Lecture 8. Inner Product Spaces
October 2022

Length and orthogonality are basic concepts in geometry and, in $\mathbb{R}^2$ and $\mathbb{R}^3$, they both can be defined using the dot product. In this lecture we extend the dot product to vectors in $\mathbb{R}^n$, and so endow $\mathbb{R}^n$ with Euclidean geometry. We then introduce the idea of an orthogonal basis — one of the most useful concepts in linear algebra.

Inner Product and Gram Matrix

DEFINITION 8.1 An **inner product** on a real vector space V is a function $\langle , \rangle : V \times V \to \mathbb{R}$ that assigns a real number $\langle \mathbf{x}, \mathbf{y} \rangle$ to every pair $\mathbf{x}$, $\mathbf{y}$ of vectors in V in such a way that the following axioms are satisfied.

P1. $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ for all $\mathbf{x} \neq \mathbf{0}$ (**positive definiteness**)
P2. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ for all $\mathbf{x}, \mathbf{y} \in V$ (**symmetry**)
P3. $\langle r\mathbf{x} + s\mathbf{y}, \mathbf{z} \rangle = r\langle \mathbf{x}, \mathbf{z} \rangle + s\langle \mathbf{y}, \mathbf{z} \rangle$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $r, s \in \mathbb{R}$ (**linearity in the first coordinate**)

A real vector space V with an inner product $\langle , \rangle$ will be called an **inner product space**. Note that every subspace of an inner product space is again an inner product space using the same inner product.

EXAMPLE 8.2 $\mathbb{R}^n$ is an inner product space with the dot product as inner product:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n,$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ are two vectors in $\mathbb{R}^n$.

This is also called the **Euclidean** inner product, and $\mathbb{R}^n$ equipped with the dot product, is called **Euclidean n-space**. Observe that if $\mathbf{x}$ and $\mathbf{y}$ are written as columns then $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}$ is a matrix product.

EXAMPLE 8.3 If A and B are $m \times n$ matrices, define $\langle A, B \rangle = \operatorname{tr}(AB^T)$, where $\operatorname{tr}(X)$ is the trace of the square matrix X. Show that $\langle , \rangle$ is an inner product in M_{mn}.

Solution. Since $\operatorname{tr}(P) = \operatorname{tr}(P^T)$ for every $m \times m$ matrix P, we have P2. Next, P3 follows because trace is a linear transformation $M_{m\times m} \to \mathbb{R}$. Turning to P1, let $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ denote the rows of the matrix A. Then the (i, j)-entry of AA^T is $\mathbf{r}_i \cdot \mathbf{r}_j$, so

$$\langle A, A \rangle = \operatorname{tr}(AA^T) = \mathbf{r}_1 \cdot \mathbf{r}_1 + \mathbf{r}_2 \cdot \mathbf{r}_2 + \cdots + \mathbf{r}_m \cdot \mathbf{r}_m$$

But $\mathbf{r}_j \cdot \mathbf{r}_j$ is the sum of the squares of the entries of $\mathbf{r}_j$, so this shows that $\langle A, A \rangle$ is the sum of the squares of all entries of A. Axiom P1 follows.

EXAMPLE 8.4 Let $C[a, b]$ denote the vector space of continuous functions from $[a, b]$ to $\mathbb{R}$. Show that

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,dx$$

defines an inner product on $C[a, b]$.

Solution. Axiom P2 is clear. Axiom P3 follows from linearity of definite integral. Finally, theorems of calculus show that $\int_a^b f^2\,dx \geqslant 0$ and, if f is continuous, that this is zero if and only if f is the zero function. This gives axiom P1.

Given an inner product $\langle , \rangle$ on a finite-dimensional vector space V, let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be a basis of V. If $\mathbf{x} = \sum\limits_{i=1}^{n} x_i\mathbf{e}_i$ and $\mathbf{y} = \sum\limits_{j=1}^{n} y_j\mathbf{e}_j$ are two vectors in V, compute $\langle \mathbf{x}, \mathbf{y}\rangle$ by adding the inner product of each term $x_i\mathbf{e}_i$ to each term $y_j\mathbf{e}_j$. The result is a double sum

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\langle x_i\mathbf{e}_i, y_j\mathbf{e}_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\langle \mathbf{e}_i, \mathbf{e}_j\rangle y_j.$$

As the reader can verify, this is a matrix product:

$$\langle \mathbf{x}, \mathbf{y}\rangle = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} \langle \mathbf{e}_1, \mathbf{e}_1\rangle & \langle \mathbf{e}_1, \mathbf{e}_2\rangle & \dots & \langle \mathbf{e}_1, \mathbf{e}_n\rangle \\ \langle \mathbf{e}_2, \mathbf{e}_1\rangle & \langle \mathbf{e}_2, \mathbf{e}_2\rangle & \dots & \langle \mathbf{e}_2, \mathbf{e}_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1\rangle & \langle \mathbf{e}_n, \mathbf{e}_2\rangle & \dots & \langle \mathbf{e}_n, \mathbf{e}_n\rangle \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \mathbf{x}^T G\mathbf{y}.$$

DEFINITION 8.5 Given a finite-dimensional Euclidean vector space V with a basis $B = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$, the matrix

$$G = \begin{pmatrix} \langle \mathbf{e}_1, \mathbf{e}_1\rangle & \langle \mathbf{e}_1, \mathbf{e}_2\rangle & \dots & \langle \mathbf{e}_1, \mathbf{e}_n\rangle \\ \langle \mathbf{e}_2, \mathbf{e}_1\rangle & \langle \mathbf{e}_2, \mathbf{e}_2\rangle & \dots & \langle \mathbf{e}_2, \mathbf{e}_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1\rangle & \langle \mathbf{e}_n, \mathbf{e}_2\rangle & \dots & \langle \mathbf{e}_n, \mathbf{e}_n\rangle \end{pmatrix}$$

is called the **Gram matrix** in V.

The fact that $\langle \mathbf{e}_i, \mathbf{e}_j\rangle = \langle \mathbf{e}_j, \mathbf{e}_i\rangle$ shows that G is symmetric. In particular, the dot product corresponds to the identity matrix I_n.

Norm and Distance

DEFINITION 8.6 If $\langle , \rangle$ is an inner product on a space V, the **norm** $\|\mathbf{x}\|$ of a vector in V is defined by

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}.$$

We define the **distance** between vectors **v** and **w** in an inner product space V to be

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

A vector $\mathbf{v} \in \mathbf{V}$ is called a **unit vector** if $\|\mathbf{x}\| = 1$. If $\mathbf{x} \neq \mathbf{0}$, then $\|\mathbf{x}\| \neq 0$ and it follows easily that $\frac{1}{\|\mathbf{x}\|}\mathbf{x}$ is a unit vector, a fact that we shall use later.

The next theorem reveals an important and useful fact about the relationship between norms and inner products.

THEOREM 8.7 (Cauchy–Bunyakovsky–Schwarz Inequality) If **x** and **y** are vectors in V, then

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \leqslant \|\mathbf{x}\|\|\mathbf{y}\|.$$

Moreover $|\langle \mathbf{x}, \mathbf{y}\rangle| = \|\mathbf{x}\|\|\mathbf{y}\|$ if and only if one of $\mathbf{x}$ and $\mathbf{y}$ is a multiple of the other.

Proof. The inequality holds if $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$ (in fact it is equality). Otherwise, let $a = \|\mathbf{x}\|^2 > 0$, $b = 2\langle \mathbf{x}, \mathbf{y}\rangle$, and $c = \|\mathbf{y}\|^2$ for convenience, and let t be any real number. Since the inner product of any vector with itself is nonnegative, it follows that

$$0 \leqslant \langle t\mathbf{x} + \mathbf{y}, t\mathbf{x} + \mathbf{y}\rangle = \|\mathbf{x}\|^2 t^2 + 2\langle \mathbf{x}, \mathbf{y}\rangle t + \|\mathbf{y}\|^2 = at^2 + bt + c$$

This inequality implies that the quadratic polynomial $at^2 + bt + c$ has either no real roots or a repeated real root. Therefore, its discriminant must satisfy the inequality $b^2 - 4ac \leqslant 0$. Expressing the coefficients a, b, and c in terms of the vectors $\mathbf{x}$ and $\mathbf{y}$ gives

$$4\langle \mathbf{x}, \mathbf{y}\rangle^2 - 4\|\mathbf{x}\|^2\|\mathbf{y}\|^2 \leqslant 0$$

or, equivalently,

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \leqslant \|\mathbf{x}\|\|\mathbf{y}\|.$$

which completes the proof. $\square$

EXAMPLE 8.8 If f and g are continuous functions on the interval $[a, b]$, then

$$\left(\int_a^b f(x)g(x)\,dx\right)^2 \leqslant \int_a^b f(x)^2\,dx \int_a^b g(x)^2\,dx.$$

There is an important consequence of the CBS inequality.

COROLLARY 8.9 (Triangle Inequality) If $\mathbf{x}$ and $\mathbf{y}$ are vectors in V, then

$$\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|.$$

REMARK 8.10 The reason for the name comes from the observation that in $\mathbb{R}^3$ the inequality asserts that the sum of the lengths of two sides of a triangle is not less than the length of the third side.

Proof.

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2 \leqslant \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2 \quad \square$$

DEFINITION 8.11 Let V be a vector space. A function $\|\cdot\| : V \to \mathbb{R}_0^+$ is said to be a **norm** in V if:

1. $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$ (positivity);
2. $\|\lambda\mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$ for every $\lambda \in \mathbb{R}$ and $\mathbf{x} \in V$ (scaling property);
3. $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$ for every $\mathbf{x}$, $\mathbf{y} \in V$ (triangle inequality).

We then say that the pair $(V, \|\cdot\|)$ is a **normed space**.

DEFINITION 8.12 Let M be a set. A function $d : M \times M \to \mathbb{R}_0^+$ is said to be a **distance** in M if:

1. $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$;
2. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for every $\mathbf{x}, \mathbf{y} \in M$ (**symmetry**);

3. $d(\mathbf{x},\mathbf{z}) \leqslant d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z})$ for every $\mathbf{x}$, $\mathbf{y},\mathbf{z} \in M$ (**triangle inequality**).

We then say that the pair (M, d) is a **metric space**.

THEOREM 8.13

- If V is an inner product space, then it is a normed space with the norm defined by

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}, \qquad \forall \mathbf{x} \in V.$$

- If V is a normed space, then it is a metric space with the distance defined by

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|, \qquad \forall \mathbf{x},\mathbf{y} \in V.$$