In Sections 7.1-7.5 we have studied diagonalization of matrices. Given an $n \times n$ matrix A, the task is to find a diagonal matrix D and an invertible matrix S such that

$$A = SDS^{-1}.$$

Diagonalization is useful, because diagonal matrices are often easier to manipulate than non-diagonal matrices. For example, products of diagonal matrices are simple:

$$\begin{bmatrix} D_{11} & 0 & \cdots & 0 \\ 0 & D_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & D_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & 0 & \cdots & 0 \\ 0 & C_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & C_{nn} \end{bmatrix} = \begin{bmatrix} D_{11}C_{11} & 0 & \cdots & 0 \\ 0 & D_{22}C_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & D_{nn}C_{nn} \end{bmatrix}.$$

From this fact it follows that powers are also simple:

$$\begin{bmatrix} D_{11} & 0 & \cdots & 0 \\ 0 & D_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & D_{nn} \end{bmatrix}^k = \begin{bmatrix} D_{11}^k & 0 & \cdots & 0 \\ 0 & D_{22}^k & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & D_{nn}^k \end{bmatrix}.$$

This means that the power of any diagonalizable matrix is also easy to compute:

$$\left(SDS^{-1}\right)^k = \left(SDS^{-1}\right)\left(SDS^{-1}\right)\cdots\left(SDS^{-1}\right) = SD^kS^{-1}.$$

Here is a more sophisticated use of diagonalization. In analogy with the exponential function e^x from calculus, which is

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!}x^k = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots,$$

we can define a matrix exponential for any $n \times n$ matrix A, by

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!}A^k = I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \cdots.$$

This power series converges for any A (proving this fact is a good exercise in linear algebra and calculus), and hence e^A is well-defined. The matrix exponential is interesting and useful for a variety of reasons, including solution of ordinary differential equations. Unfortunately, it is complicated to compute. We can work around this complication using diagonalization. First, it is not difficult to show that the exponential respects similarity:

$$e^{SBS^{-1}} = Se^BS^{-1}.$$

Second, exponentials of diagonal matrices are simple — namely, if D is diagonal $n \times n$, then

$$e^D = \begin{bmatrix} e^{D_{11}} & 0 & \cdots & 0 \\ 0 & e^{D_{22}} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{D_{nn}} \end{bmatrix}.$$

Combining these two facts gives us an easy way to exponentiate any diagonalizable matrix $A = SDS^{-1}$:

$$e^{SDS^{-1}} = S \begin{bmatrix} e^{D_{11}} & 0 & \cdots & 0 \\ 0 & e^{D_{22}} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{D_{nn}} \end{bmatrix} S^{-1}.$$

I hope that you are convinced that diagonal matrices are convenient, and hence that diagonalization is useful. Unfortunately, not all matrices are diagonalizable. ("Almost all" are, but the exceptions seem to be disproportionately important in applications.) There are various ways to work around this problem. You can find a catalogue of them at Wikipedia's *Matrix decomposition* article and related articles. In the remainder of this paper we describe the Jordan decomposition.

First, we say that a matrix J is a *Jordan block* if all of its diagonal entries are equal, all of its entries immediately above the diagonal are 1, and all of its other entries are 0. Here are three examples, for the 1×1, 2×2, and 3×3 cases:

$$\begin{bmatrix} 5 \end{bmatrix}, \qquad \begin{bmatrix} -\sqrt{11} & 1 \\ 0 & -\sqrt{11} \end{bmatrix}, \qquad \begin{bmatrix} 0.7 & 1 & 0 \\ 0 & 0.7 & 1 \\ 0 & 0 & 0.7 \end{bmatrix}.$$

It is not difficult to prove that the eigenvalues of an $n \times n$ Jordan block are its diagonal entries. That is, the same eigenvalue is repeated n times. However, the corresponding eigenspace is not n-dimensional, but merely 1-dimensional; the only eigenvector is $[1\ 0\ \cdots\ 0]^\top$.

We say that a matrix J is a *Jordan matrix* if it breaks into Jordan blocks along its diagonal. Here are some 3×3 examples:

$$\begin{bmatrix} 1.3 & 1 & 0 \\ 0 & 1.3 & 1 \\ 0 & 0 & 1.3 \end{bmatrix}, \qquad \begin{bmatrix} 0.81 & 1 & 0 \\ 0 & 0.81 & 0 \\ 0 & 0 & 0.6 \end{bmatrix}, \qquad \begin{bmatrix} 0.41 & 0 & 0 \\ 0 & 0.79 & 0 \\ 0 & 0 & 0.71 \end{bmatrix}.$$

The first example consists of a single Jordan block with eigenvalue 1.3. The second consists of a 2×2 block with eigenvalue 0.81 and a 1×1 block with eigenvalue 0.6. The third consists of three 1×1 blocks. A Jordan matrix made entirely of 1×1 blocks is the same thing as a diagonal matrix. Just so we're clear, there is no requirement that the eigenvalues in the various Jordan blocks be distinct. The matrix

$$\begin{bmatrix} 0.81 & 0 & 0 \\ 0 & 0.81 & 0 \\ 0 & 0 & 0.6 \end{bmatrix}$$

is a legitimate Jordan matrix of three blocks. (It is a useful exercise to compare it to the middle matrix just above, which differs in only one entry. How are the eigenvalues and eigenvectors of the two matrices different?)

The following theorem says that all matrices are "Jordanizable".

Theorem 1 *For any $n \times n$ matrix A, there exists an invertible matrix S and a Jordan matrix J such that $A = SJS^{-1}$. Furthermore, J is unique, up to permutations of its Jordan blocks.*

J is called the *Jordan canonical form* of A, and $A = SJS^{-1}$ is called a *Jordan decomposition* of A. The matrix S is not unique; even in this simple 2×2 example of Jordan decomposition, b could be anything:

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

For another illuminating example, let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \lambda I + \frac{1}{ad - bc} \begin{bmatrix} -ac & a^2 \\ -c^2 & ac \end{bmatrix}.$$

The only eigenvector for A is $[a\ c]^\top$, and the only eigenvalue is λ (repeated). Holding a, c, and λ to some fixed values, we can still vary b and d freely (as long as $ad - bc \neq 0$) to produce different matrices A. So while diagonalizable matrices are completely determined by their eigenvectors and eigenvalues, nondiagonalizable matrices are not.

In our class, you are not responsible for knowing how to compute the Jordan decomposition of a matrix. However, I do recommend you play around with some matrices in *Mathematica*; see the `JordanDecomposition` function. Also, you are responsible for understanding how the structure of the Jordan canonical form reflects the eigenspaces of the original matrix. To that end, here are some exercises.

1. Compute this matrix's eigenvalues and eigenvectors. For each eigenvalue, give its algebraic multiplicity and geometric multiplicity.

$$\begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

2. Give a Jordan matrix that has the following eigenvalues with the specified multiplicities.

Eigenvalue	Algebraic Multiplicity	Geometric Multiplicity
5	1	1
7	2	1
8	2	2
9	3	2

3. Find a 4×4 Jordan matrix that has eigenvalue 6 with algebraic multiplicity 4 and geometric multiplicity 2. Then find *another* 4×4 Jordan matrix with the same eigenvalue with the same multiplicities, but with different Jordan block structure — meaning that the blocks are of different sizes. (The lesson here is that knowing the eigenvalues with their multiplicities is not enough to know the Jordan canonical form.)