Orthogonal projection

Theorem: Let V be an inner product space, let v be a vector in V, and let W be a finite dimensional linear subspace of V. Then there is a unique vector $w := \pi_W(v) \in W$ such that $v - w \in W^\perp$. Furthermore:

- $\pi_W(v)$ is the vector in W which is closest to $\mathbf{v}$.
- The map $\pi_W : V \to W$ is a linear transformation.

Let $\pi_W^\perp(\mathbf{v}) := v - \pi_W(\mathbf{v})$. Then

$$\mathbf{v} = \pi_W(\mathbf{v}) + \pi_W^\perp(\mathbf{v}),$$

and $\pi_W(\mathbf{v}) \in W$ and $\pi_W^\perp(\mathbf{v}) \in W^\perp$.

The map $\pi_W^\perp : V \to W^\perp$ is also a linear transformation.

Formulas

If $A \in M_{mn}$, $X \in M_{n1}$, and $Y \in M_{m1}$,

$$(A^T Y|X) = (Y|AX).$$

Let $A \in M_{mn}$ and $Y \in M_{n1}$. Then $Y \in CS(A)$ if and only if the equation $Y = AX$ is has a solution.

Theorem: For any Y, there is an $\overline{X}$ such that $A^T Y = A^T A\overline{X}$. For any such $\overline{X}$, $\overline{Y} := A\overline{X}$ is the orthogonal projection of Y onto $CS(A)$.

Lemma:

1. $NS(A^T A) = NS(A)$.

2. $rank(A^T A) = rank(A) = rank(A^T)$.

3. $CS(A^T A) = CS(A^T)$.

Proof: If $X \in NS(A)$, $AX = 0$, hence $A^TAX = 0$, hence $X \in NS(A^TA)$. If $X \in NS(A^TA)$, then $(AX|AX) = (A^TAX|X) = (\mathbf{0}X|X) = 0$, hence $AX = 0$. Thus $X \in NS(A)$. This proves (1). Since A^TA and A both have n columns, their rank is n minus the dimension of their null spaces. Hence they have the same rank. The rank of A is the dimension of its column space, which is the same as the dimension of its row space, which is the rank of A^T. This proves (2). If $Z \in CS(A^TA)$, then there exists an X such that $Z = (A^TA)(X)$. But then $Z = A^T(AX) \in CS(A^T)$. Hence $CS(A^TA) \subseteq CS(A)$. Since the spaces have the same dimension, they must be equal.

Proof of theorem: Suppose $Y \in M_{m1}$. Then $A^TY \in CS(A^T) = CS(A^TA)$. Hence there exists $\overline{X}$ with $A^TA\overline{X} = A^TA$. Let $\overline{Y} := A\overline{X}$. Then $A^T\overline{Y} = A^TY$. Hence for any X,

$$(Y|AX) = (A^TY|X) = (A^T\overline{Y}|X) = (\overline{Y}|AX).$$

This implies that $(Y - \overline{Y}|AX) = 0$ for any X, so $Y - \overline{Y} \in (CS(A))^\perp$.