MAT 67 **Homework 3**
August 25, 2014 Indrajit Jana

Problem 6.1: Let $dim(U) = m$ and $dim(V) = n$, where $m \leq n$. Let $\{u_1, \ldots, u_m\}$ be a basis of U. We can extend it to a basis of V, say $\{u_1, \ldots, u_m, v_1, \ldots, v_{n-m}\}$. We know that $S : U \to W$ is a linear map, i.e, $S(u_1), \ldots, S(u_m)$ are known.

We would like to find a linear map $T : V \to W$ such that $T(u) = S(u)$ for all $u \in U$. It is enough to define T on the basis of V, i.e., it is enough to specify $T(u_1), \ldots T(u_m), T(v_1), \ldots, T(v_{n-m})$. Let us define $T : V \to W$ as

$$\begin{aligned} T(u_i) &= S(u_i) \quad \forall\, i = 1, \ldots, m \\ \text{and } T(v_j) &= 0_W \quad \forall\, j = 1, \ldots, n-m. \end{aligned} \tag{1}$$

Then $T : V \to W$ is a well defined linear map (because we have specified the action of T on the basis vectors of V). Also the action of T on the basis of U is same as that of S i.e., $T(u_i) = S(u_i)$ for all $i = 1, \ldots, m$. Now if $u \in U$, then we can write $u = c_1 u_1 + \cdots + c_m u_m$ for some scalars $c_i \in \mathbb{F}$, $i = 1, \ldots, m$. Then from the definition of T we can see that

$$\begin{aligned} T(u) &= T(c_1 u_1 + \cdots + c_m u_m) \\ &= c_1 T(u_1) + \cdots + c_m T(u_m) \\ &= c_1 S(u_1) + \cdots + c_m S(u_m) \quad \text{(from (1))} \\ &= S(c_1 u_1 + \cdots + c_m u_m) \\ &= S(u), \end{aligned}$$

i.e, $T(u) = S(u)$ for all $u \in U$.

Remark: You can call the above result as *Operator extension theorem*. Basically we had an operator $S : U \to W$ defined on a subspace (i.e., U) of V. Now we have extended it to an operator T which is defined on the big space V such that $T|_U = S$.

Problem 6.2: Consider the following equation

$$c_1 T(v_1) + c_2 T(v_2) + \cdots + c_n T(v_n) = 0_W. \tag{2}$$

Now if we want to prove that $T(v_1), \ldots, T(v_n)$ are linearly independent, then we need to show that $c_i = 0$ for all $i = 1, \ldots, n$. We can rewrite the equation (2) as

$$T(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n) = 0_W.$$

The above implies that $c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \in null(T)$. But it is given that T is injective, therefore $null(T) = \{0_V\}$. Hence

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = 0_V.$$

But it is also given that $v_1, v_2, \ldots, v_n$ are linearly independent. Therefore we must have $c_i = 0$ for all $i = 1, \ldots, n$. We are done.

Problem 6.4: Since $T : V \to W$ is a linear map, we have $T(v_i) \in W$ for all $i = 1, \ldots, n$ and hence $span\{T(v_1), \ldots, T(v_n)\} \subset W$.

Conversely, let us take $w \in W$. We need to show that $w \in span\{T(v_1), \ldots, T(v_n)\}$. Since $T : V \to W$ is surjective, there must exist $v \in V$ such that $T(v) = w$. We know that $span\{v_1, \ldots, v_n\} = V$, therefore we can write $v = c_1 v_1 + \cdots + c_n v_n$ for some scalars $c_i \in \mathbb{F}$, $i = 1, \ldots, n$. Consequently

$$\begin{aligned} w &= T(v) \\ &= T(c_1 v_1 + \cdots + c_n v_n) \\ &= c_1 T(v_1) + \cdots + c_n T(v_n). \end{aligned}$$

The above implies that $w \in span\{T(v_1), \ldots, T(v_n)\}$. Hence the proof.

Problem 6.8: Suppose both T and S are invertible. Then T^{-1} and S^{-1} exists and they are linear maps from V to V. Therefore we can define $S^{-1} \circ T^{-1} : V \to V$. Now we see that

$$\begin{aligned} & (S^{-1} \circ T^{-1}) \circ (T \circ S) = S^{-1} \circ (T^{-1} \circ T) \circ S = S^{-1} \circ I_V \circ S = S^{-1} \circ S = I_V \\ \text{and} \quad & (T \circ S) \circ (S^{-1} \circ T^{-1}) = T \circ (S \circ S^{-1}) \circ T^{-1} = T \circ T^{-1} = I_V. \end{aligned}$$

Therefore $S^{-1} \circ T^{-1}$ is the inverse of $T \circ S$. In other words $T \circ S$ is invertible.

Conversely, suppose $T \circ S$ is invertible. We want to show that both S and T are invertible. Since $S \in \mathcal{L}(V, V)$ and $T \in \mathcal{L}(V, V)$, using theorem 6.7.6 it is enough to show that both S and T are injective.

Suppose S is not injective. There exist a nonzero vector $v \in V$ such that $S(v) = 0$. But then $(T \circ S)(v) = T(S(v)) = T(0) = 0$, which implies that $T \circ S$ is also not injective - a contradiction. Because we know that $T \circ S$ is invertible, therefore it must be injective.

Suppose T is not injective then T is not surjective either (using theorem 6.7.6) i.e., $range(T) \subsetneqq V$. In other words, $dim(range(T)) < dim(V)$. Now we know that $range(T \circ S) \subset range(T)$. [1] Therefore $range(T \circ S) \subsetneqq V$, i.e., $T \circ S$ is not surjective - a contradiction. Because we know that $T \circ S$ is invertible, therefore it must be surjective.

Problem 7.3: Suppose λ is an eigenvalue of T, and v is a corresponding eigenvector. Then we have $T(v) = \lambda v$. Since T is invertible, we have

$$\begin{aligned} & T^{-1}(T(v)) = T^{-1}(\lambda v) \\ i.e., \quad & v = \lambda T^{-1}(v) \quad (\text{since } T^{-1} \in \mathcal{L}(V, V),\ T^{-1}(cv) = cT^{-1}(v) \text{ for any } c \in \mathbb{F}) \\ i.e., \quad & \frac{1}{\lambda} v = T^{-1}(v). \end{aligned}$$

[1] This follows simply from the definition of *range*. Let $T(S(v)) \in range(T \circ S)$, then $v \in V$. Now since $v \in V$ and $S : V \to V$, we have $S(v) \in V$. Therefore $T(S(v)) \in range(T)$

Therefore $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}.

Conversely, let $\frac{1}{\lambda}$ be an eigenvalue of T^{-1} and w be a corresponding eigenvector. Then

$$\begin{aligned} & T^{-1}(w) = \frac{1}{\lambda}w \\ i.e., \quad & T(T^{-1}(w)) = T\left(\frac{1}{\lambda}w\right) \\ i.e., \quad & w = \frac{1}{\lambda}T(w) \\ i.e., \quad & \lambda w = T(w). \end{aligned}$$

Therefore λ is an eigenvalue of T.

Problem 7.4: Since every vector $v \in V$ is an eigenvector of T, there exist scalars $\lambda_v \in \mathbb{F}$ such that $T(v) = \lambda_v v$ [2]. Note that the scalars λ_v may depend on the vector v. We need to show that all λ_v are same. In other words, we have to prove that $\lambda_v = \lambda_w$ even if $v \neq w$.

Let us consider two independent vectors $v, w \in V$. Then we know that $v - w \in V$, and therefore $v - w$ is an eigenvector of T (since it is given that all vectors in V are eigenvectors of T). So there exists a scalar $\lambda_{v-w} \in \mathbb{F}$ such that $T(v-w) = \lambda_{v-w}(v-w)$. But we know that $T(v) = \lambda_v v$ and $T(w) = \lambda_w w$. Therefore

$$\begin{aligned} & T(v-w) = \lambda_{v-w}(v-w) \\ \Rightarrow \quad & T(v) - T(w) = \lambda_{v-w}(v-w) \\ \Rightarrow \quad & \lambda_v v - \lambda_w w = \lambda_{v-w}(v-w) \\ \Rightarrow \quad & (\lambda_v - \lambda_{v-w})v + (\lambda_{v-w} - \lambda_w)w = 0. \end{aligned}$$

But v and w are chosen to be independent, therefore both $\lambda_v - \lambda_{v-w} = 0$ and $\lambda_{v-w} - \lambda_w = 0$. Which implies that $\lambda_v = \lambda_{v-w} = \lambda_w$ i.e., $\lambda_v = \lambda_w$. Therefore all λ_v s are the same. Let us rewrite the 'same' λ_v as λ [3]. Then we have $T(v) = \lambda v$ for all $v \in V$. Or we can say that $T = \lambda I_V$ i.e., T is a scalar multiple[4] of the identity map on V.

Problem 7.6: Let $\mu \in \mathbb{C}$ be an eigenvalue of T. Then there exists a nonzero vector $v \in V$ such that $T(v) = \mu v$. Therefore

$$\begin{aligned} & T^2(v) = T(T(v)) = T(\mu v) = \mu T(v) = \mu^2 v \\ & \vdots \\ & T^k(v) = \mu^k v \quad \forall\, k \in \mathbb{N}. \end{aligned}$$

Now if $p(z) \in \mathbb{C}[z]$ is a polynomial of degree n i.e., $p(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$,

[2] Notice that I have denoted the scalars as λ_v. I put the subscript v in order to indicate that the scalar λ_v may depend on v. In other words the same scalar may not work for two different vectors.

[3] It means that λ_v s are no longer v dependent. They are all same.

[4] The scalar is λ.

then using the above equations we have

$$\begin{aligned} p(T)v &= (a_nT^n + a_{n-1}T^{n-1} + \cdots + a_1T + a_0I)(v) \\ &= a_nT^n(v) + a_{n-1}T^{n-1}(v) + \cdots + a_1T(v) + a_0v \\ &= a_n\mu^n v + a_{n-1}\mu^{n-1}v + \cdots + a_1\mu v + a_0v \\ &= (a_n\mu^n + a_{n-1}\mu^{n-1} + \cdots + a_1\mu + a_0)v \\ &= p(\mu)v. \end{aligned}$$

The above equation suggests that $p(\mu)$ is an eigenvalue of $p(T)$.

Conversely, suppose $\lambda \in \mathbb{C}$ is an eigenvalue of $p(T)$. We want to show that there exists $\mu \in \mathbb{C}$ such that $p(\mu) = \lambda$ and μ is an eigenvalue of T. Since λ is an eigenvalue of $p(T)$, there exists a nonzero vector $v \in V$ such that $p(T)v = \lambda v$ i.e., $(p(T) - \lambda I)v = 0_V$. Suppose $p(z) = a_nz^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$. Then using the fundamental theorem of algebra we can say that there exist $\mu_1, \mu_2, \ldots, \mu_n \in \mathbb{C}$ such that

$$p(z) - \lambda = a_nz^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0 - \lambda = a_n(z - \mu_1)(z - \mu_2)\cdots(z - \mu_n). \tag{3}$$

So

$$\begin{aligned} & (p(T) - \lambda I)v = 0_V \\ \Rightarrow \quad & (T - \mu_1 I)(T - \mu_2 I)\cdots(T - \mu_n I)v = 0_V. \end{aligned}$$

Since $v \neq 0_V$, the above implies that $(T - \mu_1 I)(T - \mu_2 I)\cdots(T - \mu_n I)$ is not injective, hence not invertible. Therefore at least one of $(T - \mu_1 I), (T - \mu_2 I), \ldots, (T - \mu_n I)$ is not injective [5]. Suppose $(T - \mu_k I)$ is non injective. Then there exists a non zero vector $w \in V$ such that

$$\begin{aligned} & (T - \mu_k I)w = 0_V \\ \Rightarrow \quad & T(w) = \mu_k w. \end{aligned}$$

In other words μ_k is an eigenvalue of T. Also from (3) we notice that $p(\mu_k) - \lambda = 0$ i.e., $p(\mu_k) = \lambda$.

[5] Because if all of them are injective i.e., invertible, then by problem 6.8 we can say that $(T - \mu_1 I)(T - \mu_2 I)\cdots(T - \mu_n I)$ is invertible