Orthogonal Similarity and Triangularization

Recall:

- Diagonalization finds a basis $\mathcal{B}$ for a linear operator L such that $[L]_{\mathcal{B}}$ is diagonal.
- That is, diagonalization finds a geometrically natural basis for L.

What could be better than a geometrically natural basis? An *orthonormal* geometrically natural basis!

Unfortunately, not every matrix is diagonalizable, and so not every linear mapping has a geometrically natural basis, let alone an orthonormal geometrically natural basis.

In the next two lectures:

- We will look at the theory of orthogonal similarity and orthogonal diagonalization, and find which matrices are orthogonally diagonalizable.

Orthogonal Similarity and Triangularization

Definition: Two matrices A and B are said to be orthogonally similar if there exists an orthogonal matrix P such that

$$P^TAP = B$$

Notes:

- Since P is orthogonal, we have that $P^T = P^{-1}$. Hence, if A and B are orthogonally similar, then they are similar.
- Therefore, all the properties of similar matrices still hold:

$$\begin{aligned} \operatorname{rank} A &= \operatorname{rank} B \\ \operatorname{tr} A &= \operatorname{tr} B \\ \det A &= \det B \\ \det(A - \lambda I) &= \det(B - \lambda I) \end{aligned}$$

Orthogonal Similarity and Triangularization

Theorem 10.1.1 - Triangularization Theorem

If A is an $n \times n$ matrix with all real eigenvalues, then A is orthogonally similar to an upper triangular matrix T.

Proof

We prove the result by induction on n.

Base Case:
If $n = 1$, then A is upper triangular.
Therefore, we can take P to be the orthogonal matrix $P = [1]$ and the result follows.

Inductive Hypothesis:
Now assume the result holds for all $(n-1) \times (n-1)$ matrices.

Inductive Step:
Consider an $n \times n$ matrix A with all real eigenvalues.
Let λ_1 be an eigenvalue of A with corresponding unit eigenvector $\vec{v}_1$.
Extend the set $\{\vec{v}_1\}$ to a basis for $\mathbb{R}^n$ and apply the Gram-Schmidt procedure to produce an orthonormal basis $\{\vec{v}_1, \vec{w}_2, \dots, \vec{w}_n\}$ for $\mathbb{R}^n$.
There is no relationship between the vectors $\vec{w}_2, \dots, \vec{w}_n$ and the matrix A.

Orthogonal Similarity and Triangularization

Theorem 10.1.1 - Triangularization Theorem

If A is an $n \times n$ matrix with all real eigenvalues, then A is orthogonally similar to an upper triangular matrix T.

Proof

Inductive Step (Continued):
The matrix $P_1 = \begin{bmatrix} \vec{v}_1 & \vec{w}_2 & \dots & \vec{w}_n \end{bmatrix}$ is orthogonal and we have that

$$P_1^T A P_1 = \begin{bmatrix} \vec{v}_1^T \\ \vec{w}_2^T \\ \vdots \\ \vec{w}_n^T \end{bmatrix} A \begin{bmatrix} \vec{v}_1 & \vec{w}_2 & \dots & \vec{w}_n \end{bmatrix} = \begin{bmatrix} \vec{v}_1 \cdot A\vec{v}_1 & \vec{v}_1 \cdot A\vec{w}_2 & \cdots & \vec{v}_1 \cdot A\vec{w}_n \\ \vec{w}_2 \cdot A\vec{v}_1 & \vec{w}_2 \cdot A\vec{w}_2 & \dots & \vec{w}_2 \cdot A\vec{w}_n \\ \vdots & & \ddots & \vdots \\ \vec{w}_n \cdot A\vec{v}_1 & \vec{w}_n \cdot A\vec{w}_2 & \dots & \vec{w}_n \cdot A\vec{w}_n \end{bmatrix}$$

Consider the entries in the first column. Since $A\vec{v}_1 = \lambda_1 \vec{v}_1$ we have that

$$\vec{w}_i \cdot A\vec{v}_1 = \vec{w}_i \cdot (\lambda_1 \vec{v}_1) = \lambda_1(\vec{w}_i \cdot \vec{v}_1) = 0 \qquad \text{for } 2 \leq i \leq n$$

Similarly,

$$\vec{v}_1 \cdot A\vec{v}_1 = \vec{v}_1 \cdot (\lambda_1 \vec{v}_1) = \lambda_1(\vec{v}_1 \cdot \vec{v}_1) = \lambda_1$$

Hence,

$$P_1^T A P_1 = \begin{bmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & A_1 \end{bmatrix}$$

where $\vec{0}$ is the zero vector in $\mathbb{R}^{n-1}$, $\vec{b} \in \mathbb{R}^{n-1}$, and A_1 is an $(n-1) \times (n-1)$ matrix.

Orthogonal Similarity and Triangularization

Theorem 10.1.1 - Triangularization Theorem

If A is an $n \times n$ matrix with all real eigenvalues, then A is orthogonally similar to an upper triangular matrix T.

Proof

Inductive Step (Continued):

We have seen that A is orthogonally similar to $\begin{bmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & A_1 \end{bmatrix}$. As such, all the eigenvalues of A_1 are eigenvalues of A (which are real by assumption), and hence all the eigenvalues of A_1 are real.
Therefore, by the inductive hypothesis, there exists an $(n-1) \times (n-1)$ orthogonal matrix Q such that $Q^T A_1 Q = T_1$ is upper triangular.

We need a way to combine the matrices Q and P_1 so that we can show that A is orthogoanlly similar to an upper triangular matrix.

Because Q and P_1 are different sizes, we "make Q larger" by defining the matrix $P_2 = \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix}$.
Since the columns of Q are orthonormal, the columns of P_2 are also orthonormal and hence P_2 is an orthogonal matrix.
Consequently, the matrix $P = P_1 P_2$ is also orthogonal by Theorem 9.2.7.

Orthogonal Similarity and Triangularization

Theorem 10.1.1 - Triangularization Theorem

If A is an $n \times n$ matrix with all real eigenvalues, then A is orthogonally similar to an upper triangular matrix T.

Proof

Inductive Step (Continued):
Then by block multiplication we get

$$\begin{aligned} P^T AP &= (P_1P_2)^T A(P_1P_2) = P_2^T P_1^T A P_1 P_2 \\ &= \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q^T \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & A_1 \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & \vec{b}^T Q \\ \vec{0} & Q^T A_1 Q \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & \vec{b}^T Q \\ \vec{0} & T_1 \end{bmatrix} \end{aligned}$$

This is an upper triangular matrix since T_1 is upper triangular.

Thus, we have completed the inductive step, and so the result now follows by induction. □

Orthogonal Similarity and Triangularization

Notes:

- If A is orthogonally similar to an upper triangular matrix T, then A and T must share the same eigenvalues. Thus, since T is upper triangular, the eigenvalues of A must appear along the main diagonal of T.
- The proof of the theorem gives us a method for finding an orthogonal matrix P such that $P^TAP = T$ is upper triangular. Since the proof is by induction, this leads to a recursive algorithm.

Orthogonal Similarity and Triangularization

Example

Let $A = \begin{bmatrix} -1 & 3 & -1 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$. Find an orthogonal matrix P and upper triangular matrix T such that $P^TAP = T$.

Solution

Observe that -1 is an eigenvalue of A with unit eigenvector $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

We can extend $\{\vec{e}_1\}$ to the standard basis for $\mathbb{R}^3$, and so we take $P_1 = I$.

Thus, we get that

$$P_1^TAP_1 = \begin{bmatrix} -1 & \vec{b}^T \\ \vec{0} & A_1 \end{bmatrix} \qquad \text{where } \vec{b} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \text{ and } A_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

Orthogonal Similarity and Triangularization

Example

Let $A = \begin{bmatrix} -1 & 3 & -1 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$. Find an orthogonal matrix P and upper triangular matrix T such that $P^T AP = T$.

Solution

We repeat the procedure on the submatrix $A_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$:

We have that

$$C(\lambda) = \det(A_1 - \lambda I) = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1)$$

For $\lambda = 3$ we get

$$A_1 - 3I = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \Rightarrow \vec{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

We need to extend $\{\vec{v}_1\}$ to an orthonormal basis for $\mathbb{R}^2$, so we take $\vec{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$.

Hence,

$$Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Orthogonal Similarity and Triangularization

Example

Let $A = \begin{bmatrix} -1 & 3 & -1 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$. Find an orthogonal matrix P and upper triangular matrix T such that $P^T AP = T$.

Solution

Next, we take $P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$.

Finally, $P = P_1 P_2 = IP_2 = P_2$, and we get that

$$T = P^T AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} -1 & 3 & -1 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} -1 & \sqrt{2} & 2\sqrt{2} \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Note:

We have seen before that we can use examples to help us figure out proofs.

In this lecture, we used the proof to teach us how to solve the computational problem.