Video Lecture G5: Orthogonal Projection 2

Tom Roby

Outline & Objectives

- Demonstrate the *Best Approximation Theorem* that the orthogonal projection $\widehat{y} = \mathrm{Proj}_W \vec{y}$ is the closest point in W to $\vec{y}$.

- Analyze how our formula for $\widehat{y} = \mathrm{Proj}_W \vec{y}$ simplifies when we have an orthonormal basis for W, and rewrite this in terms of orthogonal matrices.

Best Approximation Theorem

Recall: $\widehat{y} = \operatorname{Proj}_W \vec{y} = \frac{\vec{y}\cdot\vec{u_1}}{\vec{u_1}\cdot\vec{u_1}}\vec{u_1} + \cdots + \frac{\vec{y}\cdot\vec{u_p}}{\vec{u_p}\cdot\vec{u_p}}\vec{u_p}$

Theorem (Best Approximation Theorem)

Let W be a subspace of $\mathbb{R}^n$ and $\vec{y} \in \mathbb{R}^n$, and set $\widehat{y} = \operatorname{Proj}_W \vec{y}$.
Then $\quad \|\vec{y} - \widehat{y}\| < \|\vec{y} - \vec{v}\| \quad \forall\, \vec{v} \neq \widehat{y}$ *in* W.

Pf: $\vec{y} - \vec{v} = (\vec{y} - \widehat{y}) + (\widehat{y} - \vec{v})$. Since $(\vec{y} - \widehat{y}) \perp (\widehat{y} - \vec{v})$ (why?), get $\|\vec{y} - \vec{v}\|^2 = \|(\vec{y} - \widehat{y}\|^2 + \|\widehat{y} - \vec{v}\|^2$. ■

$$W = \operatorname{Span}\left\{\begin{bmatrix}3\\-1\\-2\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\}, \vec{y} = \begin{bmatrix}6\\4\\-7\end{bmatrix} \implies$$

$$\widehat{y} = \frac{28}{14}\begin{bmatrix}3\\-1\\-2\end{bmatrix} + \frac{3}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}7\\-1\\-3\end{bmatrix}$$

Projecting onto Orthonormal basis

Recall: $\widehat{y} = \mathrm{Proj}_W \vec{y} = \frac{\vec{y}\cdot\vec{u_1}}{\vec{u_1}\cdot\vec{u_1}}\vec{u_1} + \cdots + \frac{\vec{y}\cdot\vec{u_p}}{\vec{u_p}\cdot\vec{u_p}}\vec{u_p}$

Theorem

Let $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p}\}$ be an orthonormal basis for a subspace W of $\mathbb{R}^n$. Then $\widehat{y} = \mathrm{Proj}_W \vec{y} = (\vec{y}\cdot\vec{u_1})\vec{u_1} + \cdots + (\vec{y}\cdot\vec{u_p})\vec{u_p}$. If $U = [\vec{u_1} \quad \ldots \quad \vec{u_p}]$, then $\widehat{y} = \mathrm{Proj}_W \vec{y} = UU^T\vec{y}$ for all $\vec{y} \in \mathbb{R}^n$.

Pf: Write out $U^T\vec{y}$ in terms of $\vec{u_i}'s$ as row vectors... ■

Here we must have $p \leq n$. *(Why?)*

How does U^TU compare with UU^T?