Math 2270 - Lecture 25 : Gram-Schmidt

Dylan Zwick

Fall 2012

This lecture finishes *section 4.4* from the textbook.

In our last lecture we discussed matrices with orthonormal columns, and why they're so virtuous. However, you're not always so lucky as to be given a matrix with orthonormal columns. If you're just given a matrix A and asked to project $\mathbf{b}$ onto the column space of A, the columns of A might be orthonormal, but you're probably not that fortunate. However, as with so many things in life, you can get over a little bad luck with some work. The work I'm talking about is you can construct a new matrix Q with orthonormal columns and the same column space as A. So, $\mathbf{C}(Q) = \mathbf{C}(A)$. Once constructed, the projection of $\mathbf{b}$ onto the column space of Q will be the same as the projection onto the column space of A, but calculating the projection using Q will be much easier than with A.

The assigned problems for this section are:

Section 4.4 - 2, 6, 11, 18, 21

1 Gram-Schmidt

Suppose we start with three independent vectors $\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$. We want to use these vectors to build three orthonormal vectors $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$. These three vectors will span the same space as $\mathbf{a}, \mathbf{b}, \mathbf{c}$, but will, in general, be much easier to work with.

We first construct three orthogonal vectors $\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$. Once you've got them, making them all length 1 is easy.

The first step is the easiest. We just set $\mathbf{A} = \mathbf{a}$. No need to worry about orthogonality on step 1. Now, we want to construct a vector $\mathbf{B}$ that is orthogonal to $\mathbf{A}$, and can be written as a linear combination of $\mathbf{a}$ and $\mathbf{b}$. We do this by projecting $\mathbf{b}$ onto the subspace spanned by $\mathbf{A}$, and then subtracting away this projection. Stated specifically,

$$\mathbf{B} = \mathbf{b} - \frac{\mathbf{A}^T\mathbf{b}}{\mathbf{A}^T\mathbf{A}}\mathbf{A}.$$

If we take the dot product of $\mathbf{A}$ and $\mathbf{B}$ we get 0:

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{A}^T\mathbf{B} = \mathbf{A}^T\mathbf{b} - \frac{\mathbf{A}^T\mathbf{b}}{\mathbf{A}^T\mathbf{A}}\mathbf{A}^T\mathbf{A} = \mathbf{A}^T\mathbf{b} - \mathbf{A}^T\mathbf{b} = 0.$$

So, $\mathbf{A}$ and $\mathbf{B}$ are orthogonal.

This, really, is the general idea. To get the next vector in our orthogonal set we just subtract away the projection onto the subspace spanned by the vectors we've already constructed. We calculate $\mathbf{C}$ by:

$$\mathbf{C} = \mathbf{c} - \frac{\mathbf{A}^T\mathbf{c}}{\mathbf{A}^T\mathbf{A}}\mathbf{A} - \frac{\mathbf{B}^T\mathbf{c}}{\mathbf{B}^T\mathbf{B}}\mathbf{B}.$$

You can check that **C** is orthogonal to **A** and **B**. If we had more than three vectors we would just keep going in this way.

To get the orthonormal vector $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$ we just divide **A**, **B**, and **C** by their respective lengths. Easy.

$$\mathbf{q}_1 = \frac{\mathbf{A}}{||\mathbf{A}||}, \mathbf{q}_2 = \frac{\mathbf{B}}{||\mathbf{B}||}, \mathbf{q}_3 = \frac{\mathbf{C}}{||\mathbf{C}||}.$$

Example - Find an orthonormal basis for the column space of A.

$$A = \begin{pmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{pmatrix}.$$

2   QR Factorization

We started with a matrix A, whose columns were $\mathbf{a}, \mathbf{b}, \mathbf{c}$. We ended with a matrix Q, whose columns are $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$. We build the columns of Q from the columns of A, and we can build the columns of A from the columns of

Q if we wish. So, there must be some matrix that relates them. This matrix is the square, upper-triangular matrix R. In general if we construct Q from A using the Gram-Schmidt process, the two matrices will be related by $A = QR$, where R is a square, upper-triangular matrix.

The reason the matrix R is upper-triangular is that at each step of Gram-Schmidt, the only vectors that are involved are the one in question, and the ones we've already dealt with. So, when we figured out $\mathbf{q}_1$, the only one of our original vectors involved was $\mathbf{a}$. When we figured out $\mathbf{q}_2$, the only original vectors involved were $\mathbf{a}$ and $\mathbf{b}$, and so on. In fact, we can express this relation as:

$$
\begin{pmatrix} \mathbf{a} & \mathbf{b} & \mathbf{c} \end{pmatrix} = \begin{pmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \end{pmatrix} \begin{pmatrix} \mathbf{q}_1^T \mathbf{a} & \mathbf{q}_1^T \mathbf{b} & \mathbf{q}_1^T \mathbf{c} \\ & \mathbf{q}_2^T \mathbf{b} & \mathbf{q}_2^T \mathbf{c} \\ & & \mathbf{q}_3^T \mathbf{c} \end{pmatrix}.
$$

The matrix on the far-right is the matrix R.

Example - Find the square upper-triangular matrix R connecting the matrices A and Q from the earlier example.