§6.3 Orthogonal Projection

Thm: (The Orthogonal Decomposition Theorem)

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

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

where $\hat{\vec{y}}$ is in W and $\vec{z}$ is in $W^{\perp}$. In fact, if $\{\vec{u}_1, \cdots, \vec{u}_p\}$ is any orthogonal basis of W, then

$$\hat{y} = \frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\vec{u}_1 + \cdots + \frac{\vec{y}\cdot\vec{u}_p}{\vec{u}_p\cdot\vec{u}_p}\vec{u}_p$$

and $\vec{z} = \vec{y} - \hat{\vec{y}}$.

$\hat{\vec{y}}$ is called orthogonal projection of $\vec{y}$ onto W, denoted by $\text{proj}_W \vec{y}$

* Properties of Orthogonal Projections

Let $W = \text{span}\{\vec{u}_1, \cdots, \vec{u}_p\}$, where $\{\vec{u}_1, \cdots, \vec{u}_p\}$ is an orthogonal basis for W.

If $\vec{y} \in W$, then $\text{Proj}_W \vec{y} = \vec{y}$.

Thm: The Best Approximation Theorem

Let W be a subspace of $\mathbb{R}^n$, let $\vec{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{y}$ be the orthogonal projection of $\vec{y}$ onto W. Then $\hat{y}$ is the closest point in W to $\vec{y}$, in the sense that

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

for all $\vec{v}$ in W distinct from $\hat{\vec{y}}$.

Remark: $\hat{\vec{y}}$ is called the best approximation to $\vec{y}$ by elements of W.

Proof: $\vec{y} - \vec{v} = \vec{y} - \hat{\vec{y}} + \hat{\vec{y}} - \vec{v}$

Note that $\vec{y} - \hat{\vec{y}} \in W^{\perp}$ and $\hat{\vec{y}} - \vec{v} \in W$.

Hence $\vec{y} - \hat{\vec{y}} \perp \hat{\vec{y}} - \vec{v}$

By Pythagorean Theorem, we have

$$\|\vec{y} - \vec{v}\|^2 = \|\vec{y} - \hat{\vec{y}}\|_2 + \|\hat{\vec{y}} - \vec{v}\|$$

Hence $\|\vec{y} - \hat{\vec{y}}\|_2 < \|\vec{y} - \vec{v}\|_2$

Def: The distance from a point $\vec{y}$ in $\mathbb{R}^n$ to a subspace W is defined to be the distance from $\vec{y}$ to the nearest point in W.

Ex 1: Find the distance from $\vec{y}$ to $W = \text{Span}\{\vec{u}_1, \vec{u}_2\}$, where

$$\vec{y} = \begin{pmatrix} -1 \\ -5 \\ 10 \end{pmatrix}, \quad \vec{u}_1 = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Solution: By the Best Approximation Theorem, the distance from $\vec{y}$ to W is $\|\vec{y} - \hat{\vec{y}}\|$, where $\hat{\vec{y}} = \text{proj}_W \vec{y}$.

Since $\{\vec{u}_1, \vec{u}_2\}$ is an orthogonal basis for W,

$$\hat{\vec{y}} = \frac{\langle \vec{y}, \vec{u}_1 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle} \vec{u}_1 + \frac{\langle \vec{y}, \vec{u}_2 \rangle}{\langle \vec{u}_2, \vec{u}_2 \rangle} \vec{u}_2$$

$$= \frac{-1 \cdot 5 + (-5) \cdot (-2) + 10 \cdot 1}{5^2 + (-2)^2 + 1^2} \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix} + \frac{-1 \cdot 1 + (-5) \cdot 2 + 10 \cdot (-1)}{1^2 + 2^2 + (-1)^2} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

$$= \frac{15}{30}\begin{pmatrix}5\\-2\\1\end{pmatrix} + \frac{-21}{6}\begin{pmatrix}1\\2\\-1\end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix}5\\-2\\1\end{pmatrix} - \frac{7}{2}\begin{pmatrix}1\\2\\-1\end{pmatrix}$$

$$= \begin{pmatrix}-1\\-8\\4\end{pmatrix}$$

Hence $\vec{y} - \hat{\vec{y}} = \begin{pmatrix}-1\\-5\\10\end{pmatrix} - \begin{pmatrix}-1\\-8\\4\end{pmatrix} = \begin{pmatrix}0\\3\\6\end{pmatrix}$ and

$$\|\vec{y} - \hat{\vec{y}}\| = \sqrt{3^2+6^2} = \sqrt{45} = 3\sqrt{5}$$

The distance from $\vec{y}$ to W is $3\sqrt{5}$.

Thm: If $\{\vec{u}_1, \cdots, \vec{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W \vec{y} = (\vec{y}\cdot\vec{u}_1)\vec{u}_1 + (\vec{y}\cdot\vec{u}_2)\vec{u}_2 + \cdots + (\vec{y}\cdot\vec{u}_p)\vec{u}_p$$

If $U = [\vec{u}_1 \ \vec{u}_2 \ \cdots \ \vec{u}_p]$, then

$$\text{proj}_W \vec{y} = UU^T\vec{y} \quad \text{for all } \vec{y} \text{ in } \mathbb{R}^n.$$

Proof: This Theorem follows directly from The Orthogonal Decomposition Theorem.

Ex: Let W be a subspace of $\mathbb{R}^n$. Let $\vec{x}$ and $\vec{y}$ be vectors in $\mathbb{R}^n$ and let $\vec{z} = \vec{x} + \vec{y}$. If $\vec{u}$ is the projection of

$\vec{x}$ onto W and $\vec{v}$ is the projection of $\vec{y}$ onto W, show that $\vec{u}+\vec{v}$ is the projection of $\vec{z}$ onto W.

Solution: Let U be a matrix whose columns consist of an orthonormal basis for W. Then

$$\text{proj}_W \vec{z} = UU^T\vec{z} = UU^T(\vec{x}+\vec{y})$$
$$= UU^T\vec{x} + UU^T\vec{y}$$
$$= \text{proj}_W \vec{x} + \text{proj}_W \vec{y}$$
$$= \vec{u} + \vec{v}$$

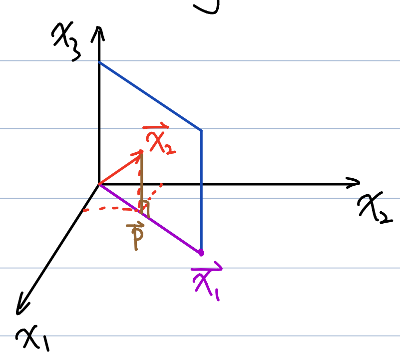
§6.4 The Gram-Schmidt Process

The Gram-Schmidt process is a simple algorithm for producing an orthogonal or orthogonal basis for any nonzero subspace of $\mathbb{R}^n$.

Example 1: Let $W = \text{Span}\{\vec{x}_1, \vec{x}_2\}$, where $\vec{x}_1 = \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}$ and $\vec{x}_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$.

Construct an orthogonal basis $\{\vec{v}_1, \vec{v}_2\}$ for W.

Solution:



Let $\vec{p}$ be the projection of $\vec{x}_2$ onto $\vec{x}_1$. The component of $\vec{x}_2$ orthogonal to $\vec{x}_1$ is $\vec{x}_2 - \vec{p}$.

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

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

$$= \begin{pmatrix}1\\2\\2\end{pmatrix} - \frac{1\cdot 3 + 6\cdot 2 + 0\cdot 2}{3^2+6^2+0^2}\begin{pmatrix}3\\6\\0\end{pmatrix}$$

$$= \begin{pmatrix}1\\2\\2\end{pmatrix} - \frac{15}{45}\begin{pmatrix}3\\6\\0\end{pmatrix} = \begin{pmatrix}0\\0\\2\end{pmatrix}$$

Then $\{\vec{v}_1, \vec{v}_2\}$ is an orthogonal set of nonzero vectors in W. Since $\dim W = 2$, the set $\{\vec{v}_1, \vec{v}_2\}$ is a basis for W.

Thm: The Gram-Schmidt Process

Given a basis $\{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

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

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

$$\vec{v}_3 = \vec{x}_3 - \frac{\vec{x}_3\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\cdot\vec{v}_1 - \frac{\vec{x}_3\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2}\vec{v}_2$$

$$\vdots$$

$$\vec{v}_p = \vec{x}_p - \frac{\vec{x}_p\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\cdot\vec{v}_1 - \cdots - \frac{\vec{x}_p\cdot\vec{v}_{p-1}}{\vec{v}_{p-1}\cdot\vec{v}_{p-1}}\vec{v}_{p-1}$$

Then $\{\vec{v}_1, \cdots, \vec{v}_p\}$ is an orthogonal basis for W.

In addition

$$\text{Span}\{\vec{v}_1, \cdots, \vec{v}_k\} = \text{Span}\{\vec{x}_1, \cdots, \vec{x}_k\}, \quad 1\le k\le p.$$

Proof: For $1\le k\le p$, let $W_k = \text{Span}\{\vec{x}_1, \cdots, \vec{x}_k\}$.

Then $W_1 = \text{Span}\{\vec{x}_1\} = \text{Span}\{\vec{v}_1\}$.

$W_2 = \text{Span}\{\vec{x}_1, \vec{x}_2\} = \text{Span}\{\vec{v}_1, \vec{v}_2\}$ since

$\vec{v}_1 = \vec{x}_1$, $\vec{v}_2 = \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \cdot \vec{v}_1$ are in W_2 and orthogonal.

Suppose $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is orthogonal set, then $\vec{x}_{k+1}$ has an orthogonal decomposition

$$\vec{x}_{k+1} = \hat{\vec{x}}_{k+1} + (\vec{x}_{k+1} - \hat{\vec{x}}_{k+1})$$

where $\hat{\vec{x}}_{k+1} = \frac{\vec{v}_1 \cdot \vec{x}_{k+1}}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 + \cdots + \frac{\vec{v}_k \cdot \vec{x}_{k+1}}{\vec{v}_k \cdot \vec{v}_k} \vec{v}_k \in W_k$

and
$$\begin{aligned} \vec{v}_{k+1} &= \vec{x}_{k+1} - \hat{\vec{x}}_{k+1} \\ &= \vec{x}_{k+1} - \frac{\vec{v}_1 \cdot \vec{x}_{k+1}}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 - \cdots - \frac{\vec{v}_k \cdot \vec{x}_{k+1}}{\vec{v}_k \cdot \vec{v}_k} \vec{v}_k \end{aligned}$$

is orthogonal to W_k.

Clearly, $\{\vec{v}_1, \cdots, \vec{v}_{k+1}\}$ is an orthogonal set.

and $\vec{v}_1, \cdots, \vec{v}_{k+1} \in W_{k+1}$.

Therefore $W_{k+1} = \text{Span}\{\vec{v}_1, \cdots, \vec{v}_{k+1}\}$.

* Orthonormal Bases

An orthonormal basis is constructed easily from an orthogonal basis $\{\vec{v}_1, \cdots, \vec{v}_p\}$: normalize all $\vec{v}_k$.

Example 2: Let $\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, and $\vec{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$. Then

$\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ is clearly linearly independent and thus is a basis for a subspace W of R^4.

Construct an orthonormal basis for W.

Solution: $\vec{v}_1 = \vec{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

$$\vec{v}_1' = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{1^2+1^2+1^2+1^2}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

$$\vec{v}_2 = \vec{x}_2 - (\vec{x}_2 \cdot \vec{v}_1')\vec{v}_1' = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \left(0\cdot\frac{1}{2} + 1\cdot\frac{1}{2} + 1\cdot\frac{1}{2} + 1\cdot\frac{1}{2}\right) \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{3}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}$$

$$\vec{v}_2' = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \frac{1}{\sqrt{\left(-\frac{3}{4}\right)^2 + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2}} \begin{pmatrix} -\frac{3}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix} = \frac{2}{\sqrt{3}} \begin{pmatrix} -\frac{3}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix} = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \end{pmatrix}$$

$$\vec{v}_3 = \vec{x}_3 - (\vec{x}_3 \cdot \vec{v}_1')\vec{v}_1' - (\vec{x}_3 \cdot \vec{v}_2') \cdot \vec{v}_2'$$

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \left(0\cdot\frac{1}{2} + 0\cdot\frac{1}{2} + 1\cdot\frac{1}{2} + 1\cdot\frac{1}{2}\right) \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} - \left(0 + 0 + \frac{1}{2\sqrt{3}} + \frac{1}{2\sqrt{3}}\right) \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} - \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

$$\vec{v}_3' = \frac{\vec{v}_3}{\|\vec{v}_3\|} = \frac{1}{\sqrt{0^2 + (-\frac{2}{3})^2 + (\frac{1}{3})^2 + (\frac{1}{3})^2}} \begin{pmatrix} 0 \\ -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}$$

$\{\vec{v}_1', \vec{v}_2', \vec{v}_3'\}$ is an orthonormal basis for span$\{x_1, x_2, x_3\}$.

Thm (The QR Factorization)

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for $\mathrm{Col}\,A$ and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Proof: The columns of A form a basis $\{\vec{x}_1, \cdots, \vec{x}_n\}$ for $\mathrm{Col}\,A$. Construct an orthonormal basis $\{\vec{u}_1, \cdots, \vec{u}_n\}$ for $W = \mathrm{Col}\,A$ with property

$\text{Span}\{\vec{u}_1, \cdots, \vec{u}_k\} = \text{Span}\{\vec{x}_1, \cdots, \vec{x}_k\}$, $1 \le k \le n$.

Let $Q = [\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_n]$.

For $k = 1, \cdots, n$, $\vec{x}_k \in \text{Span}\{\vec{u}_1, \cdots, \vec{u}_k\}$. So there are constants $r_{1k}, \cdots, r_{kk}$ such that

$$\vec{x}_k = r_{1k}\vec{u}_1 + \cdots + r_{kk}\vec{u}_k + 0\cdot\vec{u}_{k+1} + \cdots + 0\cdot\vec{u}_n$$

We assume $r_{kk} > 0$. (If $r_{kk} < 0$, multiply both r_{kk} and $\vec{u}_k$ by -1).

This shows that $\vec{x}_k = Q\vec{r}_k$, where

$$\vec{r}_k = \begin{pmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad k = 1, \cdots, n.$$

Let $R = [\vec{r}_1, \cdots, \vec{r}_n]$. Then

$$A = [\vec{x}_1, \cdots, \vec{x}_n] = [Q\vec{r}_1, \cdots, Q\vec{r}_n] = QR$$

Since the columns of A are linearly independent, R is invertible. In fact, if R is not invertible, there exists nonzero vector $c = [c_1, \cdots, c_n]$ such that $Rc = \vec{0}$. Hence $Ac = QRc = \vec{0}$ which implies the columns of A are linearly dependent.

It contradict with the fact that A are linearly independent. Hence R is invertible.

Since R is clearly uptriangular, its nonnegative diagonal entries must be positive.

Example: Find a QR factorization of $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

Solution: The columns of A are the vectors $\vec{x}_1$, $\vec{x}_2$, $\vec{x}_3$ in Example 2. The orthonormal basis are

$$\vec{u}_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} 0 \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}$$

Let $Q = [\vec{u}_1, \vec{u}_2, \vec{u}_3] = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}$

Note that $Q^TQ = I$, we have

$$R = Q^TQR = Q^TA$$

$$= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & \frac{3}{2} & 1 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$$

Ex: Suppose $A = QR$, where Q is an $m \times n$ matrix with orthonormal columns and R is an $n \times n$ matrix. Show that if the columns of A are linearly dependent, then R cannot be invertible.

Solution: If the columns of A are linearly dependent, then there exists $c_1, \cdots, c_n$ not all equal to zero, such that $Ac = \vec{0}$, where $c = (c_1, \cdots, c_n)^T$.

Since Q is an $m \times n$ matrix with orthonormal columns, $Q^TQ = I$.

Hence $Rc = Q^TQRc = Q^TAc = \vec{0}$.

That is $R\vec{x} = \vec{0}$ has nontrivial solution which implies R is not invertible.