

# Linear Algebra Notes

Cruz Godar

Math 306 and 406  
Professor Mendes  
Cal Poly, Winter 2017–Spring 2017

## I — Vector Spaces

**Definition 1.1:** Let  $k$  be a field. A **vector space** over  $k$  is a set  $V$  equipped with two binary operations  $+$  and  $\cdot$  such that for all  $u, v, w \in V$  and  $c, d \in k$ ,

1.  $u + v \in V$  and  $cv \in V$ .
2.  $u + v = v + u$ .
3.  $u + (v + w) = (u + v) + w$  and  $c(dv) = (cd)v$ .
4.  $c(u + v) = cu + cv$  and  $(c + d)v = cv + dv$ .
5. There is a vector  $0 \in V$  that satisfies  $v + 0 = v$  for all  $v \in V$ .
6. There is a vector  $1 \in V$  that satisfies  $1v = v$  for all  $v \in V$ .
7. For all  $v \in V$ , there is a vector  $-v \in V$  such that  $v + (-v) = 0$ .

**Proposition 1.2:** The element  $0$  is unique.

**Proof:** Suppose there were two elements  $0, 0'$  satisfying  $v + 0 = v + 0' = v$  for all  $v \in V$ . Then  $0 + 0' = 0$ , but  $0 + 0' = 0' + 0 = 0'$ , so  $0 = 0'$ .

**Proposition 1.3:** For each  $v \in V$ ,  $-v$  is unique.

**Proof:** Suppose there were two elements  $-v, (-v)'$  satisfying  $v + (-v) = v + (-v)' = 0$ . Then  $-v = -v + (v + (-v)')$ , so  $-v = (-v)'$ .

**Proposition 1.4:** For all  $v \in V$ ,  $0v = 0$  and  $(-1)v = -v$ .

**Proof:** We have

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

and

$$\begin{aligned} (-1)v &= (-1)v + 0 \\ &= (-1)v + v + (-v) \\ &= (-1 + 1)v + (-v) \\ &= -v. \end{aligned}$$

**Definition 1.5:** Let  $V$  be a vector space. A **subspace** of  $V$  is a nonempty set  $U \subseteq V$  such that

1.  $u + v \in U$  for all  $u, v \in U$ .
2.  $cu \in U$  for all  $u \in U$  and  $c \in k$ .

**Definition 1.6:** Let  $V$  be a vector space and  $U$  and  $W$  subspaces of  $V$ . The sum of  $U$  and  $W$  is  $U + W = \{u + w \mid u \in U, w \in W\}$ . If each element of  $V$  can be expressed uniquely as an element of  $U + W$ , we say that  $U + W$  is a **direct sum**, and we write  $U \oplus W$ .

**Proposition 1.7:**  $U + W$  is a direct sum if and only if the only expression of 0 in  $U + W$  is  $0 + 0$ .

**Proof:** ( $\Rightarrow$ ) If  $U + W$  is direct, then since  $0 = 0 + 0$  is one expression of 0, it must be the only one.

( $\Leftarrow$ ) Let  $v \in U + W$  and suppose  $v = u + w = u' + w'$  for some  $u, u' \in U$  and  $w, w' \in W$ . Then  $0 = v - v = (u + w) - (u' + w') = (u - u') + (w - w')$ . Thus  $u - u' = w - w' = 0$ , so  $u = u'$  and  $w = w'$ .

**Proposition 1.8:**  $U + W$  is a direct sum if and only if  $U \cap W = \{0\}$ .

**Proof:** ( $\Rightarrow$ ) Suppose  $U + W$  is direct and let  $v \in U \cap W$ . Then  $0 = v + (-v)$ , so by the previous result,  $v = -v = 0$ .

( $\Leftarrow$ ) Assume  $U \cap W = \{0\}$  and suppose  $u + w = 0$ . Then  $u = -w \in W$ , so  $u \in U \cap W$  and is therefore 0. Thus  $u = w = 0$ , so the previous proposition gives that  $U + W$  is direct.

## II — Bases and Dimension

**Definition 2.1:** Vectors  $v_1, \dots, v_n \in V$  are **linearly independent** if  $c_1v_1 + \dots + c_nv_n = 0$  for  $c_i \in k$  implies  $c_1 = \dots = c_n = 0$ , and **linearly dependent** if not (i.e.  $c_1v_1 + \dots + c_nv_n = 0$  for some  $c_i \in k$  not all zero).

**Definition 2.2:** The **span** of  $v_1, \dots, v_n \in V$  is the set  $\text{span}\{v_1, \dots, v_n\} = \{c_1v_1 + \dots + c_nv_n \mid c_i \in k\}$ . A set of vectors  $\{v_1, \dots, v_n\} \subseteq V$  **spans**  $V$  if  $\text{span}\{v_1, \dots, v_n\} = V$ .

**Proposition 2.3:** Let  $v_1, \dots, v_n \in V$ . Then  $\text{span}\{v_1, \dots, v_n\}$  is a subspace of  $V$ .

**Proof:** First,  $0 \in \text{span}\{v_1, \dots, v_n\}$ , so the set is nonempty. Next,  $(c_1v_1 + \dots + c_nv_n) + (d_1v_1 + \dots + d_nv_n) = (c_1 + d_1)v_1 + \dots + (c_n + d_n)v_n \in \text{span}\{v_1, \dots, v_n\}$ , so the set is closed under addition, and finally,  $c(c_1v_1 + \dots + c_nv_n) = (cc_1)v_1 + \dots + (cc_n)v_n \in \text{span}\{v_1, \dots, v_n\}$ , so it is closed under scalar multiplication. Thus  $\text{span}\{v_1, \dots, v_n\}$  is a subspace of  $V$ .

**Definition 2.4:** A **basis** for a vector space  $V$  is a set of vectors  $\{v_1, \dots, v_n\} \subseteq V$  that are linearly independent and span  $V$ .

**Definition 2.5:** A vector space is **finite-dimensional** if it has a finite basis.

**Theorem 2.6:** Let  $V$  be a finite-dimensional vector space. If  $v_1, \dots, v_k \in V$  are linearly independent, then they can be extended to form a basis for  $V$ .

**Proof:** Suppose  $\text{span}\{w_1, \dots, w_n\} = V$ . Then  $\text{span}\{v_1, \dots, v_k, w_1, \dots, w_n\} = V$ , so if  $v_1, \dots, v_k, w_1, \dots, w_n$  are linearly independent,  $\{v_1, \dots, v_k, w_1, \dots, w_n\}$  is a basis. Otherwise,  $c_1v_1 + \dots + c_kv_k + d_1w_1 + \dots + d_nw_n = 0$  for some  $c_i, d_i \in k$ . Not all the  $d_i$  can be zero, since then  $v_1, \dots, v_k$  would be linearly dependent, so there is a  $d_j \neq 0$ . Then  $\text{span}\{v_1, \dots, v_k, w_1, \dots, w_{j-1}, w_{j+1}, \dots, w_n\} = \text{span}\{v_1, \dots, v_k, w_1, \dots, w_n\}$ , since we can create  $w_j$  from the other vectors. Continue removing  $w_i$  until the set is linearly independent (this will terminate, since at most we will have  $v_1, \dots, v_k$  once again).

**Theorem 2.7:** Every basis for a finite-dimensional vector space has the same number of elements.

**Proof:** Let  $\{v_1, \dots, v_n\}$  and  $\{w_1, \dots, w_m\}$  be bases for  $V$ . Then by definition,  $v_1 = c_1w_1 + \dots + c_mw_m$  for some  $c_i \in k$  not all zero. Without loss of generality, assume  $c_1 \neq 0$ . Then  $w_1 = (-c_1)^{-1}(-v_1 + c_2w_2 + \dots + c_mw_m) \in \text{span}\{v_1, w_2, \dots, w_m\}$ , so  $\text{span}\{v_1, w_2, \dots, w_m\} = \text{span}\{w_1, \dots, w_m\} = V$ . Repeat this process until we have  $V = \text{span}\{v_1, \dots, v_n, w_{n+1}, \dots, w_m\}$ . If  $n > m$ , then  $V = \text{span}\{v_1, \dots, v_m\}$ , but then  $v_1, \dots, v_n$  are not linearly independent. Thus  $n \leq m$ , and repeating the proof by eliminating the  $v_i$  gives that  $m \leq n$ , so  $n = m$ .

---

**Definition 2.8:** Let  $V$  be a vector space. The **dimension** of  $V$ , denoted  $\dim V$ , is the number of elements in a basis for it.

**Proposition 2.9:** If  $v_1, \dots, v_n \in V$  are linearly independent and  $\dim V = n$ , then  $\{v_1, \dots, v_n\}$  is a basis for  $V$ .

**Proof:** Suppose not. Extend  $\{v_1, \dots, v_n\}$  to form a basis for  $V$ . But then that basis would have more than  $n$  elements.  $\nexists$

**Proposition 2.10:** If  $v_1, \dots, v_n \in V$  span  $V$  and  $\dim V = n$ , then  $\{v_1, \dots, v_n\}$  is a basis for  $V$ .

---

### III — Linear Maps

**Definition 3.1:** Let  $V$  and  $W$  be vector spaces. A **linear map** from  $V$  to  $W$  is a function  $T : V \rightarrow W$  such that

1. For all  $u, v \in V$ ,  $T(u + v) = Tu + Tv$ .
2. For all  $u \in V$  and  $c \in k$ ,  $T(cu) = cTu$ .

We write  $Tu$  to mean  $T(u)$ . The set of all linear maps from  $V$  to  $W$  is denoted  $\mathcal{L}(V, W)$ .

**Proposition 3.2:**  $\mathcal{L}(V, W)$  is a vector space under function addition and composition.

**Proposition 3.3:** Let  $\{v_1, \dots, v_n\}$  be a basis for  $V$  and let  $w_1, \dots, w_n \in W$ . Then there is a unique linear map  $T \in \mathcal{L}(V, W)$  such that  $Tv_i = w_i$  for each  $i$ .

**Proof:** Such a  $T$  exists, since we can define it by  $T(c_1v_1 + \dots + c_nv_n) = c_1w_1 + \dots + c_nw_n$ , and it follows that every linear map  $S$  with  $Sv_i = w_i$  is equal to  $T$  by the properties of linear maps.

**Definition 3.4:** Let  $T \in \mathcal{L}(V, W)$ . The **null space** of  $T$  is the set  $\text{null } T = \{x \in V \mid Tx = 0\}$ , and the **range** of  $T$  is the set  $\text{range } T = \{Tv \mid v \in V\}$ .

**Proposition 3.5:** Let  $T \in \mathcal{L}(V, W)$ . Then  $\text{null } T$  is a subspace of  $V$  and  $\text{range } T$  is a subspace of  $W$ .

**Theorem 3.6: (The Fundamental Theorem of Linear Maps)** Let  $T \in \mathcal{L}(V, W)$ . Then  $\dim V = \dim \text{null } T + \dim \text{range } T$ .

**Proof:** Let  $\{v_1, \dots, v_k\}$  be a basis for  $\text{null } T$  and extend it to  $\{v_1, \dots, v_n\}$  to form a basis for  $V$ . We claim that  $\{Tv_{k+1}, \dots, Tv_n\}$  is a basis for  $\text{range } T$ .

Suppose  $c_{k+1}Tv_{k+1} + \dots + c_nTv_n = 0$ . Then  $T(c_{k+1}v_{k+1} + \dots + c_nv_n) = 0$ , so  $c_{k+1}v_{k+1} + \dots + c_nv_n \in \text{null } T$ . Since  $\{v_1, \dots, v_k\}$  is a basis for  $\text{null } T$ ,  $c_{k+1}v_{k+1} + \dots + c_nv_n = c_1v_1 + \dots + c_kv_k$  for some  $c_1, \dots, c_k \in k$ . Since  $v_1, \dots, v_n$  are linearly independent,  $c_1 = \dots = c_n = 0$ , so in particular,  $c_{k+1} = \dots + c_n = 0$ . Thus  $Tv_{k+1}, \dots, Tv_n$  are linearly independent.

Let  $w \in \text{range } T$ . Then  $T(c_1v_1 + \dots + c_nv_n) = w$  for some  $c_1, \dots, c_n \in k$ , and since  $c_1v_1 + \dots + c_kv_k \in \text{null } T$ ,  $T(c_1v_1 + \dots + c_kv_k) = 0$ , so  $T(c_{k+1}v_{k+1} + \dots + c_nv_n) = w$ . Then  $w = c_{k+1}Tv_{k+1} + \dots + c_nTv_n$ , so  $w \in \text{span}\{Tv_{k+1}, \dots, Tv_n\}$ . Thus  $Tv_{k+1}, \dots, Tv_n$  span  $\text{range } T$ .

Thus  $\{Tv_{k+1}, \dots, Tv_n\}$  is a basis for  $\text{range } T$ , so in particular,  $\dim \text{range } T = n - k$  and  $\dim V = n = \dim \text{null } T + \dim \text{range } T = k + (n - k)$ .

**Proposition 3.7:** A linear map  $T \in \mathcal{L}(V, W)$  is injective if and only if  $\text{null } T = \{0\}$ .

**Proof:** ( $\Rightarrow$ ) Let  $x \in \text{null } T$ . Then  $Tx = T0 = 0$ , so  $x = 0$ , since  $T$  is injective.

( $\Leftarrow$ ) Suppose  $Tu = Tv$  for  $u, v \in V$ . Then  $T(u - v) = 0$ , so  $u - v = 0$ , since  $\text{null } T = \{0\}$ . Thus  $T$  is injective.

**Proposition 3.8:** Let  $T \in \mathcal{L}(V, W)$ . If  $\dim V > \dim W$ , then  $\text{null } T \neq 0$ .

**Proof:**  $\dim \text{null } T = \dim V - \dim \text{range } T \geq \dim V - \dim W > 0$ .

**Definition 3.9:** Let  $T \in \mathcal{L}(V, W)$ . An **inverse linear map** to  $T$  is a  $T^{-1} \in \mathcal{L}(V, W)$  such that  $T^{-1}T = I_V$  and  $TT^{-1} = I_W$ . If such a  $T^{-1}$  exists, we call  $T$  **invertible**.

**Proposition 3.10:** Let  $T \in \mathcal{L}(V, W)$ . Then  $T^{-1}$  is unique.

**Proof:** Suppose  $T_1^{-1}$  and  $T_2^{-1}$  are both inverses to  $T$ . Then  $T_1^{-1} = T_1^{-1}TT_2^{-1} = T_2^{-1}$ .

**Definition 3.11:** An **isomorphism** from  $V$  to  $W$  is an invertible linear map  $T \in \mathcal{L}(V, W)$ .  $V$  and  $W$  are **isomorphic**, denoted  $V \simeq W$ , if there is an isomorphism from  $V$  to  $W$ .

**Proposition 3.12:** If  $\dim V = \dim W$ , then  $V \simeq W$ .

**Proof:** Let  $\{v_1, \dots, v_n\}$  and  $\{w_1, \dots, w_n\}$  be bases for  $V$  and  $W$  and define  $T \in \mathcal{L}(V, W)$  by  $Tv_i = w_i$ . Then  $T$  is injective and surjective, so it is an isomorphism.

**Definition 3.13:** Let  $\{v_1, \dots, v_n\}$  and  $\{w_1, \dots, w_m\}$  be bases for  $V$  and  $W$  and let  $T \in \mathcal{L}(V, W)$ . The **matrix** of  $T$  with respect to the chosen bases is the  $m \times n$  rectangle of numbers

$$M(T) = \begin{bmatrix} | & & | \\ Tv_1 & \cdots & Tv_n \\ | & & | \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix}$$

where  $Tv_i = c_{1i}w_1 + \cdots + c_{mi}w_m$ . Notice that if  $v = c_1v_1 + \cdots + c_nv_n$ , then

$$M(T) \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} | & & | \\ Tv_1 & \cdots & Tv_n \\ | & & | \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1Tv_1 + \cdots + c_nTv_n = Tv.$$

**Definition 3.14:** Let  $A \in M_{m,n}(k)$  and  $v_1, \dots, v_k \in k^n$ . We define **matrix multiplication** by

$$A \begin{bmatrix} | & & | \\ v_1 & \cdots & v_k \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ Av_1 & \cdots & Av_k \\ | & & | \end{bmatrix}.$$

Notice that if  $T \in \mathcal{L}(V, W)$  and  $S \in \mathcal{L}(W, U)$ , then  $M(S)M(T) = M(ST)$ .

**Theorem 3.15:** Let  $V$  and  $W$  be vector spaces. Then  $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$ .

**Proof:** Let  $\{v_1, \dots, v_n\}$  and  $\{w_1, \dots, w_m\}$  be bases for  $V$  and  $W$ . We claim that  $\mathcal{L}(V, W) \simeq M_{n,m}(k)$ . Define  $M : \mathcal{L}(V, W) \rightarrow M_{n,m}(k)$  by sending a linear map to its matrix.

( $\hookrightarrow$ ) Suppose  $M(T) = M(S)$ . Then the columns of each are equal, so  $Tv_i = Sv_i$  for all  $i$ . Thus  $T = S$ , so  $M$  is injective.

( $\rightarrow$ ) Let  $A \in M_{n,m}(k)$  and define  $T \in \mathcal{L}(V, W)$  by  $Tv_i = c_1w_1 + \cdots + c_mw_m$ , where the  $c_j$  form the  $i$ th column of  $A$ . Then  $M(T) = A$ , so  $M$  is surjective.

**Definition 3.16:** Let  $V$  be a vector space. The **dual space** to  $V$  is the vector space  $V' = \mathcal{L}(V, k)$ . By the previous theorem,  $\dim V' = \dim V$ .

**Definition 3.17:** Let  $\{v_1, \dots, v_n\}$  be a basis for  $V$ . The dual basis to  $\{v_1, \dots, v_n\}$  is  $\{\varphi_{v_1}, \dots, \varphi_{v_n}\}$ , where

$$\varphi_{v_i}(v_j) = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases}.$$

**Proposition 3.18:** Let  $\{v_1, \dots, v_n\}$  be a basis for  $V$ . Then  $\{\varphi_{v_1}, \dots, \varphi_{v_n}\}$  is a basis for  $V'$ .

**Definition 3.19:** Let  $T \in \mathcal{L}(V, W)$ . The dual map  $T' \in \mathcal{L}(W', V')$  is defined by  $T'\varphi = \varphi T$ .

**Theorem 3.20:** Let  $T \in \mathcal{L}(V, W)$ . Then  $M(T') = M(T)^T$ .

**Proof:**  $M(T')_{ij}$  is the coefficient of  $\varphi_{v_i}$  in  $T'\varphi_{w_j} = \varphi_{w_j}T$ . If  $\varphi_{w_j}T = c_1\varphi_{v_1} + \dots + c_n\varphi_{v_n}$ , then  $M(T')_{ij} = c_i$ . But by the definition of  $\varphi_{w_j}$ ,  $\varphi_{w_j}Tv_i$  is the coefficient of  $w_j$  in the expression of  $Tv_i$ , which is the definition of  $M(T)_{ji}$ . Thus  $M(T)_{ji} = M(T')_{ij}$ , so  $M(T') = M(T)^T$ .

**Definition 3.21:** Let  $U \subseteq V$  (not necessarily a subspace). The **annihilator** of  $U$  is the set  $U^0 = \{\varphi \in V' \mid \varphi u = 0 \text{ for all } u \in U\}$ .

**Proposition 3.22:** Let  $U$  be a subspace of  $V$ . Then  $\dim V = \dim U + \dim U^0$ .

**Proof:** Let  $\{v_1, \dots, v_k\}$  be a basis for  $U$  and extend it to  $\{v_1, \dots, v_n\}$  to form a basis for  $V$ . Then  $\{\varphi_{v_1}, \dots, \varphi_{v_n}\}$  is a basis for  $V'$ , so  $\varphi_{v_{k+1}}, \dots, \varphi_{v_n}$  are linearly independent. Since  $\text{span}\{\varphi_{v_{k+1}}, \dots, \varphi_{v_n}\} = U^0$ ,  $\dim U^0 = n - k$ , so  $\dim V = n = \dim U + \dim U^0 = k + (n - k)$ .

**Proposition 3.23:** Let  $T \in \mathcal{L}(V, W)$ . Then  $\text{null } T' = (\text{range } T)^0$ .

**Proof:** We have  $\varphi \in \text{null } T'$  if and only if  $T'\varphi = 0$ , if and only if  $\varphi T = 0$ , if and only if  $\varphi Tv = 0$  for all  $v \in V$ , if and only if  $\varphi w = 0$  for all  $w \in \text{range } T$ , if and only if  $\varphi \in (\text{range } T)^0$ .

**Proposition 3.24:** Let  $T \in \mathcal{L}(V, W)$ . Then  $\text{range } T' = (\text{null } T)^0$ .

**Proposition 3.25:** Let  $T \in \mathcal{L}(V, W)$ . Then  $T'$  is injective if and only if  $T$  is surjective.

**Proof:**  $T'$  is injective if and only if  $\text{null } T' = \{0\}$ , if and only if  $(\text{range } T)^0 = \{0\}$ , if and only if  $\dim (\text{range } T)^0 = 0$ , if and only if  $\dim \text{range } T = \dim W$ , if and only if  $\text{range } T = W$ .

**Corollary 3.25.1:** Let  $T \in \mathcal{L}(V, W)$ . Then  $\dim \text{range } T' = \dim \text{range } T$ .

## IV — Eigenvalues and Eigenvectors

**Definition 4.1:** An **eigenvalue** of a linear map  $T \in \mathcal{L}(V) = \mathcal{L}(V, V)$  is an element  $\lambda \in k$  such that  $Tv = \lambda v$  for some nonzero  $v \in V$ . This  $v$  is called the **eigenvector** corresponding to  $\lambda$ .

**Proposition 4.2:** Let  $T \in \mathcal{L}(V)$  and  $\lambda \in k$ . Then  $\lambda$  is an eigenvalue of  $T$  if and only if  $T - \lambda I$  is not invertible.

**Proof:** We have that  $\lambda$  is an eigenvalue of  $T$  if and only if  $Tv = \lambda v$  for some  $v \neq 0$ , if and only if  $(T - \lambda I)v = 0$  for some  $v \neq 0$ , if and only if  $\text{null } (T - \lambda I) \neq \{0\}$ , if and only if  $T - \lambda I$  is not invertible.

**Theorem 4.3:** If  $\lambda_1, \dots, \lambda_k$  are distinct eigenvalues of  $T$ , then the corresponding eigenvectors  $v_1, \dots, v_k$  are linearly independent.

**Proof:** Suppose not. Then there is a minimum  $j$  for which  $v_1, \dots, v_j$  are linearly dependent, so  $v_j = c_1 v_1 + \dots + c_{j-1} v_{j-1}$  for some  $c_1, \dots, c_{j-1} \in k$ . Then

$$\lambda_j v_j = \lambda_j (c_1 v_1 + \dots + c_{j-1} v_{j-1}) = c_1 \lambda_j v_1 + \dots + c_{j-1} \lambda_j v_{j-1}.$$

But we also have

$$\lambda_j v_j = T v_j = T(c_1 v_1 + \dots + c_{j-1} v_{j-1}) = c_1 \lambda_1 v_1 + \dots + c_{j-1} \lambda_{j-1} v_{j-1},$$

so

$$c_1 (\lambda_1 - \lambda_j) v_1 + \dots + c_{j-1} (\lambda_{j-1} - \lambda_j) v_{j-1} = 0.$$

Since  $j$  was minimal,  $v_1, \dots, v_{j-1}$  are linearly independent, so  $c_i (\lambda_i - \lambda_j) = 0$  for all  $i \in \{1, \dots, j-1\}$ . Not every  $c_i = 0$ , since then  $v_j = 0$ , so some  $c_i \neq 0$ , and therefore  $\lambda_i = \lambda_j$ . But then the eigenvalues are not distinct.  $\nmid$



**Definition 4.4:** A linear map  $T \in \mathcal{L}(V)$  is **diagonalizable** if there is a basis of eigenvectors of  $T$  for  $V$  — that is, a basis such that

$$M(T) = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

**Proposition 4.5:** If  $\dim V = n$  and  $T \in \mathcal{L}(V)$  has  $n$  distinct eigenvalues, then  $T$  is diagonalizable.

**Definition 4.6:** A matrix  $A \in M_n(k)$  is **upper triangular** if it has the form

$$A = \begin{bmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & * & * \\ 0 & 0 & \cdots & 0 & * \end{bmatrix},$$

where the  $*$  are elements of  $k$ .

**Definition 4.7:** Let  $T \in \mathcal{L}(V)$ . A subspace  $U$  of  $V$  is **T-invariant** if  $Tu \in U$  for all  $u \in U$ .

**Theorem 4.8:** Let  $V$  be a vector space over an algebraically closed field  $k$  with  $\dim V = n$  and let  $T \in \mathcal{L}(V)$ . Then there is a basis for  $V$  such that  $M(T)$  is upper triangular.

**Proof:** We will proceed by induction. The base case is trivial, since every  $1 \times 1$  matrix is upper triangular.

Assume that every linear map in  $\mathcal{L}(V)$  has such a basis if  $\dim V < n$ . Let  $T \in \mathcal{L}(V)$  and let  $\lambda$  be an eigenvalue of  $T$  (This exists, since we can choose any basis for  $V$  and perform elementary row operations on  $M(T)$  to eliminate every element of a non-leading-zero column below the top one). Let  $U = \text{range } (T - \lambda I)$ . Then  $U$  is  $T$ -invariant, since  $T(Tv - \lambda v) = T(Tv) - \lambda(Tv) \in U$ , so  $T|_U \in \mathcal{L}(U)$ . Since the eigenvector corresponding to  $\lambda$  is an element of  $\text{null } (T - \lambda I)$ ,  $U \neq V$ . Thus  $\dim U < \dim V$ , so by assumption, there is a basis  $\{u_1, \dots, u_k\}$  for  $U$  such that  $M(T|_U)$  is upper triangular. Extend this to  $\{u_1, \dots, u_k, v_1, \dots, v_j\}$  to form a basis for  $V$ . Then  $Tv_i = Tv_i - \lambda v_i + \lambda v_i = c_1 u_1 + \cdots + c_k u_k + \lambda v_i$  for some  $c_1, \dots, c_k \in k$ , and so

$$M(T) = \begin{bmatrix} T & * \\ 0 & \lambda I \end{bmatrix},$$

where  $T$  is a  $k \times k$  upper triangular matrix,  $*$  is unspecified,  $0$  is the zero matrix, and  $\lambda I$  is a  $j \times j$  diagonal matrix. Thus  $M(T)$  is upper triangular.

**Theorem 4.9:** If  $M(T)$  is upper triangular with respect to the basis  $v_1, \dots, v_n$  and has diagonal entries  $\lambda_1, \dots, \lambda_n$ , then  $T$  is invertible if and only if no  $\lambda_i = 0$ .

**Proof:** ( $\Rightarrow$ ) Assume  $T^{-1}$  exists and suppose some  $\lambda_i = 0$ . Let  $U = \text{span}\{v_1, \dots, v_i\}$ . Then  $U$  is  $T$ -invariant, but  $T|_U$  is not surjective, so it is not invertible, and therefore neither is  $T$ .  $\nexists$

( $\Leftarrow$ ) It is enough to show that  $\text{null } T = \{0\}$ , so suppose  $T(c_1v_1 + \dots + c_nv_n) = 0$ . Then  $c_1Tv_1 + \dots + c_nTv_n = 0$ . Since  $Tv_i \in \text{span}\{v_1, \dots, v_i\}$ ,  $c_n = 0$ , since  $v_n$  appears only in  $Tv_n$  and  $\lambda_n \neq 0$ . Similarly,  $c_1 = \dots = c_{n-1} = 0$ . Thus  $\text{null } T = \{0\}$ .

**Theorem 4.10:** If  $M(T)$  is upper triangular with diagonal entries  $\lambda_1, \dots, \lambda_n$ , then  $T$  has eigenvalues  $\lambda_1, \dots, \lambda_n$ .

**Proof:** If  $\lambda$  is an eigenvalue of  $T$ , then  $T - \lambda I$  is not invertible. Then  $\lambda_i - \lambda = 0$  for some  $i$ , since

$$M(T - \lambda I) = \begin{bmatrix} \lambda_1 - \lambda & * & \cdots & * \\ 0 & \lambda_2 - \lambda & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n - \lambda \end{bmatrix}.$$

Repeat for all  $i$ .

## V — Inner Product Spaces

**Definition 5.1:** Let  $V$  be a vector space over  $k = \mathbb{R}$  or  $\mathbb{C}$ . An **inner product** on  $V$  is a function  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$  such that

1.  $\langle v, v \rangle \in \mathbb{R}^+$  for all nonzero  $v \in V$  and  $\langle v, v \rangle = 0$  if and only if  $v = 0$ .
2.  $\langle cu + v, w \rangle = c\langle u, w \rangle + \langle v, w \rangle$  for all  $u, v, w \in V$  and  $c \in k$ .
3.  $\langle u, v \rangle = \overline{\langle v, u \rangle}$  for all  $u, v \in V$ .

An **inner product space** is a vector space equipped with an inner product.

**Definition 5.2:** The **norm** of an element  $v \in V$  is  $\|v\| = \sqrt{\langle v, v \rangle}$ .

**Definition 5.3:** The **distance** between two vectors  $u, v \in V$  is  $\|u - v\|$ .

**Proposition 5.4:** For all  $v \in V$  and  $c \in k$ ,  $\|cv\| = |c| \cdot \|v\|$ .

**Proof:** Since  $\|cv\|^2 = \langle cv, cv \rangle = c\bar{c}\langle v, v \rangle = |c|^2\|v\|^2$ ,  $\|cv\| = |c| \cdot \|v\|$ .

**Definition 5.5:** Two vectors  $u, v \in V$  are **orthogonal** if  $\langle u, v \rangle = 0$ .

**Proposition 5.6: (The Pythagorean Theorem)** Let  $u, v \in V$  be orthogonal. Then  $\|u + v\|^2 = \|u\|^2 + \|v\|^2$ .

**Proof:** We have  $\|u + v\|^2 = \langle u + v, u + v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle = \langle u, u \rangle + \langle v, v \rangle = \|u\|^2 + \|v\|^2$ .

**Proposition 5.7: (The Cauchy-Schwarz Inequality)** For all  $u, v \in V$ ,  $\|u\| \cdot \|v\| \geq |\langle u, v \rangle|$ .

**Proof:** Let  $c = \frac{\langle u, v \rangle}{\langle v, v \rangle}$ . Then  $\|u\|^2\|v\|^2 = \|u - cv + cv\|^2\|v\|^2$ , and since  $u - cv$  is orthogonal to  $cv$ ,  $\|u\|^2\|v\|^2 = (\|u - cv\|^2 + \|cv\|^2)\|v\|^2 \geq \|cv\|^2\|v\|^2 = |c|^2\|v\|^4 = |\langle u, v \rangle|^2$ .

**Proposition 5.8:** For all  $u, v \in V$ ,  $\|u\| + \|v\| \geq \|u + v\|$ .

**Lemma 5.8.1:** For all  $z \in \mathbb{C}$ ,  $2|z| \geq z + \bar{z}$ .

**Proof:** If  $z = a + bi$ , then  $2|z| = 2|a + bi| = 2\sqrt{a^2 + b^2} \geq 2\sqrt{a^2} = 2a = z + \bar{z}$ .

**Proof:** We have  $(\|u\| + \|v\|)^2 = \|u\|^2 + 2\|u\| \cdot \|v\| + \|v\|^2 \geq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \geq \|u\|^2 + \langle u, v \rangle + \overline{\langle u, v \rangle} + \|v\|^2 = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle = \langle u + v, u + v \rangle = \|u + v\|^2$ .

**Definition 5.9:** Vectors  $e_1, \dots, e_k \in V$  are **orthonormal** if  $\|e_i\| = 1$  for all  $i$  and  $\langle e_i, e_j \rangle = 0$  for all  $i \neq j$ .

**Proposition 5.10:** If  $e_1, \dots, e_k \in V$  are orthonormal, then  $\|c_1e_1 + \dots + c_ke_k\|^2 = |c_1|^2 + \dots + |c_k|^2$ .

**Proof:** We will induct upon  $k$ . The base case is obvious, since  $\|c_1e_1\|^2 = |c_1|^2\|e_1\|^2 = |c_1|^2$ . For the induction step, assume  $\|c_1e_1 + \dots + c_ke_k\|^2 = |c_1|^2 + \dots + |c_k|^2$ . Since  $c_1e_1 + \dots + c_ke_k$  and  $c_{k+1}e_{k+1}$  are orthogonal,  $\|c_1e_1 + \dots + c_{k+1}e_{k+1}\|^2 = \|c_1e_1 + \dots + c_ke_k\|^2 + \|c_{k+1}e_{k+1}\|^2 = |c_1|^2 + \dots + |c_k|^2 + |c_{k+1}|^2$ .

**Proposition 5.11:** Orthonormal vectors are linearly independent.

**Proof:** Suppose  $e_1, \dots, e_k \in V$  are orthonormal and  $c_1 e_1 + \dots + c_k e_k = 0$ . Then  $\|c_1 e_1 + \dots + c_k e_k\|^2 = |c_1|^2 + \dots + |c_k|^2 = 0$ , so  $c_1 = \dots = c_k = 0$ .

**Proposition 5.12:** Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis for  $V$  and let  $v \in V$ . Then  $v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n$ .

**Proof:** If  $v = c_1 e_1 + \dots + c_n e_n$ , then  $\langle v, e_i \rangle = \langle c_1 e_1 + \dots + c_n e_n, e_i \rangle = \langle c_i e_i, e_i \rangle = c_i$ .

**Theorem 5.13: (The Gram-Schmidt Process)** Every finite-dimensional inner product space has an orthonormal basis.

**Proof:** Let  $\{v_1, \dots, v_n\}$  be a basis for  $V$ . Let  $e'_1 = v_1$  and  $e_1 = \frac{e'_1}{\|e'_1\|}$ . Then for each  $i \in \{2, \dots, n\}$ , let

$$e'_i = v_i - (\langle v_i, e_1 \rangle e_1 + \dots + \langle v_i, e_{i-1} \rangle e_{i-1})$$

and  $e_i = \frac{e'_i}{\|e'_i\|}$ . Then  $\{e_1, \dots, e_n\}$  is an orthonormal basis for  $V$ .

**Theorem 5.14: (Riesz Representation)** Let  $\varphi_u \in V'$  be defined by  $\varphi_u v = \langle v, u \rangle$ . Then for each  $T \in V'$ , there is a unique  $u \in V$  such that  $T = \varphi_u$ .

**Proof:** Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis for  $V$  and let  $u = \overline{T e_1} e_1 + \dots + \overline{T e_n} e_n$ . Then if  $v = c_1 e_1 + \dots + c_n e_n$ ,

$$\begin{aligned} \varphi_u v &= \langle v, u \rangle \\ &= \langle c_1 e_1 + \dots + c_n e_n, \overline{T e_1} e_1 + \dots + \overline{T e_n} e_n \rangle \\ &= c_1 \overline{T e_1} + \dots + c_n \overline{T e_n} \\ &= T(c_1 e_1 + \dots + c_n e_n) \\ &= T v. \end{aligned}$$

**Definition 5.15:** Let  $U \subseteq V$ . The **orthogonal complement** to  $U$  is the set  $U^\perp = \{v \in V \mid \langle u, v \rangle = 0 \text{ for all } u \in U\}$ .

**Proposition 5.16:** If  $U$  is a subspace of  $V$ , then so is  $U^\perp$ .

**Theorem 5.17:** If  $U$  is a finite-dimensional subspace of  $V$ , then  $V = U \oplus U^\perp$ .

**Proof:** Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis for  $U$ , let  $v \in V$ , and let  $u = c_1e_1 + \dots + c_ne_n \in U$ . Then  $v = u + (v - u)$ . If  $v - u \in U^\perp$ , this will be an expression of  $v$  in  $U + U^\perp$ . For  $v - u$  to be in  $U^\perp$ ,  $\langle v - u, e_i \rangle = 0$  for all  $i$ , so  $c_i = \langle u, e_i \rangle = \langle v, e_i \rangle$  for all  $i$ . Thus  $u$  is completely determined by  $v$ , so the expression of  $v$  as  $u + (v - u)$  is unique. Thus  $V = U \oplus U^\perp$ .

**Corollary 5.17.1:** If  $U$  is a finite-dimensional subspace of  $V$ , then  $\dim V = \dim U + \dim U^\perp$ .

**Proposition 5.18:** Let  $U \subseteq V$ . Then  $(U^\perp)^\perp = U$ .

**Proof:** Let  $u \in U$  and  $v \in U^\perp$ . Then  $\langle u, v \rangle = 0$  by definition, so  $u \in (U^\perp)^\perp$ . Thus  $U \subseteq (U^\perp)^\perp$ . Also,  $\dim U + \dim U^\perp = \dim V = \dim U^\perp + \dim (U^\perp)^\perp$ , so  $\dim U = \dim (U^\perp)^\perp$ . Thus  $U = (U^\perp)^\perp$ .

**Definition 5.19:** The **projection** of  $V$  onto a subspace  $U$  is the linear map  $P_U \in \mathcal{L}(V, U)$  given by  $P_U v = u$ , where  $v = u + u' \in U \oplus U^\perp$ .

**Proposition 5.20:** Let  $U$  be a subspace of a vector space  $V$  with  $\dim U = k$  and  $\dim V = n$ , and let  $\{u_1, \dots, u_k, u'_{k+1}, \dots, u'_n\}$  be a basis for  $V$  composed of bases for  $U$  and  $U^\perp$ . Then

$$M(P_U) = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

**Theorem 5.21:** Let  $V$  be a vector space and  $U$  a subspace. Then for all  $v \in V$  and  $u \in U$ ,  $\|v - P_U v\| \leq \|v - u\|$ ; that is, the closest vector to  $v$  in  $U$  is  $P_U v$ .

**Proof:** Since  $v - P_U v \notin U$ ,  $v - P_U v \in U^\perp$ , so  $v - P_U v$  and  $P_U v - u$  are orthogonal. Then  $\|v - u\|^2 = \|v - P_U v + P_U v - u\|^2 = \|v - P_U v\|^2 + \|P_U v - u\|^2 \geq \|v - P_U v\|^2$ .

## VI — Linear Maps and Inner Products

**Definition 6.1:** Let  $T \in \mathcal{L}(V, W)$ . The **adjoint** of  $T$  is the linear map  $T^* \in \mathcal{L}(W, V)$  such that  $\langle Tv, w \rangle_W = \langle v, T^*w \rangle_V$  for all  $v \in V$  and  $w \in W$ .

**Proposition 6.2:** Let  $T \in \mathcal{L}(U, V)$ ,  $S \in \mathcal{L}(V, W)$ , and  $c \in k$ . Then

1.  $(cT + S)^* = \bar{c}T^* + S^*$ .
2.  $(T^*)^* = T$ .
3.  $I^* = I$ .
4.  $(ST)^* = T^*S^*$ .

**Theorem 6.3:** Let  $\{e_1, \dots, e_n\}$  and  $\{f_1, \dots, f_m\}$  be orthonormal bases for  $V$  and  $W$  and let  $T \in \mathcal{L}(V, W)$ . Then  $M(T^*) = \overline{M(T)}^T$ .

**Proof:** The  $j$ th column of  $M(T^*)$  is  $T^*f_j$  expressed in the basis  $\{e_1, \dots, e_n\}$ . Since this is orthonormal,  $T^*f_j = \langle T^*f_j, e_1 \rangle e_1 + \dots + \langle T^*f_j, e_n \rangle e_n$ , so  $M(T^*)_{ij} = \langle T^*f_j, e_i \rangle$ . But  $M(T)_{ji} = \langle Te_i, f_j \rangle = \langle e_i, T^*f_j \rangle = \overline{\langle T^*f_j, e_i \rangle} = \overline{M(T^*)_{ij}}$ , so  $M(T^*) = \overline{M(T)}^T$ .

**Definition 6.4:** A linear map  $T \in \mathcal{L}(V)$  is **self-adjoint** if  $T^* = T$ .

**Proposition 6.5:** Let  $T \in \mathcal{L}(V)$  be self-adjoint. Then if  $\langle Tv, v \rangle = 0$  for all  $v \in V$ ,  $T = 0$ .

**Definition 6.6:** A linear map  $T \in \mathcal{L}(V)$  is **normal** if  $T^*T = TT^*$ .

**Proposition 6.7:** A linear map  $T \in \mathcal{L}(V)$  is **normal** if and only for all  $v \in V$ ,  $\|Tv\| = \|T^*v\|$ .

**Proof:** ( $\Rightarrow$ ) If  $T$  is normal, then  $\langle Tv, Tv \rangle = \langle v, T^*Tv \rangle = \langle v, TT^*v \rangle = \langle T^*v, T^*v \rangle$ .

( $\Leftarrow$ ) Suppose  $\langle Tv, Tv \rangle = \langle T^*v, T^*v \rangle$  for all  $v \in V$ . Then  $\langle TT^*v - T^*Tv, v \rangle = 0$  for all  $v \in V$ , so  $TT^* - T^*T = 0$ .

**Proposition 6.8:** If  $T$  is normal and  $Tv = \lambda v$  for some  $v \neq 0$ , then  $T^*v = \bar{\lambda}v$ .

**Proof:**  $(T - \lambda I)^*(T - \lambda I)$  is normal, since  $(T - \lambda I)^*(T - \lambda I) = T^*T - \bar{\lambda}IT - \lambda IT + \lambda\bar{\lambda}I = (T - \lambda I)(T - \lambda I)^*$ . Then  $0 = \|(T - \lambda I)v\| = \|(T - \lambda I)^*v\| = \|(T^* - \bar{\lambda})v\|$ , so  $T^*v = \bar{\lambda}v$ .

**Proposition 6.9:** Let  $T \in \mathcal{L}(V)$  be normal. If  $v$  and  $w$  are eigenvectors of  $T$  with distinct eigenvalues  $\lambda_1$  and  $\lambda_2$ , then  $v$  and  $w$  are orthogonal.

**Proof:** Since  $\lambda_1 \neq \lambda_2$ ,  $\lambda_1 - \lambda_2 \neq 0$ . Then  $(\lambda_1 - \lambda_2)\langle v, w \rangle = \langle \lambda_1 v - \lambda_2 v, w \rangle = \langle Tv, w \rangle - \langle v, \overline{\lambda_2} w \rangle = \langle Tv, w \rangle - \langle v, T^* w \rangle = \langle Tv, w \rangle - \langle Tv, w \rangle = 0$ , so  $\langle v, w \rangle = 0$ .

**Theorem 6.10: (Complex Spectral)** Let  $V$  be a finite-dimensional vector space over  $\mathbb{C}$  and let  $T \in \mathcal{L}(V)$ . Then  $T$  is normal if and only if there is an orthonormal basis of eigenvectors of  $T$  for  $V$ .

**Proof:** ( $\Rightarrow$ ) We will induct on  $n = \dim V$ . The base case is trivial, since if  $\dim V = 1$  and  $T \in \mathcal{L}(V)$ , then any nonzero unit vector in  $V$  constitutes an orthonormal basis of eigenvectors of  $T$ .

Suppose the theorem holds for  $n - 1$ -dimensional vector spaces and let  $T \in \mathcal{L}(V)$  be normal with  $\dim V = n$ . Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis for  $V$  such that  $M(T)$  is upper triangular (this is possible, since the Gram-Schmidt process preserves upper triangularity). Then we have

$$M(T) = \begin{bmatrix} \lambda_1 & *_{1,2} & \cdots & *_{1,n} \\ 0 & \lambda_2 & \cdots & *_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

where  $*_i$  is the vector of the first  $i - 1$  entries in column  $i$  of  $M(T)$ . Consider the first column of  $M(T)$  and  $M(T^*) = \overline{M(T)}^T$ . We have  $Te_1 = \lambda_1 e_1$  and  $T^*e_1 = \overline{\lambda_1}e_1 + \overline{*_{1,2}}e_2 + \cdots + \overline{*_{1,n}}e_n$ , but  $T$  is normal, so  $\|Te_1\| = \|T^*e_1\|$ . Thus  $|\lambda_1|^2 = |\overline{\lambda_1}|^2 + |\overline{*_{1,2}}|^2 + \cdots + |\overline{*_{1,n}}|^2$ , so  $|\overline{*_{1,2}}|^2 + \cdots + |\overline{*_{1,n}}|^2 = 0$ , and therefore  $*_{1,2} = \cdots = *_{1,n} = 0$ . Thus the first row and column of  $M(T)$  are zero, except for  $\lambda_1$ , and similarly for  $M(T^*)$ . By restricting  $T$  to  $\text{span}\{e_2, \dots, e_n\}$ , which has dimension  $n - 1$ , we are done by induction.

( $\Leftarrow$ ) If  $\{e_1, \dots, e_n\}$  is an orthonormal basis of eigenvectors of  $T$ , then

$$M(T)M(T^*) = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \overline{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \overline{\lambda_n} \end{bmatrix} = \begin{bmatrix} \overline{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \overline{\lambda_n} \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = M(T^*)M(T),$$

so  $T$  is normal.

**Theorem 6.11: (Real Spectral)** Let  $V$  be a finite-dimensional vector space over  $\mathbb{R}$  and let  $T \in \mathcal{L}(V)$ . Then  $T$  is self-adjoint if and only if there is an orthonormal basis of eigenvectors of  $T$  with real eigenvalues.

**Proof:** By the Complex Spectral Theorem, there is an orthonormal basis for  $V$  of eigenvectors of  $T$ . Since  $T$  is self-adjoint,  $M(T) = M(T^*) = \overline{M(T)}^T = \overline{M(T)}$ . Thus each  $\lambda_i = \overline{\lambda_i}$ , so all of  $T$ 's eigenvalues are real.

**Definition 6.12:** A linear map  $T \in \mathcal{L}(V)$  is **positive** if  $T$  is self-adjoint and  $\langle Tv, v \rangle \geq 0$  for all  $v \in V$ .

**Proposition 6.13:** Let  $T \in \mathcal{L}(V)$  be self-adjoint. Then  $T$  is positive if and only if every eigenvalue of  $T$  is nonnegative.

**Proof:** Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis for  $V$  of eigenvectors of  $T$  with eigenvalues  $\lambda_1, \dots, \lambda_n$ . Then  $T$  is positive if and only if  $\langle T(c_1e_1 + \dots + c_ne_n), c_1e_1 + \dots + c_ne_n \rangle \geq 0$  for all  $c_1, \dots, c_n \in \mathbb{C}$ , if and only if  $\langle c_1\lambda_1e_1 + \dots + c_n\lambda_ne_n, c_1e_1 + \dots + c_ne_n \rangle \geq 0$  for all  $c_1, \dots, c_n \in \mathbb{C}$ , if and only if  $|c_1|^2\lambda_1 + \dots + |c_n|^2\lambda_n \geq 0$  for all  $c_1, \dots, c_n \in \mathbb{C}$ , if and only if each  $\lambda_i \geq 0$  (for each  $i$ , choose  $c_i = 1$  and  $c_j = 0$  for  $j \neq i$ ).

**Definition 6.14:** Let  $T \in \mathcal{L}(V)$ . A **square root** of  $T$  is a linear map  $R \in \mathcal{L}(V)$  such that  $R^2 = T$ .

**Theorem 6.15:** Let  $T \in \mathcal{L}(V)$  be positive. Then there is a unique positive square root of  $T$ .

**Proof:** We will only show existence — the proof of uniqueness is difficult, tedious, and unenlightening. Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis for  $V$  of eigenvectors of  $T$  with eigenvalues  $\lambda_1, \dots, \lambda_n$ . Then each  $\lambda_i \geq 0$ , so the map  $R \in \mathcal{L}(V)$  defined by  $Re_i = \sqrt{\lambda_i}e_i$  is positive, and clearly  $R^2 = T$ . Thus  $T$  has a positive square root.

**Definition 6.16:** Let  $T \in \mathcal{L}(V)$  be positive. The unique positive square root of  $T$  is denoted  $\sqrt{T}$ .

**Definition 6.17:** A linear map  $T \in \mathcal{L}(V)$  is an **isometry** if  $\|Tv\| = \|v\|$  for all  $v \in V$ .

**Proposition 6.18:** A linear map  $T \in \mathcal{L}(V)$  is an isometry if and only if  $T^*T = I$ .

**Proof:**  $T$  is an isometry if and only if  $\langle Tv, Tv \rangle = \langle v, v \rangle$  for all  $v \in V$ , if and only if  $\langle T^*Tv, v \rangle - \langle Iv, v \rangle = 0$  for all  $v \in V$ , if and only if  $T^*T - I = 0$ , since  $T^*T - I$  is self-adjoint.

**Theorem 6.19:** A linear map  $T \in \mathcal{L}(V)$  is an isometry if and only if there is an orthonormal basis of eigenvectors of  $T$  with eigenvalues  $\lambda_1, \dots, \lambda_n$  such that  $|\lambda_i| = 1$ .

**Proof:** ( $\Rightarrow$ ) If  $T$  is an isometry, then it is normal, so there is an orthonormal basis of eigenvectors  $\{e_1, \dots, e_n\}$  with eigenvalues  $\lambda_1, \dots, \lambda_n$  by the Complex Spectral Theorem. Then  $|\lambda_i| = \|\lambda_ie_i\| = \|Te_i\| = \|e_i\| = 1$ .

( $\Leftarrow$ ) Let  $v = c_1e_1 + \dots + c_ne_n \in V$ . Then  $\|Tv\| = \|c_1\lambda_1e_1 + \dots + c_n\lambda_ne_n\| = \|c_1e_1 + \dots + c_ne_n\| = \|v\|$ , so  $T$  is an isometry.



**Theorem 6.20: (Polar Decomposition)** Let  $T \in \mathcal{L}(V)$ . Then there is an isometry  $S \in \mathcal{L}(V)$  such that  $T = S\sqrt{T^*T}$ .

**Proof:** Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis of eigenvectors of  $T^*T$  with eigenvalues  $\lambda_1, \dots, \lambda_n$  and suppose without loss of generality that  $\lambda_1 = \dots = \lambda_k = 0$ . Let  $\{f_1, \dots, f_k\}$  be an orthonormal basis for  $(\text{range } T)^\perp$  (the dimension is  $k$  since  $\dim \text{range } T = \dim \text{null } T^*$ ). Then define  $S$  by

$$Se_i = \begin{cases} f_i, & i \leq k \\ \frac{1}{\sqrt{\lambda_i}} Te_i, & i > k \end{cases}.$$

It follows that  $S$  is an isometry and  $T = S\sqrt{T^*T}$ .

**Definition 6.21:** The **singular values** of a linear map  $T \in \mathcal{L}(V, W)$  are  $\sigma_1, \dots, \sigma_k$ , where  $\sigma_i = \sqrt{\lambda_i}$  and  $\lambda_1, \dots, \lambda_k$  are the nonzero eigenvalues of  $T^*T$ .

**Theorem 6.22: (Singular Value Decomposition)** Let  $V$  and  $W$  be vector spaces with  $\dim V = n$  and  $\dim W = m$ . Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis of eigenvectors of  $T^*T$  with eigenvalues  $\lambda_1 \geq \dots \geq \lambda_k > 0 = \lambda_{k+1} = \dots = \lambda_n$ . Then there is an orthonormal basis  $\{f_1, \dots, f_m\}$  for  $W$  such that

$$Tv = \sigma_1 \langle v, e_1 \rangle f_1 + \dots + \sigma_k \langle v, e_k \rangle f_k,$$

or equivalently,

$$M(T) = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_k & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

**Proof:** Let  $f_i = \frac{1}{\sigma_i} Te_i$  for all  $i \leq k$  and extend and orthonormalize to form a basis for  $W$ .

**Theorem 6.23:** Let  $A \in M_{m \times n}(k)$ . Then there are isometries  $U \in M_m(k)$  and  $V \in M_n(k)$  such that  $A = U\Sigma V^*$ , where  $\Sigma \in M_{m,n}$  contains the singular values of  $A$ .

**Proof:**

$$\text{Let } U = \begin{bmatrix} | & & | \\ f_1 & \cdots & f_m \\ | & & | \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_k & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}, \text{ and } V = \begin{bmatrix} | & & | \\ e_1 & \cdots & e_n \\ | & & | \end{bmatrix}.$$

**Definition 6.24:**

$$\text{Let } \Sigma = \begin{bmatrix} \sigma_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_k & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}. \text{ The pseudoinverse to } \Sigma \text{ is } \Sigma^+ = \begin{bmatrix} \sigma_1^{-1} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_k^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}^T.$$

**Definition 6.25:** Let  $A = U\Sigma V^*$ . The **pseudoinverse** to  $A$  is  $A^+ = V\Sigma^+U^*$ .

**Proposition 6.26:** Let  $A \in M_n(k)$  be invertible. Then  $A^+ = A^{-1}$ .

**Proof:** Since  $A$  is invertible, no entry along the diagonal of  $\Sigma$  is zero, so  $\Sigma^+ = \Sigma^{-1}$ . Since  $U$  and  $V$  are isometries,  $V^*V = UU^* = I$ , so  $AA^+ = U\Sigma V^*V\Sigma^+U^* = U\Sigma\Sigma^+U^* = UU^* = I$ . Thus  $A^+ = A^{-1}$ .

**Theorem 6.27:** Let  $A \in M_{m,n}(k)$ . Then the map given by  $AA^+$  is the projection onto range  $A$ , so the vector  $\mathbf{x}$  closest to a solution to  $A\mathbf{x} = \mathbf{b}$  is  $\mathbf{x} = A^+\mathbf{b}$ .

**Proof:** Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis of eigenvectors of  $A^*A$ , let  $\sigma_1, \dots, \sigma_k$  be the singular values of  $A$ , and let  $\{f_1, \dots, f_m\}$  be the orthonormal basis given by the Singular Value Decomposition of  $A$ . Then

$$\begin{aligned} AA^+v &= A\left(\frac{1}{\sigma_1}\langle v, f_1\rangle e_1 + \cdots + \frac{1}{\sigma_k}\langle v, f_k\rangle e_k\right) \\ &= \sigma_1\left\langle \frac{1}{\sigma_1}\langle v, f_1\rangle e_1, e_1\right\rangle f_1 + \cdots + \sigma_k\left\langle \frac{1}{\sigma_k}\langle v, f_k\rangle e_k, e_k\right\rangle f_k \\ &= \langle v, f_1\rangle f_1 + \cdots + \langle v, f_k\rangle f_k, \end{aligned}$$

so if  $v = c_1f_1 + \cdots + c_mf_m$ , then  $AA^+v = c_1f_1 + \cdots + c_kf_k$ . Since  $\text{range } A = \text{span}\{f_1, \dots, f_k\}$ ,  $AA^+ = P_{\text{range } A}$ .

**Definition 6.28:**

$$\text{Let } \Sigma = \begin{bmatrix} \sigma_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_k & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}. \text{ The rank } r \text{ approximation to } \Sigma \text{ is } \Sigma_r = \begin{bmatrix} \sigma_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

**Definition 6.29:** Let  $A = U\Sigma V^*$ . The **rank  $r$  approximation** to  $A$  is  $A_r = U\Sigma_r V^*$ .

**Theorem 6.30:** Let  $A \in M_{m,n}(k)$ . Then  $A_r$  is the rank  $r$  matrix closest to  $A$  — that is, it minimizes  $\|A - X\|$ , where  $\langle A, X \rangle = \text{trace}(X^* A)$ .

## VII — Determinants

**Definition 7.1:** The **symmetric group**  $S_n$  is the group  $\{\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}\}$ , with composition given by composition of functions. The elements of  $S_n$  are called **permutations** and are written as  $\sigma = \sigma_1 \cdots \sigma_n$ , where  $\sigma(i) = \sigma_i$ .

**Definition 7.2:** Let  $\sigma$  be a permutation. The **inversion** of  $\sigma$ , denoted  $\text{inv } \sigma$ , is the number of pairs  $(i, j)$  with  $i < j$  and  $\sigma_i > \sigma_j$ .

**Definition 7.3:** The **sign** of a permutation  $\sigma$  is  $\text{sign } \sigma = (-1)^{\text{inv } \sigma}$ .

**Proposition 7.4:** Let  $\sigma$  be a permutation and  $\hat{\sigma}$  be a permutation identical to  $\sigma$ , except with  $\sigma_i$  and  $\sigma_j$  interchanged. Then  $\text{sign } \hat{\sigma} = -\text{sign } \sigma$ .

**Proof:** Suppose  $\sigma = \dots i \dots j \dots$ . Then  $\hat{\sigma} = \dots j \dots i \dots$ . Since any inversion that does not involve either  $i$  or  $j$  is unchanged from  $\sigma$  to  $\hat{\sigma}$ , we need only consider those do. Any inversion of the form  $(x, i)$  or  $(j, x)$  is unchanged, since if  $x < i$ , then  $x < j$ , and if  $x > j$ , then  $x > i$ . Thus we only need to consider the  $x$  that lie between  $i$  and  $j$ . Each one causes two inversions in  $\hat{\sigma}$  —  $(j, x)$  and  $(x, i)$  — and therefore does not affect  $\text{sign } \hat{\sigma}$ . But we have not accounted for the inversion  $(j, i)$ . Thus  $\text{sign } \hat{\sigma} = -\text{sign } \sigma$ .

**Definition 7.5:** Let  $A = [a_{ij}] \in M_{m,n}(k)$ . The **determinant** of  $A$  is given by

$$\det A = \sum_{\sigma \in S_n} (\text{sign } \sigma)(a_{\sigma_1,1}) \cdots (a_{\sigma_n,n}).$$

**Theorem 7.6:** The set

$\mathcal{A} = \{f : (\mathbb{R}^n)^n \rightarrow \mathbb{R} \mid f(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = -f(a_1, \dots, a_j, \dots, a_i, \dots, a_n), f \text{ is coordinate-wise linear}\}$   
has dimension 1 (and therefore, one basis is  $\{\det\}$ ).

**Lemma 7.6.1:** Let  $\{v_1, \dots, v_n\}$  be a basis for  $\mathbb{R}^n$  and let  $f \in \mathcal{A}$ . Then if  $f(v_1, \dots, v_n) = 0$ ,  $f(w_1, \dots, w_n) = 0$  for all  $w_1, \dots, w_n \in \mathbb{R}^n$ .

**Proof:** Expand each  $w_i$  as  $w_i = c_{i1}v_1 + \dots + c_{in}v_n$ . Since  $f$  is linear in each coordinate, we have

$$f(w_1, \dots, w_n) = f(c_{11}v_1 + \dots + c_{1n}v_n, \dots, c_{n1}v_1 + \dots + c_{nn}v_n) = \sum c_i f(v_{i_1}, \dots, v_{i_n}).$$

Now any term of this sum with some  $v_{i_j} = v_{i_k}$  will have  $f(v_{i_1}, \dots, v_{i_n}) = -f(v_{i_1}, \dots, v_{i_n}) = 0$  by the previous result, and for the rest, we can rearrange the terms to get  $f(v_{i_1}, \dots, v_{i_n}) = \pm f(v_1, \dots, v_n) = 0$ . Thus  $f(w_1, \dots, w_n) = 0$ .

**Proof:** Let  $f, g \in \mathcal{A}$  with  $g \neq 0$ , let  $\{v_1, \dots, v_n\}$  be a basis for  $\mathbb{R}^n$ , and let  $c = \frac{f(v_1, \dots, v_n)}{g(v_1, \dots, v_n)}$  ( $g(v_1, \dots, v_n) \neq 0$ , since otherwise  $g = 0$  by the lemma). Then  $(f - cg)(v_1, \dots, v_n) = 0$ , so by the lemma,  $f - cg = 0$ . Thus  $f = cg$ , so every function in  $\mathcal{A}$  is a multiple of another.

**Theorem 7.7:** Let  $A, B \in M_n(k)$ . Then  $\det AB = (\det A)(\det B)$ .

**Proof:** Define  $f \in \mathcal{A}$  by  $f(C) = \det AC$ . By the previous result, there is a  $c$  such that  $f = c \cdot \det$ . Since  $f(I) = \det A$  and  $f(I) = c \cdot \det I = c$ ,  $c = \det A$ . Then  $f(B) = \det AB = c \cdot \det B = (\det A)(\det B)$ .

**Theorem 7.8:** Let  $A \in M_n(k)$  with eigenvalues  $\lambda_1, \dots, \lambda_n$ . Then  $\det A = \lambda_1 \cdots \lambda_n$ .

**Proof:** Let  $\{v_1, \dots, v_n\}$  be a basis for  $\mathbb{R}^n$  such that  $A$  is upper triangular. Then  $A = SUS^{-1}$ , where

$$S = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \text{ and } U = \begin{bmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

Then  $\det A = \det SUS^{-1} = (\det S)(\det U)(\det S^{-1}) = \det U = \lambda_1 \cdots \lambda_n$ .