Linear Algebra Theorems and Definitions

Alexander J. Clarke

February 20, 2024

All theorems, corollaries, lemmas, remarks, and asides are direct quotes from
Linear Algebra, 4th Edition, by Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence

Contents

Chapter 0

List of Symbols

A_{ij}	the ij-th entry of the matrix A
A^{-1}	the inverse of the matrix A
$A^{\dagger}$	the pseudoinverse of the matrix A
A^{*}	the adjoint of the matrix A
$\tilde{A}_{ij}$	the matrix A with row i and column j deleted
A^{t}	the transpose of the matrix A
$(A\|B)$	the matrix A augmented by the matrix B
$B_1 \bigoplus \cdots \bigoplus B_k$	the direct sum of matrices B_1 through B_k
$\mathcal{B}(V)$	the set of bilinear forms on V
β^{*}	the dual basis of β
β_x	the T-cyclic basis generated by x
$\mathbb{C}$	the field of complex numbers
$\mathbb{C}_i$	the ith Gerschgorin disk
$\mathrm{cond}\,(A)$	the condition number of the matrix A
$C^n(\mathbb{R})$	set of functions f on $\mathbb{R}$ with $f^{(n)}$ continuous
C^{∞}	set of functions with derivatives of every order
$C(\mathbb{R})$	the vector space of continuous functions on $\mathbb{R}$
$C([0,1])$	the vector space of continuous functions on $[0,1]$
C_x	the T-cyclic subspaces generated by x
D	the derivative operator on C^{∞}
$\det\,(A)$	the determinant of the matrix A
δ_{ij}	the Kronecker delta
$\dim\,(V)$	the dimension of V
e^{A}	$\lim_{m\to\infty}\left(I + A + \frac{A^2}{2!} + \cdots + \frac{A^m}{m!}\right)$
e_i	the ith standard vector of $\mathbb{F}^n$

E_λ	the eigenspace of T corresponding to λ
$\mathbb{F}$	a field
$f(A)$	the polynomial $f(x)$ evaluated at the matrix A
F^n	the set of n-tuples with entries in a field $\mathbb{F}$
$f(T)$	the polynomial $f(x)$ evaluated at the operator T
$\mathcal{F}(S, \mathbb{F})$	the set of functions from S to a field $\mathbb{F}$
H	space of continuous complex functions on $[0, 2\pi]$
I_n or I	the $n \times n$ identity matrix
$\mathbb{I}_V$ or $\mathbb{I}$	the identity operator on V
K_λ	generalized eigenspace of T corresponding to λ
K_ϕ	$\{x \mid (\phi(T))^p(x) = 0$, for some positive integer $p\}$
L_A	left-multiplication transformation by matrix A
$\lim\limits_{m\to\infty} A_m$	the limit of a sequence of matrices
$\mathcal{L}(V)$	the space of linear transformations from V to V
$\mathcal{L}(V, W)$	the space of linear transformations from V to W
$M_{m\times n}(\mathbb{F})$	the set of $m \times n$ matrices with entries in $\mathbb{F}$
$v(A)$	the column sum of the matrix A
$v_j(A)$	the jth column sum of the matrix A
$N(T)$	the null space of T
$\operatorname{nullity}(T)$	the dimension of the null space of T
O	the zero matrix
$\operatorname{per}(M)$	the permanent of the 2×2 matrix M
$P(\mathbb{F})$	the space of polynomials with coefficients in $\mathbb{F}$
$P_n(\mathbb{F})$	the polynomials in $P(\mathbb{F})$ of degree at most n
ϕ_β	the standard representation with respect to basis β
$\mathbb{R}$	the field of real numbers
$\operatorname{rank}(A)$	the rank of the matrix A
$\operatorname{rank}(T)$	the rank of the linear transformation T
$\rho(A)$	the row sum of the matrix A
$\rho_i(A)$	the ith row sum of the matrix A
$R(T)$	the range of the linear transformation T
$S_1 + S_2$	the sum of sets S_1 and S_2
$\operatorname{span}(S)$	the span of the set S
$S^\perp$	the orthogonal complement of the set S
$[T]_\beta$	the matrix representation of T in basis β
$[T]_\beta^\gamma$	the matrix representation of T in bases β and γ
T^{-1}	the inverse of the linear transformation T

$T^{\dagger}$	the pseudoinverse of the linear transformation T
T^*	the adjoint of the linear operator T
T_0	the zero transformation
T^t	the transpose of the linear transformation T
T_θ	the rotation transformation by θ
T_W	the restriction of T to a subspace W
$\mathrm{tr}\,(A)$	the trace of the matrix A
V^*	the dual space of the vector space V
V/W	the quotient space of V modulo W
$W_1 + \cdots + W_k$	the sum of subspaces W_1 through W_k
$\sum_{i=1}^{k} W_i$	the sum of subspaces W_i through W_k
$W_1 \bigoplus W_2$	the direct sum of subspaces W_1 and W_2
$W_1 \bigoplus \cdots \bigoplus W_k$	the direct sum of subspaces W_1 through W_k
$\|x\|$	the norm of the vector $\vec{x}$
$[x]_\beta$	the coordinate vector of x relative to β
$\langle x, y \rangle$	the inner product of $\vec{x}$ and $\vec{y}$
$\mathbb{Z}_2$	the field consisting of 0 and 1
$\overline{\vec{z}}$	the complex conjugate of $\vec{z}$
$\vec{0}$	the zero vector

Chapter 1

Vector Spaces

1.1 Introduction

Theorem 1.1 (**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.

Definition 1.1.
Two nonzero vectors x and y are called **parallel** if $y = tx$ for some nonzero real number t. (Thus nonzero vectors having the same or opposite directions are parallel.)

1.2 Vector Spaces

Definition 1.2.
A **vector space** (or **linear space**) V over a field $\mathbb{F}$ consists of a set on which two operations (called **addition** and **scalar multiplication**, respectively) are defined so that for each pair of elements x and y in V there is a unique element a in $\mathbb{F}$ 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 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 $\mathbb{F}$ and each element x in V, $(ab)x = a(bx)$.

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

(VS 8) For each pair of elements a, b in $\mathbb{F}$ 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 $\mathbb{F}$ are called **scalars** and the elements of the vector space V are called **vectors**.

Note: The reader should not confuse this use of the word "vector" with the physical entity discussed in section 1.1: the word "vector" is now being used to describe any element of a vector space.

Definition 1.3.
An object of the form $(a_1, a_2, \ldots, a_n)$, where the entries $a_1, a_2, \ldots, a_n$ are elements of a field $\mathbb{F}$, is called an ***n*-tuple** with entries from $\mathbb{F}$. The elements $a_1, a_2, \ldots, a_n$ are called the **entries** or **components** of the n-tuple. Two n-tuples $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ with entries from a field $\mathbb{F}$ are called **equal** if $a_i = b_i$ for $i = 1, 2, \ldots, n$.

Definition 1.4.
Vectors in $\mathbb{F}^n$ may be written as **column vectors**

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

rather than as **row vectors** $(a_1, a_2, \ldots, a_n)$. Since a 1-tuple whose only entry is from $\mathbb{F}$ can be regarded as an element of $\mathbb{F}$, we usually write $\mathbb{F}$ rather than $\mathbb{F}^1$ for the vector space of 1-tuples with entry from $\mathbb{F}$.

Definition 1.5.
An $m \times n$ **matrix** with entries from a field $\mathbb{F}$ is a rectangular array of the form

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix},$$

where each entry a_{ij} $(1 \leq i \leq m,\ 1 \leq j \leq n)$ is an element of $\mathbb{F}$. We call the entries a_{ij} with $i = j$ the **diagonal entries** of the matrix. The entries $a_{i1}, a_{i2}, \ldots, a_{in}$ compose the ***i*th row** of the matrix, and the entries $a_{1j}, a_{2j}, \ldots, a_{mj}$ compose the ***j*th column** of the matrix. The rows of the preceding matrix are regarded as vectors in $\mathbb{F}^n$, and the columns are regarded as vectors in $\mathbb{F}^m$. The $m \times n$ matrix in which each entry equals zero is called the **zero matrix** and is denoted by O.

In this book, we denote matrices by capital italic letters (e.g. A, B, and C), and we denote the entry of a matrix A that lies in row i and column j by A_{ij}. In addition, if the number of rows and columns of a matrix are equal, the matrix 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 1.6.
The set of all $m \times n$ matrices with entries from a field $\mathbb{F}$ is a vector space which we denote by $M_{m\times n}(\mathbb{F})$, with the following operations of **matrix addition** and **scalar multiplication**: For $A, B \in M_{m\times n}(\mathbb{F})$ and $c \in \mathbb{F}$,

$$(A+B)_{ij} = A_{ij} + B_{ij} \quad \text{and} \quad (cA)_{ij} = cA_{ij}$$

for $1 \leq i \leq m$ and $1 \leq j \leq n$.

Definition 1.7.
Let S be any nonempty set and $\mathbb{F}$ be any field, and let $\mathcal{F}(S, \mathbb{F})$ denote the set of all functions from S to $\mathbb{F}$. Two functions f and g in $\mathcal{F}(S, \mathbb{F})$ are called **equal** if $f(s) = g(s)$ for each $s \in S$. The set $\mathcal{F}(S, \mathbb{F})$ is a vector space with the operations of addition and scalar multiplication defined for $f, g \in \mathcal{F}(S, \mathbb{F})$ and $c \in \mathbb{F}$ defined by

$$(f+g)(s) = f(s) + g(s) \quad \text{and} \quad (cf)(s) = c[f(s)]$$

for each $s \in S$. Note that these are the familiar operations of addition and scalar multiplication for functions used in algebra and calculus.

Definition 1.8.
A **polynomial** with coefficients from a field $\mathbb{F}$ is an expression of the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

where n is a nonnegative integer and each a_k, called the **coefficient** of x^k, is in $\mathbb{F}$. If $f(x) = 0$, that is, if $a_n = a_{n-1} = \cdots = a_0 = 0$, then $f(x)$ is called the **zero polynomial** and, for convenience, its degree is defined to be -1; otherwise, the **degree** of a polynomial is defined to be the largest exponent of x that appears in the representation

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

with a nonzero coefficient. Note that the polynomials of degree zero may be written in the form $f(x) = c$ for some nonzero scalar c. Two polynomials,

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

and

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0,$$

are called **equal** if $m = n$ and $a_i = b_i$ for $i = 1, 2, \ldots, n$.

Definition 1.9.
Let $\mathbb{F}$ be any field. A **sequence** in $\mathbb{F}$ is a function σ from the positive integers into $\mathbb{F}$. In this book, the sequence σ such that $\sigma(n) = a_n$ for $n = 1, 2, \ldots$ is denoted $\{a_n\}$. Let V consist of all sequences $\{a_n\}$ in $\mathbb{F}$ that have only a finite number of nonzero terms a_n. If $\{a_n\}$ and $\{b_n\}$ are in V and $t \in \mathbb{F}$, define

$$\{a_n\} + \{b_n\} = \{a_n + b_n\} \quad \text{and} \quad t\{a_n\} = \{ta_n\}$$

Theorem 1.2 (**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 1.1.
The vector 0 described in (VS 3) is unique.

Corollary 1.2.
The vector y described in (VS 4) is unique.

Definition 1.10.
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 1.3.
In any vector space V, the following statements are true:

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

Definition 1.11.
Let $V = \{0\}$ consist of a single vector 0 and define $0 + 0 = 0$ and $c0 = 0$ for each scalar $c \in \mathbb{F}$. Then V is called the **zero vector space**.

Definition 1.12.
A real-valued function f defined on the real line is called an **even function** if $f(-t) = f(t)$ for each real number t, and is called an **odd function** if $f(-t) = -f(t)$ for each real number t.

Definition 1.13.
If S_1 and S_2 are nonempty subsets of a vector space V, then the **sum** of S_1 and S_2, denoted $S_1 + S_2$, is the set $\{x + y \mid x \in S_1, \text{ and } y \in S_2\}$.

Definition 1.14.
A vector space V is called the **direct sum** of W_1 and W_2 if W_1 and W_2 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 W_1 and W_2 by writing $V = W_1 \oplus W_2$.

Definition 1.15.
A matrix M is called **skew-symmetric** if $M^t = -M$.

Definition 1.16.
Let W be a subspace of a vector space V over a field $\mathbb{F}$. For any $v \in V$, the set $\{v\} + W = \{v + w \mid w \in W\}$ is called the **coset of** W **containing** v. It is customary to denote this coset by $v + W$ rather than $\{v\} + W$.

Definition 1.17.
Let W be a subspace of a vector space V over a field $\mathbb{F}$, and let $S := \{v + W \mid v \in V\}$ be the set of all cosets of W. Then S is called the **quotient space of V modulo W**, and is denoted by V/W. Addition and scalar multiplication by the scalars of $\mathbb{F}$ can be defined as follows:

$$(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W$$

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

$$a(v + W) = av + W$$

for all $v \in V$ and $a \in \mathbb{F}$.

1.3 Subspaces

Definition 1.18.
A subset W of a vector space V over a field $\mathbb{F}$ is called a **subspace** of V if W is a vector space over $\mathbb{F}$ 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.

Fortunately, it is not necessary to verify all of the vector space properties to prove that a subset is a subspace. Because properties (VS 1), (VS 2), (VS 5), (VS 6), (VS 7) and (VS 8) hold for all vectors in the vector space, these properties automatically hold for the vectors in any subset. Thus a subset W of a vector space V is a subspace of V if and only if the following four properties hold:

1. $x + y \in W$ whenever $x \in W$ and $y \in W$. (W is **closed under addition**).

2. $cx \in W$ whenever $c \in \mathbb{F}$ and $x \in W$. (W is **closed under scalar multiplication**).

3. W has a zero vector.

4. Each vector in W has an additive inverse in W.1

Theorem 1.4.
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.

1. $0 \in W$.

2. $x + y \in W$ whenever $x \in W$ and $y \in W$.

3. $cx \in W$ whenever $c \in \mathbb{F}$ and $x \in W$.

Definition 1.19.
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, $(A^t)_{ij} = A_{ji}$.

Definition 1.20.
A **symmetric matrix** is a matrix A such that $A^t = A$.

Definition 1.21.
An $n \times n$ matrix M is called a **diagonal matrix** if $M_{ij} = 0$ whenever $i \neq j$; that is, if all nondiagonal entries are zero.

Definition 1.22.
The **trace** of an $n \times n$ matrix M, denoted $\mathrm{tr}(M)$, is the sum of the diagonal entries of M; that is,

$$\mathrm{tr}(M) = M_{11} + M_{22} + \cdots + M_{nn}.$$

Theorem 1.5.
Any intersection of subspaces of a vector space V is a subspace of V.

Definition 1.23.
An $m \times n$ matrix A is called **upper triangular** if all entries lying below the diagonal entries are zero; that is, if $A_{ij} = 0$ whenever $i > j$.

1.4 Linear Combinations and Systems of Linear Equations

Definition 1.24.
Let V be a vector space and S a nonempty subset of V. A vector $v \in V$ is called a **linear combination** of vectors of S if there exist a finite number of vectors $v_1, v_2, \ldots, v_n$ in S and scalars $a_1, a_2, \ldots, a_n$ in $\mathbb{F}$ such that $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. In this case we also say that v is a linear combination of $v_1, v_2, \ldots, v_n$ and call $a_1, a_2, \ldots, a_n$ the **coefficients** of the linear combination.

Definition 1.25.
Let S be a nonempty subset of a vector space V. The **span** of S, denoted $\mathrm{span}\,(S)$, is the set consisting of all linear combinations of the vectors in S. For convenience, we define $\mathrm{span}\,(\emptyset) = \{0\}$.

Theorem 1.6.
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 1.26.
A subset S of a vector space V **generates** (or **spans**) V if $\mathrm{span}(S) = V$. In this case, we also say that the vectors of S generate (or span) V.

1.5 Linear Dependence and Linear Independence

Definition 1.27.
A subset S of a vector space V is called **linearly dependent** if there exist a finite number of distinct vectors $v_1, v_2, \dots, v_n$ in S and scalars $a_1, a_2, \dots, a_n$ not all zero, such that

$$a_1 v_2 + a_2 v_2 + \cdots + a_n v_n = 0$$

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

For any vectors $v_1, v_2, \dots, v_n$, we have $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$ if $a_1 = a_2 = \cdots = a_n = 0$. We call this the **trivial representation** of 0 as a linear combination of $v_1, v_2, \dots, v_n$. Thus, for a set tot be linearly dependent, there must exist a nontrivial representation of 0 as a linear combination of vectors in the set. Consequently, any subset of a vector space that contains the zero vector is linearly dependent, because $0 = 1 \cdot 0$ is a nontrivial representation of 0 as a linear combination of vectors in the set.

Definition 1.28.
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 $\{v\}$ is linearly dependent, then $av = 0$ for some nonzero scalar a. thus

$$v = a^{-1}(av) = a^{-1}0 = 0.$$

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 1.7.
Let V be a vector space, and let $S_1 \subseteq S_2 \subseteq V$. If S_1 is linearly dependent, then S_2 is linearly dependent.

Corollary 1.3.
Let V be a vector space, and let $S_1 \subseteq S_2 \subseteq V$. If S_2 is linearly independent, then S_1 is linearly independent.

Theorem 1.8.
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 $v \in \operatorname{span}(S)$.

1.6 Bases and Dimension

Definition 1.29.
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 1.9.
Let V be a vector space and $\beta = \{v_1, v_2, \ldots, v_n\}$ 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

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$$

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

Theorem 1.10.
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 1.11 (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 1.4. Let V be a vector space having a finite basis. Then every basis for V contains the same number of vectors.

Definition 1.30.
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 1.5.
Let V be a vector space with dimension n.

1. 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.

2. Any linearly independent subset of V that contains exactly n vectors is a basis for V.

3. Every linearly independent subset of V can be extended to a basis for V.

Theorem 1.12.
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 1.6.
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.

Definition 1.31 (**The Lagrange Interpolation Formula**).
Corollary 2 of the replacement theorem can be applied to obtain a useful formula. Let $c_0, c_1, \ldots, c_n$ be distinct scalars in an infinite field $\mathbb{F}$. The polynomials $f_0(x), f_1(x), \ldots, f_n(x)$ defined by

$$f_i(x) = \frac{(x-c_0)\ldots(x-c_{i-1})(x-c_{i+1})\ldots(x-c_n)}{(c_i-c_0)\ldots(c_i-c_{i-1})(c_i-c_{i+1})\ldots(c_i-c_n)} = \prod_{\substack{k=0 \\ k\neq i}}^{n} \frac{x-c_k}{c_i-c_k}$$

are called the **Lagrange polynomials** (associated with $c_0, c_1, \ldots, c_n$). Note that each $f_i(x)$ is a polynomial of degree n and hence is in $P_n(\mathbb{F})$. By regarding $f_i(x)$ as a polynomial function $f_i : \mathbb{F} \to \mathbb{F}$, we see that

$$f_i(c_j) = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \tag{1.1}$$

This property of Lagrange polynomials can be used to show that $\beta = \{f_0, f_1, \ldots, f_n\}$ is a linearly independent subset of $P_n(\mathbb{F})$. Suppose that

$$\sum_{i=0}^{n} a_i f_i = 0 \quad \text{for some scalars } a_0, a_1, \ldots, a_n,$$

where 0 denotes the zero function. Then

$$\sum_{i=0}^{n} a_i f_i(c_j) = 0 \quad \text{for } j = 0, 1, \ldots, n.$$

But also

$$\sum_{i=0}^{n} a_i f_i(c_j) = a_j$$

by (1.1). Hence $a_j = 0$ for $j = 0, 1, \ldots, n$; so β is linearly independent. Since the dimension of $P_n(\mathbb{F})$ is $n+1$, it follows from Corollary 2 of the replacement theorem that β is a basis for $P_n(\mathbb{F})$.

Because β is a basis for $P_n(\mathbb{F})$, every polynomial function g in $P_n(\mathbb{F})$ is a linear combination of polynomial functions of β, say,

$$g = \sum_{i=0}^{n} b_i f_i.$$

It follows that

$$g(c_j) = \sum_{i=0}^{n} b_i f_i(c_j) = b_j;$$

so

$$g = \sum_{i=0}^{n} g(c_i) f_i$$

is the unique representation of g as a linear combination of elements of β. This representation is called the **Lagrange interpolation formula**. Notice that the preceding argument shows that if $b_0, b_1, \ldots, b_n$ are any $n+1$ scalars in $\mathbb{F}$ (not necessarily distinct), then the polynomial function

$$g = \sum_{i=0}^{n} b_i f_i$$

is the unique polynomial in $P_n(\mathbb{F})$ such that $g(c_j) = b_j$. Thus we have found the unique polynomial of degree not exceeding n that has specified values b_j at given points c_j in its domain $(j = 0, 1, \ldots, n)$.

An important consequence of the Lagrange interpolation formula is the following result: If $f \in P_n(\mathbb{F})$ and $f(c_i) = 0$, for $n+1$ distinct scalars $c_0, c_1, \ldots, c_n$ in $\mathbb{F}$, then f is the zero function.

1.7 Maximal Linearly Independent Subsets

Definition 1.32.
Let $\mathcal{F}$ be a family 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 1.33.
Let $\mathcal{F}$ be the family of all subsets of a nonempty set S. This family $\mathcal{F}$ is called the **power set** of S.

Definition 1.34.
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 $B \subseteq A$.

Definition 1.35 (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.

Note: The *Maximal Principle* is logically equivalent to the *Axiom of Choice*, which is an assumption in most axiomatic developments of set theory.

Definition 1.36.
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.

Corollary 1.7.
Every vector space has a basis.

Chapter 2

Linear Transformations and Matrices

2.1 Linear Transformations, Null Spaces, and Ranges

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

1. $T(x + y) = T(x) + T(y)$, and

2. $T(cx) = cT(x)$

If the underlying field $\mathbb{F}$ is the field of rational numbers, then (1) implies (2), but, in general (1) and (2) are logically independent.

We often simply call T **linear**.

Remark 2.1.
Let V and W be vector spaces (over $\mathbb{F}$). Let $T : V \to W$ be a linear transformation. Then the following properties hold:

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 $c \in \mathbb{F}$.

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 $x_1, x_2, \ldots, x_n \in V$ and $a_1, a_2, \ldots, a_n \in \mathbb{F}$, we have

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

We generally use property 2 to prove that a given transformation is linear.

Definition 2.2.
For any angle θ, define $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ by the rule: $T_\theta(a_1, a_2)$ is the vector obtained by rotating (a_1, a_2) counterclockwise by θ if $(a_1, a_2) \neq (0,0)$, and $T_\theta(0,0) = (0,0)$. Then $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation that is called the **rotation by** θ.

Definition 2.3.
Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(a_1, a_2) = (a_1, -a_2)$. T is called the **reflection about the x-axis**.

Definition 2.4.
For vector spaces V and W (over $\mathbb{F}$), we define the **identity transformation** $I_V : V \to V$ by $I_V(x) = x$ for all $x \in V$.

We define the **zero transformation** $T_0 : V \to W$ by $T_0(x) = 0$ for all $x \in V$.

Note: We often write I instead of I_V.

Definition 2.5.
Let V and W be vector spaces, and let $T : V \to W$ be linear. We define the **null space** (or **kernel**) $\mathrm{N}(T)$ to be the set of all vectors $x \in V$ such that $T(x) = 0$; that is,
$\mathrm{N}(T) = \{x \in V \mid T(x) = 0\}$.

We define the **range** (or **image**) $\mathrm{R}(T)$ of T to be the subset of W consisting of all images (under T) of vectors in V; that is, $\mathrm{R}(T) = \{T(x) \mid x \in V\}$.

Theorem 2.1.
Let V and W be vector spaces and $T : V \to W$ be linear. Then $\mathrm{N}(T)$ and $\mathrm{R}(T)$ are subspaces of V and W, respectively.

Theorem 2.2.
Let V and W be vector spaces, and let $T : V \to W$ be linear. If $\beta = \{v_1, v_2, \ldots, v_n\}$ is a basis for V, then

$$\mathrm{R}(T) = \mathrm{span}(T(\beta)) = \mathrm{span}(\{T(v_1), T(v_2), \ldots, T(v_n)\}).$$

Definition 2.6.
Let V and W be vector spaces, and let $T : V \to W$ be linear. If $\mathrm{N}(T)$ and $\mathrm{R}(T)$ are finite-dimensional, then we define the **nullity** of T, denoted $\mathrm{nullity}(T)$, and the **rank** of T, denoted $\mathrm{rank}(T)$, to be the dimensions of $\mathrm{N}(T)$ and $\mathrm{R}(T)$, respectively.

Theorem 2.3 (Dimension Theorem).
Let V and W be vector spaces, and let $T : V \to W$ be linear. If V is finite-dimensional, then

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

Theorem 2.4.
Let V and W be vector spaces, and let $T : V \to W$ be linear. Then T is one-to-one if and only if $\mathrm{N}(T) = \{0\}$.

Theorem 2.5.
Let V and W be vector spaces of equal (finite) dimension, and let $T : V \to W$ be linear. Then the following are equivalent.

1. T is one-to-one.

2. T is onto.

3. $\text{rank}\,(T) = \dim\,(V)$.

Theorem 2.6.
Let V and W be vector spaces over $\mathbb{F}$, 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 2.1.
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 $i = 1, 2, \ldots, n$, then $U = T$.

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 $T : V \to V$ is called the **projection on** W_1 **along** W_2 if, for $x = x_1 + x_2$ with $x_1 \in W$ and $x_2 \in W_2$, we have $T(x) = x_1$.

Definition 2.8.
Let V be a vector space, and let $T : V \to W$ 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, $T(W) \subseteq W$. 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 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.

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

Definition 2.10.
Let $\beta = \{v_1, v_2, \ldots, v_n\}$ 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 scalar values such that

$$x = \sum_{i=1}^{n} a_i v_i.$$

We define the **coordinate vector of x relative to β**, denoted by $[x]_\beta$, by

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

Notice that $[v_i]_\beta = e_i$ in the preceding definition. It can be shown that the correspondence $x \to [x]_\beta$ provides us with a linear transformation from V to $\mathbb{F}^n$.

Notation 2.1.
The following notation is used to construct a matrix representation of a linear transformation in the following definition.

Suppose that V and W are finite-dimensional vector spaces with ordered bases $\beta = \{v_1, v_2, \ldots, v_n\}$ 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 $a_{ij} \in \mathbb{F}$, $1 \leq i \leq m$, such that

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

Definition 2.11.
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 $A = [T]_\beta$.

Notice that the jth column of A is simply $[T(v_j)]_\gamma$. Also observe that if $U : V \to W$ is a linear transformation such that $[U]_\beta^\gamma = [T]_\beta^\gamma$, then $U = T$ by the corollary to Theorem 2.6 (Corollary 2.1).

Definition 2.12.
Let $T, U : V \to W$ be arbitrary functions, where V and W are vector spaces over $\mathbb{F}$, and let $a \in \mathbb{F}$. 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 $\mathbb{F}$, and let $T, U : V \to W$ be linear.

1. For all $a \in \mathbb{F}$, $aT + U$ is linear.
2. 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 $\mathbb{F}$.

Definition 2.13.
Let V and W be vector spaces over $\mathbb{F}$. We denote the vector space of all linear transformations from V to 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

1. $[T+U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$ and

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

2.3 Compositions of Linear Transformations and Matrix Multiplication

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

Theorem 2.10.
Let V be a vector space. Let $T, U_1, U_2 \in \mathcal{L}(V)$. 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.

Definition 2.14.
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

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

Notice that $(AB)_{ij}$ is the sum of products of corresponding entries from the ith row of A and the jth column of B.

The reader should observe that in order for the product AB to be defined, there are restrictions regarding the relative sizes of A and B. The following mnemonic device is helpful: "$(m \times n) \cdot (n \times p) = (m \times p)$"; that is, in order for the product AB to be defined, the two "inner" dimensions must be equal, and the two "outer" dimensions yield the size of the product.

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

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

Corollary 2.2.
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 2.15.
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 2.12.
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. Then

1. $A(B + C) = AB + AC$ and $(D + E)A = DA + EA$.

2. $a(AB) = (aA)B = A(aB)$ for any scalar a.

3. $I_m A = A = AI_n$.

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

Corollary 2.3.
Let A be an $m \times n$ matrix, $B_1, B_2, \ldots, B_k$ be $n \times p$ matrices, $C_1, C_2, \ldots, C_k$ be $q \times m$ matrices, and $a_1, a_2, \ldots, a_k$ be scalars. Then

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

and

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

Theorem 2.13.
Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. For each j $(1 \leq j \leq p)$ let u_j and v_j denote the jth columns of AB and B, respectively. Then

1. $u_j = Av_j$.

2. $v_j = Be_j$, where e_j is the jth standard vector of $\mathbb{F}^p$.

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 2.16.
Let A be an $m \times n$ matrix with entries from a field $\mathbb{F}$. We denote $L_A : \mathbb{F}^n \to \mathbb{F}^m$ defined by $L_A(x) = Ax$ (the matrix product of A and x) for each column vector $x \in \mathbb{F}^n$. We call L_A a **left-multiplication transformation**.

Theorem 2.15.
Let A be an $m \times n$ matrix with entries from $\mathbb{F}$. Then the left-multiplication transformation $L_A : \mathbb{F}^n \to \mathbb{F}^m$ is linear. Furthermore, if B is any other $m \times n$ matrix (with entries from $\mathbb{F}$) and β and γ are the standard ordered bases for $\mathbb{F}^n$ and $\mathbb{F}^m$, respectively, then we have the following properties.

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 $a \in \mathbb{F}$.
4. If $T : \mathbb{F}^n \to \mathbb{F}^m$ 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 $L_{I_n} = I_{\mathbb{F}^n}$.

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.

Definition 2.17.
An **incidence matrix** is a square matrix in which all the entries are either zero or one and, for convenience, all the diagonal entries are zero. If we have a relationship on a set of n objects that we denote $1, 2, \ldots, n$, then we define the associated incidence matrix A by $A_{ij} = 1$ if i is related to j, and $A_{ij} = 0$ otherwise.

Definition 2.18.
A relationship among a group of people is called a **dominance relation** if the associated incidence matrix A has the property that for all distinct pairs i and j, $A_{ij} = 1$ if and only if $A_{ji} = 0$, that is, given any two people, exactly one of them *dominates* the other.

2.4 Invertibility and Isomorphisms

Definition 2.19.
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**. 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 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 finite-dimensional vector spaces of equal dimension. then T is invertible if and only if $\text{rank}\,(T) = \dim\,(T)$.

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 2.20.
Let A be an $n \times n$ matrix. Then A is **invertible** if there exists an $n \times n$ matrix B such that $AB = BA = I$.

If A is invertible, then the matrix B such that $AB = BA = I$ is unique. (If C were another such matrix, then $C = CI = C(AB) = (CA)B = IB = B$.) The matrix B is called the **inverse** of A and is denoted by A^{-1}.

Lemma 2.1.
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 2.4.
Let V be a finite-dimensional vector space with an ordered bases 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 2.5.
Let A be and $n \times n$ matrix. Then A is invertible if and only if L_A is invertible. Furthermore, $(L_A)^{-1} = L_{A^{-1}}$.

Definition 2.21.
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 2.6.
Let V be a vector space over $\mathbb{F}$. Then V is isomorphic to $\mathbb{F}^n$ if and only if $\dim\,(V) = n$.

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

Corollary 2.7.
Let V and W be finite-dimensional vector spaces of dimension n and m, respectively. Then $\mathcal{L}(V, W)$ is finite-dimensional of dimension mn.

Definition 2.22.
Let β be an ordered basis for an n-dimensional vector space V over the field $\mathbb{F}$. The **standard representation of V with respect to β** is the function $\phi_\beta : V \to \mathbb{F}^n$ defined by $\phi_\beta(x) = [x]_\beta$ for each $x \in V$.

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 pace 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 2.23.
The matrix $Q = [I_V]_{\beta'}^\beta$, defined in Theorem 2.22, is called a **change of coordinate matrix**. Because of part (2) of the theorem, we say that Q **changes β'-coordinates into β-coordinates**.

Definition 2.24.
A linear transformation that maps a vector space V into itself 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 2.8.
Let $A \in M_{n\times n}(\mathbb{F})$, and let γ be an ordered basis for $\mathbb{F}^n$. 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 2.25.
Let A and B be matrices in $M_{n\times n}(\mathbb{F})$. We say that B is **similar** to A if there exists an invertible matrix Q such that $B = Q^{-1}AQ$.

Notice that the relation of similarity is an equivalence relation. So we need only say that A and B are similar.

2.6 Dual Spaces

Definition 2.26.
A linear transformation from a vector space V into its field of scalars $\mathbb{F}$, which is itself a vector space of dimension 1 over $\mathbb{F}$, is called a **linear functional** on V. We generally use the letters $\mathsf{f}, \mathsf{g}, \mathsf{h}, \dots$. to denote linear functionals.

Definition 2.27.
Let V be a vector space of continuous real-valued functions on the interval $[0, 2\pi]$. Fix a function $g \in V$. The function $\mathsf{h} : V \to \mathbb{R}$, defined by

$$\mathsf{h}(x) = \frac{1}{2\pi}\int_0^{2\pi} x(t)g(t)dt$$

is a linear functional on V. In the cases that $g(t)$ equals $\sin(nt)$ or $\cos(nt)$, $\mathsf{h}(x)$ is often called the ***n*th Fourier coefficient of** x.

Definition 2.28.
Let V be a finite dimensional vector space, and let $\beta = \{x_1, x_2, \dots, x_n\}$ be an ordered basis for V. For each $i = 1, 2, \dots, n$, define $\mathsf{f}_i(x) = a_i$, 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 is a linear function on V called the ***i*th coordinate function with respect to the basis** β. Note that $\mathsf{f}_i(x_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. These linear functionals play an important role in the theory of dual spaces (see Theorem 2.24).

Definition 2.29.
For a vector space V over $\mathbb{F}$, we define the **dual space** of V to be the vector space $\mathcal{L}(V, \mathbb{F})$, denoted by V^*.

Thus V^* is the vector space consisting of all linear functionals on V with the operations of addition and scalar multiplication. Note that if V is finite-dimensional, then by Corollary 2.7

$$\dim(V^*) = \dim(\mathcal{L}(V, \mathbb{F})) = \dim(V) \cdot \dim(\mathbb{F}) = \dim(V).$$

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^*. In Theorem 2.26, we show, in fact, that there is a natural identification of V and V^{**} in the case that V is finite-dimensional.

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 $\beta^* = \{\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_n\}$. Then β^* is an ordered basis for V^*, and, for any $\mathsf{f} \in V^*$, we have

$$\mathsf{f} = \sum_{i=1}^{n} \mathsf{f}(x_i)\mathsf{f}_i.$$

Definition 2.30.
Using the notation of Theorem 2.24, we call the ordered basis $\beta^* = \{\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_n\}$ of V^* that satisfies $\mathsf{f}_i(x_j) = \delta_{ij}$ $(1 \leq i,\ j \leq n)$ the **dual basis** of β.

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

Definition 2.31.
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 $[U]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$.

Definition 2.32.
For a vector x in a finite-dimensional vector space V, we define the linear functional $\hat{x} : V^* \to \mathbb{F}$ on V^* by $\hat{x}(\mathsf{f}) = \mathsf{f}(x)$ for every $\mathsf{f} \in V^*$. Since $\hat{x}$ is a linear functional on V^*, $\hat{x} \in V^{**}$.

The correspondence $x \leftrightarrow \hat{x}$ allows us to define the desired isomorphism between V^* and V^{**}.

Lemma 2.2.
Let V be a finite-dimensional vector space, and let $x \in V$. If $\hat{x}(\mathsf{f}) = 0$ for all $\mathsf{f} \in V^*$, 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 2.9.
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 V.

Definition 2.33.
Let V be a finite-dimensional vector space over $\mathbb{F}$. For every subset S of V, define the **annihilator** S^0 of S as

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

2.7 Homogeneous Linear Differential Equations with Constant Coefficients

Definition 2.34.
A **differential equation** in an unknown function $y = y(t)$ is an equation involving y, t, and derivatives of y. If the differential equation is of the form

$$a_n y^{(n)} + a_{n-1} y^{(n-)} + \cdots + a_1 y^{(1)} + a_0 y = f, \tag{2.1}$$

where $a_0, a_1, \ldots, a_n$ and f are functions of t and $y^{(k)}$ denotes the kth derivative of y, then the equation is said to be **linear**. The functions a_i are called the **coefficients** of the differential equation. When f is identically zero, (2.1) is called **homogeneous**.

If $a_n \neq 0$, we say that differential equation (2.1) is of **order** $\boldsymbol{n}$. In this case, we divide both sides by a_n to obtain a new, but equivalent, equation

$$y^{(n)} + b_{n-1} y^{(n-1)} + \cdots + b_1 y^{(1)} + b_0 y = 0,$$

where $b_i = a_i / a_n$ for $i = 0, 1, \ldots, n-1$. Because of this observation, we always assume that the coefficient a_n in (2.1) is 1.

A **solution** to (2.1) is a function that when substituted for y reduces (2.1) to an identity.

Definition 2.35.
Given a complex-valued function $x \in \mathcal{F}(\mathbb{R}, \mathbb{C})$ of a real variable t (where $\mathcal{F}(\mathbb{R}, \mathbb{C})$ is the vector space defined in Definition 1.7), there exist unique real-valued functions x_1 and x_2 of t, such that

$$x(t) = x_1(t) + i x_2(t) \quad \text{for} \quad t \in \mathbb{R},$$

where i is the imaginary number such that $i^2 = -1$. We call x_1 the **real part** and x_2 the **imaginary part** of x.

Definition 2.36.
Given a function $x \in \mathcal{F}(\mathbb{R}, \mathbb{C})$ with real part x_1 and imaginary part x_2, we say that x is **differentiable** if x_1 and x_2 are differentiable. If x is differentiable, we define the **derivative** x' of x by

$$x' = x_1' + i x_2'$$

Theorem 2.27.
Any solution to a homogeneous linear differential equation with constant coefficients has derivatives of all orders; that is, if x is a solution to such an equation, then $x^(k)$ exists for every positive integer k.

Definition 2.37.
We use $\mathbb{C}^\infty$ to denote the set of all functions in $\mathcal{F}(\mathbb{R}, \mathbb{C})$ that have derivatives of all orders.

Definition 2.38.
For any polynomial $p(t)$ over $\mathbb{C}$ of positive degree, $p(D)$ is called a **differential operator**. The **order** of the differential operator $p(D)$ is the degree of the polynomial $p(t)$.

Definition 2.39.
Given the differential equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y^{(1)} + a_0 y = 0,$$

the complex polynomial

$$p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

is called the **auxiliary polynomial** associated with the equation.

Theorem 2.28.
The set of all solutions to a homogeneous linear differential equation with constant coefficients coincides with the null space of $p(D)$ where $p(t)$ is the auxiliary polynomial associated with the equation.

Corollary 2.10.
The set of all solutions to a homogeneous linear differential equation with constant coefficients is a subspace of $\mathbb{C}^\infty$.

Definition 2.40.
We call the set of solutions to a homogeneous linear differential equation with constant coefficients the **solution space** of the equation.

Definition 2.41.
Let $c = a + ib$ be a complex number with real part a and imaginary part b. Define

$$e^c = e^a(\cos(b) + i\sin(b)).$$

The special case

$$e^{ib} = \cos(b) + i\sin(a)$$

is called **Euler's formula**.

Definition 2.42.
A function $f : \mathbb{R} \to \mathbb{C}$ defined by $f(t) = e^{ct}$ for a fixed complex number c is called an **exponential function**.

Theorem 2.29.
For any exponential function $f(t) = e^{ct}$, $f'(t) = ce^{ct}$.

Theorem 2.30.
Recall that the **order** of a homogeneous linear differential equation is the degree of its auxiliary polynomial. Thus, an equation of order 1 is of the form

$$y' + a_0 y = 0. \tag{2.2}$$

The solution space for (2.2) is of dimension 1 and has $\{e^{-a_0 t}\}$ as a basis.

Corollary 2.11.
For any complex number c, the null space of the differential operator $D - c\mathsf{I}$ has $\{e^{ct}\}$ as a basis.

Theorem 2.31.
Let $p(t)$ be the auxiliary polynomial for a homogeneous linear differential equation with constant coefficients. For any complex number c, if c is a zero of $p(t)$, then e^{ct} is a solution to the differential equation.

Theorem 2.32.
For any differential operator $p(D)$ of order n, the null space of $p(D)$ is an n-dimensional subspace of $\mathbb{C}^\infty$.

Lemma 2.3.
The differential operator $D - c\mathsf{I} : \mathbb{C}^\infty \to \mathbb{C}^\infty$ is onto for any complex number c.

Lemma 2.4.
Let V be a vector space, and suppose that T and U are linear operators on V such that U is onto and the null spaces of T and U are finite-dimensional. Then the null space of TU is finite-dimensional, and

$$\dim(\mathrm{N}(TU)) = \dim(\mathrm{N}(T)) + \dim(\mathrm{N}(U))$$

Corollary 2.12.
The solution space of any nth-order homogeneous linear differential equation with constant coefficients is an n-dimensional subspace of $\mathbb{C}^\infty$.

Theorem 2.33.
Given n distinct complex numbers $c_1, c_2, \ldots, c_n$, the set of exponential functions $\{e^{c_1 t}, e^{c_2 t}, \ldots, e^{c_n t}\}$ is linearly independent.

Corollary 2.13.
For any nth-order homogeneous linear differential equation with constant coefficients, if the auxiliary polynomial has n distinct zeros $c_1, c_2, \ldots, c_n$, then $\{e^{c_1 t}, e^{c_2 t}, \ldots, e^{c_n t}\}$ is a basis for the solution space of the differential equation.

Lemma 2.5.
For a given complex number c and a positive integer n, suppose that $(t - c)^n$ is the auxiliary polynomial of a homogeneous linear differential equation with constant coefficients. Then the set

$$\beta = \{e^{ct}, te^{ct}, \ldots, t^{n-1}e^{ct}\}$$

is a basis for the solution space of the equation.

Theorem 2.34.
Given a homogeneous linear differential equation with constant coefficients and auxiliary polynomial

$$(t - c_1)^{n_1}(t - c_2)^{n_2} \ldots (t - c_k)^{n_k},$$

where $n_1, n_2, \ldots, n_k$ are positive integers and $c_1, c_2, \ldots, c_k$ are distinct complex numbers, the following set is a basis for the solution space of the equation:

$$\{e^{c_1 t}, te^{c_1 t}, \ldots, t^{n_1 - 1}e^{c_1 t}, \ldots, e^{c_k t}, te^{c_k t}, \ldots, t^{n_k - 1}e^{c_k t}\}$$

Definition 2.43.
A differential equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y^{(1)} + a_0 y = x$$

is called a **nonhomogeneous** linear differential equation with constant coefficients if the a_i's are constant and x is a function that is not identically zero.

Chapter 3

Elementary Matrix Operations and Systems of Linear Equations

3.1 Elementary Matrix Operations and Elementary Matrices

Definition 3.1.
Let A be an $m \times n$ 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 are 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 3.2.
An $n \times n$ **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 I_n is a type 1, 2, or 3 operation, respectively.

Theorem 3.1.
Let $A \in M_{m\times n}(\mathbb{F})$, and suppose that B is obtained from A by performing an elementary row [column] operation. Then there exists an $m \times m$ [$n \times n$] 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 $m \times m$ [$n \times n$] matrix, then EA [AE] is the matrix obtained from A by performing the same elementary row [column] operation as that which produces E from I_m [I_n].

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 3.3.
If $A \in M_{m\times n}(\mathbb{F})$, we define the **rank** of A, denoted $\operatorname{rank}(A)$, to be the rank of the linear transformation $L_A : \mathbb{F}^n \to \mathbb{F}^m$.

Theorem 3.3.
Let $T : V \to W$ be a linear transformation between finite-dimensional vector spaces, and let β and γ be ordered bases for V and W, respectively. Then $\operatorname{rank}(T) = \operatorname{rank}\left([T]_\beta^\gamma\right)$.

Theorem 3.4.
Let A be an $m \times n$ matrix. if P and Q are invertible $m \times m$ and $n \times n$ matrices, respectively, then

1. $\operatorname{rank}(AQ) = \operatorname{rank}(A)$,

2. $\operatorname{rank}(PA) = \operatorname{rank}(A)$,
and therefore

3. $\operatorname{rank}(PAQ) = \operatorname{rank}(A)$.

Corollary 3.1.
Elementary row and column operations on a matrix are rank preserving.

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 $m \times n$ 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 the zero matrices. Thus $D_{ii} = 1$ for $i \leq r$ and $D_{ij} = 0$ otherwise.

Corollary 3.2.
Let A be an $m \times n$ matrix of rank r. Then there exist invertible matrices B and C of sizes $m \times m$ and $n \times n$, respectively, such that $D = BAC$, where

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

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

Corollary 3.3.
Let A be an $m \times n$ matrix. Then

1. $\operatorname{rank}(A^t) = \operatorname{rank}(A)$.

2. 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.

3. The rows and columns of any matrix generate subspaces of the same dimension, numerically equal to the rank of the matrix.

Corollary 3.4.
Every invertible matrix is a product of elementary matrices.

Theorem 3.7.
Let $T : V \to W$ and $U : W \to Z$ 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

1. $\operatorname{rank}(UT) \leq \operatorname{rank}(U)$.

2. $\operatorname{rank}(UT) \leq \operatorname{rank}(T)$.

3. $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$.

4. $\operatorname{rank}(AB) \leq \operatorname{rank}(B)$.

Definition 3.4.
Let A and B be $m \times n$ and $m \times p$ matrices, respectively. By the **augmented matrix** $(A|B)$, we mean the $m \times (n \times p)$ 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.

3.3 Systems of Linear Equations – Theoretical Aspects

Definition 3.5.
The system of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned} \tag{S}$$

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

The $m \times n$ matrix

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

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

If we let

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

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

$$Ax = b.$$

To exploit the results that we have developed, we often consider a system of linear equations as a single matrix equation.

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

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

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**.

Definition 3.6.
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**.

Any homogeneous system has at least one solution, namely, the zero vector.

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

Corollary 3.5.
If $m < n$, the system $Ax = 0$ has a nonzero solution.

Definition 3.7.
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 $\mathsf{K_H}$ be the solution set of the corresponding homogeneous system $Ax = 0$. Then for any solution s to $Ax = b$

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

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.

Definition 3.8.
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 $\text{rank}\,(A) = \text{rank}\,(A|b)$.

Definition 3.9. Consider a system of linear equations

$$\begin{aligned} a_{11}p_1 + a_{12}p_2 + \cdots + a_{1m}p_m &= p_1 \\ a_{21}p_1 + a_{22}p_2 + \cdots + a_{2m}p_m &= p_2 \\ &\cdots \\ a_{n1}p_1 + a_{n2}p_2 + \cdots + a_{nm}p_m &= p_m \end{aligned}$$

This system can be written as $Ap = p$, where

$$p = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{pmatrix}$$

and A is the coefficient matrix of the system. In this context, A is called the **input-ouput (or consumption) matrix**, and $Ap = p$ is called the **equilibrium condition**.

For vectors $b = (b_1, b_2, \ldots, b_n)$ and $c = (c_1, c_2, \ldots, c_n)$ in $\mathbb{R}^n$, we use the notation $b \geq c$ [$b > c$] to mean $b_i \geq c_i$ [$b_i > c_i$] for all i. The vector b is called **nonnegative** [**positive**] if $b \geq 0$ [$b > 0$].

Theorem 3.12.
Let A be an $n \times n$ input-output matrix having the form

$$A = \begin{pmatrix} B & C \\ D & E \end{pmatrix},$$

where D is a $1 \times (n-1)$ positive vector and C is an $(n-1) \times 1$ positive vector. Then $(I - A)x = 0$ has a one-dimensional solution set that is generated by a nonnegative vector.

3.4 Systems of Linear Equations – Computational Aspects

Definition 3.10.
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 $m \times n$ matrix. Then the system $(CA)x = Cb$ is equivalent to $Ax = b$.

a

Corollary 3.6.
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 3.11.
A matrix is said to be in **reduced row echelon form** if the following three conditions are satisfied.

1. Any row containing a nonzero entry precedes any row in which all the entries are zero (if any).

2. The first nonzero entry in each row is the only nonzero entry in its column.

3. 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.

Definition 3.12.
The following procedure 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*, 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 in the preceding row.

2. In the *backward pass* or *back-substitution*, 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.

Definition 3.13. A solution to a system of equations of the form

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

where r is the number of nonzero solutions in A' ($r \leq m$), is called a **general solution** of the system $Ax = b$. It expresses an arbitrary solution s of $Ax = b$ in terms of $n - r$ parameters.

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

1. $\text{rank}\,(A) = r$.

2. 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 s_0 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

1. The number of nonzero rows in B is r.

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

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

4. 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 3.7.
The reduced row echelon form of a matrix is unique.

Chapter 4

Determinants

4.1 Determinants of Order 2

Definition 4.1.
If

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

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

Theorem 4.1.
The function $\det : M_{2\times 2}(\mathbb{F}) \to \mathbb{F}$ 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 $\mathbb{F}^2$ 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 $A \in M_{2\times 2}(\mathbb{F})$. Then the determinant of A is nonzero if and only if A is invertible. Moreover, if A is invertible, then

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

Definition 4.2.
By the **angle** between two vectors in $\mathbb{R}^2$, we mean the angle with measure θ ($0 \leq \theta < \pi$) that is formed by the vectors having the same magnitude and direction as the given vectors by emanating from the origin.

Definition 4.3.
If $\beta = \{u, v\}$ is an ordered basis for $\mathbb{R}^2$, we define the **orientation** of β to be the real number

$$O\begin{pmatrix} u \\ v \end{pmatrix} = \frac{\det\begin{pmatrix} u \\ v \end{pmatrix}}{\left|\det\begin{pmatrix} u \\ v \end{pmatrix}\right|}$$

(The denominator of this fraction is nonzero by Theorem 4.2).

Definition 4.4.
A coordinate system $\{u, v\}$ is called **right-handed** if u can be rotated in a counterclockwise direction through an angle θ $(0 < \theta < \pi)$ to coincide with v. Otherwise, $\{u, v\}$ is called a **left-handed** system.

Definition 4.5.
Any ordered set $\{u, v\}$ in $\mathbb{R}^2$ determines a parallelogram in the following manner. Regarding u and v as arrows emanating from the origin of $\mathbb{R}^2$, we call the parallelogram having u and v as adjacent sides the **parallelogram determined by** u **and** v.

4.2 Determinants of Order n

Notation 4.1.
Given $A \in M_{n\times n}(\mathbb{F})$, for $n \geq 2$, denote the $(n-1)\times(n-1)$ matrix obtained from A by deleting row i and column j by $\tilde{A}_{ij}$. Thus for

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in M_{3\times 3}(\mathbb{R})$$

we have

$$\tilde{A}_{11} = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}, \qquad \tilde{A}_{13} = \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}, \qquad \tilde{A}_{32} = \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix}$$

and for

$$B = \begin{pmatrix} 1 & -1 & 2 & -1 \\ -3 & 4 & 1 & -1 \\ 2 & -5 & -3 & 8 \\ -2 & 6 & -4 & 1 \end{pmatrix}$$

we have

$$\tilde{B}_{23} = \begin{pmatrix} 1 & -1 & -1 \\ 2 & -5 & 8 \\ -2 & 6 & 1 \end{pmatrix} \qquad \text{and} \qquad \tilde{B}_{42} = \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & -1 \\ 2 & -3 & 8 \end{pmatrix}$$

Definition 4.6.
Let $A \in M_{n\times n}(\mathbb{F})$. 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.

Definition 4.7.
Letting

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

denote the cofactor of the row i, column j entry of A, we can express the formula for the determinant of A as

$$\det(A) = A_{11}c_{11} + A_{12}c_{12} + \cdots + A_{1n}c_{1n}.$$

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

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

wherever k is a scalar and u, v and each a_i are row vectors in $\mathbb{F}^n$.

Corollary 4.1.
If $A \in M_{n\times n}(\mathbb{F})$ has a row consisting entirely of zeros, then $\det(A) = 0$.

Lemma 4.1.
Let $B \in M_{n\times n}(\mathbb{F})$, 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 $A \in M_{n\times n}(\mathbb{F})$, then for any integer i ($1 \leq i \leq n$),

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

Corollary 4.2.
If $A \in M_{n\times n}(\mathbb{F})$ has two identical rows, then $\det(A) = 0$.

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

Theorem 4.6.
Let $A \in M_{n\times n}(\mathbb{F})$, 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 4.3.
If $A \in M_{n\times n}(\mathbb{F})$ has rank less than n, then $\det(A) = 0$.

Remark 4.1.
The following rules summarize the effect of an elementary row operation on the determinant of a matrix $A \, inM_{n\times n}(\mathbb{F})$.

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

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

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

Lemma 4.2.
The determinant of an upper triangular matrix is the product of its diagonal entries.

4.3 Properties of Determinants

Remark 4.2.
Because the determinant of the $n \times n$ matrix is 1, we can interpret Remark 4.1 as the following facts about the determinants of elementary matrices.

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

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

3. 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 $A, B \in M_{n\times n}(\mathbb{F})$, $\det(AB) = \det(A) \cdot \det(B)$.

Corollary 4.4.
A matrix $A \in M_{n\times n}(\mathbb{F})$ is invertible if and only if $\det(A) \neq 0$. Furthermore, if A is invertible, then $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

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

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.

Definition 4.8.
It is possible to interpret the determinant of a matrix $A \in M_{n\times n}(\mathbb{R})$ geometrically. 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 are in $\mathbb{R}^2$ and volume in $\mathbb{R}^3$) of the parallelepiped having the vectors $a_1, a_2, \ldots, a_n$ as adjacent sides.

Definition 4.9.
A matrix $M \in M_{n\times n}(\mathbb{C})$ is called **nilpotent** if, for some positive integer k, $M^k = O$, where O is the $n \times n$ zero matrix.

Definition 4.10.
A matrix $M \in M_{n\times n}(\mathbb{C})$ is called **skew-symmetric** if $M^t = -M$.

Definition 4.11.
A matrix $Q \in M_{n\times n}(\mathbb{R})$ is called **orthogonal** if $QQ^t = I$.

Definition 4.12.
A matrix $Q \in M_{n\times n}(\mathbb{C})$ is called **unitary** if $QQ^* = I$, where $Q^* = \overline{Q^t}$.

Definition 4.13.
A matrix $A \in M_{n\times n}(\mathbb{F})$ is called **lower triangular** if $A_{ij} = 0$ for $1 \le i < j \le n$.

Definition 4.14.
A matrix of the form

$$\begin{pmatrix} 1 & c_0 & c_0^2 & \dots & c_0^n \\ 1 & c_1 & c_1^2 & \dots & c_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & c_n & c_n^2 & \dots & c_n^n \end{pmatrix}$$

is called a **Vandermonde matrix**.

Definition 4.15.
Let $A \in M_{n\times n}(\mathbb{F})$ be nonzero. For any m $(1 \leq m \leq n)$, and $m \times m$ **submatrix** is obtained by deleting any $n - m$ rows and any $n - m$ columns of A.

Definition 4.16.
The **classical adjoint** of a square matrix A is the transpose of the matrix whose ij-entry is the ij-cofactor of A.

Definition 4.17.
Let $y_1, y_2, \dots, y_n$ be linearly independent function in $\mathbb{C}^\infty$. For each $y \in \mathbb{C}^\infty$, define $T(y) \in \mathbb{C}^\infty$ by

$$[T(y)](t) = \det \begin{pmatrix} y(t) & y_1(t) & y_2(t) & \dots & y_n(t) \\ y'(t) & y_1'(t) & y_2'(t) & \dots & y_n'(t) \\ \vdots & \vdots & \vdots & & \vdots \\ y^{(n)}(t) & y_1^{(n)}(t) & y_2^{(n)}(t) & \dots & y_n^{(n)}(t) \end{pmatrix}$$

The preceding determinant is called the **Wronskian** of $y, y_1, \dots, y_n$.

4.4 Summary – Important Facts about Determinants

Definition 4.18.
The **determinant** of an $n \times n$ matrix A having entries from a field $\mathbb{F}$ is a scalar in $\mathbb{F}$, denoted by $\det(A)$ or $|A|$, and can be computed in the following manner:

1. If A is 1×1, then $\det(A) = A_{11}$, the single entry of A.

2. If A is 2×2, then $\det(A) = A_{11}A_{22} - A_{12}A_{21}$.

3. If A is $n \times n$ for $n > 2$, then

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

(if the determinant is evaluated by the entries of row i of A) or

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

(if the determinant is evaluated by the entries of column j of A), where $\tilde{A}_{ij}$ is the $(n-1)\times(n-1)$ matrix obtained by deleting row i and column j from A.

In the formulas above, the scalar $(-1)^{i+j}\det(\tilde{A}_{ij})$ is called the **cofactor** of the row i column j of A.

Definition 4.19 (Properties of the Determinant).

1. If B is a matrix obtained by interchanging any two rows or interchanging any two columns of an $n\times n$ matrix A, then $\det(B) = -\det(A)$.
2. If B is a matrix obtained by multiplying each entry of some row or column of an $n\times n$ matrix A by a scalar k, then $\det(B) = k\cdot\det(A)$.
3. If B is a matrix obtained from an $n\times n$ matrix A by adding a multiple of row i to row j or a multiple of column i to column j for $i\neq j$, then $\det(B) = \det(A)$.

4.5 A Characterization of the Determinant

Definition 4.20.
A function $\delta : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ is called an ***n*-linear function** if it is a linear function of each row of an $n\times n$ matrix when the remaining $n-1$ rows are held fixed, that is, δ is n-linear if, for every $r = 1, 2, \ldots, n$, we have

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

whenever k is a scalar and u, v and each a_i are vectors in $\mathbb{F}^n$.

Definition 4.21.
An n-linear function $\delta : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ is called **alternating** if, for each $A\ inM_{n\times n}(\mathbb{F})$, we have $\delta(A) = 0$ whenever two adjacent rows of A are identical.

Theorem 4.10.
Let $\delta : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ be an alternating n-linear function.

1. If $A \in M_{n\times n}(\mathbb{F})$ and B is a matrix obtained from A by interchanging any two rows of A, then $\delta(B) = -\delta(A)$.
2. If $A \in M_{n\times n}(\mathbb{F})$ has two identical rows, then $\delta(A) = 0$.

Corollary 4.5.
Let $\delta : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ be an alternating n-linear function. If B is a matrix obtained from $A \in M_{n\times n}(\mathbb{F})$ by adding a multiple of some row of A to another row, then $\delta(B) = \delta(A)$.

Corollary 4.6.
Let $\delta : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ be an alternating n-linear function. if $M \in M_{n\times n}(\mathbb{F})$ has rank less than n, then $\delta(M) = 0$.

Corollary 4.7.
Let $\delta : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ be an alternating n-linear function, and let E_1, E_2 and E_3 in $M_{n\times n}(\mathbb{F})$ be elementary matrices of types 1, 2, and 3, respectively. Suppose that E_2 is obtained by multiplying some row of I by the nonzero scalar k. Then $\delta(E_1) = -\delta(I)$, $\delta(E_2) = k \cdot \delta(I)$, and $\delta(E_3) = \delta(I)$.

Theorem 4.11.
Let $\delta : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ be an alternating n-linear function such that $\delta(I) = 1$. For any $A, B \in M_{n\times n}(\mathbb{F})$, we have $\delta(AB) = \delta(A) \cdot \delta(B)$.

Theorem 4.12.
If $\delta : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ is an alternating n-linear function such that $\delta(I) = 1$, then $\delta(A) = \det(A)$ for every $A \in M_{n\times n}(\mathbb{F})$.

Chapter 5

Diagonalization

5.1 Eigenvalues and Eigenvectors

5.2 Diagonalizability

5.3 Matrix Limits and Markov Chains

5.4 Invariant Subspaces and the Cayley-Hamilton Theorem

Chapter 6

Inner Product Spaces

6.1 Inner Products and Norms

6.2 The Gram-Schmidt Orthogonalization Process and Orthogonal Complements

6.3 The Adjoint of a Linear Operator

6.4 Normal and Self-Adjoing Operators

6.5 Unitary and Orthogonal Operators and Their Matrices

6.6 Orthogonal Projections and the Spectral Theorem

6.7 The Singular Value Decomposition and the Pseudoinverse

6.8 Bilinear and Quadratic Forms

6.9 Einstein's Special Theory of Relativity

6.10 Conditioning and the Rayleigh Quotient

6.11 The Geometry of Orthogonal Operators

Chapter 7

Canonical Forms

7.1 The Jordan Canonical Form I

7.2 The Jordan Canonical Form II

7.3 The Minimal Polynomial

7.4 The rational Canonical Form