i}{2} \\ -\frac{i}{3} & \frac{3+i}{15} & \frac{3+i}{5} & 0 & 0 & 1 \end{array}\right]$$

6. Eliminate above the third pivot.

- $R_1 \longrightarrow R_1 - \frac{i}{2}R_3$.
- $R_2 \longrightarrow R_2 - \frac{i}{2}R_3$.

$$\longrightarrow \left[\begin{array}{ccc|ccc} \frac{1}{3} & \frac{1-3i}{30} & \frac{1-3i}{10} & 1 & 0 & 0 \\ 0 & -\frac{3+i}{10} & \frac{1-3i}{10} & 0 & 1 & 0 \\ -\frac{i}{3} & \frac{3+i}{15} & \frac{3+i}{5} & 0 & 0 & 1 \end{array}\right]$$

Therefore,

$$\boxed{R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} \frac{1}{3} & \frac{1-3i}{30} & \frac{1-3i}{10} \\ 0 & -\frac{3+i}{10} & \frac{1-3i}{10} \\ -\frac{i}{3} & \frac{3+i}{15} & \frac{3+i}{5} \end{bmatrix}.}$$

Exercise 1.6.15. For each of the two matrices

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

use elementary row operations to discover whether it is invertible, and to find the inverse in case it is.

Solution 1.6.16. Let

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

Let us form the augmented matrix $[A \mid I]$.

$$\left[\begin{array}{ccc|ccc} 2 & 5 & -1 & 1 & 0 & 0 \\ 4 & -1 & 2 & 0 & 1 & 0 \\ 6 & 4 & 1 & 0 & 0 & 1 \end{array}\right].$$

Perform row operations:

- $R_2 \to R_2 - 2R_1$.
- $R_3 \to R_3 - 3R_1$.

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

Now, $R_3 \to R_3 - R_2$.

$$\to \left[\begin{array}{ccc|ccc} 2 & 5 & -1 & 1 & 0 & 0 \\ 0 & -11 & 4 & -2 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{array}\right].$$

The left block has a row of zeros, therefore, A is not invertible.

Let

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

Let us form the augmented matrix $[B \mid I]$.

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

Perform row operations:

- $R_2 \longrightarrow R_2 - 3R_1$.

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

- Swap $R_2 \iff R_3$.

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

- $R_1 \longrightarrow R_1 + R_2$.
- $R_3 \longrightarrow R_3 - 5R_2$.

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 8 & -3 & 1 & -5 \end{array}\right].$$

- $R_3 \longrightarrow \frac{1}{8}R_3$.

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 1 & -\frac{3}{8} & \frac{1}{8} & -\frac{5}{8} \end{array}\right].$$

- $R_2 \longrightarrow R_2 + 2R_3$.

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -\frac{3}{4} & \frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & -\frac{3}{8} & \frac{1}{8} & -\frac{5}{8} \end{array}\right].$$

The right block is the inverse of B, therefore,

$$\boxed{B^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ -\frac{3}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{8} & \frac{1}{8} & -\frac{5}{8} \end{bmatrix}.}$$

Exercise 1.6.17. Let

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

For which X does there exist a scalar c such that $AX = cX$?

Solution 1.6.18. We are given the matrix

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

and seek all vectors X such that there exists a scalar c satisfying

$$AX = cX.$$

This is equivalent to

$$(A - cI)X = 0,$$

which means X is an eigenvector of A with eigenvalue c. Since A is a lower triangular, its eigenvalues are the diagonal entries. Thus, the only eigenvalue is

$$c = 5.$$

If $c \neq 5$, then $A - cI$ is invertible, and the only solution is $X = 0$. So for non-zero X, we must have $c = 5$.

We solve

$$(A - 5I)X = 0.$$

Compute

$$A - 5I = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Let $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Then

$$(A - 5I)X = \begin{bmatrix} 0 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This gives:

$$x_1 = 0, \quad x_2 = 0.$$

No condition is imposed on x_3, so it is free. The non-zero vectors X for which there exists a scalar c such that $AX = cX$ are exactly the multiples of

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

That is,

$$X = k \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \text{ for any scalar } k.$$

Exercise 1.6.19. Discover whether

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

is invertible, and find A^{-1} if it exists.

Solution 1.6.20. The matrix

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

is upper triangular. The determinant of an upper triangle matrix is the product of its diagonal entries:

$$\det(A) = 1 \cdot 2 \cdot 3 \cdot 4 = 24.$$

Since the determinant is non-zero, A is invertible. We seek a matrix $B = A^{-1}$ such that $AB = I$. Since A is upper triangular, its inverse B is also upper triangular. Let

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & b_{22} & b_{23} & b_{24} \\ 0 & 0 & b_{33} & b_{34} \\ 0 & 0 & 0 & b_{44} \end{bmatrix}.$$

From the equation $AB = I$, we derive the following system:

- From column 1:

$$
\begin{aligned}
1 \cdot b_{11} &= 1 \implies b_{11} = 1 \\
0 \cdot b_{11} + 2 \cdot 0 &= 0 \\
0 \cdot b_{11} + 0 \cdot 0 + 3 \cdot 0 &= 0 \\
0 \cdot b_{11} + 0 \cdot 0 + 0 \cdot 0 + 4 \cdot 0 &= 0
\end{aligned}
$$

- From column 2:

$$
\begin{aligned}
1 \cdot b_{12} + 2 \cdot b_{22} &= 0 \\
0 \cdot b_{12} + 2 \cdot b_{22} &= 1 \implies b_{22} = \frac{1}{2} \\
0 \cdot b_{12} + 0 \cdot b_{22} + 3 \cdot 0 &= 0 \\
0 \cdot b_{12} + 0 \cdot b_{22} + 0 \cdot 0 + 4 \cdot 0 &= 0 \\
b_{12} + 2 \cdot \frac{1}{2} &= 0 \implies b_{12} = -1
\end{aligned}
$$

- From column 3:

$$
\begin{aligned}
1 \cdot b_{13} + 2 \cdot b_{23} + 3 \cdot b_{33} &= 0 \\
0 \cdot b_{13} + 2 \cdot b_{23} + 3 \cdot b_{33} &= 0 \\
0 \cdot b_{13} + 0 \cdot b_{23} + 3 \cdot b_{33} &= 1 \implies b_{33} = \frac{1}{3} \\
0 \cdot b_{13} + 0 \cdot b_{23} + 0 \cdot b_{33} + 4 \cdot 0 &= 0 \\
2b_{23} + 3 \cdot \frac{1}{3} = 0 \implies 2b_{23} + 1 = 0 \implies b_{23} = -\frac{1}{2} \\
b_{13} + 2 \cdot \left(-\frac{1}{2}\right) + 3 \cdot \frac{1}{3} = 0 \implies b_{13} - 1 + 1 = 0 \implies b_{13} = 0
\end{aligned}
$$

- From column 4:

$$\begin{aligned}
1 \cdot b_{14} + 2 \cdot b_{24} + 3 \cdot b_{34} + 4 \cdot b_{44} &= 0 \\
0 \cdot b_{14} + 2 \cdot b_{24} + 3 \cdot b_{34} + 4 \cdot b_{44} &= 0 \\
0 \cdot b_{14} + 0 \cdot b_{24} + 3 \cdot b_{34} + 4 \cdot b_{44} &= 0 \\
0 \cdot b_{14} + 0 \cdot b_{24} + 0 \cdot b_{34} + 4 \cdot b_{44} &= 1 \implies b_{44} = \frac{1}{4}
\end{aligned}$$

$$3b_{34} + 4 \cdot \frac{1}{4} = 0 \implies 3b_{34} + 1 = 0 \implies b_{34} = -\frac{1}{3}$$

$$2b_{24} + 3 \cdot \left(-\frac{1}{3}\right) + 4 \cdot \frac{1}{4} = 0 \implies 2b_{24} - 1 + 1 = 0 \implies b_{24} = 0$$

$$b_{14} + 2 \cdot 0 + 3 \cdot \left(-\frac{1}{3}\right) + 4 \cdot \frac{1}{4} = 0 \implies b_{14} - 1 + 1 = 0 \implies b_{14} = 0$$

Therefore,

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

Exercise 1.6.21. Suppose A is a 2×1 matrix and that B is a 1×2 matrix. Prove that $C = AB$ is not invertible.

Solution 1.6.22. Let

$$A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

and

$$B = \begin{bmatrix} b_1 & b_2 \end{bmatrix}.$$

Their product $C = AB$ is a 2×2 matrix.

$$C = AB = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix} = \begin{bmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{bmatrix}.$$

Let us now find the determinant of C.

$$\det(C) = (a_1 b_1)(a_2 b_2) - (a_1 b_2)(a_2 b_1) = a_1 a_2 b_1 b_2 - a_1 a_2 b_1 b_2 = 0.$$

Since the determinant is zero, C is not invertible.

Exercise 1.6.23. Let A be an $n \times n$ (square) matrix. Prove the following two statements:

1. If A is invertible and $AB = 0$ for some $n \times n$ matrix B, then $B = 0$.
2. If A is not invertible, then there exists an $n \times n$ matrix B such that $AB = 0$ but $B \neq 0$.

Solution 1.6.24. (1) Assume A is invertible and $AB = 0$. Since A is invertible, there exists A^{-1} such that $AA^{-1} = I$. Multiply both sides of $AB = 0$ on the left by A^{-1}:

$$A^{-1}(AB) = A^{-1} \cdot 0.$$

Using associativity and the property of the identity matrix,

$$(A^{-1}A)B = 0 \implies IB = 0 \implies B = 0.$$

Therefore, if A is invertible, then $B = 0$.

(2) Assume A is not invertible. Then the columns of A are linearly dependent, so there exists a non-zero vector x such that

$$Ax = 0.$$

Now, define the $n \times n$ matrix B by taking x as every column:

$$B = [x \mid x \mid \cdots \mid x].$$

Then,

$$AB = A[x \mid x \mid \cdots \mid x] = [Ax \mid Ax \mid \cdots \mid Ax] = [0 \mid 0 \mid \cdots \mid 0] = 0.$$

Since $x \neq 0$, the matrix B is non-zero. Therefore, if A is not invertible, then there exists a non-zero matrix B such that $AB = 0$.

Exercise 1.6.25. Let
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$
Prove, using elementary row operations, that A is invertible if and only if $(ad - bc) \neq 0$.

Solution 1.6.26. Let
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$
We aim to prove using elementary row operations, that A is invertible if and only if $ad - bc \neq 0$. Elementary row operations preserve invertibility so we will reduce A to row-echelon form.

1. Case 1: $a \neq 0$

   - Replace R_2 with $R_2 - \frac{c}{a}R_1$:
   $$A \longrightarrow \begin{bmatrix} a & b \\ 0 & d - \frac{c}{a}b \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & \frac{ad-bc}{a} \end{bmatrix}$$
   - If $ad - bc \neq 0$, scale R_2 by $\frac{a}{ad-bc}$:
   $$\longrightarrow \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}.$$
   - Scale R_1 by $\frac{1}{a}$:
   $$\longrightarrow \begin{bmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{bmatrix}.$$
   - Replace R_1 with $R_1 - \frac{b}{a}R_2$:
   $$\longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

   Thus, if $ad - bc \neq 0$, A reduces to I, so A is invertible.

2. Case 2: $a = 0$

Then $ad - bc = -bc$. Consider two subcases:

- Subcase $a : c \neq 0$.
  - Swap R_1 and R_2 :
$$A = \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} \to \begin{bmatrix} c & d \\ 0 & b \end{bmatrix}.$$
  - If $b \neq 0$, then $ad - bc = -bc \neq 0$, and the matrix reduces to I as above.
  - If $b = 0$, then $ad - bc = 0$, and the matrix becomes
$$\begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix},$$
    which is not invertible.
- Subcase $b : c = 0$

  Then $ad - bc = 0$, and
$$A = \begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix},$$
  which has a zero row and therefore is not invertible.

Exercise 1.6.27. An $n \times n$ matrix A is called **upper-triangular** if $A_{ij} = 0$ for $i > j$, that is, if every entry below the main diagonal is 0. Prove that an upper-triangular (square) matrix is invertible if and only if every entry on its main diagonal is different from 0.

Solution 1.6.28. An $n \times n$ matrix A is upper-triangular if $A_{ij} = 0$ for all $i > j$. That is, all entries below the main diagonal are zero. We aim to prove that such a matrix A is invertible if and only if every entry on its main diagonal is non-zero.

($\Rightarrow$) If A is invertible, then all diagonal entries are non-zero.

Assume A is invertible. Then its row-reduced echelon form is the identity matrix I. During row-reduction, the diagonal entries become pivots. Suppose, for the sake of contradiction, that some diagonal entry $a_{ii} = 0$. Since A is upper-triangular, all entries in column i below row i are also zero. Thus, no row below row i can provide a non-zero pivot in column

i. Hence, column i lacks a pivot, and the row-reduced echelon form of A cannot be I, contradicting the invertibility of A. Therefore, all diagonal entries must be non-zero.

($\Leftarrow$) If all diagonal entries are non-zero, then A is invertible.

Assume $a_{ii} \neq 0$ for all i. We show that A can be reduced to I using elementary row operations:

1. Scale each row i by $\frac{1}{a_{ii}}$ to make the diagonal entry 1.
2. Eliminate above diagonal entries: Starting from the last row, use row i to eliminate entries above it in column i by subtracting suitable multiples of row i from rows above.

This process transforms A into the identity matrix I. Since A is row-equivalent to I, it is invertible.

Exercise 1.6.29. Prove that if A is an $m \times n$ matrix, and B is an $n \times m$ matrix and $n < m$, then AB is not invertible.

Solution 1.6.30. We are given the A is an $m \times n$ matrix, and B is an $n \times m$ matrix where $n < m$. Then AB is an $m \times m$ matrix. Since B is an $n \times m$ matrix with $n < m$, it has more columns than rows. By the rank-nullity theorem, the null space of B has dimension at least $m - n > 0$. Thus, there exists a non-zero vector $x \in \mathbb{R}^m$ such that

$$Bx = 0.$$

Now consider

$$(AB)x = A(Bx) = A0 = 0.$$

So, x is a non-zero vector in the null space of AB. Therefore, AB is not invertible.

Exercise 1.6.31. Let A be an $m \times n$ matrix. Show that by means of a finite number of elementary row and/or column operations, one can pass from A to a matrix R which is both 'row-reduced echelon' and 'column-reduced' echelon,' i.e., $R_{ij} = 0$ if $i \neq j$, $R_{ii} = 1$,

$1 \leq i \leq r$, $R_{ii} = 0$ if $i > r$. Show that $R = PAQ$, where P is an invertible $m \times m$ matrix and Q is an invertible $n \times n$ matrix.

Solution 1.6.32. We start with an $m \times n$ matrix A. The aim is to show that by applying a finite number of elementary row and column operations, we can transform A into a matrix R of the form

$$R = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

where I_r is the $r \times r$ identity matrix ($r \leq \min(m, n)$), and all other entries are zero. That is, $R_{ij} = 0$ for $i \neq j$, $R_{ii} = 1$ for $1 \leq i \leq r$, and $R_{ii} = 0$ for $i > r$. Furthermore, we want to show that there exists invertible matrices P and Q such that

$$R = PAQ.$$

Since elementary matrices are invertible, the product of such matrices are invertible. Thus, if we perform a sequence of row operations represented by P and column operations represented by Q, we obtain

$$R = PAQ$$

where P and Q are invertible.

We describe and inductive procedure to transform A into R.
Base Case: If $A = 0$, then $R = 0$, and we are done.
Inductive Step: Assume $A \neq 0$.

1. Find a non-zero entry: Locate a non-zero element a_{ij}. Using row and column swaps, bring it to the $(1, 1)$ position.

2. Normalize the pivot: Divide the first row by a_{11} so that the $(1, 1)$ entry becomes 1.

3. Clear column 1 : Use row operations to make all other entries in column 1 zero.

4. Clear row 1 : Use column operations to make all other entries in row 1 zero.

After these steps, the matrix becomes:

$$\begin{bmatrix} 1 & 0 \\ 0 & A' \end{bmatrix}$$

where A' is an $(m-1) \times (n-1)$ matrix.
Step 5: Apply induction: Apply the process on A'. Eventually, we obtain a matrix R of the desired form.

Each elementary operation corresponds to multiplication by an elementary matrix. Let:

- $P_1, P_2, \dots, P_k$ be elementary matrices for row operations.
- $Q_1, Q_2, \dots, Q_l$ be the elementary matrices for column operations.

Then,

$$P = P_k \cdots P_2 P_1, \quad Q = Q_l \cdots Q_2 Q_1$$

are invertible and

$$R = PAQ.$$

Chapter 2

Vector Spaces

2.1 Vector Spaces

Definition 2.1.1. A **vector space** (or linear space) consists of the following:

1. a field F of scalars;
2. a set V of objects called vectors;
3. a rule (or operation) called vector addition, which associates with each pair of vectors α, β in V a vector $\alpha + \beta$ in V, called the sum of α and β, in such a way that
   (a) addition is commutative, $\alpha + \beta = \beta + \alpha$;
   (b) addition is associative, $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$;
   (c) there is a unique vector 0 in V, called the zero vector, such that $\alpha + 0 = \alpha$ for all α in V;
   (d) for each vector α in V there is a unique vector $-\alpha$ in V such that $\alpha + (-\alpha) = 0$;
4. a rule (or operation), called vector multiplication, which associates with each scalar c in F and vector α in V a vector $c\alpha$ in V, called the product of c and α, in such a way that
   (a) $1\alpha = \alpha$ for every α in V ;
   (b) $(c_1 c_2)\alpha = c_1(c_2\alpha)$;
   (c) $c(\alpha + \beta) = c\alpha + c\beta$;
   (d) $(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha$.

A vector space is a composite object consisting of a field, a set of vectors, and two operations with certain special properties. The same set of vectors may be part of a number of distinct vector spaces. We may simply refer to the vector space as V, or when it is desirable ot specify the field, we shall say V is a **vector space over the field** F .

Example 2.1.2. The n-tuple space, F^n. Let F be any field, and let V be the set of all n-tuples $\alpha = (x_1, x_2, \ldots, x_n)$ of scalars x_i in F. If $\beta = (y_1, y_2, \ldots, y_n)$ with y_i in F, the sum of α and β is defined by

$$\alpha + \beta = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

The product of a scalar c and a vector α is defined by

$$c\alpha = (cx_1, cx_2, \ldots, cx_n).$$

Definition 2.1.3. A vector $\beta \in V$ is said to be a **linear combination** of the vectors $\alpha_1, \ldots, \alpha_n \in V$ provided there exists scalars $c_1, \ldots, c_n \in F$ such that

$$\begin{aligned} \beta &= c_1\alpha_1 + \cdots + c_n\alpha_n \\ &= \sum_{i=1}^{n} c_i\alpha_i. \end{aligned}$$

Exercises

Exercise 2.1.4. If F is a field, verify that F^n is a vector space over the field F.

Solution 2.1.5. Let F be a field. Then F^n is the set of all ordered n-tuples from F :

$$F^n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in F \text{ for } i = 1, 2, \ldots, n\}.$$

We define two operations,

- Vector addition:

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n).$$

- Scalar multiplication for $c \in F$.

$$c \cdot (a_1, \ldots, a_n) = (ca_1, \ldots, ca_n).$$

We now verify that F^n satisfies all the axioms of a vector space over F.

1. Closure under addition.

   If $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots v_n)$ are in F^n, then

   $$u + v = (u_1 + v_1, \ldots, u_n + v_n).$$

   Since F is closed under addition, each $u_i + v_i \in F$, so $u + v \in F^n$.

2. Associativity of addition.

   For $u, v, w \in F^n$:

   $$(u + v) + w = u + (v + w).$$

   This follows from the associativity of addition in F.

3. Commutativity of addition.

   For $u, v \in F^n$,

   $$u + v = v + u.$$

   This follows from the commutativity of addition in F.

4. Existence of an additive identity.

   The zero vector is

   $$0 = (0, 0, \ldots, 0).$$

   For any $u \in F^n$,

   $$u + 0 = (u_1 + 0, u_2 + 0, \ldots, u_n + 0) = (u_1, u_2, \ldots, u_n) = u.$$

5. Existence of additive inverses.

   For any $u = (u_1, u_2, \ldots, u_n) \in F^n$, define

   $$-u = (-u_1, -u_2, \ldots, -u_n).$$

Then,

$$u + (-u) = (u_1 + (-u_1), u_2 + (-u_2), \dots, u_n + (-u_n)) = (0, 0, \dots, 0) = 0.$$

6. Closure under scalar multiplication.

   For $c \in F$ and $u \in F^n$,

   $$c \cdot u = (cu_1, cu_2, \dots, cu_n).$$

   Since F is closed under multiplication, each $c \cdot u_i \in F$, so $c \cdot u \in F^n$.

7. Distributivity of scalar multiplication over vector addition.

   For $c \in F$ and $u, v \in F^n$,

   $$c \cdot (u + v) = c \cdot (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$$

   $$= (c(u_1+v_1), c(u_2+v_2), \dots, c(u_n+v_n)) = (cu_1+cv_1, cu_2+cv_2, \dots, cu_n+cv_n) = c\cdot u+c\cdot v.$$

8. Distributivity of scalar multiplication over field addition.

   For $c, d \in F$ and $u \in F^n$,

   $$(c + d) \cdot u = ((c + d)u_1, (c + d)u_2, \dots, (c + d)u_n)$$

   $$= (cu_1 + du_1, cu_2 + du_2, \dots, cu_n + du_n) = c \cdot u + c \cdot d.$$

9. Associativity of scalar multiplication.

   For $c, d \in F$ and $u \in F^n$,

   $$c \cdot (d \cdot u) = c \cdot (du_1, du_2, \dots, du_n)$$

   $$= (c(du_1), c(du_2), \dots, c(du_n)) = ((cd)u_1, (cd)u_2, \dots, (cd)u_n)$$

   $$= (cd) \cdot u.$$

10. Identity element for scalar multiplication.

    For $u \in F^n$,

    $$1 \cdot u = (1u_1, 1u_2, \dots, 1u_n) = (u_1, u_2, \dots, u_n) = u$$

where 1 is the multiplicative identity in F.

Since all the vector space axioms are satisfied, F^n is a vector space over the field F.

Exercise 2.1.6. If V is a vector space over the field F, verify that

$$(\alpha_1 + \alpha_2) + (\alpha_3 + \alpha_4) = [\alpha_2 + (\alpha_3 + \alpha_1)] + \alpha_4$$

for vectors $\alpha_1, \alpha_2, \alpha_3$ and α_4 in V.

Solution 2.1.7. In a vector space V over a field F, vector addition satisfies:

- Commutativity: $\alpha + \beta = \beta + \alpha$.
- Associativity: $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

These properties allow us to rearrange and regroup vectors in sums. We want to verify that

$$(\alpha_1 + \alpha_2) + (\alpha_3 + \alpha_4) = [\alpha_2 + (\alpha_3 + \alpha_1)] + \alpha_4.$$

Let us start with the right hand side:

$$[\alpha_2 + (\alpha_3 + \alpha_1)] + \alpha_4.$$

1. Apply commutativity inside the brackets.

   Since $\alpha_3 + \alpha_1 = \alpha_1 + \alpha_3$, we have

   $$\alpha_2 + (\alpha_3 + \alpha_1) = \alpha_2 + (\alpha_1 + \alpha_3).$$

2. Apply associativity.

   $$\alpha_2 + (\alpha_1 + \alpha_3) = (\alpha_2 + \alpha_1) + \alpha_3.$$

   So the expression becomes,

   $$[(\alpha_2 + \alpha_1) + \alpha_3] + \alpha_4.$$

3. Apply associativity again,

$$[(\alpha_2 + \alpha_1) + \alpha_3] + \alpha_4 = (\alpha_2 + \alpha_1) + (\alpha_3 + \alpha_4).$$

4. Apply commutativity,

$$(\alpha_2 + \alpha_1) + (\alpha_3 + \alpha_4) = (\alpha_1 + \alpha_2) + (\alpha_3 + \alpha_4).$$

We have shown that

$$[\alpha_2 + (\alpha_3 + \alpha_1)] + \alpha_4 = (\alpha_1 + \alpha_2) + (\alpha_3 + \alpha_4)$$

as desired.

Exercise 2.1.8. If $\mathbb{C}$ is the field of complex numbers, which vectors in $\mathbb{C}^3$ are linear combinations of $(1, 0, -1)$, $(0, 1, 1)$ and $(1, 1, 1)$?

Solution 2.1.9. Let the vectors be,

$$v_1 = (1, 0, -1),$$

$$v_2 = (0, 1, 1),$$

$$v_3 = (1, 1, 1).$$

We want to determine which vectors in $\mathbb{C}^3$ can be written as linear combinations of these three vectors. To span, the vectors must be linearly independent. Form the matrix with these vectors as rows:

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

Compute the determinant:

$$\det(A) = 1 \cdot \det \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - 0 \cdot \det \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} + (-1) \cdot \det \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$= 1 \cdot (1 \cdot 1 - 1 \cdot 1) - 0 + (-1) \cdot (0 \cdot 1 - 1 \cdot 1) = 0 + (-1)(-1) = 1.$$

Since $\det(A) \neq 0$, the vectors are linearly independent. Three linearly independent form a basis for the space. Therefore, any vector in $\mathbb{C}^3$ can be expressed as linear combinations of v_1, v_2 and v_3.

Exercise 2.1.10. Let V be the set of all pairs (x, y) of real numbers, and let F be the field of real numbers. Define

$$(x, y) + (x_1, y_1) = (x + x_1, y + y_1)$$

$$c(x, y) = (cx, y).$$

Is V, with these operations, a vector space over the field of real numbers?

Solution 2.1.11. The set V consists of all pairs (x, y) of real numbers. The operations are defined as:

$$(x, y) + (x_1, y_1) = (x + x_1, y + y_1)$$

$$c(x, y) = (cx, y).$$

We verify whether V satisfies the axioms of a vector space over $\mathbb{R}$.

Addition Axioms:

1. Closure under addition:

   If $(x, y), (x_1, y_1) \in V$, then $(x + x_1, y + y_1) \in V$.

2. Associativity of addition:

   $$[(x, y) + (x_1, y_1)] + (x_2, y_2) = (x + x_1 + x_2, y + y_1 + y_2) = (x, y) + [(x_1, y_1) + (x_2, y_2)].$$

3. Commutativity of addition:

$$(x, y) + (x_1, y_1) = (x + x_1, y + y_1) = (x_1 + x, y_1 + x) = (x_1, y_1) + (x, y).$$

4. Additive identity:

   The zero vector is $(0, 0)$, since

$$(x, y) + (0, 0) = (x + 0, y + 0) = (x, y).$$

5. Additive inverses:

   For any $(x, y) \in V$, the inverse is $(-x, -y)$ since

$$(x, y) + (-x, -y) = (x + (-x), y + (-x)) = (0, 0).$$

Scalar Multiplication Axioms:

1. Closure under scalar multiplication:

   For $c \in \mathbb{R}$ and $(x, y) \in V$,

$$c(x, y) = (cx, y) \in V.$$

2. Identity element:

$$1 \cdot (x, y) = (1x, y) = (x, y).$$

3. Distributivity of scalar multiplication over vector addition:

$$c[(x, y) + (x_1, y_1)] = c(x + x_1, y + y_1) = (c(x + x_1), y + y_1) = (cx + cx_1, y + y_1).$$

$$c(x, y) + c(x_1, y_1) = (cx, y) + (cx_1, y_1) = (cx + cx_1, y + y_1).$$

4. Distributivity of scalar multiplication over field addition:

   Let $c, d \in \mathbb{R}$. Then:

$$(c + d)(x, y) = ((c + d)x, y) = (cx + dx, y).$$

$$c(x, y) + d(x, y) = (cx, y) + (dx, y) = (cx + dx, y + y) = (cx + dx, 2y).$$

These are equal only if $y = 0$.

The axiom of distributivity of scalar multiplication over field addition fails for any $y \neq 0$. Therefore, V is not a vector space over $\mathbb{R}$.

Exercise 2.1.12. On $\mathbb{R}^n$, define two operations

$$\alpha \oplus \beta = \alpha - \beta$$

$$c \cdot \alpha = -c\alpha.$$

The operations on the right are the usual ones. Which of the axioms for a vector space are satisfied by $(\mathbb{R}^n, \oplus, \cdot)$?

Solution 2.1.13. Let $V = \mathbb{R}^n$ and define two operations:

$$\alpha \oplus \beta = \alpha - \beta$$

$$c \cdot \alpha = -c\alpha.$$

We now check which axioms of a vector space are satisfied by $(V, \oplus, \cdot)$.
Axiom 1: Closure under addition.
For any $\alpha, \beta \in V$,

$$\alpha \oplus \beta = \alpha - \beta \in V,$$

which is satisfied.
Axiom 2: Associativity of addition.

$$(\alpha \oplus \beta) \oplus \gamma = (\alpha - \beta) - \gamma = \alpha - \beta - \gamma.$$

$$\alpha \oplus (\beta \oplus \gamma) = \alpha - (\beta - \gamma) = \alpha - \beta + \gamma.$$

These are not equal in general and therefore is not satisfied.
Axiom 3: Commutavivity of addition.

$$\alpha \oplus \beta = \alpha - \beta.$$

$$\beta \oplus \alpha = \beta - \alpha.$$

These are not equal in general and therefore is not satisfied.
Axiom 4: Existence of an additive identity.
Suppose there exists $0 \in V$ such that

$$\alpha \oplus 0 = \alpha \text{ and } 0 \oplus \alpha = \alpha.$$

Then,

$$\alpha - 0 = \alpha \implies 0 = 0$$

but,

$$0 \oplus \alpha = 0 - \alpha = -\alpha \neq \alpha.$$

Therefore, it is not satisfied.
Axiom 5: Existence of additive inverses.
If an additive identity 0 existed, then for each α, there should exist $-\alpha$ such that

$$\alpha \oplus (-\alpha) = 0.$$

But since no additive identity exists, this axiom fails.
Axiom 6: Closure under scalar multiplication.
For any $c \in \mathbb{R}, \alpha \in V$,

$$c \cdot \alpha = -c\alpha \in V$$

which is satisfied.
Axiom 7: Distributivity of scalar multiplication over vector addition.

$$c \cdot (\alpha \oplus \beta) = c \cdot (\alpha - \beta) = -c(\alpha - \beta) = -c\alpha + c\beta.$$

$$c \cdot \alpha \oplus c \cdot \beta = (-c\alpha) \oplus (-c\beta) = -c\alpha + c\beta.$$

Therefore it is satisfied.
Axiom 8: Distributivity of scalar multiplication over field addition.

$$(c + d) \cdot \alpha = -(c + d)\alpha = -c\alpha - d\alpha.$$

$$c\alpha \oplus d\alpha = (-c\alpha) \oplus (-d\alpha) = -c\alpha + d\alpha.$$

These are not equal in general and therefore is not satisfied.
Axiom 9: Commutativity of scalar multiplication.

$$(cd) \cdot \alpha = -(cd)\alpha = -cd\alpha.$$

$$c \cdot (d \cdot \alpha) = c \cdot (-d\alpha) = -c(-d\alpha) = cd\alpha.$$

These are not equal in general and therefore is not satisfied.
Axiom 10: Identity element of scalar multiplication.

$$1 \cdot \alpha = -1\alpha = -\alpha \neq \alpha.$$

Therefore this is not satisfied.

Exercise 2.1.14. Let V be the set of all complex-valued functions f on the real line such that (for all t in $\mathbb{R}$)

$$f(-t) = \overline{f(t)}.$$

The bar denotes complex conjugation. Show that V, with the operations

$$(f + g)(t) = f(t) + g(t)$$

$$(cf)(t) = cf(t)$$

is a vector space over the field of real numbers. Give an example of a function in V which is not real-valued.

Solution 2.1.15. Let V be the set of all complex valued functions $f : \mathbb{R} \to \mathbb{C}$ such that for all $t \in \mathbb{R}$,

$$f(-t) = \overline{f(t)}.$$

We define addition and scalar multiplication pointwise:

- $(f + g)(t) = f(t) + g(t)$.

- $(cf)(t) = cf(t)$.

We check the vector space axioms.

1. Closure under addition.

   If $f, g \in V$, then for all $t \in \mathbb{R}$:

   $$(f+g)(-t) = f(-t) + g(-t) = \overline{f(t)} + \overline{g(t)} = \overline{f(t) + g(t)} = \overline{(f+g)(t)}.$$

   So $f + g \in V$.

2. Closure under scalar multiplication.

   If $f \in V$ and $c \in \mathbb{R}$,then:

   $$(cf)(-t) = cf(-t) = c\overline{f(t)} = \overline{cf(t)} = \overline{(cf)(t)}.$$

   So $cf \in V$.

3. Zero vector.

   The zero function $f(t) = 0$ satisfies:

   $$f(-t) = 0 = \overline{0} = \overline{f(t)},$$

   so $0 \in V$.

4. Additive inverses.

   If $f \in V$, define $-f$ by $(-f)(t) = -f(t)$. Then:

   $$(-f)(-t) = -f(-t) = -\overline{f(t)} = \overline{-f(t)} = \overline{(-f)(t)},$$

   so $-f \in V$.

The remaining vector space axioms follow from the pointwise definitions and the fact that $\mathbb{C}$ is a field. Since the scalars are real, and the operations are defined pointwise, V is a vector space over $\mathbb{R}$.

Now we want a function $f \in V$ such that $f(t) \notin \mathbb{R}$ for some t. Consider,

$$f(t) = it.$$

$$f(-t) = i(-t) = -it, \quad \overline{f(t)} = \overline{it} = -it.$$

So, $f(-t) = \overline{f(t)}$, hence $f \in V$.

Exercise 2.1.16. Let V be the set of pairs (x, y) of real numbers and let F be the field of real numbers. Define

$$(x, y) + (x_1, y_1) = (x + x_1, 0)$$

$$c(x, y) = (cx, 0).$$

Is V, with these operations, a vector space.

Solution 2.1.17. The set V of pairs of real numbers with the given operations is not a vector space over the field of real numbers because it fails to satisfy the axioms for an additive identity and for scalar multiplication by 1.

Additive identity axiom: There must exist a vector $0 \in V$ such that for any $v = (x, y) \in V$, $v + 0 = v$. Suppose $0 = (a, b)$. Then $v + 0 = (x + a, 0)$. For this to equal (x, y), we must have $x + a = x$ and $0 = y$. The first equation implies $a = 0$, but the second equation requires $y = 0$ for all v, which is not true since y can be any real number. Thus, no additive identity exists.

Scalar multiplication by 1 axiom: For any $v = (x, y) \in V$, $1 \cdot v = v$. However, $1 \cdot (x, y) = (1 \cdot x, 0) = (x, 0)$, which is not equal to (x, y) unless $y = 0$. Thus, the axiom fails.

Since these axioms are violated, V with the given operations is not a vector space.

2.2 Subspaces

Definition 2.2.1. Let V be a vector space over the field F. A **subspace** of V is a subset W of V which is itself a vector space over F with the operations of vector addition and scalar multiplication on V.

A direct check of the axioms for a vector space shows that the subset W of V is a subspace if for each α and β in W the vector $\alpha + \beta$ is again in W ; the 0 vector is in W ; for each α in W the vector $(-\alpha)$ is in W ; for each α in W and each scalar c the vector $c\alpha$ is in W.

Theorem 2.2.2. *A non-empty subset W of V is a subspace of V if and only if for each pair of vectors α, β in W and each scalar c in F the vector $c\alpha + \beta$ is again in W.*

Proof. Suppose that W is a non-empty subset of V such that $c\alpha + \beta$ belongs to W for all vectors α, β in W and all scalars c in F. Since W is nonempty, there is a vector ρ in W, and hence $(-1)\rho + \rho = 0$ is in W. Then if α is any vector in W and c any scalar, the vector $c\alpha = c\alpha + 0$ is in W. In particular, $(-1)\alpha = -\alpha$ is in W. Finally, if $\alpha + \beta$ are in W, then $\alpha + \beta = 1\alpha + \beta$ is in W. Thus W is a subspace of V.
Conversely, if W is a subspace of V, α and β are in W, and c is a scalar, certainly $c\alpha + \beta$ is in W.

$\square$

Some people prefer to use the $c\alpha + \beta$ property in Theorem 2.2.2 as the definition of a subspace. The important point is that, if W is a non-empty subset of V such that $c\alpha + \beta$ is in V for all α, β in W and all c in F, then (with the operations inherited from V) W is a vector space.

Example 2.2.3.

1. If V is any vector space, V is a subspace of V ; the subset consisting of the zero vector alone is a subspace of V, called the **the zero subspace** of V.
2. In F^n, the set of n-tuples $(x_1, \ldots, x_n)$ with $x_1 = 0$ is a subspace; however, the set of n-tuples with $x_1 = 1 + x_2$ is not a subspace $(n \geq 2)$.
3. The space of polynomial functions over the field F is a subspace of the space of all functions from F into F.

4. An $n \times n$ (square) matrix A over the field F is **symmetric** if $A_{ij} = A_{ji}$ for each i and j. The symmetric matrices form a subspace of the space of all $n \times n$ matrices over F.

5. An $n \times n$ (square) matrix A over the field $\mathbb{C}$ of complex numbers is **Hermitian** (or **self-adjoint**) if
$$A_{jk} = \overline{A_{kj}}$$
for each j, k, the bar denoting the complex conjugation.

Lemma 2.2.4. *If A is an $m \times n$ matrix over F and B, C are $n \times p$ matrices over F then*

$$A(dB + C) = d(AB) + AC \tag{2.1}$$

for each scalar d in F.

Proof.

$$\begin{aligned}
[A(dB + C)]_{ij} &= \sum_k A_{ik}(dB + C)_{kj} \\
&= \sum_k (dA_{ik}B_{kj} + A_{ik}C_{kj}) \\
&= d\sum_k A_{ik}B_{kj} + \sum_k A_{ik}C_{kj} \\
&= d(AB)_{ij} + (AC)_{ij} \\
&= [d(AB) + AC]_{ij}
\end{aligned}$$

□

Theorem 2.2.5. *Let V be a vector space over the field F. The intersection of any collection of subspaces of V is a subspace of V.*

Proof. Let $\{W_a\}$ be a collection of subspaces of V, and let $W = \bigcap_a W_a$ be their intersection. Recall that W is defined as the set of all elements belonging to every W_a. Since each W_a is a subspace, each contains the zero vector. Thus the zero vector is in the intersection W, and W is non-empty. Let α and β be vectors in W and let c be a scalar. By definition of W, both α and β belong to each W_a, and because each W_a is a subspace, the vector $(c\alpha + \beta)$ is in every W_a. Thus $(c\alpha + \beta)$ is again in W. By Theorem 2.2.2, W is a subspace of V.

□

From Theorem 2.2.5 it follows that if S is any collection of vectors in V, then there is a smallest subspace of V which contains S, that is, a subspace which contains S and which is contained in every other subspace containing S.

Definition 2.2.6. Let S be a set of vectors in a vector space V. The **subspace spanned** by S is defined to be the intersection W of all subspaces of V which contain S. When S is a finite set of vectors, $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, we shall simply call W the **subspace spanned by the vectors** $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Theorem 2.2.7. *The subspace spanned by a non-empty subset S of a vector space V is the set of all linear combinations of vectors in S.*

Proof. Let W be the subspace spanned by S. Then each linear combination

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_m\alpha_m$$

of vectors $\alpha_1, \alpha_2, \ldots, \alpha_m$ in S is clearly in W. Thus W contains the set L of all linear combinations of vectors in S. The set L, on the other hand, contains S and is non-empty. If α, β belong to L then α is a linear combination,

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_m\alpha_m$$

of vectors α_i in S, and β is a linear combination,

$$\beta = y_1\beta_1 + y_2\beta_2 + \cdots + y_n\beta_n$$

of vectors β_j in S. For each scalar c,

$$c\alpha + \beta = \sum_{i=1}^{m}(cx_i)\alpha_i + \sum_{j=1}^{n} y_i\beta_j.$$

Hence $c\alpha + \beta$ belongs to L. Thus L is a subspace of V.
Now we have shown that L is a subspace of V which contains S, and also that any subspace which contains S contains L. It follows that L is the intersection of all subspaces containing S, i.e., that L is the subspace spanned by the set S.

□

Definition 2.2.8. If $S_1, S_2, \ldots, S_k$ are subsets of a vector space V, the set of all sums

$$\alpha_1 + \alpha_2 + \cdots + \alpha_k$$

of vectors α_i in S_i is called the **sum** of the subsets $S_1, S_2, \ldots, S_k$ and is denoted by

$$S_1 + S_2 + \cdots + S_k$$

or by

$$\sum_{i=1}^{k} S_i.$$

Exercises

Exercise 2.2.9. Which of the following sets of vectors $\alpha = (a_1, \ldots, a_n)$ in $\mathbb{R}^n$ are subspaces of $\mathbb{R}^n$ $(n \geq 3)$?

1. all α such that $a_1 \geq 0$;
2. all α such that $a_1 + 3a_2 = a_3$;
3. all α such that $a_2 = a_1^2$;
4. all α such that $a_1 a_2 = 0$;
5. all α such that a_2 is rational.

Solution 2.2.10. A subset $V \subseteq \mathbb{R}^n$ is a subspace if it satisfies:

- Zero vector: $0 \in V$.
- Closed under addition: If $v, w \in V$, then $v + w \in V$.
- Closed under scalar multiplication: If $v \in V$ and $c \in \mathbb{R}$, then $cv \in V$.

1. all α such that $a_1 \geq 0$.

   - Zero vector: $(0, 0, \ldots, 0)$ has $a_1 = 0 \geq 0$.
   - Scalar multiplication:
     Let $v = (1, 0, \ldots, 0) \in V$. Then $-1 \cdot v = (-1, 0, \ldots, 0)$, but $a_1 = -1 < 0$.

   Therefore, this is not a subspace.

2. all α such that $a_1 + 3a_2 = a_3$.

   - Zero vector: $(0, 0, \ldots, 0)$ satisfies $0 + 3 \cdot 0 = 0$.
   - Addition:
     Let $v = (a_1, \ldots, a_n)$, $w = (b_1, \ldots, b_n) \in V$. Then $a_1 + 3a_2 = a_3, b_1 + 3b_2 = b_3$.
     So $(a_1 + b_1) + 3(a_2 + b_2) = a_3 + b_3$.
   - Scalar multiplication:
     For $c \in \mathbb{R}$, $ca_1 + 3(ca_2) = c(a_1 + 3a_2) = ca_3$.

   Therefore, this is a subspace.

3. all α such that $a_2 = a_1^2$.

   - Zero vector: $(0, 0, \ldots, 0)$ satisfies $0 = 0^2$.
   - Addition:
     Let $v = (1, 1, 0, \ldots, 0), w = (2, 4, 0, \ldots, 0) \in V$. Then $v + w = (3, 5, 0, \ldots, 0)$, but $5 \neq 3^2 = 9$.

   Therefore, this is not a subspace.

4. all α such that $a_1 a_2 = 0$.

- Zero vector: $(0, 0, \ldots, 0)$ satisfies $0 \cdot 0 = 0$.
- Addition:
  Let $v = (1, 0, \ldots, 0), w = (0, 1, 0, \ldots, 0) \in V$. Then $v + w = (1, 1, 0, \ldots, 0)$, but $1 \cdot 1 = 1 \neq 0$.

Therefore, this is not a subspace.

5. all α such that a_2 is rational.

- Zero vector: $(0, 0, \ldots, 0)$ has $a_2 = 0 \in \mathbb{Q}$.
- Scalar multiplication:
  Let $v = (0, 1, 0, \ldots, 0) \in V$. Then $\pi \cdot v = (0, \pi, 0, \ldots, 0)$, but $\pi \notin \mathbb{Q}$.

Therefore, this is not a subspace.

Our solution is done.

Exercise 2.2.11. Let V be the (real) vector space of all functions f from $\mathbb{R}$ into $\mathbb{R}$. Which of the following sets of functions are subsapces of V?

1. all f such that $f(x^2) = f(x)^2$;
2. all f such that $f(0) = f(1)$;
3. all f such that $f(3) = 1 + f(-5)$;
4. all f such that $f(-1) = 0$;
5. all f which are continuous.

Solution 2.2.12. A subset $W \subseteq V$ is a subspace if it satisfies:

- The zero function $0(x) = 0$ is in W.
- W is closed under addition: if $f, g \in W$, then $f + g \in W$.
- W is closed under scalar multiplication: if $f \in W$ and $c \in \mathbb{R}$, then $cf \in W$.

1. All f such that $f(x^2) = f(x)^2$.

- Zero function: $f(x) = 0$ satisfies $0 = 0^2$, so $0 \in W$.
- Addition:

  Let $f, g \in W$. Then,

  $$(f+g)(x^2) = f(x^2) + g(x^2) = f(x)^2 + g(x)^2.$$

  But,

  $$(f+g)(x)^2 = (f(x)+g(x))^2 = f(x)^2 + 2f(x)g(x) + g(x)^2.$$

  For these to be equal, we need $2f(x)g(x) = 0$ for all x, which is not generally true.

Therefore, this is not a subspace.

2. all f such that $f(0) = f(1)$.

   - Zero function: $f(0) = 0 = f(1)$, so $0 \in W$.
   - Addition: If $f(0) = f(1)$ and $g(0) = g(1)$, then

     $$(f+g)(0) = f(0) + g(0) = f(1) + g(1) = (f+g)(1).$$

   - Scalar multiplication:

     $$(cf)(0) = cf(0) = cf(1) = (cf)(1).$$

   Therefore, this is a subspace.

3. all f such that $f(3) = 1 + f(-5)$.

   - Zero function:

     $$f(3) = 0 \text{ but } 1 + f(-5) = 1 + 0 = 1.$$

     So the zero function is not included.

   Therefore, this is not a subspace.

4. all f such that $f(-1) = 0$.

- Zero function: $f(-1) = 0$ so $0 \in W$.
- Addition: If $f(-1) = 0$ and $g(-1) = 0$, then

$$(f+g)(-1) = 0 + 0 = 0.$$

- Scalar multiplication:

$$(cf)(-1) = cf(-1) = c \cdot 0 = 0.$$

Therefore, this is a subspace.

5. all f which are continuous.

- Zero function: Constant function is continuous.
- Addition: Sum of continuous functions is continuous.
- Scalar multiplication: Scalar multiple of a continuous function is continuous.

Therefore, this is a subspace.

Our solution is done.

Exercise 2.2.13. Is the vector $(3, -1, 0, -1)$ in the subspace of $\mathbb{R}^4$ spanned by the vectors $(2, -1, 3, 2)$, $(-1, 1, 1, -3)$ and $(1, 1, 9, -5)$?

Solution 2.2.14. We want to determine whether there exists scalars $a, b, c \in \mathbb{R}$ such that,

$$a(2, -1, 3, 2) + b(-1, 1, 1, -3) + c(1, 1, 9, -5) = (3, -1, 0, -1).$$

This leads to the following system of equations:

$$\begin{aligned}
2a - b + c &= 3 \\
-a + b + c &= -1 \\
3a + b + 9c &= 0 \\
2a - 3b - 5c &= -1
\end{aligned}$$

The augmented matrix for the system is,

$$\left[\begin{array}{ccc|c}
2 & -1 & 1 & 3 \\
-1 & 1 & 1 & -1 \\
3 & 1 & 9 & 0 \\
2 & -3 & -5 & -1
\end{array}\right].$$

We use Gaussian elimination to simplify the matrix.

1. $R_2 \longrightarrow -R_2$.
2. $R_1 \iff R_2$.
3. $R_2 \longrightarrow R_2 - 2R_1$.
4. $R_3 \longrightarrow R_3 - 3R_1$.
5. $R_4 \longrightarrow R_4 - 2R_1$.
6. $R_1 \longrightarrow R_1 + R_2$.
7. $R_3 \longrightarrow R_3 - 4R_2$.
8. $R_4 \longrightarrow R_4 + R_2$.

We get the final matrix:

$$\left[\begin{array}{ccc|c}
1 & 0 & 2 & 2 \\
0 & 1 & 3 & 1 \\
0 & 0 & 0 & -7 \\
0 & 0 & 0 & -2
\end{array}\right].$$

The third and fourth rows now represent the equations:

$$0a + 0b + 0c = -7 \implies 0 = -7$$
$$0a + 0b + 0c = -2 \implies 0 = -2.$$

These are contradictions, meaning the system has no solution.

Exercise 2.2.15. Let W be the set of all $(x_1, x_2, x_3, x_4, x_5)$ in $\mathbb{R}^5$ which satisfy

$$\begin{aligned} 2x_1 - x_2 + \frac{4}{3}x_3 - x_4 &= 0 \\ x_1 + \frac{2}{3}x_3 - x_5 &= 0 \\ 9x_1 - 3x_2 + 6x_3 - 3x_4 - 3x_5 &= 0. \end{aligned}$$

Find a finite set of vectors which span W.

Solution 2.2.16. The subspace W consists of all vectors $(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5$ satisfying:

$$\begin{aligned} 2x_1 - x_2 + \frac{4}{3}x_3 - x_4 &= 0 \\ x_1 + \frac{2}{3}x_3 - x_5 &= 0 \\ 9x_1 - 3x_2 + 6x_3 - 3x_4 - 3x_5 &= 0. \end{aligned}$$

To eliminate fractions, multiply each equation by 3:

$$\begin{aligned} 6x_1 - 3x_2 + 4x_3 - 3x_4 &= 0 \\ 3x_1 + 2x_3 - 3x_5 &= 0 \\ 9x_1 - 3x_2 + 6x_3 - 3x_4 - 3x_5 &= 0. \end{aligned}$$

The augmented matrix for the system is,

$$\left[\begin{array}{ccccc|c} 6 & -3 & 4 & -3 & 0 & 0 \\ 3 & 0 & 2 & 0 & -3 & 0 \\ 9 & -3 & 6 & -3 & -3 & 0 \end{array}\right].$$

We use Gaussian elimination to simplify the matrix.

1. $R_1 \longrightarrow R_1 - 2R_2$.
2. $R_3 \longrightarrow R_3 - 3R_2$.

The matrix becomes,

$$\left[\begin{array}{ccccc|c} 0 & -3 & 0 & -3 & 6 & 0 \\ 3 & 0 & 2 & 0 & -3 & 0 \\ 0 & -3 & 0 & -3 & 6 & 0 \end{array}\right].$$

Since row 1 and row 3 are identical, we can discard row 3. The system reduces to:

$$\begin{aligned} \text{(A) } -3x_2 - 3x_4 + 6x_5 &= 0 \\ \text{(B) } 3x_1 + 2x_3 - 3x_5 &= 0 \end{aligned}$$

Divide equation (A) by -3 :

$$x_2 + x_4 - 2x_5 = 0.$$

So the system simplifies to :

$$\begin{aligned} \text{(A') } 3x_1 + 2x_3 - 3x_5 &= 0 \\ \text{(B') } x_2 + x_4 - 2x_5 &= 0 \end{aligned}$$

We have 5 variables and 2 independent equations $\implies$ 3 free variables. Choose x_3, x_4, x_5 as free parameters. From (A'):

$$3x_1 = 3x_5 - 2x_3 \implies x_1 = x_5 - \frac{2}{3}x_3.$$

From (B'):

$$x_2 = 2x_5 - x_4.$$

So the general solution is:

$$(x_1, x_2, x_3, x_4, x_5) = (x_5 - \frac{2}{3}x_3, 2x_5 - x_4, x_3, x_4, x_5).$$

This can be written as a linear combination:

$$= x_3(-\frac{2}{3}, 0, 1, 0, 0) + x_4(0, -1, 0, 1, 0) + x_5(1, 2, 0, 0, 1).$$

To avoid fractions, multiply the first vector by 3:

$$(-2, 0, 3, 0, 0).$$

A finite set of vectors that span W is:

$$(-2, 0, 3, 0, 0), \quad (0, -1, 0, 1, 0), \quad (1, 2, 0, 0, 1).$$

Exercise 2.2.17. Let F be a field and let n be a positive integer ($n \geq 2$). Let V be the vector space of all $n \times n$ matrices over F. Which of the following sets of matrices A in V are subspaces of V?

1. all invertible A ;
2. all non-invertible A ;
3. all A such that $AB = BA$, where B is some fixed matrix in V ;
4. all A such that $A^2 = A$.

Solution 2.2.18.

1. all invertible A.

    - The zero matrix 0 is not invertible.

    Therefore, this is not a subspace.

2. all non-invertible A.

- The zero matrix is non-invertible, so $0 \in W$.
- Consider two non-invertible matrices whose sum is invertible.

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

  Both are non-invertible, but,

$$A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

  is invertible.

Therefore, this is not a subspace.

3. all A such that $AB = BA$, where B is some fixed matrix in V.

- Zero vector: $0B = B0 = 0$.
- Addition: If $A_1B = BA_1$ and $A_2B = BA_2$, then

$$(A_1 + A_2)B = A_1B + A_2B = BA_1 + BA_2 = B(A_1 + A_2).$$

- Scalar multiplication: If $AB = BA$, then

$$(cA)B = c(AB) = c(BA) = B(cA).$$

Therefore, this is a subspace.

4. all A such that $A^2 = A$.

- Zero vector: $0^2 = 0$.
- If $A^2 = A$ and $B^2 = B$, then,

$$(A + B)^2 = A^2 + AB + BA + B^2 = A + AB + BA + B.$$

  For this to equal $A + B$, we need $AB + BA = 0$, which is generally not true.

Therefore, this is not a subspace.

Our solution is done.

Exercise 2.2.19.

1. Prove that the only subspaces of $\mathbb{R}^1$ are $\mathbb{R}^1$ and the zero subspace.
2. Prove that a subspace of $\mathbb{R}^2$ is $\mathbb{R}^2$, or the zero subspace, or consists of all scalar multiples of some fixed vector in $\mathbb{R}^2$.
3. Can you describe the subspaces of $\mathbb{R}^3$?

Solution 2.2.20.

1. Let $W \subseteq \mathbb{R}^1$ be a subspace. If $W = \{0\}$, then it is the zero subspace. If W contains a non-zero vector v, then since $\mathbb{R}^1$ is one-dimensional, every vector in $\mathbb{R}^1$ is of the form αv for some $\alpha \in \mathbb{R}$. By closure under scalar multiplication, $W = \mathbb{R}^1$.

   Therefore, the only subspaces of $\mathbb{R}^1$ are $\mathbb{R}^1$ itself and the zero subspace.

2. Let $W \subseteq \mathbb{R}^2$ be a subspace. If $W = \{0\}$, then it is the zero subspace. If W contains a non-zero vector v, then by closure under scalar multiplication, W contains all vectors αv, i.e., the line through the origin in the direction of v. If W contains another vector w that is not a scalar multiple of v, then v and w are linearly independent. Since $\mathbb{R}^2$ is two-dimensional, $\text{span}\{v, w\} = \mathbb{R}^2$. By closure under linear combinations, $W = \mathbb{R}^2$.

   Therefore the subspaces of $\mathbb{R}^2$ are:

   - The zero subspace.
   - Lines through the origin.
   - $\mathbb{R}^2$ itself.

3. Let $W \subseteq \mathbb{R}^3$ be a subspace. If $W = \{0\}$, then it is the zero subspace. If W contains a nonzero vector v, then by closure under scalar multiplication, W contains all vectors αv, i.e., the line through the origin in the direction of v. If W contains two linearly independent vectors v and w, then $\text{span}\{v, w\}$ is a plane through the origin. By closure under linear combinations, W contains this entire plain. If W contains three linearly independent vectors, then $W = \mathbb{R}^3$. Therefore, the subspaces of $\mathbb{R}^3$ are:

- The zero subspace.
- Lines through the origin.
- Planes through the origin.
- $\mathbb{R}^3$ itself.

Our solution is done.

2.3 Bases and Dimension

Definition 2.3.1. Let V be a vector space over F. A subset S of V is said to be **linearly dependent** if there exist distinct vectors $\alpha_1, \ldots, \alpha_n$ in S and scalars $c_1, c_2, \ldots, c_n$ in F, not all of which are 0, such that

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0.$$

A set which is not linearly dependent is called **linearly independent**. If the set S contains only finitely many vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$, we sometimes say that $\alpha_1, \alpha_2, \ldots, \alpha_n$ are dependent (or independent) instead of saying S is dependent (or independent).

The following are easy consequences of the definition.

1. Any set which contains a linearly dependent set is linearly dependent.
2. Any subset of a linearly independent set is linearly independent.
3. Any set which contains the 0 vector is linearly dependentl for $1 \cdot 0 = 0$.
4. A set S of vectors is linearly independent if and only if each finite subset subset of S is linearly independent, i.e., if and only if for any distinct vectors $\alpha_1, \ldots, \alpha_n$ of S, $c_1\alpha_1 + \cdots + c_n\alpha_n = 0$ implies each $c_i = 0$.

Definition 2.3.2. Let V be a vector space. A **basis** for V is a linearly independent set of vectors in V which spans the space V. The space V is **finite dimensional** if it has a finite basis.

Theorem 2.3.3. *Let V be a vector space which is spanned by a finite set of vectors $\beta_1, \beta_2, \ldots, \beta_m$. Then any independent set of vectors in V is finite and contains no more than m elements.*

Proof. To prove this theorem it suffices to show that every subset S of V which contains more than m vectors is linearly dependent. Let S be such a set. In S there are distinct vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$ where $n > m$. Since $\beta_1, \ldots, \beta_m$ span V, there exist scalars A_{ij} in F such that

$$\alpha_j = \sum_{i=1}^{m} A_{ij}\beta_i.$$

For any n scalars $x_1, x_2, \ldots, x_n$ we have

$$\begin{aligned} x_1\alpha_1 + \cdots + x_n\alpha_n &= \sum_{j=1}^{n} x_j\alpha_j \\ &= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} A_{ij}\beta_i \\ &= \sum_{j=1}^{n} \sum_{i=1}^{m} (A_{ij}x_j)\beta_i \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A_{ij}x_j \right) \beta_i. \end{aligned}$$

Since $n > m$, Theorem 1.4.4 of Chapter 1 implies that there exists scalars $x_1, x_2, \ldots, x_n$ not all 0 such that

$$\sum_{j=1}^{n} A_{ij}x_j = 0, \quad 1 \le i \le m.$$

Hence $x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = 0$. This shows that S is a linearly dependent set.

□

Corollary 2.3.4. *Let V be a finite-dimensional vector space and let $n = \dim V$. Then*

1. *any subset of V which contains more than n vectors is linearly dependent.*

2. *no subset of V which contains fewer than n vectors can span V.*

Lemma 2.3.5. *Let S be a linearly independent subset of a vector space V. Suppose β is a vector in V which is not in the subspace spanned by S. Then the set obtained by adjoining β to S is linearly independent.*

Proof. Suppose $\alpha_1, \ldots, \alpha_m$ are distinct vectors in S and that

$$c_1\alpha_1 + \cdots + c_m\alpha_m + b\beta = 0.$$

Then $b = 0$; otherwise,

$$\beta = \left(-\frac{c_1}{b}\right)\alpha_1 + \cdots + \left(-\frac{c_m}{b}\right)\alpha_m$$

and β is in the subspace spanned by S. Thus $c_1\alpha_1 + \cdots + c_m\alpha_m = 0$, and since S is a linearly independent set each $c_i = 0$.

□

2.4 To be Continued...