Mathematical Methods for Computer Science

R.J. Gibbens

Computer Laboratory
University of Cambridge

Michaelmas Term 2008/9
(Last revised on 22 Sep 2008)

Inner product spaces

Introduction

In this section we shall consider what it means to represent a function $f(x)$ in terms of other, perhaps simpler, functions. One example is Fourier series of the form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)] \ .$$

How are the coefficients a_n and b_n related to the choice of $f(x)$ and what other representations can we use?
We shall take a quite general approach to these questions and derive the necessary framework that underpins a wide range of applications.

Linear space

Definition (Linear space)

A non-empty set V of vectors is a linear space over a field $\mathbb{F}$ of scalars if the following are satisfied.

1. Binary operation $+$ such that if $u, v \in V$ then $u + v \in V$
2. $+$ is associative: for all $u, v, w \in V$ then $(u + v) + w = u + (v + w)$
3. There exists a zero vector, written $\vec{0} \in V$, such that $\vec{0} + v = v$ for all $v \in V$.
4. For all $v \in V$, there exists an inverse vector, written $-v$, such that $v + (-v) = \vec{0}$
5. $+$ is commutative: for all $u, v \in V$ then $u + v = v + u$
6. For all $v \in V$ and $a \in \mathbb{F}$ then $av \in V$ is defined
7. For all $a \in \mathbb{F}$ and $u, v \in V$ then $a(u + v) = au + av$
8. For all $a, b \in \mathbb{F}$ and $v \in V$ then $(a + b)v = av + bv$ and $a(bu) = (ab)u$
9. For all $v \in V$ then $1v = v$, where $1 \in \mathbb{F}$ is the unit scalar.

Choice of scalars

Two common choices of scalar fields, $\mathbb{F}$, are the real numbers, $\mathbb{R}$, and the complex numbers, $\mathbb{C}$, giving rise to real and complex linear spaces, respectively.

The term vector space is a synonym for linear space.

Linear subspace

Definition (Linear subspace)

A subset $W \subset V$ is a linear subspace of V if the W is again a linear space over the same field of scalars.

Thus W is a linear subspace if $W \neq \emptyset$ and for all $u, v \in W$ and $a, b \in \mathbb{F}$ we have that $au + bv \in W$.

Linear combinations and spans

Definition (Linear combinations)

If V is a linear space and $v_1, v_2, \ldots, v_n \in V$ are vectors in V then $u \in V$ is a linear combination of $v_1, v_2, \ldots, v_n$ if there exist scalars $a_1, a_2, \ldots, a_n \in \mathbb{F}$ such that

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

We also define the span of a set of vectors as

$$\operatorname{span}\{v_1, v_2, \ldots, v_n\} = \{u \in V \,:\, u \text{ is a linear combination of } v_1, v_2, \ldots, v_n\} \,.$$

Thus, $W = \operatorname{span}\{v_1, v_2, \ldots, v_n\}$ is a linear subspace of V.

Linear independence

Definition (Linear independence)

Let V be a linear space. The vectors $v_1, v_2, \dots, v_n \in V$ are linearly independent if whenever

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

then $a_1 = a_2 = \cdots = a_n = 0$

The vectors $v_1, v_2, \dots, v_n$ are linearly dependent otherwise.

Bases

Definition (Basis)

A finite set of vectors $v_1, v_2, \ldots v_n \in V$ is a basis for the linear space V if $v_1, v_2, \ldots, v_n$ are linearly independent and $V = \text{span}\{v_1, v_2, \ldots, v_n\}$. The number n is called the dimension of V, written $n = \dim(V)$.

A result from linear algebra is that while there are infinitely many choices of basis vectors any two bases will always consist of the same number of element vectors. Thus, the dimension of a linear space is well-defined.

Inner products and inner product spaces

Suppose that V is either a real or complex linear space (that is, the scalars $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$).

Definition (Inner product)

The inner product of two vectors $u, v \in V$, written $\langle u, v \rangle \in \mathbb{F}$, is a scalar value satisfying

1. For each $v \in V$, $\langle v, v \rangle$ is a non-negative real number, so $\langle v, v \rangle \geq 0$
2. For each $v \in V$, $\langle v, v \rangle = 0$ if and only if $v = \vec{0}$
3. For all $u, v, w \in V$ and $a, b \in \mathbb{F}$, $\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$
4. For all $u, v \in V$ then $\langle u, v \rangle = \overline{\langle v, u \rangle}$.

A linear space together with an inner product is called an inner product space.
Here, $\overline{\langle v, u \rangle}$ denotes the complex conjugate of the complex number $\langle v, u \rangle$. Note that for a real linear space (so, $\mathbb{F} = \mathbb{R}$) the complex conjugate is redundant so the last condition above just says that $\langle u, v \rangle = \overline{\langle v, u \rangle} = \langle v, u \rangle$.

Useful properties of the inner product

Before looking at some examples of inner products there are several consequences of the definition of an inner product that are useful in calculations.

1. For all $v \in V$ and $a \in \mathbb{F}$ then $\langle av, av \rangle = |a|^2 \langle v, v \rangle$
2. For all $v \in V$, $\langle \vec{0}, v \rangle = 0$
3. For all $v \in V$ and finite sequences of vectors $u_1, u_2, \ldots, u_n \in V$ and scalars $a_1, a_2, \ldots, a_n$ then

$$\left\langle \sum_{i=1}^{n} a_i u_i, v \right\rangle = \sum_{i=1}^{n} a_i \langle u_i, v \rangle$$

$$\left\langle v, \sum_{i=1}^{n} a_i u_i \right\rangle = \sum_{i=1}^{n} \overline{a_i} \langle v, u_i \rangle$$

Inner product: examples

Example (Euclidean space, $\mathbb{R}^n$)

$V = \mathbb{R}^n$ with the usual operations of vector addition and multiplication by a real-valued scalar is a linear space over $\mathbb{R}$. Given two vectors $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$ we can define an inner product by

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

Often this inner product is known as the dot product and is written $x \cdot y$.

Example

Similarly, for $V = \mathbb{C}^n$, we can define an inner product by

$$\langle x, y \rangle = x \cdot y = \sum_{i=1}^{n} x_i \overline{y_i} \,.$$

Example (Space of continuous functions)

$V = C[a,b]$, the space of continuous functions $f : [a,b] \to \mathbb{C}$ with the standard operations of the sum of two functions and multiplication by a scalar, is a linear space over $\mathbb{C}$ and we can define an inner product for $f, g \in C[a,b]$ by

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

Norms

The concept of a norm is closely related to an inner product and we shall see that there is a natural way to define a norm given an inner product.

Definition (Norm)

Let V be a real or complex linear space so that, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A norm on V is a function from V to $\mathbb{R}_+$, written $||v||$, that satisfies

1. For all $v \in V$, $||v|| \geq 0$
2. $||v|| = 0$ if and only if $v = \vec{0}$
3. For each $v \in V$ and $a \in \mathbb{F}$, $||av|| = |a| \, ||v||$
4. For all $u, v \in V$, $||u + v|| \leq ||u|| + ||v||$ (the triangle inequality).

A norm can be thought of as a generalisation of the notion of distance, where for any two vectors $u, v \in V$ the number $||u - v||$ is the distance between u and v.

Norms: examples

Example (Eucidean norm)

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

$$||x|| = +\sqrt{\sum_{i=1}^{n} |x_i|^2}\,.$$

Example (Uniform norm)

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

$$||x||_\infty = \max\left\{|x_i| \,:\, i = 1, 2, \ldots, n\right\}\,.$$

Example (Uniform norm)

If $V = C[a, b]$ then for each function $f \in V$, define

$$||f||_\infty = \max\left\{|f(x)| \,:\, x \in [a, b]\right\}\,.$$

Cauchy-Schwarz inequality

Theorem (Cauchy-Schwarz inequality)

Let V be a real or complex inner product space with scalars $\mathbb{F}$ then for all $u, v \in V$

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle \langle v, v\rangle .$$

Proof.

If $v = \vec{0}$ then the result holds trivially. Now assume $v \neq \vec{0}$ so that $\langle v, v\rangle \neq 0$ and let $\lambda \in \mathbb{F}$ then

$$0 \leq \langle u - \lambda v, u - \lambda v\rangle = \langle u, u\rangle - \overline{\lambda}\langle u, v\rangle - \lambda\langle v, u\rangle + |\lambda|^2\langle v, v\rangle$$

Now set $\lambda = \frac{\langle u,v\rangle}{\langle v,v\rangle}$ so that

$$0 \leq \langle u, u\rangle - \frac{|\langle u, v\rangle|^2}{\langle v, v\rangle}$$

and hence

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle\langle v, v\rangle .$$

$\square$

Inner products and norms

Given an inner product space, V, with inner product $\langle \cdot, \cdot \rangle$ there is a natural choice of norm, namely, for all $v \in V$

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

Most of the properties that make this a norm follow simply from the properties of the inner product but we shall use the Cauchy-Schwarz inequality to establish the triangle inequality. We have,

$$\begin{aligned} ||u+v||^2 &= \langle u+v, u+v \rangle \\ &= ||u||^2 + \langle u, v \rangle + \langle v, u \rangle + ||v||^2 \\ &\leq ||u||^2 + 2|\langle u, v \rangle| + ||v||^2 \\ &\leq ||u||^2 + 2||u||\,||v|| + ||v||^2 \\ &= (||u|| + ||v||)^2\,. \end{aligned}$$

Hence, the triangle inequality, $||u+v|| \leq ||u|| + ||v||$ holds.

Orthogonal and orthonormal systems

Let V be an inner product space and take the natural choice of norm.

Definition (Orthogonality)

We say that $u, v \in V$ are orthogonal (written $u \perp v$) if $\langle u, v \rangle = 0$.

Definition (Orthogonal system)

A finite or infinite sequence of vectors (u_i) in V is an orthogonal system if

1. $u_i \neq \vec{0}$ for all such vectors u_i
2. $u_i \perp u_j$ for all $i \neq j$.

Definition (Orthonormal system)

An orthogonal system is called an orthonormal system if, in addition, $||u_i|| = 1$ for all such vectors u_i.

A vector $v \in V$ such that $||v|| = 1$ is called a unit vector.

Theorem

Suppose that $\{e_1, e_2, \ldots, e_n\}$ is an orthonormal system in the inner product space V. If $u = \sum_{i=1}^{n} a_i e_i$ then $a_i = \langle u, e_i \rangle$.

Proof.

$$\begin{aligned} \langle u, e_i \rangle &= \langle a_1 e_1 + a_2 e_2 + \cdots + a_n e_n, e_i \rangle \\ &= a_1 \langle e_1, e_i \rangle + a_2 \langle e_2, e_i \rangle + \cdots + a_n \langle e_n, e_i \rangle \\ &= a_i \, . \end{aligned}$$

$\square$

Hence, if $\{e_1, e_2, \ldots, e_n\}$ is an orthonormal system then for all $u \in \mathrm{span}\{e_1, e_2, \ldots, e_n\}$ we have

$$u = \sum_{i=1}^{n} a_i e_i = \sum_{i=1}^{n} \langle u, e_i \rangle e_i \, .$$

Fourier coefficients

Let V be an inner product space and $e_1, e_2, \ldots, e_n$ an orthonormal system (n being finite or infinite).

Definition (Generalized Fourier coefficients)

Given a vector $u \in V$, the scalars $\langle u, e_i \rangle$ ($i = 1, 2, \ldots, n$) are called the Generalized Fourier coefficients of u with respect to the given orthonormal system.

These coefficients are generalized in the sense that they refer to a general orthonormal system.

Let V be an inner product space and $e_1, e_2, \ldots, e_n$ an orthonormal system. If $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ are any sequences of scalars then

$$\left\langle \sum_{i=1}^{n} a_i e_i, \sum_{i=1}^{n} b_i e_i \right\rangle = \sum_{i=1}^{n} a_i \overline{b_i} \,.$$

Equivalently, for $u, v \in \operatorname{span}\{e_1, e_2, \ldots, e_n\}$

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

A consequence of these relations is the following theorem.

Theorem (Generalized Pythagorean Theorem)

Suppose that $\{u_1, u_2, \ldots, u_n\}$ is an orthogonal system in V and $a_1, a_2, \ldots, a_n$ are scalars then

$$|| \sum_{i=1}^{n} a_i u_i ||^2 = \sum_{i=1}^{n} |a_i|^2 \, ||u_i||^2 \,.$$

Orthogonal projections

Suppose that V is an inner product space and $e_1, e_2, \ldots, e_n$ an orthonormal system. Define $W = \text{span}\{e_1, e_2, \ldots, e_n\}$ and let $u \in V$ be any vector. We have seen that for $u \in W$

$$u = \sum_{i=1}^{n} \langle u, e_i \rangle e_i$$

but if $u \notin W$ then certainly

$$u \neq \sum_{i=1}^{n} \langle u, e_i \rangle e_i$$

since u is not a linear combination of the vectors $e_1, e_2, \ldots, e_n$. Nevertheless, there is a close connection between u and the expression $\sum_{i=1}^{n} \langle u, e_i \rangle e_i$.

Definition (Orthogonal projection)

For all $u \in V$ we define the orthogonal projection of u in W, $\tilde{u}$, by

$$\tilde{u} = \sum_{i=1}^{n} \langle u, e_i \rangle e_i \, .$$

Theorem

For each $u \in V$ and for all $w \in W$

1. $\langle u - \tilde{u}, w \rangle = 0$
2. $||u - w||^2 = ||u - \tilde{u}||^2 + ||\tilde{u} - w||^2.$

Proof

First $\langle u - \tilde{u}, e_j \rangle = 0$ for all $j = 1, 2, \ldots, n$ since

$$\begin{aligned}\langle u - \tilde{u}, e_j \rangle &= \langle u, e_j \rangle - \left\langle \sum_{i=1}^{n} \langle u, e_i \rangle e_i, e_j \right\rangle = \langle u, e_j \rangle - \sum_{i=1}^{n} \langle u, e_i \rangle \langle e_i, e_j \rangle \\ &= \langle u, e_j \rangle - \langle u, e_j \rangle \langle e_j, e_j \rangle = \langle u, e_j \rangle - \langle u, e_j \rangle = 0\,.\end{aligned}$$

So take any $w \in W$ with $w = \sum_{j=1}^{n} b_j e_j$ for some scalars $b_1, b_2, \ldots, b_n$ and

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

Now $(u - \tilde{u}) \perp w$ for all $w \in W$ and so since $\tilde{u} - w \in W$ $(u - \tilde{u}) \perp (\tilde{u} - w)$. Hence,

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

$\square$

Best approximation

Theorem

Let V be an inner product space and $\{e_1, e_2, \ldots, e_n\}$ an orthonormal system. Let $W = span\{e_1, e_2, \ldots, e_n\}$ and $u \in V$ be any vector then $\tilde{u} = \sum_{i=1}^{n} \langle u, e_i \rangle e_i$ is the closest vector to u in W. Moreover, $\tilde{u}$ is the unique such vector in W.

Proof.

For all $w \in W$,

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

and so $||u - \tilde{u}|| \leq ||u - w||$ for all $w \in W$.
To show uniqueness, suppose that $||u - \tilde{u}|| = ||u - w||$ for some $w \in W$ then $||\tilde{u} - w|| = 0$ and so $w = \tilde{u}$. □

Infinite orthonormal systems

We now consider the situation of an inner product space, V, with $\dim(V) = \infty$ and consider orthonormal systems $\{e_1, e_2, \dots\}$ consisting of infinitely many vectors.

Definition (Convergence in norm)

Let $\{u_1, u_2, \dots\}$ be an infinite sequence of vectors in the normed linear space V and let $\{a_1, a_2, \dots\}$ be a sequence of scalars. We say that the series

$$\sum_{n=1}^{\infty} a_n u_n$$

converges in norm to $w \in V$ if

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

Closure and completeness

Two further properties are defined for an infinite orthonormal system $\{e_1, e_2, \ldots\}$ in an inner product space V.

Definition (Closed)

The system is called closed in V if for all $u \in V$

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

Definition (Complete)

The system is called complete in V if the zero vector $u = \vec{0}$ is the only solution to the set of equations

$$\langle u, e_n \rangle = 0 \quad n = 1, 2, \ldots\,.$$

Remarks on closure and completeness

- It can be shown that a closed infinite orthonormal system $\{e_1, e_2, \ldots\}$ is necessarily complete (but not the converse).
- If a system is not closed then there must exist some $u \in V$ such that the linear combination

$$\sum_{n=1}^{m} \langle u, e_n \rangle e_n$$

cannot be made arbitrarily close to u, for all choices of m.
- If the system is closed it may still be that the required number of terms in the above linear combination for a “good” approximation is too great for practical purposes.
- Seeking alternative closed systems of orthonormal vectors may produce “better” approximations in the sense of requiring fewer terms for a given accuracy.

Fourier series

Representing functions

In seeking to represent functions as linear combinations of simpler functions we shall need to consider spaces of functions with closed orthonormal systems.

Definition (piecewise continuous)

A function is piecewise continuous if it is continuous, except at a finite number of points and at each such point of discontinuity, the right and left limits exists and are finite.

The space, E, of piecewise continuous functions $f : [-\pi, \pi] \to \mathbb{C}$ is seen to be a linear space, under the convention that we regard two functions in E as identical if they are equal at all but a finite number of points.
For $f, g \in E$, then

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

defines an inner product on E.

A closed infinite orthonormal system for E

An important result is that

$$\left\{ \frac{1}{\sqrt{2}}, \sin(x), \cos(x), \sin(2x), \cos(2x), \sin(3x), \cos(3x), \ldots \right\}$$

is a closed infinite orthonormal system in the space E.
Here we shall just demonstrate orthonormality and omit establishing that this system is closed.

Writing

$$||f|| = +\sqrt{< f, f >}$$

as the norm associated with our inner product, it can be establish that

$$||\frac{1}{\sqrt{2}}||^2 = 1$$

and similarily that for each $n = 1, 2, \ldots$

$$||\sin(nx)||^2 = ||\cos(nx)||^2 = 1$$

and that for $m, n \in \mathbb{N}$

- $\langle \frac{1}{\sqrt{2}}, \sin(nx) \rangle = 0$
- $\langle \frac{1}{\sqrt{2}}, \cos(nx) \rangle = 0$
- $\langle \sin(mx), \cos(nx) \rangle = 0$
- $\langle \sin(mx), \sin(nx) \rangle = 0$, $m \neq n$
- $\langle \cos(mx), \cos(nx) \rangle = 0$, $m \neq n$.

Fourier series

From our knowledge of closed orthonormal systems $\{e_1, e_2, \ldots\}$ we know that we can represent any function $f \in E$ by a linear combination

$$\sum_{n=1}^{\infty} \langle f, e_n \rangle e_n \,.$$

We now turn to consider the individual terms $\langle f, e_n \rangle e_n$ in the case of the closed orthonormal system

$$\left\{ \frac{1}{\sqrt{2}}, \sin(x), \cos(x), \sin(2x), \cos(2x), \sin(3x), \cos(3x), \ldots \right\} .$$

There are three cases, either $e_n = \frac{1}{\sqrt{2}}$ or $\sin(nx)$ or $\cos(nx)$. Recall that the vectors e_n are actually functions
in $E = \{f : [-\pi, \pi] \to \mathbb{C} : f \text{ is piecewise continuous}\}$

If $e_n = 1/\sqrt{2}$ then

$$\langle f, e_n \rangle e_n = \frac{1}{\pi} \left(\int_{-\pi}^{\pi} f(t) \frac{1}{\sqrt{2}} dt \right) \frac{1}{\sqrt{2}} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) dt \,.$$

If $e_n = \sin(nx)$ then

$$\langle f, e_n \rangle e_n = \frac{1}{\pi} \left(\int_{-\pi}^{\pi} f(t) \sin(nt) \, dt \right) \sin(nx) \,.$$

If $e_n = \cos(nx)$ then

$$\langle f, e_n \rangle e_n = \frac{1}{\pi} \left(\int_{-\pi}^{\pi} f(t) \cos(nt) \, dt \right) \cos(nx) \,.$$

Fourier coefficients

Thus the linear combination

$$\sum_{n=1}^{\infty} \langle f, e_n \rangle e_n$$

becomes the familiar Fourier series for a function f, namely

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)]$$

where

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx)\, dx, \quad n = 0, 1, 2, \ldots$$
$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx)\, dx, \quad n = 1, 2, 3, \ldots .$$

Note how the constant term is written $a_0/2$ where $a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx$.

Periodic functions

Our Fourier series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)]$$

defines a function, $g(x)$, say, that is 2π-periodic in the sense that

$$g(x + 2\pi) = g(x), \qquad \text{for all } x \in \mathbb{R}\,.$$

Hence, it is convenient to extend $f \in E$ to a 2π-periodic function defined on $\mathbb{R}$ instead of being restricted to $[-\pi, \pi]$.

Even and odd functions

A particularly useful simplification occurs when the function $f \in E$ is either an even function, that is, for all x,

$$f(-x) = f(x)$$

or an odd function, that is, for all x,

$$f(-x) = -f(x)\,.$$

The following properties can be easily verified.

1. If f, g are even then fg is even
2. If f, g are odd then fg is even
3. If f is even and g is odd then fg is odd
4. If g is odd then for any $h > 0$ then $\int_{-h}^{h} g(x)dx = 0$
5. If g is even then for any $h > 0$ then $\int_{-h}^{h} g(x)dx = 2\int_{0}^{h} g(x)dx$.

Even functions and cosine series

Recall that the Fourier coefficients are given by

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx, \quad n = 0, 1, 2, \ldots$$
$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx, \quad n = 1, 2, 3, \ldots$$

so if f is even then they become

$$a_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\cos(nx)\,dx, \quad n = 0, 1, 2, \ldots$$
$$b_n = 0, \quad n = 1, 2, 3, \ldots .$$

Odd functions and sine series

Similarly, the Fourier coefficients

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx)\, dx, \quad n = 0, 1, 2, \ldots$$
$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx)\, dx, \quad n = 1, 2, 3, \ldots,$$

for the case where f is an odd function become

$$a_n = 0, \quad n = 0, 1, 2, \ldots$$
$$b_n = \frac{2}{\pi} \int_{0}^{\pi} f(x) \sin(nx)\, dx, \quad n = 1, 2, 3, \ldots.$$

Fourier series: examples I

Consider $f(x) = x$ for $x \in [-\pi, \pi]$ then f is clearly odd and so we need to calculate a sine series with coefficients, b_n, $n = 1, 2, \ldots$ given by

$$\begin{aligned} b_n &= \frac{2}{\pi}\int_0^{\pi} x\sin(nx)\,dx = \frac{2}{\pi}\left\{\left[-x\frac{\cos(nx)}{n}\right]_0^{\pi} + \int_0^{\pi}\frac{\cos(nx)}{n}dx\right\} \\ &= \frac{2}{\pi}\left\{-\pi\frac{(-1)^n}{n} + \left[\frac{\sin(nx)}{n^2}\right]_0^{\pi}\right\} \\ &= \frac{2}{\pi}\left\{-\pi\frac{(-1)^n}{n} + 0\right\} = \frac{2(-1)^{n+1}}{n}\,. \end{aligned}$$

Hence the Fourier series of $f(x) = x$ is

$$\sum_{n=1}^{\infty}\frac{2(-1)^{n+1}}{n}\sin(nx)\,.$$

Observe that the series does not agree with $f(x)$ at $x = \pm\pi$ — a matter that we shall return to later.

Fourier series: examples II

Now suppose $f(x) = |x|$ for $x \in [-\pi, \pi]$ which is clearly an even function so we need to construct a cosine series with coefficients

$$a_0 = \frac{2}{\pi}\int_0^\pi x dx = \frac{2}{\pi}\frac{\pi^2}{2} = \pi$$

and for $n = 1, 2, \ldots$

$$a_n = \frac{2}{\pi}\int_0^\pi x\cos(nx)\,dx = \frac{2}{\pi}\left\{\left[\frac{x\sin(nx)}{n}\right]_0^\pi - \int_0^\pi \frac{\sin(nx)}{n}dx\right\}$$

$$= \frac{2}{\pi}\left\{\left[\frac{\cos(nx)}{n^2}\right]_0^\pi\right\} = \frac{2}{\pi}\left\{\frac{(-1)^n - 1}{n^2}\right\} = \begin{cases} -\frac{4}{\pi n^2} & n \text{ is odd} \\ 0 & n \text{ is even} \end{cases}.$$

Hence, the Fourier series of $f(x) = |x|$ is

$$\frac{\pi}{2} - \sum_{k=1}^{\infty} \frac{4}{\pi(2k-1)^2}\cos\left((2k-1)x\right).$$

Complex Fourier series I

We have used real-valued functions $\sin(nx)$ and $\cos(nx)$ as our orthonormal system for the linear space E but we can also use complex-valued functions. In this case, we should amend our inner product to

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx\,.$$

A suitable orthonormal system in this case is the collection of functions

$$\left\{1, e^{ix}, e^{-ix}, e^{i2x}, e^{-i2x}, \dots\right\}\,.$$

Then if $f \in E$ we have a representation, known as the complex Fourier series of $f \in E$, given by

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

where

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx, \quad n = 0, \pm 1, \pm 2, \dots\,.$$

Complex Fourier series II

Euler's formula ($e^{ix} = \cos(x) + i\sin(x)$) gives for $n = 1, 2, \ldots$ that

$$e^{inx} = \cos(nx) + i\sin(nx)$$
$$e^{-inx} = \cos(nx) - i\sin(nx)$$

and $e^{i0x} = 1$. Using these relations it can be shown that
for $n = 1, 2, \ldots$

$$c_n = \frac{a_n - ib_n}{2}, \qquad c_{-n} = \frac{a_n + ib_n}{2}.$$

Hence,

$$a_n = c_n + c_{-n}, \qquad b_n = i(c_n - c_{-n})$$

and

$$c_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-i0x}dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx = \frac{a_0}{2}.$$

Pointwise convergence and Dirichlet's conditions

The closure property of the trigonometric orthonormal system guarantees that the Fourier series for any function $f \in E$ converges in norm to f. That is,

$$\lim_{m\to\infty} ||f(x) - \left(\frac{a_0}{2} + \sum_{n=1}^{m} [a_n \cos(nx) + b_n \sin(nx)]\right)|| = 0$$

or, equivalently,

$$\lim_{m\to\infty} \int_{-\pi}^{\pi} \left| f(x) - \left(\frac{a_0}{2} + \sum_{n=1}^{m} [a_n \cos(nx) + b_n \sin(nx)]\right)\right|^2 dx = 0\,.$$

As we have already seen in the example of $f(x) = x$, this does not imply convergence to $f(x)$ at every point x.

The Dirichlet conditions

We now consider conditions on the space of functions that allow us to determine how the Fourier series behaves at individual points x.

Definition (Dirichlet conditions)

We define a subspace, E', of E by the Dirichlet conditions:

1. $f \in E$
2. For all $x \in [-\pi, \pi)$ both the left and right derivatives exist (and are finite).

Recall, that in the space E each function has a left and right limit at every point. Let these values be $f(x-)$ and $f(x+)$, respectively.

Theorem (Dirichlet's theorem)

For all $x \in [-\pi, \pi]$ the Fourier series of a function $f \in E'$ converges to the value of the expression

$$\frac{f(x-) + f(x+)}{2} .$$

- Here we should consider f not just defined on $[-\pi, \pi]$ but also make it 2π-periodic to handle the end points $\pm\pi$ correctly.
- Recall that functions $f \in E$ can have at most a finite number of points of discontinuity (that is, points where $f(x-)$ and $f(x+)$ differ).
- Hence, we can conclude that if a function f satisfies the Dirichlet conditions then the function's Fourier series converges to f at all points where f is continuous and at points of discontinuity it converges to the average of the left and right hand limits. This was indeed the case in our earlier example where $f(x) = x$.

General intervals

We have so far considered functions defined on the interval $[-\pi, \pi]$ but we may readily extend our approach to a general interval of the form $[a, b]$ (for any $a < b$). If we define $E[a, b]$ to be the space of piecewise continuous functions $f : [a, b] \to \mathbb{C}$ then we may define the Fourier series of $f \in E[a, b]$ as

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{2n\pi x}{(b-a)}\right) + b_n \sin\left(\frac{2n\pi x}{(b-a)}\right) \right]$$

where

$$a_n = \frac{2}{(b-a)} \int_a^b f(x) \cos\left(\frac{2n\pi x}{(b-a)}\right) dx, \quad n = 0, 1, 2, \ldots$$

$$b_n = \frac{2}{(b-a)} \int_a^b f(x) \sin\left(\frac{2n\pi x}{(b-a)}\right) dx, \quad n = 1, 2, 3, \ldots .$$

This may be justified by showing, for example, that

$$\left\{ \frac{1}{\sqrt{2}}, \cos\left(\frac{2n\pi x}{(b-a)}\right), \sin\left(\frac{2n\pi x}{(b-a)}\right) \quad \text{for } n = 1, 2, \ldots \right\}$$

is an infinite orthonormal system for functions in $E[a, b]$ with respect to the inner product

$$\langle f, g \rangle = \frac{2}{(b-a)} \int_a^b f(x)\overline{g(x)}dx \, .$$

Exercise: establish the corresponding details for the case of the complex Fourier series representation and a general interval $[a, b]$.

Fourier transforms

Introduction

- We have seen how functions $f : [-\pi, \pi] \to \mathbb{C}, f \in E$ can be represented in alternative ways using closed orthonormal systems, such as
$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$
where
$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \quad n = 0, \pm 1, \pm 2, \dots .$$
The domain $[-\pi, \pi]$ can be swapped for a general interval $[a, b]$ and the function can be regarded as L-periodic and defined for all $\mathbb{R}$, where $L = (b - a) < \infty$ is the length of the interval.
- We shall now consider the situation where $f : \mathbb{R} \to \mathbb{C}$ may be a non-periodic function.

Fourier transform

Definition (Fourier transform)

For $f : \mathbb{R} \to \mathbb{C}$ define the Fourier transform of f to be the function $F : \mathbb{R} \to \mathbb{C}$ given by

$$F(\omega) = \mathcal{F}_{[f]}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx$$

whenever the integral exists.

We shall use the notation $F(\omega)$ or $\mathcal{F}_{[f]}(\omega)$ as convenient. The notation $\hat{f}(\omega)$ is also seen widely in the literature.

For functions $f : \mathbb{R} \to \mathbb{C}$ define the two properties

1. piecewise continuous: if f is piecewise continuous on every finite interval. Thus f may have an infinite number of discontinuities but only a finite number in any subinterval.
2. absolutely integrable: if

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty$$

Let $G(\mathbb{R})$ be the collection of all functions $f : \mathbb{R} \to \mathbb{C}$ that are piecewise continuous and absolutely integrable.

Immediate properties

It may be shown that $G(\mathbb{R})$ is a linear space over the scalars $\mathbb{C}$ and that for $f \in G(\mathbb{R})$

1. $F(\omega)$ is defined for all $\omega \in \mathbb{R}$
2. F is a continuous function
3. $\lim_{\omega \to \pm\infty} F(\omega) = 0$

Examples

For $a > 0$, let $f(x) = e^{-a|x|}$ then

$$\begin{aligned}
F(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-a|x|}e^{-i\omega x}dx \\
&= \frac{1}{2\pi}\left\{\int_{0}^{\infty} e^{-ax}e^{-i\omega x}dx + \int_{-\infty}^{0} e^{ax}e^{-i\omega x}dx\right\} \\
&= \frac{1}{2\pi}\left\{-\left[\frac{e^{-(a+i\omega)x}}{a+i\omega}\right]_0^{\infty} + \left[\frac{e^{(a-i\omega)x}}{a-i\omega}\right]_{-\infty}^{0}\right\} \\
&= \frac{1}{2\pi}\left\{\frac{1}{a+i\omega} + \frac{1}{a-i\omega}\right\} \\
&= \frac{a}{\pi(a^2+\omega^2)}\,.
\end{aligned}$$

Properties

Several properties of the Fourier transform are very helpful in calculations.
First, note that by the linearity of integrals we have that if $f, g \in G(\mathbb{R})$ and $a, b \in \mathbb{C}$ then

$$\mathcal{F}_{[af+bg]}(\omega) = a\mathcal{F}_{[f]}(\omega) + b\mathcal{F}_{[g]}(\omega)$$

and $af + bg \in G(\mathbb{R})$.
Secondly, if f is real-valued then

$$F(-\omega) = \overline{F(\omega)}\,.$$

Even and odd real-valued functions

Theorem

If $f \in G(\mathbb{R})$ is an even real-valued function then F is even and real-valued. If f is an odd real-valued function then F is odd and purely imaginary.

Proof.

Suppose that f is even and real-valued then

$$\begin{aligned} F(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-i\omega x}dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\left[\cos(\omega x) - i\sin(\omega x)\right]dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\cos(\omega x)dx\,. \end{aligned}$$

Hence, F is real-valued and even (the imaginary part has vanished and both f and $\cos(\omega x)$ are themselves even functions). The second part follows similarly. □

Shift and scale properties

Theorem

Let $f \in G(\mathbb{R})$ *and* $a, b \in \mathbb{R}$ *with* $a \neq 0$ *and define*

$$g(x) = f(ax + b)$$

then $g \in G(\mathbb{R})$ *and*

$$\mathcal{F}_{[g]}(\omega) = \frac{1}{|a|} e^{i\omega b/a} \mathcal{F}_{[f]}\left(\frac{\omega}{a}\right)$$

Proof

Set $y = ax + b$ so for $a > 0$ then

$$\mathcal{F}_{[g]}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) e^{-i\omega\left(\frac{y-b}{a}\right)} \frac{dy}{a}$$

and for $a < 0$

$$\mathcal{F}_{[g]}(\omega) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) e^{-i\omega\left(\frac{y-b}{a}\right)} \frac{dy}{a}\,.$$

Hence,

$$\mathcal{F}_{[g]}(\omega) = \frac{1}{|a|} e^{i\omega b/a} \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) e^{-i\omega y/a} dy = \frac{1}{|a|} e^{i\omega b/a} \mathcal{F}_{[f]}\left(\frac{\omega}{a}\right)\,.$$

$\square$

Special cases

Two special cases are worth highlighting.

1. Suppose that $b = 0$ so $g(x) = f(ax)$ and so

$$\mathcal{F}_{[g]}(\omega) = \frac{1}{|a|}\mathcal{F}_{[f]}\left(\frac{\omega}{a}\right) .$$

2. Suppose that $a = 1$ so $g(x) = f(x + b)$ and so

$$\mathcal{F}_{[g]}(\omega) = e^{i\omega b}\mathcal{F}_{[f]}(\omega) .$$

Theorem

For $f \in G(\mathbb{R})$ and $c \in \mathbb{R}$ then

$$\mathcal{F}_{[e^{icx} f(x)]}(\omega) = \mathcal{F}_{[f]}(\omega - c)\,.$$

Proof.

$$\begin{aligned}
\mathcal{F}_{[e^{icx} f(x)]}(\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{icx} f(x) e^{-i\omega x} dx \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-i(\omega - c)x} dx \\
&= \mathcal{F}_{[f]}(\omega - c)\,.
\end{aligned}$$

$\square$

Modulation property

Theorem

For $f \in G(\mathbb{R})$ and $c \in \mathbb{R}$ then

$$\mathcal{F}_{[f(x)\cos(cx)]}(\omega) = \frac{\mathcal{F}_{[f]}(\omega - c) + \mathcal{F}_{[f]}(\omega + c)}{2}$$
$$\mathcal{F}_{[f(x)\sin(cx)]}(\omega) = \frac{\mathcal{F}_{[f]}(\omega - c) - \mathcal{F}_{[f]}(\omega + c)}{2i}\,.$$

Proof.

We have that

$$\begin{aligned}\mathcal{F}_{[f(x)\cos(cx)]}(\omega) &= \mathcal{F}_{\left[f(x)\frac{e^{icx}+e^{-icx}}{2}\right]}(\omega)\\ &= \frac{1}{2}\mathcal{F}_{[f(x)e^{icx}]}(\omega) + \frac{1}{2}\mathcal{F}_{[f(x)e^{-icx}]}(\omega)\\ &= \frac{\mathcal{F}_{[f]}(\omega - c) + \mathcal{F}_{[f]}(\omega + c)}{2}\,.\end{aligned}$$

Similarly, for $\mathcal{F}_{[f(x)\sin(cx)]}(\omega)$. □

Derivatives

There are further properties relating to the Fourier transform of derivatives that we shall state here but omit further proofs.

Theorem

If f is such that both $f, f' \in G(\mathbb{R})$ then

$$\mathcal{F}_{[f']}(\omega) = i\omega\mathcal{F}_{[f]}(\omega)\,.$$

Inverse Fourier transform

We have studied the Fourier transform. There is also an inverse operation of recovering a function f given the function $F(\omega) = \mathcal{F}_{[f]}(\omega)$ which takes the form

$$f(x) = \int_{-\infty}^{\infty} \mathcal{F}_{[f]}(\omega) e^{i\omega x} d\omega\,.$$

More precisely, and recalling Dirichlet's theorem for Fourier series, the following holds.

Theorem (Inverse Fourier transform)

If $f \in G(\mathbb{R})$ then for every point $x \in \mathbb{R}$ where the one-sided derivatives exist

$$\frac{f(x-) + f(x+)}{2} = \lim_{M\to\infty} \int_{-M}^{M} \mathcal{F}_{[f]}(\omega) e^{i\omega x} d\omega\,.$$

Convolution

An important operation between two functions in signal processing applications is convolution defined as follows.

Definition (Convolution)

If f and g are two functions $\mathbb{R} \to \mathbb{C}$ then the convolution function, written $f * g$, is given by

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy$$

whenever the integral exists.

Exercise: show that the convolution operation is commutative, that is $f * g = g * f$.

Fourier transforms and convolutions

The importance of Fourier transform techniques in signal processing rests, in part, on the following result that leads to much simpler descriptions and mathematical formulae in the Fourier domain.

Theorem (Convolution theorem)

For $f, g \in G(\mathbb{R})$ *then*

$$\mathcal{F}_{[f*g]}(\omega) = 2\pi \mathcal{F}_{[f]}(\omega) \cdot \mathcal{F}_{[g]}(\omega)\,.$$

Proof

We have that

$$\begin{aligned}
\mathcal{F}_{[f*g]}(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}(f*g)(x)e^{-i\omega x}dx \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}f(x-y)g(y)dy\right)e^{-i\omega x}dx \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x-y)e^{-i\omega(x-y)}g(y)e^{-i\omega y}dxdy \\
&= \int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}f(x-y)e^{-i\omega(x-y)}dx\right)g(y)e^{-i\omega y}dy \\
&= \mathcal{F}_{[f]}(\omega)\int_{-\infty}^{\infty}g(y)e^{-i\omega y}dy \\
&= 2\pi\mathcal{F}_{[f]}(\omega)\cdot\mathcal{F}_{[g]}(\omega)\,.
\end{aligned}$$

□

Some signal processing applications

We first note two types of limitations on functions.

Definition (Time-limited)

A function f is time-limited if

$$f(x) = 0 \quad \text{for all } |x| \geq M$$

for some constant M.

Definition (Band-limited)

A function $f \in G(\mathbb{R})$ is band-limited if

$$\mathcal{F}_{[f]}(\omega) = 0 \quad \text{for all } |\omega| \geq L$$

for some constant L.

Let us first calculate the Fourier transform of

$$f(x) = \begin{cases} 1 & a \leq x \leq b \\ 0 & \text{otherwise}. \end{cases}$$

We have that

$$F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx = \frac{1}{2\pi} \int_a^b e^{-i\omega x} dx .$$

So, for $\omega \neq 0$,

$$F(\omega) = \left[\frac{1}{2\pi} \left(\frac{e^{-i\omega x}}{-i\omega} \right) \right]_a^b = \frac{e^{-i\omega a} - e^{-i\omega b}}{2\pi i\omega} .$$

However, for $\omega = 0$ we have that $F(0) = \frac{1}{2\pi} \int_a^b dx = \frac{(b-a)}{2\pi}$.
For the special case when $a = -b$ with $b > 0$ then

$$F(\omega) = \begin{cases} \frac{e^{i\omega b} - e^{-i\omega b}}{2\pi i\omega} = \frac{\sin(\omega b)}{\omega\pi} & \omega \neq 0 \\ \frac{b}{\pi} & \omega = 0 . \end{cases}$$

Low-pass filters

Suppose that $f \in G(\mathbb{R})$ with Fourier transform $F(\omega)$ and choose a positive constant $L > 0$. Define

$$F_L(\omega) = \begin{cases} F(\omega) & |\omega| \leq L \\ 0 & |\omega| > L\,. \end{cases}$$

We wish to find f_L such that $\mathcal{F}_{[f_L]} = F_L$, that is, a function band-limited by L whose Fourier transform equals F in $[-L, L]$.
Rewrite $F_L(\omega) = F(\omega)G_L(\omega)$ where

$$G_L(\omega) = \begin{cases} 1 & |\omega| \leq L \\ 0 & |\omega| > L\,. \end{cases}$$

We will now use the convolution theorem to find f_L.

By the inverse transform theorem we have that for $|x| \neq L$

$$G_L(x) = \int_{-\infty}^{\infty} \frac{\sin \omega L}{\omega \pi} e^{i\omega x} d\omega$$

But G_L is clearly an even function so

$$G_L(x) = G_L(-x) = \int_{-\infty}^{\infty} \frac{\sin \omega L}{\omega \pi} e^{-i\omega x} d\omega$$

and if we interchange the variables x and ω we have

$$G_L(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{2 \sin Lx}{x} e^{-i\omega x} dx \, .$$

This says that if $g_L(x) = \frac{2 \sin Lx}{x}$ then $\mathcal{F}_{[g_L]}(\omega) = G_L(\omega)$.

In terms of convolutions we have

$$
\begin{aligned}
f_L &= \frac{1}{2\pi}(f * g_L) \\
f_L(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)\frac{2\sin(L(x-y))}{x-y}dy \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(y)\sin(L(x-y))}{x-y}dy
\end{aligned}
$$

In particular, if $f \in G(\mathbb{R})$ is such that $\mathcal{F}_{[f]}(\omega) = 0$ for $|\omega| \geq L$ then f satisfies

$$
f(x) = f_L(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(y)\sin(L(x-y)))}{x-y}dy\,.
$$

Shannon sampling theorem

Theorem (Shannon sampling theorem)

If $f \in G(\mathbb{R})$ is band-limited by the constant L then

$$f(x) = \sum_{n=-\infty}^{\infty} f\left(\frac{n\pi}{L}\right) \frac{\sin(Lx - n\pi)}{Lx - n\pi}\,.$$

Proof

Set $F(\omega) = \mathcal{F}_{[f]}(\omega)$ and use the inverse Fourier transform theorem to give

$$f(x) = \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega = \int_{-L}^{L} F(\omega) e^{i\omega x} d\omega\,.$$

So, taking $x = \frac{n\pi}{L}$ for $n \in \mathbb{Z}$ we get

$$f\left(\frac{n\pi}{L}\right) = \int_{-L}^{L} F(\omega) e^{i\omega n\pi/L} d\omega\,.$$

Consider the complex Fourier series of $F(\omega)$ restricted to $\omega \in [-L, L]$ given by

$$\sum_{n=-\infty}^{\infty} c_n e^{-in\pi\omega/L}$$

where the coefficients, c_n, are

$$c_n = \langle F, e^{-in\pi\omega/L} \rangle = \frac{1}{2L} \int_{-L}^{L} F(\omega) e^{in\pi\omega/L} d\omega = \frac{1}{2L} f\left(\frac{n\pi}{L}\right)$$

Thus, since f is band-limited by L

$$F(\omega) = \left(\sum_{n=-\infty}^{\infty} c_n e^{-in\pi\omega/L} \right) G_L(\omega) .$$

Hence,

$$F(\omega) = \frac{1}{2L} \sum_{n=-\infty}^{\infty} f\left(\frac{n\pi}{L}\right) \left(e^{-in\pi\omega/L} G_L(\omega) \right) .$$

But we have seen that $G_L(\omega) = \mathcal{F}_{[\frac{2\sin Lx}{x}]}(\omega)$ hence using the shift formula

$$e^{-in\pi\omega/L} G_L(\omega) = \mathcal{F}_{[g_{L,n}]}(\omega)$$

where

$$g_{L,n}(x) = \frac{2\sin(Lx - n\pi)}{x - \frac{n\pi}{L}} .$$

Putting this all together we have that

$$F(\omega) = \frac{1}{2L} \sum_{n=-\infty}^{\infty} f\left(\frac{n\pi}{L}\right) \mathcal{F}_{[g_{L,n}]}(\omega)$$

and taking inverse transforms

$$f(x) = \frac{1}{2L} \sum_{n=-\infty}^{\infty} f\left(\frac{n\pi}{L}\right) g_{L,n}(x) = \sum_{n=-\infty}^{\infty} f\left(\frac{n\pi}{L}\right) \frac{\sin(Lx - n\pi)}{Lx - n\pi} .$$

Remarks on Shannon's sampling theorem

- The theorem says that band-limited functions by a constant L (that is, $\mathcal{F}_{[f]}(\omega) = 0$ for $|\omega| > L$) are completely determined by their values at evenly spaced points a distance $\frac{\pi}{L}$ apart.
- Moreover, we may recover the function exactly given only it's values at this sequence of points.
- It may be shown that the functions

$$\frac{\sin(Lx - n\pi)}{Lx - n\pi}$$

for $n \in \mathbb{Z}$ form an orthonormal system with inner product

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

Discrete Fourier Transforms

We now shift attention from functions defined on intervals or on the whole of $\mathbb{R}$ to sequences of values $f[0], f[1], \ldots, f[N-1]$ and consider how we might represent them.

An important result in this area of discrete transforms is that the vectors $\{e_0, e_1, \ldots, e_{N-1}\}$ form an orthogonal system in the space $\mathbb{C}^N$ with the usual inner product where the n^{th} component of e_k is given by

$$(e_k)_n = e^{2\pi i n k/N} \quad n = 0, 1, 2, \ldots, N-1 \,.$$

and $k = 0, 1, 2, \ldots, N-1$.

Applying the usual inner product

$$\langle u, v\rangle = \sum_{n=0}^{N-1} u[n]\overline{v[n]}$$

we shall see that

$$||e_k||^2 = \langle e_k, e_k\rangle = N\,.$$

In fact, using $\{e_0, e_1, \ldots, e_{N-1}\}$ we can represent any sequence $f = (f[0], f[1], \ldots, f[N-1]) \in \mathbb{C}^N$ by

$$f = \frac{1}{N}\sum_{k=0}^{N-1}\langle f, e_k\rangle e_k\,.$$

Recall the generalized Fourier coefficients that we studied earlier.

Orthogonality

We shall show orthogonality of the vectors e_k by considering the N distinct complex roots of the equation $z^N = 1$. Put $w = e^{2\pi i/N}$ then the N distinct roots z_j ($j = 0, 1, \dots, N-1$) of $z^N = 1$ are

$$z_j = e^{2\pi i j/N} = w^j \,.$$

Now for an arbitrary integer n

$$\begin{aligned}
\frac{1}{N}\sum_{k=0}^{N-1} e^{2\pi i n k/N} &= \frac{1}{N}\sum_{k=0}^{N-1} w^{nk} \\
&= \begin{cases} 1 & \text{if } n \text{ is an integer multiple of } N \\ \frac{1}{N}\frac{1-w^{nN}}{1-w^n} = 0 & \text{otherwise}\,. \end{cases}
\end{aligned}$$

Thus,

$$\begin{aligned}
\langle e_a, e_b \rangle &= \sum_{k=0}^{N-1} e^{2\pi ika/N} e^{-2\pi ikb/N} \\
&= \sum_{k=0}^{N-1} e^{2\pi ik(a-b)/N} \\
&= \begin{cases} N & \text{if } (a-b) \text{ is a multiple of } N \\ 0 & \text{otherwise}\,. \end{cases}
\end{aligned}$$

So, indeed, we have that

$$||e_k||^2 = \langle e_k, e_k \rangle = N\,.$$

Definition (Discrete Fourier Transform/DFT)

The sequence $F[k]$, $k \in \mathbb{Z}$, defined by

$$F[k] = \langle f, e_k \rangle = \sum_{n=0}^{N-1} f[n] e^{-2\pi i n k / N}$$

is called the N-point Discrete Fourier Transform of $f[n]$

Thus, for $n = 0, 1, 2, \ldots, N-1$, we have the inverse transform

$$f[n] = \frac{1}{N} \sum_{k=0}^{N-1} F[k] e^{2\pi i n k / N}.$$

Periodicity

Note that the sequence $F[k]$ has period N since

$$F[k+N] = \sum_{n=0}^{N-1} f[n] e^{-2\pi i n(k+N)/N} = \sum_{n=0}^{N-1} f[n] e^{-2\pi i n k/N} = F[k]$$

using the relation

$$e^{-2\pi i n(k+N)/N} = e^{-2\pi i n k/N} e^{-2\pi i n} = e^{-2\pi i n k/N}.$$

Properties of the DFT

The DFT satisfies a range of similar properties to those of the FT relating to linearity, and shifts in either the n or k domain. However, the convolution operation is defined a little differently.

Definition (Cyclical convolution)

The cyclical convolution of two periodic sequences $f[n]$ and $g[n]$ of period N is defined as

$$(f * g)[n] = \sum_{m=0}^{N-1} f[m]g[n-m]\,.$$

It can then be shown that the DFT of $f * g$ is the product $F[k]G[k]$ where F and G are the DFTs of f and g, respectively.

Fast Fourier Transform algorithm

Fast Fourier Transform

The Fast Fourier Transform is not a new transform but a particular numerical algorithm for computing the DFT.
Since

$$\begin{aligned} F[k] &= \sum_{n=0}^{N-1} f[n] e^{-2\pi i n k/N} \\ &= f[0] + f[1] e^{-2\pi i k/N} + \cdots + f[N-1] e^{-2\pi i k(N-1)/N} \end{aligned}$$

we can see that in order to compute $F[k]$ we need to do about $2N$ (complex) additions and multiplications. To compute $F[k]$ in this way for all $k = 0, 1, 2, \ldots, N-1$ would require about $2N^2$ such operations. In practice, where DFTs are computed for a large number of points N, faster algorithms have been developed. Most approaches are based on the factorization of N into prime factors and are known collectively as Fast Fourier Transforms (FFT). In most popular methods N is supposed to be a power of 2.

Fast algorithms for the DFT

In 1965, James W. Cooley and John W. Tukey published a new and substantially faster algorithm for computing the DFT than the direct N^2 approach.

They showed that when N is a composite number with $N = P_1 P_2 \cdot P_m$ then it is possible to reduce the cost of computing the DFT of a vector of length N from

$$N^2 = N(P_1 P_2 \cdots P_m) \quad \text{to} \quad N((P_1 - 1) + (P_2 - 1) + \cdots + (P_m - 1))$$

complex operations. In the case when $P_1 = P_2 = \cdots = P_m = 2$ then this reduces from $N^2 = 2^{2m}$ to $2^m \cdot m = N \log_2 N$.

For example, if $N = 1024 = 2^{10}$ then there is a roughly a 100 fold improvement from $N^2 = 1,048,576$ down to $N \log_2 N = 10,240$.

See: J.W. Cooley and J.W. Tukey. (1965) An algorithm for the machine computation of complex Fourier series, *Math. Comp*, 19, 297–301.

We shall not derive any of the details here but instead give an impression of how the method operates.

First, the task of computing the DFT can be represented with matrices as

$$F = Af$$

but where the $N \times N$ matrix, A, has a great deal of internal structure. Cooley and Tukey exploited this structure in the case when $N = 2^m$ (so $m = \log_2 N$) to rewrite A as a product of matrices each of which is sparse

$$A = M_m M_{m-1} \cdots M_1 B\,.$$

Since each of these matrices contains only a small number of non-zero entries the effective number of complex operations is much reduced compared to working with A itself.

Wavelet Transforms

Wavelets

Wavelets are a further method of representing functions that has received much interest in applied fields over the last several decades. The approach fits into the general scheme of expansion using orthonormal functions. Here we expand functions $f(x)$ in terms of a doubly-infinite series

$$f(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_{jk} \Psi_{jk}(x)$$

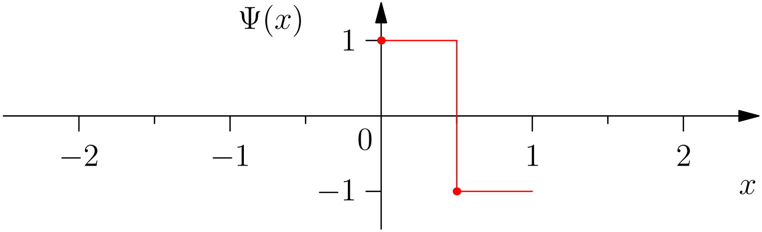
where $\Psi_{jk}(x)$ are the orthonormal functions.
The orthonormal functions arise from shifting and scaling operations applied to a single function, $\Psi(x)$, known as the mother wavelet.
The orthonormal functions are given for integers j and k by

$$\Psi_{jk}(x) = 2^{j/2}\Psi(2^j x - k)$$

The Haar wavelet

A common example is the Haar wavelet whose mother function is both localised and oscillatory defined by

$$\Psi(x) = \begin{cases} 1 & \text{if } 0 \le x < \frac{1}{2}\,, \\ -1 & \text{if } \frac{1}{2} \le x < 1\,, \\ 0 & \text{otherwise}\,. \end{cases}$$



Wavelet dilations and translations

The Haar mother wavelet oscillates and has a width (or scale) of one. The dyadic dilates of $\Psi(x)$, namely,

$$\ldots, \Psi(2^{-2}x), \Psi(2^{-1}x), \Psi(x), \Psi(2x), \Psi(2^2x), \ldots$$

have widths

$$\ldots, 2^2, 2^1, 1, 2^{-1}, 2^{-2}, \ldots$$

respectively. Since the dilate $\Psi(2^j x)$ has width 2^{-j}, its translates

$$\Psi(2^j x - k) = \Psi(2^j(x - k2^{-j})), \quad k = 0, \pm 1, \pm 2, \ldots$$

will cover the whole x-axis. The collection of coefficients d_{jk} are termed the Discrete wavelet transform, or DWT, of the function $f(x)$. Just as with Fourier transforms there are fast implementations that exploit structure.

Interpretation of d_{jk}

How should we intrepret the values d_{jk}?
Since the Haar wavelet function $\Psi(2^j x - k)$ vanishes except when

$$0 \le 2^j x - k < 1\,, \quad \text{that is} \quad k2^{-j} \le x < (k+1)2^{-j}$$

we see that d_{jk} gives us information about the behaviour of f near the point $x = k2^{-j}$ measured on the scale of 2^{-j}.
For example, the coefficients $d_{-10,k}$, $k = 0, \pm 1, \pm 2, \ldots$ correspond to variations of f that take place over intervals of length $2^{10} = 1024$ while the coefficients $d_{10,k}$ $k = 0, \pm 1, \pm 2, \ldots$ correspond to fluctuations of f over intervals of length 2^{-10}.
These observations help explain how the discrete wavelet transform can be an exceptionally efficient scheme for representing functions.

Comparison with Fourier analysis

Some of the practical motivations underlying the use of the orthonormal functions such as Fourier analysis or wavelet analysis are

- improved understanding,
- denoising signals, and
- data compression.

By representation of signals or functions in other forms these tasks become easier or more effective.
The approach taken with Fourier analysis represents signals in terms of trigonometric functions and as such is particularly suited to situations where the signal is relatively smooth and is not of limited extent.

Properties of naturally arising data

Much naturally arising data has been found to be better represented using wavelets which are better able to cope with discontinuities and where the signal is of local extent. Generally, the efficiency of the representation depends on the types of signal involved. If your signal contains

- discontinuities (in both the signal and its derivatives), or
- varying frequency behaviour

then wavelets are likely to represent the signal more efficiently than is possible with Fourier analysis.

Other classes of wavelets

- One of the most useful features of wavelets is the ease with which a scientist can select the wavelet functions adapted for the given problem.
- In fact, the Haar mother wavelet is perhaps the simplest of a very wide class of possible wavelet systems used in practice today.
- Many applied fields have started to make use of wavelets including astronomy, acoustics, signal and image processing, neurophysiology, music, magnetic resonance imaging, speach discrimination, optics, fractals, turbulence, earthquake prediction, radar, human vision, etc.