# 5.2 Independence and Dimension

Some spanning sets are better than others. If $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is a subspace of $\mathbb{R}^n$, then every vector in $U$ can be written as a linear combination of the $\mathbf{x}_i$ in at least one way. Our interest here is in spanning sets where each vector in $U$ has a *exactly one* representation as a linear combination of these vectors.

## Linear Independence

Given $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ in $\mathbb{R}^n$, suppose that two linear combinations are equal:

$$r_1\mathbf{x}_1 + r_2\mathbf{x}_2 + \cdots + r_k\mathbf{x}_k = s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + \cdots + s_k\mathbf{x}_k$$

We are looking for a condition on the set $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ of vectors that guarantees that this representation is *unique*; that is, $r_i = s_i$ for each $i$. Taking all terms to the left side gives

$$(r_1 - s_1)\mathbf{x}_1 + (r_2 - s_2)\mathbf{x}_2 + \cdots + (r_k - s_k)\mathbf{x}_k = \mathbf{0}$$

so the required condition is that this equation forces all the coefficients $r_i - s_i$ to be zero.

**Definition 5.3 Linear Independence in $\mathbb{R}^n$**

*With this in mind, we call a set $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ of vectors **linearly independent** (or simply **independent**) if it satisfies the following condition:*

$$\text{If } t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0} \text{ then } t_1 = t_2 = \cdots = t_k = 0$$

We record the result of the above discussion for reference.

**Theorem 5.2.1**

*If $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is an independent set of vectors in $\mathbb{R}^n$, then every vector in $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ has a **unique** representation as a linear combination of the $\mathbf{x}_i$.*

It is useful to state the definition of independence in different language. Let us say that a linear combination **vanishes** if it equals the zero vector, and call a linear combination **trivial** if every coefficient is zero. Then the definition of independence can be compactly stated as follows:

> A set of vectors is independent if and only if the only linear combination that vanishes is the trivial one.

Hence we have a procedure for checking that a set of vectors is independent:

**Independence Test**

*To verify that a set* $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ *of vectors in* $\mathbb{R}^n$ *is independent, proceed as follows:*

1. *Set a linear combination equal to zero:* $t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0}$.
2. *Show that* $t_i = 0$ *for each* $i$ *(that is, the linear combination is trivial).*

*Of course, if some nontrivial linear combination vanishes, the vectors are not independent.*

**Example 5.2.1**

Determine whether $\{(1, 0, -2, 5), (2, 1, 0, -1), (1, 1, 2, 1)\}$ is independent in $\mathbb{R}^4$.

**Solution.** Suppose a linear combination vanishes:

$$r(1, 0, -2, 5) + s(2, 1, 0, -1) + t(1, 1, 2, 1) = (0, 0, 0, 0)$$

Equating corresponding entries gives a system of four equations:

$$r + 2s + t = 0, \; s + t = 0, \; -2r + 2t = 0, \; \text{and } 5r - s + t = 0$$

The only solution is the trivial one $r = s = t = 0$ (verify), so these vectors are independent by the independence test.

**Example 5.2.2**

Show that the standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of $\mathbb{R}^n$ is independent.

**Solution.** The components of $t_1\mathbf{e}_1 + t_2\mathbf{e}_2 + \cdots + t_n\mathbf{e}_n$ are $t_1, t_2, \dots, t_n$ (see the discussion preceding Example 5.1.6) So the linear combination vanishes if and only if each $t_i = 0$. Hence the independence test applies.

**Example 5.2.3**

If $\{\mathbf{x}, \mathbf{y}\}$ is independent, show that $\{2\mathbf{x} + 3\mathbf{y}, \mathbf{x} - 5\mathbf{y}\}$ is also independent.

**Solution.** If $s(2\mathbf{x} + 3\mathbf{y}) + t(\mathbf{x} - 5\mathbf{y}) = \mathbf{0}$, collect terms to get $(2s + t)\mathbf{x} + (3s - 5t)\mathbf{y} = \mathbf{0}$. Since $\{\mathbf{x}, \mathbf{y}\}$ is independent this combination must be trivial; that is, $2s + t = 0$ and $3s - 5t = 0$. These equations have only the trivial solution $s = t = 0$, as required.

**Example 5.2.4**

Show that the zero vector in $\mathbb{R}^n$ does not belong to any independent set.

**Solution.** No set $\{\mathbf{0},\ \mathbf{x}_1,\ \mathbf{x}_2,\ \dots,\ \mathbf{x}_k\}$ of vectors is independent because we have a vanishing, nontrivial linear combination $1 \cdot \mathbf{0} + 0\mathbf{x}_1 + 0\mathbf{x}_2 + \cdots + 0\mathbf{x}_k = \mathbf{0}$.

**Example 5.2.5**

Given $\mathbf{x}$ in $\mathbb{R}^n$, show that $\{\mathbf{x}\}$ is independent if and only if $\mathbf{x} \neq \mathbf{0}$.

**Solution.** A vanishing linear combination from $\{\mathbf{x}\}$ takes the form $t\mathbf{x} = \mathbf{0}$, $t$ in $\mathbb{R}$. This implies that $t = 0$ because $\mathbf{x} \neq \mathbf{0}$.

The next example will be needed later.

**Example 5.2.6**

Show that the nonzero rows of a row-echelon matrix $R$ are independent.

**Solution.** We illustrate the case with 3 leading 1s; the general case is analogous. Suppose $R$ has the form $R = \left[ \begin{array}{cccccc} 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$ where $*$ indicates a nonspecified number. Let $R_1$, $R_2$, and $R_3$ denote the nonzero rows of $R$. If $t_1R_1 + t_2R_2 + t_3R_3 = 0$ we show that $t_1 = 0$, then $t_2 = 0$, and finally $t_3 = 0$. The condition $t_1R_1 + t_2R_2 + t_3R_3 = 0$ becomes

$$(0,\ t_1,\ *,\ *,\ *,\ *) + (0,\ 0,\ 0,\ t_2,\ *,\ *) + (0,\ 0,\ 0,\ 0,\ t_3,\ *) = (0,\ 0,\ 0,\ 0,\ 0,\ 0)$$

Equating second entries show that $t_1 = 0$, so the condition becomes $t_2R_2 + t_3R_3 = 0$. Now the same argument shows that $t_2 = 0$. Finally, this gives $t_3R_3 = 0$ and we obtain $t_3 = 0$.

A set of vectors in $\mathbb{R}^n$ is called **linearly dependent** (or simply **dependent**) if it is *not* linearly independent, equivalently if some nontrivial linear combination vanishes.

**Example 5.2.7**

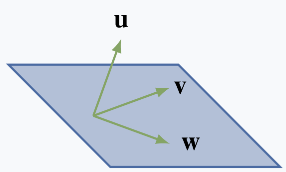
If $\mathbf{v}$ and $\mathbf{w}$ are nonzero vectors in $\mathbb{R}^3$, show that $\{\mathbf{v},\ \mathbf{w}\}$ is dependent if and only if $\mathbf{v}$ and $\mathbf{w}$ are parallel.

**Solution.** If $\mathbf{v}$ and $\mathbf{w}$ are parallel, then one is a scalar multiple of the other (Theorem 4.1.4), say $\mathbf{v} = a\mathbf{w}$ for some scalar $a$. Then the nontrivial linear combination $\mathbf{v} - a\mathbf{w} = \mathbf{0}$ vanishes, so $\{\mathbf{v},\ \mathbf{w}\}$ is dependent.

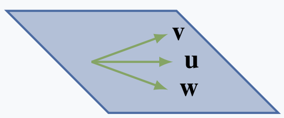
Conversely, if $\{\mathbf{v},\ \mathbf{w}\}$ is dependent, let $s\mathbf{v} + t\mathbf{w} = \mathbf{0}$ be nontrivial, say $s \neq 0$. Then $\mathbf{v} = -\frac{t}{s}\mathbf{w}$ so $\mathbf{v}$ and $\mathbf{w}$ are parallel (by Theorem 4.1.4). A similar argument works if $t \neq 0$.

With this we can give a geometric description of what it means for a set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ in $\mathbb{R}^3$ to be independent. Note that this requirement means that $\{\mathbf{v}, \mathbf{w}\}$ is also independent ($a\mathbf{v}+b\mathbf{w}=\mathbf{0}$ means that $0\mathbf{u}+a\mathbf{v}+b\mathbf{w}=\mathbf{0}$), so $M = \text{span}\,\{\mathbf{v}, \mathbf{w}\}$ is the plane containing $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{0}$ (see the discussion preceding Example 5.1.4). So we assume that $\{\mathbf{v}, \mathbf{w}\}$ is independent in the following example.

**Example 5.2.8**



$M$
$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ independent



$M$
$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ not independent

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be nonzero vectors in $\mathbb{R}^3$ where $\{\mathbf{v}, \mathbf{w}\}$ independent. Show that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is independent if and only if $\mathbf{u}$ is not in the plane $M = \text{span}\,\{\mathbf{v}, \mathbf{w}\}$. This is illustrated in the diagrams.

**Solution.** If $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is independent, suppose $\mathbf{u}$ is in the plane $M = \text{span}\,\{\mathbf{v}, \mathbf{w}\}$, say $\mathbf{u} = a\mathbf{v}+b\mathbf{w}$, where $a$ and $b$ are in $\mathbb{R}$. Then $1\mathbf{u}-a\mathbf{v}-b\mathbf{w}=\mathbf{0}$, contradicting the independence of $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$. On the other hand, suppose that $\mathbf{u}$ is not in $M$; we must show that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is independent. If $r\mathbf{u}+s\mathbf{v}+t\mathbf{w}=\mathbf{0}$ where $r$, $s$, and $t$ are in $\mathbb{R}^3$, then $r=0$ since otherwise $\mathbf{u} = -\frac{s}{r}\mathbf{v}+\frac{-t}{r}\mathbf{w}$ is in $M$. But then $s\mathbf{v}+t\mathbf{w}=\mathbf{0}$, so $s=t=0$ by our assumption. This shows that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is independent, as required.

By the inverse theorem, the following conditions are equivalent for an $n\times n$ matrix $A$:

1. *$A$ is invertible.*
2. *If $A\mathbf{x}=\mathbf{0}$ where $\mathbf{x}$ is in $\mathbb{R}^n$, then $\mathbf{x}=\mathbf{0}$.*
3. *$A\mathbf{x}=\mathbf{b}$ has a solution $\mathbf{x}$ for every vector $\mathbf{b}$ in $\mathbb{R}^n$.*

While condition 1 makes no sense if $A$ is not square, conditions 2 and 3 are meaningful for any matrix $A$ and, in fact, are related to independence and spanning. Indeed, if $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$ are the columns of $A$, and if we write $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then

$$A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n$$

by Definition 2.5. Hence the definitions of independence and spanning show, respectively, that condition 2 is equivalent to the independence of $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ and condition 3 is equivalent to the requirement that $\text{span}\,\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\} = \mathbb{R}^m$. This discussion is summarized in the following theorem:

**Theorem 5.2.2**

*If $A$ is an $m\times n$ matrix, let $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ denote the columns of $A$.*

1. *$\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ is independent in $\mathbb{R}^m$ if and only if $A\mathbf{x}=\mathbf{0}$, $\mathbf{x}$ in $\mathbb{R}^n$, implies $\mathbf{x}=\mathbf{0}$.*
2. *$\mathbb{R}^m = \text{span}\,\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ if and only if $A\mathbf{x}=\mathbf{b}$ has a solution $\mathbf{x}$ for every vector $\mathbf{b}$ in $\mathbb{R}^m$.*

For a *square* matrix $A$, Theorem 5.2.2 characterizes the invertibility of $A$ in terms of the spanning and independence of its columns (see the discussion preceding Theorem 5.2.2). It is important to be able to discuss these notions for *rows*. If $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ are $1 \times n$ rows, we define $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ to be the set of all linear combinations of the $\mathbf{x}_i$ (as matrices), and we say that $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is linearly independent if the only vanishing linear combination is the trivial one (that is, if $\{\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_k^T\}$ is independent in $\mathbb{R}^n$, as the reader can verify).[6]

**Theorem 5.2.3**

*The following are equivalent for an $n \times n$ matrix $A$:*

1. *$A$ is invertible.*
2. *The columns of $A$ are linearly independent.*
3. *The columns of $A$ span $\mathbb{R}^n$.*
4. *The rows of $A$ are linearly independent.*
5. *The rows of $A$ span the set of all $1 \times n$ rows.*

**Proof.** Let $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$ denote the columns of $A$.

(1) $\Leftrightarrow$ (2). By Theorem 2.4.5, $A$ is invertible if and only if $A\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$; this holds if and only if $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ is independent by Theorem 5.2.2.

(1) $\Leftrightarrow$ (3). Again by Theorem 2.4.5, $A$ is invertible if and only if $A\mathbf{x} = \mathbf{b}$ has a solution for every column $B$ in $\mathbb{R}^n$; this holds if and only if $\text{span}\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\} = \mathbb{R}^n$ by Theorem 5.2.2.

(1) $\Leftrightarrow$ (4). The matrix $A$ is invertible if and only if $A^T$ is invertible (by Corollary 2.4.1 to Theorem 2.4.4); this in turn holds if and only if $A^T$ has independent columns (by (1) $\Leftrightarrow$ (2)); finally, this last statement holds if and only if $A$ has independent rows (because the rows of $A$ are the transposes of the columns of $A^T$).

(1) $\Leftrightarrow$ (5). The proof is similar to (1) $\Leftrightarrow$ (4). □

**Example 5.2.9**

Show that $S = \{(2, -2, 5), (-3, 1, 1), (2, 7, -4)\}$ is independent in $\mathbb{R}^3$.

**Solution.** Consider the matrix $A = \left[ \begin{array}{rrr} 2 & -2 & 5 \\ -3 & 1 & 1 \\ 2 & 7 & -4 \end{array} \right]$ with the vectors in $S$ as its rows. A routine computation shows that $\det A = -117 \neq 0$, so $A$ is invertible. Hence $S$ is independent by Theorem 5.2.3. Note that Theorem 5.2.3 also shows that $\mathbb{R}^3 = \text{span } S$.

[6]It is best to view columns and rows as just two different *notations* for ordered $n$-tuples. This discussion will become redundant in Chapter 6 where we define the general notion of a vector space.

## Dimension

It is common geometrical language to say that $\mathbb{R}^3$ is 3-dimensional, that planes are 2-dimensional and that lines are 1-dimensional. The next theorem is a basic tool for clarifying this idea of "dimension". Its importance is difficult to exaggerate.

**Theorem 5.2.4: Fundamental Theorem**

*Let $U$ be a subspace of $\mathbb{R}^n$. If $U$ is spanned by $m$ vectors, and if $U$ contains $k$ linearly independent vectors, then $k \leq m$.*

This proof is given in Theorem 6.3.2 in much greater generality.

**Definition 5.4 Basis of $\mathbb{R}^n$**

*If $U$ is a subspace of $\mathbb{R}^n$, a set $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ of vectors in $U$ is called a **basis** of $U$ if it satisfies the following two conditions:*

1. *$\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ is linearly independent.*
2. *$U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$.*

The most remarkable result about bases[7] is:

**Theorem 5.2.5: Invariance Theorem**

*If $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ and $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$ are bases of a subspace $U$ of $\mathbb{R}^n$, then $m = k$.*

Proof. We have $k \leq m$ by the fundamental theorem because $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ spans $U$, and $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$ is independent. Similarly, by interchanging $\mathbf{x}$'s and $\mathbf{y}$'s we get $m \leq k$. Hence $m = k$. □

The invariance theorem guarantees that there is no ambiguity in the following definition:

**Definition 5.5 Dimension of a Subspace of $\mathbb{R}^n$**

*If $U$ is a subspace of $\mathbb{R}^n$ and $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ is any basis of $U$, the number, $m$, of vectors in the basis is called the **dimension** of $U$, denoted*

$$\dim U = m$$

The importance of the invariance theorem is that the dimension of $U$ can be determined by counting the number of vectors in *any* basis.[8]

[7] The plural of "basis" is "bases".

[8] We will show in Theorem 5.2.6 that every subspace of $\mathbb{R}^n$ does indeed *have* a basis.

Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ denote the standard basis of $\mathbb{R}^n$, that is the set of columns of the identity matrix. Then $\mathbb{R}^n = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ by Example 5.1.6, and $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is independent by Example 5.2.2. Hence it is indeed a basis of $\mathbb{R}^n$ in the present terminology, and we have

**Example 5.2.10**

$\dim(\mathbb{R}^n) = n$ and $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is a basis.

This agrees with our geometric sense that $\mathbb{R}^2$ is two-dimensional and $\mathbb{R}^3$ is three-dimensional. It also says that $\mathbb{R}^1 = \mathbb{R}$ is one-dimensional, and $\{1\}$ is a basis. Returning to subspaces of $\mathbb{R}^n$, we define

$$\dim\{\mathbf{0}\} = 0$$

This amounts to saying $\{\mathbf{0}\}$ has a basis containing *no* vectors. This makes sense because $\mathbf{0}$ cannot belong to *any* independent set (Example 5.2.4).

**Example 5.2.11**

Let $U = \left\{ \begin{bmatrix} r \\ s \\ r \end{bmatrix} \mid r, s \text{ in } \mathbb{R} \right\}$. Show that $U$ is a subspace of $\mathbb{R}^3$, find a basis, and calculate $\dim U$.

**Solution.** Clearly, $\begin{bmatrix} r \\ s \\ r \end{bmatrix} = r\mathbf{u} + s\mathbf{v}$ where $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. It follows that $U = \text{span}\{\mathbf{u}, \mathbf{v}\}$, and hence that $U$ is a subspace of $\mathbb{R}^3$. Moreover, if $r\mathbf{u} + s\mathbf{v} = \mathbf{0}$, then $\begin{bmatrix} r \\ s \\ r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ so $r = s = 0$. Hence $\{\mathbf{u}, \mathbf{v}\}$ is independent, and so a **basis** of $U$. This means $\dim U = 2$.

**Example 5.2.12**

Let $B = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ be a basis of $\mathbb{R}^n$. If $A$ is an invertible $n \times n$ matrix, then $D = \{A\mathbf{x}_1, A\mathbf{x}_2, \dots, A\mathbf{x}_n\}$ is also a basis of $\mathbb{R}^n$.

**Solution.** Let $\mathbf{x}$ be a vector in $\mathbb{R}^n$. Then $A^{-1}\mathbf{x}$ is in $\mathbb{R}^n$ so, since $B$ is a basis, we have $A^{-1}\mathbf{x} = t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_n\mathbf{x}_n$ for $t_i$ in $\mathbb{R}$. Left multiplication by $A$ gives $\mathbf{x} = t_1(A\mathbf{x}_1) + t_2(A\mathbf{x}_2) + \cdots + t_n(A\mathbf{x}_n)$, and it follows that $D$ spans $\mathbb{R}^n$. To show independence, let $s_1(A\mathbf{x}_1) + s_2(A\mathbf{x}_2) + \cdots + s_n(A\mathbf{x}_n) = \mathbf{0}$, where the $s_i$ are in $\mathbb{R}$. Then $A(s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + \cdots + s_n\mathbf{x}_n) = \mathbf{0}$ so left multiplication by $A^{-1}$ gives $s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + \cdots + s_n\mathbf{x}_n = \mathbf{0}$. Now the independence of $B$ shows that each $s_i = 0$, and so proves the independence of $D$. Hence $D$ is a basis of $\mathbb{R}^n$.

While we have found bases in many subspaces of $\mathbb{R}^n$, we have not yet shown that *every* subspace *has* a basis. This is part of the next theorem, the proof of which is deferred to Section 6.4 (Theorem 6.4.1) where it will be proved in more generality.

**Theorem 5.2.6**

*Let $U \neq \{\mathbf{0}\}$ be a subspace of $\mathbb{R}^n$. Then:*

1. *$U$ has a basis and $\dim U \leq n$.*
2. *Any independent set in $U$ can be enlarged (by adding vectors from the standard basis) to a basis of $U$.*
3. *Any spanning set for $U$ can be cut down (by deleting vectors) to a basis of $U$.*

**Example 5.2.13**

Find a basis of $\mathbb{R}^4$ containing $S = \{\mathbf{u},\ \mathbf{v}\}$ where $\mathbf{u} = (0,\ 1,\ 2,\ 3)$ and $\mathbf{v} = (2,\ -1,\ 0,\ 1)$.

**Solution.** By Theorem 5.2.6 we can find such a basis by adding vectors from the standard basis of $\mathbb{R}^4$ to $S$. If we try $\mathbf{e}_1 = (1,\ 0,\ 0,\ 0)$, we find easily that $\{\mathbf{e}_1,\ \mathbf{u},\ \mathbf{v}\}$ is independent. Now add another vector from the standard basis, say $\mathbf{e}_2$.
Again we find that $B = \{\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{u},\ \mathbf{v}\}$ is independent. Since $B$ has $4 = \dim \mathbb{R}^4$ vectors, then $B$ must span $\mathbb{R}^4$ by Theorem 5.2.7 below (or simply verify it directly). Hence $B$ is a basis of $\mathbb{R}^4$.

Theorem 5.2.6 has a number of useful consequences. Here is the first.

**Theorem 5.2.7**

*Let $U$ be a subspace of $\mathbb{R}^n$ where $\dim U = m$ and let $B = \{\mathbf{x}_1,\ \mathbf{x}_2,\ \dots,\ \mathbf{x}_m\}$ be a set of $m$ vectors in $U$. Then $B$ is independent if and only if $B$ spans $U$.*

**Proof.** Suppose $B$ is independent. If $B$ does not span $U$ then, by Theorem 5.2.6, $B$ can be enlarged to a basis of $U$ containing more than $m$ vectors. This contradicts the invariance theorem because $\dim U = m$, so $B$ spans $U$. Conversely, if $B$ spans $U$ but is not independent, then $B$ can be cut down to a basis of $U$ containing fewer than $m$ vectors, again a contradiction. So $B$ is independent, as required. □

As we saw in Example 5.2.13, Theorem 5.2.7 is a "labour-saving" result. It asserts that, given a subspace $U$ of dimension $m$ and a set $B$ of exactly $m$ vectors in $U$, to prove that $B$ is a basis of $U$ it suffices to show either that $B$ spans $U$ or that $B$ is independent. It is not necessary to verify both properties.

**Theorem 5.2.8**

*Let $U \subseteq W$ be subspaces of $\mathbb{R}^n$. Then:*

1. $\dim U \leq \dim W$.
2. *If $\dim U = \dim W$, then $U = W$.*

**Proof.** Write $\dim W = k$, and let $B$ be a basis of $U$.

1. If $\dim U > k$, then $B$ is an independent set in $W$ containing more than $k$ vectors, contradicting the fundamental theorem. So $\dim U \leq k = \dim W$.

2. If $\dim U = k$, then $B$ is an independent set in $W$ containing $k = \dim W$ vectors, so $B$ spans $W$ by Theorem 5.2.7. Hence $W = \text{span } B = U$, proving (2). □

It follows from Theorem 5.2.8 that if $U$ is a subspace of $\mathbb{R}^n$, then $\dim U$ is one of the integers $0, 1, 2, \dots, n$, and that:

$$\begin{array}{lll} \dim U = 0 & \text{if and only if} & U = \{\mathbf{0}\}, \\ \dim U = n & \text{if and only if} & U = \mathbb{R}^n \end{array}$$

The other subspaces of $\mathbb{R}^n$ are called **proper**. The following example uses Theorem 5.2.8 to show that the proper subspaces of $\mathbb{R}^2$ are the lines through the origin, while the proper subspaces of $\mathbb{R}^3$ are the lines and planes through the origin.

**Example 5.2.14**

1. If $U$ is a subspace of $\mathbb{R}^2$ or $\mathbb{R}^3$, then $\dim U = 1$ if and only if $U$ is a line through the origin.

2. If $U$ is a subspace of $\mathbb{R}^3$, then $\dim U = 2$ if and only if $U$ is a plane through the origin.

**Proof.**

1. Since $\dim U = 1$, let $\{\mathbf{u}\}$ be a basis of $U$. Then $U = \text{span}\{\mathbf{u}\} = \{t\mathbf{u} \mid t \text{ in } \mathbb{R}\}$, so $U$ is the line through the origin with direction vector $\mathbf{u}$. Conversely each line $L$ with direction vector $\mathbf{d} \neq \mathbf{0}$ has the form $L = \{t\mathbf{d} \mid t \text{ in } \mathbb{R}\}$. Hence $\{\mathbf{d}\}$ is a basis of $U$, so $U$ has dimension 1.

2. If $U \subseteq \mathbb{R}^3$ has dimension 2, let $\{\mathbf{v}, \mathbf{w}\}$ be a basis of $U$. Then $\mathbf{v}$ and $\mathbf{w}$ are not parallel (by Example 5.2.7) so $\mathbf{n} = \mathbf{v} \times \mathbf{w} \neq \mathbf{0}$. Let $P = \{\mathbf{x} \text{ in } \mathbb{R}^3 \mid \mathbf{n} \cdot \mathbf{x} = 0\}$ denote the plane through the origin with normal $\mathbf{n}$. Then $P$ is a subspace of $\mathbb{R}^3$ (Example 5.1.1) and both $\mathbf{v}$ and $\mathbf{w}$ lie in $P$ (they are orthogonal to $\mathbf{n}$), so $U = \text{span}\{\mathbf{v}, \mathbf{w}\} \subseteq P$ by Theorem 5.1.1. Hence

   $$U \subseteq P \subseteq \mathbb{R}^3$$

   Since $\dim U = 2$ and $\dim(\mathbb{R}^3) = 3$, it follows from Theorem 5.2.8 that $\dim P = 2$ or 3, whence $P = U$ or $\mathbb{R}^3$. But $P \neq \mathbb{R}^3$ (for example, $\mathbf{n}$ is not in $P$) and so $U = P$ is a plane through the origin.

   Conversely, if $U$ is a plane through the origin, then $\dim U = 0$, 1, 2, or 3 by Theorem 5.2.8. But $\dim U \neq 0$ or 3 because $U \neq \{\mathbf{0}\}$ and $U \neq \mathbb{R}^3$, and $\dim U \neq 1$ by (1). So $\dim U = 2$. □

Note that this proof shows that if $\mathbf{v}$ and $\mathbf{w}$ are nonzero, nonparallel vectors in $\mathbb{R}^3$, then $\text{span}\{\mathbf{v}, \mathbf{w}\}$ is the plane with normal $\mathbf{n} = \mathbf{v} \times \mathbf{w}$. We gave a geometrical verification of this fact in Section 5.1.

# Exercises for 5.2

In Exercises 5.2.1-5.2.6 we write vectors $\mathbb{R}^n$ as rows.

**Exercise 5.2.1** Which of the following subsets are independent? Support your answer.

a. $\{(1,\ -1,\ 0),\ (3,\ 2,\ -1),\ (3,\ 5,\ -2)\}$ in $\mathbb{R}^3$

b. $\{(1,\ 1,\ 1),\ (1,\ -1,\ 1),\ (0,\ 0,\ 1)\}$ in $\mathbb{R}^3$

c. $\{(1,\ -1,\ 1,\ -1),\ (2,\ 0,\ 1,\ 0),\ (0,\ -2,\ 1,\ -2)\}$ in $\mathbb{R}^4$

d. $\{(1,\ 1,\ 0,\ 0),\ (1,\ 0,\ 1,\ 0),\ (0,\ 0,\ 1,\ 1),\ (0,\ 1,\ 0,\ 1)\}$ in $\mathbb{R}^4$

**Exercise 5.2.2** Let $\{\mathbf{x},\ \mathbf{y},\ \mathbf{z},\ \mathbf{w}\}$ be an independent set in $\mathbb{R}^n$. Which of the following sets is independent? Support your answer.

a. $\{\mathbf{x}-\mathbf{y},\ \mathbf{y}-\mathbf{z},\ \mathbf{z}-\mathbf{x}\}$

b. $\{\mathbf{x}+\mathbf{y},\ \mathbf{y}+\mathbf{z},\ \mathbf{z}+\mathbf{x}\}$

c. $\{\mathbf{x}-\mathbf{y},\ \mathbf{y}-\mathbf{z},\ \mathbf{z}-\mathbf{w},\ \mathbf{w}-\mathbf{x}\}$

d. $\{\mathbf{x}+\mathbf{y},\ \mathbf{y}+\mathbf{z},\ \mathbf{z}+\mathbf{w},\ \mathbf{w}+\mathbf{x}\}$

**Exercise 5.2.3** Find a basis and calculate the dimension of the following subspaces of $\mathbb{R}^4$.

a. $\operatorname{span}\{(1,\ -1,\ 2,\ 0),\ (2,\ 3,\ 0,\ 3),\ (1,\ 9,\ -6,\ 6)\}$

b. $\operatorname{span}\{(2,\ 1,\ 0,\ -1),\ (-1,\ 1,\ 1,\ 1),\ (2,\ 7,\ 4,\ 1)\}$

c. $\operatorname{span}\{(-1,\ 2,\ 1,\ 0),\ (2,\ 0,\ 3,\ -1),\ (4,\ 4,\ 11,\ -3),\ (3,\ -2,\ 2,\ -1)\}$

d. $\operatorname{span}\{(-2,\ 0,\ 3,\ 1),\ (1,\ 2,\ -1,\ 0),\ (-2,\ 8,\ 5,\ 3),\ (-1,\ 2,\ 2,\ 1)\}$

**Exercise 5.2.4** Find a basis and calculate the dimension of the following subspaces of $\mathbb{R}^4$.

a. $U = \left\{ \left[ \begin{array}{c} a \\ a+b \\ a-b \\ b \end{array} \right] \middle|\ a \text{ and } b \text{ in } \mathbb{R} \right\}$

b. $U = \left\{ \left[ \begin{array}{c} a+b \\ a-b \\ b \\ a \end{array} \right] \middle|\ a \text{ and } b \text{ in } \mathbb{R} \right\}$

c. $U = \left\{ \left[ \begin{array}{c} a \\ b \\ c+a \\ c \end{array} \right] \middle|\ a,\ b, \text{ and } c \text{ in } \mathbb{R} \right\}$

d. $U = \left\{ \left[ \begin{array}{c} a-b \\ b+c \\ a \\ b+c \end{array} \right] \middle|\ a,\ b, \text{ and } c \text{ in } \mathbb{R} \right\}$

e. $U = \left\{ \left[ \begin{array}{c} a \\ b \\ c \\ d \end{array} \right] \middle|\ a+b-c+d = 0 \text{ in } \mathbb{R} \right\}$

f. $U = \left\{ \left[ \begin{array}{c} a \\ b \\ c \\ d \end{array} \right] \middle|\ a+b = c+d \text{ in } \mathbb{R} \right\}$

**Exercise 5.2.5** Suppose that $\{\mathbf{x},\ \mathbf{y},\ \mathbf{z},\ \mathbf{w}\}$ is a basis of $\mathbb{R}^4$. Show that:

a. $\{\mathbf{x}+a\mathbf{w},\ \mathbf{y},\ \mathbf{z},\ \mathbf{w}\}$ is also a basis of $\mathbb{R}^4$ for any choice of the scalar $a$.

b. $\{\mathbf{x}+\mathbf{w},\ \mathbf{y}+\mathbf{w},\ \mathbf{z}+\mathbf{w},\ \mathbf{w}\}$ is also a basis of $\mathbb{R}^4$.

c. $\{\mathbf{x},\ \mathbf{x}+\mathbf{y},\ \mathbf{x}+\mathbf{y}+\mathbf{z},\ \mathbf{x}+\mathbf{y}+\mathbf{z}+\mathbf{w}\}$ is also a basis of $\mathbb{R}^4$.

**Exercise 5.2.6** Use Theorem 5.2.3 to determine if the following sets of vectors are a basis of the indicated space.

a. $\{(3,\ -1),\ (2,\ 2)\}$ in $\mathbb{R}^2$

b. $\{(1,\ 1,\ -1),\ (1,\ -1,\ 1),\ (0,\ 0,\ 1)\}$ in $\mathbb{R}^3$

c. $\{(-1,\ 1,\ -1),\ (1,\ -1,\ 2),\ (0,\ 0,\ 1)\}$ in $\mathbb{R}^3$

d. $\{(5,\ 2,\ -1),\ (1,\ 0,\ 1),\ (3,\ -1,\ 0)\}$ in $\mathbb{R}^3$

e. $\{(2, 1, -1, 3), (1, 1, 0, 2), (0, 1, 0, -3), (-1, 2, 3, 1)\}$ in $\mathbb{R}^4$

f. $\{(1, 0, -2, 5), (4, 4, -3, 2), (0, 1, 0, -3), (1, 3, 3, -10)\}$ in $\mathbb{R}^4$

**Exercise 5.2.7** In each case show that the statement is true or give an example showing that it is false.

a. If $\{\mathbf{x}, \mathbf{y}\}$ is independent, then $\{\mathbf{x}, \mathbf{y}, \mathbf{x}+\mathbf{y}\}$ is independent.

b. If $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ is independent, then $\{\mathbf{y}, \mathbf{z}\}$ is independent.

c. If $\{\mathbf{y}, \mathbf{z}\}$ is dependent, then $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ is dependent for any $\mathbf{x}$.

d. If all of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ are nonzero, then $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is independent.

e. If one of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ is zero, then $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is dependent.

f. If $a\mathbf{x}+b\mathbf{y}+c\mathbf{z}=\mathbf{0}$, then $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ is independent.

g. If $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ is independent, then $a\mathbf{x}+b\mathbf{y}+c\mathbf{z}=\mathbf{0}$ for some $a$, $b$, and $c$ in $\mathbb{R}$.

h. If $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is dependent, then $t_1\mathbf{x}_1+t_2\mathbf{x}_2+\cdots+t_k\mathbf{x}_k=\mathbf{0}$ for some numbers $t_i$ in $\mathbb{R}$ not all zero.

i. If $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is independent, then $t_1\mathbf{x}_1+t_2\mathbf{x}_2+\cdots+t_k\mathbf{x}_k=\mathbf{0}$ for some $t_i$ in $\mathbb{R}$.

j. Every non-empty subset of a linearly independent set is again linearly independent.

k. Every set containing a spanning set is again a spanning set.

**Exercise 5.2.8** If $A$ is an $n\times n$ matrix, show that $\det A = 0$ if and only if some column of $A$ is a linear combination of the other columns.

**Exercise 5.2.9** Let $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ be a linearly independent set in $\mathbb{R}^4$. Show that $\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{e}_k\}$ is a basis of $\mathbb{R}^4$ for some $\mathbf{e}_k$ in the standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$.

**Exercise 5.2.10** If $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6\}$ is an independent set of vectors, show that the subset $\{\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_5\}$ is also independent.

**Exercise 5.2.11** Let $A$ be any $m\times n$ matrix, and let $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \dots, \mathbf{b}_k$ be columns in $\mathbb{R}^m$ such that the system $A\mathbf{x}=\mathbf{b}_i$ has a solution $\mathbf{x}_i$ for each $i$. If $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \dots, \mathbf{b}_k\}$ is independent in $\mathbb{R}^m$, show that $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_k\}$ is independent in $\mathbb{R}^n$.

**Exercise 5.2.12** If $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_k\}$ is independent, show $\{\mathbf{x}_1, \mathbf{x}_1+\mathbf{x}_2, \mathbf{x}_1+\mathbf{x}_2+\mathbf{x}_3, \dots, \mathbf{x}_1+\mathbf{x}_2+\cdots+\mathbf{x}_k\}$ is also independent.

**Exercise 5.2.13** If $\{\mathbf{y}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_k\}$ is independent, show that $\{\mathbf{y}+\mathbf{x}_1, \mathbf{y}+\mathbf{x}_2, \mathbf{y}+\mathbf{x}_3, \dots, \mathbf{y}+\mathbf{x}_k\}$ is also independent.

**Exercise 5.2.14** If $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is independent in $\mathbb{R}^n$, and if $\mathbf{y}$ is not in $\operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$, show that $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k, \mathbf{y}\}$ is independent.

**Exercise 5.2.15** If $A$ and $B$ are matrices and the columns of $AB$ are independent, show that the columns of $B$ are independent.

**Exercise 5.2.16** Suppose that $\{\mathbf{x}, \mathbf{y}\}$ is a basis of $\mathbb{R}^2$, and let $A = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$.

a. If $A$ is invertible, show that $\{a\mathbf{x}+b\mathbf{y}, c\mathbf{x}+d\mathbf{y}\}$ is a basis of $\mathbb{R}^2$.

b. If $\{a\mathbf{x}+b\mathbf{y}, c\mathbf{x}+d\mathbf{y}\}$ is a basis of $\mathbb{R}^2$, show that $A$ is invertible.

**Exercise 5.2.17** Let $A$ denote an $m\times n$ matrix.

a. Show that $\operatorname{null} A = \operatorname{null}(UA)$ for every invertible $m\times m$ matrix $U$.

b. Show that $\dim(\operatorname{null} A) = \dim(\operatorname{null}(AV))$ for every invertible $n\times n$ matrix $V$. [*Hint*: If $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is a basis of $\operatorname{null} A$, show that $\{V^{-1}\mathbf{x}_1, V^{-1}\mathbf{x}_2, \dots, V^{-1}\mathbf{x}_k\}$ is a basis of $\operatorname{null}(AV)$.]

**Exercise 5.2.18** Let $A$ denote an $m\times n$ matrix.

a. Show that $\operatorname{im} A = \operatorname{im}(AV)$ for every invertible $n\times n$ matrix $V$.

b. Show that $\dim(\operatorname{im} A) = \dim(\operatorname{im}(UA))$ for every invertible $m\times m$ matrix $U$. [*Hint*: If $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$ is a basis of $\operatorname{im}(UA)$, show that $\{U^{-1}\mathbf{y}_1, U^{-1}\mathbf{y}_2, \dots, U^{-1}\mathbf{y}_k\}$ is a basis of $\operatorname{im} A$.]

**Exercise 5.2.19** Let $U$ and $W$ denote subspaces of $\mathbb{R}^n$, and assume that $U\subseteq W$. If $\dim U = n-1$, show that either $W=U$ or $W=\mathbb{R}^n$.

**Exercise 5.2.20** Let $U$ and $W$ denote subspaces of $\mathbb{R}^n$, and assume that $U\subseteq W$. If $\dim W = 1$, show that either $U=\{\mathbf{0}\}$ or $U=W$.

# 5.3 Orthogonality

Length and orthogonality are basic concepts in geometry and, in $\mathbb{R}^2$ and $\mathbb{R}^3$, they both can be defined using the dot product. In this section we extend the dot product to vectors in $\mathbb{R}^n$, and so endow $\mathbb{R}^n$ with euclidean geometry. We then introduce the idea of an orthogonal basis—one of the most useful concepts in linear algebra, and begin exploring some of its applications.

## Dot Product, Length, and Distance

If $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$ are two $n$-tuples in $\mathbb{R}^n$, recall that their **dot product** was defined in Section 2.2 as follows:

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

Observe that if $\mathbf{x}$ and $\mathbf{y}$ are written as columns then $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}$ is a matrix product (and $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}\mathbf{y}^T$ if they are written as rows). Here $\mathbf{x} \cdot \mathbf{y}$ is a $1 \times 1$ matrix, which we take to be a number.

**Definition 5.6 Length in $\mathbb{R}^n$**

*As in $\mathbb{R}^3$, the **length** $\|\mathbf{x}\|$ of the vector is defined by*

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

*Where $\sqrt{(\quad)}$ indicates the positive square root.*

A vector $\mathbf{x}$ of length 1 is called a **unit vector**. If $\mathbf{x} \neq \mathbf{0}$, then $\|\mathbf{x}\| \neq 0$ and it follows easily that $\frac{1}{\|\mathbf{x}\|}\mathbf{x}$ is a unit vector (see Theorem 5.3.6 below), a fact that we shall use later.

**Example 5.3.1**

If $\mathbf{x} = (1, -1, -3, 1)$ and $\mathbf{y} = (2, 1, 1, 0)$ in $\mathbb{R}^4$, then $\mathbf{x} \cdot \mathbf{y} = 2 - 1 - 3 + 0 = -2$ and $\|\mathbf{x}\| = \sqrt{1+1+9+1} = \sqrt{12} = 2\sqrt{3}$. Hence $\frac{1}{2\sqrt{3}}\mathbf{x}$ is a unit vector; similarly $\frac{1}{\sqrt{6}}\mathbf{y}$ is a unit vector.

These definitions agree with those in $\mathbb{R}^2$ and $\mathbb{R}^3$, and many properties carry over to $\mathbb{R}^n$:

**Theorem 5.3.1**

*Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ denote vectors in $\mathbb{R}^n$. Then:*

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.
2. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$.
3. $(a\mathbf{x}) \cdot \mathbf{y} = a(\mathbf{x} \cdot \mathbf{y}) = \mathbf{x} \cdot (a\mathbf{y})$ *for all scalars $a$.*