Stat 5102 Notes: Gram-Schmidt

Charles J. Geyer

February 5, 2007

The Gram-Schmidt orthogonalization process is not a topic of the course. It is a fact we assume known from linear algebra. It is used, among other places in statistics, in the proof that the sample mean and sample variance are independent when the population is normal (Section 7.4 in DeGroot and Schervish).

We would say no more, if the instructor (me) had not blathered about it incorrectly in class in answer to a question. Here is the correct version. Read if you are curious, otherwise ignore. You will not need to know this for any exam or homework problem.

Given a set of vectors $\mathcal{A}$ and another set of orthonormal vectors $\mathcal{U}$, the following algorithm, called Gram-Schmidt, finds an orthonormal basis for the vector space spanned by $\mathcal{A} \cup \mathcal{U}$ that contains $\mathcal{U}$.

1. If $\mathcal{A}$ is empty, stop. $\mathcal{U}$ is the result.
2. Remove an element $\mathbf{a}$ from $\mathcal{A}$.
3. Set
$$\mathbf{b} = \mathbf{a} - \sum_{\mathbf{u} \in \mathcal{U}} (\mathbf{u}'\mathbf{a})\mathbf{u}$$
[$\mathbf{b}$ is orthogonal to all $\mathbf{u} \in \mathcal{U}$].
4. If $\mathbf{b} = 0$ return to step 1.
5. Set
$$\mathbf{c} = \frac{\mathbf{b}}{\|\mathbf{b}\|}$$
[$\mathbf{c}$ is orthogonal to all $\mathbf{u} \in \mathcal{U}$ and has norm one].
6. Add $\mathbf{c}$ to $\mathcal{U}$ and return to step 1.

We do one example. Let us start with $\mathcal{U}$ containing the single vector

$$\mathbf{u}_1 = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

and let $\mathcal{A}$ contain the four vectors that are the usual basis for four-dimensional Euclidean space

$$\mathbf{a}_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \ \mathbf{a}_2 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \ \mathbf{a}_3 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \ \mathbf{a}_4 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

In the first iteration of Gram-Schmidt in step 2 we remove $\mathbf{a}_1$ from $\mathcal{A}$, and in step 3 we calculate

$$\begin{aligned}
\mathbf{b} &= \mathbf{a}_1 - (\mathbf{u}_1'\mathbf{a}_1)\mathbf{u}_1 \\
&= \begin{pmatrix}1\\0\\0\\0\end{pmatrix} - \frac{1}{4}\left[\begin{pmatrix}1 & 1 & 1 & 1\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\right]\begin{pmatrix}1\\1\\1\\1\end{pmatrix} \\
&= \begin{pmatrix}1\\0\\0\\0\end{pmatrix} - \frac{1}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} \\
&= \frac{1}{4}\begin{pmatrix}3\\-1\\-1\\-1\end{pmatrix}
\end{aligned}$$

(This was the step botched in class.)

In step 5 we calculate

$$\begin{aligned}
\|\mathbf{b}\|^2 &= \left(\frac{3}{4}\right)^2 + \left(-\frac{1}{4}\right)^2 + \left(-\frac{1}{4}\right)^2 + \left(-\frac{1}{4}\right)^2 \\
&= \frac{12}{16}
\end{aligned}$$

and

$$\mathbf{c} = \frac{1}{4}\begin{pmatrix}3\\-1\\-1\\-1\end{pmatrix} \cdot \frac{4}{\sqrt{12}}$$

In step 6 we add this to $\mathcal{U}$, so call it

$$\mathbf{u}_2 = \frac{1}{\sqrt{12}}\begin{pmatrix}3\\-1\\-1\\-1\end{pmatrix}$$

Now we are ready for another iteration. In the second iteration of Gram-Schmidt in step 2 we remove $\mathbf{a}_2$ from $\mathcal{A}$, and in step 3 we calculate

$$\begin{aligned}
\mathbf{b} &= \mathbf{a}_2 - (\mathbf{u}_1'\mathbf{a}_2)\mathbf{u}_1 - (\mathbf{u}_2'\mathbf{a}_2)\mathbf{u}_2 \\
&= \begin{pmatrix}0\\1\\0\\0\end{pmatrix} - \frac{1}{4}\left[\begin{pmatrix}1 & 1 & 1 & 1\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right]\begin{pmatrix}1\\1\\1\\1\end{pmatrix} \\
&\quad - \frac{1}{12}\left[\begin{pmatrix}3 & -1 & -1 & -1\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right]\begin{pmatrix}3\\-1\\-1\\-1\end{pmatrix} \\
&= \begin{pmatrix}0\\1\\0\\0\end{pmatrix} - \frac{1}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + \frac{1}{12}\begin{pmatrix}3\\-1\\-1\\-1\end{pmatrix} \\
&= \frac{1}{12}\begin{pmatrix}0\\8\\-4\\-4\end{pmatrix}
\end{aligned}$$

In step 5 we calculate

$$\begin{aligned}
\|\mathbf{b}\|^2 &= 0^2 + \left(\frac{2}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 \\
&= \frac{6}{9}
\end{aligned}$$

and

$$\mathbf{c} = \frac{1}{3}\begin{pmatrix}0\\2\\-1\\-1\end{pmatrix} \cdot \frac{3}{\sqrt{6}}$$

In step 6 we add this to $\mathcal{U}$, so call it

$$\mathbf{u}_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}0\\2\\-1\\-1\end{pmatrix}$$

Now we are ready for another iteration. In the third iteration of Gram-

Schmidt in step 2 we remove $\mathbf{a}_3$ from $\mathcal{A}$, and in step 3 we calculate

$$\begin{aligned}
\mathbf{b} &= \mathbf{a}_3 - (\mathbf{u}_1'\mathbf{a}_3)\mathbf{u}_1 - (\mathbf{u}_2'\mathbf{a}_3)\mathbf{u}_2 - (\mathbf{u}_3'\mathbf{a}_3)\mathbf{u}_3 \\
&= \begin{pmatrix}0\\0\\1\\0\end{pmatrix} - \frac{1}{4}\left[\begin{pmatrix}1 & 1 & 1 & 1\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right]\begin{pmatrix}1\\1\\1\\1\end{pmatrix} \\
&\quad - \frac{1}{12}\left[\begin{pmatrix}3 & -1 & -1 & -1\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right]\begin{pmatrix}3\\-1\\-1\\-1\end{pmatrix} \\
&\quad - \frac{1}{6}\left[\begin{pmatrix}0 & 2 & -1 & -1\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right]\begin{pmatrix}0\\2\\-1\\-1\end{pmatrix} \\
&= \begin{pmatrix}0\\0\\1\\0\end{pmatrix} - \frac{1}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + \frac{1}{12}\begin{pmatrix}3\\-1\\-1\\-1\end{pmatrix} + \frac{1}{6}\begin{pmatrix}0\\2\\-1\\-1\end{pmatrix} \\
&= \frac{1}{12}\begin{pmatrix}0\\0\\6\\6\end{pmatrix}
\end{aligned}$$

In step 5 we calculate

$$\begin{aligned}
\|\mathbf{b}\|^2 &= 0^2 + 0^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 \\
&= \frac{2}{4}
\end{aligned}$$

and

$$\mathbf{c} = \frac{1}{2}\begin{pmatrix}0\\0\\1\\-1\end{pmatrix} \cdot \frac{2}{\sqrt{2}}$$

In step 6 we add this to $\mathcal{U}$, so call it

$$\mathbf{u}_4 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\-1\end{pmatrix}$$

Now we are ready for another iteration. In the fourth iteration of Gram-Schmidt in step 2 we remove $\mathbf{a}_4$ from $\mathcal{A}$, and in step 3 we calculate

$$
\begin{aligned}
\mathbf{b} &= \mathbf{a}_4 - (\mathbf{u}_1'\mathbf{a}_4)\mathbf{u}_1 - (\mathbf{u}_2'\mathbf{a}_4)\mathbf{u}_2 - (\mathbf{u}_3'\mathbf{a}_4)\mathbf{u}_3 - (\mathbf{u}_4'\mathbf{a}_4)\mathbf{u}_4 \\
&= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{4}\left[\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right]\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\
&\quad - \frac{1}{12}\left[\begin{pmatrix} 3 & -1 & -1 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right]\begin{pmatrix} 3 \\ -1 \\ -1 \\ -1 \end{pmatrix} \\
&\quad - \frac{1}{6}\left[\begin{pmatrix} 0 & 2 & -1 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right]\begin{pmatrix} 0 \\ 2 \\ -1 \\ -1 \end{pmatrix} \\
&\quad - \frac{1}{2}\left[\begin{pmatrix} 0 & 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right]\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{12}\begin{pmatrix} 3 \\ -1 \\ -1 \\ -1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 0 \\ 2 \\ -1 \\ -1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \\
&= \frac{1}{12}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

Since $\mathbf{b}$ is zero we go back to step 1, and since $\mathcal{A}$ is now empty (we have removed each of its original four elements, one in each iteration) we are done. The orthonormal basis is

$$
\mathbf{u}_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{u}_2 = \frac{1}{\sqrt{12}}\begin{pmatrix} 3 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \ \mathbf{u}_3 = \frac{1}{\sqrt{6}}\begin{pmatrix} 0 \\ 2 \\ -1 \\ -1 \end{pmatrix}, \ \mathbf{u}_4 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}
$$

The matrix whose rows are these vectors

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{\sqrt{12}} & -\frac{1}{\sqrt{12}} & -\frac{1}{\sqrt{12}} & -\frac{1}{\sqrt{12}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

is one possibility for the **A** that is used in the proof if the big theorem in Section 7.4 of DeGroot and Schervish.