ME599-004: Data-Driven Methods for Control Systems

Instructors: Uduak Inyang-Udoh, Yuxin Tong

January 8, 2025

Contents

Chapter 1: Data Analysis & Machine Learning Preliminaries

1 Fourier Transform and Its Applications

1.1 Inner Products

Consider 2 real finite vectors $\bar{f}, \bar{g} \in \mathbb{R}^n$,

$$\vec{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \quad \vec{g} = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{bmatrix}.$$

Their dot product, that is, the projection of one vector on another to get a magnitude is given by

$$\vec{f} \cdot \vec{g} = \vec{f}^T \vec{g} = \sum_{i=1}^{n} f_i g_i.$$

We can generalize this idea of projecting one vector onto the other to all vector spaces (e.g. complex, infinite dimensional), and we refer to this as the *inner product*.

Given a vector space V and a field F ($\mathbb{R}$ or $\mathbb{C}$), the inner product is the mapping $\langle \cdot, \cdot \rangle : V \times V \to F$ such that $\forall \alpha, \beta \in F$ and $f, g, h \in V$, the following rules/properties hold:

1. Linearity: $\langle \alpha f + \beta g, h \rangle = \alpha \langle f, h \rangle + \beta \langle g, h \rangle$
2. Conjugate Symmetry: $\langle f, g \rangle = \overline{\langle g, f \rangle}$; $\overline{\langle \cdot \rangle}$ denotes complex conjugation
3. Positive definiteness: $\langle f, f \rangle \geqslant 0$ and $\langle f, f \rangle = 0$ iff $f = \overrightarrow{0}$

Two interesting points to note here:

First, the inner product induces a norm when we find the inner product of a vector with itself, that is, it turns out that for the vector $f \in V$, the inner product $\langle f, f \rangle$ satisfies the definition of a norm:

$$\langle f, f \rangle = \|f\|^2$$

Remark: This is not the 1-norm, ∞-norm or some other p-norm. This is "the norm induced by the inner product". It turns out that this norm is equivalent to the 2-norm (see Polarization Identity on Wikipedia for more information).

Second, if f and g are infinite dimensional vectors, or function, say, defined over some interval $[a, b]$ (Fig. 1).

We know that

$$\langle (f_1, \cdots f_n), (g_1, \cdots g_n) \rangle = \sum_{i=1}^{n} f(x_i)\, g(x_i) \tag{1}$$

We can normalize (1) by noting $\Delta x = \frac{b-a}{n-1}$:

$$\frac{b-a}{n-1} \langle \vec{f}, \vec{g} \rangle = \sum_{i=1}^{n} f(x_i)\, g(x_i)\, \Delta x_j$$

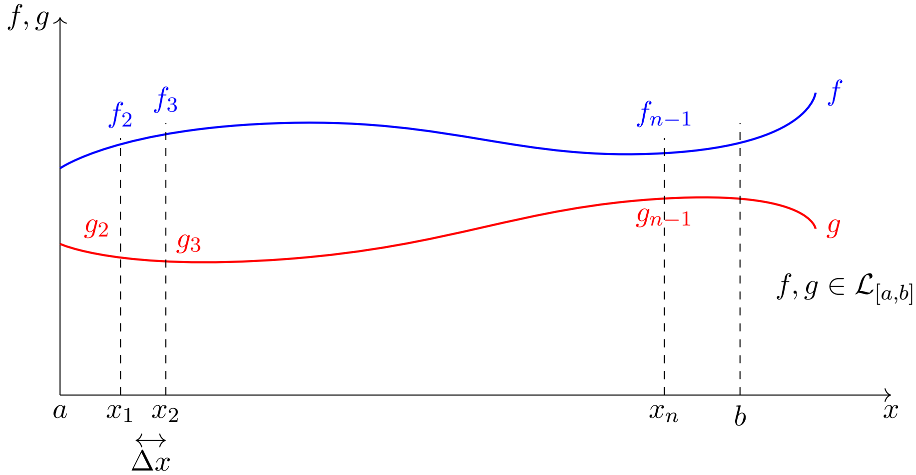


Figure 1: Functions f and g over the interval $[a, b]$.

and taking the limit,

$$\left.\begin{aligned} \lim_{\Delta x \to 0} \sum_{i=1}^{n} f(x_i)\, g(x_i)\, \Delta x &= \int_a^b f(x) g(x) dx \\ &= \langle f(x), g(x) \rangle \end{aligned}\right\} \text{assuming } f, g \text{ (now functions) are Reinmann integrable}^1$$

If you think about it, we use the inner product $\langle \cdot, \cdot \rangle$ in finite-dimensional vector spaces to find the projection of a vector to some bases[2] We can do the same for the inner product of functions by projecting them to an orthogonal (and possibly orthonormal) set of sine and cosine functions with integer periods on the domain $[a, b]$. Thus is exactly what the Fourier Series is.

Note: For complex functions, we can more generally write the inner product as

$$\langle f(x), g(x) \rangle_{[a,b]} = \int_a^b f(x) \bar{g}(x) dx.$$

[1] The limit exists

[2] Recall that a basis is a set of linear independent vectors that span a vector space log $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$, that vs respective of the vector $v \in \mathbb{R}^2$ you give me $\exists \alpha, \beta$ such that

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

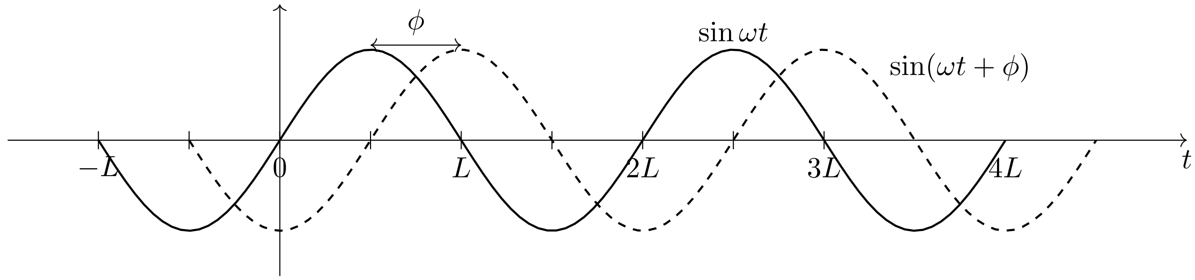
A basis is orthonormal of each element yoelds a zero inner product with other element, and has a unit induced norm. erg the standard basis for $\mathbb{R}^3$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Note that (1) the dimension of the vector space = # elements in a basis (2) The basis for a vector space is not unique

1.2 Fourier Series

Assuming I have a sine wave oscillating at angular velocity $\omega = \frac{2\pi}{2L}$, that is, it takes $2L$ units of time to complete a cycle:



Suppose I have another sine wave $\sin(\omega t + \phi)$, and I want to find what the projection of $\sin(\omega t + \phi)$ is on $\sin \omega t$, that is, the **Hilbert space inner product** $\mathcal{H}^*$.

$$\int_{-L}^{L} \sin(\omega t + \phi) \sin(\omega t)\, dt = L \cos \phi$$

The inner product is a function of ϕ, and for each multiple of $\frac{\pi}{2}$, it is zero. That is, $\sin \omega t$ and $\sin(\omega t + \frac{\pi}{2})$ (or $\cos(\omega t)$) are orthogonal. Any sinusoid oscillating at ω can be written in terms of $\sin \omega t$ and $\cos \omega t$. In fact, these can act as a basis:$\{\sin \omega t,\, \cos \omega t\}$.

How about a sine wave oscillating at higher frequencies?

$$\langle \sin(k\omega t), \sin(\omega t)\rangle = \int_{-L}^{L} \sin(k\omega t) \sin(\omega t)\, dt \quad \text{for } k = 0, 2, 3, \ldots, \infty$$

$$= \frac{-2L}{\pi(k^2 - 1)} \quad \text{for } k = 0, 2, ..., \infty$$

which means

$$\{\sin(\omega t), \sin(2\omega t), \ldots, \cos(\omega t), \cos(2\omega t), \ldots\} \quad \text{or} \quad \{1, \sin k\omega t, \cos k\omega t\}_{k=1}^{\infty}$$

forms a set of orthogonal functions.

Moreover, just like we write vectors in terms of standard, orthogonal vectors e.g. in 3D Euclidean space, for $f \in \mathbb{R}^3$:

$$\vec{f} = \frac{\langle f, e_x\rangle}{\langle e_x, e_x\rangle} e_x + \frac{\langle f, e_y\rangle}{\langle e_y, e_y\rangle} e_y + \frac{\langle f, e_z\rangle}{\langle e_z, e_z\rangle} e_z,$$

we can write:

$$f(t) = \sum_{k=0}^{\infty} \frac{\langle f, \cos(k\omega t)\rangle}{\|\cos(k\omega t)\|^2} \cos(k\omega t) + \sum_{k=1}^{\infty} \frac{\langle f, \sin(k\omega t)\rangle}{\|\sin(k\omega t)\|^2} \sin(k\omega t).$$

for all square integrable functions in $L_2([-L, L])$. Note that the cosine series starts from $k = 0$ because $\cos(0 \cdot \omega t) = 1$, representing the constant (DC) component, while the sine series starts from $k = 1$ because $\sin(0 \cdot \omega t) = 0$.

Let

$$a_k = \frac{\langle f, \cos(k\omega t)\rangle}{\|\cos(k\omega t)\|^2} = \frac{1}{\|\cos(k\omega t)\|^2} \int_{-L}^{L} f(t) \cos(k\omega t)\, dt \quad \text{for } k = 1, 2, \ldots, \infty$$

$$= \begin{cases} \frac{1}{2L} \int_{-L}^{L} f(t)\, dt & \text{if } k = 0, \\ \frac{1}{L} \int_{-L}^{L} f(t) \cos(k\omega t)\, dt & \text{else} \end{cases}$$

Similarly, let

$$b_k = \frac{\langle f, \sin(k\omega t) \rangle}{\| \sin(k\omega t) \|^2} = \frac{1}{\| \sin(k\omega t) \|^2} \int_{-L}^{L} f(t) \sin(k\omega t)\, dt \quad \text{for } k = 1, 2, \ldots, \infty$$

$$= \frac{1}{L} \int_{-L}^{L} f(t) \sin(k\omega t)\, dt.$$

Hence,

$$f(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(k\omega t) + b_k \sin(k\omega t)$$

which is the Fourier Series.

1.3 Fourier Transform

Based on Euler's formula,

$$e^{ik\omega t} = \cos(k\omega t) + i \sin(k\omega t),$$

we can write the Fourier series as:

$$f(t) = \sum_{k=0}^{\infty} c_k e^{ik\pi t/L}$$

where

$$c_k = \frac{1}{2L} \langle f(t), \psi_k \rangle = \frac{1}{2L} \int_{-L}^{L} f(t) e^{-ik\pi t/L}\, dt,$$

where

$$\psi_k = e^{ik\pi t/L}.$$

Let us assume $f(t)$ is not just a periodic function defined in the domain $[-L, L)$, and then repeat it self outside the domain. Instead, let $L \to \infty$, and thus π/L, which we now denote with $\Delta\omega$, tend to 0 ($\Delta\omega = \pi/L \to 0$).

$$f(t) = \lim_{\Delta\omega \to 0} \sum_{k=-\infty}^{\infty} \left[\frac{\Delta\omega}{2\pi} \underbrace{\int_{-\pi/\Delta\omega}^{\pi/\Delta\omega} f(\xi) e^{-ik\Delta\omega\xi}\, d\xi}_{\langle f(t), \psi_k(t) \rangle} \right] e^{ik\Delta\omega t}$$

In the limit, $\langle f(t), \psi_k(t) \rangle$ becomes the Fourier transform, and the entire summation can be written as the Riemann integral:

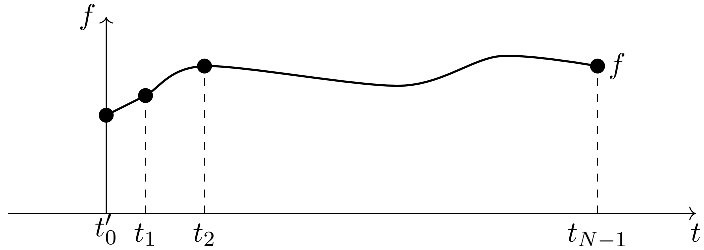
$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t}\, dt \triangleq \mathcal{F}(f(t))$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i\omega t}\, d\omega \triangleq \mathcal{F}^{-1}(\hat{f}(\omega))$$

These are regarded as the Fourier transform pair.

1.4 Discrete Fourier Transform

Typically, we have our function (in-time) defined by a set of data points, not analytically:



The smallest frequency we can detect as a basis for this set of points that constitute our discrete function (fundamental frequency) is $\frac{1}{N}$, or angular frequency $\frac{2\pi}{N}$.

So we can now write the **Discrete Fourier Transform (DFT)** as:

$$\hat{f}_k = \sum_{j=0}^{N-1} f_j e^{-ik\frac{2\pi}{N}j}$$

and the **Inverse Discrete Fourier Transform (IDFT)**:

$$f_j = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}_k e^{ij\frac{2\pi}{N}k}$$

NOTE: We could have written $\frac{1}{\sqrt{N}}$ before the DFT and IDFT.

Now, we can express $\{\hat{f}_k\}$ for $k = 0, \ldots, n-1$ as a **linear operator** (matrix) acting on $\{f_j\}$ for $j = 0, \ldots, n-1$ by denoting $e^{-i\frac{2\pi}{N}}$ by ω_N.

$$\begin{array}{c} k=0, \\ 1, \\ \vdots \\ N-1 \end{array} \left\{ \begin{bmatrix} \hat{f}_0 \\ \hat{f}_1 \\ \hat{f}_2 \\ \vdots \\ \hat{f}_{N-1} \end{bmatrix} \right. = \underbrace{\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_N^1 & \omega_N^2 & \cdots & \omega_N^{N-1} \\ 1 & \omega_N^2 & \omega_N^4 & \cdots & \omega_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_N^{N-1} & \omega_N^{2(N-1)} & \cdots & \omega_N^{(N-1)^2} \end{bmatrix}}_{\text{DFT matrix, } j=0,1,2,\ldots,N-1} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_{N-1} \end{bmatrix}$$

It should be obvious that this DFT matrix F is symmetric; In fact, it is *unitary* (see below). In addition, it is a Vandermonde matrix, i.e.,

$$F_{i,j} = \left(\omega_N^i\right)^j, \quad i, j = 0, \ldots, N-1$$

NOTE: Actually, $F^{-1} = \frac{1}{N} F^H$.

However, if we had defined the DFT with the normalization $\frac{1}{\sqrt{N}}$, i.e.,

$$\hat{f}_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} f_j e^{-ik\frac{2\pi}{N}j}, \quad f_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{f}_k e^{ij\frac{2\pi}{N}k}$$

we would have had a unitary matrix $F_{\sqrt{N}}$ for which ${F_{\sqrt{N}}}^{-1} = {F_{\sqrt{N}}}^H$, where $F_{\sqrt{N}} = \frac{1}{\sqrt{N}} F$.

1.5 Fast Fourier Transform

The DFT requires $\mathcal{O}(N^2)$ operations. But given the symmetry of the DFT matrix, it is interesting to note that:

$$\begin{aligned}\hat{f}_{k+N} &= \sum_{j=0}^{N-1} f_j e^{-i(k+N)\frac{2\pi}{N}j} \\ &= \sum_{j=0}^{N-1} f_j e^{-ik\frac{2\pi}{N}j} = \hat{f}_k\end{aligned}$$

In fact,

$$\hat{f}_{k+nN} = \hat{f}_k \quad \text{for any integer } n.$$

So,

$$\begin{aligned}\hat{f}_k &= \sum_{j=0}^{N/2-1} f_{2j} e^{-ik\frac{2\pi}{N}(2j)} + \sum_{j=0}^{N/2-1} f_{2j+1} e^{-ik\frac{2\pi}{N}(2j+1)} \\ &= \sum_{j=0}^{N/2-1} \underbrace{f_{2j}}_{even} e^{-ik\frac{2\pi}{N/2}j} + e^{-ik\frac{2\pi}{N}} \sum_{j=0}^{N/2-1} \underbrace{f_{2j+1}}_{odd} e^{-ik\frac{2\pi}{N/2}j}\end{aligned}$$

$$\begin{bmatrix} \hat{f}_k \\ \hat{f}_{N/2+k} \end{bmatrix} = \begin{bmatrix} \sum_{j=0}^{N/2-1} f_{2j} e^{-ik\frac{2\pi}{N/2}j} + e^{-ik\frac{2\pi}{N}} \sum_{j=0}^{N/2-1} f_{2j+1} e^{-ik\frac{2\pi}{N/2}j} \\ \underbrace{\sum_{j=0}^{N/2-1} f_{2j} e^{-ik\frac{2\pi}{N/2}j}}_{(1)} - \underbrace{e^{-ik\frac{2\pi}{N}}}_{(2)} \sum_{j=0}^{N/2-1} f_{2j+1} e^{-ik\frac{2\pi}{N/2}j} \end{bmatrix}$$

NOTE: (1) is just another FT for which the routine can be repeated, so we only need to evaluate (2). Hence, it is much quicker to find the FT using the recursive algorithm as follows (fast FT or FFT)

Algorithm 1: `fft`

Data: Signal vector $\mathbf{f} = \{f_0, \cdots, f_{N-1}\}$
Result: $\hat{\mathbf{f}} = \{\hat{f}_0, \cdots, \hat{f}_{N-1}\}$
if $N = 1$ **then**
  $\hat{\mathbf{f}} \leftarrow \mathbf{f}$
else
  $\hat{\mathbf{f}}_{even} \leftarrow \texttt{fft}(\{f_0, f_2, \cdots, f_{N-1}\})$;
  $\hat{\mathbf{f}}_{odd} \leftarrow \texttt{fft}(\{f_1, f_3, \cdots, f_{N-2}\})$;
  $\mathbf{n} \leftarrow \{1, 2, \cdots, \frac{N}{2} - 1\}$;
  $\omega \leftarrow e^{i\frac{2\pi}{N}\mathbf{n}}$;
  $\hat{\mathbf{f}} \leftarrow [\hat{\mathbf{f}}_{even} + \omega\hat{\mathbf{f}}_{odd}, \hat{\mathbf{f}}_{even} - \omega\hat{\mathbf{f}}_{odd}]$
end

Figure 2: FFT Algorithm

In terms of matrices, we can express one step of the ‘divide-and-conquer’ notion as:

$$\left[\hat{f}_k' s\right] = \begin{bmatrix} I_{N/2} & D_{N/2} \\ I_{N/2} & -D_{N/2} \end{bmatrix} \begin{bmatrix} F_{N/2} & 0 \\ 0 & F_{N/2} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \left[f_j{}' s\right],$$

where:

$$D = \begin{bmatrix} (e^{-i\frac{2\pi}{N}})^{\theta} & & \\ & \ddots & \\ & & \left(e^{-i\frac{2\pi}{N}}\right)^{(N/2-1)} \end{bmatrix}$$

Permutation matrices P_1 and P_2 are used to separate f_{even} and f_{odd}, $I_{N/2}$ is the identity matrix.

The partitioning of the Fourier transform can be repeated recursively.

By finding the Fourier transform in this manner, we can cut the number of arithmetic operations from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$, which makes the Fourier transformation for very large N feasible computationally.

References

[1] SL. Brunton , JN. Kutz *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press; 2019.

[2] *The Most Important Algorithm Of All Time* [Youtube Video - Veritasium]

[3] G. Wolberg, Fast Fourier Transform: A Review *Technical Report CUCS-388-88*, 1988.