Inner Product Spaces

Reference §7.1, 7.2 in Poole.

An arbitrary vector space V does not have any operation that mimics the dot product on $\mathbf{R}^n$. Thus, on V, we have no idea of "length" or "angles"; in particular, there is no concept of orthogonality. These concepts have proven useful so we wish to introduce them on a vector space if possible. We first need to abstract the (relevant) properties of the dot product. It should give a real number (and the dot product of a vector with itself should be positive unless the vector is the zero vector). It should also respect the vector space operations of addition and scalar multiplication.

Inner Product Spaces

Definition 15.1

Let V be a real vector space. An inner product on V is an operation that assigns, to every pair of vectors $\mathbf{u}$ and $\mathbf{v} \in V$, a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ such that

$$\begin{aligned} &\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \\ &\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle \\ &\langle c\,\mathbf{u}, \mathbf{v} \rangle = c\,\langle \mathbf{u}, \mathbf{v} \rangle \\ &\langle \mathbf{u}, \mathbf{u} \rangle \geqslant 0 \text{ and } \langle \mathbf{u}, \mathbf{u} \rangle = 0 \text{ if and only if } \mathbf{u} = \mathbf{0} \end{aligned}$$

for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and scalars c. A vector space with an inner product is called an inner product space.

Note that the definition for a complex vector space is different. An inner product is structure that is additional to that present for all vector spaces.

Inner Product Spaces

$\mathbf{R}^n$ is an inner product space with the dot product; that is

$$\langle\, \mathbf{u},\, \mathbf{v}\, \rangle = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^{\mathrm{T}}\, \mathbf{v}.$$

However this is not the only inner product that can be defined on $\mathbf{R}^n$.

Example

Show that the weighted dot product

$$\langle\, \mathbf{u},\, \mathbf{v}\, \rangle = w_1\, u_1\, v_1 + w_2\, u_2\, v_2 + \cdots + w_n\, u_n\, v_n$$

where $w_1,\ w_2,\ \ldots,\ w_n$ are positive scalars defines an inner product on $\mathbf{R}^n$.

Inner Product Spaces

Solution: Note that we can write this inner product as

$$\langle\, \mathbf{u},\, \mathbf{v}\, \rangle = \mathbf{u}^{\mathrm{T}}\, \mathrm{D}\, \mathbf{v}$$

where D is the diagonal matrix

$$\mathrm{D} = \begin{bmatrix} w_1 & & \\ & \ddots & \\ & & w_n \end{bmatrix}.$$

Thus

$$\langle\, \mathbf{u},\, \mathbf{v}\, \rangle = \mathbf{u}^{\mathrm{T}}\, \mathrm{D}\, \mathbf{v} = (\mathbf{u}^{\mathrm{T}}\, \mathrm{D}\, \mathbf{v})^{\mathrm{T}} = \mathbf{v}^{\mathrm{T}}\, \mathrm{D}^{\mathrm{T}}\, \mathbf{u} = \mathbf{v}^{\mathrm{T}}\, \mathrm{D}\, \mathbf{u} = \langle\, \mathbf{v},\, \mathbf{u}\, \rangle$$

since D is diagonal. Also

$$\langle\, \mathbf{u},\, \mathbf{v} + \mathbf{w}\, \rangle = \mathbf{u}^{\mathrm{T}}\, \mathrm{D}\, (\mathbf{v} + \mathbf{w}) = \mathbf{u}^{\mathrm{T}}\, \mathrm{D}\, \mathbf{v} + \mathbf{u}^{\mathrm{T}}\, \mathrm{D}\, \mathbf{w} = \langle\, \mathbf{u},\, \mathbf{v}\, \rangle + \langle\, \mathbf{u},\, \mathbf{w}\, \rangle$$

Inner Product Spaces

and

$$\langle c\,\mathbf{u},\, \mathbf{v} \rangle = (c\,\mathbf{u})^T D\,\mathbf{v} = c\,(\mathbf{u}^T D\,\mathbf{v}) = c\,\langle \mathbf{u},\, \mathbf{v} \rangle\,.$$

Finally

$$\langle \mathbf{u},\, \mathbf{u} \rangle = \mathbf{u}^T D\,\mathbf{u} = w_1\,u_1^2 + w_2\,u_2^2 + \cdots + w_n u_n^2 \geqslant 0$$

since $w_i > 0$ and so $\langle \mathbf{u},\, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

This example can be generalised. Let A be a symmetric matrix with strictly positive eigenvalues (a so-called positive definite matrix) then

$$\langle \mathbf{u},\, \mathbf{v} \rangle = \mathbf{u}^T A\,\mathbf{v}$$

defines an inner product (exercise).

Inner Product Spaces

Example

Show that

$$\langle f,\, g \rangle = \int_a^b f(x)\,g(x)\;dx$$

where $f,\; g \in \mathcal{C}[a,\, b]$ defines an inner product on $\mathcal{C}[a,\, b]$, the vector space of continuous functions on $[a,\, b]$.

Solution: We have

$$\langle f,\, g \rangle = \int_a^b f(x)\,g(x)\;dx = \int_a^b g(x)\,f(x)\;dx = \langle g,\, f \rangle\,,$$

$$\begin{aligned}\langle f,\, g+h \rangle &= \int_a^b f(x)\,(g(x)+h(x))\;dx \\ &= \int_a^b f(x)\,g(x)\;dx + \int_a^b f(x)\,h(x)\;dx = \langle f,\, g \rangle + \langle f,\, h \rangle\,,\end{aligned}$$

Inner Product Spaces

$$\langle c\,f,\ g\rangle = \int_a^b c\,f(x)\,g(x)\ dx = c\int_a^b f(x)\,g(x)\ dx = c\,\langle f,\ g\rangle$$

for $h \in \mathcal{C}[a,\ b]$ and scalar c. Finally

$$\langle f,\ f\rangle = \int_a^b [f(x)]^2\ dx \geqslant 0$$

and, since f is continuous, $\langle f,\ f\rangle = 0$ if and only if $f = 0$.

Inner Product Spaces

Definition 15.2

Let V be an inner product space. For $\mathbf{u},\ \mathbf{v} \in V$,

- *the norm (or length) of $\mathbf{v}$ is*

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v},\ \mathbf{v}\rangle};$$

- *the distance between $\mathbf{u}$ and $\mathbf{v}$ is*

$$d(\mathbf{u},\ \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|;$$

- *$\mathbf{u}$ and $\mathbf{v}$ are orthogonal if*

$$\langle \mathbf{u},\ \mathbf{v}\rangle = 0.$$

A unit vector is a vector whose norm is 1.

Inner Product Spaces

The concepts that were introduced during our discussion of orthogonality in $\mathbf{R}^n$ can now be generalized to inner product spaces. For example, an orthogonal set of vectors in an inner product space V is a (possibly infinite) set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$ such that $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ whenever $i \neq j$ and so on.

The Gram-Schmidt process allowed us to construct an orthogonal basis for any subspace of $\mathbf{R}^n$. We can mimic this construction to show that any finite dimensional subspace of an inner product space has an orthogonal basis. All we need to do is replace the dot product by the inner product. Thus, at the kth stage, we have

$$\mathbf{v}_k = \mathbf{x}_k - \frac{\langle \mathbf{x}_k, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{x}_k, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \cdots - \frac{\langle \mathbf{x}_k, \mathbf{v}_{k-1} \rangle}{\|\mathbf{v}_{k-1}\|^2}\mathbf{v}_{k-1}$$

Inner Product Spaces

Example

Find an orthogonal basis for $\mathcal{P}_2$, polynomials of degree 2 or less, with respect to the inner product

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

Solution: Using the standard basis $(1, x, x^2)$ for $\mathcal{P}_2$, we have

$$\mathbf{v}_1 = \mathbf{x}_1 = 1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 = x - \frac{\int_{-1}^{1} x\, dx}{\int_{-1}^{1} 1\, dx} 1 = x - \frac{0}{2} 1 = x$$

Inner Product Spaces

$$
\begin{aligned}
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2 \\
&= x^2 - \frac{\int_{-1}^{1} x^2 \, dx}{\int_{-1}^{1} 1 \, dx} 1 - \frac{\int_{-1}^{1} x^3 \, dx}{\int_{-1}^{1} x^2 \, dx} x \\
&= x^2 - \frac{\frac{2}{3}}{2} 1 - \frac{0}{\frac{2}{3}} x = x^2 - \tfrac{1}{3}.
\end{aligned}
$$

Therefore $(1,\ x,\ x^2 - \frac{1}{3})$ is an orthogonal basis for $\mathcal{P}_2$ on the interval $[-1,\ 1]$.

The polynomials 1, x and $x^2 - \frac{1}{3}$ are the first three Legendre polynomials. They occur frequently in applications involving partial differential equations.

Inner Product Spaces

As before we define the orthogonal projection $\text{proj}_W(\mathbf{v})$ of a vector $\mathbf{v}$ onto a subspace W of an inner product space. If $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_k\}$ is an orthogonal basis for W then

$$\text{proj}_W(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 + \frac{\langle \mathbf{v}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{v}_k \rangle}{\langle \mathbf{v}_k, \mathbf{v}_k \rangle}\mathbf{v}_k$$

with the orthogonal component given by

$$\text{perp}_W(\mathbf{v}) = \mathbf{v} - \text{proj}_W(\mathbf{v}).$$

Projections allow us to construct "best approximations". Frequently we are faced with a vector from an inner product space which is not easy to handle (say from a vector space that does not have a countable basis). One approach would be to approximate it with a vector in some subspace which might, say, be finite dimensional. We could then introduce coordinates and use MATLAB to aid in the solution to the problem at hand.

Best Approximation

Definition 15.3

Let W be a subspace of an inner product space V then, for $\mathbf{v} \in V$, the best approximation to $\mathbf{v}$ in W is the vector $\overline{\mathbf{v}} \in W$ such that

$$\|\mathbf{v} - \overline{\mathbf{v}}\| < \|\mathbf{v} - \mathbf{w}\|$$

for every $\mathbf{w} \in W$ different from $\overline{\mathbf{v}}$.

Note that the requirement that V be an inner product space is an overkill. All we need is the concept of a norm (see Poole section 7.2 for a discussion of normed linear spaces). We have seen that, for $\mathbf{R}^n$, the orthogonal projection gives the vector in W that is the "shortest distance" away from $\mathbf{v}$. The same is true in general.

Best Approximation

Theorem 15.4 (Best Approximation)

Let W be a finite dimensional subspace of an inner product space V. For $\mathbf{v} \in V$, the best approximation to $\mathbf{v}$ in W is $\text{proj}_W(\mathbf{v})$.

Proof: See Poole Theorem 7.8

Example

Find the best approximation of $\sin \pi x$ in $\mathcal{P}_3[-1,\ 1]$ with the inner product

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

Solution: We have an orthogonal basis for $\mathcal{P}_2[-1,\ 1]$; namely $(1,\ x,\ x^2 - \frac{1}{3})$. This basis extends to $\mathcal{P}_3[-1,\ 1]$ with

$$\begin{aligned}\mathbf{v}_4 &= x^3 - \frac{\langle x^3, 1 \rangle}{\langle 1, 1 \rangle} 1 - \frac{\langle x^3, x \rangle}{\langle x, x \rangle} x - \frac{\langle x^3, x^2 - \frac{1}{3} \rangle}{\langle x^2 - \frac{1}{3}, x^2 - \frac{1}{3} \rangle} (x^2 - \tfrac{1}{3}) \\ &= x^3 - \tfrac{3}{5} x.\end{aligned}$$

Best Approximation

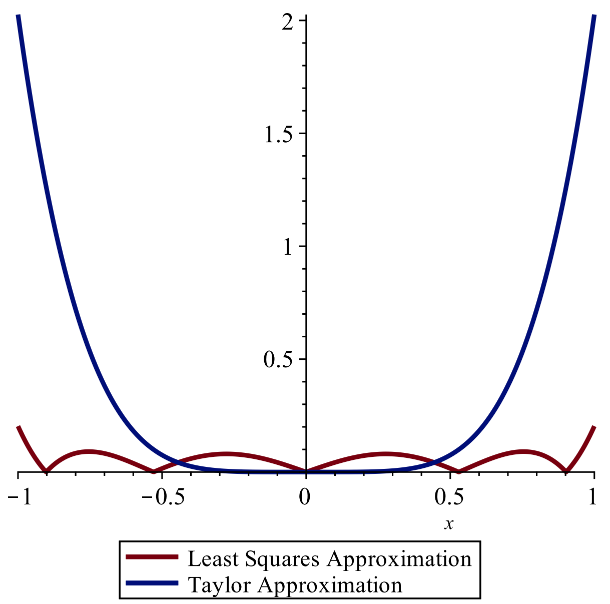
Note that $\sin \pi x$ is an odd function, $x^{2n} \sin \pi x$ is odd and so will integrate to zero on $[-1,\ 1]$. Thus $\langle \sin \pi x,\ 1 \rangle = \langle \sin \pi x,\ x^2 - \frac{1}{3} \rangle = 0$ and so the best approximation in $\mathcal{P}_3[-1,\ 1]$ will be

$$\begin{aligned}
\operatorname{proj}_{\mathcal{P}_3[-1,\ 1]}(\sin \pi x) &= \frac{\langle \sin \pi x,\ x \rangle}{\langle x,\ x \rangle}\, x + \frac{\langle \sin \pi x,\ x^3 - \frac{3}{5}x \rangle}{\langle x^3 - \frac{3}{5}x,\ x^3 - \frac{3}{5}x \rangle}\,(x^3 - \tfrac{3}{5}x) \\
&= \frac{\int_{-1}^{1} x \sin \pi x \, dx}{\frac{2}{3}}\, x \\
&\quad + \frac{\int_{-1}^{1} (x^3 - \frac{3}{5}x) \sin \pi x \, dx}{\frac{8}{175}}\,(x^3 - \tfrac{3}{5}x) \\
&= \frac{3}{2}\frac{2}{\pi}\, x + \frac{175}{8}\frac{4(\pi^2 - 15)}{5\pi^3}\,(x^3 - \tfrac{3}{5}x) \\
&= \frac{3}{\pi}\, x + \frac{35(\pi^2 - 15)}{2\pi^3}\,(x^3 - \tfrac{3}{5}x)
\end{aligned}$$

Best Approximation

By contrast, the Taylor polynomial of degree 3 is given by

$$\sin \pi x \sim \pi x - \frac{\pi^3}{6} x^3$$



Best Approximation

The Taylor polynomial provides a better approximation near 0. However this approximation breaks down as $x \to \pm 1$. On the other hand $\text{proj}_{\mathcal{P}_3[-1,\,1]}(\sin \pi x)$ sacrifices accuracy near 0 but is a much better approximation as $x \to \pm 1$.
For example, at $x = \frac{1}{2}$, the Taylor approximation gives $\sin \frac{1}{2}\pi \approx 0.925$ whereas the projection approximation gives $\sin \frac{1}{2}\pi \approx 0.984$. In fact

$$\begin{aligned} \| \sin \pi x - \text{proj}_{\mathcal{P}_3[-1,\,1]}(\sin \pi x) \| &\approx 0.094 \\ \| \sin \pi x - (\pi x - \tfrac{\pi^3}{6} x^3) \| &\approx 0.895 \end{aligned}$$

confirming that the projection is a better approximation (using this definition of "better").

Fourier Series

The space of square integrable functions on $[a,\, b]$, $L^2[a,\, b]$, consists of functions f such that

$$\int_a^b [f(x)]^2 \, dx < \infty.$$

This is clearly an infinite dimensional vector space. The space of continuous functions $\mathcal{C}[a,\, b]$ is a subspace of this space. It is also an inner product space with the inner product given by

$$\langle f,\, g \rangle = \int_a^b f(x)\, g(x)\; dx$$

The functions

$$\mathcal{B} = \{1,\, \sin x,\, \cos x,\, \sin 2x,\, \cos 2x,\, \ldots\}$$

form an orthogonal set for $L^2[0,\, 2\pi]$ (check). In fact this set forms a basis; that is, for any $f \in L^2[0,\, 2\pi]$, we have

$$f = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos nx + b_n \sin nx \right).$$

Fourier Series

Let S_N be the Nth partial sum of this series. The $=$ is the above formula is convergence in the norm; that is

$$\|f - S_N\| = \sqrt{\int_0^{2\pi} (f - S_N)^2 \, dx} \to 0$$

as $N \to \infty$. This is known as L^2 convergence or root mean square convergence. Since $\mathcal{B}$ is orthogonal, the coefficients are easy to compute:

$$a_0 = \frac{\langle f, 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{2\pi} \int_0^{2\pi} f(x) \, dx$$

$$a_n = \frac{\langle f, \cos nx \rangle}{\langle \cos nx, \cos nx \rangle} = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx \, dx$$

$$b_n = \frac{\langle f, \sin nx \rangle}{\langle \sin nx, \sin nx \rangle} = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx \, dx;$$

the well-known Fourier series.

Complex Fourier Series

In many applications, it is more convenient to use the *complex* form of the Fourier series.For a function defined on$[0, 2p]$, let

$$\omega = \frac{2\pi}{2p} = \frac{\pi}{p};$$

the *fundamental angular frequency*. We write the basis for $L^2[0, 2p]$

$$\mathcal{B} = \{\ldots, e^{-2i\omega x}, e^{-i\omega x}, 1, e^{i\omega x}, e^{2i\omega x}, \ldots\}$$

where $i = \sqrt{-1}$ (frequently denoted by j in electrical engineering).Now the Fourier formula reads

$$f = \sum_{n=-\infty}^{\infty} c_n \, e^{in\omega x}$$

with

$$c_n = \frac{1}{2p} \int_0^{2p} f(x) \, e^{-in\omega x} \, dx.$$

Complex Fourier Series

This formula may look a little strange. The basis $\mathcal{B}$ spans a *complex* vector space. The inner product for a complex vector space requires a complex conjugate to be taken on one argument so that the length (norm) is a non-negative real number. Therefore

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

Now

$$\begin{aligned} \left\langle e^{in\omega x}, e^{im\omega x} \right\rangle &= \int_0^{2p} e^{in\omega x}\, e^{-im\omega x}\, dx = \int_0^{2p} e^{i(n-m)\omega x}\, dx \\ &= \begin{cases} 2p & n = m \\ 0 & n \neq m \end{cases} \end{aligned}$$

and so $\mathcal{B}$ is an orthogonal basis.Thus

$$c_n = \frac{\left\langle f, e^{in\omega x} \right\rangle}{\left\langle e^{in\omega x}, e^{in\omega x} \right\rangle} = \frac{1}{2p} \int_0^{2p} f(x)\, e^{-in\omega x}\, dx.$$

Fourier Transform

We can view the coefficients $c_n = c(n)$ as functions of n. If we expand the interval $[0, 2p]$ to the entire real line, then $\omega \to 0$ and $\alpha = n\omega$ becomes a continuous variable. We can now view c not as a *discrete* function of n but as a *continuous* function of α.Thus

Definition 15.5 (Fourier transform)

The Fourier transform of $f(x)$ *is given by*

$$\mathcal{F}\{f(x)\} = \int_{-\infty}^{\infty} f(x)\, e^{i\alpha x}\, dx = F(\alpha)$$

and the Inverse Fourier Transform is given by

$$\mathcal{F}^{-1}\{F(\alpha)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\alpha)\, e^{-i\alpha x}\, d\alpha = f(x).$$

Note that different authors use different numerical coefficients in front of the integrals. However the product of these coefficients must be $\frac{1}{2\pi}$.

Discrete Fourier Transform

Consider a signal $f(x)$ that is sampled at equally spaced points $x = nT$ where T is the sampling rate.The sample at nT is

$$f(nT) = f(x)\,\delta(x - nT)$$

where $\delta(x)$ is the Dirac delta function. We can represent the *discrete* version of f as the sum of these impulses

$$\sum_{n=-\infty}^{\infty} f(x)\,\delta(x - nT).$$

The Fourier transform of the discrete signal is given by

$$F(\alpha) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(x)\,\delta(x - nT)\,e^{i\alpha x}\,dx = \sum_{n=-\infty}^{\infty} f(nT)\,e^{i\alpha nT}$$

by the shift property of the Dirac delta function. This is called the *discrete* Fourier transform. Note that $e^{i\alpha x}$ is periodic in α;

$$e^{i\alpha T} = e^{i(\alpha T + 2\pi)} = e^{i(\alpha + 2\pi/T)T}$$

and so we only need to consider $F(\alpha)$ in the range $[0, 2\pi/T]$. Because we sample only over one period, this reduces the sum to a *finite* sum.

Discrete Fourier Transform

Consider the function values $f(x)$ at N equally spaced points $x = nT$, $n = 0, 1, \ldots, (N-1)$ in the interval $[0, 2\pi]$. Thus $T = 2\pi/N$. The sampled signal will be

$$f = \sum_{n=0}^{N-1} c_n\,e^{inx}.$$

Let $f_k = f(kT) = f(2k\pi/N)$, the observed samples with

$$\mathbf{f} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{N-1} \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{N-1} \end{bmatrix}.$$

Discrete Fourier Transform

Now

$$
\begin{aligned}
f_k &= c_0 + c_1\, e^{i2k\pi/N} + c_2\, e^{i4k\pi/N} + \cdots + c_{N-1}\, e^{i2(N-1)k\pi/N} \\
&= c_0 + c_1\, \omega_N^k + c_2\, \omega_N^{2k} + \cdots + c_{N-1}\, \omega_N^{k(N-1)}
\end{aligned}
$$

where

$$\omega_N = e^{i2\pi/N}.$$

Thus we obtain a matrix equation

$$
\begin{aligned}
\mathbf{f} &= \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_N & \omega_N^2 & \cdots & \omega_N^{N-1} \\ 1 & \omega_N^2 & \omega_N^4 & \cdots & \omega_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_N^{N-1} & \omega_N^{2(N-1)} & \cdots & \omega_N^{(N-1)^2} \end{bmatrix} \mathbf{c} \\
&= \mathbf{F}_N\, \mathbf{c},
\end{aligned}
$$

say.

Discrete Fourier Transform

Is there any easy way to solve this system for the Fourier coefficients $\mathbf{c}$? Note that

$$\omega_N\, \overline{\omega}_N = 1$$

and so

$$\mathbf{F}_N\, \overline{\mathbf{F}}_N = \overline{\mathbf{F}}_N\, \mathbf{F} = N\, I$$

where $\overline{\mathbf{F}}_N$ is the matrix of complex conjugates.Therefore

$$\mathbf{F}_N^{-1} = \frac{1}{N} \overline{\mathbf{F}}_N.$$

We have another case where it is easy to compute the inverse!Thus

$$\mathbf{c} = \frac{1}{N}\, \overline{\mathbf{F}}_N\, \mathbf{f}.$$

Fast Fourier Transform

The computation of $\mathbf{c}$ requires $O(N^2)$ computations. Can we do better? We have

$$N\,c_k = \sum_{n=0}^{N-1} f_n\,\overline{\omega}_N^{kn}.$$

Suppose $N = 2M$ is an *even* integer. We can then split this sum into the even and odd numbered indicies:

$$N\,c_k = \sum_{n=0}^{M-1} f_{2n}\,\overline{\omega}_N^{2kn} + \sum_{n=0}^{M-1} f_{2n+1}\,\overline{\omega}_N^{k(2n+1)}.$$

The second term has a common factor $\overline{\omega}_N^k$ and so

$$N\,c_k = \sum_{n=0}^{M-1} f_{2n}\,\overline{\omega}_N^{2kn} + \overline{\omega}_N^k \sum_{n=0}^{M-1} f_{2n+1}\,\overline{\omega}_N^{2kn}.$$

Fast Fourier Transform

Next we observe

$$\overline{\omega}_N^2 = e^{-i4\pi/N} = e^{-i4\pi/(2M)} = e^{-i2\pi/M} = \overline{\omega}_M$$

Therefore we can rewrite the sums as

$$\begin{aligned} N\,c_k &= \sum_{n=0}^{M-1} f_{2n}\,\overline{\omega}_M^{kn} + \overline{\omega}_N^k \sum_{n=0}^{M-1} f_{2n+1}\,\overline{\omega}_M^{kn} \\ &= E_k + \overline{\omega}_N^k\,O_k \end{aligned}$$

say. E_k is the discrete Fourier transform of the *even-indexed* inputs and O_k is the discrete Fourier transform of the *odd-indexed* inputs.

Fast Fourier Transform

Furthermore we observe

$$
\begin{aligned}
N\, c_{k+M} &= \sum_{n=0}^{M-1} f_{2n}\, \overline{\omega}_M^{(k+M)n} + \overline{\omega}_N^{k+M} \sum_{n=0}^{M-1} f_{2n+1}\, \overline{\omega}_M^{(k+M)n} \\
&= \sum_{n=0}^{M-1} f_{2n}\, \overline{\omega}_M^{kn}\, \overline{\omega}_M^{Mn} + \overline{\omega}_N^{k}\, \overline{\omega}_N^{M} \sum_{n=0}^{M-1} f_{2n+1}\, \overline{\omega}_M^{kn}\, \overline{\omega}_M^{Mn}
\end{aligned}
$$

Note that

$$\overline{\omega}_M^{Mn} = e^{-i2Mn\pi/M} = e^{-i2n\pi} = 1$$

and

$$\overline{\omega}_N^{M} = e^{-i2M\pi/N} = e^{-i\pi} = -1.$$

Therefore

$$N\, c_{k+M} = \sum_{n=0}^{M-1} f_{2n}\, \overline{\omega}_M^{kn} - \overline{\omega}_N^{k} \sum_{n=0}^{M-1} f_{2n+1}\, \overline{\omega}_M^{kn} = E_k - \overline{\omega}_N^{k}\, O_k.$$

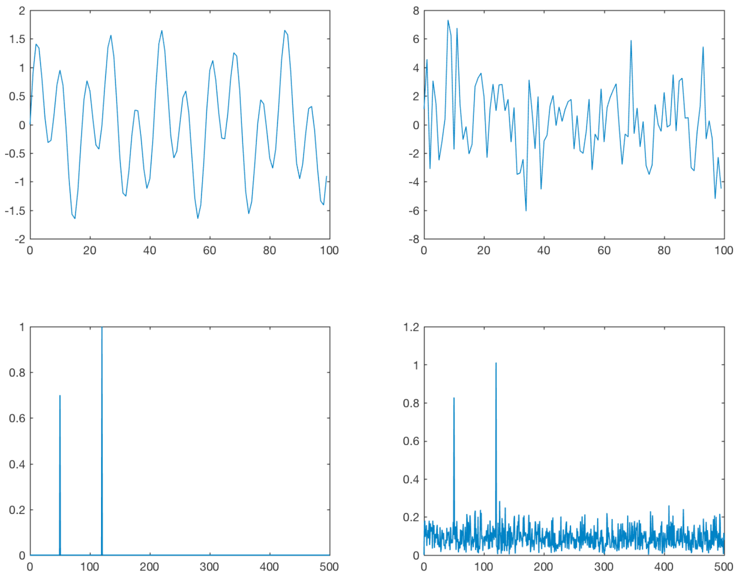
Fast Fourier Transform

We can rewrite the Fourier coefficients

$$
\begin{aligned}
N\, c_k &= E_k + \overline{\omega}_N^{k}\, O_k \\
N\, c_{k+M} &= E_k - \overline{\omega}_N^{k}\, O_k.
\end{aligned}
$$

Thus we can find a discrete Fourier transform of length N in terms of two discrete Fourier transforms of length $N/2$. If $N = 2^P$, we can recursively apply this split to rewrite the Fourier coefficients in terms of shorter and shorter Fourier transforms. This is the Fast Fourier Transform (FFT). It was “first” published by Cooley and Tukey in 1965 and has revolutionised many parts of computational mathematics. Interestingly, the algorithm can be traced back to Gauss (though he did not analyse the computational costs).

Fast Fourier Transform



Fast Fourier Transform

Cooley and Tukey showed that the asymptotic cost of this algorithm is $O(N \log N)$ (remember the naive approach is $O(N^2)$). Thus

N	N^2	$N \log N$
256	65536	1400
1024	1.0e6	7100
4096	1.6e7	34000
. 16384	2.7e8	1.6e5

An audio CD can hold about 2×10^8 samples (74 minutes at 44.1 kHz). A naive Fourier transform of this data would take about 4×10^{16} operations. With the fast Fourier transform, we need to pad the data out to the next power of 2; that is 2^{28}. The fast Fourier transform will take about 10^9 operations. The fastest Intel processors can perform in the teraflop range (10^{12} flops/sec). A teraflop machine would take 40000 seconds (that is about 11 hours) to compute the Fourier transform of the data on a CD by the naive method. However it would take only 0.001 seconds to compute the same Fourier transform by the fast method!