THE CLASSIFICATION OF MATRICES

It is a fundamental result of Linear Algebra that, in any dimension n, there is a small number of simple 'standard forms' so that every $n \times n$ matrix is equivalent to one in standard form. 'Equivalent' is understood in the usual sense of the change of basis formula: if Λ is the standard form of A, there is an invertible matrix P so that $P^{-1}AP = \Lambda$. (Λ is uniquely defined given A-except for the order in which the eigenvalues appear- but many different P will work).

This notion of equivalence is natural- it means that, with respect to the basis of $\mathbb{R}^n$ given by the columns of P, the action of A is very simple. In addition, reduction to standard form is an extremely useful procedure- for example, to solve systems of linear differential equations or difference equations.

The simplest situation occurs if A is diagonalizable with real eigenvalues- then the standard form is purely diagonal, with the eigenvalues themselves as diagonal entries. Allowing complex eigenvalues (that is, when $\mathbb{C}^n$ has a basis consisting of eigenvectors of A) introduces the minor complication of 2×2 blocks on the diagonal, of the form $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.When $n = 2$, there is only one other possibility (A has a double real eigenvalue and is not diagonalizable- example 1 below). When $n = 3$, there are 3 non-diagonalizable 'standard forms'. The purpose of this handout is to describe the cases $n = 2, 3$ completely. The general case is dealt with in 'advanced' courses, but follows the same pattern. But keep in mind that, if you pick an $n \times n$ matrix 'at random', the probability of picking a non-diagonalizable one is ZERO. (It is like trying to hit (0,0) by throwing darts at $\mathbb{R}^2$- you'd have to be very unlucky.)

Example 1.

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

The characteristic polynomial is $p(\lambda) = (\lambda - 2)^2$, so 2 is the only eigenvalue. To compute its eigenspace, we solve the homogeneous system:

$$(A - 2I)v = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0,$$

and obtain:

$$E(2) = \{t(1, 1); t \in \mathbb{R}\}.$$

So $E(2)$ is one-dimensional; since there are no other eigenvalues, A is not diagonalizable. To find a basis of $\mathbb{R}^2$ in which A acts in a simple way, we need another vector. We find one by picking an eigenvector (say, $v = (1, 1)$ itself) and solving the non-homogeneous system for w:

$$(A - 2I)w = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = v = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

which is really the single equation $x - y = 1$. Any solution will do; take, say, $w = (1, 0)$. Then the action of A on the basis $\mathcal{B} = \{v, w\}$ is extremely simple:

$$Av = 2v; \quad Aw = 2w + v.$$

This means the matrix of A in this basis is:

$$[A]_{\mathcal{B}} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} := \Lambda.$$

This Λ is the 'standard form' of A. A matrix P satisfying $P^{-1}AP = \Lambda$ is given by the vectors of $\mathcal{B}$ as columns: (*Watch the order!*)

$$P = [v|w] = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Note that, although $(A - 2I)w \neq 0$ (since w is not an eigenvector), we do have:

$$(A - 2I)^2 w = (A - 2I)(A - 2I)w = (A - 2I)v = 0,$$

since v *is* an eigenvector. w is said to be a 'generalized eigenvector of order 2'.

The general fact behind the classification is that, even when there is no basis of $\mathbb{R}^n$ consisting of eigenvectors of A, there is always a (computable!) basis consisting of *generalized* eigenvectors (that is v and λ satisfying $(A - \lambda I)^k v = 0$ for some $k \geq 0$ (k is at most n). This very simple and general fact should be known to the average educated citizen, but usually gets buried under a mountain of impenetrable jargon in an 'advanced' course. (Behind the jargon are important concepts that make a proof for arbitrary n feasible.)

In summary, there are three 'standard forms' in the 2×2 case:

$$\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \quad \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

In the first one, A is diagonalizable with real eigenvalues (λ may be equal to μ, but only if A is just λ times the identity matrix, already in the standard basis). In the second, we take $b > 0$; A has a complex-conjugate pair of eigenvalues: $\lambda = a \pm ib$. The third case is non-diagonalizable: λ has algebraic multiplicity two (i.e., is a double root of the characteristic polynomial), but 'geometric multiplicity' 1 (dimension of the eigenspace). The corresponding matrices $P = [v|w]$ are easily computable: in the first case, v (resp. w) is an eigenvector for λ (resp. μ); in the second, v (resp. w) is the real part (resp. imaginary part) of an eigenvector for $a + ib$ (where $b > 0$). In the last case, v is an eigenvector for λ and w is any solution of the non-homogeneous system $(A - \lambda I)w = v$.

Moving on to $n = 3$, the diagonalizable (or complex-diagonalizable) cases are easy; their standard forms are:

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \gamma \end{bmatrix}, \quad \begin{bmatrix} a & b & 0 \\ -b & a & 0 \\ 0 & 0 & \gamma \end{bmatrix}.$$

In the first case, A is diagonalizable with real eigenvalues; in fact, two of λ, μ, γ may coincide, or even all three (if A is already of the form λ times the identity). In any case, the dimensions of the eigenspaces add up to 3. In the second, A has a real eigenvalue γ and a complex-conjugate pair of eigenvalues $a \pm ib$. The diagonalizing matirx P is given by $P = [v|w|u]$, where u is an eigenvector for γ and v (resp. w) is the real part (resp. imaginary part) of an eigenvector for $a + ib$ (where $b > 0$). Just as in two dimensions.

There are three non-diagonalizable cases (depending on how you count!) The first one has standard form:

$$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix}.$$

(*We assume* $\lambda \neq \mu$!) This is easily understood as a 'non-diagonalizable 2×2 block' for the eigenvalue λ, followed by a real eigenvalue μ. Here we have: dim(E(λ))=dim(E(μ))=1. The difference is that μ has algebraic multiplicity 1- it is a simple root of the characteristic polynomial-while λ is a double root (i.e., it has algebraic multiplicity 2). The 'diagonalizing matrix' $P = [v|w|u]$ consists of any eigenvectors v (resp. u) for λ (resp. μ), plus a 'generalized eigenvector' w for λ, satisfying $(A - \lambda I)w = v$. The action of A on the basis $\mathcal{B} = \{v, w, u\}$ is:

$$Av = \lambda v, \quad Aw = \lambda w + v, \quad Au = \mu u.$$

The two other cases are best explained using explicit examples.

Example 2.

$$A = \begin{bmatrix} -3 & 9 & 1 \\ -4 & 9 & 1 \\ 0 & 0 & 3 \end{bmatrix}.$$

Here the characteristic polynomial is $p(\lambda) = (\lambda - 3)^3$. So 3 is the only eigenvalue, with algebraic multiplicity 3. To compute its eigenspace, we solve the homogeneous system:

$$(A - 3I)v = \begin{bmatrix} -6 & 9 & 1 \\ -4 & 6 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0.$$

The first two equations give:

$$-2x + 3y + (1/3)z = 0, \quad -2x + 3y + (1/2)z = 0,$$

so it follows that $z = 0$ and $2x = 3y$. Hence $E(3)$ is one-dimensional:

$$E(3) = \{t(3, 2, 0), t \in \mathbb{R}\}.$$

To get a basis for $\mathbb{R}^3$ we need two more vectors, which will be 'generalized eigenvectors' for $\lambda = 3$. We use the same method as when $n = 2$; pick a (non-zero) eigenvector (say, $v = (3, 2, 0)$ itself), and solve the non-homogeneous system:

$$(A - 3I)w = \begin{bmatrix} -6 & 9 & 1 \\ -4 & 6 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = v = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}.$$

This reduces to the equations:

$$-2x + 3y + (1/3)z = 1, \quad -2x + 3y + (1/2)z = 1,$$

so again it follows that $z = 0$, and x, y satisfy: $-2x + 3y = 1$ (*). Any solution of this will do; say, $x = y = 1$, leading to $w = (1, 1, 0)$.

Now, here you might think that one could get a third vector u simply by taking another (linearly independent) solution of (*)- say, $x = 0, y = 1/3$, but this won't work. The reason is that any two pairs (x, y) solving (*) differ by a solution of $-2x + 3y = 0$, exactly the equation defining the eigenspace. So if we did this, $u - w$ would be a multiple of v, and the matrix $P = [v|w|u]$ would not be invertible. We have to try something else!

A natural thing to try is to simply repeat the procedure that generated w, that is, solve the non-homogeneous system:

$$(A-3I)u = \begin{bmatrix} -6 & 9 & 1 \\ -4 & 6 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = w = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

This time we get:

$$-2x + 3y + (1/3)z = 1/3, \quad -2x + 3y + (1/2)z = 1/2,$$

and subtracting yields $(1/6)z = 1/6$, or $z = 1$. Then (x, y) can be *any* solution of $-2x+3y = 0$, including $x = y = 0$, which gives $u = (0, 0, 1)$. The procedure worked: we have for the action of A on the basis $\mathcal{B} = \{v, w, u\}$:

$$Av = 3v, \quad Aw = 3w + v, \quad Au = 3u + w,$$

so the 'standard form' of A is:

$$\Lambda = [A]_{\mathcal{B}} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}.$$

The matrix $P = [v|w|u]$ brings A to standard form: $P^{-1}AP = \Lambda$. Note that w and u are 'generalized eigenvectors' of orders 2 and 3 (respectively):

$$(A-3I)^3u = (A-3I)(A-3I)(A-3I)u = (A-3I)(A-3I)w = (A-3I)v = 0.$$

There is one possibility left when $n = 3$.

Example 3.

$$A = \begin{bmatrix} -3 & 9 & 0 \\ -4 & 9 & 0 \\ -6 & 9 & 3 \end{bmatrix}.$$

Again, the characteristic polynomial is $p(\lambda) = (\lambda - 3)^3$, so $\lambda = 3$ is the only eigenvalue. As always, we begin by computing its eigenspace:

$$(A-3I)v = \begin{bmatrix} -6 & 9 & 0 \\ -4 & 6 & 0 \\ -6 & 9 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0.$$

This amounts to the single equation: $-2x + 3y = 0$ (with z unconstrained), so this time we have a two-dimensional eigenspace:

$$E(3) = \{s(3, 2, 0) + t(0, 0, 1); s, t \in \mathbb{R}\}.$$

We only need one more vector w to build a basis, and (given recent experience) to find it we solve the non-homogeneous system:

$$(A - 3I)w = \begin{bmatrix} -6 & 9 & 0 \\ -4 & 6 & 0 \\ -6 & 9 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = v,$$

where v is an eigenvector. Previously we've had essentially only one choice (up to constant multiples), but now we must face the question:

Which eigenvector do we pick?

The choice is not at all arbitrary, if we want our non-homogeneous system to be consistent. So pick a generic vector $v = s(3, 2, 0) + t(0, 0, 1)$ in $E(3)$, and try to solve:

$$\begin{bmatrix} -6 & 9 & 0 \\ -4 & 6 & 0 \\ -6 & 9 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = v = \begin{bmatrix} 3s \\ 2s \\ t \end{bmatrix}.$$

This leads to:

$$-2x + 3y = s, \quad -2x + 3y = s, \quad -2x + 3y = t/3,$$

so we need $s = t/3$ to get a consistent system. Any non-zero values will do; say $t = 3, s = 1$, which gives $v = (3, 2, 3)$, and the equation for w: $-2x + 3y = 1$ (z unconstrained). Any solution will do here- say, $x = y = 1$,$z = 0$, and then $w = (1, 1, 0)$.

Having picked $v \in E(3)$ and w satisfying $(A - 3I)w = v$, for u we may take any vector in $E(3)$ linearly independent with w- say, $u = (0, 0, 1)$. The action of A on the basis $\mathcal{B} = \{v, w, u\}$ is then:

$$Av = 3v, \quad Aw = 3w + v, \quad Au = 3u,$$

and we have the standard form for A:

$$\Lambda = [A]_{\mathcal{B}} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

(This looks like the first non-diagonalizable 3×3 case listed above before example 2, when $\lambda = \mu$, but it is best to count it separately). The matrix P taking A to standard form is (of course) $P = [v|w|u]$, and $P^{-1}AP = \Lambda$.

So, to complete the list, the two 3×3 non-diagonalizable standard forms just described are:

$$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}.$$

The dimensions of the eigenspaces are: dim $E(\lambda)$=1 in the first case, dim $E(\lambda) = 2$ in the second. The algebraic multiplicity of λ is 3 in both cases.

EXERCISES. This topic is not difficult, but you won't learn it unless you get some practice. So here are **10** problems for you. In each case, a matrix A is given, and also its eigenvalues (listed with *algebraic* multiplicities). Your mission (which you would choose not to accept at your peril) is to find the standard form in each case, as well as the matrix that reduces A to its standard form.

1. $\begin{bmatrix} 3 & 4 \\ -1 & 7 \end{bmatrix}$ eigenvalues: $5, 5$.

2. $\begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$ eigenvalues: $2, 3$.

3. $\begin{bmatrix} -7 & 15 \\ -6 & 11 \end{bmatrix}$ eigenvalues: $2 \pm 3i$.

4. $\begin{bmatrix} -4 & 9 & 0 \\ -4 & 8 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ eigenvalues: $2, 2, 5$.

5. $\begin{bmatrix} -4 & 9 & 0 \\ -6 & 11 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ eigenvalues: $2, 2, 5$.

6. $\begin{bmatrix} 1 & 12 & 0 \\ -1 & 8 & 0 \\ 1 & -3 & 4 \end{bmatrix}$ eigenvalues: $4, 4, 5$.

7. $\begin{bmatrix} 1 & 12 & 0 \\ -1 & 8 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ eigenvalues: $4, 5, 5$.

8. $\begin{bmatrix} -3 & 24 & 12 \\ -2 & 11 & 4 \\ 4 & -16 & 5 \end{bmatrix}$ eigenvalues: $3, 5 \pm 4i$.

9. $\begin{bmatrix} 7 & -9 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 4 \end{bmatrix}$ eigenvalues: $4, 4, 4$.

10. $\begin{bmatrix} 4 & 0 & 3 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}$ eigenvalues: $4, 4, 4$.

Answers. Note that (when $n = 2$ or 3) to find the standard form it is enough to know the dimensions of the eigenspaces; this is false in dimensions 4 and higher. The matrices below are given by **rows**, in the following notation:

$$[[a, b][c, d]] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Recall that there is a unique correct answer for Λ, except for the order of the eigenvalues along the diagonal (which affects the order of the columns of P), but many possible correct answers for P, other than the one given. The first matrix given is Λ; the second is P.

1.[[5,1][0,5]];[[2,-1][1,0]]**2.**[[2,0][0,3]];[[1,2][1,1]];**3.**[[2,3][-3,2]];[[5,0][3,1]];
4.[[2,1,0][0,2,0][0,0,5]];[[3,1,0][2,1,0][0,0,1]];**5.**[[2,0,0][0,2,0][0,0,5]];[[3,0,1][2,0,1][0,1,1]];
6.[[4,1,0][0,4,0][0,0,5]];[[0,4,3][0,1,1][1,0,0]];**7.**[[5,0,0][0,5,0][0,0,4]];[[3,0,4][1,0,1][0,1,0];
8.[[5,4,0][-4,5,0][0,0,3]];[[-12,13,4][-3,1,1][1,0,0]]**9.**[[4,1,0][0,4,1][0,0,4]];[[0,3,4][0,1,1][1,0,0]];
10.[[4,1,0][0,4,1][0,0,4]];[[3,0,1][1,1,1][0,0,1]]