Assignment 7

Due: Thursday, April 7.

From the textbook exercises for Chapter 13:

- 1.3
- 2.2
- 3.1
- 3.3
- 10.1

More on determinants. The first part of this assignment is a weird approach to *Sylvester's theorem* (1). We write $M_{nn}(R)$ for the set of $n \times n$ matrices with coefficients in the ring R. If $A \in M_{nn}(\mathbb{C})$, and A has n linearly independent eigenvectors v_k with eigenvalues λ_k, then you can check that $f(A)v_k = f(\lambda_k)v_k = 0$ (because f is the characteristic polynomial and λ_k is a root). If $B \in M_{nn}(\mathbb{F})$ (entries in a field) and $Bv_k = 0$ for a set of basis vectors v_k, then $B = 0$ as a matrix (theorem of linear algebra). This proves Sylvester's theorem for "most" matrices $A \in M_{nn}(\mathbb{C})$, an open and dense set of matrices, it turns out.

We then show that the formula in (1) is true for any $A \in M_{nn}(\mathbb{C})$ using a *continuity* argument, which is straightforward if cumbersome analysis. Finally, we show that if it's true for complex matrices, then it's true for matrices with entries in any ring. There should be a strictly algebraic proof of this algebraic fact, but the present argument illustrates how ideas from different parts of mathematics can be combined.

The idea for going from $\mathbb{C}$ to a general R is that the (i,j) entry of $f(A)$ is a polynomial in the entries of A whose coefficients don't depend on which R the entries come from. More precisely, there are *universal* polynomials $P_{ij} \in \mathbb{Z}[X_{11}, \ldots, X_{nn}]$ (polynomials in n^2 variables with integer coefficients) so that the (i,j) entry of $f(A)$ is $(f(A))_{ij} = P_{ij}(a_{11}, \ldots, a_{nn})$. The left side uses the notation $(B)_{ij}$ for the (i,j) entry of B. The $\mathbb{C}$ version of Sylverster's theorem implies that $P_{ij}(z_{11}, \ldots, z_{nn}) = 0$ for any n^2 set of complex numbers z_{ij}. But a complex polynomial in m complex arguments that is always zero must be the zero polynomial (an argument shows). Thus, each P_{ij} is the zero polynomial, so $f(A) = 0$.

Continuity arguments rely on norms, or distance functions. In finite dimensions, any two distance functions are *equivalent* in that if a sequence converges using one distance function, then it converges using the other. Here are specific distance functions that seem to be easy to define and work with. For polynomials with complex coefficients, $f(X) = a_n X^n + \cdots + a_0$, define $\|f\| = \sum_{k=0}^{n} |a_k|$. For matrices $A \in M_{nn}(\mathbb{C})$, define $\|A\| = \sum_i \sum_j |a_{ij}|$. The distance between polynomials is $\|f - g\|$. A sequence of polynomials f_k converges to f if $\|f_k - f\| \to 0$ as $k \to \infty$. This is equivalent to the convergence of each of the coefficients $a_{k,j} \to a_j$ as $k \to \infty$. Matrix distance and convergence is analogous. Be aware that $\|A\|$ is not a matrix norm related to a vector norm, as is usual in linear algebra. For example, $\|I\| = n$. A matrix norm coming from a vector norm would define the same convergent sequences.

1. (Sylverster for $\mathbb{C}$, distinct eigenvalues). You may use the fact from linear algebra that if $A \in M_n(\mathbb{C})$ has distinct eigenvalues, then the corresponding eigenvectors are a basis of $\mathbb{C}^n$. Give the proof of Sylvester's theorem for this case.

2. (Start of the continuity argument for polynomials)

   (a) Suppose $f(z) = (z - \lambda_1) \cdot \cdots \cdot (z - \lambda_n)$, with $\lambda_j \in \mathbb{C}$. Show that the coefficients a_j of f are continuous functions of the numbers λ_j.

   (b) Suppose $A \in M_{nn}(\mathbb{C})$ and $f(z) = \det(A - zI)$. Show that the coefficients of f are continuous functions of the entries of A.

3. (Next step in the continuity argument)

   (a) Show that there is a monic polynomial $f \in \mathbb{C}[Z]$ and an $\epsilon > 0$ so that if $\|g - f\| \leq \epsilon$, then g has n distinct roots.

   (b) Show that there is a $B \in M_n(\mathbb{C})$ and an $\epsilon > 0$ so that if $\|B - A\| \leq \epsilon$ then A has distinct eigenvalues.

4. (Universal polynomials) Let $f(\lambda)$ be the characteristic polynomial of $A \in M_{nn}(R)$. Show that there are polynomials $P_{ij}(X_{11}, \ldots, X_{nn})$ so that $(f(A))_{ij} = P_{ij}(a_{11}, \ldots, a_{nn})$. Show that each P_{ij} has integer coefficients. Show that $\deg(P_{ij}) \leq n$ for each (i,j). [Definition of degree: The degree of a monomial is $\deg(X_1^{\alpha_1} \cdot \cdots \cdot X_m^{\alpha_m}) = \alpha_1 + \cdots + \alpha_m$. The degree of a polynomial in m variables is the maximum degree of its monomials.]

5. (Continuity argument, using universal polynomials) Show that if $A \in M_{nn}(\mathbb{C})$ and $\|B - A\| \leq \epsilon$ as in Exercise (3b), then $P_{ij}(a_{11}, \ldots, a_{nn}) = 0$.

6. (Finish the continuity argument for complex entries) Suppose $P \in \mathbb{C}[Z_1, \cdots, Z_m]$ "vanishes in the neighborhood of a point", which means that there is a point $b \in \mathbb{C}^m$ and an $\epsilon > 0$ so that if $\|b - z\| = |b_1 - z_1| + \cdots + |b_m - z_m| \leq \epsilon$ then $P(z) = 0$. Show that P is the zero polynomial. *Hint.* If $\alpha = (\alpha_1, \ldots, \alpha_m)$ is a "multi-index" with $\alpha_1 + \cdots + \alpha_n = \deg(P)$, and $a_\alpha z_1^{\alpha_1} \cdots z_m^{\alpha_m}$ is the corresponding monomial of maximal degree in P, then

   $$a_\alpha = C_\alpha \, \partial_{z_1}^{\alpha_1} \cdots \partial_{z_m}^{\alpha_m} P = 0 \; , \quad C_\alpha \neq 0 \; .$$

   These are partial derivatives as defined in multivariate calculus.

7. Sylvester's theorem is that if A is an $n \times n$ matrix with entries in a ring R, then A is a "root" of its characteristic polynomial:

$$f(A) = 0 \,, \quad \text{where } f(\lambda) = \det(A - \lambda I) \,. \tag{1}$$

(It would be a good idea to verify this directly for $n = 2$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, but don't hand in that calculation.)