Lecture 24

Math 22 Summer 2017
August 11, 2017

Just for today

- §6.3 Orthogonal projections
- §6.5 Least squares problems (start)

Last time

Last time

Theorem

Last time

Theorem

Suppose U is an orthogonal matrix

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$).

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$.

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$. But $U^{-1} = U^T$, so U has orthonormal columns.

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$. But $U^{-1} = U^T$, so U has orthonormal columns. What about the rows?

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$. But $U^{-1} = U^T$, so U has orthonormal columns. What about the rows? Well, the rows of U are the columns of U^T.

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$. But $U^{-1} = U^T$, so U has orthonormal columns. What about the rows? Well, the rows of U are the columns of U^T. So we win if we can show U^T has orthonormal columns.

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$. But $U^{-1} = U^T$, so U has orthonormal columns. What about the rows? Well, the rows of U are the columns of U^T. So we win if we can show U^T has orthonormal columns. But to do that we can use the above theorem again with U^T instead of U.

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$. But $U^{-1} = U^T$, so U has orthonormal columns. What about the rows? Well, the rows of U are the columns of U^T. So we win if we can show U^T has orthonormal columns. But to do that we can use the above theorem again with U^T instead of U. How?

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$. But $U^{-1} = U^T$, so U has orthonormal columns. What about the rows? Well, the rows of U are the columns of U^T. So we win if we can show U^T has orthonormal columns. But to do that we can use the above theorem again with U^T instead of U. How? Well,

$$(U^T)^T(U^T) = U(U^T)$$

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$. But $U^{-1} = U^T$, so U has orthonormal columns. What about the rows? Well, the rows of U are the columns of U^T. So we win if we can show U^T has orthonormal columns. But to do that we can use the above theorem again with U^T instead of U. How? Well,

$$(U^T)^T(U^T) = U(U^T)$$

But $U^T = U^{-1}$.

Last time

Theorem

Suppose U is an orthogonal matrix (i.e. U is square and $U^{-1} = U^T$). Then U has orthonormal columns and rows.

Proof.

We proved last time that any matrix U has orthonormal columns if and only if $U^T U = I_n$. But $U^{-1} = U^T$, so U has orthonormal columns. What about the rows? Well, the rows of U are the columns of U^T. So we win if we can show U^T has orthonormal columns. But to do that we can use the above theorem again with U^T instead of U. How? Well,

$$(U^T)^T(U^T) = U(U^T)$$

But $U^T = U^{-1}$. So $(U^T)^T(U^T) = U(U^T) = UU^{-1} = I_n$. □

§6.3 Theorem 8

Theorem

§6.3 Theorem 8

Theorem

Let W *be a subspace of* $\mathbb{R}^n$.

Theorem

Let W be a subspace of $\mathbb{R}^n$. Suppose $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W.

Theorem

Let W be a subspace of $\mathbb{R}^n$. Suppose $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W. Then every $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$

Theorem

Let W be a subspace of $\mathbb{R}^n$. Suppose $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W. Then every $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}} \in W$ and $\mathbf{z} \in W^{\perp}$.

Theorem

Let W *be a subspace of* $\mathbb{R}^n$*. Suppose* $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ *is an orthogonal basis for* W*. Then every* $\mathbf{y} \in \mathbb{R}^n$ *can be written uniquely as* $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ *where* $\hat{\mathbf{y}} \in W$ *and* $\mathbf{z} \in W^\perp$*. In particular,*

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p, \quad \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Theorem

Let W be a subspace of $\mathbb{R}^n$. Suppose $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W. Then every $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}} \in W$ and $\mathbf{z} \in W^\perp$. In particular,

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p, \quad \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

We call $\hat{\mathbf{y}}$ the **orthogonal projection of $\mathbf{y}$ onto** W.

Theorem

Let W be a subspace of $\mathbb{R}^n$. Suppose $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W. Then every $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}} \in W$ and $\mathbf{z} \in W^\perp$. In particular,

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p, \quad \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

We call $\hat{\mathbf{y}}$ the **orthogonal projection of $\mathbf{y}$ onto** W.

$\hat{\mathbf{y}}$ is also denoted by $\operatorname{proj}_W \mathbf{y}$.

Theorem

Let W be a subspace of $\mathbb{R}^n$. Suppose $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W. Then every $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely as $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}} \in W$ and $\mathbf{z} \in W^{\perp}$. In particular,

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p, \quad \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

We call $\hat{\mathbf{y}}$ the **orthogonal projection of $\mathbf{y}$ onto** W.

$\hat{\mathbf{y}}$ is also denoted by $\operatorname{proj}_W \mathbf{y}$.

So what's the proof?

§6.3 Proof of Theorem 8

§6.3 Proof of Theorem 8

Proof.

§6.3 Proof of Theorem 8

Proof.

Certainly $\hat{\mathbf{y}} \in W$.

§6.3 Proof of Theorem 8

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^{\perp}$?

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^{\perp}$? Well,

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

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^{\perp}$? Well,

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

What justifies the last equality?

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^\perp$? Well,

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

What justifies the last equality? So $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \in W + W^\perp$.

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^{\perp}$? Well,

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

What justifies the last equality? So $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \in W + W^{\perp}$.

How do we show uniqueness?

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^\perp$? Well,

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

What justifies the last equality? So $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \in W + W^\perp$.

How do we show uniqueness? Suppose $\boxed{\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1}$ with $\hat{\mathbf{y}}, \hat{\mathbf{y}}_1 \in W$ and $\mathbf{z}, \mathbf{z}_1 \in W^\perp$.

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^{\perp}$? Well,

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

What justifies the last equality? So $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \in W + W^{\perp}$.

How do we show uniqueness? Suppose $\boxed{\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1}$ with $\hat{\mathbf{y}}, \hat{\mathbf{y}}_1 \in W$ and $\mathbf{z}, \mathbf{z}_1 \in W^{\perp}$. The boxed equation implies

$$\underbrace{\hat{\mathbf{y}} - \hat{\mathbf{y}}_1}_{\in W} = \underbrace{\mathbf{z}_1 - \mathbf{z}}_{\in W^{\perp}}.$$

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^\perp$? Well,

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

What justifies the last equality? So $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \in W + W^\perp$.

How do we show uniqueness? Suppose $\boxed{\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1}$ with $\hat{\mathbf{y}}, \hat{\mathbf{y}}_1 \in W$ and $\mathbf{z}, \mathbf{z}_1 \in W^\perp$. The boxed equation implies

$$\underbrace{\hat{\mathbf{y}} - \hat{\mathbf{y}}_1}_{\in W} = \underbrace{\mathbf{z}_1 - \mathbf{z}}_{\in W^\perp}.$$

Let $\mathbf{v}$ denote this vector.

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^\perp$? Well,

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

What justifies the last equality? So $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \in W + W^\perp$.

How do we show uniqueness? Suppose $\boxed{\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1}$ with $\hat{\mathbf{y}}, \hat{\mathbf{y}}_1 \in W$ and $\mathbf{z}, \mathbf{z}_1 \in W^\perp$. The boxed equation implies

$$\underbrace{\hat{\mathbf{y}} - \hat{\mathbf{y}}_1}_{\in W} = \underbrace{\mathbf{z}_1 - \mathbf{z}}_{\in W^\perp}.$$

Let $\mathbf{v}$ denote this vector. What does this equation tell us about $\mathbf{v}$?

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^\perp$? Well,

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

What justifies the last equality? So $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \in W + W^\perp$.

How do we show uniqueness? Suppose $\boxed{\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1}$ with $\hat{\mathbf{y}}, \hat{\mathbf{y}}_1 \in W$ and $\mathbf{z}, \mathbf{z}_1 \in W^\perp$. The boxed equation implies

$$\underbrace{\hat{\mathbf{y}} - \hat{\mathbf{y}}_1}_{\in W} = \underbrace{\mathbf{z}_1 - \mathbf{z}}_{\in W^\perp}.$$

Let $\mathbf{v}$ denote this vector. What does this equation tell us about $\mathbf{v}$?
$\mathbf{v} \cdot \mathbf{v} = 0$.

§6.3 Proof of Theorem 8

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^{\perp}$? Well,

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

What justifies the last equality? So $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \in W + W^{\perp}$.

How do we show uniqueness? Suppose $\boxed{\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1}$ with $\hat{\mathbf{y}}, \hat{\mathbf{y}}_1 \in W$ and $\mathbf{z}, \mathbf{z}_1 \in W^{\perp}$. The boxed equation implies

$$\underbrace{\hat{\mathbf{y}} - \hat{\mathbf{y}}_1}_{\in W} = \underbrace{\mathbf{z}_1 - \mathbf{z}}_{\in W^{\perp}}.$$

Let $\mathbf{v}$ denote this vector. What does this equation tell us about $\mathbf{v}$? $\mathbf{v} \cdot \mathbf{v} = 0$. So what?

Proof.

Certainly $\hat{\mathbf{y}} \in W$. Is $\mathbf{z} \in W^{\perp}$? Well,

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

What justifies the last equality? So $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \in W + W^{\perp}$.

How do we show uniqueness? Suppose $\boxed{\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1}$ with $\hat{\mathbf{y}}, \hat{\mathbf{y}}_1 \in W$ and $\mathbf{z}, \mathbf{z}_1 \in W^{\perp}$. The boxed equation implies

$$\underbrace{\hat{\mathbf{y}} - \hat{\mathbf{y}}_1}_{\in W} = \underbrace{\mathbf{z}_1 - \mathbf{z}}_{\in W^{\perp}}.$$

Let $\mathbf{v}$ denote this vector. What does this equation tell us about $\mathbf{v}$? $\mathbf{v} \cdot \mathbf{v} = 0$. So what? $\mathbf{v} = \mathbf{0}$. □

§6.3 Example

§6.3 Example

Let

$$\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}.$$

Let W be the span of $\mathbf{u}_1$ and $\mathbf{u}_2$. Find the projection of $\mathbf{y}$ onto W and the distance from $\mathbf{y}$ to W.

§6.3 Example

Let

$$\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}.$$

Let W be the span of $\mathbf{u}_1$ and $\mathbf{u}_2$. Find the projection of $\mathbf{y}$ onto W and the distance from $\mathbf{y}$ to W.

Solution:

§6.3 Example

Let

$$\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}.$$

Let W be the span of $\mathbf{u}_1$ and $\mathbf{u}_2$. Find the projection of $\mathbf{y}$ onto W and the distance from $\mathbf{y}$ to W.

Solution: Well,

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = \frac{6}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{7}{14}\begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 7/2 \\ 1 \end{bmatrix}$$

§6.3 Example

Let

$$\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}.$$

Let W be the span of $\mathbf{u}_1$ and $\mathbf{u}_2$. Find the projection of $\mathbf{y}$ onto W and the distance from $\mathbf{y}$ to W.

Solution: Well,

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = \frac{6}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{7}{14}\begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 7/2 \\ 1 \end{bmatrix}$$

and

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -5/2 \\ 1/2 \\ 2 \end{bmatrix} \implies \|\mathbf{z}\| = \sqrt{4 + 26/4} = 3.2015621187164\ldots$$

§6.3 Theorem 9 (Best Approximation)

§6.3 Theorem 9 (Best Approximation)

Theorem

Theorem

Let W be a subspace of $\mathbb{R}^n$.

Theorem

Let W be a subspace of $\mathbb{R}^n$. Let $\mathbf{y} \in \mathbb{R}^n$.

Theorem

Let W be a subspace of $\mathbb{R}^n$. Let $\mathbf{y} \in \mathbb{R}^n$. Let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W.

§6.3 Theorem 9 (Best Approximation)

Theorem

Let W be a subspace of $\mathbb{R}^n$. Let $\mathbf{y} \in \mathbb{R}^n$. Let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the point in W that is closest to $\mathbf{y}$.

§6.3 Theorem 9 (Best Approximation)

Theorem

Let W be a subspace of $\mathbb{R}^n$. Let $\mathbf{y} \in \mathbb{R}^n$. Let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the point in W that is closest to $\mathbf{y}$. More precisely, for every $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{y}}$ we have the strict inequality

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

§6.3 Theorem 9 (Best Approximation)

Theorem

Let W be a subspace of $\mathbb{R}^n$. Let $\mathbf{y} \in \mathbb{R}^n$. Let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the point in W that is closest to $\mathbf{y}$. More precisely, for every $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{y}}$ we have the strict inequality

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

Proof.

§6.3 Theorem 9 (Best Approximation)

Theorem

Let W be a subspace of $\mathbb{R}^n$. Let $\mathbf{y} \in \mathbb{R}^n$. Let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the point in W that is closest to $\mathbf{y}$. More precisely, for every $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{y}}$ we have the strict inequality

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

Proof.

What's the proof in 2 words?

§6.3 Theorem 9 (Best Approximation)

Theorem

Let W be a subspace of $\mathbb{R}^n$. Let $\mathbf{y} \in \mathbb{R}^n$. Let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the point in W that is closest to $\mathbf{y}$. More precisely, for every $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{y}}$ we have the strict inequality

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

Proof.

What's the proof in 2 words? Pythagorean Theorem.

§6.3 Theorem 9 (Best Approximation)

Theorem

Let W be a subspace of $\mathbb{R}^n$. Let $\mathbf{y} \in \mathbb{R}^n$. Let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the point in W that is closest to $\mathbf{y}$. More precisely, for every $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{y}}$ we have the strict inequality

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

Proof.

What's the proof in 2 words? Pythagorean Theorem.

$$\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \underbrace{\|\hat{\mathbf{y}} - \mathbf{v}\|^2}_{>0} .$$

§6.3 Theorem 9 (Best Approximation)

Theorem

Let W be a subspace of $\mathbb{R}^n$. Let $\mathbf{y} \in \mathbb{R}^n$. Let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the point in W that is closest to $\mathbf{y}$. More precisely, for every $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{y}}$ we have the strict inequality

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

Proof.

What's the proof in 2 words? Pythagorean Theorem.

$$\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \underbrace{\|\hat{\mathbf{y}} - \mathbf{v}\|^2}_{>0} .$$

Draw a picture! □

§6.5 Some motivation for doing this

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W\, \mathbf{b}$.

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$.

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$. So what?

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$. So what? Well, in light of the previous theorem, we have $\boxed{\left\|\mathbf{b} - \hat{\mathbf{b}}\right\| < \|\mathbf{b} - \mathbf{v}\|}$ for any $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{b}}$.

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$. So what? Well, in light of the previous theorem, we have $\boxed{\left\|\mathbf{b} - \hat{\mathbf{b}}\right\| < \|\mathbf{b} - \mathbf{v}\|}$ for any $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{b}}$.

Now $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ and $\mathbf{v} \in W = \mathrm{Col} A$ mean $\mathbf{v} = A\mathbf{x}$ for some $\mathbf{x}$.

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col}A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col}A$. So what? Well, in light of the previous theorem, we have $\boxed{\left\|\mathbf{b} - \hat{\mathbf{b}}\right\| < \|\mathbf{b} - \mathbf{v}\|}$ for any $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{b}}$.

Now $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ and $\mathbf{v} \in W = \mathrm{Col}A$ mean $\mathbf{v} = A\mathbf{x}$ for some $\mathbf{x}$. So the boxed equation becomes

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| < \|\mathbf{b} - A\mathbf{x}\|, \quad \text{for any } \mathbf{x} \in \mathbb{R}^n.$$

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$. So what? Well, in light of the previous theorem, we have $\boxed{\left\|\mathbf{b} - \hat{\mathbf{b}}\right\| < \|\mathbf{b} - \mathbf{v}\|}$ for any $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{b}}$.

Now $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ and $\mathbf{v} \in W = \mathrm{Col} A$ mean $\mathbf{v} = A\mathbf{x}$ for some $\mathbf{x}$. So the boxed equation becomes

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| < \|\mathbf{b} - A\mathbf{x}\|, \quad \text{for any } \mathbf{x} \in \mathbb{R}^n.$$

A solution $\hat{\mathbf{x}}$ of $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ is called a **least squares solution** of $A\mathbf{x} = \mathbf{b}$.

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$. So what? Well, in light of the previous theorem, we have $\boxed{\left\| \mathbf{b} - \hat{\mathbf{b}} \right\| < \| \mathbf{b} - \mathbf{v} \|}$ for any $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{b}}$.

Now $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ and $\mathbf{v} \in W = \mathrm{Col} A$ mean $\mathbf{v} = A\mathbf{x}$ for some $\mathbf{x}$. So the boxed equation becomes

$$\| \mathbf{b} - A\hat{\mathbf{x}} \| < \| \mathbf{b} - A\mathbf{x} \| , \quad \text{for any } \mathbf{x} \in \mathbb{R}^n.$$

A solution $\hat{\mathbf{x}}$ of $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ is called a **least squares solution** of $A\mathbf{x} = \mathbf{b}$. If $A\mathbf{x} = \mathbf{b}$ is consistent, then $\| A\mathbf{x} - \mathbf{b} \| = 0$,

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$. So what? Well, in light of the previous theorem, we have $\boxed{\left\|\mathbf{b} - \hat{\mathbf{b}}\right\| < \|\mathbf{b} - \mathbf{v}\|}$ for any $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{b}}$.

Now $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ and $\mathbf{v} \in W = \mathrm{Col} A$ mean $\mathbf{v} = A\mathbf{x}$ for some $\mathbf{x}$. So the boxed equation becomes

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| < \|\mathbf{b} - A\mathbf{x}\|, \quad \text{for any } \mathbf{x} \in \mathbb{R}^n.$$

A solution $\hat{\mathbf{x}}$ of $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ is called a **least squares solution** of $A\mathbf{x} = \mathbf{b}$. If $A\mathbf{x} = \mathbf{b}$ is consistent, then $\|A\mathbf{x} - \mathbf{b}\| = 0$, but if the system is inconsistent, then $\|A\mathbf{x} - \mathbf{b}\| > 0$,

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$. So what? Well, in light of the previous theorem, we have $\boxed{\left\|\mathbf{b} - \hat{\mathbf{b}}\right\| < \|\mathbf{b} - \mathbf{v}\|}$ for any $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{b}}$.

Now $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ and $\mathbf{v} \in W = \mathrm{Col} A$ mean $\mathbf{v} = A\mathbf{x}$ for some $\mathbf{x}$. So the boxed equation becomes

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| < \|\mathbf{b} - A\mathbf{x}\|, \quad \text{for any } \mathbf{x} \in \mathbb{R}^n.$$

A solution $\hat{\mathbf{x}}$ of $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ is called a **least squares solution** of $A\mathbf{x} = \mathbf{b}$. If $A\mathbf{x} = \mathbf{b}$ is consistent, then $\|A\mathbf{x} - \mathbf{b}\| = 0$, but if the system is inconsistent, then $\|A\mathbf{x} - \mathbf{b}\| > 0$, and this positive number represents the error in being able to find a solution.

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$. So what? Well, in light of the previous theorem, we have $\boxed{\left\|\mathbf{b} - \hat{\mathbf{b}}\right\| < \|\mathbf{b} - \mathbf{v}\|}$ for any $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{b}}$.

Now $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ and $\mathbf{v} \in W = \mathrm{Col} A$ mean $\mathbf{v} = A\mathbf{x}$ for some $\mathbf{x}$. So the boxed equation becomes

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| < \|\mathbf{b} - A\mathbf{x}\|, \quad \text{for any } \mathbf{x} \in \mathbb{R}^n.$$

A solution $\hat{\mathbf{x}}$ of $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ is called a **least squares solution** of $A\mathbf{x} = \mathbf{b}$. If $A\mathbf{x} = \mathbf{b}$ is consistent, then $\|A\mathbf{x} - \mathbf{b}\| = 0$, but if the system is inconsistent, then $\|A\mathbf{x} - \mathbf{b}\| > 0$, and this positive number represents the error in being able to find a solution.

The least squares solution $\hat{\mathbf{x}}$ minimizes this error.

§6.5 Some motivation for doing this

Consider an inconsistent linear system $A\mathbf{x} = \mathbf{b}$.
Let $W = \mathrm{Col} A$ and $\hat{\mathbf{b}} := \mathrm{proj}_W \mathbf{b}$. Then there is a solution $\hat{\mathbf{x}}$ to $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ since $\hat{\mathbf{b}} \in \mathrm{Col} A$. So what? Well, in light of the previous theorem, we have $\boxed{\left\|\mathbf{b} - \hat{\mathbf{b}}\right\| < \|\mathbf{b} - \mathbf{v}\|}$ for any $\mathbf{v} \in W$ with $\mathbf{v} \neq \hat{\mathbf{b}}$.

Now $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ and $\mathbf{v} \in W = \mathrm{Col} A$ mean $\mathbf{v} = A\mathbf{x}$ for some $\mathbf{x}$. So the boxed equation becomes

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| < \|\mathbf{b} - A\mathbf{x}\|, \quad \text{for any } \mathbf{x} \in \mathbb{R}^n.$$

A solution $\hat{\mathbf{x}}$ of $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ is called a **least squares solution** of $A\mathbf{x} = \mathbf{b}$. If $A\mathbf{x} = \mathbf{b}$ is consistent, then $\|A\mathbf{x} - \mathbf{b}\| = 0$, but if the system is inconsistent, then $\|A\mathbf{x} - \mathbf{b}\| > 0$, and this positive number represents the error in being able to find a solution.

The least squares solution $\hat{\mathbf{x}}$ minimizes this error.

More of this next week.

§6.3 Theorem 10

§6.3 Theorem 10

Recall theorem 8 from today.

Recall theorem 8 from today. What happens if we insist that our basis for W be orthonormal instead of just orthogonal?

§6.3 Theorem 10

Recall theorem 8 from today. What happens if we insist that our basis for W be orthonormal instead of just orthogonal?

Theorem

Recall theorem 8 from today. What happens if we insist that our basis for W be orthonormal instead of just orthogonal?

Theorem

Let $\mathbf{y} \in \mathbb{R}^n$.

Recall theorem 8 from today. What happens if we insist that our basis for W be orthonormal instead of just orthogonal?

Theorem

Let $\mathbf{y} \in \mathbb{R}^n$. Let W be a subspace of $\mathbb{R}^n$ with orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$.

Recall theorem 8 from today. What happens if we insist that our basis for W be orthonormal instead of just orthogonal?

Theorem

Let $\mathbf{y} \in \mathbb{R}^n$. *Let* W *be a subspace of* $\mathbb{R}^n$ *with orthonormal basis* $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$. *Then*

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

Recall theorem 8 from today. What happens if we insist that our basis for W be orthonormal instead of just orthogonal?

Theorem

Let $\mathbf{y} \in \mathbb{R}^n$. Let W be a subspace of $\mathbb{R}^n$ with orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$. Then

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

Moreover, if we let $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_p]$,

Recall theorem 8 from today. What happens if we insist that our basis for W be orthonormal instead of just orthogonal?

Theorem

Let $\mathbf{y} \in \mathbb{R}^n$*. Let* W *be a subspace of* $\mathbb{R}^n$ *with orthonormal basis* $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$*. Then*

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

Moreover, if we let $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_p]$*, then*

$$\operatorname{proj}_W \mathbf{y} = UU^T\mathbf{y}.$$

§6.3 Proof of Theorem 10

§6.3 Proof of Theorem 10

Proof.

Proof.

The equation $\boxed{\operatorname{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p}$ follows immediately from Theorem 8 since $\{\mathbf{u}_i\}$ is orthonormal.

Proof.

The equation $\boxed{\operatorname{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p}$ follows immediately from Theorem 8 since $\{\mathbf{u}_i\}$ is orthonormal.

Since U is the matrix whose columns are the $\mathbf{u}_i$, the boxed expression is a linear combination of the columns of U with weights $\mathbf{y} \cdot \mathbf{u}_i$.

Proof.

The equation $\boxed{\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p}$ follows immediately from Theorem 8 since $\{\mathbf{u}_i\}$ is orthonormal.

Since U is the matrix whose columns are the $\mathbf{u}_i$, the boxed expression is a linear combination of the columns of U with weights $\mathbf{y} \cdot \mathbf{u}_i$. These weights are $\mathbf{y} \cdot \mathbf{u}_i = \mathbf{u}_i \cdot \mathbf{y} = \mathbf{u}_i^T \mathbf{y}$.

Proof.

The equation $\boxed{\text{proj}_W\, \mathbf{y} = (\mathbf{y}\cdot\mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y}\cdot\mathbf{u}_p)\mathbf{u}_p}$ follows immediately from Theorem 8 since $\{\mathbf{u}_i\}$ is orthonormal.

Since U is the matrix whose columns are the $\mathbf{u}_i$, the boxed expression is a linear combination of the columns of U with weights $\mathbf{y}\cdot\mathbf{u}_i$. These weights are $\mathbf{y}\cdot\mathbf{u}_i = \mathbf{u}_i\cdot\mathbf{y} = \mathbf{u}_i^T\mathbf{y}$. Thus,

$$U^T\mathbf{y} = \begin{bmatrix} \mathbf{y}\cdot\mathbf{u}_1 \\ \vdots \\ \mathbf{y}\cdot\mathbf{u}_p \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{y} \\ \vdots \\ \mathbf{u}_p^T\mathbf{y} \end{bmatrix}.$$

Proof.

The equation $\boxed{\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p}$ follows immediately from Theorem 8 since $\{\mathbf{u}_i\}$ is orthonormal.

Since U is the matrix whose columns are the $\mathbf{u}_i$, the boxed expression is a linear combination of the columns of U with weights $\mathbf{y} \cdot \mathbf{u}_i$. These weights are $\mathbf{y} \cdot \mathbf{u}_i = \mathbf{u}_i \cdot \mathbf{y} = \mathbf{u}_i^T \mathbf{y}$. Thus,

$$U^T \mathbf{y} = \begin{bmatrix} \mathbf{y} \cdot \mathbf{u}_1 \\ \vdots \\ \mathbf{y} \cdot \mathbf{u}_p \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T \mathbf{y} \\ \vdots \\ \mathbf{u}_p^T \mathbf{y} \end{bmatrix}.$$

Thus
$\text{proj}_W \mathbf{y} = (\mathbf{u}_1^T \mathbf{y})\mathbf{u}_1 + \cdots + (\mathbf{u}_p^T \mathbf{y})\mathbf{u}_p = U(U^T \mathbf{y}) = UU^T \mathbf{y}.$ □

Proof.

The equation $\boxed{\text{proj}_W\, \mathbf{y} = (\mathbf{y}\cdot\mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y}\cdot\mathbf{u}_p)\mathbf{u}_p}$ follows immediately from Theorem 8 since $\{\mathbf{u}_i\}$ is orthonormal.

Since U is the matrix whose columns are the $\mathbf{u}_i$, the boxed expression is a linear combination of the columns of U with weights $\mathbf{y}\cdot\mathbf{u}_i$. These weights are $\mathbf{y}\cdot\mathbf{u}_i = \mathbf{u}_i\cdot\mathbf{y} = \mathbf{u}_i^T\mathbf{y}$. Thus,

$$U^T\mathbf{y} = \begin{bmatrix} \mathbf{y}\cdot\mathbf{u}_1 \\ \vdots \\ \mathbf{y}\cdot\mathbf{u}_p \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{y} \\ \vdots \\ \mathbf{u}_p^T\mathbf{y} \end{bmatrix}.$$

Thus
$\text{proj}_W\, \mathbf{y} = (\mathbf{u}_1^T\mathbf{y})\mathbf{u}_1 + \cdots + (\mathbf{u}_p^T\mathbf{y})\mathbf{u}_p = U(U^T\mathbf{y}) = UU^T\mathbf{y}.$ □

Let's finish with an example.

§6.3 Example

§6.3 Example

Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, $W = \{\mathbf{v}_1, \mathbf{v}_2\}$, and $\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

§6.3 Example

Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, $W = \{\mathbf{v}_1, \mathbf{v}_2\}$, and $\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.
Find $\text{proj}_W \mathbf{y}$ and find a vector orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.

Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, $W = \{\mathbf{v}_1, \mathbf{v}_2\}$, and $\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Find $\text{proj}_W \mathbf{y}$ and find a vector orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.

$$\text{proj}_W \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{y} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \frac{-2}{6}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + \frac{2}{30}\begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/5 \\ 4/5 \end{bmatrix}.$$

§6.3 Example

Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, $W = \{\mathbf{v}_1, \mathbf{v}_2\}$, and $\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Find $\text{proj}_W \mathbf{y}$ and find a vector orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.

$$\text{proj}_W \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{y} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \frac{-2}{6}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + \frac{2}{30}\begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/5 \\ 4/5 \end{bmatrix}.$$

Thus $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -2/5 \\ 4/5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/5 \\ 1/5 \end{bmatrix}$ is a vector orthogonal to $\mathbf{v}_1$ and $\mathbf{v}_2$.

§6.3 Example

Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, $W = \{\mathbf{v}_1, \mathbf{v}_2\}$, and $\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Find $\text{proj}_W \, \mathbf{y}$ and find a vector orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.

$$\text{proj}_W \, \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 + \frac{\mathbf{y} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 = \frac{-2}{6} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + \frac{2}{30} \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/5 \\ 4/5 \end{bmatrix}.$$

Thus $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -2/5 \\ 4/5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/5 \\ 1/5 \end{bmatrix}$ is a vector orthogonal to $\mathbf{v}_1$ and $\mathbf{v}_2$. Note that we don't have an orthonormal basis for W.

§6.3 Example

Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, $W = \{\mathbf{v}_1, \mathbf{v}_2\}$, and $\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.
Find $\text{proj}_W \mathbf{y}$ and find a vector orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.

$$\text{proj}_W \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{y} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \frac{-2}{6}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + \frac{2}{30}\begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/5 \\ 4/5 \end{bmatrix}.$$

Thus $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -2/5 \\ 4/5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/5 \\ 1/5 \end{bmatrix}$ is a vector orthogonal to $\mathbf{v}_1$ and $\mathbf{v}_2$. Note that we don't have an orthonormal basis for W. How do we obtain one?

§6.3 Example continued

§6.3 Example continued

Let $\mathbf{u}_1, \mathbf{u}_2$ be $\mathbf{v}_1, \mathbf{v}_2$ normalized.

§6.3 Example continued

Let $\mathbf{u}_1, \mathbf{u}_2$ be $\mathbf{v}_1, \mathbf{v}_2$ normalized. Then

$$U = [\mathbf{u}_1\ \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{6} & 5/\sqrt{30} \\ 1/\sqrt{6} & -1/\sqrt{30} \\ -2/\sqrt{6} & 2/\sqrt{30} \end{bmatrix},$$

§6.3 Example continued

Let $\mathbf{u}_1, \mathbf{u}_2$ be $\mathbf{v}_1, \mathbf{v}_2$ normalized. Then

$$U = [\mathbf{u}_1 \ \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{6} & 5/\sqrt{30} \\ 1/\sqrt{6} & -1/\sqrt{30} \\ -2/\sqrt{6} & 2/\sqrt{30} \end{bmatrix},$$

and

$$\begin{aligned} UU^T &= \begin{bmatrix} 1/\sqrt{6} & 5/\sqrt{30} \\ 1/\sqrt{6} & -1/\sqrt{30} \\ -2/\sqrt{6} & 2/\sqrt{30} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \\ 5/\sqrt{30} & -1/\sqrt{30} & 2/\sqrt{30} \end{bmatrix} \\ &= \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/15 & -2/5 \\ -1/3 & -2/5 & 4/5 \end{bmatrix}. \end{aligned}$$

§6.3 Example continued

Let $\mathbf{u}_1, \mathbf{u}_2$ be $\mathbf{v}_1, \mathbf{v}_2$ normalized. Then

$$U = [\mathbf{u}_1 \ \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{6} & 5/\sqrt{30} \\ 1/\sqrt{6} & -1/\sqrt{30} \\ -2/\sqrt{6} & 2/\sqrt{30} \end{bmatrix},$$

and

$$\begin{aligned} UU^T &= \begin{bmatrix} 1/\sqrt{6} & 5/\sqrt{30} \\ 1/\sqrt{6} & -1/\sqrt{30} \\ -2/\sqrt{6} & 2/\sqrt{30} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \\ 5/\sqrt{30} & -1/\sqrt{30} & 2/\sqrt{30} \end{bmatrix} \\ &= \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/15 & -2/5 \\ -1/3 & -2/5 & 4/5 \end{bmatrix}. \end{aligned}$$

Now that we have computed UU^T, what do you think we should check?

§6.3 Example continued

Let $\mathbf{u}_1, \mathbf{u}_2$ be $\mathbf{v}_1, \mathbf{v}_2$ normalized. Then

$$U = [\mathbf{u}_1 \ \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{6} & 5/\sqrt{30} \\ 1/\sqrt{6} & -1/\sqrt{30} \\ -2/\sqrt{6} & 2/\sqrt{30} \end{bmatrix},$$

and

$$UU^T = \begin{bmatrix} 1/\sqrt{6} & 5/\sqrt{30} \\ 1/\sqrt{6} & -1/\sqrt{30} \\ -2/\sqrt{6} & 2/\sqrt{30} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \\ 5/\sqrt{30} & -1/\sqrt{30} & 2/\sqrt{30} \end{bmatrix}$$

$$= \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/15 & -2/5 \\ -1/3 & -2/5 & 4/5 \end{bmatrix}.$$

Now that we have computed UU^T, what do you think we should check? That $UU^T\mathbf{y}$ matches our computation of $\text{proj}_W \mathbf{y}$ from before!

§6.3 Example concluded

§6.3 Example concluded

We now verify that $\mathrm{proj}_W \mathbf{y} = UU^T\mathbf{y}$.

§6.3 Example concluded

We now verify that $\text{proj}_W\,\mathbf{y} = UU^T\mathbf{y}$.

$$UU^T\mathbf{y} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/15 & -2/5 \\ -1/3 & -2/5 & 4/5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/15 \\ 4/5 \end{bmatrix} = \text{proj}_W\,\mathbf{y}.$$

§6.3 Example concluded

We now verify that $\text{proj}_W\,\mathbf{y} = UU^T\mathbf{y}$.

$$UU^T\mathbf{y} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/15 & -2/5 \\ -1/3 & -2/5 & 4/5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/15 \\ 4/5 \end{bmatrix} = \text{proj}_W\,\mathbf{y}.$$

At this point, we've exhibited some of the ways in which we can use an orthogonal/orthonormal basis.

§6.3 Example concluded

We now verify that $\text{proj}_W \mathbf{y} = UU^T\mathbf{y}$.

$$UU^T\mathbf{y} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/15 & -2/5 \\ -1/3 & -2/5 & 4/5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/15 \\ 4/5 \end{bmatrix} = \text{proj}_W \mathbf{y}.$$

At this point, we've exhibited some of the ways in which we can use an orthogonal/orthonormal basis. However, we have not shown how we get them, or if they even exist in all cases!

§6.3 Example concluded

We now verify that $\text{proj}_W \mathbf{y} = UU^T\mathbf{y}$.

$$UU^T\mathbf{y} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/15 & -2/5 \\ -1/3 & -2/5 & 4/5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/15 \\ 4/5 \end{bmatrix} = \text{proj}_W \mathbf{y}.$$

At this point, we've exhibited some of the ways in which we can use an orthogonal/orthonormal basis. However, we have not shown how we get them, or if they even exist in all cases! But don't be alarmed, in the next class (Monday) Emma Hartman will give a guest lecture showing that an orthogonal/orthonormal basis can always be obtained.

§6.3 Example concluded

We now verify that $\text{proj}_W \mathbf{y} = UU^T\mathbf{y}$.

$$UU^T\mathbf{y} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/15 & -2/5 \\ -1/3 & -2/5 & 4/5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/15 \\ 4/5 \end{bmatrix} = \text{proj}_W \mathbf{y}.$$

At this point, we've exhibited some of the ways in which we can use an orthogonal/orthonormal basis. However, we have not shown how we get them, or if they even exist in all cases! But don't be alarmed, in the next class (Monday) Emma Hartman will give a guest lecture showing that an orthogonal/orthonormal basis can always be obtained. Better yet, she will show you an explicit algorithm to compute it!

§6.3 Example concluded

We now verify that $\text{proj}_W \mathbf{y} = UU^T\mathbf{y}$.

$$UU^T\mathbf{y} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/15 & -2/5 \\ -1/3 & -2/5 & 4/5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/15 \\ 4/5 \end{bmatrix} = \text{proj}_W \mathbf{y}.$$

At this point, we've exhibited some of the ways in which we can use an orthogonal/orthonormal basis. However, we have not shown how we get them, or if they even exist in all cases! But don't be alarmed, in the next class (Monday) Emma Hartman will give a guest lecture showing that an orthogonal/orthonormal basis can always be obtained. Better yet, she will show you an explicit algorithm to compute it! If you simply can't wait, take a look at §6.4 in the textbook over the weekend.