NOTES ON THE SYMMETRIC MATRIX THEOREM FOR MATH 304

1. Introduction

We will assume familiarity with Chapters 1–6 of the textbook.

We state the symmetric matrix theorem in Section 2. We discuss calculations of the diagonal form in Section 3. In Section 4 we discuss problems requiring deduction of easy consequences of the theorem.

Here is what is required knowledge for the final exam regarding the Symmetric Matrix Theorem:

- The statement of the symmetric matrix theorem.
- The easy part of the proof.
- Simple consequences of the theorem.
- Given a symmetric matrix A, diagonalize it: find an orthonormal matrix P and a diagonal matrix D such that $A = PDP^T$.

In addition proofs of Lemmas 3.2 and 3.1 are simple enough to make a good question for the final exam.

After looking at the computations in Section 3, you will be wondering how you can be asked to do such a problem on the final exam. The answer is that some of the steps can be given to you in advance. For example, you could be told explicitly what the eigenvalues are, and possibly also some of the eigenvectors.

You will be relieved to know that in practice, all of this can be done by a computer. Once you get past this exam, the theoretical knowledge that lets you tell when something has gone wrong will be more important than the ability to perform all these calculations by hand in a short time.

2. Statement of the Symmetric Matrix Theorem

The informal version of the symmetric matrix theorem is "symmetric matrices can be diagonalized". The formal statement is closer to "a real matrix is symmetric iff it can be orthogonally diagonalized":

Theorem (Symmetric Matrix Theorem)**.** *A is an $n \times n$ symmetric matrix with real entries if and only if there is a diagonal matrix D and an orthogonal matrix P such that $P^{-1}AP = D$.*

Remark 2.1. In this class we have only considered matrices with real entries, so the "real" is redundant. I include it in the statement because subsequent courses will not all be so restrictive, and the adaptation to matrices with complex entries is not entirely obvious.

The proof of one direction is easy: if $D = P^{-1}AP$, or $A = PDP^{-1}$, with D diagonal and P orthogonal, then $P^{-1} = P^T$ and $D = D^T$. Since taking transposes reverses the order of multiplication, we have $A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$, so $A = A^T$ and A is symmetric. The informal statement only bothers to state the hard part, and eliminates the "orthogonal".

Note that the matrix P is a "*change of basis matrix*" and $P^T = P^{-1}$ changes basis back. As noted in Chapter 5, A is diagonalizable iff there is a basis of $\mathbb{R}^n$ consisting of eigenvectors for A. These eigenvectors are the columns of the matrix P. See Theorem 5.3.15. Furthermore, the eigenvalues of D and A are the same (Lemma 5.3.20), and they are the diagonal entries of D.

3. How to Diagonalize a Matrix

The statement of the theorem leads directly to the following type of problem: given a symmetric matrix A, find the orthogonal matrix P that diagonalizes it, and find D. In fact, in Chapters 1–6 of the book all the tools to do this are already developed.

Here is an outline of how to proceed.

- First, find the eigenvalues of A, using the characteristic polynomial, as in §5.3.
- Next, find the eigenvectors correponding to these eigenvalues, again as in §5.3. Because A is diagonalizable, there will be a basis consisting of these eigenvectors.
- The next step is to construct an orthonormal basis consisting of eigenvectors. Eigenvectors corresponding to different eigenvalues of a symmetric matrix are automatically orthogonal, (see Lemma 3.2 below), and can be normalized to have length 1, so the only case to worry about is when the dimension of an eigenspace is more than 1, i.e. we have more than one linearly independent eigenvector corresponding to a single eigenvalue. Here we can apply Gram-Schmidt to the collection of eigenvectors corresponding to this eigenvalue.
- Then P is the orthogonal matrix of orthonormal eigenvectors, and the diagonal entries of D are the eigenvalues corresponding to the columns of P, in that order.

A summary in outline form: how to find the diagonal form D of a symmetric matrix A and the diagonalizing orthogonal matrix P.

- Find the eigenvalues of A.
- Find the corresponding eigenvectors.
- Make these eigenvectors into an orthonormal basis.
- Form the matrix P whose columns are the vectors in this orthonormal basis.
- The diagonal entries of D are the eigenvalues corresponding to the columns of P.

We'll work through all of this in a couple of examples. But before we do, let's prove the statement made above, that eigenvectors corresponding to different eigenvalues are orthogonal. First a preliminary lemma.

Lemma 3.1. *Let A be a symmetric $n \times n$ matrix and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Then $A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A\mathbf{v}$.*

Proof. Use the characterization of dot product as matrix multiplication, and the fact that $A = A^T$:

$$A\mathbf{u} \cdot \mathbf{v} = (A\mathbf{u})^T\mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \mathbf{u}^T(A\mathbf{v}) = \mathbf{u} \cdot A\mathbf{v}$$

□

Now we can show that eigenvectors corresponding to distinct eigenvalues are orthogonal.

Lemma 3.2. *Let A be a symmetric $n \times n$ matrix and $\mathbf{u}, \mathbf{v}$ eigenvectors with eigenvalues $\lambda \neq \mu$ respectively. Then $\mathbf{u} \cdot \mathbf{v} = 0$.*

Proof. We have $A\mathbf{u} = \lambda\mathbf{u}$, and $A\mathbf{v} = \mu\mathbf{v}$. Now we may use the properties of the dot product: $\mu\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mu\mathbf{v} = \mathbf{u} \cdot A\mathbf{v} = A\mathbf{u} \cdot \mathbf{v}$, using the previous lemma. But $A\mathbf{u} \cdot \mathbf{v} = \lambda\mathbf{u} \cdot \mathbf{v}$. So we have $\mu\mathbf{u} \cdot \mathbf{v} = \lambda\mathbf{u} \cdot \mathbf{v}$, and $\lambda \neq \mu$, so this is possible only if $\mathbf{u} \cdot \mathbf{v} = 0$. □

Remark 3.3. These proofs are an exercise in properties of the dot product and matrix multiplication.

Example 3.4. We now turn to our first example. Given the symmetric matrix

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

orthogonally diagonalize it, that is, find an orthogonal matrix P and a diagonal matrix D so that $P^{-1}AP = D$.

Solution. Using the type of analysis we did in Chapter 5, the characteristic polynomial of A is

$$\det(A - \lambda I) = -\lambda^3 + 17\lambda^2 - 90\lambda + 144 = -(\lambda - 8)(\lambda - 6)(\lambda - 3),$$

and the eigenvalues of A are 8, 6 and 3, all of multiplicity 1.

Form the matrix

$$A - 8I = \begin{bmatrix} -2 & -2 & -1 \\ -2 & -2 & -1 \\ -1 & -1 & -3 \end{bmatrix}$$

This matrix has rank 2, and its null space has dimension 1. Using the techniques of Chapter 3, find a basis of the kernel consisting of the single vector $\mathbf{v}_1 = (-1, 1, 0)$.

Form the matrix

$$A - 6I = \begin{bmatrix} 0 & -2 & -1 \\ -2 & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

and find its kernel. A basis of the kernel is the single vector $\mathbf{v}_2 = (-1, -1, 2)$.

Form the matrix

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

and find its kernel, as in Chapter 3. A basis of the kernel is the single vector $\mathbf{v}_3 = (1, 1, 1)$.

These vectors are automatically orthogonal, as noted earlier, and so linearly independent. So they form an orthogonal basis of $\mathbb{R}^3$, and to make them into an orthonormal basis we need only normalize them to have length 1. We have $||\mathbf{v}_1|| = \sqrt{2}$, $||\mathbf{v}_2|| = \sqrt{6}$ and $||\mathbf{v}_3 = \sqrt{3}$. This gives us

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

So, the matrices P and D are

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

and we have $P^{-1} = P^T$, $A = PDP^T$. □

In the previous example, we did not have to do the "Gram-Schmidt" step, because all the eigenvalues were distinct. Let's try an example where we do have to do this extra step.

Example 3.5. Given the symmetric matrix

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

orthogonally diagonalize it, that is, find an orthogonal matrix P and a diagonal matrix D so that $P^{-1}AP = D$.

Solution. Using the type of analysis we did in Chapter 5, the characteristic polynomial of A is

$$\det(A - \lambda I) = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2)$$

and the eigenvalues of A are -2 (multiplicity 1) and 7 (multiplicity 2). Form the matrix

$$A - (-2)I = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}$$

It has rank 2, and its kernel has dimension 1, with basis consisting of the single vector $\mathbf{v}_1 = (-1, -1/2, 1)$.

Form the matrix

$$A - 7I = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix}$$

This matrix has rank 1, so its kernel has dimension 2. The two vectors $\mathbf{v}_2 = (1, 0, 1)$, $\mathbf{v}_3 = (-1/2, 1, 0)$ are a basis of the kernel. Note that $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent, but not orthogonal. Using the Gram-Schmidt procedure, we form $\mathbf{u}_2 = \mathbf{v}_2$ and

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_2 \cdot \mathbf{v}_3}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = (-1/4, 1, 1/4)$$

Now normalize $\mathbf{v}_1, \mathbf{u}_2, \mathbf{u}_3$ to vectors of length 1 and use them as the columns of the matrix P. So we have

$$P = \begin{bmatrix} -2/3 & 1/\sqrt{2} & -1/\sqrt{18} \\ -1/3 & 0 & 4/\sqrt{18} \\ 2/3 & 1/\sqrt{2} & 1/\sqrt{18} \end{bmatrix} \qquad D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

Again, note that $P^{-1} = P^T$, and $A = PDP^T$. □

4. Simple Consequences of the Symmetric Matrix Theorem

Here we consider a few problems that require short deductions based on the theorem.

(**True** | **False**) An $n \times n$ matrix that is orthogonally diagonalizable must be symmetric.

True. This is the easy half of the Symmetric Matrix Theorem.

(**True** | **False**) An $n \times n$ symmetric matrix has n distinct real eigenvalues.

False. The $n \times n$ identity matrix I is diagonalizable, and has only one eigenvalue, namely 1. The 3×3 matrix A in Example 3.5 above has only two distinct eigenvalues.

(**True** | **False**) If A is a symmetric, invertible matrix, then $A^{-1} = A^T$

False. The given condition defines an orthogonal matrix. Not every symmetric matrix is orthogonal. Consider, for example $2I$.

Now let's try a multiple choice question:

If A is an $n \times n$ real symmetric matrix, then which of the following are true?

(a) Each eigenvalue of A is real.
(b) If A is invertible, then its inverse is also symmetric.
(c) If $Ax = 2x$ and $Ay = 3y$ then $x \cdot y = 0$.
(d) If λ_1 and λ_2 are two different eigenvalues of A and W_1 and W_2 are the corresponding eigenspaces, then W_1 and W_2 are orthogonal sets

Answer: All of them. Note that (c) and (d) are about orthogonality of eigenvalues, discussed earlier in these notes. (a) is part of the Symmetric Matrix Theorem. For part (b), by the Symmetric Matrix Theorem, $A = PDP^T$, with P orthogonal (and hence invertible with $P^{-1} = P^T$). We can write $A^{-1} = (PDP^T)^{-1} = (P^T)^{-1}D^{-1}P^{-1} = PD^{-1}P^T$. Note that D^{-1} is also a diagonal matrix and so equal to its transpose. Now compute

$$(A^{-1})^T = (PD^{-1}P^T)^T = (P^T)^T(D^{-1})^T P^T = PD^{-1}P^T = A^{-1}$$

showing that A^{-1} is symmetric.

Note the similarity of this computation to the work for the "easy half" of the Symmetric Matrix Theorem.