22m:033 Notes:
6.2 Orthogonal Sets

Dennis Roseman
University of Iowa
Iowa City, IA

http://www.math.uiowa.edu/∼roseman

April 29, 2010

1 Orthogonal Basis

Definition 1.1 *A* ***set of vectors is an orthogonal set of vectors*** *if every pair of (distinct) vectors is orthogonal.*

Example 1.2 We can quickly check that

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$$

is an orthogonal set

Proposition 1.3 *A set of n orthogonal non-zero vectors gives a basis for R^n.*

Definition 1.4 *An* ***orthogonal basis for a subspace*** *W is a set of orthogonal vectors of W which is also a basis for W.*

If we have any basis $\{\overrightarrow{u_1}, \ldots, \overrightarrow{u_p}\}$ for a subspace W, and if $\overrightarrow{y} \in W$ we can write

$$\overrightarrow{y} = c_1\overrightarrow{u_1} + \cdots + c_p\overrightarrow{u_p}.$$

For an orthogonal basis there is a nice way of calculating the c_i:

Proposition 1.5 *If* $\{\overrightarrow{u_1}, \ldots, \overrightarrow{u_p}\}$ *is an orthogonal basis for subspace* W *in* R^n *then any* $\overrightarrow{y} \in W$ *can be written as the following linear combination:*

$$\overrightarrow{y} = \left(\frac{\overrightarrow{y} \cdot \overrightarrow{u_1}}{\overrightarrow{u_1} \cdot \overrightarrow{u_1}}\right)\overrightarrow{u_1} + \cdots + \left(\frac{\overrightarrow{y} \cdot \overrightarrow{u_p}}{\overrightarrow{u_p} \cdot \overrightarrow{u_p}}\right)\overrightarrow{u_p}$$

.

In other words, $c_i = \left(\frac{\overrightarrow{y} \cdot \overrightarrow{u_i}}{\overrightarrow{u_i} \cdot \overrightarrow{u_i}}\right)$

Example 1.6 Suppose $W = R^3$ and we use the basis of Example 1.2.

Consider $y = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$. We want to express y as a linear combination of these basis vectors. We can do this in a straight forward way as we did in our Chapter 1 (do you recall how?).

Or we can use the fact that this basis is orthogonal and quickly calculate:

$$\begin{pmatrix}1\\2\\3\end{pmatrix} \cdot \begin{pmatrix}1\\0\\1\end{pmatrix} = 4 \text{ and } \begin{pmatrix}1\\0\\1\end{pmatrix} \cdot \begin{pmatrix}1\\0\\1\end{pmatrix} = 2$$

$$\begin{pmatrix}1\\2\\3\end{pmatrix} \cdot \begin{pmatrix}0\\1\\0\end{pmatrix} = 2 \text{ and } \begin{pmatrix}0\\1\\0\end{pmatrix} \cdot \begin{pmatrix}0\\1\\0\end{pmatrix} = 1$$

$$\begin{pmatrix}1\\2\\3\end{pmatrix} \cdot \begin{pmatrix}1\\0\\-1\end{pmatrix} = -2 \text{ and } \begin{pmatrix}1\\0\\-1\end{pmatrix} \cdot \begin{pmatrix}1\\0\\-1\end{pmatrix} = 2$$

So we see that

$$\begin{pmatrix}1\\2\\3\end{pmatrix} = \frac{4}{2}\begin{pmatrix}1\\0\\1\end{pmatrix} + \frac{2}{1}\begin{pmatrix}0\\1\\0\end{pmatrix} + \frac{-2}{2}\begin{pmatrix}1\\0\\-1\end{pmatrix}.$$

We can quickly verify that this last equation is correct.

2 Orthogonal Projection

The formula $c_i = \left(\frac{\overrightarrow{y} \cdot \overrightarrow{u_i}}{\overrightarrow{u_i} \cdot \overrightarrow{u_i}}\right)$ involves just two vectors and is the basis for the following definition:

Definition 2.1 *Let $\overrightarrow{y}$ and $\overrightarrow{u}$ be any two vectors in R^n with $\overrightarrow{u} \neq \overrightarrow{0}$. The* ***projection of $\overrightarrow{y}$ onto $\overrightarrow{u}$*** *is*

$$proj_{\overrightarrow{u}}\overrightarrow{y} = \left(\frac{\overrightarrow{y} \cdot \overrightarrow{u}}{\overrightarrow{u} \cdot \overrightarrow{u}}\right)\overrightarrow{u}$$

Remark 2.2 Two important things: the projection of projection of $\overrightarrow{y}$ onto $\overrightarrow{u}$ is a multiple of $\overrightarrow{u}$.

Also there is generally a big difference between projection of $\overrightarrow{y}$ onto $\overrightarrow{u}$ and projection of $\overrightarrow{y}$ onto $\overrightarrow{u}$.

For example if $\overrightarrow{y} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\overrightarrow{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$$proj_{\overrightarrow{u}}\overrightarrow{y} = \left(\frac{\overrightarrow{y} \cdot \overrightarrow{u}}{\overrightarrow{u} \cdot \overrightarrow{u}}\right)\overrightarrow{u} = \frac{6}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

while

$$proj_{\overrightarrow{y}}\overrightarrow{u} = \left(\frac{\overrightarrow{u} \cdot \overrightarrow{y}}{\overrightarrow{y} \cdot \overrightarrow{y}}\right)\overrightarrow{y} = \frac{6}{14}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} \frac{3}{7} \\ \frac{6}{7} \\ \frac{9}{7} \end{pmatrix}$$

Remark 2.3 Our text actually has two definitions. Again given $\overrightarrow{y}$ and $\overrightarrow{u}$ be any two vectors in R^n with $\overrightarrow{u} \neq \overrightarrow{0}$. The vector $\overrightarrow{u}$ spans a subspace L which is a line in R^n. They then define projection onto this *subspace* L rather than the *vector* $\overrightarrow{u}$:

$$proj_{\overrightarrow{u}}\overrightarrow{L} = \left(\frac{\overrightarrow{y} \cdot \overrightarrow{u}}{\overrightarrow{u} \cdot \overrightarrow{u}}\right)\overrightarrow{u}.$$

This is actually a good way to do this. The distinction is minor. However, most math, science and engineering texts use the notation and Definition 2.1, so we will do so also in the class in addition to the "subspace" version.

3 Orthonormal Sets

Definition 3.1 *An* ***orthonormal set*** *is a set of orthogonal unit vectors. If a basis of a subspace W is orthonormal it is call an* ***orthonormal basis for*** *W.*

Remark 3.2 A key word here is "unit".

Also, "orthogonal" and "orthonormal" are very similar words. Be careful not to confuse them.

Example 3.3

Consider

$$\overrightarrow{u_1} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \qquad \overrightarrow{u_2} = \begin{pmatrix} \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{6}} \\ \frac{-1}{\sqrt{6}} \end{pmatrix}, \qquad \overrightarrow{u_3} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix}.$$

We can check that this is an orthonormal basis for R^3 by verifying that:

$$\overrightarrow{u_1} \cdot \overrightarrow{u_1} = 1, \qquad \overrightarrow{u_2} \cdot \overrightarrow{u_2} = 1, \qquad \overrightarrow{u_3} \cdot \overrightarrow{u_3} = 1$$

and

$$\overrightarrow{u_1} \cdot \overrightarrow{u_2} = 0, \qquad \overrightarrow{u_1} \cdot \overrightarrow{u_3} = 0, \qquad \overrightarrow{u_2} \cdot \overrightarrow{u_3} = 0.$$

Proposition 3.4 *An $m \times n$ matrix U has orthogonal columns if and only if $U^T U = I$.*

Proposition 3.5 *If $m \times n$ matrix U has orthogonal columns and $\overrightarrow{x}, \overrightarrow{y} \in R^n$ then*

1. $||U\overrightarrow{x}|| = ||\overrightarrow{x}||$

2. $U\overrightarrow{x} \cdot U\overrightarrow{y} = \overrightarrow{x} \cdot \overrightarrow{y}$

3. $U\overrightarrow{x} \cdot U\overrightarrow{y} = 0$*if an only if*$\overrightarrow{x} \cdot \overrightarrow{y} = 0$

Remark 3.6 In Proposition 3.5, the key part is part 2, since the other two parts follow from this.

It is important to think of the linear transformation corresponding to U. Proposition 3.5 part 1 says this linear transformation preserves length.

Proposition 3.5 part 3 says this linear transformation preserves right angles. In fact from Proposition 3.5 part 2 we see that all angles are preserved.

In terms of physics and engineering, we say that this transformation is a "rigid motion".

Definition 3.7 *An $n \times n$ invertible matrix U is an orthogonal matrix if $U^{-1} = U^T$.*

Remark 3.8 If U is is a square $n \times n$ matrix then then U is an orthogonal matrix if and only if the column vectors of U is an orthonormal basis for R^n. (Also the row vectors will be an orthonormal basis for R^n)

Remark 3.9 If U is a square orthogonal $n \times n$ matrix $det(U) = \pm 1$.

This follows since

$$det(U) = \frac{1}{det(U^{-1})} = \frac{1}{det(U^T)} = \frac{1}{det(U)}.$$

So we see that $(det(U))^2 = 1$ and so $det(U) = \pm 1$.

Also note that the converse is not true since $\left| \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix} \right| =$ 1 and this matrix is clearly not orthogonal.

Remark 3.10 The product of orthogonal matrices is orthogonal. If A and B are $n \times n$ orthogonal matrices, then

$$(AB)^{-1} = B^{-1}A^{-1} = B^T A^T = (AB)^T.$$

In terms of corresponding linear transformations, this says that a composition of rigid motions is a rigid motion, as we should expect.