Fourier Series
and Integral Transforms

Allan Pinkus & Samy Zafrany

Technion, Israel Institute of Technology

Contents

Chapter 1

Background:

Inner Product Spaces

0. Introduction

The main topics to be studied in this chapter are orthogonal and orthonormal systems in a vector space with inner product, as well as various related concepts. These topics are sometimes, but not always, discussed in a basic course in linear algebra. Of central importance is the subject of infinite orthonormal systems which we present at the end of this chapter. These results will be applied in the next chapter on Fourier series. The first four sections of this chapter are a condensed review of some concepts and basic ideas (with proofs) from linear algebra. We use these facts in developing the different topics of this book. The reader will hopefully find in these sections a helpful synopsis and review of his knowledge of the area.

1. Linear and Inner Product Spaces

The basic algebraic structure which we use is the *linear space* (often called *vector space*) over a *field of scalars*. Our "field of scalars" will always be either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. Elements of a linear space are called *vectors*. Formally, a non-empty set V is called a *linear space* over a field F if it satisfies the following conditions:

1. **Vector Addition:** There exists an operation, generally denoted by "+", such that for any two vectors $u, v \in V$, the "sum" $u + v$ is also a vector in V.

2. **Associativity:** For every $u, v, w \in V$, we have $(u + v) + w = u + (v + w)$.

3. **Zero Vector:** There exists a vector which we denote by $\vec{0}$ and call the "zero vector" satisfying $\vec{0} + v = v$, for all $v \in V$.

4. **Inverse vector:** For every vector $v \in V$ there exists a vector, which we denote by $-v$, and call "minus v", such that $v + (-v) = \vec{0}$.

5. **Commutativity:** For every $u, v \in V$ we have $u + v = v + u$.

6. **Multiplication by a Scalar:** Multiplication by scalars is permissible. That is, for each $v \in V$ and scalar $a \in F$, there is defined $av \in V$.

7. For every $a \in F$ and $u, v \in V$, $a(u + v) = au + av$.

8. For each $a, b \in F$ and $u \in V$, $(a + b)u = au + bu$ and $a(bu) = (ab)u$.

9. For the unit scalar 1 of F and every $v \in V$, $1 \cdot v = v$.

If V is a linear space over the field of reals $\mathbb{R}$, then we say that V is a *real linear space*. If V is a linear space over the field of complex numbers $\mathbb{C}$, then we call V a *complex linear space*. A subset W of V ($W \subseteq V$) is said to be a *linear subspace* of the space V if all the above conditions hold for W over the same field of scalars as for V. Of course, the operations of vector addition and multiplication by a scalar must be the same in W as in V. A well-known criterion for checking if W is a subspace is the following: $W \neq \emptyset$ and for every $u, v \in W$ and $a, b \in F$ we have $au + bv \in W$. In other words, W is a linear subspace of V if and only if W is a non-empty subset of V, which is closed under the operation of vector addition and multiplication by scalars. We assume in what follows, unless stated otherwise, that all our linear subspaces are complex. We now quickly review a number of important concepts related to linear spaces.

Definition 1.1: *Let V be a linear space and $v_1, \ldots, v_n \in V$. The vector u is said to be a* linear combination *of the vectors v_1, ..., v_n if there exist scalars $a_1, \ldots, a_n \in F$ such that*

$$u = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$

The collection of all vectors u which are linear combinations of v_1, ..., v_n is called the span *of $v_1, \ldots, v_n$ and will be denoted by* span$\{v_1, \ldots, v_n\}$.

We also say that v_1, ..., v_n *span W*. Note that $W = \text{span}\{v_1, \ldots, v_n\}$ is a linear subspace of V.

Definition 1.2: *Let V be a linear space. The vectors $v_1, \ldots, v_n \in V$ are said to be* linearly independent *if the equation*

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = \vec{0}, \qquad a_1, a_2, \ldots, a_n \in F,$$

is satisfied only by the scalars $a_1 = \cdots = a_n = 0$. Otherwise we say that the vectors $v_1, \ldots, v_n$ are linearly dependent.

From the above two definitions it immediately follows that the vectors v_1, ..., v_n are linearly independent if and only if for each i, $1 \leq i \leq n$, the vector v_i is not a linear combination of the other vectors of the set.

Definition 1.3: *A finite set of vectors v_1, ..., v_n is said to be a* basis *for the linear space V if the set of vectors v_1, ..., v_n is linearly independent and $V = $ span$\{v_1, \ldots, v_n\}$. The natural number n is called the dimension of V and we write $n = \dim(V)$.*

The reader will recall from his knowledge of linear algebra that every real or complex non-trivial linear space has an infinite number of different bases. However, any two bases have the same number of elements and thus the definition of dimension is in fact independent of any specific choice of basis (not necessarily finite in number). Every linear space has a dimension (which may be infinite).

Another important concept is that of an inner product. Note that the definition of a linear space does not include an operation of "multiplication" between vectors. The inner product could be considered as such an operation. However, linear spaces do not in general possess an inner product.

Definition 1.4: **(Inner Product)** *Let V be a real or complex linear space. An* inner product *is an operation between two elements of V which results in a scalar (and not a vector!). We denote it by $\langle u, v \rangle$, i.e., $\langle u, v \rangle \in \mathbb{C}$. It satisfies:*

1. *For each $v \in V$, $\langle v, v \rangle$ is a non-negative real number, i.e., $\langle v, v \rangle \geq 0$.*

2. *For each $v \in V$, $\langle v, v \rangle = 0$ if and only if $v = \vec{0}$.*

3. *For each $u, v, w \in V$ and $a, b \in F$, $\langle au + bv, w \rangle = a \langle u, w \rangle + b \langle v, w \rangle$.*

4. *For each $u, v \in V$ we have $\langle u, v \rangle = \overline{\langle v, u \rangle}$.*

A linear space with an inner product defined on it is called an inner product space.

The expression $\overline{\langle v, u \rangle}$ denotes the complex conjugate of the complex number $\langle v, u \rangle$. If our field of scalars is $\mathbb{R}$, then in place of Condition 4 we write $\langle u, v \rangle = \langle v, u \rangle$. There are numerous consequences of the above four conditions. Here are a few of them.

(a) For each $u, v, w \in V$ and $a, b \in \mathbb{C}$, $\langle u, av + bw \rangle = \overline{a} \langle u, v \rangle + \overline{b} \langle u, w \rangle$.

(b) For each $v \in V$ and $a \in \mathbb{C}$, $\langle av, av \rangle = |a|^2 \langle v, v \rangle$.

(c) For each $v \in V$, $\left\langle \vec{0}, v \right\rangle = 0$.

(d)   In general, for each finite sequence of vectors $\{u_k\}_{k=1}^n$, scalars $\{a_k\}_{k=1}^n$, and every vector v,

$$\left\langle \textstyle\sum_{k=1}^n a_k u_k, v \right\rangle = \textstyle\sum_{k=1}^n a_k \left\langle u_k, v \right\rangle,$$

$$\left\langle v, \textstyle\sum_{k=1}^n a_k u_k \right\rangle = \textstyle\sum_{k=1}^n \overline{a_k} \left\langle v, u_k \right\rangle.$$

We now consider some typical examples of inner product spaces.

Example 1.1:   The Euclidean space $V = \mathbb{R}^n$ with the usual operations of vector addition and multiplication by a scalar is a linear space over $\mathbb{R}$. Let $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ be a vector of strictly positive numbers, i.e., $r_k > 0, 1 \leq k \leq n$. We define an inner product $\langle \cdot, \cdot \rangle_{\mathbf{r}}$ on $\mathbb{R}^n$ in the following way: For each pair of vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$, set

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{r}} = \sum_{k=1}^n r_k x_k y_k.$$

The vector $\mathbf{r}$ is called the *weight vector* of the inner product. If $r_k = 1, 1 \leq k \leq n$, then the inner product is denoted $\mathbf{x} \cdot \mathbf{y}$,

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n,$$

and is said to be the *standard inner product* (or *dot product*) on $\mathbb{R}^n$.

Example 1.2:   Analogously to the previous example, $V = \mathbb{C}^n$ with the usual vector addition and multiplication by scalars is a linear space over $\mathbb{C}$. Let $\mathbf{r}$ be as in Example 1.1. For each pair $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, we define

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{r}} = \sum_{k=1}^n r_k x_k \overline{y_k}.$$

It is not difficult to prove that this is an inner product on $\mathbb{C}^n$. As in the previous example, the *standard inner product* on $\mathbb{C}^n$ is

$$\mathbf{x} \cdot \mathbf{y} = x_1 \overline{y_1} + x_2 \overline{y_2} + \cdots + x_n \overline{y_n}.$$

Example 1.3:   Let $V = C[a, b]$ be the space of continuous functions $f : [a, b] \to \mathbb{C}$ with the usual operations of sum of functions and multiplication by scalars. This is a linear space over $\mathbb{C}$. For each pair of functions $f, g \in C[a, b]$, we define

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} \, dx.$$

It is easy to verify that this is an inner product on $C[a, b]$.

Example 1.4: Set

$$\ell_2 = \left\{ \mathbf{x} \ \middle| \ \mathbf{x} = (x_1, x_2, \ldots), \ x_n \in \mathbb{C}, \ \sum_{n=1}^{\infty} |x_n|^2 < \infty \right\}.$$

That is, ℓ_2 is the space of infinite sequences of complex numbers $\{x_n\}_{n=1}^{\infty}$ for which the series $\sum_{n=1}^{\infty} |x_n|^2$ converges. Vector addition in ℓ_2 is the usual vector addition of two sequences, and multiplication by a scalar is also the standard one. We will later prove that ℓ_2 is a linear space over $\mathbb{C}$ (the fact that ℓ_2 is closed under vector addition is not obvious). For each $\mathbf{x}, \mathbf{y} \in \ell_2$ we define

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = \sum_{n=1}^{\infty} x_n \overline{y_n}.$$

It is easily shown that $\langle \cdot, \cdot \rangle$ satisfies all four conditions in the definition of the inner product. It is more difficult to prove that the series $\sum_{n=1}^{\infty} x_n \overline{y_n}$ converges for every $\mathbf{x}, \mathbf{y} \in \ell_2$. This fact is a consequence of the Cauchy-Schwarz inequality which we prove in the next section.

Exercises

1. Let $V_1, V_2, \ldots, V_n$ be linear subspaces of a linear space U over $\mathbb{C}$. Prove that $V = \bigcap_{k=1}^{n} V_k$ is a linear space. What can you say about $W = \bigcup_{k=1}^{n} V_k$?

2. Prove that the set

   $$V = \{ f \mid f : \mathbb{R} \to \mathbb{R}, \ f \text{ is absolutely integrable over } \mathbb{R} \}$$

   is a linear space over $\mathbb{R}$.

3. Let $C[-1, 2]$ denote the space of continuous complex-valued functions $f : [-1, 2] \to \mathbb{C}$. Which of the following define an inner product on $C[-1, 2]$, and which do not? Explain.
   (a) $\langle f, g \rangle = \int_{-1}^{2} |f(t) + g(t)| \, dt$
   (b) $\langle f, g \rangle = \int_{-1}^{2} f(t) \overline{g(t)} \, dt + f(-\frac{1}{2}) \overline{g(-\frac{1}{2})}$
   (c) $\langle f, g \rangle = 3 \int_{-1}^{2} f(t) \overline{g(t)} \, dt$
   (d) $\langle f, g \rangle = f(0) \overline{g(0)} + f(1) \overline{g(1)}$

4. Let $V = C^2[-\pi, \pi]$ be the space of real-valued twice continuously differentiable functions defined on the interval $[-\pi, \pi]$. Set

   $$\langle f, g \rangle = f(-\pi) g(-\pi) + \int_{-\pi}^{\pi} f''(x) g''(x) \, dx.$$

   Is this an inner product on V?

5. Let $C^1[0, 1]$ denote the space of continuous functions $f : [0, 1] \to \mathbb{C}$ with a continuous first derivative on $[0, 1]$. Set

$$\langle f, g \rangle = f(0)\overline{g(0)} + f'(0)\overline{g'(0)} + f(1)\overline{g(1)}.$$

(a) Is this an inner product on the subspace $\mathcal{P}_2 = \text{span}\{1, x, x^2\}$?

(b) Is it an inner product on $C^1[0, 1]$?

6. Let $C^1[0, 1]$ be as in Exercise 5. Which of the following define an inner product on $C^1[0, 1]$?

(a) $\langle f, g \rangle = f(0)\overline{g(0)} + \int_0^1 f'(t)\overline{g'(t)}\, dt$

(b) $\langle f, g \rangle = f(0)\overline{g(0)} + f'(1)\overline{g'(1)}$

(c) $\langle f, g \rangle = 2\int_0^{\frac{1}{2}} f(t)\overline{g(t)}\, dt - \int_{\frac{1}{2}}^1 f(t)\overline{g(t)}\, dt$

2. The Norm

There is a connection between the concept of a norm and that of an inner product. The definition of a norm is in no way dependent upon that of an inner product. However, in every inner product space one can always define a norm in a very natural way.

Definition 1.5: *Let V be a linear space. A norm on V is a function from V to $\mathbb{R}_+$ which we denote by $\| \cdot \|$, and which satisfies the following properties:*

(1) *For each $v \in V$, $\|v\| \geq 0$.*

(2) *$\|v\| = 0$ if and only if $v = \vec{0}$.*

(3) *For each $v \in V$ and $a \in \mathbb{C}$, $\|av\| = |a| \cdot \|v\|$.*

(4) *For every $u, v \in V$, $\|u + v\| \leq \|u\| + \|v\|$ (the triangle inequality).*

The concept of a norm is a generalization of the concept of *size* or *distance* (from the zero vector). For every $u, v \in V$, we may consider the number $\|u - v\|$ as the distance between u and v. Hence $\|u\|$ is the distance of u from $\vec{0}$, or the size of u.

The simplest examples of norms are the absolute value on $\mathbb{R}$ and $\mathbb{C}$. Here are some more:

Example 1.5: If $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$, then for each $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in V$ we define

$$\|\mathbf{x}\| = \sqrt{\sum_{k=1}^n |x_k|^2}.$$

We will soon see that this formula defines a norm on V. It is called the *Euclidean norm*.

Example 1.6: Let V be as in the previous example. We define thereon a different norm. Let

$$\|\mathbf{x}\|_\infty = \max\Big\{ |x_k| \ \Big| \ k = 1, 2, \ldots, n \Big\}.$$

This is called the *uniform norm*.

Example 1.7: On the linear spaces $V = \mathbb{R}^n$ and $V = \mathbb{C}^n$ we can define many different norms. One common norm thereon, other than the previous two, is

$$\|\mathbf{x}\|_1 = \sum_{k=1}^{n} |x_k|.$$

It is easy to check that this is in fact a norm.

Example 1.8: If $V = C[a, b]$ then for each $f \in V$,

$$\|f\|_\infty = \max\Big\{ |f(x)| \ \Big| \ a \le x \le b \Big\}$$

is a norm. This norm is also called the *uniform norm*.

If V is an inner product space, then there is a natural norm defined thereon. It is given by

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

To prove that this is indeed a norm on V, we first prove the following important inequality.

Theorem 1.6: (**Cauchy-Schwarz Inequality**) *Let V be an inner product space. For each $u, v \in V$ we have*

$$|\langle u, v \rangle| \le \|u\| \cdot \|v\|.$$

Proof: If $\langle u, v \rangle = 0$, then there is nothing to prove. So let us assume that $\langle u, v \rangle \ne 0$ (and thus $u, v \ne \vec{0}$). For convenience set $a = \langle u, v \rangle$ (a may be a complex number). Thus for every real number λ we have

$$0 \le \|u - \lambda a v\|^2 = \langle u - \lambda a v, u - \lambda a v \rangle$$

$$= \langle u, u \rangle - \lambda \langle u, av \rangle - \lambda \langle av, u \rangle + \lambda^2 \langle av, av \rangle$$

$$= \|u\|^2 - \lambda \bar{a} \langle u, v \rangle - \lambda a \langle v, u \rangle + \lambda^2 a \bar{a} \langle v, v \rangle$$

$$= \|u\|^2 - \lambda \bar{a} a - \lambda a \bar{a} + \lambda^2 a \bar{a} \|v\|^2$$

$$= \|u\|^2 - 2\lambda |a|^2 + \lambda^2 |a|^2 \|v\|^2.$$

We consider the last expression as a quadratic polynomial in λ which is non-negative for every real λ. For this to happen the discriminant must be non-positive, from which the result follows. Equivalently, setting $\lambda = \frac{1}{\|v\|^2}$ (the minimal value of the polynomial) in the above expression leads to

$$0 \leq \|u\|^2 \|v\|^2 - |a|^2.$$

Thus $|a| \leq \|u\| \|v\|$. Since $a = \langle u, v \rangle$ this proves the desired inequality.   ■

Theorem 1.7: *Let V be an inner product space. For each $v \in V$ the equation $\|v\| = \sqrt{\langle v, v \rangle}$ defines a norm on V.*

Proof: By the definition of an inner product, the value $\langle v, v \rangle$ is non-negative for every $v \in V$. This being so, the expression $\sqrt{\langle v, v \rangle}$ is well defined (we take the non-negative square root). Now $\|v\| = 0$ if and only if $\langle v, v \rangle = 0$. By the definition of the inner product this condition is equivalent to $v = \vec{0}$. Thus Conditions 1 and 2 are satisfied. To prove Condition 3, let $a \in \mathbb{C}$. Then

$$\|av\|^2 = \langle av, av \rangle = |a|^2 \langle v, v \rangle = |a|^2 \cdot \|v\|^2$$

and thus $\|av\| = |a| \cdot \|v\|$.

It remains to prove the triangle inequality. To this end, let $u, v \in V$. Then

$$\|u + v\|^2 = \langle u + v, u + v \rangle$$

$$= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle$$

$$= \|u\|^2 + \langle u, v \rangle + \overline{\langle u, v \rangle} + \|v\|^2.$$

The expression $\langle u, v \rangle + \overline{\langle u, v \rangle}$ is a real number. From the Cauchy-Schwarz inequality it follows that

$$\left| \langle u, v \rangle + \overline{\langle u, v \rangle} \right| \leq 2 |\langle u, v \rangle| \leq 2 \|u\| \cdot \|v\|.$$

Thus

$$\|u + v\|^2 \leq \|u\|^2 + 2 \|u\| \cdot \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2.$$

This proves the triangle inequality and the theorem.   ■

Among the various examples we gave of norms, only the Euclidean norm is a norm defined via an inner product. The inner product on which it is based is the standard inner product of Example 1.2. From Theorem 1.7 it therefore follows that the Euclidean norm is really and truly a norm.

From the Cauchy-Schwarz and the triangle inequalities we can derive other more specific inequalities.

Example 1.9: If we consider $\mathbb{R}^n$ with the standard inner product, then for each $\mathbf{x} \in \mathbb{R}^n$ the norm of $\mathbf{x}$ is given by

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\sum_{k=1}^{n} x_k^2}.$$

This is also called the *distance* from $\mathbf{x}$ to $\mathbf{0}$ (recall the Pythagorean Theorem). From the Cauchy-Schwarz inequality we obtain, for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\left| \sum_{k=1}^{n} x_k y_k \right| \leq \sqrt{\sum_{k=1}^{n} x_k^2} \sqrt{\sum_{k=1}^{n} y_k^2}.$$

Example 1.10: If we consider $\mathbb{C}^n$ with the standard inner product then as above, the norm of $\mathbf{x} \in \mathbb{C}^n$ is given by

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\sum_{k=1}^{n} |x_k|^2}.$$

From the Cauchy-Schwarz inequality we obtain, for every $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$,

$$(1.1) \qquad \left| \sum_{k=1}^{n} x_k \overline{y_k} \right| \leq \sqrt{\sum_{k=1}^{n} |x_k|^2} \sqrt{\sum_{k=1}^{n} |y_k|^2}.$$

As a result of the triangle inequality we have, for each $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$,

$$(1.2) \qquad \sqrt{\sum_{k=1}^{n} |x_k + y_k|^2} \leq \sqrt{\sum_{k=1}^{n} |x_k|^2} + \sqrt{\sum_{k=1}^{n} |y_k|^2}.$$

Example 1.11: We now use the previous results to prove that ℓ_2 is in fact a linear space, and that the inner product defined in Example 1.4 is indeed an inner product. Let $\mathbf{x}, \mathbf{y} \in \ell_2$. We first prove that the inner product $\mathbf{x} \cdot \mathbf{y}$ is well defined. We recall that

$$\mathbf{x} \cdot \mathbf{y} = \sum_{n=1}^{\infty} x_n \overline{y_n} = \lim_{m \to \infty} \sum_{n=1}^{m} x_n \overline{y_n}.$$

We prove that the above series converges absolutely, i.e., the series $\sum_{n=1}^{\infty} |x_n \overline{y_n}|$ converges. For each natural integer m we have from (1.1)

$$\sum_{n=1}^{m} |x_n \overline{y_n}| \leq \sqrt{\sum_{n=1}^{m} |x_n|^2} \sqrt{\sum_{n=1}^{m} |y_n|^2} \leq \sqrt{\sum_{n=1}^{\infty} |x_n|^2} \sqrt{\sum_{n=1}^{\infty} |y_n|^2} = \|\mathbf{x}\| \cdot \|\mathbf{y}\|.$$

The partial sums $S_m = \sum_{n=1}^{m} |x_n \overline{y_n}|$ are bounded above by $\|\mathbf{x}\| \cdot \|\mathbf{y}\|$. Since these partial sums $\{S_m\}_{m=1}^{\infty}$ are monotonically increasing, they must converge. Thus

the series $\sum_{n=1}^{\infty} x_n \overline{y_n}$ converges, and the inner product on ℓ_2 is well defined. It remains to prove that ℓ_2 is closed under addition. That is, we show that if $\mathbf{x}, \mathbf{y} \in \ell_2$ then $\mathbf{x} + \mathbf{y} \in \ell_2$. We must verify that for $\mathbf{x}, \mathbf{y} \in \ell_2$ the series $\sum_{n=1}^{\infty} |x_n + y_n|^2$ converges. From the inequality (1.2) we have that for each m

$$\sum_{n=1}^{m} |x_n + y_n|^2 \leq \left[\left(\sum_{n=1}^{m} |x_n|^2 \right)^{\frac{1}{2}} + \left(\sum_{n=1}^{m} |y_n|^2 \right)^{\frac{1}{2}} \right]^2$$

$$\leq (\|\mathbf{x}\| + \|\mathbf{y}\|)^2.$$

The partial sums $\sum_{n=1}^{m} |x_n + y_n|^2$ are monotonically increasing and bounded above by the finite value $(\|\mathbf{x}\| + \|\mathbf{y}\|)^2$. Thus the series converges and $\mathbf{x} + \mathbf{y} \in \ell_2$.

Example 1.12:   With the inner product of Example 1.3, the Cauchy-Schwarz inequality on the space of continuous functions $C[a, b]$ is given by

$$\left| \int_a^b f(x) \overline{g(x)} \, dx \right|^2 \leq \left(\int_a^b |f(x)|^2 dx \right) \left(\int_a^b |g(x)|^2 dx \right).$$

Exercises

1.   Prove that the "norms" defined in Examples 1.6–1.8 are in fact norms.

2.   **(a)**   Prove that for any $f, g \in C[a, b]$, with inner product

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} \, dx,$$

   we have

$$\frac{1}{2} \int_a^b \int_a^b |f(x)g(y) - g(x)f(y)|^2 \, dx \, dy = \|f\|^2 \cdot \|g\|^2 - |\langle f, g \rangle|^2.$$

   (b)   Use the equality in (a) to prove the Cauchy-Schwarz inequality on $C[a, b]$.

3.   Let V be an inner product space, $u, v \in V$, $u, v \neq \vec{0}$.
   (a)   Show that $\langle u, v \rangle = \|u\| \cdot \|v\|$ if and only if $u = av$ for some $a \in \mathbb{C}$.
   (b)   Show that $\|u + v\| = \|u\| + \|v\|$ if and only if $u = av$ for some $a \geq 0$.

4.   Let V be an inner product space. Prove that for all $u, v \in V$ the "Parallelogram law"

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$$

   holds.

5.   Let V be a real inner product space. Prove that for each $u, v \in V$

$$\langle u, v \rangle = \frac{1}{4} \|u + v\|^2 - \frac{1}{4} \|u - v\|^2.$$

6.   Let V be a complex inner product space. Prove that for each $u, v \in V$

$$\langle u, v \rangle = \frac{1}{4}\|u + v\|^2 - \frac{1}{4}\|u - v\|^2 + \frac{i}{4}\|u + iv\|^2 - \frac{i}{4}\|u - iv\|^2.$$

7.   On the basis of the previous two exercises prove that in every normed linear space for which the parallelogram law holds we can define an inner product associated with the given norm.

8.   Prove that for any natural number n and real numbers $x_1, x_2, \ldots, x_n$ we have the inequality

$$\left| \frac{1}{n} \sum_{k=1}^{n} x_k \right| \leq \left(\frac{1}{n} \sum_{k=1}^{n} x_k^2 \right)^{\frac{1}{2}}.$$

3. Orthogonal and Orthonormal Systems

Definition 1.8: *Let V be an inner product space and $u, v \in V$. We say that u and v are* perpendicular *to each other, or* orthogonal, *if $\langle u, v \rangle = 0$. We denote this fact by $u \perp v$.*

Definition 1.9: *Let V be an inner product space. A finite sequence $\{u_k\}_{k=1}^{n}$ or an infinite sequence $\{u_k\}_{k=1}^{\infty}$ of vectors in V is called an* orthogonal system *if $u_k \neq \vec{0}$ for each k and $u_k \perp u_j$ for all $k \neq j$. If, in addition, $\|u_k\| = 1$ for every k, then we say that we have an* orthonormal system.

Every vector of length 1 (i.e., for which $\|u\| = 1$) is called a *unit vector*. This being so, an orthonormal system is an orthogonal system where each of its vectors is a unit vector. If we are given an orthogonal system $\{u_k\}_{k=1}^{n}$ (where n is finite or infinite) then we can easily obtain an orthonormal system by "normalizing" each vector of the system. For each k, set $e_k = \frac{u_k}{\|u_k\|}$. It follows that $\|e_k\| = 1$ for each k, and $e_k \perp e_j$ for $k \neq j$. Thus $\{e_k\}_{k=1}^{n}$ is an orthonormal system of the same size as the orthogonal system $\{u_k\}_{k=1}^{n}$. More importantly, $\mathrm{span}\{e_k\}_{k=1}^{n} = \mathrm{span}\{u_k\}_{k=1}^{n}$.

Proposition 1.10: *Let $\{u_k\}_{k=1}^{n}$ be a finite orthogonal system in an inner product space V. Then the vectors $\{u_k\}_{k=1}^{n}$ are linearly independent.*

Proof: Assume that

$$a_1 u_1 + a_2 u_2 + \cdots + a_n u_n = \vec{0}, \qquad a_1, a_2, \ldots, a_n \in \mathbb{C}.$$

For each $k \in \{1, 2, \ldots, n\}$

$$0 = \left\langle \vec{0}, u_k \right\rangle = \left\langle a_1 u_1 + a_2 u_2 + \cdots + a_n u_n, u_k \right\rangle$$

$$= a_1 \left\langle u_1, u_k \right\rangle + a_2 \left\langle u_2, u_k \right\rangle + \cdots + a_k \left\langle u_k, u_k \right\rangle + \cdots + a_n \left\langle u_n, u_k \right\rangle$$

$$= a_1 {\cdot} 0 + \cdots + a_{k-1} {\cdot} 0 + a_k {\cdot} \left\langle u_k, u_k \right\rangle + a_{k+1} {\cdot} 0 + \cdots + a_n {\cdot} 0$$

$$= a_k \left\langle u_k, u_k \right\rangle .$$

Since $\langle u_k, u_k \rangle = \|u_k\|^2 \neq 0$, we must have $a_k = 0$. Thus $a_k = 0$ for each k, and the vectors $\{u_k\}_{k=1}^n$ are linearly independent.   ∎

Let $\{v_k\}_{k=1}^n$ be any finite system of linearly independent vectors in an inner product space V. Does there then exist an orthonormal system $\{e_k\}_{k=1}^n$ for which

$$\mathrm{span}\{e_k\}_{k=1}^n = \mathrm{span}\{v_k\}_{k=1}^n \ ?$$

And if yes, is there a simple method of constructing such a system $\{e_k\}_{k=1}^n$ from the original system $\{v_k\}_{k=1}^n$? The answer to both questions is yes. One known method of constructing an orthonormal system from the original system is called the *Gram-Schmidt process*. We will quickly sketch this process in the next section.

One of the many advantages of an orthonormal system is the relative ease with which we can determine the coefficients of any vector in its linear span. The formula for the coefficients is to be found in this next proposition.

Proposition 1.11: *Assume V is an inner product space, and $\{e_1, \ldots, e_n\}$ an orthonormal system therein. If $u = \sum_{k=1}^n a_k e_k$ then for each k we have $a_k = \langle u, e_k \rangle$.*

Proof:

$$\langle u, e_k \rangle = \langle a_1 e_1 + a_2 e_2 + \cdots + a_n e_n, e_k \rangle$$

$$= a_1 \langle e_1, e_k \rangle + a_2 \langle e_2, e_k \rangle + \cdots + a_k \langle e_k, e_k \rangle + \cdots + a_n \langle e_n, e_k \rangle$$

$$= a_1 {\cdot} 0 + a_2 {\cdot} 0 + \cdots + a_k {\cdot} 1 + \cdots + a_n {\cdot} 0$$

$$= a_k .$$
                                                                                                                        ∎

Thus if $\{e_1, \ldots, e_n\}$ is an orthonormal system, then for each u in its span we have

$$u = \sum_{k=1}^n a_k e_k = \sum_{k=1}^n \langle u, e_k \rangle e_k.$$

The coefficient a_k is uniquely determined by the formula $a_k = \langle u, e_k \rangle$. Note that a_k is only dependent upon u and e_k. It is not dependent upon any of the other basis vectors e_j, $j \neq k$. In general, if $u \in \text{span}\{v_1, \ldots, v_n\}$, where $\{v_1, \ldots, v_n\}$ is not an orthonormal system, then each of the coefficients a_k in the representation of u as a linear combination of $v_1, \cdots, v_n$ will depend upon every one of $v_1, \ldots, v_n$ in some complicated manner. The above coefficients are of such importance that they have a name.

Definition 1.12: *Let V be an inner product space and $\{e_k\}_{k=1}^n$ an orthonormal system therein (n may be finite or infinite). Let $u \in V$. The numbers $\langle u, e_k \rangle$ are called the* generalized Fourier coefficients *of the vector u with respect to the given orthonormal system.*

An additional property of an orthonormal system is presented in the next proposition.

Proposition 1.13: *Let V be an inner product space and $\{e_1, \ldots, e_n\}$ an orthonormal system therein. If $\{a_k\}_{k=1}^n$ and $\{b_k\}_{k=1}^n$ are any sequences of scalars, then*

$$\left\langle \sum_{k=1}^n a_k e_k, \sum_{k=1}^n b_k e_k \right\rangle = \sum_{k=1}^n a_k \overline{b_k},$$

i.e., for $u, v \in \text{span}\{e_1, \ldots, e_n\}$,

$$\langle u, v \rangle = \sum_{k=1}^n \langle u, e_k \rangle \, \overline{\langle v, e_k \rangle}.$$

The proof of this proposition is similar to the proofs of Propositions 1.10 and 1.11. Note that there is also an analog of the last formula for calculating $\langle u, v \rangle$ in the case when our basis is not orthonormal. But it contains n^2 rather than only n terms.

This next theorem may be viewed as a generalization of the Pythagorean Theorem in an inner product space.

Theorem 1.14: (Generalized Pythagorean Theorem) *Let V be an inner product space.*

(a) *Let $\{u_1, \ldots, u_n\}$ be an orthogonal system in V, and $a_1, \ldots, a_n$ scalars. Then*

$$\left\| \sum_{k=1}^n a_k u_k \right\|^2 = \sum_{k=1}^n |a_k|^2 \|u_k\|^2.$$

(b) Let $\{e_1, \ldots, e_n\}$ be an orthonormal system in V. Then for every $u \in$ span$\{e_1, \ldots, e_n\}$

$$\|u\|^2 = \sum_{k=1}^{n} |\langle u, e_k \rangle|^2.$$

Proof: **(a)** This follows from the definitions.

$$\|\textstyle\sum_{k=1}^{n} a_k u_k\|^2 = \left\langle \textstyle\sum_{k=1}^{n} a_k u_k, \sum_{j=1}^{n} a_j u_j \right\rangle = \textstyle\sum_{k=1}^{n} \sum_{j=1}^{n} a_k \overline{a}_j \langle u_k, u_j \rangle$$

$$= \textstyle\sum_{k=1}^{n} |a_k|^2 \|u_k\|^2.$$

(b) This is an immediate consequence of (a) or of Proposition 1.13. ∎

We should consider (b) of the above theorem as a natural generalization of the Euclidean norm which was defined on $\mathbb{R}^n$ and $\mathbb{C}^n$ (see Example 1.5). From the isomorphism which exists between $W = \text{span}\{e_1, \ldots, e_n\}$ and $\mathbb{C}^n$ (here we assume that V is a complex linear space) we can identify each vector $u \in W$ with its sequence of generalized Fourier coefficients $(a_1, a_2, \ldots, a_n) \in \mathbb{C}^n$. Theorem 1.14 effectively says that $\|u\| = \|(a_1, \ldots, a_n)\|$, where the norm on the right-hand side is the Euclidean norm on $\mathbb{C}^n$. Thus there is not, in some sense, a significant difference between W and $\mathbb{C}^n$.

Exercises

1.   Let $C[-1, 1]$ be the space of continuous functions $f : [-1, 1] \to \mathbb{C}$ with inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\overline{g(x)}\, dx.$$

   (a) Let $P_0(x) = 1$, $P_1(x) = x$, and $P_2(x) = 1 - 3x^2$. Prove that this set of polynomials is orthogonal in $C[-1, 1]$.

   (b) Find constants a, b, and c such that the polynomial

$$P_3(x) = a + bx + cx^2 + x^3$$

   is orthogonal (perpendicular) to each of the previous polynomials.

2.   Let $\mathcal{P}_2$ be the space of real polynomials of degree less than or equal to 2. To each $f, g \in \mathcal{P}_2$, define

$$\langle f, g \rangle = \int_{0}^{\infty} f(x)g(x)e^{-x} dx.$$

    (a) Prove that this is an inner product on $\mathcal{P}_2$.

    (b) Show that the set $\{1,\ 1-x,\ 1-2x+\frac{1}{2}x^2\}$ is an orthonormal system with respect to this inner product.

3.   Let $C^1[0,1]$ be the space of continuous functions $f : [0,1] \to \mathbb{C}$ with a continuous first derivative on $[0,1]$.

    (a) Prove that

$$\langle f,g\rangle = f(0)\cdot\overline{g(0)} + \int_0^1 f'(x)\overline{g'(x)}\,dx$$

    is an inner product on $C^1[0,1]$.

    (b) Find an orthonormal system $\{h_1,h_2,h_3\}$ in $C^1[0,1]$, with respect to this inner product, for which

$$\operatorname{span}\{h_1,h_2,h_3\} = \operatorname{span}\{1,x,x^2\}.$$

4. Orthogonal Projections and Approximation in the Mean

Let V be an inner product space, and $\{e_1,\ldots,e_n\}$ an orthonormal system therein. Set $W = \operatorname{span}\{e_1,\ldots,e_n\}$. Let u be an arbitrary vector in V. In the previous section we defined the generalized Fourier coefficients of u to be the $\langle u,e_k\rangle$, $k = 1,\ldots,n$. If $u \notin W$ then $u \neq \sum_{k=1}^n \langle u,e_k\rangle e_k$, since u is not a linear combination of the e_1, ..., e_n. Nevertheless there exists an important connection between u and $\sum_{k=1}^n \langle u,e_k\rangle e_k$. In this section we study this link in some detail.

For each $u \in V$, we set $\tilde{u} = \sum_{k=1}^n \langle u,e_k\rangle e_k$. The vector $\tilde{u}$ is said to be the *orthogonal projection of u on W*.

Proposition 1.15: *For each $u \in V$,*
(a) $\langle u - \tilde{u}, w\rangle = 0$ *for all $w \in W$.*
(b) $\|u - w\|^2 = \|u - \tilde{u}\|^2 + \|\tilde{u} - w\|^2$ *for all $w \in W$.*

Remark: Part (a) of this proposition says that the vector $u - \tilde{u}$ is orthogonal to every vector $w \in W$. This being so, we sometimes say that $u - \tilde{u}$ is orthogonal to the subspace W, and write $u - \tilde{u} \perp W$.

Proof: **(a)** We first prove that $\langle u - \tilde{u}, e_j\rangle = 0$ for every $j = 1,2,\ldots,n$.

$$\langle u - \tilde{u}, e_j\rangle = \langle u, e_j\rangle - \left\langle \sum_{k=1}^n \langle u,e_k\rangle e_k, e_j \right\rangle = \langle u,e_j\rangle - \sum_{k=1}^n \langle u,e_k\rangle \langle e_k,e_j\rangle$$

$$= \langle u,e_j\rangle - \langle u,e_j\rangle \langle e_j,e_j\rangle = \langle u,e_j\rangle - \langle u,e_j\rangle = 0.$$

We now take an arbitrary $w \in W$. Thus $w = \sum_{j=1}^{n} b_j e_j$ for some scalars b_1, b_2, ..., b_n, and

$$\langle u - \tilde{u}, w \rangle = \left\langle u - \tilde{u}, \sum_{j=1}^{n} b_j e_j \right\rangle = \sum_{j=1}^{n} \overline{b_j} \langle u - \tilde{u}, e_j \rangle = \sum_{j=1}^{n} \overline{b_j} \cdot 0 = 0.$$

(b)

From part (a) we have $(u - \tilde{u}) \perp w$ for every $w \in W$. Thus $(u - \tilde{u}) \perp (\tilde{u} - w)$ since $\tilde{u} - w \in W$. From part (a) of Theorem 1.14 we obtain

$$\|u - w\|^2 = \|u - \tilde{u} + \tilde{u} - w\|^2 = \|u - \tilde{u}\|^2 + \|\tilde{u} - w\|^2$$

and this proposition is proved. ∎

We present consequences of this last result after the definition of distance (which we have already met in Section 1.2).

Definition 1.16: *Let V be any normed linear space. For each $u, v \in V$, the distance between u and v is the non-negative number $\|u - v\|$.*

To justify the term "distance" we list a number of basic properties which our definition of distance satisfies:

(a) For each $u, v \in V$, we have $\|u - v\| = \|v - u\|$. That is, the distance between u and v equals the distance between v and u.

(b) For each $u \in V$, $\|u - u\| = 0$. That is, u is distant zero from itself.

(c) For every $u, v \in V$ we have $\|u - v\| = 0$ only if $u = v$. That is, if the distance between u and v is zero then u equals v.

(d) For every $u, v, w \in V$, we have $\|u - w\| \leq \|u - v\| + \|v - w\|$. That is, the distance between u and w is always less than or equal to the sum of the distances from u and w to any third point v.

Properties (a)–(c) are direct consequences of the definition of a norm. Property (d) follows from the triangle inequality

$$\|u - w\| = \|u - v + v - w\| \leq \|u - v\| + \|v - w\|.$$

The main result of this section is the following characterization of the vector closest to u in W.

Theorem 1.17: *Let V be an inner product space and $\{e_1, \ldots, e_n\}$ an orthonormal system therein. Set $W = \text{span}\{e_1, \ldots, e_n\}$, and let $u \in V$. The vector*

$\tilde{u} = \sum_{k=1}^{n} \langle u, e_k \rangle \, e_k$ *is a closest vector to u in W. In addition, $\tilde{u}$ is the unique vector in W whose distance from u is minimal.*

Proof: We must show that $\|u - \tilde{u}\| \leq \|u - w\|$ for all $w \in W$. This is the meaning of the expression "a closest vector to u from W". This is an immediate consequence of part (b) of Proposition 1.15. For each $w \in W$,

$$\|u - w\|^2 = \|u - \tilde{u}\|^2 + \|\tilde{u} - w\|^2$$

and thus $\|u - \tilde{u}\| \leq \|u - w\|$ for each $w \in W$. The uniqueness of $\tilde{u}$ as the closest vector is a result of this same equality. If $\|u - \tilde{u}\| = \|u - w\|$ for some $w \in W$, then $\|\tilde{u} - w\| = 0$, which implies that $w = \tilde{u}$. ■

As we see $\tilde{u}$ has a simple affinity to u. It is the unique closest vector to u from W. Naturally if $u \in W$ then $\tilde{u} = u$. We will further consider this relationship when we deal with how $\tilde{u}$ might "represent" u and why.

Example 1.13: For the linear space $C[-1, 1]$, the norm

$$\|f\| = \left(\int_{-1}^{1} |f(x)|^2 \, dx \right)^{\frac{1}{2}}$$

comes from the inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x) \overline{g(x)} \, dx.$$

In order to determine the closest function to $f \in C[-1, 1]$ in the subspace $W = \text{span}\{1, x\}$ we must find scalars a^* and b^* for which

$$\|f - (a^* + b^* x)\| \leq \|f - (a + bx)\|$$

for all $a, b \in \mathbb{C}$. The functions $e_1(x) = \frac{1}{\sqrt{2}}$ and $e_2(x) = \sqrt{\frac{3}{2}} x$ form an orthonormal basis for the space W. Thus our problem is equivalent to that of finding scalars c^* and d^* for which

$$\|f - (c^* e_1 + d^* e_2)\| \leq \|f - (c e_1 + d e_2)\|$$

for all $c, d \in \mathbb{C}$. According to Theorem 1.17 there is a unique solution to this problem and it is given by

$$c^* = \langle f, e_1 \rangle, \qquad d^* = \langle f, e_2 \rangle.$$

If, for example, $f(x) = x^3$ then

$$c^* = \langle f, e_1 \rangle = \int_{-1}^{1} x^3 \frac{1}{\sqrt{2}} \, dx = 0,$$

$$d^* = \langle f, e_2 \rangle = \int_{-1}^{1} x^3 \sqrt{\frac{3}{2}} x \, dx = \frac{\sqrt{6}}{5},$$

and thus

$$c^* e_1(x) + d^* e_2(x) = \frac{\sqrt{6}}{5} \sqrt{\frac{3}{2}} x = \frac{3}{5} x$$

is the closest function to x^3 in W with respect to the given norm.

One additional consequence of Proposition 1.15 is the following.

Proposition 1.18: *Let* $\{e_1, \ldots, e_n\}$ *be an orthonormal system in an inner product space* V. *Then for each* $u \in V$ *the following inequality holds:*

$$\sum_{k=1}^{n} |\langle u, e_k \rangle|^2 \leq \|u\|^2.$$

Proof: Setting $w = \vec{0}$ in part (b) of Proposition 1.15 we obtain

$$\|u\|^2 = \|u - \tilde{u}\|^2 + \|\tilde{u}\|^2.$$

Thus $\|\tilde{u}\|^2 \leq \|u\|^2$. From part (b) of Theorem 1.14

$$\|\tilde{u}\|^2 = \sum_{k=1}^{n} |\langle u, e_k \rangle|^2$$

and the inequality follows.   ∎

It is not difficult to establish that $\sum_{k=1}^{n} |\langle u, e_k \rangle|^2 = \|u\|^2$ if and only if $u \in$ span$\{e_1, \ldots, e_n\}$.

The Gram-Schmidt Process

Let V be an inner product space and $\{v_1, \ldots, v_n\}$ a system of linearly independent vectors in V. We will describe a method whereby we obtain an orthonormal system $\{e_1, \ldots, e_n\}$ for which

$$\text{span}\{v_1, \ldots, v_n\} = \text{span}\{e_1, \ldots, e_n\}.$$

The process is an n-step method whereby at step k, $1 \leq k \leq n$, we build the vector e_k in such a way that

$$\text{span}\{v_1, \ldots, v_k\} = \text{span}\{e_1, \ldots, e_k\}.$$

Step 1: From the fact that the system $\{v_1, \ldots, v_n\}$ is linearly independent, we have that $v_1 \neq \vec{0}$. We define e_1 by

$$e_1 = \frac{v_1}{\|v_1\|}.$$

It is clear that $\|e_1\| = 1$ and that $\text{span}\{v_1\} = \text{span}\{e_1\}$.

Step 2: Let $W_1 = \text{span}\{e_1\}$ and let $\tilde{v}_2 = \langle v_2, e_1 \rangle e_1$ be the orthogonal projection of v_2 on W_1. From Proposition 1.15(a), $v_2 - \tilde{v}_2 \perp e_1$. In addition $v_2 - \tilde{v}_2 \neq \vec{0}$ since otherwise we would have $v_2 \in W_1$, contradicting the fact that the system $\{v_1, v_2\}$ is linearly independent. Thus we may define

$$e_2 = \frac{v_2 - \tilde{v}_2}{\|v_2 - \tilde{v}_2\|}.$$

It follows that $e_2 \perp e_1$, $\|e_2\| = 1$, and $\text{span}\{e_1, e_2\} = \text{span}\{v_1, v_2\}$. In other words, the system $\{e_1, e_2\}$ is an orthonormal system and it spans the same subspace as that spanned by $\{v_1, v_2\}$.

Step k: Let $W_{k-1} = \text{span}\{e_1, \ldots, e_{k-1}\}$ and let $\tilde{v}_k = \sum_{j=1}^{k-1} \langle v_k, e_j \rangle e_j$ be the orthogonal projection of v_k on W_{k-1}. From Proposition 1.15(a), $v_k - \tilde{v}_k \perp W_{k-1}$. In addition, $v_k - \tilde{v}_k \neq \vec{0}$. Thus we may define

$$e_k = \frac{v_k - \tilde{v}_k}{\|v_k - \tilde{v}_k\|}.$$

As a consequence $\text{span}\{e_1, \ldots, e_k\} = \text{span}\{v_1, \ldots, v_k\}$, and $\{e_1, \ldots, e_k\}$ is an orthonormal system.

We continue this process until the nth step where we obtain the desired orthonormal system $\{e_1, \ldots, e_n\}$.

Exercises

1. Let $f \in C[-\pi, \pi]$. For each $\alpha, \beta, \gamma \in \mathbb{C}$ define

$$F(\alpha, \beta, \gamma) = \frac{1}{\pi} \int_{-\pi}^{\pi} |f(x) - \alpha - \beta \cos x - \gamma \cos 10x|^2 dx.$$

Prove that F attains its minimum at a unique point $(\alpha_0, \beta_0, \gamma_0)$ and find this point when

     (a) $f(x) = \cos^2 x$ **(b)** $f(x) = x^3$ **(c)** $f(x) = \sin x$

     (d) $f(x) = 1 - 2\cos x$ **(e)** $f(x) = |x|$ **(f)** $f(x) = |\sin x|$

2. On the linear space $C[0, 2\pi]$ we define the inner product

$$\langle f, g \rangle = \int_0^{2\pi} f(x)\overline{g(x)}\, dx.$$

     (a) Prove that the set $S = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}} \right\}$ is an orthonormal system.

(b)   Let W be the subspace spanned by S, and let $f(x) = x$ on the interval $[0, 2\pi]$. Find the function g in S which is closest to f (that is to say, for which $\|f - g\|$ is minimal).

3.   Consider the space $C[-\pi, \pi]$ with the inner product

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}\, dx.$$

Let $W = \text{span}\{1,\ \sin x,\ \cos x,\ x\}$, and let $f(x) = |x|$. Find the function $g \in W$ for which $\|f - g\|$ is as small as possible.

4.   For each $\alpha, \beta, \gamma \in \mathbb{C}$ define

$$F(\alpha, \beta, \gamma) = \frac{1}{\pi} \int_{-\pi}^{\pi} |x - \alpha - \beta \cos x - \gamma \sin 2x|^2 dx.$$

Show that F attains its minimum at exactly one point $(\alpha_0, \beta_0, \gamma_0)$, and find this point.

5.   In the space $C[-1, 1]$ we are given the functions

$$f_0(x) = 1,\ \ f_1(x) = x + a,\ \ f_2(x) = x^2 + bx + c,\ \ f_3(x) = x^3 + Ax^2 + Bx + C$$

and it is known that $\{f_0, f_1, f_2, f_3\}$ is an orthonormal system in $C[-1, 1]$ with respect to the inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\overline{g(x)}\, dx.$$

(a)   Calculate a, b, c.

(b)   For each $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ define

$$F(\alpha, \beta, \gamma, \delta) = \int_{-1}^{1} |x^4 - \alpha f_0(x) - \beta f_1(x) - \gamma f_2(x) - \delta f_3(x)|^2 dx.$$

Prove that F attains its minimum at exactly one point $(\alpha_0, \beta_0, \gamma_0, \delta_0)$, and find this point.

5. Infinite Orthonormal Systems

Let V be an inner product space. In this section we assume that $\dim(V) = \infty$. Let $\{e_1, e_2, \ldots\}$ be an orthonormal system with an infinite number of vectors. We remark that the concept of a basis in an infinite dimensional linear space is problematic, to say the least. This being so, we should not assume a priori

that the given orthonormal system is a basis or even a spanning set for V. We will later return to this troublesome problem.

Theorem 1.19: (Bessel's Inequality) *For each $u \in V$, the series $\sum_{n=1}^{\infty} |\langle u, e_n \rangle|^2$ converges. In addition, the inequality*

$$\sum_{n=1}^{\infty} |\langle u, e_n \rangle|^2 \leq \|u\|^2$$

holds.

Proof: This result is an immediate consequence of Proposition 1.18. For each m,

$$S_m = \sum_{n=1}^{m} |\langle u, e_n \rangle|^2 \leq \|u\|^2.$$

That is, the sequence of partial sums $\{S_m\}_{m=1}^{\infty}$ is bounded above by $\|u\|^2$. Since $\{S_m\}_{m=1}^{\infty}$ is a monotonically increasing sequence, it converges to a finite sum. Thus

$$\lim_{m \to \infty} \sum_{n=1}^{m} |\langle u, e_n \rangle|^2 \leq \|u\|^2.$$ ∎

If we have the equality $\sum_{n=1}^{\infty} |\langle u, e_n \rangle|^2 = \|u\|^2$ then we say that *Parseval's identity* (sometimes called *Parseval's equality*) holds for u. An immediate consequence of Bessel's inequality is

Theorem 1.20: (Riemann-Lebesgue Lemma) *Let $\{e_1, e_2, \ldots\}$ be an infinite orthonormal system in an inner product space V. Let $u \in V$. Then*

$$\lim_{n \to \infty} \langle u, e_n \rangle = 0.$$

Proof: From Bessel's inequality, the series $\sum_{n=1}^{\infty} |\langle u, e_n \rangle|^2$ converges. If a series converges then the nth coefficient tends to zero as n tends to infinity. ∎

As stated at the beginning of this section, the concept of a basis for an infinite dimensional linear space is problematic. This whole question must be treated carefully. One of the first problems which arises has to do with the proper definition of an "infinite linear combination". Let us be more specific. We are given an inner product space V, an infinite sequence of vectors $\{u_1, u_2, \ldots\}$ therein, and a sequence of scalars $\{a_n\}_{n=1}^{\infty}$. Can we give any meaning to the expression "$\sum_{n=1}^{\infty} a_n u_n$"? We are talking about the infinite sum of vectors

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

and not an infinite sum of numbers! Even if we succeed in giving some meaning to the infinite sum of vectors, it will also be necessary to check that with this meaning a number of properties of a basis are preserved. To this end, we make use of the concept of convergence in norm.

Definition 1.21: *Let $\{w_m\}_{m=1}^{\infty}$ be an infinite sequence of vectors in a normed linear space V. We say that the sequence* converges in norm *to the vector $w \in V$ if*

$$\lim_{m\to\infty} \|w - w_m\| = 0.$$

That is to say, for each $\epsilon > 0$ there exists an $m(\epsilon)$ such that for all $m \geq m(\epsilon)$ we have $\|w - w_m\| < \epsilon$.

We can now give some meaning to what we mean by an "infinite linear combination" of vectors.

Definition 1.22: *Let $\{u_1, u_2, \ldots\}$ be an infinite sequence of vectors in the normed linear space V. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of scalars. We say that the series $\sum_{n=1}^{\infty} a_n u_n$* converges in norm *to the vector $w \in V$, and write $w = \sum_{n=1}^{\infty} a_n u_n$, if the partial sums $w_m = \sum_{n=1}^{m} a_n u_n$ converge in norm to w. In other words, the series $\sum_{n=1}^{\infty} a_n u_n$ converges in norm to w if*

$$\lim_{m\to\infty} \left\| w - \sum_{n=1}^{m} a_n u_n \right\| = 0.$$

The interpretation of the expression "the vector w is contained in the span of the infinite sequence $\{u_1, u_2, \ldots\}$" is that there exists a sequence of scalars $\{a_n\}_{n=1}^{\infty}$ such that the linear combination $a_1 u_1 + \cdots + a_m u_m$ approaches w, as m grows. The "nearness" of vectors, in a linear space with a norm, is measured by the distance between them.

We can now formulate further desired properties of infinite orthonormal systems.

Definition 1.23: *Let $\{e_1, e_2, \ldots\}$ be an infinite orthonormal system in an inner product space V. We will say that the system is* closed *in V if for every $u \in V$ we have*

$$\lim_{m\to\infty} \left\| u - \sum_{n=1}^{m} \langle u, e_n \rangle e_n \right\| = 0.$$

Recall that $\sum_{n=1}^{m} \langle u, e_n \rangle e_n$ is the vector closest to u in span$\{e_1, \ldots, e_m\}$. Thus a system is closed in V if for each element of V there is some infinite linear combination from the system which converges in norm to the element.