Math 304–504

Linear Algebra

Lecture 28:
Orthogonal sets.
The Gram-Schmidt process.

Orthogonal sets

Let V be an inner product space with an inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\|\cdot\|$.

Definition. A nonempty set $S \subset V$ of nonzero vectors is called an **orthogonal set** if all vectors in S are mutually orthogonal. That is, $\mathbf{0} \notin S$ and $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ 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 $\mathbf{e}_1 = (1, 0, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $\mathbf{e}_n = (0, 0, 0, \ldots, 1)$.

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 $\mathbf{w}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$, $\mathbf{w}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}$, $\mathbf{w}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|}$.

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

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

$$\langle f_m, f_n\rangle = \int_{-\pi}^{\pi} \sin(mx)\sin(nx)\,dx$$

$$= \int_{-\pi}^{\pi} \frac{1}{2}\big(\cos(mx - nx) - \cos(mx + nx)\big)\,dx.$$

$$\int_{-\pi}^{\pi} \cos(kx)\,dx = \frac{\sin(kx)}{k}\,\Big|_{x=-\pi}^{\pi} = 0 \ \text{ if } \ k \in \mathbb{Z},\ k \neq 0.$$

$$k = 0 \implies \int_{-\pi}^{\pi} \cos(kx)\,dx = \int_{-\pi}^{\pi} dx = 2\pi.$$

$$\langle f_m, f_n \rangle = \frac{1}{2} \int_{-\pi}^{\pi} \Big(\cos(m-n)x - \cos(m+n)x \Big)\, dx$$

$$= \begin{cases} \pi & \text{if } \; m = n \\ 0 & \text{if } \; m \neq n \end{cases}$$

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

It is orthonormal with respect to a scaled inner product

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

Orthogonality $\Longrightarrow$ linear independence

Theorem Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are nonzero vectors that form an orthogonal set. Then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ 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 $t_1, t_2, \ldots, t_k \in \mathbb{R}$.

Then 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

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be 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 $x_i, y_j \in \mathbb{R}$. Then

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

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

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

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

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^{n} x_iy_i.$$

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be a basis for an inner product space V.

Theorem If the basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is an orthogonal set then for any $\mathbf{x} \in V$

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

If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is an orthonormal set then

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

Proof: We have that $\mathbf{x} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$.

$\implies \langle \mathbf{x}, \mathbf{v}_i \rangle = \langle x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n, \mathbf{v}_i \rangle, \ 1 \le i \le n.$

$\implies \langle \mathbf{x}, \mathbf{v}_i \rangle = x_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + \cdots + x_n\langle \mathbf{v}_n, \mathbf{v}_i \rangle$

$\implies \langle \mathbf{x}, \mathbf{v}_i \rangle = x_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle.$

Let V be a vector space with an inner product. Suppose that $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ are nonzero vectors that form an orthogonal set. Given $\mathbf{x} \in V$, let

$$\mathbf{p} = \frac{\langle \mathbf{x}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 + \cdots + \frac{\langle \mathbf{x}, \mathbf{v}_n \rangle}{\langle \mathbf{v}_n, \mathbf{v}_n \rangle} \mathbf{v}_n, \quad \mathbf{o} = \mathbf{x} - \mathbf{p}.$$

Let W denote the span of $\mathbf{v}_1, \dots, \mathbf{v}_n$.

Theorem (a) $\mathbf{o} \perp \mathbf{w}$ for all $\mathbf{w} \in W$ (denoted $\mathbf{o} \perp W$).

(b) $\|\mathbf{o}\| = \|\mathbf{x} - \mathbf{p}\| = \min_{\mathbf{w} \in W} \|\mathbf{x} - \mathbf{w}\|$.

Thus $\mathbf{p}$ is the **orthogonal projection** of the vector $\mathbf{x}$ on the subspace W. Also, $\mathbf{p}$ is closer to $\mathbf{x}$ than any other vector in W, and $\|\mathbf{o}\| = \mathrm{dist}(\mathbf{x}, \mathbf{p})$ is the **distance** from $\mathbf{x}$ to W.

Orthogonalization

Let V be a vector space with an inner product. Suppose $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ is a basis for V. Let

$$\mathbf{v}_1 = \mathbf{x}_1,$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1,$$

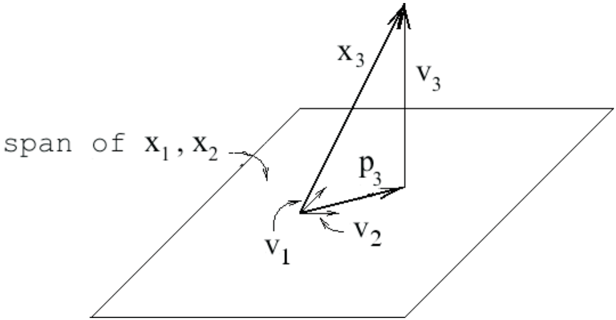
$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\mathbf{v}_n = \mathbf{x}_n - \frac{\langle \mathbf{x}_n, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \cdots - \frac{\langle \mathbf{x}_n, \mathbf{v}_{n-1} \rangle}{\langle \mathbf{v}_{n-1}, \mathbf{v}_{n-1} \rangle} \mathbf{v}_{n-1}.$$

Then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is an orthogonal basis for V.

The orthogonalization of a basis as described above is called the **Gram-Schmidt process**.


x3
v3
span of x1, x2
p3
v2
v1

Normalization

Let V be a vector space with an inner product.
Suppose $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is an orthogonal basis for V.

Let $\mathbf{w}_1 = \dfrac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$, $\mathbf{w}_2 = \dfrac{\mathbf{v}_2}{\|\mathbf{v}_2\|}$,..., $\mathbf{w}_n = \dfrac{\mathbf{v}_n}{\|\mathbf{v}_n\|}$.

Then $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ is an orthonormal basis for V.

Theorem Any finite-dimensional vector space with an inner product has an orthonormal basis.

Remark. An infinite-dimensional vector space with an inner product may or may not have an orthonormal basis.

Problem. Let Π be the plane in $\mathbb{R}^3$ spanned by vectors $\mathbf{x}_1 = (1, 2, 2)$ and $\mathbf{x}_2 = (-1, 0, 2)$.

(i) Find an orthonormal basis for Π.

(ii) Extend it to an orthonormal basis for $\mathbb{R}^3$.

$\mathbf{x}_1, \mathbf{x}_2$ is a basis for the plane Π. We can extend it to a basis for $\mathbb{R}^3$ by adding one vector from the standard basis. For instance, vectors $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3 = (0, 0, 1)$ form a basis for $\mathbb{R}^3$ because

$$\begin{vmatrix} 1 & 2 & 2 \\ -1 & 0 & 2 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ -1 & 0 \end{vmatrix} = 2 \neq 0.$$

Using the Gram-Schmidt process, we orthogonalize the basis $\mathbf{x}_1 = (1, 2, 2)$, $\mathbf{x}_2 = (-1, 0, 2)$, $\mathbf{x}_3 = (0, 0, 1)$:

$$\mathbf{v}_1 = \mathbf{x}_1 = (1, 2, 2),$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 = (-1, 0, 2) - \frac{3}{9}(1, 2, 2)$$

$$= (-4/3, -2/3, 4/3),$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2$$

$$= (0, 0, 1) - \frac{2}{9}(1, 2, 2) - \frac{4/3}{4}(-4/3, -2/3, 4/3)$$

$$= (2/9, -2/9, 1/9).$$

Now $\mathbf{v}_1 = (1, 2, 2)$, $\mathbf{v}_2 = (-4/3, -2/3, 4/3)$, $\mathbf{v}_3 = (2/9, -2/9, 1/9)$ is an orthogonal basis for $\mathbb{R}^3$ while $\mathbf{v}_1, \mathbf{v}_2$ is an orthogonal basis for Π. It remains to normalize these vectors.

$$\langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 9 \implies \|\mathbf{v}_1\| = 3$$
$$\langle \mathbf{v}_2, \mathbf{v}_2 \rangle = 4 \implies \|\mathbf{v}_2\| = 2$$
$$\langle \mathbf{v}_3, \mathbf{v}_3 \rangle = 1/9 \implies \|\mathbf{v}_3\| = 1/3$$

$$\mathbf{w}_1 = \mathbf{v}_1/\|\mathbf{v}_1\| = (1/3, 2/3, 2/3) = \tfrac{1}{3}(1, 2, 2),$$
$$\mathbf{w}_2 = \mathbf{v}_2/\|\mathbf{v}_2\| = (-2/3, -1/3, 2/3) = \tfrac{1}{3}(-2, -1, 2),$$
$$\mathbf{w}_3 = \mathbf{v}_3/\|\mathbf{v}_3\| = (2/3, -2/3, 1/3) = \tfrac{1}{3}(2, -2, 1).$$

$\mathbf{w}_1, \mathbf{w}_2$ is an orthonormal basis for Π.
$\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ is an orthonormal basis for $\mathbb{R}^3$.