Linear Algebra

Anji Yu

2010 *Mathematics Subject Classification.* Primary

Abstract.

Contents

CHAPTER 1

Vector Space

1.1. Fields

Definition 1.1. A division ring is a set of element F together with a pair of binary operators $(+, \cdot)$ referred as addition and multiplication respectively such that there is 0 and 1 in F satisfying F is an Abelian group under $+$ with identity 0; $F^* \equiv F\backslash\{0\}$ is a group under $\cdot$ with identity 1; and the two operators satisfy 2-sided distribution law,

$$\begin{aligned}(a+b)\cdot c &= a\cdot c + b\cdot c,\\ c\cdot(a+b) &= c\cdot a + c\cdot b.\end{aligned}$$

A field is a division ring whose multiplication is Abelian.

Note that if we set $b = 0$ to above equations we get $a \cdot 0 = 0 \cdot a = 0$ for every $a \in F$.

1.2. Vector space

Definition 1.2. An Abelian group V with a field F is called a vector space if

When there is no ambiguity, we might omit the base field F, and only call a vector space over F a vector space.

1.3. Linear dependence

This is a good place to talk about some conventions. When we write $\{v_i\}$ we assume there exists an corresponding index set I, and i runs over I. The order of I will be specified if necessary. If $|I| = n$, then I is regarded as a set of natural number form 1 to n. We shall carefully distinguish index letter such as $\{v_i\}$ and $\{u_j\}$, using same (resp. different) letter means taking index from same (resp. different) index set. We will adopt Einstein summation convention, which states that dummy index in an expression such as $a_i v_i$ is summation through the index set I, i.e., $a_i v_i \equiv \sum_{i\in I} a_i v_i$.

Definition 1.3. Let $\{v_i\}$ be a finite set of vectors, $\{v_i\}$ *linearly independent* if it is empty or for any set of scalars $\{a_i\}$ satisfying $\sum_i a_i v_i = 0$ implies $a_i = 0$ for each i. Otherwise $\{v_i\}$ is *linearly dependent.* A vector v is a *linear combination* of $\{v_i\}$ if there exists a set of scalars $\{a_i\}$ such that $v = a_i v_i$.

Union a linearly dependent set with any other vector set results in a linearly dependent set. On the contrary, any subset of linearly independent set is linearly independent.

Theorem 1.4. *Let $X \equiv \{x_i\}$ be a set of n vectors and $Y \equiv \{y_j\}$ be a set of $n+1$ vectors whose elements are all linear combinations of X, then Y is linearly dependent.*

Proof. If $n = 1$, we can find scalars α and β such that $y_1 = \alpha x_1$ and $y_2 = \beta x_1$, apparently they are linearly dependent. Suppose $n > 1$. Let $X' \equiv X \backslash \{x_n\}$ and $\{a_j\}$ be a set of scalars such that $y'_j \equiv y_j - a_j x_n$ is a linear combination of X' for each j. If $a_j = 0$ for each j, by induction hypothesis on $n-1$ we are done. Otherwise WLOG we assume $a_{n+1} = 1$. x_n is linear combination of $X' \cup \{y_{n+1}\}$. Let $y''_i = y_i - a_i y_{n+1}$, then $\{y''_i\}$ is a linear combination of X' for each $i \in \{1, 2 \ldots n\}$, due to induction hypothesis we have a set of not-all-zero scalars $\{b_i\}$ such that

$$\sum_{i=1}^{n} b_i y''_i = 0 \Rightarrow \sum_{i=1}^{n} b_i y_i - \sum_{i=1}^{n} b_i a_i y_{n+1} = 0.$$

Hence Y is linearly dependent. □

A more useful version of the theorem is as follows, which will be used to define dimension of a vector space.

Corollary 1.4.1. If Y is linearly independent such that in which each element is a linear combination of another set X, then $|Y| \leq |X|$.

1.4. Basis and Dimension

Definition 1.5. An ordered set B of vectors of V is a *basis* if B is linearly independent and every vector $v \in V$ is a linear combination of B. V is *finitely generated* if V has a finite basis.

Suppose V is finitely generated and B is a finite basis, let B' be another basis of V, due to the fact that B' is linearly independent and every element in it is a linear combination of B, we have $|B'| \leq |B|$, so B' is finite. Similarly, $|B| \leq |B'|$. Therefore in a finitely generated vector space every basis has the same order. Which leads to the following definition,

Definition 1.6. The order of a basis of a finitely generated vector space V is called the *dimension* of V, denoted by $\dim V$.

From now on we shall call a finite generated vector space a finite space, or simply finite, and we shall only deal with finite vector space unless otherwise stated.

To find a basis of V starting by an linearly independent set B (possibly empty), we can keep adding on B by vectors which can not be written as linear combination of B step by step until there is none could be added, i.e., every vector in V is a linear combination of B and B is linearly independent. Note this process must terminate for a finite space. Therefore we find a maximal independent set, which is a basis. This process yields the fact that any linearly independent set can be extended to a basis. We will review this process in 1.5.1

Dimension of a vector space is invariant under different choices of basis. On the other hand, a vector space is determined by its dimension up to isomorphism. Not everything in vector space is basis invariant. For example, suppose $\{v_i\}$ is a basis of vector space V, every vector $u \in V$ can be written as a linear combination of $\{v_i\}$ in a unique way $u = a_i v_i$. $\{a_i\}$ is called the *coordinate* of u under basis $\{v_i\}$. We might also just call it the coordinate if the basis is specified without ambiguity. Apparently coordinates are not basis invariant.

1.5. Local structure

1.5.1. Subspace.

Definition 1.7. A subset W of V over F is a *subspace* if W is a vector space over F. If W and U are subspace of V, the sum of them $W + U$ is defined as $\{w + u \mid w \in W, u \in U\}$.

Set-theoretical intersection of two subspace is again a subspace, the sum of two subspace is also a subspace. We prove the latter and leave former as an exercise.

Theorem 1.8. *Let W and U be two subspaces of V, $W + U$ is a subspace.*

Proof. Let $v_i = w_i + u_i \in W + U$ for $i = 1, 2$, then for any scalar a, b, $av_1 + bv_2 = aw_1 + au_1 + bw_2 + bu_2 \in W + U$. □

Suppose S is a subset of V, the set of all linear combination of vectors in S is call the *span* of it, denoted by $\langle S \rangle$. By convention the span of $\emptyset$ is $\{0\}$. It is a routine to check S' is a vector space (over the same field), hence a subspace of V. Note the span is defined by a bottom-up fashion, it could equivalently defined in a top-down way. We shall proof the following

Theorem 1.9. *Let S be a subset of V and $\{W_i\}$ be the set of all the subspace containing S, then $\langle S \rangle = \cap_i W_i$*

Proof. $\langle S \rangle \subseteq W_i$ for each i, it follows that $\langle S \rangle \subseteq \cap_i W_i$. Since $\langle S \rangle \in \{W_i\}$, we obtain the other inclusion. □

The dimension of a subspace is no more than the original space.

1.5.2. Quotient space.

1.6. Isomorphism

Definition 1.10. Two vector spaces V and W over F are isomorphic if there exists a bijection $\tau \colon V \to W$, $v \to v^\tau$ such that for all vectors $u, v \in V$ and scalars a, b,

$$(au + bv)^\tau = au^\tau + bv^\tau$$

in this case we call τ an isomorphism.

Apply τ^{-1} on the above equation, we obtain that τ^{-1} is also linear. Hence a bijection τ is an isomorphism if and only if τ^{-1} is an isomorphism

CHAPTER 2

Linear operator

2.1. Linear map

Linear maps are those maps between vector spaces preserving linearity. We remark that we shall define the linear map as a right action in an exponential way, which is against most of modern text which adopt left action. Effectively many of the results we obtained from now on will be the transpose version of the corresponding results in those text.

Definition 2.1. Let V and W be vector spaces over F. A function τ is a *linear map* if for all vectors $u, v \in V$ and scalars a, b,

$$(au + bv)^\tau = au^\tau + bv^\tau.$$

The set of all linear maps from V to W is denoted by $L(V, W)$.

An identity is a map maps every vector to itself (so the target space is the same as domain space) and we denote identity of V by I_V, and sometimes I if there is no ambiguity. A zero map is a map maps every vector to zero. It is easy to see they are linear maps.

Two linear maps are equal if and only if they map every element to the same element. Hence identical linear maps map a basis to the same elements. But the inverse is also true,

Theorem 2.2. *A linear map is completely determined by its images on a basis of domain space.*

Proof. Suppose τ_i is a linear map $\tau \colon V \to W$, $v \to v^\tau$ for $i = 1, 2$ such that there exists a basis $\{v_i\}$ of V with the property $v_i^{\tau_1} = v_i^{\tau_2}$ for each v_i. Let $v \in V$ with coordinate $\{a_i\}$, we have

$$v^{\tau_1} = a_i v_i^{\tau_1} = a_i v_i^{\tau_2} = v^{\tau_2}.$$

$v^{\tau_1} = v^{\tau_2}$ for every $v \in V$, therefore $\tau_1 = \tau_2$. □

Suppose we have a map τ (not linear yet) from a basis $\{e_i\}$ of V to W, we can linearise τ in the way that for any $v = a_i e_i \in V$,

$$v^\tau = a_i e_i^\tau.$$

It is easy to verify that τ is indeed a linear map and τ is uniquely determined. More over, if τ maps to a basis $\{e'_j\}$ of W, we may define $\tau^{-1} \colon \{e'_j\} \to \{e_i\}$, $e_i^\tau \to e_i$ then linearise it.

Let τ be a linear map $\tau \colon V \to W$, we use $\ker \tau$ to denotes $\{v \in V \mid v^\tau = 0\}$, $\mathrm{im}\tau = \{v^\tau \mid v \in V\}$. It is easy to see they are vector space.

Definition 2.3. Let $\tau, \sigma \in L(V, W)$, We define $a\tau + b\sigma$ to be the map

$$v \to v^{a\tau + b\sigma} \equiv av^{\tau} + bv^{\sigma},$$

for each element $v \in V$.

Linearity of the above map could be verified as follows

$$\begin{aligned}(cv + du)^{a\tau + b\sigma} &= a(cv + du)^{\tau} + b(cv + dv)^{\sigma} \\ &= acv^{\tau} + adu^{\tau} + bcv^{\sigma} + bdu^{\sigma} \\ &= cv^{a\tau + b\sigma} + du^{a\tau + b\sigma}.\end{aligned}$$

Hence every linear combination of linear map is a linear map, $L(V, W)$ is a vector space.

2.2. Matrix representation of a linear map

Suppose τ is a linear map from V of dimension n to W of dimension m and we fix a basis $\{e_i\}$ and $\{e'_j\}$ on V and W respectively. A matrix representation of τ under the pair of basis is a set of scalars in a form of rectangular array

$$T = \begin{bmatrix} t_{11} & t_{12} & t_{13} & \dots & t_{1m} \\ t_{21} & t_{22} & t_{23} & \dots & t_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & t_{n3} & \dots & t_{nm} \end{bmatrix}$$

uniquely determined by the basis in the way that $e_i^{\tau} = t_{ij} e'_j$. In particular for fixed i, $\{t_{ij}\}$ is the coordinates of e_i^{τ} under basis $\{e'_j\}$.

Let's see a special representation of a linear map τ: Suppose $\{u_{r+1}, \dots, u_n\}$ is a basis of $\ker \tau$, then we extent it to a basis of V as $\{u_i\}$, next we find the corresponding element in W by $\{v_1^{\tau}, \dots, v_k^{\tau}\}$ and finally extent it to a basis of W. By doing that, τ has a representation matrix

$$T = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

where I_r denotes the identity matrix.

Since $L(V, W)$ is a vector space, every element in it also has a matrix representation. Suppose $\tau, \sigma \in L(V, W)$, for any scalar a, b, we have

$$e_i^{a\tau + b\sigma} = ae_i^{\tau} + be_i^{\sigma} = a\tau_{ij} e'_j + b\sigma_{ij} e'_j = (a\tau_{ij} + b\sigma_{ij}) e'_j,$$

therefore $(a\tau + b\sigma)_{ij} = a\tau_{ij} + b\sigma_{ij}$. If we consider a linear map composition $\tau \in L(V, W)$ and $\tau \in L(W, U)$, with basis representation under basis $\{e_i\}$, $\{e'_j\}$ and $\{e''_k\}$ in V, W and U respectively, then we have

$$e_i^{\tau\sigma} = \tau_{ik} e_k^{\prime\sigma} = \tau_{ik} \sigma_{kj} e''_j.$$

therefore $(\tau\sigma)_{ij} = \tau_{ik}\sigma_{kj}$.

Although these calculation rules come from linear map, but they have been made standard to the theory of matrix.

2.3. Linear operator

Most of the times we are interested in those linear maps that maps a space to itself. Formally,

DEFINITION 2.4. *Linear operators* of V are those linear maps in $L(V, V)$. We denote $L(V, V)$ as $\mathrm{End}(V)$.

If $\tau \in \mathrm{End}(V)$, we say τ acting on V.

Let $\tau, \sigma \in \mathrm{End}(V)$, the function composition of them is again in $\mathrm{End}(V)$, i.e., $\tau\sigma \in \mathrm{End}(V)$. Therefore additional to be as a vector space, we can have another binary operation defined on $\mathrm{End}(V)$. This is a hint that the endomorphisms has richer properties that usual linear maps.

Almost all the properties of linear map can be inherited by linear operators. But there is a specific thing we would like remark on. Consider a linear operator $\tau \in \mathrm{End}(V)$ and a basis $\{e_i\}$ with the corresponding matrix representation $e_i^\tau = \tau_{ij} e_j$. The matrix representation in this case only relates to one basis. With this in mind, suppose we have another basis $\{e_i'\}$ with a basis transformation $e_i' = \phi_{ij} e_j$, then τ acting on $\{e_i'\}$ as

$$e_i'^\tau = \phi_{ij} e_j^\tau = \phi_{ij} \tau_{jk} e_k = \phi_{ij} \tau_{jk} \phi_{kl}^{-1} e_l'.$$

A subspace W of V is *τ-invariant* if $w^\tau \in W$ for every $w \in W$. The *restriction* of τ in W is the action of τ on W, denoted by $\tau|_W$.

DEFINITION 2.5. A linear operator τ is *nilpotent* if there exists a natural number k such that $\tau^k = 0$.

CHAPTER 3

Multilinear form

CHAPTER 4

Associative algebra

CHAPTER 5

Projective geometry

CHAPTER 6

Lie algebra

6.1. Root system

Suppose E is an euclidean space, i.e., a finite dimensional vector space V over $\mathbb{R}$ endowed with a positive definite symmetric bilinear form $(,)$. The **dual** of $\alpha \in E$ is defined as $\alpha^\vee = 2\alpha/(\alpha, \alpha)$. For any given vector $\alpha \in E$, we define a hyperplane $P_\alpha = \{\beta \in E \mid (\alpha, \beta) = 0\}$, which is the perpendicular complement of α with respect to E. The reflection with respect to α is the linear map acts trivially on P_α and sends v to $-v$. Explicitly, $\sigma_\alpha(\beta) = \beta - (\beta, \alpha^\vee)\alpha$. Since $(\beta, \alpha^\vee)$ occurs frequently, we denote it by $\langle \beta, \alpha \rangle$.

A subset Φ of an euclidean space E is called a (reduced) root system on E if it satisfied

1. Φ is finite, excludes 0 and spans E.
2. $\mathbb{R}\alpha \cap \Phi = \pm\alpha$ for every $\alpha \in \Phi$.
3. $\sigma_\alpha(\beta) \in \Phi$ for every $\alpha, \beta \in \Phi$.
4. $\langle \alpha, \beta \rangle \in \mathbb{Z}$.

And denote (Φ, E) as the root system with respect to E. Implicitly there is an underlying vector space V corresponds to E. The dimension of V is called the rank of Φ.

Let $\alpha, \beta \in \Phi$, $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\cos^2\theta$ where θ is the angle between α and β in E. Hence $4\cos^2\theta$ is limited to 0, 1, 2, 3, 4. The case 4 is trivial. For the case 1, 2 and 3, θ is given as $\pi/6(5\pi/6)$, $\pi/4(3\pi/4)$ and $\pi/3(2\pi/3)$ respectively. And as divisor of these numbers, at least one of $\langle \alpha, \beta \rangle$ or $\langle \beta, \alpha \rangle$ is ± 1. Due to this observation, we have the following

Proposition 6.1. *Suppose α and β are non-proportional roots, $(\alpha, \beta) < 0$ (resp. $(\alpha, \beta) > 0$), then $\alpha + \beta \in \Phi$ (resp. $\alpha - \beta \in \Phi$).*

Proof. Suppose $(\alpha, \beta) < 0$ and $\langle \alpha, \beta \rangle = -1$, $\alpha + \beta = \sigma_\beta(\alpha) \in \Phi$; similarly for $\langle \beta, \alpha \rangle$. Similarly for $(\alpha, \beta) > 0$. □

As an application, let α and β be non-proportional roots. Consider the β-string through α, i.e., the set $S = \{\alpha + i\beta \mid i \in \mathbb{Z}\}$, which is invariant under σ_β, hence $S \cap \Phi$ is also invariant under σ_β. Let $q,\ r \in \mathbb{Z}$ be maximal that $\alpha + q\beta,\ \alpha - r\beta \in \Phi$, we must have $\sigma_\beta(\alpha - r\beta) = \alpha - \langle \alpha, \beta \rangle \beta + r\beta = \alpha + q\beta$, $\langle \alpha, \beta \rangle = r - q \leq 3$, so the length of the string is at most 4. We next show the string $S \cap \Phi$ is unbroken. Suppose otherwise, there is $i \in \mathbb{Z}$ and $-r < i < q$ such that the string is broken at $\alpha + i\beta$, we have $(\alpha + i_1\beta, \beta) \geq 0$ and $(\alpha + i_2\beta, \beta) \leq 0$ for some $i_1 < i < i_2$ due to proposition 6.1, which is absurd.

A root system Φ is **reducible** if it can be partitioned to nonempty orthogonal components, otherwise it is **irreducible**.

6.2. Weyl group

The following fact is useful,

LEMMA 6.2. *Let Φ be a finite set spans E, an euclidean space on V, in which all the reflections σ_α permute Φ. Let $\sigma \in GL(V)$ acts trivially on a hyperplane P in E and sends some $\alpha \in E$ to its negative. Then $\sigma = \sigma_\alpha$.*

PROOF. Let $\tau = \sigma\sigma_\alpha$. Both σ and σ_α acts as identity on $E/\mathbb{R}\alpha$, so does τ. Hence the minimal polynomial $m(T)$ of τ on E divides $(T-1)^{(\dim E)}$. Since τ permutes Φ, $\tau^{n!}$ acts trivially on Φ, where $n = |\Phi|$, it follows that $\tau^{n!} = 1$, i.e., $m(T)$ divides g.c.d$((T^{n!}-1), (T-1)^{(\dim E)}) = T-1$, $\tau = 1$. □

Let (Φ, E) be a root system, the dual of Φ is the set of all the dual of elements in Φ, denoted by $\Phi^\vee$. It is a root system in E. Call root system (Φ, E) and (Φ', E') **isomorphic** if there exists a vector space isomorphism $\tau\colon V \to V'$ sending $\Phi \to \Phi'$ and $\langle\alpha, \beta\rangle = \langle\tau(\alpha), \tau(\beta)\rangle$ for each pair of roots $\alpha, \beta \in \Phi$. τ preserves $\langle\ ,\ \rangle$ while is not necessary a isometry. It follows at once that $\tau(\sigma_\alpha(\beta)) = \sigma_{\tau(\alpha)}(\tau(\beta))$. When there is no ambiguity, we denote $\mathrm{Aut}(\Phi)$ the group of all automorphisms of (Φ, E). Each reflection $\sigma_\alpha(\alpha \in \Phi)$ is an automorphism as well as isometry. The group $\mathcal{W}(\Phi)$ generated by all the reflections σ_α is called **Weyl group** of the Φ. There is a canonical isomorphism between $\mathcal{W}(\Phi)$ and $\mathcal{W}(\Phi^\vee)$ even though Φ is sometimes not isomorphic to Φ'.

Suppose $\sigma \in GL(V)$ permutes Φ. Then $\sigma\sigma_\alpha\sigma^{-1}$ acts on $\sigma(\beta)$ as $\sigma\sigma_\alpha\sigma^{-1}\sigma(\beta)$ $= \sigma(\sigma_\alpha(\beta)) = \sigma(\beta) - \langle\alpha, \beta\rangle\sigma(\alpha)$. So $\sigma\sigma_\alpha\sigma^{-1}$ acts trivially on $\sigma(P_\alpha)$ and sends $\sigma(\alpha)$ to its negative, which is exactly the reflection $\sigma_{\sigma(\alpha)}$ by lemma 6.2, from which we conclude that $\mathcal{W}(\Phi) \triangleleft \mathrm{Aut}(\Phi)$. Moreover, $\sigma_{\sigma(\alpha)}\sigma(\beta) = \sigma(\beta) - \langle\sigma(\alpha), \sigma(\beta)\rangle\sigma(\alpha)$, it follows that $\langle\alpha, \beta\rangle = \langle\sigma(\alpha), \sigma(\beta)\rangle$, $\sigma \in \mathrm{Aut}(\Phi)$.

APPENDIX A

Set relation

APPENDIX B

Algebraic structure

Before given the definitions, we point out that we only concentrate on the binary operators. A algebraic structure is a set with one or more binary operators.

Bibliography