# 22m:033 Notes: 5.3 Diagonalization

Dennis Roseman University of Iowa Iowa City, IA

 $http://www.math.uiowa.edu/{\sim}roseman$ 

April 19, 2010

### 1 Diagonalizable matrices

**Remark 1.1** The motivation for considering all that follows in this section is found in Section 5.4, where the geometric meaning of "similar" is made clear.

**Definition 1.2** A square matrix is **diagonalizable** if it is similar to a diagonal matrix

**Remark 1.3** In other words, A is diagonalizable if there is an invertible matrix P such that  $A = PDP^{-1}$  where D is a diagonal matrix.

Another way of putting this is: an invertible matrix Q such that  $D = QAQ^{-1}$  where D is a diagonal matrix since if we let  $Q = P^{-1}$  then

$$QAQ^{-1} = P^{-1}AP = P^{-1}(PDP^{-1})P =$$
  
 $(P^{-1}P)D(P^{-1}P) = IDI = D.$ 

So how can one find diagonalizable matrices? One way is by the following theorem. **Proposition 1.4** If A is an  $n \times n$  matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact,  $A = PDP^{-1}$  where D is a diagonal matrix if and only if the column vectors of P are n linearly independent eigenvectors of A.

The next result give us a better handle on how to use Proposition 1.4.

**Proposition 1.5** If A is an  $n \times n$  matrix A with distinct eigenvalues  $\lambda_1, \ldots, \lambda_p$ .

- 1. For  $1 \leq k \leq p$  the dimension of the eigenspace for  $\lambda_k$  is less then or equal to the algebraic multiplicity of the eigenvalue  $\lambda_k$  as a zero of the characteristic polynomial of A.
- 2. A is diagonalizable if and only if the sum of the dimensions of the distinct eigenspaces equals n and this happens if and only if, for each k the dimension of the eigenspace corresponding to  $\lambda_k$  is exactly equal to the algebraic multiplicity of the eigenvalue  $\lambda_k$ .

3. If A is diagonalizable and  $\mathcal{B}_k$  is a basis for the eigenspace corresponding to  $\lambda_k$ , then the total collection of the vectors  $\mathcal{B}_1, \ldots, \mathcal{B}_p$  forms an eigenvector basis for  $\mathbb{R}^n$ .

### 2 Some examples on mutiplicity

**Example 2.1** Consider the polynomial

 $\lambda^7 + 5\lambda^6 - 6\lambda^5 - 50\lambda^4 + 5\lambda^3 + 153\lambda^2 - 108 = (\lambda - 1)(\lambda + 1)(\lambda - 2)^2(\lambda + 3)^3.$ 

In this polynomial, +1 and -1 have multiplicity 1, 2 has multiplicity 2 and -3 has multiplicity 3.

If this were the characteristic polynomial for some  $7 \times 7$  matrix, we could conclude there is an eigenspace of dimension one for eigenvalues 1 and -1. But for eigenvalue 2, this dimension could be 1 or 2; for eigenvalue -3 it could be 1, 2 or 3.

The following two examples show that multiplicity does not give a clue to the dimensions of the eigenspace. **Example 2.2** If  $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  then the characteristic polynomial is  $(1 - \lambda)^2$ . Since this is the identity matrix it is clear that the eigenspace for  $\lambda = 1$  is two dimensional. For a basis of eigenvectors, we can take

$$\{e_1, e_2\} = \{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \}.$$

Also any basis for  $R^2$  will be a basis for this eigenspace.

In summary, in this example we have an eigenvalue of multiplicity 2 whose corresponding eigenspace has dimension 2.

Contrast the above with this next example which is a simple shear.

**Example 2.3** If  $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  then the characteristic polynomial is  $(1 - \lambda)^2$ .

The eigenspace for this value is the null space of

$$\left(\begin{array}{rrr} 1-1 & 1\\ 0 & 1-1 \end{array}\right) = \left(\begin{array}{rrr} 0 & 1\\ 0 & 0 \end{array}\right)$$

For a basis of eigenvectors, we can take

$$\{e_1\} = \{ \begin{pmatrix} 1\\ 0 \end{pmatrix} \}.$$

In summary, in this example we have an eigenvalue of multiplicity 2 whose corresponding eigenspace has dimension 1.

### 3 Algorithm for diagonalizing matrices

Given a square matrix A here is how to investigate diagonalizability of A using Proposition 1.4 and Proposition 1.5.

- 1. Find the eigenvalues of A.
  - (a) If there are none, A is *not* diagonalizable and we can stop.
  - (b) If there are n distinct eigenvalues, then A is diagonalizable. If we are further expected to find P we continue.
  - (c) If there are some eigenvalues, but less than n, we do not yet know whether or not A is diagonalizable and we must continue.

- 2. For each eigenvalue determine a basis for the eigenspace. Note: if we are in case 1b, we know each eigenspace is one-dimensional.
- 3. If the total number of basis vectors found in step 2 is less than n, then A is *not* diagonalizable and we can stop.
- 4. At this point we have determined if A is diagonalizable or not.

If it is diagonalizable, we might want to know two additional things: a matrix D and a matrix P. Except for a few special cases, neither of these is uniquely determined by A. In the case that we have less eigenvalues than n, there will be infinitely many possible candidates for P. (If you know P you can simply calculate a corresponding D using  $D = PAP^{-1}$ ).

- (a) If we only want to have a diagonal matrix D which is similar to A, we can let D be a diagonal matrix whose non-zero entries are the eigenvalues for A. The number of times an eigenvalue will repeat is exactly equal to the dimension of the corresponding eigenspace.
- (b) If we need to find a matrix P we construct one whose columns are all the basis vectors found in

step 2.

#### 4 Examples to illustrate algorithm

- 1. (a) The matrix  $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$  has characteristic polynomial  $\lambda^2 + 1$ . This has no (real) eigenvalues and thus is *not* diagonalizable. Note: geometrically this is clear since this is a rotation of  $\frac{\pi}{2}$ .
  - (b) The matrix  $\begin{pmatrix} 2 & 3 \\ 0 & 5 \end{pmatrix}$  has characteristic polynomial  $(2 \lambda)(5 \lambda)$ . Here n = 2 and we have two distinct eigenvalues, this matrix *is* diagonalizable.
  - (c) The matrix  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  has characteristic polynomial  $(1 \lambda)^2$  and so has only one eigenvalue, so we can make no conclusion so far (were it not for the fact we have already analyzed this as Example 2.3).
- 2. To do this step we find, for each eigenvalue  $\lambda_k$  a basis for the null space of  $A - \lambda_k I$ .

3. Suppose our matrix is  $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ . Note: this is a

higher dimensional version of Example 2.3. Clearly the characteristic polynomial is  $(1 - \lambda)^3$  and we have only one eigenvalue, 1. The eigenspace correspond-

ing to this value is the null space of  $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ . If  $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$  is in this null space we see that we must have z = 0, y = z = 0, and x is a free variable. So this null space is one dimensional with basis  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ . We

conclude this matrix is not diagonalizable.

4. (a) If somehow you knew that you had a  $5 \times 5$  matrix A which was diagonalizable with one eigenvalue of 2 with multiplicity 3 and one eigenvalue with eigenvalue -5 with multiplicity 2, then you could immediately conclude that A is similar to

$$\begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & -5 \end{pmatrix}$$
. We could also conclude  
that A is similar to  
$$\begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 & 0 \\ 0 & 0 & -5 & 0 & 0 \\ 0 & 0 & -5 & 0 & 0 \\ 0 & 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & -5 \end{pmatrix}, \text{ etc}$$

(b) If we somehow knew we had a matrix A with two eigenvalues, say 3 and -5 and that a basis for the eigenspace for eigenvalue 3 is

$$\left\{ \begin{pmatrix} 1\\0\\1\\0\\1 \end{pmatrix}, \begin{pmatrix} 0\\1\\1\\2\\0 \end{pmatrix} \right\}$$

and a basis for the eigenspace for eigenvalue -5 is

$$\left\{ \begin{pmatrix} 1\\2\\2\\1\\1\\1 \end{pmatrix}, \begin{pmatrix} 0\\0\\1\\0\\2 \end{pmatrix}, \begin{pmatrix} 1\\0\\2\\1\\0 \end{pmatrix} \right\}$$

Then we could use these vectors as columns and

define 
$$P = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 0 & 0 \\ 1 & 1 & 2 & 1 & 2 \\ 0 & 2 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 & 0 \end{pmatrix}$$
. We would then  
know that  $PAP^{-1} = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{pmatrix}$  with-

out having to calculate  $P^{-1}$ , which by the way is  $\begin{pmatrix} 11 & 2 & 8 & 5 & 4 \end{pmatrix}$ 

$$\begin{pmatrix} \frac{11}{7} & -\frac{2}{7} & -\frac{8}{7} & \frac{5}{7} & \frac{4}{7} \\ \frac{2}{7} & -\frac{1}{7} & -\frac{4}{7} & \frac{6}{7} & \frac{2}{7} \\ -\frac{1}{7} & \frac{4}{7} & \frac{2}{7} & -\frac{3}{7} & -\frac{1}{7} \\ -\frac{5}{7} & -\frac{1}{7} & \frac{3}{7} & -\frac{1}{7} & \frac{2}{7} \\ -\frac{3}{7} & -\frac{2}{7} & \frac{6}{7} & -\frac{2}{7} & -\frac{3}{7} \end{pmatrix}.$$

## 5 Diagonalization and high powers of matrices

It is easy to calculate high powers of a diagonal matrix:

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix}$$
$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 27 & 0 & 0 \\ 0 & 0 & 64 & 0 \\ 0 & 0 & 0 & 125 \end{pmatrix}$$

Suppose A is diagonalizable, so that  $A = PDP^{-1}$ .

If we actually know the matrix P we can use it to

calculate high powers of A as follows:

$$\begin{array}{rcl} A^2 &=& (PDP^{-1})(PDP^{-1}) \\ &=& (PD)(P^{-1}P)(DP^{-1}) \\ &=& (PD)I(DP^{-1}) \\ &=& P(DID)P^{-1} \\ &=& PD^2P^{-1} \end{array}$$

It is clear that similarly

$$A^{3} = (PDP^{-1})(PDP^{-1})(PDP^{-1})$$
  
=  $PD^{3}P^{-1}$ 

and in general for any n,

$$A^n = PD^n P^{-1}$$

Here is Exercise 3 in this section:

**Example 5.1** Given that

$$A = \begin{pmatrix} a & 0 \\ 3(b-a) & b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}$$

We can conclude that since  $D^k = \begin{pmatrix} a^k & 0 \\ 0 & b^k \end{pmatrix}$ ,  $A^k = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} a^k & 0 \\ 0 & b^k \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} a^k & 0 \\ 3a^k - 3b^k & b^k \end{pmatrix}$ .

#### 6 A large example

**Example 6.1** Lets try and diagonalize

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 2 & 0 & 2 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 \end{pmatrix}$$

Since this is triangular, we quickly calculate that the characteristic polynomial of A is

$$(\lambda - 1)(\lambda + 1)(\lambda - 2)^2(\lambda + 3)^3.$$

By the way this is the polynomial considered in Example 2.1.

The eigenspace with eigenvalue 1 must be one-dimensional and we can easily find that  $e_1$  is an eigenvector. Also, the eigenspace with eigenvalue -1 must be one-dimensional and we can easily find that  $e_2$  is an eigenvector.

For the eigenspace for eigenvalue 2, we look at the null space of:

(1-2)	0	1	0	1	0	1	
0	-1 - 2	0	2	0	2	0	
0	0	2 - 2	0	1	0	1	
0	0	0	2 - 2	0	0	0	=
0	0	0	0	-3 - 2	0	0	
0	0	0	0	0	-3 - 2	0	
0	0	0	0	0	0	-3-2	
	$ \left(\begin{array}{c} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right) $	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$     \begin{array}{c}       0 \\       2 \\       0 \\       0 \\       0 \\       5 \\       0 \\       -5 \\       0 \\       0 \\       -5 \\       0 \\       0 \\       -5 \\       0 \\       0 \\       -5 \\       0 \\       0 \\       -5 \\       0 \\ $	$ \begin{array}{c} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ -5 \end{array} $		

We row reduce this and get

•

Letting  $x_4 = t$  and  $x_3 = s$  we see the null space is

$$\begin{pmatrix} s \\ \frac{2}{3}t \\ s \\ t \\ 0 \\ 0 \\ 0 \end{pmatrix} = s \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ \frac{2}{3} \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

We can multiply the second vector by 3 and see that a basis for this eigenspace is given by:

$$\left\{ \begin{pmatrix} 0 \\ 2 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

•

Next for the eigenvalue  $\mbox{-}3,$  we look at the null space of

$$\begin{pmatrix} 1+3 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & -1+3 & 0 & 2 & 0 & 2 & 0 \\ 0 & 0 & 2+3 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2+3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3+3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3+3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3+3 \end{pmatrix} =$$

17

This row reduces to

So we know from the fourth row that  $x_4 = 0$ . For the third row, if we let  $x_7 = s$  and  $x_5 = t$ , we see that  $x_3 = -\frac{s}{5} - \frac{t}{5}$ . For the second row if we let  $x_6 = r$  then  $x_2 = -r$ . Finally from the top row we get:  $x_1 = -\frac{s}{5} - \frac{t}{5}$ . So the null space is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} -\frac{s}{5} - \frac{t}{5} \\ -r \\ -\frac{s}{5} - \frac{t}{5} \\ 0 \\ t \\ r \\ s \end{pmatrix} = s \begin{pmatrix} -\frac{1}{5} \\ 0 \\ -\frac{1}{5} \\ 0 \\ -\frac{1}{5} \\ 0 \\ 0 \\ 1 \end{pmatrix} + r \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -\frac{1}{5} \\ 0 \\ -\frac{1}{5} \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Clearing the fractions, we see that a basis for the eigenspace with eigenvalue -3 is:

$$\left\{ \begin{pmatrix} -1\\ 0\\ -1\\ 0\\ 0\\ 0\\ 0\\ 5 \end{pmatrix}, \begin{pmatrix} 0\\ -1\\ 0\\ 0\\ 0\\ 1\\ 0 \end{pmatrix}, \begin{pmatrix} -1\\ 0\\ -1\\ 0\\ 5\\ 0\\ 5 \end{pmatrix} \right\}.$$

We next assemble our matrix  ${\cal P}$  using these eigenvectors as columns and get

As a check, we can calculate the inverse  $P^{-1}$  (details

omitted—do you recall how to do this?):

$$\left(\begin{array}{cccccccccc} 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -\frac{2}{3} & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{1}{5} & 0 & \frac{1}{5} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} \\ 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 \end{array}\right),$$

and then check that: