

MATH 311

Topics in Applied Mathematics I

Lecture 28:

**Orthogonality in inner product spaces.
The Gram-Schmidt process.**

Norm

The notion of *norm* generalizes the notion of length of a vector in \mathbb{R}^n .

Definition. Let V be a vector space. A function $\alpha : V \rightarrow \mathbb{R}$, usually denoted $\alpha(\mathbf{x}) = \|\mathbf{x}\|$, is called a **norm** on V if it has the following properties:

- (i) $\|\mathbf{x}\| \geq 0$, $\|\mathbf{x}\| = 0$ only for $\mathbf{x} = \mathbf{0}$ (positivity)
- (ii) $\|r\mathbf{x}\| = |r| \|\mathbf{x}\|$ for all $r \in \mathbb{R}$ (homogeneity)
- (iii) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (triangle inequality)

A **normed vector space** is a vector space endowed with a norm. The norm defines a distance function on the normed vector space: $\text{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

Examples. $V = \mathbb{R}^n$, $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

- $\|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|, \dots, |x_n|)$.
- $\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{1/p}$, $p \geq 1$.

Examples. $V = C[a, b]$, $f : [a, b] \rightarrow \mathbb{R}$.

- $\|f\|_\infty = \max_{a \leq x \leq b} |f(x)|$.
- $\|f\|_p = \left(\int_a^b |f(x)|^p dx \right)^{1/p}$, $p \geq 1$.

Inner product

The notion of *inner product* generalizes the notion of dot product of vectors in \mathbb{R}^n .

Definition. Let V be a vector space. A function $\beta : V \times V \rightarrow \mathbb{R}$, usually denoted $\beta(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle$, is called an **inner product** on V if it is positive, symmetric, and bilinear. That is, if

- (i) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ only for $\mathbf{x} = \mathbf{0}$ (positivity)
- (ii) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ (symmetry)
- (iii) $\langle r\mathbf{x}, \mathbf{y} \rangle = r\langle \mathbf{x}, \mathbf{y} \rangle$ (homogeneity)
- (iv) $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ (distributive law)

An **inner product space** is a vector space endowed with an inner product.

Examples. $V = \mathbb{R}^n$.

- $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$.

- $\langle \mathbf{x}, \mathbf{y} \rangle = d_1x_1y_1 + d_2x_2y_2 + \cdots + d_nx_ny_n$,

where $d_1, d_2, \dots, d_n > 0$.

Examples. $V = C[a, b]$.

- $\langle f, g \rangle = \int_a^b f(x)g(x) dx$.

- $\langle f, g \rangle = \int_a^b f(x)g(x)w(x) dx$,

where w is bounded, piecewise continuous, and $w > 0$ everywhere on $[a, b]$.

Theorem Suppose $\langle \mathbf{x}, \mathbf{y} \rangle$ is an inner product on a vector space V . Then

$$\langle \mathbf{x}, \mathbf{y} \rangle^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle \quad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Proof: For any $t \in \mathbb{R}$ let $\mathbf{v}_t = \mathbf{x} + t\mathbf{y}$. Then

$$\begin{aligned} \langle \mathbf{v}_t, \mathbf{v}_t \rangle &= \langle \mathbf{x} + t\mathbf{y}, \mathbf{x} + t\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} + t\mathbf{y} \rangle + t\langle \mathbf{y}, \mathbf{x} + t\mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + t\langle \mathbf{x}, \mathbf{y} \rangle + t\langle \mathbf{y}, \mathbf{x} \rangle + t^2\langle \mathbf{y}, \mathbf{y} \rangle. \end{aligned}$$

Assume that $\mathbf{y} \neq \mathbf{0}$ and let $t = -\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}$. Then

$$\langle \mathbf{v}_t, \mathbf{v}_t \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + t\langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\langle \mathbf{y}, \mathbf{y} \rangle}.$$

Since $\langle \mathbf{v}_t, \mathbf{v}_t \rangle \geq 0$, the desired inequality follows. In the case $\mathbf{y} = \mathbf{0}$, we have $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{y} \rangle = 0$.

Cauchy-Schwarz Inequality:

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}.$$

Corollary 1 $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Equivalently, for all $x_i, y_i \in \mathbb{R}$,

$$(x_1y_1 + \cdots + x_ny_n)^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2).$$

Corollary 2 For any $f, g \in C[a, b]$,

$$\left(\int_a^b f(x)g(x) dx \right)^2 \leq \int_a^b |f(x)|^2 dx \cdot \int_a^b |g(x)|^2 dx.$$

Norms induced by inner products

Theorem Suppose $\langle \mathbf{x}, \mathbf{y} \rangle$ is an inner product on a vector space V . Then $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ is a norm.

Proof: Positivity is obvious. Homogeneity:

$$\|r\mathbf{x}\| = \sqrt{\langle r\mathbf{x}, r\mathbf{x} \rangle} = \sqrt{r^2 \langle \mathbf{x}, \mathbf{x} \rangle} = |r| \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

Triangle inequality (follows from Cauchy-Schwarz's):

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &\leq \langle \mathbf{x}, \mathbf{x} \rangle + |\langle \mathbf{x}, \mathbf{y} \rangle| + |\langle \mathbf{y}, \mathbf{x} \rangle| + \langle \mathbf{y}, \mathbf{y} \rangle \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

Examples. • The length of a vector in \mathbb{R}^n ,

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2},$$

is the norm induced by the dot product

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

• The norm $\|f\|_2 = \left(\int_a^b |f(x)|^2 dx \right)^{1/2}$ on the vector space $C[a, b]$ is induced by the inner product

$$\langle f, g \rangle = \int_a^b f(x)g(x) dx.$$

Angle

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

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \| \mathbf{x} \| \| \mathbf{y} \|$$

for all $\mathbf{x}, \mathbf{y} \in V$ (the Cauchy-Schwarz inequality).

Therefore we can define the **angle** between nonzero vectors in V by

$$\angle(\mathbf{x}, \mathbf{y}) = \arccos \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\| \mathbf{x} \| \| \mathbf{y} \|}.$$

Then $\langle \mathbf{x}, \mathbf{y} \rangle = \| \mathbf{x} \| \| \mathbf{y} \| \cos \angle(\mathbf{x}, \mathbf{y})$.

In particular, vectors \mathbf{x} and \mathbf{y} are **orthogonal** (denoted $\mathbf{x} \perp \mathbf{y}$) if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

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 $\mathbf{x} \in S$.

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

It is an orthonormal set.

Example

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

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

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

Thus the set $\{f_1, f_2, f_3, \dots\}$ 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 \implies 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 $t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \dots + t_k\mathbf{v}_k = \mathbf{0}$ for some $t_1, t_2, \dots, t_k \in \mathbb{R}$.

Then for any index $1 \leq i \leq k$ we have

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

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

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

Orthonormal basis

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

Theorem 1 Let $\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_n\mathbf{v}_n$ and $\mathbf{y} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \dots + y_n\mathbf{v}_n$, where $x_i, y_j \in \mathbb{R}$.

Then

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

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

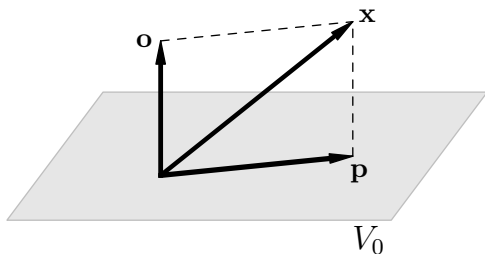
Theorem 2 For any vector $\mathbf{x} \in V$,

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

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 .



The projection \mathbf{p} is closer to \mathbf{x} than any other vector in V_0 . Hence the distance from \mathbf{x} to V_0 is $\|\mathbf{x} - \mathbf{p}\| = \|\mathbf{o}\|$.

Let \mathbf{p} be the orthogonal projection of a vector $\mathbf{x} \in V$ onto a finite-dimensional subspace V_0 .

If V_0 is a one-dimensional subspace spanned by a vector \mathbf{v} then $\mathbf{p} = \frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}$.

If V_0 admits an orthogonal basis $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ then

$$\mathbf{p} = \frac{\langle \mathbf{x}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 + \frac{\langle \mathbf{x}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 + \dots + \frac{\langle \mathbf{x}, \mathbf{v}_k \rangle}{\langle \mathbf{v}_k, \mathbf{v}_k \rangle} \mathbf{v}_k.$$

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

$$\implies \langle \mathbf{x} - \mathbf{p}, \mathbf{v}_i \rangle = 0 \implies \mathbf{x} - \mathbf{p} \perp \mathbf{v}_i \implies \mathbf{x} - \mathbf{p} \perp V_0.$$

The Gram-Schmidt orthogonalization process

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

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

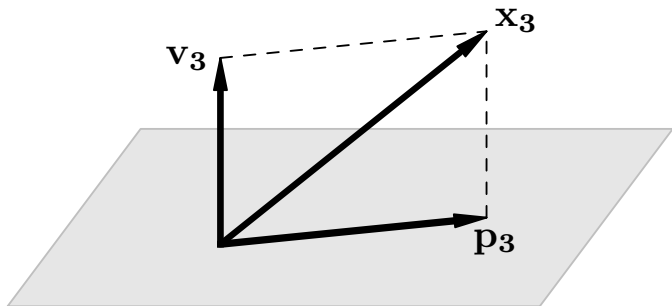
$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1,$$

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

.....

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

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



$$\text{Span}(\mathbf{v}_1, \mathbf{v}_2) = \text{Span}(\mathbf{x}_1, \mathbf{x}_2)$$

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

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 spanned by vectors $\mathbf{x}_1 = (1, 1, 0)$ and $\mathbf{x}_2 = (0, 1, 1)$.

- (i) Find the orthogonal projection of the vector $\mathbf{y} = (4, 0, -1)$ onto the plane Π .
- (ii) Find the distance from \mathbf{y} to Π .

First we apply the Gram-Schmidt process to the basis $\mathbf{x}_1, \mathbf{x}_2$:

$$\mathbf{v}_1 = \mathbf{x}_1 = (1, 1, 0),$$

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

Now that $\mathbf{v}_1, \mathbf{v}_2$ is an orthogonal basis for Π , the orthogonal projection of \mathbf{y} onto Π is

$$\begin{aligned} \mathbf{p} &= \frac{\langle \mathbf{y}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 + \frac{\langle \mathbf{y}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 = \frac{4}{2}(1, 1, 0) + \frac{-3}{3/2}(-1/2, 1/2, 1) \\ &= (2, 2, 0) + (1, -1, -2) = (3, 1, -2). \end{aligned}$$

The distance from \mathbf{y} to Π is $\|\mathbf{y} - \mathbf{p}\| = \|(1, -1, 1)\| = \sqrt{3}$.

Problem. Approximate the function $f(x) = e^x$ on the interval $[-1, 1]$ by a quadratic polynomial.

The best approximation would be a polynomial $p(x)$ that minimizes the distance relative to the uniform norm:

$$\|f - p\|_{\infty} = \max_{|x| \leq 1} |f(x) - p(x)|.$$

However there is no analytic way to find such a polynomial. Instead, one can find a **“least squares”** approximation that minimizes the integral norm

$$\|f - p\|_2 = \left(\int_{-1}^1 |f(x) - p(x)|^2 dx \right)^{1/2}.$$

The norm $\| \cdot \|_2$ is induced by the inner product

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

Therefore $\|f - p\|_2$ is minimal if p is the orthogonal projection of the function f on the subspace \mathcal{P}_3 of quadratic polynomials.

We should apply the Gram-Schmidt process to the polynomials $1, x, x^2$, which form a basis for \mathcal{P}_3 .

This would yield an orthogonal basis p_0, p_1, p_2 .

Then

$$p(x) = \frac{\langle f, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0(x) + \frac{\langle f, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1(x) + \frac{\langle f, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2(x).$$