

MATH-224 I: Linear Algebra

Vector Spaces

References: D. C. Lay , Linear Algebra and its Application, 4th ed.

A **vector space** is a nonempty set V of objects, called *vectors*, on which are defined two operations, called *addition* and *multiplication by scalars* (real numbers), subject to the ten axioms (or rules) listed below.¹ The axioms must hold for all vectors \mathbf{u} , \mathbf{v} , and \mathbf{w} in V and for all scalars c and d .

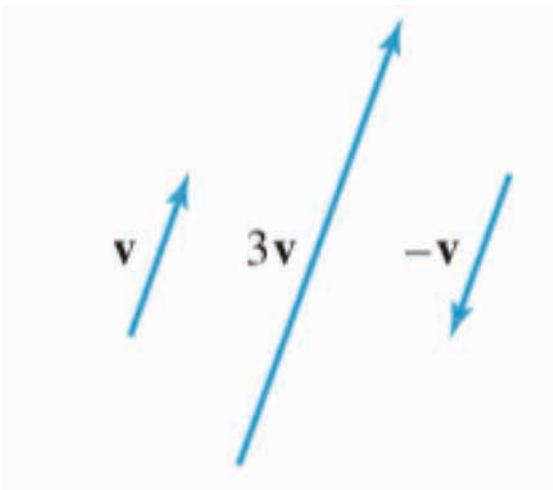
1. The sum of \mathbf{u} and \mathbf{v} , denoted by $\mathbf{u} + \mathbf{v}$, is in V .
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. There is a **zero** vector $\mathbf{0}$ in V such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. For each \mathbf{u} in V , there is a vector $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. The scalar multiple of \mathbf{u} by c , denoted by $c\mathbf{u}$, is in V .
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

For each \mathbf{u} in V and scalar c ,

$$0\mathbf{u} = \mathbf{0} \tag{1}$$

$$c\mathbf{0} = \mathbf{0} \tag{2}$$

$$-\mathbf{u} = (-1)\mathbf{u} \tag{3}$$



A **subspace** of a vector space V is a subset H of V that has three properties:

- The zero vector of V is in H .²
- H is closed under vector addition. That is, for each \mathbf{u} and \mathbf{v} in H , the sum $\mathbf{u} + \mathbf{v}$ is in H .
- H is closed under multiplication by scalars. That is, for each \mathbf{u} in H and each scalar c , the vector $c\mathbf{u}$ is in H .

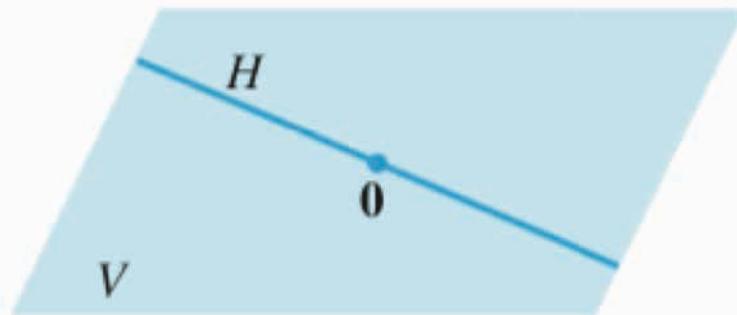


FIGURE 6
A subspace of V .

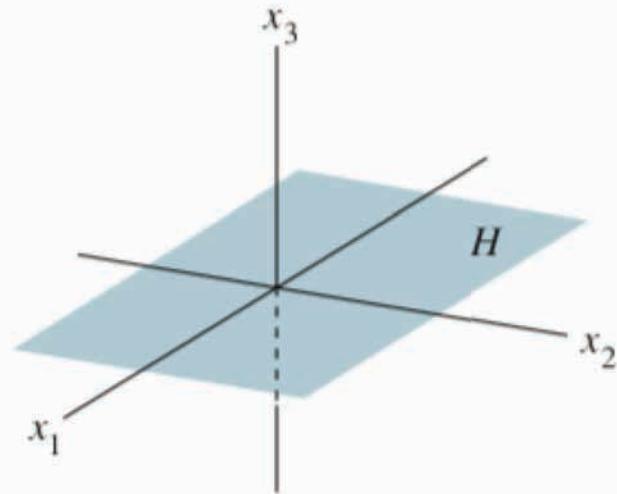


FIGURE 7
The x_1x_2 -plane as a subspace of \mathbb{R}^3 .

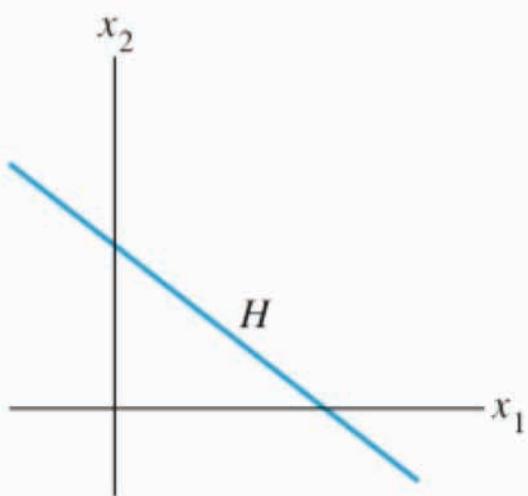


FIGURE 8
A line that is not a vector space.

A Subspace Spanned by a Set

The next example illustrates one of the most common ways of describing a subspace. As in Chapter 1, the term **linear combination** refers to any sum of scalar multiples of vectors, and $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ denotes the set of all vectors that can be written as linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_p$.

EXAMPLE 10 Given \mathbf{v}_1 and \mathbf{v}_2 in a vector space V , let $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Show that H is a subspace of V .

SOLUTION The zero vector is in H , since $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$. To show that H is closed under vector addition, take two arbitrary vectors in H , say,

$$\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 \quad \text{and} \quad \mathbf{w} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

By Axioms 2, 3, and 8 for the vector space V ,

$$\begin{aligned}\mathbf{u} + \mathbf{w} &= (s_1\mathbf{v}_1 + s_2\mathbf{v}_2) + (t_1\mathbf{v}_1 + t_2\mathbf{v}_2) \\ &= (s_1 + t_1)\mathbf{v}_1 + (s_2 + t_2)\mathbf{v}_2\end{aligned}$$

So $\mathbf{u} + \mathbf{w}$ is in H . Furthermore, if c is any scalar, then by Axioms 7 and 9,

$$c\mathbf{u} = c(s_1\mathbf{v}_1 + s_2\mathbf{v}_2) = (cs_1)\mathbf{v}_1 + (cs_2)\mathbf{v}_2$$

which shows that $c\mathbf{u}$ is in H and H is closed under scalar multiplication. Thus H is a subspace of V . ■

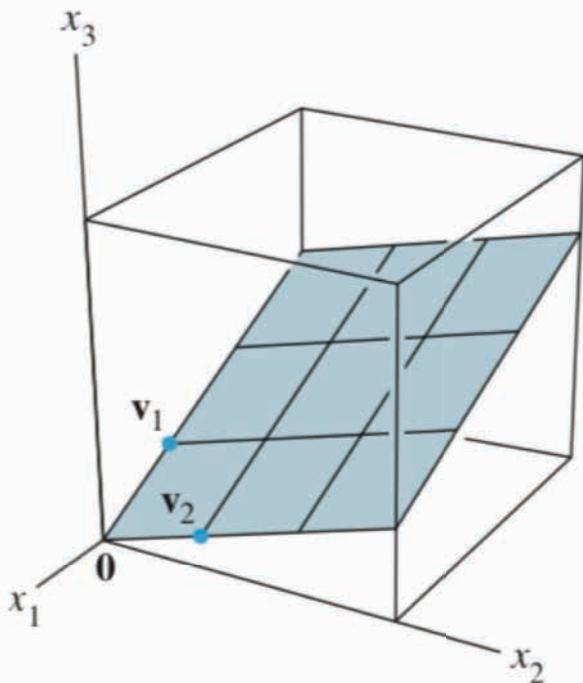


FIGURE 9

An example of a subspace.

THEOREM 1

If $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in a vector space V , then $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is a subspace of V .

We call $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ the **subspace spanned** (or **generated**) by $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$. Given any subspace H of V , a **spanning** (or **generating**) **set** for H is a set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in H such that $H = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$.

EXAMPLE 11 Let H be the set of all vectors of the form $(a - 3b, b - a, a, b)$, where a and b are arbitrary scalars. That is, let $H = \{(a - 3b, b - a, a, b) : a \text{ and } b \text{ in } \mathbb{R}\}$. Show that H is a subspace of \mathbb{R}^4 .

SOLUTION Write the vectors in H as column vectors. Then an arbitrary vector in H has the form

$$\begin{bmatrix} a - 3b \\ b - a \\ a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

\uparrow \uparrow
 \mathbf{v}_1 \mathbf{v}_2

This calculation shows that $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, where \mathbf{v}_1 and \mathbf{v}_2 are the vectors indicated above. Thus H is a subspace of \mathbb{R}^4 by Theorem 1. ■

The Null Space of a Matrix

Consider the following system of homogeneous equations:

$$\begin{aligned}x_1 - 3x_2 - 2x_3 &= 0 \\-5x_1 + 9x_2 + x_3 &= 0\end{aligned}$$

In matrix form, this system is written as $A\mathbf{x} = \mathbf{0}$, where

$$A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \tag{2}$$

Recall that the set of all \mathbf{x} that satisfy (1) is called the **solution set** of the system (1). Often it is convenient to relate this set directly to the matrix A and the equation $A\mathbf{x} = \mathbf{0}$. We call the set of \mathbf{x} that satisfy $A\mathbf{x} = \mathbf{0}$ the **null space** of the matrix A .

THEOREM 2

The null space of an $m \times n$ matrix A is a subspace of \mathbb{R}^n . Equivalently, the set of all solutions to a system $A\mathbf{x} = \mathbf{0}$ of m homogeneous linear equations in n unknowns is a subspace of \mathbb{R}^n .

PROOF Certainly $\text{Nul } A$ is a subset of \mathbb{R}^n because A has n columns. We must show that $\text{Nul } A$ satisfies the three properties of a subspace. Of course, $\mathbf{0}$ is in $\text{Nul } A$. Next, let \mathbf{u} and \mathbf{v} represent any two vectors in $\text{Nul } A$. Then

$$A\mathbf{u} = \mathbf{0} \quad \text{and} \quad A\mathbf{v} = \mathbf{0}$$

To show that $\mathbf{u} + \mathbf{v}$ is in $\text{Nul } A$, we must show that $A(\mathbf{u} + \mathbf{v}) = \mathbf{0}$. Using a property of matrix multiplication, compute

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Thus $\mathbf{u} + \mathbf{v}$ is in $\text{Nul } A$, and $\text{Nul } A$ is closed under vector addition. Finally, if c is any scalar, then

$$A(c\mathbf{u}) = c(A\mathbf{u}) = c(\mathbf{0}) = \mathbf{0}$$

which shows that $c\mathbf{u}$ is in $\text{Nul } A$. Thus $\text{Nul } A$ is a subspace of \mathbb{R}^n . ■

An Explicit Description of $\text{Nul } A$

There is no obvious relation between vectors in $\text{Nul } A$ and the entries in A . We say that $\text{Nul } A$ is defined *implicitly*, because it is defined by a condition that must be checked. No explicit list or description of the elements in $\text{Nul } A$ is given. However, *solving* the equation $A\mathbf{x} = \mathbf{0}$ amounts to producing an *explicit* description of $\text{Nul } A$. The next example reviews the procedure from Section 1.5.

EXAMPLE 3 Find a spanning set for the null space of the matrix

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

SOLUTION The first step is to find the general solution of $A\mathbf{x} = \mathbf{0}$ in terms of free variables. Row reduce the augmented matrix $[A \quad \mathbf{0}]$ to *reduced echelon form* in order to write the basic variables in terms of the free variables:

$$\left[\begin{array}{cccccc} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right], \quad \begin{aligned} x_1 - 2x_2 - x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

The general solution is $x_1 = 2x_2 + x_4 - 3x_5$, $x_3 = -2x_4 + 2x_5$, with x_2 , x_4 , and x_5 free. Next, decompose the vector giving the general solution into a linear combination of vectors where *the weights are the free variables*. That is,

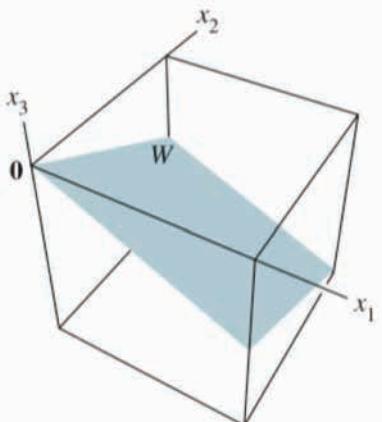
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

$\uparrow \mathbf{u}$ $\uparrow \mathbf{v}$ $\uparrow \mathbf{w}$

$$= x_2 \mathbf{u} + x_4 \mathbf{v} + x_5 \mathbf{w} \quad (3)$$

Every linear combination of \mathbf{u} , \mathbf{v} , and \mathbf{w} is an element of $\text{Nul } A$. Thus $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a spanning set for $\text{Nul } A$. ■

EXAMPLE 4 Find a matrix A such that $W = \text{Col } A$.



$$W = \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a, b \text{ in } \mathbb{R} \right\}$$

SOLUTION First, write W as a set of linear combinations.

$$W = \left\{ a \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} : a, b \text{ in } \mathbb{R} \right\} = \text{Span} \left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Second, use the vectors in the spanning set as the columns of A . Let $A = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ -7 & 0 \end{bmatrix}$.

Then $W = \text{Col } A$, as desired. ■

EXAMPLE 5 Let

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

- If the column space of A is a subspace of \mathbb{R}^k , what is k ?
- If the null space of A is a subspace of \mathbb{R}^k , what is k ?

EXAMPLE 6 With A as in Example 5, find a nonzero vector in $\text{Col } A$ and a nonzero vector in $\text{Nul } A$.

SOLUTION It is easy to find a vector in $\text{Col } A$. Any column of A will do, say, $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$.

To find a nonzero vector in $\text{Nul } A$, row reduce the augmented matrix $[A \quad \mathbf{0}]$ and obtain

$$[A \quad \mathbf{0}] \sim \begin{bmatrix} 1 & 0 & 9 & 0 & 0 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Thus, if \mathbf{x} satisfies $A\mathbf{x} = \mathbf{0}$, then $x_1 = -9x_3$, $x_2 = 5x_3$, $x_4 = 0$, and x_3 is free. Assigning a nonzero value to x_3 —say, $x_3 = 1$ —we obtain a vector in $\text{Nul } A$, namely, $\mathbf{x} = (-9, 5, 1, 0)$. ■

EXAMPLE 7 With A as in Example 5, let $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

- Determine if \mathbf{u} is in $\text{Nul } A$. Could \mathbf{u} be in $\text{Col } A$?
- Determine if \mathbf{v} is in $\text{Col } A$. Could \mathbf{v} be in $\text{Nul } A$?

Contrast Between Nul A and Col A for an $m \times n$ Matrix A

Nul A	Col A
1. Nul A is a subspace of \mathbb{R}^n .	1. Col A is a subspace of \mathbb{R}^m .
2. Nul A is implicitly defined; that is, you are given only a condition ($A\mathbf{x} = \mathbf{0}$) that vectors in Nul A must satisfy.	2. Col A is explicitly defined; that is, you are told how to build vectors in Col A .
3. It takes time to find vectors in Nul A . Row operations on $[A \quad \mathbf{0}]$ are required.	3. It is easy to find vectors in Col A . The columns of A are displayed; others are formed from them.
4. There is no obvious relation between Nul A and the entries in A .	4. There is an obvious relation between Col A and the entries in A , since each column of A is in Col A .
5. A typical vector \mathbf{v} in Nul A has the property that $A\mathbf{v} = \mathbf{0}$.	5. A typical vector \mathbf{v} in Col A has the property that the equation $A\mathbf{x} = \mathbf{v}$ is consistent.
6. Given a specific vector \mathbf{v} , it is easy to tell if \mathbf{v} is in Nul A . Just compute $A\mathbf{v}$.	6. Given a specific vector \mathbf{v} , it may take time to tell if \mathbf{v} is in Col A . Row operations on $[A \quad \mathbf{v}]$ are required.
7. Nul $A = \{\mathbf{0}\}$ if and only if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.	7. Col $A = \mathbb{R}^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every \mathbf{b} in \mathbb{R}^m .
8. Nul $A = \{\mathbf{0}\}$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.	8. Col $A = \mathbb{R}^m$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps \mathbb{R}^n onto \mathbb{R}^m .

Kernel and Range of a Linear Transformation

Subspaces of vector spaces other than \mathbb{R}^n are often described in terms of a linear transformation instead of a matrix. To make this precise, we generalize the definition given in Section 1.8.

A **linear transformation** T from a vector space V into a vector space W is a rule that assigns to each vector \mathbf{x} in V a unique vector $T(\mathbf{x})$ in W , such that

- (i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all \mathbf{u}, \mathbf{v} in V , and
- (ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for all \mathbf{u} in V and all scalars c .

The **kernel** (or **null space**) of such a T is the set of all \mathbf{u} in V such that $T(\mathbf{u}) = \mathbf{0}$ (the zero vector in W). The **range** of T is the set of all vectors in W of the form $T(\mathbf{x})$ for some \mathbf{x} in V . If T happens to arise as a matrix transformation—say, $T(\mathbf{x}) = A\mathbf{x}$ for some matrix A —then the kernel and the range of T are just the null space and the column space of A , as defined earlier.

It is not difficult to show that the kernel of T is a subspace of V . The proof is essentially the same as the one for Theorem 2. Also, the range of T is a subspace of W . See Fig. 2 and Exercise 30.

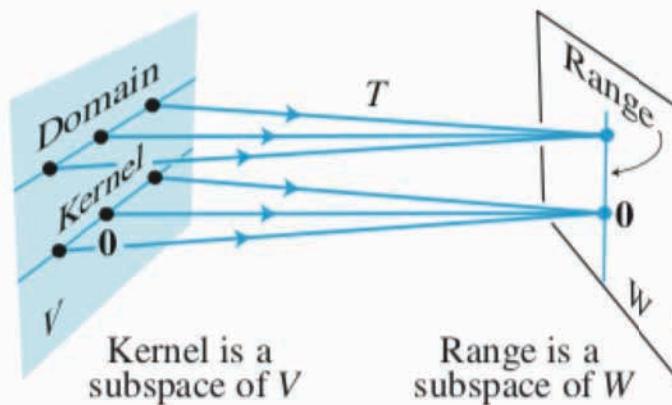


FIGURE 2 Subspaces associated with a linear transformation.

LINEARLY INDEPENDENT SETS; BASES

In this section we identify and study the subsets that span a vector space V or a subspace H as “efficiently” as possible. The key idea is that of linear independence, defined as in \mathbb{R}^n .

An indexed set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in V is said to be **linearly independent** if the vector equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \quad (1)$$

has *only* the trivial solution, $c_1 = 0, \dots, c_p = 0$.¹

The set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is said to be **linearly dependent** if (1) has a nontrivial solution, that is, if there are some weights, c_1, \dots, c_p , *not all zero*, such that (1) holds. In such a case, (1) is called a **linear dependence relation** among $\mathbf{v}_1, \dots, \mathbf{v}_p$.

Just as in \mathbb{R}^n , a set containing a single vector \mathbf{v} is linearly independent if and only if $\mathbf{v} \neq \mathbf{0}$. Also, a set of two vectors is linearly dependent if and only if one of the vectors is a multiple of the other. And any set containing the zero vector is linearly dependent. The following theorem has the same proof as Theorem 7 in Section 1.7.

An indexed set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ of two or more vectors, with $\mathbf{v}_1 \neq \mathbf{0}$, is linearly dependent if and only if some \mathbf{v}_j (with $j > 1$) is a linear combination of the preceding vectors, $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}$.

Let H be a subspace of a vector space V . An indexed set of vectors $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ in V is a **basis** for H if

- (i) \mathcal{B} is a linearly independent set, and
- (ii) the subspace spanned by \mathcal{B} coincides with H ; that is,

$$H = \text{Span}\{\mathbf{b}_1, \dots, \mathbf{b}_p\}$$

EXAMPLE 3 Let A be an invertible $n \times n$ matrix—say, $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$. Then the columns of A form a basis for \mathbb{R}^n because they are linearly independent and they span \mathbb{R}^n , by the Invertible Matrix Theorem. ■

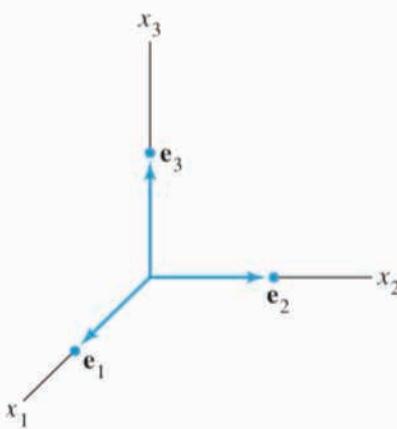


FIGURE 1

The standard basis for \mathbb{R}^3 .

EXAMPLE 4 Let $\mathbf{e}_1, \dots, \mathbf{e}_n$ be the columns of the $n \times n$ identity matrix, I_n . That is,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

The set $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is called the **standard basis** for \mathbb{R}^n (Fig. 1). ■

EXAMPLE 5 Let $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$. Determine if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for \mathbb{R}^3 .

SOLUTION Since there are exactly three vectors here in \mathbb{R}^3 , we can use any of several methods to determine if the matrix $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3]$ is invertible. For instance, two row replacements reveal that A has three pivot positions. Thus A is invertible. As in Example 3, the columns of A form a basis for \mathbb{R}^3 . ■

The Spanning Set Theorem

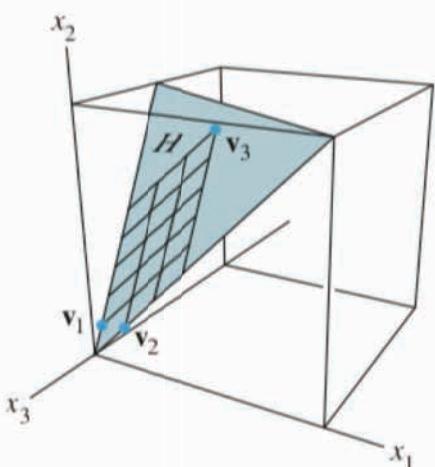
As we will see, a basis is an “efficient” spanning set that contains no unnecessary vectors. In fact, a basis can be constructed from a spanning set by discarding unneeded vectors.

EXAMPLE 7

Let

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 16 \\ -5 \end{bmatrix}, \quad \text{and} \quad H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}.$$

Note that $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, and show that $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Then find a basis for the subspace H .



SOLUTION Every vector in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ belongs to H because

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + 0\mathbf{v}_3$$

Now let \mathbf{x} be any vector in H —say, $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Since $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, we may substitute

$$\begin{aligned}\mathbf{x} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3(5\mathbf{v}_1 + 3\mathbf{v}_2) \\ &= (c_1 + 5c_3)\mathbf{v}_1 + (c_2 + 3c_3)\mathbf{v}_2\end{aligned}$$

Thus \mathbf{x} is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, so every vector in H already belongs to $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. We conclude that H and $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ are actually the same set of vectors. It follows that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of H since $\{\mathbf{v}_1, \mathbf{v}_2\}$ is obviously linearly independent. ■

The Spanning Set Theorem

Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ be a set in V , and let $H = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$.

- If one of the vectors in S —say, \mathbf{v}_k —is a linear combination of the remaining vectors in S , then the set formed from S by removing \mathbf{v}_k still spans H .
- If $H \neq \{\mathbf{0}\}$, some subset of S is a basis for H .

PROOF

- By rearranging the list of vectors in S , if necessary, we may suppose that \mathbf{v}_p is a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_{p-1}$ —say,

$$\mathbf{v}_p = a_1 \mathbf{v}_1 + \cdots + a_{p-1} \mathbf{v}_{p-1} \quad (3)$$

Given any \mathbf{x} in H , we may write

$$\mathbf{x} = c_1 \mathbf{v}_1 + \cdots + c_{p-1} \mathbf{v}_{p-1} + c_p \mathbf{v}_p \quad (4)$$

for suitable scalars c_1, \dots, c_p . Substituting the expression for \mathbf{v}_p from (3) into (4), it is easy to see that \mathbf{x} is a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_{p-1}$. Thus $\{\mathbf{v}_1, \dots, \mathbf{v}_{p-1}\}$ spans H , because \mathbf{x} was an arbitrary element of H .

- b. If the original spanning set S is linearly independent, then it is already a basis for H . Otherwise, one of the vectors in S depends on the others and can be deleted, by part (a). So long as there are two or more vectors in the spanning set, we can repeat this process until the spanning set is linearly independent and hence is a basis for H . If the spanning set is eventually reduced to one vector, that vector will be nonzero (and hence linearly independent) because $H \neq \{\mathbf{0}\}$. ■

EXAMPLE 8 Find a basis for $\text{Col } B$, where

$$B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_5] = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

SOLUTION Each nonpivot column of B is a linear combination of the pivot columns. In fact, $\mathbf{b}_2 = 4\mathbf{b}_1$ and $\mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3$. By the Spanning Set Theorem, we may discard \mathbf{b}_2 and \mathbf{b}_4 , and $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ will still span $\text{Col } B$. Let

$$S = \{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Since $\mathbf{b}_1 \neq 0$ and no vector in S is a linear combination of the vectors that precede it, S is linearly independent (Theorem 4). Thus S is a basis for $\text{Col } B$. ■

The pivot columns of a matrix A form a basis for $\text{Col } A$.

PROOF The general proof uses the arguments discussed above. Let B be the reduced echelon form of A . The set of pivot columns of B is linearly independent, for no vector in the set is a linear combination of the vectors that precede it. Since A is row equivalent to B , the pivot columns of A are linearly independent as well, because any linear dependence relation among the columns of A corresponds to a linear dependence relation among the columns of B . For this same reason, every nonpivot column of A is a linear combination of the pivot columns of A . Thus the nonpivot columns of A may be discarded from the spanning set for $\text{Col } A$, by the Spanning Set Theorem. This leaves the pivot columns of A as a basis for $\text{Col } A$. ■

COORDINATE SYSTEMS

An important reason for specifying a basis \mathcal{B} for a vector space V is to impose a “coordinate system” on V . This section will show that if \mathcal{B} contains n vectors, then the coordinate system will make V act like \mathbb{R}^n . If V is already \mathbb{R}^n itself, then \mathcal{B} will determine a coordinate system that gives a new “view” of V .

The existence of coordinate systems rests on the following fundamental result.

The Unique Representation Theorem

Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis for a vector space V . Then for each \mathbf{x} in V , there exists a unique set of scalars c_1, \dots, c_n such that

$$\mathbf{x} = c_1 \mathbf{b}_1 + \cdots + c_n \mathbf{b}_n \quad (1)$$

PROOF Since \mathcal{B} spans V , there exist scalars such that (1) holds. Suppose \mathbf{x} also has the representation

$$\mathbf{x} = d_1 \mathbf{b}_1 + \cdots + d_n \mathbf{b}_n$$

for scalars d_1, \dots, d_n . Then, subtracting, we have

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1) \mathbf{b}_1 + \cdots + (c_n - d_n) \mathbf{b}_n \quad (2)$$

Since \mathcal{B} is linearly independent, the weights in (2) must all be zero. That is, $c_j = d_j$ for $1 \leq j \leq n$. ■

Suppose $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis for V and \mathbf{x} is in V . The **coordinates of \mathbf{x} relative to the basis \mathcal{B}** (or the **\mathcal{B} -coordinates of \mathbf{x}**) are the weights c_1, \dots, c_n such that $\mathbf{x} = c_1\mathbf{b}_1 + \dots + c_n\mathbf{b}_n$.

If c_1, \dots, c_n are the \mathcal{B} -coordinates of \mathbf{x} , then the vector in \mathbb{R}^n

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

is the **coordinate vector of \mathbf{x} (relative to \mathcal{B})**, or the **\mathcal{B} -coordinate vector of \mathbf{x}** . The mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is the **coordinate mapping (determined by \mathcal{B})**.¹

EXAMPLE 1 Consider a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ for \mathbb{R}^2 , where $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Suppose an \mathbf{x} in \mathbb{R}^2 has the coordinate vector $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Find \mathbf{x} .

SOLUTION The \mathcal{B} -coordinates of \mathbf{x} tell how to build \mathbf{x} from the vectors in \mathcal{B} . That is,

$$\mathbf{x} = (-2)\mathbf{b}_1 + 3\mathbf{b}_2 = (-2)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$

EXAMPLE 2 The entries in the vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$ are the coordinates of \mathbf{x} relative to the *standard basis* $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, since

$$\begin{bmatrix} 1 \\ 6 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 6 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{e}_1 + 6 \cdot \mathbf{e}_2$$

If $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, then $[\mathbf{x}]_{\mathcal{E}} = \mathbf{x}$.

A Graphical Interpretation of Coordinates

A coordinate system on a set consists of a one-to-one mapping of the points in the set into \mathbb{R}^n . For example, ordinary graph paper provides a coordinate system for the plane when one selects perpendicular axes and a unit of measurement on each axis. Figure 1 shows the standard basis $\{\mathbf{e}_1, \mathbf{e}_2\}$, the vectors $\mathbf{b}_1 (= \mathbf{e}_1)$ and \mathbf{b}_2 from Example 1, and the vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$. The coordinates 1 and 6 give the location of \mathbf{x} relative to the standard basis: 1 unit in the \mathbf{e}_1 direction and 6 units in the \mathbf{e}_2 direction.

Figure 2 shows the vectors \mathbf{b}_1 , \mathbf{b}_2 , and \mathbf{x} from Fig. 1. (Geometrically, the three vectors lie on a vertical line in both figures.) However, the standard coordinate grid was erased and replaced by a grid especially adapted to the basis \mathcal{B} in Example 1. The coordinate vector $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ gives the location of \mathbf{x} on this new coordinate system: -2 units in the \mathbf{b}_1 direction and 3 units in the \mathbf{b}_2 direction.

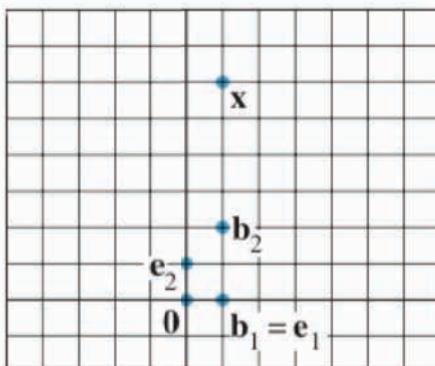


FIGURE 1 Standard graph paper.

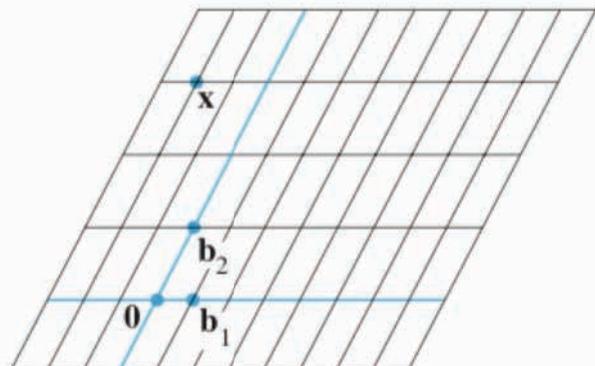


FIGURE 2 \mathcal{B} -graph paper.

Coordinates in \mathbb{R}^n

When a basis \mathcal{B} for \mathbb{R}^n is fixed, the \mathcal{B} -coordinate vector of a specified \mathbf{x} is easily found, as in the next example.

EXAMPLE 4 Let $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Find the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ of \mathbf{x} relative to \mathcal{B} .

SOLUTION The \mathcal{B} -coordinates c_1, c_2 of \mathbf{x} satisfy

$$c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$\mathbf{b}_1 \qquad \mathbf{b}_2 \qquad \mathbf{x}$

or

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \tag{3}$$

$\mathbf{b}_1 \quad \mathbf{b}_2 \qquad \mathbf{x}$

This equation can be solved by row operations on an augmented matrix or by using the inverse of the matrix on the left. In any case, the solution is $c_1 = 3$, $c_2 = 2$. Thus $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2$, and

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

See Fig. 4.

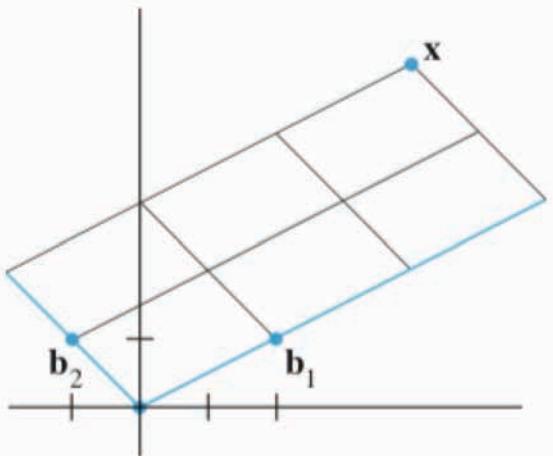


FIGURE 4

The \mathcal{B} -coordinate vector of x is $(3, 2)$.

If a vector space V has a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, then any set in V containing more than n vectors must be linearly dependent.

PROOF Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be a set in V with more than n vectors. The coordinate vectors $[\mathbf{u}_1]_{\mathcal{B}}, \dots, [\mathbf{u}_p]_{\mathcal{B}}$ form a linearly dependent set in \mathbb{R}^n , because there are more vectors (p) than entries (n) in each vector. So there exist scalars c_1, \dots, c_p , not all zero, such that

$$c_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_p[\mathbf{u}_p]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{The zero vector in } \mathbb{R}^n$$

Since the coordinate mapping is a linear transformation,

$$[c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

The zero vector on the right displays the n weights needed to build the vector $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ from the basis vectors in \mathcal{B} . That is, $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = 0 \cdot \mathbf{b}_1 + \cdots + 0 \cdot \mathbf{b}_n = \mathbf{0}$. Since the c_i are not all zero, $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is linearly dependent.¹ ■

If a vector space V has a basis of n vectors, then every basis of V must consist of exactly n vectors.

PROOF Let \mathcal{B}_1 be a basis of n vectors and \mathcal{B}_2 be any other basis (of V). Since \mathcal{B}_1 is a basis and \mathcal{B}_2 is linearly independent, \mathcal{B}_2 has no more than n vectors, by Theorem 9. Also, since \mathcal{B}_2 is a basis and \mathcal{B}_1 is linearly independent, \mathcal{B}_2 has at least n vectors. Thus \mathcal{B}_2 consists of exactly n vectors. ■

If V is spanned by a finite set, then V is said to be **finite-dimensional**, and the **dimension** of V , written as $\dim V$, is the number of vectors in a basis for V . The dimension of the zero vector space $\{\mathbf{0}\}$ is defined to be zero. If V is not spanned by a finite set, then V is said to be **infinite-dimensional**.

EXAMPLE 1 The standard basis for \mathbb{R}^n contains n vectors, so $\dim \mathbb{R}^n = n$. The standard polynomial basis $\{1, t, t^2\}$ shows that $\dim \mathbb{P}_2 = 3$. In general, $\dim \mathbb{P}_n = n + 1$. The space \mathbb{P} of all polynomials is infinite-dimensional (Exercise 27). ■

EXAMPLE 2 Let $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. Then H is the plane studied in Example 7 in Section 4.4. A basis for H is $\{\mathbf{v}_1, \mathbf{v}_2\}$, since \mathbf{v}_1 and \mathbf{v}_2 are not multiples and hence are linearly independent. Thus $\dim H = 2$. ■

Rank

With the aid of vector space concepts, this section takes a look *inside* a matrix and reveals several interesting and useful relationships hidden in its rows and columns.

For instance, imagine placing 2000 random numbers into a 40×50 matrix A and then determining both the maximum number of linearly independent columns in A and the maximum number of linearly independent columns in A^T (rows in A). Remarkably, the two numbers are the same. As we'll soon see, their common value is the *rank* of the matrix. To explain why, we need to examine the subspace spanned by the rows of A .

The Row Space

If A is an $m \times n$ matrix, each row of A has n entries and thus can be identified with a vector in \mathbb{R}^n . The set of all linear combinations of the row vectors is called the **row space** of A and is denoted by $\text{Row } A$. Each row has n entries, so $\text{Row } A$ is a subspace of \mathbb{R}^n . Since the rows of A are identified with the columns of A^T , we could also write $\text{Col } A^T$ in place of $\text{Row } A$.

EXAMPLE 1 Let

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix} \quad \text{and} \quad \begin{aligned} \mathbf{r}_1 &= (-2, -5, 8, 0, -17) \\ \mathbf{r}_2 &= (1, 3, -5, 1, 5) \\ \mathbf{r}_3 &= (3, 11, -19, 7, 1) \\ \mathbf{r}_4 &= (1, 7, -13, 5, -3) \end{aligned}$$

The row space of A is the subspace of \mathbb{R}^5 spanned by $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. That is, $\text{Row } A = \text{Span } \{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. It is natural to write row vectors horizontally; however, they may also be written as column vectors if that is more convenient. ■

EXAMPLE 2 Find bases for the row space, the column space, and the null space of the matrix

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$$

SOLUTION To find bases for the row space and the column space, row reduce A to an echelon form:

$$A \sim B = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

By Theorem 13, the first three rows of B form a basis for the row space of A (as well as for the row space of B). Thus

Basis for Row A : $\{(1, 3, -5, 1, 5), (0, 1, -2, 2, -7), (0, 0, 0, -4, 20)\}$

For the column space, observe from B that the pivots are in columns 1, 2, and 4. Hence columns 1, 2, and 4 of A (not B) form a basis for Col A :

$$\text{Basis for Col } A: \left\{ \begin{bmatrix} -2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 11 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \\ 5 \end{bmatrix} \right\}$$

Notice that any echelon form of A provides (in its nonzero rows) a basis for Row A and also identifies the pivot columns of A for Col A . However, for Nul A , we need the *reduced echelon form*. Further row operations on B yield

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

The equation $A\mathbf{x} = \mathbf{0}$ is equivalent to $C\mathbf{x} = \mathbf{0}$, that is,

$$\begin{aligned} x_1 + x_3 + x_5 &= 0 \\ x_2 - 2x_3 + 3x_5 &= 0 \\ x_4 - 5x_5 &= 0 \end{aligned}$$

So $x_1 = -x_3 - x_5$, $x_2 = 2x_3 - 3x_5$, $x_4 = 5x_5$, with x_3 and x_5 free variables. The usual calculations (discussed in Section 4.2) show that

$$\text{Basis for Nul } A: \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} \right\}$$

Observe that, unlike the basis for Col A , the bases for Row A and Nul A have no simple connection with the entries in A itself.¹

The **rank** of A is the dimension of the column space of A .

The Rank Theorem

The dimensions of the column space and the row space of an $m \times n$ matrix A are equal. This common dimension, the rank of A , also equals the number of pivot positions in A and satisfies the equation

$$\text{rank } A + \dim \text{Nul } A = n$$

PROOF By Theorem 6 in Section 4.3, $\text{rank } A$ is the number of pivot columns in A . Equivalently, $\text{rank } A$ is the number of pivot positions in an echelon form B of A . Furthermore, since B has a nonzero row for each pivot, and since these rows form a basis for the row space of A , the rank of A is also the dimension of the row space.

From Section 4.5, the dimension of $\text{Nul } A$ equals the number of free variables in the equation $A\mathbf{x} = \mathbf{0}$. Expressed another way, the dimension of $\text{Nul } A$ is the number of columns of A that are *not* pivot columns. (It is the number of these columns, not the columns themselves, that is related to $\text{Nul } A$.) Obviously,

$$\left\{ \begin{array}{l} \text{number of} \\ \text{pivot columns} \end{array} \right\} + \left\{ \begin{array}{l} \text{number of} \\ \text{nonpivot columns} \end{array} \right\} = \left\{ \begin{array}{l} \text{number of} \\ \text{columns} \end{array} \right\}$$

This proves the theorem. ■

The **rank** of A is the dimension of the column space of A .

The Rank Theorem

The dimensions of the column space and the row space of an $m \times n$ matrix A are equal. This common dimension, the rank of A , also equals the number of pivot positions in A and satisfies the equation

$$\text{rank } A + \dim \text{Nul } A = n$$

PROOF By Theorem 6 in Section 4.3, $\text{rank } A$ is the number of pivot columns in A . Equivalently, $\text{rank } A$ is the number of pivot positions in an echelon form B of A . Furthermore, since B has a nonzero row for each pivot, and since these rows form a basis for the row space of A , the rank of A is also the dimension of the row space.

From Section 4.5, the dimension of $\text{Nul } A$ equals the number of free variables in the equation $Ax = \mathbf{0}$. Expressed another way, the dimension of $\text{Nul } A$ is the number of columns of A that are *not* pivot columns. (It is the number of these columns, not the columns themselves, that is related to $\text{Nul } A$.) Obviously,

$$\left\{ \begin{array}{l} \text{number of} \\ \text{pivot columns} \end{array} \right\} + \left\{ \begin{array}{l} \text{number of} \\ \text{nonpivot columns} \end{array} \right\} = \left\{ \begin{array}{l} \text{number of} \\ \text{columns} \end{array} \right\}$$

This proves the theorem. ■

CHANGE OF BASIS

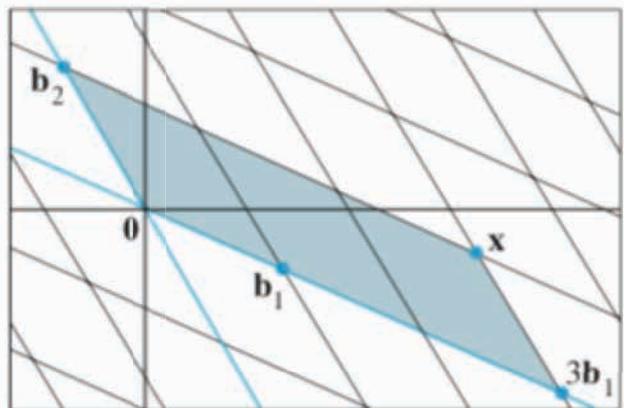
When a basis \mathcal{B} is chosen for an n -dimensional vector space V , the associated coordinate mapping onto \mathbb{R}^n provides a coordinate system for V . Each \mathbf{x} in V is identified uniquely by its \mathcal{B} -coordinate vector $[\mathbf{x}]_{\mathcal{B}}$.¹

In some applications, a problem is described initially using a basis \mathcal{B} , but the problem's solution is aided by changing \mathcal{B} to a new basis \mathcal{C} . (Examples will be given in Chapters 5 and 7.) Each vector is assigned a new \mathcal{C} -coordinate vector. In this section, we study how $[\mathbf{x}]_{\mathcal{C}}$ and $[\mathbf{x}]_{\mathcal{B}}$ are related for each \mathbf{x} in V .

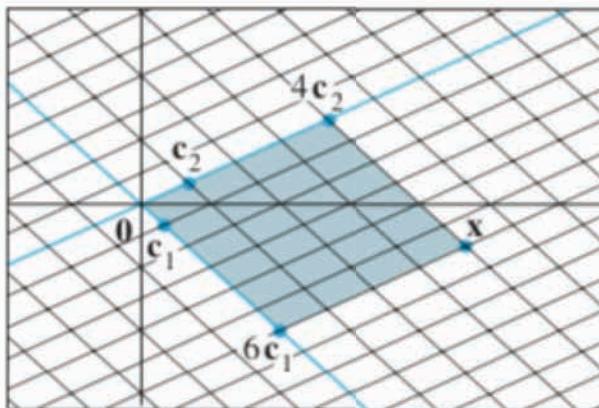
To visualize the problem, consider the two coordinate systems in Fig. 1. In Fig. 1(a), $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2$, while in Fig. 1(b), the same \mathbf{x} is shown as $\mathbf{x} = 6\mathbf{c}_1 + 4\mathbf{c}_2$. That is,

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad [\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Our problem is to find the connection between the two coordinate vectors. Example 1 shows how to do this, provided we know how \mathbf{b}_1 and \mathbf{b}_2 are formed from \mathbf{c}_1 and \mathbf{c}_2 .



(a)



(b)

FIGURE 1 Two coordinate systems for the same vector space.

EXAMPLE 1 Consider two bases $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ for a vector space V , such that

$$\mathbf{b}_1 = 4\mathbf{c}_1 + \mathbf{c}_2 \quad \text{and} \quad \mathbf{b}_2 = -6\mathbf{c}_1 + \mathbf{c}_2 \quad (1)$$

Suppose

$$\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2 \quad (2)$$

That is, suppose $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Find $[\mathbf{x}]_{\mathcal{C}}$.

SOLUTION Apply the coordinate mapping determined by \mathcal{C} to \mathbf{x} in (2). Since the coordinate mapping is a linear transformation,

$$\begin{aligned} [\mathbf{x}]_{\mathcal{C}} &= [3\mathbf{b}_1 + \mathbf{b}_2]_{\mathcal{C}} \\ &= 3[\mathbf{b}_1]_{\mathcal{C}} + [\mathbf{b}_2]_{\mathcal{C}} \end{aligned}$$

We can write this vector equation as a matrix equation, using the vectors in the linear combination as the columns of a matrix:

$$[\mathbf{x}]_{\mathcal{C}} = \left[[\mathbf{b}_1]_{\mathcal{C}} \quad [\mathbf{b}_2]_{\mathcal{C}} \right] \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad (3)$$

This formula gives $[\mathbf{x}]_{\mathcal{C}}$, once we know the columns of the matrix. From (1),

$$[\mathbf{b}_1]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad \text{and} \quad [\mathbf{b}_2]_{\mathcal{C}} = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$$

Thus (3) provides the solution:

$$[\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

The \mathcal{C} -coordinates of \mathbf{x} match those of the \mathbf{x} in Fig. 1.



Theorem

Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of a vector space V . Then there is a unique $n \times n$ matrix ${}_{\mathcal{C}}P_{\mathcal{B}}$ such that

$$[\mathbf{x}]_{\mathcal{C}} = {}_{\mathcal{C}}P_{\mathcal{B}} [\mathbf{x}]_{\mathcal{B}} \quad (4)$$

The columns of ${}_{\mathcal{C}}P_{\mathcal{B}}$ are the \mathcal{C} -coordinate vectors of the vectors in the basis \mathcal{B} . That is,

$${}_{\mathcal{C}}P_{\mathcal{B}} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} & [\mathbf{b}_2]_{\mathcal{C}} & \cdots & [\mathbf{b}_n]_{\mathcal{C}} \end{bmatrix} \quad (5)$$

The matrix ${}_{\mathcal{C} \leftarrow \mathcal{B}}^P$ in Theorem 15 is called the **change-of-coordinates matrix from \mathcal{B} to \mathcal{C}** . Multiplication by ${}_{\mathcal{C} \leftarrow \mathcal{B}}^P$ converts \mathcal{B} -coordinates into \mathcal{C} -coordinates.² Figure 2 illustrates the change-of-coordinates equation (4).

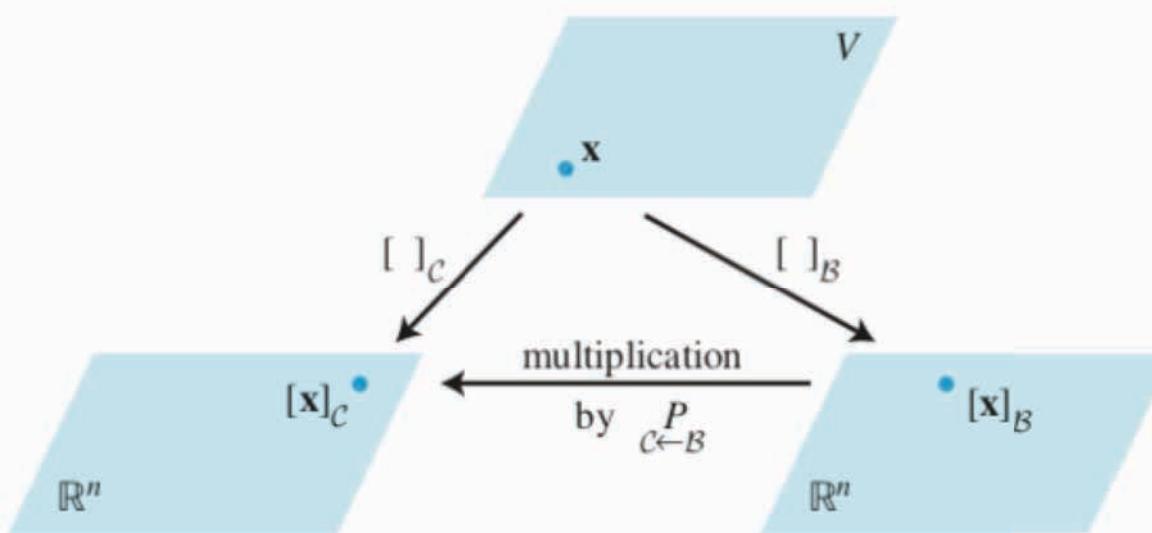


FIGURE 2 Two coordinate systems for V .

The columns of ${}_{\mathcal{C} \leftarrow \mathcal{B}}^P$ are linearly independent because they are the coordinate vectors of the linearly independent set \mathcal{B} . (See Exercise 25 in Section 4.4.) Since ${}_{\mathcal{C} \leftarrow \mathcal{B}}^P$ is square, it must be invertible, by the Invertible Matrix Theorem. Left-multiplying both sides of equation (4) by $({}_{\mathcal{C} \leftarrow \mathcal{B}}^P)^{-1}$ yields

$$({}_{\mathcal{C} \leftarrow \mathcal{B}}^P)^{-1}[x]_C = [x]_B$$

Thus $(_{\mathcal{C}} \xleftarrow{\cdot} \mathcal{B})^{-1}$ is the matrix that converts \mathcal{C} -coordinates into \mathcal{B} -coordinates. That is,

$$({}_{\mathcal{C}} \xleftarrow{\cdot} \mathcal{B})^{-1} = {}_{\mathcal{B}} \xleftarrow{\cdot} \mathcal{C} \quad (6)$$

Change of Basis in \mathbb{R}^n

If $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and \mathcal{E} is the *standard basis* $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ in \mathbb{R}^n , then $[\mathbf{b}_1]_{\mathcal{E}} = \mathbf{b}_1$, and likewise for the other vectors in \mathcal{B} . In this case, ${}_{\mathcal{E}} \xleftarrow{\cdot} \mathcal{B}$ is the same as the change-of-coordinates matrix $P_{\mathcal{B}}$ introduced in Section 4.4, namely,

$$P_{\mathcal{B}} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n]$$

To change coordinates between two nonstandard bases in \mathbb{R}^n , we need Theorem 15. The theorem shows that to solve the change-of-basis problem, we need the coordinate vectors of the old basis relative to the new basis.

Change of Basis in \mathbb{R}^n

If $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and \mathcal{E} is the *standard basis* $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ in \mathbb{R}^n , then $[\mathbf{b}_1]_{\mathcal{E}} = \mathbf{b}_1$, and likewise for the other vectors in \mathcal{B} . In this case, ${}_{\mathcal{E} \leftarrow \mathcal{B}} P$ is the same as the change-of-coordinates matrix $P_{\mathcal{B}}$ introduced in Section 4.4, namely,

$$P_{\mathcal{B}} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n]$$

To change coordinates between two nonstandard bases in \mathbb{R}^n , we need Theorem 15. The theorem shows that to solve the change-of-basis problem, we need the coordinate vectors of the old basis relative to the new basis.

EXAMPLE 2 Let $\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, and consider the bases for \mathbb{R}^2 given by $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$. Find the change-of-coordinates matrix from \mathcal{B} to \mathcal{C} .

SOLUTION The matrix ${}_{\mathcal{C}} \leftarrow {}_{\mathcal{B}} P$ involves the \mathcal{C} -coordinate vectors of \mathbf{b}_1 and \mathbf{b}_2 . Let $[\mathbf{b}_1]_{\mathcal{C}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $[\mathbf{b}_2]_{\mathcal{C}} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Then, by definition,

$$[\mathbf{c}_1 \quad \mathbf{c}_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}_1 \quad \text{and} \quad [\mathbf{c}_1 \quad \mathbf{c}_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{b}_2$$

To solve both systems simultaneously, augment the coefficient matrix with \mathbf{b}_1 and \mathbf{b}_2 , and row reduce:

$$[\mathbf{c}_1 \quad \mathbf{c}_2 \mid \mathbf{b}_1 \quad \mathbf{b}_2] = \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array} \right] \sim \left[\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array} \right] \quad (7)$$

Thus

$$[\mathbf{b}_1]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \quad \text{and} \quad [\mathbf{b}_2]_{\mathcal{C}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

EXAMPLE 3 Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$, and consider the bases for \mathbb{R}^2 given by $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$.

- Find the change-of-coordinates matrix from \mathcal{C} to \mathcal{B} .
- Find the change-of-coordinates matrix from \mathcal{B} to \mathcal{C} .

SOLUTION

- Notice that ${}_{\mathcal{B} \leftarrow \mathcal{C}}^P$ is needed rather than ${}_{\mathcal{C} \leftarrow \mathcal{B}}^P$, and compute

$$[\mathbf{b}_1 \quad \mathbf{b}_2 \mid \mathbf{c}_1 \quad \mathbf{c}_2] = \left[\begin{array}{cc|cc} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array} \right] \sim \left[\begin{array}{cc|cc} 1 & 0 & 5 & 3 \\ 0 & 1 & 6 & 4 \end{array} \right]$$

So

$${}_{\mathcal{B} \leftarrow \mathcal{C}}^P = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}$$

- By part (a) and property (6) above (with \mathcal{B} and \mathcal{C} interchanged),

$${}_{\mathcal{C} \leftarrow \mathcal{B}}^P = ({}_{\mathcal{B} \leftarrow \mathcal{C}}^P)^{-1} = \frac{1}{2} \begin{bmatrix} 4 & -3 \\ -6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix} \blacksquare$$

