Math 425 Applied & Comput. Lin. Algebra Fall 2024 Lecture

Inner Products and Orthogonal Bases

Inner Products. In this section we will try to make sense of "length" in certain vector spaces. We start with the observation that the dot product of two vectors in $\mathbb{R}^n$ is really a (matrix) product of a row vector with a column vector:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = (v_1\, v_2\, \ldots\, v_n) \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \mathbf{v}^T \mathbf{w}.$$

The connection to the length or magnitude of the vector is clear: $\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. We will call this quantity the *Euclidean norm* and we will denote it by $||\mathbf{v}||$. So

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

Because of the squares of the components of $\mathbf{v}$ above we see that the only way for a vector to have zero norm is if it is the zero vector. If $\mathbf{v} \neq \mathbf{0}$ then $||\mathbf{v}|| > 0$.

Example 1. If $\mathbf{v} = (1\, 3\, -2\, 0)^T$, then $||\mathbf{v}|| = \sqrt{1^2 + 3^2 + (-2)^2 + 0^2} = \sqrt{14}$.

Definition 1. An *inner product* on a real vector space V is a pairing that takes two vectors $\mathbf{v}, \mathbf{w} \in V$ and produces a real number, denoted by $\langle \mathbf{v}, \mathbf{w} \rangle$, which satisfies the following three axioms for every $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and scalars $c, d \in \mathbb{R}$.

- *Bilinearity*: $\langle c\mathbf{u} + d\mathbf{v}, \mathbf{w} \rangle = c\langle \mathbf{u}, \mathbf{w} \rangle + d\langle \mathbf{v}, \mathbf{w} \rangle$ and $\langle \mathbf{u}, c\mathbf{v} + d\mathbf{w} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle + d\langle \mathbf{u}, \mathbf{w} \rangle$
- *Symmetry*: $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
- *Positivity*: $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ whenever $\mathbf{v} \neq \mathbf{0}$, and if $\mathbf{v} = \mathbf{0}$ then $\langle \mathbf{v}, \mathbf{v} \rangle = 0$

We will write $||\mathbf{v}||$ for $\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ and will call it the *norm* with respect to the inner product. A vector space equipped with an inner product is called an *inner product space*.

Orthogonal and Orthonormal Bases. In this section we start out with an inner product space V. We have already encountered the word *orthogonal*. In V, two vectors $\mathbf{v}$ and $\mathbf{w}$ are called orthogonal (with respect to the inner product) if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

Definition 2. A basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V is called *orthogonal* if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for all $i \neq j$. The basis is called *orthonormal* if, in addition, $||\mathbf{v}_i|| = 1$ for all $i = 1, \ldots, n$.

Example 2. With respect to the Euclidean inner product in $\mathbb{R}^n$ the standard basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is an orthonormal basis: $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$ for $i \neq j$, and $\langle \mathbf{e}_i, \mathbf{e}_i \rangle = 1$.

A little remark: If we have an orthogonal basis of V, it is easy to turn it into an orthonormal basis: just divide each $\mathbf{v}_i$ by its norm $||\mathbf{v}_i||$.

Example 3.

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \; \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \; \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

form an orthogonal basis of $\mathbb{R}^3$. And since $||\mathbf{v}_1|| = \sqrt{3}$, $||\mathbf{v}_2|| = \sqrt{2}$, and $||\mathbf{v}_3|| = \sqrt{6}$, the following is an orthonormal basis of $\mathbb{R}^3$:

$$\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

The next two results are the most important things you need to know about orthogonal/orthonormal bases. The first one says that if a set of nonzero vectors are mutually orthogonal then they are automatically linearly independent. The second one gives an efficient way of computing coordinates with respect to an orthogonal/orthonormal basis.

Proposition 1. Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be nonzero vectors in V such that $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for all $i \neq j$. Then these vectors are linearly independent.

Proof. Let $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$. When we compute $0 = \langle \mathbf{0}, \mathbf{v}_j \rangle = \langle \sum_{i=1}^k c_i \mathbf{v_i}, \mathbf{v_j} \rangle$, using bilinearity we get $\sum_{i=1}^k c_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle$. Now because of the orthogonality hypothesis this sum is equal to $c_j \langle \mathbf{v}_j, \mathbf{v}_j \rangle = 0$, and this implies $c_j = 0$ since $\langle \mathbf{v}_j, \mathbf{v}_j \rangle \neq 0$. □

Theorem 1. Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be an orthogonal basis of V. Then the coordinates of any vector $\mathbf{v}$

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

are computed by

$$c_i = \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{||\mathbf{v}_i||^2}.$$

In case the basis is an orthonormal basis one has $c_i = \langle \mathbf{v}, \mathbf{v}_i \rangle$.

Proof. As we have done in the proof of the above proposition we compute $\langle \mathbf{v}, \mathbf{v}_i \rangle = \langle \sum_{j=1}^n c_j \mathbf{v}_j, \mathbf{v}_i \rangle = c_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle$. From here we see that $c_i = \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{||\mathbf{v}_i||^2}$. If the basis is orthonormal then $||\mathbf{v}_i||^2 = 1$, and the result follows. □

Example 4. Lets take

$$\mathbf{v}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{v}_3 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

as an orthonormal basis of $\mathbb{R}^3$. Then the coordinates of $\mathbf{v} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ are $c_1 = \langle \mathbf{v}, \mathbf{v}_1 \rangle = \frac{2}{\sqrt{3}}$, $c_2 = \langle \mathbf{v}, \mathbf{v}_2 \rangle = \frac{3}{\sqrt{2}}$, and $c_3 = \langle \mathbf{v}, \mathbf{v}_3 \rangle = -\frac{5}{\sqrt{6}}$.