Lecture 2k

Generalized Definitions

(pages 350-352)

All the things that we previously defined with respect to the dot product we can now define in any inner product space.

Definition: Let $\mathbb{V}$ be an inner product space. Then, for any $\mathbf{v} \in \mathbb{V}$, we define the **norm** (or **length**) of $\mathbf{v}$ to be

$$||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Definition: For any vectors $\mathbf{v}, \mathbf{w} \in \mathbb{V}$, the **distance** between $\mathbf{v}$ and $\mathbf{w}$ is $||\mathbf{v} - \mathbf{w}||$.

I prefer to use the term "norm" instead of length, as it helps to emphasize that we may be dealing with something other than the usual length measurement. Unfortunately, there is no other term used to describe the general notion of the distance between two vectors. The best parallel for this situation that I can think of is that you are still measuring the distance between the two vectors, but your units of measurement have changed. If you were measuring the distance between two cities, there are many different quantities that might interest you, other than simply the kilometers between them. Maybe you want to know how many gas stations there are between them, or how many stops on the train there are between them. So just keep in mind that when we now talk about the distance between two vectors, we may not be talking about distance in the traditional sense, but we do still have some sort of distance in mind.

Definition: A vector $\mathbf{v}$ in an inner product space $\mathbb{V}$ is called a **unit vector** if $||\mathbf{v}|| = 1$.

Example: Normalize the polynomial $p(x) = 1 - 3x + x^2$ in P_2 under the inner product $\langle p, q \rangle = 2p(0)q(0) + 3p(1)q(1) + 4p(2)q(2)$.

First we note that $p(0) = 1 - 3(0) + (0)^2 = 1$, $p(1) = 1 - 3(1) + (1)^2 = 1 - 3 + 1 = -1$, and $p(2) = 1 - 3(2) + (2)^2 = 1 - 6 + 4 = -1$. So we have that

$$\begin{aligned} ||p||^2 &= \langle p, p \rangle \\ &= 2p(0)p(0) + 3p(1)p(1) + 4p(2)p(2) \\ &= 2(1)^2 + 3(-1)^2 + 4(-1)^2 \\ &= 9 \end{aligned}$$

And thus, $||p|| = \sqrt{9} = 3$. So $q(x) = (1/3) - x + (1/3)x^2$ is a unit vector that is scalar multiple of p.

Next, we look at the generalized definition of orthogonal.

Definition Let $\mathbb{V}$ be an inner product space with inner product $\langle \; , \; \rangle$. Then two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{V}$ are said to be **orthogonal** if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

Definition The set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ in $\mathbb{V}$ is said to be **orthogonal** if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for all $i \neq j$. The set is said to be **orthonormal** if we also have $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$ for all i.

Example: Consider the inner product space $M(2,2)$ with $\langle A, B \rangle = \mathrm{tr}(B^T A)$. Then the matrices $A = \begin{bmatrix} 1 & -5 \\ -3 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$ are orthogonal, since

$$\left\langle \begin{bmatrix} 1 & -5 \\ -3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} \right\rangle = (1)(1)+(-5)(2)+(-3)(-2)+(1)(3) = 1-10+6+3 = 0$$

Note also that $A = \begin{bmatrix} 1 & -5 \\ -3 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 7 & 0 \\ 2 & -1 \end{bmatrix}$ are orthogonal, since

$$\left\langle \begin{bmatrix} 1 & -5 \\ -3 & 1 \end{bmatrix}, \begin{bmatrix} 7 & 0 \\ 2 & -1 \end{bmatrix} \right\rangle = (1)(7)+(-5)(0)+(-3)(2)+(1)(-1) = 7+0-6-1 = 0$$

and $B = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$ and $C = \begin{bmatrix} 7 & 0 \\ 2 & -1 \end{bmatrix}$ are orthogonal, since

$$\left\langle \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}, \begin{bmatrix} 7 & 0 \\ 2 & -1 \end{bmatrix} \right\rangle = (1)(7)+(2)(0)+(-2)(2)+(3)(-1) = 7+0-4-3 = 0$$

This means that the set $\left\{ \begin{bmatrix} 1 & -5 \\ -3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}, \begin{bmatrix} 7 & 0 \\ 2 & -1 \end{bmatrix} \right\}$ is orthogonal. It is not orthonormal, since

$$\begin{aligned} \langle A, A \rangle &= 1^2 + (-5)^2 + (-3)^2 + 1^2 = 36 \\ \langle B, B \rangle &= 1^2 + 2^2 + (-2)^2 + 3^2 = 18 \\ \langle C, C \rangle &= 7^2 + 0^2 + 2^2 + (-1)^2 = 54 \end{aligned}$$

But this means that the set $\left\{ \frac{1}{6}\begin{bmatrix} 1 & -5 \\ -3 & 1 \end{bmatrix}, \frac{1}{3\sqrt{2}}\begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}, \frac{1}{3\sqrt{6}}\begin{bmatrix} 7 & 0 \\ 2 & -1 \end{bmatrix} \right\}$ is orthonormal.

Example: Consider the inner product space of P_2 with inner product defined by $\langle p, q \rangle = 2p(0)q(0) + 3p(1)q(1) + 4p(2)q(2)$. Then the polynomials $p(x) = 3 + x - 2x^2$ and $q(x) = 2x - 3x^2$ are not orthogonal. Note first that

$$\begin{array}{lll} p(0) = 3 & p(1) = 3 + 1 - 2 = 2 & p(2) = 3 + 2 - 2(4) = -3 \\ q(0) = 0 & q(1) = 2 - 3 = -1 & q(2) = 4 - 3(4) = -8 \end{array}$$

This give us

$$\langle p, q\rangle = 2(3)(0) + 3(2)(-1) + 4(-3)(-8) = 90$$

So $\langle p, q\rangle \neq 0$.

But what if we wanted to find a vector that is orthogonal to $p(x) = 3 + x - 2x^2$? Even more so, we might want a polynomial r that is orthogonal to p and such that $\text{Span}\{p, r\} = \text{Span}\{p, q\}$. Well, that's what our Gram-Schmidt process was able to do in $\mathbb{R}^n$, and we can simply extend that process to general inner product spaces. While we are at it, let's define $\text{proj}_{\mathbb{S}}$ and $\text{perp}_{\mathbb{S}}$ too.

Definition: If $\mathbb{V}$ is an inner product space, and if $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthogonal basis for a subspace $\mathbb{S}$, then for any $\mathbf{x} \in \mathbb{V}$ we have

$$\text{proj}_{\mathbb{S}}\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{v}_1\rangle}{\langle \mathbf{v}_1, \mathbf{v}_1\rangle}\mathbf{v}_1 + \cdots + \frac{\langle \mathbf{x}, \mathbf{v}_k\rangle}{\langle \mathbf{v}_k, \mathbf{v}_k\rangle}\mathbf{v}_k$$

and

$$\text{perp}_{\mathbb{S}}\mathbf{x} = \mathbf{x} - \text{proj}_{\mathbb{S}}\mathbf{x}$$

Example Let P_2 be an inner product space with $\langle p, q\rangle = 2p(0)q(0) + 3p(1)q(1) + 4p(2)q(2)$, and let $\mathbb{S} = \text{Span}\{3 + x - 2x^2\}$. Then

$$\text{proj}_{\mathbb{S}}(2x-3x^2) = (\langle 2x-3x^2, 3+x-2x^2\rangle)/(\langle 3+x-2x^2, 3+x-2x^2\rangle)(3+x-2x^2)$$

If we set $p(x) = 3+x-2x^2$ and $q(x) = 2x-3x^2$ (as in our previous example), then we already know that $\langle 2x - 3x^2, 3 + x - 2x^2\rangle = 90$. And, using the values for p that we calculated in the previous example ($p(0) = 3$, $p(1) = 2$, and $p(2) = -3$), we can compute that $\langle 3 + x - 2x^2, 3 + x - 2x^2\rangle = 2(3)^2 + 3(2)^2 + 4(-3)^2 = 18 + 12 + 36 = 66$. And so we have that

$$\text{proj}_{\mathbb{S}}(2x - 3x^2) = \frac{90}{66}(3 + x - 2x^2) = \frac{45}{11} + \frac{15}{11}x - \frac{30}{11}x^2$$

and

$$\text{perp}_{\mathbb{S}}(2x - 3x^2) = 2x - 3x^2 - \left(\frac{45}{11} + \frac{15}{11}x - \frac{30}{11}x^2\right) = -\frac{45}{11} + \frac{7}{11}x - \frac{3}{11}x^2$$

If we set $r(x) = \text{perp}_{\mathbb{S}}(2x - 3x^2) = -(45/11) + (7/11)x - (3/11)x^2$, and we note that $r(0) = -45/11$, $r(1) = -(45/11) + (7/11) - (3/11) = -41/11$, and $r(2) = -(45/11) + (7/11)(2) - (3/11)(4) = -43/11$, then we see that

$\langle p, r\rangle = 2(3)(-45/11)+3(2)(-41/11)+4(-3)(-43/11) = (1/11)(-270-246+516) = 0$

and so we have $\langle p, \text{perp}_{\mathbb{S}} q\rangle = 0$ as desired.

Moreover, since r is of the form $p+sq$ for some $s \in \mathbb{R}$, we know that $\text{Span}\{p, q\} = \text{Span}\{p, r\}$, and so $\{p, r\}$ is an orthogonal basis for $\text{Span}\{p, q\}$.

This example doubles as an easy example of the generalized Gram-Schmidt Procedure. In general, all of our original arguments for the Gram-Schmidt Procedure translate into the world of inner product spaces. That is, if $\mathbb{V}$ is an inner product space, and if $\mathcal{A} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is a basis for a subspace $\mathbb{S}$ of $\mathbb{V}$, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ defined by

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{w}_1 \\
\mathbf{v}_2 &= \mathbf{w}_2 - \tfrac{\langle \mathbf{w}_2, \mathbf{v}_1\rangle}{\langle \mathbf{v}_1, \mathbf{v}_1\rangle}\mathbf{v}_1 \\
&\vdots \\
\mathbf{v}_n &= \mathbf{w}_n - \tfrac{\langle \mathbf{w}_n, \mathbf{v}_1\rangle}{\langle \mathbf{v}_1, \mathbf{v}_1\rangle}\mathbf{v}_1 - \cdots - \tfrac{\langle \mathbf{w}_n, \mathbf{v}_{n-1}\rangle}{\langle \mathbf{v}_{n-1}, \mathbf{v}_{n-1}\rangle}\mathbf{v}_{n-1}
\end{aligned}$$

is an orthogonal basis for $\mathbb{S}$. Also, as before, we can loosen our conditions and simply have $\mathcal{A}$ be a spanning set, as long as we throw out any zero vectors that are created by our algorithm.

There are two more results about orthogonal and orthonormal sets that we want to translate into this general world of inner product spaces. The first is that any orthogonal set that does not contain the zero vector is linearly independent (and thus all orthonormal sets are linearly independent).

The second is that if $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthogonal basis for an inner product space $\mathbb{V}$, and if $\mathbf{v}$ is any element of $\mathbb{V}$, then the coordinates $\vec{b}$ of $\mathbf{v}$ with respect to $\mathcal{B}$ satisfy

$$b_i = \frac{\langle \mathbf{v}, \mathbf{v}_i\rangle}{\langle \mathbf{v}_i, \mathbf{v}_i\rangle}$$

If $\mathcal{B}$ is an orthonormal basis, then we have $\langle \mathbf{v}_i, \mathbf{v}_i\rangle = 1$, so we simply have that $b_i = \langle \mathbf{v}, \mathbf{v}_i\rangle$.

Example: The set $\mathcal{B} = \left\{ \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} \end{bmatrix} \right\}$ is an orthonormal basis in $M(2,2)$ with the inner product $\langle A, B\rangle = \text{tr}(B^T A)$. Find the coordinates of $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ with respect to $\mathcal{B}$.

We have

$$
\begin{aligned}
b_1 &= \left\langle \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \right\rangle = (1)(1/2) + (2)(1/2) + (0)(1/2) + (1)(1/2) = 2 \\
b_2 &= \left\langle \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix} \right\rangle = (1)(1/2) + (2)(-1/2) + (0)(1/2) + (1)(-1/2) = -1 \\
b_3 &= \left\langle \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \end{bmatrix} \right\rangle = (1)(-1/\sqrt{2}) + (2)(0) + (0)(1/\sqrt{2}) + (1)(0) = -1/\sqrt{2} \\
b_4 &= \left\langle \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} \end{bmatrix} \right\rangle = (1)(0) + (2)(1/\sqrt{2}) + (0)(0) + (1)(-1/\sqrt{2}) = 1/\sqrt{2}
\end{aligned}
$$

So the coordinates of $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ with respect to $\mathcal{B}$ are $\begin{bmatrix} 2 \\ -1 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.