# 7.1 Diagonalization of Symmetric Matrices

Matrix transpose reminder:

The transpose of $AB$ is $(AB)^T = B^T A^T$

The transpose of $A^{-1}$ is $(A^{-1})^T = (A^T)^{-1}$

Definition:

A *symmetric matrix* is a matrix $A$ such that $A^T = A$.

A symmetric matrix is necessarily square.

A symmetric matrix may not have an inverse but if an inverse exists, it is also symmetric.

Here is an example of a symmetric matrix:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & -7 \end{bmatrix}$$

Multiplying any matrix $A$ by $A^T$ gives a symmetric matrix.

Note, our previous $LU$ factorization misses symmetry but a factorization $LDL^{-1}$ captures it!

Suppose $A = A^T$ can be factored into $A = LDU$ without row interchanges. Then $U$ is the transpose of $L$.

The *symmetric factorization* becomes $A = LDL^T$

The transpose of $A = LDU$ gives $A^T = U^T D^T L^T$.

Since $A = A^T$, we have a factorization of $A$ into lower triangular times diagonal times upper triangular.

$L^T$ is upper triangular with ones on the diagonal just like $U$. Since this factorization is unique, $L^T$ must be identical to $U$.

Example:

$L^T = U$ and $A = LDL^T$.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = LDL^T.$$

Elimination and Symmetric Matrices.

$A = A^T$ is an advantage. The smaller matrices stay symmetric as elimination proceeds. So we only need to work with half the matrix! The lower right-hand corner remains symmetric.

$$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \rightarrow \begin{bmatrix} a & b & c \\ 0 & d - \dfrac{b^2}{a} & e - \dfrac{bc}{a} \\ 0 & e - \dfrac{bc}{a} & f - \dfrac{c^2}{a} \end{bmatrix}.$$

The elimination effort is reduced from $\frac{n^3}{3}$ operations to $\frac{n^3}{6}$.

In addition, there is no need to store entries from both sides of the diagonal, or to store both $L$ and $U$.

**EXAMPLE 2** If possible, diagonalize the matrix $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$.

**SOLUTION** The characteristic equation of $A$ is

$$0 = -\lambda^3 + 17\lambda^2 - 90\lambda + 144 = -(\lambda - 8)(\lambda - 6)(\lambda - 3)$$

Standard calculations produce a basis for each eigenspace:

$$\lambda = 8: \ \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = 6: \ \mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}; \qquad \lambda = 3: \ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

These three vectors form a basis for $\mathbb{R}^3$. In fact, it is easy to check that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an *orthogonal* basis for $\mathbb{R}^3$. Experience from Chapter 6 suggests that an *orthonormal* basis might be useful for calculations, so here are the normalized (unit) eigenvectors.

$$\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Let

$$P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}, \quad D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Then $A = PDP^{-1}$, as usual. But this time, since $P$ is square and has orthonormal columns, $P$ is an *orthogonal* matrix, and $P^{-1}$ is simply $P^T$. (See Section 6.2.) ■

Theorem

If $A$ is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.

**PROOF** Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors that correspond to distinct eigenvalues, say, $\lambda_1$ and $\lambda_2$. To show that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, compute

$$\begin{aligned}
\lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 &= (\lambda_1 \mathbf{v}_1)^T \mathbf{v}_2 = (A\mathbf{v}_1)^T \mathbf{v}_2 && \text{Since } \mathbf{v}_1 \text{ is an eigenvector} \\
&= (\mathbf{v}_1^T A^T)\mathbf{v}_2 = \mathbf{v}_1^T (A\mathbf{v}_2) && \text{Since } A^T = A \\
&= \mathbf{v}_1^T (\lambda_2 \mathbf{v}_2) && \text{Since } \mathbf{v}_2 \text{ is an eigenvector} \\
&= \lambda_2 \mathbf{v}_1^T \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2
\end{aligned}$$

Hence $(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. But $\lambda_1 - \lambda_2 \neq 0$, so $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. ■

Definition

A matrix $A$ is *orthogonally diagonalizable* if there are an orthogonal matrix $P$, with $P^{-1} = P^T$, and a digonal matrix $D$ such that $A = PDP^T = PDP^{-1}$

If $A$ is orthogonally diagonalizable, then

$$A^T = \left(PDP^T\right)^T = P^{TT} D^T P^T = A.$$

So $A$ is symmetric.

**Theorem**
An $n \times n$ matrix $A$ is orthogonally diagonalizable if and only if $A$ is a symmetric matrix.

**EXAMPLE 3** Orthogonally diagonalize the matrix $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, whose characteristic equation is

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2)$$

**SOLUTION** The usual calculations produce bases for the eigenspaces:

$$\lambda = 7: \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = -2: \mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

Although $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent, they are not orthogonal. Recall from Section 6.2 that the projection of $\mathbf{v}_2$ onto $\mathbf{v}_1$ is $\dfrac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$, and the component of $\mathbf{v}_2$ orthogonal to $\mathbf{v}_1$ is

$$\mathbf{z}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \frac{-1/2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}$$

Then $\{\mathbf{v}_1, \mathbf{z}_2\}$ is an orthogonal set in the eigenspace for $\lambda = 7$. (Note that $\mathbf{z}_2$ is a linear combination of the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$, so $\mathbf{z}_2$ is in the eigenspace. This construction of $\mathbf{z}_2$ is just the Gram–Schmidt process of Section 6.4.) Since the eigenspace is two-dimensional (with basis $\mathbf{v}_1, \mathbf{v}_2$), the orthogonal set $\{\mathbf{v}_1, \mathbf{z}_2\}$ is an *orthogonal basis* for the eigenspace, by the Basis Theorem. (See Section 2.9 or 4.5.)

Normalize $\mathbf{v}_1$ and $\mathbf{z}_2$ to obtain the following orthonormal basis for the eigenspace for $\lambda = 7$:

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$$

An orthonormal basis for the eigenspace for $\lambda = -2$ is

$$\mathbf{u}_3 = \frac{1}{\|2\mathbf{v}_3\|} 2\mathbf{v}_3 = \frac{1}{3} \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

By Theorem 1, $\mathbf{u}_3$ is orthogonal to the other eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$. Hence $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal set. Let

$$P = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}, \quad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Then $P$ orthogonally diagonalizes $A$, and $A = PDP^{-1}$. ■

The Spectral Theorem for Symmetric Matrices:
An $n \times n$ matrix $A$ is orthogonally diagonalizable if and only if $A$ is a symmetric matrix.

The Spectral Theorem for symmetric matrices:
An $n \times n$ symmetric matrix $A$ has the following properties:

a. $A$ has $n$ real eigenvalues, counting multiplicities.
b. The dimension of the eigenspace for each eigenvalue $\lambda$ equals the multiplicity of $\lambda$ as a root of the characteristic equation.
c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.
d. $A$ is orthogonally diagonalizable.

Spectral Decomposition.

Suppose $A = PDP^{-1}$, where the columns of $P$ are orthonormal eigenvectors $\vec{u}_1, \ldots, \vec{u}_n$ of $A$ and the corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ are in the diagonal matrix $D$. Then, since $P^{-1} = P^T$,

$$A = PDP^T = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{u}_1^T \\ \vdots \\ \vec{u}_n^T \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \vec{u}_1 & \cdots & \lambda_n \vec{u}_n \end{bmatrix} \begin{bmatrix} \vec{u}_1^T \\ \vdots \\ \vec{u}_n^T \end{bmatrix}$$

So we have $A = \lambda_1 \vec{u}_1 \vec{u}_1^T + \lambda_2 \vec{u}_2 \vec{u}_2^T + \cdots + \lambda_n \vec{u}_n \vec{u}_n^T$

This representation of $A$ is called a spectral decomposition of $A$ because it breaks up $A$ into pieces determined by the spectrum (the eigenvalues) of $A$.

Each term, $\vec{u}_j \vec{u}_j^T$, is an $n \times n$ matrix of rank 1.

For example, every column of $\lambda_1 \vec{u}_1 \vec{u}_1^T$ is a multiple of $\vec{u}_1$.

In addition, each matrix $\vec{u}_j \vec{u}_j^T$ is a *projection matrix* in the sense that for each $\vec{x}$ in $\mathbb{R}^n$, the vector $\left(\vec{u}_j \vec{u}_j^T\right)\vec{x}$ is the orthogonal projection of $\vec{x}$ onto the subspace spanned by $\vec{u}_j$.

## NUMERICAL NOTE

When $A$ is symmetric and not too large, modern high-performance computer algorithms calculate eigenvalues and eigenvectors with great precision. They apply a sequence of similarity transformations to $A$ involving orthogonal matrices. The diagonal entries of the transformed matrices converge rapidly to the eigenvalues of $A$. (See the Numerical Notes in Section 5.2.) Using orthogonal matrices generally prevents numerical errors from accumulating during the process. When $A$ is symmetric, the sequence of orthogonal matrices combines to form an orthogonal matrix whose columns are eigenvectors of $A$.

A nonsymmetric matrix cannot have a full set of orthogonal eigenvectors, but the algorithm still produces fairly accurate eigenvalues. After that, nonorthogonal techniques are needed to calculate eigenvectors.

Example: Let $A=\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ and $\vec{v}=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Verify that 5 is an eigenvalue of $A$ and $\vec{v}$ is an eigenvector. Then orthogonally diagonalize $A$.

Solution:

Let $A=\begin{bmatrix}3&1&1\\1&3&1\\1&1&3\end{bmatrix}$. Since each row of $A$ sums to 5,

$$A\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}3&1&1\\1&3&1\\1&1&3\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}5\\5\\5\end{bmatrix}=5\begin{bmatrix}1\\1\\1\end{bmatrix}$$ and 5 is an eigenvalue of $A$.

The eigenvector $\begin{bmatrix}1\\1\\1\end{bmatrix}$ may be normalized to get $\vec{u}_1=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$.

Also,

$$A\begin{bmatrix}-1\\1\\0\end{bmatrix}=\begin{bmatrix}3&1&1\\1&3&1\\1&1&3\end{bmatrix}\begin{bmatrix}-1\\1\\0\end{bmatrix}=\begin{bmatrix}-2\\2\\0\end{bmatrix}=2\begin{bmatrix}-1\\1\\0\end{bmatrix},$$ so 2 is a repeated eigenvalue of $A$

associated with the eigenvector $\begin{bmatrix}-1\\1\\0\end{bmatrix}$.

For $\lambda=2,$ a basis for the eigenspace is $\left\{\begin{bmatrix}-1\\1\\0\end{bmatrix},\begin{bmatrix}-1\\-1\\2\end{bmatrix}\right\}$. This basis is an orthogonal basis for the eigenspace, and these vectors can be normalized ...

We get $\vec{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$ and $\vec{u}_3 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$.

Let

$$P = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix} \text{ and } D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Then $P$ orthogonally diagonalizes $A$, and $A = PDP^{-1}$

Theorem:

If $A$ is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.

Definition:

A matrix $A$ is said to be *orthogonally diagonalizable* if there are an orthogonal matrix $P$ and a diagonal matrix $D$ such that

$$A = PDP^T = PDP^{-1}$$

To orthogonally diagonalize an $n \times n$ matrix, we need to find $n$ linearly independent and orthonormal vectors.

If $A$ is orthogonally diagonalizable, then

$$A^T = \left(PDP^T\right)^T = P^T P^T D^T P^T = PDP^{-1} = A$$

Thus $A$ is symmetric. It turns out that every symmetric matrix is orthogonally diagonalizable.

Matrix Factorizations

1. $A = LU = \begin{pmatrix}\text{lower triangular } L \\ \text{ones on the diagonal}\end{pmatrix}\begin{pmatrix}\text{upper triangular } U \\ \text{pivots on the diagonal}\end{pmatrix}$

   Requirements: No row exchanges, as Gaussian Elimination reduces $A$ to $U$.

2. $A = S\Lambda S^{-1}$

   $= (\text{eigenvectors in } S)(\text{eigenvalues in } \Lambda)(\text{left eigenvectors in } S^{-1})$

   Requirements: $A$ must have $n$ linearly independent eigenvectors.

Matrix Factorizations continued

3. $A = Q\Lambda Q^T = (\text{orthogonal matrix } Q)(\text{real eigenvalues in } \Lambda)(Q^T \text{ is } Q^{-1})$

Requirements: $A$ is symmetric. This is the spectral theorem.

4. $A = U\Sigma V^T = \begin{pmatrix} \text{orthogonal} \\ U \text{ is } m\times m \end{pmatrix} \begin{pmatrix} m\times n \text{ matrix } \Sigma \\ \sigma_1,\ldots,\sigma_r \text{ on diagonal} \end{pmatrix} \begin{pmatrix} \text{orthogonal} \\ V \text{ is } n\times n \end{pmatrix}$

Requirements: None. The singular value decomposition (SVD) has the eigenvectors of $AA^T$ in $U$ and of $A^TA$ in $V$.

$$\sigma_i = \sqrt{\lambda_i\left(A^TA\right)} = \sqrt{\lambda_i\left(AA^T\right)}$$

## Reminder: Diagonalization of a Matrix

*The eigenvectors diagonalize a matrix.*

Suppose the $n \times n$ matrix $A$ has $n$ linearly independent eigenvectors. If these eigenvectors are the columns of a matrix $S$, then $S^{-1}AS$ is a diagonal matrix $\Lambda$.

The eigenvalues of $A$ are on the diagonal of $\Lambda$:

$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Proof: Put the eigenvectors $\vec{v}_i$ in the columns of $S$, and compute $AS$ by columns:

$$AS = A\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \cdots & \lambda_n\vec{v}_n \end{bmatrix}.$$

Then the last matrix can be written

$$\begin{bmatrix} \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \cdots & \lambda_n\vec{v}_n \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = S\Lambda$$

The order of the matrices is important.

If $\Lambda$ came before $S$ then $\lambda_1$ would multiply the entries in the first row. We want $\lambda_1$ to appear in the first column.

Therefore, $AS = S\Lambda$, or $S^{-1}AS = S\Lambda \Rightarrow A = S\Lambda S^{-1}$.

End presentation