Math 341 - Linear Algebra
§6.1 - 6.7

Fall 2019

Seek ye diligently and teach one another words of wisdom; yea, seek ye out of the best books words of wisdom; seek learning, even by study and also by faith.

- D&C 88:118

§6.1 Inner Product, Length, and Orthogonality

Definition (Inner (Dot) Product)

Let $\mathbf{u}$ and $\mathbf{w}$ be two vectors in $\mathbb{R}^n$. Then the *inner product* (sometimes called the *dot product*) of $\mathbf{u}$ and $\mathbf{w}$ is

$$\mathbf{u} \cdot \mathbf{w} = \mathbf{u}^T \mathbf{w}.$$

§6.1 Inner Product, Length, and Orthogonality

Definition (Inner (Dot) Product)

Let $\mathbf{u}$ and $\mathbf{w}$ be two vectors in $\mathbb{R}^n$. Then the *inner product* (sometimes called the *dot product*) of $\mathbf{u}$ and $\mathbf{w}$ is

$$\mathbf{u} \cdot \mathbf{w} = \mathbf{u}^T \mathbf{w}.$$

It is a 1×1 matrix (just a real number).

§6.1 Inner Product, Length, and Orthogonality

Definition (Inner (Dot) Product)

Let $\mathbf{u}$ and $\mathbf{w}$ be two vectors in $\mathbb{R}^n$. Then the *inner product* (sometimes called the *dot product*) of $\mathbf{u}$ and $\mathbf{w}$ is

$$\mathbf{u} \cdot \mathbf{w} = \mathbf{u}^T \mathbf{w}.$$

It is a 1×1 matrix (just a real number).

Another way of thinking of it is

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = u_1 w_1 + u_2 w_2 + \cdots + u_n w_n.$$

§6.1 Inner Product, Length, and Orthogonality

Example: If $\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}$, find $\mathbf{u} \cdot \mathbf{w}$ and $\mathbf{w} \cdot \mathbf{u}$.

§6.1 Inner Product, Length, and Orthogonality

Example: If $\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}$, find $\mathbf{u} \cdot \mathbf{w}$ and $\mathbf{w} \cdot \mathbf{u}$.

Theorem

Let $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and $c \in \mathbb{R}$ a scalar.

§6.1 Inner Product, Length, and Orthogonality

Example: If $\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}$, find $\mathbf{u} \cdot \mathbf{w}$ and $\mathbf{w} \cdot \mathbf{u}$.

Theorem

Let $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and $c \in \mathbb{R}$ a scalar. Then

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$

§6.1 Inner Product, Length, and Orthogonality

Example: If $\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}$, find $\mathbf{u} \cdot \mathbf{w}$ and $\mathbf{w} \cdot \mathbf{u}$.

Theorem

Let $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and $c \in \mathbb{R}$ a scalar. Then

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

§6.1 Inner Product, Length, and Orthogonality

Example: If $\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}$, find $\mathbf{u} \cdot \mathbf{w}$ and $\mathbf{w} \cdot \mathbf{u}$.

Theorem

Let $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and $c \in \mathbb{R}$ a scalar. Then

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
3. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$

§6.1 Inner Product, Length, and Orthogonality

Example: If $\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}$, find $\mathbf{u} \cdot \mathbf{w}$ and $\mathbf{w} \cdot \mathbf{u}$.

Theorem

Let $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and $c \in \mathbb{R}$ a scalar. Then

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
3. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
4. $\mathbf{u} \cdot \mathbf{u} \geq 0$,

§6.1 Inner Product, Length, and Orthogonality

Example: If $\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}$, find $\mathbf{u} \cdot \mathbf{w}$ and $\mathbf{w} \cdot \mathbf{u}$.

Theorem

Let $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and $c \in \mathbb{R}$ a scalar. Then

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

§6.1 Inner Product, Length, and Orthogonality

Definition (Length of a vector)

The **length** (or **norm**) of a vector $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2} \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}.$$

§6.1 Inner Product, Length, and Orthogonality

Definition (Length of a vector)

The **length** (or **norm**) of a vector $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2} \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}.$$

Fact: For any $\mathbf{v}$ and any scalar c, we have

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

Example: If

$$\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{w} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix},$$

compute $\|\mathbf{u}\|$ and $\|\mathbf{w}\|$, and find **unit vectors** in the same directions (i.e. vectors that point in the same direction but have length 1).

§6.1 Inner Product, Length, and Orthogonality

Example: If

$$\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{w} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix},$$

compute $\|\mathbf{u}\|$ and $\|\mathbf{w}\|$, and find **unit vectors** in the same directions (i.e. vectors that point in the same direction but have length 1).

The above procedure is called **normalizing** the vectors $\mathbf{u}$ and $\mathbf{w}$.

Definition

For $\mathbf{u}$ and $\mathbf{w}$ in $\mathbb{R}^n$, the **distance between u and w** is the length of the vector $\mathbf{u} - \mathbf{w}$, i.e.

$$\operatorname{dist}(\mathbf{u}, \mathbf{w}) = \|\mathbf{u} - \mathbf{w}\|.$$

Definition

For $\mathbf{u}$ and $\mathbf{w}$ in $\mathbb{R}^n$, the **distance between u and w** is the length of the vector $\mathbf{u} - \mathbf{w}$, i.e.

$$\operatorname{dist}(\mathbf{u}, \mathbf{w}) = \|\mathbf{u} - \mathbf{w}\|.$$

Example: For $\mathbf{u}, \mathbf{w}$ as above, find $\operatorname{dist}(\mathbf{u}, \mathbf{w})$.

§6.1 Inner Product, Length, and Orthogonality

Definition (Orthogonality)

Vectors $\mathbf{u}$ and $\mathbf{w}$ in $\mathbb{R}^n$ are said to be **orthogonal** if $\mathbf{u} \cdot \mathbf{w} = 0$.

§6.1 Inner Product, Length, and Orthogonality

Definition (Orthogonality)

Vectors $\mathbf{u}$ and $\mathbf{w}$ in $\mathbb{R}^n$ are said to be **orthogonal** if $\mathbf{u} \cdot \mathbf{w} = 0$.

Think of them as being perpendicular.

§6.1 Inner Product, Length, and Orthogonality

Definition (Orthogonality)

Vectors $\mathbf{u}$ and $\mathbf{w}$ in $\mathbb{R}^n$ are said to be **orthogonal** if $\mathbf{u} \cdot \mathbf{w} = 0$.

Think of them as being perpendicular.

Theorem

Two vectors $\mathbf{u}$ and $\mathbf{w}$ are orthogonal if and only if

$$\|\mathbf{u} + \mathbf{w}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{w}\|^2.$$

Definition (Orthogonal complement)

Let W be a subspace of $\mathbb{R}^n$. Then a vector $\mathbf{u} \in \mathbb{R}^n$ is said to be **orthogonal to** W if it is orthogonal to every vector in W.

Definition (Orthogonal complement)

Let W be a subspace of $\mathbb{R}^n$. Then a vector $\mathbf{u} \in \mathbb{R}^n$ is said to be **orthogonal to** W if it is orthogonal to every vector in W.

The set of all vectors in $\mathbb{R}^n$ which are orthogonal to W is called the **orthogonal complement of** W, and is denoted $W^\perp$.

§6.1 Inner Product, Length, and Orthogonality

Definition (Orthogonal complement)

Let W be a subspace of $\mathbb{R}^n$. Then a vector $\mathbf{u} \in \mathbb{R}^n$ is said to be **orthogonal to** W if it is orthogonal to every vector in W.

The set of all vectors in $\mathbb{R}^n$ which are orthogonal to W is called the **orthogonal complement of** W, and is denoted $W^\perp$.

Example: Let W be a line in $\mathbb{R}^3$. What is $W^\perp$?

Definition (Orthogonal complement)

Let W be a subspace of $\mathbb{R}^n$. Then a vector $\mathbf{u} \in \mathbb{R}^n$ is said to be **orthogonal to** W if it is orthogonal to every vector in W.

The set of all vectors in $\mathbb{R}^n$ which are orthogonal to W is called the **orthogonal complement of** W, and is denoted $W^\perp$.

Example: Let W be a line in $\mathbb{R}^3$. What is $W^\perp$?

Example: Let W be a plane in $\mathbb{R}^3$. What is $W^\perp$?

Theorem

If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span W, then a vector is in $W^\perp$ if and only if it is orthogonal to each of the $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Theorem

If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span W, then a vector is in $W^\perp$ if and only if it is orthogonal to each of the $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Theorem

$W^\perp$ is a subspace of $\mathbb{R}^n$.

§6.1 Inner Product, Length, and Orthogonality

Theorem

Let A be an $m \times n$ matrix. Then

§6.1 Inner Product, Length, and Orthogonality

Theorem

Let A be an $m \times n$ matrix. Then

$$(\text{Row } A)^{\perp} = \text{Nul } A \qquad \text{and} \qquad (\text{Col } A)^{\perp} = \text{Nul } A^T.$$

Theorem

Let A be an $m \times n$ matrix. Then

$$(\text{Row } A)^{\perp} = \text{Nul } A \qquad \text{and} \qquad (\text{Col } A)^{\perp} = \text{Nul } A^T.$$

Fact: If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in $\mathbb{R}^3$, then

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

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

§6.1 Inner Product, Length, and Orthogonality

Theorem

Let A be an $m \times n$ matrix. Then

$$(\text{Row } A)^{\perp} = \text{Nul } A \qquad \text{and} \qquad (\text{Col } A)^{\perp} = \text{Nul } A^T.$$

Fact: If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in $\mathbb{R}^3$, then

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

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

Example: Find the angle between the vectors $\mathbf{u}$ and $\mathbf{w}$ (from above).

§6.2 Orthogonal Sets

Definition

A set of vectors is said to be an **orthogonal set** if every vector is orthogonal to every other vector in the set.

Definition

A set of vectors is said to be an **orthogonal set** if every vector is orthogonal to every other vector in the set.

Example: The vectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix}$$

form an orthogonal set, because

Definition

A set of vectors is said to be an **orthogonal set** if every vector is orthogonal to every other vector in the set.

Example: The vectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix}$$

form an orthogonal set, because

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{0},$$

§6.2 Orthogonal Sets

Definition

A set of vectors is said to be an **orthogonal set** if every vector is orthogonal to every other vector in the set.

Example: The vectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix}$$

form an orthogonal set, because

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{0}, \qquad \mathbf{v}_2 \cdot \mathbf{v}_3 = \mathbf{0},$$

§6.2 Orthogonal Sets

Definition

A set of vectors is said to be an **orthogonal set** if every vector is orthogonal to every other vector in the set.

Example: The vectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix}$$

form an orthogonal set, because

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{0}, \qquad \mathbf{v}_2 \cdot \mathbf{v}_3 = \mathbf{0}, \qquad \text{and} \qquad \mathbf{v}_1 \cdot \mathbf{v}_3 = \mathbf{0}.$$

Theorem

A set of *nonzero* orthogonal vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is linearly independent, and hence forms a basis of the subspace it spans.

§6.2 Orthogonal Sets

Theorem

A set of *nonzero* orthogonal vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is linearly independent, and hence forms a basis of the subspace it spans.

Definition

An **orthogonal basis** for a subspace W is a basis for W that is also an orthogonal set.

§6.2 Orthogonal Sets

Theorem

A set of *nonzero* orthogonal vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is linearly independent, and hence forms a basis of the subspace it spans.

Definition

An **orthogonal basis** for a subspace W is a basis for W that is also an orthogonal set.

Theorem

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be an orthogonal basis for a subspace W. For each $\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_m\mathbf{u}_m$, the weight c_j is given by

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}.$$

Example: The vectors $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix}$$

form an orthogonal basis for $\mathbb{R}^3$.

Example: The vectors $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix}$$

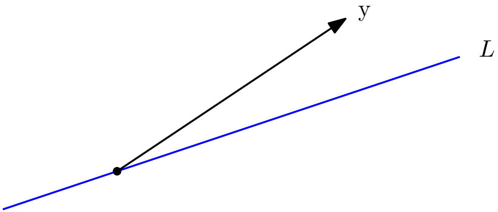
form an orthogonal basis for $\mathbb{R}^3$. Find the coordinates of

$$\mathbf{y} = \begin{bmatrix} 2 \\ 3 \\ -5 \end{bmatrix}$$

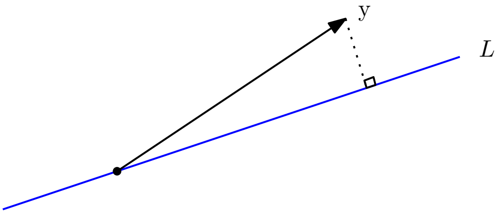
relative to the basis $\mathcal{B}$.

§6.2 Orthogonal Sets

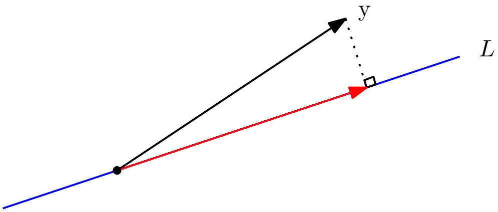
Suppose we have a line L and a vector $\mathbf{y}$, how do we find the point on L which is closest to $\mathbf{y}$?



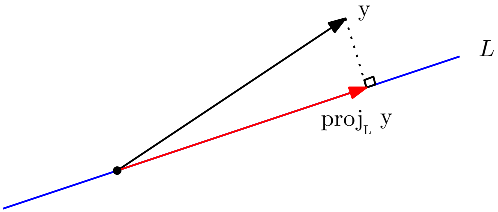
Suppose we have a line L and a vector $\mathbf{y}$, how do we find the point on L which is closest to $\mathbf{y}$?



Suppose we have a line L and a vector $\mathbf{y}$, how do we find the point on L which is closest to $\mathbf{y}$?

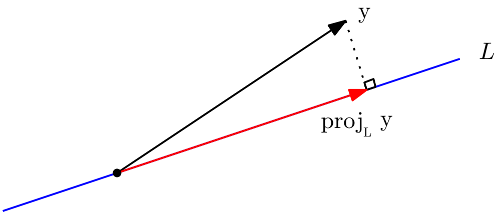


Suppose we have a line L and a vector $\mathbf{y}$, how do we find the point on L which is closest to $\mathbf{y}$?



§6.2 Orthogonal Sets

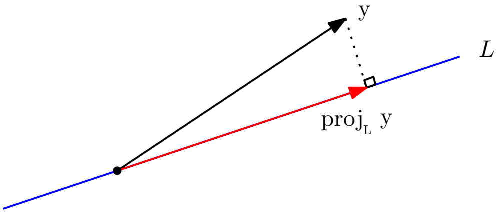
Suppose we have a line L and a vector $\mathbf{y}$, how do we find the point on L which is closest to $\mathbf{y}$?



Suppose the line L is spanned by some vector $\mathbf{u}$.

§6.2 Orthogonal Sets

Suppose we have a line L and a vector $\mathbf{y}$, how do we find the point on L which is closest to $\mathbf{y}$?



Suppose the line L is spanned by some vector $\mathbf{u}$.

Then the vector on L which is closest to $\mathbf{y}$ is called the **orthogonal projection** of $\mathbf{y}$ onto L, and is given by

$$\text{proj}_L\ \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}.$$

Notice also, that the vector

$$\mathbf{y} - \operatorname{proj}_{\mathbf{u}} \mathbf{y} = \mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}$$

is orthogonal to $\mathbf{u}$:

Notice also, that the vector

$$\mathbf{y} - \mathrm{proj}_{\mathbf{u}}\, \mathbf{y} = \mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}$$

is orthogonal to $\mathbf{u}$:

$$\mathbf{u} \cdot \left(\mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} \right)$$

Notice also, that the vector

$$\mathbf{y} - \operatorname{proj}_{\mathbf{u}} \mathbf{y} = \mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u}$$

is orthogonal to $\mathbf{u}$:

$$\mathbf{u} \cdot \left(\mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u}\right) = \mathbf{u} \cdot \mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u} \cdot \mathbf{u}$$

Notice also, that the vector

$$\mathbf{y} - \operatorname{proj}_{\mathbf{u}} \mathbf{y} = \mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u}$$

is orthogonal to $\mathbf{u}$:

$$\mathbf{u} \cdot \left(\mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u}\right) = \mathbf{u} \cdot \mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u} \cdot \mathbf{u} = \mathbf{u} \cdot \mathbf{y} - \mathbf{y} \cdot \mathbf{u}$$

Notice also, that the vector

$$\mathbf{y} - \operatorname{proj}_{\mathbf{u}} \mathbf{y} = \mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$$

is orthogonal to $\mathbf{u}$:

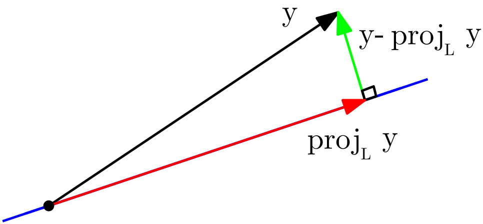
$$\mathbf{u}\cdot\left(\mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}\right) = \mathbf{u}\cdot\mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}\cdot\mathbf{u} = \mathbf{u}\cdot\mathbf{y} - \mathbf{y}\cdot\mathbf{u} = 0.$$

Notice also, that the vector

$$\mathbf{y} - \mathrm{proj}_{\mathbf{u}}\, \mathbf{y} = \mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$$

is orthogonal to $\mathbf{u}$:

$$\mathbf{u}\cdot\left(\mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}\right) = \mathbf{u}\cdot\mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}\cdot\mathbf{u} = \mathbf{u}\cdot\mathbf{y} - \mathbf{y}\cdot\mathbf{u} = 0.$$

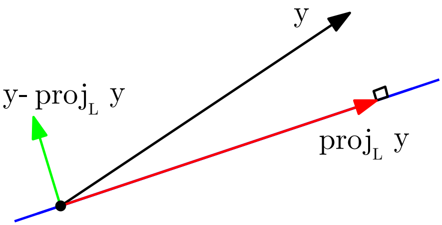


Notice also, that the vector

$$\mathbf{y} - \mathrm{proj}_{\mathbf{u}}\, \mathbf{y} = \mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$$

is orthogonal to **u**:

$$\mathbf{u}\cdot\left(\mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}\right) = \mathbf{u}\cdot\mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}\cdot\mathbf{u} \;= \mathbf{u}\cdot\mathbf{y} - \mathbf{y}\cdot\mathbf{u} = 0.$$



Thus we can write $\mathbf{y}$ as

$$\mathbf{y} = \operatorname{proj}_{\mathbf{u}} \mathbf{y} + (\mathbf{y} - \operatorname{proj}_{\mathbf{u}} \mathbf{y})$$

Thus we can write $\mathbf{y}$ as

$$\mathbf{y} = \mathrm{proj}_{\mathbf{u}}\, \mathbf{y} + (\mathbf{y} - \mathrm{proj}_{\mathbf{u}}\, \mathbf{y})$$

the sum of a vector parallel to $\mathbf{u}$ and a vector orthogonal to $\mathbf{u}$.

Thus we can write $\mathbf{y}$ as

$$\mathbf{y} = \mathrm{proj}_{\mathbf{u}}\, \mathbf{y} + (\mathbf{y} - \mathrm{proj}_{\mathbf{u}}\, \mathbf{y})$$

the sum of a vector parallel to $\mathbf{u}$ and a vector orthogonal to $\mathbf{u}$.

Example: Let

$$\mathbf{u} = \begin{bmatrix} 3 \\ 6 \\ -1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

Thus we can write $\mathbf{y}$ as

$$\mathbf{y} = \mathrm{proj}_{\mathbf{u}}\, \mathbf{y} + (\mathbf{y} - \mathrm{proj}_{\mathbf{u}}\, \mathbf{y})$$

the sum of a vector parallel to $\mathbf{u}$ and a vector orthogonal to $\mathbf{u}$.

Example: Let

$$\mathbf{u} = \begin{bmatrix} 3 \\ 6 \\ -1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

Write the vector $\mathbf{y}$ as the sum of two vectors, one orthogonal to $\mathbf{u}$ and one in $\mathrm{Span}\{\mathbf{u}\}$.

Definition

A set of orthogonal vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is called **orthonormal** if each $\mathbf{v}_j$ is a unit vector

Definition

A set of orthogonal vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is called **orthonormal** if each $\mathbf{v}_j$ is a unit vector (i.e. each $\|\mathbf{v}_j\| = 1$).

§6.2 Orthogonal Sets

Definition

A set of orthogonal vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is called **orthonormal** if each $\mathbf{v}_j$ is a unit vector (i.e. each $\|\mathbf{v}_j\| = 1$).

Example: Write the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix}$$

as an orthonormal set.

Let

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & -1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}.$$

Let

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & -1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}.$$

(the columns of U are the orthonormal set from the previous slide).

Let

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & -1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}.$$

(the columns of U are the orthonormal set from the previous slide).

What do you notice about the matrix products UU^T and U^TU?

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

If U is square ($n \times n$) then this implies that U is invertible, and $U^{-1} = U^T$.

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

If U is square ($n \times n$) then this implies that U is invertible, and $U^{-1} = U^T$.

In this case, we say that U is an orthogonal matrix.

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

If U is square ($n \times n$) then this implies that U is invertible, and $U^{-1} = U^T$.

In this case, we say that U is an orthogonal matrix.

Theorem

Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$.

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

If U is square ($n \times n$) then this implies that U is invertible, and $U^{-1} = U^T$.

In this case, we say that U is an orthogonal matrix.

Theorem

Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then

1. $\|U\mathbf{x}\| = \|\mathbf{x}\|$

§6.2 Orthogonal Sets

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

If U is square $(n \times n)$ then this implies that U is invertible, and $U^{-1} = U^T$.

In this case, we say that U is an orthogonal matrix.

Theorem

Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then

1. $\|U\mathbf{x}\| = \|\mathbf{x}\|$
2. $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$

Theorem
An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

If U is square ($n \times n$) then this implies that U is invertible, and $U^{-1} = U^T$.

In this case, we say that U is an orthogonal matrix.

Theorem
Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then

1. $\|U\mathbf{x}\| = \|\mathbf{x}\|$
2. $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$
3. $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.

Example: Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_5\}$ be an orthogonal basis for $\mathbb{R}^5$,

Example: Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_5\}$ be an orthogonal basis for $\mathbb{R}^5$, and let $W = \text{Span}\ \{\mathbf{u}_1, \mathbf{u}_2\}$.

Example: Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_5\}$ be an orthogonal basis for $\mathbb{R}^5$, and let $W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2\}$. Write the vector

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_5\mathbf{u}_5$$

as the sum of a vector in W and a vector in $W^\perp$.

Theorem (Orthogonal Decomposition Theorem)

Let W be a subspace of $\mathbb{R}^n$. Then any $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$.

§6.3 Orthogonal Projections

Theorem (Orthogonal Decomposition Theorem)

Let W be a subspace of $\mathbb{R}^n$. Then any $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$.

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W,

Theorem (Orthogonal Decomposition Theorem)

Let W be a subspace of $\mathbb{R}^n$. Then any $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$.

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = \mathrm{proj}_W \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p,$$

Theorem (Orthogonal Decomposition Theorem)

Let W be a subspace of $\mathbb{R}^n$. Then any $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as

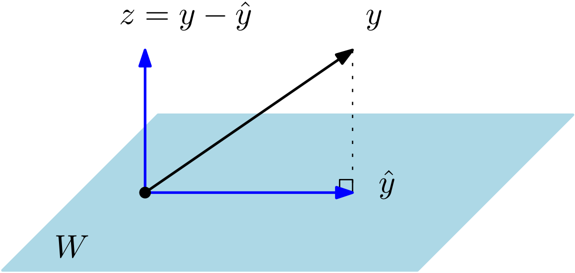
$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$.

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = \mathrm{proj}_W \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p,$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.


$z = y - \hat{y}$
y
$\hat{y}$
W

Example: Let

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} 5 \\ 7 \\ -2 \end{bmatrix}.$$

Example: Let

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} 5 \\ 7 \\ -2 \end{bmatrix}.$$

Let $W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2\}$.

Example: Let

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} 5 \\ 7 \\ -2 \end{bmatrix}.$$

Let $W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector in $W^{\perp}$.

Theorem (Best Approximation Theorem)

If W is a subspace of $\mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$, with $\hat{\mathbf{y}}$ the orthogonal projection of $\mathbf{y}$ onto W.

Theorem (Best Approximation Theorem)

If W is a subspace of $\mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$, with $\hat{\mathbf{y}}$ the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$,

Theorem (Best Approximation Theorem)

If W is a subspace of $\mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$, with $\hat{\mathbf{y}}$ the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, i.e.

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

for any vector $\mathbf{v} \in W$ not equal to $\hat{\mathbf{y}}$.

Theorem (Best Approximation Theorem)

If W is a subspace of $\mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$, with $\hat{\mathbf{y}}$ the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, i.e.

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

for any vector $\mathbf{v} \in W$ not equal to $\hat{\mathbf{y}}$.

We sometimes call $\hat{\mathbf{y}}$ the **best approximation** to $\mathbf{y}$ by an element of W.

Theorem (Best Approximation Theorem)

If W is a subspace of $\mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$, with $\hat{\mathbf{y}}$ the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, i.e.

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

for any vector $\mathbf{v} \in W$ not equal to $\hat{\mathbf{y}}$.

We sometimes call $\hat{\mathbf{y}}$ the **best approximation** to $\mathbf{y}$ by an element of W.

$\|\mathbf{y} - \hat{\mathbf{y}}\|$ is called the **distance** from $\mathbf{y}$ to W.

Example: Let

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} 5 \\ 7 \\ -2 \end{bmatrix}$$

Example: Let

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} 5 \\ 7 \\ -2 \end{bmatrix}$$

Let $W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2\}$.

Example: Let

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} 5 \\ 7 \\ -2 \end{bmatrix}$$

Let $W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2\}$. Find the best approximation to $\mathbf{y}$ by an element of W, and find the distance from $\mathbf{y}$ to W.

Theorem

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, then

$$\operatorname{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p.$$

Theorem

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, then

$$\operatorname{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p.$$

If $U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ]$,

§6.3 Orthogonal Projections

Theorem

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W\ \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p.$$

If $U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ]$, then

$$\text{proj}_W\ \mathbf{y} = UU^T\mathbf{y} \qquad \text{for all } \mathbf{y} \in \mathbb{R}^n.$$

Theorem

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W\ \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p.$$

If $U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ]$, then

$$\text{proj}_W\ \mathbf{y} = UU^T\mathbf{y} \qquad \text{for all } \mathbf{y} \in \mathbb{R}^n.$$

Recall: If U has orthonormal columns then $U^TU = I$,

Theorem

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W\ \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p.$$

If $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_p \end{bmatrix}$, then

$$\text{proj}_W\ \mathbf{y} = UU^T\mathbf{y} \qquad \text{for all } \mathbf{y} \in \mathbb{R}^n.$$

Recall: If U has orthonormal columns then $U^TU = I$, while the theorem says that UU^T gives the projection map onto Col U.

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_p \end{bmatrix}.$$

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let

$$U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ].$$

Then $W = \text{Span}\ \{\mathbf{u}_1, \ldots, \mathbf{u}_p\} = \text{Col}\ U$.

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let

$$U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ].$$

Then $W = \text{Span}\ \{\mathbf{u}_1, \ldots, \mathbf{u}_p\} = \text{Col}\ U$. Then

$$U^T U = I$$

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let

$$U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ].$$

Then $W = \text{Span}\ \{\mathbf{u}_1, \ldots, \mathbf{u}_p\} = \text{Col}\ U$. Then

$$U^T U = I \qquad \text{and} \qquad UU^T\mathbf{y} = \text{proj}_W\ \mathbf{y} \qquad \text{for all } \mathbf{y} \in \mathbb{R}^n$$

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let

$$U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ].$$

Then $W = \text{Span}\ \{\mathbf{u}_1, \ldots, \mathbf{u}_p\} = \text{Col}\ U$. Then

$$U^T U = I \qquad \text{and} \qquad U U^T \mathbf{y} = \text{proj}_W\ \mathbf{y} \qquad \text{for all } \mathbf{y} \in \mathbb{R}^n$$

If $p = n$,

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let

$$U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ].$$

Then $W = \text{Span}\ \{\mathbf{u}_1, \ldots, \mathbf{u}_p\} = \text{Col}\ U$. Then

$$U^T U = I \qquad \text{and} \qquad U U^T \mathbf{y} = \text{proj}_W\ \mathbf{y} \qquad \text{for all } \mathbf{y} \in \mathbb{R}^n$$

If $p = n$, then $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis for all of $\mathbb{R}^n = W = \text{Col}\ U$

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let

$$U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ].$$

Then $W = \text{Span}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_p\} = \text{Col}\ U$. Then

$$U^T U = I \qquad \text{and} \qquad UU^T\mathbf{y} = \text{proj}_W\ \mathbf{y} \qquad \text{for all } \mathbf{y} \in \mathbb{R}^n$$

If $p = n$, then $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis for all of $\mathbb{R}^n = W = \text{Col}\ U$

$\Rightarrow \text{proj}_W\ \mathbf{y} = \mathbf{y}$ for all $\mathbf{y} \in \mathbb{R}^n$

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let

$$U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\ ].$$

Then $W = \text{Span}\ \{\mathbf{u}_1, \ldots, \mathbf{u}_p\} = \text{Col}\ U$. Then

$$U^T U = I \qquad \text{and} \qquad U U^T \mathbf{y} = \text{proj}_W\ \mathbf{y} \qquad \text{for all } \mathbf{y} \in \mathbb{R}^n$$

If $p = n$, then $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis for all of $\mathbb{R}^n = W = \text{Col}\ U$

$\Rightarrow \text{proj}_W\ \mathbf{y} = \mathbf{y}$ for all $\mathbf{y} \in \mathbb{R}^n$

$\Rightarrow UU^T = I$.

Example: For the following orthonormal vectors

$$\mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \\ -\frac{1}{2} \\ \\ \frac{1}{2} \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 0 \\ \\ \frac{1}{\sqrt{2}} \\ \\ \frac{1}{\sqrt{2}} \end{bmatrix},$$

Example: For the following orthonormal vectors

$$\mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix},$$

use the theorem to compute $\text{proj}_W\ \mathbf{y}$, where $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, and

$$\mathbf{y} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}.$$

§6.3 Orthogonal Projections

Warm up: For the following orthonormal vectors

$$\mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

compute $\mathrm{proj}_W \mathbf{y}$, where $W = \mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, and

$$\mathbf{y} = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}.$$

Question: Suppose we have a subspace W of $\mathbb{R}^n$.

§6.4 Gram-Schmidt Process

Question: Suppose we have a subspace W of $\mathbb{R}^n$. Can we find an orthogonal basis for W?

§6.4 Gram-Schmidt Process

Question: Suppose we have a subspace W of $\mathbb{R}^n$. Can we find an orthogonal basis for W? If so, how?

§6.4 Gram-Schmidt Process

Question: Suppose we have a subspace W of $\mathbb{R}^n$. Can we find an orthogonal basis for W? If so, how?

What if we start with a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ of W.

§6.4 Gram-Schmidt Process

Question: Suppose we have a subspace W of $\mathbb{R}^n$. Can we find an orthogonal basis for W? If so, how?

What if we start with a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ of W. Can we modify it so that if becomes orthogonal?

§6.4 Gram-Schmidt Process

Question: Suppose we have a subspace W of $\mathbb{R}^n$. Can we find an orthogonal basis for W? If so, how?

What if we start with a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ of W. Can we modify it so that if becomes orthogonal?

Example: Suppose we start with $\mathbf{u}, \mathbf{w} \in \mathbb{R}^3$ which aren't orthogonal.

§6.4 Gram-Schmidt Process

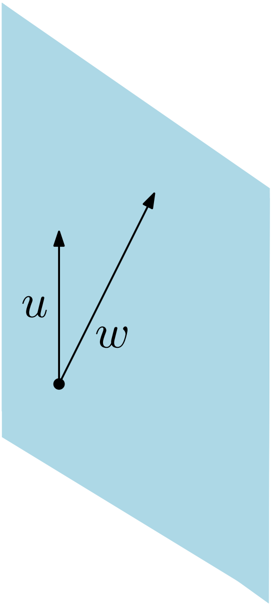
Question: Suppose we have a subspace W of $\mathbb{R}^n$. Can we find an orthogonal basis for W? If so, how?

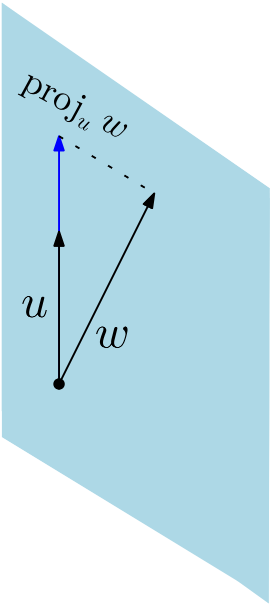
What if we start with a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ of W. Can we modify it so that if becomes orthogonal?

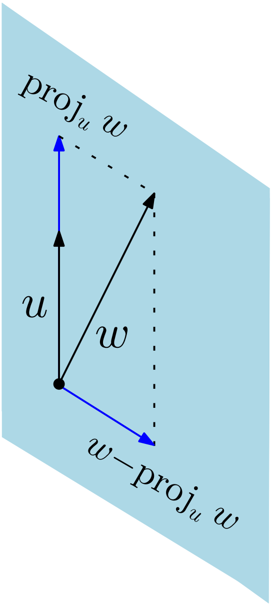
Example: Suppose we start with $\mathbf{u}, \mathbf{w} \in \mathbb{R}^3$ which aren't orthogonal.

How can we change them so that they become orthogonal and so that their span does not change?

§6.4 Gram-Schmidt Process




$\mathrm{proj}_u\, w$
u
w



$\text{proj}_u\, w$
u
w
$w - \text{proj}_u\, w$

Theorem (The Gram-Schmidt Process)

Given a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$$

etc.

§6.4 Gram-Schmidt Process

Theorem (The Gram-Schmidt Process)

Given a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$$

etc. Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for W.

Theorem (The Gram-Schmidt Process)

Given a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$$

etc. Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for W. In addition,

$$\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \mathrm{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \quad \text{for } 1 \le k \le p.$$

Example: Let W be the subspace of $\mathbb{R}^4$ spanned by

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 2 \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Example: Let W be the subspace of $\mathbb{R}^4$ spanned by

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 2 \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Find an orthogonal basis for W using the Gram-Schmidt process.

Example: Let W be the subspace of $\mathbb{R}^4$ spanned by

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 2 \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Find an orthogonal basis for W using the Gram-Schmidt process.

Example: Find an orthonormal basis of the subspace W.

Applying Gram-Schmidt gives an orthogonal basis for W:

$$\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 2 \end{bmatrix}$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} 8/7 \\ 1/7 \\ -1/7 \\ -5/7 \end{bmatrix}$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \begin{bmatrix} 2/13 \\ 10/13 \\ 16/13 \\ 2/13 \end{bmatrix}.$$

Normalizing this orthogonal basis gives an orthonormal basis for W:

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{7} \\ 1/\sqrt{7} \\ -1/\sqrt{7} \\ 2/\sqrt{7} \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 8/\sqrt{91} \\ 1/\sqrt{91} \\ -1/\sqrt{91} \\ -5/\sqrt{91} \end{bmatrix},$$

and

$$\mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{91} \\ 5/\sqrt{91} \\ 8/\sqrt{91} \\ 1/\sqrt{91} \end{bmatrix}.$$

Theorem (The QR Factorization)

Let A be an $m \times n$ matrix with linearly independent columns.

Theorem (The QR Factorization)

Let A be an $m \times n$ matrix with linearly independent columns. Then A can be factored as

$$A = QR,$$

where Q is an $m \times n$ matrix whose columns form an orthonormal basis for Col A, and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Theorem (The QR Factorization)

Let A be an $m \times n$ matrix with linearly independent columns. Then A can be factored as

$$A = QR,$$

where Q is an $m \times n$ matrix whose columns form an orthonormal basis for Col A, and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

A factorization of this form is called a **QR factorization of the matrix** A.

The columns of the matrix Q can be obtained by performing Gram-Schmidt on the columns of A.

The columns of the matrix Q can be obtained by performing Gram-Schmidt on the columns of A.

Since the columns of Q are orthonormal, we have that $Q^T Q = I$.

The columns of the matrix Q can be obtained by performing Gram-Schmidt on the columns of A.

Since the columns of Q are orthonormal, we have that $Q^TQ = I$.

Since $A = QR$, we can use this to find the matrix R as

$$Q^TA$$

The columns of the matrix Q can be obtained by performing Gram-Schmidt on the columns of A.

Since the columns of Q are orthonormal, we have that $Q^T Q = I$.

Since $A = QR$, we can use this to find the matrix R as

$$Q^T A = Q^T QR$$

The columns of the matrix Q can be obtained by performing Gram-Schmidt on the columns of A.

Since the columns of Q are orthonormal, we have that $Q^TQ = I$.

Since $A = QR$, we can use this to find the matrix R as

$$Q^TA = Q^TQR = IR = R.$$

Example: Find a QR factorization of the matrix

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

We've already done the hard work using Gram-Schmidt to find an orthonormal basis for Col A. This was the basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ we found on the previous slides. Then

$$Q = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{7}} & \frac{8}{\sqrt{91}} & \frac{1}{\sqrt{91}} \\ \frac{1}{\sqrt{7}} & \frac{1}{\sqrt{91}} & \frac{5}{\sqrt{91}} \\ -\frac{1}{\sqrt{7}} & -\frac{1}{\sqrt{91}} & \frac{8}{\sqrt{91}} \\ \frac{2}{\sqrt{7}} & -\frac{5}{\sqrt{91}} & \frac{1}{\sqrt{91}} \end{bmatrix}.$$

Finding the R matrix is now easy-peasy:

$$R = Q^T A = \begin{bmatrix} \sqrt{7} & -\frac{1}{\sqrt{7}} & \frac{2}{\sqrt{7}} \\ 0 & \sqrt{\frac{13}{7}} & -\frac{5}{\sqrt{91}} \\ 0 & 0 & 2\sqrt{\frac{7}{13}} \end{bmatrix}.$$

We can check that indeed $A = QR$, and that both Q and R have the properties stated in the theorem (Q has orthonormal columns with the same span as the columns of A, and R is an upper triangular invertible matrix with positive entries on the diagonal).

Example: Find a QR factorization of the matrix

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

§6.5 Least-Squares Problems

Many real world problems involve finding solutions $\mathbf{x}$ for equations of the form

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

where A is an $m \times n$ matrix, and $\mathbf{b} \in \mathbb{R}^m$.

§6.5 Least-Squares Problems

Many real world problems involve finding solutions $\mathbf{x}$ for equations of the form

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

where A is an $m \times n$ matrix, and $\mathbf{b} \in \mathbb{R}^m$.

Often, the matrix A and vector $\mathbf{b}$ are obtained by some observations (and may be prone to experimental or observational error).

§6.5 Least-Squares Problems

Many real world problems involve finding solutions $\mathbf{x}$ for equations of the form

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

where A is an $m \times n$ matrix, and $\mathbf{b} \in \mathbb{R}^m$.

Often, the matrix A and vector $\mathbf{b}$ are obtained by some observations (and may be prone to experimental or observational error).

Many times no solution $\mathbf{x}$ might exist.

§6.5 Least-Squares Problems

If we can't find a solution $\mathbf{x}$ which satisfies $A\mathbf{x} = \mathbf{b}$ exactly, we can try to find one that makes $A\mathbf{x}$ as close as possible to $\mathbf{b}$,

§6.5 Least-Squares Problems

If we can't find a solution $\mathbf{x}$ which satisfies $A\mathbf{x} = \mathbf{b}$ exactly, we can try to find one that makes $A\mathbf{x}$ as close as possible to $\mathbf{b}$, i.e.

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

as close to $\mathbf{0}$ as possible.

§6.5 Least-Squares Problems

If we can't find a solution $\mathbf{x}$ which satisfies $A\mathbf{x} = \mathbf{b}$ exactly, we can try to find one that makes $A\mathbf{x}$ as close as possible to $\mathbf{b}$, i.e.

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

as close to $\mathbf{0}$ as possible.

Definition (Least squares solution)

A **least-squares solution** of the equation $A\mathbf{x} = \mathbf{b}$ is a vector $\hat{\mathbf{x}} \in \mathbb{R}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

for all $\mathbf{x} \in \mathbb{R}^n$.

§6.5 Least-Squares Problems

If we can't find a solution $\mathbf{x}$ which satisfies $A\mathbf{x} = \mathbf{b}$ exactly, we can try to find one that makes $A\mathbf{x}$ as close as possible to $\mathbf{b}$, i.e.

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

as close to $\mathbf{0}$ as possible.

Definition (Least squares solution)

A **least-squares solution** of the equation $A\mathbf{x} = \mathbf{b}$ is a vector $\hat{\mathbf{x}} \in \mathbb{R}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

for all $\mathbf{x} \in \mathbb{R}^n$.

i.e. $\hat{\mathbf{x}}$ is the best possible approximation to a solution of $A\mathbf{x} = \mathbf{b}$.

§6.5 Least-Squares Problems

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution

§6.5 Least-Squares Problems

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff \mathbf{b} \in \text{Col } A$.

§6.5 Least-Squares Problems

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

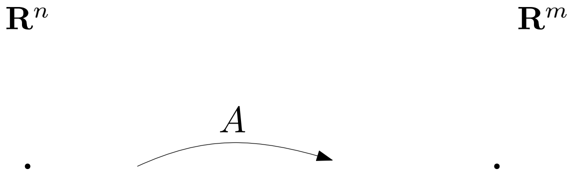
Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff \mathbf{b} \in \text{Col } A$.

$\mathbf{R}^n$ $\mathbf{R}^m$

§6.5 Least-Squares Problems

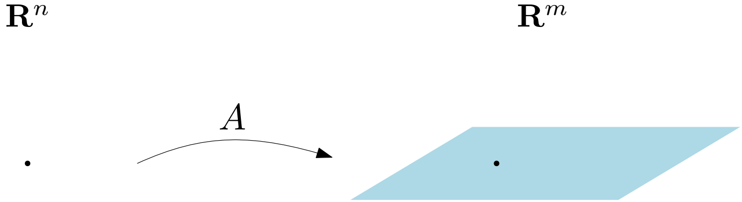
Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

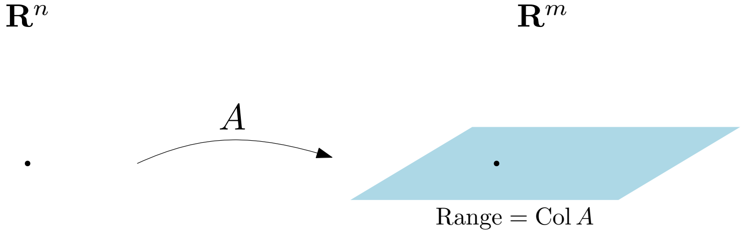
Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



§6.5 Least-Squares Problems

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

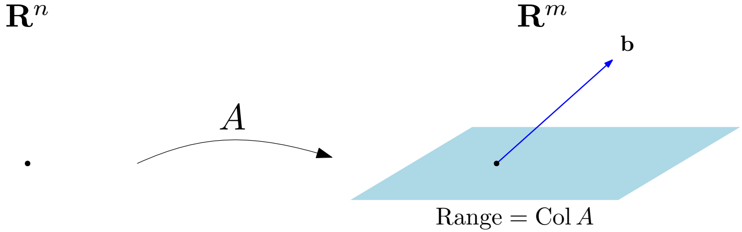
Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



§6.5 Least-Squares Problems

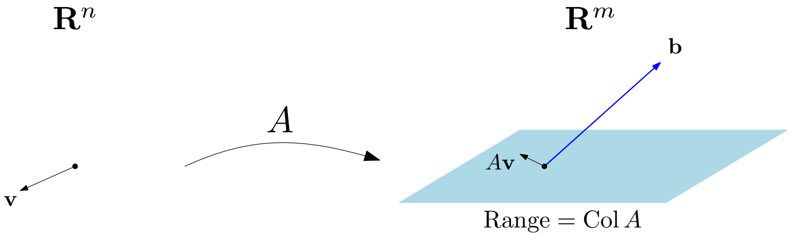
Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



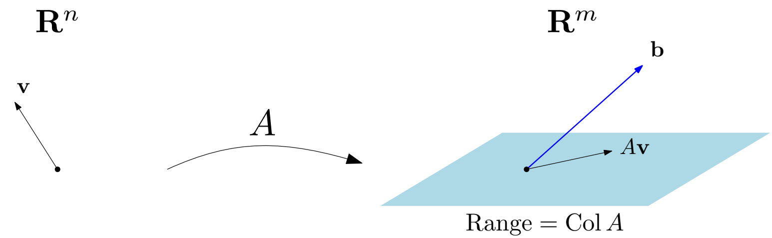
Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff \mathbf{b} \in$ Col A.



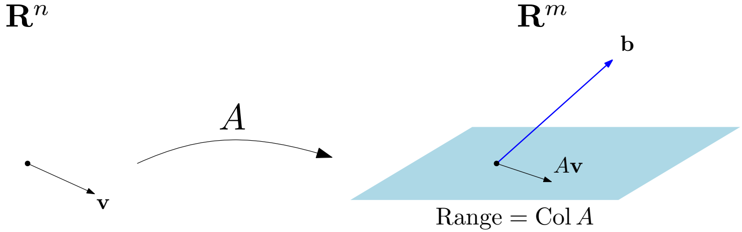
Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

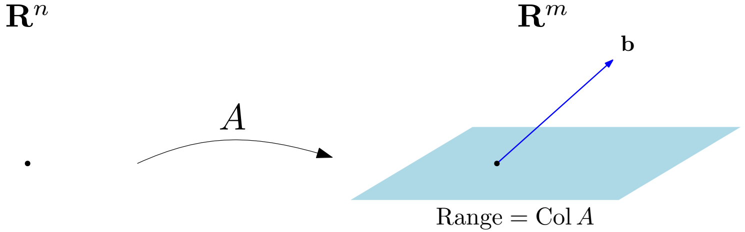
Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\Longleftrightarrow$ $\mathbf{b} \in$ Col A.



§6.5 Least-Squares Problems

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

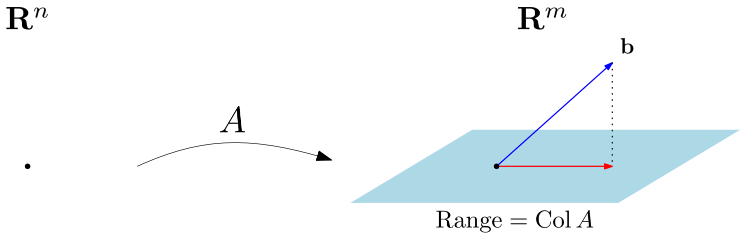
Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



§6.5 Least-Squares Problems

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

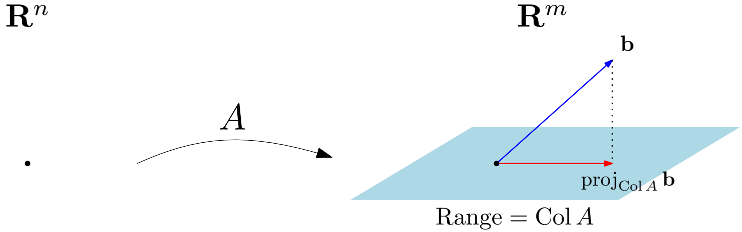
Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff \mathbf{b} \in$ Col A.



§6.5 Least-Squares Problems

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

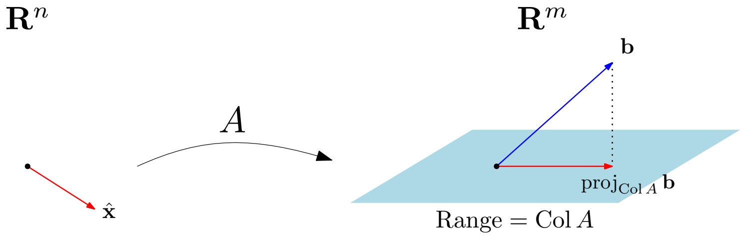
Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



§6.5 Least-Squares Problems

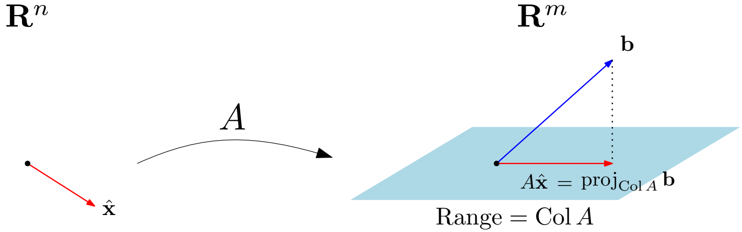
Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

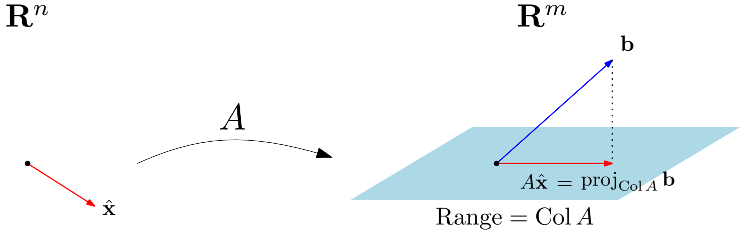
Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



§6.5 Least-Squares Problems

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.

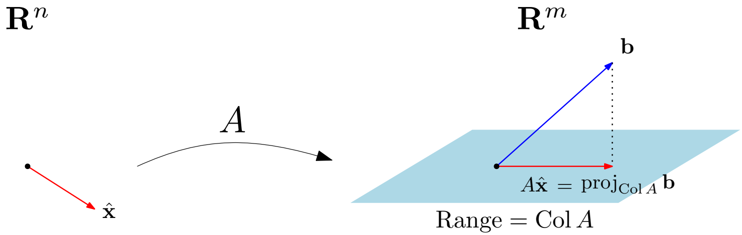


So a least-squares solution $\hat{\mathbf{x}}$ is a vector so that $A\hat{\mathbf{x}}$ is the closest vector in Col A to $\mathbf{b}$. I.e. instead of solving $A\mathbf{x} = \mathbf{b}$, we solve:

§6.5 Least-Squares Problems

Notice that for any vector $\mathbf{x}$, $A\mathbf{x}$ is always in Col A.

Thus $A\mathbf{x} = \mathbf{b}$ has an exact solution $\iff$ $\mathbf{b} \in$ Col A.



So a least-squares solution $\hat{\mathbf{x}}$ is a vector so that $A\hat{\mathbf{x}}$ is the closest vector in Col A to $\mathbf{b}$. I.e. instead of solving $A\mathbf{x} = \mathbf{b}$, we solve:

$$A\hat{\mathbf{x}} = \text{proj}_{\text{Col } A}\ \mathbf{b}.$$

Theorem

The set of least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the nonempty set of solutions to the equation

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

Theorem

The set of least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the nonempty set of solutions to the equation

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

The equations $A^T A\mathbf{x} = A^T\mathbf{b}$ are called the **normal equations** of the system $A\mathbf{x} = \mathbf{b}$.

The scalar $\|A\hat{\mathbf{x}} - \mathbf{b}\|$ is called the **least-squares error**.

The scalar $\|A\hat{\mathbf{x}} - \mathbf{b}\|$ is called the **least-squares error**.

Example: Find the least-squares approximation and error of the inconsistent system $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 2 & 0 \\ 1 & 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 5 \\ 2 \end{bmatrix}.$$

Theorem

Let A be an $m \times n$ matrix. Then the following are equivalent:

Theorem

Let A be an $m \times n$ matrix. Then the following are equivalent:

1. $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution for each $\mathbf{b} \in \mathbb{R}^m$

Theorem

Let A be an $m \times n$ matrix. Then the following are equivalent:

1. $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution for each $\mathbf{b} \in \mathbb{R}^m$
2. The columns of A are linearly independent

Theorem

Let A be an $m \times n$ matrix. Then the following are equivalent:

1. $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution for each $\mathbf{b} \in \mathbb{R}^m$
2. The columns of A are linearly independent
3. $A^T A$ is invertible.

Theorem

Let A be an $m \times n$ matrix. Then the following are equivalent:

1. $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution for each $\mathbf{b} \in \mathbb{R}^m$
2. The columns of A are linearly independent
3. $A^T A$ is invertible.

When these statements are true, the unique least-squares solution is given by

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

Computational short-cut

Computational short-cut

Example: Find the least-squares solution to $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ -1 & 2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 5 \\ -4 \end{bmatrix}.$$

Computational short-cut

Example: Find the least-squares solution to $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ -1 & 2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 5 \\ -4 \end{bmatrix}.$$

Hint: the columns of A are orthogonal. Use this fact to compute $\text{proj}_{\text{Col } A}\, \mathbf{b}$ and then to find $\hat{\mathbf{x}}$.

Computational short-cut

Computational short-cut

Theorem
Given an $m \times n$ matrix A with linearly independent columns, let $A = QR$ be a QR factorization of A.

Computational short-cut

Theorem

Given an $m \times n$ matrix A with linearly independent columns, let $A = QR$ be a QR factorization of A.

Then for each $\mathbf{b} \in \mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has unique least-squares solution

$$\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}.$$

§6.6 Applications to linear models

Problem: Suppose we have observed data points

$$(x_1, y_1), \ldots, (x_n, y_n)$$

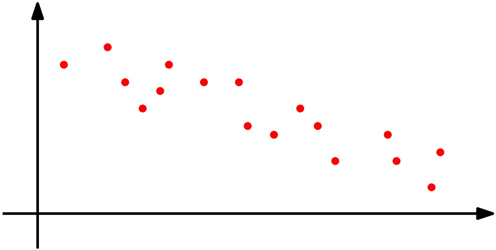
that seem to follow some linear relation.

§6.6 Applications to linear models

Problem: Suppose we have observed data points

$$(x_1, y_1), \ldots, (x_n, y_n)$$

that seem to follow some linear relation.

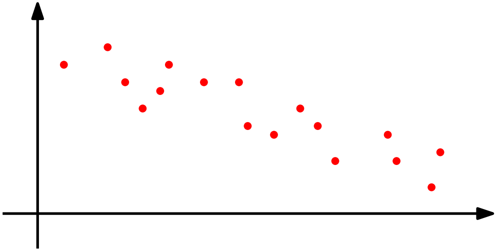


§6.6 Applications to linear models

Problem: Suppose we have observed data points

$$(x_1, y_1), \ldots, (x_n, y_n)$$

that seem to follow some linear relation.

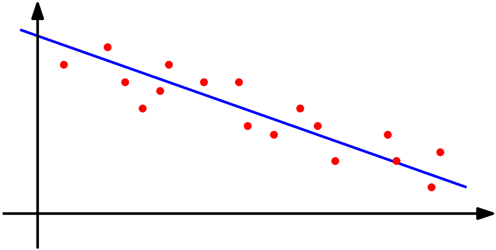


Can we find the line that best fits the data?

Problem: Suppose we have observed data points

$$(x_1, y_1), \ldots, (x_n, y_n)$$

that seem to follow some linear relation.



Can we find the line that best fits the data?

In other words, can we find parameters β_0 and β_1 so that the line

$$y = \beta_0 + \beta_1 x$$

best fits the data points (x_j, y_j)?

In other words, can we find parameters β_0 and β_1 so that the line

$$y = \beta_0 + \beta_1 x$$

best fits the data points (x_j, y_j)?

Let $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$.

Then we want to find β_0 and β_1 so that

$X\beta$

Then we want to find β_0 and β_1 so that

$$X\beta = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

Then we want to find β_0 and β_1 so that

$$X\beta = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{bmatrix}$$

Then we want to find β_0 and β_1 so that

$$X\beta = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{bmatrix} \quad \text{equals} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

We want solution to $X\beta = \mathbf{y}$

Then we want to find β_0 and β_1 so that

$$X\beta = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{bmatrix} \quad \text{equals} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

We want solution to $X\beta = \mathbf{y}$ (no solution unless points lie exactly on line)

Then we want to find β_0 and β_1 so that

$$X\beta = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{bmatrix} \quad \text{equals} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

We want solution to $X\beta = \mathbf{y}$ (no solution unless points lie exactly on line)

$\Longrightarrow$ look for least-squares solution to $X\beta = \mathbf{y}$.

Example: If we have the points

$$(1, -1), \quad (2, 4), \quad (3, 12), \quad \text{and} \quad (4, 20).$$

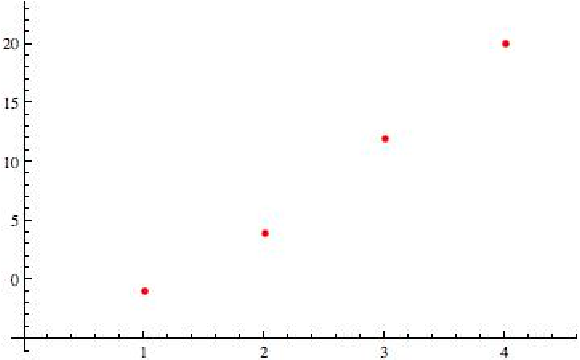
Find the least-squares problem and the normal equations.

§6.6 Applications to linear models

Example: If we have the points

$$(1, -1), \quad (2, 4), \quad (3, 12), \quad \text{and} \quad (4, 20).$$

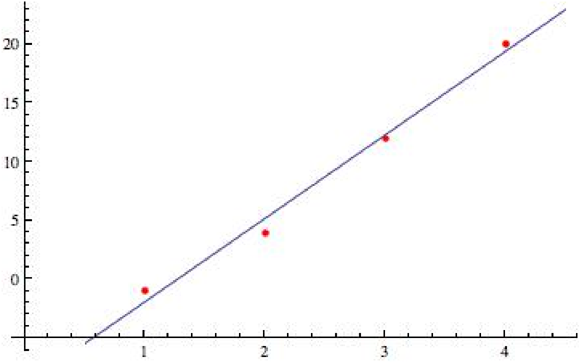
Find the least-squares problem and the normal equations.



Example: If we have the points

$$(1,-1), \quad (2,4), \quad (3,12), \quad \text{and} \quad (4,20).$$

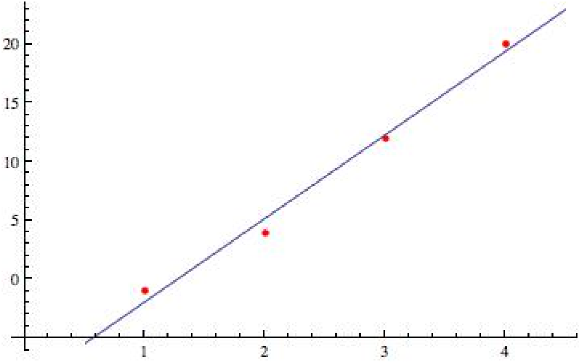
Find the least-squares problem and the normal equations.



Example: If we have the points

$$(1,-1), \quad (2,4), \quad (3,12), \quad \text{and} \quad (4,20).$$

Find the least-squares problem and the normal equations.



$$y = \frac{71}{10}x - 9$$

§6.6 Applications to linear models

Suppose we don't want to match a line to the data points, but a curve of the form

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 \sin x.$$

Suppose we don't want to match a line to the data points, but a curve of the form

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 \sin x.$$

Then we find the least-squares solution of $X\beta = \mathbf{y}$, where

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \sin x_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \sin x_n \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Definition

An **inner product** on a vector space V is a function that assigns to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a real number $\langle \mathbf{u}, \mathbf{v} \rangle$

§6.7 Inner Product Spaces

Definition

An **inner product** on a vector space V is a function that assigns to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ such that the following axioms hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $c \in \mathbb{R}$:

§6.7 Inner Product Spaces

Definition

An **inner product** on a vector space V is a function that assigns to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ such that the following axioms hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $c \in \mathbb{R}$:

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

§6.7 Inner Product Spaces

Definition

An **inner product** on a vector space V is a function that assigns to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ such that the following axioms hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $c \in \mathbb{R}$:

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

Definition

An **inner product** on a vector space V is a function that assigns to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ such that the following axioms hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $c \in \mathbb{R}$:

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

§6.7 Inner Product Spaces

Definition

An **inner product** on a vector space V is a function that assigns to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ such that the following axioms hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $c \in \mathbb{R}$:

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

Definition

An **inner product** on a vector space V is a function that assigns to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ such that the following axioms hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $c \in \mathbb{R}$:

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

Definition

An **inner product** on a vector space V is a function that assigns to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ such that the following axioms hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $c \in \mathbb{R}$:

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

A space with an inner product is called an **inner product space**.

Example: Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$. Show that the function

$\langle \mathbf{u}, \mathbf{w} \rangle = \mathbf{u}^T A \mathbf{w}$

Example: Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$. Show that the function

$$\langle \mathbf{u}, \mathbf{w} \rangle = \mathbf{u}^T A \mathbf{w} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

Example: Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$. Show that the function

$$\langle \mathbf{u}, \mathbf{w} \rangle = \mathbf{u}^T A \mathbf{w} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 2u_1w_1 + 5u_2w_2$$

defines an inner product on $\mathbb{R}^2$.

Example: Let $t_0, \ldots, t_n$ be $n+1$ distinct real numbers.

Example: Let $t_0, \ldots, t_n$ be $n+1$ distinct real numbers. Then for $\mathbf{p}$ and $\mathbf{q}$ in $\mathbb{P}_n$ define

$$\langle \mathbf{p}, \mathbf{q} \rangle = \mathbf{p}(t_0)\mathbf{q}(t_0) + \mathbf{p}(t_1)\mathbf{q}(t_1) + \cdots + \mathbf{p}(t_n)\mathbf{q}(t_n).$$

Example: Let $t_0, \ldots, t_n$ be $n+1$ distinct real numbers. Then for $\mathbf{p}$ and $\mathbf{q}$ in $\mathbb{P}_n$ define

$$\langle \mathbf{p}, \mathbf{q} \rangle = \mathbf{p}(t_0)\mathbf{q}(t_0) + \mathbf{p}(t_1)\mathbf{q}(t_1) + \cdots + \mathbf{p}(t_n)\mathbf{q}(t_n).$$

Show that this defines an inner product on $\mathbb{P}_n$.

Example: Let $t_0, \ldots, t_n$ be $n+1$ distinct real numbers. Then for $\mathbf{p}$ and $\mathbf{q}$ in $\mathbb{P}_n$ define

$$\langle \mathbf{p}, \mathbf{q} \rangle = \mathbf{p}(t_0)\mathbf{q}(t_0) + \mathbf{p}(t_1)\mathbf{q}(t_1) + \cdots + \mathbf{p}(t_n)\mathbf{q}(t_n).$$

Show that this defines an inner product on $\mathbb{P}_n$.

Let $V = \mathbb{P}_2$, and $t_0 = 0$, $t_1 = 1$, and $t_2 = 2$.

Example: Let $t_0, \ldots, t_n$ be $n+1$ distinct real numbers. Then for $\mathbf{p}$ and $\mathbf{q}$ in $\mathbb{P}_n$ define

$$\langle \mathbf{p}, \mathbf{q} \rangle = \mathbf{p}(t_0)\mathbf{q}(t_0) + \mathbf{p}(t_1)\mathbf{q}(t_1) + \cdots + \mathbf{p}(t_n)\mathbf{q}(t_n).$$

Show that this defines an inner product on $\mathbb{P}_n$.

Let $V = \mathbb{P}_2$, and $t_0 = 0$, $t_1 = 1$, and $t_2 = 2$. Compute

$$\langle 1 + 2t - t^2, 2 + t^2 \rangle.$$

Definition

If V is an inner product space, then the **norm** (or **length**) of a vector is

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

Definition

If V is an inner product space, then the **norm** (or **length**) of a vector is

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

A unit vector is a vector with $\|\mathbf{v}\| = 1$,

Definition

If V is an inner product space, then the **norm** (or **length**) of a vector is

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

A unit vector is a vector with $\|\mathbf{v}\| = 1$, the distance between $\mathbf{u}$ and $\mathbf{v}$ is $\|\mathbf{u} - \mathbf{v}\|$.

§6.7 Inner Product Spaces

Definition

If V is an inner product space, then the **norm** (or **length**) of a vector is

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

A unit vector is a vector with $\|\mathbf{v}\| = 1$, the distance between $\mathbf{u}$ and $\mathbf{v}$ is $\|\mathbf{u} - \mathbf{v}\|$. Vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

§6.7 Inner Product Spaces

Definition

If V is an inner product space, then the **norm** (or **length**) of a vector is

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

A unit vector is a vector with $\|\mathbf{v}\| = 1$, the distance between $\mathbf{u}$ and $\mathbf{v}$ is $\|\mathbf{u} - \mathbf{v}\|$. Vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Can also do orthogonal projections, Gram-Schmidt, best approximations etc. as before.

Example: Let $V = \mathbb{P}_4$, with inner product defined on previous slide (by evaluation at $-2, -1, 0, 1$ and 2).

Example: Let $V = \mathbb{P}_4$, with inner product defined on previous slide (by evaluation at $-2, -1, 0, 1$ and 2). Let $W = \mathbb{P}_2$ be thought of as a subspace of V.

Example: Let $V = \mathbb{P}_4$, with inner product defined on previous slide (by evaluation at $-2, -1, 0, 1$ and 2). Let $W = \mathbb{P}_2$ be thought of as a subspace of V. Then use Gram-Schmidt on the basis

$$\{1, t, t^2\}$$

of $\mathbb{P}_2$ to find an orthogonal basis for $\mathbb{P}_2$.

Example: Using the above orthogonal basis of $\mathbb{P}_2$ (as a subspace of the inner product space $\mathbb{P}_4$) to find the best approximation to

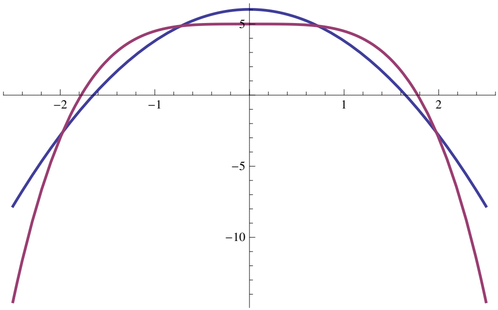
$$\mathbf{p}(t) = 5 - \frac{1}{2}t^4$$

by a polynomial in $\mathbb{P}_2$.

Example: Using the above orthogonal basis of $\mathbb{P}_2$ (as a subspace of the inner product space $\mathbb{P}_4$) to find the best approximation to

$$\mathbf{p}(t) = 5 - \frac{1}{2}t^4$$

by a polynomial in $\mathbb{P}_2$.



Theorem (Cauchy-Schwarz Inequality)

For all $\mathbf{u}, \mathbf{w}$ in V, we have

$$|\langle \mathbf{u}, \mathbf{w} \rangle| \leq \|\mathbf{u}\| \|\mathbf{w}\|.$$

Theorem (Cauchy-Schwarz Inequality)

For all $\mathbf{u}, \mathbf{w}$ in V, we have

$$|\langle \mathbf{u}, \mathbf{w} \rangle| \leq \|\mathbf{u}\| \|\mathbf{w}\|.$$

Theorem (Triangle Inequality)

For all $\mathbf{u}, \mathbf{w}$ in V, we have

$$\|\mathbf{u} + \mathbf{w}\| \leq \|\mathbf{u}\| + \|\mathbf{w}\|.$$

Example: Define an inner product on $C[a,b]$ (the space of continuous functions defined on the interval $[a,b]$) by

$$\langle f, g\rangle = \int_a^b f(t)g(t)\ dt.$$

Example: Define an inner product on $C[a, b]$ (the space of continuous functions defined on the interval $[a, b]$) by

$$\langle f, g \rangle = \int_a^b f(t)g(t)\ dt.$$

Show that this defines an inner product.

Example: Define an inner product on $C[a, b]$ (the space of continuous functions defined on the interval $[a, b]$) by

$$\langle f, g \rangle = \int_a^b f(t)g(t)\ dt.$$

Show that this defines an inner product.

Example: If $V = C[0, 1]$, let W be the subspace spanned by $\mathbf{p}_1(t) = 1, \mathbf{p}_2(t) = x$, and $\mathbf{p}_3(t) = e^x$.

Example: Define an inner product on $C[a,b]$ (the space of continuous functions defined on the interval $[a,b]$) by

$$\langle f, g\rangle = \int_a^b f(t)g(t)\, dt.$$

Show that this defines an inner product.

Example: If $V = C[0,1]$, let W be the subspace spanned by $\mathbf{p}_1(t) = 1, \mathbf{p}_2(t) = x$, and $\mathbf{p}_3(t) = e^x$. Use Gram-Schmidt to find an orthogonal basis for W.

§6.7 Inner Product Spaces

$\langle f, g \rangle = \int_0^1 f(t)g(t)\ dt$

§6.7 Inner Product Spaces

$$\langle f, g \rangle = \int_0^1 f(t)g(t)\ dt \quad \Longrightarrow \quad \text{inner product on } C[0,1].$$

$\langle f, g \rangle = \int_0^1 f(t)g(t)\; dt \quad \Longrightarrow \quad$ inner product on $C[0,1]$.

Let $W = \text{Span}\,\{1, x, e^x\}$.

§6.7 Inner Product Spaces

$\langle f, g \rangle = \int_0^1 f(t)g(t)\ dt \quad \Longrightarrow \quad$ inner product on $C[0,1]$.

Let $W = \text{Span}\ \{1, x, e^x\}$. Use Gram-Schmidt to find an orthogonal basis for W.

$\langle f, g\rangle = \int_0^1 f(t)g(t)\ dt \quad \Longrightarrow \quad$ inner product on $C[0,1]$.

Let $W = \text{Span}\ \{1, x, e^x\}$. Use Gram-Schmidt to find an orthogonal basis for W.

$\langle 1, 1\rangle = \int_0^1 1\ dt = 1$

$\langle 1, x\rangle = \int_0^1 x\ dt = \frac{1}{2}$

$\langle 1, e^x\rangle = \int_0^1 e^x\ dt = e - 1$

$$\langle f, g \rangle = \int_0^1 f(t)g(t)\ dt \quad \Longrightarrow \quad \text{inner product on } C[0,1].$$

Let $W = \text{Span}\ \{1, x, e^x\}$. Use Gram-Schmidt to find an orthogonal basis for W.

$$\langle 1, 1 \rangle = \int_0^1 1\ dt = 1$$

$$\langle 1, x \rangle = \int_0^1 x\ dt = \tfrac{1}{2}$$

$$\langle 1, e^x \rangle = \int_0^1 e^x\ dt = e - 1$$

By Gram-Schmidt we get:

$\langle f, g \rangle = \int_0^1 f(t)g(t)\, dt \quad \Longrightarrow \quad$ inner product on $C[0,1]$.

Let $W = \text{Span}\,\{1, x, e^x\}$. Use Gram-Schmidt to find an orthogonal basis for W.

$$\langle 1, 1 \rangle = \int_0^1 1\, dt = 1$$

$$\langle 1, x \rangle = \int_0^1 x\, dt = \tfrac{1}{2}$$

$$\langle 1, e^x \rangle = \int_0^1 e^x\, dt = e - 1$$

By Gram-Schmidt we get:

$\mathbf{v}_1 = 1,$

$\langle f, g\rangle = \int_0^1 f(t)g(t)\ dt \quad \Longrightarrow \quad$ inner product on $C[0,1]$.

Let $W = \text{Span}\ \{1, x, e^x\}$. Use Gram-Schmidt to find an orthogonal basis for W.

$$\langle 1, 1\rangle = \int_0^1 1\ dt = 1$$

$$\langle 1, x\rangle = \int_0^1 x\ dt = \tfrac{1}{2}$$

$$\langle 1, e^x\rangle = \int_0^1 e^x\ dt = e - 1$$

By Gram-Schmidt we get:

$\mathbf{v}_1 = 1, \qquad \mathbf{v}_2 = x - \tfrac{1}{2},$

§6.7 Inner Product Spaces

$\langle f, g\rangle = \int_0^1 f(t)g(t)\ dt \quad \Longrightarrow \quad$ inner product on $C[0,1]$.

Let $W = \text{Span}\ \{1, x, e^x\}$. Use Gram-Schmidt to find an orthogonal basis for W.

$\langle 1, 1\rangle = \int_0^1 1\ dt = 1$	$\langle x - \frac{1}{2}, e^x\rangle = \int_0^1 (x - \frac{1}{2})e^x\ dt = \frac{3-e}{2}$
$\langle 1, x\rangle = \int_0^1 x\ dt = \frac{1}{2}$	$\langle x - \frac{1}{2}, x - \frac{1}{2}\rangle = \int_0^1 (x - \frac{1}{2})^2\ dt = \frac{1}{12}$
$\langle 1, e^x\rangle = \int_0^1 e^x\ dt = e - 1$	

By Gram-Schmidt we get:

$\mathbf{v}_1 = 1, \qquad \mathbf{v}_2 = x - \frac{1}{2},$

§6.7 Inner Product Spaces

$\langle f, g\rangle = \int_0^1 f(t)g(t)\ dt \quad \Longrightarrow \quad$ inner product on $C[0,1]$.

Let $W = \text{Span}\ \{1, x, e^x\}$. Use Gram-Schmidt to find an orthogonal basis for W.

$\langle 1, 1\rangle = \int_0^1 1\ dt = 1$	$\langle x - \frac{1}{2}, e^x\rangle = \int_0^1 (x - \frac{1}{2})e^x\ dt = \frac{3-e}{2}$
$\langle 1, x\rangle = \int_0^1 x\ dt = \frac{1}{2}$	$\langle x - \frac{1}{2}, x - \frac{1}{2}\rangle = \int_0^1 (x - \frac{1}{2})^2\ dt = \frac{1}{12}$
$\langle 1, e^x\rangle = \int_0^1 e^x\ dt = e - 1$	

By Gram-Schmidt we get:

$$\mathbf{v}_1 = 1, \qquad \mathbf{v}_2 = x - \tfrac{1}{2}, \qquad \mathbf{v}_3 = e^x - (e-1) - 6(3-e)(x - \tfrac{1}{2}).$$

What are they good for?

What are they good for?

What are they good for?

What are they good for?

What are they good for?

What are they good for?

What are they good for?

What are they good for?

What are they good for?