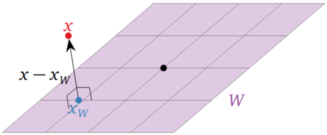
Orthogonal Projection

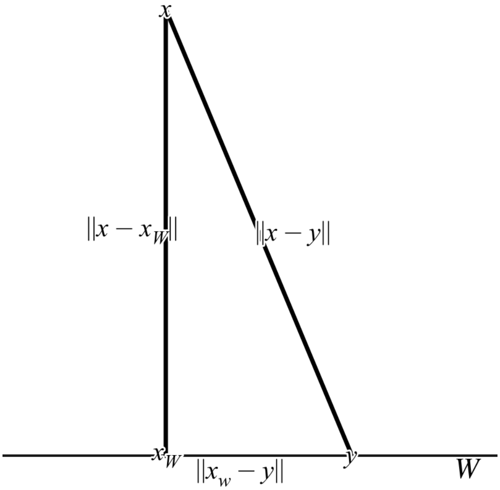


In some applications, we have a vector $\mathbf{x}$ that's not in a space we want, and can sometimes be content with the "closest" vector to $\mathbf{x}$ that lives in our space W.

Definition

Orthogonal Projection: We call this vector $\mathbf{x}_W$ to be the orthogonal projection of $\mathbf{x}$ onto the space W.

Why call it Orthogonal? An $\mathbb{R}^2$ Figure



If we take any other point as $\mathbf{x}_W$, then we see that it would be further from $\mathbf{x}$. See that the vector $\mathbf{x} - \mathbf{x}_W$ is orthogonal to W!

Orthogonal Decomposition

Let's suppose we can compute this $\mathbf{x}_W$, and note something.

Definition

Orthogonal Decomposition: Let W be a subspace of $\mathbb{R}^n$, and $\mathbf{x} \in \mathbb{R}^n$. Then, we can write $\mathbf{x}$ as

$$\mathbf{x} = \mathbf{x}_W + \mathbf{x}_{W^\perp}$$

This is called the orthogonal decomposition of $\mathbf{x}$. Where $\mathbf{x}_W$ is the orthogonal projection of $\mathbf{x}$ onto W and $\mathbf{x}_{W^\perp} = \mathbf{x} - \mathbf{x}_W$

Computing an Orthogonal Projection

Theorem

Let $A \in \mathbb{R}^{m \times n}$, $W = \mathrm{Col}\,(A)$, and $\mathbf{x} \in \mathbb{R}^m$. Then the system of linear equations given by

$$A^\top A\mathbf{c} = A^\top \mathbf{x}$$

is consistent and $\mathbf{x}_W = A\mathbf{c}$ where $\mathbf{c}$ is some solution.

Note: We sometimes call this equation the "normal equations", which is particularly important for statistics applications when finding covariances of random variables.
Note that if $n = 1$, then we have inner products instead of matrix multiplications!

Finding Orthogonal Projection Example

Let $W = \text{Span}\left(\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right)$. Find an orthogonal projection of $\mathbf{x} = \begin{bmatrix}2\\1\\4\end{bmatrix}$

We will first solve $A^\top A\mathbf{c} = A^\top \mathbf{x}$

$$A^\top A = \begin{bmatrix}1 & 0 & -1\\1 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\0 & 1\\-1 & 0\end{bmatrix} = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix} \qquad A^\top \mathbf{x} = \begin{bmatrix}1 & 0 & -1\\1 & 1 & 0\end{bmatrix}\begin{bmatrix}2\\1\\4\end{bmatrix} = \begin{bmatrix}-2\\3\end{bmatrix}$$

$$\left[\begin{array}{c|c} A^\top A & A^\top \mathbf{x}\end{array}\right] = \left[\begin{array}{cc|c} 2 & 1 & -2\\ 1 & 2 & 3\end{array}\right] \xrightarrow{R_2 = R_2 - \frac{1}{2}R_1} \left[\begin{array}{cc|c} 2 & 1 & -2\\ 0 & \frac{3}{2} & 4\end{array}\right] \xrightarrow{R_1 = R_1 - \frac{2}{3}R_2} \left[\begin{array}{cc|c} 2 & 0 & -\frac{14}{3}\\ 0 & \frac{3}{2} & 4\end{array}\right]$$

$$\xrightarrow[R_2 = \frac{2}{3}R_2]{R_1 = \frac{1}{2}R_1} \left[\begin{array}{cc|c} 1 & 0 & -\frac{7}{3}\\ 0 & 1 & \frac{8}{3}\end{array}\right] \qquad \mathbf{x}_W = A\mathbf{c} = \frac{1}{3}\begin{bmatrix}1 & 1\\0 & 1\\-1 & 0\end{bmatrix}\begin{bmatrix}-7\\8\end{bmatrix} = \frac{1}{3}\begin{bmatrix}1\\8\\7\end{bmatrix}$$

Finding Orthogonal Projection Practice

Let $W = \text{Span}\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)$. Find an orthogonal projection of $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$

Orthogonal Projection as a Linear Transformation

Let's define this orthogonal projection to be the transformation T.

$$T : \mathbb{R}^n \to W \qquad T(\mathbf{x}) = \mathbf{x}_W$$

Theorem

T is a linear transformation

Proof.

We will show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $a \in \mathbb{R}$, we have that $T(a\mathbf{x} + \mathbf{y}) = aT(\mathbf{x}) + T(\mathbf{y})$. For our convenience, we define $\mathbf{z} = a\mathbf{x} + \mathbf{y}$. Remember that $\mathbf{z}_W = A\mathbf{c}_z$ where $\mathbf{c}_z$ is a solution to $A^\top A\mathbf{c}_z = A^\top \mathbf{z}$, and similarly for $\mathbf{x}, \mathbf{y}$, so we need only show that $\mathbf{c}_z = a\mathbf{c}_x + \mathbf{c}_y$ is a solution to our system above.

$$A^\top A\mathbf{c}_z = A^\top A(a\mathbf{c}_x + \mathbf{c}_y) = aA^\top A\mathbf{c}_x + A^\top A\mathbf{c}_y = aA^\top \mathbf{x} + A^\top \mathbf{y} = A^\top(a\mathbf{x} + \mathbf{y}) = A^\top \mathbf{z}$$

□

Properties of Orthogonal Projection

Let T be our orthogonal projection as defined in the previous slide, then the following properties are true

1. $T(\mathbf{x}) = \mathbf{x}$ if and only if $\mathbf{x} \in W$
2. $T(\mathbf{x}) = \mathbf{0}$ if and only if $\mathbf{x} \in W^{\perp}$
3. $T \circ T = T$
4. T is surjective.