MATH 423

Linear Algebra II

Lecture 29: Orthogonal sets.

Orthogonality

Let V be an inner product space with an inner product $\langle \cdot, \cdot \rangle$.

Definition 1. Vectors $\mathbf{x}, \mathbf{y} \in V$ are said to be **orthogonal** (denoted $\mathbf{x} \perp \mathbf{y}$) if $\boxed{\langle \mathbf{x}, \mathbf{y} \rangle = 0.}$

Definition 2. A vector $\mathbf{x} \in V$ is said to be **orthogonal** to a nonempty set $Y \subset V$ (denoted $\mathbf{x} \perp Y$) if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for any $\mathbf{y} \in Y$.

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

Orthogonal sets

Let V be an inner product space with an inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.

Definition. A nonempty set $S \subset V$ is called an **orthogonal set** if all vectors in S are mutually orthogonal. That is, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for any $\mathbf{x}, \mathbf{y} \in S$, $\mathbf{x} \neq \mathbf{y}$. An orthogonal set $S \subset V$ is called **orthonormal** if $\|\mathbf{x}\| = 1$ for any $\mathbf{x} \in S$.

Remark. Vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in V$ form an orthonormal set if and only if

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Examples. • $V = \mathbb{R}^n$, $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y}$.

The standard basis $\mathbf{e}_1 = (1, 0, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)$, ..., $\mathbf{e}_n = (0, 0, 0, \dots, 1)$.

It is an orthonormal set.

• $V = \mathbb{R}^3$, $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y}$.

$\mathbf{v}_1 = (3, 5, 4)$, $\mathbf{v}_2 = (3, -5, 4)$, $\mathbf{v}_3 = (4, 0, -3)$.

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 0, \quad \mathbf{v}_1 \cdot \mathbf{v}_3 = 0, \quad \mathbf{v}_2 \cdot \mathbf{v}_3 = 0,$$
$$\mathbf{v}_1 \cdot \mathbf{v}_1 = 50, \quad \mathbf{v}_2 \cdot \mathbf{v}_2 = 50, \quad \mathbf{v}_3 \cdot \mathbf{v}_3 = 25.$$

Thus the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is orthogonal but not orthonormal. An orthonormal set is formed by normalized vectors $\mathbf{w}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$, $\mathbf{w}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}$, $\mathbf{w}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|}$.

- $V = C[-\pi, \pi]$, $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx$.

$f_1(x) = \sin x$, $f_2(x) = \sin 2x$, $\ldots$, $f_n(x) = \sin nx$, $\ldots$

$$\langle f_m, f_n \rangle = \int_{-\pi}^{\pi} \sin(mx)\sin(nx)\,dx = \begin{cases} \pi & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases}$$

Thus the set $\{f_1, f_2, f_3, \ldots\}$ is orthogonal but not orthonormal.

It is orthonormal with respect to a scaled inner product

$$\langle\langle f, g \rangle\rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)\,dx.$$

- $V = C([-\pi,\pi],\mathbb{C})$, $\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$.

$h_n(x) = e^{inx}$, $n \in \mathbb{Z}$.

$h_n(x) = \cos(nx) + i\sin(nx)$,
$\overline{h_n(x)} = \cos(nx) - i\sin(nx) = e^{-inx} = h_{-n}(x)$.

$$\langle h_m, h_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{imx}\,\overline{e^{inx}}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{imx}e^{-inx}\,dx$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)x}\,dx = \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases}$$

Thus the functions $\dots, h_{-2}, h_{-1}, h_0, h_1, h_2, \dots$ form an orthonormal set. One can show that this is a maximal orthonormal set in $C([-\pi,\pi],\mathbb{C})$.

Orthogonality $\Longrightarrow$ linear independence

Theorem Suppose $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are nonzero vectors that form an orthogonal set. Then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent.

Proof: Suppose $t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k = \mathbf{0}$ for some scalars $t_1, t_2, \dots, t_k$. We have to show that all those scalars are zeros.

For any index $1 \leq i \leq k$ we have

$$\langle t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k, \mathbf{v}_i\rangle = \langle \mathbf{0}, \mathbf{v}_i\rangle = 0$$

$$\Longrightarrow\ t_1\langle \mathbf{v}_1, \mathbf{v}_i\rangle + t_2\langle \mathbf{v}_2, \mathbf{v}_i\rangle + \cdots + t_k\langle \mathbf{v}_k, \mathbf{v}_i\rangle = 0.$$

By orthogonality, $t_i\langle \mathbf{v}_i, \mathbf{v}_i\rangle = 0 \Longrightarrow t_i = 0.$

Orthonormal bases

Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is an orthonormal basis for an inner product space V.

Theorem Let $\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$ and $\mathbf{y} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \cdots + y_n\mathbf{v}_n$, where $x_i, y_j \in \mathbb{C}$. Then

(i) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$,

(ii) $\|\mathbf{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$.

Proof: (ii) follows from (i) when $\mathbf{y} = \mathbf{x}$.

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

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y_j}\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^{n} x_i\overline{y_i}.$$

Fourier coefficients

Suppose $S = \{\mathbf{v}_\alpha\}_{\alpha\in\mathcal{A}}$ is an orthogonal subset of an inner product space V such that $\mathbf{0} \notin S$. For any $\mathbf{x} \in V$, a collection of scalars $c_\alpha = \dfrac{\langle \mathbf{x}, \mathbf{v}_\alpha \rangle}{\langle \mathbf{v}_\alpha, \mathbf{v}_\alpha \rangle}$, $\alpha \in \mathcal{A}$, is called the **Fourier coefficients** of the vector **x** relative to S.

Remark. Classical Fourier coefficients were the coefficients of a function $f \in C([-\pi, \pi], \mathbb{C})$ relative to the orthogonal set $1, \sin x, \cos x, \sin 2x, \cos 2x, \dots$ or the orthonormal set $\dots, e^{-2ix}, e^{-ix}, 1, e^{ix}, e^{2ix}, \dots$.

Theorem If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthogonal basis for V, then the Fourier coefficients of any vector $\mathbf{x} \in V$ relative to S coincide with the coordinates of **x** relative to S. In other words,

$$\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 + \frac{\langle \mathbf{x}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{x}, \mathbf{v}_n \rangle}{\langle \mathbf{v}_n, \mathbf{v}_n \rangle}\mathbf{v}_n.$$

Theorem If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthogonal basis for V, then the Fourier coefficients of any vector $\mathbf{x} \in V$ relative to S coincide with the coordinates of $\mathbf{x}$ relative to S. In other words,

$$\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 + \frac{\langle \mathbf{x}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{x}, \mathbf{v}_n \rangle}{\langle \mathbf{v}_n, \mathbf{v}_n \rangle}\mathbf{v}_n.$$

Proof: Let $\mathbf{p}$ denote the right-hand side of the above formula. For any index $1 \leq i \leq n$,

$$\langle \mathbf{p}, \mathbf{v}_i \rangle = \sum_{j=1}^{n} \frac{\langle \mathbf{x}, \mathbf{v}_j \rangle}{\langle \mathbf{v}_j, \mathbf{v}_j \rangle} \langle \mathbf{v}_j, \mathbf{v}_i \rangle = \frac{\langle \mathbf{x}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \langle \mathbf{v}_i, \mathbf{v}_i \rangle = \langle \mathbf{x}, \mathbf{v}_i \rangle.$$

Hence $\langle \mathbf{x}-\mathbf{p}, \mathbf{v}_i \rangle = \langle \mathbf{x}, \mathbf{v}_i \rangle - \langle \mathbf{p}, \mathbf{v}_i \rangle = 0$. That is, $\mathbf{x}-\mathbf{p} \perp \mathbf{v}_i$. Any vector $\mathbf{y} \in V$ is represented as $\mathbf{y} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$ for some scalars r_i. Then

$$\langle \mathbf{x}-\mathbf{p}, \mathbf{y} \rangle = \overline{r_1} \langle \mathbf{x}-\mathbf{p}, \mathbf{v}_1 \rangle + \cdots + \overline{r_n} \langle \mathbf{x}-\mathbf{p}, \mathbf{v}_n \rangle = 0.$$

Therefore $\mathbf{x}-\mathbf{p} \perp V$. In particular, $\mathbf{x}-\mathbf{p} \perp \mathbf{x}-\mathbf{p}$, which is only possible if $\mathbf{x}-\mathbf{p} = \mathbf{0}$.

Fourier series: linear algebra meets calculus

Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \ldots$ are nonzero vectors in an inner product space V that form an orthogonal set S. Given $\mathbf{x} \in V$, the **Fourier series** of the vector $\mathbf{x}$ relative to the orthogonal set S is a series $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n + \cdots$, where $c_1, c_2, \ldots$ are the Fourier coefficients of $\mathbf{x}$ relative to S.

The set S is called a **Hilbert basis** for V if any vector $\mathbf{x} \in V$ can be expanded into a series $\mathbf{x} = \sum_{n=1}^{\infty} \alpha_n \mathbf{v}_n$, where α_n are some scalars.

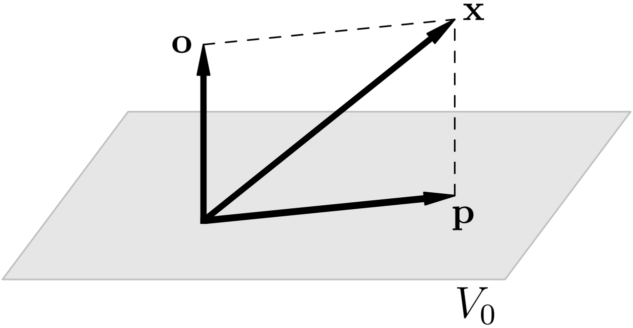
Theorem 1 If S is a Hilbert basis for V, then the above expansion is unique for any vector $\mathbf{x} \in V$. Namely, it coincides with the Fourier series of $\mathbf{x}$ relative to S.

Theorem 2 The sets $1, \sin x, \cos x, \sin 2x, \cos 2x, \ldots$ and $\{e^{inx}\}_{n\in\mathbb{Z}}$ are two Hilbert bases for the space $C([-\pi, \pi], \mathbb{C})$.

Remark. Convergence of functions in the inner product space $C([-\pi, \pi], \mathbb{C})$ need not imply pointwise convergence.

Orthogonal projection

Theorem Let V be an inner product space and V_0 be a finite-dimensional subspace of V. Then any vector $\mathbf{x} \in V$ is uniquely represented as $\mathbf{x} = \mathbf{p} + \mathbf{o}$, where $\mathbf{p} \in V_0$ and $\mathbf{o} \perp V_0$.



The component $\mathbf{p}$ is called the **orthogonal projection** of the vector $\mathbf{x}$ onto the subspace V_0.