Linear Algebra Notes

Chapter 9

MULTIPLE EIGENVALUES AND NILPOTENT MATRICES

Before taking up multiple eigenvalues, we begin with a

Prelude: Take a matrix A, and an invertible matrix B. We get a new matrix $B^{-1}AB$, which is called **conjugate** to A. We have seen that conjugate matrices have many features in common. For example, the formulas

$$\det(XY) = \det(YX), \qquad \operatorname{tr}(XY) = \operatorname{tr}(YX),$$

valid for any two matrices X, Y, imply that

$$\det(B^{-1}AB) = \det(A), \qquad \operatorname{tr}(B^{-1}AB) = \operatorname{tr}(A).$$

Now A and $B^{-1}AB$ have even more features in common. For example, the characteristic polynomial is made out of the trace and determinant, so A and $B^{-1}AB$ have the same characteristic polynomials. This means they have the same eigenvalues. They don't have the same eigenvectors, in general. However, if $A\mathbf{u} = \lambda\mathbf{u}$, then

$$B^{-1}AB(B^{-1}\mathbf{u}) = B^{-1}Au = \lambda B^{-1}\mathbf{u},$$

so $B^{-1}\mathbf{u}$ is an eigenvector for $B^{-1}AB$ with eigenvalue λ. In summary,

Proposition 1. *Conjugate matrices A and $B^{-1}AB$ have the same trace and determinant, the same characteristic polynomial, and the same eigenvalues. If $\mathbf{u}$ is a λ-eigenvector for A, then $B^{-1}\mathbf{u}$ is a λ-eigenvector for $B^{-1}AB$.*

end of Prelude.

Now to business. We have seen that if A has distinct real eigenvalues λ, μ, then we can find B so that

$$B^{-1}AB = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}. \tag{9a}$$

By Proposition 1, the diagonal matrix in (9a) shares many properties with A. However, being diagonal, it is for many purposes simpler than A. We seek an equation like (9a) in the remaining cases: Multiple eigenvalues, and Complex eigenvalues. In this chapter we consider multiple eigenvalues. We will always assume that the entries of A are real.

We get multiple eigenvalues when

$$P_A(x) = (x - \lambda)^2 = x^2 - 2\lambda x + \lambda^2.$$

This means

$$2\lambda = \operatorname{tr}(A) \qquad \text{and} \qquad \lambda^2 = \det(A),$$

so

$$4\det(A) = [\mathrm{tr}(A)]^2 \tag{9b}$$

The matrices with only one eigenvalue are those whose trace and determinant satisfy (9b). Note that $\lambda = \frac{1}{2}\,\mathrm{tr}(A)$ will be real, since the entries of A are real.

There are then two possibilities for A. Either $A = \lambda I$, or $A \neq \lambda I$. Assume $A \neq \lambda I$. Then, in contrast to previous situations, we cannot conjugate A to a diagonal matrix. But we can do almost as well:

Proposition 2. *If A has only one eigenvalue λ, and $A \neq \lambda I$, then there is an invertible matrix B such that*

$$B^{-1}AB = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

The difference from the distinct-eigenvalue case is the 1 in the upper right entry, which cannot be made zero, no matter how we choose B.

Proof. We give a recipe for finding B. As before, first find a λ-eigenvector $\mathbf{u}$, for A. Now take any vector $\mathbf{v}$ which is not proportional to $\mathbf{u}$. Let B_1 be the matrix defined by

$$B_1\mathbf{e}_1 = \mathbf{u}, \qquad B_1\mathbf{e}_2 = \mathbf{v}.$$

We do a familiar computation

$$B_1^{-1}AB_1\mathbf{e}_1 = B_1^{-1}A\mathbf{u} = \lambda B_1^{-1}\mathbf{u} = \lambda\mathbf{e}_1.$$

So

$$B_1^{-1}AB_1 = \begin{bmatrix} \lambda & g \\ 0 & h \end{bmatrix},$$

for some numbers g, h that we don't yet know. But in fact, $h = \lambda$. To see this, note that

$$2\lambda = \mathrm{tr}(A) = \mathrm{tr}(B_1^{-1}AB_1) = \mathrm{tr}\begin{bmatrix} \lambda & g \\ 0 & h \end{bmatrix} = \lambda + h,$$

(we used Proposition 1 at the second equality) so indeed, $\lambda = h$. So we in fact have

$$B_1^{-1}AB_1 = \begin{bmatrix} \lambda & g \\ 0 & \lambda \end{bmatrix}.$$

Now $g \neq 0$, since if $g = 0$ we would have $A = \lambda I$, but we are assuming $A \neq \lambda I$. Now conjugate further using $\begin{bmatrix} g & 0 \\ 0 & 1 \end{bmatrix}$, and get

$$\begin{bmatrix} g^{-1} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \lambda & g \\ 0 & \lambda \end{bmatrix}\begin{bmatrix} g & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

This shows that if we take

$$B = B_1\begin{bmatrix} g & 0 \\ 0 & 1 \end{bmatrix},$$

then

$$B^{-1}AB = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix},$$

as claimed in Proposition 2. $\square$

Now if you want to compute A^n, you could first note that

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^n = \begin{bmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{bmatrix} \tag{9b}$$

so

$$A^n = (B \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} B^{-1})^n = B \begin{bmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{bmatrix} B^{-1}. \tag{9c}$$

Actually, there is a much better formula for A^n, but first let's have an example to illustrate what we've seen so far.

Example:

$$A = \begin{bmatrix} 0 & 4 \\ -1 & 4 \end{bmatrix}.$$

We find $P_A(x) = (x-2)^2$ so $\lambda = 2$ is a multiple eigenvalue of A. Our formula for eigenvectors gives $(b, \lambda - a) = (4, 2)$. Scaling, we can take the eigenvector to be $\mathbf{u} = (2, 1)$. Now we choose any another vector $\mathbf{v}$ which is not proportional to $\mathbf{u}$. Let us take $\mathbf{v} = \mathbf{e}_1$, so we have

$$B_1 = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}.$$

We then compute

$$B_1^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix},$$

and

$$B_1^{-1}AB_1 = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 0 & 4 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}.$$

So $g = -1$ and we take

$$B = B_1 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -1 & 0 \end{bmatrix}.$$

To check our calculations we compute

$$B^{-1}AB = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

So $B = \begin{bmatrix} -2 & 1 \\ -1 & 0 \end{bmatrix}$ does the job, as predicted by Proposition 2.

Continuing further using equation (9c), we have $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}^n = \begin{bmatrix} 2^n & n2^{n-1} \\ 0 & 2^n \end{bmatrix}$, so

$$A^n = B \begin{bmatrix} 2^n & n2^{n-1} \\ 0 & 2^n \end{bmatrix} B^{-1} = \begin{bmatrix} 2^n - n2^n & n2^{n+1} \\ -n2^{n-1} & 2^n + n2^n \end{bmatrix}.$$

We could write this as a sum

$$A^n = \begin{bmatrix} 2^n & 0 \\ 0 & 2^n \end{bmatrix} + n2^{n-1} \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix}.$$

The first matrix in this sum is the diagonal matrix whose diagonal entries are the powers of the eigenvalue of A. The second matrix (considered without the factor $n2^{n-1}$), is almost A, but with the eigenvalue 2 subtracted from the diagonal entries of A. Let us call this matrix A_0. So

$$A_0 = \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0-2 & 4 \\ -1 & 4-2 \end{bmatrix} = A - 2I.$$

Thus, we could write our formula for A^n as

$$A^n = 2^n I + n2^{n-1} A_0.$$

This is a formula is much simpler than (9c), and it holds in general.

Proposition 3. *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *be a matrix with* $P_A = (x-\lambda)^2$*. Define the new matrix*

$$A_0 = \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix}.$$

Then $A_0^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, *and*

$$A^n = \lambda^n I + n\lambda^{n-1} A_0.$$

The point of Proposition 3 is that it allows us to compute A^n in a much easier way than equation (9c). We don't have to find B to use Proposition 3. However, we will use B to prove Proposition 3. The key point is that $A_0^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Proof. We use our earlier formula

$$A = B \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} B^{-1},$$

from Proposition 2. Note that

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = \lambda I + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

so

$$A = B \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} B^{-1} = B(\lambda I + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix})B^{-1} = B(\lambda I)B^{-1} + B \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} B^{-1}.$$

Since λI commutes with every matrix, we have $B(\lambda I)B^{-1} = \lambda I$, so

$$A = \lambda I + B \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} B^{-1}.$$

But also $A = \lambda I + A_0$, so

$$\lambda I + A_0 = A = \lambda I + B \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} B^{-1}.$$

Subtracting λI from both sides, we get

$$A_0 = B \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} B^{-1}.$$

Hence

$$\begin{aligned} A_0^2 &= B \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} B^{-1} B \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} B^{-1} \\ &= B \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 B^{-1} \\ &= B \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} B^{-1} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

because $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. That was the key point. Now we have

$$A^n = (\lambda I + A_0)^n.$$

Inside the parentheses, we have two matrices. One of them, λI, commutes with all matrices, and the other one, A_0, squares to zero.

Now let us pause, and temporarily empty our minds of all thoughts of A and A_0. Clear the table. On this empty table, we put two matrices X, Y with two properties. First, they commute with each other, and second, $Y^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Now we add them and raise to powers. The second power is

$$(X + Y)^2 = (X + Y)(X + Y) = XX + XY + YX + YY = X^2 + 2XY,$$

because $XY = YX$ and $Y^2 = 0$. In the general power

$$(X + Y)^n = (X + Y)(X + Y) \cdots (X + Y)$$

we get each term by choosing either X or Y in each factor. Since they commute, we can put the Y's together, which will give zero unless we have at most one power of Y. So we either choose X in each factor (one such term), or we choose Y in one factor and X in all the others (n such terms). Thus

$$(X + Y)^n = X^n + nX^{n-1}Y.$$

Now we resurrect A and A_0, and finish our calculation: Since $(\lambda I)^n = \lambda^n I$, we get

$$A^n = (\lambda I + A_0)^n = \lambda^n I + n\lambda^{n-1} A_0,$$

as claimed. $\square$

A matrix is called **nilpotent** if some power of it is zero. It turns out, for 2×2 matrices Y as we are considering here, that if $Y^n = 0$, then already $Y^2 = 0$ (see exercise 9.4). So a 2×2 matrix is nilpotent if either it is already zero, or if it squares to zero.

We may summarize the main points in this chapter as follows. Any matrix A with a multiple eigenvalue λ is a scalar matrix λI plus a nilpotent matrix $A_0 = A - \lambda I$. If A is not itself a scalar matrix, then $A_0 \neq 0$. In that case A is conjugate to $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$, and A_0 is conjugate to $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. This last matrix does not involve λ. In particular, all matrices A with multiple eigenvalues have their nilpotent parts A_0 being conjugate to the same matrix $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Exercise 9.1. *For each of the following matrices A, find the eigenvalue λ, and a matrix B such that $B^{-1}AB = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$.*

(a) $A = \begin{bmatrix} 6 & 1 \\ -1 & 8 \end{bmatrix}$
(b) $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$
(c) $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$
(d) $A = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}$

(To make life easier for the grader, please choose $\mathbf{v} = \mathbf{e}_1$ in each problem. Your B's will all be the same, having 1 in the upper right corner.)

Exercise 9.2. *Show that every eigenvector of $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ is proportional to $\mathbf{e}_1$. (Hint: Just calculate.)*

Exercise 9.3. *Let A be a matrix with multiple eigenvalue λ, and let $\mathbf{u}$ be an eigenvector of A. Assume that A is not a scalar matrix. Prove the following statements.*

(a) *Every eigenvector of A is proportional to $\mathbf{u}$. (Hint: Proposition 1.)*
(b) $A_0\mathbf{u} = (0, 0)$
(c) *If $\mathbf{v}$ is a vector for which $A_0\mathbf{v} = (0, 0)$, then $\mathbf{v}$ is an eigenvector of A.*

(Taken together, these facts mean that A has exactly one eigenline, which consists precisely of the vectors sent to $(0, 0)$ by A_0.)

Exercise 9.4. *Suppose $A^n = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ for some $n \geq 0$.*

(a) *What are the eigenvalues of A? (Hint: Apply A repeatedly to both sides of the equation $A\mathbf{u} = \lambda\mathbf{u}$.)*
(b) *Compute A^2. (Hint: Use Proposition 3, and the value of λ from part (a).)*

Exercise 9.5. *Suppose $\lambda = 0$ is the only eigenvalue of A. Does this imply that A is nilpotent?*

Exercise 9.6. *Show that the nilpotent matrices are precisely those of the form*

$$A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \quad \textit{with} \quad a^2 + bc = 0.$$

(You have to show two things: First, that every nilpotent matrix has this form. Second, that every matrix of this form is nilpotent.)