APPENDIX B

THE JORDAN FORM

Given a square matrix A, we want to choose M so that $M^{-1}AM$ is as nearly diagonal as possible. In the simplest case, A has a complete set of eigenvectors and they become the columns of M otherwise known as S. The Jordan form is $J = M^{-1}AM = \Lambda$, it is constructed entirely from 1 by 1 blocks $J_i = \lambda_i$, and the goal of a diagonal matrix is completely achieved. In the more general and more difficult case, some eigenvectors are missing and a diagonal form is impossible. That case is now our main concern.

We repeat the theorem that is to be proved:

5S If a matrix has s linearly independent eigenvectors, then it is similar to a matrix which is in ***Jordan form*** with s square blocks on the diagonal:

$$J = M^{-1}AM = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix}.$$

Each block has one eigenvector, one eigenvalue, and 1's just above the diagonal:

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & 1 \\ & & & \lambda_i \end{bmatrix}.$$

An example of such a Jordan matrix is

$$J = \begin{bmatrix} 8 & 1 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 8 & 1 \\ 0 & 8 \end{bmatrix} & & \\ & \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & \\ & & [0] \end{bmatrix} = \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & J_3 \end{bmatrix}.$$

The double eigenvalue $\lambda = 8$ has only a single eigenvector, in the first coordinate direction $e_1 = (1, 0, 0, 0, 0)$; as a result $\lambda = 8$ appears only in a single block J_1. The triple eigenvalue $\lambda = 0$ has two eigenvectors, e_3 and e_5, which correspond to the two Jordan blocks J_2 and J_3. If A had 5 eigenvectors, all blocks would be 1 by 1 and J would be diagonal.

The key question is this: *If A is some other 5 by 5 matrix, under what conditions will its Jordan form be this same matrix J? When will there exist an M such that $M^{-1}AM = J$?* As a first requirement, any similar matrix A must share the same eigenvalues 8, 8, 0, 0, 0. But this is far from sufficient—the diagonal matrix with these eigenvalues is not similar to J—and our question really concerns the eigenvectors.

To answer it, we rewrite $M^{-1}AM = J$ in the simpler form $AM = MJ$:

$$A \begin{bmatrix} & & & & \\ x_1 & x_2 & x_3 & x_4 & x_5 \\ & & & & \end{bmatrix} = \begin{bmatrix} & & & & \\ x_1 & x_2 & x_3 & x_4 & x_5 \\ & & & & \end{bmatrix} \begin{bmatrix} 8 & 1 & & & \\ 0 & 8 & & & \\ & & 0 & 1 & \\ & & 0 & 0 & \\ & & & & 0 \end{bmatrix}.$$

Carrying out the multiplications a column at a time,

$$Ax_1 = 8x_1 \quad \text{and} \quad Ax_2 = 8x_2 + x_1 \tag{1}$$

$$Ax_3 = 0x_3 \quad \text{and} \quad Ax_4 = 0x_4 + x_3 \quad \text{and} \quad Ax_5 = 0x_5. \tag{2}$$

Now we can recognize the conditions on A. It must have three genuine eigenvectors, just as J has. The one with $\lambda = 8$ will go into the first column of M, exactly as it would have gone into the first column of S: $Ax_1 = 8x_1$. The other two, which will be named x_3 and x_5, go into the third and fifth columns of M: $Ax_3 = Ax_5 = 0$. Finally there must be two other special vectors, the "*generalized eigenvectors*" x_2 and x_4. We think of x_2 as belonging to a ***string of vectors***, headed by x_1 and described by equation (1). In fact x_2 is the only other vector in the string, and the corresponding block J_1 is of order two. Equation (2) describes *two different strings*, one in which x_4 follows x_3, and another in which x_5 is alone; the blocks J_2 and J_3 are 2 by 2 and 1 by 1.

The search for the Jordan form of A becomes a search for these strings of vectors, each one headed by an eigenvector: For every i,

$$\text{either} \qquad Ax_i = \lambda_i x_i \qquad \text{or} \qquad Ax_i = \lambda_i x_i + x_{i-1}. \tag{3}$$

The vectors x_i go into the columns of M, and each string produces a single block in J. Essentially, we have to show how these strings can be constructed for every matrix A. Then if the strings match the particular equations (1) and (2), our J will be the Jordan form of A.

I think Filippov's idea, published in Volume 26 of the Moscow University Vestnik, makes the construction as clear and simple as possible. It proceeds by mathematical induction, starting from the fact that every 1 by 1 matrix is already in its Jordan form. We may assume that the construction is achieved for all matrices of order less than n—this is the "induction hypothesis"—and then explain the steps for a matrix of order n. There are three steps, and after a general description we apply them to a specific example.

Step 1 If we assume A is singular, then its column space has dimension $r < n$. Looking only within this smaller space, the induction hypothesis guarantees that a Jordan form is possible—there must be r independent vectors w_i in the column space such that

$$\text{either} \qquad Aw_i = \lambda_i w_i \qquad \text{or} \qquad Aw_i = \lambda_i w_i + w_{i-1}. \tag{4}$$

Step 2 Suppose the nullspace and the column space of A have an intersection of dimension p. Of course, every vector in the nullspace is an eigenvector corresponding to $\lambda = 0$. Therefore, there must have been p strings in step 1 which started from this eigenvalue, and we are interested in the vectors w_i that come at the end of these strings. Each of these p vectors is in the column space, so each one is a combination of the columns of A: $w_i = Ay_i$ for some y_i.

Step 3 The nullspace always has dimension $n - r$. Therefore, independent from its p-dimensional intersection with the column space, it must contain $n - r - p$ additional basis vectors z_i lying *outside* that intersection.

Now we put these steps together to give Jordan's theorem:

The r vectors w_i, the p vectors y_i, and the $n - r - p$ vectors z_i form Jordan strings for the matrix A, and these vectors are linearly independent. They go into the columns of M, and $J = M^{-1}AM$ is in Jordan form.

If we want to renumber these vectors as $x_1, \ldots, x_n$, and match them to equation (3), then each y_i should be inserted immediately after the w_i it came from; it completes a string in which $\lambda_i = 0$. The z's come at the very end, each one alone in its own string; again the eigenvalue is zero, since the z's lie in the nullspace. The blocks with nonzero eigenvalues are already finished at step 1, the blocks with zero eigenvalues grow by one row and column at step 2, and step 3 contributes any 1 by 1 blocks $J_i = [0]$.

Now we try an example, and to stay close to the previous pages we take the eigenvalues to be 8, 8, 0, 0, 0:

$$A = \begin{bmatrix} 8 & 0 & 0 & 8 & 8 \\ 0 & 0 & 0 & 8 & 8 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{bmatrix}.$$

Step 1 The column space has dimension $r = 3$, and its spanned by the coordinate vectors e_1, e_2, e_5. To look within this space we ignore the third and fourth rows and columns of A; what is left has eigenvalues 8, 8, 0, and its Jordan form comes from the vectors

$$w_1 = \begin{bmatrix} 8 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad w_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \qquad w_3 = \begin{bmatrix} 0 \\ 8 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

The w_i are in the column space, they complete the string for $\lambda = 8$, and they start the string for $\lambda = 0$:

$$Aw_1 = 8w_1, \qquad Aw_2 = 8w_2 + w_1, \qquad Aw_3 = 0w_3. \tag{5}$$

Step 2 The nullspace of A contains e_2 and e_3, so its intersection with the column space is spanned by e_2. Therefore $p = 1$ and, as expected, there is one string in equation (5) corresponding to $\lambda = 0$. The vector w_3 comes at the end (as well as the beginning) of that string, and $w_3 = A\,(e_4 - e_1)$. Therefore $y = e_4 - e_1$.

Step 3 The example has $n - r - p = 5 - 3 - 1 = 1$, and the nullvector $z = e_3$ is outside the column space. It will be this z that produces a 1 by 1 block in J.

If we assemble all five vectors, the full strings are

$$Aw_1 = 8w_1, \qquad Aw_2 = 8w_2 + w_1, \qquad Aw_3 = 0w_3, \qquad Ay = 0y + w_3, \qquad Az = 0z.$$

Comparing with equations (1) and (2), we have a perfect match -- the Jordan form of our example will be exactly the J we wrote down earlier. Putting the five vectors into the columns of M must give $AM = MJ$, or $M^{-1}AM = J$:

$$M = \begin{bmatrix} 8 & 0 & 0 & -1 & 0 \\ 0 & 1 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

We are sufficiently trustful of mathematics (or sufficiently lazy) not to multiply out $M^{-1}AM$.

In Filippov's construction, the only technical point is to verify the independence of the whole collection w_i, y_i, and z_i. Therefore, we assume that some combination is zero:

$$\sum c_i w_i + \sum d_i y_i + \sum g_i z_i = 0. \tag{6}$$

Multiplying by A, and using equations (4) for the w_i as well as $Az_i = 0$,

$$\sum c_i \begin{bmatrix} \lambda_i w_i \\ \text{or} \\ \lambda_i w_i + w_{i-1} \end{bmatrix} + \sum d_i A y_i = 0. \tag{7}$$

The Ay_i are the special w_i at the end of strings corresponding to $\lambda_i = 0$, so they cannot appear in the first sum. (They are multiplied by zero in $\lambda_i w_i$.) Since (7) is some combination of the w_i, which were independent by the induction hypothesis—they supplied the Jordan form within the column space—we conclude that *each d_i must be zero.* Returning to (6), this leaves $\sum c_i w_i = -\sum g_i z_i$, and the left side is in the column space. Since the z's were independent of that space, each g_i must be zero. Finally $\sum c_i w_i = 0$, and the independence of the w_i produces $c_i = 0$.

If the original A had not been singular, the three steps would have been applied instead to $A' = A - cI$. (The constant c is chosen to make A' singular, and it can be any one of the eigenvalues of A.) The algorithm puts A' into its Jordan form $M^{-1}A'M = J'$, by producing the strings x_i from the w_i, y_i and z_i. Then the Jordan form for A uses the same strings and the same M:

$$M^{-1}AM = M^{-1}A'M + M^{-1}cM = J' + cI = J.$$

This completes the proof that every A is similar to some Jordan matrix J. Except for a reordering of the blocks, ***it is similar to only one such*** J; there is a unique Jordan form for A. Thus the set of all matrices is split into a number of families, with the following property: *All the matrices in the same family have the same Jordan form, and they are all similar to each other* (and to J), *but no matrices in different families are similar.* In every family J is the most beautiful—if you like matrices to be nearly diagonal. With this classification into families, we stop.

EXAMPLE

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{with } \lambda = 0, 0, 0.$$

This matrix has rank $r = 2$ and only one eigenvector. Within the column space there is a single string w_1, w_2, which happens to coincide with the last two columns:

$$A \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 0 \qquad \text{and} \qquad A \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

or

$$Aw_1 = 0 \qquad \text{and} \qquad Aw_2 = 0w_2 + w_1.$$

The nullspace lies entirely within the column space, and it is spanned by w_1. Therefore $p = 1$ in step 2, and the vector y comes from the equation

$$Ay = w_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \qquad \text{whose solution is} \qquad y = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Finally the string w_1, w_2, y goes into the matrix M:

$$M = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \text{and} \qquad M^{-1}AM = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = J.$$

Application to $du/dt = Au$ As always, we simplify the problem by uncoupling the unknowns. This uncoupling is complete only when there is a full set of eigenvectors, and $u = Sv$; the best change of variables in the present case is $u = Mv$. This produces the new equation $M\,dv/dt = AMv$, or $dv/dt = Jv$, which is as simple as the circumstances allow. It is coupled only by the off-diagonal 1's within each Jordan block. In the example just above, which has a single block, $du/dt = Au$ becomes

$$\frac{dv}{dt} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} v \quad \text{or} \quad \begin{aligned} da/dt &= b \\ db/dt &= c \\ dc/dt &= 0 \end{aligned} \quad \text{or} \quad \begin{aligned} a &= a_0 + b_0 t + c_0 t^2/2 \\ b &= \phantom{a_0 + {}} b_0 + c_0 t \\ c &= \phantom{a_0 + b_0 t + {}} c_0 \end{aligned}$$

The system is solved by working upward from the last equation, and a new power of t enters at every step. (An l by l block has powers as high as t^{l-1}.) The exponentials of J, in this case and in the earlier 5 by 5 example, are

$$e^{Jt} = \begin{bmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad e^{Jt} = \begin{bmatrix} e^{8t} & te^{8t} & 0 & 0 & 0 \\ 0 & e^{8t} & 0 & 0 & 0 \\ 0 & 0 & 1 & t & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

You can see how the coefficients of a, b, and c appear in the first exponential. And in the second example, you can identify all five of the "special solutions" to $du/dt = Au$. Three of them are the pure exponentials $u_1 = e^{8t}x_1$, $u_3 = e^{0t}x_3$, and $u_5 = e^{0t}x_5$, formed as usual from the three eigenvectors of A. The other two involve the generalized eigenvectors x_2 and x_4:

$$u_2 = e^{8t}(tx_1 + x_2) \quad \text{and} \quad u_4 = e^{0t}(tx_3 + x_4). \tag{8}$$

The most general solution to $du/dt = Au$ is a combination $c_1u_1 + \cdots + c_5u_5$, and the combination which matches u_0 at time $t = 0$ is again

$$u_0 = c_1x_1 + \cdots + c_5x_5, \quad \text{or} \quad u_0 = Mc, \quad \text{or} \quad c = M^{-1}u_0.$$

This only means that $u = Me^{Jt}M^{-1}u_0$, and that the S and Λ in the old formula $Se^{\Lambda t}S^{-1}u_0$ have been replaced by M and J.

EXERCISES

B.1 Find the Jordan forms (in three steps!) of

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

B.2 Show that the special solution u_2 in Equation (8) does satisfy $du/dt = Au$, exactly because of the string $Ax_1 = 8x_1$, $Ax_2 = 8x_2 + x_1$.

B.3 For the matrix B above, use $Me^{Jt}M^{-1}$ to compute the exponential e^{Bt}, and compare it with the power series $I + Bt + (Bt)^2/2! + \cdots$.

B.4 Show that each Jordan block J_i is similar to its transpose, $J_i^{\mathrm{T}} = P^{-1}J_iP$, using the permutation matrix P with ones along the cross-diagonal (lower left to upper right). Deduce that every matrix is similar to its transpose.

B.5 Find "by inspection" the Jordan forms of

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}.$$

B.6 Find the Jordan form J and the matrix M for A and B (B has eigenvalues 1, 1, 1, -1):

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & -1 & 0 & -1 \\ 0 & 2 & 0 & 1 \\ -2 & 1 & -1 & 1 \\ 2 & -1 & 2 & 0 \end{bmatrix}$$

What is the solution to $du/dt = Au$ and what is e^{At}?

B.7 Suppose that $A^2 = A$. Show that its Jordan form $J = M^{-1}AM$ satisfies $J^2 = J$. Since the diagonal blocks stay separate, this means $J_i^2 = J_i$ for each block; show by direct computation that J_i can only be a 1 by 1 block, $J_i = [0]$ or $J_i = [1]$. Thus A is similar to a diagonal matrix of zeros and ones.

Note This is typical case of our closing theorem: *The matrix A can be diagonalized if and only if the product $(A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_p I)$, without including any repetitions of the λ's, is zero.* One extreme case is a matrix with distinct eigenvalues; the Cayley-Hamilton theorem in Exercise 5.6.23 says that with n factors $A - \lambda I$ we always get zero. The other extreme is the identity matrix, also diagonalizable ($p = 1$ and $A - I = 0$). The nondiagonalizable matrix $A = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ satisfies not $A = 0$ but only $A^2 = 0$— an equation with a repeated root.