Orthogonal Diagonalization

Last Lecture

- We saw that every square matrix with real eigenvalues is orthogonally triangularizable.

In this Lecture

- We will determine which $n \times n$ matrices have the property of being orthogonally similar to a diagonal matrix. That is, they have an orthonormal basis for $\mathbb{R}^n$ of eigenvectors.

Orthogonal Diagonalization

Definition: A matrix A is said to be orthogonally diagonalizable if it is orthogonally similar to a diagonal matrix.

Our goal is to figure out which matrices are orthogonally diagonalizable and see what properties produce them.

Assume that A is orthogonally diagonalizable. Then there exists an orthogonal matrix P and diagonal matrix D such that $D = P^T AP$.
Given $P^T = P^{-1}$, solving for A gives $A = PDP^T$.
We know that A and D are similar, so we need to see what properties D has and see if A has those same properties.
Since D is diagonal, then $D^T = D$, which implies

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

Hence, the property $A = A^T$ is shared between A and D.

□

This proves the following:

Theorem 10.2.1

If A is orthogonally diagonalizable, then $A^T = A$.

We want to know if this is sufficient. That is, we want to determine if every matrix A that satisfies the property $A^T = A$ is orthogonally diagonalizable.

Orthogonal Diagonalization

Theorem 10.2.3 - The Principal Axis Theorem

If A is a matrix such that $A^T = A$, then A is orthogonally diagonalizable.

We don't have enough information to prove this just yet, but we can potentially use the Triangularization Theorem. Before we can prove the theorem and use the Triangularization Theorem, we need to prove the following:

Lemma 10.2.2

If A is a matrix such that $A^T = A$, then all eigenvalues of A are real.

The proof of this will be looked at later.

Orthogonal Diagonalization

Theorem 10.2.3 - The Principal Axis Theorem

If A is a matrix such that $A^T = A$, then A is orthogonally diagonalizable.

Proof

By Lemma 10.2.2 all eigenvalues of A are real, and we can apply the Triangularization Theorem to get that there exists an orthogonal matrix P such that $P^T AP = T$ is upper triangular.
Following Theorem 10.2.1, we will show that T has the same properties as A since they are similar, and we get

$$T^T = (P^T AP)^T = P^T A^T P = P^T AP = T$$

Since T is an upper triangular matrix, then T^T is lower triangular, and so T is both upper and lower triangular and hence T is diagonal.

$\square$

Definition: A matrix A such that $A^T = A$ is called symmetric.

We have now proven that a matrix A is orthogonally diagonalizable if and only if it is symmetric.

Orthogonal Diagonalization

Theorem 10.2.4

A matrix A is symmetric if and only if $\vec{x} \cdot (A\vec{y}) = (A\vec{x}) \cdot \vec{y}$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Proof

Suppose that A is symmetric, then for any $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have

$$\vec{x} \cdot (A\vec{y}) = \vec{x}^T A\vec{y} = \vec{x}^T A^T \vec{y} = (A\vec{x})^T \vec{y} = (A\vec{x}) \cdot \vec{y}$$

On the other hand, if $\vec{x} \cdot (A\vec{y}) = (A\vec{x}) \cdot \vec{y}$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$, then

$$\vec{x}^T A\vec{y} = (A\vec{x})^T \vec{y} = \vec{x}^T A^T \vec{y}$$

Since this is valid for all $\vec{y} \in \mathbb{R}^n$, by Theorem 3.1.4, we get that

$$\vec{x}^T A = \vec{x}^T A^T$$

Taking the transpose of both sides gives

$$A^T \vec{x} = A\vec{x}$$

Since this is also valid for all $\vec{x} \in \mathbb{R}^n$, we get $A^T = A$ by Theorem 3.1.4.

□

Orthogonal Diagonalization

Theorem 10.2.5

If A is a symmetric matrix with eigenvectors $\vec{v}_1, \vec{v}_2$ corresponding to distinct eigenvalues λ_1, λ_2 then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof

We need to prove that $\vec{v}_1 \cdot \vec{v}_2 = 0$.
We have

$$\begin{aligned} \lambda_1(\vec{v}_1 \cdot \vec{v}_2) &= (\lambda_1\vec{v}_1) \cdot \vec{v}_2 \\ &= (A\vec{v}_1) \cdot \vec{v}_2 \\ &= \vec{v}_1 \cdot (A\vec{v}_2) \\ &= \vec{v}_1 \cdot (\lambda_2\vec{v}_2) \\ &= \lambda_2(\vec{v}_1 \cdot \vec{v}_2) \end{aligned}$$

Since $\lambda_1 \neq \lambda_2$, this is only possible if $\vec{v}_1 \cdot \vec{v}_2 = 0$.

□

Note: This theorem states that eigenvectors corresponding to different eigenvalues of a symmetric matrix A must be orthogonal.

Orthogonal Diagonalization

Example

Orthogonally diagonalize the symmetric $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 1 \end{bmatrix}$.

Solution

We have

$$C(\lambda) = \begin{vmatrix} 4-\lambda & 0 & 0 \\ 0 & 1-\lambda & -2 \\ 0 & -2 & 1-\lambda \end{vmatrix} = -(\lambda - 4)(\lambda - 3)(\lambda + 1)$$

The eigenvalues are $\lambda_1 = 4$, $\lambda_2 = 3$, and $\lambda_3 = -1$ each with algebraic multiplicity 1.

Orthogonal Diagonalization

Example

Orthogonally diagonalize the symmetric $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 1 \end{bmatrix}$.

Solution

For $\lambda_1 = 4$ we get

$$A - \lambda_1 I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & -2 \\ 0 & -2 & -3 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, a basis for the eigenspace of λ_1 is $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$.

For $\lambda_2 = 3$ we get

$$A - \lambda_2 I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & -2 \\ 0 & -2 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, a basis for the eigenspace of λ_2 is $\left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$.

Orthogonal Diagonalization

Example

Orthogonally diagonalize the symmetric $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 1 \end{bmatrix}$.

Solution

For $\lambda_3 = -1$ we get

$$A - \lambda_3 I = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, a basis for the eigenspace of λ_3 is $\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$.

As Theorem 10.2.5 predicted, the eigenvectors $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$, and $\vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ form an orthogonal set.

We get that A is diagonalized by the orthogonal matrix $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ to $D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

Note: It is important that you are good at diagonalization. You need to be particularly good at finding all eigenvalues of a given matrix, and finding a basis for the eigenspace of an eigenvalue λ by just looking at $A - \lambda I$.

Orthogonal Diagonalization

Example

Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{bmatrix}$.

Solution

We have

$$C(\lambda) = \begin{vmatrix} 5-\lambda & -4 & -2 \\ -4 & 5-\lambda & -2 \\ -2 & -2 & 8-\lambda \end{vmatrix} = -\lambda(\lambda - 9)^2$$

The eigenvalues are $\lambda_1 = 9$ with algebraic multiplicity 2 and $\lambda_2 = 0$ with algebraic multiplicity 1.

Orthogonal Diagonalization

Example

Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{bmatrix}$.

Solution

For $\lambda_1 = 9$ we get

$$A - \lambda_1 I = \begin{bmatrix} -4 & -4 & -2 \\ -4 & -4 & -2 \\ -2 & -2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, a basis for the eigenspace of λ_1 is $\{\vec{w}_1, \vec{w}_2\} = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \right\}$.

However, observe that these eigenvectors are not orthogonal to each other.

Theorem 10.2.5 only guarantees that eigenvectors corresponding to different eigenvalues are necessarily orthogonal.

We know that the eigenspace of λ_1 is a subspace of $\mathbb{R}^3$, so we first need to find an orthogonal basis for the eigenspace of λ_1.

We can do this by applying the Gram-Schmidt procedure to this set.

Orthogonal Diagonalization

Example

Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{bmatrix}$.

Solution

For $\lambda_1 = 9$ we get

$$A - \lambda_1 I = \begin{bmatrix} -4 & -4 & -2 \\ -4 & -4 & -2 \\ -2 & -2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, a basis for the eigenspace of λ_1 is $\{\vec{w}_1, \vec{w}_2\} = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \right\}$.

Pick $\vec{v}_1 = \vec{w}_1$. Then $\mathbb{S}_1 = \operatorname{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$ and

$$\vec{v}_2 = \operatorname{perp}_{\mathbb{S}_1} \vec{w}_2 = \vec{w}_2 - \frac{\vec{w}_2 \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \vec{v}_1 = \frac{1}{2} \begin{bmatrix} -1 \\ -1 \\ 4 \end{bmatrix}$$

Then, an orthogonal basis for the eigenspace of λ_1 is $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 4 \end{bmatrix} \right\}$.

Orthogonal Diagonalization

Example

Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{bmatrix}$.

Solution

For $\lambda_2 = 0$ we get

$$A - \lambda_2 I = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, a basis for the eigenspace of λ_2 is $\{\vec{v}_3\} = \left\{ \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \right\}$.

Observe that $\vec{v}_3$ is orthogonal to the orthogonal basis for the eigenspace of λ_1, as predicted by Theorem 10.2.5.

Hence we now have an orthogonal basis for $\mathbb{R}^3$ of eigenvectors of A.

Normalizing $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$, we get that A is diagonalized by the orthogonal matrix

$$P = \begin{bmatrix} 2/3 & -1/\sqrt{2} & -1/\sqrt{18} \\ 2/3 & 1/\sqrt{2} & -1/\sqrt{18} \\ 1/3 & 0 & 4/\sqrt{18} \end{bmatrix}$$

to $P^T A P = D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$.

Orthogonal Diagonalization

The examples demonstrate that the procedure for orthogonally diagonalizing a symmetric matrix is exactly the same as normal diagonalization except for the following key points:

- If we have a symmetric matrix A, then by the Principal Axis Theorem, we know that A is diagonalizable.
- Theorem 10.2.5 only guarantees that eigenvalues corresponding to different eigenvectors are orthogonal. If you encounter an eigenvalue with geometric multiplicity greater than 1, if necessary, apply the Gram-Schmidt procedure to the basis for that eigenspace to get an orthogonal basis for the eigenspace.
- When creating the diagonalizing matrix P, make sure you have an orthonormal basis of eigenvectors for $\mathbb{R}^n$ of A. That is, make sure to normalize the eigenvectors.

Much of the rest of the course is about orthogonal diagonalization and its uses.