CHAPTER 7, 8. LINEAR DEPENDENCE AND INDEPENDENCE, SPANNING SETS OF A SUBSPACE

1. INTRODUCTION

I. Motivation

We see that a subspace W of a vector space V may have different spanning sets, and these spanning sets may consist of different number of vectors. Now we want to see if all vectors in a spanning set of a subspace are necessary. In other words, can we remove some of the vectors in this set without changing the span?

Consider a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ in a vector space V. If one of the vectors in this set is a linear combination of the others, then this vector can be removed from this set, and the span of this set remains unchanged. In particular, the zero vector can always be removed from a set without changing the span.

For instance, suppose that $\mathbf{v}_1$ is a linear combination of the other vectors $\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_m$. Then every vector in the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ is in span $\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_m\}$, and, trivially, every vector in the set $\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_m\}$ is in span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$. This means

$$\text{span}\,\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_m\} = \text{span}\,\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}.$$

Therefore, removing $\mathbf{v}_1$ from the spanning set does not change the subspace.

Hence, if we want to "generate" a subspace using a set of vectors, it would better to check if any vector in this set is a linear combination of the others. If there is one, we can remove it from the set without changing the result.

The next problem is how to check, in a set of vectors, whether there is one vector that is a linear combination of the others. Suppose, vector $\mathbf{v}_1$ is a linear combination of the vectors $\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_m$, then

$$\mathbf{v}_1 = a_2\mathbf{v}_2 + a_3\mathbf{v}_3 + \ldots + a_m\mathbf{v}_m, \text{ or}$$

$\mathbf{v}_1 - a_2\mathbf{v}_2 - \underline{a}_3\mathbf{v}_3 - \ldots - a_m\mathbf{v}_m = \mathbf{0}$. (Note that the right-hand side is the zero vector).

The coefficients $a_2, a_3, \ldots, a_m$ are real numbers, which may or may not be zero. The coefficient of $\mathbf{v}_1$ is 1, which is not zero. We may multiply this equation by any non-zero number c. The equation becomes

$c\mathbf{v}_1 - ca_2\mathbf{v}_2 - c\underline{a}_3\mathbf{v}_3 - \ldots - ca_m\mathbf{v}_m = \mathbf{0}$¸ $c \neq 0$.

This means that $\mathbf{v}_1$ is a linear combination of $\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_m$ if and only if there is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ that equals the zero vector and the coefficient of $\mathbf{v}_1$ is not zero.

Because we don't actually know which vector is a linear combination of the others, we know that, one of the vectors in the set $\{v_1, v_2, \ldots, v_m\}$ is a linear combination of the others if and only if we have the identity

$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m = \mathbf{0}$,

where $c_1, c_2, \ldots, c_m$ are not all zero.

II. Definitions

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ is *linearly dependent* if there exists coefficients $c_1, c_2, \ldots, c_m$, which are not all zero, and

$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m = \mathbf{0}$.

This identity is called the *dependence relation* of the vectors.

A set is linear dependent means that one of the vectors can be expressed as a linear combination of the others, and, removing this vector from this set does not change the span of this set.

Note that, in a linearly dependent set of vectors, ONE of the vectors can be expressed as a linear combination of the others. This does not mean EVERY vector can be expressed as a linear combination of the others!

Vector $\mathbf{v}_i$ IS a linear combination of the others if the coefficient c_i in a dependence relation is not zero.

Note that the converse of this statement is not true: If we a dependence relation in which the coefficient of a vector **v** is zero, this does not mean that this vector cannot be expressed as a linear combination of the others. For example, in the set {(1, 1), (2, 2), (3,3)}, we have a dependence relation 2(1, 1) + (−1)(2, 2) + 0(3, 3)= (0, 0). But (3, 3) = 3(1, 1) + 0(2, 2) is a linear combination of (1, 1) and (2, 2).

If a set of vectors is not linearly dependent, then it is linearly independent.

By definition, a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ is *linearly independent* if

$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m = \mathbf{0}$

implies $c_1 = c_2 = \ldots = c_m = 0$.

In other words, if a set of vectors $\{\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_m\}$ is linearly independent, then none of the vectors in this set can be expressed as a linear combination of the other.

Obviously, two vectors **u** and **v** in a vector space V are linearly dependent if and only if one is a multiple of the other.

2. EXAMPLES

1. Let S be a set of vectors that contains the zero vector. Since we may choose a nonzero coefficient for the zero vector and let the other coefficients be the zero, then the result is the zero vector. Hence, a set containing the zero vector is linearly dependent.

2. Let $\mathbf{v} = (x_1, x_2, \ldots, x_n)$ be a nonzero vector in $\mathbb{R}^n$. If $c\mathbf{v} = (cx_1, cx_2, \ldots, cx_n) = \mathbf{0}$, we must have $cx_1 = 0, cx_2 = 0, \ldots, cx_n = 0$. Because **v** is not the zero vector, we have some i, $1 \le i \le n$, such that $x_i \neq 0$. Since $cv_i = 0$, we must have $c = 0$. The set $B = \{\mathbf{v}\}$ with a single nonzero vector **v** is linearly independent.

3. Let $S = \{(1, -2, 3), (5, 6, -1), (3, 2, 1)\}$. Determine whether this set of vectors is linearly dependent or linearly independent.

Let $x(1, -2, 3) + y(5, 6, -1) + z(3, 2, 1) = (0, 0, 0)$, or

$$\begin{cases} x+5y+3z=0, \\ -2x+6y+2z=0, \\ 3x-y+z=0. \end{cases}$$

If $(x, y, z) = (0, 0, 0)$ is the only solution to this system, then this set is linearly independent; otherwise, this set is linearly dependent.

Solving this system, we find that this system has infinitely many solutions. One of the solutions is $x = 1$, $y = 1$, and z −2. This set is linearly dependent and we have a dependence relation

$(1, -2, 3) + (5, 6, 1) + (-2)(3, 2, 1) = (0, 0, 0)$.

Since $z \neq 0$, vector $(3, 2, 1)$ is a linear combination of $(1, -2, 3)$ and $(5, 6, -1)$. Vector $(3, 2, 1)$ can be removed from this set without changing the spanning subspace.

span $\{(1, -2, 3), (5, 6, -1), (3, 2, 1)\}$ = span $\{(1, -2, 3), (5, 6, -1)\}$.

Of course, since x and y are not zero, we may also have removed vector $(1, -2, 3)$ or $(5, 6, -1)$ to have a different spanning set to span the same subspace.

Since neither of $(1, -2, 3)$ and $(5, 6, -1)$ is a multiple of the other, this set is linearly independent, and any one of these two vectors cannot be removed further without changing the spanned subspace.

4. Let $S = \left\{ \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix} \right\}$ be a set of 2 by 2 matrices in $\mathbf{M}_{2\times 2}$. Determine whether this set is linearly dependent of linearly independent.

Let $a\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} + b\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} + c\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + d\begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Then

$$\begin{cases} a+c+2d=0, \\ 2a-b+c+2d=0, \\ b+c+2d=0, \\ -a+b+c+d=0. \end{cases}$$

This system has infinitely many solutions. One of the solutions is $a = b = c = 1$, $d = -1$. Hence, this system is linearly dependent and we have a dependence relation

$$\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + (-1)\begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since the coefficient $d \neq 0$, matrix $\begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}$ is a linear combination of the other three matrices.

$$\text{span}\left\{\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}\right\} = \text{span}\left\{\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}.$$

Now let $a\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} + b\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} + c\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

We have

$$\begin{cases} a \quad\quad + c = 0, \\ 2a - b + c = 0, \\ \quad\quad b + c = 0, \\ -a + b + c = 0. \end{cases}$$

This system has only the zero solution. These three vectors are linearly independent, and no more vectors can be removed from this set without changing the spanned subset.

5. Let $S = \{1, 2x - 1, x^2 + x + 1, x^3\}$ be a set of four polynomials in $\mathbf{P}_3$, the vector space of all polynomials with degree at most 3. Determine whether this set is linearly dependent of linearly independent.

Let $a + b(2x - 1) + c(x^2 + x + 1) + dx^3 = 0$. Then $(a - b + c) + (2b + c)x + cx^2 + dx^3 = 0$, or

$$\begin{cases} a - b + c = 0, \\ \quad 2b + c = 0, \\ \quad\quad c = 0, \\ \quad\quad d = 0. \end{cases}$$

Obviously, the only solution to this system is the zero solution. Hence, set S is linearly independent.

6. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ be a set of vectors in $\mathbb{R}^n$, where $\mathbf{u}_i = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \ldots \\ a_{in} \end{bmatrix}$, $i = 1, 2, \ldots, k$. If $a_{ii} = 1$, $i = 1, 2, \ldots, k$, and $a_{ij} = 0$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, k$, $i \neq j$, then set S is linearly independent.

Construct a linear combination

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_k\mathbf{u}_k = c_1\begin{bmatrix} 1 \\ 0 \\ 0 \\ \ldots \\ a_{1n} \end{bmatrix} + c_2\begin{bmatrix} 0 \\ 1 \\ 0 \\ \ldots \\ a_{2n} \end{bmatrix} + c_3\begin{bmatrix} 0 \\ 0 \\ 1 \\ \ldots \\ a_{1n} \end{bmatrix} + \ldots + c_k\begin{bmatrix} 0 \\ 0 \\ 0 \\ \ldots \\ a_{1n} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \ldots \\ c_1a_{1n} + c_2a_{2n} + \ldots + +c_ka_{kn} \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \\ \ldots \\ 0 \end{bmatrix}.$$

Hence, $c_i = 0$, $i = 1, 2, \ldots, k$. This set S is linearly dependent.

In general, if every vector has a non-zero component such that, in any other vectors in this set, this component is zero, then this set of vectors is linearly independent.

For example, since a leading one is the only non-zero entry in it column, the set of nonzero rows in a matrix of RREF is linearly independent.

7. If $\mathbf{a}$, $\mathbf{b}$ are two linearly independent vectors in a vector space V, show that $\mathbf{a} + 2\mathbf{b}$, and $2\mathbf{a} - \mathbf{b}$ are linearly independent.

Let $x(\mathbf{a} + 2\mathbf{b}) + y(2\mathbf{a} - \mathbf{b}) = \mathbf{0}$. Then $(x + 2y)\mathbf{a} + (2x - y)\mathbf{b} = \mathbf{0}$. Since $\mathbf{a}$ and $\mathbf{b}$ are linearly independent, we must have

$$\begin{cases} x + 2y = 0, \\ 2x - \quad y = 0. \end{cases}$$

The only solution to this system is $x = 0$, $y = 0$, and $\mathbf{a} + 2\mathbf{b}$ and $2\mathbf{a} - \mathbf{b}$ are linearly independent.

3. Properties of Linearly Dependent Sets and Linearly Independent Sets

Let V be a vector space. $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ is a set of vectors in V.

1. S is linearly dependent if and only if there exists $\mathbf{v}_i$, $1 \leq i \leq m$, such that

span S = span $(S - \{\mathbf{v}_i\})$.

2. If S has only one vector v, then S is linearly dependent if and only if $\mathbf{v} = \mathbf{0}$.

3. If S contains the zero vector, then S is linearly dependent.

4. If S is linearly dependent, and a subset T of V contains all vectors in S, then T is also linearly dependent.

Indeed, we already have a linear combination of vectors in S that equals the zero vector with not-all-zero coefficients. By adding vectors in $T - S$ to this linear combination with zero coefficients, we have a linear combination of vectors in W that equals the zero vector with not-all-zero coefficients.

As an immediate consequence, we have the following:

5. If S in V is linearly independent, then any subset of S is linearly independent.

6. If span $S \neq V$, and $\mathbf{v} \in V -$ span S, then $S \cup \{\mathbf{v}\}$ is linearly independent.

Indeed, if we have a linear combination

$$a\mathbf{v} + a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_m\mathbf{v}_m = \mathbf{0},$$

since $\mathbf{v}$ is not in span S, $a = 0$. Then we have

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_m\mathbf{v}_m = \mathbf{0}.$$

Since S is linearly independent, we must have $a_1 = a_2 = \ldots = a_m = 0$.

CHAPTER 9, 10. BASES AND THE DIMENSION OF A VECTOR SPACE

1. Bases of a Vector Space

I. Size of a Spanning Set

We have seen that every set of two non-collinear (i.e., linearly independent) vectors spans $\mathbb{R}^2$, and every set of three non-coplanar (i.e., linearly independent) vectors spans $\mathbb{R}^3$. Now we want to look at the general case.

Let V be a vector space. Suppose $U = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ is a spanning set of V. We want to show that any set with more than m vectors is linearly dependent.

Let $W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ be a set of vectors in V, $k > m$. Since U is a spanning set of V, every vector in W is a linear combination of vectors in U:

$\mathbf{w}_1 = a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \ldots + a_{1m}\mathbf{v}_m,$
$\mathbf{w}_2 = a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \ldots + a_{2m}\mathbf{v}_m,$
$\ldots$
$\mathbf{w}_k = a_{k1}\mathbf{v}_1 + a_{k2}\mathbf{v}_2 + \ldots + a_{km}\mathbf{v}_m.$

Suppose we have a linear combination of vectors in W that equals the zero vector. Then

$c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \ldots + c_k\mathbf{w}_k = \mathbf{0}.$

Express $\mathbf{w}_i$, $i = 1, 2, \ldots, k$, as linear combinations of vectors in U. We have

$c_1(a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \ldots + a_{1m}\mathbf{v}_m) + c_2(a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \ldots + a_{2m}\mathbf{v}_m) + \ldots$
$+ c_k(a_{k1}\mathbf{v}_1 + a_{k2}\mathbf{v}_2 + \ldots + a_{km}\mathbf{v}_m) = (c_1a_{11} + c_2a_{21} + \ldots + c_ka_{k1})\mathbf{v}_1 + (c_1a_{12} + c_2a_{22} + \ldots + c_ka_{k2})\mathbf{v}_2$
$+ \ldots + (c_1a_{1m} + c_2a_{2m} + \ldots + c_ka_{km})\mathbf{v}_m = (0, 0, \ldots, 0).$

Therefore,

$c_1a_{11} + c_2a_{21} + \ldots + c_ka_{k1} = 0,$

$c_1a_{12} + c_2a_{22} + \ldots + c_ka_{k2} = 0,$
$\ldots$
$c_1a_{1m} + c_2a_{2m} + \ldots + c_ka_{km} = 0.$

This is a homogeneous system with m equations and k variables. Since $k > m$, this system has infinitely many solutions. In other words, it has nonzero solutions. Then W is linearly dependent.

***Conclusion*: Suppose that a vector space V has a spanning set with size k and $S \subseteq V$ is a set of vectors with size m. Then**

(i) if $m > k$, then S is linearly dependent; or equivalently,

(ii) if S is linearly independent, then $k \leq m$.

II. Basis

Let B be a finite spanning set of V. By removing "redundant" vectors from this set, we may assume that this set is linearly independent. A linearly independent spanning set of V is called a *basis* of V.

Let B_1 and B_2 be two bases of V. Because B_1 is linear independent and B_2 spans V, B_1 cannot have more vectors than B_2 does. Similarly, B_2 cannot have more vectors than B_1 does.

Conclusion: A**ll bases have the same number of vectors**.

2. The Dimension of a Vector Space

I. Definition

The number of vectors in any basis of V is called the *dimension* of V, denoted by dim (V).

The dimension of a vector space is not always finite. Some vector spaces have the infinite dimension. The vectors space of all polynomials and the vector space of all one-variable functions are examples of vector spaces with infinite dimension. However, in this course, we will study only vector spaces with finite dimensions.

Examples

1. Since the set $\{\mathbf{0}\}$ with only the zero vector is linearly dependent, the vector space $\{\mathbf{0}\}$ with only the zero vector has dimension 0. On the other hand, if a vector space V has at least one nonzero vector, since the set $\{\mathbf{v}\}$ with only the nonzero element is linearly independent, and the dimension of V is greater than 0.

2. In $\mathbb{R}^n$, the set of vectors

$B = \{(1, 0, 0, \ldots, 0), (0, 1, 0, \ldots, 0), (0, 0, 1, \ldots, 0), \ldots, (0, 0, 0, \ldots, 1)\}$

is obviously a linearly independent set, and it spans $\mathbb{R}^n$. B is a basis of $\mathbb{R}^n$, and dim ($\mathbb{R}^n$) = n.

This basis is called the *standard basis* of $\mathbb{R}^n$.

3. Let **M** be the vector space of symmetric 2 by 2 matrices. A 2 by 2 symmetric matrix has the form $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Since $\begin{bmatrix} a & b \\ b & c \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, the set
$B = \left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right\}$ is a spanning set of **M**.

If $a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, we must have $a = b = c = 0$. Then set B is linearly independent.

Hence, B is a basis of **M**, and dim **M** = 3.

II. Construct a Basis Starting from a Spanning Set or a Linearly Independent Set

Now assume V is a vector space with finite dimension d.

Let S be a subset of V. If we want S to be a basis, it must satisfy two conditions: S is linearly independent and S spans V. Now we look at those sets that satisfy only one of the conditions:

Suppose S is a (finite) spanning set of V. If it is linearly independent, then it IS a basis, and its size is dim(V); otherwise, we can construct a basis from S by removing "redundant" vectors (i.e., those that can be expressed as a linear combination of the remaining ones) one by one until we have a linearly independent spanning set. Then we have a basis. This proves the following result:

Every spanning set of *V* contains a basis and the size of every spanning set of *V* is greater than or equal to dim (*V*). If the size of a spanning set *S* equals the dim (*V*), then *S* is linearly independent, and it is a basis.

On the other hand, suppose *W* is a linearly independent subset of *V*. If it spans *V*, it IS a basis, and its size is dim (*V*); otherwise, there exists a vector **v** in *V* that is not in span *W*. By adding **v** to *W*, we have a linearly independent set with one more vector. Keep doing this. We will eventually have dim (*V*) vectors in an independent set, which is a basis. This proves the following result:

Every linearly independent set is contained in a basis and the size of every linearly independent set is less than or equal to dim (*V*). If the size of a linearly independent set *S* equals dim (*V*), then *S* is a spanning set of *V*, and it is a basis.

In other words, for a set *S* in a vector space *V*, if *S* satisfies any two of the following conditions, then is must satisfy the third condition, and it is a basis:

(i) *S* is linearly independent;

(ii) span (*S*) = *V*;

(iii) The size of *S* is dim (*V*).

Examples

1. Consider the set of vectors $S = \{(1, 0, 1), (0, 1, 1), (1, 1, 1), (1, 1, 0)\}$ in $\mathbb{R}^3$. We show that span $S = \mathbb{R}^3$.

A general vector in $\mathbb{R}^3$ has the form (a, b, c). To show that spam $S = \mathbb{R}^3$, we want to find x, y, z, and w, such that

$$x(1, 0, 1) + y(0, 1, 1) + z(1, 1, 1) + w(1, 1, 0) = (a, b, c).$$

Then

$$\begin{aligned} x \quad\;\; + z + w &= a, \\ y + z + w &= b, \\ x + y + z \quad\;\; &= c. \end{aligned}$$

This system has infinitely many solutions. One of them is $x = c - b$, $y = c - a$, $z = a + b - c$, and $w = 0$. Hence, span $S = \mathbb{R}^3$.

We know that dim ($\mathbb{R}^3$) = 3. S has four vectors, so it is linearly dependent.

We find a dependence relation: $(1, 0, 1) + (0, 1, 1) - 2(1, 1, 1) + (1, 1, 0) = (0, 0, 0)$.

We can remove any of them to have a smaller spanning set of $\mathbb{R}^3$. For instance, $T = \{(1, 0, 1), (0, 1, 1), (1, 1, 1)\}$ spans $\mathbb{R}^3$. Since the dimension of $\mathbb{R}^3$ is 3, T is a basis of $\mathbb{R}^3$, and it is linearly independent.

2. Consider the set of vectors $S = \{(1, 2, 3), (4, 5, 6)\}$ in $\mathbb{R}^3$. Since these two vectors are not collinear, S is linearly independent. Since the size of S is 2 < dim ($\mathbb{R}^3$) = 3, there exists a basis of $\mathbb{R}^3$ that contains S. We want to find a vector that is not in span S. We know that the subspace spanned by S is a plane in $\mathbb{R}^3$ going through the origin. We want a vector that not in this plane. An easy way is to find a vector that is perpendicular to this plane. This vector can be obtained by taking the cross product of these two vectors.

$(1, 2, 3) \times (4, 5, 6) = (-3, 6, -3) = -3(1, -2, 1)$.

Vector $(1, -2, 1)$ is not in span S. Then the set $T = \{(1, 2, 3), (4, 5, 6), (1, -2, 1)\}$ is linearly independent. Since the size of this set equals the dimension of $\mathbb{R}^3$, T is a basis of $\mathbb{R}^3$, and it spans $\mathbb{R}^3$.

III. Dimension of a Subspace

Let U be a subspace of a vector space V. A basis B of U is a linearly independent set of V. If B is also a basis of V, then dim (V) = dim (U), and span $B = U = V$. If B is not a basis of V, then we have a basis A of V that contains B. Hence, the size of A is greater than the size of B. Since the size of A is the dimension of V, and the size of B is the dimension of U, dim (V) > dim (U). We have the following result:

If U is a subspace of a vector space V, then dim (U) ≥ dim (V). The equality holds if and only if $U = V$.

3. ORDERED BASIS AND COORDINATES

I. Definition

If we give an order of the vectors in a basis of a vector space, the basis is called an *ordered basis*.

For example, $B_1 = \{(1, 2), (2, 1)\}$ and $B_2 = \{(2, 1), (1, 2)\}$ are two basis of $\mathbb{R}^2$, and there are actually the same basis because they contain the same set of vectors. However, as ordered basis, they are different: The first vector of B_1 is (1, 2), and the first vector of B_2 is (2, 1).

Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an ordered basis of a vector space V. Note that V does not have to be a subspace of $\mathbb{R}^n$! Since B has n vectors, dim $(V) = n$. As a basis, B is linearly independent, and every vector in V can be expressed as a linear combination of vectors in B.

Let $\mathbf{u}$ be a vector in V. Then

$\mathbf{u} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n$.

The ordered n-tuple $(x_1, x_2, \ldots, x_n)$ of scalars is called the *coordinates* of $\mathbf{u}$ relative to the ordered basis B.

II. Uniqueness of Coordinates

An important result of coordinates is the following:

The coordinates of any vector u relative to an ordered basis in a vector space is unique.

Assume we have

$\mathbf{u} = x_1 \mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n$, and
$\mathbf{u} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \ldots + y_n\mathbf{v}_n$.

Taking the difference,

$\mathbf{0} = (x_1 - y_1)\mathbf{v}_1 + (x_2 - y_2)\mathbf{v}_2 + \ldots + (x_n - y_n)\mathbf{v}_n$.

Since B is linearly independent, we must have $x_1 - y_1 = x_2 - y_2 = \ldots = x_n - y_n = 0$, or $x_1 = y_1$, $x_2 = y_2$, $\ldots$, $x_n = y_n$.

By this result, we can establish a one-to-one correspondence between the vectors in V and their coordinates, which are vectors in $\mathbb{R}^n$.

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an ordered basis of a vector space V with dimension n, then every vector $\mathbf{u}$ in V can be expressed as $\mathbf{u} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_m\mathbf{v}_m$, where $(x_1, x_2, \ldots, x_m)$ are coordinates of $\mathbf{u}$.

Obviously, if $\mathbf{u}_1$ has coordinates $(x_1, x_2, \ldots, x_n)$ and $\mathbf{u}_2$ has coordinates $(y_1, y_2, \ldots, y_n)$, the $c\mathbf{u}_1$ has coordinates $(cx_1, cx_2, \ldots, cx_n)$, and $\mathbf{u}_1 + \mathbf{u}_2$ has coordinates $(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$. We say that the relation between the vectors in V and their coordinates *preserves* the addition and scalar multiplication operations.

From this point of view, we see that a vector space of dimension n is "essentially the same" as $\mathbb{R}^n$, i.e., they have the same properties and behaviour as vector spaces.

For example, consider "vectors" $p_1 = 1 + x$, $p_2 = x - x^3$, $p_3 = x^2 \in \mathbf{P}_3$, the vector space of polynomials of degree at most 3. The set $S = \{p_1, p_2\}$ is linearly independent. They form a basis of a subspace U = span S of $\mathbf{P}_3$.

Polynomial $q_1 = 2p_1 + p_2 - p_3 = 2 + 3x - x^2 - x^3$ has coordinates $\mathbf{v}_1 = (1, 2, -1)$, polynomial $q_2 = -p_1 + 2p_2 + p_3 = -1 + x + x^2 - 2x^3$ has coordinates $\mathbf{v}_2 = (-1, 2, 1)$, and polynomial $q_3 = 2p_1 - 2p_3 = 2 + 2x - 2x^2$ has coordinates $\mathbf{v}_3 = (2, 0, -2)$.

Since $\mathbf{v}_3 = \mathbf{v}_1 - \mathbf{v}_2$, the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent. Then we know that the set $Q = \{q_1, q_2, q_3\}$ is linearly dependent with dependence relation $q_3 = q_1 - q_2$.

CHAPTER 14. MATRIX MULTIPLICATION

1. DEFINITION

Let $M = [a_{ij}]$ be an $m \times n$ matrix and let $N = [b_{jk}]$ be an $n \times p$ matrix. Then the product MN is defined to be $MN = [c_{ik}]$, where c_{ik} is the dot product of the i-th row of M and the k-th column of N:

$$c_{ik} = [a_{i1}, a_{i2}, \ldots, a_{in}] \cdot \begin{bmatrix} b_{1k} \\ b_{2k} \\ \cdots \\ b_{nk} \end{bmatrix} = a_{i1}b_{1k} + a_{i2}k_{2k} + \ldots + a_{in}b_{nk},\ 1 \le i \le m,\ 1 \le k \le p.$$

Example

$$\begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 & 1 \\ -1 & 1 & 0 & 2 \\ 0 & 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 & 5 \\ -2 & 1 & -2 & -1 \end{bmatrix}.$$

Remarks

1. The product MN is defined only if the number of columns of M equals the number of rows of N. Therefore, the dimension of row vectors in M equals the dimension of column vectors in N.

2. The product of an $m \times n$ matrix and an $n \times p$ matrix is an $m \times p$ matrix.

3. An n-dimensional row vector is regarded as a 1 by n matrix, and an n-dimensional column vector is regarded as a n by 1 matrix. Therefore, the product of an n-dimensional row vector multiplied by an n-dimensional column vector is a 1 by 1 matrix. An m-dimensional column multiplied by an n-dimensional row vector is an m by n matrix.

In particular, let $A = [a_{ij}]$ be an $m \times n$ matrix, and let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}$ be column vectors in $\mathbb{R}^n$. Then the linear system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m, \end{cases}$$

can be denoted by the *vector form* $x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \dots + x_n\mathbf{c}_n = \mathbf{b}$, where $\mathbf{c}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \dots \\ a_{ni} \end{bmatrix}$ is the i-th column of A, or the *matrix form*

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}, \text{ or } A\mathbf{x} = \mathbf{b}.$$

4. If A is a square matrix, denote $AA = A^2$, $A^2A = A^3$. In general, the n-th *power* of A is $A^n = A^{n-1}A$. We also define $A^0 = I$.

Examples

(1) $A = \begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix}$, $A^2 = \begin{bmatrix} 3 & -2 \\ -1 & 2 \end{bmatrix}$, $A^3 = \begin{bmatrix} -5 & 6 \\ 3 & -2 \end{bmatrix}$.

(2) A square matrix $A = [a_{ij}]$ is a *diagonal matrix* if $a_{ij} = 0$, for all $i \neq j$. If $A = [a_{ij}]$ is a diagonal matrix, then $A^k = [a_{ij}{}^k]$.

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}^k = \begin{bmatrix} a_{11}{}^k & 0 & \dots & 0 \\ 0 & a_{22}{}^k & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn}{}^k \end{bmatrix}$$

5. Matrix multiplication does not satisfy the commutative law even if both MN and NM are defined.

Example

$$\begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}.$$

6. Although we know that, if a product of two number $xy = 0$, then either $x = 0$ or $y = 0$. However, if the product of two matrices MN is the zero matrix, it is possible that neither M nor N is a zero matrix.

Example

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

7. Although we know that, if there are three numbers a, b and c, such that $ac = bc$, then either $c = 0$, or we can "cancel" c on both sides of the equality to have $a = b$. However, if there are three matrices A, B, and C such that $AC = BC$, we cannot conclude that either C is the zero matrix or $A = B$. (The cancellation law does not hold for matrix multiplication)!

Example

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

2. Properties of Matrix Multiplication

Let A, B, and C be matrices, and let k be a scalar. Assume all the products are defined. Then

(1) $(AB)C = A(BC)$. (associativity)

(2) $A(B + C) = AC + AC$. (distributivityon the right)

(3) $(A + B)C = AC + BC$. (ditributivity on the left)

(4) $k(AB) = (kA)B$.

(5) $(AB)^T = B^T A^T$.

(6) If M is an $m \times n$ matrix, then $I_m M = M\, I_n = M$, where I_m is the identity matrix of dimension m, and I_n is the identity matrix of dimension n.

(7) Denote an $s \times t$ zero matrix by $0_{s,\,t}$. If M is an $m \times n$ matrix, then

$0_{s,\,m}M = 0_{s,\,n}$, $M0_{n,\,t} = 0_{m,\,t}$.

These properties can be verified directly by calculation using the definition.

Proof of the associativity. Let $A = [a_{ij}]$ be an m by n matrix, let $B = [b_{js}]$ be an n by p matrix, and let $C = [c_{st}]$ be a p by q matrix.

The (i, s)-entry of AB is $d_{is} = a_{i1}b_{1s} + a_{i2}b_{2s} + \ldots + a_{in}b_{ns}$, $1 \le i \le m$, $1 \le s \le p$.

The (i, t)-entry of $(AB)C$ is
$e_{it} = d_{i1}c_{1t} + d_{i2}c_{2t} + \ldots + d_{ip}c_{pt}$
$= (a_{i1}b_{11} + a_{i2}b_{21} + \ldots + a_{in}b_{n1})c_{1t}$
$+ (a_{i1}b_{12} + a_{i2}b_{22} + \ldots + a_{in}b_{n2})\, c_{2t}$
$+ \ldots$
$+ (a_{i1}b_{1p} + a_{i2}b_{2p} + \ldots + a_{in}b_{np})\, c_{pt}$
$= a_{i1}(b_{11}c_{1t} + b_{12}c_{2t} + \ldots + b_{1p}c_{pt})$
$+ a_{i2}(b_{21}c_{1t} + b_{22}c_{2t} + \ldots + b_{2p}c_{pt})$
$+ \ldots$
$+ a_{in}(b_{n1}c_{1t} + b_{n2}c_{2t} + \ldots + b_{np}c_{pt})$.

On the other hand,

A (j, t)-entry in BC is $f_{jt} = b_{j1}c_{1t} + b_{j2}c_{2t} + \ldots + b_{jp}c_{pt}$, $1 \le j \le n$, $1 \le t \le q$.

An (i, t) entry in $A(BC)$ is

$g_{it} = a_{i1}f_{1t} + a_{i2}f_{2t} + \ldots + a_{in}f_{nt}$
$= a_{i1}(b_{11}c_{1t} + b_{12}c_{2t} + \ldots + b_{1p}c_{pt})$
$+ a_{i2}(b_{21}c_{1t} + b_{22}c_{2t} + \ldots + b_{2p}c_{pt})$
$+ \ldots$
$+ a_{in}(b_{n1}c_{1t} + b_{n2}c_{2t} + \ldots + b_{np}c_{pt}) = f_{it}$.

Hence, $(AB)C = A(BC)$.

Examples

1. $(A + B)(C + D) = AC + BC + AD + BD$.

2. $(A + B)(A – B) = A^2 + BA – AB – B^2$.

3. Block Multiplication

A matrix may be separated into *blocks*. The following is a 5×6 matrix separated in to four blocks of dimensions 3×4, 3×2, 2×4, and 2×2, respectively:

$$M = \left[\begin{array}{cccc:cc} 1 & 0 & 2 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ \hdashline 2 & 0 & 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right].$$

If two matrices have compatible dimensions and separated into compatible blocks, then we can multiply these two matrices block by block as if each block is a number.

Suppose matrix A is separate into pq blocks:

$$A = \begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1q} \\ A_{21} & A_{22} & \ldots & A_{2q} \\ & & \ldots & \\ A_{p1} & A_{p2} & \ldots & A_{pq} \end{bmatrix},$$

and matrix B is separated into qs blocks:

$$A = \begin{bmatrix} B_{11} & B_{12} & \dots & B_{1s} \\ B_{21} & B_{22} & \dots & B_{2s} \\ & \dots & & \\ B_{q1} & B_{q2} & \dots & B_{qs} \end{bmatrix},$$

If the number columns in A_{it} equals the number of rows in B_{tj} for all $t = 1, 2, \dots, q$, then the product consists of ps blocks, $AB = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1s} \\ C_{21} & C_{22} & \dots & C_{2s} \\ & \dots & & \\ C_{p1} & C_{p2} & \dots & C_{ps} \end{bmatrix}$, where

$C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \dots + A_{iq}B_{qj}$, $1 \le i \le p$, $1 \le j \le s$.

Example

$$A = \left[\begin{array}{cccc|cc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ \hline 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right], B = \left[\begin{array}{cc|cc} 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ \hline -1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right].$$

Then $A_{11} = I_4$, $A_{12} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$, $A_{21} = [0, 1, 0, 1]$, $A_{22} = [0, 1]$, $B_{11} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix}$, $B_{12} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 2 & 0 \\ 0 & 1 \end{bmatrix}$, $B_{21} = \begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix}$, $B_{22} = I_2$.

Then $C_{11} = A_{11}B_{11} + A_{12}B_{21} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 1 & 0 \\ 2 & 2 \\ 1 & 1 \end{bmatrix}$,

$$C_{12} = A_{11}B_{12} + A_{12}B_{22} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 3 & 1 \\ 0 & 2 \end{bmatrix},$$

$$C_{21} = A_{21}B_{11} + A_{22}B_{21} = [0, 1, 0, 1] \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} + [0, 1]\begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix} = [1, 1] + [1, 0] = [2, 1],$$

$$C_{22} = A_{21}B_{12} + A_{22}B_{22} = [0, 1, 0, 1] \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} + [0,1] = [1, 2] + [0, 1] = [1, 3].$$

Finally, the product

$$AB = \left[\begin{array}{cc:cc} -1 & 3 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 2 & 3 & 1 \\ 1 & 1 & 0 & 2 \\ \hdashline 2 & 1 & 1 & 3 \end{array}\right]$$

If a square matrix M can be expressed in the block form as $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, where A and B are both square matrices, then by the formula of block multiplication, we have $M^2 = \begin{bmatrix} A^2 & 0 \\ 0 & B^2 \end{bmatrix}$. In general, for any positive integer r, $M^r = \begin{bmatrix} A^r & 0 \\ 0 & B^r \end{bmatrix}$.