Orthogonal Sets
Lay 6.2

1 Orthogonality for sets

We say that a set $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is **orthogonal** if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ for all $i \neq j$.

Example 1.1. Consider the set $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} \subseteq \mathbb{R}^3$, where

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

Then S is orthogonal because each pair of distinct vectors in S is orthogonal. For instance, $\mathbf{u}_1 \cdot \mathbf{u}_3 = 0$ (note that we have a total of 3 combinations to check).

2 Orthogonality and independence

Theorem 2.1. *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. In particular, S is a basis for* $\operatorname{span} S$.

The proof of this theorem is included because it is useful practice for working with orthogonality and the ideas associated to inner products.

Proof. Take a set S as in the statement of the theorem. Assume that $c_1, \ldots, c_p$ are scalars such that

$$c_1\mathbf{u}_1 + \ldots + c_p\mathbf{u}_p = \mathbf{0}. \tag{1}$$

We will show that each c_i is zero, proving the claim. Take the inner product of both sides of (1) with $\mathbf{u}_1$:

$$\begin{aligned}\mathbf{u}_1 \cdot (c_1\mathbf{u}_1 + \ldots + c_p\mathbf{u}_p) &= \mathbf{u}_1 \cdot \mathbf{0} \\ c_1\mathbf{u}_1 \cdot \mathbf{u}_1 + \ldots c_p\mathbf{u}_1 \cdot \mathbf{u}_p &= 0 \\ c_1\|\mathbf{u}_1\|^2 &= 0.\end{aligned}$$

Now, since $\mathbf{u}_1$ is not zero, this implies $c_1 = 0$. Repeating this argument for c_2, c_3, etc. proves the claim. □

Note that orthogonality is a *pairwise* condition, whereas checking linear independence requires we deal with all of S at once. So this is one of the situations where orthogonal vectors make things much easier to handle.

Definition 2.2. An **orthogonal basis** for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

Orthogonal bases are much nicer than non-orthogonal bases because the weights in linear combinations can be computed easily.

Theorem 2.3. *If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for a subspace W of $\mathbb{R}^n$, then for each $\mathbf{y} \in \mathbb{R}^n$, the weights in the linear combination*

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

are given by

$$c_i = \frac{\mathbf{y} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}.$$

The reason this formula holds is similar to the reason the independence theorem holds (try dotting with the basis vectors on both sides; Lay also works the details completely).

Example 2.4. Consider the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ from the first example. This is orthogonal and so, since it consists of three independent vectors, is a basis for $\mathbb{R}^3$. Now consider the vector $\mathbf{y} = \begin{bmatrix} 3 \\ 3 \\ 4 \end{bmatrix}$.

$$\mathbf{y} \cdot \mathbf{u}_1 = 6, \quad \mathbf{y} \cdot \mathbf{u}_2 = 0, \quad \mathbf{y} \cdot \mathbf{u}_3 = 4.$$

Therefore, $\mathbf{y} = 3\mathbf{u}_1 + 4\mathbf{u}_3$ is the unique linear combination that gives $\mathbf{y}$. You can check and see that this works!

3 Orthogonal Projections

We will be treating these in more detail in the next lecture, so the notes here are somewhat brief (see Lay for more detail).

Basically, if $\mathbf{u}$ is a nonzero vector in $\mathbb{R}^n$ and $\mathbf{y}$ is another nonzero vector, we want to decompose $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is a multiple of $\mathbf{u}$ and $\mathbf{z}$ is orthogonal to $\mathbf{u}$. This is the same as projecting onto the line $L = \{\alpha \mathbf{u} : \alpha \in \mathbb{R}\}$, which is the subspace spanned by $\mathbf{u}$.

It turns out that $\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$ (see Figure 2 in Lay and the associated discussion).

We will also frequently write $\text{proj}_L \mathbf{y}$ for $\hat{\mathbf{y}}$ (this weird notation will become clearer).

4 Orthonormal sets

An orthogonal set S is called **orthonormal** if all its vectors are unit vectors. Note that we can always produce an orthonormal set from an orthogonal set by dividing each vector by its norm. Orthonormal sets are useful because $\mathbf{u} \cdot \mathbf{u} = 1$ for each $\mathbf{u}$ in the set, making finding the coefficients in linear combinations easier (note the formula above is simpler if each vector has norm 1).

There is a nice matrix-based way to check orthonormality.

Theorem 4.1. *An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.*

Matrices with orthonormal columns define length-preserving and inner product-preserving linear transformations:

Theorem 4.2. *If U is an $m \times n$ matrix with orthonormal columns and $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^n$,*

- $\|U\mathbf{x}\| = \|x\|$;
- $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$.

Example 4.3. Consider

$$U = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

where θ is some real number. This is the rotation matrix for rotation about the origin (counterclockwise) by the angle θ. Note that

$$U^T U = \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{bmatrix} = I$$

by the trigonometric identity $\cos^2\theta + \sin^2\theta = 1$. Thus, U preserves lengths and inner products. This makes sense, because rotations preserve length and angle between vectors.