Chapter 6 Linear Dependence of a Set of Vectors

Definition 6.1 Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ be vectors in $\mathbb{R}^m$. A ***linear combination of vectors*** $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ is the summation,

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n,$$

for some scalars $c_1, c_2, \ldots, c_n$. So a vector $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ if there exist scalars $c_1, c_2, \ldots, c_n$ such that,

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n = \mathbf{b},$$

or equivalently, if there exists a solution $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{R}^n$ to the system of linear equations,

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]\mathbf{c} = \mathbf{b}$$

When the vector $\mathbf{b} = \mathbf{0}$, setting all the scalars $c_1, c_2, \ldots, c_n$ to zeros always solves the equation. The question whether there exists a nontrivial linear combination of a set of vectors being equal to zero vector turns out to be of great importance. It defines the notion of linear dependence below and links the ideas of rank and echelon forms.

6.1 Linear Dependence and Independence

Definition 6.2 Vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ in $\mathbb{R}^m$ are ***linearly dependent*** if there exists scalars $c_1, c_2, \ldots, c_n$ at least one of which being nonzero, such that,

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n = \mathbf{0},$$

otherwise, they are ***linearly independent***.

That is, the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are linearly independent if $c_1 = 0, c_2 = 0, \ldots, c_n = 0$ are the only scalars such that $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n = \mathbf{0}$.

Example The vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are linearly dependent if at least one of them is a zero vector.

Problem Show that the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are linearly dependent if there is a pair of identical vectors. Show that a set of distinct nonzero vectors can be linearly dependent.

Problem Show that the vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ are linearly dependent if there is a vector that is the sum of a pair of other two vectors.

Problem Simon & Blume [1994], page, 243, #11.2 and 11.3.

2. Which of the following pairs or triplets of vectors are linearly independent?
   a) $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$
   b) $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -4 \\ -2 \end{bmatrix},$
   c) $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$
   d) $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$
3. Determine whether or not each of the following collections of vectors in $\mathbb{R}^4$ are linearly independent:
   a) $\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$
   b) $\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

6.2 Linear Independence and System of Linear Equations

With the definition of linear combination, we can recast the questions about the solutions of a system of linear equations $\mathbf{Ax} = \mathbf{b}$. The questions become: Is there a linear combination of the columns of $\mathbf{A}$ that is equal to right-hand-side vector $\mathbf{b}$? If there is, is such a linear combination unique?

We can say that the columns of the matrix $\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]$ are linearly dependent if, and only if, there exists a nontrivial solution $\mathbf{x} \neq \mathbf{0}$ to the homogeneous system of linear equations $\mathbf{Ax} = \mathbf{0}$. In other words, the columns of the matrix $\mathbf{A}$ are linearly independent if, and only if, the only solution to the system is the trivial one.

The next theorem will link the linear independence of the columns of a matrix $\mathbf{A}$ to the uniqueness of the solution of the system of linear equations $\mathbf{Ax} = \mathbf{b}$.

Theorem 6.1 Let $\mathbf{A}$ be a square matrix in $\mathbb{R}^{n\times n}$. The columns of $\mathbf{A}$ are linearly independent if, and only if the system of linear equations $\mathbf{Ax} = \mathbf{b}$ has a unique solution for any right-hand-side vector $\mathbf{b}$.

Proof Suppose the system $\mathbf{Ax} = \mathbf{b}$ has a unique solution for each $\mathbf{b}$. The only solution to the system $\mathbf{Ax} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$. Thus by definition, the columns of $\mathbf{A}$ is linearly independent.

Now let the columns of $\mathbf{A}$ be linearly independent and suppose that there exists a vector $\mathbf{b}$ such that the system $\mathbf{Ax} = \mathbf{b}$ does have two distinct solutions $\mathbf{x}_1$ and $\mathbf{x}_2$. That is, $\mathbf{Ax}_1 = \mathbf{b}$, $\mathbf{Ax}_2 = \mathbf{b}$ and $\mathbf{x}_1 \neq \mathbf{x}_2$. So $\mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$. The linear combination of the columns of $\mathbf{A}$, using the elements of the nonzero vector $\mathbf{x}_1 - \mathbf{x}_2$ as scalars, is a zero vector. So the columns of $\mathbf{A}$ are linearly dependent. This is a contradiction. So the theorem holds.❑

We are now ready to show the equivalence relationships of rank, determinant, inverse, linear independence of columns and rows of $\mathbf{A}$ and the uniqueness of the solution of the system .

Corollary 6.1 Let $\mathbf{A}$ be a square matrix in $\mathbb{R}^{n\times n}$. The following statements are equivalent.

1. The columns of $\mathbf{A}$ are linearly independent.
2. The rows of $\mathbf{A}$ are linearly independent.
3. $\mathbf{A}^{-1}$ exists.
4. $|\mathbf{A}| \neq 0$, i.e., $\mathbf{A}$ is nonsingular.
5. $rank\ \mathbf{A} = rank\ \mathbf{A}^{\mathrm{T}} = n$.
6. $\mathbf{Ax} = \mathbf{b}$ has a unique solution for any given vector $\mathbf{b} \in \mathbb{R}^n$.

Proof We have shown in Theorem 6.1 that (1) is equivalent to (6) and the rest follows directly from Corollary 5.3 of the previous chapter.❑

In the next chapter, we will show a more general result that the number of independent rows equals the number of independent columns, even when the matrix $\mathbf{A}$ is not square or not full rank. The next important result to be shown here is that the rank of a matrix is equal to the number of linearly independent rows. We first need an alternative definition of linear dependence.

6.3 Alternative Definition of Linear Dependence

An alternative definition of the linear dependence can be given by the following proposition.

Proposition 6.1 Vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ in $\mathbb{R}^m$ are ***linearly dependent*** if, and only if, there exists some $j, j = 1,2,\dots,n$, such that $\mathbf{a}_j$ can be written as a linear combination of the rest of the vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n$. Similarly, the vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ are ***linearly independent*** if, and only if, none of the vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ can be written as a linear combination of the rest.

Proof We will prove only the alternate definition of linear dependence. If $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ are linearly dependent, there exist scalars $c_1, c_2, \dots, c_n$ such that $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n = \mathbf{0}$,, with $c_j \neq 0$, for some $j = 1,2,\dots,n$. We have,

$$c_j\mathbf{a}_j = -\big(c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_{j-1}\mathbf{a}_{j-1} + c_{j+1}\mathbf{a}_{j+1} + \cdots + c_n\mathbf{a}_n\big)$$

$$\mathbf{a}_j = -\frac{c_1}{c_j}\mathbf{a}_1 - \frac{c_2}{c_j}\mathbf{a}_2 - \cdots - \frac{c_{j-1}}{c_j}\mathbf{a}_{j-1} - \frac{c_{j+1}}{c_j}\mathbf{a}_{j+1} - \cdots - \frac{c_n}{c_j}\mathbf{a}_n$$

The reverse can be as easily proved.❑

The following corollary demonstrates the close relationship between the linear dependence and the possibility of a row being reduced into a zero row by a series of elementary row operations.

Corollary 6.2 Let $\mathbf{A} = \begin{bmatrix} \mathbf{a}_{1\cdot} \\ \mathbf{a}_{2\cdot} \\ \vdots \\ \mathbf{a}_{m\cdot} \end{bmatrix}$ be a matrix in $\mathbb{R}^{m\times n}$. The rows of a matrix **A** are linearly dependent if, and only if,

a. at least one of the rows can be reduced to a zero row by a series of elementary row operations.
b. the row echelon form of matrix **A** contains a zero row.

Proof For part (a), suppose that the rows $\mathbf{a}_{1\cdot}, \mathbf{a}_{2\cdot}, \dots, \mathbf{a}_{m\cdot}$ are linearly dependent. If some of the rows is zero the Corollary 6.2 is trivially true. Otherwise, by Proposition 6.1, we can write, for some constants $d_1, d_2, \dots d_{i-1}, d_{i+1}, \dots, d_m$, such that

$$\mathbf{a}_{i\cdot} = d_1\mathbf{a}_{1\cdot} + d_2\mathbf{a}_{2\cdot} + \cdots + d_{i-1}\mathbf{a}_{i-1\cdot} + d_{i+1}\mathbf{a}_{i+1\cdot} + \cdots + d_m\mathbf{a}_{m\cdot}$$

This means that the row i can be reduced into a zero row by multiplying row k by the constant d_k and subtract it from row i, $k = 1,2,\dots,i-1,i+1,\dots,m$. The reverse can be similarly shown. Part (b) follows immediately from part (a).❑

The rows of any matrix **A** can then be divided into two groups. The first group consists of rows that can be reduced to rows of zeros, i.e., written as a linear combination of the other rows of **A**, and the second are the other rows that cannot be reduced to rows of zeros. We will show next that we cannot reduce a row to zeros in the second group by elementary row operations using rows of these second group.

6.4 Linear Dependence and Rank

For a given matrix, we can separate the rows into two sets: the set of linearly independent rows and the rest. The number of rows in this set of linearly independent rows will be shown to be equal to the rank of the matrix in Theorem 6.3 below. But first we have to show how we obtain the set of linearly independent rows.

Let **A** be a matrix in $\mathbb{R}^{m\times n}$, where $\mathbf{A} = \begin{bmatrix} \mathbf{a}_{1\cdot} \\ \mathbf{a}_{2\cdot} \\ \vdots \\ \mathbf{a}_{m\cdot} \end{bmatrix}$.

Step 1: Set $t = 1, \boldsymbol{J} \leftarrow \{1\}$. Assuming $\mathbf{a}_{1\cdot} \neq \mathbf{0}$, set $\mathbf{B} \leftarrow \{\mathbf{a}_{1\cdot}\}$. **B** is full rank.
Step 2: If $t = m$, then **STOP**, otherwise set $t \leftarrow t + 1$, and $\mathbf{B}' = \begin{bmatrix} \mathbf{B} \\ \mathbf{a}_{t\cdot} \end{bmatrix}$.
Step 3: If $\mathbf{B}'$ is full rank, then $\boldsymbol{J} \leftarrow \boldsymbol{J} \cup \{t\}$, $\mathbf{B} \leftarrow \mathbf{B}'$.
Step 4: Go to Step 2.

As the result of this algorithm, the number of elements in $\boldsymbol{J}$ is the number of linearly independent rows of **A**. The matrix **B** is just a submatrix of **A** and we can assume that its rows are just the first k rows of **A**, and write

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} \\ \mathbf{a}_{k+1\cdot} \\ \vdots \\ \mathbf{a}_{m\cdot} \end{bmatrix}$$

Note that according to the algorithm, each row $\mathbf{a}_{i\cdot}$, $i = k + 1, k + 2, \dots, m$, the last row added in can be written as a linear combination of the rows in **B**. Thus we can perform a number of elementary row operations until these last $m-k$ rows are reduced to zero rows and, by the contrapositive statement of Corollary 6.2, the row echelon form of matrix **B** does not contain a zero row. Thus we have proved the next theorem.

Theorem 6.2 The rank of a matrix **A** in $\mathbb{R}^{m\times n}$ is equal to the number of its linearly independent rows.

We can state the following results.

Corollary 6.3 The elementary row operation cannot change the number of linearly independent rows of a given matrix.

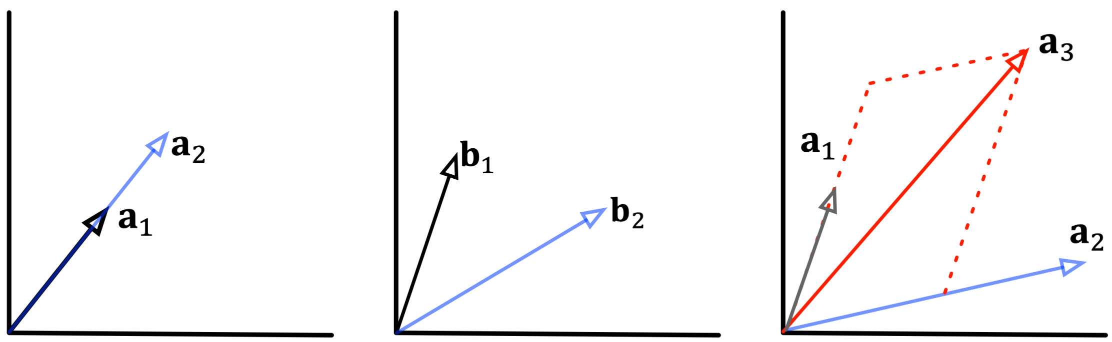
Proof By direct application of Theorem 6.2 and the detail is left as an exercise.❑

***Problem** Provide the detail of the proof of Corollary 6.3.

Problem Show that nonzero rows of row echelon matrix are linearly independent.

6.5 Graphical Illustration of Linear Dependence

See the Appendix of this chapter for the graphical interpretation of the scalar multiplication and addition of vectors.



(a) **(b)** **(c)**

Figure 6.1 Illustration of Linear Dependence and Linear Independence.

In Figure 6.1 **(a)**, the vectors $\mathbf{a}_1, \mathbf{a}_2$ are linearly dependent as one is just a multiple of the other. In Figure 6.1 **(b)**, the vectors $\mathbf{b}_1, \mathbf{b}_2$ are linearly independent as the only linear combination of $\mathbf{b}_1, \mathbf{b}_2$ that is the origin has to have all scalars being zeros. In Figure 6.1 **(c)**, the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are linearly dependent because we can find some scalars

c_1, c_2 such that $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 = \mathbf{a}_3$. Note that any pair of vectors out of these three vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are linearly independent. We can prove in general that if we have $n + 1$ or more vectors in $\mathbb{R}^n$, then these vectors are always linearly dependent We first need the result of the following lemma whose proof is left as an exercise.

Lemma 6.1 If vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ are linearly independent, then $\mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n$ are also linearly independent. If vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ are linearly dependent, then the vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{a}_{n+1}$ are also linearly dependent.

Proof Left as an exercise.❑

Theorem 6.3 Any $n + 1$ or more vectors in $\mathbb{R}^n$ are linearly dependent.

Proof By Lemma 6.1, it is sufficient to show only for $n + 1$ vectors in $\mathbb{R}^n$. Select the first n vectors out of these arbitrary $n + 1$ vectors. If these n vectors are linearly dependent, by Lemma 6.1, so are the $n + 1$ vectors. Now suppose these n vectors are linearly independent. Form a matrix by using these n vectors $\mathbf{a}_1, \mathbf{a}_2, \dots \mathbf{a}_n$ as columns and call this matrix $\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]$. This matrix $\mathbf{A}$ is full rank because of Corollary 6.1.

If the vector $\mathbf{a}_{n+1} = \mathbf{0}$, then the $n + 1$ vectors $\mathbf{a}_1, \mathbf{a}_2, \dots \mathbf{a}_n, \mathbf{a}_{n+1}$ are linearly dependent. Assume now that $\mathbf{a}_{n+1} \neq \mathbf{0}$. There is always a solution $\mathbf{x} \neq \mathbf{0}$ to the system of linear equations $\mathbf{Ax} = \mathbf{a}_{n+1}$. (Why?) That is, $\mathbf{a}_{n+1}$ can be written as a linear combination of the first n vectors. So the $n + 1$ vectors are linearly dependent.❑

Problem Answer the question (Why?) in the proof of the previous theorem.

Problem **Leon** [1994], page 141, #14. Let $\mathbf{x}_1, \mathbf{x}_2, \dots \mathbf{x}_k$ be linearly independent vectors in $\mathbb{R}^n$ and let $\mathbf{A}$ be a nonsingular $n \times n$ matrix. Define $\mathbf{y}_i = \mathbf{Ax}_i$ for $i = 1,2, \dots, k$. Show that $\mathbf{y}_1, \mathbf{y}, \dots \mathbf{y}_k$ are linearly independent

Appendix

Graphical Representation of Scalar Multiplication and Addition of Vectors

If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$, and c is a scalar, then

$$c\mathbf{x} = c\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n,$$

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \in \mathbb{R}^n.$$

The scalar multiplication and addition of vectors always produce a new vector in the same $\mathbb{R}^n$. We can illustrate graphically the scalar multiplication and vector addition in $\mathbb{R}^2$ as follows.

Let $\mathbf{x}$ be a vector in $\mathbb{R}^2$. This vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ can be represented by a point on the two dimensional Euclidian plane, with the value on the horizontal axis being x_1 and the value on the vertical axis being x_2. Of course, this representation can be extend to vectors of any order n.

To visualize a vector, it is convenient to use an arrow stemming from the point of origin with its arrow head pointing at $\mathbf{x}$. With this representation, an arrow not only shows how far away the vector $\mathbf{x}$ is from the origin, but also conveys the direction of the vector away from the origin.

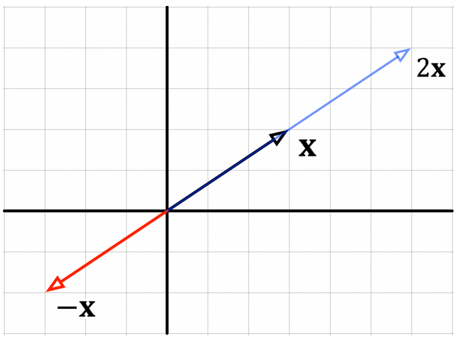


Figure 6.2 Scalar multiplication of a vector in $\mathbb{R}^2$.

When $\mathbf{x}$ is multiplied by 2, for example, the direction of the vector $2\mathbf{x}$ does not change but its arrow is extended and doubles in length. If it is multiplied by some positive number less than one, the direction does not change but the arrow is shortened proportionately. If $\mathbf{x}$ is multiplied by -1, the arrow representing $-\mathbf{x}$ points to the opposite direction of $\mathbf{x}$ but with equal length.

The figure below shows the vectors $\mathbf{x}, \mathbf{y}$ and the addition $\mathbf{x} + \mathbf{y}$. To get to the point $\mathbf{x} + \mathbf{y}$, we can start at point $\mathbf{x}$ and then, taking $\mathbf{x}$ as the new origin, draw the arrow of $\mathbf{y}$. The resulting arrow point will be the arrow point of the vector $\mathbf{x}+\mathbf{y}$. By the same reasoning, we can draw the arrow of vector $\mathbf{x}$ at the arrow point of $\mathbf{y}$ to arrive at the same arrow point of vector $\mathbf{x}+\mathbf{y}$.

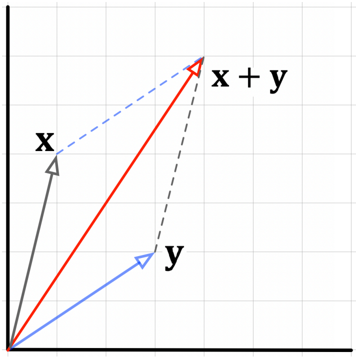


Figure 6.3 Addition of Vectors in $\mathbb{R}^2$.

Problem Draw the arrows of the vectors $\mathbf{x}-\mathbf{y}$, $\mathbf{x}+2\mathbf{y}$ and $3\mathbf{x}-2\mathbf{y}$.