

Linear algebra notes (based on Friedberg's text)

Abiyaz Chowdhury

August 11, 2018

1 Vector Spaces

1.1 Vector Spaces

Definition 1.1 A vector space (or linear space) V over a field F consists of a set on which two operations (called addition and scalar multiplication, respectively) are defined so that for each pair of elements x, y in V there is a unique element $x + y$ in V , and for each element a in F , and each element $x \in V$ there is a unique element ax in V , such that the following conditions hold:

1. For all $x, y \in V$, $x + y = y + x$. (commutativity of addition)
2. For all $x, y, z \in V$, $(x + y) + z = x + (y + z)$. (associativity of addition)
3. There exists an element in V denoted by 0 such that $x + 0 = x$ for each $x \in V$. (identity under addition)
4. For each element $x \in V$ there is an element $y \in V$ such that $x + y = 0$. (existence of additive inverse)
5. For each element $x \in V$, $1x = x$. (identity under scalar multiplication)
6. For each pair of elements $a, b \in F$ and each element $x \in V$, $(ab)x = a(bx)$. (associativity under scalar multiplication)
7. For each element $a \in F$ and each pair of elements $x, y \in V$, $a(x + y) = ax + ay$. (scalar multiplication distributes over addition)
8. For each pair of elements $a, b \in F$ and each element $x \in V$, $(a + b)x = ax + bx$. (scalar multiplication distributes over addition)

The elements $x + y$ and ax are called the sum of x and y and the product of a and x , respectively. The elements of the field F are called scalars and the elements of the vector space V are called vectors. The element 0 as defined in the third item is called the zero vector of V .

Definition 1.2 An object of the form (a_1, a_2, \dots, a_n) where the entries a_1, a_2, \dots, a_n are elements of a field F , is called an n -tuple with entries from F . The elements a_1, a_2, \dots, a_n are called the entries or components of the n -tuple. Two n -tuples (a_1, a_2, \dots, a_n) and (b_1, b_2, \dots, b_n) with entries from a field F are called equal if $a_i = b_i$ for all $i = 1, 2, \dots, n$.

Definition 1.3 The set of all n -tuples with entries from a field F is denoted by F^n .

Definition 1.4 The sum of two n -tuples (a_1, a_2, \dots, a_n) and (b_1, b_2, \dots, b_n) in F^n yields the n -tuple $(c_1, c_2, \dots, c_n) \in F^n$ where $c_i = a_i + b_i$ for all $i = 1, 2, \dots, n$.

Definition 1.5 The scalar product of an n -tuple (a_1, a_2, \dots, a_n) in F^n and a scalar $c \in F$ yields the n -tuple $(c_1, c_2, \dots, c_n) \in F^n$ where $c_i = c[a_i]$ for all $i = 1, 2, \dots, n$.

Theorem 1.1 F^n , under the operations defined above, is a vector space over F , with zero vector $(0, 0, \dots, 0)$ where 0 is the additive identity of F .

Definition 1.6 An $m \times n$ matrix with entries from a field F is a rectangular array of the form

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

where each entry a_{ij} ($1 \leq i \leq m, 1 \leq j \leq n$) is an element of F . We call the entries a_{ij} with $i = j$ the diagonal entries of the matrix. The entries $a_{i1}, a_{i2}, \dots, a_{in}$ compose the i th row of the matrix. The rows of the preceding matrix are regarded as vectors in F^n , and columns are regarded as vectors in F^m . The $m \times n$ matrix in which each entry equals zero is called the zero matrix and is denoted by O .

Definition 1.7 In the preceding matrix, m is the number of rows and n is the number of columns of the matrix. The entry of matrix A that lies in row i and column j is denoted by A_{ij} . If a matrix has an equal number of rows and columns, it is called square. Two $m \times n$ matrices A and B are called equal if all their corresponding entries are equal, that is, if $A_{ij} = B_{ij}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.

Definition 1.8 An $m \times n$ matrix A where $A_{ij} = 0$ whenever $i \neq j$ is called a diagonal matrix.

Definition 1.9 An $m \times n$ matrix A where $A_{ij} = 0$ whenever $i > j$ is called an upper triangular matrix.

Definition 1.10 An $m \times n$ matrix A where $A_{ij} = 0$ whenever $i < j$ is called a lower triangular matrix.

Definition 1.11 The set of all $m \times n$ matrices with entries from a field F is denoted by $M_{m \times n}(F)$.

Definition 1.12 Addition of two $m \times n$ matrices A and B yields the $m \times n$ matrix C where $C_{ij} = A_{ij} + B_{ij}$.

Definition 1.13 Scalar Multiplication of an $m \times n$ matrix A with entries from F and a scalar $c \in F$ yields an $m \times n$ matrix C with entries from F where $C_{ij} = cA_{ij}$.

Theorem 1.2 $M_{m \times n}(F)$, under the operations defined above, is a vector space over F , with zero vector A where $A_{ij} = 0$ for all $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$ and 0 is the additive identity of F .

Definition 1.14 Matrix multiplication of an $m \times n$ matrix A and a $n \times p$ matrix B yields the $m \times p$ matrix C where $C_{ij} = \sum_{k=1}^n A_{ik}B_{kj}$.

Definition 1.15 Let S be any nonempty set and F be any field, and let $\mathcal{F}(S, F)$ denote the set of all functions from S to F . Two functions f and g in $\mathcal{F}(S, F)$ are called equal if $f(s) = g(s)$ for each $s \in S$. We define the sum of two functions $f, g \in \mathcal{F}(S, F)$ as the function $h \in \mathcal{F}(S, F)$ where $h(s) = f(s) + g(s)$. We define the scalar multiplication of a function $f \in \mathcal{F}(S, F)$ by a scalar $c \in F$ as the function $h \in \mathcal{F}(S, F)$ where $h(s) = c[f(s)]$.

Theorem 1.3 $\mathcal{F}(S, F)$ is a vector space, under the operations defined above, with zero vector $f \in \mathcal{F}(S, F)$ where $f(s) = 0$ for all $s \in S$ is the zero vector in $\mathcal{F}(S, F)$.

Definition 1.16 A polynomial with coefficients from a field F is an expression of the form $f(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ where n is a nonnegative integer and each a_k , called the coefficient of x^k , is in F . If $f(x) = 0$, that is, if $a_n = a_{n-1} = \dots = a_0 = 0$, then $f(x)$ is called the zero polynomial, and the degree is defined as -1 . Otherwise, the degree is defined as the largest exponent of x that appears in the representation $f(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ with a nonzero coefficient. A nonzero scalar in F therefore is a polynomial of degree 0. Two polynomials $f(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ and $g(x) = b_nx^n + b_{n-1}x^{n-1} + \dots + b_1x + b_0$ are called equal if $m = n$ and $a_i = b_i$ for $i = 0, 1, \dots, n$. A polynomial f can be evaluated at a point x_0 by replacing in the polynomial every occurrence of the polynomial's variable with x_0 . The result thus obtained is an element of F , and is written as $f(x_0)$.

Example 1.1 The sum of two polynomials $f, g \in \mathcal{F}(S, F)$ having degrees m, n respectively

Definition 1.17 A sequence in F is a function σ from the positive integers into F . A sequence is typically represented as an infinite series of numbers, i.e. a_1, a_2, \dots where the entries are elements in F . We can represent this sequence as simply $\{a_n\}$. The sum of two sequences $\{a_n\}$ and $\{b_n\}$ yields a sequence $\{c_n\}$ in F where $c_i = a_i + b_i$ and the scalar product of a sequence $\{a_n\}$ in F and a scalar $c \in F$ yields a sequence $\{c_n\}$ in F where $c_i = c[a_i]$.

Theorem 1.4 The set of all sequences in F , forms a vector space over F under the operations defined above.

Theorem 1.5 If V is a vector space and $x, y, z \in V$ such that $x + z = y + z$, then $x = y$.

Theorem 1.6 The zero vector of any vector space V is unique.

Theorem 1.7 The additive identity of any vector is unique. We typically denote the additive identity of x by $-x$.

Theorem 1.8 For a vector space V , the following are true:

1. $0x = 0$ for all $x \in V$.
2. $(-a)x = -(ax) = a(-x)$ for all $a \in F$ and each $x \in V$.
3. $a0 = 0$ for each $a \in F$.

Theorem 1.9 For a vector space V , the following are true:

1. $0x = 0$ for all $x \in V$.
2. $(-a)x = -(ax) = a(-x)$ for all $a \in F$ and each $x \in V$.
3. $a0 = 0$ for each $a \in F$.

Example 1.2 Define the set $V = \{0\}$ where $0 + 0 = 0$ and $c0 = 0$ for all $c \in F$. Then V is called the zero vector space.

1.2 Subspaces

Definition 1.18 A subset W of a vector space V over a field F is called a subspace of V if W is also a vector space over F with the same operations of addition and scalar multiplication as defined on V .

Example 1.3 For any vector space V , note that V and $\{0\}$ are subspaces of V , where 0 is the zero vector of V .

Theorem 1.10 For a vector space V , a subset W of V is a subspace of V if and only if the following are all true:

1. $x + y \in W$ if $x \in W$ and $y \in W$.
2. $cx \in W$ if $x \in W$ and $c \in F$.
- 3.

Definition 1.19 The transpose A^t of an $m \times n$ matrix A is the $n \times m$ matrix defined by $A_{ij}^t = A_{ji}$.

Definition 1.20 A symmetric matrix is one that equals its own transpose. Note that symmetric matrices are always square.

Example 1.4 The set of all symmetric $n \times n$ matrices is a subspace of $M_{n \times n}(F)$.

Example 1.5 The set of all diagonal $n \times n$ matrices is a subspace of $M_{n \times n}(F)$.

Example 1.6 The set of all upper triangular $m \times n$ matrices as well as the set of all lower triangular $m \times n$ matrices are both subspaces of $M_{m \times n}(F)$.

Definition 1.21 The trace of an $n \times n$ matrix M , denoted by $\text{tr}(M)$, is the sum of the diagonal entries of M , i.e.

$$\text{tr}(M) = \sum_{i=1}^n M_{ii}$$

Example 1.7 The set of all $n \times n$ matrices having trace equal to 0 is a subspace of $M_{n \times n}(F)$.

Theorem 1.11 Any intersection of subspaces of a vector space V is also a subspace of V .

Definition 1.22 If W_1 and W_2 are both subspaces of V , then $W_1 + W_2$ is defined as the set $\{w_1 + w_2 | w_1 \in W_1, w_2 \in W_2\}$. This set is called the set sum or simply the sum of W_1, W_2 . If it is also true that $W_1 \cap W_2 = \{0\}$, then $W_1 + W_2$ is called the direct sum of W_1 and W_2 and we use the stricter notation $W_1 \oplus W_2$. Note that the direct sum is only defined for subspaces whose intersection is only at the zero vector but the set sum is defined for any two subspaces of the same vector space.

Theorem 1.12 If W_1 and W_2 are both subspaces of V , then $W_1 \oplus W_2$ is the direct sum of W_1 and W_2 if and only if every vector in $W_1 \oplus W_2$ can be uniquely represented as the sum of a vector in W_1 and a vector in W_2 .

Example 1.8 A matrix M is skew-symmetric if $M^t = -M$. A skew-symmetric matrix must be square. The set of all skew-symmetric $n \times n$ matrices is a subspace of $M_{n \times n}(F)$.

Example 1.9 The direct sum of the set of all skew-symmetric $n \times n$ matrices and the set of all symmetric $n \times n$ matrices is $M_{n \times n}(F)$.

Definition 1.23 If W is a subspace of V over F , then for $v \in V$, the set $v + W = \{v + w | w \in W\}$ is called the coset of W containing V .

Example 1.10 The coset $v + W$ is a subspace of V if and only if $v \in W$.

Example 1.11 If $v_1, v_2 \in V$, then $v_1 + W = v_2 + W$ if and only if $v_1 - v_2 \in W$.

Definition 1.24 The sum of two cosets $v_1 + W$ and $v_2 + W$ is defined as $(v_1 + v_2) + W$. Likewise, the multiplication of a coset $v + W$ by a scalar $c \in F$ is defined as $cv + W$.

Example 1.12 The collection of all cosets of W with the operations defined above is a vector space, known as the quotient space of V modulo W and is denoted by V/W .

1.3 Linear combinations and systems of linear equations

Theorem 1.13 If V is a vector space over F , and $v_1, \dots, v_n \in V$ and $c_1, \dots, c_n \in F$, then $\sum_{i=1}^n c_i v_i$ is in V and is called a linear combination of the vectors v_1, \dots, v_n . The scalars c_1, \dots, c_n are called the coefficients of the linear combination.

Definition 1.25 If S is a nonempty subset of a vector space V , then we define the span of S , denoted $\text{span}(S)$ as the set of all linear combinations of the vectors in S . We define $\text{span}(\emptyset) = \{0\}$.

Theorem 1.14 If S is a nonempty subset of a vector space V , $v \in \text{span}(S)$ if and only if v can be represented as a linear combination of finitely many vectors in S .

Theorem 1.15 If S is a nonempty subset of a vector space V , then $\text{span}(S)$ is equivalent to the intersection of all subspaces of V that contain S .

Theorem 1.16 If S is a nonempty subset of a vector space V , then $\text{span}(S)$ is a subspace of V . Furthermore, any subspace that contains S must contain $\text{span}(S)$.

Definition 1.26 A subset S of a vector space V generates (or spans) V if $\text{span}(S) = V$.

Example 1.13 If W is a subspace of a vector space V , then W is a subspace of V if and only if $\text{span}(W) = W$.

Example 1.14 (Idempotence of span) For any subset S of a vector space V , $\text{span}(\text{span}(S)) = \text{span}(S)$.

Example 1.15 If $S_1, S_2 \subseteq V$, and $S_1 \subseteq S_2$, then $\text{span}(S_1) \subseteq \text{span}(S_2)$. In particular, if $S_1 \subseteq S_2$, and $\text{span}(S_1) = V$, then $\text{span}(S_2) = V$.

Example 1.16 If $S_1, S_2 \subseteq V$, and $S_1 \subseteq S_2$, then $\text{span}(S_1) \subseteq \text{span}(S_2)$. In particular, if $S_1 \subseteq S_2$, and $\text{span}(S_1) = V$, then $\text{span}(S_2) = V$.

Example 1.17 If S_1 and S_2 are arbitrary subsets of a vector space V , then $\text{span}(S_1 \cup S_2) = \text{span}(S_1) + \text{span}(S_2)$.

Example 1.18 If S_1 and S_2 are arbitrary subsets of a vector space V , then $\text{span}(S_1 \cap S_2) \subseteq \text{span}(S_1) \cap \text{span}(S_2)$.

1.4 Linear dependence and linear independence

Definition 1.27 A nonempty subset S of a vector space V is called linearly dependent (or just dependent) if there exists a finite number of distinct vectors $u_1, \dots, u_n \in S$ and scalars a_1, \dots, a_n not all zero, such that $a_1 u_1 + \dots + a_n u_n = 0$. The vectors of S are also said to be linearly dependent. A set that is not linearly dependent is linearly independent (or just independent). Note that the empty set is linearly independent, since dependent sets must be nonempty.

Theorem 1.17 Sets containing the zero vector are dependent, and a singleton set consisting of a nonzero vector is independent.

Theorem 1.18 A subset S of a vector space V is independent if and only if every representation of $0 \in V$ as a linear combination of vectors in S has all its coefficients equal to 0.

Theorem 1.19 A subset S of a vector space V is independent if and only if every finite subset of S is independent.

Theorem 1.20 A subset S of a vector space V is dependent if and only if at least one of its vectors can be represented as a linear combination of finitely many other vectors in S .

Theorem 1.21 Let V be a vector space, and let $S_1 \subseteq S_2 \subseteq V$. Then if S_1 is dependent, so is S_2 . Stating the contrapositive, if S_2 is independent, so is S_1 .

Theorem 1.22 Let S be an independent subset of a vector space V and let v be a vector not in S . Then $S \cup \{v\}$ is dependent if and only if $v \in \text{span}(S)$.

1.5 Bases and dimension

Definition 1.28 A basis β for a vector space V is a linearly independent subset of V that generates V . If β is a basis for V , we also say that the vectors of β form a basis for V .

Example 1.19 In F^n , let $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, ..., $e_n = (0, \dots, 0, 1)$. Then the set $\{e_1, e_2, \dots, e_n\}$ is a basis for F^n and is called the standard basis for F^n .

Example 1.20 The set $\{1, x, x^2, \dots\}$ is a basis for the set of all polynomials in x .

Theorem 1.23 Let V be a vector space and $\beta = \{u_1, u_2, \dots, u_n\}$ be a subset of V . Then β is a basis for V if and only if each $v \in V$ can be uniquely expressed as a linear combination of vectors in β .

Theorem 1.24 If a vector space V is generated by a finite set S , then some subset of S is a basis for V . Hence V has a finite basis.

Theorem 1.25 (Replacement Theorem) Let V be a vector space that is generated by a set G containing exactly n vectors, and let L be an independent subset of V containing exactly m vectors. Then $m \leq n$ and there exists a subset H of G containing exactly $n - m$ vectors such that $L \cup H$ generates V .

Theorem 1.26 Let V be a vector space having a finite basis. Then every basis for V contains the same number of vectors. The number of vectors in such a basis is called the dimension of the vector space, and is denoted by $\dim(V)$.

Definition 1.29 A vector space is called finite-dimensional if it has a finite basis. A vector space that is not finite-dimensional is said to be infinite-dimensional.

Theorem 1.27 A vector space is infinite-dimensional if and only if it contains an infinite linearly independent subset.

Theorem 1.28 Let V be a vector space with dimension n . The following are all true:

1. Any finite spanning set for V contains at least n vectors, and any spanning set for V that contains exactly n vectors is a basis for V .
2. Any independent subset of V contains at most n vectors, and any independent subset of V that contains exactly n vectors is a basis for V .
3. Every independent subset of V can be extended to a basis for V (via the Replacement Theorem).
4. Every spanning set for V can be reduced to a basis for V , that is, if $\text{span}(S) = V$, then some subset of S is a basis for V .

Theorem 1.29 Let V be a vector space with dimension n and let S be a subset of V . If any two of the following three statements are true, the third statement must also be true:

1. S is linearly independent.
2. $\text{span}(S) = V$.
3. S contains n vectors.

Theorem 1.30 Let W be a subspace of a finite-dimensional vector space V . Then W is finite-dimensional and $\dim(W) \leq \dim(V)$. Moreover, if $\dim(W) = \dim(V)$, then $V = W$.

Theorem 1.31 If W is a subspace of a finite-dimensional vector space V , then any basis for W can be extended to a basis for V .

Example 1.21 If W_1, W_2 are m and n dimensional subspaces respectively of a vector space V , where $m \geq n$, then $\dim(W_1 \cap W_2) \leq n$, and $\dim(W_1 + W_2) \leq m + n$.

Example 1.22 If W_1, W_2 are subspaces of a vector space V such that $V = W_1 \oplus W_2$, and if β_1, β_2 are bases for W_1, W_2 respectively, then $\beta_1 \cap \beta_2 = \emptyset$ and $\beta_1 \cup \beta_2$ is a basis for V . Conversely, if β_1, β_2 are disjoint bases for the subspaces W_1, W_2 of the vector space V , then if $\beta_1 \cup \beta_2$ is a basis for V , then $V = W_1 \oplus W_2$.

Definition 1.30 Let c_0, c_1, \dots, c_n be distinct vectors in an infinite field F . The polynomials $f_0(x), f_1(x), \dots, f_n(x)$ defined by

$$f_i(x) = \frac{(x - c_0) \dots (x - c_{i-1})(x - c_{i+1}) \dots (x - c_n)}{(c_i - c_0) \dots (c_i - c_{i-1})(c_i - c_{i+1}) \dots (c_i - c_n)}$$

are called the Lagrange polynomials associated with the scalars c_0, c_1, \dots, c_n .

Theorem 1.32 The polynomials $f_0(x), f_1(x), \dots, f_n(x)$ as defined above form a basis for $P_n(F)$. Any polynomial $g \in P_n(F)$ can be represented uniquely as the linear combination $g = \sum_{i=0}^n g(c_i)f_i$.

Theorem 1.33 Given a set of $n+1$ points $(c_0, b_0), (c_1, b_1), \dots, (c_n, b_n) \in F^2$ (where F is an infinite field) with distinct c_0, c_1, \dots, c_n , the Lagrange interpolation of these points is the polynomial $g = \sum_{i=0}^n b_i f_i$ where $f_0(x), f_1(x), \dots, f_n(x)$ are the Lagrange polynomials associated with c_0, c_1, \dots, c_n . Note that $g(c_i) = b_i$ for all $i = 0, 1, 2, \dots, n$, meaning that g passes through the given $n+1$ points, and that there is no other polynomial in $P_n(F)$ that also passes through all of these $n+1$ points.

Example 1.23 If W_1, W_2 are finite-dimensional subspaces of a vector space V , then

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

1.6 Maximal linearly independent subsets

Definition 1.31 Let \mathcal{F} be a collection of sets. A member M of \mathcal{F} is called maximal (with respect to set inclusion) if M is contained in no member of \mathcal{F} other than M itself.

Definition 1.32 A collection of sets \mathcal{C} is called a chain (or nest or tower) if for each pair of sets A and B in \mathcal{C} , either $A \subseteq B$ or $B \subseteq A$.

Theorem 1.34 Given a chain \mathcal{C} , and a finite collection of sets $A_1, \dots, A_n \in \mathcal{C}$, one of these sets contains all the others; that is, there is some $i \in \{1, 2, \dots, n\}$ such that $A_k \subseteq A_i$ for $k = 1, 2, \dots, n$.

Axiom 1 (Maximal Principle) Let \mathcal{F} be a family of sets. If, for each chain $\mathcal{C} \subseteq \mathcal{F}$, there exists a member of \mathcal{F} that contains each member of \mathcal{C} , then \mathcal{F} contains a maximal member.

Definition 1.33 Let S be a subset of a vector space V . A maximal linearly independent subset of S is a subset B of S satisfying both of the following conditions:

1. B is linearly independent.
2. The only linearly independent subset of S that contains B is B itself.

Theorem 1.35 Let V be a vector space and S a subset that generates V . If β is a maximal linearly independent subset of S if and only if β is a basis for V .

Theorem 1.36 Let S be a linearly independent subset of a vector space V . There exists a maximal linearly independent subset of V that contains S .

Theorem 1.37 Every vector space has a basis.

Example 1.24 If W is a subspace of a vector space V , then every basis for W is a subset of some basis for V .

2 Linear Transformations and Matrices

2.1 Linear transformations, null spaces, and ranges

Definition 2.1 Let V and W be vector spaces (over F). We call a function $T : V \rightarrow W$ a linear transformation from V to W if, for all $x, y \in V$, and $c \in F$, we have:

1. $T(x + y) = T(x) + T(y)$
2. $T(cx) = cT(x)$

Theorem 2.1 Properties of linear transformations:

1. If T is linear, $T(0) = 0$.
2. T is linear if and only if $T(cx + y) = cT(x) + T(y)$ for all $x, y \in V$, and $c \in F$.
3. If T is linear, then $T(x - y) = T(x) - T(y)$ for all $x, y \in V$
4. T is linear if and only if for all $x_1, x_2, \dots, x_n \in V$ and $a_1, a_2, \dots, a_n \in F$, we have

$$T\left(\sum_{i=1}^n a_i x_i\right) = \sum_{i=1}^n a_i T(x_i)$$

Example 2.1 The transformation $T_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T_\theta(a_1, a_2) = (a_1 \cos \theta - a_2 \sin \theta, a_1 \sin \theta + a_2 \cos \theta)$ is a linear transformation, and is called rotation by θ .

Example 2.2 The transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(a_1, a_2) = (a_1, -a_2)$ is a linear transformation, and is called reflection about the x-axis. Reflection about the y-axis is defined similarly, and is also a linear transformation.

Example 2.3 The transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(a_1, a_2) = (a_1, 0)$ is a linear transformation, and is called projection on the x-axis. Projection on the y-axis is defined similarly, and is also a linear transformation.

Definition 2.2 The transformation $I_V : V \rightarrow V$ defined by $I_V(x) = x$ for all $x \in V$ is called the identity transformation in V . The transformation $T_0 : V \rightarrow W$ defined by $T_0(x) = 0$ for all $x \in V$ is called the zero transformation from V to W . Both of these are linear transformations, and I_V is often simply written as I .

Definition 2.3 Let V and W be vector spaces, and let $T : V \rightarrow W$ be linear. The null space (or kernel) of T , denoted as $N(T)$, is the subset of V that is mapped by T onto $0 \in W$.

$$N(T) = \{v \in V | T(v) = 0\}$$

The range of T , denoted by $R(T)$, is the subset of W that is mapped onto by some vector in V .

$$R(T) = \{w \in W | \exists v \in V : T(v) = w\}$$

Theorem 2.2 Let V and W be vector spaces and $T : V \rightarrow W$ be linear. Then $N(T)$ and $R(T)$ are subspaces of V and W , respectively.

Theorem 2.3 Let V and W be vector spaces and $T : V \rightarrow W$ be linear. If $\beta = \{v_1, v_2, \dots, v_n\}$ is a basis for V , then $R(T) = \text{span}(T(\beta)) = \text{span}(\{T(v_1), T(v_2), \dots, T(v_n)\})$.

Definition 2.4 Let V and W be vector spaces and $T : V \rightarrow W$ be linear. If $N(T)$ and $R(T)$ are finite-dimensional, then we define the nullity of T , denoted $\text{nullity}(T)$, and the rank of T , denoted $\text{rank}(T)$, to be the dimensions of $N(T)$ and $R(T)$, respectively.

Theorem 2.4 (Dimension Theorem/Rank Nullity Theorem) Let V and W be vector spaces and $T : V \rightarrow W$ be linear. If V is finite-dimensional, then

$$\text{nullity}(T) + \text{rank}(T) = \dim(V)$$

Definition 2.5 A function $f : S \rightarrow T$ is said to be one-to-one (or injective) if $f(x) = f(y)$ implies that $x = y$. Such a function is said to be an injection.

Definition 2.6 A function $f : S \rightarrow T$ is said to be onto (or surjective) if its range is equal to T . $T = \{y \in T | \exists x \in S : f(x) = y\}$. Such a function is said to be a surjection.

Theorem 2.5 Let V and W be vector spaces and $T : V \rightarrow W$ be linear. T is one-to-one if and only if $N(T) = \{0\}$.

Theorem 2.6 Let V and W be vector spaces of equal (finite) dimension, and let $T : V \rightarrow W$ be linear. The following statements are equivalent (all true or all false):

1. T is one-to-one.
2. T is onto.
3. $\text{rank}(T) = \dim(V)$.

Theorem 2.7 Let V and W be vector spaces over F and suppose that $\{v_1, v_2, \dots, v_n\}$ is a basis for V . For vectors w_1, w_2, \dots, w_n in W , there exists exactly one linear transformation $T : V \rightarrow W$ such that $T(v_i) = w_i$ for $i = 1, 2, \dots, n$.

Theorem 2.8 (Transformations that agree on a basis are equivalent) Let V and W be vector spaces over F and suppose that $\{v_1, v_2, \dots, v_n\}$ is a basis for V . If $U, T : V \rightarrow W$ are both linear, and $U(v_i) = T(v_i)$ for $i = 1, 2, \dots, n$, then $U = T$.

Example 2.4 Let V and W be vector spaces and $T : V \rightarrow W$ be linear.

1. T is one-to-one if and only if T carries linearly independent subsets of V onto linear independent subsets of W .
2. Suppose that T is one-to-one and that S is a subset of V . Then S is linearly independent if and only if $T(S)$ is linearly independent.
3. Suppose β is a basis for V and T is one-to-one and onto. Then $T(\beta)$ is a basis for W .

Example 2.5 Let V and W be finite-dimensional vector spaces and $T : V \rightarrow W$ be linear.

1. If T is onto, then $\dim(V) \geq \dim(W)$.
2. If T is one-to-one, then $\dim(V) \leq \dim(W)$.

Example 2.6 Let V and W be vector spaces with subspaces V_1 and W_1 , respectively. If $T : V \rightarrow W$ is linear, then $T(V_1)$ is a subspace of W and $\{x \in V : T(x) \in W_1\}$ is a subspace of V .

Definition 2.7 Let V be a vector space and W_1 and W_2 be subspaces of V such that $V = W_1 \oplus W_2$. A function $T : V \rightarrow V$ is called the projection on W_1 along W_2 if, for $x = x_1 + x_2$ with $x_1 \in W_1, x_2 \in W_2$, we have $T(x) = x_1$.

Example 2.7 Using the notation in the definition above, assume that $T : V \rightarrow V$ is the projection on W_1 along W_2 :

1. T is linear, and $W_1 = \{x \in V : T(x) = x\}$.
2. $W_1 = R(T)$ and $W_2 = N(T)$.

Definition 2.8 Let V be a vector space and let $T : V \rightarrow V$ be linear. A subspace W of V is said to be T -invariant if $T(x) \in W$ for all $x \in W$, i.e. $T(W) \subseteq W$. If W is T -invariant, we define the restriction of T on W to be the function $T_W : W \rightarrow W$ defined by $T_W(x) = T(x)$ for all $x \in W$.

Example 2.8 If W is a subspace of a vector space V , and $T : V \rightarrow V$ is linear, then the following are all true:

1. $\{0\}$, V , $R(T)$ and $N(T)$ are all T -invariant.
2. If W is T -invariant, then T_W is linear.
3. If T is the projection on W along some subspace W' , then W is T -invariant and $T_W = I_W$.
4. If $V = R(T) \oplus W$ and W is T -invariant, then $W \subseteq N(T)$. Furthermore, if V is finite-dimensional, then $W = N(T)$.
5. If W is T -invariant, then $N(T_W) = N(T) \cap W$ and $R(T_W) = T(W)$.

Example 2.9 Let V be finite-dimensional vector space and $T : V \rightarrow V$ be linear.

1. If $V = R(T) + N(T)$, then $V = R(T) \oplus N(T)$.
2. If $R(T) \cap N(T) = \{0\}$, then $V = R(T) \oplus N(T)$.

2.2 Matrix representation of a linear transformation

Definition 2.9 Let V be a finite-dimensional vector space. An ordered basis for V is a basis for V endowed with a specific order; that is, an ordered basis for V is a finite sequence of linearly independent vectors in V that generates V .

Definition 2.10 For the vector space F^n , we call $\{e_1, e_2, \dots, e_n\}$ the standard basis for F^n . Similarly, for the vector space $P_n(F)$, we call $\{1, x, \dots, x^n\}$ the standard ordered basis for $P_n(F)$. The vectors of the standard basis are called the standard vectors.

Definition 2.11 Let $\beta = \{u_1, u_2, \dots, u_n\}$ be an ordered basis for a finite-dimensional vector space V . For $x \in V$, let a_1, a_2, \dots, a_n be the unique scalars such that $x = \sum_{i=1}^n a_i u_i$. We define the coordinate vector of x relative to β , denoted $[x]_\beta$, by $[x]_\beta = (a_1, a_2, \dots, a_n)$.

Definition 2.12 Let V, W be finite-dimensional vector spaces and let $T : V \rightarrow W$ be linear. If $\beta = \{v_1, v_2, \dots, v_n\}$ and $\gamma = \{w_1, w_2, \dots, w_m\}$ are ordered bases for V, W respectively, then for each j , $1 \leq j \leq n$ there exist unique scalars $a_{ij} \in F$, $1 \leq i \leq m$, such that

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i, 1 \leq j \leq n$$

The $m \times n$ matrix A defined by $A_{ij} = a_{ij}$ is called the matrix representation of T in the ordered bases β and γ and write $A = [T]_\beta^\gamma$. If $V = W$ and $\beta = \gamma$, we write $A = [T]_\beta$.

Definition 2.13 Let $T, U : V \rightarrow W$ be arbitrary functions, where V, W are vector spaces over F , and let $a \in F$. Define $T + U : V \rightarrow W$ by $(T + U)(x) = T(x) + U(x)$ for all $x \in V$, and $aT : V \rightarrow W$ by $(aT)(x) = aT(x)$ for all $x \in V$.

Theorem 2.9 Let V and W be vector spaces over a field F , and let $T, U : V \rightarrow W$ be linear. Then:

1. For all $a \in F$, $aT + U$ is linear.
2. Using the operations of addition and scalar multiplication in the preceding definition, the collection of linear transformations from V to W is a vector space over F .

Definition 2.14 Let V, W be vector spaces over F . The vector space of all linear transformations from V to W is denoted as $\mathcal{L}(V, W)$. In the case that $V = W$, we simply write $\mathcal{L}(V)$ instead of $\mathcal{L}(V, W)$.

Theorem 2.10 Let V and W be finite-dimensional vector spaces with ordered bases β and γ respectively, and let $T, U : V \rightarrow W$ be linear transformations. Then:

1. $[T + U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$
2. $[aT]_\beta^\gamma = a[T]_\beta^\gamma$ for all $a \in F$

2.3 Composition of linear transformations and matrix multiplication

Theorem 2.11 Let V, W, Z be vector spaces over the same field F , and let $T : V \rightarrow W$ and $U : W \rightarrow Z$ be linear. Then $UT : V \rightarrow Z$ is linear.

Theorem 2.12 Let V be a vector space. Let $T, U_1, U_2 \in \mathcal{L}(V)$. Then:

1. $T(U_1 + U_2) = TU_1 + TU_2$ and $(U_1 + U_2)T = U_1T + U_2T$
2. $T(U_1U_2) = (TU_1)U_2$
3. $TI = IT = T$
4. $a(U_1U_2) = (aU_1)U_2 = U_1(aU_2)$ for all scalars $a \in F$.

Definition 2.15 Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. We define the product of A and B , denoted AB , to be the $m \times p$ matrix such that

$$(AB)_{ij} = \sum_{k=1}^n A_{ik}B_{kj}, 1 \leq i \leq m, 1 \leq j \leq p$$

Theorem 2.13 Let V, W, Z be finite-dimensional vector spaces over the same field F , with ordered bases α, β, γ respectively, and let $T : V \rightarrow W$ and $U : W \rightarrow Z$ be linear. Then $[UT]_\alpha^\gamma = [U]_\beta^\gamma[T]_\alpha^\beta$.

Theorem 2.14 Let V be a finite-dimensional vector space with an ordered basis β . Let $T, U \in \mathcal{L}(V)$. Then $[UT]_\beta = [U]_\beta[T]_\beta$.

Definition 2.16 We define the Kronecker delta δ_{ij} by $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. The $n \times n$ identity matrix I_n , is defined by $(I_n)_{ij} = \delta_{ij}$.

Theorem 2.15 Let A be an $m \times n$ matrix, B and C be $n \times p$ matrices, and D and E be $q \times m$ matrices, all with entries from a field F . Then:

1. $A(B + C) = AB + AC$ and $(D + E)A = DA + EA$.
2. $a(AB) = (aA)B = A(aB)$ for all scalars $a \in F$.
3. $I_m A = A = A I_n$
4. If V is an n -dimensional vector space with an ordered basis β , then $[I_V]_\beta = I_n$.

Theorem 2.16 Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. For each $1 \leq j \leq p$, let u_j and v_j denote the j th columns of AB and B , respectively. Then:

1. $u_j = Av_j$
2. $v_j = Be_j$ where e_j is the j th standard vector of F^p .

Theorem 2.17 Let V and W be finite-dimensional vector spaces having ordered bases β and γ respectively and let $T : V \rightarrow W$ be linear. Then, for each $u \in V$, we have

$$[T(u)]_\gamma = [T]_\beta^\gamma[u]_\beta$$

Definition 2.17 Let A be an $m \times n$ matrix with entries from a field F . We denote by L_A the mapping $L_A : F^n \rightarrow F^m$ denoted by $L_A(x) = Ax$ for each column vector $x \in F^n$. We call L_A a left-multiplication transformation.

Theorem 2.18 Let A be an $m \times n$ matrix with entries from a field F . Then the left-multiplication transformation $L_A : F^n \rightarrow F^m$ is linear. Furthermore, if B is any other $m \times n$ matrix (with entries from F) and β and γ are the standard ordered bases for F^n and F^m respectively, then we have the following:

1. $[L_A]_{\beta}^{\gamma} = A$
2. $L_A = L_B$ if and only if $A = B$.
3. $L_{A+B} = L_A + L_B$ and $L_{aA} = aL_A$ for all $a \in F$.
4. If $T : F^n \rightarrow F^m$ is linear, then there exists a unique $m \times n$ matrix C such that $T = L_C$.
In fact, $C = [T]_{\beta}^{\gamma}$.
5. If E is an $n \times p$ matrix, then $L_{AE} = L_A L_E$.
6. If $m = n$, then $L_{I_n} = I_{F^n}$.

Theorem 2.19 Let A, B, C be matrices such that $A(BC)$ is defined. Then $(AB)C$ is also defined and $A(BC) = (AB)C$. That is, matrix multiplication is associative.

Example 2.10 Let M and A be matrices for which the product MA is defined. If the j th column of A is a linear combination of a set of columns of A , then the j th column of MA is a linear combination of the corresponding columns of MA with the same corresponding coefficients.

2.4 Invertibility and isomorphisms

Definition 2.18 Let V and W be vector spaces, and let $T : V \rightarrow W$ be linear. A function $U : V \rightarrow W$ is said to be an inverse of T if $TU = I_W$ and $UT = I_V$. If T has an inverse, then T is said to be invertible, and its inverse is denoted by T^{-1} .

Theorem 2.20 Let V and W be vector spaces, and let $T : V \rightarrow W$ be linear. If T is invertible, its inverse is unique.

Theorem 2.21 Let V and W be vector spaces, and let $T : V \rightarrow W$ and $U : V \rightarrow W$ be linear. If T and U are invertible, then:

1. $(TU)^{-1} = U^{-1}T^{-1}$
2. T^{-1} is also invertible, and in particular, $(T^{-1})^{-1} = T$.

Theorem 2.22 Let V and W be vector spaces, and let $T : V \rightarrow W$ be linear. Then T is invertible if and only if it is both one-to-one and onto. Also, T is invertible if and only if $\text{rank}(T) = \dim(V)$.

Theorem 2.23 Let V and W be vector spaces, and let $T : V \rightarrow W$ be linear and invertible. Then T^{-1} is linear.

Definition 2.19 Let A be an $n \times n$ matrix. Then A is invertible if there exists an $n \times n$ matrix B such that $AB = BA = I$.

Theorem 2.24 Let V and W be vector spaces, and let $T : V \rightarrow W$ be linear and invertible. Then V is finite-dimensional if and only if W is finite-dimensional. In this case, $\dim(V) = \dim(W)$.

Theorem 2.25 Let V and W be finite-dimensional vector spaces with ordered bases β and γ , respectively. Let $T : V \rightarrow W$ be linear. Then T is invertible if and only if $[T]_{\beta}^{\gamma}$ is invertible. Furthermore, $[T^{-1}]_{\gamma}^{\beta} = ([T]_{\beta}^{\gamma})^{-1}$.

Theorem 2.26 Let V be a finite-dimensional vector space with an ordered basis β . Let $T : V \rightarrow V$ be linear. Then T is invertible if and only if $[T]_\beta$ is invertible. Furthermore, $[T^{-1}]^\beta = ([T]^\beta)^{-1}$.

Theorem 2.27 Let A be an $n \times n$ matrix. Then A is invertible if and only if L_A is invertible. Furthermore, $(L_A)^{-1} = L_{A^{-1}}$.

Definition 2.20 Let V and W be vector spaces. We say that V is isomorphic to W if there exists a linear transformation $T : V \rightarrow W$ that is invertible. Such a linear transformation is called an isomorphism from V onto W .

Theorem 2.28 Let V and W be finite-dimensional vector spaces (over the same field). Then V is isomorphic to W if and only if $\dim(V) = \dim(W)$.

Theorem 2.29 Let V be a vector space over a field F . Then V is isomorphic to F^n if and only if $\dim(V) = n$.

Theorem 2.30 Let V and W be finite-dimensional vector spaces over F of dimensions n and m , respectively, and let β and γ be ordered bases for V and W , respectively. Then the function $\phi : \mathcal{L}(V, W) \rightarrow M_{m \times n}(F)$ defined by $\phi(T) = [T]_\beta^\gamma$ is an isomorphism.

Theorem 2.31 Let V and W be finite-dimensional vector spaces over F of dimensions n and m , respectively. Then $\mathcal{L}(V, W)$ is finite-dimensional of dimension mn .

Definition 2.21 Let β be an ordered basis for an n -dimensional vector space V over the field F . The standard representation of V with respect to β is the function $\phi_\beta : V \rightarrow F^n$ defined by $\phi_\beta(x) = [x]_\beta$ for each $x \in V$.

Theorem 2.32 For any finite-dimensional vector space V with ordered basis β , ϕ_β is an isomorphism.

Example 2.11 Let V and W be finite-dimensional vector spaces, and let $T : V \rightarrow W$ be an isomorphism. Let V_0 be a subspace of V . Then:

1. $T(V_0)$ is a subspace of W .
2. $\dim(V_0) = \dim(T(V_0))$.

2.5 The change of coordinate matrix

Theorem 2.33 Let β and β' be two ordered bases for a finite-dimensional vector space V , and let $Q = [I_V]_{\beta'}^\beta$. Then:

1. Q is invertible.
2. For any $v \in V$, $[v]_\beta = Q[v]_{\beta'}$.

Definition 2.22 Using the notation above, the matrix $Q = [I_V]_{\beta'}^\beta$ is called the change of coordinate matrix that changes β' coordinates into β coordinates.