April 22 Math 3260 sec. 52 Spring 2024

Section 6.3: Orthogonal Projections

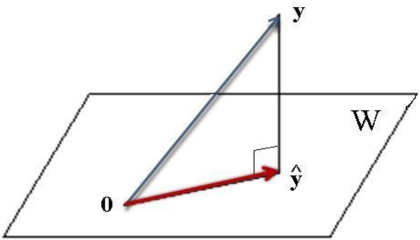

Figure: Given a subspace W of $\mathbb{R}^n$ and a vector $\mathbf{y}$ in $\mathbb{R}^n$, we can express $\mathbf{y}$ uniquely as a sum $\text{proj}_W \mathbf{y} + \mathbf{z}$ where $\text{proj}_W \mathbf{y}$ is in W and $\mathbf{z}$ is in $W^\perp$. Note that $\text{proj}_W \mathbf{y}$ is the point in W *closest* to $\mathbf{y}$ and $\|\mathbf{z}\|$ is the *distance* between $\mathbf{y}$ and W.

Orthogonal Decomposition Theorem

Let W be a subspace of $\mathbb{R}^n$. Each vector $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely as a sum

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$.

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is **any orthogonal basis** for W, then

$$\hat{\mathbf{y}} = \sum_{j=1}^{p} \left(\frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \right) \mathbf{u}_j, \quad \text{and} \quad \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Remark: The formula for $\hat{\mathbf{y}}$ is the sum of the projections of $\mathbf{y}$ onto each line $\mathrm{Span}\{\mathbf{u}_j\}$.

$$\mathrm{proj}_W \mathbf{y} = \mathrm{proj}_{\mathbf{u}_1} \mathbf{y} + \mathrm{proj}_{\mathbf{u}_2} \mathbf{y} + \cdots + \mathrm{proj}_{\mathbf{u}_p} \mathbf{y}.$$

Example

Let $\mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

We verified that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is orthogonal basis for W, and we found that

$$\text{proj}_W\, \mathbf{y} = \left(\frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right)\mathbf{u}_2 = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}.$$

The orthogonal part was

$$\mathbf{y} - \text{proj}_W\, \mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix}.$$

The distance between $\mathbf{y}$ and W was found to be

$$\text{dist}(W, \mathbf{y}) = \|\mathbf{y} - \text{proj}_W\, \mathbf{y}\| = \sqrt{2^2 + 4^2 + (-4)^2} = 6.$$

Computing Orthogonal Projections

Theorem

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis of a subspace W of $\mathbb{R}^n$, and $\mathbf{y}$ is any vector in $\mathbb{R}^n$ then

$$\mathrm{proj}_W\, \mathbf{y} = \sum_{j=1}^{p} (\mathbf{y} \cdot \mathbf{u}_j)\, \mathbf{u}_j.$$

And, if U is the matrix $U = [\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_p]$, then the above is equivalent to

$$\mathrm{proj}_W\, \mathbf{y} = UU^T\mathbf{y}.$$

Remark: In general, U is not square; it's $n \times p$. So even though UU^T will be a square matrix, it is not the same matrix as U^TU and it is not the identity matrix.

Example

Let $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ and $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Find an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ for W. Then compute the matrices $U^T U$ and UU^T where $U = [\mathbf{u}_1 \;\; \mathbf{u}_2]$.

Set $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1$ and $\vec{u}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2$

$\|\vec{v}_1\|^2 = 2^2 + 1^2 + 2^2 = 9 \quad \Rightarrow \quad \|\vec{v}_1\| = \|\vec{v}_2\| = 3$

$\|\vec{v}_2\|^2 = (-2)^2 + 2^2 + 1^2 = 9$

$\vec{u}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, \quad \vec{u}_2 = \frac{1}{3}\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$

$$U = \begin{bmatrix} 2/3 & -2/3 \\ 1/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & -2 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}$$

$$\underbrace{\underset{2\times 3}{U^T}\ \underset{3\times 2}{U}}_{2\times 2} = \frac{1}{3}\cdot\frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}\cdot\begin{bmatrix} 2 & -2 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\underbrace{\underset{3\times 2}{U}\ \underset{2\times 3}{U^T}}_{3\times 3} = \frac{1}{3}\cdot\frac{1}{3}\begin{bmatrix} 2 & -2 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}$$

Note: $(uu^T)^T = (u^T)^T u^T = uu^T$

Such matrices are called symmetric.

* $(AB)^T = B^T A^T$.

Example

$$W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$$

Use the matrix formulation to find $\text{proj}_W \, \mathbf{y}$.

$$\text{proj}_W \vec{y} = UU^T \vec{y}$$

$$= \frac{1}{9} \begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$$

$$= \frac{1}{9} \begin{bmatrix} 8(4) - 2(8) + 2(1) \\ -2(4) + 5(8) + 4(1) \\ 2(4) + 4(8) + 5(1) \end{bmatrix}$$

$$= \frac{1}{9}\begin{bmatrix} 18 \\ 36 \\ 45 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}.$$

same as before.

Section 6.4: Gram-Schmidt Orthogonalization

Big Question:

Given any-old basis for a subspace W of $\mathbb{R}^n$, can we construct an orthogonal basis for that same space?

Example: Let $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2\} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix} \right\}$. Find an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ that spans W.

Since $\vec{v}_1$ and $\vec{v}_2$ are in W, we can write them in terms of the basis.

$\vec{v}_1 = a_1 \vec{x}_1 + a_2 \vec{x}_2$

$\vec{v}_2 = b_1 \vec{x}_1 + b_2 \vec{x}_2$

Let's set $a_1 = 1$ and $a_2 = 0$

Let's set $b_2 = 1$.

So far, $\vec{v}_1 = \vec{x}_1$ and $\vec{v}_2 = b_1 \vec{x}_1 + \vec{x}_2$

Now, we need $\vec{v}_1 \cdot \vec{v}_2 = 0$

$$\vec{v}_1 \cdot \vec{v}_2 = \vec{x}_1 \cdot (b_1 \vec{x}_1 + \vec{x}_2) = 0$$

$$= b_1 \vec{x}_1 \cdot \vec{x}_1 + \vec{x}_1 \cdot \vec{x}_2 = 0$$

$$\Rightarrow b_1 \vec{x}_1 \cdot \vec{x}_1 = -\vec{x}_1 \cdot \vec{x}_2$$

$$\Rightarrow b_1 = -\frac{\vec{x}_1 \cdot \vec{x}_2}{\vec{x}_1 \cdot \vec{x}_1} = \frac{-\vec{v}_1 \cdot \vec{x}_2}{\vec{v}_1 \cdot \vec{v}_1}$$

$\vec{v}_1 = \vec{x}_1$

$\vec{v}_2 = \vec{x}_2 - \frac{\vec{v}_1 \cdot \vec{x}_2}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 \leftarrow \mathrm{proj}_{\vec{v}_1} \vec{x}_2$

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix}$$

$\vec{v}_1 \cdot \vec{x}_2 = -2$

$\vec{v}_1 \cdot \vec{v}_1 = 3$

$$\vec{v}_1 = \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix} - \frac{-2}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ -1/3 \\ -1/3 \end{bmatrix}$$

The new, orthogonal basis is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2/3 \\ -1/3 \\ -1/3 \end{bmatrix} \right\}$$

Theorem: Gram Schmidt Process

Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ be any basis for the nonzero subspace W of $\mathbb{R}^n$. Define the set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ via

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \left(\frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \left(\frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \sum_{j=1}^{p-1}\left(\frac{\mathbf{x}_p \cdot \mathbf{v}_j}{\mathbf{v}_j \cdot \mathbf{v}_j}\right)\mathbf{v}_j.
\end{aligned}$$

$\mathrm{Span}\{\vec{v}_1, \vec{v}_2\} = \mathrm{Span}\{\vec{x}_1, \vec{x}_2\}$

$\mathrm{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \mathrm{Span}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for W. Moreover, for each $k = 1, \ldots, p$

$$\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \mathrm{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}.$$