

MATH 304

Linear Algebra

Lecture 29:

Orthogonal bases.

The Gram-Schmidt orthogonalization process.

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. Nonzero vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in V$ form an **orthogonal set** if they are orthogonal to each other: $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for $i \neq j$.

If, in addition, all vectors are of unit norm, $\|\mathbf{v}_i\| = 1$, then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ is called an **orthonormal set**.

Theorem Any orthogonal set is linearly independent.

Orthonormal bases

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

Theorem Let $\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_n\mathbf{v}_n$ 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) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2 + \dots + x_ny_n$,

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

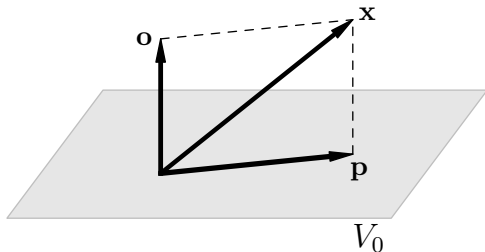
Proof: (ii) follows from (i) when $\mathbf{y} = \mathbf{x}$.

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

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 V be an inner product space. 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 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is an orthogonal basis for V_0 then

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

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

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

Coordinates relative to an orthogonal basis

Theorem If $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is an orthogonal basis for an inner product space V , then

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

for any vector $\mathbf{x} \in V$.

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

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

for any vector $\mathbf{x} \in V$.

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

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

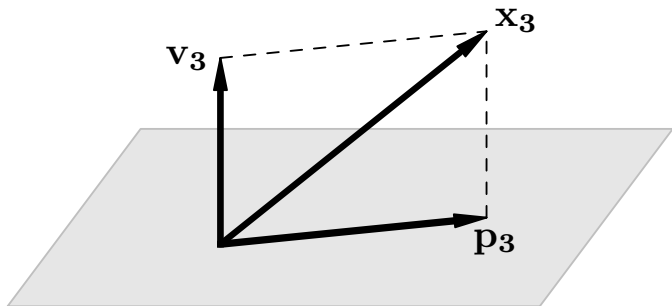
$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1,$$

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

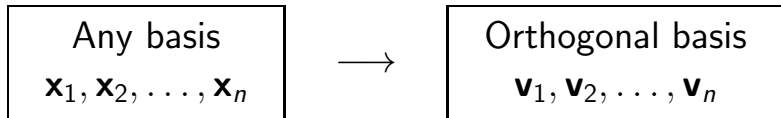
.....

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

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



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



Properties of the Gram-Schmidt process:

- $\mathbf{v}_k = \mathbf{x}_k - (\alpha_1 \mathbf{x}_1 + \dots + \alpha_{k-1} \mathbf{x}_{k-1})$, $1 \leq k \leq n$;
- the span of $\mathbf{v}_1, \dots, \mathbf{v}_k$ is the same as the span of $\mathbf{x}_1, \dots, \mathbf{x}_k$;
- \mathbf{v}_k is orthogonal to $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}$;
- $\mathbf{v}_k = \mathbf{x}_k - \mathbf{p}_k$, where \mathbf{p}_k is the orthogonal projection of the vector \mathbf{x}_k on the subspace spanned by $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}$;
- $\|\mathbf{v}_k\|$ is the distance from \mathbf{x}_k to the subspace spanned by $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}$.

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.

Orthogonalization / Normalization

An alternative form of the Gram-Schmidt process combines orthogonalization with normalization.

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

$$\mathbf{v}_1 = \mathbf{x}_1, \quad \mathbf{w}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|},$$

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

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

.....

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

$$\mathbf{w}_n = \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}.$$

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

Problem. Let V_0 be a subspace of dimension k in \mathbb{R}^n . Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ be a basis for V_0 .

(i) Find an orthogonal basis for V_0 .

(ii) Extend it to an orthogonal basis for \mathbb{R}^n .

Approach 1. Extend $\mathbf{x}_1, \dots, \mathbf{x}_k$ to a basis $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ for \mathbb{R}^n . Then apply the Gram-Schmidt process to the extended basis. We shall obtain an orthogonal basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ for \mathbb{R}^n . By construction, $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \text{Span}(\mathbf{x}_1, \dots, \mathbf{x}_k) = V_0$. It follows that $\mathbf{v}_1, \dots, \mathbf{v}_k$ is a basis for V_0 . Clearly, it is orthogonal.

Approach 2. First apply the Gram-Schmidt process to $\mathbf{x}_1, \dots, \mathbf{x}_k$ and obtain an orthogonal basis $\mathbf{v}_1, \dots, \mathbf{v}_k$ for V_0 . Secondly, find a basis $\mathbf{y}_1, \dots, \mathbf{y}_m$ for the orthogonal complement V_0^\perp and apply the Gram-Schmidt process to it obtaining an orthogonal basis $\mathbf{u}_1, \dots, \mathbf{u}_m$ for V_0^\perp . Then $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{u}_1, \dots, \mathbf{u}_m$ is an orthogonal basis for \mathbb{R}^n .

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 $\mathbf{x}_1 = (1, 2, 2)$, $\mathbf{x}_2 = (-1, 0, 2)$, $\mathbf{x}_3 = (0, 0, 1)$:

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

$$\begin{aligned}\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 = (-1, 0, 2) - \frac{3}{9}(1, 2, 2) \\ &= (-4/3, -2/3, 4/3),\end{aligned}$$

$$\begin{aligned}\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 \\ &= (0, 0, 1) - \frac{2}{9}(1, 2, 2) - \frac{4/3}{4}(-4/3, -2/3, 4/3) \\ &= (2/9, -2/9, 1/9).\end{aligned}$$

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.

$$\langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 9 \implies \|\mathbf{v}_1\| = 3$$

$$\langle \mathbf{v}_2, \mathbf{v}_2 \rangle = 4 \implies \|\mathbf{v}_2\| = 2$$

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

$$\mathbf{w}_1 = \mathbf{v}_1 / \|\mathbf{v}_1\| = (1/3, 2/3, 2/3) = \frac{1}{3}(1, 2, 2),$$

$$\mathbf{w}_2 = \mathbf{v}_2 / \|\mathbf{v}_2\| = (-2/3, -1/3, 2/3) = \frac{1}{3}(-2, -1, 2),$$

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

$\mathbf{w}_1, \mathbf{w}_2$ is an orthonormal basis for Π .

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