MATH 323
Linear Algebra

Lecture 21:
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 $\|v\| = \sqrt{\langle v, v \rangle}$.

*Definition.* Nonzero vectors $v_1, v_2, \ldots, v_k \in V$ form an *orthogonal set* if they are orthogonal to each other: $\langle v_i, v_j \rangle = 0$ for $i \neq j$.

If, in addition, all vectors are of unit norm, $\|v_i\| = 1$, then $v_1, v_2, \ldots, v_k$ is called an *orthonormal set*.

*Theorem* Any orthogonal set is linearly independent.
Orthogonal projection

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

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

The projection $p$ is closer to $x$ than any other vector in $V_0$. Hence the distance from $x$ to $V_0$ is $\|x - p\| = \|o\|$. 
Let $V$ be an inner product space. Let $p$ be the orthogonal projection of a vector $x \in V$ onto a finite-dimensional subspace $V_0$.

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

If $v_1, v_2, \ldots, v_n$ is an orthogonal basis for $V_0$ then

$$p = \frac{\langle x, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \frac{\langle x, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 + \cdots + \frac{\langle x, v_n \rangle}{\langle v_n, v_n \rangle} v_n.$$ 

Indeed, $\langle p, v_i \rangle = \sum_{j=1}^{n} \frac{\langle x, v_j \rangle}{\langle v_j, v_j \rangle} \langle v_j, v_i \rangle = \frac{\langle x, v_i \rangle}{\langle v_i, v_i \rangle} \langle v_i, v_i \rangle = \langle x, v_i \rangle$

$$\implies \langle x-p, v_i \rangle = 0 \implies x-p \perp v_i \implies x-p \perp V_0.$$
The Gram-Schmidt orthogonalization process

Let $V$ be a vector space with an inner product. Suppose $x_1, x_2, \ldots, x_n$ is a basis for $V$. Let

$$v_1 = x_1,$$

$$v_2 = x_2 - \frac{\langle x_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1,$$

$$v_3 = x_3 - \frac{\langle x_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle x_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2,$$

$$\ldots$$

$$v_n = x_n - \frac{\langle x_n, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \cdots - \frac{\langle x_n, v_{n-1} \rangle}{\langle v_{n-1}, v_{n-1} \rangle} v_{n-1}.$$

Then $v_1, v_2, \ldots, v_n$ is an orthogonal basis for $V$. 
\[
\text{Span}(\mathbf{v}_1, \mathbf{v}_2) = \text{Span}(\mathbf{x}_1, \mathbf{x}_2)
\]
Any basis \( \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \) \( \rightarrow \) Orthogonal basis \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \)

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 \( \mathbf{v}_1, \ldots, \mathbf{v}_{k-1} \) is the same as the span of \( \mathbf{x}_1, \ldots, \mathbf{x}_{k-1} \);
- \( \mathbf{v}_k \) 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 \( \mathbf{x}_1, \ldots, \mathbf{x}_{k-1} \).
Normalization

Let $V$ be a vector space with an inner product. Suppose $v_1, v_2, \ldots, v_n$ is an orthogonal basis for $V$. Let $w_1 = \frac{v_1}{\|v_1\|}, w_2 = \frac{v_2}{\|v_2\|}, \ldots, w_n = \frac{v_n}{\|v_n\|}$. Then $w_1, w_2, \ldots, w_n$ 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.
An alternative form of the Gram-Schmidt process combines orthogonalization with normalization.

Suppose \( x_1, x_2, \ldots, x_n \) is a basis for an inner product space \( V \). Let

\[
\begin{align*}
v_1 &= x_1, \\
w_1 &= \frac{v_1}{\|v_1\|}, \\
v_2 &= x_2 - \langle x_2, w_1 \rangle w_1, \\
w_2 &= \frac{v_2}{\|v_2\|}, \\
v_3 &= x_3 - \langle x_3, w_1 \rangle w_1 - \langle x_3, w_2 \rangle w_2, \\
w_3 &= \frac{v_3}{\|v_3\|}, \\
&\quad \vdots \\
v_n &= x_n - \langle x_n, w_1 \rangle w_1 - \cdots - \langle x_n, w_{n-1} \rangle w_{n-1}, \\
w_n &= \frac{v_n}{\|v_n\|}.
\end{align*}
\]

Then \( w_1, w_2, \ldots, w_n \) is an orthonormal basis for \( V \).
Problem. Let $V_0$ be a subspace of dimension $k$ in $\mathbb{R}^n$. Let $x_1, x_2, \ldots, x_k$ 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 $x_1, \ldots, x_k$ to a basis $x_1, x_2, \ldots, x_n$ for $\mathbb{R}^n$. Then apply the Gram-Schmidt process to the extended basis. We shall obtain an orthogonal basis $v_1, \ldots, v_n$ for $\mathbb{R}^n$. By construction, $\text{Span}(v_1, \ldots, v_k) = \text{Span}(x_1, \ldots, x_k) = V_0$. It follows that $v_1, \ldots, v_k$ is a basis for $V_0$. Clearly, it is orthogonal.

Approach 2. First apply the Gram-Schmidt process to $x_1, \ldots, x_k$ and obtain an orthogonal basis $v_1, \ldots, v_k$ for $V_0$. Secondly, find a basis $y_1, \ldots, y_m$ for the orthogonal complement $V_0^\perp$ and apply the Gram-Schmidt process to it obtaining an orthogonal basis $u_1, \ldots, u_m$ for $V_0^\perp$. Then $v_1, \ldots, v_k, u_1, \ldots, u_m$ is an orthogonal basis for $\mathbb{R}^n$. 
Problem. Let $\Pi$ be the plane in $\mathbb{R}^3$ spanned by vectors $x_1 = (1, 2, 2)$ and $x_2 = (-1, 0, 2)$.

(i) Find an orthonormal basis for $\Pi$.

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

$x_1, x_2$ is a basis for the plane $\Pi$. We can extend it to a basis for $\mathbb{R}^3$ by adding one vector from the standard basis. For instance, vectors $x_1, x_2,$ and $x_3 = (0, 0, 1)$ 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 $v_1 = (1, 2, 2)$, $v_2 = (-4/3, -2/3, 4/3)$, $v_3 = (2/9, -2/9, 1/9)$ is an orthogonal basis for $\mathbb{R}^3$ while $v_1, v_2$ is an orthogonal basis for $\Pi$. It remains to normalize these vectors.

$$\langle v_1, v_1 \rangle = 9 \implies \|v_1\| = 3$$
$$\langle v_2, v_2 \rangle = 4 \implies \|v_2\| = 2$$
$$\langle v_3, v_3 \rangle = 1/9 \implies \|v_3\| = 1/3$$

$w_1 = v_1/\|v_1\| = (1/3, 2/3, 2/3) = \frac{1}{3}(1, 2, 2)$,

$w_2 = v_2/\|v_2\| = (-2/3, -1/3, 2/3) = \frac{1}{3}(-2, -1, 2)$,

$w_3 = v_3/\|v_3\| = (2/3, -2/3, 1/3) = \frac{1}{3}(2, -2, 1)$.

$w_1, w_2$ is an orthonormal basis for $\Pi$.

$w_1, w_2, w_3$ is an orthonormal basis for $\mathbb{R}^3$. 
Problem. Find the distance from the point $y = (0, 0, 0, 1)$ to the subspace $V \subset \mathbb{R}^4$ spanned by vectors $x_1 = (1, -1, 1, -1)$, $x_2 = (1, 1, 3, -1)$, and $x_3 = (-3, 7, 1, 3)$.

First we apply the Gram-Schmidt process to vectors $x_1, x_2, x_3$ to obtain an orthogonal basis $v_1, v_2, v_3$ for the subspace $V$. Next we compute the orthogonal projection $p$ of the vector $y$ onto $V$:

$$p = \frac{\langle y, v_1 \rangle}{\langle v_1, v_1 \rangle}v_1 + \frac{\langle y, v_2 \rangle}{\langle v_2, v_2 \rangle}v_2 + \frac{\langle y, v_3 \rangle}{\langle v_3, v_3 \rangle}v_3.$$ 

Then the distance from $y$ to $V$ equals $\|y - p\|$.

Alternatively, we can apply the Gram-Schmidt process to vectors $x_1, x_2, x_3, y$. We should obtain an orthogonal system $v_1, v_2, v_3, v_4$. By construction, $v_4 = y - p$ so that the desired distance will be $\|v_4\|$.
\[
\mathbf{x}_1 = (1, -1, 1, -1), \quad \mathbf{x}_2 = (1, 1, 3, -1), \\
\mathbf{x}_3 = (-3, 7, 1, 3), \quad \mathbf{y} = (0, 0, 0, 1).
\]

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

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

\[
= (0, 2, 2, 0),
\]

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

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

\[
= (0, 0, 0, 0).
\]
The Gram-Schmidt process can be used to check linear independence of vectors! It failed because the vector $x_3$ is a linear combination of $x_1$ and $x_2$. $V$ is a plane, not a 3-dimensional subspace. To fix things, it is enough to drop $x_3$, i.e., we should orthogonalize vectors $x_1, x_2, y$.

$$\tilde{v}_3 = y - \frac{\langle y, v_1 \rangle}{\langle v_1, v_1 \rangle}v_1 - \frac{\langle y, v_2 \rangle}{\langle v_2, v_2 \rangle}v_2$$

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

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

$$|\tilde{v}_3| = \left| \left( \frac{1}{4}, -\frac{1}{4}, \frac{1}{4}, \frac{3}{4} \right) \right| = \frac{1}{4} \left| (1, -1, 1, 3) \right| = \frac{\sqrt{12}}{4} = \frac{\sqrt{3}}{2}. $$
Problem. Find the distance from the point \( z = (0, 0, 1, 0) \) to the plane \( \Pi \) that passes through the point \( x_0 = (1, 0, 0, 0) \) and is parallel to the vectors \( v_1 = (1, -1, 1, -1) \) and \( v_2 = (0, 2, 2, 0) \).

The plane \( \Pi \) is not a subspace of \( \mathbb{R}^4 \) as it does not pass through the origin. Let \( \Pi_0 = \text{Span}(v_1, v_2) \). Then \( \Pi = \Pi_0 + x_0 \).

Hence the distance from the point \( z \) to the plane \( \Pi \) is the same as the distance from the point \( z - x_0 \) to the plane \( \Pi - x_0 = \Pi_0 \).

We shall apply the Gram-Schmidt process to vectors \( v_1, v_2, z - x_0 \). This will yield an orthogonal system \( w_1, w_2, w_3 \). The desired distance will be \( \|w_3\| \).
\( \mathbf{v}_1 = (1, -1, 1, -1), \quad \mathbf{v}_2 = (0, 2, 2, 0), \quad \mathbf{z} - \mathbf{x}_0 = (-1, 0, 1, 0). \)

\[ \mathbf{w}_1 = \mathbf{v}_1 = (1, -1, 1, -1), \]
\[ \mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 = \mathbf{v}_2 = (0, 2, 2, 0) \text{ as } \mathbf{v}_2 \perp \mathbf{v}_1. \]

\[ \mathbf{w}_3 = (\mathbf{z} - \mathbf{x}_0) - \frac{\langle \mathbf{z} - \mathbf{x}_0, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{z} - \mathbf{x}_0, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 \]
\[ = (-1, 0, 1, 0) - \frac{0}{4} (1, -1, 1, -1) - \frac{2}{8} (0, 2, 2, 0) \]
\[ = (-1, -1/2, 1/2, 0). \]

\[ |\mathbf{w}_3| = \left| \left( -1, -\frac{1}{2}, \frac{1}{2}, 0 \right) \right| = \frac{1}{2} |(-2, -1, 1, 0)| = \frac{\sqrt{6}}{2} = \sqrt{\frac{3}{2}}. \]
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 polynomials of degree at most 2.

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