Orthogonal Projections

Matteo Paz, Dylan Rupel

April 17, 2024

The orthogonal projection is a way to find the closest point in a subspace to any point in the whole space. You only need an orthonormal basis of the subspace to carry out the operation.

Given a basis for a subspace $W \in \mathbb{R}^n$, the Gram-Schmidt process produces an orthonormal basis $\mathbf{w}_1, \ldots, \mathbf{w}_k$.

Definition 1 *The orthonormal projection* $\text{proj}_W(\mathbf{x})$ *for* $\mathbf{x} \in \mathbb{R}^n$ *is*

$$\mathbf{w}_1(\mathbf{w}_1 \cdot \mathbf{x}) + \cdots + \mathbf{w}_k(\mathbf{w}_k \cdot \mathbf{x})$$

Which can be rewritten as

$$\mathbf{w}_1\mathbf{w}_1^T\mathbf{x} + \cdots + \mathbf{w}_k\mathbf{w}_k^T\mathbf{x}$$

$$(\mathbf{w}_1\mathbf{w}_1^T + \cdots + \mathbf{w}_k\mathbf{w}_k^T)\mathbf{x}$$

By this definition, we then know that proj **is a linear transformation, with matrix** QQ^T **where** $Q = \left[\ \mathbf{w}_1 | \cdots | \mathbf{w}_k\ \right]$ **and** $\mathbf{w}_i$ **are column vectors.**

Proposition 1 $\text{proj}_W(\mathbf{x})$ *is the vector in* W *closest to* $\mathbf{x}$

Proof.
Firstly notice that $\text{proj}_W(\mathbf{x}) - \mathbf{x} \in W^\perp$:

$$\mathbf{w}_i \cdot (\mathbf{x} - \mathit{Proof.}_W(\mathbf{x}))$$

$$\mathbf{w}_i \cdot \mathbf{x} - (\mathbf{w}_i \cdot \mathbf{w}_i)(\mathbf{w}_i \cdot \mathbf{x})$$

$$\mathbf{w}_i \cdot \mathbf{x} - \mathbf{w}_i \cdot \mathbf{x} = 0$$

Thus $\text{proj}_W(\mathbf{x}) - \mathbf{x} \in W^\perp$. Since this vector is perpendicular to W, we can use the pythagorean theorem. Choose any other vector $\mathbf{v} \in W$. Now:

$$|\mathbf{x} - \mathbf{v}|^2 = |\mathbf{x} - \text{proj}_W(\mathbf{x})|^2 + |\text{proj}_W(\mathbf{x}) - \mathbf{v}|^2$$

Thus given $v \neq \text{proj}_W(\mathbf{x})$, we have $|\mathbf{x} - \mathbf{v}|^2 > |\mathbf{x} - \text{proj}_W(\mathbf{x})|^2$. So $\text{proj}_W(\mathbf{x})$ is closer to $\mathbf{x}$ than any other element of W.

Normal Equations

Lets say that we want to find a solution to the equation $A\mathbf{x} = b$. However, it might not be the case that $b \in \text{im}A$. Thus there might not exist any exact solution to this, but what is the closest we can get?

Let $\hat{b}$ be this point. Thus $\hat{b} = \text{proj}_{\text{im}(A)}(b)$. Consider the difference $b - \hat{b}$. As proven earlier, this must be in the orthogonal complement of the subspace:

$$b - \hat{b} \in \text{im}(A)^{\perp}$$

This means that it also lies in the null space of the transpose.

$$A^T(b - \hat{b}) = 0$$

$$A^T b = A^T \hat{b}$$

If x is the least-squares solution, then we should get

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

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

$$A^T A\mathbf{x} = A^T b \tag{1}$$

$$\mathbf{x} = (A^T A)^{-1} A^T b \tag{2}$$

Equation number (1) are called the *normal equations*, and are general. Equation number (2) allows for the evaluation of the exact solution but requires that $A^T A$ is invertible.