Math 304–504

Linear Algebra

Lecture 24:
Orthogonal subspaces.

Scalar product in $\mathbb{R}^n$

Definition. The **scalar product** of vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$ is

$$\boxed{\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.}$$

Properties of scalar product:

$\mathbf{x} \cdot \mathbf{x} \geq 0$, $\mathbf{x} \cdot \mathbf{x} = 0$ only if $\mathbf{x} = \mathbf{0}$ (positivity)

$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ (symmetry)

$z(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$ (distributive law)

$(r\mathbf{x}) \cdot \mathbf{y} = r(\mathbf{x} \cdot \mathbf{y})$ (homogeneity)

In particular, $\mathbf{x} \cdot \mathbf{y}$ is a **bilinear** function (i.e., it is both a linear function of $\mathbf{x}$ and a linear function of $\mathbf{y}$).

Orthogonality

Definition 1. Vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are said to be **orthogonal** (denoted $\mathbf{x} \perp \mathbf{y}$) if $\boxed{\mathbf{x} \cdot \mathbf{y} = 0.}$

Definition 2. A vector $\mathbf{x} \in \mathbb{R}^n$ is said to be **orthogonal** to a nonempty set $Y \subset \mathbb{R}^n$ (denoted $\mathbf{x} \perp Y$) if $\mathbf{x} \cdot \mathbf{y} = 0$ for any $\mathbf{y} \in Y$.

Definition 3. Nonempty sets $X, Y \subset \mathbb{R}^n$ are said to be **orthogonal** (denoted $X \perp Y$) if $\mathbf{x} \cdot \mathbf{y} = 0$ for any $\mathbf{x} \in X$ and $\mathbf{y} \in Y$.

Examples in $\mathbb{R}^3$. • The line $x = y = 0$ is orthogonal to the line $y = z = 0$.

Indeed, if $\mathbf{v} = (0, 0, z)$ and $\mathbf{w} = (x, 0, 0)$ then $\mathbf{v} \cdot \mathbf{w} = 0$.

- The line $x = y = 0$ is orthogonal to the plane $z = 0$.

Indeed, if $\mathbf{v} = (0, 0, z)$ and $\mathbf{w} = (x, y, 0)$ then $\mathbf{v} \cdot \mathbf{w} = 0$.

- The line $x = y = 0$ is not orthogonal to the plane $z = 1$.

The vector $\mathbf{v} = (0, 0, 1)$ belongs to both the line and the plane, and $\mathbf{v} \cdot \mathbf{v} = 1 \neq 0$.

- The plane $z = 0$ is not orthogonal to the plane $y = 0$.

The vector $\mathbf{v} = (1, 0, 0)$ belongs to both planes and $\mathbf{v} \cdot \mathbf{v} = 1 \neq 0$.

Proposition 1 If $X, Y \in \mathbb{R}^n$ are orthogonal sets then either they are disjoint or $X \cap Y = \{\mathbf{0}\}$.

Proof: $\mathbf{v} \in X \cap Y \implies \mathbf{v} \perp \mathbf{v} \implies \mathbf{v} \cdot \mathbf{v} = 0 \implies \mathbf{v} = \mathbf{0}$.

Proposition 2 Let V be a subspace of $\mathbb{R}^n$ and S be a spanning set for V. Then for any $\mathbf{x} \in \mathbb{R}^n$

$$\mathbf{x} \perp S \implies \mathbf{x} \perp V.$$

Proof: Any $\mathbf{v} \in V$ is represented as $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$, where $\mathbf{v}_i \in S$ and $a_i \in \mathbb{R}$. If $\mathbf{x} \perp S$ then

$$\mathbf{x} \cdot \mathbf{v} = a_1(\mathbf{x} \cdot \mathbf{v}_1) + \cdots + a_k(\mathbf{x} \cdot \mathbf{v}_k) = 0 \implies \mathbf{x} \perp \mathbf{v}.$$

Example. The vector $\mathbf{v} = (1, 1, 1)$ is orthogonal to the plane spanned by vectors $\mathbf{w}_1 = (2, -3, 1)$ and $\mathbf{w}_2 = (0, 1, -1)$ (because $\mathbf{v} \cdot \mathbf{w}_1 = \mathbf{v} \cdot \mathbf{w}_2 = 0$).

Orthogonal complement

Definition. Let $S \subset \mathbb{R}^n$. The **orthogonal complement** of S, denoted $S^\perp$, is the set of all vectors $\mathbf{x} \in \mathbb{R}^n$ that are orthogonal to S. That is, $S^\perp$ is the largest subset of $\mathbb{R}^n$ orthogonal to S.

Theorem 1 $S^\perp$ is a subspace of $\mathbb{R}^n$.

Note that $S \subset (S^\perp)^\perp$, hence $\mathrm{Span}(S) \subset (S^\perp)^\perp$.

Theorem 2 $(S^\perp)^\perp = \mathrm{Span}(S)$. In particular, for any subspace V we have $(V^\perp)^\perp = V$.

Example. Consider a line $L = \{(x, 0, 0) \mid x \in \mathbb{R}\}$ and a plane $\Pi = \{(0, y, z) \mid y, z \in \mathbb{R}\}$ in $\mathbb{R}^3$. Then $L^\perp = \Pi$ and $\Pi^\perp = L$.

Fundamental subspaces

Definition. Given an $m \times n$ matrix A, let

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\},$$
$$R(A) = \{\mathbf{b} \in \mathbb{R}^m \mid \mathbf{b} = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^n\}.$$

$R(A)$ is the range of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$, $L(\mathbf{x}) = A\mathbf{x}$. $N(A)$ is the kernel of L.

Also, $N(A)$ is the nullspace of the matrix A while $R(A)$ is the column space of A. The row space of A is $R(A^T)$.

The subspaces $N(A), R(A^T) \subset \mathbb{R}^n$ and $R(A), N(A^T) \subset \mathbb{R}^m$ are **fundamental subspaces** associated to the matrix A.

Theorem $N(A) = R(A^T)^\perp$, $N(A^T) = R(A)^\perp$.
That is, the nullspace of a matrix is the orthogonal complement of its row space.

Proof: The equality $A\mathbf{x} = \mathbf{0}$ means that the vector $\mathbf{x}$ is orthogonal to rows of the matrix A. Therefore $N(A) = S^\perp$, where S is the set of rows of A. It remains to note that $S^\perp = \mathrm{Span}(S)^\perp = R(A^T)^\perp$.

Corollary Let V be a subspace of $\mathbb{R}^n$. Then $\dim V + \dim V^\perp = n$.

Proof: Pick a basis $\mathbf{v}_1, \dots, \mathbf{v}_k$ for V. Let A be the $k \times n$ matrix whose rows are vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$. Then $V = R(A^T)$ and $V^\perp = N(A)$. Consequently, $\dim V$ and $\dim V^\perp$ are rank and nullity of A. Therefore $\dim V + \dim V^\perp$ equals the number of columns of A, which is n.

Direct sum

Definition. Let U, V be subspaces of a vector space W. We say that W is a **direct sum** of U and V (denoted $W = U \oplus V$) if any $\mathbf{w} \in W$ is uniquely represented as $\mathbf{w} = \mathbf{u} + \mathbf{v}$, where $\mathbf{u} \in U$ and $\mathbf{v} \in V$.

Remark. Given subspaces $U, V \subset W$, we can define a set $U + V = \{\mathbf{u} + \mathbf{v} \mid \mathbf{u} \in U, \mathbf{v} \in V\}$, which is also a subspace. However $U \oplus V$ may not be well defined.

Proposition The direct sum $U \oplus V$ is well defined if and only if $U \cap V = \{\mathbf{0}\}$.

Proof: $U \oplus V$ is well defined if for any $\mathbf{u}_1, \mathbf{u}_2 \in U$ and $\mathbf{v}_1, \mathbf{v}_2 \in V$ we have $\mathbf{u}_1 + \mathbf{v}_1 = \mathbf{u}_2 + \mathbf{v}_2 \implies \mathbf{u}_1 = \mathbf{u}_2$ and $\mathbf{v}_1 = \mathbf{v}_2$. Now note that $\mathbf{u}_1 + \mathbf{v}_1 = \mathbf{u}_2 + \mathbf{v}_2 \iff \mathbf{u}_1 - \mathbf{u}_2 = \mathbf{v}_2 - \mathbf{v}_1$.

Theorem $\dim U \oplus V = \dim U + \dim V$.

Proof: Pick a basis $\mathbf{u}_1, \ldots, \mathbf{u}_k$ for U and a basis $\mathbf{v}_1, \ldots, \mathbf{v}_m$ for V. Then $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_m$ is a spanning set for $U \oplus V$. Linear independence of this set follows from the fact that $U \cap V = \{\mathbf{0}\}$.

Theorem Let V be a subspace of $\mathbb{R}^n$. Then $\mathbb{R}^n = V \oplus V^\perp$.

Proof: $V \perp V^\perp \implies V \cap V^\perp = \{\mathbf{0}\} \implies V \oplus V^\perp$ is well defined. Since $\dim V \oplus V^\perp = \dim V + \dim V^\perp = n$, it follows that $V \oplus V^\perp$ is the entire space $\mathbb{R}^n$.

Given a vector $\mathbf{x} \in \mathbb{R}^n$ and a subspace $V \subset \mathbb{R}^n$, there exists a unique representation $\mathbf{x} = \mathbf{p} + \mathbf{o}$ such that $\mathbf{p} \in V$ while $\mathbf{o} \perp V$. The vector $\mathbf{p}$ is called the **orthogonal projection** of $\mathbf{x}$ onto V.

Problem. Find the orthogonal projection of the vector $\mathbf{x} = (2, 1, 0)$ onto the plane Π spanned by vectors $\mathbf{v}_1 = (1, 0, -2)$ and $\mathbf{v}_2 = (0, 1, 1)$.

We have $\mathbf{x} = \mathbf{p} + \mathbf{o}$, where $\mathbf{p} \in \Pi$ and $\mathbf{o} \perp \Pi$. Then $\mathbf{p} = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$ for some $\alpha, \beta \in \mathbb{R}$. Also, $\mathbf{o} \cdot \mathbf{v}_1 = \mathbf{o} \cdot \mathbf{v}_2 = 0$. Note that

$$\mathbf{o}\cdot\mathbf{v}_i = (\mathbf{x}-\alpha\mathbf{v}_1-\beta\mathbf{v}_2)\cdot\mathbf{v}_i = \mathbf{x}\cdot\mathbf{v}_i-\alpha(\mathbf{v}_1\cdot\mathbf{v}_i)-\beta(\mathbf{v}_2\cdot\mathbf{v}_i).$$

$$\begin{cases} \alpha(\mathbf{v}_1 \cdot \mathbf{v}_1) + \beta(\mathbf{v}_2 \cdot \mathbf{v}_1) = \mathbf{x} \cdot \mathbf{v}_1 \\ \alpha(\mathbf{v}_1 \cdot \mathbf{v}_2) + \beta(\mathbf{v}_2 \cdot \mathbf{v}_2) = \mathbf{x} \cdot \mathbf{v}_2 \end{cases}$$

$$\iff \begin{cases} 5\alpha - 2\beta = 2 \\ -2\alpha + 2\beta = 1 \end{cases} \iff \begin{cases} \alpha = 1 \\ \beta = 3/2 \end{cases}$$

Thus $\mathbf{p} = \mathbf{v}_1 + \frac{3}{2}\mathbf{v}_2 = (1, \frac{3}{2}, -\frac{1}{2})$.