**Math 561 - Homework 8**

SOLUTIONS

*Prof. Arturo Magidin*

1. Let $\mathbf{V}$ be a vector space over $\mathbb{F}$, and let $T$ be a linear operator on $\mathbf{V}$. Let $q(t)$ be a polynomial with coefficients in $\mathbb{F}$. We saw in class that $q(T)$ is a linear operator on $\mathbf{V}$.

   (i) Prove that if $\mathbf{W}$ is a $T$-invariant subspace of $\mathbf{V}$, then $\mathbf{W}$ is also $q(T)$-invariant.

   **Proof.** Write $q(t) = a_0 + a_1t + \cdots + a_kt^k$.

   If $\mathbf{W}$ is $T$-invariant, then it is $T^m$-invariant for any $m$: this holds for $m = 1$, and if $\mathbf{W}$ is both $T$ and $T^k$ invariant and $\mathbf{w} \in \mathbf{W}$, then $T^{k+1}(\mathbf{w}) = T(T^k(\mathbf{w})) \in T(\mathbf{W}) \subseteq \mathbf{W}$.

   If $\mathbf{W}$ is $U$-invariant, then it is also $aU$-invariant for any scalar $a$, since given $\mathbf{w} \in \mathbf{W}$, $U(\mathbf{w}) \in \mathbf{W}$, and therefore $a(U(\mathbf{w})) = (aU)(\mathbf{w}) \in \mathbf{W}$, since $\mathbf{W}$ is a subspace.

   Finally, if $\mathbf{W}$ is both $U$-invariant and $R$-invariant, then it is $U + R$ invariant: if $\mathbf{w} \in \mathbf{W}$, we have $U(\mathbf{w}) \in \mathbf{W}$ and $R(\mathbf{w}) \in \mathbf{W}$, and because $\mathbf{W}$ is a subspace, $(U+R)(\mathbf{w}) = U(\mathbf{w}) + R(\mathbf{w})$ is in $\mathbf{W}$.

   So if $\mathbf{W}$ is $T$-invariant, then it is $a_1T$-, $a_2T^2$-, ..., and $a_kT^k$-invariant, and also $a_0I_{\mathbf{V}}$-invariant (since every subspace is $I_{\mathbf{V}}$-invariant), and therefore also

   $$q(T) = a_0I_{\mathbf{V}} + a_1T + \cdots + a_kT^k$$

   invariant, as desired. □

   (ii) Prove that $\mathbf{R}(q(T))$, the image of $q(T)$, is a $T$-invariant subspace of $\mathbf{V}$.

   **Proof.** We know the range of a linear transformation is a subspace. To prove that it is $T$-invariant, let $\mathbf{x} \in \mathbf{R}(q(T))$. Then there exists $\mathbf{v} \in \mathbf{V}$ such that $\mathbf{x} = q(T)(\mathbf{v})$. Then

   $$T(\mathbf{x}) = T(q(T)(\mathbf{v})) = (Tq(T))(\mathbf{v}) = (q(T)T)(\mathbf{v}) = q(T)(T(\mathbf{v})) \in \mathbf{R}(q(T)),$$

   since we saw in class that for any two polynomials $p(t)$ and $q(t)$, we have $p(T)q(T) = q(T)p(T)$. So if $\mathbf{x} \in \mathbf{R}(q(T))$, then $T(\mathbf{x}) \in \mathbf{R}(q(T))$, proving that $\mathbf{R}(q(T))$ is $T$-invariant, as claimed. □

2. Let $T\colon \mathbb{R}^4 \to \mathbb{R}^4$ be the linear operator given by $T(a,b,c,d) = (a+b, b-c, a+c, a+d)$. (You may take for granted that $T$ is a linear operator). Find $C(\mathbf{v})$, where $\mathbf{v} = (1,0,0,0)$, and determine the characteristic polynomial of $T|_{C(\mathbf{v})}$.

   **Answer.** We evaluate $T^k(\mathbf{v})$ until we obtain a vector that is a linear combination of the previous vectors. We have:

   $$\begin{aligned}
   \mathbf{v} &= (1,0,0,0)\\
   T(\mathbf{v}) &= (1,0,1,1)\\
   T^2(\mathbf{v}) &= (1,-1,2,2)\\
   T^3(\mathbf{v}) &= (0,-3,3,3) = 0(1,0,0,0) - 3(1,0,1,1) + 3(1,-1,2,2)\\
   &= 0\mathbf{v} - 3T(\mathbf{v}) + 3T^2(\mathbf{v}).
   \end{aligned}$$

   So

   $$C(\mathbf{v}) = \operatorname{span}\Big\{(1,0,0,0), (1,0,1,1), (1,-1,2,2)\Big\}.$$

   Notice that this subspace contains $(1,0,0,0)$, $(0,0,1,1)$, and $(0,1,0,0)$, which are linearly independent; and since $\dim(C(\mathbf{v})) = 3$, they also form a basis. So

   $$C(\mathbf{v}) = \Big\{(a,b,c,c) \;\Big|\; a,b,c \in \mathbb{R}\Big\}.$$

The characteristic polynomial of $T|_{C(\mathbf{v})}$ is given by the way in which $T^3(\mathbf{v})$ is expressed as a linear combination of the other vectors, Since

$$0\mathbf{v} + 3T(\mathbf{v}) - 3T^2(\mathbf{v}) + T^3(\mathbf{v}) = (0,0,0,0),$$

we have that the characteristic polynomial is $p(t) = (-1)^3(3t - 3t^2 + t^3)$.

3. Let $\mathbf{V} = \mathsf{M}_{2\times 2}(\mathbb{R})$ be the vector space of $2 \times 2$ matrices with real coefficients, and let $T\colon \mathbf{V} \to \mathbf{V}$ be the linear transformation given by

$$T(A) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} A.$$

(You do not need to prove that $T$ is a linear operator). Find the $T$-cyclic subspace generated by $z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

**Answer.** We have:

$$\begin{aligned} z &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ T(z) &= \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \\ T^2(z) &= \begin{pmatrix} 3 & 3 \\ 6 & 6 \end{pmatrix} = 3T(z). \end{aligned}$$

So the $T$-cyclic subspace generated by $z$ is

$$\operatorname{span}\left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \right\}.$$

4. Let $\mathbf{V}$ be a finite dimensional vector space, and let $T$ be a linear operator on $\mathbf{V}$. Let $\mathbf{W}$ be a $T$-invariant subspace of $\mathbf{V}$. Prove that if the characteristic polynomial of $T$ splits, then the characteristic polynomial of $T|_{\mathbf{W}}$ also splits.

**Proof.** Polynomials enjoy unique factorization into irreducible factors. If $p(t)$ is the characteristic polynomial of $T$ and $q(t)$ is the characteristic polynomial of $T|_{\mathbf{W}}$, we know that $q(t)$ divides $p(t)$. So every irreducible factor of $q(t)$ is an irreducible factor of $p(t)$, which is therefore linear. Thus, $q(t)$ factors into linear terms; that is, splits. □

5. Let $\mathbf{V}$ be a finite dimensional vector space, and let $T$ be a linear operator on $\mathbf{V}$. Prove that if the characteristic polynomial of $T$ splits, and $\mathbf{W}$ is a nonzero $T$-invariant subspace of $\mathbf{V}$, then $\mathbf{W}$ contains an eigenvector of $T$.

**Proof.** Let $q(t)$ be the characteristic polynomial of $T|_{\mathbf{W}}$. Then $q(t)$ has a root (since it splits and has degree $\dim(\mathbf{W}) \geq 1$) which is also a root of the characteristic polynomial of $T$ and therefore an eigenvalue of $T$. Thus, $T|_{\mathbf{W}}$ has an eigenvalue, and therefore an eigenvector $\mathbf{w} \in \mathbf{W}$. But if $\mathbf{w} \neq \mathbf{0}$ satisfies $T|_{\mathbf{W}}(\mathbf{w}) = \lambda\mathbf{w}$, then $T(\mathbf{w}) = \lambda\mathbf{w}$ so $\mathbf{w}$ is an eigenvector of $T$. Therefore, $\mathbf{W}$ contains an eigenvector of $T$, as claimed. □

6. Let $A$ be an $n \times n$ matrix. Prove that $\dim(\operatorname{span}(I_n, A, A^2, A^3, \ldots)) \leq n$.

**Proof.** Let $p(t)$ be the characteristic polynomial of $A$. Then $p(t)$ has degree $n$, and by the Cayley-Hamilton Theorem for matrices, we know that $p(A) = 0$, the zero $n \times n$ matrix. That means that $A^n$ is a linear combination of $I_n$, $A$, $A^2, \ldots, A^{n-1}$, and therefore that

$$\operatorname{span}(I_n, A, A^2, A^3, \ldots) \subseteq \operatorname{span}(I_n, A, A^2, \ldots, A^{n-1}).$$

Since the space is spanned by $n$ vectors, it has dimension at most $n$, so

$$\dim(\text{span}(I_n, A, A^2, A^3, \ldots)) \leq n,$$

as claimed. □

7. Let $A$ be an $n \times n$ matrix with characteristic polynomial

$$f(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0.$$

(i) Prove that $A$ is invertible if and only if $a_0 \neq 0$.
**Proof.** We have:

$$\begin{aligned} A \text{ is invertible} &\iff 0 \text{ is not an eigenvalue of } A \\ &\iff 0 \text{ is not a root of } f(t) \\ &\iff f(0) \neq 0 \\ &\iff a_0 \neq 0. \qquad \square \end{aligned}$$

(ii) Prove that if $A$ is invertible, then

$$A^{-1} = -\frac{1}{a_0}\left((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2 A + a_1 I_n\right).$$

**Proof.** By the Cayley-Hamilton Theorem, we know that $f(A) = 0$, so

$$0 = (-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n.$$

Multiplying through by $A^{-1}$, we have:

$$\begin{aligned} 0 &= (-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n + a_0 A^{-1} \\ -a_0 A^{-1} &= (-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n \\ A^{-1} &= -\frac{1}{a_0}\left((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n\right), \end{aligned}$$

as claimed. □

8. Let $T$ be a linear operator on $\mathbf{V}$, let $\mathbf{v} \neq \mathbf{0}$, and let $C(\mathbf{v})$ be the $T$-cyclic subspace generated by $\mathbf{v}$. Prove that for any $\mathbf{x} \in \mathbf{V}$, $\mathbf{x} \in C(\mathbf{v})$ if and only if there exists a polynomial $g(t)$ such that $\mathbf{x} = g(T)(\mathbf{v})$.

**Proof.** Since $C(\mathbf{v})$ is spanned by $\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots$, $\mathbf{x} \in C(\mathbf{v})$ if and only if there exists an $n \geq 0$ and scalars $a_0, \ldots, a_n$ such that

$$\mathbf{x} = a_0\mathbf{v} + a_1 T(\mathbf{v}) + a_2 T^2(\mathbf{v}) + \cdots + a_n T^n(\mathbf{v}).$$

Thus, if $\mathbf{x} \in C(\mathbf{v})$ then we can let $g(t) = a_0 + a_1 t + \cdots + a_n t^n$, and then $\mathbf{x} = g(T)(\mathbf{v})$. Conversely, if there exists $g(t) = b_0 + b_1 t + \cdots + b_m t^m$ with $g(T)(\mathbf{v}) = \mathbf{x}$, then

$$\mathbf{x} = g(T)(\mathbf{v}) = b_0\mathbf{v} + b_1 T(\mathbf{v}) + \cdots + b_m T^m(\mathbf{v}) \in C(\mathbf{v}),$$

as desired. □

9. Let $T$ be a linear operator on a two-dimensional vector space $\mathbf{V}$. Prove that either $\mathbf{V} = C(\mathbf{v})$ for some $\mathbf{v} \in \mathbf{V}$, or else $T = cI$ for some scalar $c$.

   **Proof.** If $\mathbf{V} = C(\mathbf{v})$ for some $\mathbf{v} \in \mathbf{V}$, we are done. So assume this is not the case. That means that for every $\mathbf{v} \neq \mathbf{0}$, we have that $\mathbf{v}, T(\mathbf{v})$ is linearly dependent, so $T(\mathbf{v}) = \lambda\mathbf{v}$ for some scalar $\lambda$; that is, every nonzero vector is an eigenvector of $T$.

   Let $\mathbf{v}$ and $\mathbf{w}$ be two linearly independent vectors in $\mathbf{V}$, so that they form a basis; since they are nonzero, they are eigenvectors. Let $\lambda$ be the eigenvalue corresponding to $\mathbf{v}$ and $\mu$ the eigenvalue corresponding to $\mathbf{w}$. We claim that $\lambda = \mu$.

   Indeed, since $\mathbf{v}+\mathbf{w} \neq \mathbf{0}$, then $\mathbf{v}+\mathbf{w}$ is also an eigenvector of $T$; let $\kappa$ be the eigenvalue corresponding to $\mathbf{v} + \mathbf{w}$. Then

   $$\lambda\mathbf{v} + \mu\mathbf{w} = T(\mathbf{v}) + T(\mathbf{w}) = T(\mathbf{v} + \mathbf{w}) = \kappa(\mathbf{v} + \mathbf{w}) = \kappa\mathbf{v} + \kappa\mathbf{w}.$$

   That means that $(\lambda - \kappa)\mathbf{v} = (\kappa - \mu)\mathbf{w}$. But since $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, this forces $\lambda - \kappa = \kappa - \mu = 0$, so $\mu = \kappa = \lambda$.

   So for any vector $\mathbf{x} \in \mathbf{V}$, writing $\mathbf{x} = a\mathbf{v} + b\mathbf{w}$ we have

   $$T(\mathbf{x}) = aT(\mathbf{v}) + bT(\mathbf{w}) = a\lambda\mathbf{v} + b\lambda\mathbf{w} = \lambda(a\mathbf{v} + b\mathbf{w}) = \lambda\mathbf{x}.$$

   Thus, $T = \lambda I$, as claimed. □

10. Let $T$ be a linear operator on a vector space $\mathbf{V}$, and suppose that $\mathbf{V} = C(\mathbf{v})$ for some $\mathbf{v} \in \mathbf{V}$. Prove that if $U$ is a linear operator on $\mathbf{V}$, then $TU = UT$ if and only if $U = g(T)$ for some polynomial $g(t)$. HINT: Use Problem 8.

    **Proof.** If $U = g(T)$, then $UT = g(T)T = Tg(T) = TU$; this holds regardless of whether $\mathbf{V}$ is $T$-cyclic or not.

    Now assume that $\mathbf{V}$ is $T$-cyclic, $\mathbf{V} = C(\mathbf{v})$, and $UT = TU$. If $\dim(\mathbf{V}) = n$, then

    $$\beta = [\mathbf{v}, T(\mathbf{v}), \ldots, T^{n-1}(\mathbf{v})]$$

    is a basis for $\mathbf{V}$.

    Since $U(\mathbf{v}) \in \mathbf{V} = C(\mathbf{v})$, by Problem 8 there is a polynomial $g(t)$ such that $U(\mathbf{v}) = g(T)(\mathbf{v})$. Then

    $$\begin{aligned}
    U(\mathbf{v}) &= g(T)(\mathbf{v}) \\
    U(T(\mathbf{v})) &= T(U(\mathbf{v})) && \text{(since } UT = TU\text{)} \\
    &= T(g(T)(\mathbf{v})) = g(T)(T(\mathbf{v})) \\
    U(T^2(\mathbf{v})) &= T(U(T(\mathbf{v})) && \text{(since } UT = TU\text{)} \\
    &= T(g(T)(T(\mathbf{v})) = g(T)(T^2(\mathbf{v})).
    \end{aligned}$$

    Assuming that $U(T^k(\mathbf{v})) = g(T)(T^k(\mathbf{v}))$, we have

    $$\begin{aligned}
    U(T^{k+1}(\mathbf{v})) &= UT(T^k(\mathbf{v})) \\
    &= T(UT^k(\mathbf{v})) && \text{(since } UT = TU\text{)} \\
    &= T(g(T)(T^k(\mathbf{v}))) \\
    &= g(T)(T^{k+1}(\mathbf{v})).
    \end{aligned}$$

    Thus, $U(T^m(\mathbf{v})) = g(T)(T^m(\mathbf{v}))$ for all $m \geq 0$. This means that $U$ and $g(T)$ agree on $\beta$, and therefore agree on all of $\mathbf{V}$; we conclude that $U = g(T)$, as desired. □