1 Length and Dot Product in $\mathbb{R}^n$

1.1 Reviewing $\mathbb{R}^2$

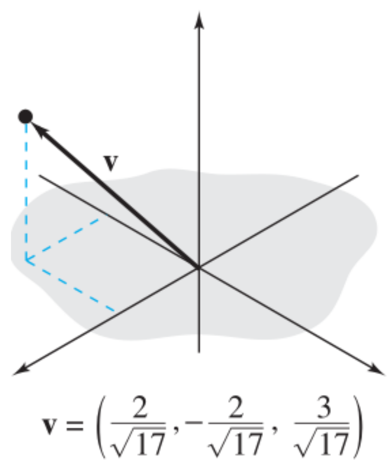
If $\mathbf{v} = (v_1, v_2)$, then **length** (or **magnitute**) of $\mathbf{v}$ is

$$||\mathbf{v}|| = \sqrt{v_1^2 + v_2^2}$$

Length of a Vector in $\mathbb{R}^n$ The **length**, or **magnitute** of a vector $\mathbf{v}$= $(v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ is

$$||\mathbf{v}|| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

REMARK. The length of a vector is also called its **norm**. If $||\mathbf{v}|| = 1$, then $\mathbf{v}$ is a **unit vector**.



Each vector in the *standard basis* for $\mathbb{R}^n$ has length 1 and is called **standard unit vector** in $\mathbb{R}^n$. 2 vectors are *parallel* if one is a scalar multiple of the other.

Length of a Scalar Multiple.

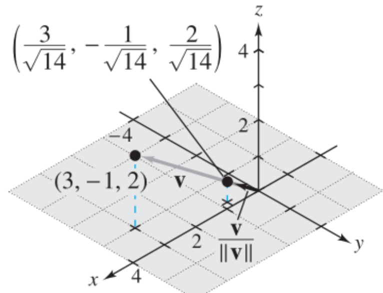
$$||c\mathbf{v}|| = |c|.||\mathbf{v}||$$

THEOREM 5.2 **Unit Vector in the Direction of v.** If $\mathbf{v}$ is a nonzero vector in $\mathbb{R}^n$, then

$$\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||}$$

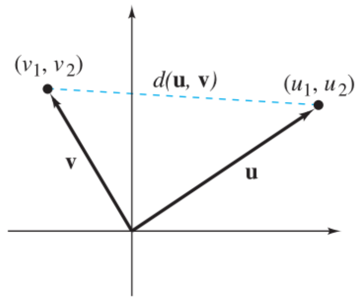
is the **unit vector in the direction of v** *(has length* 1 *and same direction as* $\boldsymbol{v}$*).*

Example 2. Finding a Unit Vector for $\mathbf{v}$= $(3, -1, 2)$.



1.2 Distance Between 2 Vectors in $\mathbb{R}^n$

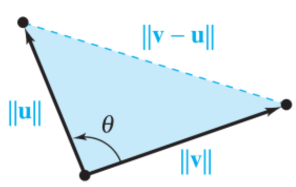
$\mathbb{R}^2$ as a model.
The distance between 2 point $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ is $d = \sqrt{(u_1 - v_1)^2 + (u_2 - v^2)^2}$.



Distance Between 2 Vectors. The **distance between 2 vectors u** and **v** in $\mathbb{R}^n$ is

$$d(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||$$

1.3 Dot Product and the Angle Between 2 Vectors



Angle Between Two Vectors

To find the angle $\theta (0 \leq \theta \leq \pi)$ of 2 vectors in $\mathbb{R}^2$, the *Law of Cosines* can be applied

$$||\mathbf{v} - \mathbf{u}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2 - 2\|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$$

Expanding and solving for $\cos\theta$ yields

$$\cos\theta = \frac{u_1 v_1 + u_2 v_2}{\|\mathbf{u}\|\|\mathbf{v}\|}$$

The numerator of the quotient above is defined as the **dot product** of **u** and **v**

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2$$

Dot Product in $\mathbb{R}^n$. The **dot product** of **u** and **v** is the *scalar* quantity

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Notice. The dot product of 2 vectors is a *scalar*, not another vector.

THEOREM 5.3 **Properties of the Dot Product**

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
3. $(\mathbf{u} \cdot \mathbf{v}) = (c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v})$
4. $\mathbf{v} \cdot \mathbf{v} = \|v\|^2$
5. $\mathbf{v} \cdot \mathbf{v} \geq 0$, $\mathbf{v} \cdot \mathbf{v} = 0$ if and only if $\mathbf{v} = 0$.

THEOREM 5.4 **The Cauchy-Schwarz Inequality**
If **u** and **v** are vectors in $\mathbb{R}^n$, then

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|,$$

wher $|\mathbf{u} \cdot \mathbf{v}|$ denotes the *absolute value* of $\mathbf{u} \cdot \mathbf{v}$.

The Angle Between 2 Vectors in $\mathbb{R}^n$. The angle θ between 2 nonzero vectors **u** and **v** in $\mathbb{R}^n$,

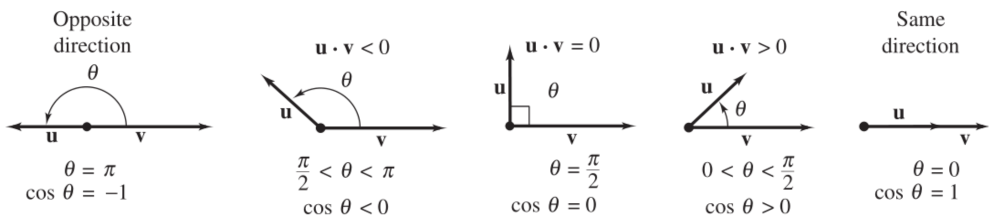
$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}, \quad 0 \leq \theta \leq \pi$$

Example. The angle between **u**$= (-4, 0, 2, -2)$ and $\mathbf{v} = (2, 0, -1, 1)$ is

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = -\frac{12}{\sqrt{24}\sqrt{6}} = -1$$

Consequently, $\theta = \pi$. It makes sense that $\mathbf{u}$ and $\mathbf{v}$ should have opposite direction, because $\mathbf{u} = -2\mathbf{v}$.

Note. Because $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$ are always positive, $\mathbf{u} \cdot \mathbf{v}$ and $\cos\theta$ will always have the same sign. The sign of the *dot product* can be used to determine whether θ is *acute* or *obtuse*.



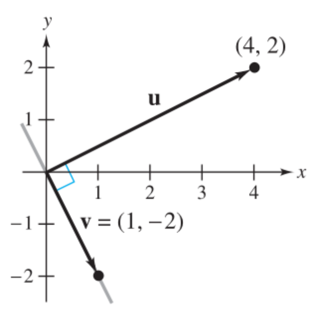
Orthogonal Vectors. Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are **orthogonal** (or *perpendicular*) if

$$\mathbf{u} \cdot \mathbf{v} = 0$$

REMARK. The vector $\mathbf{0}$ is orthogonal to every vector.

Example. Finding Orthogonal Vectors.

Determine all vectors in $\mathbb{R}^2$ that are orthogonal to $\mathbf{u}= (4, 2)$.



SOLUTION. Let $\mathbf{v}= (v_1, v_2)$ be orthogonal to $\mathbf{u}$. Then

$$\mathbf{u} \cdot \mathbf{v} = 4v_1 + 2v_2 = 0$$

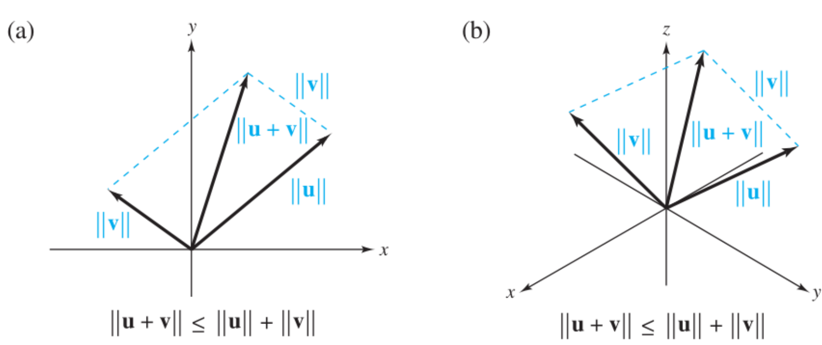
which implies that $2v_2 = -4v_1$ and $v_2 = -2v_1$. So every vector that is orthogonal to $(4, 2)$ is of the form

$$\mathbf{v} = (t, -2t) = t(1, -2)$$

where t is a real number.

THEOREM 5.5 **The Triangle Inequality**
If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, then

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$



REMARK. Equality occurs in the Triangle Inequality if and only if $\mathbf{u}$ and $\mathbf{v}$ have the same direction.

THEOREM 5.6 **The Pythagorean Theorem**
If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, then they are orthogonal if and only if

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Proof. $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2(\mathbf{u} \cdot \mathbf{v})$, and their dot product is zero.

1.4 The Dot Product and Matrix Multiplication

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. Then the **dot product** of 2 vectors is

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 v_1 + u_2 v_2 + \cdots + u_3 v_3 \end{bmatrix}$$

2 Inner Product Spaces

The previous *dot product* in $\mathbb{R}^n$ is an example of **inner product -** called **Euclidean inner product**. To distinguish between the standard inner product and other possible inner products,

$$\mathbf{u} \cdot \mathbf{v} = \text{dot product (Euclidean inner product for } \mathbb{R}^n)$$

$$\langle \mathbf{u} \cdot \mathbf{v} \rangle = \text{general inner product for vector space } V$$

Inner Product. Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be vectors in V. An **inner product** on V is a function that associates a real numer $\langle \mathbf{u} \cdot \mathbf{v} \rangle$ with each pair of $(\mathbf{u}, \mathbf{v})$ satisfies these axioms.

1. $\langle \mathbf{u} \cdot \mathbf{v} \rangle = \langle \mathbf{v} \cdot \mathbf{u} \rangle$
2. $\langle \mathbf{u} \cdot \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u} \cdot \mathbf{v} \rangle + \langle \mathbf{u} \cdot \mathbf{w} \rangle$
3. $c\langle \mathbf{u} \cdot \mathbf{v} \rangle = \langle c\mathbf{u} \cdot \mathbf{v} \rangle$
4. $\langle \mathbf{v} \cdot \mathbf{v} \rangle \geq 0$, and $\langle \mathbf{v} \cdot \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

REMARK. A vector space V with an inner product is called an **inner product space**.

(1) **A Different Inner Product for** $\mathbb{R}^2$

Show that the following function defines an inner product on $\mathbb{R}^2$

$$\langle \mathbf{u} \cdot \mathbf{v} \rangle = u_1 v_1 + 2 u_2 v_2$$

This example can be generalize to show that

$$\langle \mathbf{u} \cdot \mathbf{v} \rangle = c_1 u_1 v_1 + c_2 u_2 v_2 + \cdots + c_n u_n v_n, \quad c_i > 0$$

is an inner product of $\mathbb{R}^n$. The positive constants $c_1, c_2, \ldots, c_n$ are **weights**. If any $c_i \leq 0$, then this function does not define an inner product.

(2) **A Function That is Not an Inner Product**

$$\langle \mathbf{u} \cdot \mathbf{v} \rangle = u_1 v_1 - 2 u_2 v_2 + u_3 v_3$$

(3) **An Inner Product on** $M_{2,2}$

Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ be matrices in the vector space $M_{2,2}$.

The function $\langle A, B \rangle = a_{11} b_{11} + a_{21} b_{21} + a_{12} b_{12} + a_{22} b_{22}$ is an inner product on $M_{2,2}$.

(4) **An Inner Product Defined by a Definite Integral (Calculus)**

Let f and g be real-valued continuous function in the vector space $C[a, b]$. Show that

$$\langle f, g \rangle = \int_a^b f(x) g(x)\, dx$$

defines an inner product on $C[a, b]$.

THEOREM 5.7 **Properties of Inner Product**
Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be vectors in an inner product space V, and $c \in \mathbb{R}$.

1. $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
3. $\langle \mathbf{u}, c\mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$

Definition of Norm, Distance, and Angle.
Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in an inner product space V.

1. The **norm** (or **length**) of $\mathbf{u}$ is $\|u\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$.
2. The **distance** between $\mathbf{u}$ amd $\mathbf{v}$ is $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$.
3. The **angle** between 2 nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ is given by
$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}, \quad 0 \leq \theta \leq \pi$$
4. $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

REMARK. If $\|\mathbf{v}\| = 1$, then $\mathbf{v}$ is called a **unit vector**. Moreover, $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ is the **unit vector in the direction of v**.

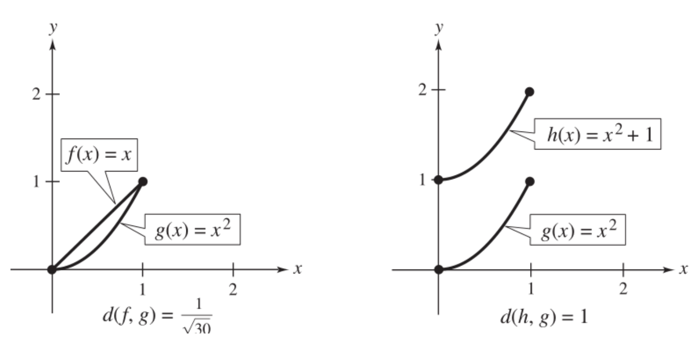
(6) **Finding Inner Products**

For polynomials $p = a_0 + a_1x + \cdots + a_nx^n$ and $q = b_0 + b_1x + \cdots + b_nx^n$ in the vector space P_n.

The fuction $\langle p, q \rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n$ is an inner product.

Note. Orthogonality depends on the particular inner product used. That is, 2 vectors may be orthogonal with respect to one inner product but not to another.

Example 7. Using Inner Product on $C[0, 1]$ (Calculus) (p.316)



If $\mathbf{u}$ and $\mathbf{v}$ are vectors in an inner product space, then

Properties of Norm

1. $\|\mathbf{u}\| \geq 0$
2. $\|\mathbf{u}\| = 0$ if and only if $\mathbf{u} = \mathbf{0}$.
3. $\|c\mathbf{u}\| = |c|\,\|\mathbf{u}\|$

Properties of Distance

1. $d(\mathbf{u}, \mathbf{v}) \geq 0$
2. $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$.
3. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$

THEOREM 5.8

1. Cauchy-Schwarz Inequality: $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$
2. Triangle Inequality: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$
3. Pythagorean Theorem: $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if
$$\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

2.1 Orthogonal Projections in Inner Product Spaces

Let **u** and **v** be vectors in the plane. If **v** is nonzero, then **u** can be orthogonally projected onto **v**. This projection is denoted by

$$\mathrm{proj}_v \mathbf{u} = a\mathbf{v}$$

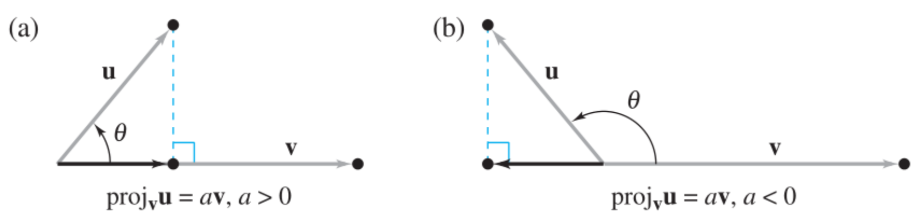
If $a > 0$, then $\cos\theta > 0$ and the length of $\mathrm{proj}_v \mathbf{u}$ is

$$\|a\mathbf{v}\| = a\|\mathbf{v}\| = \|\mathbf{u}\|\cos\theta = \frac{\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta}{\|\mathbf{v}\|} = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|}$$

which implies that $a = (\mathbf{u}\cdot\mathbf{v})/\|\mathbf{v}\|^2 = (\mathbf{u}\cdot\mathbf{v})/(\mathbf{v}\cdot\mathbf{v})$. So

$$\mathrm{proj}_{\mathbf{v}} \mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}$$

If $a < 0$, we obtain the same formula.

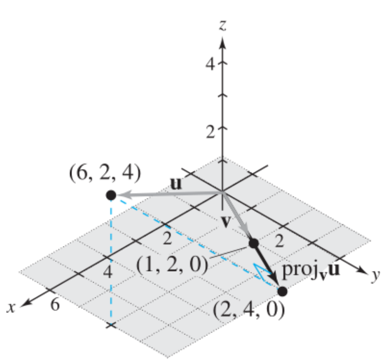


Definition of Orthogonal Projection. Let **u** and $\mathbf{v} \neq 0$ be vectors in an inner product space V, then the **orthogonal projection** of **u** onto **v** is

$$\mathrm{proj}_{\mathbf{v}} \mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{v}, \mathbf{v}\rangle}\mathbf{v}$$

REMARK. If **v** is an unit vector, then $\langle \mathbf{v}, \mathbf{v}\rangle = \|\mathbf{v}\|^2 = 1$. The formula takes the simpler form

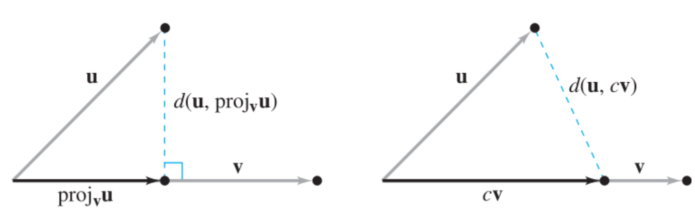
$$\mathrm{proj}_{\mathbf{v}} \mathbf{u} = \langle \mathbf{u}, \mathbf{v}\rangle \mathbf{v}$$



REMARK. **u** - $\mathrm{proj}_{\mathbf{v}}\mathbf{u}$ is orthogonal to **v**.
Of all possible scalar multiples of **v**, the orthogonal projection of **u** onto **v** is the one *closest* to **u**.

THEOREM 5.9 **Orthogonal Projection and Distance**
Let **u** and $\mathbf{v} \neq 0$ be 2 vectors in an inner product space V, then

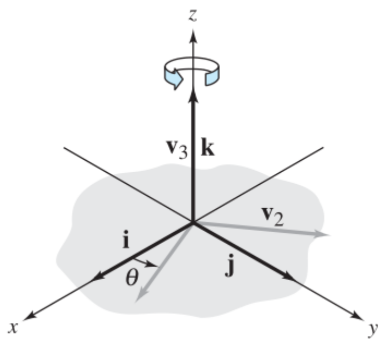
$$d(\mathbf{u}, \mathrm{proj}_{\mathbf{v}}\mathbf{u}) < d(\mathbf{u}, c\mathbf{v}), \quad c \neq \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{v}, \mathbf{v}\rangle}$$



3 Orthonormal Bases: Gram-Schmidt Process

A vector space can have many different bases, but certain bases are more *convenient* than others. For example, $\mathbb{R}^3$ has the convenient standard basis $B = \{(1,0,0), (0,1,0), (0,0,1)\}$. It has special characteristics that are particularly useful. First, 3 vectors in the basis are *mutually orthogonal*. Second, they are all *unit* vector.

Definition of Orthogonal and Orthonormal Sets
A set S of vectors in an inner product space V is **orthogonal** if every pair of vectors in S is orthogonal. If, in addition, each vector in the set is a *unit* vector, then S is **orthonormal**.



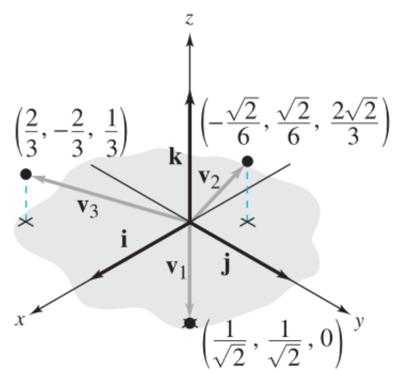
If S is a *basis*, then it is called an **orthogonal basis** or an **orthonormal basis**, respectively.
The standard basis for $\mathbb{R}^n$ is orthonormal, but it is not the only one. For instance, rotating the standard basis in $\mathbb{R}^3$ about the z-axis to form

$$B = \{(\cos\theta, \sin\theta, 0), (-\sin\theta, \cos\theta, 0), (0, 0, 1)\}$$

A Nonstandard Orthonormal Basis for $\mathbb{R}^3$

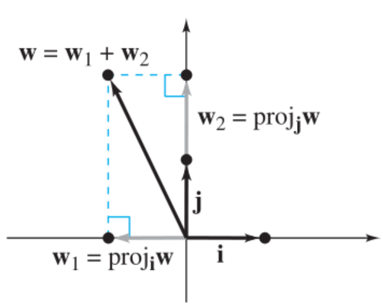
$$S = \left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(-\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}\right), \left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right)\right\}$$

3 vectors are *mutually orthogonal*. Each vector is of length 1.



THEOREM 5.10 **Orthogonal Sets Are Linearly Independent**
If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space V, then S is *linearly independent*.

COROLLARY TO THEOREM 5.10
If V is an inner profuct space of dimension n, then any orthogonal set of n nonzero vectors is a basis for V.



$\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2 = c_1\mathbf{i} + c_2\mathbf{j}$

THEOREM 5.11 **Coordinates Relative to an Orthonormal Basis**
If $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an *orthonormal* basis for an inner product space V, then the coordinate representation of a vector $\mathbf{w}$ with respect to B is

$$\mathbf{w} = \langle \mathbf{w}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{w}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{w}, \mathbf{v}_n \rangle \mathbf{v}_n$$

The coordinates of $\mathbf{w}$ relative to the *orthonormal* basis B are called the **Fourier coefficents** of $\mathbf{w}$ relative to B. The corresponding coordinate matrix of $\mathbf{w}$ relative to B is

$$[\mathbf{w}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \langle \mathbf{w}, \mathbf{v}_1 \rangle \\ \langle \mathbf{w}, \mathbf{v}_2 \rangle \\ \vdots \\ \langle \mathbf{w}, \mathbf{v}_n \rangle \end{bmatrix}$$

3.1 Gram-Schmidt Orthonormalization Process

One of the advantages of orthonormal bases is the **straighforwardness** of coordinate representation. We will now look at a procedure called the **Gram-Schmidt orthonormalization process** to find such a basis.

It has three steps.

1. Begin with a basis for the inner prodcut space. (no need to be orthorgonal nor orthonormal)
2. Convert the basis to an orthogonal basis.
3. Normalize each vector to form an orthonormal basis.

REMARK. This process leads to a matrix factorization similar to the LU-factorization.
The ***QR*-factorization** is in *Project 1*.

THEOREM 5.12 **Gram-Schmidt Orthonormalization Process**

1. Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for an inner product space V.
2. Let $B' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$, where $\mathbf{w}_i$ is given by

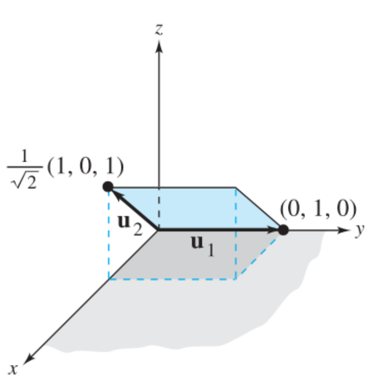
$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 \\
\mathbf{w}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2 \\
&\vdots \\
\mathbf{w}_n &= \mathbf{w}_n - \frac{\langle \mathbf{v}_n, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \frac{\langle \mathbf{v}_n, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2 - \cdots - \frac{\langle \mathbf{v}_n, \mathbf{w}_{n-1} \rangle}{\langle \mathbf{w}_{n-1}, \mathbf{w}_{n-1} \rangle}\mathbf{w}_{n-1}
\end{aligned}$$

3. Let $\mathbf{u}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}$. Then the set $B'' = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is an *orthonormal* basis for V.
   Moreover, span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ = span $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ for $k = 1, 2, \ldots, n$.

REMARK. An orthonormal set derived by the Gram-Schmidt orthonormalization process depends on the *order of the vectors* in the basis.

This process works equally well for a subspace of an inner product space.

Example 8. Applying the Gram-Schmidt Orthonormalization Process



The vectors $\mathbf{v}_1 = (0, 1, 0)$ and $\mathbf{v}_2 = (1, 1, 1)$ span a plane in $\mathbb{R}^3$. Find an orthonormal basis for this subspace.
SOLUTION. Applying the Gram-Schmidt Orthonormalization Process produces $\{(0, 1, 0), (\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2})\}$.

Example 9. Applying the Gram-Schmidt orthonormalization process to the basis $B = \{1, x, x^2\}$ in P_2, using the inner product

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\, dx$$

SOLUTION. Let $B = \{1, x, x^2\} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Then we have

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 = 1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 = x - \frac{0}{2}(1) = x \\
\mathbf{w}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2 \\
&= x^2 - \frac{2/3}{2}(1) - \frac{0}{2/3}(x) \\
&= x^2 - \frac{1}{3}
\end{aligned}$$

By normalizing these above vectors, we have $B'' = \{\frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{\sqrt{2}}x, \frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1)$.

REMARK. The polynomials $\{\frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{\sqrt{2}}x, \frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1)\}$ are called the first three **normalized Legendre polynomials**.

The computations in the Gram-Schmidt orthonormalization process are sometimes simpler when each vector $\mathbf{w}_i$ is normalized *before* it is used to determine the next vector.

The alternative form of the Gram-Schmidt orthonormalization process.

$$\begin{aligned}
\mathbf{w}_1 &= \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \\
\mathbf{w}_2 &= \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \text{where } \mathbf{w}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1 \rangle \mathbf{w}_1 \\
\mathbf{w}_3 &= \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|}, \text{where } \mathbf{w}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{w}_1 \rangle \mathbf{w}_1 - \langle \mathbf{v}_3, \mathbf{w}_2 \rangle \mathbf{w}_2 \\
&\vdots \\
\mathbf{w}_n &= \frac{\mathbf{w}_n}{\|\mathbf{w}_n\|}, \text{where } \mathbf{w}_n = \mathbf{w}_n - \langle \mathbf{v}_n, \mathbf{w}_1 \rangle \mathbf{w}_1 - \cdots - \langle \mathbf{v}_n, \mathbf{w}_{n-1} \rangle \mathbf{w}_{n-1}
\end{aligned}$$

4 Mathematical Models and Least Squares Analysis

This section is about *inconsistent* systems of linear equations and learn to find the "best possible solution" of such a system.

Example 1. Least Square Regression Line

Look for an $\mathbf{x}$ that *minimizes* the norm of the error $\|A\mathbf{x} - b\|$.

The solution $\mathbf{x} = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$ of this minimization problem is called **least square regression line** $y = c_0 + c_1 x$.

To begin, consider $A\mathbf{x} = \mathbf{b}$, where A is an $m \times n$ matrix and $\mathbf{b}$ is a column vector in $\mathbb{R}^m$. If the system is *consistent*, use Gaussian elimination with back-substitution to solve for $\mathbf{x}$. If not, however, find the "best possible" solution, which difference between $A\mathbf{x}$ and $\mathbf{b}$ is smallest.

Least Squares Problem. Given an $m \times n$ matrix A and a vector $\mathbf{b}$ in $\mathbb{R}^m$, find $\mathbf{x}$ in $\mathbb{R}^n$ such that $\|A\mathbf{x} - \mathbf{b}\|^2$ is minimized.

4.1 Orthogonal Subspaces

Definition of Orthogonal Subspaces.
The subspace S_1 and S_2 of $\mathbb{R}^n$ are **orthogonal** if $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ for all $\mathbf{v}_1$ in S_1 and $\mathbf{v}_2$ in S_2.

Example 2. $S_1 = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)$ and $S_2 = \text{span}\left(\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right)$

Notice. The zero vector is the only common vector to both S_1 and S_2. It's the only intersection.

Definition of Orthogonal Complement
If S is a subspace of $\mathbb{R}^n$, then the **orthogonal complement of S** is the set

$$S^{\perp} = \{\mathbf{u} \in \mathbb{R}^n : \mathbf{v} \cdot \mathbf{u} = 0 \quad \forall \mathbf{v} \in S\}$$

The orthogonal complement of the trivial subspace $\{\mathbf{0}\}$ is all of $\mathbb{R}^n$, and conversely, the orthogonal complement of $\mathbb{R}^n$ is the trivial subspace $\{\mathbf{0}\}$. In general, the orthogonal complement of a subspace of $\mathbb{R}^n$ is itself a *subspace* of $\mathbb{R}^n$.

Example 3. Finding the Orthogonal Complement

Find the orthogonal complement of the subspace S of $\mathbb{R}^4$ spanned by 2 column vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ of the matrix A.

$$A = \begin{array}{c} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \begin{array}{cc} \mathbf{v}_1 & \mathbf{v}_2 \end{array} \end{array}$$

SOLUTION. The orthogonal complement of S consists all the vectors $\mathbf{u}$ such that $A^T\mathbf{u} = \mathbf{0}$.

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

That is, the orthogonal complement of S is the *nullspace* of A^T.

$$S^{\perp} = N(A^T)$$

Using the techniques for solving homogeneous linear systems,you can find that a possible basis for the orthogonal complement can consist of the 2 vectors

$$\mathbf{u}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Notice. $\mathbb{R}^4$ here is split into 2 *subspaces*, $S = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ and $S^{\perp} = \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$. In fact, the 4 vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{u}_1$ and $\mathbf{u}_2$ form a basis for $\mathbb{R}^4$.

Definition of Direct Sum.
Let S_1 and S_2 be 2 *subspace* of $\mathbb{R}^n$. If each vector $\mathbf{x} \in \mathbb{R}^n$ can be *uniquely* written as a sum of a vector $\mathbf{s}_1$ from S_1 and a vector $\mathbf{s}_2$ from S_2, $\mathbf{x} = \mathbf{s}_1 + \mathbf{s}_2$, then $\mathbb{R}^n$ is **direct sum** of S_1 and S_2.

$$\mathbb{R}^n = S_1 \oplus S_2$$

Example.

(a) $S_1 = \operatorname{span}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)$ and $S_2 = \operatorname{span}\left(\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right)$, $S_1 \oplus S_2 = \mathbb{R}^2$

(b) $S = \operatorname{span}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right)$ and $S^{\perp} = \operatorname{span}\left(\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}\right)$, $S \oplus S^{\perp} = \mathbb{R}^4$

THEOREM 5.13 **Properties of Orthogonal Subspaces**
Let S be a subspace of $\mathbb{R}^n$. Then the following properties are true.

1. $\dim(S) + \dim(S^\perp) = n$
2. $\mathbb{R}^n = S \oplus S^\perp$
3. $(S^\perp)^\perp = S$

Proof. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t\}$ be a basis for S. Let A be the $n \times t$ matrix whose columns are the the basis vectors. Then $S = R(A)$, which implies that $S^\perp = N(A^T)$, where A^T is a $t \times n$ matrix of rank t. Hence, $\dim(S^\perp) = n - t$, proof complete.

Now, let's move on to projections of a vector $\mathbf{v}$ onto a subspace S. Because $\mathbb{R}^n = S \oplus S^\perp$, every vector $\mathbf{v} \in \mathbb{R}^n$ can be written as a sum of a vector from S and a vector from $S^\perp$.

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2, \quad \mathbf{v}_1 \in S, \quad \mathbf{v}_2 \in S^\perp$$

The vector $\mathbf{v}_1$ is the **projection** of $\mathbf{v}$ onto the subspace S: $\mathbf{v}_1 = \text{proj}_S\mathbf{v}$. So,

$$\mathbf{v}_2 = \mathbf{v} - \mathbf{v}_1 = \mathbf{v} - \text{proj}_S\mathbf{v}$$

which implies that the vector $\mathbf{v} - \text{proj}_S\mathbf{v}$ is orthogonal to the subspace S.

Provided with a subspace S of $\mathbb{R}^n$, we can use Gram-Schmidt orthonormalization process to calculate an orthogonal basis for S. Then compute the projection of a vector $\mathbf{v}$ is easy.

THEOREM 5.14 **Projection onto a Subspace**
If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_t\}$ is an orthonormal basis for the subspace S of $\mathbb{R}^n$, and $\mathbf{v} \in \mathbb{R}^n$, then

$$\text{proj}_S\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v} \cdot \mathbf{u}_n)\mathbf{u}_n$$

Example. **Projection onto a Subspace**

Find the projection of $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$ on to the subspace S of $\mathbb{R}^3$ spanned by the vectors

$$\mathbf{w}_1 = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_2 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

SOLUTION. By normalizing $\mathbf{w}_1$ and $\mathbf{w}_2$, we obtain an orthonormal basis for S.

$$\{\mathbf{u}_1, \mathbf{u}_2\} = \left\{ \frac{1}{\sqrt{10}}\mathbf{w}_1, \frac{1}{2}\mathbf{w}_2 \right\} = \left\{ \begin{bmatrix} 0 \\ \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

Then the projection of $\mathbf{v}$ onto S is

$$\begin{aligned} \text{proj}_S\mathbf{v} &= (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 \\ &= \frac{6}{\sqrt{10}} \begin{bmatrix} 0 \\ \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{9}{5} \\ \frac{3}{3} \end{bmatrix} \end{aligned}$$

THEOREM 5.15 **Orthogonal Projection and Distance** Let S be a subspace of $\mathbb{R}^n$, and $\mathbf{v} \in \mathbb{R}^n$. Then for all $\mathbf{u} \in S, \mathbf{u} \neq \text{proj}_S\mathbf{v}$

$$\|\mathbf{v} - \text{proj}_S\mathbf{v}\| < \|\mathbf{v} - \mathbf{u}\|$$

4.2 Fundamental Subspaces of a Matrix

Recall that if A is a $m \times n$ matrix, the column space of A is a subspace of $\mathbb{R}^m$ consisting of all vectors of the form $A\mathbf{x}, \mathbf{x} \in \mathbb{R}^n$. The **4 fundamental subspaces** of A are

$$\begin{aligned} N(A) &= \text{nullspace of } A \quad & N(A^T) &= \text{nullspace of } A^T \\ R(A) &= \text{column space of } A \quad & R(A^T) &= \text{column space of } A^T \end{aligned}$$

Example 6. Fundamental Subspaces

Find the four fundamental subspaces of the matrix

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

SOLUTION. The column space is simply the span of the first and third columns.

$$R(A) = \text{span}\left(\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}\right)$$

The column space of A^T is equivalent to the row space of A, which is spanned by the first 2 rows

$$R(A^T) = \text{span}\left(\begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right)$$

The nullspace of A is a solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$.

$$N(A) = \text{span}\left(\begin{bmatrix}-2\\1\\0\end{bmatrix}\right)$$

Finally, the nullspace of A^T is a solution space of the homogeneous system whose coefficent matrix is A^T

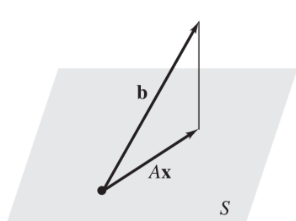
$$N(A^T) = \text{span}\left(\begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right)$$

THEOREM 5.16
Fundamental Subspaces of a Matrix

If A is an $m \times n$ matrix, then

1. $R(A)$ and $N(A^T)$ are orthogonal subspaces of $\mathbb{R}^m$.
2. $R(A^T)$ and $N(A)$ are orthogonal subspaces of $\mathbb{R}^n$.
3. $R(A) \oplus N(A^T) = \mathbb{R}^m$.
4. $R(A^T) \oplus N(A) = \mathbb{R}^n$.

4.3 Least Squares



Recall that, we are going to find a vector $\mathbf{x}$ that minimize $\|A\mathbf{x} - \mathbf{b}\|$, where A is an $m \times n$ matrix and $\mathbf{b}$ is a column vector in $\mathbb{R}^m$. Let S be the column space of A: $S = R(A)$. Assume $\mathbf{b} \notin S$, otherwise, the system $A\mathbf{x} = \mathbf{b}$ would be *consistent*.

From Theorem 5.15, we know that the desired vector is the projection of $\mathbf{b}$ onto S.

Letting $A\hat{\mathbf{x}} = \text{proj}_S\mathbf{b}$ be that projection $\implies A\hat{\mathbf{x}} - \mathbf{b} = \text{proj}_S\mathbf{b} - \mathbf{b}$ is orthogonal to $S = R(A)$.

This implies that $A\hat{\mathbf{x}} - \mathbf{b} \in R(A)^{\perp}$, which equals $N(A^T)$. This is the crucial observation: $A\hat{\mathbf{x}} - \mathbf{b}$ is in the *nullspace* of A^T.

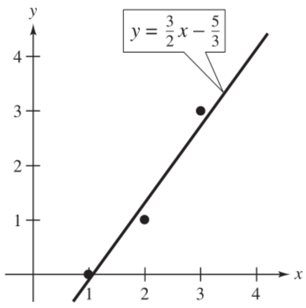
$$\begin{aligned} A^T(A\hat{\mathbf{x}} - \mathbf{b}) &= \mathbf{0} \\ A^TA\hat{\mathbf{x}} - A^T\mathbf{b} &= \mathbf{0} \\ A^TA\hat{\mathbf{x}} &= A^T\mathbf{b} \end{aligned}$$

The solution of the least squares problem comes down to solving the $n \times n$ linear S. Eq $A^TA\mathbf{x} = A^T\mathbf{b}$. These equations are called the **normal equations** of the least squares problem $A\mathbf{x} = \mathbf{b}$.

Example 7. **Solving the Normal Equations** Find the solution of the least squares problem

$$A\mathbf{x} = \mathbf{b}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$



SOLUTION. Begin by calculating the matrix products shown below

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \end{bmatrix}$$

The normal equations are

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \end{bmatrix}$$

The solution of this system of equations is $\mathbf{x} = \begin{bmatrix} -\frac{5}{3} \\ \frac{3}{2} \end{bmatrix}$, which implies that the least squares regression line for the data is $y = \frac{3}{2}x - \frac{5}{3}$.

REMARK. For an $m \times n$ matrix A, the normal equation form an $n \times n$ system of linear equations. This system is always consistent, but it may have an *infinite number of solutions*. There is a unique solution if the rank of A is n.

Example 8. **Orthogonal Projection onto a Subspace**

Find the orthogonal projection of the vector $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$ onto the column space S of the matrix

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

SOLUTION.

$$A^T A = \begin{bmatrix} 0 & 3 & 1 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 3 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 10 & 0 \\ 0 & 4 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 0 & 3 & 1 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

The normal equations are

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

$$\begin{bmatrix} 10 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

The solution of these equations is $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ \frac{1}{2} \end{bmatrix}$.

Finally, the projection of $\mathbf{b}$ onto S is

$$A\mathbf{x} = \begin{bmatrix} 0 & 2 \\ 3 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{3}{5} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{9}{5} \\ \frac{3}{5} \end{bmatrix}$$

which agrees with the solution obtained in **Example 5** (find an orthonormal basis of $R(A)$ and applying the formula).

4.4 Mathematical Modeling

Example 9. World Population

This table shows the world population for 6 different years.

Year	1980	1985	1990	1995	2000	2005
Population (y)	4.5	4.8	5.3	5.7	6.1	6.5

Let $x = 0$ represent the year 1980. Find the least squares regression quadratic polynominal $y = c_0 + c_1 x + c_2 x^2$ for these data and use the model to estimate the population for the year 2010.

SOLUTION. By substituting the data points $(0, 4.5), (5, 4.8), (10, 5.3), (15, 5.7), (20, 6.1)$ and $25, 6.5$ into the quadratic polynominal $y = c_0 + c_1 x + c_2 x^2$, we obtain the following system of linear equations.

$$\begin{cases} c_0 & = 4.5 \\ c_0 + \ \ 5c_1 + \ \ 25c_2 = 4.8 \\ c_0 + 10c_1 + 100c_2 = 5.3 \\ c_0 + 15c_1 + 225c_2 = 5.7 \\ c_0 + 20c_1 + 400c_2 = 6.1 \\ c_0 + 25c_1 + 625c_2 = 6.5 \end{cases}$$

This produces the least squares problem

$$A\mathbf{x} = \mathbf{b}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 5 & 25 \\ 1 & 10 & 100 \\ 1 & 15 & 225 \\ 1 & 20 & 400 \\ 1 & 25 & 625 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4.5 \\ 4.8 \\ 5.3 \\ 5.7 \\ 6.1 \\ 6.5 \end{bmatrix}$$

The normal equations are

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

$$\begin{bmatrix} 6 & 75 & 1375 \\ 75 & 1375 & 28,125 \\ 1375 & 28,125 & 611,875 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 32.9 \\ 447 \\ 8435 \end{bmatrix}$$

and their solution is $\mathbf{x} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} \approx \begin{bmatrix} 4.5 \\ 0.08 \\ 0 \end{bmatrix}$.

Note that $c_2 \approx 0$. So, the least squares polynominal for these data is the linear polynominal

$$y = 4.5 + 0.08x$$

Evaluating this polynominal at $x = 30$ gives the estimate of the world population for the year 2010:

$$y = 4.5 + 0.08(30) \approx 6.9 \text{ billion}$$

Example 10. Application to Astronomy

Planet	*Mercury*	*Venus*	*Earth*	*Mars*	*Jupiter*	*Saturn*
Distance, x	0.387	0.723	1.0	1.523	5.203	9.541
Period, y	0.241	0.615	1.0	1.881	11.861	29.457

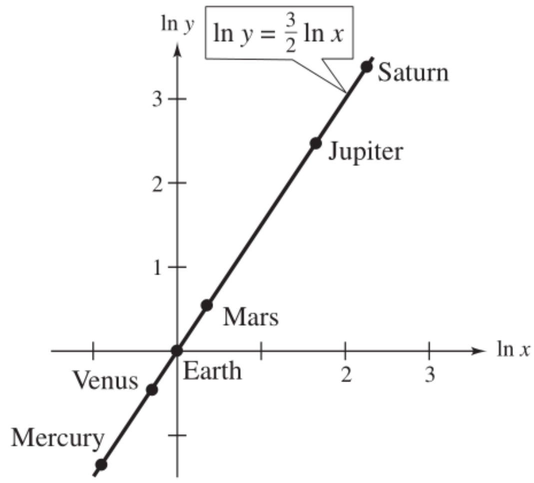
This table shows the mean distances x and the periods y of the 6 planets that are closest to the Sun.

If you plot the data as the shown, they do not seem to lie in a straight line. By taking the logarithm of each coordinate, however, you obtain points of the form $(\ln x, \ln y)$, as below

Planet	*Mercury*	*Venus*	*Earth*	*Mars*	*Jupiter*	*Saturn*
$\ln x$	–0.949	–0.324	0.0	0.421	1.649	2.256
$\ln y$	–1.423	–0.486	0.0	0.632	2.473	3.383

Using the techniques of this section, we can find the equation of the line is

$$\ln y = \frac{3}{2} \ln x \quad \text{or} \quad y = x^{3/2}$$



MORE EXAMPLE. (p.351)