Chapter 3

Linear Maps

3.A The Vector Space of Linear Maps

9/5: • Linear map (from V to W): A function $T:V\to W$ with the following properties. Also known as linear transformation.

additivity

T(u+v) = Tu + Tv for all $u, v \in V$.

homogeneity

 $T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbb{F}$ and all $v \in V$.

- Note that for linear maps, Tv means the same as the more standard functional notation T(v).
- $\mathcal{L}(V, W)$: The set of all linear maps from V to W.
- **Zero map**: The function $0 \in \mathcal{L}(V, W)$ that takes each element of some vector space to the additive identity of another vector space. Defined by

$$0v = 0$$

• Identity map: The function $I \in \mathcal{L}(V, V)$ on some vector space that takes each element to itself. Defined by

$$Iv = v$$

- We can also formalize more esoteric linear processes as linear maps.
 - For example, $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R}))$ can be thought of as the differentiation map Dp = p'. This formalizes the fact that (f+g)' = f' + g' and $(\lambda f)' = \lambda f'$.
 - We can do the same with integration: Let $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathbb{R})$ be described by $Tp = \int_0^1 p(x) dx$. This formalizes the fact that integrals are additive and homogeneous.
 - Axler (2015) gives a number more examples.
- We now prove that there exists a unique linear map from a vector space of dimension n to any n vectors in another vector space.

Theorem 3.1. Suppose v_1, \ldots, v_n is a basis of V and $w_1, \ldots, w_n \in W$. Then there exists a unique linear map $T: V \to W$ such that $Tv_j = w_j$ for each $j = 1, \ldots, n$.

Proof. First, we define a function $T: V \to W$. We then show that T satisfies the specified property. After that, we show that it is a linear map. Lastly, we show that it is unique. Let's begin.

Let $T: V \to W$ be defined by

$$T(c_1v_1 + \dots + c_nv_n) = c_1w_1 + \dots + c_nw_n$$

for all $c_1v_1 + \cdots + c_nv_n \in V$. Note that this definition is valid since, by Theorem 2.5, each $v \in V$ can be written in the form $c_1v_1 + \cdots + c_nv_n$ where $c_1, \ldots, c_n \in \mathbb{F}$.

To prove that $Tv_j = w_j$ for all j = 1, ..., n, let each c_i in the above definition equal 0 save c_j , which we set equal to 1. Then we have

$$T(0v_1 + \dots + 0v_{j-1} + 1v_j + 0v_{j+1} + \dots + 0v_n) = 0w_1 + \dots + 0w_{j-1} + 1w_j + 0w_{j+1} + \dots + 0w_n$$
$$T(v_j) = w_j$$

as desired.

To prove that T is a linear map, it will suffice to verify additivity and homogeneity, which we may do as follows. Let $u, v \in V$ with $u = a_1v_1 + \cdots + a_nv_n$ and $v = c_1v_1 + \cdots + c_nv_n$, and let $\lambda \in \mathbb{F}$. Then

$$T(u+v) = T((a_1 + c_1)v_1 + \dots + (a_n + c_n)v_n)$$

= $(a_1 + c_1)w_1 + \dots + (a_n + c_n)w_n$
= $Tu + Tv$

and

$$T(\lambda v) = T(\lambda c_1 v_1 + \dots + \lambda c_n v_n)$$

= $\lambda c_1 w_1 + \dots + \lambda c_n w_n$
= $\lambda T v$

as desired.

Now suppose $\tilde{T} \in \mathcal{L}(V,W)$ satisfies $\tilde{T}v_j = w_j$ for all $j = 1, \ldots, n$. To prove that $T = \tilde{T}$, it will suffice to show that $\tilde{T}(c_1v_1 + \cdots + c_nv_n) = T(c_1v_1 + \cdots + c_nv_n)$ for all $c_1v_1 + \cdots + c_nv_n \in V$. Let $c_1v_1 + \cdots + c_nv_n \in V$ be arbitrary. We know that $\tilde{T}(v_j) = w_j$ for all $j = 1, \ldots, n$. It follows since \tilde{T} is a linear map (specifically, since it's homogenous) that $c_jw_j = c_j\tilde{T}(v_j) = \tilde{T}(c_jv_j)$ for all $j = 1, \ldots, n$. Similarly, the additivity of \tilde{T} implies that

$$T(c_1v_1 + \dots + c_nv_n) = c_1w_1 + \dots + c_nw_n$$
$$= \tilde{T}(c_1v_1) + \dots + \tilde{T}(c_nv_n)$$
$$= \tilde{T}(c_1v_1 + \dots + c_nv_n)$$

as desired.

- Sum (of $S, T \in \mathcal{L}(V, W)$): The linear map $(S + T) \in \mathcal{L}(V, W)$ defined by (S + T)(v) = Sv + Tv for all $v \in V$.
- **Product** (of $T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$): The linear map $(\lambda T) \in \mathcal{L}(V, W)$ defined by $(\lambda T)(v) = \lambda(Tv)$ for all $v \in V$.
- It follows that, under these definitions of addition and multiplication, $\mathcal{L}(V, W)$ is a vector space.
- **Product** (of $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$): The linear map $ST \in \mathcal{L}(U, W)$ defined by (ST)(u) = S(Tu) for all $u \in U$.
 - Note that the product is just function composition, but most mathematicians do write ST instead of $S \circ T$.
- Linear maps (with the correct corresponding domains) satisfy the associativity, identity, and distributive properties. However, multiplication of linear maps is not necessarily commutative.

- $(T_1T_2)T_3 = T_1(T_2T_3)$. - $TI_V = I_WT = T$ (note that if $T \in \mathcal{L}(V, W)$, $I_V \in \mathcal{L}(V, V)$ and $I_W \in \mathcal{L}(W, W)$). - $(S_1 + S_2)T = S_1T + S_2T$ and $S(T_1 + T_2) = ST_1 + ST_2$.
- Linear maps send 0 to 0.

Theorem 3.2. Suppose $T \in \mathcal{L}(V, W)$. Then T(0) = 0.

Proof. By additivity, we have

$$T(0) = T(0+0) = T(0) + T(0)$$
$$0 = T(0)$$

as desired.

3.B Null Spaces and Ranges

• Null space (of $T \in \mathcal{L}(V, W)$): The subset of V consisting of those vectors that T maps to 0. Also known as kernel. Denoted by null T. Given by

$$\operatorname{null} T = \{ v \in V : Tv = 0 \}$$

• The null space is a subspace.

Theorem 3.3. Suppose $T \in \mathcal{L}(V, W)$. Then null T is a subspace of V.

Proof. To prove that null T is a subspace of V, it will suffice to show that $0 \in \text{null } T$, $u, v \in \text{null } T$ implies that $u + v \in \text{null } T$, and $u \in \text{null } T$ and $\lambda \in \mathbb{F}$ imply $\lambda u \in \text{null } T$. Let's begin.

By Theorem 3.2, T(0) = 0. Therefore, $0 \in \text{null } T$, as desired.

Let $u, v \in \text{null } T$ be arbitrary. Then by additivity

$$T(u+v) = Tu + Tv = 0 + 0 = 0$$

so $u + v \in \text{null } T$, as desired.

Let $u \in \text{null } T$ and $\lambda \in \mathbb{F}$ be arbitrary. Then by homogeneity,

$$T(\lambda u) = \lambda T u = \lambda 0 = 0$$

so $\lambda u \in \text{null } T$, as desired.

- Injective (function): A function $T: V \to W$ such that Tu = Tv implies u = v. Also known as one-to-one.
- If 0 is the only vector that gets mapped to 0, then T is injective.

Theorem 3.4. Let $T \in \mathcal{L}(V, W)$. Then T is injective if and only if null $T = \{0\}$.

Proof. Suppose first that T is injective. To prove that $\text{null } T = \{0\}$, it will suffice to show that $0 \in \text{null } T$ and for every $v \in \text{null } T$, v = 0. By Theorem 3.3, $0 \in \text{null } T$. Now let $v \in \text{null } T$ be arbitrary. By the definition of the null space, we have Tv = 0. By Theorem 3.2, we have T(0) = 0. Thus, by transitivity, we have that Tv = T(0). It follows by injectivity that v = 0, as desired.

Now suppose that $\operatorname{null} T = \{0\}$. To prove that T is injective, it will suffice to show that if Tu = Tv, then u = v. Suppose $u, v \in V$ satisfy Tu = Tv. Then

$$0 = Tu - Tv = T(u - v)$$

so $(u-v) \in \text{null } T = \{0\}$. It follows that u-v=0, i.e., that u=v, as desired.

• Range (of $T \in \mathcal{L}(V, W)$): The subset of W consisting of those vectors that are of the form Tv for some $v \in V$. Also known as **image**. Denoted by **range** T. Given by

$$\operatorname{range} T = \{Tv : v \in V\}$$

• The range is a subspace.

Theorem 3.5. Suppose $T \in \mathcal{L}(V, W)$. Then range T is a subspace of W.

Proof. To prove that range T is a subspace of W, it will suffice to show that $0 \in \text{range } T$, $w_1, w_2 \in \text{range } T$ implies that $(w_1 + w_2) \in \text{range } T$, and $w \in \text{range } T$ and $\lambda \in \mathbb{F}$ imply $\lambda w \in \text{range } T$. Let's begin.

By the definition of a vector space, $0 \in V$. By Theorem 3.2, T(0) = 0. Therefore, $0 \in \text{range } T$, as desired.

Let $w_1, w_2 \in \text{range } T$ be arbitrary. Then there exist $v_1, v_2 \in V$ such that $Tv_1 = w_1$ and $Tv_2 = w_2$. It follows by additivity that

$$T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2$$

Therefore, since $v_1 + v_2 \in V$, we have that $(w_1 + w_2) \in \text{range } T$, as desired.

Let $w \in \operatorname{range} T$ and $\lambda \in \mathbb{F}$ be arbitrary. Then there exists $v \in V$ such that Tv = w. It follows by homogeneity that

$$T(\lambda v) = \lambda T v = \lambda w$$

Therefore, since $\lambda v \in V$, we have that $\lambda w \in \operatorname{range} T$, as desired.

- Surjective (function): A function $T: V \to W$ such that range T = W. Also known as onto.
- We now prove a very important theorem.

Theorem 3.6 (Fundamental Theorem of Linear Maps). Suppose V is finite-dimensional and $T \in \mathcal{L}(V,W)$. Then range T is finite-dimensional and

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$$

Proof. By Theorem 3.3, null T is a subspace of V finite-dimensional. Thus, by Theorem 2.4, null T is finite-dimensional. It follows by Theorem 2.7 that we may let u_1, \ldots, u_m be a basis of null T. As a basis of a subspace of V, u_1, \ldots, u_m is a linearly independent list of vectors in V. Consequently, by Theorem 2.8, we may extend it to a basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ of V.

Having established this terminology, we can now see that to prove that range T is finite-dimensional, it will suffice to show that Tv_1, \ldots, Tv_n spans it. To show that $\operatorname{span}(Tv_1, \ldots, Tv_n) = \operatorname{range} T$, it will suffice to show that every $b_1Tv_1 + \cdots + b_nTv_n \in \operatorname{span}(Tv_1, \ldots, Tv_n)$ is an element of range T and that every $Tv \in \operatorname{range} T$ is an element of $\operatorname{span}(Tv_1, \ldots, Tv_n)$. Let $b_1Tv_1 + \cdots + b_nTv_n \in \operatorname{span}(Tv_1, \ldots, Tv_n)$ be arbitrary. Then

$$b_1 T v_1 + \dots + b_n T v_n = T(b_1 v_1 + \dots + b_n v_n)$$

= $T(0u_1 + \dots + 0u_m + b_1 v_1 + \dots + b_n v_n)$

Therefore, since $0u_1 + \cdots + 0u_m + b_1v_1 + \cdots + b_nv_n \in V$ by V's closure under addition and scalar multiplication, we have that $b_1Tv_1 + \cdots + b_nTv_n \in \text{range } T$, as desired. Now let $Tv \in \text{range } T$ be arbitrary. Since $v \in V$ and $u_1, \ldots, u_m, v_1, \ldots, v_n$ is a basis of V, Theorem 2.5 implies that $v = a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_nv_n$ for some $a_1, \ldots, a_m, b_1, \ldots, b_n \in \mathbb{F}$. Therefore,

$$Tv = T(a_1u_1 + \dots + a_mu_m + b_1v_1 + \dots + b_nv_n)$$

$$= T(a_1u_1 + \dots + a_mu_m) + T(b_1v_1 + \dots + b_nv_n)$$

$$= a_1Tu_1 + \dots + a_mTu_m + b_1Tv_1 + \dots + v_nTv_n$$

$$= a_10 + \dots + a_m0 + b_1Tv_1 + \dots + v_nTv_n$$

$$= b_1Tv_1 + \dots + v_nTv_n$$

where each $Tu_j = 0$ because each $u_j \in \text{null } T$, so $Tv \in \text{span}(Tv_1, \dots, Tv_n)$, as desired.

Before we can verify the equation from the theorem, we need to establish one last fact: that Tv_1, \ldots, Tv_n is linearly independent. Suppose $c_1, \ldots, c_n \in \mathbb{F}$ make

$$c_1Tv_1 + \dots + c_nTv_n = 0$$

$$T(c_1v_1 + \dots + c_nv_n) = 0$$

It follows that $c_1v_1 + \cdots + c_nv_n \in \text{null } T$. Thus, since u_1, \ldots, u_m is a basis of null T by Theorem 2.5, we have that

$$c_1v_1 + \dots + c_nv_n = d_1u_1 + \dots + d_mu_m$$

$$0 = d_1u_1 + \dots + d_mu_m - c_1v_1 - \dots - c_nv_n$$

for some $d_1, \ldots, d_m \in \mathbb{F}$. But since $u_1, \ldots, u_m, v_1, \ldots, v_n$ is linearly independent as the basis of V, the above equation implies that $c_1 = \cdots = c_n = 0$, as desired.

Having established that $u_1, \ldots, u_m, v_1, \ldots, v_n$ is a basis of V, u_1, \ldots, u_m is a basis of null T, and Tv_1, \ldots, Tv_n spans range T and is linearly independent in range T (i.e., is a basis of range T), we have that

$$\dim V = m + n$$

$$= \dim \operatorname{null} T + \dim \operatorname{range} T$$

as desired.

• We can now prove that a linear map to a "smaller" vector space cannot be injective.

Theorem 3.7. Suppose V and W are finite-dimensional vector spaces such that $\dim V > \dim W$. Then no linear map from V to W is injective.

Proof. Let $T \in \mathcal{L}(V, W)$. Then

$$\dim \operatorname{null} T = \dim V - \dim \operatorname{range} T$$
 Fundamental Theorem of Linear Maps
$$\geq \dim V - \dim \operatorname{range} T$$
 Theorem 2.11
$$> 0$$

It follows that null T has a basis consisting of a list of one or more vectors. As a linearly independent list, naturally none of these vectors will be equal to the zero vector. Thus, since null T contains vectors other than 0, Theorem 3.4 implies that T is not injective.

• Similarly, we can prove that a linear map to a "bigger" vector space cannot be surjective.

Theorem 3.8. Suppose V and W are finite-dimensional vector spaces such that $\dim V < \dim W$. Then no linear map from V to W is surjective.

Proof. Let $T \in \mathcal{L}(V, W)$. Then

$$\dim \operatorname{range} T = \dim V - \dim \operatorname{null} T \qquad \qquad \operatorname{Fundamental\ Theorem\ of\ Linear\ Maps} \\ \leq \dim V \qquad \qquad <\dim W$$

Therefore, range $T \neq W$, so T cannot be surjective.

• Theorems 3.7 and 3.8 allow us to express questions about systems of linear equations in terms of linear maps.

- For example, a question about a **homogenous** system of linear equations could be, "does there exist a nonzero solution to the homogenous system $\sum_{k=1}^{n} A_{1,k} x_k = 0, \dots, \sum_{k=1}^{n} A_{m,k} x_k = 0$?"
- If we define $T: \mathbb{F}^n \to \mathbb{F}^m$ by

$$T(x_1, \dots, x_n) = \left(\sum_{k=1}^n A_{1,k} x_k, \dots, \sum_{k=1}^n A_{m,k} x_k\right)$$

we can express the system of equations as $T(x_1, \ldots, x_n) = 0$ and ask instead, "is dim null T > 0?"

- **Homogenous** (system of linear equations): A system of m linear equations $\sum_{k=1}^{n} A_{1,k} x_k = c_1$ through $\sum_{k=1}^{n} A_{m,k} x_k = c_m$ such that the constant term $c_j = 0$ for all $j = 1, \ldots, m$.
- Continuing with the linear equations example, we can rigorously show the following.

Theorem 3.9. A homogenous system of linear equations with more variables than equations has nonzero solutions.

Proof. In terms of the above, $T: \mathbb{F}^n \to \mathbb{F}^m$ where n > m. Thus, by Theorem 3.7, T is not injective. Consequently, by Theorem 3.4, dim null T > 0. Therefore, the system has nonzero solutions.

Theorem 3.10. An inhomogenous system of linear equations with more equations than variables has no solution for some choice of constant terms.

Proof. In terms of the above, $T: \mathbb{F}^n \to \mathbb{F}^m$ where m > n. We want to know if there exists $(c_1, \ldots, c_m) \in \mathbb{F}^m$ such that $T(x_1, \ldots, x_n) \neq (c_1, \ldots, c_m)$ for any $(x_1, \ldots, x_n) \in \mathbb{F}^n$. In other words, we want to know if there exists $(c_1, \ldots, c_m) \in \mathbb{F}^m$ such that $(c_1, \ldots, c_m) \notin \text{range } T$, i.e., if range $T \neq \mathbb{F}^m$. But since n < m, Theorem 3.8 asserts that T is not surjective, meaning that range $T \neq W$, as desired.

 Note that while the past two results are typically proven with Gaussian elimination, the abstract approach taken here leads to cleaner proofs.

3.C Matrices

• m-by-n matrix: A rectangular array A of elements of \mathbb{F} with m rows and n columns

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}$$

where m and n are positive integers.

- The notation $A_{j,k}$ denotes the entry in row j, column k of A. In other words, the first index refers to the row number and the second index refers to the column number.
- Matrix (of $T \in \mathcal{L}(V, W)$ with respect to the bases v_1, \ldots, v_n of V and w_1, \ldots, w_m of W): The m-by-n matrix $\mathcal{M}(T)$ whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}w_1 + \dots + A_{m,k}w_m$$

- If the bases are not clear from context, then the notation $\mathcal{M}(T,(v_1,\ldots,v_n),(w_1,\ldots,w_m))$ is used.
- Another way of wording the definition states that the k^{th} column of $\mathcal{M}(T)$ consists of the scalars needed to write Tv_k as a linear combination of w_1, \ldots, w_m .
- Assuming standard bases, we "can think of the k^{th} column of $\mathcal{M}(T)$ as the T applied to the k^{th} standard basis vector" (Axler, 2015, p. 71).

- Sum (of two m-by-n matrices A, C): The m-by-n matrix A + C defined by $(A + C)_{j,k} = A_{j,k} + C_{j,k}$.
 - Symbolically,

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} + \begin{pmatrix} C_{1,1} & \cdots & C_{1,n} \\ \vdots & & \vdots \\ C_{m,1} & \cdots & C_{m,n} \end{pmatrix} = \begin{pmatrix} A_{1,1} + C_{1,1} & \cdots & A_{1,n} + C_{1,n} \\ \vdots & & \vdots \\ A_{m,1} + C_{m,1} & \cdots & A_{m,n} + C_{m,n} \end{pmatrix}$$

- Suppose $S, T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.
- **Product** (of an m-by-n matrix A and $\lambda \in \mathbb{F}$): The m-by-n matrix λA defined by $(\lambda A)_{j,k} = \lambda A_{j,k}$.
 - Symbolically,

$$\lambda \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda A_{1,1} & \cdots & \lambda A_{1,n} \\ \vdots & & \vdots \\ \lambda A_{m,1} & \cdots & \lambda A_{m,n} \end{pmatrix}$$

- Suppose $\lambda \in \mathbb{F}$ and $T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.
- $\mathbb{F}^{m,n}$: The set of all m-by-n matrices with entries in \mathbb{F} , where m and n are positive integers.
- We have that dim $\mathbb{F}^{m,n} = mn$.
 - Note that a basis of $\mathbb{F}^{m,n}$ is the set of all m-by-n matrices that have 0s everywhere save a 1 in a single place.
- **Product** (of an m-by-n matrix A and an n-by-p matrix C): The m-by-p matrix AC defined by $(AC)_{j,k} = \sum_{r=1}^{n} A_{j,r} C_{r,k}$.
 - We may derive this by noting that if $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = C$, $T: U \to V$ and $S: V \to W$, and $u_1, \ldots, u_p, v_1, \ldots, v_n$, and w_1, \ldots, w_m are bases, then

$$(ST)u_k = S\left(\sum_{r=1}^n C_{r,k}v_r\right)$$

$$= \sum_{r=1}^n C_{r,k}Sv_r$$

$$= \sum_{r=1}^n C_{r,k}\sum_{j=1}^m A_{j,r}w_j$$

$$= \sum_{j=1}^m \left(\sum_{r=1}^n A_{j,r}C_{r,k}\right)w_j$$

- Matrix multiplication is not commutative, but is distributive and associative.
- Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$. Then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.
- If A is an m-by-n matrix, then...
 - We let $A_{j,}$ denote the 1-by-n matrix consisting of row j of A;
 - We let $A_{\cdot,k}$ denote the *m*-by-1 matrix consisting of column k of A.
- Thus, if A is an m-by-n matrix and C is an n-by-p matrix, then $(AC)_{j,k} = A_{j,.}C_{.,k}$ for all $1 \leq j \leq m$ and $1 \leq k \leq p$.
- Similarly, $(AC)_{\cdot,k} = AC_{\cdot,k}$.

- Lastly, suppose A is an m-by-n matrix and $c=(c_1,\ldots,c_n)$ is an n-by-1 matrix. Then $Ac=c_1A_{\cdot,1}+\cdots+c_nA_{\cdot,n}$.
 - In other words, Ac is a linear combination of the columns of A, with the scalars that multiply the columns coming from c.