Linear Algebra Notes

Wuwei Yuan

October 27, 2020

Contents

1 Vector Spaces

1.1 Introduction

Parallelogram Law for Vector Addition. *The sum of two vectors x and y that act at the same point P is the vector beginning at P that is represented by the diagonal of parallelogram having x and y as adjacent sides.*

1.2 Vector Spaces

Definition (Fieid). *Formally, a field is a set F together with two binary operations on F called addition and multiplication. A binary operation on F is a mapping $F \times F \to F$, that is, a correspondence that associates with each ordered pair of elements of F a uniquely determined element of F. The result of the addition of a and b is called the sum of a and b, and is denoted $a + b$. Similarly, the result of the multiplication of a and b is called the product of a and b, and is denoted ab or $a \cdot b$. These operations are required to satisfy the following properties, referred to as field axioms. In these axioms, a, b, and c are arbitrary elements of the field F.*

- *Associativity of addition and multiplication: $a + (b + c) = (a + b) + c$, and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.*

- *Commutativity of addition and multiplication: $a + b = b + a$, and $a \cdot b = b \cdot a$.*

- *Additive and multiplicative identity: there exist two different elements 0 and 1 in F such that $a + 0 = a$ and $a \cdot 1 = a$.*

- *Additive inverses: for every a in F, there exists an element in F, denoted $-a$, called the additive inverse of a, such that $a + (-a) = 0$.*

- *Multiplicative inverses: for every $a \neq 0$ in F, there exists an element in F, denoted by a^{-1} or $1/a$, called the multiplicative inverse of a, such that $a \cdot a^{-1} = 1$.*

- *Distributivity of multiplication over addition: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.*

Definitions. *A vector space (or linear space) V over a field 2 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.*

(VS 1) *For all x, y in V, $x + y = y + x$ (commutativity of addition).*

(VS 2) *For all x, y, z in V, $(x + y) + z = x + (y + z)$ (associativity of addition).*

(VS 3) *There exists an element in V denoted by 0 such that $x + 0 = x$ for each x in V.*

(VS 4) For each element x in V there exists an element y in V such that $x + y = 0$

(VS 5) For each element x in V, $1x = x$.

(VS 6) For each pair of elements a, b in F and each element x in V, $(ab)x = a(bx)$.

(VS 7) For each element a in F and each pair of elements x, y in V, $a(x + y) = ax + ay$.

(VS 8) For each pair of elements a, b in F and each element x in V, $(a + b)x = ax + bx$.

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.

Theorem 1.1 (Cancellation Law for Vector Addition). *If x, y, and z are vectors in a vector space V such that $x + z = y + z$, then $x = y$.*

Corollary 1. *The vector 0 described in (VS 3) is unique.*

Corollary 2. *The vector y described in (VS 4) is unique.*

The vector 0 in (VS 3) is called the zero vector of V, and the vector y in (VS 4) (that is, the unique vector such that $x+y = 0$) is called the additive inverse of x and is denoted by $-x$.

Theorem 1.2. *In any vector space V, the following statements are true:*

(a) *$0x = 0$ for each $x \in V$.*

(b) *$(-a)x = -(ax) = a(-x)$ for each $a \in F$ and each $x \in V$.*

(c) *$a0 = 0$ for each $a \in F$.*

1.3   Subspaces

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

In any vector space V, note that V and {0} are subspaces. The latter is called the zero subspace of V.

Theorem 1.3. *Let V be a vector space and W a subset of V. Then W is a subspace of V if and only if the following three conditions hold for the operations defined in V.*

(a) $0 \in W$.

(b) $x + y \in W$ *whenever* $x \in W$ *and* $y \in W$.

(c) $cx \in W$ *whenever* $c \in F$ *and* $x \in W$.

The **transpose** A^t of an $m \times n$ matrix A is the $n \times m$ matrix obtained from A by interchanging the rows with the columns; that is, $(A^t)_{ij} = A_{ji}$.

Theorem 1.4. *Any intersection of subspaces of a vector space V is a subspace of V.*

Definition. *If S1 and S2 are nonempty subsets of a vector space V, then the sum of S1 and S2, denoted S1 +S2, is the set* $\{x + y : x \in S1 \text{ and } y \in S2\}$.

Definition. *A vector space V is called the direct sum of W1 and W2 if W1 and W2 are subspaces of V such that* $W_1 \cap W_2 = \{0\}$ *and* $W_1 + W_2 = V$. *We denote that V is the direct sum of W1 and W2 by writing* $V = W_1 \oplus W_2$.

Exercise. *W1 +W2 is a subspace of V that contains both W1 and W2.*

Exercise. *Any subspace of V that contains both W1 and W2 must also contain W1 + W2.*

Definitions. *Let W be a subspace of a vector space V over a field F. For any* $v \in V$ *the set* $\{v\} + W = \{v + w : w \in W\}$ *is called the* **coset** *of W* **containing** *v. It is customary to denote this coset by* $v + W$ *rather than* $\{v\} + W$.

Addition and scalar multiplication by scalars of F can be defined in the collection $S = \{v + W : v \in V\}$ *of all cosets of W as follows:*

$$(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W \tag{1}$$

for all $v_1, v_2 \in V$ *and*

$$a(v + W) = av + W \tag{2}$$

for all $v \in V$ *and* $a \in F$.

The set S is a vector space with the preceding operations. This vector space is called the **quotient space of** *V* **modulo** *W and is denoted by V/W.*

1.4   Linear Combinations and Systems of Linear Equations

Definition. *Let V be a vector space and S a nonempty subset of V. A vector v V is called a linear combination of vectors of S if there exist a finite number of vectors u1, u2,...,un in S and scalars a1, a2,...,an in F such that* $v = a1u1 + a2u2 + \cdots + anun$. *In this case we also say that v is a linear combination of u1, u2,...,un and call a1, a2,...,an the coefficients of the linear combination.*

Definition. *Let S be a nonempty subset of a vector space V. The span of S, denoted span(S), is the set consisting of all linear combinations of the vectors in S. For convenience, we define* $span(\varnothing) = \{0\}$.

Theorem 1.5. *The span of any subset S of a vector space V is a subspace of V. Moreover, any subspace of V that contains S must also contain the span of S.*

Definition. *A subset S of a vector space V generates (or spans) V if span(S) = V. In this case, we also say that the vectors of S generate (or span) V.*

1.5 Linear Dependence and Linear Independence

Definition. *A subset S of a vector space V is called linearly dependent if there exist a finite number of distinct vectors u1, u2,...,un in S and scalars a1, a2,...,an, not all zero, such that*

$$a1u1 + a2u2 + \cdots + anun = 0.$$

In this case we also say that the vectors of S are linearly dependent.

For any vectors u1, u2,...,un, we have a1u1 + a2u2 + $\cdots$ + anun = 0 if a1 = a2 = $\cdots$ = an = 0. We call this the trivial representation of 0 as a linear combination of u1, u2,...,un.

Definition. *A subset S of a vector space that is not linearly dependent is called linearly independent. As before, we also say that the vectors of S are linearly independent.*

The following facts about linearly independent sets are true in any vector space.

1. The empty set is linearly independent, for linearly dependent sets must be nonempty.

2. A set consisting of a single nonzero vector is linearly independent. For if u is linearly dependent, then au = 0 for some nonzero scalar a. Thus

$$u = a^{-1}(au) = a^{-1}0 = 0.$$

3. A set is linearly independent if and only if the only representations of 0 as linear combinations of its vectors are trivial representations.

Theorem 1.6. *Let V be a vector space, and let S1 $\subseteq$ S2 $\subseteq$ V. If S1 is linearly dependent, then S2 is linearly dependent.*

Corollary. *Let V be a vector space, and let S1 $\subseteq$ S2 $\subseteq$ V. If S2 is linearly independent, then S1 is linearly independent.*

Theorem 1.7. *Let S be a linearly independent subset of a vector space V, and let v be a vector in V that is not in S. Then S $\cup$ {v} is linearly dependent if and only if v $\in$ span(S).*

1.6 Base and Dimension

Definition. *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.*

Theorem 1.8. *Let V be a vector space and $\beta = \{u_1, u_2, \ldots, u_n\}$ 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 of β, that is, can be expressed in the form*

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$$

for unique scalars $a_1, a_2, \ldots, a_n$.

Theorem 1.9. *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 1.10 (Replacement Theorem)**.** *Let V be a vector space that is generated by a set G containing exactly n vectors, and let L be a linearly 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.*

Corollary 1. *Let V be a vector space having a finite basis. Then every basis for V contains the same number of vectors.*

Definitions. *A vector space is called **finite-dimensional** if it has a basis consisting of a finite number of vectors. The unique number of vectors in each basis for V is called the **dimension** of V and is denoted by $dim(V)$. A vector space that is not finite-dimensional is called **infinite-dimensional**.*

Corollary 2. *Let V be a vector space with dimension n.*

(a) *Any finite generating set for V contains at least n vectors, and a generating set for V that contains exactly n vectors is a basis for V.*

(b) *Any linearly independent subset of V that contains exactly n vectors is a basis for V.*

(c) *Every linearly independent subset of V can be extended to a basis for V.*

Theorem 1.11. *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 $V = W$.*

Corollary. *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.*

1.7 Maximal Linearly Independent Subsets*

Definition. *Let F be a family of sets. A member M of F is called maximal (with respect to set inclusion) if M is contained in no member of F other than M itself.*

Definition. *A collection of sets C is called a chain (or nest or tower) if for each pair of sets A and B in C, either A   B or B   A.*

Maximal Principle. *Let F be a family of sets. If, for each chain C F, there exists a member of F that contains each member of C, then F contains a maximal member.*

Definition. *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.*

  (a) B is linearly independent.

  (b) The only linearly independent subset of S that contains B is B itself

A basis   for a vector space V is a maximal linearly independent subset of V, because 1.   is linearly independent by definition. 2. If v   V and v /   , then   v is linearly dependent by Theorem 1.7 (p. 39) because span( ) = V.

Theorem 1.12. *Let V be a vector space and S a subset that generates V. If   is a maximal linearly independent subset of S, then   is a basis for V.*

Theorem 1.13. *Let S be a linearly independent subset of a vector space V. There exists a maximal linearly independent subset of V that contains S.*

Corollary. *Every vector space has a basis.*

It can be shown, analogously to Corollary 1 of the replacement theorem (p. 46), that every basis for an infinite-dimensional vector space has the same cardinality.
Also there are many important exercise. See textbook P62.

2 Linear Transformations and Matrices

2.1 Linear Transformations, Null Spaces, and Ranges

Definition. *Let V and W be vector spaces (over F). We call a function T: V → W a linear transformation from V to W if, for all $x, y \in V$ and $c \in F$, we have*
  (a) T(x + y) = T(x) + T(y) and
  (b) T(cx) = cT(x).

We often simply call T **linear**. The reader should verify the following properties of a function T: V → W. (See Exercise 7.)

  1. If T is linear, then 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   V.

4. T is linear if and only if, for x1, x2,... ,xn   V and a1, a2,... ,an   F, we have

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

For vector spaces V and W (over F), we define the **identity transformation** $I_V : V \to V$ by $I_V(x) = x$ for all $x \in V$ and the **zero transformation** $T_0 : V \to W by T_0(x) = 0$ for all $x \to V$. It is clear that both of these transformations are linear. We often write I instead of I_V.

Definitions. *Let V and W be vector spaces, and let T: V → W be linear. We define the **null space** (or **kernel**) N(T) of T to be the set of all vectors x in V such that T(x) = 0 ; that is, $N(T) = \{x \in V : T(x) = 0\}$.*
*We define the **range** (or **image**) R(T) of T to be the subset of W consisting of all images (under T) of vectors in V; that is, $R(T) = \{T(x) : x \in V\}$.*

Theorem 2.1. *Let V and W be vector spaces and T: V → W be linear. Then N(T) and R(T) are subspaces of V and W, respectively.*

Theorem 2.2. *Let V and W be vector spaces, and let T: V → W be linear. If $\beta = \{v1, v2, ..., vn\}$ is a basis for V, then $R(T) = span(T(\beta)) = span(\{T(v_1), T(v_2), ..., T(v_n)\})$*

It should be noted that Theorem 2.2 is true if β is infinite, that is, $R(T) = span(\{T(v) : v \in \beta\})$.

Definitions. *Let V and W be vector spaces, and let $T : V \to W$ be linear. If N(T) and R(T) are finite-dimensional, then we define the **nullity** of T, denoted nullity(T), and the **rank** of T, denoted rank(T), to be the dimensions of N(T) and R(T), respectively.*

Theorem 2.3. *Dimension Theorem Let V and W be vector spaces, and let T: V → W be linear. If V is finite-dimensional, then nullity(T) + rank(T) = dim(V).*

Theorem 2.4. *Let V and W be vector spaces, and let T: V → W be linear. Then T is one-to-one if and only if N(T) = 0 .*

Theorem 2.5. *Let V and W be vector spaces of equal (finite) dimension, and let T: V → W be linear. Then the following are equivalent.*
(a) T is one-to-one.
(b) T is onto.
(c) rank(T) = dim(V).

Theorem 2.6. *Let V and W be vector spaces over F, and suppose that $\{v_1, v_2, \ldots, v_n\}$ is a basis for V. For $w_1, w_2, \ldots, w_n$ in W, there exists exactly one linear transformation $T : V \to W$ such that $T(v_i) = w_i$ for $i = 1, 2, \ldots, n$.*

Corollary. *Let V and W be vector spaces, and suppose that V has a finite basis $\{v_1, v_2, \ldots, v_n\}$. If U,T: $V \to W$ are linear and $U(v_i) = T(v_i)$ for i = 1, 2,...,n, then U = T.*

Definition. *Let V be a vector space and W1 and W2 be subspaces of V such that $V = W1 \quad W2$. (Recall the definition of direct sum given in the exercises of Section 1.3.) A function T: $V \to V$ is called the projection on W1 along W2 if, for $x = x1 + x2$ with x1 W1 and x2 W2, we have $T(x) = x1$.*

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

2.2   The Matrix Representationof a Linear Transformation

Definition. *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.*

For the vector space F^n, we call $\{e_1, e_2, \ldots, e_n\}$ the **standard ordered basis** for F^n. Similarly, for the vector space $P_n(F)$, we call $\{1, x, \ldots, x_n\}$ the **standard ordered basis** for $P_n(F)$.

Definition. *Let $= \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for a finite dimensional vector space V. For $x \in V$, let $a_1, a_2, \ldots, a_n$ 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*

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

Exercise. *The correspondence $x \to [x]_\beta$ provides us with a linear transformation from V to F^n..*

Definition. *Suppose that V and W are finite-dimensional vector spaces with ordered bases $\beta = \{v_1, v_2, \beta, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$, respectively. Let*

$T: V \to W$ be linear. Then for each j, $1 \leq j \leq n$, there exist unique scalars $a_{ij} \in F$, $1 \leq i \leq m$, such that

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

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

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

Theorem 2.7. *Let V and W be vector spaces over a field F, and let $T, U: V \to W$ be linear.*

(a) For all $a \in F$, $aT + U$ is linear.

(b) Using the operations of addition and scalar multiplication in the preceding definition, the collection of all linear transformations from V to W is a vector space over F.

Definitions. *Let V and W be vector spaces over F. We denote the vector space of all linear transformations from V into W by $\mathcal{L}(V, W)$. In the case that $V = W$, we write $\mathcal{L}(V)$ instead of $\mathcal{L}(V, W)$.*

Theorem 2.8. *Let V and W be finite-dimensional vector spaces with ordered bases β and γ, respectively, and let $T, U: V \to W$ be linear transformations. Then*

(a) $[T + U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$ and

(b) $[aT]_\beta^\gamma = a[T]_\beta^\gamma$ for all scalars a.

2.3 Composition of Linear Transformations and Matrix Multiplication

Definition. *Let A, B, and C be sets and $f : A \to B$ and $g : B \to C$ be functions. By following f with g, we obtain a function $g \circ f : A \to C$ called the **composite** of g and f. Thus $(g \circ f)(x) = g(f(x))$ for all $x \in A$.*

Theorem 2.9. *Let V, W, and Z be vector spaces over the same field F, and let $T: V \to W$ and $U: W \to Z$ be linear. Then $UT: V \to Z$ is linear.*

Theorem 2.10. *Let V be a vector space. Let $T, U1, U2 \in L(V)$. Then*

(a) $T(U_1 + U_2) = TU_1 + TU_2$ and $(U_1 + U_2)T = U_1T + U_2T$

(b) $T(U_1 U_2) = (TU_1)U_2$ (c) $TI = IT = T$ (d) $a(U1U2)=(aU1)U2 = U1(aU2)$ for all scalars a

Definition. *Let A be an m × n matrix and B be an n × p matrix. We define the product of A and B, denoted AB, to be the m × p matrix such that*

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

Theorem 2.11. *Let V, W, and Z be finite-dimensional vector spaces with ordered bases α, β, and γ, respectively. Let T: V → W and U: W → Z be linear transformations. Then*

$$[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$$

Corollary. *Let V be a finite-dimensional vector space with an ordered basis β. Let T, U ∈ L(V). Then $[UT]_\beta = [U]_\beta [T]_\beta$.*

Definitions. *We define the **Kronecker delta** δ_{ij} by $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \ne j$. The n × n **identity matrix** I_n is defined by $(I_n)_{ij} = \delta_{ij}$.*

Theorem 2.12. *Let A be an m × n matrix, B and C be n × p matrices, and D and E be q × m matrices. Then*

(a) *A(B + C) = AB + AC and (D + E)A = DA + EA.*

(b) *a(AB)=(aA)B = A(aB) for any scalar a.*

(c) *$I_m A = A = A I_n$.*

(d) *If V is an n-dimensional vector space with an ordered basis β, then $[I_V]_\beta = I_n$.*

Corollary. *Let A be an m × n matrix, B1, B2,...,Bk be n × p matrices, C1, C2,...,Ck be q × m matrices, and a1, a2,...,ak be scalars. Then*

$$A \left(\sum_{i=1}^{k} a_i B_i \right) = \sum i = 1^k a_i A B_i$$

and

$$\left(\sum_{i=1}^{k} a_i C_i \right) A = \sum i = 1^n a_i C_i A$$

Theorem 2.13. *Let A be an m × n matrix and B be an n × p matrix. For each j (1 ≤ j ≤ p) let uj and vj denote the jth columns of AB and B, respectively. Then*

(a) *$u_j = A v_j$*
(b) *$v_j = B e_j$, where e_j is the jth standard vector of F^p.*

It follows (see Exercise 14) from Theorem 2.13 that column j of AB is a linear combination of the columns of A with the coefficients in the linear combination being the entries of column j of B. An analogous result holds for rows; that is, row i of AB is a linear combination of the rows of B with the coefficients in the linear combination being the entries of row i of A.

11

Theorem 2.14. *Let V and W be finite-dimensional vector spaces having ordered bases β and γ, respectively, and let T: $V \to W$ be linear. Then, for each $u \in V$, we have*

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

Definition. *Let A be an $m \times n$ matrix with entries from a field F. We denote by L_A the mapping $L_A : F_n \to F_m$ defined by $L_A(x) = Ax$ (the matrix product of A and x) for each column vector $x \in F_n$. We call L_A a **left-multiplication transformation**.*

Theorem 2.15. *Let A be an $m \times n$ matrix with entries from F. Then the left-multiplication transformation $L_A : F_n \to F_m$ 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 properties.*

(a) $[L_A]_\beta^\gamma = A$.

(b) $L_A = L_B$ *if and only if $A = B$.*

(c) $L_{A+B} = L_A + L_B$ *and $L_{aA} = aL_A$ for all $a \in F$.*

(d) *If $T : F^n \to F^m$ is linear, then there exists a unique $m \times n$ matrix C such that $T = L_C$. In fact, $C = [T]_\beta^\gamma$*

(e) *If E is an $n \times p$ matrix, then $L_{AE} = L_A L_E$.*

(f) *If $m = n$, then $L_{I_n} = I_{F_n}$.*

Theorem 2.16. *Let A, B, and 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.*

2.4   Invertibility and Isomorphisms

Definition. *Let V and W be vector spaces, and let T: $V \to W$ be linear. A function U: $W \to V$ 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**. As noted in Appendix B, if T is invertible, then the inverse of T is unique and is denoted by T^{-1}.*

The following facts hold for invertible functions T and U.
1. $(TU)^{-1} = U^{-1}T^{-1}$.
2. $(T^{-1})^{-1} = T$; in particular, T^{-1} is invertible.
We often use the fact that a function is invertible if and only if it is both one-to-one and onto. We can therefore restate Theorem 2.5 as follows.
3. Let T: $V \to W$ be a linear transformation, where V and W are finitedimensional spaces of equal dimension. Then T is invertible if and only if rank(T) = dim(V).

Theorem 2.17. *Let V and W be vector spaces, and let T: $V \to W$ be linear and invertible. Then $T^{-1} : W \to V$ is linear.*

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

The matrix B is called the **inverse** of A and is denoted by A^{-1}.

Lemma. *Let T be an invertible linear transformation from V to W. Then V is finite-dimensional if and only if W is finite-dimensional. In this case, $dim(V) = dim(W)$.*

Theorem 2.18. *Let V and W be finite-dimensional vector spaces with ordered bases β and γ, respectively. Let T: $V \to W$ be linear. Then T is invertible if and only if $[T]_\beta^\gamma$ is invertible. Furthermore, $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$.*

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

Corollary 2. *Let A be an $n \times n$ matrix. Then A is invertible if and only if L_A is invertible. Furthermore, $(L_A)^{-1} = L_{A^{-1}}$.*

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

Theorem 2.19. *Let V and W be finite-dimensional vector spaces (over the same field). Then V is isomorphic to W if and only if $dim(V) = dim(W)$.*

Corollary. *Let V be a vector space over F. Then V is isomorphic to F_n if and only if $dim(V) = n$.*

Theorem 2.20. *Let V and W be finite-dimensional vector spaces over F of dimensions n and m, respectively, and let   and   be ordered bases for V and W, respectively. Then the function $\Phi : \mathcal{L}(V, W) \to M_{m \times n}(F)$, defined by $\Phi(T) = [T]_\beta^\gamma$ for $T \in L(V, W)$, is an isomorphism.*

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

Definition. *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 $\phi_\beta : V \to F^n$ defined by $\phi_\beta(x) = [x]_\beta$ for each $x\beta V$.*

Theorem 2.21. *For any finite-dimensional vector space V with ordered basis β, ϕ_β is an isomorphism.*

2.5   The Change of Coordinate Matrix

Theorem 2.22. *Let β and β' be two ordered bases for a finite-dimensional vector space V, and let $Q = [I_V]_{\beta'}^\beta$. Then*
(a) Q is invertible.
(b) For any $v \in V$, $[v]_\beta = Q[v]_{\beta'}$.

The matrix $Q = [I_V]_{\beta'}^{\beta}$ defined in Theorem 2.22 is called a **change of co-ordinate matrix**. Because of part (b) of the theorem, we say that Q **changes β'-coordinates into β-coordinates**. Observe that if $\beta = \{x_1, x_2, \ldots, x_n\}$ and $\beta' = \{x_1', x_2', \ldots, x_n'\}$ then

$$x_j' = \sum_{i=1}^{n} Q_{ij} x_i$$

for j = 1, 2,... ,n; that is, the jth column of Q is $[x_j']_\beta$.

Notice that if Q changes β'-coordinates into β-coordinates, then Q^{-1} changes β-coordinates into β'-coordinates.

For the remainder of this section, we consider only linear transformations that map a vector space V into itself. Such a linear transformation is called a **linear operator** on V.

Theorem 2.23. *Let T be a linear operator on a finite-dimensional vector space V, and let β and β' be ordered bases for V. Suppose that Q is the change of coordinate matrix that changes β'-coordinates into β-coordinates. Then*

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q$$

Corollary. *Let $A \in M_{n \times n}(F)$, and let γ be an ordered basis for F^n. Then $[L_A]\gamma = Q^{-1}AQ$, where Q is the $n \times n$ matrix whose jth column is the jth vector of γ.*

Definition. *Let A and B be matrices in $M_{n \times n}(F)$. We say that B is **similar** to A if there exists an invertible matrix Q such that $B = Q^{-1}AQ$.*

2.6 Dual Spaces*

In this section, we are concerned exclusively with linear transformations from a vector space V into its field of scalars F, which is itself a vector space of dimension 1 over F. Such a linear transformation is called a **linear functional** on V. We generally use the letters f, g, h,... to denote linear functionals.

Let V be a finite-dimensional vector space, and let $\beta = \{x_1, x_2, \ldots, x_n\}$ be an ordered basis for V. For each i = 1, 2,...,n, define $f_i(x) = a_i$, where

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

is the coordinate vector of x relative to β. Then f_i is a linear functional on V called the *i***th coordinate function with respect to the basis** β. Note that $f_i(x_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta.

Definition. *For a vector space V over F, we define the **dual space** of V to be the vector space $\mathcal{L}(V, F)$, denoted by V^*.*

Hence by Theorem 2.19, V and V^* are isomorphic. We also define the **double dual** V^{**} of V to be the dual of V^*.

Theorem 2.24. *Suppose that V is a finite-dimensional vector space with the ordered basis $\beta = \{x_1, x_2, \ldots, x_n\}$. Let $f_i(1 \leq i \leq n)$ be the ith coordinate function with respect to β as just defined, and let $\beta^* = \{f_1, f_2, \ldots, f_n\}$. Then β^* is an ordered basis for V^*, and, for any $f \in V^*$, we have*

$$f = \sum_{i=1}^{n} f(x_i) f_i$$

Definition. *Using the notation of Theorem 2.24, we call the ordered basis $\beta^* = \{f_1, f_2, \ldots, f_n\}$ of V^* that satisfies $f_i(x_j) = \delta_{ij}(1 \leq i, j \leq n)$ the **dual basis** of β.*

Theorem 2.25. *Let V and W be finite-dimensional vector spaces over F with ordered bases β and γ, respectively. For any linear transformation $T: V \to W$, the mapping $T^t : W^* \to V^*$ defined by $T^t(g) = gT$ for all $g \in W^*$ is a linear transformation with the property that $[T^t]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$.*

The linear transformation T^t defined in Theorem 2.25 is called the **transpose** of T. It is clear that T^t is the unique linear transformation U such that $[U]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$.

For a vector $x \in V$, we define $\hat{x} : V^* \to F$ by $\hat{x}(f) = f(x)$ for every $f \in V^*$. It is easy to verify that $\hat{x}$ is a linear functional on V^*, so $\hat{x} \in V^{**}$.

Lemma. *Let V be a finite-dimensional vector space, and let $x \in V$. If $\hat{x}(f) = 0$ for all $f \in V^*$, then $x = 0$.*

Theorem 2.26. *Let V be a finite-dimensional vector space, and define $\Psi : V \to V^{**}$ by $\Psi(x) = \hat{x}$. Then Ψ is an isomorphism.*

Corollary. *Let V be a finite-dimensional vector space with dual space V^*. Then every ordered basis for V^* is the dual basis for some basis for V.*

Definition. *V denotes a finite-dimensional vector space over F. For every subset S of V, define the **annihilator** S^0 of S as*

$$S^0 = \{f \in V^* : f(x) = 0 \text{ for all } x \in S\}$$

2.7 Homogeneous Linear Differential Equations with Constant Coefficients*

Gugugu.

3 Elementary Matrix Operations and Systems of Linear Equations

3.1 Elementary Matrix Operations and Elementary Matrices

Definitions. *Let A be an m × n matrix. Any one of the following three operations on the rows [columns] of A is called an **elementary row [column] operation**:*

(1) interchanging any two rows [columns] of A;

(2) multiplying any row [column] of A by a nonzero scalar;

(3) adding any scalar multiple of a row [column] of A to another row [column].

Any of these three operations is called an **elementary operation**. Elementary operations are of **type 1**, **type 2**, or **type 3** depending on whether they are obtained by (1), (2), or (3).

Definition. *An n × n **elementary matrix** is a matrix obtained by performing an elementary operation on I_n. The elementary matrix is said to be of **type 1**, **2**, or **3** according to whether the elementary operation performed on In is a type 1, 2, or 3 operation, respectively.*

Theorem 3.1. *Let $A \in M_{m \times n}(F)$, and suppose that B is obtained from A by performing an elementary row [column] operation. Then there exists an m × m [n × n] elementary matrix E such that B = EA [B = AE]. In fact, E is obtained from $I_m[I_n]$ by performing the same elementary row [column] operation as that which was performed on A to obtain B. Conversely, if E is an elementary m × m [n × n] matrix, then EA [AE] is the matrix obtained from A by performing the same elementary row [column] operation as that which produces E from Im [In].*

Theorem 3.2. *Elementary matrices are invertible, and the inverse of an elementary matrix is an elementary matrix of the same type.*

3.2 The Rank of a Matrix and Matrix inverses

Definition. *If $A \in M_{m \times n}(F)$, we define the **rank** of A, denoted rank(A), to be the rank of the linear transformation $L_A : F_n \to F_m$.*

Many results about the rank of a matrix follow immediately from the corresponding facts about a linear transformation. An important result of this type, which follows from Fact 3 (p. 100) and Corollary 2 to Theorem 2.18 (p. 102), is that an n × n matrix is invertible if and only if its rank is n.

Theorem 3.3. *Let $T : V \to W$ be a linear transformation between finitedimensional vector spaces, and let β and γ be ordered bases for V and W, respectively. Then $rank(T) = rank([T]_\beta^\gamma)$.*

Theorem 3.4. *Let A be an m × n matrix. If P and Q are invertible m × m and n × n matrices, respectively, then*

  (a) *rank(AQ) = rank(A),*

  (b) *rank(PA) = rank(A), and therefore,*

  (c) *rank(PAQ) = rank(A).*

Corollary. *Elementary row and column operations on a matrix are rankpreserving*

Theorem 3.5. *The rank of any matrix equals the maximum number of its linearly independent columns; that is, the rank of a matrix is the dimension of the subspace generated by its columns.*

Theorem 3.6. *Let A be an m × n matrix of rank r. Then $r \leq m, r \leq n$, and, by means of a finite number of elementary row and column operations, A can be transformed into the matrix*

$$D = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$$

, where O_1, O_2, and O_3 are zero matrices. Thus $D_{ii} = 1$ for $i \leq r$ and $D_{ij} = 0$ otherwise.

Corollary 1. *Let A be an m × n matrix of rank r. Then there exist invertible matrices B and C of sizes m × m and n×n, espectively, such that $D = BAC$, where*

$$D = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$$

is the m × n matrix in which O_1, O_2, and O_3 are zero matrices.

Corollary 2. *Let A be an m × n matrix. Then*

  (a) $rank(A^t) = rank(A)$.

  (b) *The rank of any matrix equals the maximum number of its linearly independent rows; that is, the rank of a matrix is the dimension of the subspace generated by its rows.*

  (c) *The rows and columns of any matrix generate subspaces of the same dimension, numerically equal to the rank of the matrix.*

Corollary 3. *Every invertible matrix is a product of elementary matrices.*

Theorem 3.7. *Let T: V → W and U: W → Z be linear transformations on finite-dimensional vector spaces V, W, and Z, and let A and B be matrices such that the product AB is defined. Then*

(a) $rank(UT) \leq rank(U)$.

(b) $rank(UT) \leq rank(T)$.

(c) $rank(AB) \leq rank(A)$.

(d) $rank(AB) \leq rank(B)$.

The Inverse of a Matrix

Definition. *Let A and B be m × n and m × p matrices, respectively. By the* ***augmented matrix*** *(A|B), we mean the m × (n + p) matrix (A B), that is, the matrix whose first n columns are the columns of A, and whose last p columns are the columns of B.*

Being able to test for invertibility and compute the inverse of a matrix allows us, with the help of Theorem 2.18 (p. 101) and its corollaries, to test for invertibility and compute the inverse of a linear transformation. The next example demonstrates this technique.

3.3 Systems of Linear Equations - Theoretical Aspects

The system of equations

$$(S) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

where a_{ij} and b_i $(1 \leq i \leq m$ and $1 \leq j \leq n)$ are scalars in a field F and $x_1, x_2, \ldots, x_n$ are n variables taking values in F, is called a **system of m linear equations in n unknowns over the field F**.

The $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{1m} \end{pmatrix}$$

is called the **coefficient matrix** of the system (S).

If we let

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad and \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

then the system (S) may be rewritten as a single matrix equation

$$Ax = b$$

A **solution** to the system (S) is an n-tuple

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in F^n$$

such that $As = b$. The set of all solutions to the system (S) is called the **solution set** of the system. System (S) is called **consistent** if its solution set is nonempty; otherwise it is called **inconsistent**.

We see that a system of linear equations can have one, many, or no solutions.

Definitions. *A system $Ax = b$ of m linear equations in n unknowns is said to be* ***homogeneous*** *if $b = 0$. Otherwise the system is said to be* ***nonhomogeneous***.

Theorem 3.8. *Let $Ax = 0$ be a homogeneous system of m linear equations in n unknowns over a field F. Let K denote the set of all solutions to $Ax = 0$. Then $K = N(L_A)$; hence K is a subspace of F^n of dimension $n-rank(L_A) = n-rank(A)$.*

Corollary. *If m<n, the system $Ax = 0$ has a nonzero solution.*

We refer to the equation Ax = 0 as the **homogeneous system corresponding to** Ax = b.

Theorem 3.9. *Let K be the solution set of a system of linear equations $Ax = b$, and let K_H be the solution set of the corresponding homogeneous system $Ax = 0$. Then for any solution s to $Ax = b$*

$$K = \{s\} + K_H = \{s + k : k \in K_H\}$$

Theorem 3.10. *Let $Ax = b$ be a system of n linear equations in n unknowns. If A is invertible, then the system has exactly one solution, namely, $A^{-1}b$. Conversely, if the system has exactly one solution, then A is invertible.*

The matrix (A|b) is called the **augmented matrix** of the system Ax = b.

Theorem 3.11. *Let $Ax = b$ be a system of linear equations. Then the system is consistent if and only if rank(A) = rank(A|b).*

Theorem 3.12.

3.4 Systems of Linear Equations - Computational Aspects

Definition. *Two systems of linear equations are called equivalent if they have the same solution set.*

Theorem 3.13. *Let $Ax = b$ be a system of m linear equations in n unknowns, and let C be an invertible $m \times m$ matrix. Then the system $(CA)x = Cb$ is equivalent to $Ax = b$.*

Corollary. *Let $Ax = b$ be a system of m linear equations in n unknowns. If $(A'|b')$ is obtained from $(A|b)$ by a finite number of elementary row operations, then the system $A'x = b'$ is equivalent to the original system.*

Definition. *A matrix is said to be in **reduced row echelon form** if the following three conditions are satisfied.*
(a) Any row containing a nonzero entry precedes any row in which all the entries are zero (if any).
(b) The first nonzero entry in each row is the only nonzero entry in its column.
(c) The first nonzero entry in each row is 1 and it occurs in a column to the right of the first nonzero entry in the preceding row.

It can be shown (see the corollary to Theorem 3.16) that the reduced row echelon form of a matrix is unique.

The procedure described on pages 183–185 for reducing an augmented matrix to reduced row echelon form is called **Gaussian elimination**. It consists of two separate parts.

1. In the forward pass (steps 1-5), the augmented matrix is transformed into an upper triangular matrix in which the first nonzero entry of each row is 1, and it occurs in a column to the right of the first nonzero entry of each preceding row.

2. In the backward pass or back-substitution (steps 6-7), the upper triangular matrix is transformed into reduced row echelon form by making the first nonzero entry of each row the only nonzero entry of its column.

Theorem 3.14. *Gaussian elimination transforms any matrix into its reduced row echelon form.*

To solve a system for which the augmented matrix is in reduced row echelon form, divide the variables into two sets. The first set consists of those variables that appear as leftmost variables in one of the equations of the system (in this case the set is $\{x_1, x_2, x_4\}$). The second set consists of all the remaining variables (in this case, $\{x_3, x_5\}$). To each variable in the second set, assign a parametric value $t_1, t_2, \ldots (x_3 = t_1, x_5 = t_2)$, and then solve for the variables of the first set in terms of those in the second set. See page 188 of the textbook.

Theorem 3.15. *Let $Ax = b$ be a system of r nonzero equations in n unknowns. Suppose that rank(A) = rank(A|b) and that (A|b) is in reduced row echelon form. Then*

(a) rank(A) = r.

(b) If the general solution obtained by the procedure above is of the form

$$s = s_0 + t_1 u_1 + t_2 u_2 + \cdots + t_{n-r} u_{n-r},$$

then $\{u_1, u_2, \ldots, u_{n-r}\}$ is a basis for the solution set of the corresponding homogeneous system, and s0 is a solution to the original system.

Theorem 3.16. *Let A be an $m \times n$ matrix of rank r, where $r > 0$, and let B be the reduced row echelon form of A. Then*

(a) The number of nonzero rows in B is r.

(b) For each $i = 1, 2, \ldots, r$, there is a column b_{j_i} of B such that $b_{j_i} = e_i$.

(c) The columns of A numbered $j_1, j_2, \ldots, j_r$ are linearly independent.

(d) For each $k = 1, 2, \ldots n$, if column k of B is $d_1 e_1 + d_2 e_2 + \cdots + d_r e_r$, then column k of A is $d_1 a_{j_1} + d_2 a_{j_2} + \cdots + d_r a_{j_r}$.

Corollary. *The reduced row echelon form of a matrix is unique.*

4 Determinants

4.1 Determinants of Order 2

Definition. *If*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

*is a 2×2 matrix with entries from a field F, then we define the **determinant** of A, denoted $\det(A)$ or $|A|$, to be the scalar $ad - bc$.*

Since $\det(A + B) \neq \det(A) + \det(B)$, the function $\det : M_{2 \times 2}(R) \to R$ is not a linear transformation.

Theorem 4.1. *The function $\det : M_{2 \times 2}(F) \to F$ is a linear function of each row of a 2×2 matrix when the other row is held fixed. That is, if u, v, and w are in F^2 and k is a scalar, then*

$$\det \begin{pmatrix} u + kv \\ w \end{pmatrix} = \det \begin{pmatrix} u \\ w \end{pmatrix} + k \det \begin{pmatrix} v \\ w \end{pmatrix}$$

and

$$\det \begin{pmatrix} w \\ u + kv \end{pmatrix} = \det \begin{pmatrix} w \\ u \end{pmatrix} + k \det \begin{pmatrix} w \\ v \end{pmatrix}$$

Theorem 4.2. *Let $A \in M_{2 \times 2}(F)$. Then the determinant of A is nonzero if and only if A is invertible. Moreover, if A is invertible, then*

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$$

The Area of a Parallelogram

4.2   Determinants of Order n

Given $A \in M_{n \times n}(F)$, for $n \geq 2$, denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting row i and column j by $\tilde{A}_{ij}$.

Definitions. *Let $A \in M_{n \times n}(F)$. If $n = 1$, so that $A = (A_{11})$, we define $\det(A) = A_{11}$. For $n \geq 2$, we define $\det(A)$ recursively as*

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \det(\tilde{A}_{1j})$$

*The scalar $\det(A)$ is called the **determinant** of A and is also denoted by $|A|$. The scalar*

$$(-1)^{i+j} \det(\tilde{A}_{ij})$$

*is called the **cofactor** of the entry of A in row i, column j.*

Thus the determinant of A equals the sum of the products of each entry in row 1 of A multiplied by its cofactor. This formula is called **cofactor expansion along the first row** of A.

Theorem 4.3. *The determinant of an $n \times n$ matrix is a linear function of each row when the remaining rows are held fixed. That is, for $1 \leq r \leq n$, we have*

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u + kv \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u \\ ar+1 \\ \vdots \\ a_n \end{pmatrix} + k \det \begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ v \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix}$$

whenever k is a scalar and u, v, and each a_i are row vectors in F^n.

Corollary. *If $A \in M_{n \times n}(F)$ has a row consisting entirely of zeros, then $\det(A) = 0$.*

Lemma. *Let $B \in M_{n \times n}(F)$, where $n \geq 2$. If row i of B equals e_k for some $k(1 \leq k \leq n)$, then $\det(B) = (-1)^{i+k} \det(\tilde{B}_{ik})$.*

Theorem 4.4. *The determinant of a square matrix can be evaluated by cofactor expansion along any row. That is, if $A \in M_{n \times n}(F)$, then for any integer $i(1 \ i \ n)$,*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \cdot \det(\tilde{A}_{ij})$$

Corollary. *If $A \in M_{n \times n}(F)$ has two identical rows, then $\det(A) = 0$.*

Theorem 4.5. *If $A \in M_{n \times n}(F)$ and B is a matrix obtained from A by interchanging any two rows of A, then $\det(B) = -\det(A)$.*

Theorem 4.6. *Let $A \in M_{n \times n}(F)$, and let B be a matrix obtained by adding a multiple of one row of A to another row of A. Then $\det(B) = \det(A)$.*

Corollary. *If $A \in M_{n \times n}(F)$ has rank less than n, then $\det(A) = 0$.*

The following rules summarize the effect of an elementary row operation on the determinant of a matrix $A \in M_{n \times n}(F)$.

(a) If B is a matrix obtained by interchanging any two rows of A, then $\det(B) = -\det(A)$.

(b) If B is a matrix obtained by multiplying a row of A by a nonzero scalar k, then $\det(B) = k \det(A)$.

(c) If B is a matrix obtained by adding a multiple of one row of A to another row of A, then $\det(B) = \det(A)$.

The determinant of an upper triangular matrix is the product of its diagonal entries.

4.3   Properties of Determinants

Because the determinant of the n×n identity matrix is 1, we can interpret the statements on page 217 as the following facts about the determinants of elementary matrices.

(a) If E is an elementary matrix obtained by interchanging any two rows of I, then $\det(E) = -1$.

(b) If E is an elementary matrix obtained by multiplying some row of I by the nonzero scalar k, then $\det(E) = k$.

(c) If E is an elementary matrix obtained by adding a multiple of some row of I to another row, then $\det(E) = 1$.

Theorem 4.7. *For any $A, B \in M_{n \times n}(F)$, $\det(AB) = \det(A) \cdot \det(B)$.*

Corollary. *A matrix $A \in M_{n \times n}(F)$ is invertible if and only if $\det(A) \neq 0$. Furthermore, if A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$.*

Theorem 4.8. *For any $A \in M_{n \times n}(F)$, $\det(A^t) = \det(A)$.*

Theorem 4.9 (Cramer's Rule). *Let $Ax = b$ be the matrix form of a system of n linear equations in n unknowns, where $x = (x_1, x_2, \ldots, x_n)^t$. If $\det(A) \neq 0$, then this system has a unique solution, and for each $k(k = 1, 2, \ldots, n)$,*

$$x_k = \frac{\det(M_k)}{\det(A)}$$

where M_k is the $n \times n$ matrix obtained from A by replacing column k of A by b.

If the rows of A are $a_1, a_2, \ldots, a_n$, respectively, then $|\det(A)|$ is the **n-dimensional volume** (the generalization of area in R^2 and volume in R^3) of the parallelepiped having the vectors $a_1, a_2, \ldots, a_n$ as adjacent sides.

Specifically, if γ is any ordered basis for R^n and β is the standard ordered basis for R^n, then γ induces a right-handed coordinate system if and only if $\det(Q) > 0$, where Q is the change of coordinate matrix changing γ-coordinates into β-coordinates.

4.4   Summary - Important Facts about Determinants

If A and B are similar matrices, then $\det(A) = \det(B)$.

4.5   A Characterization of the Determinant*