Math 2270-1

Notes of 11/8/19

Orthogonal Projections

- We start with revisiting the idea of projecting a point $\mathbf{y}$ in $\mathbb{R}^2$ onto a line through the origin.
- The projection of $\mathbf{y}$ is the point on the line that is closest to $\mathbf{y}$.

- This idea can be generalized to projecting a point $\mathbf{y}$ in $\mathbb{R}^n$ onto a subspace of R^n.
- **Theorem 8.** Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

  where $\hat{y}$ is in W and $\mathbf{z}$ is in $W^\perp$.
- This is the **orthogonal Decomposition theorem.** The vector $\hat{\mathbf{y}}$ in (1) is called the **orthogonal projection of $\mathbf{y}$ onto W**.
- The textbook uses the notation

$$\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}.$$

- The textbook proves the Theorem by actually computing $\hat{\mathbf{y}}$ and $\mathbf{z}$:
- In fact, if $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_p}\}$ is any orthogonal basis of W, then

$$\begin{aligned}\hat{y} &= \frac{\mathbf{y} \bullet \mathbf{u}_1}{\mathbf{u}_1 \bullet \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \bullet \mathbf{u}_2}{\mathbf{u}_2 \bullet \mathbf{u}_2}\mathbf{u}_2 + \ldots + \frac{\mathbf{y} \bullet \mathbf{u}_p}{\mathbf{u}_p \bullet \mathbf{u}_p}\mathbf{u}_p \\ &= \sum_{i=1}^{p} \frac{\mathbf{y} \bullet \mathbf{u}_i}{\mathbf{u}_i \bullet \mathbf{u}_i}\mathbf{u}_i\end{aligned}$$

  and

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

- However, the uniqueness of the decomposition (1) shows that the orthogonal projections depends only on W and not on its basis.

- Example 2, textbook. Let

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Geometric Interpretation of Orthogonal Projection

- Note that if the dimension of W is one, and $\{\mathbf{u}\}$ is a basis of W then the orthogonal projection of $\mathbf{y}$ onto W is just

$$\hat{\mathbf{y}} = \frac{\mathbf{u} \bullet \mathbf{y}}{\mathbf{u} \bullet \mathbf{u}} \mathbf{u}.$$

- Thus the terms $\frac{\mathbf{u}_i \bullet \mathbf{u}_i}{\mathbf{u}_i \bullet \mathbf{u}_i} \mathbf{u}_i$ in

$$\hat{y} = \sum_{i=1}^{p} \frac{\mathbf{u}_i \bullet \mathbf{y}}{\mathbf{u}_i \bullet \mathbf{u}_i} \mathbf{u}_i$$

are just the projections of $\mathbf{y}$ onto the spaces $\text{span}\{\mathbf{u}_i\}$.

- As in our initial example, the orthogonal projection $\hat{\mathbf{y}}$ of $\mathbf{y}$ onto W is the point in W that is closest to $\mathbf{y}$. This is the contents of

- **Best Approximation Theorem** (Theorem 9, p. 352) Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point W to $\mathbf{y}$, in the sense that

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

  for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

- This is a simple consequence of the Pythagorean Theorem.

- Example 4, p. 353, textbook. Let

$$\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad \text{and} \quad W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2\}$$

  Compute the point $\hat{y}$ in W that is closest to $\mathbf{y}$ and its distance from $\mathbf{y}$.

- Formulas simplify if our basis is orthonormal, rather than just orthogonal.

- **Theorem 10**, p. 353. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\hat{y} = \mathrm{proj}_W \mathbf{y} = \sum_{i=1}^{p} (\mathbf{y} \bullet \mathbf{u}_i)\mathbf{u}_i.$$

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

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

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

- Next question: How do we get orthonormal bases?