Orthonormal Basis

Definition

Orthonormal Basis: We say that a set of vectors $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ is an orthonormal basis of some subspace W of $\mathbb{R}^m$ if

1. $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ is an orthonormal set
2. $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ is a basis for W.

Theorem

Let $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ be an orthonormal basis of $\mathbb{R}^n$. Then, the $\mathcal{B}$ coordinates for a vector $\mathbf{x} \in \mathbb{R}^n$ are given by

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \mathbf{q}_1^\top \mathbf{x} & \ldots & \mathbf{q}_n^\top \mathbf{x} \end{bmatrix}$$

This means if we have a basis like this, then our lives are a lot easier!

Orthonormal Columns

Definition

Orthonormal Columns: We say that a matrix $Q \in \mathbb{R}^{m\times n}$ has orthonormal columns if $Q^\top Q = I$. (*Note*: We don't necessarily have $QQ^\top = I$!)

This is just putting an orthonormal set into columns of a matrix!

Matrices with Orthonormal Columns and our Norm

Theorem

If $Q \in \mathbb{R}^{m \times n}$ such that Q has orthonormal columns, then for any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$$

Proof.

Let $\mathbf{x} \in \mathbb{R}^n$ and $Q \in \mathbb{R}^{m \times n}$ have orthonormal columns. Then see that

$$\|Q\mathbf{x}\|_2 = \sqrt{\mathbf{x}^\top Q^\top Q\mathbf{x}} = \sqrt{\mathbf{x}^\top \mathbf{x}} = \|\mathbf{x}\|_2$$

$\square$

Gram-Schmidt Process (Slightly Different Than Our Text!)

Another problem we want to do is take some basis of our space and convert it to a basis of orthonormal vectors. One method is the Gram-Schmidt Process, which is given below

Definition

Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be a basis for a subspace W of $\mathbb{R}^n$. Then the Gram-Schmidt Process computes $\mathbf{q}_1, \ldots, \mathbf{q}_m$ such that

$$\text{Span}\left(\mathbf{v}_1, \ldots, \mathbf{v}_m\right) = \text{Span}\left(\mathbf{q}_1, \ldots, \mathbf{q}_m\right) \text{ and } \{\mathbf{q}_1, \ldots, \mathbf{q}_m\} \text{ is an orthonormal set.}$$

We compute the $\mathbf{q}_\ell$ vectors as follows:

$$\mathbf{u}_1 = \mathbf{v}_1 \qquad \mathbf{q}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|_2}$$

For $\ell = 2, \ldots, m$:

$$\mathbf{u}_\ell = \mathbf{v}_\ell - \langle \mathbf{v}_\ell, \mathbf{q}_1 \rangle \mathbf{q}_1 - \cdots - \langle \mathbf{v}_\ell, \mathbf{q}_{\ell-1} \rangle \mathbf{q}_{\ell-1} \qquad \mathbf{q}_\ell = \frac{\mathbf{u}_\ell}{\|\mathbf{u}_\ell\|_2}$$

Note: Here, we are using $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y}$

QR Decomposition (From Linear Algebra with Applications)

Theorem

Let $A \in \mathbb{R}^{m \times n}$ be written as $A = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}$. There exists exist some $Q \in \mathbb{R}^{m \times n}$, $R \in \mathbb{R}^{n \times n}$ such that $Q^\top Q = I$ and R is upper triangular, and these matrices are of the form

$$Q = \begin{bmatrix} \mathbf{q}_1 & \dots & \mathbf{q}_n \end{bmatrix} \qquad R = \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{q}_1 \rangle & \langle \mathbf{v}_2, \mathbf{q}_1 \rangle & \dots & \langle \mathbf{v}_n, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{v}_2, \mathbf{q}_2 \rangle & \dots & \langle \mathbf{v}_n, \mathbf{q}_2 \rangle \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \langle \mathbf{v}_n, \mathbf{q}_n \rangle \end{bmatrix}$$

Where each $\mathbf{q}_\ell$ is computed via the Gram-Schmidt process

We won't be proving this formula, but if you're interested in what this would look like, `https://en.wikipedia.org/wiki/QR_decomposition#Using_the_Gram%E2%80%93Schmidt_process` has a write-up of what that proof would look like.

Alternative Orthogonal Transformation

Theorem

Let W be a subspace of $\mathbb{R}^m$ such that $W = \mathrm{Col}(A)$ for some matrix $A \in \mathbb{R}^{m\times n}$. Then $\mathbf{x}_W = QQ^\top\mathbf{x}$ where $A = QR$ from the previous slide.

Proof.

Recall that we say $\mathbf{x}_W = A\mathbf{c}$ where $\mathbf{c}$ is a solution to $A^\top A\mathbf{c} = A^\top\mathbf{x}$.. So we will show that if we replace $A\mathbf{c}$ with $QQ^\top\mathbf{x}$ we also solve this equation!

$$A^\top A\mathbf{c} = A^\top QQ^\top\mathbf{x} = R^\top Q^\top QQ^\top\mathbf{x} = R^\top Q^\top\mathbf{x} = (QR)^\top\mathbf{x} = A^\top\mathbf{x}$$

$\square$

Matrix Associated With Orthogonal Projection

There are two ways to compute a matrix associated with our orthogonal projection T.

1. Using our normal equations and projecting the standard basis vectors
2. Form $QQ^\top$ where Q is from the QR decomposition

We will demonstrate both of these methods using $W = \text{Span}\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)$

Computing Matrix Associated With Orthogonal Projection Method 1

$$W = \text{Span}\left(\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right)$$

We need to compute $T(\mathbf{e}_1), T(\mathbf{e}_2), T(\mathbf{e}_3)$. For convenience, we recall that $A^\top A = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$. We also have that

$$A^\top \mathbf{e}_1 = \begin{bmatrix}1\\1\end{bmatrix} \qquad A^\top \mathbf{e}_2 = \begin{bmatrix}0\\1\end{bmatrix} \qquad A^\top \mathbf{e}_3 = \begin{bmatrix}-1\\0\end{bmatrix}$$

We now solve our systems!

$$\left[\begin{array}{cc|c}2 & 1 & 1\\1 & 2 & 1\end{array}\right] \to \left[\begin{array}{cc|c}1 & 0 & \frac{1}{3}\\0 & 1 & \frac{1}{3}\end{array}\right] \qquad \left[\begin{array}{cc|c}2 & 1 & 0\\1 & 2 & 1\end{array}\right] \to \left[\begin{array}{cc|c}1 & 0 & -\frac{1}{3}\\0 & 1 & \frac{2}{3}\end{array}\right] \qquad \left[\begin{array}{cc|c}2 & 1 & -1\\1 & 2 & 0\end{array}\right] \to \left[\begin{array}{cc|c}1 & 0 & -\frac{2}{3}\\0 & 1 & \frac{1}{3}\end{array}\right]$$

So, we have our matrix

$$P = \frac{1}{3}\left[A\begin{bmatrix}1\\1\end{bmatrix} \quad A\begin{bmatrix}-1\\2\end{bmatrix} \quad A\begin{bmatrix}-2\\1\end{bmatrix}\right] = \frac{1}{3}\begin{bmatrix}2 & 1 & -1\\1 & 2 & 1\\-1 & 1 & 2\end{bmatrix}$$

Computing Matrix Associated With Orthogonal Projection Method 2 Part 1

$$W = \text{Span}\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)$$

Following Gram-Schmidt, we get that

$$\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \qquad \mathbf{q}_2 = \frac{\sqrt{2}}{2\sqrt{3}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

If we put these in the columns of Q and simplify we get that $Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{3}} \\ -1 & \frac{1}{\sqrt{3}} \end{bmatrix}$

Computing Matrix Associated With Orthogonal Projection Method 2 Part 2

Now, we compute $QQ^\top$!

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

$$= \frac{1}{2}\left(\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}\right) = \frac{1}{6}\begin{bmatrix} 4 & 2 & -2 \\ 2 & 4 & 2 \\ -2 & 2 & 4 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

Least-Squares Problem

What do we do if we can't solve the system below exactly?

$$A\mathbf{x} = \mathbf{b}$$

We could give up, but that's no fun! Instead, we want to get as close as possible. One way of saying this is pick an answer such that

$$\|A\mathbf{x} - \mathbf{b}\|_2$$

is as small as possible. This means we pick an $\hat{\mathbf{x}}$ such that $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to $\mathbf{b}$.
But wait, this is very similar to orthogonal projections!

The Normal Equations are a Least Squares Solution!

Theorem

Let $A \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m$. The solution to

$$A^\top A\mathbf{x} = A^\top \mathbf{b}$$

is an $\mathbf{x}$ such that

$$\|A\mathbf{x} - \mathbf{b}\|_2$$

is minimal.

QR Makes This Easier (For Computers)

Theorem

Let $A \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m$. The solution to

$$R\mathbf{x} = Q^\top \mathbf{b}$$

is an $\mathbf{x}$ such that

$$\|A\mathbf{x} - \mathbf{b}\|_2$$

is minimal where $Q \in \mathbb{R}^{m \times n}, R \in \mathbb{R}^{n \times n}$ is a QR decomposition of A.

Proof.

As we discussed previously, we can always solve $A\mathbf{x} = \mathbf{b}_{\mathrm{Col}}(A)$.So, we use the fact that $QQ^\top \mathbf{b}$ projects $\mathbf{b}$ onto the columnspace of A to get

$$A\hat{\mathbf{x}} = QQ^\top \mathbf{b} \to QR\hat{\mathbf{x}} = QQ^\top \mathbf{b} \to R\hat{\mathbf{x}} = Q^\top \mathbf{b}$$

□