MATH 4242 Applied Linear Algebra

Sylvester W. Zhang[♮]

Summer 2024

Contents

1. Systems of Linear Equations

A $m \times n$ system of linear equation is of the form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{n1}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{n2}x_n &= b_n \\ \cdots \quad \cdots & \quad \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_n \end{aligned}$$

Such equation can be represented using product of matrices.

$$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{21} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

or by an augmented matrix.

[♮] swzhang@umn.edu University of Minnesota.

$$\left[\begin{array}{cccc|c} a_{11} & a_{21} & \cdots & a_{m1} & b_1 \\ a_{21} & a_{22} & \cdots & a_{m2} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{array}\right]$$

Definition 1.1. We have three types of elementary row operations.

(1) Multiply the i-th equation (or the i-th row of the augmented matrix), then add it to the j-th equation (or the j-th row of the augmented matrix).
(2) Permute the equations (or the rows of the augmented matrix)
(3) Multiply one equation (or one row of the augmented matrix) by a non-zero number.

1.1. **Systems of $n \times n$ Equations.** Matrices considered in this sections are all $n \times n$.

Definition 1.2. A matrix is *regular* if it can be turned into a upper triangular matrix such that every entry on the diagonal is non-zero.

Proposition 1.3. *Let E be the matrix with 1's on the diagonal and $E_{ij} = k \neq 0$ is the only other non-zero entry in the lower triangular part. Then for any matrix M, EM is the matrix obtained by multiplying the j-th row of M then adding to the i-th row of M.*

Proposition 1.4. *A matrix A is regular if and only if it has an LU factorization, i.e.*

$$A = LU$$

where L is a lower uni-triangular matrix, and U is a upper triangular matrix with non-zero diagonal entries.

Definition 1.5. Let $w \in S_n$ be a permutation, then define $P_w = \{a_{ij}\}$ to be the matrix such that

$$a_{i,j} = \begin{cases} 1 & j = w(i) \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 1.6. *For any matrix M, $P_w M$ is the matrix obtained by permuting the rows of M according to the permutation w.*

Definition 1.7. A matrix A is called *non-singular* if it can be turned into a upper triangular matrix without non-zero diagonal entry via row operations of the first two types.

Proposition 1.8. *A matrix A is non-singular if and only if it has a permuted LU factorization: $PA = LU$ where P is some permutation matrix.*

Definition 1.9. Let $A = (a_{ij})$, defined transpose of A to be $A^t := (a_{ji})$.

Proposition 1.10. *Denote A^t the transpose of A. We have that $AB = (BA)^t$.*

Proposition 1.11. *A matrix A is regular iff it admits an LDV factorization, $A = LDU$ where L is lower-unitriangular matrix, D is a diagonal matrix, and U is a uni-upper triangular matrix.*

Definition 1.12. Let A be an $n \times n$ matrix. Suppose X is a matrix such that $XA = AX = I$ where I is the identity matrix. Then X is called the inverse of A and denoted by A^{-1}. A matrix is called *invertible* if A^{-1} exists.

Proposition 1.13. *A matrix is invertible if and only if it is non-singular.*

Remark 1.14. Inverse of a matrix can be found using Gauss-Jordan Elimination — see chapter 1 of Olver-Shakiban.

1.2. **Systems of $m \times n$ Equations.**

Definition 1.15. A matrix is in row echelon form if it looks like,

$$\begin{pmatrix} \bullet & * & * & * & * & * \\ 0 & \bullet & * & * & * & * \\ 0 & 0 & 0 & \bullet & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

where $\bullet$'s are non-zero entries (called *pivots*) and $*$ represent generic entries. The pivots are the first non-zero entries in each rows. We require the pivots occupy the first several rows consecutively.

Proposition 1.16. *Every matrix can be turned into a row echelon form using elementary row operations of type I and II. In other words, every matrix A has a factorization $PA = LU$ where P is a permutation matrix, L is a lower uni-triangular matrix, and U a matrix in row-echelon form.*

Definition 1.17. Since every matrix can be turned in to row-echelon form using elementary row operations, we define its *rank* to be the number of pivots.

Proposition 1.18. *A square $n \times n$ matrix is non-singular if its rank is n (full-rank).*

2. Vector Spaces

2.1. **Some Basic Setup.**

Definition 2.1. [1] A field is a set $\mathbb{F}$ with two binary operations $\times$ (multiplication) and $+$ (addition), satisfying the following axioms.

- $a + b = b + a$ and $a \times b = b \times a$ for all $a, b \in \mathbb{F}$.
- There exists an additive identity 0 such that $0 + a = a + 0 = a$ for all $a \in \mathbb{F}$.
- There exists a multiplication identity 1 such that $1 \times a = a \times 1 = a$ for all $a \in \mathbb{F}$.
- For every $a \in \mathbb{F}$, there exists an element denoted $-a$, such that $a + (-a) = 0$.
- $0 \neq 1$.
- For every $a \in \mathbb{F}$ and $a \neq 0$, there exists an element denoted a^{-1}, such that $a \times (a^{-1}) = 1$.
- For every $a, b, c \in \mathbb{F}$, $a \times (b + c) = ab + ac$.

For most part of this class, we will take $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C} = \{a + bi | a, b \in \mathbb{R} \text{ and } i^2 = -1\}$.

Definition 2.2. For a field $\mathbb{F}$, denote $\mathbb{F}[x]$ the ring[2]of polynomials over $\mathbb{F}$.

$$\mathbb{F}[x] = \{a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n | a_0, \cdots, a_n \in \mathbb{F}, n \geqslant 0, x^m x^n = x^{m+n}\}$$

Proposition 2.3. *Every polynomial $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$ with complex coefficient has at least one complex solution. Note that this is not true for real polynomials.*

Definition 2.4. A field $\mathbb{F}$ is called *algebraically closed* if every polynomial in $\mathbb{F}[x]$ has a solution in $\mathbb{F}$. (By Proposition 2.3, $\mathbb{C}$ is algebraically closed).

[1] You don't need to worry too much about the abstract structures of a field. The purpose of this definition is to make everything self-contained. You can basically think of a field as a set on which you can do some sort of arithmetic.

[2] A ring is a field, where multiplication need not to be commutative, and multiplicative identity (0) need not exists.

Proposition 2.5. *The field of complex numbers $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$. In other words, $\mathbb{C}$ is the smallest algebraically closed field that contains $\mathbb{R}$.*

2.2. **Vector Spaces and Subspaces.** Let $\mathbb{F}$ be a field.

Definition 2.6. A set V is called a vector space over $\mathbb{F}$ if there exists an addition map

$$add : V \times V \to V,$$

a scalar multiplication map

$$mult : \mathbb{F} \times V \to V,$$

and a zero vector 0 such that $v + 0 = v$ for all $v \in V$ and $\lambda 0 = 0$ for all $\lambda \in \mathbb{F}$. (Here $\times$ denote the Cartesian product of sets[3].) We will abbreviate them by $a(v_1, v_2) = v_1 + v_2$ and $mult(a, v) = av$.

Note that this definition (implicitly) requires that a vector space V is closed under addition and scalar multiplication, i.e. $v_1 + v_2 = add(v_1, v_2) \in V$ and $av = mult(a, v) \in V$.

Elements of a vector spaces are called *vectors.*

Definition 2.7. Let V be a vector space over $\mathbb{F}$. A subset U of V is a *subspace* if it is closed under addition and scalar multiplication, and contains the zero vector. (In other words, a subspace is a subset that is a vector space itself.)

Definition 2.8. Let $U_1, \cdots, U_m$ be subspaces of V. Then define their sum to be

$$U_1 + \cdots + U_m = \{u_1 + \cdots + u_m | u_1 \in U_1, \cdots, u_m \in U_m\}$$

Proposition 2.9. *Let $U_1, \cdots, U_m$ be subspaces of V. Then $U_1 + \cdots + U_m$ is also a subspace of V, furthermore, it's the smallest subspace of V that contain all of $U_1, \cdots, U_m$.*

Definition 2.10. A sum of subspaces $U_1 + \cdots + U_m$ of V is a *direct sum* if every vector $v \in U_1 + \cdots + U_m$ can be uniquely written as $v = u_1 + \cdots + u_m$ where $u_i \in U_i$ for each i. When a summation is direct, we denote it as $U_1 \oplus \cdots \oplus U_m$.

2.3. **Linear Combination, Span, and Dimension.** Let V be a vector space over $\mathbb{F}$.

Definition 2.11. Let $v_1, v_2, \cdots, v_n \in V$, a vector $v \in V$ is a linear combination of $\{v_1, \cdots, v_n\}$ if there exists $a_1, \cdots, a_n \in \mathbb{F}$ such that

$$v = a_1 v_1 + \cdots + a_n v_n$$

Definition 2.12. Let $v_1, v_2, \cdots, v_n$ be a list of vectors in V, define their span to be the set of all linear combinations of $v_1, \cdots, v_n$.

$$\text{span}(v_1, \cdots, v_n) = \{a_1 v_1 + \cdots + a_n v_n | a_1, \cdots, a_n \in \mathbb{F}\}$$

Proposition 2.13. *For a list of vectors $v_1, \cdots, v_n \in V$, $\text{span}(v_1, \cdots, v_n)$ is a subspace of V. Furthermore, it's the smallest subspace containing all of $v_1, \cdots, v_n$.*

Definition 2.14. A vector space V is said to be *finite dimensional* it it is the span of a finitely many vectors.

Definition 2.15. $v_1, \cdots, v_m \in V$ are *linearly independent* if the only way to write 0 as a linear combination of $v_1, \cdots, v_n$ is

$$0 = 0v_1 + 0v_2 + \cdots + 0v_n.$$

[3]For sets A and B, defined $A \times B = \{(a, b) | a \in A, b \in B\}$

Proposition 2.16. $v_1, \cdots, v_m \in V$ *are* linearly independent *if and only if any vector* $v \in \operatorname{span}(v_1, \cdots, v_m)$ *can be uniquely written as a linear combination of* $v_1, \cdots, v_n$.

Definition 2.17. A list of vectors $v_1, \cdots, v_n$ is a basis of V if

- $V = \operatorname{span}(v_1, \cdots, v_n)$
- $v_1, \cdots, v_n$ are linearly independent.

Proposition 2.18. *$v_1, \cdots, v_n$ is a basis of V iff every vector $v \in V$ can be uniquely written as a linear combination of $v_1, \cdots, v_n$.*

Lemma 2.19. *Let $v_1, \cdots, v_m \in V$ be a list of vectors that spans V, i.e.* $\operatorname{span}(v_1, \cdots, v_m) = V$. *Then $\{v_1, \cdots, v_m\}$ can be reduced to a basis of V. In other words, there exists a basis $\{w_1, \cdots, w_n\}$ of V such that $w_i \in \{v_1, \cdots, v_m\}$ for all i and $n \leqslant m$.*

Lemma 2.20. *Let $v_1, \cdots, v_k \in V$ be linearly independent. Then there exists a basis of V in the form*

$$\{v_1, \cdots, v_k, w_1, \cdots, w_m\}$$

Note that it's possible that $m = 0$, in the case when $\{v_1 \cdots v_k\}$ is already a basis.

Corollary 2.21. *If U is a subspace of V, then there exists another subspace W such that $V = U \oplus W$.*

Proposition 2.22. *If $v_1, \cdots, v_n$ is a basis of V and $w_1, \cdots, w_m$ is another basis of V. Then $n = m$.*

Definition 2.23. Define the dimension of a vector space to be the size of its basis.

Proposition 2.24. *If $\{v_1, \cdots, v_n\}$ linearly independent and $n = \dim(V)$, then $\{v_1, \cdots, v_n\}$ is a basis.*

Proposition 2.25. *If U is a subspace of V, then* $\dim(U) \leqslant \dim(V)$. *Furthermore,* $\dim(U) = \dim(V)$ *iff* $U = V$.

Proposition 2.26. *If* $\operatorname{span}(v_1, \cdots, v_n) = V$ *and* $n = \dim(V)$, *then* $\{v_1, \cdots, v_n\}$ *is a basis.*

Theorem 2.27. *Let V be a finite dimensional vector space and V_1, V_2 subspaces. Then*

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)$$

Corollary 2.28. *If $V_1 + V_2$ is a direct sum, then* $\dim(V_1 \oplus V_2) = \dim(V_1) + \dim(V_2)$.[4]

3. Linear Maps and Matrices

3.1. **Linear Maps.** Let V, W be vector spaces over $\mathbb{F}$.

Definition 3.1. A map $T : V \to W$ is linear if

(1) $T(u + v) = T(u) + T(v)$ for all $u, v \in V$.
(2) $T(\lambda v) = \lambda T(v)$ for all $\lambda \in \mathbb{F}$ and $v \in V$.

Definition 3.2. We denote the set of all linear maps from $V \to W$ by $\operatorname{Hom}(V, W)$. And define $\operatorname{End}(V) = \operatorname{Hom}(V, V)$.

[4] We will see later that the converse is also true.

Lemma 3.3. *Let $v_1, \cdots, v_n$ be a basis for V and $w_1, \cdots, w_n$ a basis for W (i.e. V, W same dimension). Then there exists a unique linear map $T \in \text{Hom}(V, W)$ such that $T(v_i) = w_i$ for all i. The map is given by $T(a_1v_1 + \cdots + a_nv_n) = a_1w_1 + \cdots + a_nw_n$.*

Proposition 3.4. *The set $\text{Hom}(V, W)$ is a vector space over $\mathbb{F}$, with addition and scalar multiplication given as follows.*

$$(\varphi + \psi)(v) := \varphi(v) + \psi(v)$$
$$(\lambda\varphi)(v) := \lambda\varphi(v)$$

Lemma 3.5. *Let $T \in \text{Hom}(V, W)$, then $T(0_V) = 0_W$.*

Let $T \in \text{Hom}(V, W)$.

Definition 3.6. The *kernal* (or null space) of T is $\text{Ker}(T) = \{v \in V : Tv = 0\}$

Proposition 3.7. $\text{Ker}(T)$ *is a subspace of V.*

Proposition 3.8. $\text{Ker}(T) = \{0\}$ *if and only if T is injective.*

Definition 3.9. The *image* (or range) of T is $\text{Img}(T) = \{Tv | v \in V\}$

Proposition 3.10. $\text{Img}(T)$ *is a subspace of W.*

Proposition 3.11. *T is surjective iff* $\text{Img}(T) = W$.

Theorem 3.12. $\dim(V) = \dim(\text{Ker}(T)) + \dim(\text{Img}(T))$.

Proposition 3.13.
(1) *if* $\dim(V) > \dim(W)$*, then any* $T \in \text{Hom}(V, W)$ *is not injective.*
(2) *if* $\dim(V) < \dim(W)$*, then any* $T \in \text{Hom}(V, W)$ *is not surjective.*
(3) *if there exists a bijective* $T \in \text{Hom}(V, W)$*, then* $\dim(V) = \dim(W)$.

3.2. **Matrices from Linear Maps.** Denote the set of all $m \times n$ matrix with entries in $\mathbb{F}$ by $M_{m\times n}(\mathbb{F})$. Let V, W be finite dimensional vector spaces over $\mathbb{F}$.

Definition 3.14. Suppose V has basis $v_1, \cdots, v_n$ and W has basis $w_1, \cdots, w_m$. Let $T \in \text{Hom}(V, W)$. Then define $\mathcal{M}(T)$ to be the matrix $[a_{ij}]$ such that

$$T(v_k) = a_{1k}w_1 + a_{2k}w_2 + \cdots + a_{m,k}w_m.$$

Remark 3.15. Note that the usage of $\mathcal{M}$ requires a choice of basis for V and W. In general we shall denote $\mathcal{M}_{B_1,B_2}(T)$ where B_1 is the basis for V and B_2 the basis for W. However in most case we will omit the subscript when the context is clear.

Proposition 3.16. *Let $S, T \in \text{Hom}(V, W)$, $\mathcal{M}(S) + \mathcal{M}(T) = \mathcal{M}(S + T)$*
Let $S \in \text{Hom}(U, W)$ and $T \in \text{Hom}(V, U)$, then $\mathcal{M}(S)\mathcal{M}(T) = \mathcal{M}(ST)$.

Definition 3.17. For any $A \in M_{m\times n}$, Let $A_{\bullet,k}$ denote the k-th column vector and $A_{k,\bullet}$ denote the k-th row vector.

Proposition 3.18. $(AB)_{\bullet,k} = A(B_{\bullet,k})$

Theorem 3.19. *For any $A \in M_{m\times n}$, we have*

$$\dim(\text{span}(A_{1,\bullet}, \cdots, A_{m,\bullet})) = \dim(\text{span}(A_{\bullet,1}, \cdots, A_{\bullet,n}))$$

Proposition 3.20. *The dimension of column span or row span of a matrix equals to its rank (see Definition 1.17).*

Definition 3.21. A linear map $T \in \mathrm{Hom}(V, W)$ is invertible if there exists an linear map $S \in \mathrm{Hom}(W, V)$ such that $TS = \mathrm{id}_W$ and $ST = \mathrm{id}_V$.

Proposition 3.22. *If a map T is invertible, then its inverse is unique, denote it by T^{-1}.*

Proposition 3.23. *A map T is invertible iff $\mathcal{M}(T)$ is non-singular Definition 1.7. Furthermore, $\mathcal{M}(T^{-1}) = \mathcal{M}(T)^{-1}$.*

Definition 3.24. A linear map is an *isomorphism* if it's invertible. Two vectors spaces are *isomorphic* if there exists an isomorphism between them.

Theorem 3.25. *Two vector spaces over $\mathbb{F}$ is isomorphic if and only if they have the same dimension. (In other words, vector spaces are classified by $\mathbb{N}$)*

Corollary 3.26. *Let $\dim V = n$ and $\dim W = m$. The vector space $\mathrm{Hom}(V, W) \cong M_{m\times n}(\mathbb{F})$ are isomorphic, with the map $\mathcal{M}$ being the isomorphism.*

Theorem 3.27. *Let V be a vector space with basis $B_1 = \{v_1, \cdots, v_n\}$. Suppose it has another basis $B_2 = \{w_1, \cdots, w_n\}$. Let $C = \mathcal{M}_{B_1,B_2}(\mathrm{id})$ where $\mathrm{id} \in \mathrm{Hom}(V, V)$ is the identity map. Then change of basis corresponds to conjugation by C.*

In particular, let $T \in \mathrm{Hom}(V, V)$ and $A = \mathcal{M}_{B_1,B_1}(T)$ and $B = \mathcal{M}_{B_2,B_2}(T)$. Then we have

$$A = C^{-1}BC$$

3.3. Quotient and Dual spaces.

Definition 3.28. Let $v \in V$ and $U \subseteq V$. Define $v + U := \{v + u | u \in U\}$. This is called a *coset*.

Definition 3.29. Let $U \subseteq V$. Define the quotient space V/U to be $\{v + U | v \in V\}$, with addition and scalar multiplication given by

$$(v_1 + U) + (v_2 + U) = (v_1 + v_2) + U$$
$$\lambda(v + U) = \lambda v + U$$

Definition 3.30 (alternative definition). Let $\sim$ be an equivalence relation on V. Define $[v]_\sim := \{u \in V | u \sim v\}$ the equivalence class generated by v. Then we can define quotient space $V/\sim := \{[v]_\sim | v \in V\}$.

Remark 3.31. For $U \subset V$, define an equivalence relation $\sim_U$ by $v \sim_U u \iff v - u \in U$. Then Definitions 3.29 and 3.30 agree, i.e. $V/U = V/\sim_U$.

Definition 3.32. For $U \subset V$, define the quotient map $\pi : V \to V/U$ by $\pi(v) = v + U$. Note that $\mathrm{Ker}(\pi) = U$.

Proposition 3.33. $\dim V/U = \dim V - \dim U$.

Theorem 3.34. *For any $T \in \mathrm{Hom}(V, W)$, define $\tilde{T} \in \mathrm{Hom}(V/\mathrm{Ker}(T), W)$ by $\tilde{T}(v + \mathrm{Ker}(T)) = Tv$. Then $\tilde{T}\pi = T$, and defines an isomorphism between $V/\mathrm{Ker}(T)$ and $\mathrm{Img}(T)$.*

Definition 3.35. A linear map from V to $\mathbb{F}$ is called a *linear functional*. Denote $\mathrm{Hom}(V, \mathbb{F})$ the set of all linear functionals on V.

Proposition 3.36. *$\mathrm{Hom}(V, \mathbb{F})$ is a vector space, with addition and multiplication given by $(f + g)(v) = f(v) + g(v)$ and $(\lambda f)(v) = \lambda f(v)$. This is called the* dual space *of V, and is denoted by T^*.*

Proposition 3.37. $\dim V = \dim V^*$.

Definition 3.38. Let $v_1, \cdots, v_n$ be a basis of V. Then define $v_i^* \in V^*$ to be the linear functional $v_i^*(v_j) = \delta_{i,j}$[5]. For any $v = a_1 v_1 + \cdots a_n v_n \in V$, define $v^* = a_1 v_1^* + \cdots + a_n v_n^*$.

Proposition 3.39. *Let $v_1, \cdots, v_n$ be a basis. Then $v = v_1^*(v)v_1 + \cdots + v_n^*(v)v_n$ for all $v \in V$.*

Proposition 3.40. *$v_1^*, \cdots, v_n^*$ is a basis for V^*.*

Definition 3.41. Suppose $T \in \operatorname{Hom}(V, W)$. Define the *dual linear map* $T^* \in \operatorname{Hom}(W^*, V^*)$ to be

$$T^*(f) = f \circ T$$

Proposition 3.42.
- $(S + T)^* = S^* + T^*$
- $(\lambda T)^* = \lambda T^*$.
- $(ST)^* = T^* S^*$.

Definition 3.43. For any subspace $U \subseteq V$, define its annihilator $U^0 := \{f \in V^* : f(u) = 0 \text{ for all } u \in U\}$.

Proposition 3.44. *U^0 is a subspace of V^*.*

Proposition 3.45. $\dim U^0 = \dim V - \dim U$. *Recall that this is also the dimension of V/U. In particular, there is an isomorphism $(V/U)^* \cong U^0$ given by π^*.*

Proposition 3.46. (a) $U^0 = \{0\} \iff U = V$ (b) $U^0 = V^* \iff U = \{0\}$.

Theorem 3.47. (a) $(\operatorname{Img} T)^0 = \operatorname{Ker} T^*$ (b) $(\operatorname{Ker} T)^0 = \operatorname{Img} T^*$

Corollary 3.48. $\operatorname{Ker} T^* \cong (V/\operatorname{Img} T)^*$ *and* $\operatorname{Img} T^* \cong (V/\operatorname{Ker} T)^*$.

Corollary 3.49. *T is injective iff T^* is surjective. T is surjective iff T^* is injective.*

Theorem 3.50. *Let $T \in \operatorname{Hom}(V, W)$, and $T^* \in \operatorname{Hom}(W^*, V^*)$. Then $\mathcal{M}(T)^t = \mathcal{M}(T^*)$.*

4. Inner Product Spaces

Throughout this section, let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

4.1. Inner Products and Norms.

Definition 4.1. The dot product of two vectors in $\mathbb{R}^n$ is a map from $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{F}$, defined by

$$(x_1, \cdots, x_n) \cdot (y_1, \cdots, y_n) = x_1 y_1 + \cdots + x_n y_n$$

Definition 4.2. The dot product of two vectors in $\mathbb{C}^n$ is a map from $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{F}$, defined by

$$(x_1, \cdots, x_n) \cdot (y_1, \cdots, y_n) = x_1 \overline{y}_1 + \cdots + x_n \overline{y}_n$$

where $\overline{a + bi} = a - bi$ is the complex conjugate.

Definition 4.3. Let V be vector space over $\mathbb{F}$ ($\mathbb{C}$ or $\mathbb{R}$). A inner product on V is a map $V \times V \to \mathbb{F}$ which sends (v, u) to $\langle v, u \rangle$ such that

(1) $\langle v, v \rangle \geqslant 0$.
(2) $\langle v, v \rangle = 0 \iff v = 0$.
(3) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$

[5]Here $\delta_{i,j}$ is the Kronecker delta symbol: $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise.

(4) $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$.
(5) $\langle u, v\rangle = \overline{\langle v, u\rangle}$[6]

Proposition 4.4 (Bilinearity)**.** *A inner product $\langle\ ,\ \rangle$ satisfy $\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle$ and $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$.*

Remark 4.5. A pairing satisfying (4) and (6) of Definition 4.3 and Proposition 4.4 together is known as being *bilinear*. Usually a inner product is defined to be bilinear, however, as we see here one-sided linearity is enough to imply bilinearity.

Proposition 4.6. *An inner product $\langle\ ,\ \rangle$ on V satisfy*

(1) *Fix any $u \in V$, the map $v \mapsto \langle u, v\rangle$ is a linear functional.*
(2) *$\langle v, 0\rangle = 0 = \langle 0, v\rangle$ for any $v \in V$.*

Definition 4.7. Given an inner product $\langle\ ,\ \rangle$, define the norm $\|\ \|$ to be the positive square-root $\|v\| = \sqrt{\langle v, v\rangle}$.

Proposition 4.8. *Let $I = [a, b] \subset \mathbb{R}$ be an closed interval on $\mathbb{R}$. Let $V = C^0(I)$ denote all continuous $\mathbb{R}$-valued functions defined on I (domain is I). Then*

$$\langle f, g\rangle = \int_a^b f(x)g(x)\ dx$$

defines an inner product on V. The norm $\|f\| = \sqrt{\langle f, g\rangle}$ is called the L_2 norm on $C^0(I)$.

Definition 4.9. For $z \in \mathbb{C}$, define the complex modulus to be $|z| = \sqrt{z\overline{z}}$. Note that when z is real (i.e. no imaginary part), then $|z|$ is the absolute value.

Theorem 4.10 (Cauchy-Schwartz inequality)**.** *Let V be an inner product space with inner product $\langle\ ,\ \rangle$ and norm $\|\ \|$. Then for any $u, v \in V$, we have*

$$|\langle u, v\rangle| \leqslant \|u\|\|v\|$$

Moreover, the equality occurs only when u, v are linearly independent.

Remark 4.11. The Cauchy-Schwartz inequality tells us that the ratio $\frac{|\langle u,v\rangle|}{\|u\|\|v\|}$ is in between -1 and 1. Therefore we can define the 'abstract' angle between two vectors v, u to be

$$\theta_{u,v} = \arccos \frac{|\langle u, v\rangle|}{\|u\|\|v\|}$$

Definition 4.12. We say two vectors u, v are orthogonal if $\langle u, v\rangle = 0$.

Proposition 4.13. *If v, u orthogonal in V, then $\|v\|^2 + \|u\|^2 = \|v + u\|^2$.*

Remark 4.14. Definition 4.12 generalizes the usual notion of orthogonality in $\mathbb{R}^2$ in a sense that when two vectors are orthogonal, then the angle between then is $\theta_{u,v} = \pi/2$.

Theorem 4.15. *Let V be an inner product space, and $u, v \in V$. Then $\|u + v\| \leqslant \|u\| + \|v\|$.*

Definition 4.16. We can define norms more generally without requiring an inner product. A norm on V is a map $\|\cdot\| : V \to \mathbb{R}_{\geqslant 0}$ such that

- $\|v\| \geqslant 0$ and $\|v\| = 0$ only when $v = 0$.
- $\|\lambda v\| = |\lambda|\|v\|$

[6]When $\mathbb{F} = \mathbb{R}$, the 'complex' conjugate of a real number is just itself.

- $\|v + u\| \leqslant \|v\| + \|u\|$

Proposition 4.17. *Let $K = (k_{ij})$ be the matrix whose entries are the inner product of the basis vectors, i.e. $k_{ij} = \langle e_i, e_j \rangle$. Then for $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$ we have*

$$\langle x, y \rangle = \left\langle \sum_i x_i e_i, \sum_j y_j e_j \right\rangle = \sum_{i,j} x_i y_j \langle e_i, e_j \rangle = x^t K y$$

Definition 4.18. A $n \times n$ matrix is positive-definite if $K^t = K$ and satisfy $x^t K x > 0$ for all $0 \neq x \in \mathbb{F}^n$. More generally, we say K is positive semi-definite if $K^t = K$ and $x^t K x \geqslant 0$ for all x.

Theorem 4.19. *Every inner product is given by $\langle x, y \rangle = x^t K y$ where K is a positive-definite matrix.*

Proposition 4.20. *Positive-definite matrices are non-singular (invertible).*

Definition 4.21. Given any $v_1, \cdots, v_n \in V$, we define the Gram matrix to be $K = (k_{ij})$ where $k_{ij} = \langle v_i, v_j \rangle$. In particular, Let A be the matrix whose column vectors are $v_1, \cdots, v_n$, then the Gram matrix is $K = A^t C A$, where C is the symmetric positive definite matrix defining the inner product.

Proposition 4.22. *A Gram matrix is always positive semi-definite. A Gram matrix is positive-definite if and only if $v_1, \cdots, v_n$ are linearly independent.*

Proposition 4.23. *Let A be an $m \times n$ matrix ($m \geqslant n$), then TFAE:*

- *$K = A^t A$ is positive-definite;*
- $\mathrm{Ker}(A) = 0$*;*
- *A has linearly independent columns;*
- $\mathrm{rank}(A) = n$*.*

Theorem 4.24. *Suppose $A \in M_{m \times n}$ and $K = A^T A$ is positive-definite. Then for any symmetric positive definite matrix $C \in M_{m \times m}$, the matrix $K' = A^t C A$ is also positive-definite.*

Proposition 4.25. *For $K = A^t C A$, we have* $\mathrm{Ker}(K) = \mathrm{Ker}(A)$*, and hence* $\mathrm{rank}(K) = \mathrm{rank}(A)$*.*

4.2. **Orthonormal Basis.**

Definition 4.26. Let V be a real or complex inner product space. A basis $v_1, \cdots, v_n$ of V is called *orthogonal* if $\langle v_i, v_j \rangle = \delta_{i,j}$ for all i. An orthogonal basis of unit vectors is called *orthonormal*.

Proposition 4.27. *Let $v_1, \cdots, v_n \in V$ be pair-wise orthogonal, then they must be linearly independent.*

Corollary 4.28. *Let $v_1, \cdots, v_n \in V$ be pair-wise orthogonal, then they form a basis for* $\mathrm{span}(v_1, \cdots, v_n)$.

Proposition 4.29. *If $e_1, \cdots, e_n$ is an orthonormal basis, then for any $v \in V$, we have that*

- $v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$
- $\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2$

Proposition 4.30. *If $e_1, \cdots, e_n$ is an orthonormal basis, then*

Theorem 4.31. *Given any basis $w_1, \cdots, w_n$ of V, one can construct an orthogonal basis $v_1, \cdots, v_n$ using the Gram-Schmidt process:*

- $v_1 = w_1$.
- $v_2 = w_2 - \dfrac{\langle w_2, v_1\rangle}{\|v_1\|^2} v_1$
- $v_3 = w_3 - \dfrac{\langle w_3, v_1\rangle}{\|v_1\|^2} v_1 - \dfrac{\langle w_3, v_2\rangle}{\|v_2\|^2} v_2$
- $\cdots$
- $v_k = w_k - \displaystyle\sum_{i=1}^{k-1} \dfrac{\langle w_k, v_i\rangle}{\|v_i\|^2} v_i$
- $\cdots$.

Corollary 4.32. *Every finite dimensional inner-product space has an orthonormal basis.*

Theorem 4.33. *Suppose V is finite-dimensional and T is a linear functional on V. Then there is a unique vector $v \in V$ such that $T(u) = \langle u, v\rangle$for every $u \in V$.*

Definition 4.34. A matrix A is orthogonal if $A^t A = I = AA^t$, or equivalently $A^t = A^{-1}$.

Proposition 4.35. *A matrix is orthogonal if and only if its column vectors form an orthonormal basis of $\mathbb{F}^n$ w.r.t the dot product.*

Definition 4.36. Let $W \subset V$ be a subspace. A vector $v \in V$ is said to be orthogonal to W is $\langle v, w\rangle = 0$ for all $w \in W$.

Definition 4.37. Two subspaces $W, U \subset V$ are said to be orthogonal if $\langle w, u\rangle = 0$ for all $w \in W, u \in U$.

Definition 4.38. The orthogonal complement of a subset $W \subset V$, denoted $W^\perp$, is the set of all vectors in V that are orthogonal to W.

$$W^\perp = \{v \in V | \langle v, w\rangle = 0 \text{ for all } w \in W\}$$

Proposition 4.39. *Let $U^\perp$ be the orthogonal complement of $U \subset V$.*

(1) *$U^\perp$ is always a subspace of V*[7].
(2) $\{0\}^\perp = V$.
(3) $V^\perp = \{0\}$
(4) $U^\perp \cap U \subseteq \{0\}$.
(5) *If $W \subset U \subset V$, then $U^\perp \subset W^\perp$.*

Proposition 4.40. *Let U be a finite dimensional subspace of V, then*

$$V = U^\perp \oplus U$$

And $\dim U^\perp = \dim V - \dim U$.

Proposition 4.41. $U = (U^\perp)^\perp$

[7]Even if U is not a subspace.

Definition 4.42. The orthogonal projection of v onto W, denoted $\text{Proj}_W(v)$ is the element $w \in W$ such that $v - w$ is orthogonal to W.

In other words, if we write v in the direct sum of $V = W^\perp \otimes W$ as $v = w' + w$ with $w \in W$ and $w' \in W^\perp$, then $\text{Proj}_W(v) = w'$.

Theorem 4.43. *Let $w_1, \cdots, w_n$ be an orthogonal basis for a subspace $W \subset V$. Then the orthogonal projection of v onto W is*

$$\text{Proj}_W(v) = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_n \rangle u_n$$

Definition 4.44. Define the cokernel of a linear map $T \in \text{Hom}(V, W)$ to be the quotient space $\text{coKer}(T) = W / \text{Img}(T)$ and the coimage to be $\text{coImg}(T) = V / \text{Ker}(T)$.

Theorem 4.45. *We have* $\text{Ker}(T) = \text{coImg}(T)^\perp$ *and* $\text{Img}(T) = \text{coKer}(T)^\perp$

Proposition 4.46. *The equation $Ax = b$ has a solution if b is orthogonal to the cokernel of A.*

5. Eigenvalues and Eigenvectors

Definition 5.1. Let A be an $n \times n$ complex or real matrix, then λ is an eigenvalue of A if there exists a non-zero vector v, called an eigenvector, such that $Av = \lambda v$.

Proposition 5.2. *λ is an eigenvalue of A if and only if $A - \lambda I$ is singular, i.e. there exist solutions to the equation $(A - \lambda I)v = 0$.*

Definition 5.3. The characteristic polynomial of A, denoted $P_A(x)$, is defined to be

$$P_A(x) = \det(A - xI)$$

Remark 5.4. A matrix A and its transpose A^t have the same characteristic polynomial.

Proposition 5.5. *Over the complex numbers, the characteristic polynomial can be factored into*

$$P_A(x) = (-1)^n (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$$

where $\lambda_1, \cdots, \lambda_n$ are the complex eigenvalues of A.

Proposition 5.6. $\lambda_1 + \cdots + \lambda_n = \text{tr}(A)$ *and* $\lambda_1 \lambda_2 \cdots \lambda_n = \det(A)$.

Definition 5.7. The algebraic multiplicity of an eigenvalue λ is its multiplicity as a root of $P_A(x)$.

Proposition 5.8. *If $v_1, \cdots, v_k$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \cdots, \lambda_k$, then $v_1, \cdots, v_k$ are linearly independent.*

Definition 5.9. Let λ be an eigenvalue of A, then the eigenspace corresponding to λ, denoted by V_λ is the space $\text{Ker}(A - \lambda I)$. The dimension of V_λ is called the geometric multiplicity of λ.

Definition 5.10. A matrix A is called complete if the algebraic multiplicity and geometric multiplicity of all eigenvalues equal.

Proposition 5.11. *A $n \times n$ matrix with n distinct eigenvalues is complete.*

Theorem 5.12. *Every complete matrix A can be diagonalized as follows*

$$A = SDS^{-1}$$

where $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ and S is the matrix whose columns are linearly independent eigenvectors $v_1, \cdots, v_n$.

Remark 5.13. We will more often call complete matrices diagonalizable.

Theorem 5.14. *If A is a diagonalizable (complete) matrix, then all the k dimensional complex invariant subspace of A are spanned by linearly independent eigenvectors of A.*

Proposition 5.15. *Every real symmetric matrix is diagonalizable.*

Theorem 5.16 (spectral decomposition)**.** *Let A be a symmetric matrix, then $A = QDQ^t$ where D is the diagonal matrix of eigenvalues of A, and Q is an orthogonal matrix whose columns are the orthonormal eigenvectors of A.*

Corollary 5.17. *A symmetric real matrix A is positive definite if and only if all of its eigenvalues are positive. (Note that when A is not necessary symmetric, being positive definite implies having positive eigenvalues, but the converse is not always true). A symmetric matrix A is positive semi-definite if and only if all of its eigenvalues are non-negative (possibly zero).*

Definition 5.18. Let λ be an eigenvalue of A. Then a Jordan chain of λ is a list of vectors $w_1, \cdots, w_k$ such that

$$Aw_1 = \lambda w_1, \quad Aw_2 = \lambda w_2 + w_1, \cdots, Aw_k = \lambda w_k + w_{k-1}$$

Definition 5.19. A non-zero vector w is called a generalized eigenvector of A if $(A - \lambda I)^k w = 0$ for some finite number k.

Definition 5.20. A Jordan basis of a square matrix A is a basis of $\mathbb{C}^n$ (or $\mathbb{R}^n$) consisting of Jordan chains of A.

Theorem 5.21. *Every square matrix has a Jordan basis.*

Definition 5.22. A Jordan block is a matrix such that the diagonal entries are the same number, the super-diagonal entires are either 1 or 0, and the other entires are zero.

Definition 5.23. A Jordan canonical form of a square matrix A, is the block-diagonal matrix where each diagonal block is a Jordan block with eigenvalues on the diagonal. Denote J_A the Jordan canonical form of A.

Theorem 5.24. *Any square matrix A can be written as $A = SJ_AS^{-1}$ where J_A is the Jordan canonical form of A, and S is the matrix whose columns form the Jordan basis.*

Remark 5.25. The number of Jordan blocks in J_A equals to the number of Jordan chains in the Jordan basis of A.

Definition 5.26. The singular values of a general (non-square) matrix are the square roots of the eigenvalues of the Gram matrix $K = A^tA$.

Theorem 5.27. *Every $m \times n$ matrix of rank r can be written as*

$$A = P\Lambda Q^t$$

where Λ *is a* $r \times r$ *diagonal matrix with the singular values of* A*. The columns of* A *form an orthogonal basis for* $\mathrm{Img}(A)$ *and the columns of* Q *form an basis for coImg*(A)*. In particular, the columns of* Q *are the normalized eigenvectors of the Gram matrix* $K = A^t A$ *corresponding to the non-zero eigenvalues.*

Remark 5.28. If A has no zero eigenvalue, then the SVD of A is 'the same' as the spectral decomposition of $K = A^t A$: $K = A^t A = (Q\Lambda P^t)(P\Lambda Q^t) = Q\Lambda^2 Q^t$.