25 Lecture 25: Orthogonal Diagonalization

25.1 Recall:

- A square matrix $\mathbf{A}$ is symmetric if $\mathbf{A}^T = \mathbf{A}$
- A matrix $\mathbf{A}$ is an orthogonal matrix if $\mathbf{A}$ is a square real matrix and $\mathbf{A}^T\mathbf{A} = \mathbf{I}$.
- A real square matrix $\mathbf{A}$ is orthogonal is and only if $\mathbf{A}^T = \mathbf{A}^{-1}$
- A square matrix $\mathbf{A}$ is orthogonal if and only if columns of $\mathbf{A}$ form and orthonormal set.

25.2 Orthogonal diagonalization: definitions and examples

Definition 25.1. *An $n \times n$ matrix $\mathbf{A}$ is said to be orthogonally diagonalizable if there are an orthogonal matrix $\mathbf{P}$ (with $P^T = \mathbf{P}^{-1}$) and a diagonal matrix D such that*

$$\mathbf{A} = \mathbf{PDP}^T$$

Example 25.1. *Diagonalize matrix below if possible*

$$\mathbf{A} = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$$

Solution: The characteristic polynomial of $\mathbf{A}$ is

$$p_{\mathbf{A}}(\lambda) = -\lambda^3 + 16\lambda^2 - 90\lambda + 144 = -(\lambda - 3)(\lambda - 6)(\lambda - 8).$$

$p_{\mathbf{A}}(\lambda) = 0$ has three roots, $\lambda = 3$, $\lambda = 6, \lambda = 8$. Thus $\mathbf{A}$ has three distinct real eigenvalues $\lambda = 3$, $\lambda = 6$, and $\lambda = 8$.

Solving the homogeneous system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ for each eigenvalue λ_1, λ_2, and λ_3, we obtain the corresponding eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ (for } \lambda_1 = 3); \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \text{ (for } \lambda_2 = 6), \qquad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \text{ (for } \lambda_3 = 8)$$

We observe that the set formed by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, which is $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$ is orthogonal. We now build an orthonormal set from this one by normalizing each vector in the set. Define

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \qquad \mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \begin{bmatrix} \frac{-1}{\sqrt{6}} \\ \frac{-1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}, \qquad \mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix},$$

then the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthonormal. We next define

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

then $\mathbf{A}$ has the orthogonal diagonalization

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T.$$

Note that if $\mathbf{A}$ is orthogonally diagonalizable then

$$\mathbf{A}^T = (\mathbf{P}\mathbf{D}\mathbf{P}^T)^T = (\mathbf{P}^T)^T\mathbf{D}^T\mathbf{P}^T = \mathbf{P}\mathbf{D}\mathbf{P}^T.$$

Thus, $\mathbf{A}$ is symmetric.

25.3 Eigenvalues and eigenvectors of symmetric matrices and orthogonally diagonalizable

Theorem 14. *If A is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.*

We see from Example 25.1 that matrix $\mathbf{A}$ is symmetric and three eigenvalues $\lambda = 3$, $\lambda = 6$, and $\lambda = 8$ are distinct. So, any two distinct vectors in the set of eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are orthogonal.

Theorem 15. *An $n\times$ matrix $\mathbf{A}$ is orthogonally diagonalizable if and only if $\mathbf{A}$ is a symmetric matrix.*

Example 25.2. *Orthogonally diagonalize the matrix* $\mathbf{A} = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

Solution:

The characteristic polynomial is

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

Equation $p_{\mathbf{A}}(\lambda) = 0$ has to distinct roots, $\lambda = -2$, $\lambda = 7$. Thus $\mathbf{A}$ has two distinct real eigenvalues $\lambda = -2$ and $\lambda = 7$ (of multiplicity 2).

Solving the homogeneous system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ for each eigenvalue -2 and 7, we obtain: For $\lambda = -2$, $\mathbf{A}$ has an eigenvector $\mathbf{v}_1 = \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$. For $\lambda = 7$, $\mathbf{A}$ has two linearly independent eigenvectors $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$. We see that $\mathbf{v}_1$ is orthogonal to both $\mathbf{v}_2$ and $\mathbf{v}_3$ but $\mathbf{v}_2$ is not orthogonal to $\mathbf{v}_3$ although they are linearly independent. Note that $\mathbf{v}_2$ and $\mathbf{v}_3$ form a basis for the eigenspace associated with the eigenvector $\lambda = 7$. We now need to find an orthogonal basis formed by two vectors

w_2 and $\mathbf{w}_3$, for this eigenspace using $\mathbf{v}_2$ and $\mathbf{v}_3$. To do that, we use the Gram-Schmidt process. Let $\mathbf{w}_2 = \mathbf{v}_2$ and find w_3 as

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{w}_2}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} - \frac{-1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}.$$

Then $\{\mathbf{w}_2, \mathbf{w}_3\}$ is an orthogonal set in the eigenspace for $\lambda = 7$. (Note that the number of vector in the basis equals the multiplicity of the eigenvalues in this case), and the set $\{\mathbf{v}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is the orthogonal set of eigenvectors of matrix $\mathbf{A}$. We normalize each vector in this set to obtain an orthonormal basis for the eigenspace for matrix $\mathbf{A}$, which is

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}.$$

Hence $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal set. Let

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

Then $\mathbf{P}$ orthogonal matrix, and the orthogonal diagonalization of $\mathbf{A}$ is $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T$.

25.4 Spectral Decomposition

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

$$\begin{aligned}\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T &= \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}\begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\mathbf{u}_1 & \cdots & \lambda_n\mathbf{u}_n \end{bmatrix}\begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T\end{aligned}$$

This representation of A is called a spectral decomposition of $\mathbf{A}$ because it breaks up $\mathbf{A}$ into pieces determined by the spectrum (eigenvalues) of $\mathbf{A}$. Each term $\lambda_j\mathbf{u}_j\mathbf{u}_j^T$ is an $n \times n$ matrix of rank 1. Furthermore, each matrix $\mathbf{u}_j\mathbf{u}_j^T$ is a projection matrix in the sense that for each $\mathbf{x}$ in $\mathbb{R}^n$, the vector $\left(\mathbf{u}_j\mathbf{u}_j^T\right)\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto the subspace spanned by $\mathbf{u}_j$.

Example 25.3. *Construct a spectral decomposition of the matrix A that has the orthogonal diagonalization*

$$A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}\begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Solution: Let $\mathbf{u}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$. Then

$$\mathbf{A} = 8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T$$

25.5 Suggested problems

Problem 25.1. *Orthogonally diagonalize the matrix* $\mathbf{A} = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$