Homework Notes — Week 3

Math 24 — Spring 2014

§2.1:5 To see that $T : \mathsf{P}_2(\mathbb{R}) \to \mathsf{P}_3(\mathbb{R})$ is linear, it suffices to check that

$$T(af(x) + bg(x)) = aT(f(x)) + bT(g(x))$$

for any $f(x), g(x) \in \mathsf{P}_2(\mathbb{R})$ and any real numbers a, b. From familiar properties of polynomials from algebra and calculus, we see that this is indeed the case:

$$\begin{aligned} T(af(x) + bg(x)) &= x(af(x) + bg(x)) + \tfrac{d}{dx}[af(x) + bg(x)] \\ &= (xaf(x) + xbg(x)) + (af'(x) + bg'(x)) \\ &= (xaf(x) + af'(x)) + (xbg(x) + bg'(x)) \\ &= a(xf(x) + f'(x)) + b(xg(x) + g'(x)) = aT(f(x)) + bT(g(x)). \end{aligned}$$

To find a basis for $\mathsf{N}(T)$, we need to look at the equation $T(f(x)) = 0$, or $f'(x) = -xf(x)$. Unless $f(x)$ is the zero polynomial, the degree of $f'(x)$ is one less than that of $f(x)$ and the degree of $-xf(x)$ is one more than that of $f(x)$. Therefore, the only case where this equality can hold is when $f(x)$ is the zero polynomial. Therefore $\mathsf{N}(T) = \{0\}$, and a basis for this subspace is $\varnothing$.

To find a basis for $\mathsf{R}(T)$, we can use the proof of the Dimension Theorem to guide us. Since $\{1, x, x^2\}$ is a basis for $\mathsf{P}_2(\mathbb{R})$ that extends our basis for $\mathsf{N}(T)$, $\{T(1), T(x), T(x^2)\}$ must form a basis for $\mathsf{R}(T)$. (See the Claim from the proof of the Dimension Theorem in the April 4 slides.) Therefore, $\{x, x^2 + 1, x^2 + 2x\}$ is a basis for $\mathsf{R}(T)$.

So $\mathrm{nullity}(T) = 0$, $\mathrm{rank}(T) = 3$, and since $\dim(\mathsf{P}_2(\mathbb{R})) = 3$, the Dimension Theorem is indeed true.

§2.1:14* To organize ideas, it is convenient to break the first two parts into two theorems.

One direction of part (a) corresponds to the following result.

Theorem A1. *Suppose $T : \mathsf{V} \to \mathsf{W}$ is a one-to-one linear transformation. If $S \subseteq \mathsf{V}$ is linearly independent then*

$$T(S) = \{T(x) \in \mathsf{W} : x \in \mathsf{V}\}$$

is linearly independent.

Proof. Suppose $S \subseteq \mathsf{V}$ is linearly independent and suppose

$$a_1y_1 + a_2y_2 + \cdots + a_ny_n = 0$$

where $y_1, y_2, \ldots, y_n$ are distinct elements of $T(S)$. By definition of $T(S)$, we can find elements $x_1, x_2, \ldots, x_n$ of S such that

$$y_1 = T(x_1), y_2 = T(x_2), \ldots, y_n = T(x_n).$$

Note that the elements $x_1, x_2, \ldots, x_n$ are necessarily distinct since $x_i = x_j$ implies $y_i = T(x_i) = T(x_j) = y_j$ and we know the $y_1, y_2, \ldots, y_n$ are distinct.

Because

$$\begin{aligned} 0 &= a_1y_1 + a_2y_2 + \cdots + a_ny_n \\ &= a_1T(x_1) + a_2T(x_2) + \cdots + a_nT(x_n) \\ &= T(a_1x_1 + a_2x_2 + \cdots + a_nx_n), \end{aligned}$$

we see that $a_1x_1 + a_2x_2 + \cdots + a_nx_n \in \mathsf{N}(T)$. Since T is one-to-one, we know from Theorem 2.4 that $\mathsf{N}(T) = \{0\}$ and therefore that

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0.$$

Since S is linearly independent by hypothesis, we conclude that $a_1 = a_2 = \cdots = a_n = 0$.

Because $a_1y_1+a_2y_2+\cdots+a_ny_n$ was, at the outset, an arbitray linear combination of elements of $T(S)$, we conclude that $T(S)$ is linearly independent. $\square$

For the converse of part (a), we only need to consider when S has only one element.

Theorem A2. *Suppose $T : \mathsf{V} \to \mathsf{W}$ is a linear transformation. If for any linearly independent one element set $\{x\} \subseteq \mathsf{V}$, the set $\{T(x)\}$ is also linearly independent, then T is one-to-one.*

Proof. First note that, by Theorem 2.4, T is one-to-one exactly when $\mathsf{N}(T) = \{0\}$ or, equivalently, when $\mathsf{N}(T) \subseteq \{0\}$ since we always have $\{0\} \subseteq \mathsf{N}(T)$.

To say that $\{x\}$ is linearly independent simply means that x is nonzero. Similarly, to say that $\{T(x)\}$ is linearly independent simply means that $T(x)$ is nonzero. So the statement of Theorem A2 simply says that if x is nonzero then $T(x)$ is nonzero too. Looking at the contrapositive, this is equivalent to saying that if $T(x) = 0$ then $x = 0$ or, in other words, that $\mathsf{N}(T) \subseteq \{0\}$. Therefore, for any linearly independent one element set $\{x\} \subseteq \mathsf{V}$, the set $\{T(x)\}$ is also linearly independent, then T is one-to-one. $\square$

For part (b), it may seem at first that this is the same as in part (a), but the two statements are actually quite different. We first deal with the case where the set S is finite.

Theorem B1. *Suppose $T : \mathsf{V} \to \mathsf{W}$ is a one-to-one linear transformation. A set $\{x_1, x_2, \ldots, x_n\} \subseteq \mathsf{V}$ is linearly independent if and only if $\{T(x_1), T(x_2), \ldots, T(x_n)\} \subseteq \mathsf{W}$ is linearly independent.*

Proof. The key fact is that if $a_1, a_2, \ldots, a_n$ are any scalars then

$$T(a_1x_1 + a_2x_2 + \cdots + a_nx_n) = a_1T(x_1) + a_2T(x_2) + \cdots + a_nT(x_n).$$

For the forward implication, we see that if

$$a_1T(x_1) + a_2T(x_2) + \cdots + a_nT(x_n) = 0$$

then

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \in \mathsf{N}(T).$$

By Theorem 2.4, $\mathsf{N}(T) = \{0\}$ since T is one-to-one. So,

$$a_1T(x_1) + a_2T(x_2) + \cdots + a_nT(x_n) = 0$$

implies

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0.$$

Therefore, if $\{x_1, x_2, \ldots, x_n\}$ is linearly independent, we must then have $a_1 = a_2 = \cdots = a_n = 0$. Since this is true for any scalars $a_1, a_2, \ldots, a_n$, we conclude that if $\{x_1, x_2, \ldots, x_n\}$ is linearly independent then so is $\{T(x_1), T(x_2), \ldots, T(x_n)\}$.

For the backward implication, we see that if

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

then

$$a_1T(x_1) + a_2T(x_2) + \cdots + a_nT(x_n) = 0.$$

Therefore, if $\{T(x_1), T(x_2), \ldots, T(x_n)\}$ is linearly independent, we must then have $a_1 = a_2 = \cdots = a_n = 0$. Since this is true for any scalars $a_1, a_2, \ldots, a_n$, we conclude that if $\{T(x_1), T(x_2), \ldots, T(x_n)\}$ is linearly independent then so is $\{x_1, x_2, \ldots, x_n\}$. □

For the general case, note that a (possibly infinite) set of vectors is linearly independent if and only if every finite subset of it is linearly independent. This is because a linear dependency only involves finitely many vectors from a given set. So if the finite set of vectors involved in a given linear dependency is linearly independent, then that linear dependency must be the trivial one.

Theorem B2. *Suppose $T : \mathsf{V} \to \mathsf{W}$ is a one-to-one linear transformation. A set $S \subseteq \mathsf{V}$ is linearly independent if and only if $T(S) \subseteq \mathsf{W}$ is linearly independent.*

Proof. By Theorem B1, a finite set $\{x_1, x_2, \ldots, x_n\} \subseteq S$ is linearly independent if and only if the corresponding finite set $\{T(x_1), T(x_2), \ldots, T(x_n)\} \subseteq T(S)$ is linearly independent. Therefore, every finite subset of S is linearly independent if and only if every finite subset of $T(S)$ is linearly independent. By the observation above, we conclude that S is linearly independent if and only if $T(S)$ is linearly independent. □

Theorem C. *Suppose $T : \mathsf{V} \to \mathsf{W}$ is a one-to-one and onto linear transformation. Then $\{v_1, v_2, \ldots, v_n\}$ is a basis for V if and only if $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for W.*

Proof. By Theorem B1, we know that $\{v_1, v_2, \ldots, v_n\}$ is linearly independent if and only if $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is linearly independent. So it suffices to show that (1) if $\{v_1, v_2, \ldots, v_n\}$ is a basis for V then $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ generates W, and that (2) if $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for W then $\{v_1, v_2, \ldots, v_n\}$ generates V.

Suppose first that $\{v_1, v_2, \ldots, v_n\}$ is a basis for V. By Theorem 2.2, $\mathsf{R}(T) = \operatorname{span}\{T(v_1), T(v_2), \ldots, T(v_n)\}$. Since T is onto, we also have $\mathsf{R}(T) = \mathsf{W}$. Therefore, we conclude that $\operatorname{span}\{T(v_1), T(v_2), \ldots, T(v_n)\}$ generates W.

Suppose next that $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for W. Given any $x \in \mathsf{V}$, we can find scalars $a_1, a_2, \ldots, a_n$ such that

$$T(x) = a_1 T(v_1) + a_2 T(v_2) + \cdots + a_n T(v_n) = T(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n).$$

Since T is one-to-one, we conclude that

$$x = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$$

and therefore $x \in \operatorname{span}\{v_1, v_2, \ldots, v_n\}$. Since x was an arbitrary vector in V, we conclude that $\{v_1, v_2, \ldots, v_n\}$ generates V. □

§2.1:21 To see that T is onto, given a target sequence $(b_1, b_2, \dots) \in \mathsf{V}$, note that the sequence

$$(a_1, a_2, \dots) = U(b_1, b_2, \dots) = (0, b_1, b_2, \dots)$$

has the property that

$$T(a_1, a_2, \dots) = (a_2, a_3, \dots) = (b_1, b_2, \dots).$$

To see that U is one-to-one, suppose $(a_1, a_2, \dots), (b_1, b_2, \dots) \in \mathsf{V}$ are such that

$$(0, a_1, a_2, \dots) = U(a_1, a_2, \dots) = U(b_1, b_2, \dots) = (0, b_1, b_2, \dots).$$

This holds precisely when

$$0 = 0, \quad a_1 = b_1, \quad a_2 = b_2, \quad \dots$$

Since $0 = 0$ is simply true, this means $(a_1, a_2, \dots) = (b_1, b_2, \dots)$. We have thus shown that

$$U(a_1, a_2, \dots) = U(b_1, b_2, \dots) \quad \text{implies} \quad (a_1, a_2, \dots) = (b_1, b_2, \dots),$$

i.e., that U is one-to-one.

Also note that since $TU = I_{\mathsf{V}}$, the fact that T is onto and the fact that U is one-to-one follow from the Theorem on 'One-to-One, Onto, and Inverses' from the April 16 slides.

§2.2:5bde For (b), since

$$T(1) = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \quad T(x) = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \quad T(x^2) = \begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix},$$

we see that

$$[T]_\beta^\alpha = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

For (d), since

$$T(1) = 1, \qquad T(x) = 2, \qquad T(x^2) = 4,$$

we see that

$$T_\beta^\gamma = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix}.$$

For (e), since

$$A = (1)E^{11} + (-2)E^{12} + (0)E^{21} + (4)E^{22},$$

we see that

$$[A]_\alpha = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 4 \end{pmatrix}.$$

§2.2:8* Theorem. *Let* V *be an* n*-dimensional vector space with an ordered basis* $\beta = \{v_1, v_2, \ldots, v_n\}$. *The function* $T : \mathsf{V} \to F^n$ *defined by* $T(x) = [x]_\beta$ *is a linear transformation.*

Proof. We need to check that $T(x+y) = T(x) + T(y)$ and $T(cx) = cT(x)$ for all $x, y \in \mathsf{V}$ and all scalars c.

Given $x, y \in \mathsf{V}$, write

$$T(x) = (a_1, a_2, \ldots, a_n) \quad \text{and} \quad T(y) = (b_1, b_2, \ldots, b_n),$$

so that

$$x = a_1v_1 + a_2v_2 + \cdots + a_nv_n$$

and

$$y = b_1v_1 + b_2v_2 + \cdots + b_nv_n.$$

Since

$$\begin{aligned} x + y &= (a_1v_1 + a_2v_2 + \cdots + a_nv_n) + (b_1v_1 + b_2v_2 + \cdots + b_nv_n) \\ &= (a_1 + b_1)v_1 + (a_2 + b_2)v_2 + \cdots + (a_n + b_n)v_n, \end{aligned}$$

we see that

$$T(x+y) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n) = T(x) + T(y).$$

Similarly, given any scalar c, we have

$$cx = (ca_1)v_1 + (ca_2)v_2 + \cdots + (ca_n)v_n$$

and so

$$T(cx) = (ca_1, ca_2, \ldots, ca_n) = cT(x).$$

Therefore, $T : \mathsf{V} \to F^n$ is indeed a linear transformation. □

§2.2:10 Since
$$(2,3) = -(0,1) + 3(1,1),$$
we see that
$$T(2,3) = -T(0,1) + 3T(1,1) = (-1,-4) + (6,15) = (5,11).$$

To see wether T is one-to-one, we can compute $\mathsf{N}(T)$ and use Theorem 2.4. First, note that
$$\begin{aligned} T(a,b) &= T((a-b)(1,0) + b(1,1)) \\ &= (a-b)(1,4) + b(2,5) = (a+b, 4a+b). \end{aligned}$$
Therefore $T(a,b) = (0,0)$ exactly when $a+b=0$ and $4a+b=0$. The only solution to these equations is when $a=b=0$, and thus $\mathsf{N}(T) = \{(0,0)\}$. It follows from Theorem 2.4 that T is one-to-one.

§2.3:3 (a) Since
$$U(1) = (1,0,1), \quad U(x) = (1,0,-1), \quad U(x^2) = (0,1,0),$$
we see that
$$[U]_\beta^\gamma = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Since
$$\begin{aligned} T(1) &= (0)(3+x) + 2(1) = 2, \\ T(x) &= (1)(3+x) + 2(x) = 3x+3, \\ T(x^2) &= (2x)(3+x) + 2(x^2) = 4x^2 + 6x, \end{aligned}$$
we see that
$$[T]_\beta = [T]_\beta^\beta = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 3 & 6 \\ 0 & 0 & 4 \end{pmatrix}.$$

Since
$$UT(1) = (2,0,2), \quad UT(x) = (6,0,0), \quad UT(x^2) = (6,4,-6),$$
we see that
$$[UT]_\beta^\gamma = \begin{pmatrix} 2 & 6 & 6 \\ 0 & 0 & 4 \\ 2 & 0 & -6 \end{pmatrix}.$$

Indeed, we can check that

$$\begin{pmatrix} 2 & 6 & 6 \\ 0 & 0 & 4 \\ 2 & 0 & -6 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 3 & 0 \\ 0 & 3 & 6 \\ 0 & 0 & 4 \end{pmatrix}.$$

(b) We have $[h(x)]_\beta = (1, -2, 1)$ and $[U(h(x))]_\gamma = U(h(x)) = (1, 1, 3)$. Indeed, we can check

$$\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

§2.3:4bd For (b), we have from above that

$$[T]_\beta^\alpha = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Therefore, by Theorem 2.14,

$$\begin{aligned} [T(4 - 6x + 3x^2)]_\alpha &= [T]_\beta^\alpha [4 - 6x + 3x^2]_\beta \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ -6 \\ 3 \end{pmatrix} = \begin{pmatrix} -6 \\ 2 \\ 0 \\ 6 \end{pmatrix}. \end{aligned}$$

For (d), we have from above that

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix}.$$

Therefore, by Theorem 2.14,

$$\begin{aligned} [T(6 - x + 2x^2)]_\gamma &= [T]_\beta^\gamma [6 - x + 2x^2]_\beta \\ &= \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \end{pmatrix}. \end{aligned}$$

§2.3:11* **Theorem.** *If* V *is a vector space and* $T : \mathsf{V} \to \mathsf{V}$ *is a linear transformation, then* $T^2 = T_0$ *if and only if* $\mathsf{R}(T) \subseteq \mathsf{N}(T)$.

Proof. First, we will show that if $\mathsf{R}(T) \subseteq \mathsf{N}(T)$ then $T^2 = T_0$. Given an arbitrary $x \in \mathsf{V}$, we have $T(x) \in \mathsf{R}(T)$ by definition. Since $\mathsf{R}(T) \subseteq \mathsf{N}(T)$, we also have $T(x) \in \mathsf{N}(T)$. By definition, this means that $T(T(x)) = 0$, or $T^2(x) = 0$ for short. We therefore conclude that $T^2(x) = 0$ for every $x \in \mathsf{V}$, which means that $T^2 = T_0$.

Next, we will show that if $T^2 = T_0$ then $\mathsf{R}(T) \subseteq \mathsf{N}(T)$. By definition, $y \in \mathsf{R}(T)$ means that $y = T(x)$ for some $x \in \mathsf{V}$. Since $T^2 = T_0$, we see that for any such y, we have

$$T(y) = T(T(x)) = T_0(x) = 0,$$

and hence $y \in \mathsf{N}(T)$. We therefore conclude that every element of $\mathsf{R}(T)$ is an element of $\mathsf{N}(T)$, equivalently that $\mathsf{R}(T) \subseteq \mathsf{N}(T)$. $\square$