REU APPRENTICE CLASS #17

INSTRUCTOR: LÁSZLÓ BABAI
SCRIBE: ASILATA BAPAT

Wednesday, July 20, 2011

1. Problem solutions

Today we discussed problems 113, 142, 143(a), 162, 163, 168, 169.

Problem 113. (Presented by Zihao)

Let $v_1, \dots, v_m$ be eigenvectors of a matrix A with distinct eigenvalues $\lambda_1, \dots, \lambda_m$. Suppose that the vectors are linearly dependent, and let k be the smallest integer such that $v_k \in \text{Span}\{v_1, \dots, v_{k-1}\}$. Then we can write the following equation:

$$v_k = \alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1},$$

for some scalars $\alpha_1, \dots, \alpha_{k-1}$.

Now apply the matrix A to the preceding equation, and also multiply the equation by λ_k to obtain the following equations:

$$\lambda_k v_k = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_{k-1} \lambda_{k-1} v_{k-1},$$
$$\lambda_k v_k = \alpha_1 \lambda_k v_1 + \cdots + \alpha_{k-1} \lambda_k v_{k-1}.$$

Subtracting these two equations, we see that

$$0 = \alpha_1(\lambda_1 - \lambda_k) v_1 + \cdots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1}.$$

Since $v_1, \dots, v_{k-1}$ are linearly independent, it follows that $\alpha_i(\lambda_i - \lambda_k) = 0$ for every $i < k$. However, $\lambda_i \neq \lambda_k$ for $i < k$. This means that $\alpha_i = 0$ for all $i < k$.

Since $v_k = \sum_{i=1}^{k-1} \alpha_i v_i$, we have $v_k = 0$. However, v_k was an eigenvector, so v_k cannot be zero. This is a contradiction.

Problem 142. Recall that the Cayley-Hamiton theorem says that an $n \times n$ matrix A satisfies its own characteristic polynomial f_A.

Consider A to be a matrix of variables, that is $A = (x_{ij})$. Then the equation $f_A(A) = 0$ is equivalent to a certain system of n^2 identities in the variables x_{ij}.

Let us compute an example, namely the 2×2 case. In this case, the characteristic polynomial is the following:

$$f_A(t) = (t - x_{11})(t - x_{22}) - x_{12}x_{21} = t^2 - (x_{11} + x_{22})t + x_{11}x_{22} - x_{12}x_{21}.$$

Plugging in the matrix A, we see that

$$A^2 - (x_{11} + x_{12})A + (x_{11}x_{22} - x_{12}x_{21})I = 0.$$

Let us compute the entries of the left hand side of this equation (which is a 2×2 matrix of multivariate polynomials). We see that

$$A^2 = \begin{pmatrix} x_{11}^2 + x_{12}x_{21} & x_{11}x_{12} + x_{12}x_{22} \\ x_{21}x_{11} + x_{22}x_{21} & x_{21}x_{12} + x_{22}^2 \end{pmatrix}.$$

Computing out the entries of the polynomial, we see that the top right entry is

$$(x_{11}x_{12} + x_{12}x_{22}) - (x_{11} + x_{22})x_{12},$$

which is formally zero (all terms cancel).

The top left entry becomes

$$x_{11}^2 + x_{12}x_{21} - (x_{11} + x_{22})x_{11} + (x_{11}x_{22} - x_{12}x_{21}),$$

which is also formally zero.

We know that if a degree-n polynomial of one variable vanishes at more than n places, then it is the zero polynomial. If we can prove a similar result for multivariate polynomials, then it will follow that the n^2 identities that we get from $f_A(A)$ are formal identities (all terms cancel). (Refer to problem 176 for this result.)

Assuming the result of Problem 176, we see that the n^2 identities are formal identities in the ring $\mathbb{Z}[x_{11}, \ldots, x_{nn}]$.

Since the obvious map $\mathbb{Z}[x_{11}, \ldots, x_{nn}] \to \mathbb{F}_q[x_{11}, \ldots, x_{nn}]$ is a homomorphism, the formal identities on the left map to formal identities on the right. Therefore, once we know that the Cayley-Hamiton theorem is true over the integers (which we know because we proved it over the complex numbers), we have also proved it over $\mathbb{F}_q$. In fact, the same is true if we replace $\mathbb{F}_q$ by any commutative ring with identity.

Problem 143. (Jenny)

- Find the traces of both sides. We know that $\mathrm{Tr}(AB - BA) = \mathrm{Tr}(AB) - \mathrm{Tr}(BA) = 0$. However, $\mathrm{Tr}(I) \neq 0$ in characteristic zero.

Problem 162. Recall that an $n \times n$ complex matrix A is called normal if $AA^* = A^*A$. Recall from the complex spectral theorem that a matrix A is normal if and only if it is unitarily similar to a diagonal matrix.

Suppose that $A = (\alpha_{ij})$ is an $n \times n$ normal triangular matrix. We will compute the top-left corner element of each of the two products A^*A and AA^* (which are supposed to be equal).

Let $AA^* = A^*A = (c_{ij})$. If the columns of A are denoted by $\underline{a}_i$, then we evaluate A^*A to get the following:

$$c_{11} = \underline{a}_1^* \underline{a}_1 = \|\underline{a}_1\|^2 = \sum_{i=1}^{n} |\alpha_{i1}|^2.$$

However, $\alpha_{i1} = 0$ if $i > 1$. Therefore $c_{11} = |\alpha_{11}|^2$.

Now we compute the same quantity using AA^*. Let $\underline{r}_1, \ldots, \underline{r}_n$ be the rows of A. Therefore we get

$$c_{11} = \underline{r}_1^* \underline{r}_1 = \|\underline{r}_1\|^2 = \sum_{j=1}^{n} |\alpha_{1j}|^2.$$

Setting the two computed values of c_{11} to be equal, we see that $\alpha_{1j} = 0$ for all $j > 1$, so α_{11} is the only nonzero entry in the first row. Now use induction to finish the proof.

Problem 163. Suppose that A is a unitary matrix with eigenvalue λ. This means that there is a nonzero vector v such that $Av = \lambda v$. Taking conjugate transposes, we see that

$$v^* A^* = (Av)^* = \overline{\lambda} v^*.$$

Now consider the product $(v^* A^*)(Av)$:

$$v^* A^* A v = v^* I v = v^* v = \overline{\lambda}\lambda \cdot v^* v.$$

Observe that $\overline{\lambda}\lambda = |\lambda|^2$. Also, $v^* v = \|v\|^2$, and this is a positive real number since $v \neq 0$. Dividing by this quantity on both sides, we see that $|\lambda|^2 = 1$.

Problem 168. (Presented by Nathan)

The Rayleigh quotient for a nonzero vector x is defined by $R_A(x) = (x^t A x)/(x^t x)$.

First let A be a diagonal matrix:

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Then the Rayleigh quotient of this matrix A is

$$R_A(x) = \frac{\sum_i \lambda_i x_i^2}{\sum_i x_i^2}.$$

Suppose that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

We will first show that the maximum value of $R_A(x)$ is equal to λ_1. Observe that if $x = (1, 0, \dots, 0)$, then $R_A(x) = \lambda_1$. However, $\sum_i \lambda_i x_i^2 \le \sum_i \lambda_1 x_i^2$, so $R_A(x) \le \lambda_1$. This proves that the maximum value of $R_A(x)$ is exactly λ_1. A very similar argument works to show that the minimum value of $R_A(x)$ is equal to λ_n.

To finish the proof for a general real symmetric matrix A, observe that by the spectral theorem, A has an orthonormal eigenbasis, say $e_1, \dots, e_n$ with eigenvalues $\lambda_1 \ge \cdots \ge \lambda_n$ respectively.

Let x be a nonzero vector. Write x as a linear combination of the e_i as follows: $x = \sum_i \alpha_i e_i$. Then

$$\begin{aligned} x^t x &= \sum_i \alpha_i^2, \\ x^t A x &= x^t \left(\sum_i \alpha_i A e_i \right) = \left(\sum_i \alpha_i e_i^t \right) \left(\sum_j \alpha_j \lambda_j e_j \right) \\ &= \sum_i \lambda_i \alpha_i^2. \end{aligned}$$

Now it is clear that the same proof as before works.

Problem 169. (Presented by Hannah)

Recall that a matrix B is called positive definite if B is symmetric and if $x^t B x > 0$ for all non-zero vectors x.

Let B be a positive-definite (real symmetric) matrix. Let λ be an eigenvalue of B. Then there is an eigenvector v with eigenvalue λ. So $Bv = \lambda v$, and therefore $v^t B v = v^t \lambda v = \lambda \|v\|^2$. Since $v^t B v > 0$, and $\|v\|^2 > 0$, we see that $\lambda > 0$.

Now suppose that B is an $n \times n$ real symmetric matrix with eigenvalues $\lambda_1, \dots, \lambda_n$. Suppose that $\lambda_i > 0$ for all i. We will show that B is positive definite.

Using the spectral theorem for real symmetric matrices, we know that B has an orthonormal eigenbasis, say $v_1, \dots, v_n$, such that $Bv_i = \lambda_i v_i$ for every i. Let x be some nonzero vector. Now write x as a linear combination of the v_i, that is, $x = \sum_{i=1}^n \alpha_i v_i$. Therefore $x^t = \sum_{j=1}^n \alpha_j v_j^t$. Therefore

$$\begin{aligned} x^t B x &= \left(\sum_{i=1}^n \alpha_i v_i^t \right) B \left(\sum_{j=1}^n \alpha_j v_j \right), \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j v_i^t B v_j, \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \lambda_j v_i^t v_j. \end{aligned}$$

However, $v_i^t v_j = \delta_{ij}$. Therefore we see that

$$x^t B x = \sum_{i=1}^n \alpha_i^2 \lambda_i.$$

Observe that every term is nonnegative, but at least one of the α_i is nonzero since $x \ne 0$. Therefore $x^t B x > 0$.

2. New problems

Problem 175. Show that the set $U(n) = \{A \in M_n(\mathbb{C}) \mid AA^* = I\}$ of unitary matrices forms a group.

Problem 176. Suppose that $f(x_1, \dots, x_n)$ is a multivariate polynomial over a field F such that the degree of f is at most d in each variable. (E.g., the polynomial $x^3y^3z^2 + 7xyz^3$ satisfies the degree bound $d = 3$.) Let $\alpha_0, \dots, \alpha_d$ be distinct elements of the field. Suppose that for every substitution of values $\beta_i \in \{\alpha_0, \dots, \alpha_d\}$ we have $f(\beta_1, \dots, \beta_n) = 0$. Then $f = 0$ (as a formal polynomial, i.e., all coefficients are zero.)

(Hint: Use induction on the number of variables.)

Problem 177. Show that in $\mathbb{R}^3$, every sense-preserving (orientation-preserving) congruence that fixes a point is a rotation.