Math 113 — Homework 6

Graham White

May 21, 2013

Book problems, Chapter 8:

22. Let $\{v_1, v_2, v_3, v_4\}$ be a basis of $\mathbb{C}^4$. Define a linear operator T by $T(v_1) = T(v_2) = 0$, $T(v_3) = v_3$, $T(v_4) = v_3 + v_4$. (Why is there a linear operator taking these values? Why is it unique?)

    The matrix of T with respect to this basis is
    $$\mathcal{M}(T) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$
    The matrix of T^2 with respect to the same basis is
    $$\mathcal{M}(T^2) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$
    Note that the matrices of T and T^2, together with the identity matrix, are linearly independent, so the minimal polynomial of T has degree at least three. (Which vector space is this linear independence statement referring to? Could you prove it?)

    Now, the transformation $T(T-1)^2$ is the zero map, because $(T-1)^2$ takes v_3 and v_4 to 0 and v_1 and v_2 to $\operatorname{Span}(v_1, v_2)$, and T takes v_1 and v_2 to 0. Therefore the minimal polynomial of T is $x(x-1)^2$.

    We did not explicitly reference Jordan normal forms in this solution, but you should think about how they were used to produce the construction and proof.

23. If V has a basis of eigenvectors of T, then with respect to this basis, the matrix of T is diagonal. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvales of T. Then each element of our basis is killed by $(T - \lambda_i)$ for some i, while the other elements of the basis are multiplied by some scalars. Hence $(T - \lambda_1) \cdots (T - \lambda_k) = 0$. Therefore the minimal polynomial of T is a factor of $(x - \lambda_1) \cdots (x - \lambda_k)$, so has no repeated roots.

    Consider a Jordan basis of V for T. If the minimal polynomial of T has no repeated roots, then the matrix of T with respect to this basis has no 1s above the diagonal, as if it did, then that Jordan block would be nonzero in $(T - \lambda_1) \cdots (T - \lambda_k)$, a contradiction. Therefore this matrix is diagonal. By Proposition 5.21 of Axler, this implies that V has a basis of eigenvectors of T.

27. Let $\{v_1, v_2, v_3, v_4\}$ be a basis of $\mathbb{C}^4$. Define a linear operator T by $T(v_1) = 0$, $T(v_2) = v_2$, $T(v_3) = v_3$, $T(v_4) = 3v_4$.

    The matrix of T with respect to this basis is
    $$\mathcal{M}(T) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Note that the characteristic polynomial of T is $x(x-1)^2(x-3)$. Each root of the characteristic polynomial must also be a root of the minimal polynomial, so the minimal polynomial must have degree at least three. The operator $T(T-1)(T-3)$ takes each element of our basis to 0, so is the zero operator. Therefore the minimal polynomial of T is $x(x-1)(x-3)$, so T has the required minimal and characteristic polynomials.

29. The only eigenvalue of a nilpotent operator is zero, so the characteristic polynomial of N is x^n. The minimal polynomial of N is a factor of the characteristic polynomial, so the minimal polynomial of N is x^k for some $k \leq n$. Consider the Jordan normal form of N. The operator N^k is zero if and only if $N_i^k = 0$ for each Jordan block N_i of this matrix. But direct calculation shows that if N_i is an $m \times m$ matrix with ones above the diagonal and zeros elsewhere, then $N_i^m \neq 0$ and $N_i^{m+1} = 0$.

    Combining these results, we get that the minimal polynomial of N is x^m, where the largest Jordan block of N is $m \times m$. An $m \times m$ Jordan block has $m-1$ ones above the diagonal, so this is the required result.

30. Let $n = \dim(V)$. If there is a direct sum decomposition $V = U \oplus W$ into proper subspaces of V, then the dimensions of U and W are both strictly less than n. The minimal polynomial of T is the lowest common multiple of the minimal polynomials of T restricted to U and to W, which are polynomials of degree less than n. But $(x-\lambda)^n$ is not the lowest common multiple of two polynomials of degree less than n. Therefore if there is such a direct sum decomposition of V, then the minimal polynomial of T is not $(x-\lambda)^n$.

    Let the minimal polynomial of T be $p(x)$. If $p(x)$ is not $(x-\lambda)^n$, then either $p(x) = (x-\lambda)^k$ for some $k < n$, or $p(x)$ has more than one root.

    If $p(x) = (x-\lambda)^k$ for some $k < n$, consider a Jordan basis for T. The largest Jordan block in the matrix of T with respect to this basis is $k \times k$, with $k < n$, so there is more than one Jordan block. Let U be the subspace spanned by the basis vectors corresponding to the columns of one Jordan block, and W be the subspace spanned by all of the other basis vectors. Then U and W are invariant under T and $V = U \oplus W$.

    If $p(x)$ has more than one root, then we may write $p(x) = q(x)r(x)$ as the product of two polynomials of degree at least 1 which have no common roots. Consider the subspaces $\ker(q(T))$ and $\ker(r(T))$. Any element of their intersection is in $\ker(\gcd(q,r)(T)) = \ker(1) = \{0\}$. Applying the rank-nullity theorem to $q(T)$, $r(T)$ and $p(T) = q(T)r(T)$ gives us that $\dim(\ker(q(T))) + \dim(\ker(r(T))) \geq n$, so V is the direct sum $\ker(q(T)) \oplus \ker(r(T))$.

    We have showed that there is no direct sum decomposition of V into invariant subspaces of T if and only if the minimal polynomial of T is of the form $(x-\lambda)^n$, as required.

Book problems, Chapter 9:

3. For any eigenvalue λ of A_1, consider a corresponding eigenvector. Appending an appropriate number of zeros to this vector, we get a λ–eigenvector of A. Likewise, for any eigenvector of any of the A_i, we may prepend and append appropriate numbers of zeros to produce an eigenvector of A with the same eigenvalue.

    Now, let λ be an eigenvalue of A, with a corresponding eigenvector v. The vector v is nonzero, so at least one of its entries is nonzero. Let k be such that at least one of the entries of v in the range corresponding to A_k is nonzero. Let A_k be an $n \times n$ matrix, and let v_k be the $n \times 1$ vector comprised of corresponding n entries of v. Directly calculating Av_k, we see that $A_k v_k = \lambda v_k$. Therefore λ is an eigenvalue of A_k.

    (Note that it was necessary to choose v_k nonzero to show that λ was an eigenvalue of A_k. If you try to use just A_1 and v_1, then you won't be able to do this)

    We have shown that a scalar is an eigenvalue of one of the A_i if and only if it is an eigenvalue of A, proving the required result.

10. For the sake of a contradiction, assume that there is some vector space V and some $T \in \mathcal{L}(V)$ such that $\dim(\ker(T^2 + \alpha T + \beta I)^k)$ is odd-dimensional, for some I. Choose V such that $n = \dim(V)$ is minimal among such counterexamples.

    Let v be any nonzero element of $U = \ker(T^2 + \alpha T + \beta I)^k$. Powers of T commute with one another, so $v, Tv, T^2v, \ldots$ are all elements of U. Let W be the subspace spanned by these vectors. The subspace W is even dimensional, because any linear dependence must divide the minimal polynomial of $T|_U$, which is a power of $T^2 + \alpha T + \beta I$ and thus has only even dimensional factors.

The subspace W is invariant under T, by definition, so T induces a linear operator on the quotient space V/W. The dimension of the kernel of $(T^2 + \alpha T + \beta I)^k$ on V/W is $\dim(\ker(T^2 + \alpha T + \beta I)^k) - \dim(W)$, which is odd by assumption. Therefore we have a smaller counterexample, contradicting the minimality of n.

Hence $\dim(\ker(T^2 + \alpha T + \beta I)^k)$ is always even dimensional, for any V, T and k.

Alternative:

We know that the subspace $U = \ker(T^2 + \alpha T + \beta I)^k$ is T–invariant, because powers of T commute with one another. Therefore, if U is odd-dimensional, then T has an eigenvector in U (Any linear operator on an odd-dimensional real vector space has an eigenvector). But then the corresponding eigenvalue is a root of $(x^2 + \alpha x + \beta)^k$, which has only two roots, both of which are nonreal. This is a contradiction, so U is even dimensional.

Other problems:

1. (a) In order to show that $\mathcal{L}(V \times W, X)$ is a vector space, it suffices to show that it is a subspace of the vector space of functions from $V \times W$ to X. To check this, let f, g be arbitrary bilinear maps, $\mathbf{0}$ the zero map from $V \times W$ to X and c, d, k arbitrary scalars.

   We have that

$$\begin{aligned} \mathbf{0}(v, cw + dw') &= 0 \\ &= c\mathbf{0}(v, w) + d\mathbf{0}(v, w') \end{aligned}$$

$$\begin{aligned} (f + g)(v, cw + dw') &= f(v, cw + dw') + g(v, cw + dw') \\ &= cf(v, w) + df(v, w') + cg(v, w) + dg(v, w') \\ &= c(f + g)(v, w) + d(f + g)(v, w') \end{aligned}$$

$$\begin{aligned} (kf)(v, cw + dw') &= kf(v, cw + dw') \\ &= kcf(v, w) + kdf(v, w') \\ &= c(kf)(v, w) + d(kf)(v, w') \end{aligned}$$

   The linearity in the first variable is verified similarly.

   We have showed that the set of bilinear maps contains the zero map, is closed under addition and is closed under scalar multiplication, so it is a subspace of the vector space of functions from $V \times W$ to X. Therefore it is a vector space in its own right.

   (b) Consider a bilinear function f and any element (v, w) of $V \times W$. Writing each of v and w as a linear combination of basis elements, we have that

$$\begin{aligned} f(v, w) &= f(a_1 v_1 + \ldots + a_k v_k, b_1 w_1 + \ldots + b_l w_l) \\ &= \sum_{i=1}^{k} a_i f(v_i, b_1 w_1 + \ldots + b_l w_l) \\ &= \sum_{i=1}^{k} \sum_{j=1}^{l} a_i b_j f(v_i, w_j) \end{aligned}$$

   But this expression depends only on the values of f on pairs of basis elements (v_i, w_j) and on the constants a_i and b_j, which are uniquely determined by v and w. Therefore $f(v, w)$ can be determined from the values taken by f on pairs of basis elements (v_i, w_j).

It is easily checked that the function

$$f(v,w) = \sum_{i=1}^{k}\sum_{j=1}^{l} a_i b_j f(v_i, w_j)$$

is bilinear for any set of fixed values $f(v_i, w_j)$. Therefore we have shown that fixing the value of f on each pair of basis elements arbitrarily determines a unique bilinear map.

For each $1 \leq i \leq k$ and each $1 \leq j \leq l$, let f_{ij} be the unique bilinear function which is 1 on (v_i, w_j) and 0 on other pairs of basis elements.

From the above calculation, we have that any bilinear function f is equal to $\sum_{i=1}^{k}\sum_{j=1}^{l} f(v_i, w_j) f_{ij}$, so the f_{ij} span $\mathcal{L}(V \times W, \mathbb{F})$.

Now, assume that we have a linear combination of the f_{ij} which is equal to the zero function. That is,

$$\mathbf{0} = \sum_{i=1}^{k}\sum_{j=1}^{l} a_{ij} f_{ij}.$$

Then for each i, j, we have that

$$\mathbf{0}(v_i, w_j) = \sum_{i=1}^{k}\sum_{j=1}^{l} a_{ij} f_{ij}(v_i, w_j).$$

Therefore $0 = a_{ij}$, by the definition of the f_{ij}. This is true for each of the a_{ij}, so the set of f_{ij} is linearly independent. This set is linearly independent and spans the space of bilinear maps, so is a basis for that space. Therefore the dimension of $\mathcal{L}(V \times W, \mathbb{F})$ is kl.

2. (a) As in the hint, we know that $V_{\mathbb{C}} = V \oplus V$ is a vector space over $\mathbb{R}$. In order to show that it is a vector space over $\mathbb{C}$, we just need to prove that those vector space axioms involving scalars are also true for this wider class of scalars. Namely, we must show the two distributivity axioms and the compatibility of multiplication.

   Let $(a+bi)$ and $c+di$ be arbitrary complex numbers and $v \oplus iv'$ and $w \oplus iw'$ be any elements of $V_{\mathbb{C}}$. We have that

$$\begin{aligned}(a+bi)(v \oplus iv' + w \oplus iw') &= (a+bi)(v + w \oplus i(v'+w')) \\ &= a(v+w) - b(v'+w') \oplus i(a(v'+w') + b(v+w)) \\ &= av - bv' \oplus i(av' + bv) + aw - bw' \oplus i(aw' + bw) \\ &= (a+bi)(v \oplus iv') + (a+bi)(w \oplus iw')\end{aligned}$$

$$\begin{aligned}(a+bi+c+di)(v \oplus iv') &= (a+c+(b+d)i)(v \oplus iv') \\ &= (a+c)v - (b+d)v' \oplus i((a+c)v' + (b+d)v) \\ &= av - bv' \oplus i(av' + bv) + cv - dv' \oplus i(cv' + dv) \\ &= (a+bi)(v \oplus iv') + (c+di)(v \oplus iv')\end{aligned}$$

$$\begin{aligned}((a+bi)(c+di))(v \oplus iv') &= (ac - bd + i(ad+bc))(v \oplus iv') \\ &= (ac-bd)v - (ad+bd)v' \oplus i((ac-bd)v + (ad+bc)v') \\ &= (a+bi)(cv - dv' \oplus i(cv' + dv)) \\ &= (a+bi)((c+di)(v \oplus iv'))\end{aligned}$$

   Therefore $V_{\mathbb{C}}$ is a vector space over $\mathbb{C}$.

It is easily verified that the map from $\mathbb{R}^n_{\mathbb{C}}$ to $\mathbb{C}^n$ which takes

$$(a_1, \ldots, a_n) \oplus i(b_1, \ldots, b_n) \longmapsto (a_1 + ib_1, \ldots, a_n + ib_n)$$

is an isomorphism.

(b) We know that $V_{\mathbb{C}} = V \oplus V$ is spanned as a real vector space by

$$\{(v_1, 0), \ldots, (v_n, 0), (0, v_1), \ldots, (0, v_n)\}.$$

But for each v_j, $i(v_j, 0) = (0, v_j)$, so the n elements $\{(v_1, 0), \ldots, (v_n, 0)\}$ span $V_{\mathbb{C}}$ as a complex vector space.

Now, assume that some $\mathbb{C}$–linear combination of the elements $\{(v_1, 0), \ldots, (v_n, 0)\}$ is zero. That is,

$$(a_1 + ib_1)(v_1, 0) + \ldots + (a_n + ib_n)(v_n, 0) = (0, 0).$$

Then by the definition of complex scalar multiplication, we have that

$$(a_1 v_1, b_1 v_1) + \ldots + (a_n v_n, b_n v_n) = (0, 0).$$

Therefore

$$(a_1 v_1 + \ldots + a_n v_n, b_1 v_1 + \ldots + b_n v_n) = (0, 0).$$

But $\{v_1, \ldots, v_n\}$ are linearly independent, so each a_j and each b_j are equal to zero, so each complex scalar $a_j + ib_j$ is equal to zero, so the set $\{(v_1, 0), \ldots, (v_n, 0)\}$ is $\mathbb{C}$–linearly independent.

Therefore the set $\{(v_1, 0), \ldots, (v_n, 0)\}$ is a basis for $V_{\mathbb{C}}$ as a vector space over $\mathbb{C}$.

(c) Define a map $T_{\mathbb{C}} : V_{\mathbb{C}} \longrightarrow V_{\mathbb{C}}$ by $T_{\mathbb{C}}(v, v') = (T(v), T(v'))$. Note that $T_{\mathbb{C}}(v, 0) = (T(v), 0)$, so $T_{\mathbb{C}}$ extends T.

We now check that $T_{\mathbb{C}}$ is a linear map. Let (v, v') and (w, w') be arbitrary elements of $V_{\mathbb{C}}$ and let $a + bi$ be any complex number.

$$\begin{aligned}
T_{\mathbb{C}}((v, v') + (w, w')) &= T_{\mathbb{C}}(v + w, v' + w') \\
&= (T(v + w), T(v' + w')) \\
&= (T(v) + T(w), T(v') + T(w')) \\
&= (T(v), T(w)) + (T(v'), T(w')) \\
&= T_{\mathbb{C}}(v, w) + T_{\mathbb{C}}(v', w')
\end{aligned}$$

$$\begin{aligned}
T_{\mathbb{C}}((a + bi)(v, v')) &= T_{\mathbb{C}}(av - bv', av' + bv) \\
&= (T(av - bv'), T(av' + bv)) \\
&= (aT(v) - bT(v'), aT(v') + bT(v)) \\
&= (a + bi)(T(v), T(v')) \\
&= (a + bi)T_{\mathbb{C}}(v, v')
\end{aligned}$$

Therefore the map $T_{\mathbb{C}}$ is linear, as required.

Notice that by the definition of $T_{\mathbb{C}}$, if $T(v_j) = a_{1j} v_1 + \ldots + a_{nj} v_n$, then $T_{\mathbb{C}}(v_j, 0) = a_{1j}(v_1, 0) + \ldots + a_{nj}(v_n, 0)$, so the required matrix equality holds, by the definition of the matrix of a linear transformation.

(d) We check that complex conjugation is an anti-linear map.

$$\begin{aligned}\overline{v \oplus iv' + w \oplus iw'} &= \overline{v + w \oplus i(v' + w')} \\ &= v + w \oplus -i(v' + w') \\ &= v \oplus -iv' + w \oplus -iw' \\ &= \overline{v \oplus iv'} + \overline{w \oplus iw'}\end{aligned}$$

$$\begin{aligned}\overline{(a + ib)(v \oplus iv')} &= \overline{av - bv' \oplus i(av' + bv)} \\ &= av - bv' \oplus -i(av' + bv) \\ &= (a - ib)(v \oplus iv') \\ &= \overline{(a + ib)}(v \oplus iv')\end{aligned}$$

Let A and C be complex anti-linear maps and B be a complex linear map. Then $A \circ B \circ C$ is a complex linear map, as follows

$$\begin{aligned}ABC(v + w) &= AB(C(v) + C(w)) \\ &= A(BC(v) + BC(w)) \\ &= ABC(v) + ABC(w)\end{aligned}$$

$$\begin{aligned}ABC(kv) &= AB(\overline{k}(Cv)) \\ &= A(\overline{k}BC(v)) \\ &= \overline{\overline{k}}ABC(v) \\ &= kABC(v)\end{aligned}$$

The maps $\overline{T_{\mathbb{C}}}$ and $T_{\mathbb{C}}$ agree on each $(v_i, 0)$, so they are equal.

(e) Let $\lambda = x + iy$. If $v_1 \oplus iv_2$ is a λ–eigenvector of $T_{\mathbb{C}}$, then $T_{\mathbb{C}}(v_1 \oplus iv_2) = (x + iy)(v_1 \oplus iv_2)$. Expanding this out, we get that $T(v_1) = xv_1 - yv_2$ and that $T(v_2) = xv_2 + yv_1$. Hence

$$\begin{aligned}T_{\mathbb{C}}(v_1 \oplus -iv_2) &= T(v_1) \oplus -iT(v_2) \\ &= xv_1 - yv_2 \oplus -i(xv_2 + yv_1) \\ &= (x - iy)(v_1 \oplus -iv_2)\end{aligned}$$

Therefore $v_1 \oplus -iv_2$ is an eigenvector of $T_{\mathbb{C}}$ with eigenvalue $\overline{\lambda} = x - iy$.

We know that $T(v_1) = xv_1 - yv_2$ and that $T(v_2) = xv_2 + yv_1$, so the subspace U is invariant under T. If λ is not real, then $T(v_1) \in U$ is not a real multiple of v_1, so U is at most two dimensional. By definition, U is spanned by two vectors, so U is exactly two dimensional.

We know that eigenvectors corresponding to different eigenvalues are linearly independent, so $v_1 \oplus v_2$ and $v_1 \oplus -v_2$ span U. The polynomial

$$(T - (x + iy))(T - (x - iy)) = T^2 - (\lambda + \overline{\lambda})T + \lambda\overline{\lambda}$$

sends each of these vectors to zero, so is the zero operator on U. Therefore U is a two-dimensional T–invariant subspace with associated eigenpair $(-(\lambda + \overline{\lambda}), \lambda\overline{\lambda})$, as required.

(f) In the calculations of the previous part, if $v_1 = 0$, then

$$T_{\mathbb{C}}(v_1 \oplus iv_2) = T(v_1) \oplus iT(v_2) = i(xv_2 + yv_1) = i(xv_2),$$

so $v_1 \oplus iv_2$ is an eigenvector of T with eigenvalue x. But $v_1 \oplus iv_2$ is an eigenvector of T with eigenvalue λ, so λ is real. A similar calculation shows that if $v_2 = 0$ then λ is real.

If either of v_1 or v_2 is zero, then the other is a λ–eigenvector of T. If both are nonzero, then both $v_1 \oplus iv_2$ and $v_1 \oplus -iv_2$ are λ–eigenvectors of $T_{\mathbb{C}}$, so $v_1 \oplus 0$ is a λ–eigenvector of $T_{\mathbb{C}}$, so v_1 is a λ–eigenvector of T. Likewise, v_2 is a λ–eigenvector of T. (Why can't v_1 and v_2 both be zero?)

We know that $T_{\mathbb{C}}$ has an eigenvalue, because $V_{\mathbb{C}}$ is a finite dimensional complex vector space. If this eigenvalue is complex, then the previous part gives us a two dimensional T–invariant subspace of V. If it is real, then this part shows that T has an eigenvector, so the span of that vector is a one-dimensional T–invariant subspace of V. Either way, V has a subspace which is T–invariant and of dimension either 1 or 2.

If v_1 and v_2 are linearly independent, then this part actually gives us two different one-dimensional T–invariant subspaces of V, and these span a two-dimensional T–invariant subspace.