

Orthogonal Projections

Definition: For S a subspace of \mathbb{R}^n with a given basis: $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ and a point $\mathbf{g} \in \mathbb{R}^n$ then $\mathbf{a} \in S$ is the orthogonal projection of \mathbf{g} onto S if:

$$\forall i \leq k \quad (\mathbf{g} - \mathbf{a}) \cdot \mathbf{u}_i = 0$$

Lemma 1: For a vector sub space S of \mathbb{R}^n with a given basis: $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ if some vector $\mathbf{v} \in S$ is such that for all $i \leq k$ $\mathbf{v} \cdot \mathbf{u}_i = 0$ then $\mathbf{v} = \mathbf{0}$

Proof. Since $\mathbf{v} \in S$ we can write:

$$\mathbf{v} = t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k$$

It is sufficient to show that the magnitude of \mathbf{v} is 0:

$$\begin{aligned} |\mathbf{v}|^2 &= \mathbf{v} \cdot \mathbf{v} = \mathbf{v} \cdot (t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k) = t_1 \mathbf{u}_1 \mathbf{v} + t_2 \mathbf{u}_2 \mathbf{v} + \dots + t_k \mathbf{u}_k \mathbf{v} = 0 \\ &\iff |\mathbf{v}| = 0 \\ &\iff \mathbf{v} = \mathbf{0} \end{aligned}$$

□

Proposition 1: If an orthogonal projection of \mathbf{g} onto S exists it is unique

Proof. Let \mathbf{a} and \mathbf{b} be orthogonal projections of \mathbf{g} onto S

By definition for an arbitrary $i \leq k$ we have:

$$\begin{aligned} 0 &= (\mathbf{g} - \mathbf{a}) \cdot \mathbf{u}_i = (\mathbf{g} - \mathbf{a} + \mathbf{b} - \mathbf{b}) \cdot \mathbf{u}_i = (\mathbf{g} - \mathbf{b} - (\mathbf{a} - \mathbf{b})) \cdot \mathbf{u}_i \\ &= (\mathbf{g} - \mathbf{b}) \cdot \mathbf{u}_i - (\mathbf{a} - \mathbf{b}) \cdot \mathbf{u}_i = -(\mathbf{a} - \mathbf{b}) \cdot \mathbf{u}_i \\ &\iff (\mathbf{a} - \mathbf{b}) \cdot \mathbf{u}_i = 0 \end{aligned}$$

Since $(\mathbf{a} - \mathbf{b}) \in S$ we can apply lemma 1:

$$\begin{aligned} &\Rightarrow (\mathbf{a} - \mathbf{b}) = \mathbf{0} \\ &\iff \mathbf{a} = \mathbf{b} \end{aligned}$$

□

Proposition 2: If $\mathbf{a} \in S$ is an orthogonal projection of \mathbf{g} relative to a basis: $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ then it is also orthogonal relative to any other basis

Proof. suppose that $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is also a basis of S and take an arbitrary \mathbf{v}_i

$$\begin{aligned} (\mathbf{g} - \mathbf{a}) \cdot \mathbf{v}_i &= (\mathbf{g} - \mathbf{a}) \cdot (t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k) \quad \because \mathbf{v}_i \in S \\ &= t_1 (\mathbf{g} - \mathbf{a}) \cdot \mathbf{u}_1 + t_2 (\mathbf{g} - \mathbf{a}) \cdot \mathbf{u}_2 + \dots + t_k (\mathbf{g} - \mathbf{a}) \cdot \mathbf{u}_k = 0 \\ &\Rightarrow (\mathbf{g} - \mathbf{a}) \cdot \mathbf{v}_i = 0 \end{aligned}$$

□

Proposition 3 The given definition is equivalent to saying that the orthogonal projection of \mathbf{a} onto S (denoted \mathbf{g}) then $(\mathbf{g} - \mathbf{a})$ is orthogonal to every vector in S .

Proof. (\Rightarrow) Fixing an arbitrary $\mathbf{v} \in S$ we know that \mathbf{v} can be written as $\sum_{i=1}^k a_i \mathbf{u}_i$ for each \mathbf{u}_i being a basis vector of S . Therefore:

$$(\mathbf{g} - \mathbf{a}) \cdot \mathbf{v} = (\mathbf{g} - \mathbf{a}) \cdot \sum_{i=1}^k a_i \mathbf{u}_i = \sum_{i=1}^k (\mathbf{g} - \mathbf{a}) \cdot a_i \mathbf{u}_i = 0$$

(\Leftarrow) This is trivial because if $(\mathbf{g} - \mathbf{a})$ is orthogonal to all vectors in S then in particular it is perpendicular to basis vectors of S . \square

Given the results of propositions 1 and 2, orthogonal projections are independant of choice of basis vectors and are unique. Therefore, the following notation will be introduced for \mathbf{a} being \mathbf{g} 's orthogonal projection onto S :

$$P_S(\mathbf{g}) = \mathbf{a}$$

Theorem 1 If a point $\mathbf{a} \in S$ minimizes the distance between \mathbf{g} and the space S then $P_S(\mathbf{g}) = \mathbf{a}$

Proof. We can represent vectors in S as parameterized in k variables where k is the dimension of the basis of S . Therefore distances between S and \mathbf{g} can be written as:

$$d(t_1, t_2, \dots, t_k) = ||t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k - \mathbf{g}||$$

Since d is always positive it suffices to consider d^2 when talking about local extrema. Which can be written as:

$$(t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k - \mathbf{g}) \cdot (t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k - \mathbf{g})$$

Since the product rule holds for the dot product or inner product we can write the derivate of d^2 with respect to the variable t_i as:

$$2\mathbf{u}_i \cdot (t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k - \mathbf{g})$$

To have a minimum we must have that:

$$\forall i \leq k \quad \frac{\partial d^2}{\partial t_i} = 0$$

$$\begin{aligned} \forall i \leq k \quad 2\mathbf{u}_i \cdot (t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k - \mathbf{g}) &= 0 \\ \iff \forall i \leq k \quad \mathbf{u}_i \cdot (t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k - \mathbf{g}) &= 0 \end{aligned}$$

Therefore \mathbf{a} satisfies this then it is \mathbf{g} 's orthogonal projection by my original definition. \square

Notice that if we allow points in S to be written as linear combinations of S 's basis \mathbf{u}_i we can calculate the projection point by solving the system of equations:

$$\begin{aligned} \mathbf{u}_1 \cdot (t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k - \mathbf{a}) &= 0 \\ &\dots \\ \mathbf{u}_k \cdot (t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 + \dots + t_k \mathbf{u}_k - \mathbf{a}) &= 0 \end{aligned}$$

Which is equivalent to:

$$\begin{aligned} \mathbf{u}_1 \cdot t_1 \mathbf{u}_1 + \mathbf{u}_1 \cdot t_2 \mathbf{u}_2 + \dots + \mathbf{u}_1 \cdot t_k \mathbf{u}_k &= \mathbf{u}_1 \cdot \mathbf{a} \\ &\dots \end{aligned}$$

$$\mathbf{u}_k \cdot t_1 \mathbf{u}_1 + \mathbf{u}_k \cdot t_2 \mathbf{u}_2 + \cdots + \mathbf{u}_k \cdot t_k \mathbf{u}_k = \mathbf{u}_k \cdot \mathbf{a}$$

Which is again equivalent to solving the matrix equation:

$$\begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_k \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{u}_k \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{u}_k \cdot \mathbf{u}_1 & \mathbf{u}_k \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_k \cdot \mathbf{u}_k \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ \cdots \\ t_k \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{a} \\ \mathbf{u}_2 \cdot \mathbf{a} \\ \cdots \\ \mathbf{u}_k \cdot \mathbf{a} \end{bmatrix}$$

Taking for granted that this matrix is invertable we can calculate $P_S(\mathbf{g})$'s coordinates with respect to the chosen basis as:

$$\begin{bmatrix} t_1 \\ t_2 \\ \cdots \\ t_k \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_k \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{u}_k \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{u}_k \cdot \mathbf{u}_1 & \mathbf{u}_k \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_k \cdot \mathbf{u}_k \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{a} \\ \mathbf{u}_2 \cdot \mathbf{a} \\ \cdots \\ \mathbf{u}_k \cdot \mathbf{a} \end{bmatrix}$$

This is the method employed by my program to calculate projections!