Orthogonal Projection

Big Idea. The point in a subspace $U \subseteq \mathbb{R}^n$ nearest to $\boldsymbol{x} \in \mathbb{R}^n$ is the projection $\operatorname{proj}_U(\boldsymbol{x})$ of $\boldsymbol{x}$ onto U

Projection onto a Vector

DEFINITION

The projection of a vector $\boldsymbol{x}$ onto a vector $\boldsymbol{u}$ is

$$\operatorname{proj}_{\boldsymbol{u}}(\boldsymbol{x}) = \frac{\langle \boldsymbol{x} | \boldsymbol{u} \rangle}{\langle \boldsymbol{u} | \boldsymbol{u} \rangle} \boldsymbol{u}$$

Projection onto $\boldsymbol{u}$ is given by matrix multiplication

$$\operatorname{proj}_{\boldsymbol{u}}(x) = P\boldsymbol{x} \text{ where } P = \frac{1}{\|\boldsymbol{u}\|^2} \boldsymbol{u}\boldsymbol{u}^T$$

Note that $P^2 = P, P^T$ and $\operatorname{rank}(P) = 1$

Orthogonal Bases

DEFINITION

Let $U \subseteq \mathbb{R}^n$ be a subspace.
A set of vectors $\sum_m \boldsymbol{w}_m$ is an orthogonal basis for U if it is a basis for U and the vectors in the set are orthogonal
Furthermore, if each $\boldsymbol{w}_j$ is a unit vector, $\|\boldsymbol{w}_j\| = 1$ then the set is an orthonormal basis

THEOREM

Let $\sum_m \boldsymbol{u}_m$ be a basis of the subspace $U \subseteq \mathbb{R}^n$ The **Gram-Schmidt orthogonalization algorithm** constructs an orthogonal basis of U:

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{u}_1 \\
\boldsymbol{v}_2 &= \boldsymbol{u}_2 - \operatorname{proj}_{\boldsymbol{v}_1}(\boldsymbol{u}_2) \\
\boldsymbol{v}_3 &= \boldsymbol{u}_3 - \operatorname{proj}_{\boldsymbol{v}_1}(\boldsymbol{u}_3) - \operatorname{proj}_{\boldsymbol{v}_2}(\boldsymbol{u}_3) \\
&\vdots \\
\boldsymbol{v}_m &= \boldsymbol{u}_m = \operatorname{proj}_{\boldsymbol{v}_1}(\boldsymbol{u}_m) - \operatorname{proj}_{\boldsymbol{v}_2}(\boldsymbol{u}_m) - \cdots - \operatorname{proj}_{\boldsymbol{v}_{m-1}}(\boldsymbol{u}_m)
\end{aligned}$$

Then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ is an orthogonal basis of U. Furthermore let

$$\boldsymbol{w}_k = \frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}, \quad k = 1, \ldots, m$$

Then $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ is an orthonormal basis of U

EXAMPLE

Construct an orthonormal basis of the subspace U spanned by

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \boldsymbol{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Compute the orthogonal basis of U $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$
$\boldsymbol{v}_1 = \boldsymbol{u}_1$
$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \mathrm{proj}_{\boldsymbol{v}_1}(\boldsymbol{u}_2)$
$\boldsymbol{v}_3 = \boldsymbol{u}_3 - \mathrm{proj}_{\boldsymbol{v}_1}(\boldsymbol{u}_3) - \mathrm{proj}_{\boldsymbol{v}_2}(\boldsymbol{u}_3)$
We find an orthogonal basis

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \boldsymbol{v}_3 = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

And an orthonormal basis

$$\boldsymbol{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \boldsymbol{w}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \boldsymbol{w}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

Projection onto a subspace

DEFINITION

Let $U \subseteq \mathbb{R}^n$ be a subspace and let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ be an orthogonal basis of U. The projection of a vector $\boldsymbol{x}$ onto U is

$$\mathrm{proj}_U(\boldsymbol{x}) = \frac{\langle \boldsymbol{x} | \boldsymbol{u}_1 \rangle}{\langle \boldsymbol{u}_1 | \boldsymbol{u}_1 \rangle} \boldsymbol{u}_1 + \cdots + \frac{\langle \boldsymbol{x} | \boldsymbol{u}_m \rangle}{\langle \boldsymbol{u}_m | \boldsymbol{u}_m \rangle} \boldsymbol{u}_m$$

Projection onto U is given by matrix multiplication

$$\mathrm{proj}_U(\boldsymbol{x}) = P\boldsymbol{x} \text{ where } P = \frac{1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1\boldsymbol{u}_1^T + \cdots + \frac{1}{\|\boldsymbol{u}_m^2\|}\boldsymbol{u}_m\boldsymbol{u}_m^T$$

Note that $P^2 = P, P^T = P$ and $\mathrm{rank(P)=m}$

DEFINITION

A matrix P is an orthogonal projector (or orthogonal projection matrix) if $P^2 = P$ and $P^T = P$

THEOREM

Let P be the orthogonal projection onto U. Then $I - P$ is the orthogonal projection matrix onto $U^\perp$

EXAMPLE

Find the orthogonal matrix P which projects onto the subspace spanned by the vectors

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Compute $\langle \boldsymbol{u}_1 | \boldsymbol{u}_2 \rangle = 0$ therefore the vectors are orthogonal. Compute

$$\begin{aligned} P &= \frac{1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1\boldsymbol{u}_1^T + \frac{1}{\|\boldsymbol{u}_2\|^2}\boldsymbol{u}_2\boldsymbol{u}_2^T \\ &= \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\ &= \frac{1}{6}\begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix} \end{aligned}$$

EXAMPLE

Find the orthogonal projection matrix $P_\perp$ which projects onto $U_\perp$ where U is the subspace spanned by the vectors

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

In the last example we found P so $P_\perp = I - P$

$$P_\perp = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

Note that

$$\boldsymbol{u}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

is orthogonal to $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ and is a basis of the orthogonal complement $U^\perp$. Therefore we could also compute

$$P_\perp = \frac{1}{\|\boldsymbol{u}_3\|^2}\boldsymbol{u}_3\boldsymbol{u}_3^T$$

Projection theorem

THEOREM

Let $U \subseteq \mathbb{R}^n$ be a subspace and let $\boldsymbol{x} \in \mathbb{R}^n$. Then

$$\boldsymbol{x} - \mathrm{proj}_U(\boldsymbol{x}) \in U^\perp$$

and $\mathrm{proj}_U(\boldsymbol{x})$ is the closest vector in U to $\boldsymbol{x}$ in the sense that

$$\|\boldsymbol{x} - \mathrm{proj}_U(\boldsymbol{x})\| < \|\boldsymbol{x} - \boldsymbol{y}\| \quad \forall\, \boldsymbol{y} \in U\,,\ \boldsymbol{y} \neq \mathrm{proj}_U(\boldsymbol{x})$$