MA 1B RECITATION 02/26/15

1. Inner Product Spaces, Norms, Orthogonal/Orthonormal Vectors

Recall that a real vector space V is said to have an inner product $(,): V \times V \to \mathbf{R}$ if the function $(,)$ taking ordered pairs to the scalar field satisfies

- $(x, y) = (y, x)$
- $(x, y + z) = (x, y) + (x, z)$
- $c(x, y) = (cx, y)$
- $(x, x) > 0$ if $x \neq 0$

For complex vector spaces, we usually replace the first condition by Hermitian symmetry:

$$(x, y) = \overline{(y, x)}.$$

where the bar indicates complex conjugation.

Example 1. The standard example is the dot product on $\mathbf{R}^n$:

$$\langle v, w \rangle = v \cdot w = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n.$$

The definition of inner product is just the generalization of the notion of dot product. We'll see some other examples on the homework assignments.

Remark 2. We often use $\langle , \rangle$ to denote inner products instead of $(,)$ to avoid confusion with ordered pairs or tuples.

Remark 3. Apostol uses the physics convention for inner product. Specifically, linearity in the second coordinate of an inner product is the physics convention for Hermitian inner products; it corresponds well to the bra-ket notation that you use in quantum mechanics, for example.

Usually, mathematicians use the opposite convention (linearity in the first coordinate).

Inner products allow you to define a **norm**

$$||x|| = (x, x)^{1/2}$$

which gives us a notion of length, and also allow us to talk about orthogonality. For $\mathbf{R}^n$, this just corresponds to the usual length, i.e.

$$||x|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

We say two vectors $v, w \in V$ are **orthogonal** if

$$\langle v, w \rangle = 0.$$

A subset $S \subset V$ is an **orthogonal set** if for any two distinct elements $v, w \in S$, we have $\langle v, w \rangle = 0$. An orthogonal set is said to be **orthonormal** if $||v|| = 1$ for all $v \in S$.

Date: February 26, 2015.

Here's an example of an orthonormal basis in $\mathbf{R}^4$ with respect to the dot product:

$$\left\{\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right),\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right),\left(-\frac{1}{2},\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)\right\}$$

2. Gram–Schmidt: Finding an orthonormal basis

Let V be an inner product space.

Gram–Schmidt process: takes in a finite linearly independent set $S = \{v_1, \ldots, v_n\}$ in an inner product space and generates an orthogonal set $S' = \{w_1, \ldots, w_n\}$ such that

$$\mathrm{span}(S) = \mathrm{span}(S').$$

How do we do this? Let $\vec{u} \in V$ be any vector. We define the projection operator on V by

$$\mathrm{proj}_{\vec{u}}(\vec{v}) = \frac{\langle \vec{v}, \vec{u}\rangle}{\langle \vec{u}, \vec{u}\rangle}\vec{u}.$$

We call this "the (orthogonal) projection of $\vec{v}$ onto $\vec{u}$." It projects the vector $\vec{v}$ orthogonally onto the line spanned by the vector $\vec{u}$, as you might expect.

We then generates the orthogonal set as follows:

$$\begin{aligned}
w_1 &= v_1 \\
w_2 &= v_2 - \mathrm{proj}_{w_1}(v_2) \\
w_3 &= v_3 - \mathrm{proj}_{w_1}(v_3) - \mathrm{proj}_{w_2}(v_3) \\
&\cdots \\
w_n &= v_n - \sum_{j=1}^{n-1} \mathrm{proj}_{w_j}(v_n).
\end{aligned}$$

Essentially, to get orthogonal vectors, we simply subtract off the parts that would make the new vector non-orthogonal to the old ones.

Remark 1. We can even get *orthonormal* vectors from the Gram–Schmidt process by normalizing the resulting vectors, i.e. taking $e_k = \frac{w_k}{||w_k||}$ as the vectors.

Example 2. Consider the inner product on P_2, the real vector space of (real) polynomials of degree at most 2, given by

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx.$$

What is an orthonormal basis for $(P_2, \langle,\rangle)$?

Recall that $(1, x, x^2)$ is a basis. We can use this to produce an orthonormal basis by the Gram–Schmidt process.

After some work, we get

$$(1, \sqrt{3}(-1+2x), \sqrt{5}(1-6x+6x^2)).$$

3. Orthogonal Projections

The projection operator can be defined for a general subspace as follows. Let U be a subspace of V. We have a decomposition

$$V = U \oplus U^{\perp}$$

where $U^{\perp}$ is the set of vectors orthogonal to U, that is

$$U^{\perp} = \{v \in V \mid \langle v, u\rangle = 0 \text{ for all } u \in U\}.$$

Therefore, we can write any $v \in V$ uniquely as

$$v = u + u'$$

where $u \in U$ and $u' \in U^{\perp}$. Then we have the orthogonal projection of V onto U, denoted by $\operatorname{proj}_U : V \to U$, given by

$$\operatorname{proj}_U(v) = u$$

with the notation as above.

Example 1. Consider the case $V = \mathbf{R}^2$. Let U be the subspace generated by $e_1 = (1, 0)$. Then $U^{\perp}$ is the y-axis, i.e. the subspace generated by $e_2 = (0, 1)$. Consider the point $(3, 2) \in V$. We have

$$(3, 2) = 3e_1 + 2e_2 \in V = U \oplus U^{\perp}$$

and so

$$\operatorname{proj}_U(3, 2) = 3e_1 = (3, 0).$$

Here are some basic facts about the projection operator:

- $\operatorname{proj}_U \in \mathcal{L}(V)$
- $\operatorname{range} \operatorname{proj}_U = U$
- $\operatorname{nullspace} \operatorname{proj}_U = U^{\perp}$
- $v - \operatorname{proj}_U(v) \in U^{\perp}$ for all $v \in V$
- $\operatorname{proj}_U^2 = \operatorname{proj}_U$
- $||\operatorname{proj}_U(v)|| \leq ||v||$ for all $v \in V$

The orthogonal projection allows us to solve the following minimization problem.

Question 2. *Let U be a subspace of V and $v \in V$. What is the "distance" from U to v? What's the closest point in U to v? In other words, what is the point $u \in U$ such that $||v - u||$ is minimized?*

Orthogonal projections give a simple answer to this question.

Proposition 3. *Let U be a subspace of V and $v \in V$. Then*

$$||v - \operatorname{proj}_U(v)|| \leq ||v - u||$$

for any $u \in U$. Moreover, if $u \in U$ and the inequality above is an equality, then $u = \operatorname{proj}_U(v)$.

In other words, $\operatorname{proj}_U(v)$ is always the "closest" point in U to v.

Example 4. We can see this in our simple example above. Is there any point on the x-axis closer to $(3, 2)$ than $(3, 0)$? No. For example, you can see this by forming a right triangle along vertices $(3, 2)$, $(3, 0)$, and any point $(d, 0)$ on the x-axis. The hypotenuse will always be between $(3, 2)$ and $(d, 0)$.

The same idea holds in the general case.

Proposition 5. *If* $(e_1, \ldots, e_m)$ *is an orthonormal basis of* U, *then*

$$\text{proj}_U(v) = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m.$$

Proof. We have $V = U \oplus U^\perp$. For $v \in V$, we have

$$v = (\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m) + v - \langle v, e_1 \rangle e_1 - \cdots - \langle v, e_m \rangle e_m$$

and note that the terms in the parentheses are in U. We claim that the remaining terms

$$w = v - \langle v, e_1 \rangle e_1 - \cdots - \langle v, e_m \rangle e_m$$

are in $U^\perp$. Since $(e_1, \ldots, e_m)$ is an orthonormal set, for all $j = 1, \ldots, m$, we have

$$\begin{aligned} \langle w, e_j \rangle &= \langle v - \langle v_1, e_1 \rangle - \cdots - \langle v, e_m \rangle e_m, e_j \rangle \\ &= \langle v, e_j \rangle - \langle v, e_j \rangle \langle e_j, e_j \rangle \\ &= \langle v, e_j \rangle - \langle v, e_j \rangle \\ &= 0. \end{aligned}$$

Therefore, $w \in U^\perp$. Our formula then follows. □

Example 6. In $\mathbf{R}^4$, let

$$U = \text{span}\{(1, 1, 0, 0), (1, 1, 1, 2)\}.$$

Find $u \in U$ such that $||u - (1, 2, 3, 4)||$ is as small as possible.

Solution. We first find an orthonormal basis of U by applying the Gram–Schmidt procedure to the generating set of U. We obtain

$$\begin{aligned} e_1 &= (1/\sqrt{2}, 1/\sqrt{2}, 0, 0) \\ e_2 &= (0, 0, 1/\sqrt{5}, 2/\sqrt{5}). \end{aligned}$$

Thus, $\{e_1, e_2\}$ is an orthonormal basis of U.

By the two propositions above, the closest point $u \in U$ to $(1, 2, 3, 4)$ is

$$\langle (1, 2, 3, 4), e_1 \rangle e_1 + \langle (1, 2, 3, 4), e_2 \rangle e_2 = (3/2, 3/2, 11/5, 22/5).$$

□