Math 350 Fall 2011
Food for thought 1

Here are some interesting problems about the density of diagonalizable matrices that you can think about in your free time.

1. (a) Suppose A is a 2×2 matrix with entries in $\mathbb{C}$, say

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}.$$

Let $f(t) = t^2 + a_1 t + a_0$ be the characteristic polynomial of A.

(i) Find a_1 and a_0 in terms of the entries of A.

(ii) Explain why $f(t)$ has 2 distinct roots if and only if $a_1^2 - 4a_0 \neq 0$. Hence, using also part (i), find a necessary and sufficient condition on the entries of A for which A has 2 distinct eigenvalues over $\mathbb{C}$. This is then a sufficient condition under which A would be diagonalizable over $\mathbb{C}$.

(b) Suppose B is a 2×2 matrix with entries in $\mathbb{C}$. Show that there exists a sequence B_m of 2×2 matrices with entries in $\mathbb{C}$ such that B_m is diagonalizable over $\mathbb{C}$ for all m, and that

$$\lim_{m \to \infty} B_m = B,$$

in the sense that

$$\lim_{m \to \infty} (B_m)_{i,j} = B_{i,j} \quad \text{for all } 1 \leq i, j \leq 2,$$

where $(B_m)_{i,j}$ and $B_{i,j}$ are the (i,j)-th entry of the matrices B_m and B respectively.

2. The above question can be generalized to higher dimensions. To do so, we first develop some algebraic preliminaries in this question.

First, for $j = 1, \dots, n$, we define $s_j(t_1, \dots, t_n)$ to be the coefficient of t^{n-j} in the polynomial $p(t) := (t + t_1) \dots (t + t_n)$. This s_j is usually called the j-th elementary symmetric polynomial of the n variables $t_1, \dots, t_n$.

A bijective map $\sigma \colon \{1, \dots, n\} \to \{1, \dots, n\}$ is usually called a permutation of n letters. The set of all such is denoted S_n, the symmetric group of n letters. A polynomial $P(t_1, \dots, t_n)$ of n variables is said to be symmetric if

$$P(t_{\sigma(1)}, \dots, t_{\sigma(n)}) = P(t_1, \dots, t_n) \quad \text{for all } \sigma \in S_n.$$

(a) Show that

$$s_1(t_1, \dots, t_n) = t_1 + \dots + t_n$$

and

$$s_n(t_1, \dots, t_n) = t_1 \dots t_n.$$

(b) Show that the s_j defined above is a symmetric polynomial of $t_1, \dots, t_n$, for each $j = 1, \dots, n$. Hence, any polynomial in $s_1, \dots, s_n$ is a symmetric polynomial of $t_1, \dots, t_n$.

(c) Show that if $P(t_1, \dots, t_n)$ is a symmetric polynomial of n variables, then there exists a polynomial Q of n variables such that

$$P(t_1, \dots, t_n) = Q(s_1(t_1, \dots, t_n), \dots, s_n(t_1, \dots, t_n)).$$

(Hint: Use a double induction here, first on the number of variables n, then on the total degree of the given polynomial P. In fact, suppose a symmetric polynomial $P(t_1, \dots, t_n)$ of $t_1, \dots, t_n$ is given. Then the polynomial $P(t_1, \dots, t_{n-1}, 0)$, obtained by plugging in $t_n = 0$, is a symmetric polynomial of the $n-1$ variables $t_1, \dots, t_{n-1}$. Hence by induction hypothesis, there exists a polynomial q of $n-1$ variables such that

$$P(t_1, \dots, t_{n-1}, 0) = q(s_1(t_1, \dots, t_{n-1}), \dots, s_{n-1}(t_1, \dots, t_{n-1})).$$

Now consider

$$R(t_1, \dots, t_n) := P(t_1, \dots, t_n) - q(s_1(t_1, \dots, t_n), \dots, s_{n-1}(t_1, \dots, t_n)).$$

Then R is a symmetric polynomial of n variables $t_1, \dots, t_n$. Also, by construction, $R(t_1, \dots, t_{n-1}, 0) = 0$, so t_n divides R, and it follows by the symmetry of R that $t_1 \dots t_n$ divides R, i.e. $s_n(t_1, \dots, t_n)$ divides R. Then

$$\frac{R(t_1, \dots, t_n)}{s_n(t_1, \dots, t_n)}$$

is a symmetric polynomial of n variables that has smaller total degree than $P(t_1, \dots, t_n)$, so by induction hypothesis,

$$\frac{R(t_1, \dots, t_n)}{s_n(t_1, \dots, t_n)} = r(s_1(t_1, \dots, t_n), \dots, s_n(t_1, \dots, t_n))$$

for some polynomial r of n variables. Thus

$$\begin{aligned} &P(t_1, \dots, t_n) \\ =&q(s_1(t_1, \dots, t_n), \dots, s_{n-1}(t_1, \dots, t_n)) \\ &\quad + s_n(t_1, \dots, t_n) r(s_1(t_1, \dots, t_n), \dots, s_n(t_1, \dots, t_n)), \end{aligned}$$

completing the proof of the induction step.)

3. This question is a generalization of Question 1 to n dimensions.

(a) Suppose A is an $n \times n$ matrix with entries in $\mathbb{C}$. Suppose $\lambda_1, \dots, \lambda_n$ is a listing (with multiplicities) of the eigenvalues of A. Let also

$$f(t) = (-1)^n (t^n + a_{n-1} t^{n-1} + \dots + a_0)$$

be the characteristic polynomial of A.

(i) Show that $a_0, \dots, a_{n-1}$ are polynomials in the entries of A.

(ii) Show that $a_j = (-1)^{n-j} s_{n-j}(\lambda_1, \dots, \lambda_n)$ for all $j = 0, \dots, n-1$.

(iii) Suppose

$$P(t_1, \dots, t_n) := \prod_{1 \le i < j \le n} (t_i - t_j)^2.$$

Show that $P(t_1, \dots, t_n)$ is a symmetric polynomial of the n variables $t_1, \dots, t_n$. Hence, using Question 2 and part (ii), show that $P(\lambda_1, \dots, \lambda_n)$ is a polynomial of $a_0, \dots, a_{n-1}$. It then follows from part (i) that $P(\lambda_1, \dots, \lambda_n)$ is a polynomial of the entries of A.

(iv) Suppose P is defined as in part (iii). Show that if $P(\lambda_1, \dots, \lambda_n) \neq 0$, then A is diagonalizable. Hence, using part (iii), find a sufficient condition on the entries of A such that A is diagonalizable.

(b) One can now prove the following: Suppose B is an $n \times n$ matrix with entries in $\mathbb{C}$. Show that there exists a sequence B_m of $n \times n$ matrices with entries in $\mathbb{C}$ such that B_m is diagonalizable over $\mathbb{C}$ for all m, and that

$$\lim_{m \to \infty} B_m = B,$$

in the sense that

$$\lim_{m \to \infty} (B_m)_{i,j} = B_{i,j} \quad \text{for all } 1 \le i, j \le n,$$

where $(B_m)_{i,j}$ and $B_{i,j}$ are the (i, j)-th entry of the matrices B_m and B respectively.

4. The result of Question 3(b) can be restated as follows: the set of all $n \times n$ complex matrices that are diagonalizable over $\mathbb{C}$ is dense in the set of all $n \times n$ complex matrices. On the other hand, show that the set of all *unitarily* diagonalizable matrices is NOT dense in the set of all $n \times n$ complex matrices. (Hint: A complex $n \times n$ matrix is unitarily diagonalizable if and only if it is normal, and the set of normal matrices is a proper closed subset of the set of all $n \times n$ complex matrices.) Also, the result of Question 1(b) is false if $\mathbb{C}$ is replaced by $\mathbb{R}$, since there are many 2×2 real matrices that does not even have real eigenvalues.

5. In this question we see some applications of the result in Question 3(b).

(a) We have seen that if B is a diagonalizable $n \times n$ matrix over $\mathbb{C}$, then

$$\det(I + tB) = 1 + s_1 t + \cdots + s_n t^n$$

for some coefficients $s_1, \dots, s_n$, where $s_1 = \text{trace}(B)$ and $s_n = \det(B)$. Using Question 3(b), show that the same conclusion holds without the assumption that B is diagonalizable. It follows that if B is any $n \times n$ matrix with entries in $\mathbb{C}$, then B is invertible if and only if $\det(I + tB)$ is a polynomial of degree n in t.

(b) Suppose A is an $n \times n$ complex matrix. We define the exponential, sine and cosine of A by the following formula:

$$e^A := \lim_{m \to \infty} \left(I + \frac{1}{1!} A + \frac{1}{2!} A^2 + \cdots + \frac{1}{m!} A^m \right)$$

$$\sin(A) := \lim_{m\to\infty}\left(\frac{1}{1!}A - \frac{1}{3!}A^3 + \frac{1}{5!}A^5 + \cdots + \frac{(-1)^m}{(2m+1)!}A^{2m+1}\right)$$
$$\cos(A) := \lim_{m\to\infty}\left(I - \frac{1}{2!}A^2 + \frac{1}{4!}A^4 + \cdots + \frac{(-1)^m}{(2m)!}A^{2m}\right).$$

Here the limits of the matrices are defined as in Question 3(b). These mimic the definition of the exponential, sine and cosine of a complex number by power series.

(i) Show that if A is diagonalizable over $\mathbb{C}$, say $A = PDP^{-1}$ where D is a diagonal matrix, then

$$e^A = Pe^D P^{-1},$$
$$\sin(A) = P\sin(D)P^{-1} \quad \text{and} \quad \cos(A) = P\cos(D)P^{-1}.$$

(ii) Show that if D is a diagonal matrix with entries $\lambda_1, \dots, \lambda_n$, then e^D is a diagonal matrix with entries $e^{\lambda_1}, \dots, e^{\lambda_n}$. State and prove a similar statement for $\sin(D)$ and $\cos(D)$.

(iii) Show that

$$\sin^2(D) + \cos^2(D) = I$$

for all diagonal matrices D. (Here we write $\sin^2(D)$ for $(\sin(D))^2$, etc.) (Hint: Here you may use freely the fact that $\sin^2(\lambda) + \cos^2(\lambda) = 1$ for all complex numbers λ.)

(iv) Using part (i), show that

$$\sin^2(A) + \cos^2(A) = I$$

if A is diaognalizable over $\mathbb{C}$.

(v) Hence, using Question 3(c), prove that the same conclusion holds without the assumption that A is diagonalizable over $\mathbb{C}$.

(vi) Similarly, show that $\det(e^A) = e^{\mathrm{Trace}(A)}$ for all $n \times n$ complex matrix A. (In particular, we have the conclusions of parts (v) and (vi) for all $n \times n$ real matrices A, which is not very easy to see without first passing to complex matrices!)