MATH 223. Orthonormal bases and Gram-Schmidt process. Richard Anstee

Consider a vector space V with an inner product $<,>: V \times V \rightarrow \mathbf{R}$. We are interested in finding *orthonormal* bases for vector spaces. An orthonomal basis $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_t\}$ is a basis so that

$$< \mathbf{w}_i, \mathbf{w}_j >= \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

An orthonomal basis has the basis vectors mutually orthogonal *and* of unit length.

Let U be a vector subspace of V with U having some basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$. We seek a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ which form an othonormal basis for U. The way we implement Gram-Schmidt for hand calculation, we do not normalize our vectors until the last step to avoid all the square roots.

First start with $k = 2$. Let U be a vector subspace of V with U having some basis $\{\mathbf{u}_1, \mathbf{u}_2\}$. We set

$$\mathbf{v}_1 = \mathbf{u}_1.$$

Then we do the standard projection (if you are familiar with this in Physics),

$$\mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{\mathbf{v}_1}\mathbf{u}_2$$

We readily compute that

$$< \mathbf{v}_1, \mathbf{v}_2 >=< \mathbf{v}_1, \mathbf{u}_2 > - < \mathbf{v}_1, \text{proj}_{\mathbf{v}_1}\mathbf{u}_2 >=< \mathbf{v}_1, \mathbf{u}_2 > - \frac{< \mathbf{u}_2, \mathbf{v}_1 >}{< \mathbf{v}_1, \mathbf{v}_1 >} < \mathbf{v}_1, \mathbf{v}_1 >= 0$$

Also we note that $\mathbf{v}_1, \mathbf{v}_2 \in \text{span}(\mathbf{u}_1, \mathbf{u}_2)$ and moreover, we may write the equations as

$$\mathbf{u}_1 = \mathbf{v}_1,$$

$$\mathbf{u}_2 = \mathbf{v}_2 + \text{proj}_{\mathbf{v}_2}\mathbf{u}_2.$$

Thus $\mathbf{u}_1, \mathbf{u}_2 \in \text{span}(\mathbf{v}_1, \mathbf{v}_2)$ from which we conclude $\text{span}(\mathbf{u}_1, \mathbf{u}_2) = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$. This becomes the inductive step in our proof.

Example Say we have discovered that $\text{span}\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis for an eigenspace given by the equation $3x - 2y + z = 0$. Then we can obtain an orthonormal basis for that eigenspace. Here the inner product is the dot product.

$$\mathbf{u}_1 = \begin{bmatrix} -1/3 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2/3 \\ 1 \\ 0 \end{bmatrix}$$

We clear fractions and instead use

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$$

$$\mathbf{v}_1 = \mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{\mathbf{v}_1}\mathbf{u}_2 = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}}{\begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}} \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} - \frac{-2}{10} \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{18}{10} \\ 3 \\ \frac{6}{10} \end{bmatrix} \text{ could use } \begin{bmatrix} 18 \\ 30 \\ 6 \end{bmatrix} \text{ or } \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix}$$

You may check that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ and of course $\text{span}(\mathbf{u}_1, \mathbf{u}_2) = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$. The latter isn't immediately obvious until you look at the equation determining $\mathbf{v}_2$.

The general Gram-Schmidt algorithm (where we hold off normalizing our vectors until later) can be written is as follows:

$$\begin{array}{llll} \mathbf{v}_1 = \mathbf{u}_1. & & & \\ \mathbf{v}_2 = \mathbf{u}_2 & -\text{proj}_{\mathbf{v}_1}\mathbf{u}_2 & & \\ \mathbf{v}_3 = \mathbf{u}_3 & -\text{proj}_{\mathbf{v}_1}\mathbf{u}_3 & -\text{proj}_{\mathbf{v}_2}\mathbf{u}_3 & \\ & \vdots & & \\ \mathbf{v}_k = \mathbf{u}_k & -\text{proj}_{\mathbf{v}_1}\mathbf{u}_k & -\text{proj}_{\mathbf{v}_2}\mathbf{u}_k \quad \cdots & -\text{proj}_{\mathbf{v}_{k-1}}\mathbf{u}_k \end{array}$$

Lemma 0.1 *span*$\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_t\} =$ *span*$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t\}$

Proof: We do this by induction on t. The result is easy for $t = 1, 2$ as we have done above. Now imagine we are defining $\mathbf{v}_t$ from $\mathbf{u}_t$ subtracting all the projections namely $\mathbf{v}_t = \mathbf{u}_t - \text{proj}_{\mathbf{v}_1}\mathbf{u}_t - \text{proj}_{\mathbf{v}_2}\mathbf{u}_t \cdots - \text{proj}_{v_{t-1}}\mathbf{u}_t$. We immediately have $\mathbf{v}_t \subseteq \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{t-1}, \mathbf{u}_t\}$ and so using induction that $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{t-1}\} = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{t-1}\}$ we deduce that $\{v_1, \mathbf{v}_2, \ldots, \mathbf{v}_t\} \subseteq \text{span}\{u_1, \mathbf{u}_2, \ldots, \mathbf{u}_t\}$. In a similar way we have $\mathbf{u}_t = \mathbf{v}_t + \text{proj}_{\mathbf{v}_1}\mathbf{u}_t + \text{proj}_{\mathbf{v}_2}\mathbf{u}_t \cdots + \text{proj}_{v_{t-1}}\mathbf{u}_t$ and so using $\text{span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{t-1}\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{t-1}\}$, we obtain $u_t \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t\}$. Thus $\text{span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_t\} \subseteq \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t\}$.

We may conclude $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t\} = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_t\}$. ∎

Lemma 0.2 *After we have completed Gram-Schmidt, we have* $<\mathbf{v}_i, \mathbf{v}_j>= 0$ *for* $i \neq j$.

Proof: Use induction on t so that are induction hypothesis is that $<\mathbf{v}_i, \mathbf{v}_j>= 0$ for $1 \leq i < j < t$. Assume $i < j = t$ and then

$$<\mathbf{v}_i, \mathbf{v}_t>=<\mathbf{v}_i, \mathbf{u}_t> - <\mathbf{v}_i, \text{proj}_{\mathbf{v}_1}\mathbf{u}_t> - <\mathbf{v}_i, \text{proj}_{\mathbf{v}_2}\mathbf{u}_t> \cdots - <\mathbf{v}_i, \text{proj}_{\mathbf{v}_{j-1}}\mathbf{u}_t> .$$

Using induction that $<\mathbf{v}_i, \mathbf{v}_j>= 0$ for $1 \leq i < j < t$, we can get rid of all the projection terms except the last so that

$$<\mathbf{v}_i, \mathbf{v}_t>=<\mathbf{v}_i, \mathbf{u}_t> - <\mathbf{v}_i, \text{proj}_{\mathbf{v}_i}\mathbf{u}_t>=<\mathbf{v}_i, \mathbf{u}_t> - \frac{<\mathbf{v}_i, \mathbf{u}_t>}{<\mathbf{v}_i, \mathbf{v}_i>} <\mathbf{v}_i, \mathbf{v}_i>= 0$$

This completes the proof. ∎

Example

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{v}_1 = \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{\mathbf{v}_1}\mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

$$\mathbf{v}_3 = \mathbf{u}_3 - \text{proj}_{\mathbf{v}_1}\mathbf{u}_3 - \text{proj}_{\mathbf{v}_2}\mathbf{u}_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{-1}{2} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1/2 \\ -1/2 \end{bmatrix}$$

Then the following three vectors are an orthogonal basis for $\mathbf{R}^n$.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1/2 \\ -1/2 \end{bmatrix}$$

Ther are not orthonomal but you can dvide them by their lengths to obtain an orthonormal basis for $\mathbf{R}^n$:

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2/\sqrt{6} \\ -1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix} \tag{1}$$

One application is in (2) below.

Example An important example of an orthogonal basis arises for continuous functions when we define

$$< f, g >= \int_0^{2\pi} f(x)g(x)dx.$$

One can verify that the $2n+1$ functions

$$1, \sin(x), \cos(x), \sin(2x), \cos(2x), \ldots, \sin(nx), \cos(nx)$$

are orthogonal (I can do the first few easily!). To obtain an orthonormal basis we must devide by length.

$$\int_0^{2\pi} 1dx = 2\pi$$

so

$$< \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{2\pi}} >= 1.$$

Similarly

$$\int_0^{2\pi} \sin^2(x)dx = \pi$$

and so

$$< \frac{1}{\sqrt{\pi}}\sin(x), \frac{1}{\sqrt{\pi}}\sin(x) >= 1.$$

Orthogonal Matrices

Many interesting thing happens when we have an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in $\mathbf{R}^n$. Let M be the $n \times n$ matrix formed as $M = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3\, \cdots \mathbf{v}_n]$. we compute that $M^T M = I$ since the i, j entry of $M^T M$ is the dot product $\mathbf{v}_i^T \mathbf{v}_j$. Thus $M^T = M^{-1}$.

Definition 0.3 *We say an $n \times n$ M is an orthogonal matrix if $M^T = M^{-1}$. If M is an orthogonal matrix then the rows of M form an orthonormal basis for $\mathbf{R}^n$ and the columns of M form an orthonormal basis for $\mathbf{R}^n$.*

Example

Using the orthonormal basis from (1), we obtain

$$M = \begin{bmatrix} 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix} \tag{2}$$