Advanced Linear Algebra, Fall 2011.
Solutions to the take-home part of Midterm #2.

1. Let F be a field, $A \in Mat_{m\times n}(F)$ for some m and n, and let $k = rk(A)$.

(a) Prove that if $A = A_1 + \ldots + A_l$ where $rk(A_i) = 1$ for each i, then $l \geq k$.

(b) Prove that there exist matrices $A_1, \ldots, A_k$, with $rk(A_i) = 1$ for each i, such that $A = A_1 + \ldots + A_k$.

Solution: (a) By HW# 6.1, $rk(X+Y) \leq rk(X) + rk(Y)$ for any $X, Y \in Mat_{m\times n}(F)$, and straightforward induction implies that $rk(X_1 + \ldots + X_r) \leq \sum_{i=1}^{r} rk(X_i)$ for any r. Hence

$$k = rk(A) = rk(A_1 + \ldots + A_l) \leq \sum_{i=1}^{l} rk(A_i) = l.$$

(b) Since $rk(A) = k$, there exists k integers $1 \leq j_1 < \ldots < j_k \leq n$ s.t. $col_{j_1}(A), \ldots, col_{j_k}(A)$ are linearly independent (in particular, they are all nonzero), and every column of A is a linear combination of $col_{j_1}(A), \ldots, col_{j_k}(A)$. Thus, if $J = \{j_1, \ldots, j_k\}$, then for any $s \in \{1, \ldots, n\} \backslash J$ there exist $\lambda_{s1}, \ldots, \lambda_{sk}$ s.t. $col_s(A) = \sum_{i=1}^{k} \lambda_{si} col_{j_i}(A)$.

Define the matrices $A_1, \ldots, A_k$ by specifying each column as follows:

$$col_s(A_i) = \begin{cases} col_{j_i}(A) & \text{if } s = j_i \\ 0 & \text{if } s \in J, \text{ but } s \neq j_i \\ \lambda_{si} col_{j_i}(A) & \text{if } s \notin J. \end{cases}$$

Then by construction $A = A_1 + \ldots + A_k$. We claim that each A_i has rank 1. Indeed, $A_i \neq 0$ since $col_{j_i}(A_i) = col_{j_i}(A) \neq 0$, so $rk(A_i) \geq 1$. On the other hand, every column of A_i is a multiple of $col_{j_i}(A)$, so $rk(A_i) \leq 1$.

2. Let V be a finite-dimensional vector space over a field F and let $T \in \mathcal{L}(V)$ be such that $\chi_T(x)$ splits. Let $n = \dim(V)$.

(a) Suppose that T has n distinct eigenvalues. Prove that V has precisely 2^n T-invariant subspaces.

Solution: Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A. By Corollary 16.5, T is diagonalizable, and moreover $\dim E_{\lambda_i}(T) = 1$ for all i (if $\dim E_{\lambda_i}(T) > 1$ for some i, then $\sum_{i=1}^{n} \dim E_{\lambda_i}(T)$ would exceed $n = \dim(V)$).

Since T is diagonalizable, by HW 8.6(e), a subspace W of V is T-invariant $\iff W = \oplus_{i=1}^{n}(E_{\lambda_i}(T) \cap W)$. Since each $E_{\lambda_i}(T)$ is one-dimensional, there are only two possibilities for $E_{\lambda_i}(T) \cap W$: either $E_{\lambda_i}(T) \cap W = W$ or $E_{\lambda_i}(T) \cap W = \{0\}$. Thus there are (at most) 2^n possibilities for $\oplus_{i=1}^{n}(E_{\lambda_i}(T) \cap W)$, so there are at most 2^n T-invariant subspaces.

Let us now show that there are at least 2^n T-invariant subspaces. For every subset J of $\{1, \ldots, n\}$ let $W_J = \oplus_{j \in J} E_{\lambda_j}(T)$. Then W_J is T-invariant, and moreover, $W_J \neq W_{J'}$ for $J \neq J'$. Since $\{1, \ldots, n\}$ has 2^n subsets (if one wants to construct a subset of $\{1, \ldots, n\}$, for each integer $1 \leq i \leq n$ there are two choices: either to include i in the subset or not, and choices for different i are independent of each other). Thus, we exhibited 2^n T-invariant subspaces in V.

(b) Suppose that there exists an ordered basis $\beta = \{v_1, \ldots, v_n\}$ of V s.t. $[T]_\beta$ is a Jordan block of size n corresponding to $\lambda = 0$ (equivalently $(v_1, \ldots, v_n)$ is a nilpotent T-cycle). Prove that V has precisely $n+1$ T-invariant subspaces.

Solution: For each $1 \leq k \leq n$ let $V_k = Span(v_1, \ldots, v_k)$, and put $V_0 = \{0\}$. By definition of T we have $T(v_1) = 0$ and $T(v_i) = v_{i-1}$ for $i \leq 1$. This implies that $T(v_i) \in V_k$ for $1 \leq i \leq k$, and by linearity $T(Span(v_1, \ldots, v_k)) \subseteq V_k$, so each V_k is T-invariant. Thus, we constructed $n+1$ T-invariant subspaces $\{0\} = V_0, V_1, \ldots, V_n = V$, and we need to show that there are no others.

Let W be any T-invariant subspace, and let k be the smallest integer s.t. $W \subseteq V_k$ (such k exists since W is surely contained in $V_n = V$). We will show that $W = V_k$. By definition of k there exists $v \in W \setminus V_{k-1}$, which means that $v = \sum_{i=1}^{k} \lambda_i v_i$ with $\lambda_k \neq 0$. Note that $T^{k-1}(v) = \lambda_k v_1 \neq 0$, but $T^k(v) = 0$, so $(T^{k-1}(v), T^{k-2}(v), \ldots, v)$ is a nilpotent T-cycle with nonzero initial vector. Thus, by Lemma 18.2, the vectors $T^{k-1}(v), T^{k-2}(v), \ldots, v$ are linearly independent. Since W is T-invariant and contains v, it contains all the vectors $T^{k-1}(v), T^{k-2}(v), \ldots, v$, and so $\dim(W) \geq k$. On the other hand, $W \subseteq V_k$ and $\dim(V_k) = k$. This is only possible if $W = V_k$.

(c) Give an example (with proof) where V has infinitely many T-invariant subspaces and T is not scalar, that is, $T \neq \lambda I$ for any $\lambda \in F$.

Solution: Let F be any infinite field. Choose two distinct elements $\lambda, \mu \in F$, let $A = diag(\lambda, \lambda, \mu)$ be the diagonal matrix with diagonal entries λ, λ, μ and $T = L_A : F^3 \to F^3$ be the left multiplication by A. Then T is not scalar, but $T(v) = \lambda v$ for all $v \in Span(e_1, e_2)$.

For each $\alpha \in F$, the subspace $W_\alpha = Span(e_1 + \alpha e_2)$ is a T-invariant subspace of V, and $W_\alpha \neq W_\beta$ for $\alpha \neq \beta$ since $e_1 + \alpha e_2$ and $e_1 + \beta e_2$ are not proportional for $\alpha \neq \beta$. Since F is infinite, we have constructed infinitely many T-invariant subspaces of V.

(d) Give an example where $n = 3$ and V has precisely **six** T-invariant subspaces.

Solution: First we explain a natural way to construct such an example. Again recall that by HW#8.6(e), if $T \in (V)$ is diagonalizable, then a subspace W of V is T-invariant $\iff$ $W = \oplus_{\lambda \in Spec(T)}(W \cap E_\lambda(W))$. If T is not diagonalizable, the above assertion may not be true (e.g., as part (b) shows), but the following generalization does hold:

Theorem A: *Let V be a finite-dimensional vector space and $T \in \mathcal{L}(V)$. For each $\lambda \in Spec(T)$ let $K_\lambda(T)$ be the corresponding generalized eigenspace of T. Then a subspace W of V is T-invariant $\iff$ $W = \oplus_{\lambda \in Spec(T)} W_\lambda$ where W_λ is a T-invariant subspace of $K_\lambda(T)$ for each λ.*

We shall not prove Theorem A in general, but we will establish it in a special case, which is sufficient to solve (e). Anyway, Theorem A suggests a natural way to construct the desired example: since $6 = 2 \cdot 3$, it suffices to find V of dimension 3 and $T \in \mathcal{L}(V)$ with exactly two eigenvalues λ and μ s.t. there are three T-invariant subspace inside $K_\mu(T)$ and two T-invariant subspace inside $K_\lambda(T)$. An example with these properties is easy to produce.

Let F be any field and fix a nonzero element $\lambda \in F$. Let $T : F^3 \to F^3$ be the unique linear map s.t. $T(e_1) = 0$, $T(e_2) = e_1$ and $T(e_3) = \lambda e_3$. Then $Spec(T) = \{0, \lambda\}$, $K_0(T) = Span(e_1, e_2)$ and $K_\lambda(T) = Span(e_3)$. Note that T restricted to $Span(e_1, e_2)$ acts as L_A where $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, so by part (b) there are three T-invariant subspaces inside $K_0(T)$ (namely, $\{0\}, Span(e_1)$ and $Span(e_1, e_2)$). The restriction of T to $Span(e_3)$ is simply multiplication by λ, so by part (a) there are two T-invariant subspaces inside $K_0(T)$ (namely, $\{0\}$ and $Span(e_3)$). According to Theorem A, there are precisely six T-invariant subspaces in F^3, namely

$$\begin{array}{ll} \{0\} & Span(e_3) \\ Span(e_1) & Span(e_1) + Span(e_3) = Span(e_1, e_3) \\ Span(e_1, e_2) & Span(e_1, e_2) + Span(e_3) = V \end{array}$$

(e) It remains to prove (without referring to Theorem A) that in our example from (d) there are no T-invariant subspaces besides those listed above. (This is equivalent to proving Theorem A for the specific T from part (d)).

Let T be as in part (d), and let $W \subseteq F^3$ be a T-invariant subspace. Take any $w = a_1e_1 + a_2e_2 + a_3e_3 \in W$. Then W also contains $T^2(w) = \lambda^2 a_3 e_3$, hence also contains $\frac{1}{\lambda^2}T^2(w) = a_3e_3$ and $w - a_3e_3 = a_1e_1 + a_2e_2$. Thus, $w = w_1 + w_2$ where $w_1 = a_1e_1 + a_2e_2 \in W \cap K_0(T)$ and $w_2 = a_3e_3 \in K_\lambda(T)$, so $W \subseteq (W \cap K_0(T)) \oplus (W \cap K_\lambda(T))$. The opposite inclusion $(W \cap K_0(T)) \oplus (W \cap K_\lambda(T)) \subseteq W$ is obvious, so $W = (W \cap K_0(T)) \oplus (W \cap K_\lambda(T))$.

Finally note that $W \cap K_0(T)$ is T-invariant (being an intersection of T-invariant subspaces) and contained in $K_0(T)$, so $W \cap K_0(T)$ must equal $\{0\}, Span(e_1)$ or $Span(e_1, e_2)$ by part (b), and $W \cap K_\lambda(T)$ must equal to $\{0\}$ or $Span(e_3) = K_\lambda(T)$ by part (a) (or simply because $K_\lambda(T)$ is one-dimensional). Thus, there are at $6 = 2 \cdot 3$ T-invariant subspaces.

3. Let F be a field, $A \in Mat_{n\times n}(F)$, and assume that $\chi_A(x)$ splits. Let $\lambda_1, \ldots, \lambda_t$ be the distinct eigenvalues of A, and m_i the multiplicity of λ_i.

(a) Prove that $tr(A^k) = m_1\lambda_1^k + \ldots + m_t\lambda_t^k$ for each $k \in \mathbb{Z}_{>0}$.

Solution: We start with a basic result about multiplying upper-triangular matrices which can be checked by direct computation. Below for a matrix X by X_{ij} we denote the (i,j)-entry of X.

Lemma: *Let $A, B \in Mat_{n\times n}(F)$ be upper-triangular matrices. Then AB is also upper-triangular and $(AB)_{ii} = A_{ii}B_{ii}$ for all $1 \le i \le n$.*

Using this lemma, by straightforward induction we obtain the following:

Corollary: *Let $A \in Mat_{n\times n}(F)$ be upper-triangular. Then for any $k \in \mathbb{N}$ the matrix A^k is upper-triangular and $(A^k)_{ii} = (A_{ii})^k$ for all $1 \le i \le n$.*

We now use this corollary to solve part (a).

Case 1: A is in JCF. Then A is upper-triangular, with diagonal entires λ_1 appearing m_1 times, …, λ_t appearing m_t times. By Corollary, A^k is also upper-triangular, with diagonal entires λ_1^k appearing m_1 times, …, λ_t^k appearing m_t times. Hence $\mathrm{tr}(A^k) = \sum_{i=1}^t m_i\lambda_i^k$.

General Case: Since $\chi_A(x)$ splits, there exist a matrix J in JCF and an invertible matrix Q s.t. $A = QJQ^{-1}$. We know that $\chi_A(x) = \chi_J(x)$, so A and J have the same eigenvalues with the same multiplicities. So by Case 1,

$$\mathrm{tr}(J^k) = \sum_{i=1}^{t} m_i\lambda_i^k. \qquad (***)$$

Also note that $A^k = (QJQ^{-1})^k = QJ^kQ^{-1}$. Hence by HW#7.5(b) and (***) we have $\mathrm{tr}(A^k) = \mathrm{tr}(J^k) = \sum_{i=1}^t m_i\lambda_i^k$.

(b) Assume that $F = \mathbb{R}$ (real numbers) and $\mathrm{tr}(A^k) = 0$ for all $k \in \mathbb{Z}_{>0}$. Prove that $t = 1$ and $\lambda_1 = 0$ (that is, A has just one eigenvalue and that eigenvalue is 0). Deduce that A is nilpotent.

Solution: We are given that $\mathrm{tr}(A) = 0, \mathrm{tr}(A^2) = 0, \ldots, \mathrm{tr}(A^t) = 0$, so by part (a) for each $1 \le j \le t$ we have $\sum_{i=1}^t m_i\lambda_j^i = 0$. This system of t scalar equations is equivalent to one matrix equation $\Lambda \cdot v = 0$ where $\Lambda \in Mat_{t\times t}(\mathbb{R})$ is given by $\Lambda_{ij} = \lambda_j^i$ and $v = \begin{pmatrix} m_1\lambda_1 \\ \vdots \\ m_t\lambda_t \end{pmatrix}$. Then Λ is a Vandermonde matrix, and since $\lambda_1, \ldots, \lambda_t$ are distinct, Λ is invertible. Thus, we conclude that $v = 0$, so $m_i\lambda_i = 0$ for each i. By assumption, each $m_i > 0$, so we must have $\lambda_i = 0$ for each i. Since $\lambda_1, \ldots, \lambda_t$ are distinct, this is only possible if $t = 1$ and $\lambda_1 = 0$.

Now we prove that A is nilpotent. Let $J = JCF(A)$, so that $A = QJQ^{-1}$ for some Q. Since $A^k = QJ^kQ^{-1}$ for every k, it suffices to show that J is nilpotent.

Since 0 is the only eigenvalue of A, the matrix J is a direct sum of several Jordan blocks with 0 on the diagonal: $J = \oplus_{i=1}^s J(0, n_i)$ where $\sum n_i = n$. We verified earlier that $J(0, k)^k = 0$ for every $k \in \mathbb{N}$. Since $n \geq n_i$ for each i, using the formula for multiplying block-diagonal matrices, we get

$$J^n = \oplus_{i=1}^s J(0, n_i)^n = \oplus_{i=1}^s J(0, n_i)^{n_i} \cdot J(0, n_i)^{n-n_i} = 0.$$

(c) Does the assertion of (b) remain true if $\mathbb{R}$ is replaced by an arbitrary field F? Prove or give a counterexample.

Solution: The proof from part (b) remains valid over any field F of characteristic zero, that is, any field F s.t. $\underbrace{1 + \ldots + 1}_{n \text{ times}} \neq 0$ in F for every positive integer n. The latter condition is indeed used in the proof of (b) (although it was not explicitly mentioned): when we use the equation $m_i\lambda_i = 0$ to derive that $\lambda_i = 0$, we treat m_i not as an integer, but as an element of F represented by that integer (that is, 1_F added to itself m_i times), so we can only derive that $\lambda_i = 0$ if m_i represents a nonzero element of F. This tells us that we need to consider fields of positive characteristic to find a counterexample.

It is easy to see that any field of positive characteristic can be used as a counterexample, but for simplicity we use the familiar fields $\mathbb{Z}_p$ (where p is any fixed prime). Let $A = I_p \in Mat_{p\times p}(\mathbb{Z}_p)$, the $p \times p$ identity matrix over $\mathbb{Z}_p$. Then $A^k = I_p$ for all $k \in \mathbb{N}$, so A is not nilpotent, but $\text{tr}(A^k) = p \times 1 = 0$ for all $k \in \mathbb{N}$.

Remark: As many of you pointed out, part (b) (as stated) admitted a very simple solution – it is enough to consider the equation $\text{tr}(A^2) = 0$ and use the fact that squares of nonzero real numbers are positive. This argument, however, does not work over $\mathbb{C}$ (complex numbers), while the proof presented above carries over without any changes (as explained at the beginning of the solution for (c)).