Gram-Schmidt Process

Gram-Schmidt is an algorithm that takes a set of linearly independent vectors $\{\vec{s}_1, \ldots, \vec{s}_n\}$ and generates an orthonormal set of vectors $\{\vec{q}_1, \ldots, \vec{q}_n\}$ that span the same vector space as the original set. Concretely, $\{\vec{q}_1, \ldots, \vec{q}_n\}$ satisfy the following:

- $\forall 0 < k \leq n$, span($\{\vec{s}_1, \ldots, \vec{s}_k\}$) = span($\{\vec{q}_1, \ldots, \vec{q}_k\}$)
- $\{\vec{q}_1, \ldots, \vec{q}_n\}$ is an orthonormal set of vectors

Definition: *Orthonormal*

A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is **orthonormal** if all the vectors are mutually orthogonal to each other and all are of unit length. That is:

- Orthogonal: For all pairs of vectors $\vec{v}_i, \vec{v}_j$ where $i \neq j$, $\langle \vec{v}_i, \vec{v}_j \rangle = 0$. For real vectors, this means $\vec{v}_i^T \vec{v}_j = 0$.
- Normalized: For all i, $\|\vec{v}_i\| = 1$. (This implies that $\|\vec{v}_i\| = \langle \vec{v}_i, \vec{v}_i \rangle = 1$.)

The Gram-Schmidt algorithm works by first finding the unit vector $\vec{q}_1$ such that span($\{\vec{q}_1\}$) = span($\{\vec{s}_1\}$). Subsequently, the unit vector $\vec{q}_2$ is calculated such that $\langle \vec{q}_1, \vec{q}_2 \rangle = 0$ and span($\{\vec{q}_1, \vec{q}_2\}$) = span($\{\vec{s}_1, \vec{s}_2\}$). This is continued through n vectors, resulting in the orthonormal set of vectors $\{\vec{q}_1, \ldots, \vec{q}_n\}$ that span the same vector space as $\{\vec{s}_1, \ldots, \vec{s}_n\}$.

How is this done? Finding $\vec{q}_1$ is straightforward, since it is the first vector in our new set, and therefore we must only satisfy $\|\vec{q}_1\| = 1$ and span($\{\vec{q}_1\}$) = span($\{\vec{s}_1\}$). Since span($\{\vec{s}_1\}$) is a one dimensional vector space, the unit vector that spans the same vector space would just be the unit vector in the same direction as $\vec{s}_1$. We have

$$\vec{q}_1 = \frac{\vec{s}_1}{\|\vec{s}_1\|}. \tag{1}$$

Calculating $\vec{q}_2$ requires that we satisfy:

(a) Spanning the same vector space as original set: span($\{\vec{q}_1, \vec{q}_2\}$) = span($\{\vec{s}_1, \vec{s}_2\}$)

(b) Orthogonal to previous vectors: $\langle \vec{q}_1, \vec{q}_2 \rangle = 0$

(c) Normalized: $\|\vec{q}_2\| = 1$

Using the vector $\vec{q}_1$ that we calculated above, we notice that

$$\text{span}(\{\vec{q}_1, \vec{s}_2\}) = \text{span}(\{\vec{s}_1, \vec{s}_2\}),$$

satisfying the first condition. However, $\vec{q}_1$ and $\vec{s}_2$ are not necessarily orthogonal.

We know from EE 16A that the following subspaces are equivalent for any pair of linearly independent vectors $\vec{v}_1, \vec{v}_2$:

- span$(\vec{v}_1,\, \vec{v}_2)$
- span$(\vec{v}_1,\, \alpha\vec{v}_2)$
- span$(\vec{v}_1,\, \vec{v}_1 + \vec{v}_2)$
- span$(\vec{v}_1,\, \vec{v}_1 - \vec{v}_2)$
- span$(\vec{v}_1,\, \vec{v}_2 - \alpha\vec{v}_1)$

Let us choose vector

$$\vec{z}_2 = \vec{s}_2 - \alpha\vec{q}_1,$$

which will also have the same span as $\{\vec{q}_1, \vec{s}_2\}$ (and therefore the same span as $\{\vec{s}_1, \vec{s}_2\}$).

What should α be if we would like $\vec{q}_1$ and $\vec{z}_2 = \vec{s}_2 - \alpha\vec{q}_1$ to be orthogonal to each other? We know from working with Orthogonal Matching Pursuit (OMP), that $\vec{s}_2 - \text{proj}_{\vec{q}_1}(\vec{s}_2)$ will be orthogonal to $\vec{q}_1$, where

$$\text{proj}_{\vec{q}_1}(\vec{s}_2) = \frac{\vec{q}_1^T \vec{s}_2}{\|\vec{q}_1\|^2}\vec{q}_1$$

is the projection of $\vec{s}_2$ onto $\vec{q}_1$. This makes sense because the projection of $\vec{s}_2$ onto $\vec{q}_1$ provides the component of $\vec{s}_2$ that is along $\vec{q}_1$. Subtracting off this component from $\vec{s}_2$ will only leave components of $\vec{s}_2$ that are orthogonal to $\vec{q}_1$.

Therefore, if we set

$$\alpha\vec{q}_1 = \text{proj}_{\vec{q}_1}(\vec{s}_2),$$

the resulting

$$\vec{z}_2 = \vec{s}_2 - \text{proj}_{\vec{q}_1}(\vec{s}_2) = \vec{s}_2 - \frac{\vec{q}_1^T \vec{s}_2}{\|\vec{q}_1\|^2}\vec{q}_1,$$

will be orthogonal to $\vec{q}_1$.

To back up this intuition, let's solve for $\vec{z}_2$ algebraically using the definition of orthogonality:

$$\vec{q}_1^T \vec{z}_2 = 0 \tag{2}$$

$$\vec{q}_1^T (\vec{s}_2 - \alpha\vec{q}_1) = 0 \tag{3}$$

$$\vec{q}_1^T \vec{s}_2 - \alpha\|\vec{q}_1\|^2 = 0 \tag{4}$$

$$\alpha = \frac{\vec{q}_1^T \vec{s}_2}{\|\vec{q}_1\|^2} \tag{5}$$

$$\rightarrow \vec{z}_2 = \vec{s}_2 - \frac{\vec{q}_1^T \vec{s}_2}{\|\vec{q}_1\|^2}\vec{q}_1 \tag{6}$$

Now we normalize $\vec{z}_2$ to complete the process of finding the $\vec{q}_2$ which satisfies all three conditions above:

$$\vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|}$$

In the question below, you will work through how this methodology leads to the Gram-Schmidt algorithm for calculating the orthonormal set $\{\vec{q}_1, \ldots, \vec{q}_n\}$ from n linearly independent vectors $\{\vec{s}_1, \ldots, \vec{s}_n\}$.

Questions

1. Gram-Schmidt Algorithm

Now let's see how we can do this with a set of three linearly independent vectors $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$.

(a) Find unit vector $\vec{q}_1$ such that span($\{\vec{q}_1\}$) = span($\{\vec{s}_1\}$).

(b) Given $\vec{q}_1$ from the previous step, find $\vec{q}_2$ such that $span(\{\vec{q}_1, \vec{q}_2\}) = span(\{\vec{s}_1, \vec{s}_2\})$ and $\vec{q}_2$ is orthogonal to $\vec{q}_1$.

(c) Now given $\vec{q}_1$ and $\vec{q}_2$ in the previous steps, find $\vec{q}_3$ such that $span(\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}) = span(\{\vec{s}_1, \vec{s}_2, \vec{s}_3\})$.

(d) Let's extend this algorithm to n linearly independent vectors. That is, given an input $\{\vec{s}_1, \ldots, \vec{s}_n\}$, write the algorithm to calculate the orthonormal set of vectors $\{\vec{q}_1, \ldots, \vec{q}_n\}$, where span($\{\vec{s}_1, \ldots, \vec{s}_n\}$) = span($\{\vec{q}_1, \ldots, \vec{q}_n\}$). *Hint: How would you calculate the i^{th} vector, $\vec{q}_i$?*

2. The Order of Gram-Schmidt

(a) If we are performing the Gram-Schmidt method on a set of vectors, does the order in which we take the vectors matter? Consider the set of vectors

$$\left\{\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3\right\} = \left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\} \tag{7}$$

Perform Gram-Schmidt on these vectors first in the order $\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3$. and then in the order $\vec{v}_3,\ \vec{v}_2,\ \vec{v}_1$. Do you get the same answer?

(b) Now perform Gram-Schmidt on these vectors in the order $\vec{v}_3,\ \vec{v}_2,\ \vec{v}_1$. Do you get the same result?

3. Orthonormal Matrices and Projections

An orthonormal matrix, $\mathbf{A}$, is a matrix whose columns, $\vec{a}_i$, are:

- Orthogonal (ie. $\langle \vec{a}_i, \vec{a}_j \rangle = 0$ when $i \neq j$)
- Normalized (ie. vectors with length equal to 1, $\|\vec{a}_i\| = 1$). This implies that $\|\vec{a}_i\|_2 = \langle \vec{a}_i, \vec{a}_i \rangle = 1$.

(a) Suppose that the matrix $\mathbf{A} \in \mathbb{R}^{N \times M}$ has linearly independent columns. The vector $\vec{y}$ in $\mathbb{R}^N$ is not in the subspace spanned by the columns of $\mathbf{A}$. What is the projection of $\vec{y}$ onto the subspace spanned by the columns of $\mathbf{A}$?

(b) Show if $\mathbf{A} \in \mathbb{R}^{N \times N}$ is an orthonormal matrix then the columns, $\vec{a}_i$, form a basis for $\mathbb{R}^N$.

(c) When $\mathbf{A} \in \mathbb{R}^{N \times M}$ and $N \geq M$ (i.e. tall matrices), show that if the matrix is orthonormal, then $\mathbf{A}^T\mathbf{A} = \mathbf{I}_{M \times M}$.

(d) Again, suppose $\mathbf{A} \in \mathbb{R}^{N \times M}$ where $N \geq M$ is an orthonormal matrix. Show that the projection of $\vec{y}$ onto the subspace spanned by the columns of $\mathbf{A}$ is now $\mathbf{A}\mathbf{A}^T\vec{y}$.

4. (Optional) Orthogonal Coordinate Change Examples

(a) **Transformation from Standard Basis to another Orthonormal Basis in** $\mathbb{R}^3$
Calculate the coordinate transformation between the following bases

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad V = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{8}$$

i.e. find a matrix T such that $\vec{x}_v = T\vec{x}_u$ where $\vec{x}_u$ is the coordinates of a vector in a basis of the columns of U and $\vec{x}_v$ is the coordinates of the same vector in the basis of the columns of V.

Draw a picture of the two different coordinate frames. Let $\vec{x}_u = [\ 1,\ 0,\ 0\ ]^T$. Compute $\vec{x}_v$ and compare the results with your picture. Repeat this for $\vec{x}_u = [\ 0,\ 1\ ,0\ ]^T$. Are the results intuitive?

Now let $\vec{x}_u = [\ 1,\ 2,\ 1\ ]^T$. What is $\vec{x}_v$? How would you verify that this is correct?

(b) **Transformation between Two Orthonormal Bases in** $\mathbb{R}^3$
Calculate the coordinate transformation between the following bases

$$U = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad V = \begin{bmatrix} 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} \tag{9}$$

i.e. find a matrix T such that $\vec{x}_v = T\vec{x}_u$. Draw a picture of the two different coordinate frames. Let $\vec{x}_u = [\ 1,\ 0, 0]^T$. Compute $\vec{x}_v$ and compare the results with your picture. Repeat this for $\vec{x}_u = [\ 0,\ 1, 0]^T$. Are the results intuitive?

Now let $\vec{x}_u = [\ 1/\sqrt{2},\ 1,\ 1/\sqrt{2}\ ]^T$. What is $\vec{x}_v$? How would you verify that this is correct?

5. (Extra Practice) Gram-Schmidt

(a) Use Gram-Schmidt to find an orthonormal basis for the following three vectors.

$$\vec{v}_1 = \begin{bmatrix} \sqrt{2} \\ -\sqrt{2} \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} \sqrt{2} \\ 0 \\ -\sqrt{2} \end{bmatrix}$$

(b) Express $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$ as vectors in the basis you found in part (a).

Contributors:

- EE 16A folks.
- Regina Eckert.