Math 22 –
Linear Algebra and its applications

- Lecture 22 -

Instructor: Bjoern Muetzel

GENERAL INFORMATION

- **Office hours:** **Tu 1-3 pm**, Th, Sun 2-4 pm in **KH 229**

  Tutorial: **Tu**, Th, Sun **7-9 pm** in **KH 105**

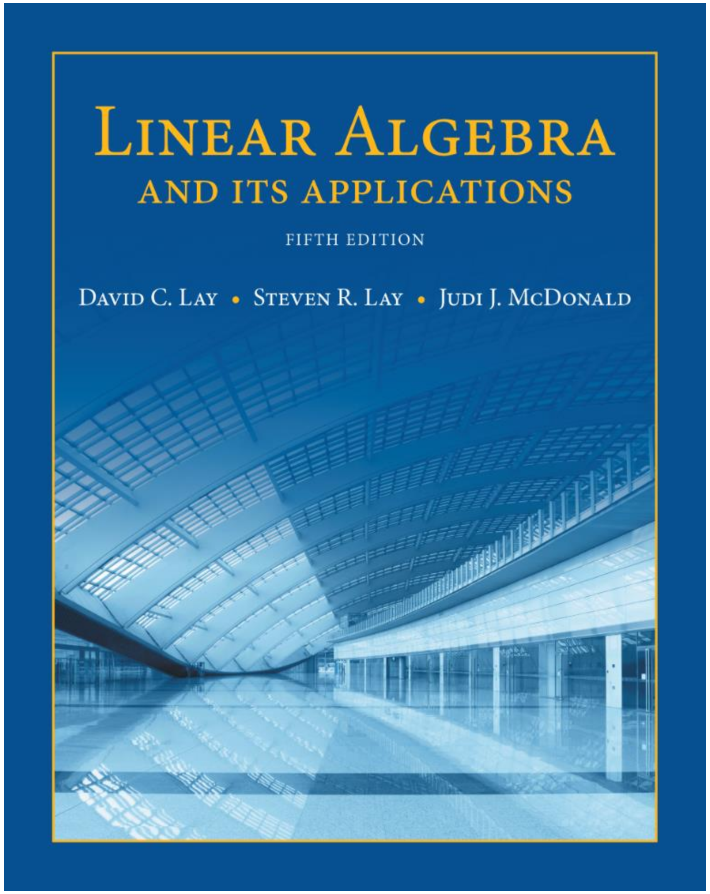
- **Homework 7**: due **Wednesday** at **4 pm** outside **KH 008**

- **Thursday:** **x-hour** will be a **lecture**

6 Orthogonality and Least Squares

6.3

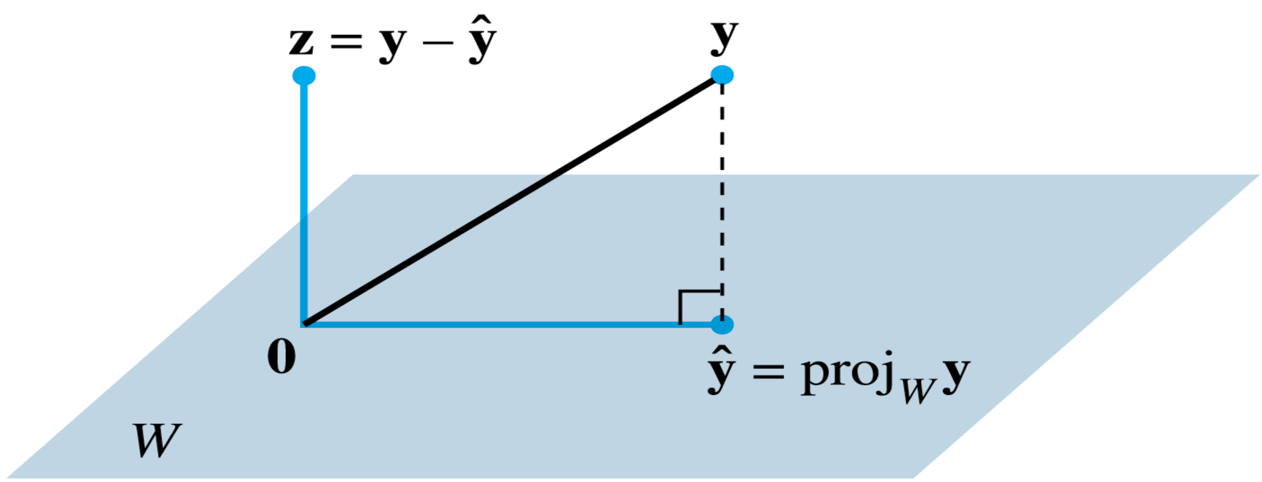
ORTHOGONAL PROJECTIONS



Summary:

1.) We can find the **orthogonal projection** of a **vector** $\boldsymbol{y}$ in $\mathbb{R}^n$ onto a **subspace** $\boldsymbol{W}$. This allows us to **approximate** the **vector** $\boldsymbol{y}$ with a vector $\hat{\boldsymbol{y}}$ in $\boldsymbol{W}$.

2.) We will see that, in a certain sense, this is the **best approximation** of a vector **y** with a vector in **W**.



The orthogonal projection of **y** onto W.

GEOMETRIC INTERPRETATION

THE ORTHOGONAL DECOMPOSITION THEOREM

- **Theorem 8:** Let ***W*** be a **subspace** of $\mathbb{R}^n$. Then each y in $\mathbb{R}^n$ can be written uniquely in the form

$$\boxed{y = \underbrace{\hat{\boldsymbol{y}}}_{\text{in } W} + \underbrace{\boldsymbol{z}}_{\text{in } W^{\perp}}} \quad . \qquad (1)$$

- In fact, if $\{u_1, u_2, \ldots, u_p\}$ is any orthogonal basis of W, then

$$\boxed{\hat{\boldsymbol{y}} = \frac{\boldsymbol{y}\cdot\boldsymbol{u_1}}{\boldsymbol{u_1}\cdot\boldsymbol{u_1}}\boldsymbol{u_1} + \cdots + \frac{\boldsymbol{y}\cdot\boldsymbol{u_p}}{\boldsymbol{u_p}\cdot\boldsymbol{u_p}}\boldsymbol{u_p}} \qquad (2) \quad \text{or}$$

$$\boxed{\hat{\boldsymbol{y}} = proj_{\boldsymbol{u_1}}(y) + \cdots + proj_{\boldsymbol{u_p}}(y) = \mathbf{proj}_W\,(\mathrm{y})}$$

and $$\boxed{\boldsymbol{z} = y - \hat{\boldsymbol{y}}} \quad .$$

Note: The vector $\hat{\boldsymbol{y}}$ = $\mathbf{proj}_W$ (y) is called the **orthogonal projection of y onto** *W*. The total projection decomposes into **line projections**.

Picture:

Proof of Theorem 8: 1.) This construction is correct

a) $\hat{y}$ in W: it can be written as a linear combination of basis vectors of W.

b) z is in $W^{\perp}$

- We know that $\boldsymbol{z} = \boldsymbol{y} - \widehat{\boldsymbol{y}}$. Since u_1 is orthogonal to $u_2, \dots, u_p$, it follows from the equation $\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \cdots + \frac{y \cdot u_p}{u_p \cdot u_p} u_p$ that

- Thus z is orthogonal to u_1. Similarly, z is orthogonal to each u_j in the basis for W. Hence z is orthogonal to every vector in W. That is, z is in $W^{\perp}$.

THE ORTHOGONAL DECOMPOSITION THEOREM

2.) Uniqueness of the decomposition:

Example 1: Let $u_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, u_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $y = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$.

1.) Show that $\{u_1, u_2\}$ is an orthogonal basis for $W = Span\{u_1, u_2\}$.

2.) Write y as the sum of a vector $\hat{y}$ in W and a vector z in $W^{\perp}$.

3.) Draw a picture of u_1, u_2, y and $\hat{y}$ in $\mathbb{R}^3$.

THE BEST APPROXIMATION THEOREM

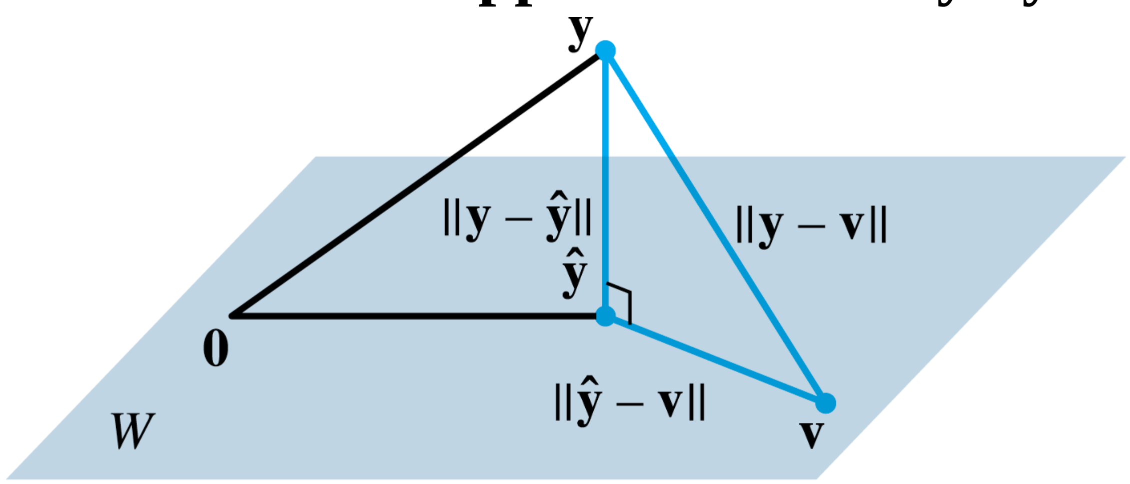
- **Theorem 9:** Let W be a subspace of $\mathbb{R}^n$ and y be a vector in $\mathbb{R}^n$. Let $\hat{y} = proj_W(y)$ be the orthogonal projection of y onto W.

Then $\hat{\boldsymbol{y}}$ is the **closest point in W to y**, i.e.

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

for all v in W distinct from $\hat{y}$. Hence $\|y - \hat{y}\| = \text{dist}(y, W)$.

- The vector $\hat{y}$ is called **the best approximation to y by elements of** W.



The orthogonal projection of **y** onto W is the closest point in W to **y**.

Note: The **distance** from *y* to *v*, given by $||\boldsymbol{y} - \boldsymbol{v}||$, can be regarded as the **"error"** of using *v* in place of *y*. The theorem says that this **error is minimized** when $\boldsymbol{v} = \hat{\boldsymbol{y}}$.

Proof of Theorem 9:

PROPERTIES OF ORTHOGONAL PROJECTIONS

- **Example 2:** Let $u_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and $y = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}$.
- Let $W = Span\{u_1, u_2\}$. Show that $\{u_1, u_2\}$ is an orthogonal basis for W. Then find the distance

$$\|y - \hat{y}\| = \text{dist}(\text{y}, \text{W}) \qquad \text{from } \mathbf{y} \text{ to } W.$$

PROPERTIES OF ORTHOGONAL PROJECTIONS

- **Theorem 10:** If $\{u_1, u_2, \ldots, u_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\hat{y} = proj_W(y) = (y \cdot u_1)u_1 + (y \cdot u_2)u_2 + \cdots + (y \cdot u_p)u_p$$

If $\mathrm{U} = [u_1, u_2, \ldots, u_p]$, then

$$proj_W(y) = UU^T y \quad \text{for all } y \text{ in } \mathbb{R}^n.$$

- **Proof:** The first part follows immediately from **Theorem 8**. For the second part we rewrite the equation in matrix notation.