Chapter 9: Orthogonality — Section 4

The Gram-Schmidt process

What you need to know already:	What you can learn here:
➤ What orthogonal and orthonormal bases are.	➤ How to construct an orthogonal or orthonormal basis from any given basis.

We have seen that orthogonal bases are, in a certain way, "*nice*" bases and that orthonormal bases are even better. So, mathematicians asked whether it is possible to construct an orthogonal or orthonormal basis for a subspace for which we have any good old basis.

One part of the question is easy to answer.

Strategy for changing an orthogonal basis to an orthonormal basis

Divide each vector in the basis by its length.

Proof

Changing the length of a vector does not change its direction, so that all that is needed is a change of length.

Changing a generic basis to an orthogonal one is not as easy, but it is not difficult either, just slightly involved. You may remember that one of the *Learning questions* in a previous section asked you to do that for a basis consisting of only two vectors.

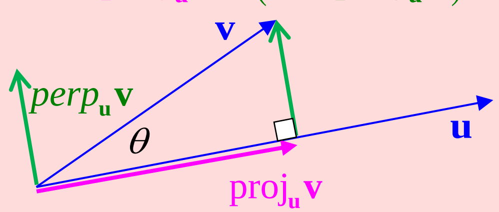
If you developed the solution to that question, this section is just a generalization of it. If not, I am about to explain how it is done!

All we need to do is use the decomposition of a vector generated by the projection operation. Here is a refresher of what this is.

Technical fact

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, $\mathbf{v}$ can be written as the sum of a vector parallel to $\mathbf{u}$ and one perpendicular to $\mathbf{u}$, as follows:

$$\mathbf{v} = proj_{\mathbf{u}}\mathbf{v} + perp_{\mathbf{u}}\mathbf{v}$$
$$= proj_{\mathbf{u}}\mathbf{v} + \left(\mathbf{v} - proj_{\mathbf{u}}\mathbf{v}\right)$$



So, if we start from any two non-parallel vectors $\{\mathbf{u}, \mathbf{v}\}$, their span can also be obtained as the span of $\{\mathbf{u}, perp_{\mathbf{u}}\mathbf{v}\}$, but these two new vectors are perpendicular to each other by construction.

Well, we can generalize this procedure in order to obtain an orthogonal basis from any starting basis.

Strategy for constructing an orthogonal basis: The Gram-Schmidt process

Given a basis $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ for a subspace $\mathcal{V}$ of $\mathbb{R}^n$, to obtain an ***orthogonal basis*** $\{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_k\}$ of $\mathcal{V}$, we can proceed as follows:

1. Let $\mathbf{w}_1 = \mathbf{v}_1$
2. Let $\mathbf{w}_2 = \mathbf{v}_2 - proj_{\mathbf{w}_1}\mathbf{v}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \bullet \mathbf{w}_1}{\mathbf{w}_1 \bullet \mathbf{w}_1}\mathbf{w}_1$
3. Let $\mathbf{w}_3 = \mathbf{v}_3 - proj_{\mathbf{w}_1}\mathbf{v}_3 - proj_{\mathbf{w}_2}\mathbf{v}_3$
   $= \mathbf{v}_3 - \frac{\mathbf{v}_3 \bullet \mathbf{w}_1}{\mathbf{w}_1 \bullet \mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{v}_3 \bullet \mathbf{w}_2}{\mathbf{w}_2 \bullet \mathbf{w}_2}\mathbf{w}_2$
4. …
5. $\mathbf{w}_k = \mathbf{v}_k - \sum_{i=1}^{k-1} \frac{\mathbf{v}_k \bullet \mathbf{w}_i}{\mathbf{w}_i \bullet \mathbf{w}_i}\mathbf{w}_i$

Huh? What are all these formulae?

The formulae we need to get the vectors of the new basis. Basically, at each step we are "*straightening*" the new vector to make it orthogonal to all of the previous ones. The proof that the formulae work and provide an orthogonal basis will hopefully shed light also on the content of their structure.

Proof

We need to check that what we obtain in this way is a set of orthogonal vectors and that it forms a basis for $\mathcal{V}$.

That the vectors are orthogonal can be checked iteratively:

$$\mathbf{w}_2 \bullet \mathbf{w}_1 = \left(\mathbf{v}_2 - \frac{\mathbf{v}_2 \bullet \mathbf{w}_1}{\mathbf{w}_1 \bullet \mathbf{w}_1}\mathbf{w}_1\right) \bullet \mathbf{w}_1 = \mathbf{v}_2 \bullet \mathbf{w}_1 - \frac{\mathbf{v}_2 \bullet \mathbf{w}_1}{\cancel{\mathbf{w}_1 \bullet \mathbf{w}_1}}\cancel{\mathbf{w}_1 \bullet \mathbf{w}_1}$$
$$= \mathbf{v}_2 \bullet \mathbf{w}_1 - \mathbf{v}_2 \bullet \mathbf{w}_1 = 0$$

Now that we know this, we have:

$$\mathbf{w}_3 \bullet \mathbf{w}_2 = \left(\mathbf{v}_3 - \frac{\mathbf{v}_3 \bullet \mathbf{w}_1}{\mathbf{w}_1 \bullet \mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{v}_3 \bullet \mathbf{w}_2}{\mathbf{w}_2 \bullet \mathbf{w}_2}\mathbf{w}_2\right) \bullet \mathbf{w}_2$$
$$= \mathbf{v}_3 \bullet \mathbf{w}_2 - 0 - \frac{\mathbf{v}_3 \bullet \mathbf{w}_2}{\cancel{\mathbf{w}_2 \bullet \mathbf{w}_2}}\cancel{\mathbf{w}_2 \bullet \mathbf{w}_2} = 0$$

All other checks proceed in the same way.

Since the vectors are all orthogonal to each other they are linearly independent. Moreover, since they were obtained though linear combinations of vectors in $\mathcal{V}$, they are still in $\mathcal{V}$. Therefore we have k linearly independent vectors of $\mathcal{V}$, that is, a basis.

And then, to get an orthonormal basis from this, all we have to do is divide each new vector by its length.

Yep. Not only that, but we can use the same idea of scalar multiplication to make our computations easier at any intermediate steps. That is why I am delaying the example on how the process works.

Knot on your finger

At any point in the Gram-Schmidt process, the ***length*** of the vectors constructed up to that point can be ***changed*** in order to make the next computations easier.

Adding this step does not change the orthogonality of the vectors constructed and may be referred to as the ***modified Gram-Schmidt*** process.

This modified process can be used, for instance, to ***eliminate fractions*** when working by hand.

So, here is an example involving the Gram-Schmidt process and some tricks to modify it efficiently.

Example: $B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

This set forms a pretty basis for $\mathbb{R}^3$, but it is not an *orthogonal* basis. To apply the Gram-Schmidt process to it, we start by letting:

$$\mathbf{w}_1 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

To get the second vector we construct:

$$\mathbf{w}_2 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} - perp_{\mathbf{w}_1} \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \bullet \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}}{\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \bullet \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} =$$

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

You can easily check that this is orthogonal to $\mathbf{w}_1$. In order to avoid having to work with fractions, we modify this second vector and instead use:

$$\mathbf{w}_2 = 3 \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & -2 \end{bmatrix}$$

Notice that this is still perpendicular to the first vector. For the third vector, we now compute:

$$\mathbf{w}_3 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \bullet \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}}{\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \bullet \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} - \frac{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \bullet \begin{bmatrix} 1 & 1 & -2 \end{bmatrix}}{\begin{bmatrix} 1 & 1 & -2 \end{bmatrix} \bullet \begin{bmatrix} 1 & 1 & -2 \end{bmatrix}} \begin{bmatrix} 1 & 1 & -2 \end{bmatrix} =$$

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

Again, it is easy to check that this vector is orthogonal to both of the previous ones. To eliminate the fractions, we modify it to be:

$$\mathbf{w}_3 = 2 \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}$$

Therefore, the orthogonal basis we obtain is:

$$B_1 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\}$$

And if we want an orthonormal basis, we simply divide each of them by its length, thus obtaining:

$$B_2 = \left\{ \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} \right\}$$

You may want to check that these vectors, put together, form an orthogonal matrix. Do you remember the quick way to do that?

Yes: multiply by the transpose!

I agree that this is not a difficult process, but it long, tedious and full of number-crunching. Any way to make it shorter?

The same method we can always use, these days, to make a repetitive process less boring.

Use a computer?

Yes! Here is a challenge for you: write a program on your graphing calculator (or by using a more advanced computing language) to implement the Gram-Schmidt process!

Summary

- The Gram-Schmidt process is a simple, but repetitive process that allows us to change any basis to an orthogonal basis.
- By using an additional modification, it can also provide an orthonormal basis, or simplify the intermediate number-crunching steps.

Common errors to avoid

- Just because the process looks simple, but boring, does not mean that you can skip the practice on its implementation by hand! Rather try it sufficiently, until you understand how to do it and how it works.

Learning questions for Section LA 9-4

Review questions:

1. Explain the purpose of the Gram-Schmidt process.
2. Describe how to implement the Gram-Schmidt process.

Memory questions:

1. What is the purpose of the Gram-Schmidt process?
2. What modification can be done at each step of the Gram-Schmidt process without compromising the final goal?
3. What method is needed to change an independent set to an orthonormal set?

Computation questions:

1. Apply the Gram-Schmidt process to the rows of the matrix $\begin{bmatrix} 1 & 0 & -1 & 1 \\ 1 & -2 & 1 & -1 \\ 2 & 0 & -1 & 0 \end{bmatrix}$

2. Apply the Gram-Schmidt process to the columns of the matrix $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$.

3. Check that the rows of the matrix $\begin{bmatrix} -1/2 & \sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ are not orthogonal and apply the Gram-Schmidt process to change them to an orthonormal basis.

4. Use the Gram-Schmidt process to construct an orthogonal basis and then an orthonormal basis for $\textbf{span}\left\{\begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \\ 0 \end{bmatrix}\right\}$

5. Given the matrix $\begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$, construct an orthogonal matrix with the same first row and whose row space is the same as that of the given matrix.

6. Apply the Gram-Schmidt process to the set $\left\{\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right\}$ and obtain an orthonormal basis for $\mathbb{R}^3$.

7. Construct an orthogonal basis for $\mathbb{R}^4$ that contains the $\begin{bmatrix} 1 & 1 & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 0 & 2 & 1 \end{bmatrix}$ and that consists of vectors whose components are integers.

8. Prove that the vectors $\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2/3 \\ -1/3 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -5/2 \end{bmatrix}$ form an orthogonal basis for $\mathbb{R}^3$ and then use them to construct an orthonormal basis.

9. Compute the projection of the vector $\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$ on the plane spanned by the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$. Also, compute its projection on the orthogonal complement to such plane.

Theory questions:

1. What basic vector operation is key to the Gram-Schmidt process?

2. What modification can be done at each step of the Gram-Schmidt process without compromising the final goal?

3. If $\{\mathbf{u}, \mathbf{v}\}$ is a basis for $\mathbb{R}^2$, is it true that any vector is equal to the sum of its projections on the two basis vectors?

4. Is it always possible to change a set of vectors into an orthonormal set by using the Gram-Schmidt process?

5. If we apply the Gram-Schmidt process to the rows of an orthogonal matrix, what do we get?

Templated questions:

1. Select three independent 4D vectors and use the modified Gram-Schmidt process to obtain from them an orthonormal basis for $\mathbb{R}^4$.

What questions do you have for your instructor?