Vector Spaces and Linear Transformations

Wei Shi, Jinan University

2017.11.1

Definition (Field)

A field $F = \{F, +, \cdot\}$ is an algebraic structure formed by a set F, and closed under binary operations $+$ (addition) and $\cdot$ (multiplication). The addition operation

- is associative, commutative
- has a unique 0 and inverse element

The multiplication operation

- is associative, commutative and distributive
- has a unique 1
- has a unique inverse for nonzero element

Example

The set of real numbers together with $+$ and $\times$ is a field.

Definition (Vector Space)

A vector or linear space V defined over a field F is the set V (vectors) that is closed under binary operations $+$ (vector addition) and $\cdot$ (scalar multiplication). Vector addition satisfies

- associative, commutative
- has a unique 0 and inverse element

The scalar multiplication satisfies

- associative, distributive
- has a unique 1

Vector Space

Theorem

X is some nonempty set. The set of all functions $f : X \to F$ with

- $(f+g)(x) = f(x) + g(x)$
- $(\alpha f)(x) = \alpha f(x)$, $\alpha \in F$

is a vector space.

Example

In the above theorem,

- $X = \{1, 2, \cdots, n\}$ and $F = \mathbb{R}$: the vector space is the set of vectors in $\mathbb{R}^n$, denoted by $V_n(\mathbb{R})$
- $X = \mathbb{N}$ and $F = \mathbb{R}$: sequences
- $X = \{\{1, \cdots, m\} \times \{1, \cdots, n\}\}$ and $F = \mathbb{R}$: set of $m \times n$ matrices

Definition (Vector Subspace)

$(V, +, \cdot)$ is a vector space over the field F. $(S, +, \cdot)$ is a vector subspace of $(V, +, \cdot)$ if $S \subset V$, and for all $\alpha, \beta \in F$, and for all $x, y \in S$, $\alpha x + \beta y \in S$.

For a set $A \subset V$. Define span(A) as the set of all vectors that are linear combinations of elements in A.

Exercise

Show that span(A) is the intersection of all vector subspaces of V that contain A.

Linear Independence

Definition (Linear Independence)

A set of vectors $\{x_1, \cdots, x_n\}$ is linearly dependent if $\sum_{i=1}^{n} \alpha_i x_i = 0$ and $\alpha_i \neq 0$ for some i. It is linearly independent if $\sum_{i=1}^{n} \alpha_i x_i = 0$ implies that $\alpha_i = 0$ for all i.

Exercise

Show that $\{x_1, \cdots, x_n\}$ is linearly dependent if a subset of vectors is linearly dependent.

Basis

Definition (Basis)

A subset W *spans* or *generates* V if for all $x \in V$, $x = \sum_{i=1}^{n} \alpha_i w_i$, $w_i \in W$ and n is finite. A *Hamel basis* for a vector space V is a set of vectors that is linearly independent and spans V. The cardinal number of the Hamel basis is the dimension of the vector space V (dim V).

Denote a Hamel basis for vector space V as $\{v_s \in V, s \in S\}$ where S is an index set. For a vector $x \in V$, it can be represented by a linear combinations of vectors in the Hamel basis in the sense that $x = \sum_{i \in S'} \alpha_i v_i$ and S' is a finite subset of S.

Theorem

Every nonzero vector has a unique representation with respect to a Hamel basis.

Basis

Exercise

Consider the set $\{(x_1, x_2, x_3) \in \mathbb{R}^3, x_1 + x_2 = x_3\}$. *Show that it is a vector subspace of* $\mathbb{R}^3$ *and find a basis. What is its dimension?*

For a finite dimensional vector space (e.g. $\mathbb{R}^3$), every basis has the same number of elements, and every linearly independent set of vectors of dimension $dim(V)$ is a basis. We show this in the following two theorems.

Theorem

If $\{v_1, \cdots, v_n\}$ is a basis of V, then no set of more than n vectors is linearly independent.

Theorem

The vector space V has dimension n. Any linearly independent set of n vectors is a basis of V.

Linear Transformations

We first consider the general case with two vectors spaces X, Y defined over the same field F.

Definition (Linear Transformation)

A transformation $T : X \to Y$ is a linear transformation if for all $x_1, x_2 \in X$ and $\alpha, \beta \in F$,

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2).$$

Denote $L(X, Y)$ as the set of all linear transformations from X to Y. The following theorem establishes that $L(X, Y)$ is itself a vector space.

Theorem

The set of all linear transformations from X to Y is a vector space.

Consider a linear function $T : X \to Y$.

Definition (Image and Kernel)

- Image or range:
  $\text{im}(T) = T(X) = \{y \in Y, y = T(x) \text{ for some } x \in X\}$.
- Kernel of null space: $\ker(T) = \{x \in X, T(x) = 0\}$.

Theorem

For a linear transformation, $T(x)$ is a vector subspace of Y. If $\{v_1, \cdots, v_n\}$ is a basis for X, then $\{T(v_1), \cdots, T(v_n)\}$ spans $T(x)$.

Similar results also hold for the null space.

Theorem

For a linear transformation, $\ker(T)$ is a vector space of X.

An important connection between the dimensions of the image, kernel and the underlying vector space is established in the following theorem.

Theorem

$T : X \rightarrow Y$ is a linear transformation and X is a finite dimensional vector space.

$$dim\ (X) = dim\ ker(T) + dim\ im(T).$$

dim im(T) is also called the rank of T.

Inverse of a Linear Transformation

Definition

$T : X \to Y$ is a linear mapping. T is invertible if there exists $S : Y \to X$, such that

$$\forall x \in X, S(T(x)) = x$$
$$\forall y \in Y, T(S(y)) = y.$$

If such S exists, it is the inverse of T: $T^{-1} = S$.

From the definition, we can see that if T is invertible, it needs to be injective (one-to-one) and surjective (onto).

Theorem

A necessary and sufficient condition for T to be injective (one-to-one) is that its null space is $\{0\}$.

Inverse of a Linear Transformation

Another important property with the inverse of a linear transformation is that it is still linear.

Theorem

If $T \in L(X, Y)$ is invertible, then $T^{-1} \in L(Y, X)$.

Matrix Representation of a T

Let X and Y be finite dimensional vector spaces with dimensions n and m, and basis $\{v_1, \cdots, v_n\}$ for X and $\{w_1, \cdots, w_m\}$ for Y. A linear mapping $T : X \to Y$ can be represented by a $m \times n$ matrix M_T, where the i-th column of M_T is given by the vector of coordinates for $T(v_i)$ with respect to basis $\{w_1, \cdots, w_m\}$, i.e.

$$\mathrm{col}_i(M_T) = \mathrm{crd}_{\mathbb{W}}(T(v_i)).$$

The matrix representation can be seen as a function from $L(X, Y)$ to $F_{m\times n}$. It can be shown that the matrix representation is

- linear
- injective and surjective.

Therefore the matrix representation captures all the essential properties of a linear mapping (isomorphic) in finite dimensional vector spaces.

Norm of a Linear Mapping

For a linear mapping $T \in L(X, y = Y)$, define its norm by

$$||T|| = \sup\{\frac{||T(x)||}{||x||}, x \in X, x \neq 0\}.$$

Theorem

Let $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ with matrix representation A. Let $\mu = \max_{i,j}\{A_{ij}\}$, the norm of T can be bounded by

$$\mu \leq ||T|| \leq \mu\sqrt{mn}.$$

Change of Basis

- Matrices A and B are *similar* if for some invertible matrix P, $P^{-1}AP = B$.
- For a linear mapping $T : V \to V$, and let $\mathbb{A} = \{a_1, \cdots, a_n\}$ and $\mathbb{B} = \{b_1, \cdots, b_n\}$ be two basis of V. The matrix representation of T with respect to $\mathbb{A}$ is in general different from the representation with respect to $\mathbb{B}$. However, these two matrices are similar.

Eigenvalues and Eigenvectors

- A is an $n \times n$ matrix. If $Av = \lambda v$ for some nonzero vector v and scalar λ, then λ is an eigenvalue of A and v the corresponding eigenvector.
- $\prod_{i=1}^{n} \lambda_i = |A|$.
- $\sum_{i=1}^{n} \lambda_i = \text{tr}(A)$.
- If A is triangular, then $\lambda_i = A_{ii}$.
- An $n \times n$ matrix with n linearly independent eigenvectors is diagonalizable, i.e. $E^{-1}AE = \Lambda$ where E is the matrix of eigenvectors and Λ is a diagonal matrix with corresponding eigenvalues in the diagonal.