Solutions to selected exercises from §2.3

1 Conventions with transformations and matrices

A few clarifications on matrices and transformations.

1.1

If A is an $n \times m$ matrix (so, n rows and m columns) then we will reference the entries of this matrix as $A = (a^i_j)$. Here, a^i_j is the ith row and jth column of the matrix A. In particular, $i = 1, \ldots, n$ and $j = 1, \ldots m$ which I usually leave implicit when referencing the matrix.

Matrix multiplication AB is defined when A is a $n \times m$ matrix and B is a $m \times p$ matrix. The result is the $n \times p$ matrix $AB = (c^i_k)$ where the entry in the kth row and ith column is

$$c^i_k = \sum_{j=1}^{m} a^i_j b^j_k. \tag{1}$$

1.2

The book uses the notation $R(\mathsf{T})$ (range) for the image $\mathrm{Im}(\mathsf{T})$ of a linear transformation. Also, the book uses $N(\mathsf{T})$ (null-space) for the kernel $\ker \mathsf{T}$. I will continue to use the notations $\mathrm{Im}\, \mathsf{T}$ and $\ker \mathsf{T}$ in class.

Question 11

Suppose that $\mathsf{T}^2 = \mathsf{T} \circ \mathsf{T} = \mathbf{0}$ is the zero transformation and let $y \in \mathrm{Im}\, \mathsf{T}$. Then $y = \mathsf{T}(x)$ for some $x \in \mathsf{V}$. By the formula $\mathsf{T}(y) = \mathsf{T}(\mathsf{T}(x)) = 0$. Thus $y \in \ker \mathsf{T}$. We have shown that $\mathrm{Im}\, \mathsf{T} \subset \ker \mathsf{T}$.

Question 12

(a) Suppose that $\mathsf{U} \circ \mathsf{T}$ is injective (one-to-one). If $\mathsf{T}(x) = \mathsf{T}(y)$ then certainly $\mathsf{U}(\mathsf{T}(x)) = \mathsf{U}(\mathsf{T}(y))$ as well. But this is the same as $(\mathsf{U} \circ \mathsf{T})(x) = (\mathsf{U} \circ \mathsf{T})(y)$. Since $\mathsf{U} \circ \mathsf{T}$ is injective we conclude that $x = y$. Thus, T is also injective.

(b) Suppose $\mathsf{U} \circ \mathsf{T}$ is surjective (onto). This means that for any $z \in \mathsf{Z}$ we can find a $x \in X$ such that $\mathsf{U}(\mathsf{T}(x)) = z$. Thus, U is also surjective.

(c) If both U, T are injective then we will show that $\mathsf{U} \circ \mathsf{T}$ is also injective. Indeed, if $\mathsf{U}(\mathsf{T}(x)) = \mathsf{U}(\mathsf{T}(y))$ then since U is injective we conclude that $\mathsf{T}(x) = \mathsf{T}(y)$. Since T is also injective we conclude furthermore that $x = y$.

Now, suppose that U, T are both surjective. If $z \in \mathsf{Z}$ is an arbitrary vector then since U is surjective there exists $w \in \mathsf{W}$ such that $\mathsf{U}(w) = z$. Since T is surjective, we can find $v \in \mathsf{V}$ such that $\mathsf{T}(v) = \mathsf{U}(w)$. In total, we see that $(\mathsf{U} \circ \mathsf{T})(v) = z$, thus $\mathsf{U} \circ \mathsf{T}$ is surjective.

Question 13

We show that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ where A, B are $n \times n$ matrices. Let $C = AB$ have entries (c^i_k). By definition of matrix multiplication, for any $i, j = 1, \ldots, n$ we have

$$c^i_k = \sum_{j=1}^{n} a^i_j b^j_k \tag{2}$$

In particular, when $i = k$ we have

$$c^i_i = \sum_{j=1}^{n} a^i_j b^j_i. \tag{3}$$

Thus

$$\mathrm{tr}(AB) = \mathrm{tr}(C) = \sum_{i=1}^{n} c^i_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a^i_j b^j_i \tag{4}$$

Now, let $\tilde{C} = BA$ have entries $(\tilde{c}^i_k)$. The, by definition of matrix multiplication we have

$$\tilde{c}^i_i = \sum_{j=1}^{n} b^i_j a^j_i \tag{5}$$

Thus

$$\mathrm{tr}(BA) = \mathrm{tr}(\tilde{C}) = \sum_{i=1}^{n} \tilde{c}^i_i = \sum_{i=1}^{n} \sum_{j=1}^{n} b^i_j a^j_i. \tag{6}$$

We see that the two expressions above are the same after swapping (i, j).

Question 16

(a) This is a generalization of the problem on the midterm exam. First, let's recall the statement from the midterm:

Theorem 1.1. *If* $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ *is a linear transformation such that* $\mathsf{T}^2 = \mathsf{T}$ *then* $\mathsf{V} = \mathrm{Im}\,\mathsf{T} \oplus \ker \mathsf{T}$.

Proof. Let's recall the argument. If $y \in \mathrm{Im}\,\mathsf{T}$ then $y = \mathsf{T}(x)$ for some x. But then $y = \mathsf{T}(x) = \mathsf{T}(\mathsf{T}(x)) = \mathsf{T}(y)$. Thus, if $\mathsf{T}(y) = 0$ then $y = 0$. This shows that the intersection $\mathrm{Im}\,\mathsf{T} \cap \ker \mathsf{T} = \{0\}$ is trivial. To see the direct sum decomposition it suffices to show that any vector $v \in \mathsf{V}$ can be written as $v = y + x$ where $y \in \mathrm{Im}\,\mathsf{T}$ and $x \in \ker \mathsf{T}$. Let $y = \mathsf{T}(v)$ and let $x = v - \mathsf{T}(v)$. Then y is clearly in the image and $\mathsf{T}(x) = 0$ since $\mathsf{T}^2 = \mathsf{T}$. □

Let's now turn to the problem at hand. We are no longer assuming that $\mathsf{T} = \mathsf{T}^2$, but only assuming the weaker condition that $\dim \mathrm{Im}\,\mathsf{T} = \dim \mathrm{Im}\,\mathsf{T}^2$. We want the same conclusion as in the theorem above. Notice immediately we can translate this to a stronger statement about subspaces.

Lemma 1.2. *For any* $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ *one has* $\mathrm{Im}(\mathsf{T}^2) \subset \mathrm{Im}(\mathsf{T})$ *and* $\ker \mathsf{T} \subset \ker \mathsf{T}^2$.

Proof. This is immediate. If $y = \mathsf{T}(\mathsf{T}(x))$ then $y \in \mathrm{Im}\,\mathsf{T}^2$. But $y = \mathsf{T}(z)$ where $z = \mathsf{T}(x)$, so $y \in \mathrm{Im}\,\mathsf{T}$ as well. □

From this lemma, and the dimension assumption, we see that $\mathrm{Im}(\mathsf{T}) = \mathrm{Im}(\mathsf{T}^2)$, as subspaces.

Next, note that from the dimension condition it *also* follows that $\dim\ker\mathsf{T} = \dim\ker\mathsf{T}^2$. And, by the lemma this implies $\ker\mathsf{T} = \ker\mathsf{T}^2$.

Now, suppose $y = \mathsf{T}(x) \in \mathrm{Im}\,\mathsf{T}$. If $y \in \ker\mathsf{T}$ then $\mathsf{T}^2(x) = 0$. So $x \in \ker\mathsf{T}^2$. But, from what we just showed this means that $x \in \ker\mathsf{T}$. Thus $y = \mathsf{T}(x) = 0$. This shows that $\mathrm{Im}\,\mathsf{T} \cap \mathrm{Im}\,\mathsf{T}^2 = \{0\}$.

(b) We will prove that $\ker\mathsf{T}^k = \ker\mathsf{T}^{k+1}$ using mathematical induction. We have just showed the base case $k = 1$. Suppose the claim is true for k, we want to show it is true for $k+1$. We automatically know that $\ker\mathsf{T}^k \subset \ker\mathsf{T}^{k+1}$. We need to show the reverse inclusion. Suppose $\mathsf{T}^{k+1}(x) = 0$, that is $x \in \mathsf{T}^{k+1}$. Then, $\mathsf{T}^k(\mathsf{T}(x)) = 0$, so $\mathsf{T}(x) \in \ker\mathsf{T}^k$. By the inductive hypothesis $\ker\mathsf{T}^k = \ker\mathsf{T}^{k-1}$. This means that $\mathsf{T}(x) \in \ker\mathsf{T}^{k-1}$. Thus $\mathsf{T}^{k-1}(\mathsf{T}(x)) = \mathsf{T}^k(x) = 0$. Thus $x \in \mathsf{T}^k$.

By the dimension theorem it follows that $\mathrm{Im}\,\mathsf{T}^k = \mathrm{Im}\,\mathsf{T}^{k+1}$ for all k as well. The result follows.

Question 17

In this problem we determine all linear transformations $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ with the property that $\mathsf{T}^2 = \mathsf{T}$. We just showed in the midterm that

$$\mathsf{V} = \mathrm{Im}\,\mathsf{T} \oplus \ker\mathsf{T}. \tag{7}$$

With respect to this decomposition, one has

$$\mathsf{T}(y, x) = y. \tag{8}$$

In other words, T is the projection onto the subspace $\mathrm{Im}\,\mathsf{T}$.

We have shown the following.

Theorem 1.3. *There is a bijective correspondence between linear transformations* $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ *such that* $\mathsf{T}^2 = \mathsf{T}$ *and subspaces* $\mathsf{W} \subset \mathsf{V}$. *The linear transformation corresponding to a subspace* W *is* $\mathsf{T}_\mathsf{W}\colon \mathsf{V} \to \mathsf{V}$ *defined by*

$$\mathsf{T}_\mathsf{W}(v) = \begin{cases} v & v \in \mathsf{W}, \\ 0 & \textit{else}. \end{cases} \tag{9}$$

Notice that this function always satisfies $\mathsf{T}_\mathsf{W}^2 = \mathsf{T}_\mathsf{W}$.