Math 110, Summer 2012 Long Homework 4 (SOME) SOLUTIONS

Due Wednesday 7/25, 10.10am, in Etcheverry 3109. Late homework will not be accepted.

Please write your answers in complete English sentences (where applicable). Make your arguments rigorous - if something is 'obvious', state why this is the case. Full credit will be awarded to those solutions that are complete and answer the question posed in a coherent manner.

1. In this problem you will show that the nilpotent matrices with one Jordan block are the *regular nilpotent matrices* - this means that the nilpotent class of such matrices is the 'largest'.

Given $A \in Mat_n(\mathbb{C})$ we denote its similarity class by

$$\mathcal{O}(A) = \{B \in Mat_n(\mathbb{C}) \mid A \text{ is similar to } B\}.$$

a) Given $A \in Mat_n(\mathbb{C})$ we define the *commutator of A* to be

$$C(A) = \{B \in Mat_n(\mathbb{C}) \mid AB = BA\}.$$

.

i) Show that $C(A) \subset Mat_n(\mathbb{C})$ is a subspace, for any $A \in Mat_n(\mathbb{C})$.
ii) Suppose that A and B are similar. **Fix** $P \in \mathrm{GL}_n(\mathbb{C})$ **such that** $P^{-1}AP = B$. Show that, for every invertible $X \in C(A)$, $Q^{-1}AQ = B$ where $Q = XP$.
iii) Let $Q \in Mat_n(\mathbb{C})$ be such that $Q^{-1}AQ = B$. Show that there is some invertible $Y \in C(A)$ such that $Q = YP$.
iv) Deduce that for every ordered basis $\mathcal{C} \subset \mathbb{C}^n$ such that $[T_A]_\mathcal{C} = B$ we can associate a unique invertible matrix $X(\mathcal{C}) \in C(A)$ such that $P_{\mathcal{S}\leftarrow\mathcal{C}} = X(\mathcal{C})P$. (*Hint: consider Corollary 1.7.7.*)

   Therefore, we have a defined a function

$$\theta : \{\mathcal{C} \subset \mathbb{C}^n \mid [T_A]_\mathcal{C} = B\} \to C(A) \cap \mathrm{GL}_n(\mathbb{C}) \ ; \ \mathcal{C} \mapsto X(\mathcal{C})$$

v) Show that θ is bijective. (*Hint: for surjectivity use Corollary 1.7.7.*)

Solution:

i) Fix $A \in Mat_n(\mathbb{C})$. Let $X, Y \in C(A), \lambda, \mu \in \mathbb{C}$. Then, we have

$$A(\lambda X + \mu Y) = \lambda AX + \mu AY = \lambda XA + \mu YA = (\lambda X + \mu Y)A.$$

   Hence, $C(A)$ is a subspace of $Mat_n(\mathbb{C})$.

ii) We fix P such that

$$P^{-1}AP = B.$$

   Let $X \in C(A)$ be invertible. Then,

$$(XP)^{-1}A(XP) = P^{-1}X^{-1}AXP = P^{-1}X^{-1}XAP = P^{-1}AP = B.$$

iii) Suppose that

$$Q^{-1}AQ = B.$$

   Then, we must have

$$Q^{-1}AQ = P^{-1}AP \implies PQ^{-1}AQP^{-1} = A,$$

   so that $Y = QP^{-1} \in C(A)$. It is easy to see that $YP = QP^{-1}P = Q$.

iv) Let $\mathcal{C} \subset \mathbb{C}^n$ be an ordered basis such that $[T_A]_{\mathcal{C}} = B$. Thus, we have

$$P_{\mathcal{C} \leftarrow \mathcal{S}} A P_{\mathcal{S} \leftarrow \mathcal{C}} = B.$$

Let $Q = P_{\mathcal{S} \leftarrow \mathcal{C}}$ and $X(\mathcal{C}) = QP^{-1}$. Then,

$$X(\mathcal{C})^{-1} A X(\mathcal{C}) = PQ^{-1}AQP^{-1} = PBP^{-1} = A,$$

so that $X(\mathcal{C}) \in C(A)$. Moreover, we have that $X(\mathcal{C})P = Q = P_{\mathcal{S} \leftarrow \mathcal{C}}$. Suppose that $X' \in C(A)$ is another invertible matrix such that

$$X'P = P_{\mathcal{S} \leftarrow \mathcal{C}} = X(\mathcal{C})P.$$

Then, we must have $X' = X(\mathcal{C})$, so that $X(\mathcal{C}) \in C(A)$ is the unique such matrix with the property that $X(\mathcal{C})P = P_{\mathcal{S} \leftarrow \mathcal{C}}$.

v) We have defined the function

$$\theta : \{\mathcal{C} \subset \mathbb{C}^n \mid [T_A]_{\mathcal{C}} = B\} \to C(A) \cap \mathrm{GL}_n(\mathbb{C}) \; ; \; \mathcal{C} \mapsto X(\mathcal{C}).$$

Let $X \in C(A)$ be invertible. Then, consider the basis $\mathcal{C} = (c_1, \dots, c_n) \subset \mathbb{C}^n$ where

$$XP = [c_1 \ \cdots \ c_n] \in \mathrm{GL}_n(\mathbb{C}).$$

Since XP is invertible then its columns form a basis. Now, we need to show that $[T_A]_{\mathcal{C}} = B$: indeed, we have

$$[T_A]_{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{S}} A P_{\mathcal{S} \leftarrow \mathcal{C}} = (XP)^{-1} A (XP),$$

since $XP = P_{\mathcal{S} \leftarrow \mathcal{C}}$, by definition. Then,

$$(XP)^{-1} A (XP) = P^{-1} X^{-1} A X P = P^{-1} A P = B,$$

as $X \in C(A)$. Thus, $[T_A]_{\mathcal{C}} = B$. We still need to show that $X(\mathcal{C}) = X$: this follows because $XP = P_{\mathcal{S} \leftarrow \mathcal{C}}$ and $X(\mathcal{C})$ is the unique matrix such that this property holds. Hence, $X = X(\mathcal{C})$. Therefore, we have just shown that $\theta(\mathcal{C}) = X$, where $\mathcal{C}$ is the basis defined above, so that θ is surjective.

To show that θ is injective we need to recall the definition of an injective function. Suppose that $\theta(\mathcal{C}) = \theta(\mathcal{C}')$. Then, this means that

$$X(\mathcal{C}) = X(\mathcal{C}') \implies P_{\mathcal{S} \leftarrow \mathcal{C}} = X(\mathcal{C})P = X(\mathcal{C}')P = P_{\mathcal{S} \leftarrow \mathcal{C}'}.$$

As the columns of $P_{\mathcal{S} \leftarrow \mathcal{B}}$ are precisely the vectors in $\mathcal{B}$, for any ordered basis $\mathcal{B}$ (and in the correct order), the above equality of matrices shows that $\mathcal{C} = \mathcal{C}'$. Hence, θ is injective.

We define the *dimension of* $\mathcal{O}(A)$ to be $n^2 - \dim_{\mathbb{C}} C(A)$.[1]

[1] The reason for this definition is (roughly) because we can consider

$$\mathcal{O}(A) = \{Q^{-1}AQ \mid Q \in \mathrm{GL}_n(\mathbb{C})\}.$$

Thus, we can define a surjective function

$$\alpha : \mathrm{GL}_n(\mathbb{C}) \to \mathcal{O}(A) \; ; \; Q \mapsto Q^{-1}AQ.$$

However, this function is not injective. In fact, for every $B \in \mathcal{O}(A)$ (say $P^{-1}AP = B$) we have

$$\alpha^{-1}(B) = \{Q \in \mathrm{GL}_n(\mathbb{C}) \mid \alpha(Q) = B\} = \{XP \mid X \in C(A)\}.$$

You have just shown that there is a bijection

$$\alpha^{-1}(B) \to C(A),$$

for any B. Thus, we could consider the measure of 'noninjectivity' to be $\dim_{\mathbb{C}} C(A)$. Then, we can consider the dimension of $\mathcal{O}(A)$ (= '$\mathrm{im}\alpha$') to be $\dim GL_n(\mathbb{C}) - \dim C(A)$. This is a sort of geometric Rank Theorem result.

b) Consider $Mat_3(\mathbb{C})$. There are three distinct nilpotent classes (as there are three partitions of 3) and any nilpotent $A \in Mat_3(\mathbb{C})$ is similar to precisely one of

$$N_{1^3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ N_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ N_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

i) Show that

$$C(N_{1^3}) = Mat_3(\mathbb{C}), \ C(N_{12}) = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \mid d = f = g = 0, a = e \right\},$$

$$C(N_3) = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \mid d = g = h = 0, a = e = i, b = f \right\} = \left\{ \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{bmatrix} \mid a, b, c \in \mathbb{C} \right\}.$$

ii) Deduce that

$$\dim C(N_{1^3}) = 9, \ \dim C(N_{12}) = 5, \ \dim C(N_3) = 3,$$

and that $\mathcal{O}(N_3)$ has the largest dimension.

Solution:

i) By definition

$$C(N_{1^3}) = \{B \in Mat_3(\mathbb{C}) \mid BN_{1^3} = N_{1^3}B\} = \{B \in Mat_3(\mathbb{C}) \mid 0_3 B = B0_3\} = Mat_3(\mathbb{C}),$$

as $0_3 B = 0_3 = B0_3$, for any $B \in Mat_3(\mathbb{C})$.

By considering an arbitrary matrix

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in Mat_3(\mathbb{C}),$$

and the equality

$$AN_{12} = N_{12}A,$$

you should find that

$$C(N_{12}) = \left\{ \begin{bmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{bmatrix} \mid a, b, c, d, e \in \mathbb{C} \right\}.$$

Similarly, we find that

$$C(N_3) = \left\{ \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{bmatrix} \mid a, b, c \in \mathbb{C} \right\}.$$

ii) It is easy to see the corresponding dimensions by counting the number of free variables we have describing each set. Hence, since

$$\dim \mathcal{O}(N_3) = 9 - \dim C(N_\pi),$$

for π a partition of 3, we see that $\dim \mathcal{O}(N_3) = 6$ is the largest possible.

We will now (partially) show that the results we have obtained for the case $n = 3$ hold in general (ie, for every n we have $\dim \mathcal{O}(N_n)$ is maximal).

The following result will be useful: let $e_{ij} \in Mat_n(\mathbb{C})$ be the matrix with 0 everywhere except a 1 in the ij-entry. Then, we have

$$e_{ij}e_{kl} = \begin{cases} e_{il}, & \text{if } j = k, \\ 0, & \text{otherwise.} \end{cases}$$

You DO NOT have to show this.

c) Consider the nilpotent matrix N_n consisting of one 0-Jordan block. Thus, we have

$$N_n = e_{12} + e_{23} + \ldots + e_{n-1,n} = \sum_{j=1}^{n-1} e_{j,j+1}.$$

i) Show that, for $1 \leq k, l \leq n$, we have

$$N_n e_{kl} - e_{kl} N_n = \begin{cases} -e_{1,l+1}, & \text{if } k = 1,\ 1 \leq l < n, \\ e_{k-1,n}, & \text{if } 1 < k \leq n,\ l = n, \\ e_{k-1,l} - e_{k,l+1}, & \text{if } k \neq 1,\ l \neq n, \\ 0, & \text{if } k = 1, l = n. \end{cases} .$$

ii) Show that, for each $i \in \{0, \ldots, n-1\}$,

$$W_i \stackrel{def}{=} \text{span}_{\mathbb{C}}\{N_n e_{j,j+i} - e_{j,j+i} N_n \mid j+i \leq n \text{ and } j \geq 1\} = \text{span}_{\mathbb{C}}\{e_{j,j+i+1} \mid j+i+1 \leq n \text{ and } j \geq 1\}.$$

Deduce that $\dim W_i = n - 1 - i$.

You have just shown that the i^{th} diagonal[2] of an arbitrary $n \times n$ matrix A is mapped <u>onto</u> the $(i+1)^{st}$ diagonal by the morphism $ad(N_n)$, for $i = 0, \ldots, n-1$.

iii) Show that, for each $i \in \{-1, \ldots, -(n-1)\}$,

$$W_i = \text{span}_{\mathbb{C}}\{N_n e_{j+|i|,j} - e_{j+|i|,j} N_n \mid j+|i| \leq n, j \geq 1\} = \text{span}_{\mathbb{C}}\{e_{j+|i|-1,j} - e_{j+|i|,j+1} \mid j+|i|-1 \leq n \text{ and } j \geq 1\}.$$

Deduce that $\dim W_i = n + i$.

Hint: show that the set $\{e_{j+|i|-1,j} - e_{j+|i|,j+1} \mid j + |i| \leq n$ and $j \geq 1\}$ is linearly independent.

You have just shown that the i^{th} diagonal of an arbitrary $n \times n$ matrix A is mapped <u>injectively into</u> the $(i+1)^{st}$ diagonal by the morphism $ad(N_n)$, for $i = -1, \ldots, -(n-1)$.

iv) Consider the morphism

$$ad(N_n) : Mat_n(\mathbb{C}) \to Mat_n(\mathbb{C})\ ;\ B \mapsto N_n B - B N_n.$$

You have just determined the image of $ad(N_n)$ in ii)-iii): we have (you DO NOT need to justify this)

$$\text{im } ad(N_n) = W_{-(n-1)} \oplus W_{-(n-2)} \oplus \cdots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_{n-1}.$$

Deduce that $\dim ad(N_n) = n(n-1)$ and, using the Rank Theorem, deduce that

$$\dim C(N_n) = n.$$

(*Hint: what is* $\ker ad(N_n)$*?*)

[2] We label the diagonals of an arbitrary $n \times n$ matrix as follows: the main diagonal is the 0^{th} diagonal and the diagonals to the right are labelled $1, \ldots, n-1$ as move we move from left to right. The diagonals to the left of the main diagonal are labelled $-1, -2, \ldots, -(n-1)$ as we move from right to left.

In fact, it can be shown that

$$C(N_n) = \left\{ \begin{bmatrix} c_1 & c_2 & \cdots & c_n \\ 0 & c_1 & & \vdots \\ \vdots & & \ddots & c_2 \\ 0 & \cdots & & c_1 \end{bmatrix} \mid c_1, \dots, c_n \in \mathbb{C} \right\}.$$

Solution:

i) We have

$$N_n = e_{12} + e_{23} + \dots + e_{n-1,n}.$$

Using

$$e_{kl}e_{ij} = \begin{cases} e_{kj}, & l = i \\ 0_n, & l \neq i, \end{cases}$$

the given expressions are easily obtained. For example,

$$N_n e_{1,l} - e_{1,l} N_n = (e_{12} + \dots + e_{n-1,n})e_{1,l} - e_{1,l}(e_{12} + \dots + e_{n-1,n}) = 0_n - e_{1,l+1} = -e_{1,l+1}.$$

The other equalities are similar.

ii) Let $i \in \{0, \dots, n-1\}$. Then,

$$N_n e_{j,j+i} - e_{j,j+i} N_n = (e_{12} + \dots + e_{n-1,n})e_{j,j+i} - e_{j,j+i}(e_{12} + \dots + e_{n-1,n}) = \begin{cases} -e_{1,i+2}, & j = 1, \\ e_{n-i-1,n}, & j = n-i, \\ e_{j-1,j+i} - e_{j,j+i+1}, & 1 < j < n-i. \end{cases}$$

Denote

$$x_j = N_n e_{j,j+i} - e_{j,j+i} N_n.$$

Then, using the results just obtained we have, for $k = 1, \dots, n-i-1$,

$$x_1 + \dots + x_k = -e_{1,i+2} + (e_{1,i+2} - e_{2,i+3}) + (e_{2,i+3} - e_{3,i+4}) + \dots + (e_{k-1,k+i} - e_{k,k+i+1}) = -e_{k,k+i+1}.$$

Hence, we see that, for each $k = 1, \dots, n-i-1$,

$$e_{k,k+i+1} \in \mathrm{span}_{\mathbb{C}}\{N_n e_{j,j+i} - e_{j,j+i} N_n \mid j+i \leq n, j \geq 1\},$$

and these are precisely the basis vectors of diagonal $(i+1)$. Hence,

$$W_i = \mathrm{span}_{\mathbb{C}}\{e_{j,j+i+1} \mid 1 \leq j \leq n = i-1\}.$$

Since the set $\{e_{j,j+i+1} \mid 1 \leq j \leq n-i-1\}$ is linearly independent, we have that

$$\dim W_i = n - i - 1.$$

iii) Let $i \in \{-1, \dots, -(n-1)\}$. Then, for each $j = 1, \dots, n-|i|$,

$$N_n e_{j+|i|,j} - e_{j+|i|,j} N_n = (e_{12} + \dots + e_{n-1,n})e_{j+|i|,j} - e_{j+|i|,j}(e_{12} + \dots + e_{n-1,n}) = e_{j+|i|-1,j} - e_{j+|i|,j+1},$$

so that

$$W_i = \mathrm{span}_{\mathbb{C}}\{e_{j+|i|-1,j} - e_{j+|i|,j+1} \mid 1 \leq j \leq n-|i|\}.$$

Then, if we denote, for $j = 1, \dots, n-|i|$,

$$y_j = e_{j+|i|-1,j} - e_{j+|i|,j+1},$$

we have that $\{y_j \mid 1 \leq j \leq n-|i|\}$ is linearly independent: indeed, suppose that

$$\lambda_1 y_1 + \ldots + \lambda_{n-|i|} y_j = 0_n,$$

then we have

$$\begin{aligned} 0_n &= \lambda_1(e_{|i|,1} - e_{|i|+1,2}) + \ldots + \lambda_{n-|i|}(e_{n-1,n-|i|} - e_{n,n-|i|+1}) \\ &= \lambda_1 e_{|i|,1} + (\lambda_2 - \lambda_1)e_{|i|+1,2} + \ldots + (\lambda_{n-|i|} - \lambda_{n-|i|-1})e_{n-1,n-|i|} - \lambda_{n-|i|}e_{n,n-|i|}, \end{aligned}$$

so that

$$\lambda_1 = 0, (\lambda_2 - \lambda_1) = 0, \ldots, (\lambda_{n-|i|} - \lambda_{n-|i|-1}) = 0, \lambda_{n-|i|} = 0.$$

This implies that

$$\lambda_1 = \ldots = \lambda_{n-|i|} = 0,$$

and $\{y_j\}$ is linearly independent. Hence, $\{y_j\}$ is a basis of W_i and

$$\dim W_i = n - |i| = n + i.$$

d) Now, suppose that π is a partition of n such that $\pi \neq n$. Then, consider the block diagonal matrix

$$N_\pi = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{bmatrix},$$

where each $J_i \in Mat_{n_i}(\mathbb{C})$ is a 0-Jordan block. (*So, we have* $\pi : n_1 + n_2 + \ldots + n_k = n$*, where* $n_1 \geq n_2 \geq \ldots \geq n_k > 0$.

Define, for each i,

$$m_i = n_1 + n_2 + \ldots + n_i, \quad \text{and } m_0 = 0.$$

Show that

$$ad(N_\pi)(e_{m_i+1,m_j}) = 0, \quad \text{for each } i = 0, \ldots, k-1 \text{ and } j = 1, \ldots, k,$$

and deduce that $\dim C(N_\pi) \geq k^2$. In particular, if $k^2 \geq n$ then

$$\dim C(N_\pi) \geq \dim C(N_n),$$

and

$$\dim \mathcal{O}(N_n) \geq \dim \mathcal{O}(N_\pi).$$

Solution: We have

$$\begin{aligned} N_\pi =& e_{12} + \ldots + e_{m_1-1,m_1} + e_{m_1+1,m_1+2} + \ldots + e_{m_2-1,m_2} + e_{m_2+1,m_2+2} + \ldots + e_{m_3-1,m_3} \\ &+ \ldots + e_{m_{k-1}+1,m_{k-1}+2} + \ldots + e_{m_k-1,m_k}, \end{aligned}$$

whenever this sum makes sense (ie if $n_i = n_{i+1} = \ldots = n_k = 1$ then the expression stops at $e_{m_{i-1}-1,m_{i-1}}$).

Then, it is now straightforward to check that

$$N_\pi e_{m_i+1,m_j} - e_{m_i+1,m_j} N_\pi = 0_n,$$

for each $i = 0, \ldots, k-1, j = 1, \ldots, k$. Thus, we have found a linearly independent subset

$$\{e_{m_i+1,m_j} \mid 1 \leq j \leq k, 0 \leq i \leq k-1\} \subset C(N_\pi),$$

so that

$$\dim C(N_\pi) \geq k^2 = |\{e_{m_i+1,m_j} \mid 1 \leq j \leq k, 0 \leq i \leq k-1\}|.$$