MATH 110 - SOLUTIONS TO PRACTICE MIDTERM LECTURE 1, SUMMER 2009

GSI: SANTIAGO CAÑEZ

1. Given vector spaces V and W, $V \times W$ is the vector space given by

$$V \times W = \{(v, w) \mid v \in V \text{ and } w \in W\},$$

with addition and scalar multiplication defined componentwise; i.e.

$$(v, w) + (v', w') := (v + v', w + w') \text{ and } a(v, w) := (av, aw).$$

Prove that a map $T : V \to W$ is linear if and only if its graph, defined by

$$\text{graph } T := \{(v, Tv) \in V \times W \mid v \in V\},$$

is a subspace of $V \times W$.

Proof. Suppose that T is linear. Since $T(0) = 0$, we know that

$$(0, 0) = (0, T(0)) \in \text{graph } T.$$

Now, let $(u, Tu), (v, Tv) \in \text{graph } T$. Then $Tu + Tv = T(u + v)$, so

$$(u, Tu) + (v, Tv) = (u + v, Tu + Tv) = (u + v, T(u + v)) \in \text{graph } T,$$

and thus graph T is closed under addition. Finally, let $a \in \mathbb{F}$ and let $(u, Tu) \in \text{graph } T$. Since $aTu = T(au)$, we have

$$a(u, Tu) = (au, aTu) = (au, T(au)) \in \text{graph } T,$$

so graph T is closed under scalar multiplication. We conclude that the graph of T is a subspace of $V \times W$.

Conversely, suppose that graph T is a subspace of $V \times W$. We must show that T preserves addition and scalar multiplication. First, let $u, v \in V$. Then

$$(u, Tu) \text{ and } (v, Tv) \text{ are in graph } T.$$

Since the graph is closed under addition, we must have

$$(u, Tu) + (v, Tv) = (u + v, Tu + Tv) \in \text{graph } T.$$

But for this to be in the graph of T, we must have $Tu + Tv = T(u + v)$ according to the definition of the graph. Thus T preserves addition. Similarly, let $a \in \mathbb{F}$ and $u \in V$. Then

$$a(u, Tu) = (au, aTu) \in \text{graph } T$$

since graph T is closed under scalar multiplication. Again, this implies that $aTu = T(au)$, so we conclude that T preserves scalar multiplication and is hence linear. □

2. Let $T \in \mathcal{L}(V)$ and suppose that U and W are T-invariant subspaces of V. Prove that $U \cap W$ and $U + W$ are also T-invariant.

Date: July 10, 2009.

Proof. First, let $x \in U \cap W$. Then $x \in U$ and $x \in W$. Since U and W are T-invariant, $Tx \in U$ and $Tx \in W$. Thus $Tx \in U \cap W$ so $U \cap W$ is T-invariant.

Now, let $x \in U + W$. Then $x = u + w$ for some $u \in U$ and $w \in W$. Hence

$$Tx = T(u + w) = Tu + Tw.$$

Since U and W are T-invariant, $Tu \in U$ and $Tw \in W$. Thus Tx can be written as something in U plus something in W, so $Tx \in U + W$ and $U + W$ is T-invariant. □

3. Axler, 2.7

Proof. Suppose first that V is infinite dimensional and let v_1 be any nonzero element of V. Then v_1 does not span V since V is infinite dimensional, so there exists $v_2 \in V$ such that $v_2 \notin \text{span}(v_1)$. Again, since V is infinite dimensional, (v_1, v_2) does not span V, so we can pick $v_3 \in V$ such that $v_3 \notin \text{span}(v_1, v_2)$. Continuing in this manner we construct a sequence $v_1, v_2, \ldots$ which satisfies the required condition. Indeed, for each n, since

$$v_n \notin \text{span}(v_1, \ldots, v_{n-1}),$$

the list $(v_1, \ldots, v_n)$ is linearly independent.

Conversely, suppose we have such a sequence of vectors. If V were finite dimensional, say of dimension n, then any list of more than n vectors would be linearly dependent, contradicting the existence of a sequence such as the one we have. Thus V is infinite dimensional □

4. Prove that $C(\mathbb{R})$ is infinite-dimensional. (Hint: Use the previous problem)

Proof. By the previous problem, it suffices to construct a sequence of functions so that for any n, the first n functions are linearly independent. Consider the functions

$$1, x, x^2, x^3, \ldots \in C(\mathbb{R}).$$

For any n, we claim $(1, x, \ldots, x^n)$ are linearly independent. Indeed, suppose that

$$a_0 + a_1 x + \cdots + a_n x^n = 0.$$

Since we are considering both sides as functions, this equation must hold for every possible value of x. In particular, setting $x = 0$ shows that $a_0 = 0$. Then, taking a derivative and again setting $x = 0$ shows that $a_1 = 0$. Similarly, it follows that all a_i are zero, so that $(1, x, \ldots, x^n)$ is linearly independent for each n. □

5. Axler, 3.8

Proof. By Proposition 2.13 of Axler (every subspace has a complimentary subspace), there exists a subspace U such that

$$V = \operatorname{null} T \oplus U.$$

We claim that this is the U that works. First, $U \cap \operatorname{null} T = \{0\}$ since the above is a direct sum. Now, let $w \in \operatorname{range} T$. Then there is some $v \in V$ so that $Tv = w$. Because of the above direct sum, we can write

$$v = x + u$$

for some $x \in \operatorname{null} T$ and $u \in U$. Note that then

$$w = Tv = T(x + u) = Tx + Tu = Tu$$

since $x \in \operatorname{null} T$. Thus $w \in \{Tu \mid u \in U\}$, showing that $\operatorname{range} T \subseteq \{Tu \mid u \in U\}$. But the opposite containment is just follows from the definition of range, so these two sets are equal as required. □

6*. Axler, 5.13 (Hint: By problem 5.12, it is enough to show that every vector in V is an eigenvector of T. Say $\dim V = n$. Let $v \in V$ be nonzero and show that $\operatorname{span}(v)$ can be written as the intersection of some number of $(n-1)$-dimensional subspaces of V, then use the second problem)

Proof. Let $\dim V = n$ and let $v \in V$ be nonzero. Then we can extend to get a basis

$$(v, v_1, \dots, v_{n-1})$$

of V. Let U_i denote the span of these vectors, excluding v_i. Then U_i is $(n-1)$-dimensional since the vectors used to form these spans are linearly independent, so they form a basis of their own span. Thus by the assumption, each U_i is T-invariant. By the second problem then, their intersection is T-invariant. Note: The second problem only deals with two subspaces at a time, but it is easy to extend the result to any number of subspaces.

We claim that this intersection is $\operatorname{span}(v)$. If so, then $\operatorname{span}(v)$ is T-invariant, implying that v is an eigenvector of T since then Tv would be a multiple of v. Thus problem 5.12 implies that T is a scalar multiple of the identity. Clearly,

$$\operatorname{span}(v) \subseteq U_1 \cap \cdots \cap U_{n-1}$$

since v itself is in each of the U_i. The tricky part is to show the opposite containment.

Let $x \in U_1 \cap \cdots \cap U_{n-1}$. Since v and the v_i's together form a basis of V, we know there is a unique way of writing x as

$$x = av + a_1v_1 + \cdots + a_nv_n.$$

Since $x \in U_1$, we can write x in a unique way as a linear combination of v and $v_2, \dots, v_{n-1}$. This implies that in the expression we have above for x, a_1 must be zero or else we would have more than one way of writing x in terms of the given basis of V. Also, since $x \in U_2$, we can write x in a unique way as a linear combination of v and $v_1, v_3, \dots, v_{n-1}$, implying that a_2 above is zero. Similarly, we conclude that

$$a_1 = \cdots = a_{n-1} = 0,$$

so that $x = av \in \operatorname{span}(v)$. This shows that

$$U_1 \cap \cdots \cap U_{n-1} \subseteq \operatorname{span}(v),$$

so the intersection of the U_i is $\operatorname{span}(v)$ as claimed. (Whew!) □

7. Axler, 2.3

Proof. Since $(v_1 + w, \dots, v_n + w)$ is linearly dependent, there exists scalars $a_1, \dots, a_n$, not all zero, such that

$$a_1(v_1 + w) + \cdots + a_n(v_n + w) = 0.$$

Rewriting this, we have

$$a_1v_1 + \cdots + a_nv_n = -(a_1 + \cdots + a_n)w.$$

Now, $a_1 + \cdots + a_n \neq 0$ since otherwise, the right side above would be zero and hence all the a_i would be zero by the linear independence of the v_i. Thus we can multiply both sides by $(a_1 + \cdots + a_n)^{-1}$, showing that w is a linear combination of the v_i's as required. □

8. Axler, 2.10

Proof. Let $(v_1, \dots, v_n)$ be a basis of V, and let $U_i = \operatorname{span}(v_i)$. Clearly each U_i is one-dimensional. Since any vector in V can be written in a unique way as a sum of elements from the U_i, i.e. as a linear combination of the v_i, V is the direct sum of the U_i according to the definition of direct sum given in the book, which is equivalent to the definition we introduced in class. □

9. Let $T : V \to W$ be a linear map and let $v_1, \dots, v_k \in V$. Prove that if $(Tv_1, \dots, Tv_k)$ is linearly independent in W, then $(v_1, \dots, v_k)$ is linearly independent in V.

Proof. Suppose that
$$a_1 v_1 + \cdots + a_k v_k = 0$$
for some scalars $a_1, \dots, a_k \in \mathbb{F}$. We must show that all a_i are zero. Applying T to both sides gives
$$a_1 T v_1 + \cdots + a_k T v_k = T(0) = 0.$$
Thus since the Tv_i are linearly independent, all the a_i are zero as was to be shown. □

10. Axler, 3.12

Proof. Suppose that $T : V \to W$ is a surjective linear map. Then $\operatorname{range} T = W$, so $\dim \operatorname{range} T = \dim W$. Thus from
$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T = \dim \operatorname{null} T + \dim W,$$
it follows that $\dim V \geq \dim W$.

Conversely, suppose that $\dim V \geq \dim W$. We must construct a surjective linear map T from V to W. Let $(v_1, \dots, v_n)$ be a basis of V and let $(w_1, \dots, w_m)$ be a basis of W. Since $m \leq n$, we can define the map T by setting
$$Tv_i = w_i \text{ for } i = 1, \dots, m, \text{ and } Tv_j = 0 \text{ for } j = m+1, \dots, n,$$
then using the fact that a linear map is determined by what it does to a basis to extend the definition of T linearly to all of V. We claim that this T is then surjective. Indeed, note that $(Tv_1 = w_1, \dots, Tv_m = w_m)$ spans W, so $\operatorname{range} T = W$ as desired. □

11. Let $S : \mathcal{L}(V) \to \mathcal{L}(V)$ be the map defined by
$$S(T) = \begin{cases} T^{-1} & \text{if } T \text{ is invertible} \\ 0 & \text{if } T \text{ is not invertible.} \end{cases}$$
Prove or disprove that S is linear. (Hint: Try to figure out a simple case first, say $V = \mathbb{R}^2$)

Proof. First, for the case of $V = \mathbb{R}^2$, note that
$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
are not invertible, but their sum I is. Hence
$$S(A + B) = S(I) = I \neq 0 + 0 = S(A) + S(B),$$
so S is not linear.

For the general case, we can either pick an isomorphism of V with $\mathbb{F}^n$ and reduce the problem to that of $n \times n$ matrices, in which case something similar to the above will work, or we can proceed as follows. Let $(v_1, \dots, v_n)$ be a basis of V. Define a linear map A by setting
$$Av_1 = v_1 \text{ and } Av_i = 0 \text{ for all other } i,$$

and a linear map B by setting

$$Bv_1 = 0 \text{ and } Bv_i = v_i \text{ for all other } i.$$

It may help to see where these come from by figuring out what their matrices with respect to the given basis would be. Then neither A nor B are injective, since their null spaces are non-trivial, so neither is invertible. Thus

$$S(A) + S(B) = 0 + 0 = 0.$$

However, one can see that $A + B$ is in fact the identity operator since $(A + B)v_i = v_i$ for all i, so $S(A + B) = S(I) = I$. Thus again, S is not linear. □

12. Axler, 5.10

Proof. Without loss of generality, it is enough to show that if λ is a nonzero eigenvalue of T, then λ^{-1} is an eigenvalue of T^{-1}. For the converse, we can simply apply the forward direction to T^{-1} and use the fact that $(T^{-1})^{-1} = T$.

Let v be a nonzero eigenvector of T corresponding to λ. Then

$$Tv = \lambda v.$$

Applying T^{-1} to both sides yields

$$v = T^{-1}(\lambda v) = \lambda T^{-1}v.$$

Thus multiplying through by $1/\lambda$ gives

$$\frac{1}{\lambda}v = T^{-1}v,$$

showing that $1/\lambda$ is an eigenvalue of T^{-1}. □

13. Recall that $T \in \mathcal{L}(V)$ is *nilpotent* if $T^k = 0$ for some positive integer k. Suppose that V is a complex vector space. Prove that $T \in \mathcal{L}(V)$ is nilpotent if and only if 0 is the only eigenvalue of T. (Hint for the backwards direction: There exists a basis of V with respect to which $\mathcal{M}(T)$ is upper- triangular. What can you say about the diagonal entries of $\mathcal{M}(T)$?)

Proof. Suppose that T is nilpotent. Since T acts on a finite-dimensional complex vector space, we know that T has an eigenvalue. We claim that this eigenvalue must be 0. Indeed, let λ be an arbitrary eigenvalue of T and let v be a nonzero eigenvector corresponding to λ. Let k be a positive integer so that $T^k = 0$. Then

$$0 = T^k v = T^{k-1}(Tv) = T^{k-1}(\lambda v) = \lambda T^{k-1}v = \lambda T^{k-2}(Tv) = \lambda^2 T^{k-2}v = \cdots = \lambda^k v.$$

Since $v \neq 0$, this implies that $\lambda^k = 0$, and thus $\lambda = 0$. Hence 0 is the only eigenvalue of T.

Conversely, suppose that 0 is the only eigenvalue of T and say $\dim V = n$. Let $(v_1, \dots, v_n)$ be a basis of V with respect to which T has an upper-triangular matrix, which exists by Theorem 5.13 in the book. Since the diagonal entries of this matrix are then the eigenvalues of T, we see that this matrix has zeroes all along the diagonal. Thus $\mathcal{M}(T)$ looks like

$$\begin{pmatrix} 0 & * & \cdots & * \\ & 0 & \ddots & \vdots \\ & & \ddots & * \\ & & & 0 \end{pmatrix},$$

where we have omitted writing the zeroes below the main diagonal. Note that if you multiply this matrix by itself, you get a similar looking matrix, but now with zeroes also above the main diagonal; taking a cubed power gives you yet another diagonal of zeroes above that, and so on. In other words, each consecutive power of such a matrix has fewer nonzero entries than the power before it. (To get an idea as to why this works, work out the cases of 2×2 and 3×3 matrices of the above form.) It follows that if this is an $n \times n$ matrix, then

$$\mathcal{M}(T)^n = 0,$$

where the 0 on the right denotes the matrix with all its entries equal to zero. By properties of the matrix of a linear map, we have

$$\mathcal{M}(T^n) = \mathcal{M}(T)^n = 0.$$

But the only operator whose matrix with respect to $(v_1, \ldots, v_n)$ is 0 is the zero operator itself, so we conclude that $T^n = 0$, showing that T is nilpotent. $\square$

14. Suppose that U, W, and W' are subspaces of V such that

$$U \oplus W = U \oplus W'.$$

For each $w \in W$, since w is then also in $U \oplus W$, this implies that we can write

$$w = u + w'$$

in a unique way with $u \in U$ and $w' \in W'$. Define a linear map $T : W \to W'$ by setting

$$Tw = w' \text{ as constructed above.}$$

Prove that T is an isomorphism. (Hint: First show that $\dim W = \dim W'$; then it suffices to show that T is injective)

Proof. Since $U \oplus W = U \oplus W'$, we have

$$\dim U + \dim W = \dim(U \oplus W) = \dim(U \oplus W') = \dim U + \dim W',$$

so $\dim W = \dim W'$. Thus, to show T is invertible, it suffices to show that T is injective. So, let $w \in \operatorname{null} T$. Then $Tw = 0$. By the definition of T, this means that to write w as an element of $U \oplus W'$, the piece in W' is zero; i.e. we can write

$$w = u + 0 \text{ for some } u \in U.$$

Thus $w = u \in W$. But U and W form a direct sum, meaning that $U \cap W = \{0\}$. Hence $w \in U \cap W$ must be zero, showing that $\operatorname{null} T = \{0\}$ and thus that T is injective.

We are done, but let us also show that T is surjective directly for completeness sake. Let $w' \in W$. Since then $w' \in U \oplus W' = U \oplus W$, we can write

$$w' = u + w \text{ for some } u \in U \text{ and } w \in W.$$

We claim that $Tw = w'$. To see what Tw is, we must first determine how to write w as an element of U plus and element of W'. But the above equation tells us that

$$w = -u + w',$$

so to write w in the required form, we must use $w' \in W'$. This shows that $Tw = w'$, so T is surjective. $\square$

15*. Recall that an operator $T \in \mathcal{L}(V)$ is *diagonalizable* if there exists a basis of V consisting of eigenvectors of T. Suppose that $T \in \mathcal{L}(V)$ is diagonalizable and that U is a nonzero, T-invariant subspace of V. Prove that $T|_U \in \mathcal{L}(U)$ is diagonalizable. [Hint: First show that if $v_1, \ldots, v_k$ are eigenvectors which correspond to distinct eigenvalues of T and $v_1 + \cdots + v_k \in U$, then $v_i \in U$ for each i. (Hint for the hint: first do the case $k = 2$)]

Proof. To start, suppose that $v_1, \ldots, v_k$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ of T and suppose that $v_1 + \cdots + v_k \in U$. We claim that each $v_i \in U$. Since U is T-invariant, we also have

$$T(v_1 + \cdots + v_k) = Tv_1 + \cdots + Tv_k = \lambda_1 v_1 + \cdots + \lambda_k v_k \in U.$$

Thus,

$$T(v_1 + \cdots + v_k) - \lambda_k(v_1 + \cdots + v_k) = (\lambda_1 - \lambda_k)v_1 + \cdots + (\lambda_{k-1} - \lambda_k)v_{k-1},$$

being the sum of two elements of U, is also in U. Call this vector u. Again, T-invariance implies that T applied to this now, which is

$$Tu = (\lambda_1 - \lambda_k)\lambda_1 v_1 + \cdots + (\lambda_{k-1} - \lambda_k)\lambda_{k-1}v_{k-1},$$

is in U. Hence

$$Tu - \lambda_{k-1}u = (\lambda_1 - \lambda_k)(\lambda_1 - \lambda_{k-1})v_1 + \cdots + (\lambda_{k-2} - \lambda_k)(\lambda_{k-2} - \lambda_{k-1})v_{k-2}$$

is in U. Notice that after each such step, we get a linear combination involving one less v_i. Continuing in this manner, we can show that

$$(\lambda_1 - \lambda_k)(\lambda_1 - \lambda_{k-1}) \cdots (\lambda_1 - \lambda_2)v_1 \in U.$$

Since the eigenvalues are distinct, none of these scalars is zero. Thus, dividing by this scalar, we conclude that $v_1 \in U$ since U is closed under scalar multiplication. Note that at the step before we ended up with the above expression only involving v_1, we would have had

$$(\lambda_1 - \lambda_k) \cdots (\lambda_1 - \lambda_3)v_1 + (\lambda_2 - \lambda_k) \cdots (\lambda_2 - \lambda_3)v_2 \in U.$$

Since we now know $v_1 \in U$, this and the fact that again all the scalars involved are nonzero implies that $v_2 \in U$. Similarly, working our way backwards, we conclude that all v_i are in U.

Now, since T is diagonalizable, there exists a basis of V consisting of eigenvectors of T. By Proposition 5.21, this is equivalent to saying that we decompose V into a direct sum of eigenspaces:

$$V = \operatorname{null}(T - \lambda_1 I) \oplus \cdots \oplus \operatorname{null}(T - \lambda_t I),$$

where $\lambda_1, \ldots, \lambda_t$ are the distinct eigenvalues of T. To show that $T|_U$ is diagonalizable, we must construct a basis of U consisting of eigenvectors of T. The point is that these eigenvectors should actually be elements of U themselves. This is where the claim we proved above will help.

Let $(u_1, \ldots, u_m)$ be any basis of U. Then for each i, we can write

$$u_i = v_{i1} + \cdots + v_{it}$$

for some vectors $v_{ij} \in \operatorname{null}(T - \lambda_j I)$ because of the existence of the above direct sum. Note that then $v_{i1}, \ldots, v_{it}$ are eigenvectors of T corresponding to distinct eigenvalues of T; by the first part of the proof, this implies that each v_{ij} is in U since their sum u_i is. Since the u_i span U and each u_i is a linear combination of the v_{ij}, this means that the collection of the v_{ij} are elements of U which span U. Since there are finitely many of the v_{ij}, we can reduce this list to a basis of U, and this is the basis consisting of eigenvectors of $T|_U$ we are looking for. Thus $T|_U$ is diagonalizable.

(As you can see, this is not an easy result. However, going through the proof and trying to understand it will be very useful in your preparation as it uses many of the ideas and notions we have been covering. Try going through step by step, and, as always, ask if something is unclear.) $\square$

16. Let $p(z) \in P(\mathbb{F})$ be a polynomial and let $T \in \mathcal{L}(V)$. Prove that $\operatorname{null} T$ and $\operatorname{range} T$ are invariant under $p(T)$.

Proof. Write $p(z)$ as $p(z) = a_0 + a_1 z + \cdots + a_n z^n$. We first show that $\operatorname{null} T$ is $p(T)$-invariant. To this end, let $x \in \operatorname{null} T$. We must show that $p(T)x \in \operatorname{null} T$. Hence we compute:

$$T(p(T)x) = T(a_0 x + a_1 Tx + \cdots + a_n T^n x) = T(a_0 x) = a_0 Tx = 0,$$

where the second and last equalities from the fact that $Tx = 0$ since $x \in \operatorname{null} T$. Thus $p(T)x \in \operatorname{null} T$ so $\operatorname{null} T$ is $p(T)$-invariant.

Now, let $w \in \operatorname{range} T$. Again, we want to show that $p(T)w \in \operatorname{range} T$. Since $w \in \operatorname{range} T$, there exists $v \in V$ so that $Tv = w$. We then have

$$\begin{aligned} p(T)w &= (a_0 I + a_1 T + \cdots + a_n T^n)(Tv) \\ &= a_0 Tv + a_1 T^2 v + \cdots + a_n T^{n+1} v \\ &= T(a_0 v + a_1 Tv + \cdots + a_n T^n v) \\ &= T(p(T)v). \end{aligned}$$

Thus, there exists an element, namely $p(T)v$, in V so that applying T to it gives $p(T)w$. This means that $p(T)w \in \operatorname{range} T$, as was to be shown. $\square$