1 Theorem 9 : The Best Approximation Theorem

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 in W to $\mathbf{y}$, in the sense that

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

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

What is the use of this theorem?

In fact, this theorem is significantly important to Statistics.

1.1 Least-Squares Problems

If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a **least-squares solution** of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

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

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

Theorem 13 says that the solution set of

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

is the same as the solution set of

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

From this fact, we know that there exists a unique solution for such problem if and only if A^TA is invertible (or rank $A = n$).

Statistical presumptions of Linear Models are based on this theorem.

$$X\beta = \mathbf{y}$$

1. Find an orthogonal basis for the column space of the matrix

$$\begin{pmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{pmatrix}$$

2. Find the orthogonal projection of $\mathbf{b}$ onto Col A and a least-squares solution of $A\mathbf{x} = \mathbf{b}$.

a.

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix}$$

b.

$$A = \begin{pmatrix} 4 & 0 & 1 \\ 1 & -5 & 1 \\ 6 & 1 & 0 \\ 1 & -1 & -5 \end{pmatrix}, \quad B = \begin{pmatrix} 9 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

3. Mark each statement True or False. Justify your answer. (A is an $m \times n$ matrix and $\mathbf{b}$ is in $\mathbb{R}^m$.)

a. A least-squares solution of $A\mathbf{x} = \mathbf{b}$ is a vector $\hat{\mathbf{x}}$ that satisfies $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, where $\hat{\mathbf{b}}$ is the orthogonal projection of $\mathbf{b}$ onto Col A.

b. If the columns of A are linearly independent, then the equation $A\mathbf{x} = \mathbf{b}$ has exactly one least-squares solution.

c. If $\mathbf{b}$ is in the column space of A, then every solution of $A\mathbf{x} = \mathbf{b}$ is a least-squares solution.

d. A least-squares solution of $A\mathbf{x} = \mathbf{b}$ is a list of weights that, when applied to the columns of A, produces the orthogonal projection of $\mathbf{b}$ onto Col A.

4. In Assignment 7, you are given a problem ; Show that Nul $A =$ Nul A^TA. Using this fact without proof, show that

$$\text{rank } A^TA = \text{rank } A$$