MAT225 Section Summary: 6.2

Orthogonal Sets

Summary

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.

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

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

are given by

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

$$\bar{y} = \underbrace{\frac{\bar{y}\cdot\bar{u}_1}{\bar{u}_1\cdot\bar{u}_1}}_{\|u_1\|^2}\bar{u}_1 + \ldots = \left(\bar{y}\cdot\frac{\bar{u}_1}{\|u_1\|}\right)\frac{\bar{u}_1}{\|u_1\|} + \ldots$$

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

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,

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

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

So

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

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

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

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 ordinary basis of vectors $\breve{\mathbf{i}}$, $\breve{\mathbf{j}}$, 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 $U^TU = I$.

$n \leq m$ $\quad [U^T][U] \neq I \quad [U^T][U] = I$

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

1. $\|\mathbf{Ux}\| = \|\mathbf{x}\|$
2. $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ $\qquad (U\bar{x}) \cdot (U\bar{y}) = (U\bar{x})^T(U\bar{y}) = \bar{x}^T U^T U \bar{y} = \bar{x}^T\bar{y} = \bar{x} \cdot \bar{y}$
3. $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.

It is easy to verify these properties. Right?!

Exercise #25, p. 393

a) $\|U\bar{x}\|^2 = (U\bar{x})^T U\bar{x} = \bar{x}^T \underbrace{U^T U}\bar{x} = \bar{x}^T I \bar{x} = \bar{x}^T\bar{x}$

$= \|\bar{x}\|^2 \Rightarrow \|U\bar{x}\| = \|\bar{x}\|$ $\quad$ Q.E.D.

(norms positive)

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:

Example: #28, p. 393

$U_{n \times n}$, orthogonal

Show that rows of U form an orthonormal basis of R^n.

$$\underline{U^T} U = I$$

row space of U is the column space of U^T.

$$\boxed{U \cdot U^T = I \ ?}$$

Greg's conjecture

Because U is square, $U^T = U^{-1}$

Greg's conjecture is true!

$$(U^T)^T U^T = I$$

So U^T is an orthogonal matrix, with n orthonormal columns, a basis of R^n.