Math 425 Applied & Comput. Lin. Algebra Fall 2024 Lecture

The Gram-Schmidt Process

In this lecture we assume that V is an inner product space with inner product $\langle \mathbf{v}, \mathbf{w} \rangle$ and $\dim V = n$. In particular, we assume that we have already computed a basis for V: $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$. The Gram-Schmidt process is an algorithm that computes an orthogonal (and hence an orthonormal) basis of V: $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. This algorithm will cycle through n iterations. In the ith iteration, it will start out with $i-1$ orthogonal vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$, and it will compute a new vector $\mathbf{v}_i$ that is orthogonal to the first $i-1$ vectors. In the first $(i = 1)$ iteration we set $\mathbf{v}_1 = \mathbf{w}_1$. If $n = 1$, then we are done. If $n > 1$. We proceed with iteration two $(i = 2)$. Because the span of $\mathbf{v}_1 = \mathbf{w}_1$ does not contain $\mathbf{w}_2$, the vectors $\mathbf{v}_1$ and $\mathbf{w}_2$ are linearly independent, and so no multiple of $\mathbf{v}_1$ can be equal to $\mathbf{w}_2$. Hence $\mathbf{v}_2 = \mathbf{w}_2 - c\mathbf{v}_1$ is nonzero for any scalar c. Can we choose a scalar c so that $\mathbf{v}_2$ and $\mathbf{v}_1$ are orthogonal ? If there were such a scalar then $0 = \langle \mathbf{v}_2, \mathbf{v}_1 \rangle = \langle \mathbf{w}_2, \mathbf{v}_1 \rangle - c\langle \mathbf{v}_1, \mathbf{v}_1 \rangle$. Therefore,

$$c = \frac{\langle \mathbf{w}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}$$

will do the job. It is important to note that $\mathrm{span}\{\mathbf{w}_1, \mathbf{w}_2\} = \mathrm{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ (do you see why?).

Now suppose that in the first $i-1$ iterations we constructed orthogonal vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$ such that the span of these $i-1$ vectors is equal to the span of $\mathbf{w}_1, \ldots, \mathbf{w}_{i-1}$. This means $\mathbf{w}_i$ is *not* in span $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}$ and therefore $\mathbf{v}_i = \mathbf{w}_i - c_1\mathbf{v}_1 - c_2\mathbf{v}_2 - \cdots - c_{i-1}\mathbf{v}_{i-1}$ is a nonzero vector for any choice of scalars $c_1, \ldots, c_{i-1}$. We choose scalars so that $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for all $j = 1, \ldots, i-1$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$ are already mutually orthogonal we see that

$$c_j = \frac{\langle \mathbf{w}_i, \mathbf{v}_j \rangle}{\langle \mathbf{v}_j, \mathbf{v}_j \rangle}$$

for $j = 1, \ldots, i-1$. Meanwhile, we maintain the invariance that $\mathrm{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_i\} = \mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$. After n iterations we will obtain n mutually orthogonal vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. These are guaranteed to be linearly independent by the result we proved in the previous lecture. Since $\dim V = n$, we have computed an orthogonal basis!

Let's see the Gram-Schmidt process is action on a simple example.

Example 1. We will use the Euclidean inner product on $\mathbb{R}^3$. The vectors

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{w}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \ \mathbf{w}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

form a basis of $\mathbb{R}^3$. In the first iteration of the Gram-Schmidt process we set $\mathbf{v}_1 = \mathbf{w}_1$. In the second iteration $c = \frac{\langle \mathbf{w}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} = \frac{1}{2}$. Therefore

$$\mathbf{v}_2 = \mathbf{w}_2 - \frac{1}{2}\mathbf{v}_1 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}.$$

In the third (and last) iteration $c_1 = \frac{\langle \mathbf{w}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} = \frac{1}{2}$ and $c_2 = \frac{\langle \mathbf{w}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} = \frac{1/2}{3/2} = \frac{1}{3}$. Therefore

$$\mathbf{v}_3 = \mathbf{w}_3 - \frac{1}{2}\mathbf{v}_1 - \frac{1}{3}\mathbf{v}_2 = \begin{pmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}.$$

The Gram-Schmidt process can be streamlined to produce an orthonormal basis *on the fly*: of course, once we have an orthogonal basis we can obtain an orthonormal basis, but we can achieve this as we go. It is quite useful since in the algorithm then all $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$, and the divisions by this quantity is easy and numerically much more stable. We are not going into the details of the most efficient Gram-Schmidt algorithm though.

Orthogonal matrices. Before we leave the topic of orthogonal vectors we will introduce *orthogonal matrices* for later use. Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in an *orthonormal* basis of $\mathbb{R}^n$. We put these vectors in a matrix $Q = (\mathbf{v}_1\, \mathbf{v}_2\, \ldots\, \mathbf{v}_n)$. We realize that $Q^T Q = I_n$. This means that the inverse of Q is its transpose.

Definition 1. An $n \times n$ matrix Q is called an orthogonal matrix if $Q^T Q = I_n$, i.e., $Q^{-1} = Q^T$. The columns of an orthogonal matrix is an orthonormal basis of $\mathbb{R}^n$.

Example 2. We already saw that

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

as an orthonormal basis of $\mathbb{R}^3$. Therefore

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

is an orthogonal matrix.