Math 433, Fall 2006

Homework 8**, Selected Solutions**

1) First note that if $R = CSC^{-1}$ then $R^n = CS^nC^{-1}$, so if $R^d = I$, but $R, R^2, ..., R^{d-1} \neq I$, then $S^d = I$, but $S, S^2, ..., S^{d-1} \neq I$.

If a matrix A over $\mathbb{C}$ is diagonalizable, then it is similar to a matrix of the form $B = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. If A has order 4 then B has order 4 so we must have $\lambda_i \in \{i, -i, 1, -1\}$. Since $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} B \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{bmatrix}$ we need not worry about the ordering of the λ's. The following pairs of values for (λ_1, λ_2) give us our matrices:

$$(i,i), (i,-i), (i,1), (i,-1), (-i,1), (-i,-1), (-i,-i).$$

If A is not diagonalizable then it is similar to a matrix of the form $B = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$. The fourth power of this matrix is $\begin{bmatrix} \lambda^4 & 4\lambda^3 \\ 0 & \lambda^4 \end{bmatrix}$ which is never the identity. All matrices over $\mathbb{C}$ of order 4 are diagonalizable.

2) If A and B have the same characteristic polynomial they have the same eigenvalues with the same multiplicities.

a) If we are in the 2×2 case and the eigenvalues are distinct, say λ_1 and λ_2. Then both A and B are similar to $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ and hence similar to one another.

If the eigenvalues are the same, say λ, then since we are in the nonscalar case, both A and B must have Jordan canonical form $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$, and so are similar to one another.

b) The matrices $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ clearly have the same characteristic polynomial and are nonscalar. They are *not* similar, as the square of the first matrix is not the zero matrix, but the square of the second matrix is the zero matrix.

3) Ok, here we assume our field is algebraically closed, so A is similar to an upper triangular matrix T with the λ_i on the diagonal. As $A = STS^{-1}$ for some S, we have $A^k = ST^kS^{-1}$. Raising T to the kth power is hard, but knowing the diagonal values of T^k is easy, they are just the λ_i^k. The trace and characteristic polynomial of A^k are those of T^k, namely $\sum \lambda_i^k$ and $\prod(x - \lambda_i^k)$ (or $\prod(\lambda_i^k - x)$ depending on your convention for defining characteristic polynomials).

4) If $C = SAS^{-1}$, and $B^2 = A$, we see $(SBS^{-1})^2 = SBS^{-1}SBS^{-1} = SB^2S^{-1} = C$, so it suffices to find a square root of a matrix similar to A, namely its Jordan canonical form (JCF). We henceforth assume A is in JCF.

Thus A is a matrix in block form with blocks as in definition 25.17. We first extract the square root of such a $n \times n$ block when $\alpha_i \neq 0$. Choose a $\beta \in \mathbb{C}$ such that $\beta^2 = \alpha_i$ (We can do this. Why?). Consider the upper triangular matrix D with ij entry d_{ij} given by $d_{ij} = \begin{cases} 0 & i \neq j-1 \\ 1 & i = j-1 \end{cases}$. Then the matrix D^r has ij entry $\begin{cases} 0 & i \neq j-r \\ 1 & i = j-r \end{cases}$. Note that since we are dealing with $n \times n$ matrices that $D^n = 0$. We assume A is as in definition 25.17 and we find complex numbers γ_k such that $(\beta I + \sum_{k=1}^{n} \gamma_k D^k)^2 = A$. The left hand side multiplies out to

$$\beta^2 I + \sum_{k=1}^{n} \left(2\beta\gamma_k + \sum_{j=1}^{k} 2\gamma_j \gamma_{k-j} \right) D^k.$$

In this expansion we need the coefficient of D to be 1 and the coefficient of D^k to be 0 for $k > 1$. Clearly γ_1 must equal $1/(2\beta)$. The coefficient of D^2 is determined now except for γ_2 which we can solve for. Continuing we see we can solve for γ_k in terms of β and $\gamma_1, ..., \gamma_{k-1}$. Thus we can take the square root of a matrix as in definition 25.17 with $\alpha_i \neq 0$.

Putting together the blocks, we have proved that if 0 is *not* an eigenvalue of A then there is a matrix B such that $B^2 = A$.

If the eigenvalues of A are (with possible multiplicities) $\lambda_1, ..., \lambda_n$ then the eigenvalues of B are sqaure roots of these numbers. If A has a square root B, then 0 must occur as an eigenvalue with the same multiplicity in A and B. So if A_0 and B_0 correspond to all blocks as in definition 25.17 with $\alpha_i = 0$, we must have $JCF(B_0^2) = A_0$. For the rest of this problem we may assume A has all eigenvalues equal to 0.

We now analyze a block D as defined above. If D is $2r \times 2r$, then with respect to the reordering $\{\vec{e}_1, \vec{e}_3, \vec{e}_5, ..., \vec{e}_{2r-1}, \vec{e}_2, \vec{e}_4, \vec{e}_6, ...\vec{e}_{2r}\}$ the JCF of D^2 is of the form $\begin{bmatrix} X_r & \mathbf{0} \\ \mathbf{0} & X_r \end{bmatrix}$ where X_r is an $r \times r$ matrix as in definition 25.17 and $\mathbf{0}$ is the $r \times r$ matrix with all zero entries.

If D is $(2r+1)\times(2r+1)$, then with respect to the reordering $\{\vec{e}_1, \vec{e}_3, \vec{e}_5, ..., \vec{e}_{2r+1}, \vec{e}_2, \vec{e}_4, \vec{e}_6, ..., \vec{e}_{2r}\}$ the JCF of D^2 is of the form $\begin{bmatrix} X_{r+1} & \mathbf{0}_{r+1,r} \\ \mathbf{0}_{r,r+1} & X_r \end{bmatrix}$ where X_s is an $s \times s$ matrix as in definition 25.17 with $\alpha_i = 0$ and $\mathbf{0}_{a,b}$ is the $a \times b$ matrix with all zero entries.

We are assuming A is of the form $A = \begin{bmatrix} A_{m_1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & A_{m_2} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_{m_3} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & A_{m_n} \end{bmatrix}$ where the A_{m_i} is the $m_i \times m_i$ matrix as in definition 25.17.

If $B^2 = A$, then the $JCF(A) = JCF(JCF(B)^2)$. We apply this to A and try to find its square root B. We necessarily have that all eigenvalues of B are 0, so $JCF(B)$ is made up of matrices as in definition 25.17 with $\alpha_i = 0$. These last matrices are of size $2a_1 \times 2a_1, 2a_2 \times 2a_2, \ldots 2a_m \times 2a_m$ and $(2b_1+1)\times(2b_1+1), (2b_2+1)\times(2b_2+1), \ldots (2b_t+1)\times(2b_t+1)$. We say the sizes of the blocks of B_0 are $2a_1, 2a_2, ..., 2a_m, 2b_1+1, ..., 2b_t+1$. Then the $JCF(JCF(B_0)^2)$ has block entries of sizes $a_1, a_1, a_2, a_2, ..., a_m, a_m, b_1+1, b_1, b_2+1, b_2, ..., b_t+1, b_t$.

Returning to our original A, we see A has a square root only if the blocks corresponding to eigenvalue 0 in A (which we assume is in JCF) can be paired such that either the sizes in a pair are the same or the sizes in a pair differ by 1. This condition is necessary. That it is sufficient is straight forward. To a pair of blocks in A of sizes (d, d) with eigenvalue 0 we include in B a block of size $2d$ with eigenvalue 0. To a pair of blocks in A of sizes $(d, d+1)$

with eigenvalue 0 we include in B a block of size $2d + 1$ with eigenvalue 0. The nonzero eigenvalues have already been addressed. We get $JCF(B^2) = A$ so A has a square root.