Lecture II Banach and Hilbert Spaces — Fourier Series and OrthogonalPolynomials

Lei Du
dulei@dlut.edu.cn

School of Mathematical Sciences
Dalian University of Technology

September, 2022

Contents

Introduction

This chapter is dedicated to normed vector spaces (n.v.s.) and pre-Hilbert spaces (or inner product spaces), or more specifically v.s. equipped with a Euclidean (for real v.s.) or Hermitian (for complex v.s.) inner product. In the case of finite-dimensional v.s., we obtain Euclidian and Hermitian spaces, respectively. In the case of infinite-dimensional v.s., Banach and Hilbert spaces will be introduced.

Introduction

This chapter is dedicated to normed vector spaces (n.v.s.) and pre-Hilbert spaces (or inner product spaces), or more specifically v.s. equipped with a Euclidean (for real v.s.) or Hermitian (for complex v.s.) inner product. In the case of finite-dimensional v.s., we obtain Euclidian and Hermitian spaces, respectively. In the case of infinite-dimensional v.s., Banach and Hilbert spaces will be introduced.

The objective of this chapter is to introduce the basic concepts underlying Banach and Hilbert spaces and to give an illustration thereof through the problem of function approximation in the form of expansions over Hilbert bases. Two types of bases will be considered: trigonometric bases leading to Fourier series and bases of orthogonal polynomials.

Introduction

This chapter is dedicated to normed vector spaces (n.v.s.) and pre-Hilbert spaces (or inner product spaces), or more specifically v.s. equipped with a Euclidean (for real v.s.) or Hermitian (for complex v.s.) inner product. In the case of finite-dimensional v.s., we obtain Euclidian and Hermitian spaces, respectively. In the case of infinite-dimensional v.s., Banach and Hilbert spaces will be introduced.

The objective of this chapter is to introduce the basic concepts underlying Banach and Hilbert spaces and to give an illustration thereof through the problem of function approximation in the form of expansions over Hilbert bases. Two types of bases will be considered: trigonometric bases leading to Fourier series and bases of orthogonal polynomials.

This chapter is not intended to make an exhaustive presentation of Hilbert spaces and of the theory of function approximation. In particular, the demonstrations of convergence properties of Fourier series will not be provided here. Our goal is rather to make the link between Hilbert spaces and function approximation through the use of expansions over Hilbert

Introduction

Representation and approximation of functions play a very important role in signal and image processing, as Fourier series constitute the most famous example. It is also possible to use Hilbert bases of orthogonal polynomials, such as Legendre, Hermite, Laguerre, or Chebyshev polynomials, for the modeling of linear systems and nonlinear systems. In addition to their application for function approximation, orthogonal polynomials are also used in combinatorics, coding and probability theory, as well as for solving interpolation problems. Other examples of Hilbert bases that are not considered in this chapter are wavelet bases used for multiresolution (or multi-scale) analysis. In image processing, this type of analysis is intended to decompose an image to extract features that enable segmentation, classification, shape recognition, or compression to be performed.

Definition of distance

A metric space is a space equipped with a distance.

Definition of distance

A metric space is a space equipped with a distance.

Let E be a set. A map $d : E \times E \to \mathbb{R}^+$is a distance, also called a metric, if it satisfies the following properties for all $x, y, z \in E$:

- Non-negativity: $d(x, y) \geq 0$
- Strict positivity: $d(x, y) = 0 \Leftrightarrow x = y$
- Symmetry: $d(x, y) = d(y, x)$
- Triangle inequality: $d(x, y) \leq d(x, z) + d(z, y)$

Definition of distance

A metric space is a space equipped with a distance.

Let E be a set. A map $d : E \times E \to \mathbb{R}^+$is a distance, also called a metric, if it satisfies the following properties for all $x, y, z \in E$:

- Non-negativity: $d(x, y) \geq 0$
- Strict positivity: $d(x, y) = 0 \Leftrightarrow x = y$
- Symmetry: $d(x, y) = d(y, x)$
- Triangle inequality: $d(x, y) \leq d(x, z) + d(z, y)$

The set E endowed with the distance d is denoted by (E, d). When the property of definite positivity is not imposed, that is, if $d(x, y)$ may be zero for a couple of distinct points x and y, it is said that d defines a semi-distance.

Definition of topology

The notion of distance, which does not require defining a structure, makes it possible to generate a topology based on the concepts of open ball and open set:

Definition of topology

The notion of distance, which does not require defining a structure, makes it possible to generate a topology based on the concepts of open ball and open set:

- For all $x \in E$, the open (or closed) ball of center x and radius r is defined as $\mathcal{B}_r(x) = \{y \in E : d(x, y) < r\}$ (or $\overline{\mathcal{B}}_r(x) = \{y \in E : d(x, y) \leq r\}$), with $r \in \mathbb{R}^+$.

Definition of topology

The notion of distance, which does not require defining a structure, makes it possible to generate a topology based on the concepts of open ball and open set:

- For all $x \in E$, the open (or closed) ball of center x and radius r is defined as $\mathcal{B}_r(x) = \{y \in E : d(x, y) < r\}$ (or $\overline{\mathcal{B}}_r(x) = \{y \in E : d(x, y) \leq r\}\big)$, with $r \in \mathbb{R}^+$.
- A subset F of a metric space (E, d) is an open set if and only if for all $x \in F$, there is an open ball $\mathcal{B}_r(x) \subset F$.

Definition of topology

The notion of distance, which does not require defining a structure, makes it possible to generate a topology based on the concepts of open ball and open set:

- For all $x \in E$, the open (or closed) ball of center x and radius r is defined as $\mathcal{B}_r(x) = \{y \in E : d(x, y) < r\}$ (or $\overline{\mathcal{B}}_r(x) = \{y \in E : d(x, y) \leq r\}\big)$, with $r \in \mathbb{R}^+$.
- A subset F of a metric space (E, d) is an open set if and only if for all $x \in F$, there is an open ball $\mathcal{B}_r(x) \subset F$.
- A subset F of a metric space (E, d) is closed if it is the complement of an open set.

Definition of topology

The notion of distance, which does not require defining a structure, makes it possible to generate a topology based on the concepts of open ball and open set:

- For all $x \in E$, the open (or closed) ball of center x and radius r is defined as $\mathcal{B}_r(x) = \{y \in E : d(x, y) < r\}$ (or $\overline{\mathcal{B}}_r(x) = \{y \in E : d(x, y) \leq r\}\big)$, with $r \in \mathbb{R}^+$.
- A subset F of a metric space (E, d) is an open set if and only if for all $x \in F$, there is an open ball $\mathcal{B}_r(x) \subset F$.
- A subset F of a metric space (E, d) is closed if it is the complement of an open set.

The topology of (E, d) is the set of open sets $F \subset E$. Open sets are used to define the concepts of convergence and continuity.

Definition of topology

In a metric space, open and closed sets satisfy the following properties:

- If $F_1, \cdots, F_N$ are open sets, their intersection $\bigcap\limits_{n=1}^{N} F_n$ is an open set.
- An arbitrary union of open sets is open.
- If $F_1, \cdots, F_N$ are closed sets, their union $\bigcup\limits_{n=1}^{N} F_n$ is a closed set.
- An arbitrary intersection of closed sets is closed.

Definition of topology

In a metric space, open and closed sets satisfy the following properties:

- If $F_1, \cdots, F_N$ are open sets, their intersection $\bigcap_{n=1}^{N} F_n$ is an open set.
- An arbitrary union of open sets is open.
- If $F_1, \cdots, F_N$ are closed sets, their union $\bigcup_{n=1}^{N} F_n$ is a closed set.
- An arbitrary intersection of closed sets is closed.

When a v.s. is endowed with a topology, it is called a topological v.s. The Banach and Hilbert spaces are two well-known examples of topological v.s.

Examples of distances

Below a few examples of distances.

- For $E = \mathbb{R}$ (or $\mathbb{C}$), the map $\mathbb{R}^2 \ni (x, y) \mapsto d(x, y) = |x - y|$ defines a distance between x and y, where $|\cdot|$ designates the absolute value (or the modulus).

Examples of distances

Below a few examples of distances.

- For $E = \mathbb{R}$ (or $\mathbb{C}$), the map $\mathbb{R}^2 \ni (x, y) \mapsto d(x, y) = |x - y|$ defines a distance between x and y, where $|\cdot|$ designates the absolute value (or the modulus).
- For $E = \mathbb{R}^N$ (or $\mathbb{C}^N$), the following map is called Hölder's distance:

$$d_p(x, y) = \left(\sum_{n=1}^{N} |x_n - y_n|^p \right)^{1/p}, 1 \leq p \leq \infty.$$

In particular, for $p = 1, 2, \infty$, we have:

$$d_1(x, y) = \sum_{n=1}^{N} |x_n - y_n|$$

$$d_2(x, y) = \sqrt{\sum_{n=1}^{N} |x_n - y_n|^2}$$

$$d_\infty(x, y) = \max_n \{|x_n - y_n|\}.$$

Examples of distances

- In data statistical analysis, the Mahalanobis distance is very commonly used. It is a means to measure the dissimilarity between two random vectors $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^N$, having the same distribution with the covariance matrix $\boldsymbol{\Sigma}$, such as:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \mathbf{y})}.$$

  If the covariance matrix is diagonal, with diagonal elements equal to σ_n^2, for $n \in \langle N \rangle$, we have the normalized Euclidean distance:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{n=1}^{N} \frac{(x_n - y_n)^2}{\sigma_n^2}}.$$

Inequalities and equivalent distances

A distance d satisfies the generalized triangle inequality for all $x_n \in E, n \in \langle N \rangle$, and the second triangle inequality for all $x, y, z \in E$:

$$d(x_1, x_N) \leq \sum_{n=1}^{N-1} d(x_n, x_{n+1})$$

$$|d(x, z) - d(z, y)| \leq d(x, y).$$

Inequalities and equivalent distances

A distance d satisfies the generalized triangle inequality for all $x_n \in E, n \in \langle N \rangle$, and the second triangle inequality for all $x, y, z \in E$:

$$d(x_1, x_N) \leq \sum_{n=1}^{N-1} d(x_n, x_{n+1})$$

$$|d(x,z) - d(z,y)| \leq d(x,y).$$

Two distances d and δ defined over the same metric space E are said to be equivalent if there exists two positive numbers a and b such that, for all $x, y \in E$, we have:

$$a\, d(x,y) \leq \delta(x,y) \leq b\, d(x,y).$$

Distance and convergence of sequences

In a metric space (E, d), a sequence $(x_n)_{n\in\mathbb{N}}$ converges to x if and only if $d\left(x_n, x\right) \to 0$ when $n \to \infty$, or equivalently:

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, d\left(x_n, x\right) < \epsilon.$$

This amounts to saying that the sequence enters and remains in the open ball $\mathcal{B}_\epsilon(x)$.

Distance and local continuity of a function

Given two metric spaces (E, d) and (F, δ), and a map $f : E \to F$, it is said that f is continuous at x if and only if for any sequence (x_n) of E that converges to x, then the sequence $(f(x_n))$ of F converges to $f(x)$, that is, for any $\epsilon > 0$, there exists $\eta > 0$ such that:

$$\forall y \in E, d(x, y) < \eta \Rightarrow \delta(f(x), f(y)) < \epsilon.$$

Isometries and Lipschitzian maps

Let $f\colon (E,d) \to (F,\delta)$ be a map between two metric spaces.

- f is an isometry (等距) if for all $x, y \in E$, we have:

$$\delta[f(x), f(y)] = d(x,y)$$

  It is then said that f preserves distances.

Isometries and Lipschitzian maps

Let $f\colon (E, d) \to (F, \delta)$ be a map between two metric spaces.

- f is an isometry (等距) if for all $x, y \in E$, we have:

$$\delta[f(x), f(y)] = d(x, y)$$

  It is then said that f preserves distances.

- f is Lipschitzian of constant C (known as C-Lipschitzian) when:

$$\delta[f(x), f(y)] \leq Cd(x, y), \forall x, y \in E$$

  This is a regularity property stronger than continuity. f is called a contracting map or a contraction if $C \in [0, 1[$.

Definition of norm and triangle inequalities

A $\mathbb{K}$-v.s. E is said to be normed if it is equipped with a norm.

Definition of norm and triangle inequalities

A $\mathbb{K}$-v.s. E is said to be normed if it is equipped with a norm.

A norm on a v.s. E is a map $E \to \mathbb{R}^+$, denoted by $\|\cdot\|$, which to any vector $x \in E$ associates the number $\|x\|$ satisfying the following properties:

- Non-negativity: $\forall x \in E : \|x\| \geq 0$
- Strict positivity: $\|x\| = 0 \Leftrightarrow x = 0$
- Homogeneity: $\forall (\lambda, x) \in \mathbb{K} \times E : \|\lambda x\| = |\lambda| \|x\|$
- Triangle inequality: $\forall (x, y) \in E \times E : \|x + y\| \leq \|x\| + \|y\|$

where $|\cdot|$ denotes the absolute value if $\mathbb{K} = \mathbb{R}$, or the modulus if $\mathbb{K} = \mathbb{C}$. A normed v.s. (n.v.s.) E equipped with the norm $\|\cdot\|$ is denoted by $(E, \|\cdot\|)$.

Definition of norm and triangle inequalities

定理

The norm also satisfies a generalization of the triangle inequality as well as a second triangle inequality:

$$\left\|\sum_{n=1}^{N} x_n\right\| \leq \sum_{n=1}^{N} \|x_n\|, \forall (x_1, \cdots, x_N) \in E^N.$$

$$|\ \|x\| - \|y\|\ | \leq \|x \pm y\|, \forall (x, y) \in E^2.$$

Examples of norms: vector norms

For $\mathbf{x} \in \mathbb{R}^N$, the following norms can be defined: (Hölder's norm or l_p norm)

$$\|\mathbf{x}\|_p = \left(\sum_{n=1}^{N} |x_n|^p \right)^{1/p}, \quad 1 \leq p \leq \infty.$$

Examples of norms: vector norms

For $\mathbf{x} \in \mathbb{R}^N$, the following norms can be defined: (Hölder's norm or l_p norm)

$$\|\mathbf{x}\|_p = \left(\sum_{n=1}^{N} |x_n|^p \right)^{1/p}, \quad 1 \leq p \leq \infty.$$

定理 (Hölder's inequality)

Suppose (p, q) is a pair of conjugated exponents, such that $\frac{1}{p} + \frac{1}{q} = 1$. For any vectors x and $\mathbf{y}$ of $\mathbb{R}^N$, Hölder's inequality is given by:

$$\left| \sum_{n=1}^{N} x_n y_n \right| \leq \left(\sum_{n=1}^{N} |x_n|^p \right)^{1/p} \left(\sum_{n=1}^{N} |y_n|^q \right)^{1/q} = \|\mathbf{x}\|_p \|\mathbf{y}\|_q.$$

Examples of norms: vector norms

定理

For all $\mathbf{x} \in \mathbb{R}^N$, *we have the following inequalities:*

$$\|\mathbf{x}\|_p \leq \|\mathbf{x}\|_q \quad , \quad 1 \leq q \leq p$$
$$\|\mathbf{x}\|_p \leq N^{1/p}\|\mathbf{x}\|_\infty, \forall p \geq 1,$$

from which one can deduce:

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{N}\|\mathbf{x}\|_\infty;$$
$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq N\|\mathbf{x}\|_\infty;$$
$$\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{N}\|\mathbf{x}\|_2.$$

Equalities occur when $\mathbf{x}$ has at most one non-zero component. Otherwise, these are strict inequalities.

Examples of norms: infinite sequences

定理

In the space $l^p(I, \mathbb{K})$, with $1 \leq p \leq \infty$ and $I = \mathbb{Z}$ or $\mathbb{N}$, of infinite scalar sequences $(x_n)_{n \in I}$, defined and p-summable over I, with values in $\mathbb{K}$, the l_p norm is written as:

$$\|x\|_p = \left(\sum_{n \in I} |x_n|^p \right)^{1/p}$$

where the p-summability property means that $\sum_{n \in I} |x_n|^p < \infty$. This norm generalizes the Hölder norm to spaces of infinite sequences.

Examples of norms: infinite sequences

定理

The l_p norm can be generalized to the case of infinite vector sequences $(\mathbf{x}_n)_{n\in I}$ of the space $\mathbf{l}^p\left(I, \mathbb{K}^N\right)$, with $1 \leq p \leq \infty$ and $I = \mathbb{Z}$ or $\mathbb{N}$, defined and p-summable over I, with values in K^N :

$$\|\mathbf{x}\|_p = \left(\sum_{n\in I} \|\mathbf{x}_n\|^p\right)^{1/p},$$

where $\|.\|$ can be any norm in K^N. In general, the l_p norm is considered, which gives the following mixed norm:

$$\|\mathbf{x}\|_p = \left(\sum_{n\in I} \|\mathbf{x}_n\|_p^p\right)^{1/p}, 1 \leq p \leq \infty.$$

Examples of norms: functions

定理

In the space $L^p(I, \mathbb{K})$ of scalar functions f, defined and p-integrable over the interval $I \subseteq \mathbb{R}$, with values in $\mathbb{K}$, the Hölder L_p norm is defined as:

$$\|f\|_p = \left(\int_I |f(t)|^p dt \right)^{1/p}, \quad 1 \leq p \leq \infty,$$

the property of p-integrability meaning that $\int_I |f(t)|^p dt < \infty$.

Examples of norms: functions

定理

In the space $L^p(I, \mathbb{K})$ of scalar functions f, defined and p-integrable over the interval $I \subseteq \mathbb{R}$, with values in $\mathbb{K}$, the Hölder L_p norm is defined as:

$$\|f\|_p = \left(\int_I |f(t)|^p dt \right)^{1/p}, \quad 1 \leq p \leq \infty,$$

the property of p-integrability meaning that $\int_I |f(t)|^p dt < \infty$.

We have the following special cases:

- For $p = 1$, L_1 norm $\|f\|_1 = \int_I |f(t)| dt$.
- For $p = 2$, L_2 norm $\|f\|_2 = \left(\int_I |f(t)|^2 dt \right)^{1/2}$.
- For $p = \infty$, L_∞ norm $\|f\|_\infty = \sup_{t \in I} |f(t)|$.

Examples of norms: functions

定理

In the space $\mathbf{L}^p\left(I, \mathbb{K}^N\right)$ of vector functions $\mathbf{f}$, defined and p integrable over the interval I, with values in $\mathbb{K}^N$, the norm L_p is given by:

$$\|\mathbf{f}\|_p = \left(\int_I \|\mathbf{f}(t)\|^p dt\right)^{1/p},$$

where $\|\cdot\|$ can be any norm in $\mathbb{K}^N$. By considering the l_p norm in $\mathbb{K}^N$, the following composite norm is defined:

$$\|\mathbf{f}\|_p = \left(\int_I \|\mathbf{f}(t)\|_p^p dt\right)^{1/p} = \left(\int \sum_{n=1}^{N} |f_n(t)|^p \, dt\right)^{1/p},$$

where $f_n(t)$ is the nth component of $\mathbf{f}(t)$.

Equivalent norms

Two norms $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ defined over the same v.s. E are said to be equivalent if there exists two positive numbers a and b such that, for all $x \in E$, we have:

$$a\|x\|_{(1)} \leq \|x\|_{(2)} \leq b\|x\|_{(1)}.$$

In a finite-dimensional v.s., all norms are equivalent. Moreover, all norms lead to the same notion of convergence, which is not the case for infinite-dimensional v.s., for which different notions of convergence can be defined.

Distance associated with a norm

The norm is used as a measure of the length of a vector. It can also be used to define the distance between two elements x and y of E as:

$$d(x, y) = \|x - y\|.$$

Therefore, a normed $\mathbb{K}$-v.s. E is a metric space for the distance (metric) induced by the norm.

Distance associated with a norm

The norm is used as a measure of the length of a vector. It can also be used to define the distance between two elements x and y of E as:

$$d(x, y) = \|x - y\|.$$

Therefore, a normed $\mathbb{K}$-v.s. E is a metric space for the distance (metric) induced by the norm.

A metric associated with a norm satisfies the following properties:

- Translation invariance : $\quad \forall t \in E : d(x + t, y + t) = d(x, y)$
- Homogeneity : $\quad \forall \lambda \in \mathbb{K} : d(\lambda x, \lambda y) = |\lambda| d(x, y).$

Distance associated with a norm

The norm is used as a measure of the length of a vector. It can also be used to define the distance between two elements x and y of E as:

$$d(x,y) = \|x-y\|.$$

Therefore, a normed $\mathbb{K}$-v.s. E is a metric space for the distance (metric) induced by the norm.

A metric associated with a norm satisfies the following properties:

- Translation invariance : $\quad \forall t \in E : d(x+t, y+t) = d(x,y)$
- Homogeneity : $\quad \forall \lambda \in \mathbb{K} : d(\lambda x, \lambda y) = |\lambda| d(x,y).$

The closed unit ball of a n.v.s. E is defined as the set: $\overline{\mathcal{B}} = \{y \in E : \|y\| \leq 1\}$. Similarly, the closed ball of center x, with radius $r > 0$, is defined as the set $\overline{\mathcal{B}}_r(x) = \{y \in E : \|y-x\| \leq r\}$.

Pre-Hilbert spaces

Pre-Hilbert spaces, also called as inner product spaces, are real or complex (incomplete) vector spaces equipped with an inner product. When these spaces are finite-dimensional, they are called Euclidean spaces in the real case and Hermitian (or unitary) spaces in the complex case.

Pre-Hilbert spaces

Pre-Hilbert spaces, also called as inner product spaces, are real or complex (incomplete) vector spaces equipped with an inner product. When these spaces are finite-dimensional, they are called Euclidean spaces in the real case and Hermitian (or unitary) spaces in the complex case.

A complete space equipped with an inner product is called a Hilbert space. Based on inner product, we can define the important notions of induced norm and distance, and also geometrical notions such as the length of a vector, the angle and orthogonality between two of them, and thus the definition of orthonormal basis.

Real pre-Hilbert spaces: Euclidean inner product

Given a finite-dimensional real v.s. E, a Euclidean inner product over E, denoted by $\langle .,. \rangle$, refers to any bilinear form:

$$E^2 \ni (x,y) \mapsto \langle x,y \rangle \in \mathbb{R}$$

which is symmetric and positive definite, that is, satisfying the following properties:

$$\langle x,y \rangle = \langle y,x \rangle, \forall (x,y) \in E^2$$
$$\langle x,x \rangle = 0 \text{ if and only if } x = 0$$
$$\langle x,x \rangle \geq 0, \forall x \in E$$

Real pre-Hilbert spaces: Euclidean inner product

Given a finite-dimensional real v.s. E, a Euclidean inner product over E, denoted by $\langle ., . \rangle$, refers to any bilinear form:

$$E^2 \ni (x, y) \mapsto \langle x, y \rangle \in \mathbb{R}$$

which is symmetric and positive definite, that is, satisfying the following properties:

$$\begin{aligned} &\langle x, y \rangle = \langle y, x \rangle, \forall (x, y) \in E^2 \\ &\langle x, x \rangle = 0 \text{ if and only if } x = 0 \\ &\langle x, x \rangle \geq 0, \forall x \in E \end{aligned}$$

Due to the bilinearity property, the inner product is linear in both variables x and y. It is said that E equipped with this inner product is a real pre-Hilbert space, and it is usually denoted by $(E, \langle ., . \rangle)$.

Examples of Euclidian norms and inner products

In the following table, we present inner products and norms over three real pre-Hilbert spaces, where $C^0([a,b],\mathbb{R})$ and $C^0_{2\pi}(\mathbb{R},\mathbb{R})$ denote the v.s. of continuous functions over $[a,b]$, with real values, and of 2π-periodic continuous functions, from $\mathbb{R}$ to $\mathbb{R}$, respectively.

Spaces	Inner products	Norms
$\mathbb{R}^N$	$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n=1}^N x_n y_n = \mathbf{y}^T \mathbf{x}$	$\Vert \mathbf{x} \Vert_2 = \left(\sum_{n=1}^N x_n^2 \right)^{1/2}$
$C^0([a,b],\mathbb{R})$	$\langle f, g \rangle = \int_a^b f(t) g(t) dt$	$\Vert f \Vert_2 = \left(\int_a^b f^2(t) dt \right)^{1/2}$
$C^0_{2\pi}(\mathbb{R},\mathbb{R})$	$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) g(t) dt$	$\Vert f \Vert_2 = \left(\frac{1}{2\pi} \int_0^{2\pi} f^2(t) dt \right)^{1/2}$

Table 1: Examples of Euclidian norms and inner products

Complex pre-Hilbert spaces: Hermitian inner product

In the case of a finite-dimensional complex v.s. E, an Hermitian inner product over E designates any positive definite sesquilinear form:

$$E^2 \ni (x, y) \mapsto \langle x, y \rangle \in \mathbb{C}$$

which is symmetric and positive definite, that is, satisfying the following properties:

$$\langle x, y \rangle = \overline{\langle y, x \rangle}, \forall (x, y) \in E^2$$
$$\langle x, x \rangle = 0 \text{ if and only if } x = 0$$
$$\langle x, x \rangle \geq 0, \forall x \in E$$

Complex pre-Hilbert spaces: Hermitian inner product

In the case of a finite-dimensional complex v.s. E, an Hermitian inner product over E designates any positive definite sesquilinear form:

$$E^2 \ni (x, y) \mapsto \langle x, y \rangle \in \mathbb{C}$$

which is symmetric and positive definite, that is, satisfying the following properties:

$$\langle x, y \rangle = \overline{\langle y, x \rangle}, \forall (x, y) \in E^2$$
$$\langle x, x \rangle = 0 \text{ if and only if } x = 0$$
$$\langle x, x \rangle \geq 0, \forall x \in E$$

Thus,

$$\forall (x_1, x_2) \in E^2, \forall \lambda \in \mathbb{C}, \langle x_1 + \lambda x_2, y \rangle = \langle x_1, y \rangle + \lambda \langle x_2, y \rangle,$$
$$\forall (y_1, y_2) \in E^2, \forall \lambda \in \mathbb{C}, \langle x, y_1 + \lambda y_2 \rangle = \langle x, y_1 \rangle + \bar{\lambda} \langle x, y_2 \rangle.$$

Examples of Hermitian norms and inner products

In the following table, we present inner products and norms over three complex pre-Hilbert spaces, where $C^0([a,b],\mathbb{R})$ and $C^0_T(\mathbb{R},\mathbb{R})$ denote the v.s. of continuous functions over $[a,b]$, with complex values, and of T-periodic continuous functions, from $\mathbb{R}$ to $\mathbb{C}$, respectively.

Spaces	Inner products	Norms
$\mathbb{C}^N$	$\langle \mathbf{x},\mathbf{y}\rangle = \sum_{n=1}^N x_n\bar{y}_n = \mathbf{y}^H\mathbf{x}$	$\Vert\mathbf{x}\Vert_2 = \left(\sum_{n=1}^N \lvert x_n\rvert^2\right)^{1/2}$
$C^0([a,b],\mathbb{C})$	$\langle f,g\rangle = \int_a^b f(t)\bar{g}(t)dt$	$\Vert f\Vert_2 = \left(\int_a^b \lvert f(t)\rvert^2 dt\right)^{1/2}$
$C^0_T(\mathbb{R},\mathbb{C})$	$\langle f,g\rangle = \frac{1}{T}\int_0^T f(t)\bar{g}(t)dt$	$\Vert f\Vert_2 = \left(\frac{1}{T}\int_0^T \lvert f(t)\rvert^2 dt\right)^{1/2}$

Table 2: Examples of Hermitian norms and inner products

Examples of inner products

For the Euclidean and Hermitian inner products, we have, respectively:

$$\langle Ax, y\rangle = y^T Ax = \langle x, A^T y\rangle \text{ for } x, y \in \mathbb{R}^N,$$
$$\langle Ax, y\rangle = y^H Ax = \langle x, A^H y\rangle \text{ for } x, y \in \mathbb{C}^N.$$

Norm induced from an inner product

The quantity $\|x\| = \sqrt{\langle x, x\rangle}$ represents the norm of the vector x, induced from the inner product $\langle ., . \rangle$. It is also known as the norm associated with the inner product. It is called Euclidean norm in the case of a real v.s. and Hermitian norm in the case of a complex v.s.

Norm induced from an inner product

The quantity $\|x\| = \sqrt{\langle x, x \rangle}$ represents the norm of the vector x, induced from the inner product $\langle ., . \rangle$. It is also known as the norm associated with the inner product. It is called Euclidean norm in the case of a real v.s. and Hermitian norm in the case of a complex v.s.

The induced norm satisfies the following properties:

- For all $(x_1, \cdots, x_N) \in E^N$, taking into account the Hermitian symmetry property $\langle x, y \rangle = \overline{\langle y, x \rangle}$ gives:

$$
\begin{aligned}
\left\| \sum_{n=1}^{N} x_n \right\|^2 &= \sum_{n=1}^{N} \|x_n\|^2 + \sum_{1 \leq i,j \leq N, i \neq j} \langle x_i, x_j \rangle \\
&= \sum_{n=1}^{N} \|x_n\|^2 + 2 \sum_{1 \leq i < j \leq N} \mathrm{Re}\left(\langle x_i, x_j \rangle \right)
\end{aligned}
$$

Norm induced from an inner product

- For $N = 2$, the previous equality becomes:

$$\|x + y\|^2 = \|x\|^2 + 2\,\mathrm{Re}(\langle x, y\rangle) + \|y\|^2, \forall (x, y) \in E^2. \tag{1}$$

Norm induced from an inner product

- For $N = 2$, the previous equality becomes:

$$\|x+y\|^2 = \|x\|^2 + 2\,\mathrm{Re}(\langle x, y\rangle) + \|y\|^2, \forall (x,y) \in E^2. \tag{1}$$

By replacing y by $-y$, we obtain:

$$\|x-y\|^2 = \|x\|^2 - 2\,\mathrm{Re}(\langle x, y\rangle) + \|y\|^2, \forall (x,y) \in E^2, \tag{2}$$

Norm induced from an inner product

- For $N = 2$, the previous equality becomes:

$$\|x+y\|^2 = \|x\|^2 + 2\,\mathrm{Re}(\langle x, y\rangle) + \|y\|^2, \forall (x, y) \in E^2. \tag{1}$$

By replacing y by $-y$, we obtain:

$$\|x-y\|^2 = \|x\|^2 - 2\,\mathrm{Re}(\langle x, y\rangle) + \|y\|^2, \forall (x, y) \in E^2, \tag{2}$$

and by summing the last two equalities member-wise, the equality (or identity) of the parallelogram can be deduced:

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right), \forall (x, y) \in E^2. \tag{3}$$

In geometry, this equality reflects the fact that, in the plane, the sum of the squares of the diagonals of a parallelogram is equal to the sum of the squares of its sides.

Norm induced from an inner product

It characterizes the existence of an inner product based on a norm. This means that an inner product cannot be associated with a norm that does not satisfy equality $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$.

Norm induced from an inner product

It characterizes the existence of an inner product based on a norm. This means that an inner product cannot be associated with a norm that does not satisfy equality $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$.

For example, for the l_p Hölder norm and for $E = \mathbb{R}^2$, with $\mathbf{x} = [0\,1]^T$ and $\mathbf{y} = [1\,0]^T$, $\|\mathbf{x}+\mathbf{y}\|_p^2 + \|\mathbf{x}-\mathbf{y}\|_p^2 = 2 \times 2^{2/p}$ and $2\left(\|\mathbf{x}\|_p^2 + \|\mathbf{y}\|_p^2\right) = 4$. As a result, the equality of the parallelogram is only satisfied for $p = 2$, which implies that an inner product can only be associated with the l_p Hölder norm, for $p = 2$.

Cauchy-Schwarz inequality

- For all $(x, y) \in E^2$, the Cauchy-Schwarz inequality is written as:

$$|\langle x, y\rangle| \leq \|x\| \, \|y\|$$

  with equality when x and y are collinear (i.e. when there exists $\lambda \in \mathbb{K}$ such that $y = \lambda x$), or if x or y is zero.

- In $\mathbb{C}^N$ and $C^0([a, b], \mathbb{C})$, the Cauchy-Schwarz inequality becomes:

$$\left| \sum_{n=1}^{N} x_n \bar{y}_n \right|^2 \leq \sum_{n=1}^{N} |x_n|^2 \sum_{n=1}^{N} |y_n|^2 ,$$

$$\left| \int_a^b f(t) \bar{g}(t) \, dt \right|^2 \leq \int_a^b |f(t)|^2 dt \int_a^b |g(t)|^2 dt.$$

Cauchy-Schwarz inequality

In $E = \mathbb{R}^N$, the Cauchy-Schwarz inequality can be interpreted in terms of Euclidean geometry with:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \|\mathbf{y}\| \cos\theta \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

where θ is the angle between vectors $\mathbf{x}$ and $\mathbf{y}$. When $\theta = \frac{\pi}{2}$, we have $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ which reflects the perpendicularity of vectors $\mathbf{x}$ and $\mathbf{y}$.

Minkowski inequality

The Minkowski inequality is written as, for all $(x, y) \in E^2$:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p, \quad 1 \leq p < \infty$$

Equality holds if x or y is the null vector. The Minkowski inequality can be seen as the triangle inequality for the l_p norm.

Minkowski inequality

The Minkowski inequality is written as, for all $(x, y) \in E^2$:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p, \quad 1 \leq p < \infty$$

Equality holds if x or y is the null vector. The Minkowski inequality can be seen as the triangle inequality for the l_p norm.

- For $E = \mathbb{C}^N$, and the norm l_p, the Minkowski inequality is written as:

$$\left(\sum_{n=1}^{N} |x_n + y_n|^p\right)^{1/p} \leq \left(\sum_{n=1}^{N} |x_n|^p\right)^{1/p} + \left(\sum_{n=1}^{N} |y_n|^p\right)^{1/p}.$$

Minkowski inequality

The Minkowski inequality is written as, for all $(x, y) \in E^2$:

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p, \quad 1 \leq p < \infty$$

Equality holds if x or y is the null vector. The Minkowski inequality can be seen as the triangle inequality for the l_p norm.

- For $E = \mathbb{C}^N$, and the norm l_p, the Minkowski inequality is written as:

$$\left(\sum_{n=1}^{N} |x_n + y_n|^p\right)^{1/p} \leq \left(\sum_{n=1}^{N} |x_n|^p\right)^{1/p} + \left(\sum_{n=1}^{N} |y_n|^p\right)^{1/p}.$$

- For two functions f and g of the space $L^p(I, \mathbb{K})$, with $I \subseteq \mathbb{R}$ and the norm L_p, the Minkowski inequality is written as:

$$\left(\int_I |f(t) + g(t)|^p dt\right)^{1/p} \leq \left(\int_I |f(t)|^p dt\right)^{1/p} + \left(\int_I |g(t)|^p dt\right)^{1/p}.$$

Polarization formulae

Since $\langle x, iy\rangle = -i\langle x, y\rangle$, with $i^2 = -1$, we have:

$$\|x \pm iy\|^2 = \|x\|^2 \pm 2\,\mathrm{Im}(\langle x, y\rangle) + \|y\|^2, \forall (x, y) \in E^2. \tag{4}$$

Polarization formulae

Since $\langle x, iy \rangle = -i\langle x, y \rangle$, with $i^2 = -1$, we have:

$$\|x \pm iy\|^2 = \|x\|^2 \pm 2\,\mathrm{Im}(\langle x, y \rangle) + \|y\|^2, \forall (x, y) \in E^2. \tag{4}$$

From relations (1)-(4), the following polarization formulae can be deduced:

$$\mathrm{Re}(\langle x, y \rangle) = \frac{\|x + y\|^2 - \|x - y\|^2}{4}, \mathrm{Im}(\langle x, y \rangle) = \frac{\|x + iy\|^2 - \|x - iy\|^2}{4}.$$

Polarization formulae

Since $\langle x, iy\rangle = -i\langle x, y\rangle$, with $i^2 = -1$, we have:

$$\|x \pm iy\|^2 = \|x\|^2 \pm 2\,\mathrm{Im}(\langle x, y\rangle) + \|y\|^2, \forall (x, y) \in E^2. \tag{4}$$

From relations (1)-(4), the following polarization formulae can be deduced:

$$\mathrm{Re}(\langle x, y\rangle) = \frac{\|x+y\|^2 - \|x-y\|^2}{4}, \mathrm{Im}(\langle x, y\rangle) = \frac{\|x+iy\|^2 - \|x-iy\|^2}{4}.$$

These formulae allow us to define the Hermitian inner product based on the norm:

$$\langle x, y\rangle = \frac{\|x+y\|^2 - \|x-y\|^2}{4} + i\frac{\|x+iy\|^2 - \|x-iy\|^2}{4}. \tag{5}$$

Polarization formulae

Since $\langle x, iy\rangle = -i\langle x, y\rangle$, with $i^2 = -1$, we have:

$$\|x \pm iy\|^2 = \|x\|^2 \pm 2\,\mathrm{Im}(\langle x, y\rangle) + \|y\|^2, \forall (x, y) \in E^2. \tag{4}$$

From relations (1)-(4), the following polarization formulae can be deduced:

$$\mathrm{Re}(\langle x, y\rangle) = \frac{\|x+y\|^2 - \|x-y\|^2}{4}, \mathrm{Im}(\langle x, y\rangle) = \frac{\|x+iy\|^2 - \|x-iy\|^2}{4}.$$

These formulae allow us to define the Hermitian inner product based on the norm:

$$\langle x, y\rangle = \frac{\|x+y\|^2 - \|x-y\|^2}{4} + i\frac{\|x+iy\|^2 - \|x-iy\|^2}{4}. \tag{5}$$

In the case of a real v.s. E, the aforementioned formula, reduced to Euclidean inner product, can be simplified as:

$$\langle x, y\rangle = \frac{\|x+y\|^2 - \|x-y\|^2}{4}, \quad \text{or} \quad \langle x, y\rangle = \frac{\|x+y\|^2 - \|x\|^2 - \|y\|^2}{2}.$$

Distance associated with an inner product

A distance can be associated with any inner product according to the following formula:

$$d(x, y) = \sqrt{\langle x - y, x - y \rangle} = \|x - y\|, \quad \forall x, y \in E,$$

where $\|x - y\|$ is the norm of $x - y$ induced from the inner product. Subsequently, any pre-Hilbert space is a metric space for the distance induced from the inner product.

Weighted inner products

For example, in the spaces $\mathbb{C}^N$ and $C^0([a, b], \mathbb{C})$, weighted inner products can be defined as:

$$\langle \mathbf{x}, \mathbf{y} \rangle_\rho = \sum_{n=1}^{N} \rho_n x_n \bar{y}_n$$

$$\langle f, g \rangle_\rho = \int_a^b \rho(t) f(t) \bar{g}(t) dt$$

where ρ_n and $\rho(t)$ are positive weightings.

Summary

Properties	Relations
Triangle inequality	$\|x+y\| \leq \|x\| + \|y\|$
Other inequality	$\lvert \|x\| - \|y\| \rvert \leq \|x \pm y\|$
Hermitian symmetry	$\langle x, y\rangle = \overline{\langle y, x\rangle}$
Induced norm	$\|x\| = \sqrt{\langle x, x\rangle}$
Cauchy-Schwarz inequality	$\lvert\langle x, y\rangle\rvert \leq \|x\| \, \|y\|$
Norm of $x \pm y$	$\|x \pm y\|^2 = \|x\|^2 \pm 2\,\mathrm{Re}(\langle x, y\rangle) + \|y\|^2$
Parallelogram equality	$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$
Norm of $x \pm iy$	$\|x \pm iy\|^2 = \|x\|^2 \pm 2\,\mathrm{Im}(\langle x, y\rangle) + \|y\|^2$
Polarization identity	$\langle x, y\rangle = \frac{\|x+y\|^2 - \|x-y\|^2}{4} + i\frac{\|x+iy\|^2 - \|x-iy\|^2}{4}$

Table 1. Properties of Hermitian norms and inner products.

Orthogonal/perpendicular vectors

- Vectors $x, y \in E$, assumed to be non-null, are said to be orthogonal, and denoted by $x \perp y$, if their inner product is zero:

$$x \perp y \Leftrightarrow \langle x, y \rangle = 0 \Leftrightarrow \mathrm{Re}(\langle x, y \rangle) = \mathrm{Im}(\langle x, y \rangle) = 0.$$

Orthogonal/perpendicular vectors

- Vectors $x, y \in E$, assumed to be non-null, are said to be orthogonal, and denoted by $x \perp y$, if their inner product is zero:

$$x \perp y \Leftrightarrow \langle x, y \rangle = 0 \Leftrightarrow \mathrm{Re}(\langle x, y \rangle) = \mathrm{Im}(\langle x, y \rangle) = 0.$$

- Vectors $x, y \in E$ are said to be perpendicular when only the real part of the inner product is zero, that is $\mathrm{Re}(\langle x, y \rangle) = 0$.

Orthogonal/perpendicular vectors

- Vectors $x, y \in E$, assumed to be non-null, are said to be orthogonal, and denoted by $x \perp y$, if their inner product is zero:

$$x \perp y \Leftrightarrow \langle x, y \rangle = 0 \Leftrightarrow \mathrm{Re}(\langle x, y \rangle) = \mathrm{Im}(\langle x, y \rangle) = 0.$$

- Vectors $x, y \in E$ are said to be perpendicular when only the real part of the inner product is zero, that is $\mathrm{Re}(\langle x, y \rangle) = 0$.

In $\mathbb{R}^N$ and $\mathbb{C}^N$, we have:

$$\mathbf{x}, \mathbf{y} \in \mathbb{R}^N, \mathbf{x} \perp \mathbf{y} \Leftrightarrow \sum_{n=1}^{N} x_n y_n = \mathbf{y}^T \mathbf{x} = 0,$$

$$\mathbf{x}, \mathbf{y} \in \mathbb{C}^N, \mathbf{x} \perp \mathbf{y} \Leftrightarrow \sum_{n=1}^{N} x_n \bar{y}_n = \mathbf{y}^H \mathbf{x} = 0,$$

where $\mathbf{y}^H = \overline{\mathbf{y}}^T = [\bar{y}_1, \bar{y}_2, \cdots \bar{y}_N]^T$ is the transconjugated vector of $\mathbf{y}$.

Pythagorean theorem

It is important to note that the notions of orthogonality and perpendicularity are equivalent in the case of real pre-Hilbert spaces since then we have $\mathrm{Re}(\langle x, y \rangle) = \langle x, y \rangle$, and $\mathrm{Re}(\langle x, y \rangle)$ equal to zero implies that the inner product is also equal to zero,

Pythagorean theorem

It is important to note that the notions of orthogonality and perpendicularity are equivalent in the case of real pre-Hilbert spaces since then we have $\mathrm{Re}(\langle x, y\rangle) = \langle x, y\rangle$, and $\mathrm{Re}(\langle x, y\rangle)$ equal to zero implies that the inner product is also equal to zero,whereas in the case of complex pre-Hilbert spaces, this equality to zero implies that $\langle x, y\rangle = i\,\mathrm{Im}(\langle x, y\rangle)$ which may be non-zero.

Pythagorean theorem

It is important to note that the notions of orthogonality and perpendicularity are equivalent in the case of real pre-Hilbert spaces since then we have $\mathrm{Re}(\langle x, y\rangle) = \langle x, y\rangle$, and $\mathrm{Re}(\langle x, y\rangle)$ equal to zero implies that the inner product is also equal to zero,whereas in the case of complex pre-Hilbert spaces, this equality to zero implies that $\langle x, y\rangle = i\,\mathrm{Im}(\langle x, y\rangle)$ which may be non-zero.Based on Eq. (1), when $\mathrm{Re}(\langle x, y\rangle) = 0$, we obtain the Pythagorean theorem:

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

Pythagorean theorem

It is important to note that the notions of orthogonality and perpendicularity are equivalent in the case of real pre-Hilbert spaces since then we have $\mathrm{Re}(\langle x, y\rangle) = \langle x, y\rangle$, and $\mathrm{Re}(\langle x, y\rangle)$ equal to zero implies that the inner product is also equal to zero,whereas in the case of complex pre-Hilbert spaces, this equality to zero implies that $\langle x, y\rangle = i\,\mathrm{Im}(\langle x, y\rangle)$ which may be non-zero.Based on Eq. (1), when $\mathrm{Re}(\langle x, y\rangle) = 0$, we obtain the Pythagorean theorem:

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

Conversely, in the case of a complex pre-Hilbert space, when the Pythagorean theorem is verified, this implies that x and y are perpendicular, but not that they are orthogonal.

Pythagorean theorem

It is important to note that the notions of orthogonality and perpendicularity are equivalent in the case of real pre-Hilbert spaces since then we have $\mathrm{Re}(\langle x, y\rangle) = \langle x, y\rangle$, and $\mathrm{Re}(\langle x, y\rangle)$ equal to zero implies that the inner product is also equal to zero,whereas in the case of complex pre-Hilbert spaces, this equality to zero implies that $\langle x, y\rangle = i\,\mathrm{Im}(\langle x, y\rangle)$ which may be non-zero.Based on Eq. (1), when $\mathrm{Re}(\langle x, y\rangle) = 0$, we obtain the Pythagorean theorem:

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

Conversely, in the case of a complex pre-Hilbert space, when the Pythagorean theorem is verified, this implies that x and y are perpendicular, but not that they are orthogonal.

It can be concluded that:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 \Leftrightarrow x \text{ and } y \text{ perpendicular}$$

which corresponds to the Pythagorean theorem in geometry.

Orthogonal subspaces and orthogonal complement

A few definitions related to the notion of orthogonality:

- Let F denote a subspace of E. The vector $x \in E$ is orthogonal to F if it is orthogonal to any vector of F.

Orthogonal subspaces and orthogonal complement

A few definitions related to the notion of orthogonality:

- Let F denote a subspace of E. The vector $x \in E$ is orthogonal to F if it is orthogonal to any vector of F.
- A family $x = \{x_1, \cdots, x_N\}$ of non-zero vectors is orthogonal if and only if $\langle x_i, x_j \rangle = 0$ for all $i, j \in \langle N \rangle$, with $i \neq j$.

Orthogonal subspaces and orthogonal complement

A few definitions related to the notion of orthogonality:

- Let F denote a subspace of E. The vector $x \in E$ is orthogonal to F if it is orthogonal to any vector of F.
- A family $x = \{x_1, \cdots, x_N\}$ of non-zero vectors is orthogonal if and only if $\langle x_i, x_j \rangle = 0$ for all $i, j \in \langle N \rangle$, with $i \neq j$.
- When vectors are mutually perpendicular, such a family of vectors satisfies the generalized Pythagorean theorem:

$$\left\| \sum_{n=1}^{N} x_n \right\|^2 = \sum_{n=1}^{N} \|x_n\|^2 .$$

Orthogonal subspaces and orthogonal complement

A few definitions related to the notion of orthogonality:

- Let F denote a subspace of E. The vector $x \in E$ is orthogonal to F if it is orthogonal to any vector of F.
- A family $x = \{x_1, \cdots, x_N\}$ of non-zero vectors is orthogonal if and only if $\langle x_i, x_j \rangle = 0$ for all $i, j \in \langle N \rangle$, with $i \neq j$.
- When vectors are mutually perpendicular, such a family of vectors satisfies the generalized Pythagorean theorem:

$$\left\| \sum_{n=1}^{N} x_n \right\|^2 = \sum_{n=1}^{N} \|x_n\|^2 .$$

- Two subspaces F and G of E are said to be orthogonal, denoted by $F \perp G$, $\Leftrightarrow \forall x \in F, \forall y \in G, x \perp y$. Subsequently, any basis of F is orthogonal to any basis of G.

Orthogonal subspaces and orthogonal complement

A few definitions related to the notion of orthogonality:

- Let F denote a subspace of E. The vector $x \in E$ is orthogonal to F if it is orthogonal to any vector of F.
- A family $x = \{x_1, \cdots, x_N\}$ of non-zero vectors is orthogonal if and only if $\langle x_i, x_j \rangle = 0$ for all $i, j \in \langle N \rangle$, with $i \neq j$.
- When vectors are mutually perpendicular, such a family of vectors satisfies the generalized Pythagorean theorem:

$$\left\| \sum_{n=1}^{N} x_n \right\|^2 = \sum_{n=1}^{N} \|x_n\|^2 .$$

- Two subspaces F and G of E are said to be orthogonal, denoted by $F \perp G, \Leftrightarrow \forall x \in F, \forall y \in G, x \perp y$. Subsequently, any basis of F is orthogonal to any basis of G.
- Given a subspace F of a finite-dimensional Euclidean v.s. E, its orthogonal space, denoted by $F^\perp$, is defined as:

$$F^\perp = \{x \in E : \forall y \in F, \langle x, y \rangle = 0\}.$$

Summary

Membership	Properties	Definitions
$x, y \in E$	Perpendicularity	$\mathrm{Re}(\langle x, y\rangle) = 0$
$x, y \in E$	Pythagorean theorem	$\|x+y\|^2 = \|x\|^2 + \|y\|^2$
$x, y \in E$	Orthogonality $(x \perp y)$	$\langle x, y\rangle = 0$
$x \in E, F \subset E$	$x \perp F$	$\langle x, y\rangle = 0, \forall y \in F$
$F, G \subset E$	$F \perp G$	$\langle x, y\rangle = 0, \forall x \in F, \forall y \in G$
$F \subset E$	$F^\perp$	$F^\perp = \{x \in E : \forall y \in F, \langle x, y\rangle = 0\}$

Table 2. Perpendicularity and orthogonality properties.

Orthogonal complement

The set $F^{\perp}$ is a subspace of E orthogonal to F. It is called the orthogonal complement of F, and it is said that $E = F \oplus F^{\perp}$ is the orthogonal direct sum of F and $F^{\perp}$, with $\dim\left(F^{\perp}\right) = \dim(E) - \dim(F)$.

We have the following properties:

- $\left(F^{\perp}\right)^{\perp} = F$.
- If G is a subset of the subspace F, then: $G \subset F \Leftrightarrow F^{\perp} \subset G^{\perp}$.

Orthogonal complement

The set $F^\perp$ is a subspace of E orthogonal to F. It is called the orthogonal complement of F, and it is said that $E = F \oplus F^\perp$ is the orthogonal direct sum of F and $F^\perp$, with $\dim\left(F^\perp\right) = \dim(E) - \dim(F)$.
We have the following properties:

- $\left(F^\perp\right)^\perp = F$.
- If G is a subset of the subspace F, then: $G \subset F \Leftrightarrow F^\perp \subset G^\perp$.

These properties do not always hold if E is infinite-dimensional.

Orthogonal complement

The set $F^{\perp}$ is a subspace of E orthogonal to F. It is called the orthogonal complement of F, and it is said that $E = F \oplus F^{\perp}$ is the orthogonal direct sum of F and $F^{\perp}$, with $\dim\left(F^{\perp}\right) = \dim(E) - \dim(F)$.
We have the following properties:

- $\left(F^{\perp}\right)^{\perp} = F$.
- If G is a subset of the subspace F, then: $G \subset F \Leftrightarrow F^{\perp} \subset G^{\perp}$.

These properties do not always hold if E is infinite-dimensional.

More generally, given N subspaces $F_1, \cdots, F_N$ of an inner product space E, it is said that E is the orthogonal direct sum of $F_1, \cdots, F_N$ if E is a direct sum $\left(E = \left(\overset{\oplus}{\underset{n=1}{N}} F_n\right)\right.$ of the subspaces $F_n, n \in \langle N\rangle$, which are pairwise orthogonal.

Orthonormal bases

For an N-dimensional Euclidean/Hermitian space E, a basis $\{e_n, n \in \langle N \rangle\}$ is said to be orthonormal if $\langle e_n, e_p \rangle = \delta_{np}, \forall n, p \in \langle N \rangle$, where δ_{np} is the Kronecker delta. This means that all its vectors are pairwise orthogonal ($\langle e_n, e_p \rangle = 0, \forall n \neq p$), and unitary ($\|e_n\| = 1, \forall n \in \langle N \rangle$).

Orthonormal bases

For an N-dimensional Euclidean/Hermitian space E, a basis $\{e_n, n \in \langle N \rangle\}$ is said to be orthonormal if $\langle e_n, e_p \rangle = \delta_{np}, \forall n, p \in \langle N \rangle$, where δ_{np} is the Kronecker delta. This means that all its vectors are pairwise orthogonal $(\langle e_n, e_p \rangle = 0, \forall n \neq p)$, and unitary $(\|e_n\| = 1, \forall n \in \langle N \rangle)$.

If the vectors of a basis $\{b_n, n \in \langle N \rangle\}$ are orthogonal but non-unitary, the basis is said to be orthogonal, and then $\left\{\frac{b_n}{\|b_n\|}, n \in \langle N \rangle\right\}$ is an orthonormal basis.

Orthogonal/unitary endomorphisms and isometries

Let E be a finite-dimensional Euclidean v.s., and $f \in \mathcal{L}(E)$. It is said that f is an orthogonal endomorphism if one of the following two equivalent conditions is satisfied:

$$\begin{gathered} \forall (x, y) \in E^2, \langle f(x), f(y) \rangle = \langle x, y \rangle, \\ \forall x \in E, \|f(x)\| = \|x\|. \end{gathered} \tag{6}$$

It is then said that f preserves the inner product and the norm. Such orthogonal endomorphism is also called an isometry or an orthogonal transformation.

Orthogonal/unitary endomorphisms and isometries

Let E be a finite-dimensional Euclidean v.s., and $f \in \mathcal{L}(E)$. It is said that f is an orthogonal endomorphism if one of the following two equivalent conditions is satisfied:

$$\begin{gathered} \forall (x, y) \in E^2, \langle f(x), f(y) \rangle = \langle x, y \rangle, \\ \forall x \in E, \|f(x)\| = \|x\|. \end{gathered} \tag{6}$$

It is then said that f preserves the inner product and the norm. Such orthogonal endomorphism is also called an isometry or an orthogonal transformation.

In the case where E is a Hermitian space, conditions (6) define a unitary endomorphism.

Orthogonal/unitary endomorphisms and isometries

- The set of orthogonal endomorphisms of E is a group (endowed with the law of map composition) called the orthogonal group of E, and denoted by $O(E)$. Similarly, in the case where E is a Hermitian space, the set of unitary endomorphisms of E is called the unitary group of E, and denoted by $\mathcal{U}(E)$.

Orthogonal/unitary endomorphisms and isometries

- The set of orthogonal endomorphisms of E is a group (endowed with the law of map composition) called the orthogonal group of E, and denoted by $O(E)$. Similarly, in the case where E is a Hermitian space, the set of unitary endomorphisms of E is called the unitary group of E, and denoted by $\mathcal{U}(E)$.
- Orthogonal endomorphisms satisfy the following properties:
  - The matrix $\mathbf{A}$ associated with an orthogonal endomorphism is orthogonal, that is, it satisfies the property: $\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}$.
  - An orthogonal endomorphism transforms an orthonormal basis into an orthonormal basis.
  - If f and g are two orthogonal endomorphisms, then their composite $g \circ f$ is itself an orthogonal endomorphism. This infers that the product of two orthogonal matrices is also an orthogonal matrix.

Orthogonal projection onto a subspace

Let F be a subspace of a finite-dimensional Euclidean v.s. E. An orthogonal projection onto F, denoted by p_F, is the projection onto F parallel to $F^\perp$. For any $x \in E$, $p_F(x)$ is called the orthogonal projection of x onto F and is characterized by:

$$p_F(x) \in F\ ,\ x - p_F(x) \in F^\perp.$$

Orthogonal projection and Fourier expansion

Given a basis $\mathcal{B} = \{b_1, \cdots, b_N\}$ of F, the orthogonal projection $p_F(x)$ is expressed as $p_F(x) = \sum_{n=1}^{N} \lambda_n b_n$. The coefficients λ_n are determined such that $x - p_F(x) \in F^\perp$, and thus $x - p_F(x) \perp b_n$, for $n \in \langle N \rangle$, which amounts to solving the linear system of equations:

$$\langle x - p_F(x), b_n \rangle = 0, n \in \langle N \rangle.$$

Orthogonal projection and Fourier expansion

Given a basis $\mathcal{B} = \{b_1, \cdots, b_N\}$ of F, the orthogonal projection $p_F(x)$ is expressed as $p_F(x) = \sum_{n=1}^{N} \lambda_n b_n$. The coefficients λ_n are determined such that $x - p_F(x) \in F^{\perp}$, and thus $x - p_F(x) \perp b_n$, for $n \in \langle N \rangle$, which amounts to solving the linear system of equations:

$$\langle x - p_F(x), b_n \rangle = 0, n \in \langle N \rangle.$$

When the basis $\{b_1, \cdots, b_N\}$ is orthogonal, then $\lambda_n = \frac{\langle x, b_n \rangle}{\|b_n\|^2}$. The orthogonal projection is then written as:

$$p_F(x) = \sum_{n=1}^{N} \frac{\langle x, b_n \rangle}{\|b_n\|^2} b_n.$$

Orthogonal projection and Fourier expansion

Given a basis $\mathcal{B} = \{b_1, \cdots, b_N\}$ of F, the orthogonal projection $p_F(x)$ is expressed as $p_F(x) = \sum_{n=1}^{N} \lambda_n b_n$. The coefficients λ_n are determined such that $x - p_F(x) \in F^\perp$, and thus $x - p_F(x) \perp b_n$, for $n \in \langle N \rangle$, which amounts to solving the linear system of equations:

$$\langle x - p_F(x), b_n \rangle = 0, n \in \langle N \rangle.$$

When the basis $\{b_1, \cdots, b_N\}$ is orthogonal, then $\lambda_n = \frac{\langle x, b_n \rangle}{\|b_n\|^2}$. The orthogonal projection is then written as:

$$p_F(x) = \sum_{n=1}^{N} \frac{\langle x, b_n \rangle}{\|b_n\|^2} b_n.$$

If the basis is orthonormal, the projector can be simplified as:

$$p_F(x) = \sum_{n=1}^{N} \langle x, b_n \rangle \, b_n = \sum_{n=1}^{N} c_n b_n, \text{ where } c_n = \langle x, b_n \rangle \,. \quad (7)$$

Orthogonal projection and Fourier expansion

The term $c_n b_n$ represents the orthogonal projection of x onto the space generated by b_n. The expansion $p_F(x)$, which expresses x in the form of a sum of N mutually orthogonal vectors, is called the Fourier expansion of x. The scalars c_n are the coordinates of x in the basis $\mathcal{B}$ and are called the Fourier coefficients.

Orthogonal projection and Fourier expansion

The term $c_n b_n$ represents the orthogonal projection of x onto the space generated by b_n. The expansion $p_F(x)$, which expresses x in the form of a sum of N mutually orthogonal vectors, is called the Fourier expansion of x. The scalars c_n are the coordinates of x in the basis $\mathcal{B}$ and are called the Fourier coefficients.

定理

The squared projection error, or more specifically, the square of the norm of the difference between x and its orthogonal projection onto F, is given by:

$$\|x - p_F(x)\|^2 = \|x\|^2 - \|p_F(x)\|^2 = \|x\|^2 - \sum_{n=1}^{N} |c_n|^2 . \tag{8}$$

Bessel's inequality and Parseval's equality

定理

Bessel's inequality and Parseval's equality are given by:

$$\sum_{n=1}^{N} |c_n|^2 \leq \|x\|^2; \quad \sum_{n=1}^{N} |c_n|^2 = \|x\|^2 \text{ when } x \in F.$$

Bessel's inequality and Parseval's equality

定理

Bessel's inequality and Parseval's equality are given by:

$$\sum_{n=1}^{N} |c_n|^2 \leq \|x\|^2; \quad \sum_{n=1}^{N} |c_n|^2 = \|x\|^2 \text{ when } x \in F.$$

证明.

The inequality directly results from Eq. (8)

$$\sum_{n=1}^{N} |c_n|^2 = \|x\|^2 - \|x - p_F(x)\|^2 \leq \|x\|^2.$$

Bessel's inequality and Parseval's equality

定理

Bessel's inequality and Parseval's equality are given by:

$$\sum_{n=1}^{N} |c_n|^2 \leq \|x\|^2; \quad \sum_{n=1}^{N} |c_n|^2 = \|x\|^2 \text{ when } x \in F.$$

证明.

The inequality directly results from Eq. (8)

$$\sum_{n=1}^{N} |c_n|^2 = \|x\|^2 - \|x - p_F(x)\|^2 \leq \|x\|^2.$$

The equality is also obtained from Eq. (8) when $x \in F$, which implies that $x = p_F(x)$, and so $x - p_F(x) = 0$. □

Some observations

- Since $p_F(x) \in F$, we have $p_F(p_F(x)) = p_F(x)$, which characterizes a projector, and the idempotence (幂等) property.

Some observations

- Since $p_F(x) \in F$, we have $p_F(p_F(x)) = p_F(x)$, which characterizes a projector, and the idempotence (幂等) property.
- In the context of the theory of estimation in the least squares sense, $p_F(x)$ represents the estimate $\widehat{x}_{MC}$ of x using the data of the subspace F and $\|x - \widehat{x}_{MC}\|^2 = \|x - p_F(x)\|^2$ is the quadratic error of estimation.

Some observations

- Since $p_F(x) \in F$, we have $p_F(p_F(x)) = p_F(x)$, which characterizes a projector, and the idempotence (幂等) property.
- In the context of the theory of estimation in the least squares sense, $p_F(x)$ represents the estimate $\widehat{x}_{MC}$ of x using the data of the subspace F and $\|x - \widehat{x}_{MC}\|^2 = \|x - p_F(x)\|^2$ is the quadratic error of estimation.
- From a geometric point of view, the orthogonal projection of x onto F minimizes the following criterion:

$$\min_{y \in F} d^2(x, y) = \min_{y \in F} \|x - y\|_2^2.$$

Some observations

- Since $p_F(x) \in F$, we have $p_F(p_F(x)) = p_F(x)$, which characterizes a projector, and the idempotence (幂等) property.
- In the context of the theory of estimation in the least squares sense, $p_F(x)$ represents the estimate $\widehat{x}_{MC}$ of x using the data of the subspace F and $\|x - \widehat{x}_{MC}\|^2 = \|x - p_F(x)\|^2$ is the quadratic error of estimation.
- From a geometric point of view, the orthogonal projection of x onto F minimizes the following criterion:

$$\min_{y \in F} d^2(x, y) = \min_{y \in F} \|x - y\|_2^2.$$

- From the point of view of the theory of function approximation, equality (7) can be extended to the expansion of a function $f(t)$ over a basis of orthonormal functions $\{e_n(t), n \in \mathbb{Z}\}$ in the following form:

$$f(t) \simeq \sum_{n \in \mathbb{Z}} c_n e_n(t).$$

Summary

Properties	Relations
Orthogonal projection	$p_{\mathcal{B}}(x) = \sum_{n=1}^{N} \langle x, b_n \rangle b_n = \sum_{n=1}^{N} c_n b_n$
Bessel's inequality	$\sum_{n=1}^{N} \lvert c_n \rvert^2 \leq \lVert x \rVert^2$
Parseval's equality	$\sum_{n=1}^{N} \lvert c_n \rvert^2 = \lVert x \rVert^2, x \in F$

Table 3. Orthogonal projection of $x \in E$ onto the orthonormal basis $\mathcal{B} = \{b_1, \cdots, b_N\}$ of $F \subset E$.

Gram–Schmidt orthonormalization process

This method, named after the Danish and German mathematicians Jörgen Pedersen Gram (1850-1916) and Erhard Schmidt (1876-1959), was in fact already known to the French mathematician Pierre-Simon Laplace (1749-1827) in 1816.

Åke Björck
Gram-Schmidt orthogonalization: 100 years and more, 2010.

Steven J. Leon & Åke Björck & Walter Gander
Gram-Schmidt orthogonalization: 100 years and more, Numer. Linear Algebra Appl., 20:492—532, 2013.

Orthonormalization of a set of functions

Consider the inner product of two functions f and $g \in C^0([a,b],\mathbb{R})$ defined as:

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

The Gram-Schmidt algorithm is a method used to build an orthonormal set of functions $\{f_n\}$, in the interval $[a,b]$, from a set of linearly independent functions $\{g_n, n \in [0, N-1]\}$. This orthonormalization algorithm is employed, in particular, to build bases of orthogonal polynomials in the interval $[a,b]$, from the set of linearly independent polynomials $\left\{1, t, t^2, t^3, \cdots, t^n, \cdots\right\}$, using a weighted inner product:

$$\langle p_n, p_k\rangle = \int_a^b w(t)p_n(t)p_k(t)dt$$

where p_n is a polynomial of degree n, with $w(t) > 0$.

Orthonormalization of a set of functions

As we will see later, different systems of orthogonal polynomials can be obtained depending on the choice of the interval $[a, b]$ and of the weighting $w(t)$. The interval $[a, b]$ may be infinite at one or both ends. In this case, we should have $\lim_{t \to \pm\infty} w(t) = 0$ as necessary (but not sufficient) condition of existence of the integral $\int_{-\infty}^{\infty} \left(\text{ or } \int_0^{\infty} \right) w(t) p_n^2(t) dt = \|p_n\|^2$.

Orthonormalization of a set of functions

As we will see later, different systems of orthogonal polynomials can be obtained depending on the choice of the interval $[a, b]$ and of the weighting $w(t)$. The interval $[a, b]$ may be infinite at one or both ends. In this case, we should have $\lim_{t \to \pm\infty} w(t) = 0$ as necessary (but not sufficient) condition of existence of the integral $\int_{-\infty}^{\infty}$ (or $\int_{0}^{\infty}$) $w(t) p_n^2(t) dt = \|p_n\|^2$.

The GS algorithm without normalization of polynomials can be written as:

$$p_n(t) = t^n - \sum_{k=0}^{n-1} \frac{\langle p_k, t^n \rangle}{\langle p_k, p_k \rangle} p_k(t), \;\; p_0(t) = 1.$$

The orthogonal polynomials thus determined are unique up to a multiplicative constant that can be chosen to have unitary $(\|p_n\| = 1)$ or monic orthogonal polynomials, such that the coefficient of the term of degree n of p_n is equal to 1.

An example

Consider the inner product $\langle f, g\rangle = \frac{1}{2}\int_{-1}^{+1} f(t)g(t)dt$ used for constructing Legendre polynomials. Application of the GS algorithm, with $p_0(t) = 1$ and $\|p_0\| = 1$, gives:

$$p_1(t) = t - \langle p_0, t\rangle\, p_0(t) = t - \frac{1}{2}\int_{-1}^{+1} tdt = t$$

$$\|p_1\|^2 = \frac{1}{2}\int_{-1}^{+1} t^2 dt = \frac{1}{3}$$

$$p_2(t) = t^2 - \left\langle p_0, t^2\right\rangle p_0(t) - 3\left\langle p_1, t^2\right\rangle p_1(t) = t^2 - \frac{1}{2}\int_{-1}^{+1} t^2 dt - \frac{3t}{2}\int_{-1}^{+1} t^3 d$$

$$= t^2 - \frac{1}{3}$$

$$\|p_2\|^2 = \frac{1}{2}\int_{-1}^{+1}\left(t^2 - \frac{1}{3}\right)^2 dt = \frac{4}{45}$$

$$p_3(t) = t^3 - \left\langle p_0, t^3\right\rangle p_0(t) - 3\left\langle p_1, t^3\right\rangle p_1(t) - \frac{45}{4}\left\langle p_2, t^3\right\rangle p_2(t) = t^3 - \frac{3t}{5}.$$

Convergence & Cauchy sequence

A unilateral sequence is a function $f\colon \mathbb{N}^* \to \mathbb{C}, \mathbb{N}^* \ni n \mapsto x_n \in \mathbb{C}$, written as $(x_n)_{n\in\mathbb{N}^*}$, or $(x_n)_{n\geq 1}$, or simply (x_n). In the case of a bilateral sequence, the domain of f is $\mathbb{Z}$ instead of $\mathbb{N}^*$.

Convergence & Cauchy sequence

A unilateral sequence is a function $f\colon \mathbb{N}^* \to \mathbb{C}, \mathbb{N}^* \ni n \mapsto x_n \in \mathbb{C}$, written as $(x_n)_{n\in\mathbb{N}^*}$, or $(x_n)_{n\geq 1}$, or simply (x_n). In the case of a bilateral sequence, the domain of f is $\mathbb{Z}$ instead of $\mathbb{N}^*$.

- It is said that a sequence $(x_n)_{n\geq 1}$ in a n.v.s. $(E, \|\cdot\|)$ converges (in norm) to the limit x if the sequence $\|x_n - x\|$ tends to 0 when n tends to infinity:
$$\lim_{n\to\infty} \|x_n - x\| = 0,$$
and it is written that $\lim\limits_{n\to\infty} x_n = x$. Equivalently, (x_n) converges to x if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\lim \|x_n - x\| < \epsilon$ for all $n \geq N$. If not, (x_n) diverges.

Convergence & Cauchy sequence

A unilateral sequence is a function $f: \mathbb{N}^* \to \mathbb{C}, \mathbb{N}^* \ni n \mapsto x_n \in \mathbb{C}$, written as $(x_n)_{n\in\mathbb{N}^*}$, or $(x_n)_{n\geq 1}$, or simply (x_n). In the case of a bilateral sequence, the domain of f is $\mathbb{Z}$ instead of $\mathbb{N}^*$.

- It is said that a sequence $(x_n)_{n\geq 1}$ in a n.v.s. $(E, \|\cdot\|)$ converges (in norm) to the limit x if the sequence $\|x_n - x\|$ tends to 0 when n tends to infinity:
$$\lim_{n\to\infty} \|x_n - x\| = 0,$$
and it is written that $\lim_{n\to\infty} x_n = x$. Equivalently, (x_n) converges to x if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\lim \|x_n - x\| < \epsilon$ for all $n \geq N$. If not, (x_n) diverges.
- In a metric space (E, d), it is said that $(x_n)_{n\geq 1}$ is a Cauchy sequence if:
$$\forall \epsilon > 0, \exists N \in \mathbb{N} : d(x_n, x_p) < \epsilon, \forall n, p \geq N.$$

Series and sequence of partial sums

A series is the sum of a sequence. Consider sequences in a n.v.s. E. Given a sequence $(x_n)_{n\geq 1}$ in E, the associated series is defined as $\sum\limits_{n=1}^{\infty} x_n$. A sequence is in turn associated with this series, called the sequence of partial sums, such that the Nth partial sum, denoted by S_N^x, is defined as $S_N^x = \sum_{n=1}^{N} x_n$.

Series and sequence of partial sums

A series is the sum of a sequence. Consider sequences in a n.v.s. E. Given a sequence $(x_n)_{n\geq 1}$ in E, the associated series is defined as $\sum_{n=1}^{\infty} x_n$. A sequence is in turn associated with this series, called the sequence of partial sums, such that the Nth partial sum, denoted by S_N^x, is defined as $S_N^x = \sum_{n=1}^{N} x_n$.

If the sequence (S_N^x) of partial sums converges to S when $N \to \infty$, then the series $\sum x_n$ converges to S. Otherwise, the series diverges.

Complete metric space and examples

A metric space E is said to be complete if any Cauchy sequence in E converges to a limit in E.

Complete metric space and examples

A metric space E is said to be complete if any Cauchy sequence in E converges to a limit in E.

- Any finite-dimensional n.v.s. equipped with the metric $d(x, y) = \|x - y\|$ associated with the norm is a complete metric space..
- Any finite-dimensional pre-Hilbert space is complete. On the other hand, any infinite-dimensional pre-Hilbert space is not complete.
- The Cartesian product of a finite number of complete metric spaces is a complete metric space.

Complete metric space and examples

A metric space E is said to be complete if any Cauchy sequence in E converges to a limit in E.

- Any finite-dimensional n.v.s. equipped with the metric $d(x, y) = \|x - y\|$ associated with the norm is a complete metric space..
- Any finite-dimensional pre-Hilbert space is complete. On the other hand, any infinite-dimensional pre-Hilbert space is not complete.
- The Cartesian product of a finite number of complete metric spaces is a complete metric space.

In the case of an incomplete metric space (E, d), its extension to a complete metric space is called the completion of E with respect to d. Any incomplete metric space can be completed by preserving its structure. Thus, a n.v.s. can be completed to a Banach space, and an inner product space to a Hilbert space.

Adherence, density and separability

A few fundamental results related to sequences and more generally to metric spaces. (x_n) is a sequence of a metric set (E, d), and F is a subset of E.

- A subsequence(or extracted sequence) of a sequence $(x_n)_{n\in\mathbb{N}}$ is a sequence $(x_{n_i})_{i\in\mathbb{N}}$, where $(n_i)_{i\in\mathbb{N}}$ is a strictly increasing sequence of elements of $\mathbb{N}$.
- If (x_n) converges, then every extracted sequence converges to the same limit.
- The limits of convergent subsequences of (x_n) are called adherence values of (x_n), and the set of sequence limits in $F \subseteq E$ is called the adherence (or closure) of F, and denoted by $\bar{F}$:

$$x \in \bar{F} \Leftrightarrow \exists\,(x_n) \in F \text{ such that } \lim_{n\to\infty} x_n = x$$

- If a Cauchy sequence has an adherence value, then it converges.
- A subspace F of a n.v.s. E is said to be closed if and only if any convergent sequence of F converges to a limit in F, namely, if $(x_n) \in F$ and $\lim_{n\to\infty} x_n = x$, then $x \in F$.

Adherence, density and separability

For a finite-dimensional n.v.s. E, any subspace F is closed. This is not the case in general, if E is infinite-dimensional.

- A subspace F of a complete metric space E is itself complete if and only if it is closed.
- A subset F of a metric space E is said to be dense in E if and only if the adherence $\bar{F}$ is E. From a topological point of view, the concept of density makes it possible to formalize the fact that for any point $x \in E$, there exists a point of F as close to x as desired.
- A metric space (E, d) is said to be separable if it contains a countable dense subset $F \subset E$, that is, for which $\bar{F} = E$.

Definitions of Banach and Hilbert spaces

A Banach space is a complete n.v.s., or in other words, a complete metric space for the distance induced by the norm.

A Hilbert space is a (real or complex) pre-Hilbert space (or inner product space) in which the norm associated with the inner product makes it a complete space and therefore a Banach space. Complete means that any sequence of functions in a Hilbert space converges to a limit belonging to the space.

Properties

- Any finite-dimensional n.v.s. is a Banach space.
- Any finite-dimensional pre-Hilbert space is a Hilbert space.
- It is important to point out that Hilbert spaces generalize Banach spaces in the sense that they are equipped with an inner product, which is not the case of Banach spaces. This is because it is not always possible to associate an inner product with a norm, while a norm can always be induced from an inner product. Subsequently, any Hilbert space is a Banach space, but the converse is not necessarily true.
- A Banach space is a Hilbert space if and only if its norm satisfies the parallelogram equality (3). There exists a single inner product built using this norm and given by (5).

Links between metric space, n.v.s., pre-Hilbert space, Banach and Hilbert spaces

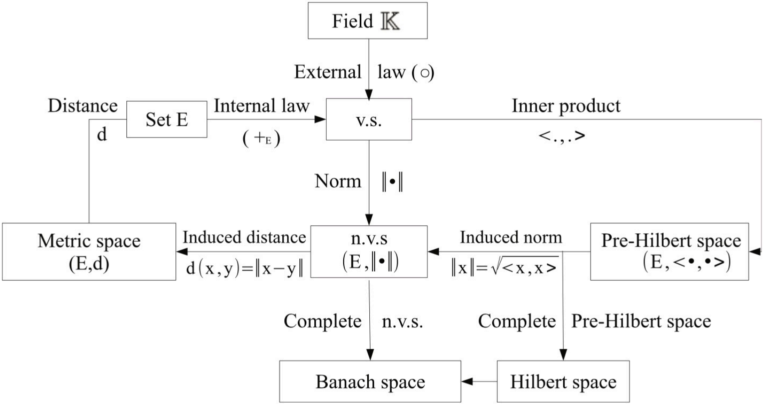

图: Banach and Hilbert spaces.

Examples of Hilbert spaces

- l^p and L^p spaces do not admit any inner product associated with l_p and L_p Hölder norms, for $p \neq 2$. Only l^2 and L^2 spaces corresponding to $p = 2$ are Hilbert spaces.
- $l^2(\mathbb{Z}, \mathbb{K})$: $l^2(\mathbb{Z}, \mathbb{K}) = \left\{ x : \mathbb{Z} \ni n \mapsto x_n \in \mathbb{K}, \sum_{n \in \mathbb{Z}} |x_n|^2 < \infty \right\}$,

$$\langle x, y \rangle = \sum_{n \in \mathbb{Z}} x_n y_n \quad (\text{ or } \sum_{n \in \mathbb{Z}} x_n \bar{y}_n).$$

- $L^2(\mathbb{R}, \mathbb{K})$: $L^2(\mathbb{R}, \mathbb{K}) = \left\{ f : \mathbb{R} \ni t \mapsto f(t) \in \mathbb{K}, \int_{\mathbb{R}} |f(t)|^2 dt < \infty \right\}$,

$$\langle f, g \rangle = \int_{\mathbb{R}} f(t) g(t) dt \quad (\text{ or } \int_{\mathbb{R}} f(t) \bar{g}(t) dt).$$

Hilbert bases: Definitions

- A family $\{b_n\}_{n\in\mathbb{N}^*}$ of vectors of a Hilbert space E is said to be complete (or total) if the space generated by the vectors $b_n, n \in \mathbb{N}^*$, that is, $\mathrm{Vect}\,(b_n, n \in \mathbb{N}^*)$, is equal to the space E. This means that $\mathrm{Vect}\,(b_n, n \in \mathbb{N}^*)$ is dense in E. Any vector $x \in E$ can then be written as the sum of a series $\sum_{n\in\mathbb{N}^*} x_n b_n$, namely:

$$\lim_{N\to\infty} \left\| x - \sum_{n=1}^{N} x_n b_n \right\| = 0$$

  where $(x_n)_{n\in\mathbb{N}^*}$ represent the coordinates of x in the basis $\{b_n\}_{n\in\mathbb{N}^*}$.
- A family of vectors $\{e_n\}_{n\in I}$ of a Hilbert space is called orthonormal, if $\langle e_n, e_p \rangle = \delta_{np}$ for all $n, p \in I$.
- A complete orthonormal family is called a Hilbert basis. The orthonormality property can be demonstrated in a similar way to the case of a finite-dimensional v.s.
- Any separable Hilbert space has a countable Hilbert basis.

Examples of Hilbert bases

- In $l^2(\mathbb{N}^*, \mathbb{C})$, with the inner product $\langle x, y\rangle = \sum_{n\in\mathbb{N}^*} x_n \bar{y}_n$, a canonical Hilbert basis is given by: $e_n = \{\underbrace{0, \cdots, 0}_{\text{(n-1) terms}}, 1, 0, \cdots\}, n \in \mathbb{N}^*$.
- In $L^2([-\pi, \pi], \mathbb{C})$, the space of 2π-periodic functions, with $\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\bar{g}(t)dt$, complex exponential functions $e_n, n \in \mathbb{Z}$, defined as $e_n(t) = e^{int}, t \in [-\pi, \pi], i^2 = -1$, form a complete orthonormal family, called Fourier Hilbert basis.
- The family of trigonometric functions $(\phi_n), n \in \mathbb{N}$ defined as:

$$\begin{aligned}\phi_0 &= 1\\ \phi_{2k}(t) &= \sqrt{2}\cos kt, k \in \mathbb{N}^*\\ \phi_{2k-1}(t) &= \sqrt{2}\sin kt, k \in \mathbb{N}^*\end{aligned}$$

is a Hilbert basis, also called trigonometric basis, in $L^2([-\pi, \pi], \mathbb{C})$.

Fourier series, Parseval's equality and Bessel's inequality

Let f denote a function of a Hilbert space $(E, \langle ., . \rangle)$ and F a complete subspace of E admitting as Hilbert basis $\{e_n, n \in \mathbb{Z}\}$. The expansion of $f \in E$ over this basis is written as:

$$p_F(f) = \sum_{n \in \mathbb{Z}} c_n(f) e_n, c_n(f) = \langle f, e_n \rangle .$$

It corresponds to the orthogonal projection of f onto F, such that $f - p_F(f) \in F^{\perp}$. Parseval's equality then becomes:

$$\sum_{n \in \mathbb{Z}} |c_n(f)|^2 = \|f\|^2$$

and for any finite subset $I \subset \mathbb{Z}$, the Bessel's inequality can be written as:

$$\sum_{n \in I} |c_n(f)|^2 \leq \|f\|^2.$$

2π-periodic functions from $\mathbb{R}$ to $\mathbb{C}$

Let f be a function of the v.s. $C^0_{2\pi}(\mathbb{R},\mathbb{C})$, namely, continuous and of period 2π, from $\mathbb{R}$ to $\mathbb{C}$. Its Fourier series expansion over the orthonormal basis $\mathcal{B} = \{e_n : t \mapsto e^{int}, n \in \mathbb{Z}\}$, with the following inner product:

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\bar{g}(t)dt,$$

is written as:

$$S^f(t) = \sum_{n\in\mathbb{Z}} c_n(f)e^{int}$$

$$c_n(f) = \langle f, e_n\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}dt$$

where $c_n(f)$ is called the exponential Fourier coefficient of rank n of function f.

2π-periodic functions from $\mathbb{R}$ to $\mathbb{C}$

Taking Euler's formula $e^{int} = \cos nt + i \sin nt$ into account, this Fourier series expansion can be rewritten as:

$$\begin{aligned} S^f(t) &= c_0 + \sum_{n \in \mathbb{N}^*} \left(c_n e^{int} + c_{-n} e^{-int} \right) \\ &= c_0 + \sum_{n \in \mathbb{N}^*} \left(c_n + c_{-n} \right) \cos nt + i \left(c_n - c_{-n} \right) \sin nt \\ &= a_0 + \sum_{n \in \mathbb{N}^*} \left(a_n \cos nt + \underline{b_n} \sin nt \right) \end{aligned}$$

with:

$$\begin{aligned} a_0 &= c_0 = \frac{1}{2\pi} \int_0^{2\pi} f(t) dt, \\ a_n &= c_n + c_{-n} = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos nt dt, \\ b_n &= i \left(c_n - c_{-n} \right) = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin nt dt. \end{aligned}$$

The Fourier series expansions $S^f(t) = \sum_{n\in\mathbb{Z}} c_n(f)e^{int}$ and $S^f(t) = a_0 + \sum_{n\in\mathbb{N}^*} (a_n \cos nt + b_n \sin nt)$ are expressed in complex exponential and sine-cosine forms, respectively. They are also called complex form and real form of the Fourier series.

The Fourier series expansions $S^f(t) = \sum_{n\in\mathbb{Z}} c_n(f)e^{int}$ and $S^f(t) = a_0 + \sum_{n\in\mathbb{N}^*} (a_n \cos nt + b_n \sin nt)$ are expressed in complex exponential and sine-cosine forms, respectively. They are also called complex form and real form of the Fourier series.

The real form of the Fourier series presents the following advantages:

- The sine-cosine form is a unilateral series ($n \in \mathbb{N}^*$), while the complex exponential form is a bilateral series ($n \in \mathbb{Z}$f).
- For a real-valued function f, the coefficients a_n and b_n are real, while the coefficients c_n are complex. In general, the real form is preferred to the complex form in the case of real-valued functions.
- The computation of the Fourier coefficients is simplified in the case of an even or odd function.

An example

Consider the 2π-periodic function defined on $[-\pi, \pi]$ as $f(t) = t$. Since this function is odd, we have $a_n = 0, \forall n \in \mathbb{N}$, and using an integration by parts, we get:

$$b_n = \frac{2}{\pi}\int_0^{\pi} t \sin ntdt = \frac{2}{\pi}\left(\left[-\frac{t\cos nt}{n}\right]_0^{\pi} + \frac{1}{n}\int_0^{\pi} \cos ntdt\right) = \frac{2}{n}(-1)^{n+1},$$

which gives:

$$S^f(t) = 2\left(\sin t - \frac{1}{2}\sin 2t + \cdots + \frac{(-1)^{n+1}}{n}\sin nt + \cdots\right)$$

Note that this expression of $S^f(t)$ is valid for $-\pi < t < \pi$, but not at the discontinuity points $\pm\pi$ of f. The Dirichlet-Jordan theorem allows the convergence of the Fourier series to be studied at these points of discontinuity.

T-periodic functions from $\mathbb{R}$ to $\mathbb{C}$

As already pointed out, the shift from a period 2π to a period T can be achieved by changing the variable of integration t in $\frac{2\pi}{T}t = \omega t$ and therefore dt in $\frac{2\pi}{T}dt = \omega dt$ in the calculation formulae of c_n, a_n and b_n.

Convergence of Fourier series

定理 (Dirichlet-Jordan theorem)

If f is a 2π-periodic function, from $\mathbb{R}$ to $\mathbb{C}$, continuous and piecewise smooth, then f has a Fourier series:

$$S_N^f(t) = \sum_{n=-N}^{N} c_n e^{int} = a_0 + \sum_{n=1}^{N} \left(a_n \cos nt + b_n \sin nt\right),$$

such that, at all points t of discontinuity of f, we have:

$$\lim_{N\to+\infty} S_N^f(t) = \frac{f(t^+) + f(t^-)}{2}$$

where $f(t^+)$and $f(t^-)$are the right and left limits of $f(s)$, as $s \to t$, and $S_N^f(t)$ converges pointwise to $f(t)$, at all points t where f is continuous:

$$\lim_{N\to+\infty} S_N^f(t) = f(t).$$

Convergence of Fourier series

The piecewise class C^1 assumption means that f is such that at any point t of discontinuity of the first kind, we have $f(t) = \frac{f(t^-)+f(t^+)}{2}$, namely, $f(t)$ is equal to the average value of its left and right limits at t.

Note that the convergence of the partial sums S_N^f in the neighborhood of a discontinuity point is characterized by the Gibbs-Wilbraham phenomenon, consisting in some overshoot with oscillations.

Examples of Fourier series

Functions	a_0	a_n	b_n
$f_1(t) = \begin{cases} -t+1 & \text{if } 0 \le t \le 1 \\ t+1 & \text{if } -1 \le t \le 0 \end{cases}$	$\frac{1}{2}$	$\frac{2}{n^2\pi^2}(1-\cos n\pi) = \begin{cases} 0 & \text{if } n = 2k \\ \frac{4}{(2k+1)^2\pi^2} & \text{if } n = 2k+1 \end{cases}$	0
$f_2(t) = \lvert t \rvert \quad , \quad \text{if } t \in [-\pi, \pi]$	$\frac{\pi}{2}$	$\frac{2}{n\pi^2}((-1)^n - 1) = \begin{cases} 0 & \text{if } n = 2k \\ -\frac{4}{(2k+1)^2\pi} & \text{si } n = 2k+1 \end{cases}$	0
$f_3(t) = t(2\pi - t)\ ,\ t \in [0, 2\pi]$	$\frac{2\pi^2}{3}$	$-\frac{4}{n^2}$	0
$f_4(t) = t \quad , \quad -\pi \le t \le \pi$	0	0	$2\frac{(-1)^{n+1}}{n}$
$f_5(t) = \begin{cases} 1 & \text{if } 0 < t < \pi \\ -1 & \text{if } -\pi < t < 0 \end{cases}$	0	0	$\frac{2}{n\pi}(1-\cos n\pi) = \begin{cases} 0 & \text{if } n = 2k \\ \frac{4}{(2k+1)\pi} & \text{if } n = 2k+1 \end{cases}$
$f_6(t) = \begin{cases} t(\pi - t)\ , & t \in [0, \pi] \\ t(\pi + t)\ , & t \in [-\pi, 0] \end{cases}$	0	0	$\frac{4}{n^3\pi}(1-\cos n\pi) = \begin{cases} 0 & \text{if } n = 2k \\ \frac{8}{(2k+1)^3\pi} & \text{if } n = 2k+1 \end{cases}$

图: Six examples of Fourier series.

Examples of Fourier series

The corresponding Fourier series are represented by the following equations:

$$S^{f_1}(t) = \frac{1}{2} + \frac{4}{\pi^2} \sum_{k \in \mathbb{N}} \frac{\cos(2k+1)\pi t}{(2k+1)^2}$$

$$S^{f_2}(t) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{k \in \mathbb{N}} \frac{\cos(2k+1)t}{(2k+1)^2}$$

$$S^{f_3}(t) = \frac{2\pi^2}{3} - 4 \sum_{n \in \mathbb{N}^*} \frac{\cos nt}{n^2}$$

$$S^{f_4}(t) = 2 \sum_{n \in \mathbb{N}^*} (-1)^{n+1} \frac{\sin nt}{n}$$

$$S^{f_5}(t) = \frac{4}{\pi} \sum_{k \in \mathbb{N}} \frac{\sin(2k+1)t}{2k+1}$$

$$S^{f_6}(t) = \frac{8}{\pi} \sum_{k \in \mathbb{N}} \frac{\sin(2k+1)t}{(2k+1)^3}$$

Examples of Fourier series

These results highlight the limitations of Fourier series expansions. These expansions are suitable for the approximation of periodic functions, with periodicity created by extension of the interval of definition, corresponding in practice to an interval of observation of finite length.

Examples of Fourier series

These results highlight the limitations of Fourier series expansions. These expansions are suitable for the approximation of periodic functions, with periodicity created by extension of the interval of definition, corresponding in practice to an interval of observation of finite length.

Nonetheless, the convergence of partial Fourier sum S_N^f is highly dependent on the shape of the signal to be analyzed. In addition, a local perturbation in the signal causes all the Fourier coefficients to be modified. This is not the case of wavelet bases that allow for a multiresolution expansion of the signal to be analyzed, by way of decomposition of the Hilbert space $L_2(R, R)$ into a sum of orthogonal subspaces, each subspace being associated with a level of resolution.

Expansions over bases of orthogonal polynomials

In the same way as trigonometric Hilbert bases allow periodic functions to be represented and analyzed using Fourier series, it is possible to use polynomial bases for the representation and analysis of functions of the Hilbert space $L_2([a, b], R)$. In the following, we consider Legendre, Hermite, Laguerre, and Chebyshev (of the first kind) orthogonal polynomials that lead to the series named after these authors.

Expansions over bases of orthogonal polynomials

In the same way as trigonometric Hilbert bases allow periodic functions to be represented and analyzed using Fourier series, it is possible to use polynomial bases for the representation and analysis of functions of the Hilbert space $L_2([a, b], R)$. In the following, we consider Legendre, Hermite, Laguerre, and Chebyshev (of the first kind) orthogonal polynomials that lead to the series named after these authors.

In the v.s. of polynomials, the set of monomials $\{1, t, \cdots, t^n, \cdots\}$ is independent. Considering the weighted inner product:

$$\langle f, g \rangle = \int_a^b w(t) f(t) g(t) dt,$$

various polynomial Hilbert bases can be built by applying the Gram-Schmidt method to the set $\{1, t, t^2, \cdots\}$, for various intervals $[a, b]$ and various weighting functions $w(t)$, associated with different spaces of functions.

Expansions over bases of orthogonal polynomials

In the following, we summarize the spaces of functions and inner products considered for the construction of orthogonal Legendre, Hermite, Laguerre, and Chebyshev polynomials, which will be denoted by $L_n(t), H_n(t), \Lambda_n(t)$, and $T_n(t)$, respectively.

Polynomials	Spaces	Inner products
Legendre	$L^2([-1,1],\mathbb{R}) = \left\{ f : [-1,1] \to \mathbb{R}, \int_{-1}^{1} f^2(t)dt < \infty \right\}$	$< f, g >= \frac{1}{2} \int_{-1}^{1} f(t)g(t)dt$
Hermite	$L^2(\mathbb{R},\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{R}, \int_{-\infty}^{+\infty} f^2(t)e^{-t^2} dt < \infty \right\}$	$< f, g >= \int_{-\infty}^{+\infty} f(t)g(t)e^{-t^2} dt$
Laguerre	$L^2(\mathbb{R}^+,\mathbb{R}) = \left\{ f : \mathbb{R}^+ \to \mathbb{R}, \int_{0}^{\infty} f^2(t)e^{-t}dt < \infty \right\}$	$< f, g >= \int_{0}^{\infty} f(t)g(t)e^{-t}dt$
Chebyshev	$L^2([-1,1],\mathbb{R}) = \left\{ f : [-1,1] \to \mathbb{R}, \int_{-1}^{1} \frac{f^2(t)}{\sqrt{1-t^2}} dt < \infty \right\}$	$< f, g >= \frac{2}{\pi} \int_{-1}^{1} \frac{f(t)g(t)}{\sqrt{1-t^2}} dt$

Spaces of polynomials and associated inner products.

Expansions over bases of orthogonal polynomials

In the following, we present Rodrigues formulae (with $n \in \mathbb{N}^*$) and three-term recurrence relations satisfied by the four families of polynomials. Rodrigues formulae allow the orthogonal polynomials to be generated from successive derivations of different functions.

Polynomials	Rodrigues formulae	Recurrence relations
Legendre	$L_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n}(t^2-1)^n$	$(n+1)L_{n+1}(t) = (2n+1)\,tL_n(t) - nL_{n-1}(t)$ $L_0(t) = 1, L_1(t) = t \;,\; \forall n \geq 1$
Hermite	$H_n(t) = (-1)^n e^{t^2} \frac{d^n}{dt^n}(e^{-t^2})$	$H_{n+1}(t) = 2tH_n(t) - 2nH_{n-1}(t)$ $H_0(t) = 1, H_1(t) = 2t \;,\; \forall n \geq 1$
Laguerre	$\Lambda_n(t) = \frac{e^t}{n!} \frac{d^n}{dt^n}(t^n e^{-t})$	$(n+1)\Lambda_{n+1}(t) = (2n+1-t)\Lambda_n(t) - n\Lambda_{n-1}(t)$ $\Lambda_0(t) = 1, \Lambda_1(t) = 1-t \;,\; \forall n \geq 1$
Chebyshev	$T_n(t) = \frac{(-1)^n 2^n n!}{(2n)!}\sqrt{1-t^2}\,\frac{d^n}{dt^n}$ $\times\left((1-t^2)^{n-\frac{1}{2}}\right)$	$T_{n+1}(t) = 2t\,T_n(t) - T_{n-1}(t)$ $T_0(t) = 1, T_1(t) = t \;,\; \forall n \geq 1$

图: Rodrigues formulae and recurrence relations.

Expansions over bases of orthogonal polynomials

Polynomials	**Differential equations, $\forall n \in \mathbb{N}$**
Legendre	$(1-t^2)L_n''(t) - 2tL_n'(t) + n(n+1)L_n(t) = 0$
Hermite	$H_n''(t) - 2tH_n'(t) + 2nH_n(t) = 0$
Laguerre	$t\Lambda_n''(t) + (1-t)\Lambda_n'(t) + n\Lambda_n(t) = 0$
Chebyshev	$(1-t^2)T_n''(t) - tT_n'(t) + n^2T_n(t) = 0$

图: Differential equations.

Expansions over bases of orthogonal polynomials

Polynomials	Orthogonality relations
Legendre	$\langle L_n, L_p \rangle = \frac{1}{2}\int_{-1}^{1} L_n(t)L_p(t)dt = \frac{1}{2n+1}\delta_{np}$
Hermite	$\langle H_n, H_p \rangle = \int_{-\infty}^{+\infty} H_n(t)H_p(t)e^{-t^2}dt = 2^n n!\sqrt{\pi}\delta_{np}$
Laguerre	$\langle \Lambda_n, \Lambda_p \rangle = \int_0^{\infty} \Lambda_n(t)\Lambda_p(t)e^{-t}dt = \delta_{np}$
Chebyshev	$\langle T_n, T_p \rangle = \frac{2}{\pi}\int_{-1}^{1} T_n(t)T_p(t)\frac{dt}{\sqrt{1-t^2}} = \begin{cases} 0, n \neq p \\ 2, n = p = 0 \\ 1, n = p \neq 0 \end{cases}$

图: Orthogonality relations.

Chebyshev polynommials: the first-kind T_n

定义

The Chebyshev polynomial $T_n(x)$ of the first kind is a polynomial in x of degree n, defined by the relation

$$T_n(x) = \cos n\theta \quad \text{when } x = \cos\theta. \tag{9}$$

Chebyshev polynommials: the first-kind T_n

定义

The Chebyshev polynomial $T_n(x)$ of the first kind is a polynomial in x of degree n, defined by the relation

$$T_n(x) = \cos n\theta \quad \text{when } x = \cos\theta. \tag{9}$$

It is well known (as a consequence of de Moivre's Theorem) that $\cos n\theta$ is a polynomial of degree n in $\cos\theta$, and indeed

$$\cos 0\theta = 1, \quad \cos 1\theta = \cos\theta, \quad \cos 2\theta = 2\cos^2\theta - 1$$
$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta, \quad \cos 4\theta = 8\cos^4\theta - 8\cos^2\theta + 1, \quad \ldots$$

定义

The Chebyshev polynomial $T_n(x)$ of the first kind is a polynomial in x of degree n, defined by the relation

$$T_n(x) = \cos n\theta \quad \text{when } x = \cos\theta. \tag{9}$$

It is well known (as a consequence of de Moivre's Theorem) that $\cos n\theta$ is a polynomial of degree n in $\cos\theta$, and indeed

$$\cos 0\theta = 1, \quad \cos 1\theta = \cos\theta, \quad \cos 2\theta = 2\cos^2\theta - 1$$
$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta, \quad \cos 4\theta = 8\cos^4\theta - 8\cos^2\theta + 1, \quad \ldots$$

Deduce from Equ. (9), that the first few Chebyshev polynomials are

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1$$
$$T_3(x) = 4x^3 - 3x, \quad T_4(x) = 8x^4 - 8x^2 + 1, \quad \ldots$$

Chebyshev polynommials T_n

In practice it is neither convenient nor efficient to work out each $T_n(x)$ from first principles. Rather by combining the trigonometric identity

$$\cos n\theta + \cos(n-2)\theta = 2\cos\theta\cos(n-1)\theta$$

with the Definition, we obtain the fundamental recurrence relation

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x), \quad n = 2, 3, \ldots$$

which together with the initial conditions

$$T_0(x) = 1, \quad T_1(x) = x$$

recursively generates all the polynomials $\{T_n(x)\}$ very efficiently.

Chebyshev polynommials T_n

One interpretation of the equation $T_n(x) = \cos(n \arccos(x))$ is the following quote from Forman S. Acton's book Numerical Methods that Work: "Chebyschev polynomials are actually cosine curves with a somewhat disturbed horizontal scale, but the vertical scale has not been touched."

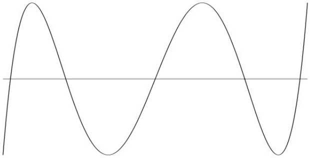

Figure 1.1: $T_5(x)$ on range $[-1, 1]$

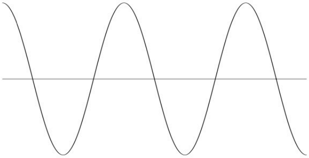

Figure 1.2: $\cos 5\theta$ on range $[0, \pi]$

Chebyshev polynommials T_n: properties

- $\max\limits_{-1\le x\le 1} |T_n(x)| = 1$. In fact there are $n+1$ local extrema, $n-1$ of these extrema are interior to $[-1,1]$, the other two extrema being at the end points $x = \pm 1$, i.e., $x = \cos(\frac{k\pi}{n})$, or in their natural order $x = \cos(\frac{(n-k)}{\pi} n)$, for $k = 0, 1, \ldots, n$.
- $x_k = \cos\left(\frac{(2k-1)\pi}{2n}\right)$, $k = 1, 2, \ldots, n$ are zeros of T_n. In particular, all the roots of T_n are real and lie in the interval $[-1,1]$.
- The leading coefficient (that of x_n) in $T_n(x)$ for $n > 1$ is 2^{n-1}.
- $T_n(-x) = (-1)^n T_n(x)$, i.e., $T_n(x)$ is an even or odd function, involving only even or odd powers of x, according as n is even or odd.
- $\frac{2}{\pi}\int_0^\pi \cos(m\theta)\cos(n\theta)d\theta = \frac{2}{\pi}\int_{-1}^1 T_m(x)T_n(x)\frac{dx}{\sqrt{1-x^2}} = \delta_{mn}$.
- A monic polynomial is a polynomial whose leading coefficient is 1. In approximation theory, it is useful to identify the nth degree monic polynomial with the smallest uniform norm on $[-1,1]$, which turns out to be $2^{1-n}T_n(x)$.
- $T_m(T_n(x)) = T_{mn}(x)$.

Chebyshev polynommials: the second-kind U_n

定义

The Chebyshev polynomial $U_n(x)$ of the second kind is a polynomial of degree n in x defined by

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta} \text{ when } x = \cos\theta. \tag{10}$$

Chebyshev polynommials U_n

Elementary formulae give

$$\sin 1\theta = \sin\theta, \quad \sin 2\theta = 2\sin\theta\cos\theta, \quad \sin 3\theta = \sin\theta\left(4\cos^2\theta - 1\right)$$
$$\sin 4\theta = \sin\theta\left(8\cos^3\theta - 4\cos\theta\right), \quad \ldots$$

so that we see that the ratio of sine functions (10) is indeed a polynomial in $\cos\theta$, and we may immediately deduce that

$$U_0(x) = 1, \quad U_1(x) = 2x, \quad U_2(x) = 4x^2 - 1$$
$$U_3(x) = 8x^3 - 4x, \quad \ldots$$

Chebyshev polynommials U_n

Elementary formulae give

$$\sin 1\theta = \sin\theta, \quad \sin 2\theta = 2\sin\theta\cos\theta, \quad \sin 3\theta = \sin\theta\left(4\cos^2\theta - 1\right)$$
$$\sin 4\theta = \sin\theta\left(8\cos^3\theta - 4\cos\theta\right), \quad \ldots$$

so that we see that the ratio of sine functions (10) is indeed a polynomial in $\cos\theta$, and we may immediately deduce that

$$U_0(x) = 1, \quad U_1(x) = 2x, \quad U_2(x) = 4x^2 - 1$$
$$U_3(x) = 8x^3 - 4x, \quad \ldots$$

By combining the trigonometric identity

$$\sin(n+1)\theta + \sin(n-1)\theta = 2\cos\theta\sin n\theta$$

we see that $U_n(x)$ satisfies the recurrence relation

$$U_n(x) = 2xU_{n-1}(x) - U_{n-2}(x), \quad n = 2, 3, \ldots$$

which together with the initial conditions

$$U_0(x) = 1, \quad U_1(x) = 2x$$

Relationship between T_n and U_n

A similar trigonometric identity

$$\sin(n+1)\theta - \sin(n-1)\theta = 2\sin\theta\cos n\theta$$

leads us to a relationship

$$U_n(x) - U_{n-2}(x) = 2T_n(x), \quad n = 2, 3, \ldots$$

between the polynomials of the first and second kinds.

Relationship between T_n and U_n

A similar trigonometric identity

$$\sin(n+1)\theta - \sin(n-1)\theta = 2\sin\theta\cos n\theta$$

leads us to a relationship

$$U_n(x) - U_{n-2}(x) = 2T_n(x), \quad n = 2, 3, \ldots$$

between the polynomials of the first and second kinds.

Note that the recurrence equation for $\{U_n(x)\}$ is identical in form to the recurrence equation for $\{T_n(x)\}$. The different initial conditions ($U_1 = 2x$, $T_1 = x$) yield the different polynomial systems.

Shifted Chebyshev polynomials T_n^*

Since the range $[0, 1]$ is quite often more convenient to use than the range $[-1, 1]$, we can map x in $[0, 1]$ to s in $[-1, 1]$ by the transformation

$$s = 2x - 1 \text{ or } x = \frac{1}{2}(1 + s),$$

and this leads to a shifted Chebyshev polynomial $T_n^*(x)$ of degree n in x on $[0, 1]$ given by

$$T_n^*(x) = T_n(s) = T_n(2x - 1).$$

Shifted Chebyshev polynomials T_n^*

Since the range $[0,1]$ is quite often more convenient to use than the range $[-1,1]$, we can map x in $[0,1]$ to s in $[-1,1]$ by the transformation

$$s = 2x - 1 \text{ or } x = \frac{1}{2}(1+s),$$

and this leads to a shifted Chebyshev polynomial $T_n^*(x)$ of degree n in x on $[0,1]$ given by

$$T_n^*(x) = T_n(s) = T_n(2x-1).$$

Thus we have the polynomials

$$T_0^*(x) = 1, \quad T_1^*(x) = 2x - 1, \quad T_2^*(x) = 8x^2 - 8x + 1,$$
$$T_3^*(x) = 32x^3 - 48x^2 + 18x - 1, \quad \ldots .$$

Then, we may deduce the recurrence relation for T_n^* in the form

$$T_n^*(x) = 2(2x-1)\,T_{n-1}^*(x) - T_{n-2}^*(x)$$

with initial conditions

$$T_0^*(x) = 1, \quad T_1^*(x) = 2x - 1.$$

Shifted Chebyshev polynomials T_n^*

The polynomials $T_n^*(x)$ have a further special property, since

$$T_{2n}(x) = \cos(2n\theta) = \cos n(2\theta) = T_n(\cos 2\theta) = T_n(2x^2 - 1) = T_n^*(x^2)$$

so that

$$T_{2n}(x) = T_n^*(x^2).$$

Shifted Chebyshev polynomials T_n^*

The polynomials $T_n^*(x)$ have a further special property, since

$$T_{2n}(x) = \cos(2n\theta) = \cos n(2\theta) = T_n(\cos 2\theta) = T_n(2x^2 - 1) = T_n^*(x^2)$$

so that

$$T_{2n}(x) = T_n^*(x^2).$$

Thus $T_n^*(x)$ is precisely $T_{2n}(\sqrt{x})$, a higher degree Chebyshev polynomial in

the square root of the argument, and this relation gives an important link between $\{T_n\}$ and $\{T_n^*\}$ which complements the shift relationship $T_n^*(x) = T_n(s) = T_n(2x - 1)$.

Shifted Chebyshev polynommials T_n^*

It is of course possible to define T_n^*, like T_n, directly by a trigonometric relation. Indeed, if we combine $T_n(x) = \cos n\theta$ and $T_{2n}(x) = T_n^*(x^2)$ we obtain

$$T_n^*(x) = \cos 2n\theta \text{ when } x = \cos^2\theta.$$

Shifted Chebyshev polynommials T_n^*

It is of course possible to define T_n^*, like T_n, directly by a trigonometric relation. Indeed, if we combine $T_n(x) = \cos n\theta$ and $T_{2n}(x) = T_n^*(x^2)$ we obtain

$$T_n^*(x) = \cos 2n\theta \text{ when } x = \cos^2\theta.$$

This relation might alternatively be rewritten, with θ replaced by $\phi/2$, in the form

$$T_n^*(x) = \cos n\phi \text{ when } x = \cos^2\phi/2 = \frac{1}{2}(1 + \cos\phi).$$

Indeed the latter formula could be obtained directly from $T_n^*(x) = T_n(s) = T_n(2x-1)$, by writing

$$T_n(s) = \cos n\phi \text{ when } s = \cos\phi.$$

Shifted Chebyshev polynommials T_n^*

It is of course possible to define T_n^*, like T_n, directly by a trigonometric relation. Indeed, if we combine $T_n(x) = \cos n\theta$ and $T_{2n}(x) = T_n^*(x^2)$ we obtain

$$T_n^*(x) = \cos 2n\theta \text{ when } x = \cos^2\theta.$$

This relation might alternatively be rewritten, with θ replaced by $\phi/2$, in the form

$$T_n^*(x) = \cos n\phi \text{ when } x = \cos^2\phi/2 = \frac{1}{2}(1+\cos\phi).$$

Indeed the latter formula could be obtained directly from $T_n^*(x) = T_n(s) = T_n(2x-1)$, by writing

$$T_n(s) = \cos n\phi \text{ when } s = \cos\phi.$$

Note that the shifted Chebyshev polynomial $T_n^*(x)$ is neither even nor odd, and indeed all powers of x from $1 = x^0$ to x^n appear in $T_n^*(x)$. The leading coefficient of x^n in $T_n^*(x)$ for $n > 0$ is 2^{2n-1}.

Chebyshev polynomials for the general range $[a, b]$

More generally we may define Chebyshev polynomials appropriate to any given finite range $[a, b]$ of x, by making this range correspond to the range $[-1, 1]$ of a new variable s under the linear transformation

$$s = \frac{2x - (a + b)}{b - a}.$$

The Chebyshev polynomials of the first kind appropriate to $[a, b]$ are thus $T_n(s)$.

Chebyshev polynomials of a complex variable

We have chosen to define the polynomials $T_n(x), U_n(x)$ with reference to the interval $[-1, 1]$. However, their expressions as sums of powers of x can of course be evaluated for any real x, even though the substitution $x = \cos\theta$ is not possible outside this interval.

Chebyshev polynomials of a complex variable

We have chosen to define the polynomials $T_n(x), U_n(x)$ with reference to the interval $[-1,1]$. However, their expressions as sums of powers of x can of course be evaluated for any real x, even though the substitution $x = \cos\theta$ is not possible outside this interval.

For x in the range $[1,\infty)$, we can make the alternative substitution

$$x = \cosh\Theta,$$

with Θ in the range $[0,\infty)$, and it is easily verified that precisely the same polynomials $T_n(x), U_n(x)$ are generated by the relations

$$T_n(x) = \cosh n\Theta,$$
$$U_n(x) = \frac{\sinh(n+1)\Theta}{\sinh\Theta}.$$

Chebyshev polynomials of a complex variable

For x in the range $(-\infty, -1]$ we can make use of the odd or even parity of the Chebyshev polynomials to deduce that, for instance,

$$T_n(x) = (-1)^n \cosh n\Theta,$$

where

$$x = -\cosh \Theta.$$

Note that none of the two kinds of Chebyshev polynomials can have any zeros or turning points in the range $[1, \infty)$. The same applies to the range $(-\infty, -1]$. $T_n(x), U_n(x)$ have n real zeros in the interval $[-1, 1]$. The Chebyshev polynomial $T_n(x)$ can be further extended into (or initially defined as) a polynomial $T_n(z)$ of a complex variable z.

Complex chebyshev polynommials: Conformal mapping of a circle to and from an ellipse

For convenience, consider not only the variable z but a related complex variable w such that

$$z = \frac{1}{2}\left(w + w^{-1}\right). \tag{11}$$

Then, if w moves on the circle $|w| = r$ (for $r > 1$) centred at the origin, we have

$$w = re^{i\theta} = r\cos\theta + ir\sin\theta,$$
$$w^{-1} = r^{-1}e^{-i\theta} = r^{-1}\cos\theta - ir^{-1}\sin\theta,$$

and from Equ. (11), we have $z = a\cos\theta + ib\sin\theta$, where
$a = \frac{1}{2}\left(r + r^{-1}\right), \quad b = \frac{1}{2}\left(r - r^{-1}\right)$.
Hence z moves on the standard ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

centred at the origin, with major and minor semi-axes a, b.

Conformal mapping of a circle to and from an ellipse: case $r > 1$

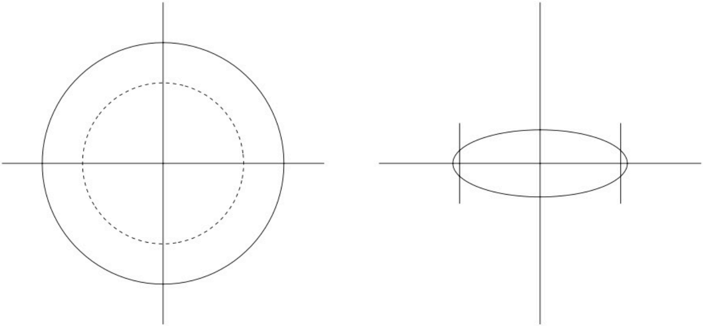

图: The circle $|w| = r = 1.5$ and its image in the z plane.

Conformal mapping of a circle to and from an ellipse: case $r = 1$

In the case $r = 1$, where w moves on the unit circle, we have $b = 0$ and the ellipse collapses into the real interval $[-1, 1]$. However, z traverses the interval twice as w moves round the circle: from -1 to 1 as θ moves from $-\pi$ to 0, and from 1 to -1 as θ moves from 0 to π.

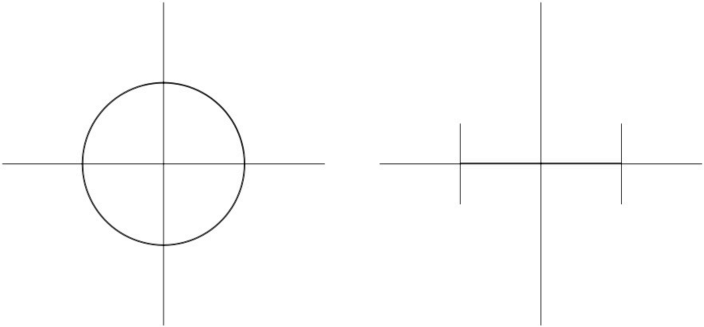

图: The circle $|w| = r = 1$ and its image in the z plane.

Conformal mapping of a circle to and from an ellipse

From Equ. (11) we readily deduce that w satisfies

$$w^2 - 2wz + 1 = 0,$$

a quadratic equation with two solutions $w = w_1, w_2 = z \pm \sqrt{z^2 - 1}$. This means that the mapping from w to z is 2 to 1 , with branch points at $z = \pm 1$. It is convenient to define the complex square root $\sqrt{z^2 - 1}$ so that it lies in the same quadrant as z (except for z on the real interval $[-1, 1]$, along which the plane must be cut), and to choose the solution

$$w = w_1 = z + \sqrt{z^2 - 1},$$

so that $|w| = |w_1| \geq 1$. Then w depends continuously on z along any path that does not intersect the interval $[-1, 1]$, and it is easy to verify that

$$w_2 = w_1^{-1} = z - \sqrt{z^2 - 1},$$

with $|w_2| \leq 1$.

Conformal mapping of a circle to and from an ellipse

If w_1 moves on $|w_1| = r$, for $r > 1$, then w_2 moves on $|w_2| = \left|w_1^{-1}\right| = r^{-1} < 1$. Hence both of the concentric circles

$$C_r := \{w : |w| = r\}, \quad C_{1/r} := \left\{w : |w| = r^{-1}\right\}$$

transform into the same ellipse defined by $z = a\cos\theta + \mathrm{i}b\sin\theta$ or $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, namely

$$E_r := \left\{z : \left|z + \sqrt{z^2 - 1}\right| = r\right\}.$$

Chebyshev polynommials in z

Defining $z = \frac{1}{2}\left(w + w^{-1}\right)$, if w lies on the unit circle $|w| = 1$, then $z = a\cos\theta + \mathrm{i}b\sin\theta$ gives

$$z = \cos\theta$$

and hence,

$$w = z + \sqrt{z^2 - 1} = \mathrm{e}^{\mathrm{i}\theta}.$$

Thus $T_n(z)$ is now a Chebyshev polynomial in a real variable and,

$$T_n(z) = \cos n\theta = \frac{1}{2}(\mathrm{e}^{\mathrm{i}n\theta} + \mathrm{e}^{-\mathrm{i}n\theta}) = \frac{1}{2}(w^n + w^{-n}).$$

This leads immediately to the general definition, for all complex z, namely

$$T_n(z) = \frac{1}{2}(w^n + w^{-n}), \text{ where } z = \frac{1}{2}(w + w^{-1}).$$

Chebyshev polynommials in z

Alternatively we may write $T_n(z)$ directly in terms of z as

$$T_n(z) = \frac{1}{2}\left\{\left(z + \sqrt{z^2 - 1}\right)^n + \left(z - \sqrt{z^2 - 1}\right)^n\right\}.$$

If z lies on the ellipse E_r, the locus of $z = \frac{1}{2}(w + w^{-1})$ when $|w| = r > 1$, then it follows from $T_n(z) = \frac{1}{2}(w^n + w^{-n})$ that we have the inequality

$$\frac{1}{2}\left(r^n - r^{-n}\right) \leq |T_n(z)| \leq \frac{1}{2}\left(r^n + r^{-n}\right).$$

Chebyshev polynommials in z

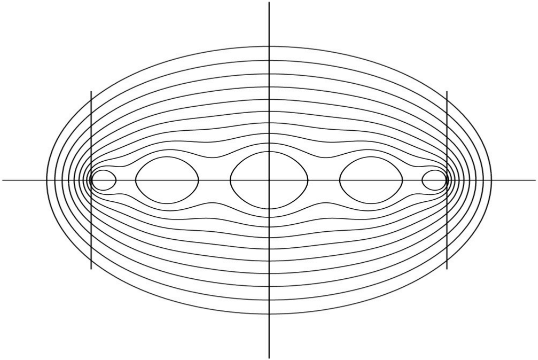

图: Contours of $|T_5(z)|$ in the complex plane.

We see the level curves of the absolute value of $T_5(z)$, and it can easily be seen how these approach an elliptical shape as the value increases.

Powers of x in terms of $\{T_n(x)\}$

Let $x = \cos\theta$, the power $x^n = \cos^n\theta$ can be expressed in terms of $\cos n\theta, \cos(n-2)\theta, \cos(n-4)\theta, \ldots$, by using the binomial theorem as follows:

$$\begin{aligned}\left(e^{i\theta} + e^{-i\theta}\right)^n =& e^{in\theta} + \binom{n}{1} e^{i(n-2)\theta} + \cdots + \binom{n}{n-1} e^{-i(n-2)\theta} + e^{-in\theta} \\ &= \left(e^{in\theta} + e^{-in\theta}\right) + \binom{n}{1}\left(e^{i(n-2)\theta} + e^{-i(n-2)\theta}\right) + \\ &+ \binom{n}{2}\left(e^{i(n-4)\theta} + e^{-i(n-4)\theta}\right) + \cdots\end{aligned}$$

Using the fact that

$$\left(e^{i\theta} + e^{-i\theta}\right)^n = (2\cos\theta)^n = 2^n\cos^n\theta$$

we have

$$x^n = 2^{1-n}\sum_{k=0}^{\lfloor n/2\rfloor}{}' \binom{n}{k} T_{n-2k}(x).$$

Powers of x in terms of $\{T_n(x)\}$: an example

例

Taking $n = 4$

$$\begin{aligned} x^4 &= 2^{-3} \sum_{k=0}^{2}{}' \binom{4}{k} T_{4-2k}(x) \\ &= 2^{-3} \left[T_4(x) + \binom{4}{1} T_2(x) + \frac{1}{2} \binom{4}{2} T_0(x) \right] \\ &= \frac{1}{8} T_4(x) + \frac{1}{2} T_2(x) + \frac{3}{8} T_0(x). \end{aligned}$$

$T_n(x)$ in terms of powers of x

It is not quite as simple to derive formulae in the reverse direction. The obvious device to use is de Moivre's Theorem:

$$\cos n\theta + \mathrm{i} \sin n\theta = (\cos\theta + \mathrm{i}\sin\theta)^n.$$

Expanding by the binomial theorem and taking the real part,

$$\cos n\theta = \cos^n\theta - \binom{n}{2}\cos^{n-2}\theta\sin^2\theta + \binom{n}{4}\cos^{n-4}\theta\sin^4\theta + \cdots$$

If $\sin^2\theta$ is replaced by $1-\cos^2\theta$ throughout, then a formula is obtained for $\cos n\theta$ in terms of $\cos^n\theta, \cos^{n-2}\theta, \cos^{n-4}\theta, \ldots$ On transforming to $x = \cos\theta$, this leads to the required formula for $T_n(x)$ in terms of $x^n, x^{n-2}, x^{n-4}, \ldots$.

$T_n(x)$ in terms of powers of x

Refer to Rivlin (1974), the relevant result is obtained in the form

$$T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \left[(-1)^k \sum_{j=k}^{\lfloor n/2 \rfloor} \binom{n}{2j} \binom{j}{k} \right] x^{n-2k}.$$

$T_n(x)$ in terms of powers of x

Refer to Rivlin (1974), the relevant result is obtained in the form

$$T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \left[(-1)^k \sum_{j=k}^{\lfloor n/2 \rfloor} \left(\begin{array}{c} n \\ 2j \end{array} \right) \left(\begin{array}{c} j \\ k \end{array} \right) \right] x^{n-2k}.$$

A rather simpler formula is given, for example, by Clenshaw (1962) and Snyder (1966) in the form

$$T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} c_k^{(n)} x^{n-2k}$$

where

$$c_k^{(n)} = (-1)^k 2^{n-2k-1} \left[2 \left(\begin{array}{c} n-k \\ k \end{array} \right) - \left(\begin{array}{c} n-k-1 \\ k \end{array} \right) \right] \quad (2k < n)$$

and

$$c_k^{(2k)} = (-1)^k \quad (k \geq 0).$$

Evaluation of a Chebyshev sum

Suppose that we wish to evaluate the sum

$$S_n = \sum_{r=0}^{n} a_r T_r(x) = a_0 T_0(x) + a_1 T_1(x) + \cdots + a_n T_n(x). \qquad (12)$$

We may write Equ. (12) in vector form as

$$S_n = \mathbf{a}^T \mathbf{t},$$

where $\mathbf{a}^T$ and $\mathbf{t}$ denote the row- and column-vectors

$$\mathbf{a}^T = (a_0, a_1, \ldots, a_n)\,, \quad \mathbf{t} = \begin{pmatrix} T_0(x) \\ T_1(x) \\ \vdots \\ T_n(x) \end{pmatrix}.$$

Evaluation of a Chebyshev sum

Then the recurrence relation

$$T_r(x) - 2xT_{r-1}(x) + T_{r-2}(x) = 0, r = 2, 3, \ldots$$

with $T_0(x) = 1$ and $T_1(x) = x$, may be written in matrix notation as

$$\begin{pmatrix} 1 & & & & & & \\ -2x & 1 & & & & & \\ 1 & -2x & 1 & & & & \\ & 1 & -2x & 1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 1 & -2x & 1 & \\ & & & & 1 & -2x & 1 \end{pmatrix} \begin{pmatrix} T_0(x) \\ T_1(x) \\ T_2(x) \\ T_3(x) \\ \vdots \\ T_{n-1}(x) \\ T_n(x) \end{pmatrix} = \begin{pmatrix} 1 \\ -x \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

or (denoting the $(n+1) \times (n+1)$ matrix by $\mathbf{A}$)

$$\mathbf{At} = \mathbf{c}.$$

Evaluation of a Chebyshev sum

Let

$$\mathbf{b}^T = (b_0, b_1, \ldots, b_n)$$

be the row vector satisfying the equation

$$(b_0, b_1, \ldots, b_n) \begin{pmatrix} 1 & & & & \\ -2x & 1 & & & \\ 1 & -2x & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2x & 1 \end{pmatrix} = (a_0, a_1, \ldots, a_n)$$

or

$$\mathbf{b}^T \mathbf{A} = \mathbf{a}^T.$$

Then we have

$$S_n = \mathbf{a}^T \mathbf{t} = \mathbf{b}^T \mathbf{A} \mathbf{t} = \mathbf{b}^T \mathbf{c} = b_0 - b_1 x.$$