Linear Transformations

Basic definitions and properties. We begin by recalling some familiar definitions.

- A *function* (or *map*, or *transformation*) F from a set $\mathcal{X}$ to a set $\mathcal{Y}$ (denoted $F : \mathcal{X} \to \mathcal{Y}$) is an assignment to each element $x \in \mathcal{X}$ a unique element $F(x) \in \mathcal{Y}$.
- A function $F : \mathcal{X} \to \mathcal{Y}$ is *one-to-one* if, for each $y \in \mathcal{Y}$, there is *at most* one $x \in \mathcal{X}$ such that $y = F(x)$. Another way of saying that F is one-to-one is that, for x and y in $\mathcal{X}$, $F(x) = F(y)$ only if $x = y$.
- A function $F : \mathcal{X} \to \mathcal{Y}$ is *onto* if, for every $y \in \mathcal{Y}$, there is *at least* one $x \in \mathcal{X}$ such that $y = F(x)$.
- If $F : \mathcal{X} \to \mathcal{Y}$ is both one-to-one *and* onto, then F has an *inverse* function $F^{-1} : \mathcal{Y} \to \mathcal{X}$, defined as follows: Suppose $y \in \mathcal{Y}$ is given. Since F is onto, there is at least one $x \in \mathcal{X}$ such that $F(x) = y$. Since F is also one-to-one, this x is unique, i.e., the only element of $\mathcal{X}$ that F maps onto y. Then we define $x = F^{-1}(y)$.

Now suppose $\mathcal{V}$ and $\mathcal{W}$ are vector spaces. We make several basic definitions

Definition 1. A map $T : \mathcal{V} \to \mathcal{W}$ is a *linear transformation* if $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ for all x and y in $\mathcal{V}$ and all scalars α and β.

Definition 2. The *nullspace* of a linear transformation $T : \mathcal{V} \to \mathcal{W}$, denoted $\mathcal{N}(T)$, is the set of all $x \in \mathcal{V}$ such that $T(x) = 0$.

Note that $\mathcal{N}(T)$ is a subspace of $\mathcal{V}$. Note also that if T is linear, then $T(0) = 0$. Consequently, $0 \in \mathcal{N}(T)$ and $\{0\}$ (the subspace consisting only of 0) is a subspace of $\mathcal{N}(T)$ for every linear transformation T.

Definition 3. The *range* of a linear transformation $T : \mathcal{V} \to \mathcal{W}$, denoted $\mathcal{R}(T)$, is the set of all $w \in \mathcal{W}$ such that $w = T(x)$ for some $x \in \mathcal{V}$.

Note that $\mathcal{R}(T)$ is a subspace of $\mathcal{W}$.

Lemma 4. Suppose that $T : \mathcal{V} \to \mathcal{W}$ is a linear transformation. Then T is one-to-one if and only if $\mathcal{N}(T) = \{0\}$, and T is onto if and only if $\mathcal{R}(T) = \mathcal{W}$.

Proof. It is immediate from the definitions that T is onto if and only if $\mathcal{R}(T) = \mathcal{W}$, so we only prove that T is one-to-one if and only if $\mathcal{N}(T) = \{0\}$. Suppose that T is one-to-one. Then, since $T(0) = 0$, we have that $T(x) = 0$ only if $x = 0$. It follows that $\mathcal{N}(T) = \{0\}$. To show the converse, suppose that $\mathcal{N}(T) = \{0\}$. If $T(x) = T(y)$ for x and y in $\mathcal{V}$, then $0 = T(x) - T(y) = T(x - y)$. It follows that $x - y \in \mathcal{N}(T)$ and, therefore, that $x - y = 0$, i.e., $x = y$. □

Linear transformations, linear independence, spanning sets and bases. Suppose that $\mathcal{V}$ and $\mathcal{W}$ are vector spaces and that $T : \mathcal{V} \to \mathcal{W}$ is linear.

LEMMA 5. If T is one-to-one and v_1, ..., v_k are linearly independent in $\mathcal{V}$, then $T(v_1)$, ..., $T(v_k)$ are linearly independent in $\mathcal{W}$.

Proof. Assume that T is one-to-one and v_1, ..., v_k are linearly independent in $\mathcal{V}$. Suppose that $\alpha_1 T(v_1) + \ldots + \alpha_k T(v_k) = 0$. Since T is linear, this implies that $T(\alpha_1 v_1 + \ldots + \alpha_k v_k) = 0$. Since T is one-to-one, it follows that $\alpha_1 v_1 + \ldots + \alpha_k v_k = 0$. Since v_1, ..., v_k are linearly independent, this implies $\alpha_1 = \ldots = \alpha_k = 0$, and we conclude that $T(v_1)$, ..., $T(v_k)$ are linearly independent in $\mathcal{W}$. $\square$

LEMMA 6. If v_1, ..., v_k span $\mathcal{V}$, then $T(v_1)$, ..., $T(v_k)$ span $\mathcal{R}(T)$.

Proof. Suppose that $w \in \mathcal{R}(T)$. Then there is some $v \in \mathcal{V}$ such that $T(v) = w$. If v_1, ..., v_k span $\mathcal{V}$, then we can write $v = \alpha_1 v_1 + \ldots + \alpha_k v_k$ for scalars α_1, ..., α_k. Since T is linear, it follows that

$$w = T(v) = T(\alpha_1 v_1 + \ldots + \alpha_k v_k) = \alpha_1 T(v_1) + \ldots + \alpha_k T(v_k),$$

and we conclude that $T(v_1)$, ..., $T(v_k)$ span $\mathcal{R}(T)$. $\square$

LEMMA 7. If $\{v_1, \ldots, v_k\}$ is a basis of $\mathcal{V}$ and T is one-to-one, then $\{T(v_1), \ldots, T(v_k)\}$ is a basis of $\mathcal{R}(T)$.

Proof. If $\{v_1, \ldots, v_k\}$ is a basis of $\mathcal{V}$ and T is one-to-one, then it follows from Lemma 5 that $T(v_1)$, ..., $T(v_k)$ are linearly independent and from Lemma 6 that they span $\mathcal{R}(T)$. Thus $\{T(v_1), \ldots, T(v_k)\}$ is a basis of $\mathcal{R}(T)$. $\square$

LEMMA 8. We have that $\dim \mathcal{R}(T) \leq \dim \mathcal{V}$, and $\dim \mathcal{R}(T) = \dim \mathcal{V}$ if and only if T is one-to-one.

Proof. Suppose that $\dim \mathcal{V} = k$, and let $\{v_1, \ldots, v_k\}$ be a basis for $\mathcal{V}$. By Lemma 6, $T(v_1)$, ..., $T(v_k)$ span $\mathcal{R}(T)$. Then a subset of $\{T(v_1), \ldots, T(v_k)\}$ is a basis for $\mathcal{R}(T)$, and it follows that $\dim \mathcal{R}(T) \leq k$. To complete the proof, note that $\dim \mathcal{R}(T) = k$ if and only if $T(v_1)$, ..., $T(v_k)$ are linearly independent. If T is one-to-one, then $T(v_1)$, ..., $T(v_k)$ are linearly independent by Lemma 5. Conversely, if T is not one-to-one, then $Tv = 0$ for some non-zero $v \in \mathcal{V}$. Writing $v = \alpha_1 v_1 + \ldots + \alpha_k v_k$, we then have $0 = T(\alpha_1 v_1 + \ldots + \alpha_k v_k) = \alpha_1 T(v_1) + \ldots + \alpha_k T(v_k)$. Since not all of the α_i's are zero, we conclude that $T(v_1)$, ..., $T(v_k)$ are linearly dependent. $\square$

LEMMA 9. If $\dim \mathcal{V} = \dim \mathcal{W}$, then T is one-to-one if and only if it is onto.

Proof. Suppose that $\dim \mathcal{V} = \dim \mathcal{W}$. Then by Lemma 8, we have that T is one-to-one if and only if $\dim \mathcal{R}(t) = \dim \mathcal{V} = \dim \mathcal{W}$, which holds if and only if $\mathcal{R}(T) = \mathcal{W}$, i.e., T is onto. $\square$

Note that if T is one-to-one, then we can define an inverse map $T^{-1} : \mathcal{R}(T) \to \mathcal{V}$ in the usual way, i.e., for each $w \in \mathcal{R}(T)$, we define $T^{-1}(w)$ to be the unique $v \in \mathcal{V}$ such that $T(v) = w$. In the particular case when T is one-to-one and $\dim \mathcal{V} = \dim \mathcal{W}$, we have $T^{-1} : \mathcal{W} \to \mathcal{V}$, i.e., T^{-1} is defined on all of $\mathcal{W}$.

Examples.

EXAMPLE 10. Let $\mathcal{P}_2$ denote the vector space of all polynomials of degree less than or equal to two. Then there is a natural map $T : \mathcal{P}_2 \to I\!R^3$ defined by

$$p(x) = a_0 + a_1 x + a_2 x^2 \in \mathcal{P}_2 \longrightarrow T(p) = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \in I\!R^3.$$

It is easy to show that T is linear, one-to-one, and onto.

EXAMPLE 11. Suppose that $A \in I\!R^{m\times n}$. Then A naturally defines a map $T : I\!R^n \to I\!R^m$ by setting $T(x) = Ax$ for $x \in I\!R^n$. It is easy to confirm that T is linear. The following are easily verified using the above lemmas and things we already know about $N(A)$, $C(A)$, and $\operatorname{rank} A$:

- $\mathcal{N}(T) = N(A)$, and T is one-to-one $\iff N(A) = \{0\} \iff \operatorname{rank} A = n$.
- $\mathcal{R}(T) = C(A)$, and T is onto $\iff C(A) = I\!R^m \iff \operatorname{rank} A = m$.

It is interesting to examine some cases and note the implications for the existence and uniqueness of solutions of $Ax = b$ for $b \in I\!R^m$.

Case 1: $m > n$. In this case, T can't be onto, since $\operatorname{rank} A \le n < m$. As noted above, T is one-to-one $\iff \operatorname{rank} A = n$. Then $Ax = b$ does not have a solution for some $b \in I\!R^m$; if a solution exists, then it is unique $\iff N(A) = \{0\} \iff \operatorname{rank} A = n$.

Case 2: $m < n$. In this case, T can't be one-to-one, since $\dim N(A) = n - \operatorname{rank} A > m - \operatorname{rank} A \ge 0$. As noted above, T is onto $\iff \operatorname{rank} A = m$. Then a solution of $Ax = b$ exists for all $b \in I\!R^n \iff \operatorname{rank} A = m$; if a solution exists, then it cannot be unique.

Case 3: $m = n$. In this case, Lemma 9 implies that T is one-to-one if and only if it is onto. A logically equivalent restatement is that T is one-to-one if and only if it is one-to-one *and* onto. Thus we conclude that the following are equivalent:

- $Ax = 0 \iff x = 0$.
- $Ax = b$ has a unique solution for every $b \in I\!R^n$.

These conclusions have, more or less, already been reached by considering the echelon form of A. It is satisfying, though, to have been able to obtain them using little more than the framework and properties of linear operators on finite-dimensional vector spaces. It is especially satisfying that *we obtained the equivalence in Case 3 ($m = n$) using only the general result in Lemma 9.*

There's more: Suppose we are in Case 3 ($m = n$) and that T is one-to-one and onto, i.e., that the equivalent conditions above hold. On the one hand, we know that T has an inverse map $T^{-1} : I\!R^n \to I\!R^n$

defined as follows: For each $y \in I\!R^n$, we set $T^{-1}(y) = x$, where x is the unique vector in $I\!R^n$ such that $T(x) = y$. On the other hand, since the two equivalent conditions hold, we know that A is nonsingular and therefore has an inverse matrix A^{-1} for which $A^{-1}A = AA^{-1} = I$ (the identity matrix). Then

$$T^{-1}(y) = x \iff y = T(x) \iff y = Ax \iff A^{-1}y = A^{-1}Ax = Ix = x,$$

and so $T^{-1}(y) = A^{-1}y$ for each $y \in I\!R^n$.

EXAMPLE 12. Let $\mathcal{P}_2$ again denote the vector space of all polynomials of degree less than or equal to two, and define a transformation T on $\mathcal{P}_2$ by $T(p) = dp/dx$, i.e., if $p(x) = a_0 + a_1x + a_2x^2 \in \mathcal{P}_2$, then $T(p)(x) = a_1 + 2a_2x$. We can regard this as a transformation from $\mathcal{P}_2$ to either $\mathcal{P}_2$ or $\mathcal{P}_1$, the space of all polynomials of degree less than or equal to one. In the second case, T is onto; in the first case, it is not. In either case $\mathcal{N}(T) = \{p \in \mathcal{P}_2 : \ p(x) = a_0\}$ (the polynomials of degree zero), a one-dimensional subspace of $\mathcal{P}_2$, so T is not one-to-one in either case.

Before introducing the next example, we need the following definition.

DEFINITION 13. The *span* of vectors $v_1, \ldots, v_k$, denoted by $\mathrm{span}\{v_1, \ldots, v_k\}$, is the set of all linear combinations $\alpha_1 v_1 + \ldots + \alpha_k v_k$ for scalars $\alpha_1, \ldots, \alpha_k$.

It is easy to verify that $\mathrm{span}\{v_1, \ldots, v_k\}$ is a subspace of the vector space in which $v_1, \ldots, v_k$ reside. Moreover, $\dim\,(\mathrm{span}\{v_1, \ldots, v_k\}) \le k$, and $\dim\,(\mathrm{span}\{v_1, \ldots, v_k\}) = k$ if and only if $v_1, \ldots, v_k$ are linearly independent.

EXAMPLE 14. Take $\mathcal{V} = \mathrm{span}\{\cos, \sin\}$, i.e., the space of all functions f such that, for some scalars α and β, $f(x) = \alpha\cos(x) + \beta\sin(x)$ for all x. We can show that $\cos$ and $\sin$ are linearly independent, as follows: If $\alpha\cos + \beta\sin = 0$, i.e., $\alpha\cos(x) + \beta\sin(x) = 0$ for all x, then in particular $0 = \alpha\cos(0) + \beta\sin(0) = \alpha$ and $0 = \alpha\cos(\pi/2) + \beta\sin(\pi/2) = \beta$. Thus $\alpha\cos + \beta\sin = 0$ only if $\alpha = \beta = 0$, and therefore $\cos$ and $\sin$ are linearly independent. It follows from the observations above that $\dim\mathcal{V} = 2$.

Note that all functions in $\mathcal{V}$ are differentiable since $\cos$ and $\sin$ are differentiable. Moreover, if $f = \alpha\cos + \beta\sin \in \mathcal{V}$, then $df/dx = -\alpha\sin + \beta\cos \in \mathcal{V}$. Then we can define a map $T : \mathcal{V} \to \mathcal{V}$ by $T(f) = df/dx$ for $f \in \mathcal{V}$. It is easy to verify that T is linear.

We can also show that T is one-to-one, as follows: If $f = \alpha\cos + \beta\sin \in \mathcal{V}$ is such that $T(f) = 0$, then $T(f)(x) = df/dx(x) = -\alpha\sin(x) + \beta\cos(x) = 0$ for all x. Since $\cos$ and $\sin$ are linearly independent, it follows that $\alpha = \beta = 0$ and thus $f = 0$.

Since T is one-to-one, it follows from Lemma 9 (with $\mathcal{W} = \mathcal{V}$) that T is also onto. Then, since T is one-to-one and onto, it has an inverse map T^{-1}. This allows us to make a nice statement about the existence and uniqueness of solutions of a differential equation: *For each g in $\mathcal{V}$, there is a unique f in $\mathcal{V}$ such that $df/dx = g$.* The underlining is intended to emphasize that the solution is unique only within $\mathcal{V}$. More generally, if $f \in \mathcal{V}$ is a solution, then so is $f + C$ for any constant C. However, $f + C$ is in $\mathcal{V}$ if and only if $C = 0$.