MAT224 Linear Algebra II
Lecture Notes

Yuchen Wang

February 6, 2019

Contents

1 Vector Spaces

1.1 Vector Spaces

Definition 1.1.1 A (real) vector space is a set V (whose elements are called vectors) together with

1. an operation called vector addition, which for each pair of vectors $\mathbf{x}, \mathbf{y} \in V$ produced another vector in V denoted $\mathbf{x} + \mathbf{y}$, and
2. an operation called multiplication by a scalar (a real number), which for each vector $\mathbf{x} \in V$, and each scalar $c \in \mathbb{R}$ produced another vector in V denoted $c\mathbf{x}$

Furthermore, the two operations must satisfy the following axioms: $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, \forall c, d \in \mathbb{R}$,

1. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
2. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
3. $\exists \mathbf{0} \in V$ s.t. $\mathbf{x} + \mathbf{0} = \mathbf{x}$ (additive identity)
4. $\exists - \mathbf{x} \in V$ s.t. $\mathbf{x} + -\mathbf{x} = \mathbf{0}$ (additive inverse)
5. $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$
6. $(c + d)\mathbf{x} = c\mathbf{x} + d\mathbf{x}$
7. $(cd)\mathbf{x} = c(d\mathbf{x})$
8. $1\mathbf{x} = \mathbf{x}$

Smooth functions C^∞
Most functions are not smooth.

1.2 Subspaces

Example $C^\infty(\mathbb{R}) < C^k(\mathbb{R}) <$ Differentiable functions $< C(\mathbb{R}) < F(\mathbb{R})$

Definition Let V be a vector space and Let $W \subseteq V$ be a subset. Then W is a (vector) subspace of V if W is a vector space itself under the operations of vector sum and scalar multiplication from V.

Theorem 1.2.8 Let V be a vector space and Let $W \subseteq V$ be a nonempty subset of V. Then W is a subspace of V iff $\forall \mathbf{x}, \mathbf{y} \in W$, and all $c \in \mathbb{R}$, we have $c\mathbf{x} + \mathbf{y} \in W$.
<u>*proof:*</u> $\rightarrow$: If W is a subspace of V, then $\forall \mathbf{x}, \mathbf{y} \in W$ and $c \in \mathbb{R}, c\mathbf{x} + \mathbf{y} \in W$ holds since W itself is a real vector space.
$\leftarrow$: If $\forall \mathbf{x}, \mathbf{y} \in W$, and all $c \in \mathbb{R}$, we have $c\mathbf{x} + \mathbf{y} \in W$
Can have $c = 1$, so $\mathbf{x} + \mathbf{y} \in W$ (close under addition)
$c = -1$ and $\mathbf{y} = \mathbf{x}$, so $-\mathbf{x} + \mathbf{x} = \mathbf{0} \in W$ (additive identity)
$\mathbf{y} = \mathbf{0}$, so $c\mathbf{x} \in W$ (close under scalar multiplication)
These implies all the axioms. ■

Examples

1. $W = \{f \in C(\mathbb{R}) | f(\pi) = 0\}$. W subspace of $C(\mathbb{R})$? -Yes
2. $W = \{f \in C(\mathbb{R}) | f(e) = e\}$. W subspace of $C(\mathbb{R})$? -No, not close under addition
3. $W = \{(x_1, ..., x_n) | x_i \geq 0 \forall i\}$. W subspace of $C(\mathbb{R})$? -No, there is no additive inverse for each item in W.

Theorem 1.2.13 Let V be a vector space. Then the intersection of any collection of subspaces of V is a subspace of V.
<u>*proof:*</u> Consider any collection of subspace of V. Note that the intersection of the subspaces is not empty since at least the zero vector from V is in it. Now let $\mathbf{x}, \mathbf{y}$ be any two vectors in the intersection, so they are in every single subspace in the collection. Therefore $c\mathbf{x} + \mathbf{y}$ is also in every single subspace in the collection, so that it is in the intersection as well. Hence the intersection is a subspace of V. ■

Application The set of all solutions of any simultaneous system of equations is a subspace of $\mathbb{R}^n$.

Corollary 1.2.14 Let $a_{ij} (1 \leq i \leq m, 1 \leq j \leq n)$ be any real numbers and let $W = \{(x_1, ..., x_n) \in \mathbb{R}^n | a_{i1}x_1 + \ldots + a_{in}x_n = 0 \text{ for all } i, 1 \leq i \leq m\}$. Then W is a subspace of $\mathbb{R}^n$.

1.3 Linear Combinations

Definition 1.3.1 Let S be a subset of a vector space V.

1. A *linear combination* of vectors in S is any sum $a_1\mathbf{x}_1 + \ldots + a_n\mathbf{x}_n$ where the $a_i \in \mathbb{R}$, and the $\mathbf{x}_i \in S$
2. If $S \neq \emptyset$, the set of all linear combinations of vectors in S is called the *span* of S, and denoted Span(S). If $S = \emptyset$, we define Span(S) $= \{\mathbf{0}\}$. (Remark: It is a mathematician convention)
3. If W = Span(S), we say S spans (or generates) W.

Theorem 1.3.4 Let V be a vector space and let S be any subset of V. Then Span(S) is a subspace of V.

proof: Span(S) is non-empty by definition. Let $\mathbf{x}, \mathbf{y} \in Span(S)$, then they are linear combinations of vectors in S. Check that $c\mathbf{x} + \mathbf{y}$ is also a linear combination of vectors in S, so $c\mathbf{x} + \mathbf{y} \in Span(S)$. Hence Span(S) is a subspace of V. ■

Definition Let W_1 and W_2 be subspaces of a vector space V. The *sum* of W_1 and W_2 is the set

$$W_1 + W_2 = \{\mathbf{x} \in V | \mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \text{ for some } \mathbf{x}_1 \in W_1 \text{ and } \mathbf{x}_2 \in W_2\}$$

Proposition 1.3.8 The basis of sum is the union of two bases Let $W_1 = Span(S_1)$ and $W_2 = Span(S_2)$ be subspaces of a vector space V. Then $W_1 + W_2 = Span(S_1 \cup S_2)$

Theorem 1.3.9 Let W_1 and W_2 be subspaces of a vector space V. Then $W_1 + W_2$ is also a subspace of V.

Proposition 1.3.11 $W_1 + W_2$ is the smallest subspace containing $W_1 \cup W_2$: Let W_1 and W_2 be subspaces of a vector space V and let W be a subspace of V such that $W_1 \cup W_2 \subseteq W$. Then $W_1 + W_2 \subseteq W$

Remark $W_1 \cup W_2$ is a subspace of V iff one is contained in another.

1.4 Linear Dependence and Linear Independence

Definitions 1.4.2 Let V be a vector space, and let S be a subset of V.

1. A *linear dependence* among the vectors of S is an equation

$$a_1\mathbf{x}_1 + \ldots + a_n\mathbf{x}_n = \mathbf{0}$$

   where the $x_i \in S$, and the $a_i \in \mathbb{R}$ are not all zero (i.e., at least one of the $a_i \neq \mathbf{0}$

2. the set S is said to be *linearly dependent* if there exists a linear dependence among the vectors in S.

Fact Let S be a set. If $\mathbf{0} \in S$, then S is dependent.

Definition 1.4.4 A subset S of a vector space V is *linearly independent* if whenever we have $a_i \in \mathbb{R}$ and $\mathbf{x}_i \in S$ such that $a_1\mathbf{x}_1 + \ldots + a_n\mathbf{x}_n = \mathbf{0}$, then $a_i = 0$ for all i.

Example In any vector space the empty subset $\emptyset$ is linearly independent.

Proposition 1.4.7

1. Let S be a linearly independent subset of a vector space V, and let S' be another subset of V that contains S. Then S' is also linearly dependent.

2. Let S be linearly independent subset of a vector space V and let S' be another subset of V that is contained in S. Then S' is also linearly independent.

1.5 Interlude on Solving Systems of Linear Equations (MAT223)

1.6 Bases And Dimension (Jan 17)

Definition A subset S of vector space V is called a *basis* of V if V = Span(S) and S is linearly independent.

Remark A basis is the maximal set of linearly independent vectors / minimal set of spanning vectors.

Examples

1. the standard basis S = $\{e_1,...,e_n\}$ in $\mathbb{R}^n$, since every vector $(a_1, ..., a_n) \in \mathbb{R}^n$ may be written as the linear combination $(a_1, ..., a_n) = a_1e_1 + ... + a_ne_n$
2. The vector space $\mathbb{R}^n$ has many other bases as well. e.g., in $\mathbb{R}^2$, consider the set $S = \{(1, 2), (1, -1)\}$, which is l.i.
3. Let $V = P_n(\mathbb{R})$ and consider $S = \{1, x, x^2, ..., x^n\}$, which is a basis of V.
   proof: It is clear that S spans V. For independence, consider
   $$a_0 + a_1x + a_2x^2 + ... + a_{n-1}x^{n-1} + a_nx^n = \mathbf{0}$$
   Take the derivative of both sides,
   $$\frac{d^n}{dx^n}(a_0 + a_1x + a_2x^2 + ... + a_{n-1}x^{n-1} + a_nx^n) = \frac{d^n}{dx^n}(0)$$
   $$n!a_n = 0 \implies a_n = 0$$
   Similarly, we have $a_i = 0$ for all i, as wanted.
4. The empty subset, $\emptyset$, is a basis of the vector space consisting only of a zero vector, $\{\mathbf{0}\}$.

Theorem 1.6.3 Let V be a vector space, and let S be a nonempty subset of V. Then S is a basis of V iff every vector $\mathbf{x} \in V$ may be written uniquely as a linear combination of the vectors in S.
Proof: $\rightarrow$: Assume S is a basis of V, then given $\mathbf{x} \in V$, there are scalars $a_i \in \mathbb{R}$ and vectors $x_i \in S$ s.t. $\mathbf{x} = a_1x_1 + ... + a_nx_n$. To show this linear combination is unique, consider a possible second linear combination of vectors in S which also adds up to $\mathbf{x}$: $x = b_1x_1 + ... + b_nx_n$. Subtracting these two expressions for $\mathbf{x}$, we find that

$$\mathbf{0} = a_1x_1 + ... + a_nx_n - (b_1x_1 + ... + b_nx_n)$$
$$= (a_1 - b_1)x_1 + ... + (a_n - b_n)x_n$$

Since S is linearly independent, the equation implies that $a_i = b_i$ for all i.

$\leftarrow$: Assume every vector $\mathbf{x} \in V$ may be written uniquely as a linear combination of the vectors in S. This implies $Span(S) = V$. We must show that S is l.i. Consider an equation

$$a_1 x_1 + ... + a_n x_n = \mathbf{0}$$

Note that it is also the case that

$$0\mathbf{x} = 0(x_1 + ... + x_n) = \mathbf{0}$$

Since we assumed that every $\mathbf{x}$ has a unique representation in S, then it must be true that $a_i = 0$ for all i. Hence S is l.i.

Theorem 1.6.6 Let V be a vector space that has a finite spanning set, and let S be a linearly independent subset of V. Then there exists a basis S' of V, with $S \subset S'$

Lemma 1.6.8 Let S be a linearly independent subset of V and let $x \in V$, but $x \notin S$. Then $S \cup \{\mathbf{x}\}$ is l.i. iff $\mathbf{x} \notin Span(S)$.

Insight the number of vectors in a basis is, in a rough sense, a measure of "how big" the space is.

Theorem 1.6.10 (Basis Theorem) Let V be a vector space and let S be a spanning set for V, which has m elements. Then no linearly independent set in V can have more than m elements.

proof: It suffices to show that every set in V with more than m elements is linearly dependent. Write $S = y_1, ..., y_m$ and suppose $S' = x_1, ..., x_n$ is a subset of V with $n > m$ vectors. Consider an equation

$$(1) a_1 x_1 + ... + a_n x_n = \mathbf{0}$$

Our goal is to show that a_i not all 0. Since S spans V, there are scalars b_{ij} s.t. for each i,

$$x_i = b_{i1} y_1 + ... + b_{im} y_m$$

Substituting these equations into (1), we get

$$a_1(b_{11} y_1 + ... + b_{1m} y_m) + ... + a_n(b_{n1} y_1 + ... + b_{nm} y_m) = \mathbf{0}$$

Collecting terms and rearranging,

$$(a_1b_{11} + ... + a_nb_{n1})y_1 + ... + (a_1b_{1m} + ... + a_nb_{nm})y_m = \mathbf{0}$$

Since S is l.i., this is equivalent to solving the system

$$b_{11}a_1 + ... + b_{n1}a_n = 0$$

$$\vdots$$

$$b_{1m}a_1 + ... + b_{nm}a_n = 0$$

But this is a system with n unknowns and m equations and $n > m$, so there must exist a non-trivial solution $\{a_1, ..., a_n\}$, which is what we wanted to show. ■

Corollary 1.6.11 Let V be a vector space and let S and S' be two bases of V, with m and m' elements, respectively. Then $m = m'$.

proof:
Since S is a spanning set of V and S' is l.i., we have $m' \leq m$. Since S' is a spanning set of V and S is l.i.m we have $m \leq m'$. Hence $m = m'$. ■

Definitions 1.6.12

1. If V is a vector space with some finite basis(possibly empty), we say V is *finite-dimentional*.
2. Let V be a finite-dimensional vector space. The dimension of V, denoted dim(V), is the number of vectors in a (hence any) basis of V.
3. If $V = \{\mathbf{0}\}$, we define dim(V) = 0.

Examples

1. For each n, dim($\mathbb{R}^n$) $= n$, since the standard basis contains n vectors.
2. dim$(P_n(\mathbb{R})) = n + 1$, since a basis for $P_n(\mathbb{R})$ contains n + 1 functions.
3. The vector spaces $P(\mathbb{R}), C^1(\mathbb{R})$ and $C(\mathbb{R})$ are not finite-dimensional. We say that such spaces are *infinite-dimentional*.
4. dim$(Span\{(1, 2, 3), (4, 5, 6), (7, 8, 9)\}) = 2$

Corollary 1.6.14 Let W be a subspace of a finite-dimensional vector space V. Then $\dim(W) \leq \dim(V)$. Furthermore, $\dim(W) = \dim(V)$ iff $W = V$.

Corollary 1.6.15 Let W be a subspace of $\mathbb{R}^n$ defined by a system of homogeneous linear equations. Then dim(W) is equal to the number of free variables in the corresponding echelon form system.

Theorem 1.6.18 Let W_1 and W_2 be finite-dimensional subspaces of a vector space V. Then

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$$

Remark Analogous to the Principle of Inclusion-Exclusion

proof: Result obvious if either W_1 or W_2 is $\{\mathbf{0}\}$.
Therefore, we assume that neither W_1 nor W_2 is $\{\mathbf{0}\}$. Starting from a basis S of $W_1 \cap W_2$. We can always find sets T_1 and T_2 (disjoint from S) such that $S \cup T_1$ is a basis for W_1 and $S \cup T_2$ is a basis for W_2. We claim that $U = S \cup T_1 \cup T_2$ is a basis for $W_1 + W_2$, since

$$U = S \cup T_1 \cup T_2 = (S \cup T_1) \cup (S \cup T_2)$$

$$Span(U) = Span((S \cup T_1) \cup (S \cup T_2)) = W_1 + W_2$$

Next, prove that U is linearly independent. Any potential linear dependence among the vectors in U must have the form

$$\mathbf{v} + \mathbf{w}_1 + \mathbf{w}_2 = \mathbf{0}$$

where $\mathbf{v} \in Span(S) = W_1 \cap W_2, \mathbf{w}_1 \in Span(T_1) \subset W_1, \mathbf{w}_2 \in Span(T_2) \subset W_2$. (slice the linear combination into the sum of the vectors from 3 vector spaces). It suffices to prove that in any such potential linear dependence, we must have $\mathbf{v} = \mathbf{w}_1 = \mathbf{w}_2 = \mathbf{0}$ (each vector is a lin comb, and equals $\mathbf{0}$).
Consider $\mathbf{w}_2 = -\mathbf{v} - \mathbf{w}_1$. Since $-\mathbf{v} - \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2$, we must have $\mathbf{w}_2 \in W_1 \cap W_2$. By definition, $\mathbf{w}_2 \in Span(T_2)$ But $S \cap T_2 = \emptyset$, hence $Span(S) \cap Span(T_2) = \{\mathbf{0}\}$. Therefore we must have $\mathbf{w}_2 = \mathbf{0}$. So then $-\mathbf{v} = \mathbf{w}_1 \in W_1 \cap W_2$. Since $S \cap T_1 = \emptyset$, $Span(S) \cap Span(T_1) = \{\mathbf{0}\}$ and we have $\mathbf{w}_1 = \mathbf{0}$, so $\mathbf{v} = \mathbf{0}$ as well. So U is independent.

$$\begin{aligned}
|U| &= |S| + |T_1| + |T_2| \\
&= \dim W_1 \cap W_2 + (\dim W_1 - \dim W_1 \cap W_2) + (\dim W_2 - \dim W_1 \cap W_2) \\
&= \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)
\end{aligned}$$

Exercises for 1.4 1.(k), 7

Exercises for 1.6 1.(d)(e)(f), 3, 4, 16

2 Linear Transformations

2.1 Linear Tranformations

A function T from V to W is denoted by $T : V \to W$. The vector $\mathbf{w} = T(\mathbf{v})$ in W is called the *image* of $\mathbf{v}$ under the function T. Loosely speaking, we want our functions to turn the algebraic operations of addition and scalar multiplication in V into addition and scalar multiplication in W.

Definition 2.1.1 A function $T : V \to W$ is called a *linear mapping* or a *linear transformation* if it satisfies

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}$ and $\mathbf{v} \in$ V
2. $T(a\mathbf{v}) = aT(\mathbf{v})$ for all $a \in \mathbb{R}$ and $\mathbf{v} \in V$

V is called the *domain* of T and W is called the *target* of T.
We say that a linear transformation preserves the operations of addition and scalar multiplication.

Property A linear mapping always takes the zero vector in the domain vector space to the zero vector in the target vector space.

Proposition 2.1.2 A function $T : V \to W$ is a linear transformation if and only if for all a and $b \in \mathbb{R}$ and all $\mathbf{u}$ and $\mathbf{v} \in V$

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$$

Corollary 2.1.3 A function $T : V \to W$ is a linear transformation if and only if for all $a_1, .., a_k \in \mathbb{R}$ and for all $\mathbf{v}_1, ..., \mathbf{v}_k \in V$:

$$T(\sum_{i=1}^{k} a_i \mathbf{v}_i) = \sum_{i=1}^{k} a_i T(\mathbf{v}_i)$$

Remark prove this by induction!

Examples

1. Let V be any vector space, and let W = V. The <u>*identity transformation*</u> $I : V \to V$ is defined by I($\mathbf{v}$) = $\mathbf{v}$ for all $\mathbf{v} \in$ V.

2. Let V and W be any vector spaces, and let $T : V \to W$ be the mapping that takes every vector in V to the zero vector in W:

$$T(\mathbf{v}) = \mathbf{0}_W$$

for all $\mathbf{v} \in$ V. T is called *zero transformation*.

3. *$T(\boldsymbol{x}) = a_1x_1 + ... + a_nx_n$*

4. *Differentiation, definite integration*

Remark The inner product plays a crucial role in linear algebra in that it provides a bridge between algebra and geometry, which is the heart of the more advanced material that appears later in the text.

Proposition 2.1.14 If $T : V \to W$ is a linear transformation and V is finite-dimensional, then T is uniquely determined by its values on the members of a basis of V.
proof: Show that if S and T are linear transformations that take the same values on each member of a basis for V, then in fact S = T.

$$\begin{aligned} T(v) &= T(a_1v_1 + ... + a_kv_k) \\ &= a_1T(v_1) + ... + a_kT(v_k) \\ &= a_1S(v_1) + ... + a_kS(v_k) \\ &= S(a_1v_1 + ... + a_kv_k) \\ &= S(v) \end{aligned}$$

Therefore, S and T are equal as mappings from V to W. ■

2.2 Linear Transformations Between Finite-Dimensional Vector Spaces

Proposition 2.2.1 Let $T : V \to W$ be a linear transformation between the finite-dimensional vector spaces V and W. If $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ is a basis for V and $\{\mathbf{w}_1, ..., \mathbf{w}_l\}$ is a basis for W, then $T : V \to W$ is uniquely determined by the $l \cdot k$ scalars used to express $T(\mathbf{v}_j), j = 1, ..., k$, in terms of $\mathbf{w}_1, ..., \mathbf{w}_l$.

Definition 2.2.6 Let $T: V \to W$ be a linear transformation between the finite-dimensional vector spaces V and W, and let $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ and $\beta = \{\mathbf{w}_1, ..., \mathbf{w}_l\}$, respectively, be any bases for V and W. Let $a_{ij}, 1 \leq i \leq l$ and $1 \leq j \leq k$ be the $l \cdot k$ scalars that determine T with respect to the bases α and β. The matrix whose entries are the scalars a_{ij}, $1 \leq i \leq l$ and $1 \leq j \leq k$, is called the *matrix of the linear transformation T with respect to the bases α for V and β for W*. This matrix is denoted by $[T]_\alpha^\beta$.

Remark The basis vectors in the domain and target spaces are written in some particular order.

Definition of coordinate vectors If $\mathbf{v} = a_1\mathbf{v}_1 + ... + a_k\mathbf{v}_k$ and $\mathbf{w} = b_1\mathbf{w}_1 + ... + b_l\mathbf{w}_l$, we can express $\mathbf{v}$ and $\mathbf{w}$ in coordinates, respectively, as a $k \times 1$ matrix and as an $l \times 1$ matrix, with respect to the chosen bases. These coordinate vectors will be denoted by $[\mathbf{v}]_\alpha$ and $[\mathbf{w}]_\beta$, respectively. Thus $[\mathbf{v}]_\alpha = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$ and $[\mathbf{w}]_\beta = \begin{bmatrix} b_1 \\ \vdots \\ b_l \end{bmatrix}$

Proposition 2.2.15 Let $T: V \to W$ be a linear transformation between vector spaces V of dimension k and W of dimension l. Let $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ be a basis for V and $\beta = \{\mathbf{w}_1, ..., \mathbf{w}_l\}$ be a basis for W. Then for each $\mathbf{v} \in V$,

$$[T(\mathbf{v})]_\beta = [T]_\alpha^\beta [\mathbf{v}]_\alpha$$

<u>*proof:*</u> Let $\mathbf{v} = x_1\mathbf{v}_1 + ... + x_k\mathbf{v}_k \in V$. Then if $T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + ... + a_{lj}\mathbf{w}_l$

$$\begin{aligned} T(\mathbf{v}) &= \sum_{j=1}^{k} x_j T(\mathbf{v}_j) \\ &= \sum_{j=1}^{k} x_j (\sum_{i=1}^{l} a_{ij}\mathbf{w}_i) \\ &= \sum_{i=1}^{l} (\sum_{j=1}^{k} x_j a_{ij}) \mathbf{w}_i \end{aligned}$$

Thus, the ith coefficient of T($\mathbf{v}$) in terms of β is $\sum_{j=1}^{k} x_j a_{ij}$ and $[T(\mathbf{v})]_\beta = \begin{bmatrix} \sum_{j=1}^{k} x_j a_{1j} \\ \vdots \\ \sum_{j=1}^{k} x_j a_{lj} \end{bmatrix}$ which is precisely $[T(\mathbf{v})]_\beta = [T]_\alpha^\beta [\mathbf{v}]_\alpha$. ■

Proposition 2.2.19 Let $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ be a basis for V and $\beta = \{\mathbf{w}_1, ..., \mathbf{w}_l\}$ be a basis for W, and let $\mathbf{v} = x_1\mathbf{v}_1 + ... + x_k\mathbf{v}_k \in V$

1. If A is an $l \times k$ matrix, then the function

$$T(\mathbf{v}) = \mathbf{w}$$

   where $[\mathbf{w}]_\beta = A[\mathbf{v}]_\alpha$ is a linear transformation.

2. If $A = [S]_\alpha^\beta$ is the matrix of a transformation $S : V \to W$, then the transformation T constructed from $[S]_\alpha^\beta$ is equal to S.

3. If T is the transformation of (1) constructed from A, then $[T]_\alpha^\beta = A$

Proposition 2.2.20 Let V and W be finite-dimensional vector spaces. Let α be a basis for V and β a basis for W. Then the assignment of a matrix to a linear transformation from V to W given by T goes to $[T]_\alpha^\beta$ is injective and surjective.

Notes

1. When proving a function T is not a linear transformation, can consider $T(\mathbf{0}) \neq \mathbf{0}$.

2.3 Kernel and Image

Definition 2.3.1 The *kernel* of T, denoted $Ker(T)$, is the subset of V consisting of all vectors $\mathbf{v} \in V$ such that $T(\mathbf{v}) = 0$.

Remark Kernel is different from null spaces. A null space is about a matrix, and it is something in $\mathbb{R}^n$.

Proposition 2.3.2 Let $T : V \to W$ be a linear transformation. Ker(T) is a subspace of V.

Examples

1. Let $V = P_3(\mathbb{R})$. Define $T : V \to V$ by $T(p(x)) = \frac{d}{dx}p(x)$. Ker(T) only consists constant polynomials.

2. Let $V = W = \mathbb{R}^2$. Let T be a rotation R_θ. Then $Ker(T) = \{\mathbf{0}\}$.

Proposition 2.3.7 Let $T : V \to W$ be a linear transformation of finite-dimensional vector spaces, and let α and β be bases for V and W, respectively. Then $\mathbf{x} \in Ker(T)$ ife lf the coordinate vector of $\mathbf{x}$, $[\mathbf{x}]_\alpha$, satisfies the system of equations

$$\begin{gathered} a_{11}x_1 + ... + a_{1k}x_k = 0 \\ \vdots \\ a_{l1}x_1 + ... + a_{lk}x_k = 0 \end{gathered}$$

where the coefficients a_{ij} are the entries of the matrix $[T]_\alpha^\beta$.

Remark This says

$$x \in \ker(T) \iff [x]_\alpha \in Nul[T]_\alpha^\beta$$

Proposition 2.3.8 Independence is Basis-Independent Let V be a finite-dimensional vector space, and let $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ be a basis for V. Then the vectors $\mathbf{x}_1, ..., \mathbf{x}_m \in V$ are linearly independent iff their corresponding coordinate vectors $[\mathbf{x}_1]_\alpha, ..., [\mathbf{x}_m]_\alpha$ are linearly independent.

Definition 2.3.10 The subset of W consisting of all vectors $\mathbf{w} \in W$ for which there exists a $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$ is called the *image* of T and is denoted by $Im(T)$.

Proposition 2.3.11 Let $T : V \to W$ be a linear transformation. The image of T is a subspace of W.

Proposition 2.3.12 If $\{\mathbf{v}_1, ..., \mathbf{v}_m\}$is any set that spans V (in particular, it could be a basis of V), then $\{T(\mathbf{v}_1), ..., T(\mathbf{v}_m)\}$ spans $Im(T)$.

Corollary 2.3.13 If $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ is a basis for V and $\beta = \{\mathbf{w}_1, ..., \mathbf{w}_l\}$ is a basis for W, then the vectors in W, whose coordinate vectors (in terms of β) are the columns of $[T]_\alpha^\beta$, span $Im(T)$.

Rank-Nullity Theorem 2.3.17 If V is finite-dimensional vector space and $T : V \to W$ is a linear transformation, then

$$\dim(Ker(T)) + \dim(Im(T)) = \dim(V)$$

Equivalently,

$$Nullity(T) + Rank(T) = \dim(V)$$

2.4 Applications of the Dimension Theorem

Proposition 2.4.2 A linear transformation $T : V \to W$ is injective iff $\dim(Ker(T)) = 0$, or $\dim(Im(T)) = \dim(V)$.

Remark Analogously, in MAT223 we said that a matrix is one-to-one if all the columns are l.i..

Corollary 2.4.3 A linear mapping $T : V \to W$ on a finite-dimensional vector space V is injective iff $\dim(Im(T)) = \dim(V)$.

Corollary 2.4.4 If $\dim(W) < \dim(V)$ and $T : V \to W$ is a linear mapping, then T is not injective.
proof:

$$\dim(Im(T)) \leq \dim(W) < \dim(V)$$
$$\implies dim(Ker(T)) > 0$$

Proposition 2.4.7 If W is finite-dimensional, then a linear mapping $T : V \to W$ is surjective iff $\dim(Im(T)) = \dim(W)$

Remark Analogously, in MAT223 we said that a matrix $A \in M_{m\times n}(\mathbb{R})$ is onto if there is a pivot in every row, or the columns of A spans $\mathbb{R}^m$.

Corollary 2.4.8 If V and W are finite-dimensional, with $\dim(V) < \dim(W)$, then there is no surjective linear mapping $T : V \to W$
proof: $\dim(Im(T)) \leq \dim(V) < \dim(W) \implies T$ is not surjective

Corollary 2.4.9 A linear mapping $T : V \to W$ can be surjective iff

$$\dim(V) \geq \dim(W)$$

Proposition 2.4.10 Let $\dim(V) = \dim(W)$. A linear transformation $T : V \to W$ is injective iff it is surjective.

Proposition 2.4.11 Let $T : V \to W$ be a linear transformation, and let $w \in Im(T)$. Let v_1 be any fixed vector with $T(v_1) = w$. Then every vector $v_2 \in T^{-1}(\{w\})$ can be written uniquely as $v_2 = v_1 + u$, where $u \in Ker(T)$

Remark In this situation $T^{-1}(\{w\})$ is a subspace of V iff $w = 0$.

Corollary 2.4.15 Let $T : V \to W$ be a linear transformation of finite-dimensional vector spaces, and let $w \in W$. Then there is a unique vector $v \in V$ such that $T(v) = w$ iff

1. $w \in Im(T)$, and
2. dim(Ker(T)) = 0

Proposition 2.4.16 With notation as before

1. The set of solutions of the system of linear equations $A\mathbf{x} = \mathbf{b}$ is the subset $T^{-1}(\{\mathbf{b}\})$ of $V = \mathbb{R}^n$
2. The set of solutions of the system of linear equations $A\mathbf{x} = \mathbf{b}$ is a subspace of V iff the system is homogeneous, in which case the set of solutions is $Ker(T)$.

Corollary 2.4.17

1. The number of free variables in the homogeneous system $A\mathbf{x} = \mathbf{0}$ (or its echelon form equivalent) is equal to $\dim(Ker(T))$
2. The number of basic variables of the system is equal to $\dim(Im(T))$

Definition 2.4.18 Given an inhomogeneous system of equations, $A\mathbf{x} = \mathbf{b}$, any single vector $\mathbf{x}$ satisfying the system (necessarily $\mathbf{x} \neq \mathbf{0}$) is called a <u>particular solution</u> of the system of equations.

Proposition 2.4.19 Let $\mathbf{x}_p$ be a particular solution of the system $A\mathbf{x} = \mathbf{b}$. Then every other solution to $A\mathbf{x} = \mathbf{b}$ is of the form $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h$, where $\mathbf{x}_h$ is a solution of the corresponding homogeneous system of equations $A\mathbf{x} = \mathbf{0}$. Furthermore, given $\mathbf{x}$ and $\mathbf{x}_p$, there is a unique $\mathbf{x}_h$ such that $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h$.

Corollary 2.4.20 The system $A\mathbf{x} = \mathbf{b}$ has a unique solution iff $\mathbf{b} \in Im(T)$ and the only solution to $A\mathbf{x} = \mathbf{0}$ is the zero vector.

2.5 Composition of Linear Transformations

Definition Let U, V, and W be vector spaces, and let $S : U \to V$ and $T : V \to W$ be linear transformations. The *composition* of S and T is denoted $TS : U \to W$ and is defined by

$$TS(\mathbf{v}) = T(S(\mathbf{v}))$$

Notice that this is well defined since the image of S is contained in V, which is the domain of T.

Proposition 2.5.1 Let $S : U \to V$ and $T : V \to W$ be linear transformations, then TS is a linear transformation.

Remark In general, ST is not equal to TS. We emphasize that the composition is well defined only if the image of the first transformation is contained in the domain of the second.

Proposition 2.5.4

1. Let $R : U \to V, S : V \to W$ and $T : W \to X$ be linear transformations of the vector space U,V,W and X as indicated. Then

$$T(SR) = (TS)R \text{ (associativity)}$$

2. Let $R : U \to V, S : V \to W$ and $T : W \to X$ be linear transformations of the vector space U,V,W and X as indicated. Then

$$T(R + S) = TR + TS \text{ (distributivity)}$$

3. Let $R : U \to V, S : V \to W$ and $T : W \to X$ be linear transformations of the vector space U,V,W and X as indicated. Then

$$(T + S)R = TR + SR \text{ (distributivity)}$$

Proposition 2.5.6 Let $S : U \to V$ and $T : V \to W$ be linear transformations. Then

1. $Ker(S) \subset Ker(TS)$
2. $Im(TS) \subset Im(T)$

proof:

1. If $\mathbf{u} \in Ker(S), S(\mathbf{u}) = \mathbf{0}$. Then $TS(\mathbf{u}) = T(\mathbf{0}) = \mathbf{0}$. Therefore $\mathbf{u} \in Ker(TS)$.
2. If $\mathbf{x} \in Im(TS)$, then $\exists \mathbf{u} \in U$ s.t. $TS(\mathbf{u}) = T(S(\mathbf{u})) = \mathbf{x}$, then $\exists \mathbf{v} = S(\mathbf{u}) \in V$ s.t. $T(\mathbf{v}) = \mathbf{x}$. Therefore $\mathbf{x} \in Im(T)$ ■

Corollary 2.5.7 Let $S : U \to V$ and $T : V \to W$ be linear transformations of finite-dimensional vector spaces. Then

1. $\dim(Ker(S)) \leq \dim(Ker(TS))$
2. $\dim(Im(TS)) \leq \dim(Im(T))$

Proposition 2.5.9 If $[S]_\alpha^\beta$ has entries a_{ij}, $i = 1, \ldots, n$ and $j = 1, \ldots, m$ and $[T]_\beta^\gamma$ has entries b_{kl}, $k = 1, \ldots, p$ and $l = 1, \ldots, n$, then the entries of $[TS]_\alpha^\gamma$ are $\sum_{l=1}^n b_{kl} a_{lj}$

Definition 2.5.10 Let A be an $n \times m$ matrix and B a $p \times n$ matrix, then the *matrix product* BA is defined to be the $p \times m$ matrix whose entries are $\sum_{l=1}^n b_{kl} a_{lj}$ for $k = 1, \ldots, p$ and $j = 1, \ldots, m$.

Proposition 2.5.13 Let $S : U \to V$ and $T : V \to W$ be linear transformations between finite-dimensional vector spaces. Let α, β and γ be bases for U,V and W, respectively. Then

$$[TS]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta$$

In words, the matrix of the composition of two linear transformations is the product of the matrices of the transformations

Proposition 2.5.14

1. Let A,B and C be $m \times n, n \times p$ and $p \times r$ matrices, then
$$(AB)C = A(BC) \text{ (associativity)}$$
2. Let A,B and C be $m \times n, n \times p$ and $p \times r$ matrices, then
$$A(B + C) = AB + AC \text{ (distributivity)}$$
3. Let A,B and C be $m \times n, n \times p$ and $p \times r$ matrices, then
$$(A + B)C = AC + BC \text{ (distributivity)}$$

3 Problem Notes

1 $S = \{\mathbf{a}\} \subseteq \mathbb{R}^2$, then we cannot determine whether S is dependent (when $\mathbf{a} = \mathbf{0}$) or independent (when $\mathbf{a} \neq \mathbf{0}$)

2 If a set in a vector space contains the zero vector, then it is linearly dependent.

3 Multiple Choice questions: check the hypothesis of theorems!