GRAM-SCHMIDT ALGORITHM AND QR FACTORIZATION

1. Motivating Problem

In many situations we have a basis of a subspace, V, but want to find an orthonormal basis. This is useful, for example, in giving a formula for proj_V. The *Gram-Schmidt* algorithm allows us to convert any basis of V to an orthonormal basis. Strategy (for two dimensional V):

(1) Scale first vector to make it unit.
(2) Project second vector onto orthogonal complement of first to make it orthogonal to first.
(3) Scale second vector to make it unit.

EXAMPLE: Find an orthonormal basis of $V = \mathrm{span}\left(\begin{bmatrix}2\\2\\1\end{bmatrix}, \begin{bmatrix}2\\0\\2\end{bmatrix}\right)$. First, set

$$\vec{u}_1 = \frac{1}{||\vec{v}_1||}\vec{v}_1 = \frac{1}{3}\begin{bmatrix}2\\2\\1\end{bmatrix} = \begin{bmatrix}2/3\\2/3\\1/3\end{bmatrix}.$$

Next, project onto orthogonal space to $\vec{u}_1$

$$\vec{v}_2^{\perp} = \vec{v}_2 - \vec{v}_2^{||} = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1)\vec{u}_1 = \begin{bmatrix}2\\0\\2\end{bmatrix} - 2\begin{bmatrix}2/3\\2/3\\1/3\end{bmatrix} = \begin{bmatrix}2/3\\-4/3\\4/3\end{bmatrix}.$$

Finally, normalize the length of this vector:

$$\vec{u}_2 = \frac{1}{||\vec{v}_2^{\perp}||}\vec{v}_2^{\perp} = \frac{1}{2}\begin{bmatrix}2/3\\-4/3\\4/3\end{bmatrix} = \begin{bmatrix}1/3\\-2/3\\2/3\end{bmatrix}$$

So $\vec{u}_1, \vec{u}_2$ is an orthonormal basis of V. Observe that

$$\vec{v}_1 = 3\vec{u}_1 \text{ and } \vec{v}_2 = 2\vec{u}_1 + 2\vec{u}_2$$

That is,

$$S_{\mathcal{B}\to\mathcal{U}} = \begin{bmatrix}3 & 2\\0 & 2\end{bmatrix}$$

where

$$\mathcal{B} = (\vec{v}_1, \vec{v}_2) \text{ and } \mathcal{U} = (\vec{u}_1, \vec{u}_2)$$

In other words,

$$\begin{bmatrix}\vec{v}_1 & | & \vec{v}_2\end{bmatrix} = \begin{bmatrix}\vec{u}_1 & | & \vec{u}_2\end{bmatrix}\begin{bmatrix}3 & 2\\0 & 2\end{bmatrix}.$$

2. Gram-Schmidt algorithm

We generalize the above procedure so that it holds any basis. Key Idea: Turn a basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_m)$ of $V \subset \mathbb{R}^n$ into a orthonormal basis $\mathcal{U} = (\vec{u}_1, \ldots, \vec{u}_m)$ of V. In order to write this most compactly, we observe that when $V = \left\{\vec{0}\right\}$, then for any $\vec{x}$,

$$\mathrm{proj}_V(\vec{x}) = \vec{0}.$$

The algorithm is:

(1) Start $i = 1$
(2) (Project) Let
$$\vec{v}_i^{\perp} := \vec{v}_i - \mathrm{proj}_{V_i}(\vec{v}_i) = \vec{v}_i - \sum_{j=1}^{i-1} (\vec{u}_j \cdot \vec{v}_i)\vec{u}_j$$
where $V_i = \mathrm{span}(\vec{v}_1, \ldots, \vec{v}_{i-1}) = \mathrm{span}(\vec{v}_1, \ldots, \vec{v}_{i-1})$.
(3) (Scale) Let $\vec{u}_i = \frac{1}{||\vec{v}_i^{\perp}||}\vec{v}_i^{\perp}$
(4) If $i < m$, then increment i and start again at (2). Otherwise, we are done.

In the above, by construction, $V_1 = \left\{\vec{0}\right\}$ so the first projection does nothing.

Some comments:

(1) One is constructing the orthonormal basis, $\vec{u}_1, \ldots, \vec{u}_{i-1}$, of V_i from the previous steps in the algorithm. This means one can compute the orthogonal projection in a straightforward manner.
(2) The fact that $\mathcal{B}$ is a basis is what ensures that $||\vec{v}_i^{\perp}|| > 0$.

EXAMPLE: Compute an orthonormal basis of

$$\ker \begin{bmatrix} 1 & 3 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

This matrix is in rref already so we have

$$\ker = \mathrm{span}\left(\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right)$$

Lets take a basis to be

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix} \text{ and } \vec{v}_2 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

We apply the Gram-Schmidt algorithm to this to obtain:

$$\vec{u}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

and

$$\vec{v}_2^{\perp} = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1)\vec{u}_1 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix} = \frac{2}{3}\begin{bmatrix} -5 \\ 3 \\ -4 \\ 2 \end{bmatrix}$$

and

$$||\vec{v}_2^{\perp}|| = \frac{2}{3}\sqrt{(-5)^2 + 3^2 + (-4)^2 + 2^2} = \frac{2}{3}\sqrt{54} = 2\sqrt{6}.$$

Hence,

$$\vec{u}_2 = \frac{1}{3\sqrt{6}} \begin{bmatrix} -5 \\ 3 \\ -4 \\ 2 \end{bmatrix} = \frac{\sqrt{6}}{18} \begin{bmatrix} -5 \\ 3 \\ -4 \\ 2 \end{bmatrix}.$$

3. QR factorization

Let $\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_m)$ be a basis of $V \subset \mathbb{R}^n$. Suppose that $\mathcal{U} = (\vec{u}_1, \dots, \vec{u}_m)$ is the basis obtained from $\mathcal{B}$ by the Gram-Schmidt algorithm. If we write

$$R = S_{\mathcal{B} \to \mathcal{U}}, M = \begin{bmatrix} \vec{v}_1 & | & \cdots & | & \vec{v}_m \end{bmatrix} \text{ and } Q = \begin{bmatrix} \vec{u}_1 & | & \cdots & | & \vec{u}_m \end{bmatrix},$$

then we have

$$M = QR$$

If one looks at how the Gram-Schmidt algorithm works, it is not hard to see that R is upper triangular and the diagonal elements are all positive. Such a factorization is called a QR-factorization and is useful in many computational settings. Notice, that by definition, if M has a QR-factorization $M = QR$, then the columns of Q form an orthonormal basis of $\mathrm{Im}\,(M)$.

EXAMPLE: Compute the QR factorization of

$$M = \begin{bmatrix} 2 & 0 \\ 1 & 6 \\ -2 & -6 \end{bmatrix}$$

Have

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \text{ and } \vec{v}_2 = \begin{bmatrix} 0 \\ 6 \\ -6 \end{bmatrix}.$$

As such,

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

and

$$\vec{v}_2^{\perp} = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1)\vec{u}_1 = \begin{bmatrix} 0 \\ 6 \\ -6 \end{bmatrix} - 6 \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \\ -2 \end{bmatrix}$$

and

$$\vec{u}_2 = \frac{1}{6} \begin{bmatrix} -4 \\ 4 \\ -2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 2/3 \\ -1/3 \end{bmatrix}.$$

Clearly, $\vec{v}_1 = 3\vec{u}_1$ and $\vec{v}_2 = 6\vec{u}_1 + 6\vec{u}_2$ and so

$$\underbrace{\begin{bmatrix} 2 & 0 \\ 1 & 6 \\ -2 & -6 \end{bmatrix}}_{M} = \underbrace{\begin{bmatrix} 2/3 & -2/3 \\ 1/3 & 2/3 \\ -2/3 & -1/3 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 3 & 6 \\ 0 & 6 \end{bmatrix}}_{R}$$