6 Orthogonality and Least Squares

6.1 Inner Product, Length, and Orthogonality

Recall that we may think of vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ as arrows with initial point at the origin. The length of a vector $\mathbf{x}$ in $\mathbb{R}^2$ or $\mathbb{R}^3$ is called the *norm* of $\mathbf{x}$, denoted $\|\mathbf{x}\|$. Thus for $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, we have $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}$. Similarly, for $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$, we have $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$. Even though we cannot draw pictures in higher dimensions, the generalization to $\mathbb{R}^n$ is obvious; we define the norm of $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ by

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

The norm is not linear on $\mathbb{R}^n$. To inject linearity into the discussion, we introduce the dot product.

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, then we regard $\mathbf{u}$ and $\mathbf{v}$ as $n \times 1$ matrices. The transpose of $\mathbf{u}$ is a $1 \times n$ matrix $\mathbf{u}^T$. Then the product $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix, which we write as a single number (a scalar) without brackets.

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, the *dot product* of $\mathbf{x}$ and $\mathbf{y}$, denoted $\mathbf{x} \cdot \mathbf{y}$, is defined by

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n,$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$. Note that the dot product of two vectors in $\mathbb{R}^n$ is a number, not a vector. Obviously, $\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2$ for all $\mathbf{x} \in \mathbb{R}^n$. In particular, $\mathbf{x} \cdot \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$, with equality if and only if $\mathbf{x} = \mathbf{0}$. Also, if $\mathbf{y} \in \mathbb{R}^n$ is fixed, then clearly tha map from $\mathbb{R}^n$ to $\mathbb{R}$ that sends $\mathbf{x} \in \mathbb{R}^n$ to $\mathbf{xy}$ is linear.

An inner product is a generalization of the dot product. An *inner product* on V is a function that takes each ordered pair $(\mathbf{u}, \mathbf{v})$ of elements in V to a number $\langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{R}$ and has the following properties (presented in the textbook as Theorem 1):

positivity $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ for all $\mathbf{v} \in V$

definiteness $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Longleftrightarrow \mathbf{v} = \mathbf{0}$

additivity $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$

homogeneity $\langle a\mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{v}, \mathbf{w} \rangle$ for all $a \in \mathbb{R}$ and all $\mathbf{v}, \mathbf{w} \in V$

symmetry $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ for all $\mathbf{v}, \mathbf{w} \in V$

An *inner product space* is a vector space V along with an inner product on V.

The most important example of an inner product space is $\mathbb{R}^n$, which is the dot product. As another example of an inner product space, consider the vector space $\mathbb{P}_m$ of all polynomials with coefficients in $\mathbb{R}$ and degree at most m. We can define an inner product on $\mathbb{P}_m$ by

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

In the definition of an inner product, the conditions of additivity and homogeneity can be combined into a requirement of linearity. More precisely, for each fixed $\mathbf{w} \in V$, the function that takes $\mathbf{v}$ to $\langle \mathbf{v}, \mathbf{w} \rangle$ is a linear map from V to $\mathbb{R}$. Becauce every linear map takes $\mathbf{0}$ to 0, we must have

$$\langle \mathbf{0}, \mathbf{w} \rangle = 0$$

for every $\mathbf{w} \in V$. Thus by the symmetry property we also have

$$\langle \mathbf{w}, \mathbf{0} \rangle = 0.$$

For $\mathbf{v} \in V$, we define the *norm* of $\mathbf{v}$, denoted $\|\mathbf{v}\|$, by

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

An example is the norm defined for $\mathbb{R}^n$ at the beginning of this section. As another example, if $\mathbf{p} \in \mathbb{P}_m$ with inner product given by (1), then

$$\|\mathbf{p}\| = \sqrt{\int_0^1 |\mathbf{p}(x)|^2 dx}.$$

Note that $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$, because $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$. Another property of norm is that $\|a\mathbf{v}\| = |a|\|\mathbf{v}\|$ for all $a \in \mathbb{R}$ and all $\mathbf{v} \in V$. Here's the proof:

$$\begin{aligned} \|a\mathbf{v}\|^2 &= \langle a\mathbf{v}, a\mathbf{v} \rangle \\ &= a\langle \mathbf{v}, a\mathbf{v} \rangle \\ &= aa\langle \mathbf{v}, \mathbf{v} \rangle \\ &= |a|^2\|\mathbf{v}\|^2; \end{aligned}$$

taking square roots now gives the desired equality. This proof illustrates a general principle: working with norms squared is usually easier than working directly with norms.

Two vectors $\mathbf{u}, \mathbf{v} \in V$ are said to be *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Note that the order of the vectors does not matter because of symmetry. Clearly $\mathbf{0}$ is orthogonal to every vector. Furthermore, $\mathbf{0}$ is the only vector that is orthogonal to itself.

The next theorem is over 2,500 years old.

Theorem 2 (Pythagorean Theorem)**.** *If $\mathbf{u}, \mathbf{v}$ are orthogonal vectors in V, then*

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2. \tag{2}$$

Proof. Suppose that $\mathbf{u}, \mathbf{v}$ are orthogonal vectors in V. Then

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2, \end{aligned}$$

as desired. □

Note that the converse of Pythagorean Theorem holds in real inner-product spaces, but not in complex inner-product spaces, where $\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle = 2\text{Re}\langle \mathbf{u}, \mathbf{v} \rangle$.

If U is a subset of V, then the *orthogonal complement* of U, denoted $U^\perp$, is the set of all vectors in V that are orthogonal to every vector in U:

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

We may verify that $U^\perp$ is always a subspace of V, that $V^\perp = \{\mathbf{0}\}$, and that $\{\mathbf{0}\}^\perp = V$. Also note that if $U_1 \subset U_2$, then $U_1^\perp \supset U_2^\perp$.

Definition. *For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, the* distance *between $\mathbf{u}$ and $\mathbf{v}$, denoted* $\text{dist}(\mathbf{u}, \mathbf{v})$*, is the length of the vector $\mathbf{u} - \mathbf{v}$:*

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Theorem 3. *Let A be an $m \times n$ matrix. The orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null spac of A^T:*

$$(\text{Row } A)^\perp = \text{Nul } A \quad \textit{and} \quad (\text{Col } A)^\perp = \text{Nul } A^T$$

Examples.