6.2 Orthonormal Sets

Def. A set \( \mathbf{u}_1, \ldots, \mathbf{u}_p \) is an orthonormal set if it is orthogonal and every vector is a unit vector.

i.e. \( \mathbf{u}_i \cdot \mathbf{u}_i = \mathbf{u}_i^T \mathbf{u}_i = 1 \)

\( \mathbf{u}_i \cdot \mathbf{u}_j = \mathbf{u}_i^T \mathbf{u}_j = 0 \) if \( i \neq j \).

for all \( i = 1, \ldots, p \),

\( j = 1, \ldots, p \).

Thm 6. An \( m \times n \) matrix \( \mathbf{U} \) has orthonormal columns if and only if \( \mathbf{U}^T \mathbf{U} = \mathbf{I} \).

pf. To simplify notation, suppose \( \mathbf{U} \) is an \( m \times 3 \) matrix, \( \mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] \).

\[
\mathbf{U}^T \mathbf{U} = \begin{bmatrix}
\mathbf{u}_1^T \\
\mathbf{u}_2^T \\
\mathbf{u}_3^T
\end{bmatrix}
\begin{bmatrix}
\mathbf{u}_1 \\
\mathbf{u}_2 \\
\mathbf{u}_3
\end{bmatrix}
= \begin{bmatrix}
\mathbf{u}_1^T \mathbf{u}_1 & \mathbf{u}_1^T \mathbf{u}_2 & \mathbf{u}_1^T \mathbf{u}_3 \\
\mathbf{u}_2^T \mathbf{u}_1 & \mathbf{u}_2^T \mathbf{u}_2 & \mathbf{u}_2^T \mathbf{u}_3 \\
\mathbf{u}_3^T \mathbf{u}_1 & \mathbf{u}_3^T \mathbf{u}_2 & \mathbf{u}_3^T \mathbf{u}_3
\end{bmatrix}
\]

That is, \( \mathbf{U}^T \mathbf{U} = \mathbf{I} \) if and only if \( \mathbf{u}_i^T \mathbf{u}_i = 1 \)

for \( i = 1, 2, 3 \).

\( \mathbf{u}_i^T \mathbf{u}_j = 0 \)

for \( i = 1, 2, 3 \),

\( j = 1, 2, 3 \).
Thm 7 Let \( U \) be an \( m \times n \) matrix with \( n \) (orthonormal) columns, and let \( x, y \in \mathbb{R}^n \).

Then

(a) \( |ux| = |x| \)

(b) \((ux) \cdot (uy) = x \cdot y\)

(c) \((ux) \cdot (uy) = 0 \) if and only if \( x \cdot y = 0 \).

Proof: (b) \((ux) (uy) = (ux)^T (uy)\)

\[ = (x^T u^T) (uy) \]

\[ = x^T (u^T u) y \]

\[ = x^T (I) y \quad \text{by Thm 6.} \]

\[ = x^T y \]

\[ = x \cdot y \]


Parts (a), (c) now follow from (b).

Def. An orthogonal matrix is an invertible matrix \( P \) such that \( P^{-1} = P^T \). Note Thm 6 implies the columns of \( P \) are orthonormal.
7.1 Diagonalization of symmetric matrices.

Def. A symmetric matrix \( A \) is a matrix such that \( A = A^T \). This implies \( A \) is square.

- \( \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \) and \( \begin{bmatrix} 3 & 4 & -4 \\ 4 & 7 & 8 \\ -4 & 8 & 9 \end{bmatrix} \) are symmetric.

- \( \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \) and \( \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \) are not symmetric.
Ex 2 If possible, diagonalize the matrix

\[ A = \begin{bmatrix} 1 & -6 & 4 \\ -6 & 2 & -2 \\ 4 & -2 & -3 \end{bmatrix}. \]

The char. eqn of \( A \) is \( \det(A - \lambda I) = (\lambda + 3)(\lambda + 6)(\lambda - 9) \)

and a corresponding basis for each eigenspace is

\( \lambda = -3 \): \( v_1 = \left( \begin{array}{c} 1 \\ 3 \\ -2 \end{array} \right) \),

\( \lambda = -6 \): \( v_2 = \left( \begin{array}{c} 1 \\ -2 \\ 1 \end{array} \right) \),

\( \lambda = 9 \): \( v_3 = \left( \begin{array}{c} 2 \\ 2 \\ -3 \end{array} \right) \).

Notice \( v_1, v_2, v_3 \) are in fact orthogonal.
Letting \( u_1 = \frac{v_1}{\|v_1\|}, \ u_2 = \frac{v_2}{\|v_2\|}, \ u_3 = \frac{v_3}{\|v_3\|} \)

and \( P = [u_1, u_2, u_3], \ D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix} \)

we see that \( A = PDP^{-1} = PDPT \) (since \( P \) is orthogonal).

Thus, if \( A \) is symmetric, then any eigenvectors from different eigenvalues are orthogonal.
PF (Thm 1) Let $V_1, V_2$ be eigenvectors which correspond to distinct eigenvalues $\lambda_1, \lambda_2$.

We have $\lambda_1 (V_1 \cdot V_2) = (\lambda_1 V_1) \cdot V_2$

$= (\lambda_1 V_1)^T V_2$

$= (AV_1)^T V_2$

$= VA^T V_2$

$= V_1^T AV_2$ (A symmetric)

$= V_1^T (\lambda_2 V_2)$

$= \lambda_2 (V_1^T V_2)$

$= \lambda_2 (V_1 \cdot V_2)$

$\Rightarrow (\lambda_1 - \lambda_2) (V_1 \cdot V_2) = 0 \Rightarrow V_1 \cdot V_2 = 0$

Since $\lambda_1 - \lambda_2 \neq 0$.

Thus $V_1, V_2$ are orthogonal, and arbitrary, which completes the proof. $\Box$
Def. A square matrix $A$ is said to be orthogonally diagonalizable if there are an orthogonal matrix $P$ and a diagonal matrix $D$ such that

$$A = PDP^{-1} = PDPT.$$ 

Notice if $A$ is orthogonally diagonalizable, then

$$A^T = (PDP^T)^T = (P^T)^TD^TP^T = PDPT = A.$$ 

That is, $A$ is symmetric.
This proves one direction of Thm 2. An \( n \times n \) matrix is orthogonally diagonalizable if and only if \( A \) is symmetric.

Ex 3: Orthogonally diagonalize the matrix

\[
A = \begin{bmatrix}
-7 & 4 & 4 \\
4 & 8 & -1 \\
4 & -1 & 8 \\
\end{bmatrix}
\]

Char. eq'n \( \det(A - \lambda I) = 0 \):

\( (2 - \lambda)(\lambda + 9) \). (\( 2 - \lambda \))

Bases for eigenspaces:

\( \lambda = -9 \) : \( V_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \)

\( \lambda = 9 \) : \( V_2 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, V_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \)

Notice \( V_2, V_3 \) are not orthogonal.
but we may find an orthogonal basis for this eigenspace using the Gram-Schmidt process,

\[ \mathbf{z}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 = \begin{pmatrix} \frac{3}{4} \\ 0 \end{pmatrix} - \frac{1}{17} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \]
\[ = \frac{4}{17} \begin{pmatrix} 3 \\ 1 \end{pmatrix}. \]

\[ \mathbf{z}_3 \text{ is the component of } \mathbf{v}_3 \text{ orthogonal to } \mathbf{v}_2. \]

Now \( \{ \mathbf{v}_2, \mathbf{z}_3 \} \) is an orthogonal basis for the eigenspace corresponding to \( \lambda = 9 \).

Normalize \( \mathbf{v}_2, \mathbf{z}_3 \) to obtain

\[ \mathbf{u}_2 = \frac{\mathbf{v}_2}{|\mathbf{v}_2|} = \sqrt{17} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \]
\[ \mathbf{u}_3 = \frac{\mathbf{z}_3}{|\mathbf{z}_3|} = \frac{1}{\sqrt{306}} \begin{pmatrix} 3 \\ 1 \end{pmatrix}. \]
So that $\{u_1, u_2, u_3\}$ is a orthonormal basis for eigenspace corresponding to $\lambda = 9$.

Let $u_i = \frac{v_i}{|v_i|} = \sqrt{18} \begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix}$.

Thm 1 $\Rightarrow \{u_1, u_2, u_3\}$ is an orthonormal basis for $\mathbb{R}^3$.

So, if $P = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$

\[ P^T P = \begin{bmatrix} \frac{4}{\sqrt{18}} & \frac{1}{\sqrt{17}} & \frac{4}{\sqrt{306}} \\ \frac{1}{\sqrt{18}} & 0 & \frac{17}{\sqrt{306}} \\ \frac{1}{\sqrt{18}} & \frac{4}{\sqrt{17}} & \frac{17}{\sqrt{306}} \end{bmatrix} \]
and $D = \begin{pmatrix} -9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}$,

then $A = PDPT$.

That is, $P$ orthogonally diagonalizes $A$.

The **Spectral Theorem**

**Def** The set of eigenvalues $\lambda$ of $A$ is called the spectrum of $A$.

**Thm 3** The spectral theorem for symmetric matrices.
The Spectral Theorem

An $n \times n$ symmetric matrix $A$ has the following properties:

(a) $A$ has $n$ real eigenvalues, counting multiplicity.

(b) The geometric multiplicity of $\lambda$ is equal to the algebraic multiplicity of $\lambda$. If $\lambda$ is an eigenvalue of $A$.

(c) The eigenspaces are mutually orthogonal, in that eigenvectors corresponding to different eigenvalues are orthogonal.

(d) $A$ is orthogonally diagonalizable.
Pf: Part (a) follows from Exercise 23, 24 in section 5.6 (uses properties of complex conjugation).

- Part (b) follows from (c) and the fact nullity \((A - \lambda I) = \text{nullity} (D - \lambda I)

if \(A = PD P^{-1}\).

(see exercise 31 in 7.1)

- Part (c) is Thm 1.

- (d) follows from (a) and the Schur factorization.

Shown in supplementary exercise 16, Ch. 6.
Spectral Decomposition

Suppose \( A = PDP^{-1} \), where \( P = [u_1, \ldots, u_n] \)
and \( \{u_1, \ldots, u_n\} \) is an orthonormal basis for \( \mathbb{R}^n \),

and \( D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \\ \vdots & \vdots \\ 0 & \lambda_n \end{bmatrix} \),

then since \( P^{-1} = P^T \),

\[
A = PD P^T = [u_1, \ldots, u_n] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \\ \vdots & \vdots \\ 0 & \lambda_n \end{bmatrix} [u_1^T, \ldots, u_n^T] = [\lambda_1 u_1, \ldots, \lambda_n u_n] [u_1^T, \ldots, u_n^T]
\]

(Thm 10.2.4) \[
= \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \ldots + \lambda_n u_n u_n^T
\]
That is,

\[ A = \lambda_1 u_1 u_1^T + \ldots + \lambda_n u_n u_n^T \]

This representation \( A \) of \( A \) is called the \textbf{spectral decomposition} of \( A \).

\textbf{Remarks}

1. Each \( u_i u_i^T \) is a matrix of rank 1, since each \( col(m) \) is a multiple of \( u_i \).

2. Each \( u_i u_i^T \) is a \textbf{projection matrix} in the sense that, for every \( x \in \mathbb{R}^n \), \((u_i u_i^T)x\) is the orthogonal projection of \( x \) onto the subspace spanned by \( u_i \).
This is because \((u_i^c u_i^c)^T x = (x \cdot u_i) u_i^c\)

So that \((x - (u_i^c u_i^c)^T x) \cdot u_i = x \cdot u_i - (x \cdot u_i) (u_i \cdot u_i^c)\)

\[= x \cdot u_i - (x \cdot u_i) (u_i \cdot u_i) = 0\] since \(u_i \cdot u_i = 1\).

**Example:** Construct a spectral decomposition of

\[
A = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 4/5 & -3/5 \\ 3/5 & 4/5 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 25 \end{bmatrix}
\]

= \(PD P^T\)

Denote columns of \(P\) by \(u_1, u_2\).

\[u_i u_i^T = \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix} \begin{bmatrix} 4/5 & 3/5 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 16 & 12 \\ 12 & 9 \end{bmatrix}\]
\[ u_2 \ u_2^T = \begin{bmatrix} -3/5 \\ 4/5 \end{bmatrix} \begin{bmatrix} -3/5 \\ 4/5 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 9 \\ -12 \\ -12 \\ 16 \end{bmatrix} \]

So that

\[ 50 u_1 u_1^T + 25 u_2 u_2^T = \begin{bmatrix} 32 \\ 24 \\ 24 \\ 18 \end{bmatrix} + \begin{bmatrix} 9 \\ -12 \\ -12 \\ 16 \end{bmatrix} = \begin{bmatrix} 41 \\ 12 \\ 12 \\ 34 \end{bmatrix} = A. \]