MATH 304
Linear Algebra

Lecture 21:
The Gram-Schmidt orthogonalization process.
Eigenvalues and eigenvectors of a matrix.
Orthogonal sets

Let $V$ be a vector space with an inner product.

**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 the **orthogonal projection** of the vector $x$ onto the subspace $V_0$. The distance from $x$ to the subspace $V_0$ is $\|o\|$.

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

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

$$v_1 = x_1,$$
$$v_2 = x_2 - \frac{\langle x_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1,$$
$$v_3 = x_3 - \frac{\langle x_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle x_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2,$$
$$\vdots$$
$$v_n = x_n - \frac{\langle x_n, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \cdots - \frac{\langle x_n, v_{n-1} \rangle}{\langle v_{n-1}, v_{n-1} \rangle} v_{n-1}.$$

Then $v_1, v_2, \ldots, v_n$ is an orthogonal basis for $V$. 
\[ \text{Span}(v_1, v_2) = \text{Span}(x_1, x_2) \]
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 \( \mathbf{x}_1, \ldots, \mathbf{x}_{k-1} \).
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.
Orthogonalization / Normalization

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

\[
\begin{align*}
v_1 &= x_1, \\
w_1 &= \frac{v_1}{\|v_1\|}, \\
v_2 &= x_2 - \langle x_2, w_1 \rangle w_1, \\
w_2 &= \frac{v_2}{\|v_2\|}, \\
v_3 &= x_3 - \langle x_3, w_1 \rangle w_1 - \langle x_3, w_2 \rangle w_2, \\
w_3 &= \frac{v_3}{\|v_3\|}, \\
&
\end{align*}
\]

\[
\begin{align*}
v_n &= x_n - \langle x_n, w_1 \rangle w_1 - \cdots - \langle x_n, w_{n-1} \rangle w_{n-1}, \\
w_n &= \frac{v_n}{\|v_n\|}. \\
\end{align*}
\]

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 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 \( x_1 = (1, 2, 2), x_2 = (-1, 0, 2), x_3 = (0, 0, 1) \):

\[ v_1 = x_1 = (1, 2, 2), \]

\[ v_2 = x_2 - \frac{\langle x_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = (-1, 0, 2) - \frac{3}{9} (1, 2, 2) \]

\[ = (-4/3, -2/3, 4/3), \]

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

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

\[ = (2/9, -2/9, 1/9). \]
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 \( \Pi \). It remains to normalize these vectors.

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

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

\( \mathbf{w}_1, \mathbf{w}_2 \) is an orthonormal basis for \( \Pi \).
\( \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3 \) 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 \( x_3 = (-3, 7, 1, 3). \)

Let us apply the Gram-Schmidt process to vectors \( x_1, x_2, x_3, y. \) We should obtain an orthogonal system \( v_1, v_2, v_3, v_4. \) The desired distance will be \( |v_4|. \)
\[ \mathbf{x}_1 = (1, -1, 1, -1), \quad \mathbf{x}_2 = (1, 1, 3, -1), \quad \mathbf{x}_3 = (-3, 7, 1, 3), \quad \mathbf{y} = (0, 0, 0, 1). \]

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

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

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

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

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

\[ = (0, 0, 0, 0). \]
The Gram-Schmidt process can be used to check linear independence of vectors!

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

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

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

\[
= (1/4, -1/4, 1/4, 3/4).
\]

\[
|\mathbf{v}_3| = \left| \left( \frac{1}{4}, -\frac{1}{4}, \frac{1}{4}, \frac{3}{4} \right) \right| = \frac{1}{4} |(1, -1, 1, 3)| = \frac{\sqrt{12}}{4} = \frac{\sqrt{3}}{2}.
\]
Problem. Find the distance from the point \( z = (0, 0, 1, 0) \) 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 \( \Pi_0 \).

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 \( |w_3| \).
\( \mathbf{v}_1 = (1, -1, 1, -1), \quad \mathbf{v}_2 = (0, 2, 2, 0), \quad \mathbf{z} - \mathbf{x}_0 = (-1, 0, 1, 0). \)

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

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

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

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

\[ = (-1, -1/2, 1/2, 0). \]

\[ |\mathbf{w}_3| = \left| \left( -1, -\frac{1}{2}, \frac{1}{2}, 0 \right) \right| = \frac{1}{2} \left| (-2, -1, 1, 0) \right| = \frac{\sqrt{6}}{2} = \sqrt{3}. \]
**Definition.** Let $A$ be an $n \times n$ matrix. A number $\lambda \in \mathbb{R}$ is called an eigenvalue of the matrix $A$ if $Av = \lambda v$ for a nonzero column vector $v \in \mathbb{R}^n$. The vector $v$ is called an eigenvector of $A$ belonging to (or associated with) the eigenvalue $\lambda$.

**Remarks.** • Alternative notation: eigenvalue = characteristic value, eigenvector = characteristic vector. • The zero vector is never considered an eigenvector.
Example. \( A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \). 

\[
\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix},
\]

\[
\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -6 \end{pmatrix} = 3 \begin{pmatrix} 0 \\ -2 \end{pmatrix}.
\]

Hence \((1, 0)\) is an eigenvector of \( A \) belonging to the eigenvalue 2, while \((0, -2)\) is an eigenvector of \( A \) belonging to the eigenvalue 3.
Example. \( A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \).

\[
\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.
\]

Hence \((1, 1)\) is an eigenvector of \( A \) belonging to the eigenvalue 1, while \((1, -1)\) is an eigenvector of \( A \) belonging to the eigenvalue \(-1\).

Vectors \( \mathbf{v}_1 = (1, 1) \) and \( \mathbf{v}_2 = (1, -1) \) form a basis for \( \mathbb{R}^2 \). Consider a linear operator \( L : \mathbb{R}^2 \to \mathbb{R}^2 \) given by \( L(\mathbf{x}) = A\mathbf{x} \). The matrix of \( L \) with respect to the basis \( \mathbf{v}_1, \mathbf{v}_2 \) is \( B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \).
Let $A$ be an $n \times n$ matrix. Consider a linear operator $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $L(x) = Ax$.

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be a nonstandard basis for $\mathbb{R}^n$ and $B$ be the matrix of the operator $L$ with respect to this basis.

**Theorem**  The matrix $B$ is diagonal if and only if vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are eigenvectors of $A$.

If this is the case, then the diagonal entries of the matrix $B$ are the corresponding eigenvalues of $A$.

$$A \mathbf{v}_i = \lambda_i \mathbf{v}_i \iff B = \begin{pmatrix} \lambda_1 & & O \\ & \lambda_2 & \\ O & \ddots & \ddots \end{pmatrix}$$