# Chapter 1: Systems of Linear Equations and Matrices (Part 2)

## 1 Matrix Operations

**Definition 1.1.** A **matrix** is a rectangular array of numbers. The numbers in the array are called the entries in the matrix.

A general $m \times n$ matrix is of the form:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

- A matrix $\mathbf{A}$ with $n$ rows and $n$ columns is called a **square matrix** of order $n$.
- The entries $a_{11}, a_{22}, ..., a_{nn}$ are said to be on the **main diagonal** of $\mathbf{A}$.

**Definition 1.2.** Two matrices are defined to be **equal** if they have the same size and their corresponding entries are equal.

**Definition 1.3.** Matrix Operations

- If $\mathbf{A}$ and $\mathbf{B}$ are matrices of the same size, then the **sum** $\mathbf{A} + \mathbf{B}$ is the matrix obtained by adding the entries of $\mathbf{B}$ to the corresponding entries of $\mathbf{A}$, and the difference $\mathbf{AB}$ is the matrix obtained by subtracting the entries of B from the corresponding entries of $\mathbf{A}$. Matrices of different sizes cannot be added or subtracted.
- If $\mathbf{A}$ is any matrix and $c$ is any scalar, then the product $c\mathbf{A}$ is the matrix obtained by multiplying each entry of the matrix $\mathbf{A}$ by $c$. The matrix $c\mathbf{A}$ is said to be a **scalar multiple** of $\mathbf{A}$.
- If $\mathbf{A}$ is an $m \times r$ matrix and $\mathbf{B}$ is an $r \times n$ matrix, then the product $\mathbf{AB}$ is the $m \times n$ matrix whose entries are determined as follows: To find the entry in row $i$ and column $j$ of $\mathbf{AB}$, single out row $i$ from the matrix $\mathbf{A}$ and column $j$ from the matrix $\mathbf{B}$. Multiply the corresponding entries from the row and column together, and then add up the resulting products.

**Definition 1.4** (Transpose)**.** If $\mathbf{A}$ is any $m \times n$ matrix, then the **transpose** of $\mathbf{A}$, denoted by $\mathbf{A}^T$, is defined to be the $n \times m$ matrix that results by interchanging the rows and columns of $\mathbf{A}$; that is, the first column of $\mathbf{A}^T$ is the first row of $\mathbf{A}$, the second column of $A^T$ is the second row of $A$, and so forth.

$$(\mathbf{A}^T)_{ij} = \mathbf{A}_{ji}, \qquad i = 1, 2, ..., m, \quad j = 1, 2, ..., n.$$

**Definition 1.5.** If $\mathbf{A}$ is a square matrix, then the **trace** of $\mathbf{A}$, denoted by $tr(A)$, is defined to be the sum of the entries on the main diagonal of $\mathbf{A}$.

Note: The trace of $\mathbf{A}$ is undefined if $\mathbf{A}$ is not a square matrix.

**Example 1.1.** Find trace of the following matrices.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & 4 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & -2 & -1 & 10 \\ 0 & 1 & 1 & -3 \end{bmatrix}.$$

## 1.1 Matrix Multiplication by Columns and by Rows

We can use matrix partitioning for finding particular rows or columns of a matrix product AB without computing the entire product. Specifically, the following formulas, whose proofs are left as exercises, show how individual column vectors of AB can be obtained by partitioning B into column vectors and how individual row vectors of AB can be obtained by partitioning A into row vectors.

**AB computed column by column**

$$\mathbf{AB} = \mathbf{A}[\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n] = [\mathbf{Ab}_1, \mathbf{Ab}_2, \dots, \mathbf{Ab}_n]$$

I.e. " the $j$-th column vector of $\mathbf{AB} = \mathbf{A}[j-\text{th column vector of} B]$"

**AB computed row by row**

$$\mathbf{AB} = \begin{bmatrix} a_1 \\ a_2 \\ a_m \end{bmatrix} \mathbf{B} = \begin{bmatrix} a_1 B \\ a_2 B \\ a_m B \end{bmatrix}.$$

I.e. " the $i$-th row vector of $\mathbf{AB} = [i - \text{th row vector of} \mathbf{A}]B$"

### 1.1.1 Matrix Products as Linear Combinations

**Definition 1.6** (Linear Combination)**.** If $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_r$ are matrices of the same size, and if $c_1, c_2, \ldots, c_r$ are scalars, then an expression of the form

$$c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \cdots + c_r\mathbf{A}_r$$

is called a **linear combination** of $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_r$ with **coefficients** $c_1, c_2, \ldots, c_r$.

**Theorem 1.1.** If $\mathbf{A}$ is an $m \times n$ matrix, and if $\mathbf{x}$ is an $n \times 1$ column vector, then the product $\mathbf{Ax}$ can be expressed as a linear combination of the column vectors of $\mathbf{A}$ in which the coefficients are the entries of $\mathbf{x}$.

**Proof:**

**Example 1.2.** (Exercise) Consider the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 11 \\ 11 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 11 \\ 11 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 11 \\ 11 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 121 \\ 221 \\ 1211 \\ 1111 \end{bmatrix},$$

Determine whether or not $\mathbf{w}$ can be expressed as a linear combination $w = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3\mathbf{v}_3$ and if so, what are the relevant values $\lambda_1, \lambda_2$, and $\lambda_3$?

**Example 7.7.** Consider the vectors

$$v_1 = \begin{bmatrix} 11 \\ 11 \\ 1 \\ 1 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 1 \\ 11 \\ 11 \\ 1 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1 \\ 1 \\ 11 \\ 11 \end{bmatrix} \qquad w = \begin{bmatrix} 121 \\ 221 \\ 1211 \\ 1111 \end{bmatrix}.$$

We ask whether $w$ can be expressed as a linear combination $w = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$, and if so, what are the relevant values $\lambda_1$, $\lambda_2$ and $\lambda_3$? Following Method 7.6, we write down the augmented matrix $[v_1|v_2|v_3|w]$ and row-reduce it:

$$\left[\begin{array}{ccc|c} 11 & 1 & 1 & 121 \\ 11 & 11 & 1 & 221 \\ 1 & 11 & 11 & 1211 \\ 1 & 1 & 11 & 1111 \end{array}\right] \xrightarrow{1} \left[\begin{array}{ccc|c} 1 & 1 & 11 & 1111 \\ 11 & 1 & 1 & 121 \\ 11 & 11 & 1 & 221 \\ 1 & 11 & 11 & 1211 \end{array}\right] \xrightarrow{2}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 11 & 1111 \\ 0 & -10 & -120 & -12100 \\ 0 & 0 & -120 & -12000 \\ 0 & 10 & 0 & 100 \end{array}\right] \xrightarrow{3} \left[\begin{array}{ccc|c} 1 & 1 & 11 & 1111 \\ 0 & 1 & 12 & 1210 \\ 0 & 0 & 1 & 100 \\ 0 & 1 & 0 & 10 \end{array}\right] \xrightarrow{4}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 11 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 100 \\ 0 & 1 & 0 & 10 \end{array}\right] \xrightarrow{5} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 100 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

(1: move the bottom row to the top; 2: subtract multiples of row 1 from the other rows; 3: divide rows 2,3 and 4 by $-10$, $-120$ and 10; 4: subtract multiples of row 3 from the other rows; 5: subtract multiples of row 2 from the other rows.)

The final matrix corresponds to the system of equations

$$\lambda_1 = 1 \qquad \lambda_2 = 10 \qquad \lambda_3 = 100 \qquad 0 = 0$$

so we conclude that

$$w = v_1 + 10v_2 + 100v_3.$$

In particular, $w$ can be expressed as a linear combination of $v_1$, $v_2$ and $v_3$. We can check the above equation directly:

$$v_1 + 10v_2 + 100v_3 = \begin{bmatrix} 11 \\ 11 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 10 \\ 110 \\ 110 \\ 10 \end{bmatrix} + \begin{bmatrix} 100 \\ 100 \\ 1100 \\ 1100 \end{bmatrix} = \begin{bmatrix} 121 \\ 221 \\ 1211 \\ 1111 \end{bmatrix} = w.$$

**Example 7.8.** Consider the vectors

$$a_1 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \qquad a_2 = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} \qquad a_3 = \begin{bmatrix} 0 \\ 3 \\ -2 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

To test whether $b$ is a linear combination of $a_1$, $a_2$ and $a_3$, we write down the relevant augmented matrix and row-reduce it:

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

$$\left[\begin{array}{ccc|c} 1 & 0 & -3 & -2 \\ 0 & 1 & 2 & -3 \\ 0 & 3 & 6 & 5 \end{array}\right] \xrightarrow{3} \left[\begin{array}{ccc|c} 1 & 0 & -3 & -2 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 14 \end{array}\right] \xrightarrow{4}$$

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

(Stage 1: move the top row to the bottom, and multiply the other two rows by $-1$; Stage 2: subtract 2 times row 1 from row 3; Stage 3: subtract 3 times row 2 from row 3; Stage 4: divide row 3 by 14; Stage 5: subtract multiples of row 3 from rows 1 and 2.)

### 1.1.2 Linear independence

**Definition 1.7.** Let $\mathcal{V} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list of vectors in $\mathbb{R}^n$.
Suppose that there is a linear relation between the vectors $\mathbf{v}_i$ is a relation of the form

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k = 0$$

where $\lambda_1, \ldots, \lambda_k$ are scalars. In other words, we can express 0 as a linear combination of $\mathcal{V}$.

- For any list we have the **trivial** linear relation

$$0v_1 + 0v_2 + \cdots + 0v_k = 0.$$

- There may or may not be any **nontrivial** linear relations.
- If the list $\mathcal{V}$ has a **nontrivial** linear relation, we say that it is a **linearly dependent** list.
- If the only linear relation on $\mathcal{V}$ is the **trivial one**, we instead say that $\mathcal{V}$ is **linearly independent** or just **independent**.

**Example 1.3.** Consider the list $\mathcal{V}$ given by

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Determine if the list $\mathcal{V}$ is linearly dependent.
**Solution:** Since there is a nontrivial linear relation $\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 - \mathbf{v}_4 = 0$, then the list $\mathcal{A}$ is dependent.

**Example 1.4.** Consider the list $\mathcal{A}$ given by

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 12 \\ 1 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Determine if there is a nontrivial linear relation on the list $\mathcal{A}$.

**Solution:** Since there is a nontrivial linear relation $3\mathbf{a}_1 + \mathbf{a}_2 + 3\mathbf{a}3 - 4\mathbf{a}_4 = 0$, then the list $\mathcal{A}$ is dependent.

**Example 1.5.** Consider the list $\mathcal{U}$ given by

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Determine if the list $\mathcal{U}$ is linearly dependent.

**Example 1.6.** Let $\mathbf{v}$ and $\mathbf{w}$ be vectors in $\mathbb{R}^n$, and suppose that $\mathbf{v} \neq 0$ and that the list $(\mathbf{v}, \mathbf{w})$ is linearly dependent.
Prove that there is a number $\alpha$ such that $\mathbf{w} = \alpha\mathbf{v}$.

**Proof**
Because the list is dependent, there is a linear relation

$$\lambda\mathbf{v} + \mu\mathbf{w} = 0$$

where $\lambda$ and $\mu$ are not both zero.
There are apparently three possibilities:
(a) $\lambda \neq 0$ and $\mu \neq 0$;
(b) $\lambda = 0$ and $\mu \neq 0$;
(c)$\lambda \neq 0$ and $\mu = 0$;
However, case (c) is not really possible. Indeed, in case (c) the equation $\lambda\mathbf{v} + \mu\mathbf{w} = 0$ would reduce to $\lambda\mathbf{v} = 0$, and we could multiply by $1/\lambda$ to get $\mathbf{v} = 0$, which contradicts that $\mathbf{v} \neq 0$ by assumption.

In case (a) or (b) we can take $\alpha = -\lambda/\mu$ and we have $\mathbf{w} = \alpha\mathbf{v}$.

### 1.1.3 Checking linear dependence by using row-reduction

There is a systematic method using row-reduction for checking linear (in)dependence, as we will explain shortly. We first need a preparatory observation
We will consider 3 case of matrix $\mathbf{B}$ of size $p \times q$.

- When $p < q$, the matrix $\mathbf{B}$ is said to be "*wide.*"
- When $p = q$, the matrix $\mathbf{B}$ is said to be "*square.*"
- When $p > q$, the matrix $\mathbf{B}$ is said to be "*tall.*"

**Lemma 1.1.** Let $\mathbf{B}$ be a $p \times q$ matrix in RREF.

(a) If $\mathbf{B}$ is wide ($p < q$) then it is impossible for every column to contain a pivot.

(b) If $\mathbf{B}$ is square ($p = q$)then the only way for every column to contain a pivot is when $\mathbf{B} = \mathbf{I}_q$.

(c) If $\mathbf{B}$ is tall ($p > q$)then the only way for every column to contain a pivot is if $\mathbf{B}$ consists of $\mathbf{I}_q$ with $p - q$ rows of zeros added at the bottom. I.e. $\mathbf{B} = \left[\begin{array}{c} \mathbf{I}_q \\ \hline \\ \mathbf{0}_{(p-q)\times q} \end{array}\right]$

<u>**Proof:**</u>

**Theorem 1.2.** **Procedure for checking linear dependence by using row-reduction**
Let $\mathcal{V} = \mathbf{v}_1, \ldots, \mathbf{v}_K$ be a list of vectors in $\mathbb{R}^n$. We can check whether this list is dependent as follows. Let

$$\mathbf{A} = \left[\begin{array}{c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_K \end{array}\right]$$

and let $\mathbf{B}$ be the RREF of $\mathbf{A}$. Then

- If every column of $\mathbf{B}$ contains a pivot (so $\mathbf{B}$ has the form discussed in the previous Lemma) then $\mathcal{V}$ is independent.
- If some column of $\mathbf{B}$ has no pivot, then the list $\mathcal{V}$ is dependent. Moreover, we can find the coeffcients $\lambda_i$ in a nontrivial linear relation by solving the vector equation $\mathbf{B}\lambda = 0$ (which is easy because B is in RREF).

**Proof:**

*Proof of correctness of Method 8.8.* Put

$$A = \left[\begin{array}{c|c|c} v_1 & \cdots & v_m \end{array}\right]$$

as in step (a) of the method, and let $B$ be the RREF form of $A$. Note that for any vector $\lambda = \begin{bmatrix}\lambda_1 & \dots & \lambda_m\end{bmatrix}^T \in \mathbb{R}^m$, we have

$$A\lambda = \left[\begin{array}{c|c|c} v_1 & \cdots & v_m \end{array}\right] \begin{bmatrix}\lambda_1 \\ \vdots \\ \lambda_m\end{bmatrix} = \lambda_1 v_1 + \cdots + \lambda_m v_m.$$

Thus, linear relations on our list are just the same as solutions to the homogeneous equation $A\lambda = 0$. By Theorem 6.8, these are the same as solutions to the equation $B\lambda = 0$, which can be found by Method 5.4. If there is a pivot in every column then none of the variables $\lambda_i$ is independent, so the only solution is $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$. Thus, the only linear relation on $\mathcal{V}$ is the trivial one, which means that the list $\mathcal{V}$ is linearly independent.

Suppose instead that some column (the $k$'th one, say) does not contain a pivot. Then in Method 5.4 the variable $\lambda_k$ will be independent, so we can choose $\lambda_k = 1$. This will give us a nonzero to solution to $B\lambda = 0$, or equivalently $A\lambda = 0$, corresponding to a nontrivial linear relation on $\mathcal{V}$. This shows that $\mathcal{V}$ is linearly dependent. □

**Remarks**

If $K \geq n$ then $\mathcal{V}$ is automatically dependent and we do not need to go through the method.
For example, any list of 5 vectors in $\mathbb{R}^3$ is automatically dependent, any list of 10 vectors in $\mathbb{R}^9$ is automatically dependent, and so on.)
Indeed, in this case the matrices **A** and **B** are wide, so it is impossible for **B** to have a pivot in every column.

**Example 1.7.** Consider the list $\mathcal{V}$ given by

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Determine if the list $\mathcal{V}$ is linearly dependent.

**Solution:**

**Example 8.10.** In example 8.2 we considered the list

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

We can write down the corresponding matrix and row-reduce it as follows:

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

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

The end result has no pivot in the last column, so the original list is dependent. To find a specific linear relation, we solve the equation

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

to get $\lambda_1 = -\lambda_4$, $\lambda_2 = -\lambda_4$ and $\lambda_3 = \lambda_4$ with $\lambda_4$ arbitrary. Taking $\lambda_4 = 1$ gives $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (-1, -1, 1, 1)$, corresponding to the relation $-v_1 - v_2 + v_3 + v_4 = 0$.

**Example 1.8.** Consider the list $\mathcal{A}$ given by

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 12 \\ 1 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Determine if there is a nontrivial linear relation on the list $\mathcal{A}$.

**Solution:**

**Example 8.11.** In Example 8.3 we considered the list

$$a_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad a_2 = \begin{bmatrix} 12 \\ 1 \end{bmatrix} \qquad a_3 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \qquad a_4 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Here we have 4 vectors in $\mathbb{R}^2$, so they must be dependent by Remark 8.9. Thus, there exist nontrivial linear relations

$$\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 + \lambda_4 a_4 = 0.$$

To actually find such a relation, we write down the corresponding matrix and row-reduce it as follows:

$$\begin{bmatrix} 1 & 12 & -1 & 3 \\ 2 & 1 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 12 & -1 & 3 \\ 0 & -23 & 1 & -5 \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & 12 & -1 & 3 \\ 0 & 1 & -1/23 & 5/23 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -11/23 & 9/23 \\ 0 & 1 & -1/23 & 5/23 \end{bmatrix}$$

We now need to solve the matrix equation

$$\begin{bmatrix} 1 & 0 & -11/23 & 9/23 \\ 0 & 1 & -1/23 & 5/23 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

As this is in RREF, we can just read off the solution: $\lambda_1 = \frac{11}{23}\lambda_3 - \frac{9}{23}\lambda_4$ and $\lambda_2 = \frac{1}{23}\lambda_3 - \frac{5}{23}\lambda_4$ with $\lambda_3$ and $\lambda_4$ arbitrary. If we choose $\lambda_3 = 23$ and $\lambda_4 = 0$ we get $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (11, 1, 23, 0)$ so we have a relation

$$11a_1 + a_2 + 23a_3 + 0a_4 = 0.$$

(You should check directly that this is correct.) Alternatively, we can choose $\lambda_3 = 3$ and $\lambda_4 = -4$. Using the equations $\lambda_1 = \frac{11}{23}\lambda_3 - \frac{9}{23}\lambda_4$ and $\lambda_2 = \frac{1}{23}\lambda_3 - \frac{5}{23}\lambda_4$ we get $\lambda_1 = 3$ and $\lambda_2 = 1$ giving a different relation

$$3a_1 + a_2 + 3a_3 - 4a_4 = 0.$$

This is the relation that we observed in Example 8.3.

**Example 1.9.** Consider the list $\mathcal{U}$ given by

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Determine if the list $\mathcal{U}$ is linearly dependent.

**Example 8.12.** In Example 8.4 we considered the list $\mathcal{U}$ given by

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We can write down the corresponding matrix and row-reduce it as follows:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 0 & 0 \end{bmatrix}$$

The final matrix has a pivot in every column, as in Lemma 8.7. It follows that the list $\mathcal{U}$ is independent.

## 2 Algebraic Properties of Matrices

**Theorem 2.1. Properties of Matrix Arithmetic**
Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid.

(a) $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$ [Commutative law for matrix addition]

(b) $\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$ [Associative law for matrix addition]

(c) $\mathbf{A}(\mathbf{BC})=(\mathbf{AB})\mathbf{C}$ [Associative law for matrix multiplication]

(d) $\mathbf{A}(\mathbf{B}+\mathbf{C})=\mathbf{AB}+\mathbf{AC}$ [Left distributive law]

(e) $(\mathbf{B}+\mathbf{C})\mathbf{A}=\mathbf{BA}+\mathbf{CA}$ [Right distributive law]

(f) $\mathbf{A}(\mathbf{B}-\mathbf{C})=\mathbf{AB}-\mathbf{AC}$

(g) $(\mathbf{B}-\mathbf{C})\mathbf{A}=\mathbf{BA}-\mathbf{CA}$

(h) $a(\mathbf{B}+\mathbf{C})=a\mathbf{B}+a\mathbf{C}$

(i) $a(\mathbf{B}-\mathbf{C})=a\mathbf{B}-a\mathbf{C}$

(j) $(a+b)\mathbf{C}=a\mathbf{C}+b\mathbf{C}$

(k) $(a-b)\mathbf{C}=a\mathbf{C}-b\mathbf{C}$

(l) $a(b\mathbf{C})=(ab)\mathbf{C}$

(m) $a(\mathbf{BC})=(a\mathbf{B})\mathbf{C}=\mathbf{B}(a\mathbf{C})$

**Proof of (d)**

***Proof (d)*** We must show that $A(B+C)$ and $AB+AC$ have the same size and that corresponding entries are equal. To form $A(B+C)$, the matrices $B$ and $C$ must have the same size, say $m\times n$, and the matrix $A$ must then have $m$ columns, so its size must be of the form $r\times m$. This makes $A(B+C)$ an $r\times n$ matrix. It follows that $AB+AC$ is also an $r\times n$ matrix and, consequently, $A(B+C)$ and $AB+AC$ have the same size.

Suppose that $A=[a_{ij}]$, $B=[b_{ij}]$, and $C=[c_{ij}]$. We want to show that corresponding entries of $A(B+C)$ and $AB+AC$ are equal; that is,

$$\big(A(B+C)\big)_{ij}=(AB+AC)_{ij}$$

for all values of $i$ and $j$. But from the definitions of matrix addition and matrix multiplication, we have

$$\begin{aligned}\big(A(B+C)\big)_{ij}&=a_{i1}(b_{1j}+c_{1j})+a_{i2}(b_{2j}+c_{2j})+\cdots+a_{im}(b_{mj}+c_{mj})\\&=(a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{im}b_{mj})+(a_{i1}c_{1j}+a_{i2}c_{2j}+\cdots+a_{im}c_{mj})\\&=(AB)_{ij}+(AC)_{ij}=(AB+AC)_{ij} \quad \blacktriangleleft\end{aligned}$$

**Remarks**

It is not true that all laws of real arithmetic carry over to matrix arithmetic.

1. **No commutative law:**

   For example, in real arithmetic it is always true that $ab = ba$, which is called the commutative law for multiplication. In matrix arithmetic, however, the equality of $\mathbf{AB}$ and $\mathbf{BA}$ can fail for three possible reasons:

   - $\mathbf{AB}$ may be defined and $\mathbf{BA}$ may not (for example, if $\mathbf{A}$ is $2 \times 3$ and $\mathbf{B}$ is $3 \times 4$).
   - $\mathbf{AB}$ and $\mathbf{BA}$ may both be defined, but they may have different sizes (for example, if $\mathbf{A}$ is $2 \times 3$ and $\mathbf{B}$ is $3 \times 2$).
   - $\mathbf{AB}$ and $\mathbf{BA}$ may both be defined and have the same size, but the two products may be different (as illustrated in the next example).

2. **No cancellation law:**

$$\mathbf{AB} = \mathbf{AC} \qquad \not\Rightarrow \qquad \mathbf{B} = \mathbf{C}$$

   Example:

3. **A Zero Product with Nonzero Factors**

$$\mathbf{AB} = 0 \qquad \not\Rightarrow \qquad \mathbf{A} = 0 \qquad \text{and} \qquad \mathbf{B} = 0$$

   Example:

▶ EXAMPLE 3 **Failure of the Cancellation Law**

Consider the matrices

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

We leave it for you to confirm that

$$AB = AC = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$$

Although $A \neq 0$, canceling $A$ from both sides of the equation $AB = AC$ would lead to the incorrect conclusion that $B = C$. Thus, the cancellation law does not hold, in general, for matrix multiplication (though there may be particular cases where it is true).

▶ EXAMPLE 4 **A Zero Product with Nonzero Factors**

Here are two matrices for which $AB = 0$, but $A \neq 0$ and $B \neq 0$:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 7 \\ 0 & 0 \end{bmatrix}$$ ◀

# 3 Inverse of Matrix

**Definition 3.1.** Let $\mathbf{A}$ be a square matrix.

- If a matrix $\mathbf{B}$ of the same size can be found such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$ , then $\mathbf{A}$ is said to be **invertible** (or **nonsingular**) and $\mathbf{B}$ is called an **inverse** of $\mathbf{A}$.
  Note: In this case, we also have that $\mathbf{B}$ is **invertible** (or **nonsingular**) and $\mathbf{A}$ is called an **inverse** of $\mathbf{B}$.
- If no such matrix $\mathbf{B}$ can be found, then $\mathbf{A}$ is said to be **singular**.

**Example 3.1.**

$$\mathbf{A} = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}.$$

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

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

Thus, $\mathbf{A}$ and $\mathbf{B}$ are invertible and each is an inverse of the other.

**Theorem 3.1** (Uniqueness of inverse ). An invertible matrix has exactly one inverse.
I.e. If $\mathbf{B}$ and $\mathbf{C}$ are both inverses of the matrix $\mathbf{A}$, then $\mathbf{B} = \mathbf{C}$.

**THEOREM 1.4.4** *If $B$ and $C$ are both inverses of the matrix $A$, then $B = C$.*

*Proof* Since $B$ is an inverse of $A$, we have $BA = I$. Multiplying both sides on the right by $C$ gives $(BA)C = IC = C$. But it is also true that $(BA)C = B(AC) = BI = B$, so $C = B$. ◀

If $\mathbf{A} \in \mathbb{R}^{n\times n}$ is invertible, then its inverse will be denoted by the symbol $\mathbf{A}^{-1}$. I.e.

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$$

## 3.1 Computing the inverse of an invertible matrix

**Theorem 3.2.** The matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if $ad - bc \neq 0$, in which case the inverse is given by the formula

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

**Example 3.2.** Determine whether each of the following matrices is invertible. If so, find its inverse.

$$(a) \begin{bmatrix} 6 & 1 \\ 5 & 2 \end{bmatrix} \qquad \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}$$

**Example 3.3.** Solve a pair of equations of the form

$$\begin{aligned} u &= ax + by \\ v &= cx + dy \end{aligned}$$

for $x$ and $y$in terms of $a, b, c, d, u$ and $v$.

**Theorem 3.3.** If $\mathbf{A}$ and $\mathbf{B}$ are invertible matrices with the same size, then $\mathbf{AB}$ is invertible and

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

*Aproduct of any number of invertible matrices is invertible, and the inverse of the product is the product of the inverses in the reverse order.*

**Proof:**

**THEOREM 1.4.6** *If $A$ and $B$ are invertible matrices with the same size, then $AB$ is invertible and*

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

***Proof*** We can establish the invertibility and obtain the stated formula at the same time by showing that

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

But

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

and similarly, $(B^{-1}A^{-1})(AB) = I$. ◀

**Example 3.4.** Show that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Consider the matrices

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

We leave it for you to show that

$$AB = \begin{bmatrix} 7 & 6 \\ 9 & 8 \end{bmatrix}, \quad (AB)^{-1} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

and also that

$$A^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix}, \quad B^{-1}A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

**Powers of a Matrix**
Let $\mathbf{A}$ be a square matrix.

- We define the nonnegative integer powers of A to be

$$\mathbf{A}^0 = \mathbf{I} \qquad \text{and} \qquad \mathbf{A}^n = \mathbf{A}\mathbf{A}\ldots\mathbf{A} \qquad [\text{ n factors}]$$

- If $\mathbf{A}$ is invertible, then we define the negative integer powers of $\mathbf{A}$ to be

$$\mathbf{A}^{-n} = (\mathbf{A}^{-1})^n = \mathbf{A}^{-1}\mathbf{A}^{-1}\ldots\mathbf{A}^{-1} \qquad [\text{ n factors}]$$

- $$\mathbf{A}^r\mathbf{A}^s = \mathbf{A}^{r+s}$$
$$(A^r)^s = A^{rs}$$

**Example 3.5.** Let $\mathbf{A}^{-3} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Find $\mathbf{A}^6$.

**Example 3.6.** Prove that if **A** is an invertible matrix, then $A^T$ is also invertible and

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T.$$

*Proof* We can establish the invertibility and obtain the formula at the same time by showing that

$$A^T(A^{-1})^T = (A^{-1})^T A^T = I$$

But from part (*e*) of Theorem 1.4.8 and the fact that $I^T = I$, we have

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

which completes the proof. ◀