12 Lecture 12: Orthogonal projections (20/03/2020)

12.1 Projection on a line

Below will follow an intuitive example of the projection of a vector in $\mathbb{R}^2$ on a line L. To find the projection of the vector on L we look for the vector which is closest to $\vec{y}$ and on the line L. This is done by using the vector $\vec{z}$ which is a vector perpendicular to the line L.

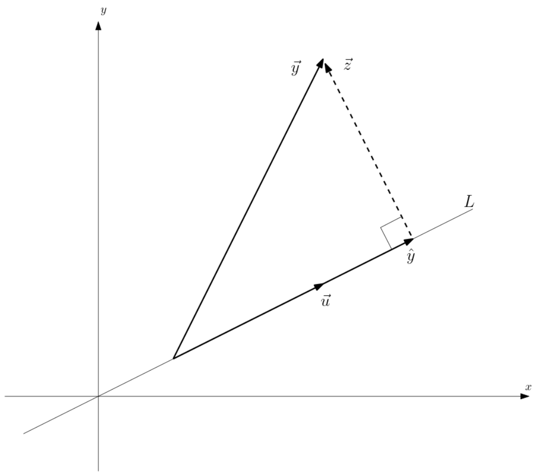

Figure 1: Projection of the vector $\vec{y}$ on the line L where $L = \text{Span}\{\vec{u}\}$, $\vec{u} \neq \vec{0}$ and $\text{proj}_L \vec{y} = \hat{y}$

$$\hat{y} \text{ on } L,\ \vec{z} \perp L \Rightarrow \begin{cases} \hat{y} = \alpha \vec{u} \\ \vec{z} \cdot \vec{u} = 0 \end{cases} \tag{1}$$

$$\vec{z} = \vec{y} - \hat{y} = \vec{y} - \alpha \vec{u} \tag{2}$$

$$0 = \vec{z} \cdot \vec{u} = (\vec{y} - \alpha \vec{u}) \cdot \alpha \tag{3}$$

$$\vec{y} \cdot \vec{u} = \alpha \vec{u} \cdot \vec{u} \tag{4}$$

$$\text{Thus: } \alpha = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \tag{5}$$

This then gives an expression for the vector $\vec{y}$:

$$\hat{y} = \alpha \vec{u} = \text{proj}_L(\vec{y}) = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u} \tag{6}$$

bra-ket notation:

$$\hat{y} = \frac{\langle \vec{y} | \vec{u} \rangle}{\langle \vec{u} | \vec{u} \rangle} \vec{u} \tag{7}$$

12.2 Orthogonal Projections

An Orthogonal projection is an extension of the idea of a perpendicular vector in $\mathbb{R}^2$. Orthogonal vectors can be in any $\mathbb{R}^n$ space. As it turns out, for every vector $\vec{y} \in \mathbb{R}^n$ there is a unique vector $\vec{z} \in W^\perp$ which will give the vector $\hat{y} \in W$.

Theorem 1 *Let W be a subspace of $\mathbb{R}^n$. Then each $\vec{y} \in \mathbb{R}^n$ can be written uniquely in the form:*

$$\vec{y} = \hat{y} + \vec{z} \tag{8}$$

where $\hat{y} = proj_W(\vec{y}) \in W$ and $\vec{z} \in W^\perp$.

The projection of $\vec{y}$ onto W is the closest vector $\hat{y}$ in W that is the closest to $\vec{y}$. This can be expressed as:

$$(\vec{y} - \hat{y}) \perp W \quad \text{or} \quad (\vec{y} - \hat{y}) \in W^{\perp} \tag{9}$$

Theorem 2 *let $\vec{y}$ be a vector in $\mathbb{R}^n$, and W a linear subspace of $\mathbb{R}^n$. There is a unique vecotr $\vec{y} \in W$ and a unique vector $\vec{z} \in W^{\perp}$ such that:*

$$\vec{y} = \hat{y} + \vec{z} \tag{10}$$

if $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is an orthogonal basis of W then:

$$\hat{y} = \sum_{i=1}^{k} \left(\frac{\langle \vec{u}_i | \vec{y} \rangle}{\langle \vec{u}_i | \vec{u}_i \rangle} \right) \tag{11}$$

Note: if $\vec{y} \in W$, then $\vec{y} = \hat{y}$. This means that if a vector $\vec{y}$ is already in the span of the subspace W, then the projection $\hat{y} = proj_W(\vec{y})$ is the vector $\vec{y}$ itself.

12.3 Subspaces with an orthonormal basis

If the basis in theorem 2 is an orthonormal basis the expression reduces to the following:

$$\hat{y} = \sum_{i=1}^{k} \langle \vec{y} | \vec{u}_i \rangle \vec{u}_i \tag{12}$$

Theorem 3 *Let $\{\vec{u}_1, \cdots, \vec{u}_p\}$ be an orthonormal basis of $W \subset \mathbb{R}^n$, then:*

1. *$proj_W(\vec{y}) = UU^T\vec{y}$, with $U = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_p \end{bmatrix}$*
2. *if $P = UU^T$ is the standard matrix of $proj_W$, then $P^2 = P = P^T$*

12.4 Best approximation

Theorem 4 *Let W be a subspace of $\mathbb{R}^n$, $\vec{y}$ a vector in $\mathbb{R}^n$ and $\hat{y}$ an orthogonal projection of $\vec{y}$ onto W. Then:*

$$||\vec{y} - \hat{y}|| \leq ||\vec{y} - \vec{v}|| \tag{13}$$

for all vectors $\vec{v} \in W$.

Theorem 4 states that the distance between the vector $\vec{y}$ and the vector $\hat{y}$ is always shorter then the distance from $\vec{y}$ to any random vector $\vec{v}$ in the subspace W. The equals symbol is for the case where $\vec{v} = \hat{y}$. The vector $\hat{y}$ is thus referred to as the closest approximation to $\vec{y}$ by elements of W. This gives the following expression:

$$\text{dist}(\vec{y}, W) = \text{dist}(\vec{y}, \hat{y}) = ||\vec{y} - \hat{y}|| \tag{14}$$