Math 110 - Fall 05 - Lectures notes # 39 - Dec 2 (Friday)

We begin Chapter 7, starting with a summary of the main points
of the chapter. We will only cover some of it.

In the last chapter we discussed using unitary similarity
transformations Q to either make a complex matrix A upper triangular:

```
   A = Q * T * Q^*   ... the Schur Factorization
```

(where Q^* * Q = I and T is upper triangular)
or to diagonalize A if possible, which is possible if and only
if A is "normal": A^* * A = A * A^*

Chapter 7 discusses the similarity transformation that is "closest"
to diagonal, for any A, diagonalizable or not.
This matrix factorization is called the Jordan Canonical Form:

```
   A = S * J * S^{-1}
```

Here J is upper triangular, with the eigenvalues on the diagonal,
and possibly with some 1s right above the diagonal, and the
other matrix entries 0:

Thm (Jordan Canonical Form, or JCF). Every square complex
matrix A is similar to a matrix in JCF: A = S * J * S^{-1},
where J has the following properties:

```
   (1) J = diag(J_1,J_2,...,J_k) is block diagonal.
       Each J_i is called a Jordan block.
   (2) Each J_i is upper triangular with an eigenvalue lambda_i
       on the diagonal, 1s right above the diagonal, and 0s elsewhere:
         J_i = [ lambda_i    1      0 0  ...         0    ]
               [    0     lambda_i  1 0  ...         0    ]
               [                    ...                   ]
               [    0        ...       lambda_i     1    ]
               [    0        ...          0     lambda_i ]
   (3) The lambda_i may or may not be distinct
```

Up to the order of the blocks J_i on the diagonal, the JCF is unique. In other
words, two matrices are similar if and only if they have the same Jordan blocks.

Ex: The following matrices have the same eigenvalues and multiplicities,
but different JCFs, so no two of them are similar:

```
    A1 = [ 1 0 0 0 ], A2 = [ 1 1 0 0 ], A3 = [ 1 0 0 0 ], A4 = [ 1 1 0 0 ]
         [ 0 1 0 0 ]       [ 0 1 0 0 ]       [ 0 1 0 0 ]       [ 0 1 0 0 ]
         [ 0 0 2 0 ]       [ 0 0 2 0 ]       [ 0 0 2 1 ]       [ 0 0 2 1 ]
         [ 0 0 0 2 ]       [ 0 0 0 2 ]       [ 0 0 0 2 ]       [ 0 0 0 2 ]
```

In fact, any 4 by 4 matrix with two eigenvalues equal to 1 and two eigenavalues equal to 2 must be similar to one of A1 through A4.

Before we prove this Theorem, we show what it is good for:

(1) Computing eigenvectors
(2) Computing A^m explicitly, for any A
(3) Deciding if the limit of A^m exists, as m -> infinity
(4) Computing any function f(A) explicitly, for any A
(5) Solving the differential equation dx(t)/dt = A*x(t) or
    dx(t)/dt = A*x(t) + f(t), explicitly, for any A

(1) Computing eigenvectors: A*x = lambda*x means S*J*S^{-1}*x = lambda*x or J*S^{-1}*x = lambda*S^{-1}*x or J*y = lambda*y where y = S^{-1}*x. So it suffices to compute eigevectors y of J to get eigenvectors x = S*y of A. This means we need to compute the null space of J - lambda*I.

Let's start with J equal to a single Jordan block with eigenvalue lambda:

```
  J - lambda*I = [ 0 1 0 ... 0 ]
                 [ 0 0 1 ... 0 ]
                 [     ...     ]
                 [ 0 0 ... 0 1 ]
                 [ 0 0 ... 0 0 ]
```

It is easy to see the e_1 is in the null space, since the first column of J - lambda*I is 0. Since the other columns of J - lambda*I are linearly independent (being e_1 through e_{n-1}), e_1 is the only eigenvector of a single Jordan block.

Now consider the general case J = diag(J_1,...,J_k), where each J_i is a Jordan block with eigenvalue lambda_i and dimension n_i. Then J - lambda*I = diag(J_1 - lambda*I,...,J_k - lambda*I) where each I has the same dimension n_i as the J_i it is being subtracted from. Then (J - lambda*I)*y = 0 means

```
  0 = diag(J_1 - lambda*I,...,J_k - lambda*I) * [ y_1 ]   n_1
                                                [ y_2 ]   n_2
                                                [ ... ]
                                                [ y_k ]   n_k
    = [ (J_1 - lambda_I)*y_1 ]   n_1
      [ (J_2 - lambda_I)*y_2 ]   n_2
      [        ...           ]
      [ (J_k - lambda_I)*y_k ]   n_k
```

so that each y_i must be in the nullspace of J_i - lambda*I.

```
When lambda neq lambda_i, then J_i - lambda*I has
lambda_i - lambda neq 0 on its diagonal, and
so is nonsingular, and so y_i = 0.
When lambda = lambda_i, then we showed above that either y_i = 0 or
y_i = [ 1, 0, ..., 0]^t.

In other words, any eigenvector for lambda is a linear combination
of the unique eigenvectors of the Jordan blocks J_i for which
lambda_i = lambda. If there are m such Jordan blocks, the dimension
of the eigenspace corresponding to lambda is therefore m.

(2) Computing A^m explicitly, for any A. Since
    A^m = (S * J * S^{-1})^m
        = S * J * S^{-1} * S * J * S^{-1} * ... * S * J * S^{-1}
                   ... m times
        = S * J * J * ... * J * S^{-1}
                   ... cancelling the m-1 pairs S^{-1} * S
        = S * J^m * S^{-1}
    it suffices to compute J^m. And since
    J^m = [ J_1  0  0 ... 0 ]^m = [ J_1^m   0    0  ...    0   ]
          [  0  J_2 0 ... 0 ]     [   0   J_2^m  0  ...    0   ]
          [       ...       ]     [              ...           ]
          [  0   0 ... 0 J_k]     [   0     0   ...  0   J_k^m ]
    it suffices to compute the m-th power of a single Jordan block.

    Let's do some examples. For simpler notation, we let x be the
    eigenvalue of the single Jordan block J.
    If J = [ x ] is 1 x 1, then of course J^m = [ x^m ].
    If J = [ x 1 ], then J^2 = [ x^2 2*x ], J^3 = [ x^3 3*x^2 ]
           [ 0 x ]             [  0  x^2 ]        [  0  x^3   ]
    J^4 = [ x^4 4*x^3 ], ...
          [  0  x^4   ]
    A natural conjecture is that J^m = [ x^m  m*x^(m-1) ]
                                       [  0   x^m       ]
    which we could prove by induction. Here is another example:


    Ex:  J = [ 0 1 0 0 0 ], J^2 = [ 0 0 1 0 0 ], J^3 = [ 0 0 0 1 0 ]
             [ 0 0 1 0 0 ]        [ 0 0 0 1 0 ]        [ 0 0 0 0 1 ]
             [ 0 0 0 1 0 ]        [ 0 0 0 0 1 ]        [ 0 0 0 0 0 ]
             [ 0 0 0 0 1 ]        [ 0 0 0 0 0 ]        [ 0 0 0 0 0 ]
             [ 0 0 0 0 0 ]        [ 0 0 0 0 0 ]        [ 0 0 0 0 0 ]
```

```
J^4=[ 0 0 0 0 1 ], J^5 = [ 0 0 0 0 0 ]
    [ 0 0 0 0 0 ]         [ 0 0 0 0 0 ]
    [ 0 0 0 0 0 ]         [ 0 0 0 0 0 ]
    [ 0 0 0 0 0 ]         [ 0 0 0 0 0 ]
    [ 0 0 0 0 0 ]         [ 0 0 0 0 0 ]

Lemma 1: Let N be an n by n Jordan block with eigenvalue = 0. Then
         N^m is zeros except for the m-th diagonal above the main
         diagonal (the m-th "superdiagonal"), which is all ones.
Proof: Write N^(m+1) = N^m * N, and note that since
       N = [ 0, e_1, e_2,... e_{n-1} ], multiplying any
       matrix A = [a_1, a_2,...,a_n] by N yields
       A*N = [A*0, A*e_1, A*e_2,..., A*e_{n-1}]
           = [ 0 ,  a_1 ,  a_2, ... , a_{n-1} ]
       i.e. it moves all the columns right by one, and makes
       the first column 0.

Ex:  J = [ 1 1 0 0 0 ], J^2 = [ 1 2 1 0 0 ], J^3 = [ 1 3 3 1 0 ]
         [ 0 1 1 0 0 ]        [ 0 1 2 1 0 ]        [ 0 1 3 3 1 ]
         [ 0 0 1 1 0 ]        [ 0 0 1 2 1 ]        [ 0 0 1 3 3 ]
         [ 0 0 0 1 1 ]        [ 0 0 0 1 2 ]        [ 0 0 0 1 3 ]
         [ 0 0 0 0 1 ]        [ 0 0 0 0 1 ]        [ 0 0 0 0 1 ]

     J^4=[ 1 4 6 4 1 ], J^5 = [ 1 5 10 10  5]
         [ 0 1 4 6 4 ]        [ 0 1  5 10 10]
         [ 0 0 1 4 6 ]        [ 0 0  1  5 10]
         [ 0 0 0 1 4 ]        [ 0 0  0  1  5]
         [ 0 0 0 0 1 ]        [ 0 0  0  0  1]

Note appearance of Pascal's triangle in the rows of these matrices:

                     1
                   1   1
                 1   2   1
               1   3   3   1
             1   4   6   4   1
           1   5  10  10   5   1
```

```
Thm 2: Let J be an n by n Jordan block with eigenvalue x.  Then
         (J^m)_ij = m!/((m-j+i)! *  (j-i)!) * x^(m-j+i)
                  = (  m  ) * x^(m-j+i)
                    ( j-i )
                  = "m choose j-i" * x^(m-j+i)
       where we interpret "m choose j-i" = 0 if j-i>m  In other words

            (m)*x^m = x^m  appears all along the main diagonal
            (0)

            (m)*x^(m-1) = m*x^{m-1} appears along superdiagonal 1
            (1)

            (m)*x^(m-2) = m*(m-1)/2*x^{m-2} appears along superdiagonal 2
            (2)

            (m)*x^(m-i) = (m!/(i! * (m-i)!)) * x^(m-i)   appears along
            (i)                                          superdiagonal i

Proof 1: We use the binomial theorem: Write J = x*I + N,
         where N contains just the offdiagonal 1s of J. Then
         J^m = (x*I + N)^m
             = sum_{i=0 to m} (m choose i)* (x*I)^(m-i) * N^i
         (The proof that this formula is true is the same
          induction on m as in the scalar case, but depends on
          the fact the x*I and N commute.)
         From Lemma 1 we know that N^i has nonzeros only
         on superdiagonal i (or is zero if i >= n). Thus the sum is
         J^m = sum_{i=0 to min(n-1,m)} (m choose i) * x^(m-i) * N^i
         So since (J^m)_ij is on diagonal j-i, the only
         nonzero part of the sum is
         (J^m)_ij = (m choose j-i) * x^(m-j+i) * (N^(j-i))_ij
                  = (m choose j-i) * x^(m-j+i)  ... as desired.

Here is a proof that does not assume the binomial theorem:

Lemma 2: Let J' by an n by n Jordan block with eigenvalue 1. Then
         (J'^m)_ij = m!/((m-j+i)! *  (j-i)!)
                  = (  m  )
                    ( j-i )
                  = "m choose j-i"
         where we interpret "m choose j-i" = 0 if j-i>m.
```

```
Proof: The proof is the same induction as for Pascal's Triangle
       or the binomial theorem.
       When j = i, (J'^m)_ij = 1 = m!/(m! * 0!) as desired.
       When j > i
        (J'^(m+1))_ij = (J' * J'^m)_ij
             = sum_{k=1 to n} J'_ik * (J'^m)_kj
                   ... definition of matrix multiply
             = sum_{k=i to i+1} (J'^m)_kj
                   ... since only nonzeros in J' are J'_ii = J'_{i,i+1} = 1
             = (J'^m)_ij + (J'^m)_{i+1,j}
             = m!/((m-j+i)! * (j-i)!)  *  m!/((m-j+i+1)! * (j-i-1)!)
             = m!/((m-j+i)! * (j-i-1)!) * ( 1/(j-i) + 1/(m-j+i+1) )
             = m!/((m-j+i)! * (j-i-1)!) * ( (m+1)/((j-i)*(m-j+i+1)) )
             = (m+1)!/((m-j+i+1)! * (j-i)!)
             = "m+1 choose j-i"   ... as desired

Different proof of Thm 2: We reduce to the case in the above example,
where Pascal's Triangle appeared. For simplicity we look at n=5;
all other dimensions are similar.  Assume the eigenvalue x neq 0.
Let D = diag(1,x,x^2,x^3,x^4).  Then we can easily confirm that

          J = [ x 1 0 0 0 ] = D * [ x x 0 0 0 ] * D^{-1}
              [ 0 x 1 0 0 ]       [ 0 x x 0 0 ]
              [ 0 0 x 1 0 ]       [ 0 0 x x 0 ]
              [ 0 0 0 x 1 ]       [ 0 0 0 x x ]
              [ 0 0 0 0 x ]       [ 0 0 0 0 x ]

            = x * D * [ 1 1 0 0 0 ] * D^{-1}
                      [ 0 1 1 0 0 ]
                      [ 0 0 1 1 0 ]
                      [ 0 0 0 1 1 ]
                      [ 0 0 0 0 1 ]

            = x * D * J' * D^{-1}
 Therefore

 J^m = x^m * D * J' * D^{-1} * D * J' * D^{-1} * ... * D * J' * D^{-1}
                  ... m "copies" of D * J' * D^{-1}
     = x^m * D * J'^m * D{-1}  ... cancelling all pairs D^{-1} * D

 where J'^m looks like the matrix in Lemma 2, with entries from
```

```
 Pascal's Triangle. Thus for j >= i
    (J^m)_ij = x^m * D_ii * (J'^m)_ij * (D_jj)^{-1}
             = x^m * x^(i-1) * ( m ) * x^(1-j)
                               (j-i)
             = ( m ) * x^(m-j+i)
               (j-i)

 Note that (m) * x^(m-k) = (d^k x^m / dx^k ) / k!
           (k)
 In other words, if we let f(x) = x^m, we can say that
    (J^m)_ij = 1/(j-i)!  *  d^(j-i) f(x) / dx^(j-i)
 We will see below that this formula is true for any function f(x),
 not just x^m.

(3) Deciding if the limit of A^m exists, or stays bounded, as m -> infinity.
    By part 2, it suffices to look at what happens to J^m, for
    J a single Jordan block. We leave the proofs of the following results
    as an exercise:
     (1) lim_{k -> infinity} A^k exists if and only if all the Jordan
         blocks of A either
             have eigenvalue |lambda| < 1, or
             have eigenvalue lambda   = 1 and are 1 by 1
         Furthermore the limit is nonzero if and only if there is
         at least one Jordan block with eigenvalue 1
     (2) A^k remains bounded for all k if and only if all the Jordan
         blocks of A either
             have eigenvalue |lambda| < 1, or
             have eigenvalue |lambda| = 1 and are 1 by 1
     (3) In all other cases, A^k is unbounded as k -> infinity:
         at least one Jordan block of A
             has eigenvalue |lambda| > 1, or
             has eigenvalue |lambda| = 1, and is 2 by 2 or larger.


(4) Computing f(A) explicitly, for any A. Suppose
        f(A) = sum_{i=0 to d} p_i*A^i.
    where d could be finite (so f(A) is a polynomial) or infinite
    (so the f(A) is given by a Taylor expansion).
    Since we know how to compute A^i, we see that all we need to do
    is compute f(J) for a Jordan block, since
       A = S * diag(J_1,...,J_k) * S^{-1}
    implies
```

```
  f(A) = sum_{i=0 to d} p_i * (S * diag(J_1,...,J_k) * S^{-1})^i
       = sum_{i=0 to d} p_i * S * diag(J_1^i,...,J_k^i) * S^{-1}
       = S * [ sum_{i=0 to d} p_i * diag(J_1^i,...,J_k^i) ] * S^{-1}
       = S * [ diag(f(J_1),...,f(J_k)) ] * S^{-1}
Continuing, for simplicity let J be a single Jordan block
with eigenvalue x. Then
   (f(J))_jj = (sum_{i=0 to d} p_i * J^i)_jj
             =  sum_{i=0 to d} p_i * (J^i)_jj
             =  sum_{i=0 to d} p_i * x^i
             = f(x)
and for k>j
   (f(J))_jk = (sum_{i=0 to d} p_i * J^i)_jk
             =  sum_{i=0 to d} p_i * (J^i)_jk
             =  sum_{i=0 to d} p_i * (i choose k-j) * x^(i-k+j)
             =  sum_{i=0 to d} p_i/(k-j)! * d^{k-j} x^i / dx^{k-j}
             =  1/(k-j)! * d^{k-j} ( sum_{i=0 to d} p_i * x^i ) / dx^{k-j}
             =  1/(k-j)! * d^{k-j} f(x) / dx^{k-j}
For example, for n=5, we get
   f(J) = [ f(x) f'(x)/1!  f''(x)/2!  f'''(x)/3!  f''''(x)/4! ]
          [ 0    f(x)      f'(x)/1!   f''(x)/2!   f'''(x)/3!  ]
          [ 0    0         f(x)       f'(x)/1!    f''(x)/2!   ]
          [ 0    0         0          f(x)        f'(x)/1!    ]
          [ 0    0         0          0           f(x)        ]
```

```
(5) Solving dx(t)/dt = A*x(t) explicitly. Write A = S * J * S^{-1} so
      dx(t)/dt = S*J*S^{-1}*x(t)
    or
      S^{-1}*dx(t)/dt = J*S^{-1}*x(t)
    or
      d(S^{-1}*x(t))/ dt = J*(S^{-1}*x(t))
    or
      dy(t)/ dt = J*y(t)    ... where y(t) = S^{-1}*x(t)
    If J = diag(J_1,...,J_k) where J_i is n_i by n_i, and
      y = [y_1]  n_1
          [y_2]  n_2
          [...]
          [y_k]  n_k
    then dy(t)/dt = J*y(t) implies that for each i we get
      dy_i(t)/dt = J_i * y_i(t)
    so it suffices to solve a differential equation with a single Jordan block.
    Earlier we showed the solution to the differential equation
      d y(t)/dt = A*y(t)
```

```
for A diagonalizable was
  y(t) = exp(t*A) * y(0)
       = [ sum_{i=0 to infinity} (t^i * A^i / i!) ] * y(0)
We can differentiate this term by term to confirm it is true for any A:
  d y(t)/dt = d [ sum_{i=0 to infinity} (t^i * A^i / i!) ] * y(0) ] /dt
            = sum_{i=0 to infinity} d(t^i * A^i / i!)/dt ] * y(0)
            = sum_{i=0 to infinity} (d(t^i)/dt * A^i / i!) ] * y(0)
            = sum_{i=0 to infinity} (i*t^(i-1)) * A^i / i!) ] * y(0)
            = sum_{i=1 to infinity} (t^(i-1)) * A^i / (i-1)!) ] * y(0)
            = sum_{i=0 to infinity} (t^i * A^(i+1) / i!) ] * y(0)
            = A * sum_{i=0 to infinity} (t^i * A^i / i!) ] * y(0)
            = A * y(t)   ... as desired.
When A = J is a single Jordan block with eigenvalue lambda, then we can write
   exp(t*A)_jk = 1/(k-j)! * t^(k-j) * exp(t*lambda)
```