Section 6.2: Orthogonal Sets

April 16, 2008

Abstract

As you may recall, often we are interested in using bases for which the basis vectors are perpendicular, or orthogonal.

Definition: *orthogonal set* A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ in $\mathbb{R}^n$is said to be an **orthogonal set** if

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0 \text{ whenever } i \neq j.$$

In many cases we like our bases to be orthogonal (that is, the vectors to be mutually perpendicular). Even better are *orthonormal* bases, in which the orthogonal vectors are of unit length.

Theorem 4: If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

Consider

$$\underline{v} = c_1\underline{u}_1 + \ldots + c_p\underline{u}_p = \underline{0}$$

$$\underline{u}_i \cdot \underline{v} = c_i \underbrace{\underline{u}_i \cdot \underline{u}_i}_{>0} = 0 \Rightarrow \boxed{c_i = 0}$$

$\Rightarrow$ S is linearly independent. $\forall i$.

Definition: *orthogonal basis* An **orthogonal basis** for a subspace W of $\mathbb{R}^n$is a basis for W that is also an orthogonal set.

Theorem 5: Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination

B

$$\mathbf{y} = c_1\mathbf{u}_1 + \ldots + c_p\mathbf{u}_p \tag{1}$$

are given by

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}$$

coordinates of the vector y in the basis: $[y]_B$.

$$\underline{u}_j \cdot \underline{y} = c_j\, \underline{u}_j \cdot \underline{u}_j$$

Whoops! We have a notational collision: the author wants to use the "hat" symbol to indicate the orthogonal projection of $\mathbf{y}$ onto another vector. To make your lives easier, I'll give up my notation, albeit unhappily. I'll indicate unit vectors by the notation $\breve{\mathbf{u}}$: hence,

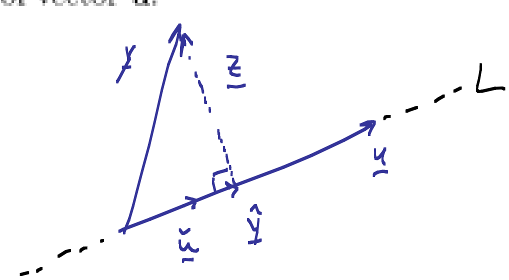
$$\breve{\mathbf{u}} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$$

I don't like this, because the vector $\hat{\mathbf{y}}$ (as defined below) is different for different vectors $\mathbf{u}$. Mathworld (maintainers of Mathematica), many other mathematicians, and I like to reserve the "hat" for unit vectors[1]

So

$$\hat{\mathbf{y}} = \text{proj}_L \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} = (\mathbf{y} \cdot \breve{\mathbf{u}})\breve{\mathbf{u}},$$

where L in this case is the subspace generated by $\mathbf{u}$. I prefer this right-most form of the projection, as it makes clear what's going on: we form a unit vector $\breve{\mathbf{u}}$ in the direction of $\mathbf{u}$, cast a shadow along this unit vector using the inner product, and then weight the normal vector $\breve{\mathbf{u}}$ by this coefficient. This corresponds to the "shadow" cast by the vector $\mathbf{y}$ onto the direction of vector $\mathbf{u}$.



Then we can write

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\mathbf{z}$ is orthogonal to $\mathbf{u}$. We can rewrite equation (1) as

$$\mathbf{y} = c_1\mathbf{u}_1 + \ldots + c_p\mathbf{u}_p = \text{proj}_{\mathbf{u}_1}\mathbf{y} + \ldots + \text{proj}_{\mathbf{u}_p}\mathbf{y}$$

[1]see http://mathworld.wolfram.com/NormalizedVector.html

which just says that we break vector $\mathbf{y}$ into its components along the orthogonal direction to represent it. This is what we do with our familiar basis of vectors in three-space, $\breve{\mathbf{i}}$, $\breve{\mathbf{j}}$, and $\breve{\mathbf{k}}$.

The fact of the matter is that orthonormal bases are used more often than orthogonal bases, so we generally are working with normalized vectors.

Theorem 6: An m x n matrix U has orthonormal columns if and only if $U^T U = I_{n \text{ x } n}$. $m \geq n$; otherwise dependence.

$u_i^T u_i = u_i \cdot u_i = 1$

$$\begin{bmatrix} u_1^T \\ \vdots \\ u_n^T \end{bmatrix} \underbrace{\begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}}_{m \times n} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ & & \ddots & \end{bmatrix} = I_{n \times n}$$

Theorem 7: Let U be an m x n matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then

(a) $\|\mathbf{Ux}\| = \|\mathbf{x}\|$

(b) $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$

(c) $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.

It is easy enough to verify these properties:

Example: Exercise #25, p. 393

(a) $\|U\underline{x}\| = \|\underline{x}\|$: $\|U\underline{x}\|^2 = (U\underline{x}) \cdot (U\underline{x}) = (U\underline{x})^T U\underline{x} = \underline{x}^T U^T U \underline{x}$
$= \underline{x}^T I \underline{x} = \underline{x}^T \underline{x} = \|x\|^2$ QED.

(b) $(U\underline{x}) \cdot (U\underline{y}) = (U\underline{x})^T U\underline{y} = \underline{x}^T U^T U y = \underline{x}^T \underline{y} = \underline{x} \cdot \underline{y}$

(c) follows immediately from (b).

Definition: *Orthogonal matrix* : a square matrix such that $U^{-1} = U^T$, having orthonormal columns. It's ironic that the name is "orthogonal", rather than "orthonormal". Feel free to call such a matrix an orthonormal matrix.

Curiously enough, orthonormal columns in an orthogonal matrix imply that the rows are also orthonormal, as we see in the following exercise:

Example: #28, p. 393

U - orthogonal, so $U^{-1} = U^T$

$$U^T U = I = U U^T = \boxed{(U^T)^T U^T = I}$$

Invoke Theorem 6, + note that columns of U^T are the rows of U.