Least Squares Approximation

Definition: If W is a subspace of $\mathbb{R}^n$ and $\mathbf{v}$ is a vector in $\mathbb{R}^n$, then the **best approximation** to $\mathbf{v}$ in W is the vector $\mathbf{w}$ in W that minimizes the Euclidean distance $\|\mathbf{v} - \mathbf{w}\|$. In other words

$$\|\mathbf{v} - \mathbf{w}\| < \|\mathbf{v} - \mathbf{u}\|$$

for every vector $\mathbf{u}$ in W different from $\mathbf{w}$.

The Best Approximation Theorem: If W is a subspace of $\mathbb{R}^n$ and $\mathbf{v}$ is a vector in $\mathbb{R}^n$, then the best approximation to $\mathbf{v}$ in W is the vector $\mathbf{w} = \text{proj}_W(\mathbf{v})$.

Definition: If A is an $m \times n$ matrix and $\mathbf{b}$ is in $\mathbb{R}^m$, a **least squares solution** of $A\mathbf{x} = \mathbf{b}$ is a vector $\mathbf{x}_{LS}$ in $\mathbb{R}^n$ such that

$$\|\mathbf{b} - A\mathbf{x}_{LS}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

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

The Least Squares Theorem: Let A be an $m \times n$ matrix and let $\mathbf{b}$ be in $\mathbb{R}^m$. Then $A\mathbf{x} = \mathbf{b}$ always has at least one least squares solution $\mathbf{x}_{LS}$. Moreover:

(a) $\mathbf{x}_{LS}$ is a least squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if $\mathbf{x}_{LS}$ is a solution of the normal equations $A^T A\mathbf{x}_{LS} = A^T\mathbf{b}$.

(b) A has linearly independent columns if and only if $A^T A$ is invertible. In this case, the least squares solution of $A\mathbf{x} = \mathbf{b}$ is unique and is given by

$$\mathbf{x}_{LS} = (A^T A)^{-1} A^T \mathbf{b}$$