Notes on Linear Algebra

Eric Liu

Contents

Chapter 1 Linear Algebra Done Outrageous _ Page 2

1.1 Determinant — 2
1.2 Dual Space — 3
1.3 Tensor Algebra — 4
1.4 Norm and Inner Product — 12
1.5 Operator Norm — 18

Chapter 1

Linear Algebra Done Outrageous

1.1 Determinant

Let $N \triangleq \{1, \dots, n\}$. Given a permutation $\sigma \in S_n$, if

$$\Big\{(x, y) \in N^2 : x < y \text{ and } \sigma(x) > \sigma(y)\Big\} \text{ has even numbers of element.}$$

We say σ is **even** and write

$$\operatorname{sgn} \sigma = 1$$

Given some n-by-n square matrix A, its determinant is by definition

$$\det A \triangleq \sum_{\sigma \in S_n} \Big(\operatorname{sgn} \sigma \prod_{k=1}^{n} A_{\sigma(k),k} \Big)$$

Or equivalently, the unique alternating multilinear map from $(\mathbb{F}^n)^n$ to $\mathbb{F}$ such that

$$\det(I) = 1$$

Theorem 1.1.1. (Determinant) Let $\{v_1, \dots, v_n\}$ be a basis for V and $\{\omega_1, \dots, \omega_n\}$ be its dual basis. We have

$$\operatorname{sgn} \sigma = \det \Big(\begin{bmatrix} \omega_1 v_{\sigma(1)} & \cdots & \omega_1 v_{\sigma(n)} \\ \vdots & \ddots & \vdots \\ \omega_n v_{\sigma(1)} & \cdots & \omega_n v_{\sigma(n)} \end{bmatrix} \Big)$$

1.2 Dual Space

Given a vector space V over $\mathbb{R}$, the map $\alpha : U \to (U^\vee)^\vee$ defined by

$$\alpha(x)(\rho) \triangleq \rho(x)$$

is a vector space isomorphism. Given a linear map $A : V \to W$, we may define its **dual map** $A^\vee : W^\vee \to V^\vee$ by

$$A^\vee(\xi)(v) \triangleq \xi \circ A(v)$$

1.3 Tensor Algebra

Abstract

In this section, by the term **ring**, we mean a ring with a multiplication identity, and by the term **real algebra**, we mean a real vector space equipped with a vector multiplication compatible with both scalar multiplication and addition. In this definition, for a real algebra A to be a ring, A must be associative. By the term **ideal**, we mean a 2-sided ideal. If we say a multi-linear map $M : V^k \to Z$ is **alternating**, we mean that M maps $(v_1, \dots, v_n)$ to 0 if two arguments coincide.

Given a finite collection $(V_1, \dots, V_n)$ of finite dimensional real vector space, by the term **tensor product of** $V_1, \dots, V_n$, we mean a real vector space usually denoted by $V_1 \otimes \cdots \otimes V_n$ and a multilinear map $\otimes : V_1 \times \cdots \times V_n \to V_1 \otimes \cdots \otimes V_n$ satisfying the universal property: If $B : V_1 \times \cdots \times V_n \to Z$ is a multilinear map, then there exists a unique linear map $\beta : V_1 \otimes \cdots \otimes V_n$ such that

$$B(v_1, \dots, v_n) = \beta(v_1 \otimes \cdots \otimes v_n)$$

In other words, we have the commutative diagram

$$\begin{tikzcd}
V_1 \times \cdots \times V_n \arrow[r, "B"] \arrow[d, "\otimes"'] & U \\
V_1 \otimes \cdots \otimes V_n \arrow[ur, dashed, "\exists ! \beta"'] &
\end{tikzcd}$$

This approach indeed define a pair of vector space and multilinear map uniquely up to isomorphism, in the sense of Theorem 1.3.1, where we define the isomorphism between tensor product.

Theorem 1.3.1. (Uniqueness of Tensor product) Given a finite collection $(V_1, \dots, V_n)$ of finite dimensional real vector space, if $V_1 \otimes \cdots \otimes V_n, V_1 \otimes' \cdots \otimes' V_n$ both satisfy the universal property, then there exists an linear isomorphism $T : V_1 \otimes \cdots \otimes V_n \to V_1 \otimes' \cdots \otimes' V_n$ such that

$$T(v_1 \otimes \cdots \otimes v_n) = v_1 \otimes' \cdots \otimes' v_n$$

Proof. Because $V_1 \otimes \cdots \otimes V_n$ satisfies the universal property, there exists a linear map $T : V_1 \otimes \cdots \otimes V_n \to V_1 \otimes' \cdots \otimes' V_n$ such that

$$\otimes' = T \circ \otimes$$

It remains to show T is bijective. Similarly, because $V_1 \otimes' \cdots \otimes' V_n$ satisfies the universal property, there exists a linear map $T' : V_1 \otimes' \cdots \otimes' V_n \to V_1 \otimes \cdots \otimes V_n$ such that

$$\otimes = T' \circ \otimes'$$

Composing the two equations, we have

$$\otimes' = T \circ T' \circ \otimes'$$

It then follows from uniqueness of the induced linear map in universal property that $T \circ T' = \mathbf{id} : V_1 \otimes' \cdots \otimes' V_n \to V_1 \otimes' \cdots \otimes' V_n$. This implies T is indeed bijective. ■

We have shown that tensor products is unique up to isomorphism. A construction further shows that if B_i are bases for V_i, then

$$\{v_1 \otimes \cdots \otimes v_n : v_i \in B_i \text{ for all } 1 \leq i \leq n\} \text{ form a basis for } V_1 \otimes \cdots \otimes V_n$$

Theorem 1.3.2. (Associativity of the Tensor product) Given three finite-dimensional real vector spaces X, Y, Z, there exists a unique linear isomorphism $F : X \otimes Y \otimes Z \to (X \otimes Y) \otimes Z$ that satisfy

$$F(x \otimes y \otimes z) = (x \otimes y) \otimes z$$

Proof. Define $f : X \times Y \times Z \to (X \otimes Y) \otimes Z$ by

$$f(x, y, z) \triangleq (x \otimes y) \otimes z$$

It follows from the universal property that there exists a unique linear map $F : X \otimes Y \otimes Z \to (X \otimes Y) \otimes Z$ such that

$$F(x \otimes y \otimes z) = f(x, y, z) = (x \otimes y) \otimes z$$

It remains to show F is bijective. For all $z \in Z$, define $h_z : X \times Y \to X \otimes Y \otimes Z$ by

$$h_z(x, y) \triangleq x \otimes y \otimes z$$

If follows from the universal property that there exists a unique linear map $H_z : X \otimes Y \to X \otimes Y \otimes Z$ such that

$$H_z(x \otimes y) = h_z(x, y) = x \otimes y \otimes z$$

Define $h : (X \otimes Y) \times Z \to X \otimes Y \otimes Z$ by

$$h(v, z) \triangleq H_z(v)$$

It is clear that h in linear in $(X \otimes Y)$. We now show h is linear in Z, that is

$$H_{c_1 z_1 + z_2} = c_1 H_{z_1} + H_{z_2}$$

By definition,

$$\begin{aligned}(c_1 H_{z_1} + H_{z_2})(x \otimes y) &= c_1 x \otimes y \otimes z_1 + x \otimes y \otimes z_2 \\ &= x \otimes y \otimes (c_1 z_1 + z_2) = h_{c_1 z_1 + z_2}(x, y)\end{aligned}$$

It then follows from the uniqueness part of the universal property that $H_{c_1 z_1 + z_2} = c_1 H_{z_1} + H_{z_2}$. (done)

We have shown h is indeed bilinear. It follows from the universal property that there exists a unique linear map $H : (X \otimes Y) \otimes Z \to X \otimes Y \otimes Z$ such that

$$H((x \otimes y) \otimes z) = h(x \otimes y, z) = H_z(x \otimes y) = x \otimes y \otimes z$$

Let $\otimes : X \times Y \times Z \to X \otimes Y \otimes Z$ denotes the tensor product, we now have

$$\otimes = H \circ F \circ \otimes$$

It then follows from universal property that $H \circ F = \mathbf{id} : X \otimes Y \otimes Z \to X \otimes Y \otimes Z$. This implies F is indeed bijective. (done) ■

Let V be a finite-dimensional real vector space. By its **tensor algebra**, we mean any real associative algebra $T(V)$ with an injective linear map $i : V \to T(V)$ that satisfies the universal property: If A is a real associative algebra and $f : V \to A$ is a linear map, then there exists a unique algebra homomorphism $F : T(V) \to A$ such that the diagram

$$\begin{tikzcd} V \arrow[r, "i"] \arrow[dr, "f"'] & T(V) \arrow[d, dashed, "F"] \\ & A \end{tikzcd}$$

commutes. The proof that such definition is indeed unique up to isomorphism is similar to that of Theorem 1.3.1 and thus omitted. We now give the most useful construction.

Let V be finite-dimensional real vector space. We use the notation

$$T^n(V) \triangleq \overbrace{V \otimes \cdots \otimes V}^{n \text{ copies}}$$

and call $T^n(V)$ the **n-th tensor power of** V or the **n-fold tensor product of** V.

Define

$$\begin{aligned} T(V) &\triangleq \bigoplus_{n=0}^{\infty} T^n(V) \\ &= \mathbb{R} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots \end{aligned}$$

and define for all $f, g \in T(V)$ the multiplication

$$(fg)(n) \triangleq \sum_{k=0}^{n} f(k)g(n-k)$$

where

$$\begin{aligned} &\Big(\sum_I a_I v_{I(1)} \otimes \cdots \otimes v_{I(k)} \Big) \Big(\sum_J b_J v_{J(1)} \otimes \cdots \otimes v_{J(l)} \Big) \\ &\triangleq \sum_{I,J} a_I b_J v_{I(1)} \otimes \cdots \otimes v_{I(k)} \otimes v_{J(1)} \otimes \cdots \otimes v_{J(l)} \end{aligned}$$

where $\{v_1, \dots, v_m\}$ is some basis for V, I run through the set of function that maps $\{1, \dots, k\}$ into $\{1, \dots, m\}$ and J run through the set of function that maps $\{1, \dots, l\}$ into $\{1, \dots, m\}$. For example, given two elements

$$(5, 0, v_1 \otimes v_2, 0, 0, \dots) \text{ and } (7, v_3, 0, 0, \dots)$$

of $T(V)$, their product is defined to be

$$(35, 5v_3, 7v_1 \otimes v_2, v_1 \otimes v_2 \otimes v_3, 0, 0, \dots)$$

Tedious effort shows that our multiplication is consistent with abuse of notation in the sense that if $f, g \in T(V)$ is defined by

$$f(k) \triangleq \begin{cases} w_1 \otimes \cdots \otimes w_n & \text{if } k = n \\ 0 & \text{if otherwise} \end{cases} \text{ and } g(k) \triangleq \begin{cases} w_{n+1} \otimes \cdots \otimes w_{n+l} & \text{if } k = l \\ 0 & \text{if otherwise} \end{cases}$$

then

$$(fg)(k) = \begin{cases} w_1 \otimes \cdots \otimes w_{n+l} & \text{if } n = k + l \\ 0 & \text{if otherwise} \end{cases}$$

does form an associative algebra with multiplication identity $1 \in \mathbb{R}$. Thus, $T(V)$ is in fact a ring. Let $I(V) \subseteq T(V)$ be the ideal generated by $\{v \otimes v : v \in V\}$. By definition, ideal $I(V)$ is a subgroup of $T(V)$. To see that $I(V)$ is closed under scalar multiplication, observe that for all $t \in \mathbb{R}$ and $x \in T(V)$, the scalar multiplication tx is identical to tx where t is treated as an element of $T(V)$, so it follows from definition of ideal that $I(V)$ is also a vector subspace of $T(V)$. Let $\{v_1, \dots, v_n\}$ be a basis for V, and let S be the set of function that maps $\{1, \dots, n\}$ into $\{1, \dots, k\}$. We know for a fact that

$$T^k(V) = \operatorname{span}\{v_{I(1)} \otimes \cdots \otimes v_{I(k)} : I \in S\}$$

If we define $I^k(V) \triangleq I(V) \cap T^k(V)$, one then have

$$I^k(V) = \operatorname{span}\{v_{I(1)} \otimes \cdots \otimes v_{I(k)} : I(j) = I(j+1) \text{ for some } j\} \tag{1.1}$$

This is proved by showing $I^0(V) \oplus I^1(V) \oplus I^2(V) \oplus \cdots$ is indeed the smallest ideal containing $\{v \otimes v : v \in V\}$. Define an equivalence class on $T(V)$ by

$$x \sim y \overset{\triangle}{\iff} x - y \in I(V)$$

Because ideal form a subgroup, we see that our definition indeed give an equivalence relation. We then can define on the set of equivalence class $T(V) \setminus I(V)$ addition, scalar multiplication and vector multiplication

$$[x] + [y] \triangleq [x + y] \text{ and } [x] \wedge [y] \triangleq [xy] \text{ and } c[x] \triangleq [cx]$$

which is well defined and form an algebra as one can check. We call this algebra $T(V) \setminus I(V)$ the **exterior algebra** $\wedge^*(V)$. Note that if we refer to $v, w \in T^k(V)$ as elements of $\wedge^*(V)$, we mean $[v], [w]$. Immediately, we see that the wedge product is **alternating** in the sense that if $v \in V$, then

$$v \wedge v = 0$$

and is **anti-symmetric** in the sense that if $v, w \in V$, then

$$v \wedge w = -w \wedge v$$

We use the notation

$$\wedge^k(V) \triangleq \left\{ [x] \in \wedge^*(V) : x \in \overbrace{V \otimes \cdots \otimes V}^{k \text{ copies}} \right\}$$

Immediately from Equation 1.1, we see that $\wedge^k(V)$ is the vector space

$$\operatorname{span}\{v_{I(1)} \wedge \cdots \wedge v_{I(k)} : I \in S\}$$

where $\{v_1, \dots, v_n\}$ is a basis for V and S is the set of function that maps $\{1, \dots, k\}$ into $\{1, \dots, n\}$. If we define the vector subspace $I^k(V) \triangleq T^k(V) \cap I(V)$, there exists a natural vector space isomorphism

$$\wedge^k(V) \underset{\text{v.s.}}{\cong} T^k(V) \setminus I^k(V); [x] \leftrightarrow [x]$$

where $T^k(V) \setminus I^k(V)$ is the quotient vector space.

Theorem 1.3.3. (Universal mapping property for alternating k-linear map) For any vector space Z over $\mathbb{R}$ and any alternating k-linear map $f : V^k \to Z$, there is a unique linear map $F : \bigwedge^k V \to Z$ such that the diagram

$$\begin{tikzcd}
\bigwedge^k V \arrow[dr, "F"] & \\
V^k \arrow[u, "\wedge"] \arrow[r, "f"'] & Z
\end{tikzcd}$$

commute, i.e.,

$$F(v_1 \wedge \cdots \wedge v_k) = f(v_1, \dots, v_k) \text{ for all } v_1, \dots, v_k \in V$$

Proof. By universal property of tensor product, there exists unique linear map $h : T^k(V) \to Z$ such that

$$h(v_1 \otimes \cdots \otimes v_k) = f(v_1, \cdots, v_k)$$

Because f is alternating, we see from the characterization of $I^k(V)$ given in Equation 1.1 that h vanishes on $I^k(V)$. We then can induce a linear map

$$F : \wedge^k(V) \cong \frac{T^k(V)}{I^k(V)} \to Z$$

by $F([x]) \triangleq h(x)$. This then give us the desired

$$F(v_1 \wedge \cdots \wedge v_k) = h(v_1 \otimes \cdots \otimes v_k) = f(v_1, \dots, v_k)$$

Note that F is unique because all such linear map take the same values on $\{v_{I(1)} \wedge \cdots \wedge v_{I(k)} : I \in S\}$ which spans $\wedge^k(V)$. ■

Let $\{w_1,\dots,w_l\}\subseteq V$ be linear independent. An immediate consequence of the universal mapping property for alternating k-linear map is that one may define alternating multilinear $f:V^l\to\mathbb{R}$ by

$$B(v_1,\dots,v_l)\triangleq\det(M)\text{ where }v_i=\sum_j M_{i,j}w_j$$

and see that $F:\wedge^l(V)\to\mathbb{R}$ take $w_1\wedge\cdots\wedge w_l$ to 1. This implies that

$$w_1\wedge\cdots\wedge w_l\neq 0$$

Theorem 1.3.4. (Anti-symmetry of wedge product) If $\alpha\in\wedge^k(V),\beta\in\wedge^l(V)$, then $\alpha\wedge\beta=(-1)^{kl}(\beta\wedge\alpha)$.

Proof. Let $v_1,\dots,v_n$ be a basis of V. Let S_k be the space of function that maps $\{1,\dots,k\}$into $\{1,\dots,n\}$, S_l be the space of function that maps $\{1,\dots,l\}$ into $\{1,\dots,n\}$. We may then write

$$\alpha=\sum_{I\in S_k}a_I(v_{I(1)}\wedge\cdots\wedge v_{I(k)})\text{ and }\beta=\sum_{J\in S_l}b_J(v_{J(1)}\wedge\cdots\wedge v_{J(l)})$$

and compute

$$\begin{aligned}\alpha\wedge\beta&=\sum_{I\in S_k,J\in S_l}a_Ib_J(v_{I(1)}\wedge\cdots\wedge v_{I(k)}\wedge v_{J(1)}\wedge\cdots\wedge v_{J(l)})\\&=\sum_{I\in S_k,J\in S_l}(-1)a_Ib_J(v_{I(1)}\wedge\cdots\wedge v_{J(1)}\wedge v_{I(k)}\cdots\wedge\cdots\wedge v_{J(l)})\\&=\sum_{I\in S_k,J\in S_l}(-1)^ka_Ib_J(v_{J(1)}\wedge v_{I(1)}\wedge\cdots\wedge v_{I(k)}\wedge v_{J(2)}\wedge\cdots\wedge v_{J(l)})\\&=\sum_{I\in S_k,J\in S_l}(-1)^{kl}a_Ib_J(v_{J(1)}\wedge\cdots\wedge v_{J(l)}\wedge v_{I(1)}\wedge\cdots\wedge v_{I(k)})=(-1)^{kl}\beta\wedge\alpha\end{aligned}$$

∎

Following from Theorem 1.3.4, Equation 1.1 and tedious effort, one can see that if $\{v_1,\dots,v_n\}$ is a basis for V, then

$$\{v_{i_1}\wedge\cdots\wedge v_{i_k}:1\leq i_1<\cdots<i_k\leq n\}$$

form a basis for $\wedge^k(V)$. If $A:V\to W$ is a linear map, we define linear map $\wedge^kA:\wedge^k(V)\to\wedge^k(W)$ by linear extension of

$$\wedge^k(A)(v_1\wedge\cdots\wedge v_n)=Av_1\wedge\cdots\wedge Av_n$$

Note that if $A : V \to V$ and $\dim(V) = n$, then $\wedge^n A : \wedge^n(V) \to \wedge^n(V)$ is given by the determinant since given basis $\{v_1, \dots, v_n\}$, we have

$$\begin{aligned}\wedge^n A(v_1 \wedge \cdots \wedge v_n) &= \Big(\sum_j A_{j,1} v_j\Big) \wedge \cdots \wedge \Big(\sum_j A_{j,n} v_j\Big) \\ &= \sum_{\sigma \in S_n} A_{\sigma(1),1} \cdots A_{\sigma(n),n} v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{\sigma(1),1} \cdots A_{\sigma(n),n} v_1 \wedge \cdots \wedge v_n\end{aligned}$$

1.4 Norm and Inner Product

This section contains

(a) definition and basic properties of the term **norm**

(b) definition and basic properties of the term **inner product**

(c) definition and basic properties of the term **positive semi-definite Hermitian form**

(d) full statement and proof of **Cauchy Schwarz Inequality** for both inner product space and positive semi-definite Hermitian form

(e) statement and proof of **SVD** (singular value decomposition).

(Norm Axiom Part)

Recall that by a **normed space** V, we mean a vector space over a sub-field $\mathbb{F}$ of $\mathbb{C}$ equipped with $\|\cdot\| : V \to \mathbb{R}_0^+$ satisfying the following axioms:

(a) $\|x\| = 0 \implies x = 0$ (positive-definiteness)

(b) $\|sx\| = |s| \cdot \|x\|$ for all $s \in \mathbb{F}$ and $x \in V$ (absolute-homogenity)

(c) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ (triangle inequality)

Observe

$$\|0\| = \|0 + x\| \leq \|0\| + \|x\| \text{ for all } x \in V$$

This shows that $\|x\| \geq 0$ for all $x \in V$. Also observe

$$\|0\| = \|0(x)\| = |0| \cdot \|x\| = 0$$

We can now rewrite the normed space axioms into

(a) $\|x\| = 0 \iff x = 0$ (positive-definiteness)

(b) $\|sx\| = |s| \cdot \|x\|$ for all $s \in \mathbb{F}$ and $x \in V$ (absolute-homogeneity)

(c) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ (triangle inequality)

(d) $\|x\| \geq 0$ for all $x \in V$ (non-negativity)

(Inner Product Axiom Part)

Recall that by an **inner product space** V, we mean a vector space over $\mathbb{R}$ or $\mathbb{C}$ equipped with $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{R}$ or $\mathbb{C}$ satisfying the following axioms

(a) $\langle x, x \rangle > 0$ for all $x \neq 0$ (Positive-definiteness)

(b) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (Conjugate symmetry)

(c) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ and $\langle cx, z \rangle = c\langle x, z \rangle$ (Linearity in the first argument)

Note that conjugate symmetry let us deduce

$$\langle x, x \rangle = \overline{\langle x, x \rangle} \implies \langle x, x \rangle \in \mathbb{R}$$

Also, one can easily use linearity in first argument to deduce

$$\langle 0, 0 \rangle = 2\langle 0, 0 \rangle \implies \langle 0, 0 \rangle = 0$$

This now let us rewrite the inner product space over $\mathbb{C}$ axioms into

(a) $\langle x, x \rangle \geq 0$ for all $x \in V$ (non-negativity)

(b) $\langle x, x \rangle = 0 \iff x = 0$ (positive-definiteness)

(c) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (conjugate symmetry)

(d) $\langle cx + y, z \rangle = c\langle x, z \rangle + \langle y, z \rangle$ and $\langle x, cy + z \rangle = \overline{c}\langle x, y \rangle + \langle x, z \rangle$ (Linearity)

Note that using $c = 1$ and $y = 0$, ($\because \langle 0, z \rangle = 0\langle x, z \rangle = 0$) one can check that the latter expression of linearity implies the first expression.

If the scalar field is $\mathbb{R}$, then conjugate symmetry is just symmetry and we also have linearity in the second argument.

This now let us rewrite the inner product space over $\mathbb{R}$ axioms into

(a) $\langle x, x \rangle \geq 0$ for all $x \in V$ (non-negativity)

(b) $\langle x, x \rangle = 0 \iff x = 0$ (positive-definiteness)

(c) $\langle x, y \rangle = \langle y, x \rangle$ (symmetry)

(d) Linearity in both arguments

If we do not require $\langle \cdot, \cdot \rangle$ to be positive-definite, but only non-negative, i.e. $\langle x, x \rangle \geq 0$ for all $x \in V$, then we have a **positive semi-definite Hermitian form**. Formally speaking, a positive semi-definite Hermitian form $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{R}$ or $\mathbb{C}$ satisfy the following axioms

(a) $\langle x, x \rangle \geq 0$ for all $x \in V$ (non-negativity)

(b) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (conjugate symmetry)

(c) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ and $\langle cx, z \rangle = c\langle x, z \rangle$ (Linearity in the first argument)

Example 1 (Example of Positive semi-definite Hermitian form)

$$\text{arbitrary } V \text{ over } \mathbb{R} \text{ or } \mathbb{C} \qquad \langle x, y \rangle \triangleq 0 \text{ for all } x, y$$

(Norm Induce Part)

Given a vector space V over $\mathbb{R}$ or $\mathbb{C}$, one can check that if V is equipped with an inner product $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{R}$ or $\mathbb{C}$, then we can induce a norm on V by

$$\|x\| \triangleq \sqrt{\langle x, x \rangle} \qquad (x \in V)$$

Note that

$$\|x\| = 0 \iff \langle x, x \rangle = 0$$

This implies that if $\langle \cdot, \cdot \rangle$ is an inner product (satisfy positive-definiteness), then $\| \cdot \|$ is also positive-definite. And if $\langle \cdot, \cdot \rangle$ is not positive-definite, then there exists $x \neq 0 \in V$ such that $\|x\| = 0$, which make $\| \cdot \|$ a **semi-norm**.

Absolute homogeneity follows from the linearity of inner product.

To check triangle inequality, we first have to prove Cauchy-Schwarz inequality.

Theorem 1.4.1. (Basic Property of Positive semi-definite Hermitian form) Given a positive semi-definite Hermitian form $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{R}$ or $\mathbb{C}$ and $x, y \in V$, we have

$$\langle x, x \rangle = 0 \implies \langle x, y \rangle = 0$$

Proof. Assume $\langle x, y\rangle \neq 0$. Fix $t > \frac{\|y\|^2}{2|\langle x,y\rangle|^2}$. Compute

$$\begin{aligned}\|y - t\langle y, x\rangle x\|^2 &= \|y\|^2 + \|(-t)\langle y, x\rangle x\|^2 + \langle -t\langle y, x\rangle x, y\rangle + \langle y, -t\langle y, x\rangle x\rangle \\ &= \|y\|^2 + t^2 \left|\langle x, y\rangle\right|^2 \|x\|^2 - t\langle y, x\rangle\langle x, y\rangle - t\langle x, y\rangle\langle y, x\rangle \\ &= \|y\|^2 - 2t \left|\langle x, y\rangle\right|^2 < 0 \text{ CaC}\end{aligned}$$

■

Theorem 1.4.2. (Cauchy-Schwarz Inequality) Given a positive semi-definite Hermitian form $\langle\cdot,\cdot\rangle : V^2 \to \mathbb{C}$ on vector space V over $\mathbb{C}$, we have

(a) $|\langle x, y\rangle| \leq \|x\| \cdot \|y\| \quad (x, y \in V)$

(b) the equality hold true if x, y are linearly dependent

(c) the equality hold true if and only if x, y are linearly dependent (provided $\langle\cdot,\cdot\rangle$ is an inner product)

Proof. We first prove

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\| \qquad (x, y \in V)$$

Fix $x, y \in V$. Theorem 1.4.1 tell us $\|x\| = 0 \implies \langle x, y\rangle = 0$. Then we can reduce the problem into proving

$$\frac{|\langle x, y\rangle|^2}{\|x\|^2} \leq \|y\|^2$$

Set $z \triangleq y - \frac{\langle y,x\rangle}{\|x\|^2}x$. We then have

$$\langle z, x\rangle = \langle y - \frac{\langle y, x\rangle}{\|x\|^2}x, x\rangle = \langle y, x\rangle - \frac{\langle y, x\rangle}{\|x\|^2}\langle x, x\rangle = 0$$

Then from $y = z + \frac{\langle y,x\rangle}{\|x\|^2}x$, we can now deduce

$$\begin{aligned}\langle y, y\rangle &= \langle z + \frac{\langle y, x\rangle}{\|x\|^2}x, z + \frac{\langle y, x\rangle}{\|x\|^2}x\rangle \\ &= \langle z, z\rangle + \left|\frac{\langle y, x\rangle}{\langle x, x\rangle}\right|^2 \langle x, x\rangle \\ &= \langle z, z\rangle + \frac{|\langle x, y\rangle|^2}{\langle x, x\rangle}\end{aligned}$$

Because $\langle z,z\rangle \geq 0$, we now have

$$\langle y,y\rangle = \langle z,z\rangle + \frac{|\langle x,y\rangle|^2}{\langle x,x\rangle} \geq \frac{|\langle x,y\rangle|^2}{\langle x,x\rangle} \text{ (done)}$$

The equality hold true if and only if $\langle z,z\rangle = 0$. This explains the other two statements regarding the equality. ∎

The proof is clearly geometrical. If one wish to remember the proof, one should see the trick we use is exactly

$$z \triangleq y - |y|\,(\cos\theta)\hat{x} \text{ is the projection of } y \text{ onto } x^{\perp}$$

CSp.jpeg

Then all we do rest is just expanding $|y|^2 = |z+\widetilde{x}|^2$, where $\widetilde{x} = y - z = |y|\,(\cos\theta)\hat{x}$, which give the answer and is easy to compute since $z \cdot \widetilde{x} = 0$.

Now, with Cauchy-Schwarz Inequality, we can check the triangle inequality

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \langle x,x\rangle + \langle y,y\rangle + 2 \text{ Re } \langle x,y\rangle \\
&\leq \|x\|^2 + \|y\|^2 + 2\,|\langle x,y\rangle| \\
&\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = \Big(\|x\| + \|y\|\Big)^2
\end{aligned}$$

(Euclidean Space Abstract Part)

By a **concrete Euclidean Space**, we mean some space of n-tuple $(x_1,\dots,x_n)$ over $\mathbb{R}$, equipped with inner product $\langle\cdot,\cdot\rangle_E$ defined by

$$\langle(x_1,\dots,x_n),(y_1,\dots,y_n)\rangle_E=\sqrt{\sum_{k=1}^{n}(y_k-x_k)^2}$$

By an **Euclidean Space**, we simply mean a finite dimensional vector space V over $\mathbb{R}$, equipped with an inner product $\langle\cdot,\cdot\rangle$ such that there exists a concrete Euclidean space E and an isomorphism $\varphi : V\to E$ such that

$$\langle x,y\rangle=\langle\varphi(x),\varphi(y)\rangle_E \qquad (x,y\in V)$$

Note that if you define $\langle\cdot,\cdot\rangle$ on the space of n-tuples $(x_1,\dots,x_n)$ over $\mathbb{R}$ by

$$\langle(x_1,\dots,x_n),(y_1,\dots,y_n)\rangle=2\sqrt{\sum_{k=1}^{n}(y_k-x_k)^2}$$

Then, the space of n-tuple is clearly not a concrete Euclidean space, and clearly an Euclidean space.

(SVD)

1.5 Operator Norm

Abstract

This section introduces the concept of the operator norm and proves some fundamental results related operator norm and finite-dimensional normed spaces. For example, we establish results such as a linear operator being bounded if and only if it is continuous and the equivalence of all norms on finite-dimensional vector spaces.

In this section, and particularly in functional analysis, we say a function T between two metric space is a **bounded operator** if T always map bounded set to bounded set. In particular, if T is a linear transformation between two normed space, we say T is a **bounded linear operator**. Now, suppose $\mathcal{X}, \mathcal{Y}$ are two normed space over $\mathbb{R}$ or $\mathbb{C}$. In space $L(\mathcal{X}, \mathcal{Y})$, alternatively, we can define

$$T \text{ is bounded } \overset{\triangle}{\iff} \exists M \in \mathbb{R}, \forall x \in \mathcal{X}, \|Tx\| \leq M\|x\|$$

The proof of equivalency is simple. For $(\longrightarrow)$, observe

$$\|Tx\| = \|x\| \cdot \|T\frac{x}{\|x\|}\| \leq \Big(\sup\{\|Ty\| : \|y\| = 1\}\Big)\|x\|$$

For $(\longleftarrow)$, observe

$$\|Tx - Ty\| = \|T(x - y)\| \leq M\|x - y\|$$

We first show that a linear transformation is continuous if and only if it is bounded.

Theorem 1.5.1. (Liner Operator is Bounded if and only if it is Continuous) Given two normed space $\mathcal{X}, \mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$ and $T \in L(\mathcal{X}, \mathcal{Y})$, we have

$$T \text{ is a bounded operator } \iff T \text{ is continuous on } \mathcal{X}$$

Proof. If T is bounded, we see that T is Lipschitz.

$$\|Tx - Ty\| \leq M\|x - y\|$$

Now, suppose T is linear and continuous at 0. Let ϵ satisfy

$$\sup_{\|y\| \leq \epsilon} \|Ty\| \leq 1$$

Observe that for all $x \in \mathcal{X}$, we have

$$\|Tx\| = \frac{\|x\|}{\epsilon}\Big\|T\frac{\epsilon x}{\|x\|}\Big\| \leq \frac{\|x\|}{\epsilon}$$

■

Here, we introduce a new terminology, which shall later show its value. Given a set X, we say two metrics d_1, d_2 on X are **equivalent**, and write $d_1 \sim d_2$, if we have

$$\exists m, M \in \mathbb{R}^+, \forall x, y \in X, md_1(x, y) \leq d_2(x, y) \leq Md_1(x, y)$$

Now, given a fixed vector space V, naturally, we say two norms $\|\cdot\|_1, \|\cdot\|_2$ on V are **equivalent** if

$$\exists m, M \in \mathbb{R}^+, \forall x \in X, m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1$$

We say two metric d_1, d_2 on X are **topologically equivalent** if the topology they induce on X are identical.

A few properties can be immediately spotted.

(a) Our definition of $\sim$ between metrics of a fixed X is an equivalence relation.

(b) Our definition of $\sim$ between norms on a fixed V is an equivalence relation.

(c) Equivalent norms induce equivalent metrics.

(d) Equivalent metrics are topologically equivalent.

We now prove if V is finite-dimensional, then all norms on V are equivalent. This property will later show its value, as used to prove linear map of finite-dimensional domain is always continuous

Theorem 1.5.2. (All Norms on Finite-dimensional space are Equivalent) Suppose V is a finite-dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$. Then

$$\text{all norms on } V \text{ are equivalent}$$

Proof. Let $\{e_1, \dots, e_n\}$ be a basis of V. Define ∞-norm $\|\cdot\|_\infty$ on V by

$$\Big\|\sum \alpha_i e_i\Big\|_\infty \triangleq \max |\alpha_i|$$

It is easily checked that $\|\cdot\|_\infty$ is indeed a norm. Fix a norm $\|\cdot\|$ on V. We reduce the problem into

$$\text{finding } m, M \in \mathbb{R}^+ \text{ such that } m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty$$

We first claim

$$M = \sum \|e_i\| \text{ suffices}$$

Compute

$$\|x\| = \Big\| \sum \alpha_i e_i \Big\| \leq \sum |\alpha_i| \, \|e_i\| \leq \|x\|_\infty \sum \|e_i\| = M\|x\|_\infty \text{ (done)}$$

Note that reverse triangle inequality give us

$$\Big| \|x\| - \|y\| \Big| \leq \|x - y\| \leq M\|x - y\|_\infty \tag{1.2}$$

Then we can check that

(a) $\|\cdot\| : \Big(V, \|\cdot\|_\infty\Big) \to \mathbb{R}$ is Lipschitz continuous because of Equation 1.2.

(b) $S \triangleq \{y \in V : \|y\|_\infty = 1\}$ is sequentially compact in $\|\cdot\|$ and non-empty.

Now, by EVT, we know $\min_{y \in S} \|y\|$ exists. Note that $\min_{y \in S} \|y\| > 0$, since $0 \notin S$. We claim

$$m = \min_{y \in S} \|y\| \text{ suffices}$$

Fix $x \in V$ and compute

$$m\|x\|_\infty = \|x\|_\infty (\min_{y \in S} \|y\|) \leq \|x\|_\infty \cdot \Big\| \frac{x}{\|x\|_\infty} \Big\| = \|x\| \text{ (done) (done)}$$

$\blacksquare$

Theorem 1.5.3. (Linear map of Finite-dimensional Domain is always Continuous) Given a finite-dimensional normed space $\mathcal{X}$ over $\mathbb{R}$ or $\mathbb{C}$, an arbitrary normed space $\mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$ and a linear transformation $T : \mathcal{X} \to \mathcal{Y}$, we have

$$T \text{ is continuous}$$

Proof. Fix $x \in \mathcal{X}, \epsilon$. We wish

$$\text{to find } \delta \text{ such that } \forall h \in \mathcal{X} : \|h\| \leq \delta, \|T(x + h) - Tx\| \leq \epsilon$$

Let $\{e_1, \dots, e_n\}$ be a basis of $\mathcal{X}$. Note that $\|\sum \alpha_i e_i\|_1 \triangleq \sum |\alpha_i|$ is a norm. Because $\mathcal{X}$ is finite-dimensional, we know $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent. Then, we can fix $M \in \mathbb{R}^+$ such that

$$\|x\|_1 \leq M\|x\| \quad (x \in V)$$

We claim

$$\delta = \frac{\epsilon}{M(\max \|Te_i\|)} \text{ suffices}$$

Fix $\|h\| \leq \delta$ and express $h = \sum \alpha_i e_i$. Compute using linearity of T

$$\begin{aligned}
\|T(x+h) - Tx\| &= \|\sum \alpha_i T e_i\| \\
&\leq \sum |\alpha_i| \, \|Te_i\| \\
&\leq \|h\|_1 (\max \|Te_i\|) \\
&\leq M\|h\|(\max \|Te_i\|) = \epsilon \text{ (done)}
\end{aligned}$$

■

We now see that, because Linear transformation is bounded if and only if it is continuous and Linear map of finite-dimensional domain is always continuous, if $\mathcal{X}$ is finite-dimensional, then all linear map of domain $\mathcal{X}$ are bounded. A counter example to the generalization of this statement is followed.

Example 2 (Differentiation is an Unbounded Linear Operator)

$$\mathcal{X} = \Big(\mathbb{R}[x]|_{[0,1]}, \|\cdot\|_\infty\Big), D(P) \triangleq P'$$

Note that $\{x^n\}_{n\in\mathbb{N}}$ is bounded in $\mathcal{X}$ and $\{D(x^n)\}_{n\in\mathbb{N}}$ is not.

Now, suppose $\mathcal{X}, \mathcal{Y}$ are two fixed normed spaces over $\mathbb{R}$ or $\mathbb{C}$. We can easily check that the set $BL(\mathcal{X}, \mathcal{Y})$ of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ form a vector space over whichever field $\mathcal{Y}$ is over.

Naturally, our definition of boundedness of linear operator derive us a norm on $BL(\mathcal{X}, \mathcal{Y})$, as followed

$$\|T\|_{\text{op}} \triangleq \inf\{M \in \mathbb{R}^+ : \forall x \in \mathcal{X}, \|Tx\| \leq M\|x\|\} \tag{1.3}$$

Before we show that our definition is indeed a norm, we first give some equivalent definitions and prove their equivalency.

Theorem 1.5.4. (Equivalent Definitions of Operator Norm) Given two fixed normed space $\mathcal{X}, \mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$, a bounded linear operator $T : \mathcal{X} \to \mathcal{Y}$, and define $\|T\|_{\text{op}}$ as in Equation 1.3, we have

$$\|T\|_{\text{op}} = \sup_{x\in\mathcal{X}, x\neq 0} \frac{\|Tx\|}{\|x\|}$$

Proof. Define $J \triangleq \{M \in \mathbb{R}^+ : \forall x \in \mathcal{X}, \|Tx\| \leq M\|x\|\}$ and observe

$$J = \{M \in \mathbb{R}^+ : M \geq \frac{\|Tx\|}{\|x\|}, \forall x \neq 0 \in \mathcal{X}\}$$

This let us conclude

$$\sup_{x \in \mathcal{X}, x \neq 0} \frac{\|Tx\|}{\|x\|} = \min J = \|T\|_{\text{op}}$$

■

It is now easy to see

$$\|T\|_{\text{op}} = \sup_{x \in \mathcal{X}, x \neq 0} \frac{\|Tx\|}{\|x\|} \tag{1.4}$$

$$= \sup_{x \in \mathcal{X}, \|x\|=1} \|Tx\| \tag{1.5}$$

It is not all in vain to introduce the equivalent definitions. See that the verification of $\|\cdot\|_{\text{op}}$ being a norm on $BL(\mathcal{X}, \mathcal{Y})$ become simple by utilizing the equivalent definitions.

(a) For positive-definiteness, fix non-trivial T and fix $x \in \mathcal{X} \setminus N(T)$. Use Equation 1.4 to show $\|T\|_{\text{op}} \geq \frac{\|Tx\|}{\|x\|} > 0$.

(b) For absolute homogenity, use Equation 1.5 and $\|Tcx\| = |c| \cdot \|Tx\|$.

(c) For triangle inequality, use Equation 1.5 and $\|(T_1 + T_2)x\| \leq \|T_1x\| + \|T_2x\|$.

Naturally, and very very importantly, Equation 1.4 give us

$$\|Tx\| \leq \|T\|_{\text{op}} \cdot \|x\| \qquad (x \in \mathcal{X})$$

This inequality will later be the best tool to help analyze the derivatives of functions between Euclidean spaces, and perhaps better, it immediately give us

$$\frac{\|T_1T_2x\|}{\|x\|} \leq \frac{\|T_1\|_{\text{op}} \cdot \|T_2\|_{\text{op}} \cdot \|x\|}{\|x\|} = \|T_1\|_{\text{op}} \cdot \|T_2\|_{\text{op}}$$

Then Equation 1.4 give us

$$\|T_1T_2\|_{\text{op}} \leq \|T_1\|_{\text{op}} \cdot \|T_2\|_{\text{op}}$$