A Short Course on
LINEAR ALGEBRA

M.Thamban Nair

Department of Mathematics
Indian Institute of Technology Madras

Contents

Preface

This is based on an elective course that the author gave to B.Tech. students of IIT Madras two times (2003-2004). During the preparation of these notes, the author benefitted from the interactions that he had with two of his colleagues, Professor S.H.Kulkarni and Professor Arindama Singh who were also co-teachers for the courses.

The aim of the course is to introduce basics of Linear Algebra and some topics in Numerical Linear Algebra and their applications.

December 2003 M. T. Nair

Present Edition

The present edition is meant for the course MA2031: "Linear Algebra for Engineers", prepared by omitting two chapters related to numerical analysis. Also, the title is changed from "A Short Course on Linear Algebra and its Applications" to *A Short Course on Linear Algebra.*

January, 2020 M. T. Nair

1

Vector Spaces

1.1 Motivation

The notion of a *vector space* is an abstraction of the familiar set of vectors in two or three dimensional Euclidian space. For example, let $\vec{x} = (x_1, x_2)$ and $\vec{y} = (y_1, y_2)$ be two vectors in the plane $\mathbb{R}^2$. Then we have the notion of addition of these vectors so as to get a new vector denoted by $\vec{x} + \vec{y}$, and it is defined by

$$\vec{x} + \vec{y} = (x_1 + y_1, x_2 + y_2).$$

This addition has an obvious geometric meaning: If O is the coordinate origin, and if P and Q are points in $\mathbb{R}^2$ representing the vectors $\vec{x}$ and $\vec{y}$ respectively, then the vector $\vec{x} + \vec{y}$ is represented by a point R in such way that OR is the diagonal of the parallelogram for which OP and OQ are adjacent sides.

Also, if α is a positive real number, then the multiplication of $\vec{x}$ by α is defined by

$$\alpha\vec{x} = (\alpha x_1, \alpha x_2).$$

Geometrically, the vector $\vec{\alpha x}$ is an elongated or contracted form of $\vec{x}$ in the direction of $\vec{x}$. Similarly, we can define $\alpha\vec{x}$ with a negative real number α, so that $\alpha\vec{x}$ represents in the negative direction. Representing the coordinate-origin by $\vec{0}$, and $-\vec{x} := (-1)\vec{x}$, we see that

$$\vec{x} + \vec{0} = \vec{x}, \qquad \vec{x} + (-\vec{x}) = \vec{0}.$$

We may denote the sum $\vec{x} + (-\vec{y})$ by $\vec{x} - \vec{y}$.

Now, abstracting the above properties of vectors in the plane, we define the notion of a vector space.

We shall denote by $\mathbb{F}$ the field of real numbers or the field of complex numbers. If special emphasis is required, then the fields of real numbers and complex numbers will be denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively.

1.2 Definition and Some Basic Properties

***Definition* 1.1 (Vector space)** A *vector space* over $\mathbb{F}$ is a set V together with two operations called (i) *addition* which associates each pair (x, y) of elements in V a unique element in V denoted by $x+y$, and (ii) *scalar multiplication* which associates each pair (α, x) with $\alpha \in \mathbb{F}$ and $x \in V$, a unique element in V denoted by αx, so that these operations satisfy the following axioms:

(a) $x + y = y + x \quad \forall\, x, y \in V$.

(b) $(x + y) + z = x + (y + z) \quad \forall\, x, y, z \in V$.

(c) $\exists\, \theta \in V$ such that $x + \theta = x \quad \forall\, x \in V$.

(d) $\forall\, x \in V$, $\exists\, \tilde{x} \in V$ such that $x + \tilde{x} = \theta$.

(e) $\alpha(x + y) = \alpha x + \alpha y \quad \forall\, \alpha \in \mathbb{F}, \forall\, x, y \in V$.

(f) $(\alpha + \beta)x = \alpha x + \beta x \quad \forall\, \alpha, \beta \in \mathbb{F}, \forall\, x \in V$.

(g) $(\alpha\beta)x = \alpha(\beta x) \quad \forall\, \alpha, \beta \in \mathbb{F}, \forall\, x \in V$.

(h) $1x = x \quad \forall\, x \in V$.

Elements of a vector space are called **vectors**, and elements of the field $\mathbb{F}$ (over which the vector space is defined) are often called **scalars**.

Proposition 1.1 *Let V be a vector space, and θ_1 and θ_2 in V be such that*

$$x + \theta_1 = x \quad \textit{and} \quad x + \theta_2 = x \quad \forall x \in X.$$

Then $\theta_1 = \theta_2$.

Proof. Using the hypothesis and axioms (a) and (c), we have

$$\theta_2 = \theta_2 + \theta_1 = \theta_1 + \theta_2 = \theta_1.$$

This completes the proof. ∎

By the above proposition, we can assert that there is exactly one element $\theta \in V$ such that $x + \theta = V$ for all $x \in V$.

***Definition* 1.2 (zero element)** Let V be a vector space. The unique element $\theta \in V$ such that $x + \theta = x$ for all $x \in V$ is called the *zero element* or simply, the *zero* in V.

Notation: The zero element in a vector space as well as the zero in the scalar field are often denoted by the same symbol 0.

Exercise 1.1 Let V be a vector space. For $x, y \in V$, show that $x + y = x$ implies $y = \theta$. ♦

Proposition 1.2 *Let V be a vector space. For $x \in V$, let x' and x'' be in V such that*

$$x + x' = \theta \quad \textit{and} \quad x + x'' = \theta.$$

Then $x' = x''$.

Proof. By hypothesis and using the axioms (a), (b), (c), it follows that

$$x' = x' + \theta = x' + (x + x'') = (x' + x) + x'' = \theta + x'' = x''.$$

This completes the proof. ∎

The above proposition shows that, for every $x \in V$, there exists only one element $\tilde{x} \in V$ such that $x + \tilde{x} = \theta$.

***Definition* 1.3 (additive inverse)** Let V be a vector space. For each $x \in V$, the unique element $\tilde{x} \in V$ such that $x + \tilde{x} = \theta$ is called the *additive inverse* of x.

Notation: For x in a vector space, the unique element $\tilde{x}$ which satisfies $x + \tilde{x} = \theta$ is denoted by $-x$.

Proposition 1.3 *Let V be a vector space. Then, for all $x \in V$,*

$$0x = \theta \quad \textit{and} \quad (-1)x = -x.$$

Proof. Let $x \in V$. Since $0x = (0 + 0)x = 0x + 0x$, it follows that $0x = \theta$. Thus, (i) is proved. Now, $x + (-1)x = [1 + (-1)]x = 0.x = \theta$ so that, by the uniqueness of the additive inverse of x, it follows that $(-1)x = -x$. ∎

Notation: For x, y in a vector space, the expression $x + (-y)$ is denoted by $x - y$.

We observe that a vector space V, by definition, cannot be an empty set. It contains at least one element, viz., the zero element.

If a vector space V contains at least one nonzero element, then it contains infinitely many nonzero elements: If x is a nonzero element in V, and if α, β are scalars such that $\alpha \neq \beta$, then $\alpha x \neq \beta x$ (see Exercise 1.2 below). This is a consequence of axiom (h).

Exercise 1.2 Show that, if $x \in V$ and $x \neq 0$, then $\alpha x \neq \beta x$ for every $\alpha, \beta \in \mathbb{F}$ with $\alpha \neq \beta$. ♦

Unless otherwise specified, we always assume that the vector space under discussion is *non-trivial* , i.e., it contains at least one nonzero element.

1.3 Examples of Vector Spaces

EXAMPLE 1.1 (Space $\mathbb{F}^n$) Consider the set $\mathbb{F}^n$ of all n–tuples of scalars, i.e.,

$$\mathbb{F}^n := \{x = (\alpha_1, \ldots, \alpha_n) : \alpha_i \in \mathbb{F},\ i = 1, \ldots, n\}.$$

For $x = (\alpha_1, \ldots, \alpha_n)$, $y = (\beta_1, \ldots, \beta_n)$ in $\mathbb{F}^n$, and $\alpha \in \mathbb{F}$, define the addition and scalar multiplication coordinate-wise as

$$x + y = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n), \quad \alpha x = (\alpha\alpha_1, \ldots, \alpha\alpha_n).$$

Then it can be seen that $\mathbb{F}^n$ is a vector space with zero element $\theta := (0, \ldots, 0)$ and additive inverse of $x = (\alpha_1, \ldots, \alpha_n)$ as $-x = (-\alpha_1, \ldots, -\alpha_n)$.

EXAMPLE 1.2 (Space $\mathcal{P}_n$) For $n \in \{0, 1, 2, \ldots\}$, let $\mathcal{P}_n$ be the set of all polynomials of degree at most n, with coefficients in $\mathbb{F}$, i.e., $x \in \mathcal{P}_n$ if and only if x is of the form

$$x = a_0 + a_1 t + \ldots + a_n t^n$$

for some scalars $a_0, a_1 \ldots, a_n$. Then $\mathcal{P}_n$ is a vector space with addition and scalar multiplication defined as follows:

For $x = a_0 + a_1 t + \ldots a_n t^n,\ y = b_0 + b_1 t + \ldots + b_n t^n$ in $\mathcal{P}_n$ and $\alpha \in \mathbb{F}$,

$$x + y = (a_0 + b_0) + (a_1 + b_1)t + \ldots + (a_n + b_n)t^n,$$

$$\alpha x = \alpha a_0 + \alpha a_1 t + \ldots + \alpha a_n t^n.$$

The zero polynomial, i.e., the polynomial with all its coefficients zero, is the zero element of the space, and

$$-x = -a_0 - a_1 t - \ldots - a_n t^n.$$

EXAMPLE 1.3 (Space $\mathcal{P}$) Let $\mathcal{P}$ be the set of all polynomials with coefficients in $\mathbb{F}$, i.e., $x \in \mathcal{P}$ if and only if $x \in \mathcal{P}_n$ for some $n \in \{0, 1, 2, \ldots\}$. For $x, y \in \mathcal{P}$, let n, m be such that $x \in \mathcal{P}_n$ and $y \in \mathcal{P}_m$. Then we have $x, y \in \mathcal{P}_k$, where $k = \max\{n, m\}$. Hence we can define $x + y$ and αx for $\alpha \in \mathbb{F}$ as in $\mathcal{P}_k$. With this addition and scalar multiplication, it follows that $\mathcal{P}$ is a vector space.

EXAMPLE 1.4 (Space $\mathbb{F}^{m\times n}$) Let $V = \mathbb{F}^{m\times n}$ be the set of all $m \times n$ matrices with entries in $\mathbb{F}$. If A is a matrix with its ij-th entry a_{ij}, then we shall write $A = [a_{ij}]$. It is seen that V is a vector space with respect to the addition and scalar multiplication defined as follows: For $A = [a_{ij}]$, $B = [b_{ij}]$ in V, and $\alpha \in \mathbb{F}$,

$$A + B := [a_{ij} + b_{ij}], \qquad \alpha A := [\alpha a_{ij}].$$

In this space, $-A = [-a_{ij}]$, and the matrix with all its entries are zeroes is the zero element.

EXAMPLE 1.5 (Space $\underline{\mathbb{F}}^k$) This example is a special case of the last one. For each $k \in \mathbb{N}$, let $\underline{\mathbb{F}}^k$ denotes the set of all column k-vectors, i.e., the set of all $k \times 1$ matrices. Obviously, $\underline{\mathbb{F}}^k$ is a vector space over $\mathbb{F}$. This vector space is in one-one correspondence with $\mathbb{F}^k$. One such correspondence is given by $T : \mathbb{F}^k \to \underline{\mathbb{F}}^k$ defined by

$$T((x_1, \ldots, x_k)) = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_k \end{bmatrix}, \qquad (x_1, \ldots, x_k) \in \mathbb{F}^k.$$

EXAMPLE 1.6 (Sequence space) Let V be the set of all scalar sequences. For (α_n) and (β_n) in V, and $\alpha \in \mathbb{F}$, we define

$$(\alpha_n) + (\beta_n) = (\alpha_n + \beta_n), \quad \alpha(\alpha_n) = (\alpha\alpha_n).$$

With this addition and scalar multiplication, V is a vector space with its zero element as the sequence of zeroes, and $-(\alpha_n) = (-\alpha_n)$.

Exercise 1.3 Verify that the sets considered in Examples 1.1 – 1.6 are indeed vector spaces with respect to the operations defined there. ♦

For the next example the reader may recall the definition of a real valued continuous function: A function $x : \Omega \to \mathbb{R}$, i.e., a real valued function function x defined on a subset Ω of $\mathbb{R}$, is said to be continuous at a point $s_0 \in \Omega$, if for every given $\varepsilon > 0$, it is possible to find a $\delta > 0$, which may depend on s_0 as well as ε, such that

$$s \in \Omega, \quad |s - s_0| < \delta \Rightarrow |x(s) - x(s_0)| < \varepsilon.$$

EXAMPLE 1.7 (Space $C(\Omega)$) Let Ω be a subset of $\mathbb{R}$ and $C(\Omega)$ be the set of all real valued continuous functions defined on Ω. For $x, y \in C(\Omega)$ and $\alpha \in \mathbb{F}$, we define $x + y$ and αx point-wise, i.e.,

$$(x + y)(t) = x(t) + y(t), \quad (\alpha x)(t) = \alpha x(t), \quad t \in \Omega.$$

Then it can be shown that $x + y, \alpha x \in C(\Omega)$, and $C(\Omega)$ is a vector space over $\mathbb{R}$ with zero element as the zero function, and additive inverse of $x \in C(\Omega)$ as the function $-x$ defined by $(-x)(t) = -x(t),\ t \in \Omega$.

NOTATION. If $\Omega = [a, b]$, we shall denote the space $C(\Omega)$ by $C[a, b]$. In case we want to emphasis the scalar field is $\mathbb{R}$, then we shall write $C([a, b], \mathbb{R})$ in place of $C[a, b]$.

EXAMPLE 1.8 (Space $\mathcal{R}[a, b]$) Let $\mathcal{R}[a, b]$ be the set of all real valued Riemann integrable functions on $[a, b]$. From the theory of Riemann integration, it follows that if $x, y \in \mathcal{R}[a, b]$ and $\alpha \in \mathbb{F}$, then $x + y$ and αx defined pointwise belongs to $\mathcal{R}[a, b]$. It is seen that (*Verify*) $\mathcal{R}[a, b]$ is a vector space over $\mathbb{R}$.

EXAMPLE 1.9 (Function space $\mathcal{F}(\Omega, \mathbb{F})$) Let Ω be a nonempty set and $\mathcal{F}(\Omega, \mathbb{F})$ be the set of all functions from Ω into $\mathbb{F}$. For $x,\ y \in \mathcal{F}(\Omega, \mathbb{F})$ and $\alpha \in \mathbb{F}$, let $x + y$ and αx be defined point-wise, i.e.,

$$(x + y)(s) = x(s) + y(s), \quad (\alpha x)(s) = \alpha x(s), \quad s \in \Omega.$$

Let $-x$ and θ be defined by

$$(-x)(s) = -x(s), \quad \theta(s) = 0, \quad s \in S.$$

Then it is easy to see that $\mathcal{F}(\Omega, \mathbb{F})$ is a vector space over $\mathbb{F}$.

For showing that $\mathcal{F}(\Omega, \mathbb{F})$ is a vector space, what one essentially requires is the linear structure on $\mathbb{F}$. Thus, in a similar fashion we can show that if W is a vector space, then $\mathcal{F}(\Omega, W)$, the set of all functions from Ω into W, is a vector space.

Consider the following particular cases of the above example.

(a) Let $\Omega = \{1, \ldots, n\}$. Then it can be seen that the function $T : \mathcal{F}(S, \mathbb{F}) \to \mathbb{F}^n$ defined by

$$T(x) = (x(1), \ldots, x(n)), \quad x \in X,$$

is bijective. We also see that for every x, y in $\mathcal{F}(S, \mathbb{F})$ and $\alpha \in \mathbb{F}$,

$$T(x + y) = T(x) + T(y), \quad T(\alpha x) = \alpha T(x).$$

Such a map is called a *linear transformation* or a *linear operator*. Linear transformations will be considered in more detail in the next chapter. Here we give only its definition.

***Definition* 1.4** Let V and W be vector spaces. Then a function $T : X \to W$ is called a **linear transformation** or a **linear operator** if

$$T(x + y) = T(x) + T(y) \quad \text{and} \quad T(\alpha x) = \alpha T(x)$$

for every $x, y \in V$ and $\alpha \in \mathbb{F}$.

A bijective linear transformation is sometimes called a **linear isomorphism**. Thus, if there is a linear isomorphism $T : V \to W$ between vector spaces V and W, then as far as their linear structures are concerned, they are indistinguishable. Hence we may regard them same, up to a linear isomorphism. Thus, if $\Omega = \{1, \ldots, n\}$, then the vector spaces $\mathcal{F}(\Omega, \mathbb{F})$ and $\mathbb{F}^n$ can be considered as same. With this identification in mind, we shall denote the j-th entry of an element $x = (\alpha_1, \ldots, \alpha_n)$ of $\mathbb{F}^n$ by $x(j)$ for $j = 1, \ldots, n$.

Similarly, if Ω is a set with n elements, say $\Omega = \{s_1, \ldots, s_n\}$, then $\mathcal{F}(\Omega, \mathbb{F})$ can be identified with $\mathbb{F}^n$ by the map

$$x \mapsto (x(s_1), \ldots, x(s_n)), \quad x \in \mathcal{F}(\Omega, \mathbb{F}).$$

(b) Next suppose $\Omega = \mathbb{N}$, the set of all positive integers. Then $\mathcal{F}(\Omega, \mathbb{F})$ can be identified with the set of all scalar sequences. The identification is given by

$$x \mapsto (x(1), x(2), \ldots), \quad x \in \mathcal{F}(\mathbb{N}, \mathbb{F}).$$

With this identification, the n–th entry α_n of a scalar sequence $x = (\alpha_n)$ is also denoted by $x(n)$.

Similarly, if Ω is a denumerable set, say $S = \{s_1, s_2, \ldots\}$, then $\mathcal{F}(S, \mathbb{F})$ can be identified with the set of all scalar sequences by the map

$$x \mapsto (x(s_1), x(s_2), \ldots)\,, \quad x \in \mathcal{F}(\Omega, \mathbb{F}).$$

(c) Consider $\Omega = \{1, \ldots, m\} \times \{1, \ldots, n\}$. Then the resulting vector space $\mathcal{F}(\Omega, \mathbb{F})$ is in one–one correspondence with the set $\mathbb{F}^{m\times n}$ of all $m \times n$ matrices with entries in $\mathbb{F}$. The bijective map, in this case, is

$$x \mapsto [x(i,j)],$$

where $[x(i,j)]$ is the $m \times n$ matrix whose ij–th entry is $x(i,j)$.

Exercise 1.4 Show that the maps considered in (a), (b) and (c) above are linear transformations. ♦

Exercise 1.5 Verify that the sets considered in Examples 1.7 – 1.9 are indeed vector spaces with respect to the operations defined there. ♦

Exercise 1.6 Find a bijective linear transformation between $\mathbb{F}^n$ and $\mathcal{P}_{n-1}$. ♦

EXAMPLE 1.10 Let J be an interval and $\mathcal{P}_n(J)$ be the set of all polynomials of degree at most n considered as functions on J. Thus, $x \in \mathcal{P}_n(J)$ if and only $x : J \to \mathbb{F}$ and there exist scalars $a_0, a_1, \ldots, a_n$ such that

$$x(t) = a_0 + a_1 t + \ldots + a_n t^n, \quad t \in J.$$

Then as in Example 1.2, $\mathcal{P}_n(J)$ is a vector space.

EXAMPLE 1.11 Let J be an interval and $\mathcal{P}(J) = \cup_{n=0}^{\infty} \mathcal{P}_n(J)$. Then as in Example 1.3, $\mathcal{P}(J)$ is a vector space.

NOTATION : If $J = [a,b]$, then we may write $\mathcal{P}_n(J)$ and $\mathcal{P}(J)$ as $\mathcal{P}_n[a,b]$ and $\mathcal{P}[a,b]$ respectively.

EXAMPLE 1.12 (Product space) Let $V_1, \ldots, V_n$ be vector spaces. Then the *cartesian product*

$$V = V_1 \times \cdots \times V_n,$$

the set of all of ordered n-tuples $(x_1, \ldots, x_n)$ with $x_j \in V_j$ for $j \in \{1, \ldots, n\}$, is a vector space with respect to the addition and scalar multiplication defined by

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n),$$

$$\alpha(x_1, \ldots, x_n) := (\alpha x_1, \ldots, \alpha x_n)$$

with zero element $(0, \ldots, 0)$ and additive inverse of $x = (x_1, \ldots, x_n)$ defined by $-x = (-x_1, \ldots, -x_n)$.

This vector space is called the **product space** of $V_1, \ldots V_n$.

As a particular example, the space $\mathbb{F}^n$ can be considered as the product space $V_1 \times \cdots \times V_n$ with $V_j = \mathbb{F}$ for $j = 1, \ldots, n$.

Exercise 1.7 In each of the following, a set is given and some operations are defined. Check whether V is a vector space with these operations:

(i) Let $V = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}$ with addition and scalar multiplication as in $\mathbb{R}^2$.

(ii) Let $V = \{x = (x_1, x_2) \in \mathbb{R}^2 : 2x_1 + 3x_2 = 0\}$ with addition and scalar multiplication as in $\mathbb{R}^2$.

(iii) Let $V = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 = 1\}$ with addition and scalar multiplication as for $\mathbb{R}^2$.

(iv) Let $V = \mathbb{R}^2$, $\mathbb{F} = \mathbb{R}$. For $x = (x_1, x_2)$, $y = (y_1, y_2)$, let $x + y := (x_1 + y_1, x_2 + y_2)$ and for all $\alpha \in \mathbb{R}$,

$$\alpha x := \begin{cases} (0, 0) & \alpha = 0, \\ (\alpha x_1, x_2/\alpha), & \alpha \neq 0. \end{cases}$$

(v) Let $V = \mathbb{C}^2$, $\mathbb{F} = \mathbb{C}$. For $x = (x_1, x_2)$, $y = (y_1, y_2)$, let

$$x + y := (x_1 + 2y_1, x_2 + 3y_2) \quad \text{and} \quad \alpha x := (\alpha x_1, \alpha x_2) \quad \forall \alpha \in \mathbb{C}.$$

(vi) Let $V = \mathbb{R}^2$, $\mathbb{F} = \mathbb{R}$. For $x = (x_1, x_2)$, $y = (y_1, y_2)$, let

$$x + y := (x_1 + y_1, x_2 + y_2) \quad \text{and} \quad \alpha x := (x_1, 0) \quad \forall \alpha \in \mathbb{R}.$$

1.4 Subspace and Span

1.4.1 Subspace

We have seen that

- $\mathcal{P}_n$ which is a subset of the vector space $\mathcal{P}$ is also a vector space,
- $C[a, b]$ which is a subset of the vector space $\mathcal{F}([a, b], \mathbb{R})$ is also a vector space,

- $V = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}$, which is a subset of $\mathbb{R}^2$ is a vector space with respect to the addition and scalar multiplication as in $\mathbb{R}^2$.

- $V = \{x = (x_1, x_2) \in \mathbb{R}^2 : 2x_1 + 3x_2 = 0\}$ which is a subset of $\mathbb{R}^2$ is a vector space with respect to the addition and scalar multiplication as in $\mathbb{R}^2$.

These examples motivate the following definition.

***Definition* 1.5 (Subspace)** Let V_0 be a subset of a vector space V. Then V_0 is called a *subspace* of V if V_0 is a vector space with respect to the operations of addition and scalar multiplication as in V.

Theorem 1.4 *Let V be a vector space, and V_0 be a subset of V. Then V_0 is a subspace of V if and only if for every pair of vectors x, y in V_0 and for every $\alpha \in \mathbb{F}$,*

$$x + y \in V_0 \quad \textit{and} \quad \alpha x \in V_0.$$

Proof. Clearly, if V_0 is a subspace of V, then $x + y \in V_0$ and $\alpha x \in V_0$ for all $x, y \in V_0$ and for all $\alpha \in \mathbb{F}$.

Conversely, suppose that $x + y \in V_0$ and $\alpha x \in V_0$ for all $x, y \in V_0$ and for all $\alpha \in \mathbb{F}$. Then, for any $x \in V_0$,

$$\theta = 0x \in V_0 \quad \text{and} \quad -x = (-1)x \in V_0.$$

Thus, axioms (c) and (d) in the definition of a vector space are satisfied for V_0. All the remaining axioms are trivially satisfied as elements of V_0 are elements of V as well. ∎

EXAMPLE 1.13 The space $\mathcal{P}_n$ is a subspace of $\mathcal{P}_m$ for $n \leq m$.

EXAMPLE 1.14 The space $C[a, b]$ is a subspace of $\mathcal{R}[a, b]$.

EXAMPLE 1.15 (Space $C^k[a, b]$) For $k \in \mathbb{N}$, let $C^k[a, b]$ be the set of all $\mathbb{F}$-valued functions defined on $[a, b]$ such that for each $j \in \{1, \ldots, k\}$, j-th derivative $x^{(j)}$ of x exists and $x^{(j)} \in C[a, b]$. It can be seen that $C^k[a, b]$ is a subspace of $C[a, b]$.

EXAMPLE 1.16 The space $\mathcal{P}[a, b]$ is a subspace of $C^k[a, b]$ for every $k \geq 1$.

Exercise 1.8 Let A be an $m \times n$ matrix of scalars. Show that the set of all $x \in \underline{\mathbb{F}}^n$ which satisfies $Ax = 0$ is a subspace of $\underline{\mathbb{F}}^n$.

EXAMPLE 1.17 Let V be the space of all scalar sequences and

$$\ell^1(\mathbb{N}) := \Big\{x \in V : \sum_{j=1}^{\infty} |x(j)| < \infty\Big\},$$

the set of all absolutely summable sequences. We show that $\ell^1(\mathbb{N})$ is a subspace of $\mathcal{F}(\mathbb{N}, \mathbb{F})$: For $x, y \in \ell^1(\mathbb{N})$ and $\alpha, \beta \in \mathbb{F}$, and $n \in \mathbb{N}$, we have

$$\begin{aligned} \sum_{j=1}^{n} |\alpha x(j) + \beta y(j)| &\leq |\alpha| \sum_{j=1}^{n} |x(j)| + |\beta| \sum_{j=1}^{n} |y(j)| \\ &\leq |\alpha| \sum_{j=1}^{\infty} |x(j)| + |\beta| \sum_{j=1}^{\infty} |y(j)|. \end{aligned}$$

By letting $n \to \infty$, we have $\sum_{j=1}^{\infty} |\alpha x(j) + \beta y(j)| < \infty$ so that $\alpha x + \beta y \in \ell^1(\mathbb{N})$.

EXAMPLE 1.18 For a nonempty set Ω, let

$$\ell^\infty(\Omega) := \big\{x \in \mathcal{F}(S, \mathbb{F}) : \sup_{s \in S} |x(s)| < \infty\big\}.$$

Note that $\ell^\infty(\Omega)$ is the set of all bounded functions on S. Thus, $x \in \ell^\infty(\Omega)$ if and only there exists $M_x > 0$ such that $|x(s)| \leq M_x$ for all $s \in S$. We show that $\ell^\infty(\Omega)$ is a subspace of $\mathcal{F}(S, \mathbb{F})$: To see this, let $x, y \in B\ell^\infty(\Omega)$ and $\alpha, \beta \in \mathbb{F}$. Suppose $M_x > 0$, $M_y > 0$ such that $|x(s)| \leq M_x$ and $|y(s)| \leq M_y$ for all $s \in S$. Then,

$$|\alpha x(s) + \beta y(s)| \leq |\alpha| M_x + |\beta| M_y \quad \forall\, s \in S.$$

Thus, $\sup_{s \in S} |\alpha x(s) + \beta y(s)| < \infty$, and hence $\alpha x + \beta y \in \ell^\infty$.

In this example, if S is a finite set, then $\ell^\infty(\Omega) = \mathcal{F}(S, \mathbb{F})$. But, if S is an infinite set, then $\ell^\infty(\Omega)$ is a proper subspace of $\mathcal{F}(S, \mathbb{F})$. To see this, let $\{s_1, s_2, \ldots\}$ be a denumerable subset of S, and let $x \in \mathcal{F}(S, \mathbb{F})$ be defined by $x(s_j) = j$ for all $j \in \mathbb{N}$. Then we see that x does not belong to $\ell^\infty(\Omega)$.

EXAMPLE 1.19 Let V be the space of all scalar sequences, and

$$c_{00} := \{(\alpha_1, \alpha_2, \ldots) \in V : \exists\, k \in \mathbb{N} \,\text{such that}\, \alpha_j = 0 \;\forall\, j \geq k\}.$$

Then it is seen that c_{00} is a subspace of V.

EXAMPLE 1.20 The set c_{00} introduced in Example 1.19 is a subspace of $\ell^1(\mathbb{N})$, and the sets

$$\begin{aligned} c_0 &:= \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : x(n) \to 0 \text{ as } n \to \infty\}, \\ c &:= \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : (x(n)) \;\text{ converges }\} \end{aligned}$$

are subspaces of $\ell^\infty(\mathbb{N})$. We observe that

$$c_{00} \subseteq \ell^1(\mathbb{N}) \subseteq c_0 \subseteq c \subseteq \ell^\infty(\mathbb{N}).$$

The above inclusions are, in fact, proper. To see this, let x, y, u, v in $\mathcal{F}(\mathbb{N}, \mathbb{F})$ be defined

$$x(j) = (-1)^j, \qquad y(j) = \frac{j}{j+1}, \qquad u(j) = \frac{1}{j}, \qquad v(j) = \frac{1}{j^2},$$

for $j \in \mathbb{N}$. Then we see that

$$x \in \ell^\infty(\mathbb{N}) \setminus c, \qquad y \in c \setminus c_0,$$

$$u \in c_0 \setminus \ell^1(\mathbb{N}), \qquad v \in \ell^1(\mathbb{N}) \setminus c_{00}.$$

Exercise 1.9 Suppose V_0 is a subspace of a vector space V, and V_1 is a subspace of V_0. Then show that V_1 is a subspace of V. ♦

Exercise 1.10 Show that

$$V_0 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0,\; x_1 + 2x_2 + 3x_3 = 0\}$$

is a subspace of $\mathbb{R}^3$. Observe that V_0 is the intersection of

$$V_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$$

and

$$V_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + 2x_2 + 3x_3 = 0\}.$$

♦

Theorem 1.5 *Suppose V_1 and V_2 are subspaces of a vector space. Then $V_1 \cap V_2$ is a subspace of V.*

Proof. Suppose $x, y \in V_1 \cap V_2$ and $\alpha \in \mathbb{F}$. Then $x, y \in V_1$ and $x, y \in V_2$. Since V_1 and V_2 are subspaces, it follows that $\alpha x, x+y \in V_1$ and $\alpha x, x+y \in V_2$ so that $\alpha x, x+y \in V_1 \cap V_2$. Thus, by Theorem 1.4, $V_1 \cap V_2$ is a subspace. ∎

Is union of two subspaces a subspace? Not necessarily:To see this consider the subspaces

$$V_1 := \{x = (x_1, x_2) : x_2 = x_1\}, \qquad V_2 := \{x = (x_1, x_2) : x_2 = 2x_1\}$$

of the space $\mathbb{R}^2$. Note that $x = (1,1) \in V_1$ and $y = (1,2) \in V_2$, but $x + y = (2,3) \notin V_1 \cup V_2$. Hence $V_1 \cup V_2$ is not a subspace of $\mathbb{R}^2$.

Theorem 1.6 *Let V_1 and V_2 be subspaces of a vector space. Then $V_1 \cup V_2$ is a subspace if and only if either $V_1 \subseteq V_2$ or $V_2 \subseteq V_1$.*

Proof. Suppose either $V_1 \subseteq V_2$ or $V_2 \subseteq V_1$. Then either $V_1 \cup V_2 = V_2$ or $V_1 \cup V_2 = V_1$; in both the cases $V_1 \cup V_2$ is a subspace. Conversely, suppose $V_1 \cup V_2$ is a subspace. Assume for a moment that $V_1 \not\subseteq V_2$ and $V_2 \not\subseteq V_1$. Then, there exists $x, y \in V$ such that $x \in V_1 \setminus V_2$ and $y \in V_2 \setminus V_1$. Now, $x, y \in V_1 \cup V_2$. Since $V_1 \cup V_2$ is a subspace, $x + y \in V_1 \cup V_2$. This implies that either $x + y \in V_1$ or $x + y \in V_2$, which in turn implies $y \in V_1$ or $x \in V_2$. This is a contradiction. Hence $V_1 \not\subseteq V_2$ and $V_2 \not\subseteq V_1$ is not possible. Hence, $V_1 \subseteq V_2$ or $V_2 \subseteq V_1$. ∎

Exercise 1.11 Suppose Λ is a set, and for each $\lambda \in \Lambda$ let V_λ be a subspace of a vector space V. Then $\cap_{\lambda \in \Lambda} V_\lambda$ is a subspace of V. ♦

Exercise 1.12 In each of the following vector space V, see if the subset V_0 is a subspace of V:

(i) $V = \mathbb{R}^2$ and $V_0 = \{(x_1, x_2) : x_2 = 2x_1 - 1\}$.
(ii) $V = \mathbb{R}^3$ and $V_0 = \{(x_1, x_2, x_3) : 2x_1 - x_2 - x_3 = 0\}$.
(iii) $V = C[-1, 1]$ and $V_0 = \{f \in V : f \text{ is an odd function}\}$.
(iv) $V = C[0, 1]$ and $V_0 = \{f \in V : f(t) \geq 0 \, \forall t \in [0, 1]\}$.
(v) $V = \mathcal{P}_3$ and $V_0 = \{a_0 + a_1 t + a_2 t^2 + a_3 t^3 : a_0 = 0\}$.
(vi) $V = \mathcal{P}_3$ and $V_0 = \{a_0 + a_1 t + a_2 t^2 + a_3 t^3 : a_2 = 0\}$.

Exercise 1.13 Prove that the only proper subspaces of $\mathbb{R}^2$ are the straight lines passing through the origin.

Exercise 1.14 Let V be a vector space and $u_1, \ldots, u_n$ are in V. Show that

$$V_0 := \{\alpha_1 u_1 + \ldots + \alpha_n u_n : \alpha_i \in \mathbb{F},\, i = 1, \ldots, n\}$$

is a subspace of V. ♦

1.4.2 Linear Combination and Span

***Definition* 1.6 (Linear combination)** Let V be a vector space and $u_1, \ldots, u_n$ are in V. Then, by a *linear combination* of $u_1, \ldots, u_n$, we mean an element in V of the form $\alpha_1 u_1 + \cdots + \alpha_n u_n$ with $\alpha_j \in \mathbb{F}$, $j = 1, \ldots, n$.

***Definition* 1.7 (Span)** Let S be a subset of V. Then the set of all linear combinations of elements of S is called the *span* of S, and is denoted by $\operatorname{span} S$.

Thus, for $S \subseteq V$, $x \in \operatorname{span} S$ if and only if there exists $x_1, \ldots, x_n$ in S and scalars $\alpha_1, \ldots, \alpha_n$ such that $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$.

As a convention, span of the empty set is taken to be the singleton set $\{0\}$.

Remember! By a linear combination, we always mean a linear combination of a *finite number* of elements in the space. An expression of the form $\alpha_1 x_1 + \alpha_n x_n + \cdots$ with $x_1, x_2, \ldots$ in V and $\alpha_1, \alpha_2, \ldots$ in $\mathbb{F}$ has no meaning in a vector space, unless there is some additional structure which allows such expression.

Theorem 1.7 *Let V be a vector space, and $S \subseteq V$. Then $\operatorname{span} S$ is a subspace of V, and $\operatorname{span} S$ is the smallest subspace containing S.*

Proof. The fact that $\operatorname{span} S$ is a subspace of V is left as an exercise (Hint: Use Theorem 1.4). It remains to show that $\operatorname{span} S$ is the smallest subspace containing S. For this, consider a subspace V_0 of V such that $S \subset V_0$. Then, as V_0 is a subspace, every linear combination of members of S has to be in V_0, that is, $\operatorname{span} S \subset V_0$. This completes the proof. ∎

Exercise 1.15 Let S be a subset of a vector space V. Show that S is a subspace if and only if $S = \operatorname{span} S$.

Exercise 1.16 Let V be a vector space. Show that the the following hold.

(i) Let S be a subset of V. Then

$$\operatorname{span} S = \bigcap \{Y : Y \text{ is a subspace of } V \text{ containing } S\}.$$

(ii) Suppose V_0 is a subspace of V and $x_0 \in V \setminus V_0$. Then for every $x \in \operatorname{span}\{x_0; V_0\} := \operatorname{span}(\{x_0\} \cup V_0)$, there exist a unique $\alpha \in \mathbb{F}$, $y \in V_0$ such that $x = \alpha x_0 + y$. ♦

1.4.3 Examples

EXAMPLE 1.21 Let $V = \mathbb{F}^n$ and for each $j \in \{1, \ldots, n\}$, let $e_j \in \mathbb{F}^n$ be the element with its j-th coordinate 1 and all other coordinates 0's. Then $\mathbb{F}^n$ is the span of $\{e_1, \ldots, e_n\}$.

EXAMPLE 1.22 For $1 \leq k < n$, let

$$V_0 := \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n : \alpha_j = 0, j = k+1, \ldots, n\}.$$

Then it is seen that V_0 is the span of $\{e_1, \ldots, e_k\}$, where $e_j(i) = \delta_{ij}$ with $j = 1, \ldots, k;\ i = 1, \ldots, n$.

EXAMPLE 1.23 Let $V = \mathcal{P}$, and $u_j(t) = t^{j-1}$ for $t \in [a, b]$, $j \in \mathbb{N}$. Then $\mathcal{P}_n$ is the span of $\{u_1, \ldots, u_{n+1}\}$, and $\mathcal{P} = \operatorname{span}\{u_1, u_2, \ldots\}$.

NOTATION: For $(i, j) \in \mathbb{N} \times \mathbb{N}$, let

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Thus, in the above example, the i-th coordinate of e_j is δ_{ij} for $i, j = 1, \ldots, n$, i.e.,

$$e_i = (\delta_{i1}, \delta_{i2}, \ldots, \delta_{in}), \qquad i \in \{1, \ldots, n\}.$$

EXAMPLE 1.24 The space c_{00} is the span of $\{e_1, e_2, \ldots\}$, where $e_j(i) = \delta_{ij}$ with $i, j \in \mathbb{N}$.

Exercise 1.17 Let $u_j(t) = t^{j-1}$, $j \in \mathbb{N}$. Show that span of $\{u_1, \ldots, u_{n+1}\}$ is $\mathcal{P}_n$, and span of $\{u_1, u_2, \ldots\}$ is $\mathcal{P}$. ♦

Exercise 1.18 Let $u_1(t) = 1$, and for $j = 2, 3, \ldots$, let $u_j(t) = 1 + t + \ldots + t^j$. Show that span of $\{u_1, \ldots, u_n\}$ is $\mathcal{P}_n$, and span of $\{u_1, u_2, \ldots\}$ is $\mathcal{P}$. ♦

1.4.4 Sums of Subsets and Subspaces

***Definition* 1.8 (Sum of subsets)** Let V be a vector space, $x \in V$, and E, E_1, E_2 be subsets of V. Then we define the following:

$$x + E := \{x + u : u \in E\},$$

$$E_1 + E_2 := \{x_1 + x_2 : x_1 \in E_1,\ x_2 \in E_2\}.$$

The set $E_1 + E_2$ is called the *sum of the subsets* E_1 and E_2.

Theorem 1.8 *Suppose V_1 and V_2 are subspaces of V. Then $V_1 + V_2$ is a subspace of V. In fact,*

$$V_1 + V_2 = \mathrm{span}\,(V_1 \cup V_2).$$

Proof. Let $x, y \in V_1 + V_2$ and $\alpha \in \mathbb{F}$. Then, there exists $x_1, y_1 \in V_1$ and $x_2, y_2 \in V_2$ such that $x = x_1 + y_1$, $y = y_1 + y_2$. Hence,

$$x + y = (x_1 + y_1) + (y_1 + y_2) = (x_1 + y_1) + (x_2 + y_2) \in V_1 + V_2,$$

$$\alpha(x + y) = \alpha(x_1 + y_1) = (\alpha x_1 + \alpha y_1) \in V_1 + V_2.$$

Thus, $V_1 + V_2$ is a subspace of V.

Now, since $V_1 \cup V_2 \subseteq V_1 + V_2$, and since $V_1 + V_2$ is a subspace, we have $\mathrm{span}\,(V_1 \cup V_2) \subseteq V_1 + V_2$. Also, since $V_1 \subseteq \mathrm{span}\,(V_1 \cup V_2)$, $V_2 \subseteq \mathrm{span}\,(V_1 \cup V_2)$, and since $\mathrm{span}\,(V_1 \cup V_2)$ is a subspace, we have $V_1 + V_2 \subseteq \mathrm{span}\,(V_1 \cup V_2)$. Thus,

$$V_1 + V_2 \subseteq \mathrm{span}\,(V_1 \cup V_2) \subseteq V_1 + V_2,$$

which proves the last part of the theorem. ∎

Exercise 1.19 Suppose V_1 and V_2 are subspaces of a vector space V such that $V_1 \cap V_2 = \{0\}$. Show that every $x \in V_1 + V_2$ can be written *uniquely* as $x = x_1 + x_2$ with $x_1 \in V_1$ and $x_2 \in V_2$. ♦

***Definition* 1.9** If V_1 and V_2 are subspaces of a vector space V such that $V_1 \cap V_2 = \{0\}$, then we write $V_1 + V_2$ as $V_1 \oplus V_2$, and call it as *direct sum* of V_1 and V_2. ◊

Exercise 1.20 Suppose V_1 and V_2 are subspaces of a vector space V. Show that $V_1 + V_2 = V_1$ if and only if $V_2 \subseteq V_1$. ♦

1.5 Basis and Dimension

***Definition* 1.10 (Linear dependence)** Let V be a vector space. A subset E of V is said to be *linearly dependent* if there exists $u \in E$ such that $u \in \operatorname{span}(E \setminus \{u\})$.

***Definition* 1.11 (Linear independence)** Let V be a vector space. A subset E of V is said to be *linearly independent* in V if it is not linearly dependent.

Thus, if E is a subset of V, then

• E is linearly dependent if and only if there exists $\{u_1, \ldots, u_n\} \subseteq E$ and scalars $\alpha_1, \ldots, \alpha_n$, with at least one of them nonzero, such that $\alpha_1 u_1 + \cdots + \alpha_n x_n = 0$, and

• E is linearly independent if and only if for every finite subset $\{u_1, \ldots, u_n\}$ of E, $\alpha_1 u_1 + \cdots + \alpha_n x_n = 0 \;\Rightarrow\; \alpha_i = 0 \quad \forall\, i = 1, \ldots, n$.

If $\{u_1, \ldots, u_n\}$ is a linearly independent (respectively, dependent) subset of a vector space V, then we may also say that $u_1, \ldots, u_n$ are linearly independent (respectively, dependent) in V.

Note that a linearly dependent set cannot be empty. In other words, the empty set is linearly independent!

Caution! If $u_1, \ldots, u_n$ are such that at least one of them is not in the span of the remaining, then we *cannot* conclude that $u_1, \ldots, u_n$ are linearly independent. For the linear independence of $\{u_1, \ldots, u_n\}$, it is required that $u_i \notin \operatorname{span}\{u_j : j \neq i\}$ for every $i \in \{1, \ldots, n\}$.

Also, if $\{u_1, \ldots, u_n\}$ are linearly dependent, then it *does not imply* that any one of them is in the span of the rest.

To illustrate the above points, consider two linearly independent vectors u_1, u_2. Then we have $u_1 \notin \operatorname{span}\{u_2, 3u_2\}$, but $\{u_1, u_2, 3u_2\}$ is linearly dependent, and $\{u_1, u_2, 3u_2\}$ is linearly dependent, but $u_1 \notin \operatorname{span}\{u_2, 3u_2\}$.

Exercise 1.21 Let V be a vector space.

(i) Show that a subset $\{u_1, \ldots, u_n\}$ of V is linearly dependent if and only if there exists a nonzero $(\alpha_1, \ldots, \alpha_n)$ in $\mathbb{F}^n$ such that $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0$.

(ii) Show that a subset $\{u_1, \ldots, u_n\}$ of V is linearly independent if and only if the function $(\alpha_1, \ldots, \alpha_n) \mapsto \alpha_1 u_1 + \cdots + \alpha_n u_n$ from $\mathbb{F}^n$ into V is injective.

(iii) Show that if $E \subseteq V$ is linearly independent in V, then $0 \notin E$.

(iv) Show that if $E \subseteq V$ is linearly dependent in V, then every superset of E is also linearly dependent.

(v) Show that if $E \subseteq V$ is linearly independent in V, then every subset of E is also linearly independent.

(vi) Show that if $\{u_1, \ldots, u_n\}$ is a linearly independent subset of V, and if Y is a subspace of V such that $\{u_1, \ldots, u_n\} \cap Y = \varnothing$, then every x in the span of $\{u_1, \ldots, u_n, Y\}$ can be written uniquely as $x = \alpha_1 u_1 + \cdots + \alpha_n u_n + y$ with $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$, $y \in Y$.

(vii) Show that if E_1 and E_2 are linearly independent subsets of V such that $(\operatorname{span} E_1 \cap (\operatorname{span} E_2) = \{0\}$, then $E_1 \cup E_2$ is linearly independent. ♦

Exercise 1.22 For each $k \in \mathbb{N}$, let $\underline{\mathbb{F}}^k$ denotes the set of all column k-vectors, i.e., the set of all $k \times 1$ matrices. Let A be an $m \times n$ matrix of scalars with columns $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n$. Show the following:

(i) The equation $A\underline{x} = \underline{0}$ has a non-zero solution if and only if $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n$ are linearly dependent.

(ii) For $\underline{y} \in \underline{\mathbb{F}}^m$, the equation $A\underline{x} = \underline{y}$ has a solution if and only if $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n, \underline{y}$ are linearly dependent, i.e., if and only if $\underline{y}$ is in the span of columns of A. ♦

***Definition* 1.12 (Basis)** A subset E of a vector space V is said to be a *basis* of V if it is linearly independent and $\operatorname{span} E = V$.

EXAMPLE 1.25 For each $j \in \{1, \ldots, n\}$, let $e_j \in \mathbb{F}^n$ be such that $e_j(i) = \delta_{ij}$, $i, j = 1, \ldots, n$. Then we have seen that $\{e_1, \ldots, e_n\}$ is linearly independent and its span is $\mathbb{F}^n$. Hence $\{e_1, \ldots, e_n\}$ is a basis of $\mathbb{F}^n$.

EXAMPLE 1.26 For each $j \in \{1, \ldots, n\}$, let $\underline{e}_j \in \underline{\mathbb{F}}^n$ be such that $\underline{e}_j(i) = \delta_{ij}$, $i, j = 1, \ldots, n$. Then it is easily seen that $\{\underline{e}_1, \ldots, \underline{e}_n\}$ is linearly independent and its span is $\underline{\mathbb{F}}^n$. Hence $\{\underline{e}_1, \ldots, \underline{e}_n\}$ is a basis of $\underline{\mathbb{F}}^n$.

***Definition* 1.13 (Standard bases of $\mathbb{F}^n$ and $\underline{\mathbb{F}}^n$)** The basis $\{e_1, \ldots, e_n\}$ of $\mathbb{F}^n$ is called the *standard basis* of $\mathbb{F}^n$, and the basis $\{\underline{e}_1, \ldots, \underline{e}_n\}$ of $\underline{\mathbb{F}}^n$ is called the *standard basis* of $\underline{\mathbb{F}}^n$.

EXAMPLE 1.27 Let $u_j(t) = t^{j-1}$, $j \in \mathbb{N}$. Then $\{u_1, \ldots, u_{n+1}\}$ is a basis of $\mathcal{P}_n$, and $\{u_1, u_2, \ldots\}$ is a basis of $\mathcal{P}$.

Exercise 1.23 Let $u_1(t) = 1$, and for $j = 2, 3, \ldots$, let $u_j(t) = 1 + t + \ldots + t^{j-1}$. Show that $\{u_1, \ldots, u_{n+1}\}$ is a basis of $\mathcal{P}_n$, and $\{u_1, u_2, \ldots\}$ is a basis of $\mathcal{P}$. ♦

EXAMPLE 1.28 For $i = 1, \ldots, m$; $j = 1, \ldots, n$, let M_{ij} be the $m \times n$ matrix with its (i,j)-th entry as 1 and all other entries 0. Then

$$\{M_{ij} : i = 1 \ldots, m;\ j = 1, \ldots, n\}$$

is a basis of $\mathbb{F}^{m \times n}$.

EXAMPLE 1.29 For $\lambda \in [a, b]$, let $u_\lambda(t) = \exp(\lambda t)$, $t \in [a, b]$. Then it is seen that $\{u_\lambda : \alpha \in [a, b]\}$ is an uncountable linearly independent subset of $C[a, b]$.

Clearly, a linearly independent subset of a subspace remains linearly independent in the whole space. Thus, the set $\{u_1, u_2, \ldots\}$ in Example 1.27(ii) is linearly independent in $C[a, b]$ and $\mathcal{F}([a, b], \mathbb{F})$.

Exercise 1.24 If $\{u_1, \ldots, u_n\}$ is a basis of a vector space V, then show that every $x \in V$, can be expressed uniquely as $x = \alpha_1 u_1 + \cdots + \alpha_n u_n$; i.e., for every $x \in V$, there exists a unique n-tuple $(\alpha_1, \ldots, \alpha_n)$ of scalars such that $x = \alpha_1 u_1 + \cdots + \alpha_n u_n$. ♦

Exercise 1.25 Consider the system of equations

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\
\ldots & + & \ldots & + & \ldots & + & \ldots & = & \ldots \\
a_{m1}x_1 & + & a_{m1}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m
\end{array}$$

Show that the above system has at most one solution if and only if the vectors

$$w_1 := \begin{bmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{bmatrix}, \quad w_2 := \begin{bmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{bmatrix}, \ldots, \quad w_n := \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{bmatrix}$$

are linearly independent. ♦

Exercise 1.26 Let $u_1, \ldots, u_n$ are linearly independent vectors in a vector space V. Let $[a_{ij}]$ be an $m \times n$ matrix of scalar, and let

$$\begin{array}{lclclclcl}
v_1 & := & a_{11}u_1 & + & a_{21}u_2 & + & \ldots & + & a_{m1}u_n \\
v_2 & := & a_{12}u_1 & + & a_{22}u_2 & + & \ldots & + & a_{m2}u_n \\
\ldots & & \ldots & + & \ldots & + & \ldots & + & \ldots \\
v_n & := & a_{1n}u_1 & + & a_{2n}u_2 & + & \ldots & + & a_{mn}u_n.
\end{array}$$

Show that the $v_1, \ldots, v_m$ are linearly independent if and only if the vectors

$$w_1 := \begin{bmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{m1} \end{bmatrix}, \quad w_2 := \begin{bmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{m2} \end{bmatrix}, \ldots, \quad w_n := \begin{bmatrix} a_{1n} \\ a_{m2} \\ \ldots \\ a_{mn} \end{bmatrix}$$

are linearly independent. ♦

Exercise 1.27 Let $p_1(t) = 1 + t + 3t^2$, $p_2(t) = 2 + 4t + t^2$, $p_3(t) = 2t + 5t^2$. Are the polynomials p_1, p_2, p_3 linearly independent? ♦

Theorem 1.9 *Let V be a vector space and $E \subseteq V$. Then the following are equivalent.*

(i) E is a basis of V

(ii) E is a maximal linearly independent set in V, i.e., E is linearly independent, and a proper superset of E cannot be linearly independent.

(iii) E is a minimal spanning set of V, i.e., span of E is V, and a proper subset of E cannot span V.

Proof. $(i) \iff (ii)$: Suppose E is a basis of V. Suppose E' is a proper superset of E. Let $x \in E' \backslash E$. Since E is a basis, $x \in \text{span}\,(E)$. This shows that E' is linearly dependent (as $E \cup \{x\} \subseteq E'$).

Conversely, suppose E is a maximal linearly independent set. If E is not a basis, then there exists $x \notin \text{span}\,(E)$. Then, $E \cup \{x\}$ is a linearly independent which is a proper superset of E – a contradiction to the maximality of E.

$(i) \iff (iii)$: Suppose E is a basis of V. Suppose E' is a proper subset of E. Then, it is clear that there exists $x \in E \setminus E'$ which is not in the span of E' (since $E' \cup \{x\} \subseteq E$). Hence, E' does not span V.

Conversely, suppose E is a minimal spanning set of V. If E is not a basis, then E is linearly dependent, and hence there exists

$x \in \text{span}(E \setminus \{x\})$. Since E spans V, it follows that $E \setminus \{x\}$, which is a proper subset of E, also spans V – a contradiction to the fact that E is a minimal spanning set of V. ∎

Exercise 1.28 Suppose V_1 and V_2 are subspaces of a finite dimensional vector space V such that $V_1 \cap V_2 = \{0\}$. If E_1 and E_2 are bases of V_1 and V_2, respectively, then prove that $E_1 \cup E_2$ is a basis of $V_1 + V_2$.

1.5.1 Dimension of a Vector Space

***Definition* 1.14 (Finite dimensional space)** A vector space V is said to be a *finite dimensional space* if there is a finite basis for V.

Recall the empty set is considered as a linearly independent set, and its span is the zero space.

***Definition* 1.15 (Infinite dimensional space)** A vector space which is not a finite dimensional space is called an *infinite dimensional space.*

Theorem 1.10 *If a vector space has a finite spanning set, then it has a finite basis. In fact, if S is a finite spanning set of V, then there exists a basis $E \subseteq S$.*

Proof. Let V be a vector space and S be a finite subset of V such that $\text{span}\, S = V$. If S itself is linearly independent, then we are through. Suppose S is not linearly independent. Then there exists $u_1 \in S$ such that $u_1 \in \text{span}(S \setminus \{u_1\})$. Let $S_1 = S \setminus \{u_1\}$. Clearly,

$$\text{span}\, S_1 = \text{span}\, S = V.$$

If S_1 is linearly independent, then we are through. Otherwise, there exists $u_2 \in S_1$ such that $u_2 \in \text{span}(S_1 \setminus \{u_2\})$. Let $S_2 = S \setminus \{u_1, u_2\}$. Then, we have

$$\text{span}\, S_2 = \text{span}\, S_1 = V.$$

If S_2 is linearly independent, then we are through. Otherwise, continue the above procedure. This procedure will stop after a finite number of steps, as the original set S is a finite set, and we end up with a subset S_k of S which is linearly independent and $\text{span}\, S_k = V$. ∎

By definition, an infinite dimensional space cannot have a finite basis. Is it possible for a finite dimensional space to have an infinite basis, or an infinite linearly independent subset? The answer is, as expected, negative. In fact, we have the following result.

Theorem 1.11 *Let V be a finite dimensional vector space with a basis consisting of n elements. Then every subset of V with more than n elements is linearly dependent.*

Proof. Let $\{u_1, \ldots, u_n\}$ be a basis of V, and $\{x_1, \ldots, x_{n+1}\} \subset V$. We show that $\{x_1, \ldots, x_{n+1}\}$ is linearly dependent.

If $\{x_1, \ldots, x_n\}$ is linearly dependent, then $\{x_1, \ldots, x_{n+1}\}$ is linearly dependent. So, let us assume that $\{x_1, \ldots, x_n\}$ is linearly independent. Now, since $\{u_1, \ldots u_n\}$ is a basis of V, there exist scalars $\alpha_1, \ldots, \alpha_n$ such that

$$x_1 = \alpha_1 u_1 + \cdots + \alpha_n u_n.$$

Since $x_1 \neq 0$, one of $\alpha_1, \ldots, \alpha_n$ is nonzero. Without loss of generality, assume that $\alpha_1 \neq 0$. Then we have $u_1 \in \operatorname{span}\{x_1, u_2, \ldots, u_n\}$ so that

$$V = \operatorname{span}\{u_1, u_2, \ldots, u_n\} = \operatorname{span}\{x_1, u_2, \ldots, u_n\}.$$

Let $\alpha_1^{(2)}, \ldots, \alpha_n^{(2)}$ be scalars such that

$$x_2 = \alpha_1^{(2)} x_1 + \alpha_2^{(2)} u_2 + \cdots + \alpha_n^{(2)} u_n.$$

Since $\{x_1, x_2\}$ is linearly independent, at least one of $\alpha_2^{(2)}, \ldots, \alpha_n^{(2)}$ is nonzero. Without loss of generality, assume that $\alpha_2^{(2)} \neq 0$. Then we have $u_2 \in \operatorname{span}\{x_1, x_2, u_3, \ldots, u_n\}$ so that

$$V = \operatorname{span}\{x_1, u_2, \ldots, u_n\} = \operatorname{span}\{x_1, x_2, u_3, \ldots, u_n\}.$$

Now, let $1 \leq k \leq n-1$ be such that

$$V = \operatorname{span}\{x_1, x_2, \ldots, x_k, u_{k+1}, \ldots, u_n\}.$$

Suppose $k < n-1$. Then there exist scalars $\alpha_1^{(k+1)}, \ldots, \alpha_n^{(k+1)}$ such that

$$x_{k+1} = \alpha_1^{(k+1)} x_1 + \cdots + \alpha_k^{(k+1)} x_k + \alpha_{k+1}^{(k+1)} u_{k+1} + \cdots + \alpha_n^{(k+1)} u_n.$$

Since $\{x_1, \ldots, x_{k+1}\}$ is linearly independent, at least one of the scalars $\alpha_{k+1}^{(k+1)}, \ldots, \alpha_n^{(k+1)}$ is nonzero. Without loss of generality, assume that

$\alpha_{k+1}^{(k+1)} \neq 0$. Then we have $u_{k+1} \in \text{span}\,\{x_1, \ldots, x_{k+1}, u_{k+2}, \ldots, u_n\}$ so that

$$\begin{aligned} V &= \text{span}\,\{x_1, \ldots, x_k, u_{k+1}, \ldots, u_n\} \\ &= \text{span}\,\{x_1, \ldots, x_{k+1}, u_{k+2}, \ldots, u_n\}. \end{aligned}$$

Thus, the above procedure leads to $V = \text{span}\,\{x_1, \ldots, x_{n-1}, u_n\}$ so that there exist scalars $\alpha_1^{(n)}, \ldots, \alpha_n^{(n)}$ such that

$$x_n = \alpha_1^{(n)} x_1 + \cdots + \alpha_{n-1}^{(n)} x_{n-1} + \alpha_n^{(n)} u_n.$$

Since $\{x_1, \ldots, x_n\}$ is linearly independent, it follows that $\alpha_n^{(n)} \neq 0$. Hence,

$$u_n \in \text{span}\,\{x_1, \ldots, x_n\}.$$

Consequently,

$$V = \text{span}\,\{x_1, x_2, \ldots, x_{n-1}, u_n\} = \text{span}\,\{x_1, x_2, \ldots, x_{n-1}, x_n\}.$$

Thus, $x_{n+1} \in \text{span}\,\{x_1, \ldots, x_n\}$, showing that $\{x_1, \ldots, x_{n+1}\}$ is linearly dependent. ∎

The following three corollaries are easy consequences of Theorem 1.11. Their proofs are left as exercises for the reader.

Corollary 1.12 *If V is a finite dimensional vector space, then any two bases of V have the same number of elements.*

Corollary 1.13 *If a vector space contains an infinite linearly independent subset, then it is an infinite dimensional space.*

Corollary 1.14 *If (a_{ij}) is an $m \times n$ matrix with $a_{ij} \in \mathbb{F}$ and $n > m$, then there exists a nonzero $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ such that*

$$a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n = 0, \qquad i = 1, \ldots, m.$$

Exercise 1.29 Assuming Corollary 1.14, give an alternate proof for Theorem 1.11. ♦

By Corollary 1.14, we see that if $\mathbf{A}$ is an $m \times n$ matrix with entries from $\mathbb{F}$ and $n > m$, then there exists an $n \times 1$ nonzero matrix $\mathbf{x}$ such that

$$\mathbf{A}\mathbf{x} = \mathbf{0},$$

where $\mathbf{0}$ is the $m \times 1$ zero matrix.

***Definition* 1.16 (n-vector)** An $n \times 1$ matrix is also called an *n-vector.*

In view of Corollary 1.12, the following definition makes sense.

***Definition* 1.17 (Dimension)** Suppose V is a finite dimensional vector space. Then the *dimension* of V is the number of elements in a basis of V, and this number is denoted by $\dim(V)$. If V is infinite dimensional, then its dimension is defined to be infinity and we write $\dim(V) = \infty$.

EXAMPLE 1.30 The spaces $\mathbb{F}^n$ and $\mathcal{P}_{n-1}$ are of dimension n.

EXAMPLE 1.31 It is seen that the set $\{e_1, e_2, \ldots,\} \subseteq \mathcal{F}(\mathbb{N}, \mathbb{F})$ with $e_j(i) = \delta_{ij}$ is a linearly independent subset of the spaces $\ell^1(\mathbb{N})$ and $\ell^\infty(\mathbb{N})$. Hence, it follows that $\ell^1(\mathbb{N})$ and $\ell^\infty(\mathbb{N})$ are infinite dimensional spaces.

EXAMPLE 1.32 We see that $\{u_1, u_2, \ldots,\}$ with $u_j(t) = t^{j-1}$, $j \in \mathbb{N}$, is linearly independent in $C^k[a,b]$ for every $k \in \mathbb{N}$. Hence, the space $C^k[a,b]$ for each $k \in \mathbb{N}$ is infinite dimensional.

EXAMPLE 1.33 Suppose S is a finite set consisting of n elements. Then $\mathcal{F}(S, \mathbb{F})$ is of dimension n. To see this, let $S = \{s_1, \ldots, s_n\}$, and for each $j \in \{1, \ldots, n\}$, define $f_j \in \mathcal{F}(S, \mathbb{F})$ by

$$f_j(s_i) = \delta_{ij}, \qquad i \in \{1, \ldots, n\}.$$

Then the set $\{f_1, \ldots, f_n\}$ is a basis of $\mathcal{F}(S, \mathbb{F})$: Clearly,

$$\sum_{j=1}^{n} \alpha_j f_j = 0 \;\Rightarrow\; \alpha_i = \sum_{j=1}^{n} \alpha_j f_j(s_i) = 0 \quad \forall\, i.$$

Thus, $\{f_1, \ldots, f_n\}$ is linearly independent. To see that it spans the space, it is enough to note that $f = \sum_{j=1}^{n} f(s_j) f_j$ for all $f \in \mathcal{F}(S, \mathbb{F})$.

Exercise 1.30 Suppose V_1 and V_2 are finite dimensional vector spaces. Show that $V_1 \times V_2$ is a finite dimensional vector space and

$$\dim(V_1 \times V_2) = \dim(V_1) + \dim(V_2).$$

(Hint: If $\{u_1, \ldots, u_m\}$ and $\{v_1, \ldots, v_n\}$ are bases of V_1 and V_2, respectively, then $\{(u_i, 0) : i = 1, \ldots, m\} \cup \{(0, v_j) : j = 1, \ldots, n\}$ is a basis of $V_1 \times V_2$.)

Exercise 1.31 Let V_1 and V_2 be subspaces of a finite dimensional vector space V such that $V_1 \cap V_2 = \{0\}$. Show that $V_1 \times V_2$ is linearly isomorphic with $V_1 \oplus V_2$.

Remark 1.1 We have seen that if there is a finite spanning set S for a vector space V, then there is a basis $E \subseteq S$. Also, we know that a linearly independent set E is a basis if and only if it is maximal, in the sense that, there is no linearly independent set properly containing E. The existence of such a maximal linearly independent set can be established using an axiom in set theory known as *Zorn's lemma.* It is called lemma, as it known to be equivalent to an axiom, called *Axiom of choice.*

1.5.2 Dimension of Sum of Subspaces

Theorem 1.15 *Suppose V_1 and V_2 are subspaces of a finite dimensional vector space V. If $V_1 \cap V_2 = \{0\}$, then*

$$\dim (V_1 + V_2) = \dim V_1 + \dim V_2.$$

Proof. Suppose $\{u_1, \ldots, u_k\}$ is a basis of V_1 and $\{v_1, \ldots, v_\ell\}$ is a basis of V_2. We show that $E := \{u_1, \ldots, u_k, v_1, \ldots, v_\ell\}$ is a basis of $V_1 + V_2$. Clearly (*Is it clear?*) $\operatorname{span} E = V_1 + V_2$. So, it is enough to show that E is linearly independent. For this, suppose $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_\ell$ are scalars such that $\alpha_1 u_1 + \ldots + \alpha_k u_k + \beta_1 v_1 + \ldots + \beta_\ell v_\ell = 0$. Then we have

$$x := \alpha_1 u_1 + \ldots + \alpha_k u_k = -(\beta_1 v_1 + \ldots + \beta_\ell v_\ell) \in V_1 \cap V_2 = \{0\}$$

so that $\alpha_1 u_1 + \ldots + \alpha_k u_k = 0$ and $\beta_1 v_1 + \ldots + \beta_\ell v_\ell = 0$. From this, by the linearly independence of u_i's and v_j's, it follows that $\alpha_i = 0$ for $i \in \{1, \ldots, k\}$ and $\beta_j = 0$ for all $j \in \{1, \ldots, \ell\}$. Hence, E is linearly independent. This completes the proof. ∎

In fact, the above theorem is a particular case of the following.

Theorem 1.16 *Suppose V_1 and V_2 are subspaces of a finite dimensional vector space V. Then*

$$\dim (V_1 + V_2) = \dim V_1 + \dim V_2 - \dim (V_1 \cap V_2).$$

For the proof of the above theorem we shall make use of the following result.

Proposition 1.17 *Let V be a finite dimensional vector space. If E_0 is a linearly independent subset of V, then there exists a basis E of V such that $E_0 \subseteq E$.*

Proof. Let $E_0 = \{u_1, \ldots, u_k\}$ be a linearly independent subset of V, and let $\{v_1, \ldots, v_n\}$ be a basis of V. Let

$$E_1 = \begin{cases} E_0 & \text{if } v_1 \in \operatorname{span}(E_0), \\ E_0 \cup \{v_1\} & \text{if } v_1 \notin \operatorname{span}(E_0). \end{cases}$$

Clearly, E_1 is linearly independent, and

$$E_0 \subseteq E_1, \qquad \{v_1\} \subseteq \operatorname{span}(E_1).$$

Then define

$$E_2 = \begin{cases} E_1 & \text{if } v_2 \in \operatorname{span}(E_1), \\ E_1 \cup \{v_2\} & \text{if } v_2 \notin \operatorname{span}(E_1). \end{cases}$$

Again, it is clear that E_2 is linearly independent, and

$$E_1 \subseteq E_2, \qquad \{v_1, v_2\} \subseteq \operatorname{span}(E_2).$$

Having defined $E_1, \ldots, E_j$, $j < n$, we define

$$E_{j+1} = \begin{cases} E_j & \text{if } v_{j+1} \in \operatorname{span}(E_j), \\ E_j \cup \{v_{j+1}\} & \text{if } v_{j+1} \notin \operatorname{span}(E_j). \end{cases}$$

Thus, we get linearly independent sets $E_1, E_2, \ldots, E_n$ such that

$$E_0 \subseteq E_1 \subseteq \ldots E_n, \qquad \{v_1, v_2, \ldots, v_n\} \subseteq \operatorname{span}(E_n).$$

Since $\{v_1, \ldots, v_n\}$ is a basis of V, it follows that $E := E_n$ is a basis of V such that $E_0 \subseteq E_n = E$. ∎

Proof of Theorem 1.16. Let $\{u_1, \ldots, u_k\}$ be a basis of the subspace $V_1 \cap V_2$. By Proposition 1.17, there exists $v_1, \ldots, v_\ell$ in V_1 and $w_1, \ldots, w_m$ in V_2 such that $\{u_1, \ldots, u_k, v_1, \ldots, v_\ell\}$ is a basis of V_1, and $\{u_1, \ldots, u_k, w_1, \ldots, w_m\}$ is a basis of V_2. We show that $E := \{u_1, \ldots, u_k, v_1, \ldots, v_\ell, w_1, \ldots, w_m\}$ is a basis of $V_1 + V_2$.

Clearly, $V_1 + V_2 = \text{span}\,(E)$. Hence, it is enough to show that E is linearly independent. For this, let $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_\ell, \gamma_1, \ldots, \gamma_m$ be scalars such that

$$\sum_{i=1}^{k} \alpha_i u_i + \sum_{i=1}^{\ell} \beta_i v_i + \sum_{i=1}^{m} \gamma_i w_i = 0. \tag{$*$}$$

Then

$$x := \sum_{i=1}^{k} \alpha_i u_i + \sum_{i=1}^{\ell} \beta_i v_i = -\sum_{i=1}^{m} \gamma_i w_i \in V_1 \cap V_2.$$

Hence, there exists scalars $\delta_1, \ldots, \delta_k$ such that

$$\sum_{i=1}^{k} \alpha_i u_i + \sum_{i=1}^{\ell} \beta_i v_i = \sum_{i=1}^{k} \delta_i u_i, \quad \textit{i.e.,} \quad \sum_{i=1}^{k} (\alpha_i - \delta_i) u_i + \sum_{i=1}^{\ell} \beta_i v_i = 0$$

Since $\{u_1, \ldots, u_k, v_1, \ldots, v_\ell\}$ is a basis of V_1, it follows that $\alpha_i = \delta_i$ for all $i = 1, \ldots, k$, and $\beta_j = 0$ for $j = 1, \ldots, \ell$. Hence, from $(*)$,

$$\sum_{i=1}^{k} \alpha_i u_i + \sum_{i=1}^{m} \gamma_i w_i = 0.$$

Now, since $\{u_1, \ldots, u_k, w_1, \ldots, w_m\}$ is a basis of V_2, it follows that $\alpha_i = 0$ for all $i = 1, \ldots, k$, and $\gamma_j = 0$ for all $j = 1, \ldots, m$.

Thus, we have shown that $\{u_1, \ldots, u_k, v_1, \ldots, v_\ell, w_1, \ldots, w_m\}$ is a basis of $V_1 + V_2$. Since $\dim\,(V_1 + V_2) = k + \ell + m$, $\dim V_1 = k + \ell$, $\dim V_2 = k + m$ and $\dim\,(V_1 \cap V_2) = k$, we get

$$\dim\,(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim\,(V_1 \cap V_2).$$

This completes the proof. ▮

1.6 Quotient Space

Let V be a vector space and V_0 be a subspace of V. For $x \in V_0$, let

$$x + V_0 := \{x + u : u \in V_0\}.$$

We note that (verify!), for $x, y \in V$,

$$x + V_0 = y + V_0 \iff x - y \in V_0,$$

so that we have

$$x + V_0 = V_0 \iff x \in V_0.$$

Let us denote

$$V/V_0 := \{x + V_0 : x \in V\}.$$

Theorem 1.18 *On the set V/V_0, consider the the operation of addition and scalar multiplication as follows: For $x, y \in V$ and $\alpha \in \mathbb{F}$,*

$$(x + V_0) + (y + V_0) := (x + y) + V_0,$$

$$\alpha(x + V_0) := \alpha x + V_0.$$

With these operations, V/V_0 is a vector space with its zero as V_0 and additive inverse of $x + V_0$ as $(-x) + V_0$.

***Definition* 1.18** The vector space V/V_0 is called the **quotient space** of V with respect to V_0. ◊

EXAMPLE 1.34 Let $V = \mathbb{R}^2$ over $\mathbb{F} = \mathbb{R}$ and V_0 be the straight line given by $V_0 := \{(x, y) : ax + by = 0\}$ for some $a, b \in \mathbb{R}$. Then for every $v \in \mathbb{R}^2$, $v + V$ is the line passing through v and parallel to V_0.

EXAMPLE 1.35 Let $V = \mathbb{R}^3$ over $\mathbb{F} = \mathbb{R}$ and V_0 be the plane given by $V_0 := \{(x, y, z) : ax + by + cz = 0\}$ for some $a, b, c \in \mathbb{R}$. Then for every $v \in \mathbb{R}^2$, $v + V$ is the plane passing through v and parallel to V_0.

Theorem 1.19 *Let V is a finite dimensional vector space and V_0 is a subspace of V. If V_1 is a subspace of V such that $V = V_0 \oplus V_1$, then V_1 is linearly isomorphic with V/V_0.*

As a corollary to the above theorem we have the following.

Theorem 1.20 *Let V is a finite dimensional vector space and V_0 is a subspace of V. Then* $\dim(V/V_0)$ *is a finite dimensional space and*

$$\dim(V) = \dim(V_0) + \dim(V/V_0).$$

2

Linear Transformations

2.1 Motivation

We may recall from the theory of matrices that if A is an $m \times n$ matrix, and if $\vec{x}$ is an n-vector, then $A\vec{x}$ is an m-vector. Moreover, for any two n-vectors $\vec{x}$ and $\vec{y}$, and for every scalar α,

$$A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}, \qquad A(\alpha\vec{x}) = \alpha A\vec{x}).$$

Also, we recall from calculus, if f and g are real-valued differentiable functions (defined on an interval J), and α is a scalar, then

$$\frac{d}{dt}(f + g) = \frac{d}{dt}f + \frac{d}{dt}g, \qquad \frac{d}{dt}(\alpha f) = \alpha\frac{d}{dt}f.$$

Note also that, if f and g are continuous real-valued functions defied on an interval $[a, b]$, then

$$\int_a^b (f+g)(t)dt = \int_a^b f(t)\,dt + \int_a^b g(t)\,dt, \qquad \int_a^b (\alpha f)(t) = \alpha \int_a^b f(t)\,dt,$$

and for every $s \in [a, b]$,

$$\int_a^s (f+g)(t)dt = \int_a^s f(t)\,dt + \int_a^s g(t)\,dt, \qquad \int_a^s (\alpha f)(t) = \alpha \int_a^s f(t)\,dt.$$

Abstracting the above operations between specific vector spaces, we define the notion of a linear transformation between general vector spaces.

2.2 Definition and Examples

***Definition* 2.1 (Linear transformation)** Let V_1 and V_2 be vector spaces (over the same scalar field $\mathbb{F}$). A a function $T : V_1 \to V_2$ is said to be a *linear transformation* or a *linear operator* from V_1 to V_2 if

$$T(x + y) = T(x) + T(y), \qquad T(\alpha x) = \alpha T(x)$$

for every $x, y \in V_1$ and for every $\alpha \in \mathbb{F}$.

For $x \in V_1$, we may denote the element $T(x)$ by Tx as well.

EXAMPLE 2.1 Let V be a vector space and λ be a scalar. Define $T : V \to V$ by $T(x) = \lambda x$, $x \in V$. Then we see that T is a linear transformation.

EXAMPLE 2.2 (Matrix as linear transformation) Let $A = (a_{ij})$ be an $m \times n$-matrix of scalars. For $x \in \underline{\mathbb{F}}^n$, let $T(x) = Ax$ for every $x \in \underline{\mathbb{F}}^n$. Then it follows that $T : \underline{\mathbb{F}}^n \to \underline{\mathbb{F}}^m$ is a linear transformation.

EXAMPLE 2.3 For each $j \in \{1, \ldots, n\}$, the function $T_j : \mathbb{F}^n \to \mathbb{F}$ defined by $T_j x = x_j$ for $x = (x_1, \ldots, x_n) \in \mathbb{F}^n$, is a linear transformation.

More generally, we have the following example.

EXAMPLE 2.4 Let V be an n-dimensional space and let $E = \{u_1, \ldots, u_n\}$ be a basis of V. For $x = \sum_{j=1}^{n} \alpha_j u_j \in V$, and for each $j \in \{1, \ldots, n\}$, define $T_j : V \to \mathbb{F}$ by

$$T_j x = \alpha_j.$$

Then T_j is a linear transformation.

EXAMPLE 2.5 (Evaluation of functions) For a given point $\tau \in [a, b]$, let $T_\tau : C[a, b] \to \mathbb{F}$ be defined by

$$T_\tau f = f(\tau), \qquad f \in C[a, b].$$

Then T_τ is a linear transformation.

More generally, we have the following example.

EXAMPLE 2.6 Given points $\tau_1, \ldots, \tau_n$ in $[a, b]$, and $\omega_1, \ldots \omega_n$ in $\mathbb{F}$, let $T : C[a, b] \to \mathbb{F}$ be defined by

$$Tf = \sum_{i=1}^{n} f(\tau_i)\omega_i, \qquad f \in C[a, b].$$

Then T is a linear transformation

EXAMPLE 2.7 (Differentiation) Let $T : C^1[a, b] \to C[a, b]$ be defined by

$$Tf = f', \qquad f \in C^1[a, b],$$

where f' denotes the derivative of f. Then T is a linear transformation.

EXAMPLE 2.8 For $\lambda, \mu \in \mathbb{F}$, the function $T : C^1[a, b] \to C[a, b]$ defined by

$$Tf = \lambda f + \mu f', \qquad f \in C^1[a, b],$$

is a linear transformation.

More generally, we have the following example.

EXAMPLE 2.9 Let T_1 and T_2 be linear transformations from V_1 to V_2 and λ and μ be scalars. Then $T : V_1 \to V_2$ defined by

$$T(x) = \lambda T_1(x) + \mu T_2(x), \quad x \in V_1,$$

is a linear transformation.

EXAMPLE 2.10 (Definite integration) Let $T : C[a, b] \to \mathbb{F}$ be defined by

$$Tf = \int_a^b f(t)\,dt, \qquad f \in C[a, b].$$

Then T is a linear transformation.

EXAMPLE 2.11 (Indefinite integration) Let $T : C[a, b] \to C[a, b]$ be defined by

$$(Tf)(s) = \int_a^s f(t)\,dt, \qquad f \in C[a, b], \quad s \in [a, b].$$

Then T is a linear transformation.

EXAMPLE 2.12 (Linear transformation induced by a matrix) Let $A = (a_{ij})$ be an $m \times n$-matrix of scalars. For $x = (\alpha_1, \ldots, \alpha_n)$ in $\mathbb{F}^n$, let

$$Tx = (\beta_1, \ldots, \beta_m), \qquad \beta_i = \sum_{j=1}^{n} a_{ij}\alpha_j, \quad i = 1, \ldots, m.$$

Then $T : \mathbb{F}^n \to \mathbb{F}^m$ is a linear transformation.

More generally, let V_1 be an n-dimensional vector space and V_2 be an m-dimensional vector space. Let $E_1 = \{u_1, \ldots, u_n\}$ and $E_2 = \{v_1, \ldots, v_m\}$ be a bases of V_1 and V_2, respectively. For $x = \sum_{j=1}^{n} \alpha_j u_j \in V$, define $T : V_1 \to V_2$ by

$$Tx = \sum_{i=1}^{m} \beta_i v_i, \quad \text{where} \quad \beta_i = \sum_{j=1}^{n} a_{ij}\alpha_j, \quad i \in \{1, \ldots, m\}.$$

Then T is a linear transformation.

EXAMPLE 2.13 Let V_1 and V_2 be vector spaces with $\dim V_1 = n < \infty$. Let $E_1 = \{u_1, \ldots, u_n\}$ be a basis of V_1 and $E_2 = \{v_1, \ldots, v_m\}$ be a subset of V_2. For $x = \sum_{j=1}^{n} \alpha_j u_j \in V_1$, define $T : V_1 \to V_2$ by

$$Tx = \sum_{i=1}^{m} \alpha_i v_i.$$

Then T is a linear transformation.

Exercise 2.1 Show that the linear transformation T in Example 2.13 is

(a) injective if and only if E_2 is linearly independent,

(b) surjective if and only if $\operatorname{span}(E_2) = V_2$. ♦

***Definition* 2.2 (Isomorphism of vector spaces)** Vector V_1 and V_2 are said to be **linearly isomorphic** if there exists a bijective linear transformation $T : V_1 \to V_2$, and in that case we write $V_1 \simeq V_2$.

Example 2.13 shows that any two finite dimensional vector spaces of the same dimension are linearly isomorphic.

If the codomain of a linear transformation is the scalar field $\mathbb{F}$, then it has a special name.

***Definition* 2.3 (Linear functional)** Let V be a vector space. A linear transformation $T : V \to \mathbb{F}$ is called a *linear functional* on V.

Linear functionals are usually denoted by small letters such as f, g, etc., whereas linear transformations between general vector spaces are denoted by capital letters A, B, T, etc.

The linear transformations given in Examples 2.3, 2.4, 2.5, and 2.10 are linear functionals.

The linear functionals $f_1, \ldots, f_n$ defined in Example 2.4 are called the **coordinate functionals** on V with respect to the basis E of V.

***Definition* 2.4** Let V be an n-dimensional vector space and let $E = \{u_1, \ldots, u_n\}$ be a basis of V. For $x = \sum_{j=1}^{n} \alpha_j u_j \in V$, and for each $j \in \{1, \ldots, n\}$, define $f_j : V \to \mathbb{F}$ by

$$f_j(x) = \alpha_j.$$

Then $f_1, \ldots, f_n$ are linear functionals on V, and they are called the **coordinate functionals** on V with respect to the basis E of V.

We observe that if $f_1, \ldots, f_n$ are the coordinate functionals on V with respect to the basis $E = \{u_1, \ldots, u_n\}$ of V, then

$$f_j(u_i) = \delta_{ij} \quad \forall\, i, j = 1, \ldots, n.$$

It is to be remarked that these linear functionals depend not only on the basis $E = \{u_1, \ldots, u_n\}$, but also on the order in which $u_1, \ldots, u_n$ appear in the representation of any $x \in V$.

2.3 Space of Linear Transformations

We shall denote the set of all linear transformations from a vector space V_1 to a vector space V_2 by $\mathcal{L}(V_1, V_2)$. If $V_1 = V_2$, then we write $\mathcal{L}(V_1, V_2)$ by $\mathcal{L}(V)$, where $V = V_1 = V_2$.

On the set $\mathcal{L}(V_1, V_2)$, define addition and scalar multiplication *pointwise*, i.e., for T, T_1, T_2 in $\mathcal{L}(V_1, V_2)$ and $\alpha \in \mathbb{F}$, linear transformations $T_1 + T_2$ and αT are defined by

$$(T_1 + T_2)(x) = T_1 x + T_2 x,$$

$$(\alpha T)(x) = \alpha T x$$

for all $x \in V$. Then it is seen that $\mathcal{L}(V_1, V_2)$ is a vector space with its zero element as the zero operator $O : V_1 \to V_2$ defined by

$$Ox = 0 \quad \forall\, x \in V_1$$

and the additive inverse $-T$ of $T \in \mathcal{L}(V_1, V_2)$ is $-T : V_1 \to V_2$ defined by

$$(-T)(x) = -Tx \quad \forall\, x \in V_1.$$

The space $\mathcal{L}(V, \mathbb{F})$ of all linear functionals on V is called the **dual** of the space V.

Theorem 2.1 *Let V be a finite dimensional vector space, and let $E = \{u_1, \ldots, u_n\}$ be a basis of V. If $f_1, \ldots, f_n$ are the coordinate functionals on V with respect to E, then we have the following:*

(i) *Every $x \in V$ can be written as $x = \sum_{j=1}^{n} f_j(x)u_j$.*

(ii) *$\{f_1, \ldots, f_n\}$ is a basis of $\mathcal{L}(V, \mathbb{F})$.*

Proof. Since $E = \{u_1, \ldots, u_n\}$ is a basis of V, for every $x \in V$, there exist unique scalars $\alpha_1, \ldots \alpha_n$ such that $x = \sum_{j=1}^{n} \alpha_j u_j$. Now, using the relation $f_i(u_j) = \delta_{ij}$, it follows that

$$f_i(x) = \sum_{j=1}^{n} \alpha_j f_i(u_j) = \alpha_i, \qquad i = 1, \ldots, n.$$

Therefore, the result in (i) follows.

To see (ii), first we observe that if $\sum_{i=1}^{n} \alpha_i f_i = 0$, then

$$\alpha_j = \sum_{i=1}^{n} \alpha_i f_i(u_j) = 0 \quad \forall\, j = 1, \ldots, n.$$

Hence, $\{f_1, \ldots, f_n\}$ is linearly independent in $\mathcal{L}(V, \mathbb{F})$. It remains to show that the span $\{f_1, \ldots, f_n\} = \mathcal{L}(V, \mathbb{F})$. For this, let $f \in \mathcal{L}(V, \mathbb{F})$ and $x \in V$. Then using the representation of x in (i), we have

$$f(x) = \sum_{j=1}^{n} f_j(x) f(u_j) = \left(\sum_{j=1}^{n} f(u_j) f_j \right)(x)$$

for all $x \in V$. Thus, $f = \sum_{j=1}^{n} f(u_j) f_j$ so that $f \in \operatorname{span}\{f_1, \ldots, f_n\}$. This completes the proof. ∎

Let V be a finite dimensional vector space and let $E = \{u_1, \ldots, u_n\}$ be a basis of V, and $f_1, \ldots, f_n$ be the associated coordinate functionals. In view of the above theorem, we say that $\{f_1, \ldots, f_n\}$ is the **dual basis** of $\mathcal{L}(V, \mathbb{F})$ with respect to the (ordered) basis E of V.

Exercise 2.2 (a) Let V_1 and V_2 be finite dimensional vector spaces, and $E_1 = \{u_1, \ldots, u_n\}$ and $E_2 = \{v_1, \ldots, v_m\}$ be bases of V_1 and V_2, respectively. Let $F_1 = \{f_1, \ldots, f_n\}$ be the dual basis of $\mathcal{L}(V_1, \mathbb{F})$ with respect to E_1 and $F_2 = \{g_1, \ldots, g_n\}$ be the dual basis of $\mathcal{L}(V_2, \mathbb{F})$ with respect to E_2. For $i = 1, \ldots, n$; $j = 1, \ldots, m$, let $T_{ij} : V \to W$ defined by

$$T_{ij}(x) = f_j(x) v_i, \qquad x \in V_1.$$

Show that $\{T_{ij} : i = 1, \ldots, n; j = 1, \ldots, m\}$ is a basis of $\mathcal{L}(V_1, V_2)$.

(b) Let V_1 and V_2 be vector spaces, and V_0 be a subspace of V_1. Let $A_0 : V_0 \to V_2$ be a linear transformation. Show that there exists a linear transformation $T : V_1 \to V_2$ such that $A_{|V_0} = A_0$. ♦

2.4 Matrix Representations

Let V_1 and V_2 be finite dimensional vector spaces, and $E_1 = \{u_1, \ldots, u_n\}$ and $E_2 = \{v_1, \ldots, v_m\}$ be bases of V_1 and V_2, respectively. Let $T : V_1 \to V_2$ be a linear transformation. Note that for every $x \in V_1$, there exists a unique $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ such that $x = \sum_{j=1}^n \alpha_j u_j$. Then, by the linearity of T, we have

$$T(x) = \sum_{j=1}^{n} \alpha_j T(u_j).$$

Since $T(u_j) \in V_2$ for each $j = 1, \ldots, n$ and $\{v_1, \ldots, v_m\}$ is a basis of V_2, Tu_j can be written as

$$T(u_j) = \sum_{i=1}^{m} a_{ij} v_i$$

for some scalars $a_{1j}, a_{2j}, \ldots, a_{mj}$. Thus,

$$T(x) = \sum_{j=1}^{n} \alpha_j T u_j = \sum_{j=1}^{n} \alpha_j \left(\sum_{i=1}^{m} a_{ij} v_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \alpha_j \right) v_i.$$

If we denote the i-th coordinate of a vector $\vec{x} \in \mathbb{F}^n$ by $\vec{x}_i$, then the above relation connecting the linear transformation T and the matrix $A = (a_{ij})$ can be written as

$$Tx = \sum_{i=1}^{m} (A\vec{x})_i v_i.$$

In view of the above representation of T, we say that the $m \times n$ matrix $A = (a_{ij})$ is the *matrix representation* of T, with respect to the ordered bases E_1 and E_2 of V_1 and V_2 respectively. This fact is written as

$$[T]_{E_2,E_1} = (a_{ij}).$$

Clearly, the above discussion also shows that for every $m \times n$ matrix $A = (a_{ij})$, there exists a linear transformation $T : V_1 \to V_2$ such that $[T]_{E_2,E_1} = (a_{ij})$. Thus, there is a one-one correspondence between $\mathcal{L}(V_1, V_2)$ onto $\mathbb{F}^{m\times n}$, namely,

$$T \mapsto [T]_{E_2,E_1}.$$

Exercise 2.3 Let V_1 and V_2 be finite dimensional vector spaces, and $E_1 = \{u_1, \ldots, u_n\}$ and $E_2 = \{v_1, \ldots, v_m\}$ be bases of V_1 and V_2, respectively. Show the following:

(a) If $\{g_1, \ldots, g_m\}$ is the ordered dual basis of $\mathcal{L}(V_1, \mathbb{F})$ with respect to the basis E_2 of V_2, then for every $T \in \mathcal{L}(V_1, V_2)$,

$$[T]_{E_2,E_1} = (g_i(Tu_j)).$$

(b) If $A, B \in \mathcal{L}(V_1, V_2)$ and $\alpha \in \mathbb{F}$, then

$$[A + B]_{E_2,E_1} = [A]_{E_2,E_1} + [B]_{E_2,E_1}, \qquad [\alpha A]_{E_2,E_1} = \alpha[A]_{E_2,E_1}.$$

(c) Suppose $\{M_{ij} : i = 1 \ldots, m; \; j = 1, \ldots, n\}$ is a basis of $\mathbb{F}^{m\times n}$. If $T_{ij} \in \mathcal{L}(V_1, V_2)$ is the linear transformation such that $[T_{ij}]_{E_2,E_1} = M_{ij}$, then $\{T_{ij} : i = 1 \ldots, m; \; j = 1, \ldots, n\}$ is a basis of $\mathcal{L}(V_1, V_2)$. (For example, M_{ij} can be takes as in Example 1.28. ♦

Exercise 2.4 Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$T(x_1, x_2, x_3) = (x_2 + x_3, x_3 + x_1, x_1 + x_2), \quad (x_1, x_2, x_3) \in \mathbb{R}^3.$$

Find the matrix representation of T with respect to the basis given in each of the following.

(a) $E_1 = \{(1,0,0),(0,1,0),(0,0,1)\}$, $E_2 = \{(1,0,0),(1,1,0),(1,1,1)\}$
(b) $E_1 = \{(1,0,0),(1,1,0),(1,1,1)\}$, $E_2 = \{(1,0,0),(0,1,0),(0,0,1)\}$
(c) $E_1 = \{(1,1,-1),(-1,1,1),(1,-1,1)\}$,
$E_2 = \{(-1,1,1),(1,-1,1),(1,1,-1)$

Exercise 2.5 Let $T : \mathcal{P}^3 \to \mathcal{P}^2$ be defined by $T(a_0 + a_1t + a_2t^2 + a_3t^3) = a_1 + 2a_2t + 3a_3t^2$. Find the matrix representation of T with respect to the basis given in each of the following.

(a) $E_1 = \{1, t, t^2, t^3\}$, $E_2 = \{1+t, 1-t, t^2\}$
(b) $E_1 = \{1, 1+t, 1+t+t^2, t^3\}$, $E_2 = \{1, 1+t, 1+t+t^2\}$
(c) $E_1 = \{1, 1+t, 1+t+t^2, 1+t+t^2+t^3\}$, $E_2 = \{t^2, t, 1\}$

Exercise 2.6 Let $T : \mathcal{P}^2 \to \mathcal{P}^3$ be defined by $T(a_0 + a_1t + a_2t^2) = (a_0t + \frac{a_1}{2}t^2 + \frac{a_2}{3}t^3)$. Find the matrix representation of T with respect to the basis given in each of the following.

(a) $E_1 = \{1+t, 1-t, t^2\}$, $E_2 = \{1, t, t^2, t^3\}$,
(b) $E_1 = \{1, 1+t, 1+t+t^2\}$, $E_2 = \{1, 1+t, 1+t+t^2, t^3\}$,
(c) $E_1 = \{t^2, t, 1\}$, $E_2 = \{1, 1+t, 1+t+t^2, 1+t+t^2+t^3\}$,

2.5 Rank and Nullity

Let V_1 and V_2 be vector spaces and $T : V_1 \to V_2$ be a linear transformation. Then it can be easily seen that the sets

$$R(T) = \{Tx : x \in V_1\}, \qquad N(T) = \{x \in V_1 : Tx = 0\}$$

are subspaces of V_1 and V_2, respectively.

***Definition* 2.5** The subspaces $R(T)$ and $N(T)$ associated with a linear transformation $T : V_1 \to V_2$ are called the **range space** of T and **null space** of T, respectively. $\Diamond$

***Definition* 2.6** The dimension of $R(T)$ is called the **rank** of T, denoted by $\operatorname{rank} T$, and the dimension of $N(T)$ is called the **nullity** of T, denoted by $\operatorname{null} T$. $\Diamond$

Let $T : V_1 \to V_2$ be a linear transformation. Clearly, T is onto or surjective if and only if $R(T) = V_2$. Using the linearity of T, it can be seen that (*Verify*)

- T is one-one if and only if $N(T) = \{0\}$.

Theorem 2.2 *Let $T : V_1 \to V_2$ be a linear transformation. If $\{u_1, \ldots, u_n\}$ is a basis of V_1, then $R(T) = \operatorname{span}\{Tu_1, \ldots, Tu_n\}$.*

Proof. Let $\{u_1, \ldots, u_n\}$ be a basis of V_1.

Clearly, $\operatorname{span}\{Tu_1, \ldots, Tu_n\} \subseteq R(T)$. To show the other-way inclusion, let $y \in R(T)$ and let $x \in V_1$ be such that $Tx = y$. Since $\{u_1, \ldots, u_n\}$ is a basis of V_1, $x = \sum_{i=1}^n \alpha_i u_i$ for some $\alpha_1, \ldots, \alpha_n$ in $\mathbb{F}$. Hence, $y = Tx = \sum_{i=1}^n \alpha_i Tu_i \in \operatorname{span}\{Tu_1, \ldots, Tu_n\}$. ∎

By the above theorem, we have

- $\operatorname{rank}(T) \leq \dim(V_1)$.

Theorem 2.3 *Let $T : V_1 \to V_2$ be a linear transformation. The we have the following.*

(1) *If T is one-one and $u_1, \ldots, u_k$ are linearly independent in V_1, then $Tu_1, \ldots, Tu_k$ are linearly independent in V_2.*

(2) *If $u_1, \ldots, u_n$ are in V_1 such that $Tu_1, \ldots, Tu_n$ are linearly independent in V_2, then $u_1, \ldots u_n$ are linearly independent in V_1.*

Proof. (1) Suppose T is one-one and $u_1, \ldots, u_k$ are linearly independent in V_1. To show that $Tu_1, \ldots, Tu_k$ are linearly independent in V_2. For this, let $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ be such that

$$\alpha_1 Tu_1 + \cdots + \alpha_k Tu_k = 0. \tag{$*$}$$

We have to show that $\alpha_i = 0, \ldots, \alpha_k = 0$. Now, $(*)$ implies that $T(\alpha_1 u_1 + \cdots + \alpha_k u_k) = 0$, so that $\alpha_1 u_1 + \cdots + \alpha_k u_k \in N(T)$. Since T is one-one, we have $N(T = \{0\}$ so that $\alpha_1 u_1 + \cdots + \alpha_k u_k = 0$. Now, linearly independence of $u_1, \ldots, u_k$ implies that $\alpha_1 = 0, \ldots, \alpha_k = 0$.

(2) Suppose $u_1, \ldots, u_n$ are in V_1 such that $Tu_1, \ldots, Tu_n$ are linearly independent in V_2. To show that $u_1, \ldots, u_n$ are linearly independent in V_1. For this, let $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ be such that

$$\alpha_1 u_1 + \cdots + \alpha_k u_k = 0. \tag{$**$}$$

We have to show that $\alpha_i = 0, \ldots, \alpha_k = 0$. From $(**)$, we have $T(\alpha_1 u_1 + \cdots + \alpha_k u_k) = 0$, i.e., $\alpha_1 Tu_1 + \cdots + \alpha_k Tu_k = 0$. Since $Tu_1, \ldots, Tu_n$ are linearly independent, $\alpha_1 = 0, \ldots, \alpha_k = 0$. ∎

As a corollary to the above two theorems, we have the following.

Theorem 2.4 *Let V_1 and V_2 be finite dimensional vector spaces of the same dimension, and let $T : V_1 \to V_2$ be a linear transformation. Then T is one-one if and only if it is onto.*

Proof. Let $u_1, \ldots, u_n$ in V_1 form a basis of V_1. Then by Theorem 2.3 (1), $\{Tu_1, \ldots, Tu_n\}$ is linearly independent in V_2. Since $\dim(V_2) = n$, $\{Tu_1, \ldots, Tu_n\}$ is a basis of V_2. Hence, by Theorem 2.2, $R(T) = \operatorname{span}\{Tu_1, \ldots, tu_n\} = V_2$. Hence, T is onto.

Conversely, suppose that T is onto, i.e., $R(T) = V_2$. By Theorem 2.2, $V_2 = R(T) = \operatorname{span}\{Tu_1, \ldots, tu_n\} = V_2$. Since $\dim(V_2) = n$, $\{Tu_1, \ldots, Tu_n\}$ is a basis of V_2. Now, let $x \in V_1$ be such that $Tx = 0$. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ be such that $x = \sum_{i=1}^n \alpha_i u_i$. Thus, we have $T(\sum_{i=1}^n \alpha_i u_i) = 0$, i.e., $\sum_{i=1}^n \alpha_i Tu_i = 0$. Since $Tu_1, \ldots, Tu_n$ are linearly independent in V_2, we have $\alpha_1 = 0, \ldots, \alpha_k = 0$. Thus, $x = 0$. Thus, we have shown that $Tx = 0$ implies $x = 0$. Hence, T is one-one. ∎

The above theorem need not be true if the spaces involved are infinite dimensional.

EXAMPLE 2.14 Let V be the vector space of all sequences. Define $T_1 : V \to V$ and $T_2 : V \to V$ by

$$T_1(\alpha_1, \alpha_2, \ldots) = (0, \alpha_1, \alpha_2, \ldots), \quad (\alpha_1, \alpha_2, \ldots) \in V,$$

$$T_2(\alpha_1, \alpha_2, \ldots) = (\alpha_2, \alpha_3, \alpha_4, \ldots), \quad (\alpha_1, \alpha_2, \ldots) \in V.$$

We observe that

- T_1 and T_2 are linear transformations;
- T_1 is one-one, but not onto;
- T_2 is onto, but not one-one.

The above T_1 is called a *right shift operator* and T_2 is called a *left shift operator* on V.

Theorem 2.5 *If V_1 and V_2 are finite dimensional vector spaces, then*

$$V_1 \simeq V_2 \iff \dim(V_1) = \dim(V_2).$$

Theorem 2.6 *Let V and W be finite dimensional vector spaces and $T : V \to V$ be a linear transformation. Let V_0 ne a subspace of V such that $V = N(T) \oplus V_0$. Then $V_0 \simeq R(T)$.*

Proof. Let $\{u_1, \ldots, u_k\}$ be a basis of $N(T)$, and let $\{v_1, \ldots, v_\ell\}$ be a basis of V_0. Define $T_0 : V_0 \to R(T)$ by

$$T_0\Big(\sum_{i=1}^{\ell} \alpha_i v_i\Big) = \sum_{i=1}^{\ell} \alpha_i T v_i, \quad (\alpha_1, \ldots, \alpha_\ell) \in \mathbb{F}^\ell.$$

It can be easily verified (*verify!*) that T_0 is a bijective linear transformation. ∎

As a corollary to the above theorem, we have the following.

Theorem 2.7 (rank-nullity theorem) *Let V and W be vector spaces and $T : V \to W$ be a linear transformation. Then*

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

***Definition* 2.7** A linear transformation $T : V \to W$ is said to be of **finite rank** if $\text{rank}\, T < \infty$. ♢

Exercise 2.7 Let $T : V_1 \to V_2$ be a linear transformation between vector spaces V_1 and V_2. Show that T is of finite rank if and only if there exists $n \in \mathbb{N}$, $\{v_1, \ldots, v_n\} \subset V_2$ and $\{f_1, \ldots, f_n\} \subset \mathcal{L}(V_1, \mathbb{F})$ such that $Ax = \sum_{j=1}^{n} f_j(x) v_j$ for all $x \in V_1$. ♦

2.5.1 Product and Inverse

The following theorem can be proved easily (*Exercise*).

Theorem 2.8 *Let V_1, V_2, V_3 be vector spaces, and let $T_1 : V_1 \to V_2$ and $T_2 : V_2 \to V_3$ be linear transformations. Then $T : V_1 \to V_3$ defined by*

$$T(x) = T_2(T_1(x)), \quad x \in V_1$$

is a linear transformation.

***Definition* 2.8** Let T_1, T_2, T be as in Theorem 2.8. Then the linear transformation T is called the **product** of T_2 and T_1, and it is denoted by T_2T_1. ♢

Note that

$$T_1 \in \mathcal{L}(V_1, V_2),\ T_2 \in \mathcal{L}(V_2, V_3) \Rightarrow T_2T_1 \in \mathcal{L}(V_1, V_3)$$

and if $V_1 = V_2 = V_3 = V$, then T_1T_2 and T_2T_1 are well-defined and they belong to $\mathcal{L}(V)$.

Theorem 2.9 *Let V_1, V_2, V_3 be finite dimensional vector spaces, and let $T_1 : V_1 \to V_2$ and $T_2 : V_2 \to V_3$ be linear transformations. Let E_1, E_2, E_3 be bases of V_1, V_2, V_3, respectively. Then*

$$[T_2T_1]_{E_3,E_1} = [T_2]_{E_3,E_2}[T_1]_{E_2,E_1}.$$

Proof. Let $E_1 := \{u_1, \ldots, u_n\}$, $E_2 := \{v_1, \ldots, v_m\}$, $E_3 := \{w_1, \ldots, w_k\}$ be ordered bases of V_1, V_2, V_3, respectively. Let

$$[T_1]_{E_2,E_1} = (a_{ij}), \quad [T_2]_{E_3,E_2} = (b_{ij}).$$

That is,

$$T_1u_j = \sum_{i=1}^{m} a_{ij}v_i, \quad T_2v_j = \sum_{i=1}^{k} b_{ij}w_i.$$

Then,

$$\begin{aligned} T_2T_1u_j &= T_2\Big(\sum_{\ell=1}^{m} a_{\ell j}v_\ell\Big) = \sum_{\ell=1}^{m} a_{\ell j}T_2v_\ell \\ &= \sum_{\ell=1}^{m} a_{\ell j}\sum_{i=1}^{k} b_{i\ell}w_i = \sum_{i=1}^{k}\Big(\sum_{\ell=1}^{m} b_{i\ell}a_{\ell j}\Big)w_i. \end{aligned}$$

Thus, $[T_2T_1]_{E_3,E_1} = (c_{ij})$, where $c_{ij} = \sum_{\ell=1}^{m} b_{i\ell}a_{\ell j}$. Hence, $[T_2T_1]_{E_3,E_1}$ is the matrix $[T_2]_{E_3,E_2}[T_1]_{E_2,E_1}$. ∎

Let V be a finite dimensional vector space and let $E = \{u_1, \ldots, u_n\}$ be an ordered basis. Then we know that, for every $x \in V$, there exists a unique $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ such that $x = \sum_{i=1}^{n} \alpha_j u_j$. Let us use the notation $[x]_E$ for the column vector with entries $\alpha_1, \ldots, \alpha_n$, that is,

$$[x]_E := [\alpha_1, \ldots, \alpha_n]^T.$$

We observe (verify!) that for $x, y \in V$ and $\alpha \in \mathbb{K}$,

$$[x] = 0 \iff x = 0, \quad [x + y]_E = [x]_E + [y]_E, \quad [\alpha x]_E = \alpha[x]_E.$$

In fact, the map $x \mapsto [x]_E$ is a linear isomorphism from V onto $\mathbb{K}^n$.

Now, if W is a finite dimensional vector space and $F = \{v_1, \ldots, v_m\}$ is an ordered basis of W, and $T : V \to W$ is a linear transformation, then we see (verify!) that

$$[Tx]_F = [T]_{F,E}[x]_E \quad \forall\, x \in V.$$

Hence, it follows that

$$[Tu_j]_F = [T]_{F,E}[u_j]_E = [T]_{F,E}e_j^T \quad \forall\, j = 1, \ldots, n.$$

Thus, $[Tu_j]_F$ is the j-th column of $[T]F, E$.

Theorem 2.10 *Let V and W be a finite dimensional vector space and let $E = \{u_1, \ldots, u_n\}$ and $F = \{v_1, \ldots, v_m\}$ be an ordered bases of V and W, respectively, and let $T : V \to W$ be a linear transformation. Then T is one-one if and only if columns of $[T]_{F,E}$ are linearly independent.*

Proof. Recall that T is one-one iff $N(T) = 0$. Hence, T is not one one-one iff $\exists\, x \neq 0$ in V such that $Tx = 0$ iff $\exists\, (\alpha_1, \ldots, \alpha_n) \neq 0$ such that $T(\alpha_1 u_1 + \cdots + \alpha_n u_n) = 0$. But,

$$[T(\alpha_1 u_1 + \cdots + \alpha_n u_n)]_F = \alpha_1 [T(u_1)]_F + \cdots + \alpha_n [T(u_n)]_F.$$

Thus, T is not one one-one iff $[T(u_1)]_F, \ldots, [T(u_n)]_F$ are linearly dependent in $\mathbb{K}^m$, that is, iff the columns of $[T]_{F,E}$ are linearly dependent. ∎

Theorem 2.11 *Let V and W be finite dimensional vector spaces, and $A : V \to W$ be linear transformation. Then A is bijective if and only if there exists linear transformation $B : W \to V$ such that*

$$BA = I_V, \quad AB = I_W,$$

where I_V and I_W are the identity operators on V and W, respectively.

Proof. Suppose A is bijective. Let $B : W \to V$ be defined by $By = x$ for $y \in W$, where $x \in V$ is the unique vector in V such that $Ax = y$. Then, it can be seen that B is a linear transformation satisfying

$$BAx = x, \quad ABy = y \quad \forall\, x \in V,\, y \in W.$$

If $\tilde{B} : W \to V$ is a linear transformation such that

$$\tilde{B}Ax = x, \quad A\tilde{B}y = y \quad \forall\, x \in V,\, y \in W,$$

then for every $y \in W$, if $x \in V$ is the unique vector in V such that $Ax = y$, then

$$\tilde{B}y = \tilde{B}Ax = x = BAx = By.$$

Hence, $\tilde{B} = B$.

Conversely, suppose $B : W \to V$ is a transformation such that

$$BAx = x, \quad ABy = y \quad \forall\, x \in V,\, y \in W.$$

Then, for $x \in V$,

$$Ax = 0 \Rightarrow x = BAx = 0$$

so that A is one-one. Also, for every $y \in W$, $ABy = y$ so that A is onto as well. ∎

***Definition* 2.9** Let $A : V \to W$ be a bijective linear transformation. Then the unique linear transformation $B : W \to V$ obtained as in Theorem 2.11 satisfying

$$BA = I_V, \quad AB = I_W,$$

is called the **inverse** of A, and it is denoted by A^{-1}. ♢

We observe that if $T_1 : V_1 \to V_2$ and $T_2 : V_2 \to V_3$ are bijective linear transformations, then

$$(T_2T_1)^{-1} = T_1^{-1}T_2^{-1}.$$

Exercise 2.8 Prove the last statement.

Theorem 2.12 *Let V and W be finite dimensional vector spaces, and let $A : V \to W$ be linear transformation. Then we have the following.*

(1) *A is one-one if and only if there exists a linear transformation $B : W \to V$ such that $BA = I_V$.*

(2) *A is onto if and only if there exists a linear transformation $B : W \to V$ such that $AB = I_W$.*

Proof. (1) Suppose A is one-one. Let $A_0 : V \to R(A)$ be defined by $A_0x = Ax$ for all $x \in V$. Then, it can be see that A_0 is a bijective linear transformation. Let W_0 be a subspace of W such that $W = R(T) \oplus W_0$. Define $B : W \to V$ by

$$By = A_0^{-1}y_0, \quad y \in W,$$

where $y_0 \in R(A)$ is the unique vector such that $y - y_0 \in W_0$. Then we have

$$BAx = A_0^{-1}Ax = A_0^{-1}A_0x = x \quad \forall\, x \in V.$$

(2) Suppose A is onto. Let V_0 be a subspace of V such that $V = N(T) \oplus V_0$. Define $A_0 : V_0 \to W$ by $A_0 x = Ax$ for all $x \in V$. Then, it can be see that A_0 is a bijective linear transformation. Define $B : W \to V$ by

$$By = A_0^{-1} y, \quad y \in W.$$

Then we have

$$ABy = AA_0^{-1} y = A_0 A_0^{-1} y = y \quad \forall\, y \in W.$$

This completes the proof. ∎

Having defined product of operators, we can define **powers of operators**.

***Definition* 2.10** For $A \in \mathcal{L}(V)$ and $n \in \mathbb{N}$, $A^n : V \to V$ is defined inductively by

$$A^n(x) = A\left(A^{n-1}(x)\right), \qquad x \in X,$$

where $A^0(x) := x$ for every $x \in V$. ◊

Using powers of operators, we can define **polynomials of an operator**.

***Definition* 2.11** For $A \in \mathcal{L}(V)$ and for a polynomial $p \in \mathcal{P}_n$, say $p(t) = a_1 + a_1 t + \cdots + a_n t^n$, the operator $p(A) : V \to V$ is defined by $p(A) = a_1 I + a_1 A + \cdots + a_n A^n$. ◊

Exercise 2.9 Let V and W be finite dimensional vector spaces of the same dimension, and let $T : V \to W$ be a linear transformation. Let E and F be ordered bases of V and W, respectively. Show that Show $\det[T]_{F,E}$ is independent of the bases E and F.

2.5.2 Change of basis

Let V be a finite dimensional vector space and let E and F be any two ordered bases of V. For $x \in V$, how are the vectors $[x]_E$ and $[x]_F$ related?

We know that there exists (λ_{ij}) such that

$$v_j = \sum_{i=1}^{n} \lambda_{ij} u_i, \quad j = 1, \ldots, n. \qquad (*)$$

If $x = \sum_{j=1}^n \alpha_j u_j = \sum_{j=1}^n \beta_j v_j$, then we have

$$\sum_{i=1}^n \alpha_i u_i = \sum_{j=1}^n \beta_j v_j = \sum_{j=1}^n \beta_j \Big(\sum_{i=1}^n \lambda_{ij} u_i \Big) = \sum_{i=1}^n \Big(\sum_{j=1}^n \lambda_{ij} \beta_j \Big) u_i.$$

Hence,

$$\alpha_i = \sum_{j=1}^n \lambda_{ij} \beta_j.$$

Thus,

$$[x]_E = J[x]_F, \quad \text{where} \quad J := (\lambda_{ij}).$$

From $(*)$, we also observe that

$$Iv_j = \sum_{i=1}^n a_{ij} u_i, \quad j = 1, \ldots, n$$

so that

$$[I]_{E,F} = (a_{ij}) = J, \quad [x]_E = [I]_{E,F}[x]_F.$$

The above matrix is the matrix corresponding the change of basis E to F, called the **change of basis matrix**.

Now, let V and W be finite dimensional vector spaces and let $T : V \to W$ be a linear transformation. Let E and $\tilde{E}$ be ordered bases of V, and F and $\tilde{F}$ be ordered bases of W. How the matrices $[T]_{F,E}$ and $[T]_{\tilde{F},\tilde{E}}$ are related?

Note that

$$[T]_{\tilde{F},\tilde{E}} = [I_2]_{\tilde{F},F}[T]_{F,E}[I_1]_{E,\tilde{E}},$$

where I_1 and I_2 are the identity operators on V and W, respectively. Also,

$$[I_1]_{\tilde{E},E}[I_1]_{E,\tilde{E}} = [I_1]_{\tilde{E},\tilde{E}}, \qquad [I_2]_{\tilde{F},F}[I_2]_{F,\tilde{F}} = [I_2]_{\tilde{F},\tilde{F}}.$$

It can be seen that, for any basis E of a finite dimensional vector space, $[I]_{EE}$ is the identity matrix (δ_{ij}). Hence,

$$[I_1]^{-1}_{E,\tilde{E}} = [I_1]_{\tilde{E},E}, \qquad [I_2]^{-1}_{\tilde{F},F} = [I_2]_{F,\tilde{F}},$$

so that

$$[T]_{F,E} = [I_2]_{F,\tilde{F}}[T]_{\tilde{F},\tilde{E}}[I_1]_{\tilde{E},E}.$$

Thus, if

$$E = \{u_1, \ldots, u_n\}, \quad \tilde{E} = \{\tilde{u}_1, \ldots, \tilde{u}_n\},$$

$$F = \{v_1, \ldots, v_m\}, \quad \tilde{F} = \{\tilde{v}_1, \ldots, \tilde{v}_m\},$$

then find the matrices J_1 and J_2 corresponding to the linear transformations

$$u_j \mapsto \tilde{u}_j \quad \text{and} \quad v_j \mapsto \tilde{v}_j$$

on V and W, respectively, then

$$[T]_{\tilde{F},\tilde{E}} = J_2^{-1}[T]_{F,E}J_1.$$

In particular, if $V = W$, $E = F$, $\tilde{E} = \tilde{F}$, then we have

$$[T]_{\tilde{E},\tilde{E}} = J_1^{-1}[T]_{E,E}J_1.$$

***Definition* 2.12** Matrices $A \in \mathbb{K}^{m\times n}$ and $B \in \mathbb{K}^{m\times n}$ are said to be **equivalent** if there exists invertible matrices $P \in \mathbb{K}^{n\times n}$ and $Q \in \mathbb{K}^{m\times m}$ such that

$$B = Q^{-1}AP.$$

Square matrices $A \in \mathbb{K}^{n\times n}$ and $B \in \mathbb{K}^{n\times n}$ are said to be **similar** if there exists an invertible matrix $P \in \mathbb{K}^{n\times n}$ such that

$$B = P^{-1}AP.$$

$\diamondsuit$

Thus, if $T : V \to V$ is a linear transformation on a finite dimensional vector space V and if E and F are bases of V, then $[T]_{EE}$ is similar to $[T]_{FF}$.

2.5.3 Eigenvalues and Eigenvectors

Let $A : V \to X$ be a linear operator on a vector space V.

A scalar λ is called an **eigenvalue** of A if there exists a nonzero vector $x \in X$ such that

$$Ax = \lambda x,$$

and in that case, x is called an **eigenvector** of A corresponding to the eigenvalue λ.

The set of all eigenvectors of A corresponding to an eigenvalue, together with the zero vector, is called an **eigenspace** of A, and the set of all eigenvalues of A is called the **eigenspectrum** of A.

We denote the eigenspectrum of A by $\sigma_{\text{eig}}(A)$.

Thus, $\lambda \in \mathbb{F}$ is an eigenvalue of A if and only if $A - \lambda I$ is not injective, and in that case, $N(A-\lambda I)$ is the corresponding eigenspace of A.

EXAMPLE 2.15 The conclusions in (i)-(vi) below can be verified easily:

(i) Let $A : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$A : (\alpha_1, \alpha_2, \alpha_3) \mapsto (\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3)$$

Then $\sigma_{\rm eig}(A) = \{1\}$ and $N(A - I) = \text{span}\,\{(0, 0, 1)\}$.

(ii) Let $A : \mathbb{F}^2 \to \mathbb{F}^2$ be defined by $A : (\alpha_1, \alpha_2) \mapsto (\alpha_2, -\alpha_1)$. If $\mathbb{F} = \mathbb{R}$, then A has no eigenvalues, i.e., $\sigma_{\rm eig}(A) = \varnothing$.

(iii) Let A be as in (ii) above. If $\mathbb{F} = \mathbb{C}$, then $\sigma_{\rm eig}(A) = \{i, -i\}$, $N(A - iI) = \text{span}\,\{(1, i)\}$ and $N(A + iI) = \text{span}\,\{(1, -i)\}$.

(iv) Let $V = c_{00}$, and let (λ_n) be a sequence of scalars. Let $A : V \to X$ be defined by

$$(Ax)(i) = \lambda_i x(i) \quad \forall\, x \in X,\ i \in \mathbb{N}.$$

Then $\sigma_{\rm eig}(A) = \{\lambda_1, \lambda_2, \ldots\}$, and for each $j \in \mathbb{N}$, e_j is an eigenvector corresponding to the eigenvalue λ_j. In case $\lambda_1, \lambda_2, \ldots$ are distinct, then $N(A - \lambda_j I) = \text{span}\,\{e_j\}$ for all $j \in \mathbb{N}$. Here, $e_j \in c_{00}$ is such that $e_j(i) = \delta_{ij}$, for $i, j \in \mathbb{N}$.

(v) Let $A : \mathcal{P} \to \mathcal{P}$ be defined by

$$(Ax)(t) = tx(t), \qquad x \in \mathcal{P}.$$

Then $\sigma_{\rm eig}(A) = \varnothing$.

(vi) Let X be $\mathcal{P}[a, b]$ and $A : V \to X$ be defined by

$$(Ax)(t) = \frac{d}{dt}x(t), \qquad x \in \mathcal{P}.$$

Then $\sigma_{\rm eig}(A) = \{0\}$ and $N(A) = \text{span}\,\{x_0\}$, where $x_0(t) = 1$ for all $t \in [a, b]$.

Existence of an eigenvalue

From the above examples we observe that in those cases in which the eigenspectrum is empty, either the scalar field is $\mathbb{R}$ or the vector space is infinite dimensional. The next result shows that if the space is finite dimensional and if the scalar field is the set of all complex numbers, then the eigenspectrum is nonempty.

Theorem 2.13 *Let X be a finite dimensional vector space over $\mathbb{C}$. Then every linear operator on X has at least one eigenvalue.*

Proof. Let X be an n-dimensional vector space over $\mathbb{C}$, and $A : V \to X$ be a linear operator. Let x be a nonzero element in X. Since $\dim X = n$, the set $\{x, Ax, A^2x, \ldots, A^nx\}$ is linearly dependent. Let $a_0, a_1 \ldots, a_n$ be scalars with at least one of them being nonzero such that

$$a_0x + a_1Ax + \cdots + a_nA^nx = 0.$$

Let $k = \max\{j : a_j \neq 0,\, j = 1, \ldots, n\}$. Then writing

$$p(t) = a_0 + a_1t + \cdots + a_kt^k, \quad p(A) = a_0I + a_1A + \cdots + a_kA^k,$$

we have

$$p(A)(x) = 0.$$

By fundamental theorem of algebra, there exist $\lambda_1, \ldots, \lambda_k$ in $\mathbb{C}$ such that

$$p(t) = a_k(t - \lambda_1)(t - \lambda_2) \ldots (t - \lambda_k).$$

Thus, we have

$$(A - \lambda_1I)(A - \lambda_2I) \ldots (A - \lambda_kI)(x) = p(A)(x) = 0.$$

The above relation shows that at least one of $A - \lambda_1I, \ldots, A - \lambda_kI$ is not injective so that at least one of $\lambda_1, \ldots, \lambda_k$ is an eigenvalue of A. ∎

Theorem 2.14 *Let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues of a linear operator $A : V \to X$ with corresponding eigenvectors $u_1, \ldots, u_n$, respectively. Then the set $\{u_1, \ldots, u_n\}$ is linearly independent.*

Proof. We prove this result by induction. The result is obvious if $n = 1$. Hence, we consider the case of $n > 1$. Let $k \in \mathbb{N}$ be such that $k < n$, and assume that $\{u_1, \ldots, u_k\}$ is linearly independent. We have to show that $\{u_1, \ldots, u_k, u_{k+1}\}$ is linearly independent. For scalars $c_1, \ldots, c_k, c_{k+1}$, let

$$x = c_1u_1 + \cdots + c_ku_k + c_{k+1}u_{k+1}.$$

We have to show that, if $x = 0$, then $c_j = 0$ for $j = 1, \ldots, k + 1$.

We note that

$$Ax = c_1\lambda_1u_1 + \cdots + c_k\lambda_ku_k + c_{k+1}\lambda_{k+1}u_{k+1}$$

so that

$$Ax - \lambda_{k+1}x = c_1(\lambda_1 - \lambda_{k+1})u_1 + \cdots + c_k(\lambda_{k+1} - \lambda_k)u_k.$$

Now suppose that $x = 0$. Then we have $Ax - \lambda_{k+1}x = 0$, i.e.,

$$c_1(\lambda_1 - \lambda_{k+1})u_1 + \cdots + c_k(\lambda_k - \lambda_{k+1})u_k = 0.$$

From this, using the fact that $\{u_1, \ldots, u_k\}$ is linearly independent in X, and $\lambda_1, \ldots, \lambda_k, \lambda_{k+1}$ are distinct, it follows that $c_j = 0$ for $j = 1, \ldots, k$. Therefore, $0 = x = c_{k+1}u_{k+1}$ so that $c_{k+1} = 0$. This completes the proof. ∎

By the above theorem we can immediately infer that if V is finite dimensional, then the eigenspectrum of every linear operator on X is a finite set.

3

Inner Product Spaces

3.1 Motivation

In Chapter 1 we defined a vector space as an abstraction of the familiar Euclidian space. In doing so, we took into account only two aspects of the set of vectors in a plane, namely, the vector addition and scalar multiplication. Now, we consider the third aspect, namely the *angle* between vectors.

Recall from plane geometry that if $\vec{x} = (x_1, x_2)$ and $\vec{y} = (y_1, y_2)$ are two *non-zero* vectors in the plane $\mathbb{R}^2$, then the angle $\theta_{x,y}$ between $\vec{x}$ and $\vec{y}$ is given by

$$\cos \theta_{x,y} := \frac{x_1 y_1 + x_2 y_2}{|\vec{x}|\,|\vec{y}|},$$

where for a vector $\vec{u} = (u_1, u_2) \in \mathbb{R}^2$, $|\vec{u}|$ denotes the *absolute value* of the vector $\vec{u}$, i.e.,

$$|\vec{u}| := \sqrt{u_1^2 + u_2^2},$$

which is the distance of the point $(u_1, u_2) \in \mathbb{R}^2$ from the coordinate origin.

We may observe that the angle $\theta_{x,y}$ between the vectors $\vec{x}$ and $\vec{y}$ is completely determined by the quantity $x_1 y_1 + x_2 y_2$, which is the *dot product* of $\vec{x}$ and $\vec{y}$. Breaking the convention, let us denote this quantity, i.e., the *dot product* of $\vec{x}$ and $\vec{y}$, by $\langle \vec{x}, \vec{y} \rangle$, i.e.,

$$\langle \vec{x}, \vec{y} \rangle = x_1 y_1 + x_2 y_2.$$

A property of the function $(\vec{x}, \vec{y}) \mapsto \langle \vec{x}, \vec{y} \rangle$ that one notices immediately is that, for every fixed $\vec{y} \in \mathbb{R}^2$, the function

$$x \mapsto \langle \vec{x}, \vec{y} \rangle, \qquad \vec{x} \in \mathbb{R}^2,$$

is a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}$, i.e.,

$$\langle \vec{x} + \vec{u}, \vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{u}, \vec{y} \rangle, \qquad \langle \alpha \vec{x}, \vec{y} \rangle = \alpha \langle \vec{x}, \vec{y} \rangle \tag{3.1}$$

for all $\vec{x}, \vec{u}$ in $\mathbb{R}^2$. Also, we see that for all $\vec{x}$, $\vec{y}$ in $\mathbb{R}^2$,

$$\langle \vec{x}, \vec{x})\rangle \geq 0, \tag{3.2}$$

$$\langle \vec{x}, \vec{x} \rangle = 0 \iff \vec{x} = \vec{0}, \tag{3.3}$$

$$\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle. \tag{3.4}$$

If we take $\mathbb{C}^2$ instead of $\mathbb{R}^2$, and if we define $\langle \vec{x}, \vec{y} \rangle = x_1 y_1 + x_2 y_2$, for $\vec{x}, \vec{y}$ in $\mathbb{C}^2$, then the above properties are not satisfied by all vectors in $\mathbb{C}^2$. In order to accommodate the complex situation, we define a *generalized dot product*, as follows: For $\vec{x}$, $\vec{y}$ in $\mathbb{F}^2$, let

$$\langle \vec{x}, \vec{y} \rangle_* = x_1 \bar{y}_1 + x_2 \bar{y}_2,$$

where for a complex number z, $\bar{z}$ denotes its complex conjugation. It is easily seen that $\langle \cdot, \cdot \rangle_*$ satisfies properties (3.1) – (3.4).

Now, we shall consider the abstraction of the above *modified dot product*.

3.2 Definition and Some Basic Properties

***Definition* 3.1 (Inner Product)** An *inner product* on a vector space V is a map $(x, y) \mapsto \langle x, y \rangle$ which associates each pair (x, y) of vectors in V, a unique scalar $\langle x, y \rangle$ which satisfies the following axioms:

(a) $\langle x, x \rangle \geq 0 \quad \forall\, x \in V$,

(b) $\langle x, x \rangle = 0 \iff x = 0$,

(c) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \qquad \forall\, x, y, z \in V$,

(d) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle \quad \forall\, \alpha \in \mathbb{F}$ and $\forall\, x, y \in V$, and

(e) $\langle x, y \rangle = \overline{\langle y, x \rangle} \quad \forall\, x, y \in V$.

***Definition* 3.2 (Inner Product Space)** A vector space together with an inner product is called an *inner product space*.

If an inner product $\langle \cdot, \cdot \rangle$ is defined on a vector space V, and if V_0 is a subspace of V, then the restriction of $\langle \cdot, \cdot \rangle$ to $V_0 \times V_0$, i.e., the map $(x, y) \mapsto \langle x, y \rangle$ for $(x, y) \in V_0 \times V_0$ is an inner product on V_0.

Before giving examples of inner product spaces, let us observe some properties of an inner product.

Proposition 3.1 *Let V be an inner product space. For a given $y \in V$, let $f : V \to \mathbb{F}$ be defined by*

$$f(x) = \langle x, y \rangle, \qquad x \in V.$$

Then f is a linear functional on V.

Proof. The result follows from axioms (c) and (d) in the definition of an inner product: Let $x, x' \in V$ and $\alpha \in \mathbb{F}$. Then, by axioms (c) and (d),

$$f(x + x') = \langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle = f(x) + fx'),$$

$$f(\alpha x) = \langle \alpha x, y \rangle = \alpha \langle x, y \rangle = \alpha f(x).$$

Hence, f is a linear transformation. ∎

Proposition 3.2 *Let V be an inner product space. Then for every x, y, u, v in V, and for every $\alpha \in \mathbb{F}$,*

$$\langle x, u + v \rangle = \langle x, u \rangle + \langle x, v \rangle, \qquad \langle x, \alpha y \rangle = \bar{\alpha} \langle x, y \rangle.$$

Proof. The result follows from axioms (c),(d) and (e) in the definition of an inner product: Let x, y, u, v in V and $\alpha \in \mathbb{F}$.

$$\langle x, u+v \rangle = \overline{\langle u + v, x \}} = \overline{\langle u, x \rangle + \langle v, x \rangle} = \overline{\langle u, x \rangle} + \overline{\langle v, x \rangle} = \langle x, u \rangle + \langle x, v \rangle,$$

$$\langle x, \alpha y \rangle = \overline{\langle \alpha y, x \rangle} = \overline{\alpha \langle y, x \rangle} = \bar{\alpha} \langle x, y \rangle.$$

This completes the proof. ∎

3.3 Examples of Inner Product Spaces

EXAMPLE 3.1 For $x = (\alpha_1, \ldots, \alpha_n)$ and $y = (\beta_1, \ldots, \beta_n)$ in $\mathbb{F}^n$, define

$$\langle x, y \rangle = \sum_{j=1}^{n} \alpha_j \bar{\beta}_j.$$

It is seen that $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{F}^n$.

The above inner product is called the **standard inner product** on $\mathbb{F}^n$.

EXAMPLE 3.2 Suppose V is a finite dimensional vector space, say of dimension n, and $E := \{u_1, \ldots, u_n\}$ is an *ordered basis* of V. For $x = \sum_{i=1}^n \alpha_i u_i$, $y = \sum_{i=1}^n \beta_i u_i$ in V, let

$$\langle x, y\rangle_E := \sum_{i=1}^n \alpha_i \bar{\beta}_i.$$

Then it is easily seen that $\langle \cdot, \cdot\rangle_E$ is an inner product on V.

More generally, if $T : V \to \mathbb{F}^n$ is a linear isomorphism, then

$$\langle x, y\rangle_T := \langle Tx, Ty\rangle_{\mathbb{F}^n}$$

defines an inner product on V. Here, $\langle \cdot, \cdot\rangle_{\mathbb{F}^n}$ is the standard inner product on $\mathbb{F}^n$.

EXAMPLE 3.3 For $f, g \in C[a, b]$, let

$$\langle f, g\rangle := \int_a^b f(t)\overline{g(t)}\,dt.$$

This defines an inner product on $C[a, b]$: Clearly,

$$\langle f, f\rangle = \int_a^b |f(t)|^2\,dt \geq 0 \quad \forall f \in C[a, b],$$

and by continuity of the function f,

$$\langle f, f\rangle := \int_a^b |f(t)|^2\,dt = 0 \iff f(t) = 0 \quad \forall\, t \in [a, b].$$

The other axioms can be verified easily.

EXAMPLE 3.4 Let $\tau_1, \ldots, \tau_{n+1}$ be distinct real numbers. For $p, q \in \mathcal{P}_n$, let

$$\langle p, q\rangle := \sum_{i=1}^{n+1} p(\tau_i)\overline{q(\tau_i)}.$$

This defines an inner product on $\mathcal{P}_n$: Clearly,

$$\langle p, p\rangle = \sum_{i=1}^{n+1} |p(\tau_i)|^2 \geq 0 \quad \forall p \in \mathcal{P}_n,$$

and by the fact that a nonzero polynomial in $\mathcal{P}_n$ cannot have more than n distinct zeros, it follows that

$$\langle p, p\rangle := \sum_{i=1}^{n+1} |p(\tau_i)|^2 = 0 \iff p = 0.$$

The other axioms can be verified easily.

3.4 Norm of a Vector

Recall that the absolute value of a vector $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$, is given by

$$|\vec{x}| = \sqrt{x_1^2 + x_2^2}$$

Denoting the standard inner product on $\mathbb{R}^2$ by $\langle x, x\rangle_2$, it follows that

$$|\vec{x}| = \sqrt{\langle x, x\rangle_2}.$$

As an abstraction of the above notion, we define the *norm* of a vector.

***Definition* 3.3 (Norm of a Vector)** Let V be an inner product space. Then for $x \in V$, then *norm* of x is defined as the non-negative square root of $\langle x, x\rangle$, and it is denoted by $\|x\|$, i.e,

$$\|x\| := \sqrt{\langle x, x\rangle}, \qquad x \in V.$$

Exercise 3.1 If x is a non-zero vector, then show that $u := x/\|x\|$ is a vector of norm 1. ♦

Recall from elementary geometry that if a, b are the lengths of the adjacent sides of a parallelogram, and if c, d are the lengths of its diagonals, then $2(a^2 + b^2) = c^2 + d^2$. This is the well-known parallelogram law. This has a generalized version in the setting of inner product spaces.

Theorem 3.3 (Parallelogram law) *For vectors x, y in an inner product space V,*

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Exercise 3.2 Verify the parallelogram law (Theorem 3.3). ♦

3.5 Orthogonality and Orthonormal Bases

Recall that the angle $\theta_{x,y}$ between vectors $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^2$ is given by

$$\cos\theta_{x,y} := \frac{\langle \vec{x}, \vec{y}\rangle_2}{|\vec{x}|\,|\vec{y}|}.$$

Hence, we can conclude that the vectors $\vec{x}$ and $\vec{y}$ are *orthogonal* if and only if $\langle \vec{x}, \vec{y}\rangle_2 = 0$. This observation motivates us to have the following definition.

***Definition* 3.4 (Orthogonal vectors)** Vectors x and y in an inner product space V are said to be *orthogonal to each other* or x *is orthogonal to* y if $\langle x, y \rangle = 0$. In this case we write $x \perp y$, and read x perpendicular to y, or x perp y.

Note that

(a) for x, y in V, x is orthogonal to y if and only if y is orthogonal to x, i.e., $x \perp y \iff y \perp x$, and

(b) the zero vector is orthogonal to every vector, i.e., $0 \perp x$ for all $x \in V$.

Theorem 3.4 *Let V be an inner product space, and $x \in V$. If $\langle x, y \rangle = 0$ for all $y \in V$, then $x = 0$.*

Proof. Clearly, if $\langle x, y \rangle = 0$ for all $y \in V$, then $\langle x, x \rangle = 0$ as well. Hence $x = 0$ ∎

As an immediate consequence of the above theorem, we have the following.

Corollary 3.5 *Let V be an inner product space, and $u_1, u_2, \ldots, u_n$ be linearly independent vectors in V. Let $x \in V$. Then*

$$\langle x, u_i \rangle = 0 \quad \forall\, i \in \{1, \ldots, n\} \iff \langle x, y \rangle = 0 \quad \forall\, y \in \operatorname{span}\{u_1, \ldots, u_n\}.$$

In particular, if $\{u_1, u_2, \ldots, u_n\}$ is a basis of V, and if $\langle x, u_i \rangle = 0$ for all $i \in \{1, \ldots, n\}$, then $x = 0$.

Exercise 3.3 If $\dim V \geq 2$, and if $0 \neq x \in V$, then find a non-zero vector which is orthogonal to x. ♦

EXAMPLE 3.5 Consider the standard inner product on $\mathbb{F}^n$. For each $j \in \{1, \ldots, n\}$, let

$$e_j = (\delta_{1j}, \delta_{2j}, \ldots, \delta_{nj}), \quad \text{where} \quad \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

It is easily seen that $e_i \perp e_j$ for every $i \neq j$. Also, $e_i + e_j \perp e_i - e_j$ for every $i, j \in \{1, \ldots, n\}$.

EXAMPLE 3.6 Consider the the vector space $C[0, 2\pi]$ with inner product defined by

$$\langle f, g \rangle := \int_0^{2\pi} f(t)\overline{g(t)}\, dt$$

for $f, g \in C[0, 2\pi]$. For $n \in \mathbb{N}$, let

$$u_n(t) := \sin(nt), \quad v_n(t) = \cos(nt), \qquad 0 \le t \le 2\pi.$$

Since

$$\int_0^{2\pi} \cos(kt)\, dt = 0 = \int_0^{2\pi} \sin(kt)\, dt \quad \forall\, k \in \mathbb{Z},$$

it follows that, for $n \neq m$,

$$\langle u_n, u_m \rangle = \langle v_n, v_m \rangle = \langle u_n, v_n \rangle = \langle u_n, v_m \rangle = 0.$$

Recall from elementary geometry that if a, b c are lengths of sides of a right angled triangle with c being the hypotenuse, then $a^2 + b^2 = c^2$. This is the *Pythagoras theorem* Here is the generalized form of it in the setting of an inner product space.

Theorem 3.6 (Pythagoras theorem) *Suppose x and y are vectors in an inner product space which are orthogonal to each other. Then*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Proof. Left as an exercise. ∎

It is easily seen that, if the scalar field is $\mathbb{R}$, then the converse of the Pythagoras theorem also holds. That is,

Theorem 3.7 *If V is an inner product space over $\mathbb{R}$, and if $x, y \in V$ are such that $\|x + y\|^2 = \|x\|^2 + \|y\|^2$, then $x \perp y$.*

However, if the scalar field is $\mathbb{C}$, then the converse of Pythagoras theorem need not be true . A simple example shows this: Let $X = \mathbb{C}$ with standard inner product, and for nonzero real numbers $\alpha, \beta \in \mathbb{R}$, let $x = \alpha$, $y = i\,\beta$. Then we have

$$\|x + y\|^2 = \|\alpha + i\,\beta\|^2 = |\alpha|^2 + |\beta|^2 = \|x\|^2 + \|y\|^2,$$

but $\langle x, y \rangle = -i\,\alpha\beta \neq 0$.

***Definition* 3.5 (Orthogonal to a set)** Suppose S is a subset of an inner product space V, and $x \in S$. Then x is said to be *orthogonal to S* if $\langle x, y \rangle = 0$ for all $y \in S$. In this case, we write $x \perp S$. The set of vectors orthogonal to S is denoted by $S^\perp$, i.e.,

$$S^\perp := \{x \in V : x \perp S\}.$$

Exercise 3.4 Let V be an inner product space.
(a) Show that $V^{\perp} = \{0\}$.
(b) If S is a basis of V, then show that $S^{\perp} = \{0\}$. ♦

***Definition* 3.6 (Orthogonal set)** Suppose S is a subset of an inner product space V. Then S is said to be an *orthogonal set* if $\langle x, y\rangle = 0$ for all distinct $x,\ y \in S$, i.e., for every $x, y \in S$, $x \neq y$ implies $x \perp y$.

Theorem 3.8 *Let S be an orthogonal set in an inner product space V. If $0 \notin S$, then S is linearly independent.*

Proof. Suppose $0 \notin S$ and $\{u_1, \ldots, u_n\} \subseteq S$. If $\alpha_1, \ldots, \alpha_n$ are scalars such that $\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n = 0$, then for every $j \in \{1, \ldots, n\}$, we have

$$0 = \left\langle \sum_{i=1}^{n} \alpha_i u_i, u_j \right\rangle = \sum_{i=1}^{n} \langle \alpha_i u_i, u_j\rangle = \sum_{i=1}^{n} \alpha_i \langle u_i, u_j\rangle = \alpha_j \langle u_j, u_j\rangle.$$

Hence, $\alpha_j = 0$ for all $j \in \{1, \ldots, n\}$. ∎

***Definition* 3.7 (Orthonormal set)** Suppose S is a subset of an inner product space V. Then S is said to be an *orthonormal set* if it is an orthogonal set and $\|x\| = 1$ for all $x \in S$.

By Theorem 3.8, it follows that every orthonormal set is linearly independent. In particular, if V is an n-dimensional inner product space and E is an orthonormal set consisting of n vectors, then E is a basis of V.

***Definition* 3.8 (Orthonormal basis)** Suppose V is a finite dimensional inner product space. An orthonormal set in V which is also a basis of V is called an orthonormal basis of V.

EXAMPLE 3.7 The set $E := \{e_1, e_2, \ldots, e_n\}$ in Example 3.5 is an orthonormal basis of $\mathbb{F}^n$ (with respect to the standard inner product).

Theorem 3.9 *Suppose V is an inner product space, and $\{u_1, \ldots, u_n\}$ is an orthonormal subset of V. Then, for every $x \in \operatorname{span}\{u_1, \ldots, u_n\}$,*

$$x = \sum_{j=1}^{n} \langle x, u_j\rangle u_j, \qquad \|x\|^2 = \sum_{j=1}^{n} |\langle x, u_j\rangle|^2.$$

Proof. Let $x \in \text{span}\,\{u_1, \ldots, u_n\}$, Then there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$x = \alpha_1 u_1 + \cdots + \alpha_n u_n.$$

Hence, for every $i \in \{1, \ldots, n\}$,

$$\langle x, u_i \rangle = \alpha_1 \langle u_1, u_i \rangle + \cdots + \alpha_n \langle u_n, u_i \rangle = \alpha_i.$$

and

$$\begin{aligned} \|x\|^2 = \langle x, x \rangle &= \left\langle \sum_{i=1}^n \alpha_i u_i, \sum_{j=1}^n \alpha_j u_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \bar{\alpha}_j \langle u_i, u_j \rangle \\ &= \sum_{i=1}^n |\alpha_i|^2 = \sum_{i=1}^n |\langle x, u_i \rangle|^2. \end{aligned}$$

This completes the proof. ∎

The proof of the following corollary is immediate from the above theorem.

Corollary 3.10 (Fourier expansion and Parseval's identity) *If $\{u_1, \ldots, u_n\}$ is an orthonormal basis of an inner product space V , then for every $x \in V$,*

$$x = \sum_{j=1}^n \langle x, u_j \rangle u_j, \qquad \|x\|^2 = \sum_{j=1}^n |\langle x, u_j \rangle|^2.$$

Another consequence of Theorem 3.9 is the following.

Corollary 3.11 (Bessel's inequality) *Suppose V is an inner product space, and $\{u_1, \ldots, u_n\}$ is an orthonormal subset of V. Then, for every $x \in V$,*

$$\sum_{j=1}^n |\langle x, u_j \rangle|^2 \leq \|x\|^2.$$

Proof. Let $x \in V$, and let

$$y = \sum_{i=1}^n \langle x, u_i \rangle u_i.$$

Since $y \in \text{span}\,\{u_1, \ldots, u_n\}$, by Theorem 3.9,

$$\|y\|^2 = \sum_{i=1}^n |\langle y, u_i \rangle|^2.$$

Note that $\langle y, u_i \rangle = \langle x, u_i \rangle$ for all $i \in \{1, \ldots, n\}$, i.e., $\langle x-y, u_i \rangle = 0$ for all $i \in \{1, \ldots, n\}$. Hence, $\langle x - y, y \rangle = 0$. Therefore, by Pythagoras theorem,

$$\|x\|^2 = \|y\|^2 + \|x-y\|^2 \geq \|y\|^2 = \sum_{i=1}^{n} |\langle x, u_i \rangle|^2.$$

This completes the proof. ∎

EXAMPLE 3.8 Let $V = C[0, 2\pi]$ with inner product $\langle x, y \rangle := \int_0^{2\pi} x(t)\overline{y(t)}dt$ for x, y in $C[0, 2\pi]$. For $n \in \mathbb{Z}$, let u_n be defined by

$$u_n(t) = e^{i\,nt}, \qquad t \in [0, 2\pi].$$

Then it is seen that

$$\langle u_n, u_m \rangle = \int_0^{2\pi} e^{i\,(n-m)t} dt = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{if } n \neq m. \end{cases}$$

Hence, $\{u_n : n \in \mathbb{Z}\}$ is an orthonormal set in $C[0, 2\pi]$. By Theorem 3.9, if $x \in \operatorname{span}\{u_j : j = -N, -N+1, \ldots, 0, i, \ldots, N\}$,

$$x = \sum_{j=-N}^{N} a_n\, e^{i\,nt} \quad \text{with} \quad a_n = \frac{1}{2\pi} \int_0^{2\pi} x(t) e^{-i\,nt} dt.$$

Now, suppose that V is an n-dimensional inner product space, and $\{u_1, \ldots, u_n\}$ be an orthonormal basis of V. Then, by Corollary 3.10, every $x \in V$ can be written as

$$x = \sum_{j=1}^{n} \langle x, u_j \rangle u_j.$$

Hence, for every linear functional $f : V \to \mathbb{F}$,

$$\begin{aligned} f(x) &= \sum_{j=1}^{n} \langle x, u_j \rangle f(u_j) \\ &= \sum_{j=1}^{n} \langle x, \overline{f(u_j)} u_j \rangle \\ &= \left\langle x, \sum_{j=1}^{n} \overline{f(u_j)} u_j \right\rangle. \end{aligned}$$

Thus, we have given a constructive proof for the following theorem.

Theorem 3.12 (Riesz representation theorem) *Let V be a finite dimensional inner product space. Then for very linear functional $f : V \to \mathbb{F}$, there exists $y_f \in V$ such that*

$$f(x) = \langle x, y_f \rangle \quad \forall\, x \in V.$$

It is easily seen that the vector y_f in the above theorem is unique. Indeed, if y_1 and y_2 are in V such that

$$f(x) = \langle x, y_1 \rangle, \quad f(x) = \langle x, y_2 \rangle \quad \forall\, x \in V,$$

then

$$\langle x, y_1 - y_2 \rangle = 0 \quad \forall\, x \in V$$

so that by Theorem 3.4, $y_1 - y_2 = 0$, i.e., $y_1 = y_2$.

Exercise 3.5 Suppose V is an n-dimensional inner product space and $\{u_1, \ldots, u_n\}$ be an orthonormal basis of V. Show that every linear functional $f : V \to \mathbb{F}$ can be written as

$$f = \sum_{j=1}^{n} f(u_j) f_j,$$

where, for each $j \in \{1, \ldots, n\}$, $f_j : V \to \mathbb{F}$ is the linear functional defined by $f_j(x) = \langle x, u_j \rangle$, $x \in V$. ♦

3.6 Gram-Schmidt Orthogonalization

A question that naturally arises is: Does every finite dimensional inner product space has an orthonormal basis? We shall answer this question affirmatively.

Theorem 3.13 (Gram-Schmidt orthogonalization) *Let V be an inner product space and $u_1, u_2, \ldots, u_n$ are linearly independent vectors in V. Then there exist orthogonal vectors $v_1, v_2, \ldots, v_n$ in V such such that*

$$\operatorname{span}\{u_1, \ldots, u_k\} = \operatorname{span}\{v_1, \ldots, v_k\} \quad \forall\, k \in \{1, \ldots, n\}.$$

In fact, the vectors $v_1, v_2, \ldots, v_n$ defined by

$$\begin{aligned} v_1 &:= u_1 \\ v_{k+1} &:= u_{k+1} - \sum_{j=1}^{k} \frac{\langle u_{k+1}, v_j \rangle}{\langle v_j, v_j \rangle} v_j, \quad k = 1, 2, \ldots, n-1, \end{aligned}$$

satisfy the requirements.

Proof. We construct orthogonal vectors $v_1, v_2, \ldots, v_n$ in V such such that $\operatorname{span}\{u_1, \ldots, u_k\} = \operatorname{span}\{v_1, \ldots, v_k\}$ for all $k \in \{1, \ldots, n\}$.

Let $v_1 = u_1$. Let us write u_2 as

$$u_2 = \alpha u_1 + v_2,$$

where α is chosen in such a way that $v_2 := u_2 - \alpha u_1$ is orthogonal to v_1, i.e., $\langle u_2 - \alpha u_1, v_1\rangle = 0$, i.e.,

$$\alpha = \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle}.$$

Thus, the vector

$$v_2 := u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1$$

is orthogonal to v_1. Moreover, using the linearly independence of u_1, u_2, it follows that $v_2 \neq 0$, and $\operatorname{span}\{u_1, u_2\} = \operatorname{span}\{v_1, v_2\}$. Next, we write

$$u_3 = (\alpha_1 v_1 + \alpha_2 v_2) + v_3,$$

where α_1, α_2 are chosen in such a way that $v_3 := u_3 - (\alpha_1 v_1 + \alpha_2 v_2)$ is orthogonal to v_1 and v_2, i.e.,

$$\langle u_3 - (\alpha_1 v_1 + \alpha_2 v_2), v_1\rangle = 0, \qquad \langle u_3 - (\alpha_1 v_1 + \alpha_2 v_2), v_2\rangle = 0.$$

That is, we take

$$\alpha_1 = \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle}, \qquad \alpha_2 = \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle}.$$

Thus, the vector

$$v_3 := u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2$$

is orthogonal to v_1 and v_2. Moreover, using the linearly independence of u_1, u_2, u_3, it follows that $v_3 \neq 0$, and

$$\operatorname{span}\{u_1, u_2, u_3\} = \operatorname{span}\{v_1, v_2, v_3\}.$$

Continuing this procedure, we obtain orthogonal vectors $v_1, v_2, \ldots, v_n$ defined by

$$v_{k+1} := u_{k+1} - \frac{\langle u_{k+1}, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_{k+1}, v_2\rangle}{\langle v_2, v_2\rangle} v_2 - \ldots - \frac{\langle u_{k+1}, v_k\rangle}{\langle v_k, v_k\rangle} v_k$$

which satisfy

$$\text{span}\,\{u_1, \ldots, u_k\} = \text{span}\,\{v_1, \ldots, v_k\}$$

for each $k \in \{1, 2, \ldots, k-1\}$. ∎

Exercise 3.6 Let V be an inner product space, and let $u_1, u_2, \ldots, u_n$ be orthonormal vectors. Define $w_1, w_2, \ldots, w_n$ iteratively as follows:

$$v_1 := u_1 \quad \text{and} \quad w_1 = \frac{v_1}{\|v_1\|}$$

and for each $k \in \{1, 2, \ldots, n-1\}$, let

$$v_{k+1} := u_{k+1} - \sum_{i=1}^{k} \langle u_{k+1}, w_i\rangle w_i \quad \text{and} \quad w_{k+1} = \frac{v_{k+1}}{\|v_{k+1}\|}.$$

Show that $\{w_1, w_2, \ldots, w_n\}$ is an orthonormal set, and

$$\text{span}\,\{w_1, \ldots, w_k\} = \text{span}\,\{u_1, \ldots, u_k\}, \qquad k = 1, 2, \ldots, n. \blacklozenge$$

From Theorem 3.13, we can conclude the following.

Theorem 3.14 *Every finite dimensional inner product space has an orthonormal basis.*

EXAMPLE 3.9 Let $V = \mathbb{F}^3$ with standard inner product. Consider the vectors $u_1 = (1, 0, 0)$, $u_2 = (1, 1, 0)$, $u_3 = (1, 1, 1)$. Clearly, u_1, u_2, u_3 are linearly independent in $\mathbb{F}^3$. Let us orthogonalize these vectors according to the Gram-Schmidth orthogonalizaion procedure:

Take $v_1 = u_1$, and

$$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1.$$

Note that $\langle v_1, v_1\rangle = 1$ and $\langle u_2, v_1\rangle = 1$. Hence, $v_2 = u_2 - v_1 = (0, 1, 0)$. Next, let

$$v_3 = u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2.$$

Note that $\langle v_2, v_2\rangle = 1$, $\langle u_3, v_1\rangle = 1$ and $\langle u_3, v_2\rangle = 1$ Hence, $v_3 = u_2 - v_1 - v_2 = (0, 0, 1)$. Thus,

$$\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

is the Gram-Schmidt orthogonalization of $\{u_1, u_2, u_3\}$.

EXAMPLE 3.10 Again let $V = \mathbb{F}^3$ with standard inner product. Consider the vectors $u_1 = (1,1,0)$, $u_2 = (0,1,1)$, $u_3 = (1,0,1)$. Clearly, u_1, u_2, u_3 are linearly independent in $\mathbb{F}^3$. Let us orthogonalize these vectors according to the Gram-Schmidth orthogonalizaion procedure:

Take $v_1 = u_1$, and

$$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1.$$

Note that $\langle v_1, v_1\rangle = 2$ and $\langle u_2, v_1\rangle = 1$. Hence,

$$v_2 = (0,1,1) - \frac{1}{2}(1,1,0) = (-1/2, 1/2, 1).$$

Next, let

$$v_3 = u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2.$$

Note that $\langle v_2, v_2\rangle = 3/2$, $\langle u_3, v_1\rangle = 1$ and $\langle u_3, v_2\rangle = 1/2$ Hence,

$$v_3 = (1,0,1) - \frac{1}{2}(1,1,0) - \frac{1}{3}(-1/2, 1/2, 1) = (-2/3, 2/3, -2/3).$$

Thus,

$$\{(1,1,0), (-1/2, 1/2, 1), (-2/3, 2/3, -2/3)\}$$

is the Gram-Schmidt orthogonalization of $\{u_1, u_2, u_3\}$.

EXAMPLE 3.11 Let $V = \mathcal{P}$ be with the the inner product

$$\langle p, q\rangle = \int_{-1}^{1} p(t)\overline{q(t)}\, dt, \qquad p, q \in V.$$

Let $u_j(t) = t^{j-1}$ for $j = 1,2,3$ and consider the linearly independent set $\{u_1, u_2, u_3\}$ in V. Now let $v_1(t) = u_1(t) = 1$ for all $t \in [-1, 1]$, and let

$$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1.$$

Note that

$$\langle v_1, v_1\rangle = \int_{-1}^{1} v_1(t)\overline{v_1(t)}\, dt = \int_{-1}^{1} dt = 2,$$

$$\langle u_2, v_1\rangle = \int_{-1}^{1} u_2(t)\overline{v_1(t)}\, dt = \int_{-1}^{1} t\, dt = 0.$$

Hence, we have $v_2(t) = u_2(t) = t$ for all $t \in [-1, 1]$. Next, let

$$v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2.$$

Here,

$$\langle u_3, v_1 \rangle = \int_{-1}^{1} u_3(t)\overline{v_1(t)}\, dt = \int_{-1}^{1} t^2\, dt = \frac{2}{3},$$

$$\langle u_3, v_2 \rangle = \int_{-1}^{1} u_3(t)\overline{v_2(t)}\, dt = \int_{-1}^{1} t^3\, dt = 0.$$

Hence, we have $v_3(t) = t^2 - \frac{1}{3}$ for all $t \in [-1, 1]$. Thus,

$$\left\{1, t, t^2 - \frac{1}{3}\right\}$$

is an orthogonal set of polynomials.

***Definition* 3.9 (Legendre polynomials)** The polynomials

$$p_o(t), p_1(t), p_2(t) \ldots$$

obtained by orthogonalizing $1, t, t^2, \ldots$ using the inner product

$$\langle p, q \rangle = \int_{-1}^{1} p(t)\overline{q(t)}\, dt, \qquad p, q \in \mathcal{P},$$

are called *Legendre polynomials.*

It is clear that the n-th Legendre polynomial $p_n(t)$ is of degree n. We have seen in Example 3.11 that

$$p_0(t) = 1, \quad p_1(t) = t, \quad p_2(t) = t^2 - \frac{1}{3}.$$

3.7 Cauchy-Schwarz Inequality and Its Consequences

Let us look at the arguments used in the construction of v_2 from u_1, u_2 in the proof of Theorem 3.13: Suppose x and y are two nonzero vectors. Then we can write x as sum of two orthogonal elements, namely, u and v, where

$$u = \frac{\langle x, y \rangle}{\langle y, y \rangle} y, \qquad v = x - \frac{\langle x, y \rangle}{\langle y, y \rangle} y.$$

The vector u can be thought of as the *projection* of the vector x onto the span of y. Using this argument we prove an important result, called *Cauchy-Schwarz inequality.*

Theorem 3.15 (Cauchy-Schwarz inequality) *Let V be an inner product space, and $x, y \in V$. Then*

$$|\langle x, y\rangle| \le \|x\| \; \|y\|.$$

Equality holds in the above inequality if and only if x and y are linearly dependent.

Proof. The result is obvious if either $x = 0$ or $y = 0$. Hence, assume that both x and y are nonzero vectors. As we have explained in the preceding paragraph, let us write $x = u + v$, where

$$u = \frac{\langle x, y\rangle}{\langle y, y\rangle} y, \qquad v = x - \frac{\langle x, y\rangle}{\langle y, y\rangle} y.$$

Then, by Pythagoras theorem,

$$\|x\|^2 = \|u\|^2 + \|v\|^2 = \frac{|\langle x, y\rangle|^2}{|\langle y, y\rangle|^2}\|y\|^2 + \|v\|^2 = \frac{|\langle x, y\rangle|^2}{\|y\|^2} + \|v\|^2.$$

Thus, $|\langle x, y\rangle| \le \|x\| \; \|y\|$. Equality holds in this inequality if and only if $v := x - \frac{\langle x,y\rangle}{\langle y,y\rangle} y = 0$, i.e., if and only if x is a scalar multiple of y if and only if x and y are linearly dependent. ∎

As a corollary of the above theorem we have the following.

Corollary 3.16 (Triangle inequality) *Suppose V is an inner product space. Then for every x, y in V,*

$$\|x + y\| \le \|x\| + \|y\|.$$

Proof. Let $x, y \in V$. Then, using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y\rangle \\
&= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\
&= \|x\|^2 + \|y\|^2 + 2 \operatorname{Re} \langle x, y\rangle \\
&\le \|x\|^2 + \|y\|^2 + 2 \, |\langle x, y\rangle| \\
&\le \|x\|^2 + \|y\|^2 + 2 \, \|x\| \, \|y\| \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

Thus, $\|x + y\| \le \|x\| + \|y\|$ for every $x, y \in V$. ∎

Exercise 3.7 Let V be an inner product space, and let $x, y \in V$. Then, show the following:

(a) $\|x\| \geq 0$.

(b) $\|x\| = 0$ iff $x = 0$.

(c) $\|\alpha x\| = |\alpha| \, \|x\|$ for all $\alpha \in \mathbb{F}$..

(d) If $\|x + y\| = \|x\| + \|y\|$, then either $y = 0$ or $x = \alpha y$ for some scalar α.

(e) $\|x + \alpha y\| = \|x - \alpha y\| \ \forall \alpha \in \mathbb{F}$ if and only if $\langle x, y \rangle = 0$. ♦

Remark 3.1 For nonzero vectors x and y in an inner product space V, by Schwarz inequality, we have

$$\frac{|\langle x, y \rangle|}{\|x\| \|y\|} \leq 1.$$

This relation motivates us to define the *angle* between any two nonzero vectors x and y in V as

$$\theta_{x,y} := \cos^{-1}\left(\frac{|\langle x, y \rangle|}{\|x\| \, \|y\|}\right).$$

Note that if $x = cy$ for some nonzero scalar c, then $\theta_{x,y} = 0$, and if $\langle x, y \rangle = 0$, then $\theta_{x,y} = \pi/2$.

***Definition* 3.10** Suppose V is a vector space. A function $\nu : V \to \mathbb{R}$ is called a *norm* on V if it satisfies the following axioms:

(a) $\nu(x) \geq 0$ for all $x \in V$, and $\nu(x) = 0$ iff $x = 0$,

(b) $\nu(x + y) \leq \nu(x) + \nu(y)$ for all $x, y \in V$, and

(c) $\nu(\alpha x) = |\alpha| \nu(x)$ for all $x \in V$ and for all $\alpha \in \mathbb{F}$.

Corollary 3.16 and Exercise 3.7 (a)-(c) shows that, in an inner product space V, the function $x \mapsto \|x\|$ is a norm.

We have seen that, in an inner product space V, the norm $\| \cdot \|$ satisfies the parallelogram law. A natural question is whether every norm ν on a vector space V satisfies parallelogram law:

$$\nu^2(x + y) + \nu^2(x - y) = 2[\nu^2(x) + \nu^2(y)]$$

The answer is, in fact, negative. To see this consider the following examples.

EXAMPLE 3.12 Let $V = \mathbb{R}^2$. For $x = (x_1, x_2) \in \mathbb{R}^2$, define $\nu(x) = |x_1| + x_2|$. Then it is easily seen that ν is a norm on $\mathbb{R}^2$. But it does not satisfy the parallelogram law: Note that for $x = e_1 + e_2$ and $y = e_1 - e_2$,

$$\nu(x+y) = 2, \quad \nu(x-y) = 2, \quad \nu(x) = \nu(y) = 2.$$

From these relations it follows that

$$\nu^2(x+y) + \nu^2(x-y) = 8, \quad \text{but} \quad 2[\nu^2(x) + \nu^2(y)] = 16.$$

Thus, the above norm does not satisfy the parallelogram law.

EXAMPLE 3.13 For $f \in C[0,1]$, let

$$\nu(f) = \int_0^1 |f(t)|dt.$$

Then it is easily seen that ν is a norm on $C[0,1]$. But it does not satisfy the parallelogram law: To see this consider

$$f(t) = t, \qquad g(t) = 1 - t \quad \forall\, t \in [a,b].$$

Then we have

$$\nu(f) = \nu(g) = \nu(f-g) = \frac{1}{2}, \qquad \nu(f+g) = 1$$

From these relations it follows that ν does not satisfy the parallelogram law.

3.8 Best Approximation

In applications one may come across functions which are too complicated to handle for computational purposes. In such cases, one would like to replace them by functions of "simpler forms" which are easy to handle. This is often done by approximating the given function by certain functions belonging to a finite dimensional space spanned by functions of simple forms. For instance, one may want to approximate a continuous function f defined on certain interval $[a, b]$ by a polynomial, say a polynomial p in $\mathcal{P}_n$ for some specified n. It is desirable to find that polynomial p such that

$$\|f - p\| \leq \|f - q\| \qquad \forall\, q \in \mathcal{P}_n.$$

Here, $\|.\|$ is a norm on $C[a, b]$. Now the question is whether such a polynomial exists, and if exists, then is it unique; and if there is a unique such polynomial, then how can we find it. These are the issues that we discuss in this section, in an abstract frame work of inner product spaces.

***Definition* 3.11** Let V be an inner product spaceand V_0 be a subspace of V. Let $x \in V$. A vector $x_0 \in V_0$ is a called a **best approximation** of x from V_0 if

$$\|x - x_0\| \leq \|x - v\| \qquad \forall\, v \in V_0.$$

Proposition 3.17 *Let V be an inner product space, V_0 be a subspace of V, and $x \in V$. If $x_0 \in V_0$ is such that $x - x_0 \perp V_0$, then x_0 is a best approximation of x, and it is the unique best approximation of x from V_0.*

Conversely, if $x_0 \in V_0$ is a best approximation of x, then $x - x_0 \perp V_0$.

Proof. Suppose $x_0 \in V_0$ is such that $x - x_0 \perp V_0$. Then, for every $u \in V_0$,

$$\begin{aligned} \|x - u\|^2 &= \|(x - x_0) + (x_0 - u)\|^2 \\ &= \|x - x_0\|^2 + \|x_0 - u\|^2. \end{aligned}$$

Hence

$$\|x - x_0\| \leq \|x - v\| \qquad \forall\, v \in V_0,$$

showing that x_0 is a best approximation.

To see the uniqueness, suppose that $v_0 \in V_0$ is another best approximation of x. Then, we have

$$\|x - x_0\| \le \|x - v_0\| \quad \text{and} \quad \|x - v_0\| \le \|x - x_0\|,$$

so that $\|x - x_0\| = \|x - v_0\|$. Therefore, using the fact that $\langle x - x_0, x_0 - v_0\rangle = 0$, we have

$$\|x - v_0\|^2 = \|x - x_0\|^2 + \|x_0 - v_0\|^2.$$

Hence, it follows that $\|x_0 - v_0\| = 0$. Thus $v_0 = x_0$.

Conversely, suppose that $x_0 \in V_0$ is a best approximation of x. Then $\|x - x_0\| \le \|x - u\|$ for all $u \in V_0$. In particular, if $v \in V_0$,

$$\|x - x_0\| \le \|x - (x_0 + \alpha v\| \quad \forall \alpha \in \mathbb{F}.$$

Hence, for every $\alpha \in \mathbb{F}$,

$$\begin{aligned} \|x - x_0\|^2 &\le \|x - (x_0 + \alpha v\|^2 \\ &= \langle (x - x_0) + \alpha v, (x - x_0) + \alpha v\rangle \\ &= \|x - x_0\|^2 - 2\mathrm{Re}\langle x - x_0, \alpha v\rangle + |\alpha|^2\|v\|^2. \end{aligned}$$

Taking $\alpha = \langle x - x_0, v\rangle / \|v\|^2$, we have

$$\langle x - x_0, \alpha v\rangle = \frac{|\langle x - x_0, v\rangle|^2}{\|v\|^2} = |\alpha|^2\|v\|^2$$

so that

$$\begin{aligned} \|x - x_0\|^2 &\le \|x - x_0\|^2 - 2\mathrm{Re}\langle x - x_0, \alpha v\rangle + |\alpha|^2\|v\|^2 \\ &= \|x - x_0\|^2 - \frac{|\langle x - x_0, v\rangle|}{\|v\|^2}. \end{aligned}$$

Hence, $\langle x - x_0, v\rangle = 0$. ∎

By the above proposition, in order to find a best approximation of $x \in V$ from V_0, it is enough to find a vector $x_0 \in V_0$ such that $x - x_0 \perp V_0$; and we know that such vector x_0 is unique.

Theorem 3.18 *Let V be an inner product space, V_0 be a finite dimensional subspace of V, and $x \in V$. Let $\{u_1, \ldots, u_n\}$ be an orthonormal basis of V_0. Then for $x \in V$, the vector*

$$x_0 := \sum_{i=1}^{n} \langle x, u_i\rangle u_i$$

is the unique best approximation of x from V_0.

Proof. Clearly, $x_0 := \sum_{i=1}^n \langle x, u_i\rangle u_i$ satisfies the hypothesis of Proposition 3.17. ∎

The above theorem shows how to find a best approximation from a finite dimensional subspace V_0, provided we know an orthonormal basis of V_0. Suppose we know only a basis of V_0. Then, we can find an orthonormal basis by Gram-Schmidt procedure. Another way to find a best approximation is to use Proposition 3.17:

Suppose $\{v_1, \ldots, v_n\}$ is a basis of V_0. By Proposition 3.17, the vector x_0 that we are looking for should satisfy $\langle x - x_0, v_i\rangle$ for every $i = 1, \ldots, n$. Thus, we have to find scalars $\alpha_1, \ldots, \alpha_n$ such that

$$\left\langle x - \sum_{j=1}^n \alpha_j v_j,\ v_i \right\rangle = 0 \qquad \forall\, i = 1, \ldots, n.$$

That is to find $\alpha_1, \ldots, \alpha_n$ such that

$$\sum_{j=1}^n \langle v_j,\ v_i\rangle \alpha_j = \langle x, v_i\rangle \qquad \forall\, i = 1, \ldots, n.$$

The above system of equations is uniquely solvable (*Why?*) to get $\alpha_1, \ldots, \alpha_n$. Note that if the basis $\{v_1, \ldots, v_n\}$ is an orthonormal basis basis of V_0, then $\alpha_j = \langle x, v_j\rangle$ for $j = 1, \ldots, n$.

Exercise 3.8 Show that, if $\{v_1, \ldots, v_n\}$ is a linearly independent subset of an inner product space V, then the columns of the matrix $M := (a_{ij})$ with $a_{ij} = \langle v_j, v_i\rangle$, are linearly independent. Deduce that, the matrix is invertible. ♦

EXAMPLE 3.14 Let $V = \mathbb{R}^2$ with usual inner product, and $V_0 = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1 = x_2\}$. Let us find the best approximation of $x = (0, 1)$ from V_0.

We have to find a vector of the form $x_0 = (\alpha, \alpha)$ such that $x - x_0 = (0, 1) - (\alpha, \alpha) = (-\alpha, 1 - \alpha)$ is orthogonal to V_0. Since V_0 is spanned by the single vector $(1, 1)$, the requirement is to find α such that $(-\alpha, 1 - \alpha)$ is orthogonal to $(1, 1)$, i.e., α has to satisfy the equation $-\alpha + (1 - \alpha = 0$, i.e., $\alpha = 1/2$. Thus the best approximation of $x = (0, 1)$ from V_0 is the vector $x_0 = (1/2, 1/2)$.

EXAMPLE 3.15 Let V be the vector space $C[0, 1]$ over $\mathbb{R}$ with the inner product: $\langle x, u\rangle = \int_0^1 x(t)u(t)dt$, and let $V_0 = \mathcal{P}_1$. Let us find the best approximation of x define by $x(t) = t^2$ from space V_0.

We have to find a vector x_0 of the form $x_0(t) = a_0 + a_1 t$ such that the function $x - x_0$ defined by $(x - x_0)(t) = t^2 - a_0 - a_1 t$ is orthogonal to V_0. Since V_0 is spanned by u_1, u_2 where $u_1(t) = 1$ and $u_2(t) = t$, the requirement is to find a_0, a_1 such that

$$\langle x - x_0, u_1 \rangle = \int_0^1 (t^2 - a_0 - a_1 t)dt = 0,$$

$$\langle x - x_0, u_2 \rangle = \int_0^1 (t^3 - a_0 t - a_1 t^2)dt = 0.$$

That is

$$\int_0^1 (t^2 - a_0 - a_1 t)dt = [t^3/3 - a_0 t - a_1 t^2/2]_0^1 = 1/3 - a_0 - a_1/2 = 0,$$

$$\int_0^1 (t^3 - a_0 t - a_1 t^2)dt = [t^4/4 - a_0 t^2/2 - a_1 t^3/3]_0^1 = 1/4 - a_0/2 - a_1/3 = 0.$$

Hence, $a_0 = -1/6$ and $a_1 = 1$, so that the best approximation x_0 of t^2 from $\mathcal{P}_1$ is given by $x_0(t) := -1/63 + t$.

Exercise 3.9 Let V be an inner product space and V_0 be a finite dimensional subspace of V. Show that for every $x \in V$, there exists a unique pair of vectors u, v with $u \in V_0$ and $v \in V_0^\perp$ satisfying $x = u + v$. In fact,

$$V = V_0 + V_0^\perp. \quad \blacklozenge$$

Exercise 3.10 Let $V = C[0,1]$ over $\mathbb{R}$ with inner product $\langle x, u \rangle = \int_0^1 x(t)u(t)dt$. Let $V_0 = \mathcal{P}_3$. Find best approximation for x from V_0, where $x(t)$ is given by

(i) e^t, (ii) $\sin t$, (iii) $\cos t$, (iv) t^4. ♦

3.9 Best Approximate Solution

In this section we shall make use of the results from the previous section to define and find a *best approximate solution* for an equation $Ax = y$ where $A : V_1 \to V_2$ is a linear transformation between vector spaces V_1 and V_2 with V_2 being an inner product space.

***Definition* 3.12** Let V_1 and V_2 be vector spaces with V_2 being an inner product space, and let $A : V_1 \to V_2$ be a linear transformation. Let $y \in V_2$. Then a vector $x_0 \in V_1$ is called a **best approximate solution** or **a least-square solution** of the equation $Ax = y$ if

$$\|Ax_0 - y\| \le \|Au - y\| \qquad \forall\, u \in V_1.$$

It is obvious that $x_0 \in V_1$ is a best approximate solution of $Ax = y$ if and only if $y_0 := Ax_0$ is a best approximation of y from the range space $R(A)$. Thus, from Proposition 3.17, we can conclude the following.

Theorem 3.19 *Let V_1 and V_2 be vector spaces with V_2 being an inner product space, and let $A : V_1 \to V_2$ be a linear transformation. If $R(A)$ is a finite dimensional subspace of V_2, then the equation $Ax = y$ has a best approximate solution. Moreover, a vector $x_0 \in V_1$ is a best approximate solution if and only if $Ax_0 - y$ is orthogonal to $R(A)$.*

Clearly, a best approximate solution is unique if and only if A is injective.

Next suppose that $A \in \mathbb{R}^{m\times n}$, i.e., A is an $m \times n$ matrix of real entries. Then we know that range space of A, viewing it as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, is the space spanned by the columns of A. Suppose $u_1, \ldots, u_n$ be the columns of A. Then, given $y \in \mathbb{R}^m$, a vector $x_0 \in \mathbb{R}^n$ is a best approximate solution of $Ax = y$ if and only if $Ax_0 - y$ is orthogonal to u_i for $i = 1, \ldots, n$, i.e., if and only if $u_i^T(Ax_0 - y) = 0$ for $i = 1, \ldots, n$, i.e., if and only if $A^T(Ax_0 - y) = 0$, i.e., if and only if

$$A^T A x_0 = A^T y.$$

EXAMPLE 3.16 Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and let $y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Clearly, the equation $Ax = y$ has no solution. It can be seen that $x_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is a solution of the equation $A^T Ax = A^T y$. Thus, x_0 is a best approximate solution of $Ax = y$.

3.10 QR-Factorization and Best Approximate Solution

Suppose that $A \in \mathbb{R}^{m\times n}$, i.e., A is an $m \times n$ matrix of real entries with $n \le m$. Assume that the columns of A are linearly independent. Then we know that, if the equation $Ax = y$ has a solution, then the solution is unique. Now, let $u_1, \ldots, u_n$ be the columns of A, and let $v_1, \ldots, v_n$ are orthonormal vectors obtained by orthonormalizing $u_1, \ldots, u_n$. Hence, we know that for each $k \in \{1, \ldots, n\}$,

$$\operatorname{span}\{u_1, \ldots, u_k\} = \operatorname{span}\{v_1, \ldots, v_k\}.$$

Hence, there exists an upper triangular $n \times n$ matrix $R := (a_{ij})$ such that $u_j = a_{1j}v_1 + a_{2j}v_2 + \ldots + a_{nj}v_j, \qquad j = 1, \ldots, n$. Thus,

$$[u_1, u_2, \ldots, u_n] = [v_1, v_2, \ldots, v_n]R.$$

Note that $A = [u_1, u_2, \ldots, u_n]$, and the matrix $Q := [v_1, v_2, \ldots, v_n]$ satisfies the relation

$$Q^T Q = I.$$

***Definition* 3.13** The factorization $A = QR$ with columns of Q being orthonormal and R being an upper triangular matrix is called a **QR-factorization** of A.

We have see that if columns of $A \in \mathbb{R}^{m \times n}$ are linearly independent, then A has a QR-factorization.

Now, suppose that $A \in \mathbb{R}^{m \times n}$ with columns of A are linearly independent, and $A = QR$ is the QR-factorization of A. Let $y \in \mathbb{R}^m$. Since columns of A are linearly independent, the equation $Ax = y$ has a unique best approximate solution, say x_0. Then we know that

$$A^T A x_0 = A^T y.$$

Using the QR-factorization $A = QR$ of A, we have

$$R^T Q^T Q R x_0 = R^T Q^T y.$$

Now, $Q^T Q = I$, and R^T is injective, so that it follows that

$$R x_0 = Q^T y.$$

Thus, if $A = QR$ is the QR-factorization of A, then the best approximate solution of $Ax = y$ is obtained by solving the equation

$$Rx = Q^T y.$$

For more details on best approximate solution one may see
http://mat.iitm.ac.in/~mtnair/LRN-Talk.pdf

4

Error Bounds and Stability of Linear Systems

4.1 Norms of Vectors and Matrices

Recall that a norm $\|\cdot\|$ on a vector space V is a function which associates each $x \in V$ a unique non-negative real number $\|x\|$ such that the following hold:

(a) For $x \in V$, $\|x\| = 0 \iff x = 0$

(b) $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$,

(c) $\|\alpha x\| = |\alpha| \, \|x\| \quad \forall \alpha \in \mathbb{F}$, $x \in V$.

We have already seen that if V is an inner product space, then the function $x \mapsto \|x\| := \langle x, x\rangle^{1/2}$ is a norm on V. It can be easily sen that for $x = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k$,

$$\|x\|_1 := \sum_{j=1}^{k} |x_j|, \qquad \|x\|_\infty := \max_{1 \leq i \leq k} |x_i|$$

define norms on $\mathbb{R}^k$. The norm induced by the standard inner product on $\mathbb{R}^k$ is denoted by $\|\cdot\|_2$, i.e.,

$$\|x\|_2 := \Big(\sum_{j=1}^{k} |x(t)|^2\Big)^{1/2}.$$

Exercise 4.1 Show that $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$ for every $x \in \mathbb{R}^k$. Compute $\|x\|_\infty$, $\|x\|_2$, $\|x\|_1$ for $x = (1, 1, 1) \in \mathbb{R}^3$.

We know that on $C[a, b]$,

$$\|x\|_2 := \langle x, x\rangle^{1/2} = \Big(\int_a^b |x(t)|^2 dt\Big)^{1/2}$$

defines a norm. It is easy to show that

$$\|x\|_1 := \int_a^b |x(t)|dt \qquad \|x\|_\infty := \max_{a \le b} |x(t)|$$

also define norms on $C[a,b]$.

Exercise 4.2 Show that there exists no constant $c > 0$ such that $\|x\|_\infty \le c\,\|x\|_1$ for all $x \in C[a,b]$.

Next we consider norms of matrices. Considering an $n \times n$ matrix as an element of $\mathbb{R}^{n^2}$, we can obtain norms of matrices. Thus, analogues to the norms $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$ on $\mathbb{R}^n$, for $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, the quantities

$$\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|, \qquad \Big(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2\Big)^{1/2}, \qquad \max_{1\le i,j\le n} |a_{ij}|$$

define norms on $\mathbb{R}^{n\times n}$.

Given a vector norm $\|\cdot\|$ on $\mathbb{R}^n$, it can be seen that

$$\|A\| := \sup_{\|x\|\le 1} \|Ax\|, \qquad A \in \mathbb{R}^{n\times n},$$

defines a norm on the space $\mathbb{R}^{n\times n}$. Since this norm is associated with the norm of the space $\mathbb{R}^n$, and since a matrix can be considered as a linear operator on $\mathbb{R}^n$, the above norm on $\mathbb{R}^{n\times n}$ is called a *matrix norm associated with a vector norm.*

The above norm has certain important properties that other norms may not have. For example, it can be seen that

- $\|Ax\| \le \|A\|\,\|x\| \quad \forall x \in \mathbb{R}^n$,
- $\|Ax\| \le c\,\|x\| \quad \forall x \in \mathbb{R}^n \;\Rightarrow\; \|A\| \le c,.$

Moreover, if $A, B \in \mathbb{R}^{n\times n}$ and if I is the identity matrix, then

- $\|AB\| \le \|A\|\,\|B\|, \qquad \|I\| = 1.$

Exercise 4.3 Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ and and $A \in \mathbb{R}^{n\times n}$. Suppose $c > 0$ is such that $\|Ax\| \le c\,\|x\|$ for all $x \in \mathbb{R}^n$, and there exists $x_0 \ne 0$ in $\mathbb{R}^n$ such that $\|Ax_0\| = c\,\|x_0\|$. Then show that $\|A\| = c$.

In certain cases operator norm can be computed from the knowledge of the entries of the matrix. Let us denote the matrix norm associated with $\|\cdot\|_1$ and $\|\cdot\|_\infty$ by the same notation, i.e., for $p \in \{1, \infty\}$,

$$\|A\|_p := \sup_{\|x\|_p \leq 1} \|Ax\|_p, \qquad A \in \mathbb{R}^{n\times n}.$$

Theorem 4.1 *If* $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, *then*

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|, \qquad \|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}|.$$

Proof. Note that for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$\begin{aligned}\|Ax\|_1 &= \sum_{i=1}^n \Big|\sum_{j=1}^n a_{ij}x_j\Big| \leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|\,|x_j| \\ &= \sum_{j=1}^n \Big(\sum_{i=1}^n |a_{ij}|\Big)|x_j| \leq \Big(\max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|\Big) \sum_{j=1}^n |x_j|.\end{aligned}$$

Thus, $\|A\|_1 \leq \max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|$. Also, note that $\|Ae_j\|_1 = \sum_{i=1}^n |a_{ij}|$ for every $j \in \{1, \ldots, n\}$ so that $\sum_{i=1}^n |a_{ij}| \leq \|A\|_1$ for every $j \in \{1, \ldots, n\}$. Hence, $\max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}| \leq \|A\|_1$. Thus, we have shown that

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|.$$

Next, consider the norm $\|\cdot\|_\infty$ on $\mathbb{R}^n$. In this case, for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we have

$$\|Ax\|_\infty = \max_{1\leq i\leq n} \Big|\sum_{j=1}^n a_{ij}x_j\Big|.$$

Since

$$\Big|\sum_{j=1}^n a_{ij}x_j\Big| \leq \sum_{j=1}^n |a_{ij}|\,|x_j| \leq \|x\|_\infty \sum_{j=1}^n |a_{ij}|,$$

it follows that

$$\|Ax\|_\infty \leq \Big(\max_{1`i\leq n} \sum_{j=1}^n |a_{ij}|\Big)\|x\|_\infty.$$

From this we have $\|A\|_\infty \leq \max_{1`i\leq n} \sum_{j=1}^{n} |a_{ij}|$. Now, let $i_0 \in \{1,\ldots,n\}$ be such that $\max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| = \sum_{j=1}^{n} |a_{i_0j}|$, and let $x_0 = (\alpha_1,\ldots,\alpha_n)$ be such that $\alpha_j = \begin{cases} |a_{i_0j}|/a_{i_0j} & \text{if}\, a_{i_0j} \neq 0, \\ 0 & \text{if}\, a_{i_0j} \neq 0. \end{cases}$ Then $\|x_0\|_\infty = 1$ and

$$\sum_{j=1}^{n} |a_{i_0j}| = \Big|\sum_{j=1}^{n} a_{i_0j}\alpha_j\Big| = |(Ax_0)_{i_o}| \leq \|Ax_0\|_\infty \leq \|A\|_\infty.$$

Thus, $\max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| = \sum_{j=1}^{n} |a_{i_0j}| \leq \|A\|_\infty$. Thus we have proved that

$$\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|.$$

This completes the proof of the theorem. ∎

What about the matrix norm

$$\|A\|_2 := \max_{\|x\|\leq 1} \|Ax\|_2, \quad A \in \mathbb{R}^{n\times n},$$

induced by $\|\cdot\|_2$ on $\mathbb{R}^n$? In fact, there is no simple representation for this in terms of the entries of the matrix. However, we have the following.

Theorem 4.2 *Suppose $A = (a_{ij}) \in \mathbb{R}^{n\times n}$. Then*

$$\|A\|_2 \leq \Big(\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2\Big)^{1/2}.$$

If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the (non-negative) eigenvalues of the matrix A^TA, then

$$\|A\|_2 = \max_{1\leq J\leq n} \sqrt{\lambda_j}.$$

Proof. Using the Cauchy-Schwarz inequality on $\mathbb{R}^n$, we have, for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$\begin{aligned} \|Ax\|_2^2 &= \sum_{i=1}^n \Big| \sum_{j=1}^n a_{ij} x_j \Big|^2 \\ &\leq \sum_{i=1}^n \Big[\Big(\sum_{j=1}^n |a_{ij}|^2 \Big) \Big(\sum_{j=1}^n |x_j|^2 \Big) \Big] \\ &\leq \Big(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \Big) \|x\|_2^2. \end{aligned}$$

Thus, $\|A\|_2 \leq \Big(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \Big)^{1/2}$.

Since $A^T A$ is a symmetric matrix, it has n real eigenvalues (may be some of the are repeated) with corresponding orthonormal eigenvectors $u_1, u_n, \ldots, u_n$. Note that, for every $j \in \{1, 2, \ldots, n\}$,

$$\lambda_j = \lambda_j \langle u_j, u_j \rangle = \langle \lambda_j u_j, u_j \rangle = \langle A^T A u_j, u_j \rangle = \langle A u_j, A u_j \rangle = \|A u_j\|^2$$

so that λ_j' s are non-negative, and $|\lambda_j| \leq \|A\|$ for all j. Thus,

$$\max_{1 \leq \mathrm{j} \leq n} \sqrt{\lambda_j} \leq \|A\|.$$

To see the reverse inequality, first we observe that $u_1, u_n, \ldots, u_n$ form an orthonormal basis of $\mathbb{R}^n$. Hence, every $x \in \mathbb{R}^n$ can be written as $x = \sum_{j=1}^n \langle x, u_j \rangle u_j$, so that

$$A^T A x = \sum_{j=1}^n \langle x, u_j \rangle A^T A u_j = \sum_{j=1}^n \langle x, u_j \rangle \lambda_j u_j.$$

Thus, we have $\|Ax\|^2 = \langle Ax, Ax \rangle = \langle ATAx, x \rangle$ so that

$$\begin{aligned} \|Ax\|^2 &= \Big\langle \sum_{j=1}^n \langle x, u_j \rangle \lambda_j u_j, \sum_{i=1}^n \langle x, u_i \rangle u_i \Big\rangle \\ &= \sum_{j=1}^n |\langle x, u_j \rangle|^2 \lambda_j \\ &\leq \Big(\max_{1 \leq j \leq n} \lambda_j \Big) \|x\|^2. \end{aligned}$$

Hence, $\|A\|_2 \leq \max_{1 \leq j \leq n} \sqrt{\lambda_j}$. This completes the proof. ∎

Exercise 4.4 Find $\|A\|_1$, $\|A\|_\infty$, for the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 2 & 1 \end{bmatrix}$.

4.2 Error Bounds for System of Equations

Given an invertible matrix $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$, consider the equation

$$Ax = b.$$

Suppose the data b is not known exactly, but a perturbed data $\tilde{b}$ is known. Let $\tilde{x} \in \mathbb{R}^n$ be the corresponding solution, i.e.,

$$A\tilde{x} = \tilde{b}.$$

Then, we have $x - \tilde{x} = A^{-1}(b - \tilde{b})$ so that

$$\|x-\tilde{x}\| \le \|A^{-1}\|b-\tilde{b}\| = \|A^{-1}\|b-\tilde{b}\|\frac{\|Ax\|}{\|b\|} \le \|A\|\,\|A^{-1}\|\frac{\|b-\tilde{b}\|}{\|b\|}\|x\|,$$

$$\|b-\tilde{b}\| \le \|A\|\,\|x-\tilde{x}\| = \|A\|\,\|x-\tilde{x}\|\frac{\|A^{-1}b\|}{\|x\|} \le \|A\|\,\|A^{-1}\|\frac{\|x-\tilde{x}\|}{\|x\|}\|b\|.$$

Thus, denoting the quantity $\|A\|\,\|A^{-1}\|$ by $\kappa(A)$,

$$\frac{1}{\kappa(A)}\frac{\|b-\tilde{b}\|}{\|b\|} \le \frac{\|x-\tilde{x}\|}{\|x\|} \le \kappa(A)\frac{\|b-\tilde{b}\|}{\|b\|}. \tag{4.1}$$

From the above inequalities, it can be inferred that if $\kappa(A)$ is large, then it can happen that for small relative error $\|b-\tilde{b}\|/\|b\|$ in the data, the relative error $\|x-\tilde{x}\|/\|x\|$ in the solution may be large. In fact, there do exist $b, \tilde{b}$ such that

$$\frac{\|x-\tilde{x}\|}{\|x\|} = \kappa(A)\frac{\|b-\tilde{b}\|}{\|b\|},$$

where $x, \tilde{x}$ are such that $Ax = b$ and $A\tilde{x} = \tilde{b}$. To see this, let x_0 and u be vectors such that

$$\|Ax_0\| = \|A\|\,\|x_0\|, \qquad \|A^{-1}u\| = \|A^{-1}\|\,\|u\|,$$

and let

$$b := Ax_0, \quad \tilde{b} := b + u, \quad \tilde{x} := x_0 + A^{-1}u.$$

Then it follows that $A\tilde{x} = \tilde{b}$ and

$$\frac{\|x_0 - \tilde{x}\|}{\|x_0\|} = \frac{\|A^{-1}u\|}{\|x_0\|} = \frac{\|A^{-1}\|\,\|u\|}{\|x_0\|} = \frac{\|A\|\,\|A^{-1}\|\,\|u\|}{\|Ax_0\|} = \kappa(A)\frac{\|b - \tilde{b}\|}{\|b\|}.$$

The quantity $\kappa(A) := \|A\|\,\|A^{-1}\|$ is called the **condition number** of the matrix A. To illustrate the observation in the preceding paragraph, let us consider

$$A = \begin{bmatrix} 1 & 1+\varepsilon \\ 1-\varepsilon & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

It can be seen that

$$A^{-1} = \frac{1}{\varepsilon^2}\begin{bmatrix} 1 & -1-\varepsilon \\ -1+\varepsilon & 1 \end{bmatrix} \quad \text{so that} \quad x = A^{-1}b = -\frac{1}{\varepsilon}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

From this, it is clear that, if ε is small, then for small $\|b\|$, $\|x\|$ can be very large. In this case, it can be seen that

$$\|A\|_\infty = 2+\varepsilon, \quad \|A^{-1}\|_\infty = \frac{1}{\varepsilon^2}(2=\varepsilon), \quad \kappa(A) = \Big(\frac{2+\varepsilon}{\varepsilon}\Big)^2 > \frac{4}{\varepsilon^2}.$$

In practice, while solving $Ax = b$ by numerically, we obtain an approximate solution $\tilde{x}$ in place of the actual solution. One would like to know how much error incurred by this procedure. We can have inference on this from (4.1), by taking $\tilde{b} := A\tilde{x}$.

Exercise 4.5 Let $A \in \mathbb{R}^{n\times n}$ be an invertible matrix. Then there exist vectors x, u such that $\|Ax_0\| = \|A\|\,\|x_0\|$ and $\|A^{-1}u\| = \|A^{-1}\|\,\|u\|$ – Justify.

Exercise 4.6

1. Suppose A, B in $\mathbb{R}^{n\times n}$ are invertible matrices, and $b, \tilde{b}$ are in $\mathbb{R}^n$. Let $x, \tilde{x}$ are in $\mathbb{R}^n$ be such that $Ax = b$ and $B\tilde{x} = \tilde{b}$. Show that
$$\frac{\|x - \tilde{x}\|}{\|x\|} \le \|A\|\,\|B^{-1}\|\Big(\frac{\|A-B\|}{\|A\|} + \frac{\|b-\tilde{b}\|}{\|b\|}\Big).$$
[Hint: Use the fact that $B(x - \tilde{x}) = (B - A)x + (b - \tilde{b})$, and use the fact that $\|(B - A)x\| \le \|B - A\|\,\|x\|$, and $\|b - \tilde{b}\| = \|b - \tilde{b}\|\|Ax\|/\|b\| \le \|b - \tilde{b}\|\|A\|\,\|x\|/\|b\|$.]

2. Let $B \in \mathbb{R}^{n\times n}$. If $\|B\| < 1$, then show that $I - B$ is invertible, and $\|(I - B)^{-1}\| \le 1/(1 - \|B\|)$.
[Hint: Show that $I - B$ is injective, by showing that for every x, $\|(I - B)x\| \ge (1 - \|B\|)\|x\|$, and then deduce the result.]

3. Let $A, B \in \mathbb{R}^{n\times n}$ be such that A is invertible, and $\|A - B\| < 1/\|A^{-1}\|$. Then, show that, B is invertible, and

$$\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A - B\|\,\|A^{-1}\|}.$$

[Hint: Observe that $B = A - (A - B) = [I - (A - B)A^{-1}]A$, and use the previous exercise.]

4. Let $A, B \in \mathbb{R}^{n\times n}$ be such that A is invertible, and $\|A - B\| < 1/2\|A^{-1}\|$. Let $b, \tilde{b}, x, \tilde{x}$ be as in Exercise 1. Then, show that, B is invertible, and

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq 2\kappa(A)\Big(\frac{\|A - B\|}{\|A\|} + \frac{\|b - \tilde{b}\|}{\|b\|}\Big).$$

[Hint: Apply conclusion in Exercise 3 to that in Exercise 1.]

5

Additional Exercises

In the following V denotes a vector space over $\mathbb{F}$ which is $\mathbb{R}$ or $\mathbb{C}$.

1. Let V be a vector space. For $x, y \in V$, show that $x + y = x$ implies $y = \theta$.

2. Suppose that $x \in V$ is a nonzero vector. Then show that $\alpha x \neq \beta x$ for every $\alpha, \beta \in \mathbb{F}$ with $\alpha \neq \beta$.

3. Let $\mathcal{R}[a, b]$ be the set of all real valued Riemann integrable functions on $[a, b]$. Show that $\mathcal{R}[a, b]$ is a vector space over $\mathbb{R}$.

4. Let V be the set of all polynomials of degree 3. Is it a vector space with respect to the usual addition and scalar multiplication?

5. Let S be a nonempty set, $s_0 \in S$. Show that the set V of all functions $f : S \to \mathbb{R}$ such that $f(s_0) = 0$ is a vector space with respect to the usual addition and scalar multiplication of functions.

6. Find a bijective linear transformation between $\mathbb{F}^n$ and $\mathcal{P}_{n+1}$.

7. In each of the following, a set set is given and some operations are defined. Check whether V is a vector space with these operations:

   (i) Let $V = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}$ with addition and scalar multiplication as in $\mathbb{R}^2$.

   (ii) Let $V = \{x = (x_1, x_2) \in \mathbb{R}^2 : 2x_1 + 3x_2 = 0\}$ with addition and scalar multiplication as in $\mathbb{R}^2$.

   (iii) Let $V = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 = 1\}$ with addition and scalar multiplication as for $\mathbb{R}^2$.

(iv) Let $V = \mathbb{R}^2$, $\mathbb{F} = \mathbb{R}$. For $x = (x_1, x_2)$, $y = (y_1, y_2)$, let $x + y := (x_1 + y_1, x_2 + y_2)$ and for all $\alpha \in \mathbb{R}$,

$$\alpha x := \begin{cases} (0,0) & \alpha = 0, \\ (\alpha x_1, x_2/\alpha), & \alpha \neq 0. \end{cases}$$

(v) Let $V = \mathbb{C}^2$, $\mathbb{F} = \mathbb{C}$. For $x = (x_1, x_2)$, $y = (y_1, y_2)$, let

$$x + y := (x_1 + 2y_1, x_2 + 3y_2) \quad \text{and} \quad \alpha x := (\alpha x_1, \alpha x_2) \quad \forall \alpha \in \mathbb{C}.$$

(vi) Let $V = \mathbb{R}^2$, $\mathbb{F} = \mathbb{R}$. For $x = (x_1, x_2)$, $y = (y_1, y_2)$, let

$$x + y := (x_1 + y_1, x_2 + y_2) \quad \text{and} \quad \alpha x := (x_1, 0) \quad \forall \alpha \in \mathbb{R}.$$

8. Let $A \in \mathbb{R}^{n\times n}$, O is the zero in $\mathbb{R}^{n\times 1}$. Show that the set V_0 of of all $n \times 1$ matrices X such that $AX = O$, is a subspace of $\mathbb{R}^{n\times 1}$.

9. Suppose V_0 is a subspace of a vector space V, and V_1 is a subspace of V_0. Then show that V_1 is a subspace of V.

10. Give an example to show that union of two subspaces need not be a subspace.

11. Let S be a subset of a vector space V. Show that S is a subspace if and only if $S = \operatorname{span} S$.

12. Let V be a vector space. Show that the the following hold.

    (i) Let S be a subset of V. Then $\operatorname{span} S$ is the intersection of all subspaces of V containing S.

    (ii) Suppose V_0 is a subspace of V and $x_0 \in V$ such that $x_0 \notin V_0$. Then for every $x \in \operatorname{span}\{x_0; X_0\}$, there exist a unique $\alpha \in \mathbb{F}$, $y \in V_0$ such that $x = \alpha x_0 + y$.

13. Show that

    (a) $\mathcal{P}_n$ is a subspace of $\mathcal{P}_m$ for $n \leq m$,

    (b) $C[a, b]$ is a subspace of $\mathcal{R}[a, b]$,

    (c) $C^k[a, b]$ is a subspace of $C[a, b]$.

14. For each λ in the open interval $(0,1)$, let $u_\lambda = (1, \lambda, \lambda^2, \ldots)$. Show that $u_\lambda \in \ell^1$ for each $\in (0,1)$, and the set $\{u_\lambda : 0 < \lambda < 1\}$ is a linearly independent in ℓ^1. Infer that every basis of the spaces c_0, c, ℓ^∞ is an uncountable set.

15. Let A be an $m \times n$ matrix, and $\mathbf{b}$ be a column m-vector. Show that the system $A\mathbf{x} = \mathbf{b}$ has a solution n-vector if and only if $\mathbf{b}$ is in the span of columns of A.

16. Let $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$. What is the span of $\{e_1 + e_2, e_2 + e_3, e_3 + e_1\}$?

17. Let S be a subset of a vector space V. Show that S is a subspace if and only if $S = \operatorname{span} S$.

18. Let V be a vector space. Show that the the following hold.

    (i) Let S be a subset of V. Then

    $$\operatorname{span} S = \bigcap \{Y : Y \text{ is a subspace of } V \text{ containing } S\}.$$

    (ii) Suppose V_0 is a subspace of V and $x_0 \in V \setminus V_0$. Then for every $x \in \operatorname{span}\{x_0; X_0\}$, there exist a unique $\alpha \in \mathbb{F}$, $y \in V_0$ such that $x = \alpha x_0 + y$. ♦

19. Consider the system of equations

    $$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \ldots & + & \ldots & + & \ldots & + & \ldots & = & \ldots \\ a_{m1}x_1 & + & a_{m1}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}$$

    Let

    $$u_1 := \begin{bmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{m1} \end{bmatrix}, u_2 := \begin{bmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{m2} \end{bmatrix}, \ldots, u_n := \begin{bmatrix} a_{1n} \\ a_{2n} \\ \ldots \\ a_{mn} \end{bmatrix}.$$

    (a) Show that the above system has a solution vector $x = [x_1, \ldots, x_n]^T$ if and only if $b = [b_1, \ldots, b_n]^T \in span(\{u_1, \ldots, u_n\}$.

    (b) Show that the above system has atmost one solution vector $x = [x_1, \ldots, x_n]^T$ if and only if $\{u_1, \ldots, u_n\}$ is linearly independent.

20. Show that every superset of a linearly dependent set is linearly dependent, and every subset of a linearly independent set is linearly independent.

21. Give an example to justify the following: E is a subset of vector space such that there exists an vector $u \in E$ which is not a linear combination of other members of E, but E is linearly dependent.

22. Is union (resp., intersection) of two linearly independent sets a linearly independent?

23. Is union (resp., intersection) of two linearly dependent sets a linearly dependent?

24. Show that vectors $u = (a, c)$, $v = (b, d)$ are linearly independent in $\mathbb{R}^2$ iff $ad - bc \neq 0$.

25. Show that $V_0 := \{x = (x_1, x_2, x_3) : x_1 + x_2 + x_3 = 0\}$ is a subspace of $\mathbb{R}^3$. Find a basis for V_0.

26. Show that $E := \{1 + t^n, t + t^n, t^2 + t^n, \ldots, t^{n-1} + t^n, t^n\}$ is a basis of $\mathcal{P}_n$.

27. Let $u_1, \ldots, u_n$ are linearly independent vectors in a vector space V. Let $[a_{ij}]$ be an $m \times n$ matrix of scalar, and let

$$\begin{array}{lclcccccc} v_1 & := & a_{11}u_1 & + & a_{21}u_2 & + & \ldots & + & a_{m1}u_n \\ v_2 & := & a_{12}u_1 & + & a_{22}u_2 & + & \ldots & + & a_{m2}u_n \\ \ldots & & \ldots & + & \ldots & + & \ldots & + & \ldots \\ v_n & := & a_{1n}u_1 & + & a_{2n}u_2 & + & \ldots & + & a_{mn}u_n. \end{array}$$

Show that the $v_1, \ldots, v_m$ are linearly independent if and only if the vectors

$$w_1 := \begin{bmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{bmatrix}, \quad w_2 := \begin{bmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{bmatrix}, \ldots, \quad w_n := \begin{bmatrix} a_{1n} \\ a_{m2} \\ \cdots \\ a_{mn} \end{bmatrix}$$

are linearly independent. ♦

28. Let $u_1(t) = 1$, and for $j = 2, 3, \ldots$, let $u_j(t) = 1 + t + \ldots + t^j$. Show that span of $\{u_1, \ldots, u_n\}$ is $\mathcal{P}_n$, and span of $\{u_1, u_2, \ldots\}$ is $\mathcal{P}$.

29. Let $p_1(t) = 1 + t + 3t^2$, $p_2(t) = 2 + 4t + t^2$, $p_3(t) = 2t + 5t^2$. Are the polynomials p_1, p_2, p_3 linearly independent? ♦

30. Show that a basis of a vector space is a minimal spanning set, and maximal linearly independent set.

31. Suppose V_1 and V_2 are subspaces of a vector space V such that $V_1 \cap V_2 = \{0\}$. Show that every $x \in V_1 + V_2$ can be written *uniquely* as $x = x_1 + x_2$ with $x_1 \in V_1$ and $x_2 \in V_2$.

32. Suppose V_1 and V_2 are subspaces of a vector space V. Show that $V_1 + V_2 = V_1$ if and only if $V_2 \subseteq V_1$.

33. Let V be a vector space.

    (i) Show that a subset $\{u_1, \ldots, u_n\}$ of V is linearly independent if and only if the function $(\alpha_1, \ldots, \alpha_n) \mapsto \alpha_1 u_1 + \cdots + \alpha_n u_n$ from $\mathbb{F}^n$ into V is injective.

    (ii) Show that if $E \subseteq V$ is linearly dependent in V, then every superset of E is also linearly dependent.

    (iii) Show that if $E \subseteq V$ is linearly independent in V, then every subset of E is also linearly independent.

    (iv) Show that if $\{u_1, \ldots, u_n\}$ is a linearly independent subset of V, and if Y is a subspace of V such that $(\mathrm{span}\,\{u_1, \ldots, u_n\}) \cap Y = \{0\}$, then every V in the span of $\{u_1, \ldots, u_n, Y\}$ can be written uniquely as $x = \alpha_1 u_1 + \cdots + \alpha_n u_n + y$ with $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$, $y \in Y$.

    (v) Show that if E_1 and E_2 are linearly independent subsets of V such that $(\mathrm{span}\, E_1) \cap (\mathrm{span}\, E_2) = \{0\}$, then $E_1 \cup E_2$ is linearly independent.

34. For each $k \in \mathbb{N}$, let $\underline{\mathbb{F}}^k$ denotes the set of all column k-vectors, i.e., the set of all $k \times 1$ matrices. Let A be an $m \times n$ matrix of scalars with columns $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n$. Show the following:

    (i) The equation $A\underline{x} = \underline{0}$ has a non-zero solution if and only if $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n$ are linearly dependent.

    (ii) For $\underline{y} \in \underline{\mathbb{F}}^m$, the equation $A\underline{x} = \underline{y}$ has a solution if and only if $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n, \underline{y}$ are linearly dependent, i.e., if and only if $\underline{y}$ is in the span of columns of A.

35. For $i = 1, \ldots, m$; $j = 1, \ldots, n$, let E_{ij} be the $m \times n$ matrix with its (i,j)-th entry as 1 and all other entries 0. Show that

$$\{E_{ij} : i = 1 \ldots, m;\; j = 1, \ldots, n\}$$

is a basis of $\mathbb{F}^{m \times n}$.

36. If $\{u_1, \ldots, u_n\}$ is a basis of a vector space V, then show that every $x \in V$, can be expressed *uniquely* as $x = \alpha_1 u_1 + \cdots + \alpha_n u_n$; i.e., for every $x \in V$, there exists a unique n-tuple $(\alpha_1, \ldots, \alpha_n)$ of scalars such that $x = \alpha_1 u_1 + \cdots + \alpha_n u_n$.

37. Suppose S is a set consisting of n elements and V is the set of all real valued functions defined on S. Show that V is a vector space of dimension n.

38. Let $t_0, t_1, \ldots, t_n$ be in $[a,b]$ such that $a = t_0 < t_1 < \ldots < t_n = b$. For $k \in \mathbb{N}$, let $X_{k,n}$ be the set of all those functions $x \in C([a,b], \mathbb{R})$ such that the restriction of x to each interval $[t_{j-1}, t_j]$ is a polynomial of degree atmost k. Then show that $X_{k,n}$ is a linear space over $\mathbb{R}$. What is the dimension of $X_{k,n}$?

39. Given real numbers $a_0, a_1, \ldots, a_k$, let X be the set of all solutions $x \in C^k[a,b]$ of the differential equation

$$a_0 \frac{d^k x}{dt^k} + a_1 \frac{d^{k-1} x}{dt^{k-1}} + \cdots + a_k x = 0.$$

Show that X is a linear space over $\mathbb{R}$. What is the dimension of X?

40. Let $t_0, t_1, \ldots, t_n$ be in $[a,b]$ such that $a = t_0 < t_1 < \ldots < t_n = b$. For each $j \in \{1, \ldots, n\}$, let u_j be in $C([a,b], \mathbb{R})$ such that

$$u_j(t_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

and the restriction of u_j to each interval $[t_{j-1}, t_j]$ is a polynomial of degree atmost 1. Show that the span of $\{u_1, \ldots, u_n\}$ is the space $X_{1,n}$ in Problem 38.

41. The spaces $c_{00}, c_0, \ell^1, \ell^\infty, \mathcal{P}, C[a,b], \mathcal{R}[a,b]$ are all infinite dimensional spaces – Why?

42. State with reason whether $T : \mathbb{R}^2 \to \mathbb{R}^2$ in each of the following is a linear transformation:

    (a) $T(x_1, x_2) = (1, x_2)$, (b) $T(x_1, x_2) = (x_1, x_2^2)$
    (c) $T(x_1, x_2) = (\sin(x_1), x_2)$ (d) $T(x_1, x_2) = (x_1, 2 + x_2)$

43. Check whether the functions T in the following are linear transformations:

    (i) $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x, y) = (2x + y, x + y^2)$.

    (ii) $T : C^1[0, 1] \to \mathbb{R}$ defined by $T(u) = \int_0^1 [u(t)]^2 dt$.

    (iii) $T : C^1[-1, 1] \to \mathbb{R}^2$ defined by $T(u) = \left(\int_{-1}^1 u(t)dt, u'(0)\right)$.

    (iii) $T : C^1[0, 1] \to \mathbb{R}$ defined by $T(u) = \int_0^1 u'(t)dt$.

44. Let $T_1 : V_1 \to V_2$ and $T_2 : V_2 \to V_3$ be linear transformations. Show that the function $T : V_1 \to V_3$ defined by $Tx = T_2(T_1 x)$, $x \in V_1$, is a linear transformation.

    [The above transformation T is called the *composition* of T_2 and T_1, and is usually denoted by $T_2 T_1$.]

45. If $T_1 : C^1[0, 1] \to C[0, 1]$ is defined by $T_1(u) = u'$, and $T_2 : C[0, 1] \to \mathbb{R}$ is defined by $T_2(v) = \int_0^1 v(t)dt$, then find $T_2 T_1$.

46. Let V_1, V_2, V_3 be finite dimensional vector spaces, and let E_1, E_2, E_3 be bases of V_1, V_2, V_3 respectively. If $T_1 : V_1 \to V_2$ and $T_2 : V_2 \to V_3$ are linear transformations. Show that $[T_2 T_1]_{E_1,E_3} = [T_2]_{E_2,E_3} [T_1]_{E_1,E_2}$.

47. If $T_1 : \mathcal{P}_n[0, 1] \to \mathcal{P}_n[0, 1]$ is defined by $T_1(u) = u'$, and $T_2 : \mathcal{P}_n[0, 1] \to \mathbb{R}$ is defined by $T_2(v) = \int_0^1 v(t)dt$, then find $[T_1]_{E_1,E_2}$, $[T_2]_{E_2,E_3}$, and $[T_2 T_1]_{E_1,E_3}$, where $E_1 = E_2 = \{1, t, t^2, \ldots, t^n\}$ and $E_3 = \{1\}$.

48. Justify the statement: Let $T_1 : V_1 \to V_2$ be a linear transformation. Then T is bijective iff there exists a linear transformation $T_2 : V_2 \to V_1$ such that $T_1 T_2 : V_2 \to V_2$ is the identity transformation on V_2 and $T_2 T_1 : V_1 \to V_1$ is the identity transformation on V_1.

49. Let V_1 and V_2 be vector spaces with $dim V_1 = n < \infty$. Let $\{u_1, \ldots, u_n\}$ be a basis of V_1 and $\{v_1, \ldots, v_n\} \subset V_2$. Find a linear transformation $T : V_1 \to V_2$ such that $T(u_j) = v_j$ for

$j = 1, \ldots, n$. Show that there is only one such linear transformation.

50. Let T be the linear transformation obtained as in the above problem. Show that

    (a) T is one-one if and only if $\{v_1, \ldots, v_n\}$ is linearly independent, and

    (b) T is onto if and only if $\operatorname{span}(\{v_1, \ldots, v_n\}) = V_2$.

51. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which satisfies $T(1,0) = (1,4)$ and $T(1,1) = (2,5)$. Find the $T(2,3)$.

52. Does there exists a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ such that $T(1,0,2) = ((1,1)$ and $T(1/2,0,1) = ((0,1)$?

53. Show that if V_1 and V_2 are finite dimensional vector spaces of the same dimension, then the there exists a bijective linear transformation from V_1 to V_2.

54. Find bases for $N(T)$ and $R(T)$ for the linear transformation T in each the following:

    (a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x_1, x_2) = (x_1 - x_2, 2x_2)$,

    (b) $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(x_1, x_2) = (x_1 + x_2, 0, 2x_3 - x_2)$,

    (c) $T : \mathbb{R}^{n\times n} \to \mathbb{R}$ defined by $T(A) = trace(A)$. (Recall that trace of a square matrix is the sum of its diagonal elements.)

55. Let $T : V_1 \to V_2$ is a linear transformation. Given reasons for the following:

    (a) $\text{rank}(T) \le \text{dim}V_1$.

    (b) T onto implies $\text{dim}V_2 \le \text{dim}V_1$,

    (c) T one-one implies $\text{dim}V_1 \le \text{dim}V_2$

    (d) Suppose $\text{dim}V_1 = \text{dim}V_2 < \infty$. Then T is one-one if and only T is onto.

56. Let V_1 and V_2 be finite dimensional vector spaces, and $E_1 = \{u_1, \ldots, u_n\}$ and $E_2 = \{v_1, \ldots, v_m\}$ be bases of V_1 and V_2, respectively. Let $F_1 = \{f_1, \ldots, f_n\}$ be the dual basis of $\mathcal{L}(V_1, \mathbb{F})$ with respect to E_1 and $F_2 = \{g_1, \ldots, g_n\}$ be the dual basis of

$\mathcal{L}(V_2, \mathbb{F})$ with respect to E_2. For $i = 1, \ldots, n$; $j = 1, \ldots, m$, let $T_{ij} : V \to W$ defined by

$$T_{ij}(x) = f_j(x)v_i, \qquad x \in V_1.$$

Show that $\{T_{ij} : i = 1, \ldots, n; j = 1, \ldots, m\}$ is a basis of $\mathcal{L}(V_1, V_2)$.

57. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$T(x_1, x_2, x_3) = (x_2 + x_3, x_3 + x_1, x_1 + x_2), \quad (x_1, x_2, x_3) \in \mathbb{R}^3.$$

Find the matrix representation of T with respect to the basis given in each of the following.

(a) $E_1 = \{(1,0,0), (0,1,0), (0,0,1)\}$, $E_2 = \{(1,0,0), (1,1,0), (1,1,1)\}$

(b) $E_1 = \{(1,0,0), (1,1,0), (1,1,1)\}$, $E_2 = \{(1,0,0), (0,1,0), (0,0,1)\}$

(c) $E_1 = \{(1,1,-1), (-1,1,1), (1,-1,1)\}$,
$E_2 = \{(-1,1,1), (1,-1,1), (1,1,-1)$

58. Let $T : \mathcal{P}^3 \to \mathcal{P}^2$ be defined by $T(a_0 + a_1 t + a_2 t^2 + a_3 t^3) = a_1 + 2a_2 t + 3a_3 t^2$. Find the matrix representation of T with respect to the basis given in each of the following.

(a) $E_1 = \{1, t, t^2, t^3\}$, $E_2 = \{1 + t, 1 - t, t^2\}$

(b) $E_1 = \{1, 1 + t, 1 + t + t^2, t^3\}$, $E_2 = \{1, 1 + t, 1 + t + t^2\}$

(c) $E_1 = \{1, 1 + t, 1 + t + t^2, 1 + t + t^2 + t^3\}$, $E_2 = \{t^2, t, 1\}$

59. Let $T : \mathcal{P}^2 \to \mathcal{P}^3$ be defined by $T(a_0 + a_1 t + a_2 t^2) = (a_0 t + \frac{a_1}{2} t^2 + \frac{a_2}{3} t^3)$. Find the matrix representation of T with respect to the basis given in each of the following.

(a) $E_1 = \{1 + t, 1 - t, t^2\}$, $E_2 = \{1, t, t^2, t^3\}$,

(b) $E_1 = \{1, 1 + t, 1 + t + t^2\}$, $E_2 = \{1, 1 + t, 1 + t + t^2, t^3\}$,

(c) $E_1 = \{t^2, t, 1\}$, $E_2 = \{1, 1 + t, 1 + t + t^2, 1 + t + t^2 + t^3\}$,

60. A linear transformation $T : V \to W$ is said to be of *finite rank* if $\operatorname{rank} T < \infty$.

Let $T : V_1 \to V_2$ be a linear transformation between vector spaces V_1 and V_2. Show that T is of finite rank if and only if there exists $n \in \mathbb{N}$, $\{v_1, \ldots, v_n\} \subset V_2$ and $\{f_1, \ldots, f_n\} \subset \mathcal{L}(V_1, \mathbb{F})$ such that $Ax = \sum_{j=1}^{n} f_j(x) v_j$ for all $x \in V_1$. ♦

61. Let V_1 and V_2 be inner product spaces with inner products $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$ respectively. One $V = V_1 \times V_2$, define

$$\langle (x_1,x_2),(y_1,y_2)\rangle_V := \langle x_1,y_1\rangle_1 + \langle x_2,y_2\rangle_2, \qquad \forall\, (x_1,x_2),(y_1,y_2) \in V.$$

Show that $\langle\cdot,\cdot\rangle_V$ is an inner product on V.

62. Let $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$ are inner products on a vector space V. Show that $\langle x,y\rangle := \langle x,y\rangle_1 + \langle x,y\rangle_2$ defines another inner product on V.

63. For x, y in an inner product space V, show that $(x+y) \perp (x-y)$ if and only if $\|x\| = \|y\|$.

64. Let V be an inner product space. For $S \subset V$, let

$$S^\perp := \{x \in V : \langle x,u\rangle = 0 \quad \forall u \in S\}.$$

Show that

(a) $S^\perp$ is a subspace of V.

(b) $V^\perp = \{0\}, \quad \{0\}^\perp = V$.

(c) $S \subset S^{\perp\perp}$.

(d) If V is finite dimensional and V_0 is a subspace of V, then $V_0^{\perp\perp} = V_0$.

65. Find the best approximation of $x \in V$ from V_0 where

(a) $V = \mathbb{R}^3,\ x := (1,2,1),\ V_0 := \operatorname{span}\{(3,1,2),1,0,1)\}$.

(b) $V = \mathbb{R}^3,\ x := (1,2,1),$ and V_0 is the set of all $(\alpha_1,\alpha_2,\alpha_3)$ in $\mathbb{R}^4$ such that $\alpha_1+\alpha_2+\alpha_3 = 0\}$.

(c) $V = \mathbb{R}^4,\ x := (1,0,-1,1)\ V_0 := \operatorname{span}\{(1,0,-1,1),\ (0,0,1,1)\}$.

(d) $V = C[-1,1],\ x(t) = e^t,\ V_0 = \mathcal{P}_3$.

66. Let $A \in \mathbb{R}^{m\times n}$ and $y \in \mathbb{R}^m$. Show that, there exists $x \in \mathbb{R}^n$ such that $\|Ax - y\| \le \|Au = y\|$ for all $u \in \mathbb{R}^n$, if and only if $A^TAx = A^Ty$.

67. Let $A \in \mathbb{R}^{m\times n}$ and $y \in \mathbb{R}^m$. If columns of A are linearly independent, then show that there exists a unique $x \in \mathbb{R}^n$ such that $A^TAx = A^Ty$.

68. Find the best approximate solution (least square solution) for the system $Ax = y$ in each of the following:

(a) $A = \begin{bmatrix} 3 & 1 \\ 1 & 2 \\ 2 & -1 \end{bmatrix}; \qquad y = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}.$

(b) $A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}; \qquad y = \begin{bmatrix} 0 \\ 1 \\ -1 \\ -2 \end{bmatrix}.$

69. Show that $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$ for every $x \in \mathbb{R}^k$.

Find $c_1, c_2, c_3, c_4 > 0$ such that

$$c_1\|x\|_2 \leq \|x\|_\infty \leq c_2\|x\|_2, \quad c_3\|x\|_1 \leq \|x\|_1 \leq c_4\|x\|_\infty \quad \forall x \in \mathbb{R}^k.$$

Compute $\|x\|_\infty$, $\|x\|_2$, $\|x\|_1$ for $x = (1, 1, 1) \in \mathbb{R}^3$.

70. Show that there exists no constant $c > 0$ such that $\|x\|_\infty \leq c\,\|x\|_1$ for all $x \in C[a, b]$.

71. Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ and and $A \in \mathbb{R}^{n\times n}$. Suppose $c > 0$ is such that $\|Ax\| \leq c\,\|x\|$ for all $x \in \mathbb{R}^n$, and there exists $x_0 \neq 0$ in $\mathbb{R}^n$ such that $\|Ax_0\| = c\,\|x_0\|$. Then show that $\|A\| = c$.

72. Find $\|A\|_1$, $\|A\|_\infty$, for the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 2 & 1 \end{bmatrix}.$

73. Suppose A, B in $\mathbb{R}^{n\times n}$ are invertible matrices, and $b, \tilde{b}$ are in $\mathbb{R}^n$. Let $x, \tilde{x}$ are in $\mathbb{R}^n$ be such that $Ax = b$ and $B\tilde{x} = \tilde{b}$. Show that

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \|A\|\,\|B^{-1}\|\Big(\frac{\|A - B\|}{\|A\|} + \frac{\|b - \tilde{b}\|}{\|b\|}\Big).$$

[Hint: Use the fact that $B(x - \tilde{x}) = (B - A)x + (b - \tilde{b})$, and use the fact that $\|(B - A)x\| \leq \|B - A\|\,\|x\|$, and $\|b - \tilde{b}\| = \|b - \tilde{b}\|\|Ax\|/\|b\| \leq \|b - \tilde{b}\|\|A\|\,\|x\|/\|b\|$.]

74. Let $B \in \mathbb{R}^{n\times n}$. If $\|B\| < 1$, then show that $I - B$ is invertible, and $\|(I - B)^{-1}\| \leq 1/(1 - \|B\|)$.

[Hint: Show that $I - B$ is injective, by showing that for every x, $\|(I - B)x\| \geq (1 - \|B\|)\|x\|$, and then deduce the result.]

75. Let $A, B \in \mathbb{R}^{n\times n}$ be such that A is invertible, and $\|A - B\| < 1/\|A^{-1}\|$. Then, show that, B is invertible, and
$$\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A - B\|\,\|A^{-1}\|}.$$
[Hint: Observe that $B = A - (A - B) = [I - (A - B)A^{-1}]A$, and use the previous problem.]

76. Let $A, B \in \mathbb{R}^{n\times n}$ be such that A is invertible, and $\|A - B\| < 1/2\|A^{-1}\|$. Let $b, \tilde{b}, x, \tilde{x}$ be as in Problem 73. Then, show that, B is invertible, and
$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq 2\kappa(A)\Big(\frac{\|A - B\|}{\|A\|} + \frac{\|b - \tilde{b}\|}{\|b\|}\Big).$$
[Hint: Apply conclusion in Problem 75 to that in Problem 73]

77. Suppose $u_1, \ldots, u_n$ are functions defined on $[a, b]$, and $t_1, \ldots, t_n$ are points in $[a, b]$. Let $\beta_1, \ldots, \beta_n$ are real numbers. Then show that there exists a unique $\varphi \in \operatorname{span}\{u_1, \ldots, u_n\}$ satisfying $\varphi(t_i) = \beta_i$ for $i = 1, \ldots, n$ if and only if the matrix $[u_j(t_i)]$ is invertible.

78. Suppose $u_1, \ldots, u_n$ are functions defined on $[a, b]$, and $t_1, \ldots, t_n$ are points in $[a, b]$. Show that, if the matrix $[u_j(t_i)]$ is invertible, then $u_1, \ldots, u_n$ are linearly independent.

    Hint: A square matrix is invertible if and only if its columns are linearly independent.

79. Suppose $u_1, \ldots, u_n$ are functions defined on $[a, b]$, and $t_1, \ldots, t_n$ are points in $[a, b]$ such that the matrix $[u_j(t_i)]$ is invertible. If $v_1, \ldots, v_n$ are linearly independent functions in $\operatorname{span}\{u_1, \ldots, u_n\}$, then show that the matrix $[v_j(t_i)]$ is also invertible.

    Hint: Let $X_0 := \operatorname{span}\{u_1, \ldots, u_n\}$ and $[u_j(t_i)]$ is invertible. Then observe that, the function $J : X_0 \to \mathbb{R}^n$ defined by $J(x) = [x(t_1), \ldots, x(t_n)]^T$ is bijective.

80. Let $t_1, \ldots, t_n$ be distinct points in $\mathbb{R}$, and for each $j \in \{1, 2, \ldots, n\}$, let $\ell_j(t) = \prod_{i\neq j} \frac{t - t_i}{t_j - t_i}$. Then show that $\{\ell_1, \ldots, \ell_n\}$ is a basis of $\mathcal{P}_{n-1}$, and it satisfies $\ell_j(t_i) = \delta_{ij}$ for all $i, j = 1, \ldots, n$. Deduce from the previous exercise that the matrix $[t_j^{i-1}]$ is invertible.