Math 304–504

Linear Algebra

Lecture 29:
The Gram-Schmidt process (continued).

Orthogonal sets

Let V be a vector space with an inner product.

Definition. Nonzero vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$ 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 $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is called an **orthonormal set**.

Theorem Any orthogonal set is linearly independent.

Orthogonal projection

Let V be an inner product space.

Let $\mathbf{x}, \mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$. Then $\mathbf{p} = \dfrac{\langle \mathbf{x}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}$ is the **orthogonal projection** of the vector $\mathbf{x}$ onto the vector $\mathbf{v}$. That is, the remainder $\mathbf{o} = \mathbf{x} - \mathbf{p}$ is orthogonal to $\mathbf{v}$.

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

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

is the **orthogonal projection** of the vector $\mathbf{x}$ onto the subspace spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_n$. That is, the remainder $\mathbf{o} = \mathbf{x} - \mathbf{p}$ is orthogonal to $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

The Gram-Schmidt orthogonalization process

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

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

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

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

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

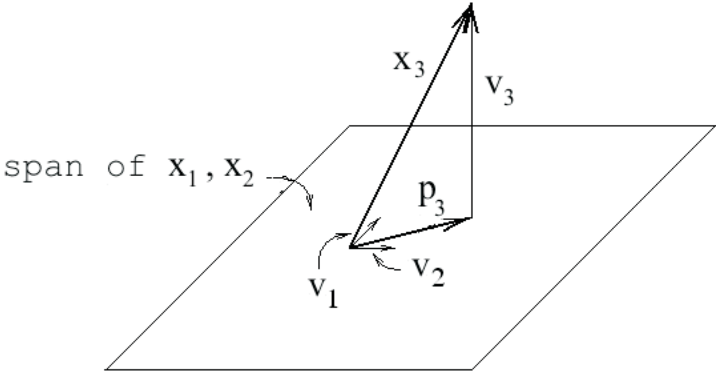
$$\mathbf{v}_n = \mathbf{x}_n - \frac{\langle \mathbf{x}_n, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{x}_n, \mathbf{v}_{n-1} \rangle}{\langle \mathbf{v}_{n-1}, \mathbf{v}_{n-1} \rangle}\mathbf{v}_{n-1}.$$

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

Any basis		Orthogonal basis
$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$	$\longrightarrow$	$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$

Properties of the Gram-Schmidt process:

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



x3
v3
span of x1, x2
p3
v1
v2

Problem. Find the distance from the point $\mathbf{y} = (0, 0, 0, 1)$ to the subspace $\Pi \subset \mathbb{R}^4$ spanned by vectors $\mathbf{x}_1 = (1, -1, 1, -1)$, $\mathbf{x}_2 = (1, 1, 3, -1)$, and $\mathbf{x}_3 = (-3, 7, 1, 3)$.

Let us apply the Gram-Schmidt process to vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{y}$. We should obtain an orthogonal set $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$. The desired distance will be $|\mathbf{v}_4|$.

$$\mathbf{x}_1 = (1, -1, 1, -1),\ \mathbf{x}_2 = (1, 1, 3, -1),$$
$$\mathbf{x}_3 = (-3, 7, 1, 3),\ \mathbf{y} = (0, 0, 0, 1).$$

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

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 = (1, 1, 3, -1) - \frac{4}{4}(1, -1, 1, -1)$$

$$= (0, 2, 2, 0),$$

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

$$= (-3, 7, 1, 3) - \frac{-12}{4}(1, -1, 1, -1) - \frac{16}{8}(0, 2, 2, 0)$$

$$= (0, 0, 0, 0).$$

The Gram-Schmidt process can be used to check linear independence of vectors!

The vector $\mathbf{x}_3$ is a linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$.
Π is a plane, not a 3-dimensional subspace.
We should orthogonalize vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}$.

$$\mathbf{v}_4 = \mathbf{y} - \frac{\langle \mathbf{y}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \frac{\langle \mathbf{y}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2$$

$$= (0, 0, 0, 1) - \frac{-1}{4}(1, -1, 1, -1) - \frac{0}{8}(0, 2, 2, 0)$$

$$= (1/4, -1/4, 1/4, 3/4).$$

$$|\mathbf{v}_4| = \left|\left(\frac{1}{4}, -\frac{1}{4}, \frac{1}{4}, \frac{3}{4}\right)\right| = \frac{1}{4}\,|(1, -1, 1, 3)| = \frac{\sqrt{12}}{4} = \frac{\sqrt{3}}{2}.$$

Problem. Find the distance from the point $\mathbf{z} = (0,0,1,0)$ to the plane Π that passes through the point $\mathbf{x}_0 = (1,0,0,0)$ and is parallel to the vectors $\mathbf{v}_1 = (1,-1,1,-1)$ and $\mathbf{v}_2 = (0,2,2,0)$.

The plane Π is not a subspace of $\mathbb{R}^4$ as it does not pass through the origin. Let $\Pi_0 = \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2)$. Then $\Pi = \Pi_0 + \mathbf{x}_0$.

Hence the distance from the point $\mathbf{z}$ to the plane Π is the same as the distance from the point $\mathbf{z} - \mathbf{x}_0$ to the plane Π_0.

We shall apply the Gram-Schmidt process to vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{z} - \mathbf{x}_0$. This will yield an orthogonal set $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$. The desired distance will be $|\mathbf{w}_3|$.

$\mathbf{v}_1 = (1, -1, 1, -1)$, $\mathbf{v}_2 = (0, 2, 2, 0)$, $\mathbf{z} - \mathbf{x}_0 = (-1, 0, 1, 0)$.

$$\mathbf{w}_1 = \mathbf{v}_1 = (1, -1, 1, -1),$$

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 = \mathbf{v}_2 = (0, 2, 2, 0) \ \text{ as } \mathbf{v}_2 \perp \mathbf{v}_1.$$

$$\mathbf{w}_3 = (\mathbf{z} - \mathbf{x}_0) - \frac{\langle \mathbf{z} - \mathbf{x}_0, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \frac{\langle \mathbf{z} - \mathbf{x}_0, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2$$

$$= (-1, 0, 1, 0) - \frac{0}{4}(1, -1, 1, -1) - \frac{2}{8}(0, 2, 2, 0)$$

$$= (-1, -1/2, 1/2, 0).$$

$$|\mathbf{w}_3| = \left|\left(-1, -\frac{1}{2}, \frac{1}{2}, 0\right)\right| = \frac{1}{2}\,|(-2, -1, 1, 0)| = \frac{\sqrt{6}}{2} = \sqrt{\frac{3}{2}}.$$

Modifications of the Gram-Schmidt process

The first modification 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

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

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

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

. . . .

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

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

Modifications of the Gram-Schmidt process

Another modification is a recursive process which is more stable to roundoff errors than the original process. Suppose $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ is a basis for an inner product space V. Let

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

. . . .

$$\mathbf{v}_n = \mathbf{x}_n - \langle \mathbf{x}_n, \mathbf{w}_1 \rangle \mathbf{w}_1.$$

Then $\mathbf{w}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a basis for V, $\|\mathbf{w}_1\| = 1$, and $\mathbf{w}_1$ is orthogonal to $\mathbf{v}_2, \ldots, \mathbf{v}_n$. Now repeat the process with vectors $\mathbf{v}_2, \ldots, \mathbf{v}_n$, and so on.

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, we are going to 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

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