

Math 2270-1

Notes of 11/25/2019

Announcements

- hw 14 opens today. It's mostly a review, consisting of old T/F questions.
- this week, discuss parts of chpt 7
- next week, review
- Final Exam, Thursday, 12/12 8:00-10:00 am, LCB 219
- Additional and optional Q&A session, Wednesday, 12/11, 10:30-12:30, LCB 215. HW 14 will close December 4.

Chapter 7: Symmetric Matrices

- The main result today is that a (real square) matrix has an **orthogonal similarity transform to diagonal form** if and only if it is symmetric. Getting there will be a little technical!
- Throughout this chapter let A be a square real matrix. Unless stated otherwise, the number of rows and columns is n .
- A square matrix A is symmetric if

$$A = A^T.$$

This means that

$$a_{ij} = a_{ji}$$

- The entries along the diagonal are arbitrary, but those off the diagonal occur in pairs.
- The concept of symmetry only applies to square matrices.

- Examples

- The set of symmetric $n \times n$ matrices forms a linear space. What is its dimension?

Review of Eigenvalues and Similarity

- A non-zero vector \mathbf{x} is an **eigenvector** of A corresponding to the eigenvalue λ if

$$A\mathbf{x} = \lambda\mathbf{x}.$$

- An eigenvector is determined only up to a non-zero factor.
- The eigenvalues of A are the roots of the characteristic equation

$$\det(A - \lambda I) = 0.$$

- A has precisely n eigenvalues, properly counting multiplicity.
- The eigenvalues of A may be complex. If there are any complex eigenvalues they occur in conjugate complex pairs.
- The eigenvalues of a symmetric matrix are real. (see notes of 10/25/19).
- For any symmetric matrix, eigenvectors corresponding to distinct eigenvalues are orthogonal.

Similarity

- Two matrices A and B are similar if there is a non-singular matrix P such that

$$B = P^{-1}AP.$$

- P is sometimes referred to as a **similarity transform**.
- Similar matrices have the same eigenvalues and the eigenvectors of B are of the form $P^{-1}\mathbf{x}$ where \mathbf{x} is an eigenvector of A :

$$B(P^{-1}\mathbf{x}) = (P^{-1}AP)P^{-1}\mathbf{x} = P^{-1}A\mathbf{x} = \lambda(P^{-1}\mathbf{x}).$$

- A matrix A is **diagonalizable** if it is similar to a diagonal (but not necessarily real) matrix.
- A is diagonalizable if and only if it has a linearly independent set of n eigenvectors.
- Suppose there is such a set, satisfying

$$A\mathbf{p}_i = \lambda_i\mathbf{p}_i, \quad i = 1, \dots, n. \quad (1)$$

- Then let

$$P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \dots \mathbf{p}_n] \quad \text{and} \quad D = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

- Equation (1) is the column version of the matrix equation

$$AP = PD$$

- Hence

$$D = P^{-1}AP$$

and so D is similar to A and the similarity transform P is the matrix of eigenvectors.

- this suggests the following procedure for diagonalizing a matrix:
 1. Find n linearly independent eigenvectors.
 2. Collect those vectors into the matrix P .
 3. compute P^{-1}
 4. Then

$$D = P^{-1}AP$$

is a diagonal matrix with the eigenvalues along the diagonal.

- **Example 3**, textbook. The (symmetric) matrix

$$A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

has the characteristic equation

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

Thus -2 is a single eigenvalue and 7 is a double eigenvalue.

- The standard technique of solving the linear system

$$A\mathbf{x} = \lambda x$$

gives the eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}$$

for $\lambda = 7$ and

$$\mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

for $\lambda = -2$

- These eigenvectors are linearly independent. Collecting them into matrix

$$P = \begin{bmatrix} 1 & -1/2 & -1 \\ 0 & 1 & -1/2 \\ 1 & 0 & 1 \end{bmatrix}$$

gives the similarity transform

$$P^{-1}AP = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

- We saw in chapter 6 that orthogonal matrices are particularly nice.

- Two matrices A and B are **orthogonally similar** if there is an orthogonal matrix P such that

$$B = P^{-1}AP = P^TAP.$$

- Suppose A has an orthonormal basis of n eigenvectors.
- Then we can collect those into an orthogonal matrix P where

$$P^{-1} = P^T$$

- We get $D = P^TAP$ and therefore

$$A = PDP^T.$$

- The eigenvectors \mathbf{v}_1 and \mathbf{v}_2 in Example 3 are linearly independent, but not orthogonal. However, they span a space of dimension 2, the eigenspace associated with the eigenvalue 7, and we can construct an orthogonal basis of that space by the Gram-Schmidt Process. We can also normalize \mathbf{v}_3 to be a unit vector. This gives the modified orthogonal similarity transform

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

with

$$Q^T A Q = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

- Returning to the general orthogonal similarity transform, note that the matrix

$$A = PDP^T$$

is symmetric. Thus we see that **a matrix that is orthogonally diagonalizable** is symmetric.

- The converse is also true: Every symmetric (real) matrix is **orthogonally diagonalizable**.
- We get **Theorem 2**. A matrix is orthogonally diagonalizable if and only if it is symmetric.
- We saw the “only if” part, the “if” part is trickier. The textbook gives more information, referring to a bunch of exercises, but it does not have a proof.
- However, here is a proof by induction.
- Induction: We show that a statement is true for $n = 1$. Then we show that it is true for n if it is true for $n - 1$, for all $n = 1, 2, 3, \dots$
- There is nothing to show for $n = 1$. A is symmetric and P is the 1×1 identity.
- So suppose that any symmetric $(n-1) \times (n-1)$ matrix A_1 does have an orthogonal similarity transform to diagonal form:

$$P_1^{-1} A_1 P_1 = D_1$$

where P_1 is orthogonal, i.e.,

$$P_1^T P_1 = I_1$$

where I_1 is the $(n-1) \times (n-1)$ identity, and D_1 is $(n-1) \times (n-1)$ diagonal.

- Now suppose A is symmetric and $n \times n$. Then it has a real eigenvalue λ_1 and corresponding real eigenvector \mathbf{v}_1 , i.e.,

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1.$$

- Since eigenvectors are determined only up to a constant factor we may assume that

$$\|\mathbf{v}_1\| = 1.$$

- We construct an orthogonal $n \times n$ matrix

$$\bar{P} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_n]$$

- This is always possible. It can be accomplished, for example, by rotating the standard coordinate system so that the first standard basis vector \mathbf{e}_1 lines up with \mathbf{v}_1 , or we could add vectors to $\{\mathbf{v}_1\}$ to get a basis of R^n , and then apply the Gram-Schmidt process.
- Then

$$A\bar{P} = [\lambda_1\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \dots \quad A\mathbf{v}_n]$$

and

$$\begin{aligned} \bar{P}^T A \bar{P} &= [\lambda_1 \bar{P}^T \mathbf{v}_1 \quad \bar{P}^T A \mathbf{v}_2 \quad \dots \quad \bar{P}^T A \mathbf{v}_n] \\ &= [\lambda_1 \mathbf{e}_1 \quad \bar{P}^T A \mathbf{v}_2 \quad \dots \quad \bar{P}^T A \mathbf{v}_n] \end{aligned}$$

since \bar{P} is orthogonal.

- Moreover, $\bar{P}^T A \bar{P}$ is symmetric since A is symmetric:

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

- Thus $\bar{P}^T A \bar{P}$ has the block structure

$$\tilde{A} = \bar{P}^T A \bar{P} = \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & A_1 \end{bmatrix}$$

where $\mathbf{0}$ is the zero vector in \mathbb{R}^{n-1} and A_1 is $(n-1) \times (n-1)$ and symmetric.

- By our induction hypothesis A_1 has an orthogonal similarity transform to diagonal form:

$$D_1 = P_1^T A_1 P_1.$$

- We now apply a standard trick and embed the matrix A_1 in the $n \times n$ identity matrix. This gives

$$\tilde{P} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & P_1 \end{bmatrix}.$$

- Note that \tilde{P} is an orthogonal matrix!
- Now observe that

$$\begin{aligned} \tilde{P}^T \tilde{A} \tilde{P} &= \tilde{P}^T \bar{P}^T A \bar{P} \tilde{P} \\ &= (\bar{P} \tilde{P})^T A (\bar{P} \tilde{P}) \\ &= \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & D_1 \end{bmatrix} \end{aligned} \quad (2)$$

- This may perhaps be seen more clearly by multiplying the matrices involved in the form we use for multiplication:

$$\begin{aligned}
 & \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & P_1 \end{bmatrix} = \tilde{P} \\
 \tilde{A} &= \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & A_1 \end{bmatrix} \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & A_1 P_1 \end{bmatrix} = A \tilde{P} \\
 \tilde{P}^T &= \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & P_1^T \end{bmatrix} \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & P_1^T A_1 P_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & D_1 \end{bmatrix}
 \end{aligned}$$

- We define

$$P = \bar{P} \tilde{P}.$$

P is the product of two orthogonal matrices and hence orthogonal:

$$P^T P = (\bar{P} \tilde{P})^T (\bar{P} \tilde{P}) = \tilde{P}^T \bar{P}^T \bar{P} \tilde{P} = \tilde{P}^T \tilde{P} = I.$$

- The matrix

$$D = \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & D_1 \end{bmatrix}$$

is diagonal, and so we get the required orthogonal transform

$$P^T A P = D.$$

- **End of Proof!**

- Recall that a matrix may be defective, i.e., it may not have a complete set of n linearly independent eigenvectors.
- However, the existence of an orthogonal similarity transform implies that symmetric matrices are never defective.
- More precisely:

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

- A has n real eigenvalues, counting multiplicities.
- The dimension of the eigenspace λ equals the multiplicity of λ as a root of the characteristic equation.
- The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to distinct eigenvalues are orthogonal.
- A is orthogonally diagonalizable.



Put a little more simply: As far as the eigenvalue problem, **symmetric matrices** are as nice as can be.



We don't know enough yet to appreciate it, but the corresponding statement for linear system is that **orthogonal matrices** are as nice as can be.