LECTURE 24: ORTHOGONALITY AND ISOMETRIES

Orthogonality

Definition 1. If $\vec{v}$ and $\vec{u}$ are vectors in an inner product space V, then $\vec{u}$ and $\vec{v}$ are orthogonal, written $\vec{u} \perp \vec{v}$, if

$$\langle \vec{u}, \vec{v} \rangle = 0.$$

Since V is an inner product space "is orthogonal to" is a symmetric relationship.

If X is a subset of V, then we can define the orthogonal complement of X via

$$X^{\perp} = \{\vec{v} \in V | \langle \vec{v}, x \rangle = 0 \quad \forall x \in X\}.$$

The following two results are clear:

Proposition 1. *For any subset X,*

$$X^{\perp} = \langle X \rangle^{\perp},$$

and

$$\langle X \rangle \subset X^{\perp\perp}.$$

If V is finite dimensional, then the last inclusion is an equality.

Orthogonality provides a way to easily compute inner products. In particular, if $\{\vec{u}_1, \ldots, \vec{u}_n\}$ is a basis such that $\vec{u}_i \perp \vec{u}_j$ for $i \neq j$, then if $\vec{v} = \sum a_i \vec{u}_i$ and $\vec{w} = \sum b_i \vec{u}_i$, then

$$\langle \vec{v}, \vec{w} \rangle = \sum a_i \bar{b}_j \langle \vec{u}_i, \vec{u}_j \rangle = \sum a_i \bar{b}_i \langle \vec{u}_i, \vec{u}_i \rangle.$$

Definition 2. A set X is orthonormal if for all $\vec{u}_i, \vec{u}_j \in X$, $\langle \vec{u}_i, \vec{u}_j \rangle = \delta_{i,j}$.

Thus for an orthonormal basis, we can very easily compute inner products. Moreover, we can easily express any vector as a linear combination of basis vectors.

Proposition 2. *If $\{\vec{u}_1, \ldots, \vec{u}_n\}$ is an orthonormal basis, then*

$$\vec{v} = \sum \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i$$

for all $\vec{v} \in V$.

Proposition 3. *If X is an orthonormal set, then X is linearly independent.*

Proof. Consider a linear dependence relation

$$a_1 \vec{u}_1 + \cdots + a_n \vec{u}_n = \vec{0}.$$

Taking the inner product with $\vec{u}_i$ yields $a_i = 0$, and the result follows immediately. □

A very important result is that given any (linearly independent set) X in an inner product space, we can find an orthogonal (or orthonormal set) U with the same span as X. In particular, given a basis, we can find an orthogonal or orthonormal basis. This is the Grahm-Schmidt orthogonalization process. If $\{\vec{v}_1, \dots, \vec{v}_m\}$ is a linearly independent set, then let $\vec{u}_1 = \vec{v}_1$, and for $i > 1$, let

$$\vec{u}_i = \vec{v}_i - \sum_{j=1}^{i-1} \frac{\langle \vec{v}_i, \vec{u}_j \rangle}{\langle \vec{u}_j, \vec{u}_j \rangle} \vec{u}_j.$$

Proposition 4. *The set $\{\vec{u}_1, \dots, \vec{u}_n\}$ defined above is orthogonal and*

$$\langle \vec{u}_1, \dots, \vec{u}_i \rangle = \langle \vec{v}_1, \dots, \vec{v}_i \rangle$$

for all i.

Proof. The second part is immediate. Each of the vectors $\vec{v}_1$ through $\vec{v}_i$ is a linear combination of the vectors $\vec{u}_1$ through $\vec{u}_i$.

For the first part, we compute directly. Since orthogonality is a symmetric relation, it will suffice to show that $\langle \vec{u}_i, \vec{u}_k \rangle = 0$ for $k < i$:

$$\langle \vec{u}_i, \vec{u}_k \rangle = \langle \vec{v}_i, \vec{u}_k \rangle - \sum_{j=1}^{i-1} \frac{\langle \vec{v}_i, \vec{u}_j \rangle}{\langle \vec{u}_j, \vec{u}_j \rangle} \langle \vec{u}_j, \vec{u}_k \rangle.$$

By induction on i, we see that the sum collapses into a single term:

$$\frac{\langle \vec{v}_i, \vec{u}_k \rangle}{\langle \vec{u}_k, \vec{u}_k \rangle} \langle \vec{u}_k, \vec{u}_k \rangle = \langle \vec{v}_i, \vec{u}_k \rangle,$$

and so $\langle \vec{u}_i, \vec{u}_k \rangle = 0$. $\square$

It is important to realize that this does not depend on the inner product space being $\mathbb{R}^n$ or $\mathbb{C}^n$. This works for any inner product space (since each expression only involves finitely many terms, we can apply this in the infinite dimensional case. We do lose control over the number of terms, however, and the span can change).

At the end of our discussion of isometries, we will see that Grahm-Schmidt actually gives a rather deep connection with matrices. For now, we will set the stage with the following:

Theorem 1. *Choosing an orthonormal basis gives an isomorphism $T \colon \mathbb{F}^n \to V$, where we send the i^{th} standard basis to the i^{th} basis vector in V, and*

$$\langle T\vec{v}, T\vec{u} \rangle = \vec{v} \cdot \vec{u}$$

for all $\vec{v}, \vec{u} \in \mathbb{F}^n$.

Isometries

Let $L \colon V \to W$ be a map of normed spaces.

Definition 3. L is an isometry if

$$||L\vec{v}|| = ||\vec{v}||$$

for all $\vec{v}$.

Thus an isometry preserves distance. In particular, an isometry is necessarily an injection, since if $\vec{v} \in \ker(L)$, then

$$0 = ||L\vec{v}|| = ||\vec{v}||,$$

and so $\vec{v} = 0$. On the other hand, if V and W are not finite dimensional, then an isometry need not be an isomorphism.

Proposition 5. *An isometry is always continuous for the metric topologies.*

Proof. In our ϵ-δ definition, we need only take δ to be ϵ. $\square$

Thus we can explain part of the statement about metric spaces and completions: any metric space admits an isometry to a complete metric space (in fact, we showed that the inclusion of M into our $\bar{M}$ was an isometry, so this is immediate).

For an inner product space, an isometry also preserves the inner product:

$$\langle \vec{v}, \vec{w} \rangle = \langle L\vec{v}, L\vec{w} \rangle.$$

This is because of the polarization identities which relate the inner product to the norms of various terms. We show the real case and leave the complex case as an easy, similar exercise.

Proposition 6. *For any $\vec{v}, \vec{u}$, we have*

$$\langle \vec{u}, \vec{v} \rangle = \frac{1}{4} \left(||\vec{u} + \vec{v}||^2 - ||\vec{u} - \vec{v}||^2 \right).$$

Proof. Expand the right-hand side and use symmetry. $\square$

This means that not only do isometries preserve distance but also they preserve angles (ie they are "conformal"). So all geometry is preserved by an isometry.

Some of the most important isometries are "self-isometries".

Definition 4. Let $O(V)$ be the set of self-isometries of V. This is a group under composition.

Proposition 7. *If $\{\vec{u}_1, \ldots, \vec{u}_n\}$ is an orthonormal set in V and $A \in O(V)$, then*

$$\{A\vec{u}_1, \ldots, A\vec{u}_n\}$$

is an orthonormal set.

Proof. This is immediate from isometries preserving angles. $\square$

Of course, the previous result is also true for isometries in general.

Thus given an ordered orthonormal basis for V, applying A gives another ordered orthonormal basis. Our usual arguments then give us an interpretation of $O(V)$: this is the collection of all "change-of-orthonormal-basis" matrices.

For $V = \mathbb{R}^n$, we have a preferred orthonormal basis: the standard basis. Let $A \in O(\mathbb{R}^n)$. Then since the columns of A are the values of A on an orthonormal basis, they form an orthonormal set.

Proposition 8. *If $A \in O(\mathbb{R}^n)$, then $A^tA = I$, and therefore the inverse to A is A^t.*

Proof. The element in position (i, j) in A^tA is the dot product of the i^{th} column of A with the j^{th} column of A. Since the columns are orthonormal, we conclude that $A^tA = I$.

Since $\mathbb{R}^n$ is finite dimensional, if $BA = I$, then $AB = I$ and B is the inverse to A. We apply this to $B = A^t$. $\square$

Elements of $O(\mathbb{R}^n)$ are called orthogonal matrices. We can restate Grahm-Schmidt in the following way.

Proposition 9. *If $M \in GL_n$ is any matrix, then we can write M as*

$$M = P \cdot R,$$

where $P \in O(\mathbb{R}^n)$ and where R is upper-triangular.

The matrix R essentially records the steps required to pass from the columns of M as a set of vectors to an orthogonal set, and P is the normalized form of the orthogonal vectors. This is the so-called "polar decomposition" of M. Singular matrices also have a polar decomposition, but there is more choice here.

If we work over $\mathbb{C}$, then we have a similar statement. The analogous result for $\mathbb{C}^1$ is that any non-zero complex number z can be written as $z = e^{i\theta}r$, where $r > 0$. The $e^{i\theta}$ part is exactly an element of $O(\mathbb{C})$, and the other part is our "upper triangular matrix".

Why would polar form be nice? Two reasons: first the operation of putting a matrix into polar form is continuous and second, the space of upper-triangular matrices is very simple. Thus we see that the topology of the space of all matrices or all invertible matrices is determined by that of the orthogonal matrices. That's a nice simplification.