Lecture 2j

Inner Product Spaces

(pages 348-350)

So far, we have taken the essential properties of $\mathbb{R}^n$ and used them to form general vector spaces, and now we have taken the essential properties of the standard basis and created orthonormal bases. But the definition of an orthonormal basis was dependent on the properties of the dot product. But what was so special about the dot product? We take the essential properties of the dot product and use then to define the more general concept of an inner product.

Definition: Let $\mathbb{V}$ be a vector space over $\mathbb{R}$. An **inner product** on $\mathbb{V}$ is a function $\langle\ ,\ \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ such that

(1) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ for all $\mathbf{v} \in \mathbb{V}$, and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$. (positive definite)

(2) $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{V}$ (symmetric)

(3) $\langle \mathbf{v}, s\mathbf{w} + t\mathbf{z} \rangle = s\langle \mathbf{v}, \mathbf{w} \rangle + t\langle \mathbf{v}, \mathbf{z} \rangle$ for all $s, t \in \mathbb{R}$ and $\mathbf{v}, \mathbf{w}, \mathbf{z} \in \mathbb{V}$ (bilinear)

Note that, if we combine properties (2) and (3), then we also have that $\langle s\mathbf{v} + t\mathbf{w}, \mathbf{z} \rangle = s\langle \mathbf{v}, \mathbf{z} \rangle + t\langle \mathbf{w}, \mathbf{z} \rangle$. In its full generality, we have that

$$\begin{aligned} \langle s_1\mathbf{x} + s_1\mathbf{v}, t_1\mathbf{w} + t_2\mathbf{z} \rangle &= s_1\langle \mathbf{x}, t_1\mathbf{w} + t_2\mathbf{z} \rangle + s_2\langle \mathbf{v}, t_1\mathbf{w} + t_2\mathbf{z} \rangle \\ &= s_1t_1\langle \mathbf{x}, \mathbf{w} \rangle + s_1t_2\langle \mathbf{x}, \mathbf{z} \rangle + s_2t_1\langle \mathbf{v}, \mathbf{w} \rangle + s_2t_2\langle \mathbf{v}, \mathbf{z} \rangle \end{aligned}$$

Definition: A vector space $\mathbb{V}$ with an inner product is called an **inner product space**.

Let's start our study of inner product spaces by looking in $\mathbb{R}^n$.

Example: Show that function $\langle\ ,\ \rangle$ defined by

$$\langle \vec{x}, \vec{y} \rangle = x_1y_1 + 4x_2y_2 + 9x_3y_3$$

is an inner product on $\mathbb{R}^3$.

To show this, we verify that $\langle\ ,\ \rangle$ satisfies the three properties of an inner product.

(1) For any $\vec{x} \in \mathbb{R}^3$, $\langle \vec{x}, \vec{x} \rangle = x_1x_1 + 4x_2x_2 + 9x_3x_3 = x_1^2 + 4x_2^2 + 9x_3^2$. Since each of x_1^2, x_2^2, and x_3^2 are greater than or equal to zero, we see that $\langle \vec{x}, \vec{x} \rangle \geq 0$. What if $\langle \vec{x}, \vec{x} \rangle = 0$? Then $x_1^2 + 4x_2^2 + 9x_3^2 = 0$. But since the terms in this sum are all greater than or equal to zero, the only way their sum can equal zero is if each of the terms equals zero. So we get $x_1^2 = 0$, $4x_2^2 = 0$, and $9x_3^2 = 0$. This means that $x_1 = 0$, $x_2 = 0$, and $x_3 = 0$, which means that $\vec{x} = \vec{0}$. And so we see that $\langle\ ,\ \rangle$ is positive definite.

(2) For any $\vec{x}, \vec{y} \in \mathbb{R}^3$, $\langle \vec{x}, \vec{y} \rangle = x_1y_1 + 4x_2y_2 + 9x_3y_3 = y_1x_1 + 4y_2x_2 + 9y_3x_3 = \langle \vec{y}, \vec{x} \rangle$, so $\langle\ ,\ \rangle$ is symmetric.

(3) For any $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ and $s, t \in \mathbb{R}$,

$$\begin{aligned} \langle \vec{x}, s\vec{y} + t\vec{z} \rangle &= x_1(sy_1 + tz_1) + 4x_2(sy_2 + tz_2) + 9x_3(sy_3 + tz_3) \\ &= sx_1y_1 + tx_1z_1 + 4sx_2y_2 + 4tx_2z_2 + 9sx_3y_3 + 9tx_3z_3 \\ &= s(x_1y_1 + 4x_2y_2 + 9x_3y_3) + t(x_1z_1 + 4x_2z_2 + 9x_3z_3) \\ &= s\langle \vec{x}, \vec{y} \rangle + t\langle \vec{x}, \vec{z} \rangle \end{aligned}$$

So $\langle\ ,\ \rangle$ is bilinear.

Example: The function $\langle\ ,\ \rangle$ defined by

$$\langle \vec{x}, \vec{y} \rangle = 2x_1y_1 + 3x_2y_2$$

is not an inner product on $\mathbb{R}^3$ because it is not positive definite. For example, we get that $\left\langle \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \right\rangle = 2(0)(0) + 3(0)(0) = 0$, even though $\begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$. Note that this function does define an inner product on $\mathbb{R}^2$, so this example is intended to show the importance of paying attention which vector space you are working in.

Example: The function $\langle\ ,\ \rangle$ defined by

$$\langle \vec{x}, \vec{y} \rangle = x_1y_1 + x_1y_2 + x_1y_3$$

is not an inner product on $\mathbb{R}^3$ because it is not symmetric. For example, $\left\langle \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \right\rangle = (1)(2) + (1)(3) + (1)(4) = 9$, but $\left\langle \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\rangle = (2)(1) + (2)(2) + (2)(3) = 12$, so $\left\langle \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \right\rangle \neq \left\langle \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\rangle$. (Our function is also not positive definite. It is bilinear.)

The dot product was only defined in $\mathbb{R}^n$, but an inner product can be defined in any vector space. What could be an inner product in a matrix space? Since our result is supposed to be a number, not a vector, matrix multiplication is not a contender. What about the determinant?

Example: The function $\langle A, B\rangle = \det(AB)$ is not an inner product on $M(2,2)$. The easiest thing to see is that it is not positive definite, since $\det\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = \det\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 0$, so $\left\langle \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle = 0$ even though $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. This counterexample easily extends to show that $\langle A, B\rangle = \det(AB)$ is not an inner product on $M(n,n)$ for any n. It is also worth noting that the determinant is not bilinear, since in general $\det(sA) = s^n \det A \neq s \det A$. So, while the determinant is useful for many things, it is not useful for defining an inner product on matrix spaces.

There is another operation associated with matrices that results in a number, though, and that is the trace. Recall that the trace of a matrix is the sum of the diagonal entries.

Example: The function $\langle A, B\rangle = tr(B^T A)$ is an inner product on $M(m,n)$. To see this, let's first look at this definition in $M(2,3)$:

$$\begin{aligned}
&\left\langle \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}, \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \right\rangle \\
&= \mathrm{tr}\left(\begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \\ b_{13} & b_{23} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \right) \\
&= \mathrm{tr}\begin{bmatrix} b_{11}a_{11} + b_{21}a_{21} & ? & ? \\ ? & b_{12}a_{12} + b_{22}a_{22} & ? \\ ? & ? & b_{13}a_{13} + b_{23}a_{23} \end{bmatrix} \\
&= b_{11}a_{11} + b_{21}a_{21} + b_{12}a_{12} + b_{22}a_{22} + b_{13}a_{13} + b_{23}a_{23}
\end{aligned}$$

Note that I left some entries in the matrix $B^T A$ blank because it was not necessary to calculate them in order to compute the trace. In fact, the entries on the diagonal are $\begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} \cdot \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$, $\begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix} \cdot \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$, and $\begin{bmatrix} b_{13} \\ b_{23} \end{bmatrix} \cdot \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}$. That is, the i-th diagonal entry is the dot product of the i-th column of B with the i-th column of A. As you may now be guessing, this will be true in our general case. Consider that our definition for matrix multiplication was that the ij-th entry of BA is the dot product of the i-th row of B with the j-th column of A. But our "B" is currently B^T, and the i-th row of B^T is the same as the i-th column of B, so the ij-th entry of $B^T A$ is the dot product of the i-th column of B and the j-th column of A. Since we are looking for the trace of $B^T A$, we need to take the sum of the diagonal entries (the ii entries), which will be the dot product of the i-th column of B with the i-th column of A. So, if $A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}$ and $B = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_n \end{bmatrix}$, then

$$\langle A, B \rangle = \operatorname{tr}(B^T A) = \sum_{i=1}^{n} \vec{b}_i \cdot \vec{a}_i$$

With this formula in hand, let's show that $\langle\ ,\ \rangle$ is an inner product.

(1) $\langle A, A \rangle = \sum_{i=1}^{n} \vec{a}_i \cdot \vec{a}_i$. Since $\vec{a}_i \cdot \vec{a}_i \geq 0$ for all i (as the dot product is positive definite), we know that $\sum_{i=1}^{n} \vec{a}_i \cdot \vec{a}_i \geq 0$. Moreover, if $\sum_{i=1}^{n} \vec{a}_i \cdot \vec{a}_i = 0$, then we need each $\vec{a}_i \cdot \vec{a}_i = 0$, and again using the fact that the dot product is positive definite, this means that $\vec{a}_i = 0$ for all i. And since the columns of A are all the zero vector, we get that A is the zero matrix. As such, we have shown that $\langle\ ,\ \rangle$ is positive definite.

(2) We have already shown that $\langle A, B \rangle = \sum_{i=1}^{n} \vec{b}_i \cdot \vec{a}_i$, so now we want to look at $\langle B, A \rangle$. By the same arguments we used for $\langle A, B \rangle$, we know that the ij-th entry of $A^T B$ is the dot product of the i-th column of A with the j-th column of B. This means that $\operatorname{tr}(A^T B) = \sum_{i=1}^{n} \vec{a}_i \cdot \vec{b}_i$. Since dot products are symmetrical, we know that $\vec{a}_i \cdot \vec{b}_i = \vec{b}_i \cdot \vec{a}_i$. And thus we have the following:

$$\begin{aligned} \langle B, A \rangle &= \operatorname{tr}(A^T B) \\ &= \textstyle\sum_{i=1}^{n} \vec{a}_i \cdot \vec{b}_i \\ &= \textstyle\sum_{i=1}^{n} \vec{b}_i \cdot \vec{a}_i \\ &= \langle A, B \rangle \end{aligned}$$

$$\text{(3)} \quad \begin{aligned} \langle A, sB + tC \rangle &= \textstyle\sum_{i=1}^{n} (s\vec{b}_i + t\vec{c}_i) \cdot \vec{a}_i \\ &= \textstyle\sum_{i=1}^{n} s\vec{b}_i \cdot \vec{a}_i + t\vec{c}_i \cdot \vec{a}_i \\ &= s\textstyle\sum_{i=1}^{n} \vec{b}_i \cdot \vec{a}_i + t\textstyle\sum_{i=1}^{n} \vec{c}_i \cdot \vec{a}_i \\ &= s\langle A, B \rangle + t\langle A, C \rangle \end{aligned}$$

Before we leave this example, let's take a further look at the formula $\langle A, B \rangle = \operatorname{tr}(B^T A) = \sum_{i=1}^{n} \vec{b}_i \cdot \vec{a}_i$. Because the dot product $\vec{b}_i \cdot \vec{a}_i$ is the sum $b_{i1}a_{i1} + \cdots + b_{im}a_{im}$. But since we cycle through all possible i values when we take the inner product, what we end up with is that $\langle A, B \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}a_{ij}$. Don't let the double sum scare you! The important fact is what is inside: that we are taking the product of the corresponding entries of A and B, and adding them together. This formula is quite reminiscent of our formula for the dot product in $\mathbb{R}^n$. (And, in fact, under the standard isomorphism from $M(m, n)$ to $\mathbb{R}^{mn}$, this is the same as the dot product.) So whenever we need to calculate $\operatorname{tr}(B^T A)$, we know it is $\sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}a_{ij}$. So, for example,

$\left\langle \begin{bmatrix} 2 & 1 \\ 3 & -5 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -3 & -2 \end{bmatrix} \right\rangle = (2)(1) + (1)(0) + (3)(-3) + (-5)(-2) = 2 + 0 - 9 + 10 = 3.$

Isn't that easier than multiplying the matrices? Note that the double sum would technically have us multiply down a column and then move to the next column and to reverse the order of the A and B terms. But thanks to the commutivity of addition and multiplication in the reals, we can multiply our pairs in any order we like, so long as we remember to do every pair. I choose to multiply across the rows simply because I frequently think of a matrix in terms of reading the entries along the rows first. (For example, the standard basis for $M(m,n)$ finds the solitary 1 traveling across rows before going down to the next column.) You can use whatever pattern feels natural to you.

Now let's look at some examples of the inner product in polynomial spaces. Again, we need to find an operation on polynomials that results in a number, not a polynomial. Most of the things we know to do with polynomials from outside linear algebra (multiplying, factoring, taking derivatives) still result in polynomials. But there is one thing we do with polynomials that results in a number–plug in a number!

Example: Show that $\langle p,q\rangle = 2p(0)q(0)+3p(1)q(1)+4p(2)q(2)$ defines an inner product on P_2.

(1) We have that $\langle p,p\rangle = 2(p(0))^2+3(p(1))^2+4(p(2))^2$ is clearly greater than or equal to zero, since all the terms in the sum are positive. Moreover, the only way to have $2(p(0))^2+3(p(1))^2+4(p(2))^2 = 0$ is to have $p(0)=0$, $p(1)=0$, and $p(2)=0$. If $p(x)=p_0+p_1x+p_2x^2$, then, $p(0)=0$ means $p_0+p_1(0)+p_2(0)=0$, so $p_0=0$. And $p(1)=0$ means $p_0+p_1(1)+p_2(1^2)=0$, which means that $p_1+p_2=0$, or that $p_1=-p_2$. Lastly, $p(2)=0$ means that $p_0+p_1(2)+p_2(2^2)=0$, so $2(-p_2)+4p_2=0$. From this we see that $p_2=0$, which means that $p_1=0$ also. So we have shown that if $\langle p,p\rangle=0$, then $p_0=p_1=p_2=0$, so p is the zero polynomial. As such, we see that $\langle\ ,\ \rangle$ is positive definite.

(2) $\langle p,q\rangle = 2p(0)q(0)+3p(1)q(1)+4p(2)q(2) = 2q(0)p(0)+3q(1)p(1)+4q(2)p(2) = \langle q,p\rangle$, so $\langle\ ,\ \rangle$ is symmetric.

(3) For any $p,q,r\in P_2$ and $s,t\in\mathbb{R}$, we have the following:

$$\begin{aligned}
\langle p, sq+tr\rangle &= 2p(0)(sq+tr)(0)+3p(1)(sq+tr)(1)+4p(2)(sq+tr)(2)\\
&= 2p(0)(sq(0)+tr(0))+3p(1)(sq(1)+tr(1))+4p(2)(sq(2)+tr(2))\\
&= 2sp(0)q(0)+2tp(0)r(0)+3sp(1)q(1)+3tp(1)r(1)+4sp(2)q(2)+4tp(2)r(2)\\
&= 2sp(0)q(0)+3sp(1)q(1)+4sp(2)q(2)+2tp(0)r(0)+3tp(1)r(1)+4tp(2)r(2)\\
&= s\langle p,q\rangle + t\langle p,r\rangle
\end{aligned}$$

So we see that $\langle\ ,\ \rangle$ is bilinear.

Example: The function $\langle p,q\rangle = 2p(0)q(0)+3p(1)q(1)+4p(2)q(2)$ is also an inner product on P_1, but is not an inner product on P_3, or P_n for any $n>2$. The reason it isn't an inner product in the larger polynomial spaces is that it fails to be positive definite. Consider for example that if $p(x)=2x-3x^2+x^3$, then $p(0)=2(0)-3(0)+0=0$, $p(1)=2(1)-3(1)+(1)=0$, and $p(2)=2(2)-3(4)+(8)=0$, so $\langle p,p\rangle=0$ even though p is not the zero polynomial.

How did I find $p(x)$? Well, I knew I needed $p(0) = 0$, $p(1) = 0$, and $p(2) = 0$. In P_3, that forms the following system of equations:

$$\begin{aligned} p_0 &= 0 \\ p_0 + p_1 + p_2 + p_3 &= 0 \\ p_0 + 2p_1 + 4p_2 + 8p_3 &= 0 \end{aligned}$$

Going ahead and plugging in $p_0 = 0$, we focus our attention on finding a solution to the system

$$\begin{aligned} p_1 + p_2 + p_3 &= 0 \\ 2p_1 + 4p_2 + 8p_3 &= 0 \end{aligned}$$

We solve this system by row reducing its coefficient matrix as follows:

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \end{bmatrix}$$

So we see that this system has non-trivial solutions. The general solution is

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = t \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$$

I set $t = 1$ to get the solution $p_1 = 2$, $p_2 = -3$, and $p_3 = 1$, which is the polynomial $2x - 3x^2 + x^3$.