Day 25 Tue 4/17/2018

Note the document posted on blackboard. This covers the math for the rest of the course. Any questions about anything?

Note to self: The file chap10.pdf was created using a modified version of the tex files in the folder C:\User Files\dan-home\classes\fall 2013\matrinomials\latex docs.

Outline for lecture: JCF for monorootic companion matrix

1. Monorootic polynomials
   a. def: monic monorootic of degree n means equal to $(x - r)^n$ for some scalar r.
   b. Companion matrix is then $C_p = \left[\begin{array}{c|c} 0 & I \\ \hline \multicolumn{2}{c}{-\mathbf{c}^T} \end{array}\right]$ where $\mathbf{c}^T$ is the vector

      $$\left[\ \binom{n}{0}(-r)^n \quad \binom{n}{1}(-r)^{n-1} \quad \binom{n}{2}(-r)^{n-2} \quad \cdots \quad \binom{n}{n-1}(-r)\ \right]$$

   c. prove using binomial theorem
2. Near Diagonalization or Bidiagonalization
   a. In the monorootic case for $n > 1$ the roots are not distinct. So C_p cannot be diagonalized.
   b. It *can* be nearly diagonalized, by reducing it to JCF.
   c. Specifically, $P^{-1}C_p\,P = J$ with $J = rI + N$ where r is the unique root of p and N is the patterned matrix we studied earlier, consisting of 1's on the super-diagonal and 0's elsewhere. We will see what P is and a verification of the equation later.
   d. Can rewrite the equation as $C_p = PJP^{-1}$ and therefore $C_p^k = PJ^kP^{-1}$
   e. We can use this in a vectorized difference equation, if we can find the matrix P and if we can easily compute J^k.
3. Powers of J.
   a. $J^k = (rI + N)^k$
   b. This can be expanded using binomial theorem. For example,
      $$(rI + N)^3 = r^3I^3 + 3r^2I^2N + 3rIN^2 + N^3 = r^3I + 3r^2N + 3rN^2 + N^3$$
   c. Note: the final expression at right is $f(N)$ where $f(t) = r^3 + 3r^2t + 3rt^2 + t^3$.
   d. Recall the special structure of $f(N)$ for a polynomial N. For example, if $f(t) = 2 + 3t - 5t^2 + 8t^3 + 11t^4 + 17t^5$ and N is the 4×4 N matrix
   e. For $k \geq n$, we see that
      $$(rI + N)^k = r^kI + \binom{k}{1}r^{k-1}N + \binom{k}{2}r^{k-2}N^2 + \cdots + \binom{k}{n-1}r^{k-n+1}N^{n-1}$$

f. This creates an upper triangular matrix with all entries on the diagonal equal to r^k, all entries on the first super-diagonal equal to $\binom{k}{1}r^{k-1}$, and so on. The four by four version looks like this:

$$(rI+N)^k = \begin{bmatrix} r^k & \binom{k}{1}r^{k-1} & \binom{k}{2}r^{k-2} & \binom{k}{2}r^{k-3} \\ & r^k & \binom{k}{1}r^{k-1} & \binom{k}{2}r^{k-2} \\ & & r^k & \binom{k}{1}r^{k-1} \\ & & & r^k \end{bmatrix}.$$

4. The P matrix

For any $n \geq 2$ in $\mathbb{N}$ and for any scalar r, we define the matrix

$$P_n(r) = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ r & 1 & 0 & 0 & \cdots & 0 \\ r^2 & 2r & 1 & 0 & \cdots & 0 \\ r^3 & 3r^2 & 3r & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ r^n & nr^{n-1} & \binom{n}{2}r^{n-2} & \binom{n}{3}r^{n-3} & \cdots & 1 \end{bmatrix}.$$

That is, $P_n(r)$ is a lower triangular $(n+1)\times(n+1)$ matrix with ij entry

$$(P_n(r))_{ij} = \binom{i-1}{j-1} r^{i-j}$$

for $j \leq i$. We write $P(r)$ in place of $P_n(r)$ when we are not concerned about the dimensions of the matrix.

Note that the nonzero entries of row i are the terms of $(r+1)^{i-1}$; the coefficients are the entries of row $i-1$ of Pascal's triangle

5. Special Cases

a. When $r = 1$ the nonzero entries make a Pascal's Triangle

b. When $r = -1$ the nonzero entries make an alternating sign Pascal's triangle

c. For example, with $n = 4$,

$$P_4(1) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \qquad P_4(-1) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ -1 & 3 & -3 & 1 & 0 \\ 1 & -4 & 6 & -4 & 1 \end{bmatrix}$$

6. Key Identity: $P(r)P(s) = P(r+s)$
   a. Provable using algebraic manipulations
   b. Simpler proof: observe that $P(r)$ is the matrix representation of a linear transformation in polynomial space, namely, substituting $t + r$ for t.
   c. Graphically that is a leftward horizontal translation of the graph by r units
   d. Using either of these aspects, it is clear that transforming first by r and then by s results in a net transformation by $r + s$.
   e. Application: $P(-r)$ is the inverse of $P(r)$.
      Proof: Compute the product $P(-r)P(r) = P(-r + r) = P(0) = I$.
   f. Reduction of C_p to JCF is simplified because we know exactly what the inverse matrix of P will be.
7. Verification of the equation $P^{-1}C_p\, P = J = rI + N$
   a. Suffices to prove that $C_p\, P = PJ = P(rI + N)$
   b. Recall from the diagonalizable case, the eigenvectors for each eigenvalue of a compantion matric can be derived from a key identity:

$$C_p\,\mathbf{v}(t) = t\mathbf{v}(t) + p(t)\mathbf{e}_n, \qquad (\star)$$

      where $\mathbf{v}(t) = [1 \quad t \quad t^2 \quad \cdots \quad t^{n-1}]^T$ and $\mathbf{e}_n = [0 \quad 0 \quad \cdots \quad 0 \quad 1]^T$.
   c. When t is a root of p the identity turns into an eigenvector/value equation.
   d. Key idea: when r is a *repeated* root of p, that means it is also a root of p'.
   e. More generally, if r is repeated m times, then p is divisible by $(t - r)^m$, and r is a root of $p(t)$ as well at the first $m - 1$ derivatives.
   f. With that as motivation, let us consider what happens when we differentiate identity ($\star$) repeatedly.
8. For the case $n = 4$, differentiate three times
   a. $C_p\,\mathbf{v}(t) = t\mathbf{v}(t) + p(t)\mathbf{e}_n$
   b. $C_p\,\mathbf{v}'(t) = t\mathbf{v}'(t) + \mathbf{v}(t) + p'(t)\mathbf{e}_n$
   c. $C_p\,\mathbf{v}''(t) = t\mathbf{v}''(t) + 2\mathbf{v}'(t) + p''(t)\mathbf{e}_n$
   d. $C_p\,\mathbf{v}'''(t) = t\mathbf{v}'''(t) + 3\mathbf{v}''(t) + p'''(t)\mathbf{e}_n$

9. When t is a root ...

   a. If t is replaced by r, the root of the characteristic polynomial, then $p(r) = p'(r) = p''(r) = p'''(r) = 0.$

   b. We find
   $$C_p\,[\mathbf{v}(r) \quad \mathbf{v}'(r) \quad \mathbf{v}''(r) \quad \mathbf{v}'''(r)\,] = r[\mathbf{v}(r) \quad \mathbf{v}'(r) \quad \mathbf{v}''(r) \quad \mathbf{v}'''(r)\,] + [0 \quad \mathbf{v}(r) \quad 2\mathbf{v}'(r) \quad 3\mathbf{v}''(r)]$$

   c. Defining $[\mathbf{v}(r) \quad \mathbf{v}'(r) \quad \mathbf{v}''(r) \quad \mathbf{v}'''(r)\,]$ to be Q, this says
   $$C_pQ = rQ + Q\begin{bmatrix} 0 & 1 & & \\ & 0 & 2 & \\ & & 0 & 3 \\ & & & 0 \end{bmatrix}$$

   d. This is tantalizingly close to $C_p\,P = PJ = P(rI + N) = rP + PN$. It inspires one to cast about algebraically for a way to make it fit correctly. Eventually one discovers that P should be defined as $P = \left[\mathbf{v}(r) \quad \mathbf{v}'(r) \quad \frac{1}{2}\mathbf{v}''(r) \quad \frac{1}{6}\mathbf{v}'''(r)\right]$

   e. $$\begin{aligned} C_pP &= C_p\left[\mathbf{v}(r) \quad \mathbf{v}'(r) \quad \tfrac{1}{2}\mathbf{v}''(r) \quad \tfrac{1}{6}\mathbf{v}'''(r)\right] \\ &= r\left[\mathbf{v}(r) \quad \mathbf{v}'(r) \quad \tfrac{1}{2}\mathbf{v}''(r) \quad \tfrac{1}{6}\mathbf{v}'''(r)\right] + \left[0 \quad \mathbf{v}(r) \quad \mathbf{v}'(r) \quad \tfrac{1}{2}\mathbf{v}''(r)\right] \\ &= rP + P\begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = rP + PN = P(rI + N)\ . \end{aligned}$$

   f. This shows $C_p\,P = PJ$ in the 4×4 case.

   g. To prove the identity holds in the general $n \times n$ case, it helps to recognize that
   $$\frac{1}{j!}(t^m)^{(j)} = \begin{cases} 0 \quad \text{if} \quad j > m \\ \frac{1}{j!}m(m-1)\cdots(m-j+1)\left(t^{m-j}\right) \quad \text{otherwise} \end{cases}$$
   and that $\frac{1}{j!}m(m-1)\cdots(m-j+1) = \frac{1}{j!}\frac{m!}{(m-j)!} = \binom{m}{j}$. Then our P matrix can be expressed in the form given earlier in paragraph 4.

End of Day