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 $T(v)$.
- $\boldsymbol{\mathcal{L}(V, W)}$: 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 $Tp = \int_0^1 p(x)\,\mathrm{d}x$. This formalizes the fact that integrals are additive and homogenous.
  - 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 $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_n \in W$. Then there exists a unique linear map $T : V \to W$ such that $Tv_j = w_j$ for each $j = 1, \dots, n$.*

  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

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

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 $c_1, \dots, c_n \in \mathbb{F}$.

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

$$\begin{aligned} T(0v_1 + \cdots + 0v_{j-1} + 1v_j + 0v_{j+1} + \cdots + 0v_n) &= 0w_1 + \cdots + 0w_{j-1} + 1w_j + 0w_{j+1} + \cdots + 0w_n \\ T(v_j) &= w_j \end{aligned}$$

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

$$\begin{aligned} T(u + v) &= T((a_1 + c_1)v_1 + \cdots + (a_n + c_n)v_n) \\ &= (a_1 + c_1)w_1 + \cdots + (a_n + c_n)w_n \\ &= Tu + Tv \end{aligned}$$

and

$$\begin{aligned} T(\lambda v) &= T(\lambda c_1v_1 + \cdots + \lambda c_nv_n) \\ &= \lambda c_1w_1 + \cdots + \lambda c_nw_n \\ &= \lambda Tv \end{aligned}$$

as desired.

Now suppose $\tilde{T} \in \mathcal{L}(V, W)$ satisfies $\tilde{T}v_j = w_j$ for all $j = 1, \dots, n$. 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 $j = 1, \dots, n$. 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 $j = 1, \dots, n$. Similarly, the additivity of $\tilde{T}$ implies that

$$\begin{aligned} T(c_1v_1 + \cdots + c_nv_n) &= c_1w_1 + \cdots + c_nw_n \\ &= \tilde{T}(c_1v_1) + \cdots + \tilde{T}(c_nv_n) \\ &= \tilde{T}(c_1v_1 + \cdots + c_nv_n) \end{aligned}$$

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

  $$\begin{aligned} T(0) = T(0+0) &= T(0) + T(0) \\ 0 &= T(0) \end{aligned}$$

  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* $\textbf{null}\,\boldsymbol{T}$. *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 $\operatorname{null} T$ is a subspace of V.*

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

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

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

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

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

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

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

  so $\lambda u \in \operatorname{null} T$, 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 $\operatorname{null} T = \{0\}$.*

  Proof. Suppose first that T is injective. To prove that $\operatorname{null} T = \{0\}$, it will suffice to show that $0 \in \operatorname{null} T$ and for every $v \in \operatorname{null} T$, $v = 0$. By Theorem 3.3, $0 \in \operatorname{null} T$. Now let $v \in \operatorname{null} T$ 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 $(u-v) \in \operatorname{null} T = \{0\}$. 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* $\textbf{range}\,\boldsymbol{T}$. *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 $0 \in \operatorname{range} T$, $w_1, w_2 \in \operatorname{range} T$ implies that $(w_1 + w_2) \in \operatorname{range} T$, and $w \in \operatorname{range} T$ and $\lambda \in \mathbb{F}$ imply $\lambda w \in \operatorname{range} T$. Let's begin.

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

  Let $w_1, w_2 \in \operatorname{range} T$ 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 $(w_1 + w_2) \in \operatorname{range} T$, 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

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

  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 $u_1, \dots, u_m$ be a basis of null T. As a basis of a subspace of V, $u_1, \dots, u_m$ is a linearly independent list of vectors in V. Consequently, by Theorem 2.8, we may extend it to a basis $u_1, \dots, u_m, v_1, \dots, v_n$ 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 $Tv_1, \dots, Tv_n$ spans it. To show that $\operatorname{span}(Tv_1, \dots, Tv_n) = \operatorname{range} T$, it will suffice to show that every $b_1Tv_1 + \cdots + b_nTv_n \in \operatorname{span}(Tv_1, \dots, Tv_n)$ is an element of range T and that every $Tv \in \operatorname{range} T$ is an element of $\operatorname{span}(Tv_1, \dots, Tv_n)$. Let $b_1Tv_1 + \cdots + b_nTv_n \in \operatorname{span}(Tv_1, \dots, Tv_n)$ be arbitrary. Then

$$\begin{aligned} b_1Tv_1 + \cdots + b_nTv_n &= T(b_1v_1 + \cdots + b_nv_n) \\ &= T(0u_1 + \cdots + 0u_m + b_1v_1 + \cdots + b_nv_n) \end{aligned}$$

  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 $b_1Tv_1 + \cdots + b_nTv_n \in \operatorname{range} T$, as desired. Now let $Tv \in \operatorname{range} T$ be arbitrary. Since $v \in V$ and $u_1, \dots, u_m, v_1, \dots, v_n$ 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 $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbb{F}$. Therefore,

$$\begin{aligned} Tv &= T(a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_nv_n) \\ &= T(a_1u_1 + \cdots + a_mu_m) + T(b_1v_1 + \cdots + b_nv_n) \\ &= a_1Tu_1 + \cdots + a_mTu_m + b_1Tv_1 + \cdots + v_nTv_n \\ &= a_10 + \cdots + a_m0 + b_1Tv_1 + \cdots + v_nTv_n \\ &= b_1Tv_1 + \cdots + v_nTv_n \end{aligned}$$

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

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

$$\begin{aligned} c_1 Tv_1 + \cdots + c_n Tv_n &= 0 \\ T(c_1 v_1 + \cdots + c_n v_n) &= 0 \end{aligned}$$

It follows that $c_1 v_1 + \cdots + c_n v_n \in \operatorname{null} T$. Thus, since $u_1, \dots, u_m$ is a basis of $\operatorname{null} T$ by Theorem 2.5, we have that

$$\begin{aligned} c_1 v_1 + \cdots + c_n v_n &= d_1 u_1 + \cdots + d_m u_m \\ 0 &= d_1 u_1 + \cdots + d_m u_m - c_1 v_1 - \cdots - c_n v_n \end{aligned}$$

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

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

$$\begin{aligned} \dim V &= m + n \\ &= \dim \operatorname{null} T + \dim \operatorname{range} T \end{aligned}$$

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

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

  It follows that $\operatorname{null} T$ 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 $\operatorname{null} T$ 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

  $$\begin{aligned} \dim \operatorname{range} T &= \dim V - \dim \operatorname{null} T && \text{Fundamental Theorem of Linear Maps} \\ &\leq \dim V && < \dim W \end{aligned}$$

  Therefore, $\operatorname{range} T \neq W$, 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 $T(x_1, \dots, x_n) = 0$ and ask instead, "is $\dim \operatorname{null} T > 0$?"

- **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 $j = 1, \dots, m$.
- 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, $\dim \operatorname{null} T > 0$. 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 $(c_1, \dots, c_m) \in \mathbb{F}^m$ such that $T(x_1, \dots, x_n) \neq (c_1, \dots, c_m)$ for any $(x_1, \dots, x_n) \in \mathbb{F}^n$. In other words, we want to know if there exists $(c_1, \dots, c_m) \in \mathbb{F}^m$ such that $(c_1, \dots, c_m) \notin \operatorname{range} T$, i.e., if $\operatorname{range} T \neq \mathbb{F}^m$. But since $n < m$, Theorem 3.8 asserts that T is not surjective, meaning that $\operatorname{range} T \neq W$, 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 nunber and the second index refers to the column number.

- **Matrix** (of $T \in \mathcal{L}(V, W)$ with respect to the bases $v_1, \dots, v_n$ of V and $w_1, \dots, w_m$ of W): The m-by-n matrix $\mathcal{M}(T)$ whose entries $A_{j,k}$ are defined by

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

- If the bases are not clear from context, then the notation $\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$ 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 $w_1, \dots, w_m$.
- 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 $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.
- $\boldsymbol{\mathbb{F}^{m,n}}$: 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 $u_1, \dots, u_p$, $v_1, \dots, v_n$, and $w_1, \dots, w_m$ are bases, then

$$\begin{aligned} (ST)u_k &= S\left(\sum_{r=1}^{n} C_{r,k} v_r\right) \\ &= \sum_{r=1}^{n} C_{r,k} S v_r \\ &= \sum_{r=1}^{n} C_{r,k} \sum_{j=1}^{m} A_{j,r} w_j \\ &= \sum_{j=1}^{m} \left(\sum_{r=1}^{n} A_{j,r} C_{r,k}\right) w_j \end{aligned}$$

  - Matrix multiplication is not commutative, but is distributive and associative.
- We now prove a relation between matrix and operator composition.

Theorem 3.11. *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 $A_{j,\cdot}$ 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 $(AC)_{j,k} = A_{j,\cdot} C_{\cdot,k}$ 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 $c = (c_1, \dots, c_n)$ is an n-by-1 matrix. Then $Ac = c_1 A_{\cdot,1} + \cdots + c_n A_{\cdot,n}$.
  - In other words, Ac is a linear combination of the columns of A, with the scalars that multiply the columns coming from c.

3.D Invertibility and Isomorphic Vector Spaces

9/6:

- **Invertible** (linear map): A linear map $T \in \mathcal{L}(V, W)$ such that there exists a linear map $S \in \mathcal{L}(V, W)$ such that ST equals the identity map on V and TS equals the identity map on W.
- **Inverse** (of $T \in \mathcal{L}(V, W)$): The linear map $S \in \mathcal{L}(W, V)$ satisfying $ST = I_V$ and $TS = I_W$. *Denoted by* $\boldsymbol{T^{-1}}$.
- We now justify the use of the word "the" in the definition of the inverse.

  Theorem 3.12. *An invertible linear map has a unique inverse.*

  Proof. Suppose $T \in \mathcal{L}(V, W)$ is invertible and S_1, S_2 are inverses of T. Then

  $$S_1 = S_1 I = S_1(TS_2) = (S_1 T)S_2 = IS_2 = S_2$$

  as desired. ∎

- We now give a criterion for invertibility.

  Theorem 3.13. *A linear map is invertible if and only if it is injective and surjective.*

  Proof. Let $T \in \mathcal{L}(V, W)$. Suppose first that T is invertible.

  To prove that T is injective, it will suffice to show that for all $u, v \in V$, $Tu = Tv$ implies that $u = v$. Let u, v be arbitrary elements of V that satisfy $Tu = Tv$. Then

  $$u = T^{-1}(Tu) = T^{-1}(Tv) = v$$

  as desired.

  To prove that T is surjective, it will suffice to show that $\operatorname{range} T = W$. Since $\operatorname{range} T \subset W$, we need only show that $W \subset \operatorname{range} T$. Let $w \in W$ be arbitrary. Since $w = T(T^{-1}w)$ where $T^{-1}w \in V$, we have that $w \in \operatorname{range} T$, as desired.

  Now suppose that T is injective and surjective. To prove that T is invertible, we will define a function $S : W \to V$, prove that it is a linear map, prove that $TS = I_W$, and prove that $ST = I_V$. Let Sw be the unique element of V such that $T(Sw) = w$ (the surjectivity of T guarantees that there exists *an* element of V that T maps to w, and the injectivity of T guarantees the uniqueness of said element).

  To prove that S is a linear map, it will suffice to show that S is additive and homogenous. To verify additivity, first note that the additivity of T implies that

  $$T(Sw_1 + Sw_2) = T(Sw_1) + T(Sw_2) = w_1 + w_2$$

  But since the above equation implies that $Sw_1 + Sw_2$ is the unique element of V that T maps to $w_1 + w_2$, we have by the definition of S that $S(w_1 + w_2) = Sw_1 + Sw_2$. The proof is symmetric for homogeneity.

  To prove that $TS = I_W$, we need only appeal to the definition of S, which states that $(TS)w = T(Sw) = w$ for all $w \in W$. It immediately follows that $TS = I_W$.

  To prove that $ST = I_V$, first note that for all $v \in V$,

  $$T((ST)v) = (TS)(Tv) = I(Tv) = Tv$$

  It follows by the injectivity of T that $(ST)v = v$, i.e., that $ST = I_V$, as desired. ∎

- **Isomorphism**: An invertible linear map.
- **Isomorphic** (vector spaces): Two vector spaces such that there exists an isomorphism from one vector space onto the other one.
- Isomorphic vector spaces have the same dimension.

Theorem 3.14. *Two finite-dimensional vector spaces over $\mathbb{F}$ are isomorphic if and only if they have the same dimension.*

Proof. Suppose V, W are isomorphic finite-dimensional vector spaces over $\mathbb{F}$. Then there exists an isomorphism $T : V \to W$. By the definition of isomorphism, T is an invertible linear map, meaning by Theorem 3.13 that T is injective and surjective. Thus, since there exists an injective linear map $T : V \to W$, the contrapositive of Theorem 3.7 asserts that $\dim V \leq \dim W$. Additionally, since there exists a surjective linear map $T : V \to W$, the contrapositive of Theorem 3.8 asserts that $\dim V \geq \dim W$. Therefore, we have that $\dim V = \dim W$, as desired.

Now suppose that $\dim V = \dim W$. Let $v_1, \dots, v_n$ be a basis of V, and let $w_1, \dots, w_n$ be a basis of W. By Theorem 3.1, there exists a unique linear map $T : V \to W$ such that $Tv_j = w_j$ for each $j = 1, \dots, n$. To prove that T is an isomorphism, Theorem 3.13 tells us that it will suffice to show that it is injective and surjective. To show that T is surjective, it will suffice to show that $\operatorname{range} T = W = \operatorname{span}(w_1, \dots, w_n)$. But since $Tv_j = w_j \in \operatorname{range} T$ for all $j = 1, \dots, n$, $\operatorname{range} T \subset W$, and $\operatorname{range} T$ is a vector space (see Theorem 3.5), we have that $\operatorname{range} T = \operatorname{span}(w_1, \dots, w_n) = W$, as desired. To prove that T is injective, Theorem 3.4 tells us that it will suffice to show that $\operatorname{null} T = \{0\}$, i.e., that $\dim \operatorname{null} T = 0$. But since $\dim \operatorname{range} T = \dim W = \dim V$, we have by the Fundamental Theorem of Linear Maps that

$$\begin{aligned} \dim \operatorname{null} T + \dim \operatorname{range} T &= \dim V \\ &= \dim W \\ &= \dim \operatorname{range} T \\ \dim \operatorname{null} T &= 0 \end{aligned}$$

as desired. ∎

- This result implies that every finite-dimensional vector space of dimension n is isomorphic to $\mathbb{F}^n$.
- It also allows us to formalize the link between linear maps from V to W and matrices in $\mathbb{F}^{m,n}$.

Theorem 3.15. *Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. Then $\mathcal{M}$ is an isomorphism between $\mathcal{L}(V, W)$ and $\mathbb{F}^{m,n}$.*

Proof. We have already established that $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$ and that $\mathcal{M}(\lambda T) = \lambda\mathcal{M}(T)$, so we already know that $\mathcal{M}$ is a linear map. To prove that it is invertible, Theorem 3.13 tells us that it will suffice to show that $\mathcal{M}$ is injective and surjective.

To show that $\mathcal{M}$ is injective, Theorem 3.4 tells us that it will suffice to verify that $\operatorname{null} \mathcal{M} = \{0\}$. Let $T \in \mathcal{L}(V, W)$ be arbitrary. If $\mathcal{M}(T) = 0$ (0 denoting the zero matrix), then

$$Tv_k = 0w_1 + \cdots + 0w_m = 0$$

for all $k = 1, \dots, n$. But since $v_1, \dots, b_n$ is a basis of V, this implies that $T = 0$ (0 denoting the zero transformation), as desired.

To show that $\mathcal{M}$ is surjective, it will suffice to verify that $\operatorname{range} \mathcal{M} = \mathbb{F}^{m,n}$. Clearly $\operatorname{range} \mathcal{M} \subset \mathbb{F}^{m,n}$, so we focus on the other direction. Let $A \in \mathbb{F}^{m,n}$ be arbitrary. Define $T \in \mathcal{L}(V, W)$ by

$$Tv_k = \sum_{j=1}^{m} A_{j,k} w_j$$

for $k = 1, \dots, n$. It follows by the definition of a matrix of a linear transformation that $\mathcal{M}(T) = A$, as desired. ∎

- We can now determine the dimension of the vector space of linear maps from one finite-dimensional vector space to another.

  Theorem 3.16. *Suppose V and W are finite-dimensional. Then $\mathcal{L}(V,W)$ is finite-dimensional and*

  $$\dim \mathcal{L}(V,W) = (\dim V)(\dim W)$$

  Proof. By Theorem 3.15, $\mathcal{L}(V,W)$ and $\mathbb{F}^{m,n}$ are isomorphic. Thus, by Theorem 3.14, $\mathcal{L}(V,W)$ and $\mathbb{F}^{m,n}$ have the same dimension. Therefore, we have that

  $$\begin{aligned}\dim \mathcal{L}(V,W) &= \dim \mathbb{F}^{m,n} \\ &= mn \\ &= (\dim V)(\dim W)\end{aligned}$$

  as desired. ■

- **Matrix** (of $v \in V$ with respect to the basis $v_1,\dots,v_n$ of V): The n-by-1 matrix $\mathcal{M}(v)$ whose entries $A_{j,1}$ are defined by

  $$v = A_{1,1}v_1 + \cdots + A_{n,1}v_n$$

- We now show that the columns of the matrix of T are directly related to the effect T has on basis vectors.

  Theorem 3.17. *Suppose $T \in \mathcal{L}(V,W)$, $v_1,\dots,v_n$ is a basis of V, and $w_1,\dots,w_m$ is a basis of W. Let $1 \leq k \leq n$. Then*

  $$\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$$

  Proof. As an element of W, $Tv_k = c_1w_1 + \cdots + c_mw_m$ for some $c_1,\dots,c_m \in \mathbb{F}$. By the definition of the matrix of T, the values in column k are $c_1,\dots,c_m$. Similarly, by the definition of the matrix of Tv_k, the values in its one column are $c_1,\dots,c_m$, as desired. ■

- Linear maps act like matrix multiplication.

  Theorem 3.18. *Suppose $T \in \mathcal{L}(V,W)$, $v \in V$, $v_1,\dots,v_n$ is a basis of V, and $w_1,\dots,w_m$ is a basis of W. Then*

  $$\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v)$$

  Proof. Let $v = c_1v_1 + \cdots + c_nv_n$. Then by the linearity of T, $Tv = c_1Tv_1 + \cdots + c_nTv_n$. It follows by the linearity of $\mathcal{M}$, Theorem 3.17, and the fact that a matrix times a vector is equal to a linear combination of columns that

  $$\begin{aligned}\mathcal{M}(Tv) &= c_1\mathcal{M}(Tv_1) + \cdots + c_n\mathcal{M}(Tv_n) \\ &= c_1\mathcal{M}(T)_{\cdot,1} + \cdots + c_n\mathcal{M}(T)_{\cdot,n} \\ &= \mathcal{M}(T)\mathcal{M}(v)\end{aligned}$$

  as desired. ■

- "Each m-by-n matrix A induces a linear map from $\mathbb{F}^{n,1}$ to $\mathbb{F}^{m,1}$, namely the matrix multiplication function that takes $x \in \mathbb{F}^{n,1}$ to $Ax \in \mathbb{F}^{m,1}$" (Axler, 2015, p. 85).
- **Operator**: A linear map from a vector space to itself.
- $\boldsymbol{\mathcal{L}(V)}$: The set of all operators on V.
  - Mathematically, $\mathcal{L}(V) = \mathcal{L}(V,V)$.

- On infinite-dimensional vector spaces, neither injectivity nor surjectivity, alone, implies invertibility.
- However, only one does on finite-dimensional vector spaces.

Theorem 3.19. *Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then the following are equivalent.*

(a) *T is invertible;*
(b) *T is injective;*
(c) *T is surjective.*

Proof. First, suppose that T is invertible. Then by Theorem 3.13, T is injective, as desired.

Second, suppose that T is injective. Then by Theorem 3.4, $\operatorname{null} T = \{0\}$. It follows by the Fundamental Theorem of Linear Maps that

$$\begin{aligned}\dim \operatorname{range} T &= \dim V - \dim \operatorname{null} T\\ &= \dim V\end{aligned}$$

Thus, since $\operatorname{range} T$ has the same dimension as V and is a subspace of V (by Theorem 3.5), $\operatorname{range} T = V$. Therefore, T is surjective, as desired.

Third, suppose that T is surjective. Then $\operatorname{range} T = V$. It follows that $\dim \operatorname{range} T = \dim V$. Thus, by the Fundamental Theorem of Linear Maps, we have that

$$\begin{aligned}\dim \operatorname{null} T &= \dim V - \dim \operatorname{range} T\\ &= 0\end{aligned}$$

Consequently, by Theorem 3.4, T is injective. Therefore, by Theorem 3.13, T is invertible, as desired. ■

Exercises

10/11: **9** Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST is invertible if and only if both S and T are invertible.

Proof. Suppose first that ST is invertible. Then by Theorem 3.19, ST is injective and surjective. We will prove that S, T are invertible in turn; in particular, we will first prove that S is surjective and then prove that T is injective. Let's begin. To prove that S is surjective, it will suffice to show that $\operatorname{range} S = V$. Let $v \in V$ be arbitrary. Since ST is surjective, there exists $v' \in V$ such that $STv' = v$. Thus, since $Tv' \in V$, the fact that $S(Tv') = v$ implies that $v \in \operatorname{range} S$. The inclusion in the other direction is obvious. Now to prove that T is injective, it will suffice to show that $Tv = Tv'$ implies $v' = v'$. Let $Tv = Tv'$. Then $STv = STv'$. It follows that $v = v'$ by the injectivity of ST, as desired.

Now suppose that S and T are invertible. Then by Theorem 3.12, there exist S^{-1}, T^{-1} such that

$$SS^{-1} = I = S^{-1}S \qquad TT^{-1} = I = T^{-1}T$$

Let $(ST)^{-1} = T^{-1}S^{-1}$. Then

$$ST(ST)^{-1} = STT^{-1}S^{-1} = SIS^{-1} = SS^{-1} = I = T^{-1}T = T^{-1}IT = T^{-1}S^{-1}ST = (ST)^{-1}ST$$

so ST is invertible, as desired. ■

10 Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that $ST = I$ if and only if $TS = I$.

Proof. Suppose first that $ST = I$. It is easy to see that defining $(ST)^{-1} = I^{-1} = I$ yields an inverse of ST. Thus, since ST is invertible, we have by Exercise 3.D.9 that S and T are invertible. It follows since $ST = I$ that $S = T^{-1}$, meaning that

$$TS = TT^{-1} = I$$

as desired.

The proof is symmetric in the other direction. ■

3.E Products and Quotients of Vector Spaces

9/6:
- **Product** (of $V_1, \dots, V_m$): The set
$$V_1 \times \cdots \times V_m = \{(v_1, \dots, v_m) : v_1 \in V_1, \dots, v_m \in V_m\}$$
  - Addition on $V_1 \times \cdots \times V_m$ is defined by
$$(u_1, \dots, u_m) + (v_1, \dots, v_m) = (u_1 + v_1, \dots, u_m + v_m)$$
  - Scalar multiplication on $V_1 \times \cdots \times V_m$ is defined by
$$\lambda(v_1, \dots, v_m) = (\lambda v_1, \dots, \lambda v_m)$$
- The product of n vector spaces over $\mathbb{F}$ is a vector space over $\mathbb{F}$, with addition and scalar multiplication defined as above.
- We can, for example, identify $\mathbb{R}^2 \times \mathbb{R}^3$ with $\mathbb{R}^5$ by constructing an isomorphism from every vector $((x_1, x_2), (x_3, x_4, x_5)) \in \mathbb{R}^2 \times \mathbb{R}^3$ to the vector $(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5$.
- The dimension of the product of vector spaces is equal to the sum of the dimensions of the component vector spaces.

  Theorem 3.20. *Suppose $V_1, \dots, V_m$ are finite-dimensional vector spaces. Then $V_1 \times \cdots \times V_m$ is finite-dimensional and*
$$\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m$$

  Proof. Choose a basis of each V_j. For each basis vector of each V_j, consider the element of $V_1 \times \cdots \times V_m$ that equals the basis vector in the j^{th} slot and 0 in the other slots. The list of all such vectors is linearly independent and spans $V_1 \times \cdots \times V_m$. Thus, it is a basis of $V_1 \times \cdots \times V_m$. The length of this basis is $\dim V_1 + \cdots + \dim V_m$, as desired. ■

- We now relate products and direct sums.

  Theorem 3.21. *Suppose that $U_1, \dots, U_m$ are subspaces of V. Define a linear map $\Gamma : U_1 \times \cdots \times U_m \to U_1 + \cdots + U_m$ by*
$$\Gamma(u_1, \dots, u_m) = u_1 + \cdots + u_m$$
  Then $U_1 + \cdots + U_m$ is a direct sum if and only if Γ is injective.

  Proof. Suppose first that Γ is injective. Then the only way to write 0 as a sum $u_1 + \cdots + u_m$, where each $u_j \in U_j$, is by taking each $u_j = 0$. It follows by the condition on direct sums that $U_1 + \cdots + U_m$ is a direct sum. The proof is symmetric in the reverse direction. ■

  - Note that since Γ is surjective by the definition of $U_1 + \cdots + U_m$, the condition that Γ is injective could be changed to the condition that Γ is invertible.

- We can now prove that the dimensions add up in a direct sum.

  Theorem 3.22. *Suppose V is finite-dimensional and $U_1, \dots, U_m$ are subspaces of V. Then $U_1 + \cdots + U_m$ is a direct sum if and only if*
$$\dim(U_1 + \cdots + U_m) = \dim U_1 + \cdots + \dim U_m$$

  Proof. Suppose first that $U_1 + \cdots + U_m$ is a direct sum. Then by Theorem 3.21, there exists an invertible linear map Γ from $U_1 \times \cdots \times U_m$ to $U_1 + \cdots + U_m$. Thus, by Theorem 3.14, $U_1 \times \cdots \times U_m$ and $U_1 + \cdots + U_m$ have the same dimension. Therefore,
$$\begin{aligned}\dim(U_1 + \cdots + U_m) &= \dim(U_1 \times \cdots \times U_m) \\ &= \dim U_1 + \cdots + \dim U_m && \text{Theorem 3.20}\end{aligned}$$
  as desired.

  The proof is symmetric in the other direction. ■

- **Sum** (of $v \in V$ and U a subspace of V): The subset of V defined by

$$v + U = \{v + u : u \in U\}$$

- **Affine subset** (of V): A subset of V of the form $v + U$ for some $v \in V$ and some subspace U of V.
- **Parallel** (subset to U): An affine subset $v + U$ of V.
- **Quotient space**: The set of all affine subsets of V parallel to U.

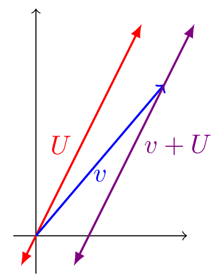

Figure 3.1: Visualizing $v + U$.

  - Symbolically,

$$V/U = \{v + U : v \in V\}$$

- Two affine subsets parallel to U are equal or disjoint.

Theorem 3.23. *Suppose U is a subspace of V and $v, w \in V$. Then the following are equivalent.*

(a) $v - w \in U$;

(b) $v + U = w + U$;

(c) $(v + U) \cap (w + U) \neq \emptyset$.

Proof. First, suppose that $v - w \in U$. Let $x \in v + U$ be arbitrary. Then $x = v + u$ for some $u \in U$. Now since $v - w \in U$, $u \in U$, and U is a subspace, we have that $v - w + u \in U$. Thus, $x = w - w + v + u = w + (v - w + u) \in w + U$. The proof is symmetric in the other direction. Therefore, $v + U = w + U$, as desired.

Second, suppose that $v + U = w + U$. Since U is nonempty ($0 \in U$ by definition), we know that $v + U \neq \emptyset \neq w + U$. Therefore, $(v + U) \cap (w + U) \supset \{0\} \neq \emptyset$, as desired.

Third, suppose that $(v + U) \cap (w + U) \neq \emptyset$. Then there exists x such that $x \in v + U$ and $x \in w + U$. It follows that $x = v + u_1$ and $x = w + u_2$ for some $u_1, u_2 \in U$. Thus, by transitivity, $v + u_1 = w + u_2$. Therefore, $v - w = u_2 - u_1 \in U$, as desired. ∎

- **Sum** (of $v + U, w + U \in V/U$): The affine subset $(v + w) + U$.
- **Product** (of $v + U \in V/U$ and $\lambda \in \mathbb{F}$): The affine subset $(\lambda v) + U$.
- We now verify that the above operations are well-defined and prove that the quotient space is a vector space.

Theorem 3.24. *Suppose U is a subspace of V. Then V/U, with the operations of addition and scalar multiplication as defined above, is a vector space.*

Proof. The way affine subsets are defined, we may have $v + U = \hat{v} + U$ and yet have $v \neq \hat{v}$. Thus, we must first guarantee that the operations of addition and scalar multiplication, as defined above, are well-defined, i.e., that if $v + U = \hat{v} + U$ and $w + U = \hat{w} + U$, then $(v + w) + U = (\hat{v} + \hat{w}) + U$ and $(\lambda v) + U = (\lambda \hat{v}) + U$. Let's begin.

To confirm that addition as defined above is a well-defined operation, let $v, \hat{v}, w, \hat{w} \in V$ be such that $v + U = \hat{v} + U$ and $w + U = \hat{w} + U$. Then by Theorem 3.23, $v - \hat{v} \in U$ and $w - \hat{w} \in U$. It follows since U is a subspace that $(v - \hat{v}) + (w - \hat{w}) \in U$. Consequently, $(v + w) - (\hat{v} + \hat{w}) \in U$, so by Theorem 3.23 again, $(v + w) + U = (\hat{v} + \hat{w}) + U$, as desired.

Similarly, $v + U = \hat{v} + U$ implies $v - \hat{v} \in U$, implies $\lambda(v - \hat{v}) = \lambda v - \lambda \hat{v} \in U$, implies $(\lambda v) + U = (\lambda \hat{v}) + U$, as desired.

The remaining proof that V/U is a vector space is straightforward; note that $0 + U$ is the identity element and $(-v) + U$ is the additive inverse. ■

- **Quotient map**: The linear map $\pi : V \to V/U$ defined by $\pi(v) = v + U$ for all $v \in V$.
- We now give a formula for the dimension of a quotient space.

Theorem 3.25. *Suppose V is finite-dimensional and U is a subspace of V. Then*

$$\dim V/U = \dim V - \dim U$$

Proof. Let π be the quotient map from V to V/U. From Theorem 3.23, we know that in order for $w + U = 0 + U$, we must have $v - 0 = v \in U$. Thus, $\pi(u) = 0$ if and only if $u \in U$, meaning $\operatorname{null} \pi = U$. Additionally, we clearly have that $\operatorname{range} \pi = V/U$. Therefore, by the Fundamental Theorem of Linear Maps, we have that

$$\begin{aligned} \dim V &= \dim \operatorname{null} \pi + \dim \operatorname{range} \pi \\ &= \dim U + \dim V/U \\ \dim V/U &= \dim V - \dim U \end{aligned}$$

as desired. ■

- Lastly, consider the fact that we can add any vector in the null space of a linear map T to an argument passed to T without changing its output. In other words, if $T \in \mathcal{L}(V, W)$, $v \in V$, and $u \in \operatorname{null} T$, then $T(v + u) = Tv + Tu = Tv$. We formalize this concept with the following definition.
- $\boldsymbol{\tilde{T}}$: The function from $V/(\operatorname{null} T)$ to W defined by $\tilde{T}(v + \operatorname{null} T) = Tv$, where $T \in \mathcal{L}(V, W)$.
- We now state a few basic results about $\tilde{T}$.

Theorem 3.26. *Suppose $T \in \mathcal{L}(V, W)$. Then*

(*a*) *$\tilde{T}$ is a linear map from $V/(\operatorname{null} T)$ to W;*

(*b*) *$\tilde{T}$ is injective;*

(*c*) $\operatorname{range} \tilde{T} = \operatorname{range} T$*;*

(*d*) *$V/(\operatorname{null} T)$ is isomorphic to* $\operatorname{range} T$*.*

3.F Duality

9/7: - **Linear functional** (on v): A linear map from V to $\mathbb{F}$.
  - In other words, a linear functional is an element of $\mathcal{L}(V, \mathbb{F})$.

- **Dual space** (of V): The vector space of all linear functionals on V. *Denoted by* $\boldsymbol{V'}$. *Also known as* $\boldsymbol{V^*}$. *Given by*

$$V' = \mathcal{L}(V, \mathbb{F})$$

- We now give a definition of the dimension of the dual space.

  Theorem 3.27. *Suppose V is finite-dimensional. Then V' is also finite-dimensional and*

$$\dim V' = \dim V$$

  Proof. By Theorem 3.16, we have that

$$\begin{aligned}\dim V' &= \dim \mathcal{L}(V, \mathbb{F})\\ &= (\dim V)(\dim \mathbb{F})\\ &= (\dim V)(1)\\ &= \dim V\end{aligned}$$

  as desired. ∎

- **Dual basis** (of a basis $v_1, \dots, v_n$ of V): The list $\varphi_1, \dots, \varphi_n$ of elements of V', where each φ_j is the linear functional on V such that

$$\varphi_j(v_k) = \begin{cases} 1 & k = j\\ 0 & k \neq j\end{cases}$$

  where $v_1, \dots, v_n$ is a basis of V.

- We now verify that the dual basis of a basis of V is actually a basis of the dual space.

  Theorem 3.28. *Suppose V is finite-dimensional. Then the dual basis of a basis of V is a basis of V'.*

  Proof. Let $v_1, \dots, v_n$ be a basis of V, and let $\varphi_1, \dots, \varphi_n$ be the corresponding dual basis. Since the dual basis has length equal to the dimension of V' (by Theorem 3.27), Theorem 2.12 tells us that it will suffice to show that $\varphi_1, \dots, \varphi_n$ is linearly independent to confirm that it is a basis of V'. To do so, suppose

$$a_1\varphi_1 + \cdots + a_n\varphi_n = 0$$

  where $a_1, \dots, a_n \in \mathbb{F}$ and 0 denotes the zero transformation. Since $(a_1\varphi_1 + \cdots + a_n\varphi_n)(v_j) = a_j$ for $j = 1, \dots, n$, we have that for any vector $c_1v_1 + \cdots + c_nv_n \in V$,

$$(a_1\varphi_1 + \cdots + a_n\varphi_n)(c_1v_1 + \cdots + c_nv_n) = c_1a_1 + \cdots + c_na_n$$

  Therefore, the only way to guarantee that $c_1a_1 + \cdots + c_na_n = 0$ is to let $a_1 = \cdots = a_n = 0$, as desired. ∎

- **Dual map** (of $T \in \mathcal{L}(V, W)$): The linear map $T' \in \mathcal{L}(W', V')$ defined by

$$T'(\varphi) = \varphi \circ T$$

  for all $\varphi \in W'$. *Also known as* $\boldsymbol{T^*}$.

- We now prove some algebraic properties of dual maps.

  Theorem 3.29.

  (a) $(S + T)' = S' + T'$ *for all* $S, T \in \mathcal{L}(V, W)$.

Proof. Let $S, T \in \mathcal{L}(V, W)$ be arbitrary. To prove that $(S+T)' = S' + T'$, it will suffice to show that $(S+T)'(\varphi) = (S' + T')(\varphi)$ for all $\varphi \in W'$. Let $\varphi \in W'$ be arbitrary. However, before we go into the main equality, it will be useful if we verify that $\varphi \circ (S+T) = \varphi \circ S + \varphi \circ T$. To do so, it will suffice to show that $(\varphi \circ (S+T))(v) = (\varphi \circ S + \varphi \circ T)(v)$ for all $v \in V$. Let $v \in V$ be arbitrary. Then

$$\begin{aligned}(\varphi \circ (S+T))(v) &= \varphi((S+T)(v)) \\ &= \varphi(S(v) + T(v)) \\ &= \varphi(S(v)) + \varphi(T(v)) \\ &= (\varphi \circ S)(v) + (\varphi \circ T)(v) \\ &= (\varphi \circ S + \varphi \circ T)(v)\end{aligned}$$

Now we can show that

$$\begin{aligned}(S+T)'(\varphi) &= \varphi \circ (S+T) \\ &= \varphi \circ S + \varphi \circ T \\ &= S'(\varphi) + T'(\varphi) \\ &= (S' + T')(\varphi)\end{aligned}$$

as desired. ■

(*b*) *$(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbb{F}$ and all $T \in \mathcal{L}(V, W)$.*

Proof. The proof is symmetric to the proof of part (a). ■

(*c*) *$(ST)' = T'S'$ for all $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$.*

Proof. Let $\varphi \in W'$ be arbitrary. Then

$$(ST)'(\varphi) = \varphi \circ (ST) = (\varphi \circ S) \circ T = T'(\varphi \circ S) = T'(S'(\varphi)) = (T'S')(\varphi)$$

as desired. ■

- **Annihilator** (of $U \subset V$): The set

$$U^0 = \{\varphi \in V' : \varphi(u) = 0 \ \forall \ u \in U\}$$

- The annihilator is a subspace.

Theorem 3.30. *Suppose $U \subset V$. Then U^0 is a subspace of V'*

Proof. To prove that U^0 is a subspace of V', it will suffice to show that $0 \in U^0$, $\varphi, \psi \in U^0$ implies $\varphi + \psi \in U^0$, and $\varphi \in U^0$ and $\lambda \in \mathbb{F}$ imply $\lambda\varphi \in U^0$. Let's begin.

Since $0(u) = 0$ for all $u \in U$, $0 \in U^0$.

Let $\varphi, \psi \in U^0$ be arbitrary. Let $u \in U$ be arbitrary. Then $(\varphi + \psi)(u) = \varphi(u) + \psi(u) = 0 + 0 = 0$, as desired.

The proof is symmetric for scalar multiplication. ■

- Dimension of the annihilator.

Theorem 3.31. *Suppose V is finite-dimensional and U is a subspace of V. Then*

$$\dim U + \dim U^0 = \dim V$$

Proof. Let $i \in \mathcal{L}(U, V)$ be the identity map $i(u) = u$ for all $u \in U$. Then $i' : V' \to U'$ is a linear map. It follows by the Fundamental Theorem of Linear Maps that

$$\dim \operatorname{range} i' + \dim \operatorname{null} i' = \dim V'$$

Since $i'(\varphi) = \varphi \circ i = \varphi$ for all $\varphi \in V'$, and $U^0 = \{\varphi \in V' : \varphi = 0\}$, we have that $i'(\varphi) = 0$ for all $\varphi \in U^0$. Thus, $U^0 = \operatorname{null} i'$. Additionally, we have that $\dim V = \dim V'$ by Theorem 3.27. Lastly, let $\psi \in U'$ be arbitrary. Define $\psi \in V'$ by

$$\psi(v) = \begin{cases} \varphi(v) & v \in U \\ 0 & v \notin U \end{cases}$$

Thus, $i'(\psi) = \psi \circ i = \varphi$. It follows that $\varphi \in \operatorname{range} i'$. Consequently, $\operatorname{range} i' = U'$, so $\dim U = \dim U' = \dim \operatorname{range} i'$ by Theorem 3.27. Therefore, we have from the first equation and the three substitutions that

$$\dim U + \dim U^0 = \dim V$$

as desired.[1] ∎

- We now describe the null space of T'.

  Theorem 3.32. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then*

  (*a*) $\operatorname{null} T' = (\operatorname{range} T)^0$.

  Proof. First, let $\varphi \in \operatorname{null} T'$ be arbitrary. Then $T'(\varphi) = \varphi \circ T = 0$. It follows that $0 = (\varphi \circ T)(v) = \varphi(Tv)$ for all $v \in V$. But this means that φ is a linear functional that maps every element of $\operatorname{range} T$ to 0, i.e., that $\varphi \in (\operatorname{range} T)^0$. The proof is symmetric in the other direction. ∎

  (*b*) $\dim \operatorname{null} T' = \dim \operatorname{null} T + \dim W - \dim V$.

  Proof. We have that

$$\begin{aligned} \dim \operatorname{null} T' &= \dim(\operatorname{range} T)^0 && \text{Theorem 3.32a} \\ &= \dim W - \dim \operatorname{range} T && \text{Theorem 3.31} \\ &= \dim W - (\dim V - \dim \operatorname{null} T) && \text{Fundamental Theorem of Linear Maps} \\ &= \dim \operatorname{null} T + \dim W - \dim V \end{aligned}$$

  as desired. ∎

  - Note that the proof of part (a) does not use the hypothesis that V, W are finite-dimensional, so the argument holds for infinite-dimensional vector spaces as well.

- T surjective is equivalent to T' injective.

  Theorem 3.33. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is surjective if and only if T' is injective.*

  Proof. Suppose first that T is surjective. Then $\operatorname{range} T = W$. It follows by Theorem 3.31 that

$$\dim(\operatorname{range} T)^0 = \dim W - \dim \operatorname{range} T = 0$$

  meaning that $(\operatorname{range} T)^0 = \{0\}$. Thus, by Theorem 3.32a, $\operatorname{null} T' = \{0\}$. Therefore, by Theorem 3.4, T' is injective, as desired.

  The proof is symmetric in the other direction. ∎

[1] Note that we may also prove this by constructing a basis of U extending it to a basis of V, and showing that the extended portion of the dual basis is a basis of U^0.

- We now describe the range space of T'.

Theorem 3.34. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V,W)$. Then*

(*a*) $\dim \operatorname{range} T' = \dim \operatorname{range} T$.

Proof. We have that

$$\begin{aligned}
\dim \operatorname{range} T' &= \dim W' - \dim \operatorname{null} T' && \text{Fundamental Theorem of Linear Maps} \\
&= \dim W - \dim \operatorname{null} T' && \text{Theorem 3.27} \\
&= \dim W - \dim(\operatorname{range} T)^0 && \text{Theorem 3.32a} \\
&= \dim \operatorname{range} T && \text{Theorem 3.31}
\end{aligned}$$

as desired. ■

(*b*) $\operatorname{range} T' = (\operatorname{null} T)^0$.

Proof. First, let $\varphi \in \operatorname{range} T'$ be arbitrary. Then there exists $\psi \in W'$ such that $\varphi = T'(\psi)$. Now let $v \in \operatorname{null} T$ be arbitrary. It follows that

$$\varphi(v) = (T'(\psi))(v) = (\psi \circ T)(v) = \psi(Tv) = \psi(0) = 0$$

Therefore, $\varphi \in (\operatorname{null} T)^0$, as desired.
Second, we have that

$$\begin{aligned}
\dim \operatorname{range} T' &= \dim \operatorname{range} T && \text{Theorem 3.34a} \\
&= \dim V - \dim \operatorname{null} T && \text{Fundamental Theorem of Linear Maps} \\
&= \dim(\operatorname{null} T)^0 && \text{Theorem 3.31}
\end{aligned}$$

Therefore, since Theorem 3.5 implies that $\operatorname{range} T'$ is a subspace of $(\operatorname{null} T)^0$ and $\dim \operatorname{range} T' = \dim(\operatorname{null} T)^0$, Exercise 2.C.1 asserts that $\operatorname{range} T' = (\operatorname{null} T)^0$, as desired. ■

- T injective is equivalent to T' surjective.

Theorem 3.35. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V,W)$. Then T is injective if and only if T' is surjective.*

Proof. Suppose first that T is injective. Then by Theorem 3.4, $\operatorname{null} T = \{0\}$. Thus, since Theorem 3.2 asserts that $\varphi(0) = 0$ for any linear functional, we have that every linear functional is in the annihilator of $\operatorname{null} T$, i.e., that $(\operatorname{null} T)^0 = V'$. It follows by Theorem 3.34b that $\operatorname{range} T' = V'$. Therefore, T' is surjective, as desired.

The proof is symmetric in the other direction. ■

9/8: - **Transpose** (of an m-by-n matrix A): The matrix obtained from A by interchanging the rows and columns. More specifically, the n-by-m matrix A^t whose entries are given by $(A^t)_{k,j} = A_{j,k}$. *Denoted by* $\boldsymbol{A^t}$.

- Properties of the transpose:

$$(A + C)^t = A^t + C^t \qquad\qquad (\lambda A)^t = \lambda A^t$$

- Transpose of a product.

Theorem 3.36. *If A is an m-by-n matrix and C is an n-by-p matrix, then*

$$(AC)^t = C^t A^t$$

Proof. We have that

$$\begin{aligned}((AC)^t)_{k,j} &= (AC)_{j,k} \\ &= \sum_{r=1}^{n} A_{j,r}C_{r,k} \\ &= \sum_{r=1}^{n} (C^t)_{k,r}(A^t)_{r,j} \\ &= (C^tA^t)_{k,j}\end{aligned}$$

for all $1 \leq k \leq p$ and $1 \leq j \leq m$, as desired. ■

- We now show that the transpose and the dual map are essentially the same object.

Theorem 3.37. *Suppose $T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(T') = (\mathcal{M}(T))^t$.*

Proof. Let $v_1, \dots, v_n$ be a basis of V, and let $\varphi_1, \dots, \varphi_n$ be the corresponding dual basis of V'. Similarly, let $w_1, \dots, w_m$ be a basis of W, and let $\psi_1, \dots, \psi_m$ be the corresponding dual basis of W'. Let $A = \mathcal{M}(T)$ and $C = \mathcal{M}(T')$. Let $1 \leq j \leq m$ and $1 \leq k \leq n$ be arbitrary. Then we have from the definition of $\mathcal{M}(T')$ that

$$T'(\psi_j) = \sum_{r=1}^{n} C_{r,j}\varphi_r$$

from the definition of T' that

$$\begin{aligned}(\psi \circ T)(v_k) &= \sum_{r=1}^{n} C_{r,j}\varphi_r(v_k) \\ &= C_{k,j}\end{aligned}$$

and from the definition of $\mathcal{M}(T)$ that

$$\begin{aligned}(\psi \circ T)(v_k) &= \psi_j(Tv_k) \\ &= \psi_j\left(\sum_{r=1}^{m} A_{r,k}w_r\right) \\ &= \sum_{r=1}^{m} A_{r,k}\psi_j(w_r) \\ &= A_{j,k}\end{aligned}$$

Therefore, from the last two results, we have by transitivity that $A_{j,k} = C_{k,j}$ for all $1 \leq j \leq m$ and $1 \leq k \leq n$. It follows that $C = A^t$, i.e., that $\mathcal{M}(T') = (\mathcal{M}(T))^t$, as desired. ■

- **Row rank** (of a matrix A): The dimension of the span of the rows of A in $\mathbb{F}^{1,n}$.
- **Column rank** (of a matrix A): The dimension of the span of the columns of A in $\mathbb{F}^{m,1}$.
- The dimension of range T equals the column rank of $\mathcal{M}(T)$.

Theorem 3.38. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then* $\dim \operatorname{range} T$ *equals the column rank of $\mathcal{M}(T)$.*

Proof. Let $v_1, \dots, v_n$ be a basis of V, and let $w_1, \dots, w_m$ be a basis of W. Since $Tv = c_1Tv_1 + \cdots + c_nTv_n$ for all $Tv \in \operatorname{range} T$ (because $v = c_1v_1 + \cdots + c_nTv_n$ for some $c_1, \dots, c_n \in \mathbb{F}$ for all $v \in V$, and T is a linear map), we have that $\operatorname{range} T = \operatorname{span}(Tv_1, \dots, Tv_n)$. Additionally, since $\mathcal{M}$ is an isomorphism from $\operatorname{span}(Tv_1, \dots, Tv_n)$ to $\operatorname{span}(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n))$, Theorem 3.14 asserts that $\dim\operatorname{span}(Tv_1, \dots, Tv_n) = \dim\operatorname{span}(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n))$. Therefore,

$$\begin{aligned}\dim\operatorname{range} T &= \dim\operatorname{span}(Tv_1, \dots, Tv_n)\\ &= \dim\operatorname{span}(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n))\end{aligned}$$

where the latter value is the column rank, as desired. ∎

- Row rank equals column rank.

Theorem 3.39. *Suppose $A \in \mathbb{F}^{m,n}$. Then the row rank of A equals the column rank of A.*

Proof. Let $T : \mathbb{F}^{n,1} \to \mathbb{F}^{m,1}$ be defined by $Tx = Ax$. It follows that $\mathcal{M}(T) = A$. Thus,

$$\begin{aligned}\text{column rank}\, A &= \text{column rank}\, \mathcal{M}(T) &&\\ &= \dim\operatorname{range} T && \text{Theorem 3.38}\\ &= \dim\operatorname{range} T' && \text{Theorem 3.34a}\\ &= \text{column rank}\, \mathcal{M}(T') && \text{Theorem 3.38}\\ &= \text{column rank}\, A^t && \text{Theorem 3.37}\\ &= \text{row rank}\, A &&\end{aligned}$$

as desired. ∎

- **Rank** (of A): The column rank of A.