Day 24 Fri 4/13/2018

Remind students about formal problem set. Handout copies if I am able to print any.
Take questions.

Topic 1: Proof of the equation for the inverse of the E matrix in prior set of lecture notes.

Topic 2: Nondiagonalizable matrices, Jordan Canonical Form, Minimal polynomials

1. Nondiagonalizable matrices
   a. Not all matrices are diagonalizable. In particular, if $p(t)$ has any repeated roots, its companion matrix is not diagonalizable
   b. We can still vectorize a difference equation and express the solution of the scalar equation in the form $a_k = [1\ \ 0\ \ 0\ \ \cdots\ \ 0]C^k[a_0\ \ a_1\ \ a_2\ \ \cdots\ a_{n-1}]^T$. But in the repeated roots case, using diagonalization to simplify C^k is not possible.
   c. In this case we can use a variation on diagonalization
   d. A key result in linear algebra states that every square matrix A can be reduced to a bidiagonal matrix J with a particular pattern of entries
   e. The matrix J is called the Jordan Canonical Form (*JCF*) of A. It has all eigenvalues on the main diagonal. Each eigenvalue appears as many times as it occurs as a root in the characteristic polynomial. For each repeated root λ, there can also be one or more 1's on the first super diagonal in between diagonal entries equal to λ. All other entries are equal to 0. If the matrix is diagonalizable, then the diagonalization ***is*** the JCF matrix J. Otherwise, J will be a diagonal matrix with at least one entry of 1 appearing on the main super diagonal.
   f. Reduction of A to its JCF J means that there exists an invertible matrix P for which $P^{-1}AP = J$. This holds iff $A = PJP^{-1}$.
2. Examples of possible JCF matrices
   a. The following are JCF matrices:

$$\begin{bmatrix} 7 & 1 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix} \quad \begin{bmatrix} 7 & 1 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix} \quad \begin{bmatrix} 7 & 1 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix} \quad \begin{bmatrix} 7 & 1 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 & 0 \\ 0 & 0 & 7 & 1 & 0 \\ 0 & 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{bmatrix}$$

   b. These matrices have a *diagonal block* structure, as illustrated below.

$$\left[\begin{array}{cc|c|c} 7 & 1 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 \\ \hline 0 & 0 & 0 & 6 \end{array}\right] \quad \left[\begin{array}{cc|cc} 7 & 1 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ \hline 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{array}\right] \quad \left[\begin{array}{ccc|c} 7 & 1 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 0 & 0 & 7 & 0 \\ \hline 0 & 0 & 0 & 6 \end{array}\right] \quad \left[\begin{array}{cc|cc|c} 7 & 1 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 & 0 \\ \hline 0 & 0 & 7 & 1 & 0 \\ 0 & 0 & 0 & 7 & 0 \\ \hline 0 & 0 & 0 & 0 & 6 \end{array}\right]$$

c. Often the zero entries outside the diagonal blocks are suppressed, to make the diagonal block structure more apparent, like so:

$$\begin{bmatrix} 7 & 1 & & \\ 0 & 7 & & \\ & & 3 & \\ & & & 6 \end{bmatrix} \quad \begin{bmatrix} 7 & 1 & & \\ 0 & 7 & & \\ & & 3 & 1 \\ & & 0 & 3 \end{bmatrix} \quad \begin{bmatrix} 7 & 1 & 0 & \\ 0 & 7 & 1 & \\ 0 & 0 & 7 & \\ & & & 6 \end{bmatrix} \quad \begin{bmatrix} 7 & 1 & & & \\ 0 & 7 & & & \\ & & 7 & 1 & \\ & & 0 & 7 & \\ & & & & 6 \end{bmatrix}$$

d. The repeated entries equal to a single eigenvalue can be broken up into diagonal blocks in all possible ways, from a single block with all superdiagonal entries equal to 1 (as for the 7's in the third example), to multiple blocks, each with super-diagonal 1's (as in the fourth example), to all 1 × 1 blocks (as shown at right). However, a given matrix *A* can only be reduced to *one* of these possibilities

$$\begin{bmatrix} 7 & & & & \\ & 7 & & & \\ & & 7 & & \\ & & & 7 & \\ & & & & 6 \end{bmatrix}$$

e. The total number of diagonal entries for each eigenvalue is equal to the multiplicity as a root of the characteristic polynomial. The number of different diagonal blocks for the eigenvalue is equal to the number of linearly independent eigenvectors – each such eigenvector gives rise to a block. The size of the largest block for a given eigenvalue is equal to the multiplicity as a root of the *minimal* polynomial. (This is easy to show by considering what happens when we substitute a *J* matrix into a polynomial.) More on this later. All of these aspects provide partial information about the JCF of a matrix, but they do not provide enough information to predict what the JCF is. For example, suppose the characteristic polynomial of some 8 × 8 matrix is $(t-7)^7(t-5)$, and there are 3 independent eigenvectors for 7. Thus there have to be seven 7's on the diagonal of the JCF, and there have to be three blocks with 7's. Suppose also that the minimal polynomial is $(t-7)^3(t-5)$. That tells us that the largest of the blocks has to be 3 × 3. But there are two possible JCF's given this information:

$$\begin{bmatrix} 7 & 1 & 0 & & & & & \\ 0 & 7 & 1 & & & & & \\ 0 & 0 & 7 & & & & & \\ & & & 7 & 1 & 0 & & \\ & & & 0 & 7 & 1 & & \\ & & & 0 & 0 & 7 & & \\ & & & & & & 7 & \\ & & & & & & & 5 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 7 & 1 & 0 & & & & & \\ 0 & 7 & 1 & & & & & \\ 0 & 0 & 7 & & & & & \\ & & & 7 & 1 & & & \\ & & & 0 & 7 & & & \\ & & & & & 7 & 1 & \\ & & & & & 0 & 7 & \\ & & & & & & & 5 \end{bmatrix}$$

f. There is an algorithm that can be used to construct the appropriate block for each independent eigenvector, but we are not going to go into that. Nor will we see a proof that JCF's as described here always exist, and have the properties described. I am mentioning all this mainly to provide some context for the one special case we will examine in detail: the case of a characteristic polynomial with only one distinct root.

3. Minimal polynomials
   a. We considered these in prior exercises (including an optional problem in formal problem set 3)
   b. In an earlier lecture, the Cayley-Hamilton theorem was mentioned. It states: If $p(t)$ is the characteristic polynomial for a square matrix A, then $p(A)$ equals the zero matrix. This is sometimes described by saying that $p(t)$ *annihilates* A.
   c. The minimal polynomial is the monic polynomial of lowest degree that annihilates A. It is unique. Moreover, it is a divisor of *every* polynomial that annhilates A. In particular, the minimal polynomial of A is a divisor of the characteristic polynomial. These facts are simple consequence of polynomial long division, and the fact that the remainder, if one occurs, has lower degree than the divisor.
   d. Problem 2b of the third formal problem set implies that for a *companion* matrix C, the characteristic polynomial is the same as the minimal polynomial. So for each eigenvalue λ of C the size of the largest block in the JCF must equal the multiplicity of λ as a root of the characteristic polynomial. But that implies that there must be exactly one block for λ. This is also implied by another fact we derived earlier: each eigenvalue has only *one* independent eigenvector, and therefore one block in the JCF.
   e. This is actually an iff theorem: The JCF of matrix A has exactly one diagonal block for each eigenvalue iff the characteristic polynomial equals the minimal polynomial, and those conditions hold iff A is similar to a companion matrix.
   f. In general, any power of a block diagonal matrix can be computed by just raising the diagonal blocks to the power. This extends to an easy calculation of a polynomial in a block diagonal matrix – we just apply the polynomial separately to each diagonal block. In particular, it is easy to see that for any $k \times k$ diagonal block B with λ's on the diagonal and 1's on the super-diagonal, the minimal polynomial is $(t - \lambda)^k$, because $B - \lambda I$ is a $k \times k$ N matrix. This shows that the minimal polynomial for the full J matrix has the same exponent on the factor $(t - \lambda)$ as the k for the largest diagonal block for λ. In other words, the largest diagonal block for λ is $k \times k$ where k is the multiplicity of λ as a root of the minimal polynomial.
4. JCF for a companion matrix
   a. We already know that there is a single independent eigenvector for every eigenvalue.
   b. The JCF must have exactly one block for each eigenvalue, so we can predict the JCF just by knowing the factorization of the characteristic polynomial

c. As in the repeated roots case, there is an explicit formulation of the *P* matrix for the JCF of a companion matrix, and it is a variant of the Vandermonde matrix we saw before

d. Also as before, there is a simple formula for expressing powers of the JCF matrix *J*. So we can compute an equation in the form
$a_k = [1 \ \ 0 \ \ 0 \ \ \cdots \ \ 0]P\, J^k\, P^{-1}[a_0 \ \ a_1 \ \ a_2 \ \ \cdots \ a_{n-1}]^T$

e. However, I know of no simple way to simplify $P^{-1}[a_0 \ \ a_1 \ \ a_2 \ \ \cdots \ a_{n-1}]^T$ as we did in the distinct roots case, even with a very special choice of initial terms.

5. Special Case: only one distinct root

   a. The characteristic polynomial would have to be of the form $p(t) = (t - \lambda)^n$.

   b. The JCF has just one $n \times n$ diagonal block, with λ's on the diagonal and all 1's on the first superdiagonal.

   c. We will see that in this case we *can* find a simple formulation for both P and P^{-1}.

End of Day