Chapter 2
Multivariate Harmonic Analysis

Multivariate harmonic analysis technique has been employed widely in determining cyclic variations of multivariate time series. Although it cannot reveal what causes the cycles, it can reveal the likely frequency, amplitude, and phase of those cycles. Moreover, it can be used to simulate atmospheric and oceanic circulation, rule out significant impacts from anthropogenic factors, and ultimately predict what will happen next under climate change scenarios. In this chapter, we will introduce basic techniques and tools in multivariate harmonic analysis, including Fourier transform, fractional Fourier transform, space–frequency representation, sparse approximation, spherical harmonics, harmonic analysis on graphs.

2.1 Fourier Transform

The multivariate Fourier transform is a systematic way to decompose multivariate time series into a superposition of trigonometric functions. It can determine how the total variance of multivariate time series is distributed in frequency. In this section, we will introduce multivariate Fourier transform, Parseval identity, Poisson summation formula, and Shannon sampling theorem.

The Fourier transform of a $d-$variate complex-valued function $f(t_1, \ldots, t_d)$ is defined as

$$\widehat{f}(\omega_1, \ldots, \omega_d) = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(t_1, \ldots, t_d)\, e^{-2\pi i(\omega_1 t_1 + \cdots + \omega_d t_d)} \mathrm{d}t_1 \cdots \mathrm{d}t_d. \qquad (2.1.1)$$

Let $\mathbf{t} = (t_1, \ldots, t_d)$, $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_d)$, $(\boldsymbol{\omega}, \mathbf{t}) = \sum_{k=1}^{d} \omega_k t_k$, and $\mathrm{d}\mathbf{t} = \mathrm{d}t_1 \cdots \mathrm{d}t_d$. Then, (2.1.1) becomes

Z. Zhang, *Multivariate Time Series Analysis in Climate and Environmental Research*, https://doi.org/10.1007/978-3-319-67340-0_2

$$\widehat{f}(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\mathbf{t})\, e^{-2\pi i(\boldsymbol{\omega},\mathbf{t})}\mathrm{d}\mathbf{t}$$

If $f(\mathbf{t})$ is integrable on $\mathbb{R}^d$, then its Fourier transform is continuous on $\mathbb{R}^d$. From $|e^{-2\pi i(\boldsymbol{\omega},\mathbf{t})}| = 1$, it follows that

$$|\widehat{f}(\boldsymbol{\omega})| \le \int_{\mathbb{R}^d} |f(\mathbf{t})|\mathrm{d}\mathbf{t} \qquad (\boldsymbol{\omega} \in \mathbb{R}^d),$$

i.e., Fourier transform $\widehat{f}(\boldsymbol{\omega})$ is a bounded function on $\mathbb{R}^d$. The Riemann–Lebesgue lemma shows that

$$\widehat{f}(\boldsymbol{\omega}) \to 0 \quad \text{as} \quad \| \boldsymbol{\omega} \| = (\omega_1^2 + \cdots + \omega_d^2)^{\frac{1}{2}} \to \infty.$$

The smoother the function $f(\mathbf{t})$ is, the quicker the Fourier transform $\widehat{f}(\boldsymbol{\omega})$ decays.

Let $f(\mathbf{t})$ be a $d-$dimensional real-valued function. If

$$f(t_1, \ldots, t_{k-1}, -t_k, t_{k+1}, \ldots, t_d) = f(t_1, \ldots, t_{k-1}, t_k, t_{k+1}, \ldots, t_d) \qquad (k = 1, \ldots, d),$$

then $f(\mathbf{t})$ is called an *even function*. Fourier transform of an even, real-valued multivariate function is also an even, real-valued multivariate function. If

$$f(\mathbf{t}) = f_1(t_1) f_2(t_2) \cdots f_d(t_d) \qquad (\mathbf{t} = (t_1, \ldots, t_d)),$$

then $f(\mathbf{t})$ is called a *separable function*. Fourier transform of a separable multivariate function is a product of univariate Fourier transforms, i.e.,

$$\begin{aligned}
\widehat{f}(\boldsymbol{\omega}) &= \textstyle\int_{\mathbb{R}^d} f(\mathbf{t})\, e^{-2\pi i(\boldsymbol{\omega},\mathbf{t})}\mathrm{d}\mathbf{t} \\
&= \left(\textstyle\int_{\mathbb{R}^d} f_1(t_1)\, e^{-2\pi i \omega_1 t_1}\mathrm{d}t_1\right) \cdots \left(\textstyle\int_{\mathbb{R}^d} f_d(t_d)\, e^{-2\pi i \omega_d t_d}\mathrm{d}t_d\right) \\
&= \widehat{f}_1(\omega_1) \cdots \widehat{f}_d(\omega_d).
\end{aligned}$$

Below, we give some examples of multivariate Fourier transforms:

Characteristic function of $[-T, T]^d$ is

$$\chi_{[-T,T]^d}(\mathbf{t}) = \begin{cases} 1, \ \mathbf{t} \in [-T, T]^d, \\ 0, \ \mathbf{t} \in \mathbb{R}^d \setminus [-T, T]^d. \end{cases}$$

The corresponding Fourier transform is

$$\widehat{\chi}_{[-T,T]^d}(\boldsymbol{\omega}) = \prod_1^d \frac{\sin(2\pi\omega_k T)}{\pi\omega_k} \qquad (\boldsymbol{\omega} = (\omega_1, \ldots, \omega_d)).$$

Delta function is defined as

$$\delta(\mathbf{t}) = \lim_{T\to 0} \frac{1}{(2T)^d} \chi_{[-T,T]^d}(\mathbf{t}).$$

The corresponding Fourier transform is

$$\widehat{\delta}(\boldsymbol{\omega}) = \lim_{T\to 0} \frac{1}{(2T)^d} \widehat{\chi}_{[-T,T]^d}(\boldsymbol{\omega}) = \prod_1^d \left(\lim_{T\to 0} \frac{\sin(2\pi\omega_k T)}{2\pi\omega_k T} \right) = 1.$$

Gaussian function is defined as $G_M(\mathbf{t}) = e^{-M|\mathbf{t}|^2}$ ($M > 0$; $|\mathbf{t}|^2 = t_1^2 + \cdots + t_d^2$). The corresponding Fourier transform is

$$\widehat{G}_M(\boldsymbol{\omega}) = \left(\frac{\pi}{M}\right)^{\frac{d}{2}} e^{-\frac{\pi^2}{M}|\boldsymbol{\omega}|^2} \qquad (|\boldsymbol{\omega}|^2 = \omega_1^2 + \cdots + \omega_d^2), \tag{2.1.2}$$

i.e., Fourier transform of Gaussian function is still Gaussian function. It plays an important role in the windowed Fourier transform in Sect. 2.6 and the normal distribution in Sect. 4.1.

Multivariate Fourier transforms have the following properties.

(a) Translation. For $\boldsymbol{\tau} \in \mathbb{R}^d$,

$$(f(\mathbf{t}-\boldsymbol{\tau}))^\wedge(\boldsymbol{\omega}) = e^{-2\pi i(\boldsymbol{\omega},\boldsymbol{\tau})} \widehat{f}(\boldsymbol{\omega}),$$

$$\left(e^{2\pi i(\boldsymbol{\tau},\mathbf{t})} f(\mathbf{t})\right)^\wedge(\boldsymbol{\omega}) = \widehat{f}(\boldsymbol{\omega}-\boldsymbol{\tau}).$$

(b) Dilation. For $a \in \mathbb{R}$ and $a \neq 0$,

$$(f(a\mathbf{t}))^\wedge(\boldsymbol{\omega}) = \frac{1}{|a|^d} \widehat{f}\left(\frac{\boldsymbol{\omega}}{a}\right).$$

(c) Derivation:

$$\frac{\partial \widehat{f}}{\partial \omega_k}(\boldsymbol{\omega}) = (-2\pi i t_k f(\mathbf{t}))^\wedge(\boldsymbol{\omega}),$$

$$\left(\frac{\partial f}{\partial t_k}\right)^\wedge(\boldsymbol{\omega}) = 2\pi i \omega_k \widehat{f}(\boldsymbol{\omega}),$$

where $\mathbf{t} = (t_1, \ldots, t_d)$ and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_d)$.

(d) Convolution. Let f and g be functions on $\mathbb{R}^d$. Define convolution of f and g as

$$h(\mathbf{t}) = (f * g)(\mathbf{t}) = \int_{\mathbb{R}^d} f(\mathbf{t}-\boldsymbol{\tau}) g(\boldsymbol{\tau}) \mathrm{d}\boldsymbol{\tau}.$$

Fourier transform of convolution is equal to the product of Fourier transforms, i.e., $\widehat{h}(\boldsymbol{\omega}) = \widehat{f}(\boldsymbol{\omega})\widehat{g}(\boldsymbol{\omega})$.

The *inverse Fourier transform* is defined as

$$f(\mathbf{t}) = \int_{\mathbb{R}^d} \widehat{f}(\boldsymbol{\omega})\, e^{2\pi i(\boldsymbol{\omega},\mathbf{t})} \mathrm{d}\boldsymbol{\omega},$$

which can reconstruct the original function $f(\mathbf{t})$ for given a Fourier transform $\widehat{f}(\boldsymbol{\omega})$.

If $f(\mathbf{t})$ is an even function, by the inverse Fourier transform formulas, it follows that

$$f(\mathbf{t}) = f(-\mathbf{t}) = \int_{\mathbb{R}^d} \widehat{f}(\boldsymbol{\omega})\, e^{-2\pi i(\boldsymbol{\omega},\mathbf{t})} \mathrm{d}\mathbf{t},$$

i.e., the even function $f(\mathbf{t})$ is the Fourier transform of $\widehat{f}(\boldsymbol{\omega})$. Again by Fourier transforms of characteristic, delta, and Gaussian functions, it follows that

(a) if $f(\mathbf{t}) = \prod_1^d \frac{\sin(2\pi t_k T)}{\pi t_k}$, then $\widehat{f}(\boldsymbol{\omega}) = \chi_{[-T,T]^d}(\boldsymbol{\omega})$;
(b) if $f(\mathbf{t}) = e^{-\frac{\pi^2}{M}|\mathbf{t}|^2}$, then $\widehat{f}(\boldsymbol{\omega}) = e^{-M|\boldsymbol{\omega}|^2}$;
(c) if $f(\mathbf{t}) = 1$, then $\widehat{f}(\boldsymbol{\omega}) = \delta(\boldsymbol{\omega})$;

by (c) and the translation formula, it follows that

(d) if $f(\mathbf{t}) = e^{2\pi i(\boldsymbol{\tau},\mathbf{t})}$, then $\widehat{f}(\boldsymbol{\omega}) = \delta(\boldsymbol{\omega} - \boldsymbol{\tau})$;

by (d) and Euler's formula, it follows that

(e) if $f(\mathbf{t}) = \cos(2\pi(\boldsymbol{\tau},\mathbf{t}))$, then $\widehat{f}(\boldsymbol{\omega}) = \frac{1}{2}(\delta(\boldsymbol{\omega}-\boldsymbol{\tau}) + \delta(\boldsymbol{\omega}+\boldsymbol{\tau}))$;
(f) if $f(\mathbf{t}) = \sin(2\pi(\boldsymbol{\tau},\mathbf{t}))$, then $\widehat{f}(\boldsymbol{\omega}) = \frac{1}{2i}(\delta(\boldsymbol{\omega}-\boldsymbol{\tau}) - \delta(\boldsymbol{\omega}+\boldsymbol{\tau}))$.

If f and g are square integrable on $\mathbb{R}^d$ (i.e., $\int_{\mathbb{R}^d} |f(\mathbf{t})|^2 \mathrm{d}\mathbf{t} < \infty$ and $\int_{\mathbb{R}^d} |g(\mathbf{t})|^2 \mathrm{d}\mathbf{t} < \infty$), then

Parseval Formula: $\int_{\mathbb{R}^d} f(\mathbf{t})\overline{g(\mathbf{t})}\mathrm{d}\mathbf{t} = \int_{\mathbb{R}^d} \widehat{f}(\boldsymbol{\omega})\overline{\widehat{g}(\boldsymbol{\omega})}\mathrm{d}\boldsymbol{\omega}$;
Plancherel Formula: $\int_{\mathbb{R}^d} |f(\mathbf{t})|^2 \mathrm{d}\mathbf{t} = \int_{\mathbb{R}^d} |\widehat{f}(\boldsymbol{\omega})|^2 \mathrm{d}\boldsymbol{\omega}$.

Suppose that $f(\mathbf{t})$ is square integrable on $[0, 1]^d$ and $f(\mathbf{t}+\mathbf{n}) = f(\mathbf{t})$ for all $\mathbf{t} \in \mathbb{R}^d$ and $\mathbf{n} \in \mathbb{Z}^d$. Then, $f(\mathbf{t})$ can be expanded into Fourier series with respect to $\{e^{2\pi i(\mathbf{n},\mathbf{t})}\}_{\mathbf{n}\in\mathbb{Z}^d}$:

$$f(\mathbf{t}) = \sum_{\mathbf{n}\in\mathbb{Z}^d} c_{\mathbf{n}}(f)\, e^{2\pi i(\mathbf{n},\mathbf{t})} \qquad (\mathbf{n} = (n_1, \ldots, n_d),\ \mathbf{t} = (t_1, \ldots, t_d)),$$

where the inner product $(\mathbf{n}, \mathbf{t}) = \sum_{k=1}^d n_k t_k$, and Fourier coefficients are

$$c_{\mathbf{n}}(f) = \int_{[0,1]^d} f(\mathbf{t})\, e^{-2\pi i(\mathbf{n},\mathbf{t})} \mathrm{d}\mathbf{t}. \tag{2.1.3}$$

Fourier coefficients have the following important properties:

Riemann–Lebesgue Lemma:

$$c_{\mathbf{n}}(f) \to 0 \text{ as } \| \mathbf{n} \| \to \infty, \text{ where } \| \mathbf{n} \| = (n_1^2 + \cdots + n_d^2)^{\frac{1}{2}}.$$

Parseval Identity: $\sum_{\mathbf{n}\in\mathbb{Z}^d} |c_{\mathbf{n}}(f)|^2 = \int_{[0,1]^d} |f(\mathbf{t})|^2 \mathrm{d}\mathbf{t}.$

Convolution Formula:

$$c_{\mathbf{n}}(f * g) = c_{\mathbf{n}}(f)\, c_{\mathbf{n}}(g), \quad \text{where } f * g = \int_{[0,1]^d} f(\mathbf{t}-\mathbf{s})g(\mathbf{s})\mathrm{d}\mathbf{s}.$$

Poisson Summation Formula: $\sum_{\mathbf{m}\in\mathbb{Z}^d} f(\mathbf{t}+\mathbf{m}) = \sum_{\mathbf{n}\in\mathbb{Z}^d} \widehat{f}(\mathbf{n})\, e^{2\pi i(\mathbf{n},\mathbf{t})}.$

Shannon Sampling Theorem. If the Fourier transform of $f(\mathbf{t})$ satisfies $\widehat{f}(\boldsymbol{\omega}) = 0\ (\boldsymbol{\omega} \notin [-\frac{1}{2}, \frac{1}{2}]^d)$, then the following interpolation formula holds:

$$f(\mathbf{t}) = \sum_{\mathbf{n}\in\mathbb{Z}^d} f(\mathbf{n}) \prod_{k=1}^{d} \frac{\sin \pi(t_k - n_k)}{\pi(t_k - n_k)}.$$

More generally, if $\widehat{f}(\boldsymbol{\omega}) = 0\ (\boldsymbol{\omega} \notin [-\frac{b}{2}, \frac{b}{2}]^d)$ and the sampling interval $\Delta \le \frac{1}{b}$, then the interpolation formula becomes

$$f(\mathbf{t}) = \sum_{\mathbf{n}\in\mathbb{Z}^d} f(\mathbf{n}\Delta) \prod_{k=1}^{d} \frac{\sin \frac{\pi}{\Delta}(t_k - n_k\Delta)}{\frac{\pi}{\Delta}(t_k - n_k\Delta).}$$

If $|\widehat{f}(\boldsymbol{\omega})| < \epsilon\ (\boldsymbol{\omega} \notin [-\frac{b}{2}, \frac{b}{2}]^d)$, then this interpolation formula holds approximately and the error is less than $\frac{\epsilon}{\pi}$.

The Shannon sampling theorem allows the replacement of a continuous band-limited signal by a discrete sequence of its samples without the loss of any information. Moreover, it specifies the lowest sampling rate to reproduce the original signal.

2.2 Discrete Fourier Transform

The discrete Fourier transform (DFT) is a specific kind of Fourier transformation. It requires an input function that is discrete and whose nonzero values have a finite duration. Such inputs are often created by sampling a continuous function.

Start from Fourier coefficient formula (2.1.3), i.e.,

$$c_{\mu_1,\ldots,\mu_d}(f) = \int_{[0,1]^d} f(t_1, \ldots, t_d)\, e^{-2\pi i(\mu_1 t_1 + \cdots + \mu_d t_d)} \mathrm{d}t_1 \cdots \mathrm{d}t_d.$$

Take $(t_1, \dots, t_d) = (\frac{\nu_1}{N_1}, \dots, \frac{\nu_d}{N_d})$ $(\nu_l = 0, \dots, N_l - 1$ $(l = 1, \dots, d))$ as the lattice distribution of $[0, 1]^d$. The numerical calculation of the Fourier coefficients is

$$X_{\mu_1,\dots,\mu_d} := \frac{1}{N_1 \cdots N_d} \sum_{\nu_1=0}^{N_1-1} \cdots \sum_{\nu_d=0}^{N_d-1} f\left(\frac{\nu_1}{N_1}, \dots, \frac{\nu_d}{N_d}\right) e^{-2\pi i(\frac{\mu_1\nu_1}{N_1}+\cdots+\frac{\mu_d\nu_d}{N_d})}.$$

Denote $x_{\nu_1,\dots,\nu_d}$ as the value of $f(t_1, \dots, t_d)$ at the lattice point $(\frac{\nu_1}{N_1}, \dots, \frac{\nu_d}{N_d})$. So

$$X_{\mu_1,\dots,\mu_d} = \frac{1}{N_1 \cdots N_d} \sum_{\nu_1=0}^{N_1-1} \cdots \sum_{\nu_d=0}^{N_d-1} x_{\nu_1,\dots,\nu_d}\, e^{-2\pi i(\frac{\mu_1\nu_1}{N_1}+\cdots+\frac{\mu_d\nu_d}{N_d})}. \tag{2.2.1}$$

The above transform from $\{x_{\nu_1,\dots,\nu_d}\}_{\nu_k=0,\dots,N_k-1(k=1,\dots,d)}$ to $\{X_{\mu_1,\dots,\mu_d}\}_{\mu_k=0,\dots,N_k-1(k=1,\dots,d)}$ is called the *$d-$variate discrete Fourier transform* (DFT), denoted by $\mathrm{DFT}(x_{\nu_1,\dots,\nu_d})$.

If $x_{\nu_1,\dots,\nu_d}$ is separable $x_{\nu_1,\dots,\nu_d} = x_{\nu_1}^{(1)} x_{\nu_2}^{(2)} \cdots x_{\nu_d}^{(d)}$, then $\mathrm{DFT}(x_{\nu_1,\dots,\nu_d})$ is

$$X_{\mu_1,\dots,\mu_d} = \left(\frac{1}{N_1}\sum_{\nu_1=0}^{N_1-1} x_{\nu_1}^{(1)}\, e^{-2\pi i\frac{\mu_1\nu_1}{N_1}}\right)\cdots\left(\frac{1}{N_d}\sum_{\nu_d=0}^{N_d-1} x_{\nu_d}^{(d)}\, e^{-2\pi i\frac{\mu_d\nu_d}{N_d}}\right) =: X_{\mu_1}^{(1)} \cdots X_{\mu_d}^{(d)},$$

where $X_{\mu_k}^{(k)}$ is DFT of $x_{\nu_k}^{(k)}(k = 1, \dots, d)$. This implies $\mathrm{DFT}(x_{\nu_1,\dots,\nu_d})$ is also separable. If $N = N_i (i = 1, \dots, d)$, Formula (2.2.1) is reduced in the form:

$$X_{\boldsymbol{\mu}} = \frac{1}{N^d} \sum_{\boldsymbol{\nu}\in\Lambda_N} x_{\boldsymbol{\nu}}\, e^{-\frac{2\pi i}{N}(\boldsymbol{\mu},\boldsymbol{\nu})},$$

where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d) \in \Lambda_N$ and $\boldsymbol{\nu} = (\nu_1, \dots, \nu_d)$, and $\Lambda_N = [0, N)^d \bigcap \mathbb{Z}^d$.

The *$d-$variate inverse discrete Fourier transform*(IDFT) is

$$x_{\nu_1,\dots,\nu_d} = \sum_{\mu_1=0}^{N_1-1} \cdots \sum_{\mu_d=0}^{N_d-1} X_{\mu_1,\dots,\mu_d}\, e^{2\pi i(\frac{\mu_1\nu_1}{N_1}+\cdots+\frac{\mu_d\nu_d}{N_d})}.$$

For simplicity of notation, we only prove the case $d = 2$.

Exchanging the order of summation gives

$$\sum_{\mu_1=0}^{N_1-1}\sum_{\mu_2=0}^{N_2-1} X_{\mu_1,\mu_2}\, e^{2\pi i(\frac{\mu_1\nu_1}{N_1}+\frac{\mu_2\nu_2}{N_2})}$$

$$= \frac{1}{N_1N_2}\sum_{\mu_1=0}^{N_1-1}\sum_{\mu_2=0}^{N_2-1}\left(\sum_{\alpha_1=0}^{N_1-1}\sum_{\alpha_2=0}^{N_2-1} x_{\alpha_1,\alpha_2}\, e^{-2\pi i\left(\frac{\mu_1\alpha_1}{N_1}+\frac{\mu_2\alpha_2}{N_2}\right)}\right) e^{2\pi i\left(\frac{\mu_1\nu_1}{N_1}+\frac{\mu_2\nu_2}{N_2}\right)} \tag{2.2.2}$$

$$= \frac{1}{N_1N_2}\sum_{\alpha_1=0}^{N_1-1}\sum_{\alpha_2=0}^{N_2-1} x_{\alpha_1,\alpha_2}\left(\sum_{\mu_1=0}^{N_1-1}\sum_{\mu_2=0}^{N_2-1} e^{2\pi i\left(\frac{\nu_1-\alpha_1}{N_1}\mu_1+\frac{\nu_2-\alpha_2}{N_2}\mu_2\right)}\right).$$

Using the known formulas $e^{i(\alpha+\beta)} = e^{i\alpha}e^{i\beta}$ and $\sum_{\lambda=0}^{N-1} e^{\lambda\mu} = \frac{1-e^{N\mu}}{1-e^{\mu}}$ $(\mu \neq 0)$, the term in brackets becomes

$$\sum_{\mu_1=0}^{N_1-1} e^{2\pi i\,\frac{\nu_1-\alpha_1}{N_1}\mu_1}\sum_{\mu_2=0}^{N_2-1} e^{2\pi i\,\frac{\nu_2-\alpha_2}{N_2}\mu_2} = \begin{cases} 0, & (\alpha_1,\alpha_2)\neq(\nu_1,\nu_2), \\ N_1N_2, & (\alpha_1,\alpha_2)=(\nu_1,\nu_2). \end{cases}$$

Then, substituting this into (2.2.2) gives

$$\sum_{\mu_1=0}^{N_1-1}\sum_{\mu_2=0}^{N_2-1} X_{\mu_1\mu_2}\, e^{2\pi i\left(\frac{\mu_1\nu_1}{N_1}+\frac{\mu_2\nu_2}{N_2}\right)} = x_{\nu_1\nu_2}.$$

This means that the inverse discrete Fourier transform can reconstruct the original data from its discrete Fourier transform.

Let $\widetilde{x}_{\nu_1,\ldots,\nu_d}$ be a $d-$dimensional sequence on $\mathbb{Z}^d$ such that $x_{\nu_1+p_1N_1,\ldots,\nu_d+p_dN_d} = x_{\nu_1,\ldots,\nu_d}$, where $(N_1,\ldots,N_d)\in\mathbb{Z}^d$ and $(p_1,\ldots,p_d)\in\mathbb{Z}^d$. Then, $\widetilde{x}_{\nu_1,\ldots,\nu_d}$ is called a *recurring sequence with period* $(N_1,\ldots,N_d)$.

Discrete Fourier transforms (DFTs) have the following properties:

(a) Linearity. If $x_{\nu_1,\ldots,\nu_d}$ and $y_{\nu_1,\ldots,\nu_d}$ are two sequences, then

$$\text{DFT}(\alpha x_{\nu_1,\ldots,\nu_d}+\beta y_{\nu_1,\ldots,\nu_d}) = \alpha\,\text{DFT}(x_{\nu_1,\ldots,\nu_d})+\beta\,\text{DFT}(y_{\nu_1,\ldots,\nu_d}) \qquad (\alpha,\beta\in\mathbb{R}).$$

(b) Translation. If $\widetilde{x}_{\nu_1,\ldots,\nu_d}$ is a recurring sequence with period $(N_1,\ldots,N_d)$, then

$$\text{DFT}\,(\widetilde{x}_{\nu_1-\tau_1,\ldots,\nu_d-\tau_d}) = \text{DFT}\,(\widetilde{x}_{\nu_1,\ldots,\nu_d})\, e^{-2\pi i\left(\frac{\mu_1\tau_1}{N_1}+\cdots+\frac{\mu_d\tau_d}{N_d}\right)} \qquad ((\tau_1,\ldots,\tau_d)\in\mathbb{Z}^d).$$

If $\widetilde{X}_{\nu_1,\ldots,\nu_d} = \text{DFT}\,(\widetilde{x}_{\nu_1,\ldots,\nu_d})$, then

$$\text{IDFT}\,(\widetilde{X}_{\nu_1-\tau_1,\ldots,\nu_d-\tau_d}) = \text{IDFT}\,(\widetilde{X}_{\nu_1,\ldots,\nu_d})\, e^{2\pi i\left(\frac{\mu_1\tau_1}{N_1}+\cdots+\frac{\mu_d\tau_d}{N_d}\right)}.$$

Moreover, if $N = N_i (i = 1,\ldots,d)$, then

$$\text{DFT}\,(\widetilde{x}_{\boldsymbol{\nu}-\boldsymbol{\tau}}) = \text{DFT}\,(\widetilde{x}_{\boldsymbol{\nu}})\, e^{-\frac{2\pi i}{N}(\boldsymbol{\mu},\boldsymbol{\tau})},$$

$$\text{IDFT}\,(\widetilde{X}_{\boldsymbol{\nu}-\boldsymbol{\tau}}) = \text{IDFT}\,(\widetilde{X}_{\boldsymbol{\nu}})\, e^{\frac{2\pi i}{N}(\boldsymbol{\mu},\boldsymbol{\tau})},$$

where $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d)$, $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)$, and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)$.
(c) Symmetry. Let $X_{\mu_1,\ldots,\mu_d}$ be DFT of $x_{\nu_1,\ldots,\nu_d}$. Then, $\text{DFT}(\overline{x}_{\nu_1,\ldots,\nu_d}) = \overline{X}_{N_1-\mu_1,\ldots,N_d-\mu_d}$.
(d) Parseval Identity. Let $X_{\mu_1,\ldots,\mu_d}$ be DFT of $x_{\nu_1,\ldots,\nu_d}$. Then,

$$\sum_{\mu_1=0}^{N_1-1} \cdots \sum_{\mu_d=0}^{N_d-1} |X_{\mu_1,\ldots,\mu_d}|^2 = N \sum_{\nu_1=0}^{N_1-1} \cdots \sum_{\nu_d=0}^{N_d-1} |x_{\nu_1,\ldots,\nu_d}|^2.$$

(e) Convolutions. Let $\widetilde{x}_{\boldsymbol{\nu}}$, $\widetilde{y}_{\boldsymbol{\nu}}$ ($\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d)$) be both recurring sequences with periods $(N_1, \ldots, N_d)$ and $(M_1, \ldots, M_d)$, respectively. If $N = N_i = M_i$ $(i = 1, \ldots, d)$, the *recurring convolution* of $\widetilde{x}_{\boldsymbol{\nu}}$ and $\widetilde{y}_{\boldsymbol{\nu}}$ is defined as

$$\widetilde{h}_{\mathbf{k}} = (\widetilde{x}_{\boldsymbol{\nu}} * \widetilde{y}_{\boldsymbol{\nu}})_{\mathbf{k}} = \sum_{\boldsymbol{\nu}\in\Lambda_N} \widetilde{x}_{\mathbf{k}-\boldsymbol{\nu}} \widetilde{y}_{\boldsymbol{\nu}} \qquad (\mathbf{k} \in \mathbb{R}^d),$$

where $\Lambda_N = [0, N)^d \bigcap \mathbb{Z}^d$. Noticing that the DFT of $\widetilde{h}_{\mathbf{k}}$ is

$$\begin{aligned}
\widetilde{H}_{\mathbf{n}} &= \sum_{\mathbf{k}\in\Lambda_N} \left(\sum_{\boldsymbol{\nu}\in\Lambda_N} \widetilde{x}_{\mathbf{k}-\boldsymbol{\nu}} \widetilde{y}_{\boldsymbol{\nu}} \right) e^{-\frac{2\pi i(\mathbf{n},\mathbf{k})}{N}} \\
&= \sum_{\mathbf{k}\in\Lambda_N} \left(\sum_{\boldsymbol{\nu}\in\Lambda_N} \widetilde{x}_{\mathbf{k}} \widetilde{y}_{\boldsymbol{\nu}} \right) e^{-\frac{2\pi i(\mathbf{n},\mathbf{k}+\boldsymbol{\nu})}{N}} \\
&= \left(\sum_{\mathbf{k}\in\Lambda_N} \widetilde{x}_{\mathbf{k}}\, e^{-\frac{2\pi i(\mathbf{n},\mathbf{k})}{N}} \right) \left(\sum_{\boldsymbol{\nu}\in\Lambda_N} \widetilde{y}_{\boldsymbol{\nu}}\, e^{-\frac{2\pi i(\mathbf{n},\boldsymbol{\nu})}{N}} \right),
\end{aligned}$$

DFT of the recurring convolution is equal to $\widetilde{X}_{\mathbf{n}} \widetilde{Y}_{\mathbf{n}}$, i.e.,

$$\text{DFT}(\widetilde{x}_{\boldsymbol{\nu}} * \widetilde{y}_{\boldsymbol{\nu}}) = \text{DFT}(\widetilde{x}_{\boldsymbol{\nu}})\text{DFT}(\widetilde{y}_{\boldsymbol{\nu}}).$$

Let $\widetilde{x}_{\nu_1,\ldots,\nu_d}$ $(\nu_l = 0, \ldots, N_l - 1)$ and $\widetilde{y}_{\nu_1,\ldots,\nu_d}$ $(\nu_l = 0, \ldots, M_l - 1)$ be both $d-$dimensional sequences, and

$$P = \max\{N_1, \ldots, N_d, M_1, \ldots, M_d\}.$$

Take $Q = 2P - 1$. Let $\widetilde{x}_{\boldsymbol{\nu}}$ and $\widetilde{y}_{\boldsymbol{\nu}}$ be both recurring sequences with period $\mathbf{Q} = (Q, \ldots, Q)$ and

$$\widetilde{x}_{\nu_1,\ldots,\nu_d} = \begin{cases} x_{\nu_1,\ldots,\nu_d}, & \nu_l = 0, \ldots, N_l - 1, \\ 0, & \nu_l = N_l, \ldots, Q - 1 \quad (l = 1, \ldots, d), \end{cases}$$

$$\widetilde{y}_{\nu_1,\ldots,\nu_d} = \begin{cases} y_{\nu_1,\ldots,\nu_d}, & \nu_l = 0,\ldots,M_l-1, \\ 0, & \nu_l = M_l,\ldots,Q-1 \quad (l=1,\ldots,d). \end{cases}$$

The *linear convolution* of x_ν and y_ν is defined as recurring convolution of $\widetilde{x}_\nu$ and $\widetilde{y}_\nu$, i.e.,

$$\widetilde{h}_{\mathbf{k}} = (x_\nu * y_\nu)_{\mathbf{k}} = (\widetilde{x}_\nu * \widetilde{y}_\nu)_{\mathbf{k}} = \sum_{\boldsymbol{\nu}\in\Lambda_N} \widetilde{x}_{\mathbf{k}-\boldsymbol{\nu}}\widetilde{y}_{\boldsymbol{\nu}},$$

where $\mathbf{k} = (k_1,\ldots,k_d)$ and $\boldsymbol{\nu} = (\nu_1,\ldots,\nu_d)$. So the DFT of $h_{\mathbf{k}}$ is $\widetilde{H}_{\mathbf{n}} = \widetilde{X}_{\mathbf{n}}\widetilde{Y}_{\mathbf{n}}$ ($\mathbf{n} \in \Lambda_N$), i.e.,

$$\mathrm{DFT}(x_\nu * y_\nu) = \mathrm{DFT}(\widetilde{x}_\nu)\mathrm{DFT}(\widetilde{y}_\nu).$$

The computation of DFT in (2.2.1) needs $\prod_{k=1}^d N_k$ multiplications and $\prod_{k=1}^d (N_k - 1)$ summations, and the total number of the computation equals approximately $(\prod_{k=1}^d N_k)^2$. The *fast Fourier transform* (FFT) is a fast algorithm to compute DFT through reducing the complexity of computing as follows.

Note that

$$e^{-2\pi i\left(\frac{\mu_1\nu_1}{N_1}+\frac{\mu_2\nu_2}{N_2}+\cdots+\frac{\mu_d\nu_d}{N_d}\right)} = e^{-2\pi i\frac{\mu_1\nu_1}{N_1}} e^{-2\pi i\frac{\mu_2\nu_2}{N_2}} \cdots e^{-2\pi i\frac{\mu_d\nu_d}{N_d}}.$$

The (2.2.1) is rewritten in the form:

$$X_{\mu_1,\ldots,\mu_d} = \frac{1}{N_1}\sum_{\nu_1=0}^{N_1-1} e^{-2\pi i\frac{\mu_1\nu_1}{N_1}} \left(\frac{1}{N_2}\sum_{\nu_2=0}^{N_2-1} e^{-2\pi i\frac{\mu_2\nu_2}{N_2}} \cdots \left(\frac{1}{N_d}\sum_{\nu_d=0}^{N_d-1} x_{\nu_1,\ldots,\nu_d} e^{-2\pi i\frac{\mu_d\nu_d}{N_d}}\right)\cdots\right). \tag{2.2.3}$$

Denote by $R_1(\nu_1,\ldots,\nu_{d-1},\mu_d)$ the last sum, i.e.,

$$R_1(\nu_1,\ldots,\nu_{d-1},\mu_d) = \frac{1}{N_d}\sum_{\nu_d=0}^{N_d-1} x_{\nu_1,\ldots,\nu_d}\, e^{-2\pi i\frac{\mu_d\nu_d}{N_d}},$$

where $\nu_l = 0,\ldots,N_l-1$ $(l=1,\ldots,d-1)$. Regard $R_1(\nu_1,\ldots,\nu_{d-1},\mu_d)$ as a function of μ_d $(\mu_d = 0,\ldots,N_d-1)$. Denote $P_{\mu_d} = R_1(\nu_1,\ldots,\nu_{d-1},\mu_d)$, $p_{\nu_d} = x_{\nu_1,\ldots,\nu_d}$, and $w_{N_d} = e^{-2\pi i\frac{1}{N_d}}$. Then,

$$P_{\mu_d} = \frac{1}{N_d}\sum_{\nu_d=0}^{N_d-1} p_{\nu_d} w_{N_d}^{\mu_d\nu_d}.$$

Without loss of generality, assume $N_d = 2^M$. Otherwise, let $2^{M-1} \le N_d < 2^M$. We extend the original sequence $p_0,\ldots,p_{N_{d-1}}$ into a new sequence $p_0,\ldots,p_{N_d-1}$, $0,\ldots,0$ by adding $2^M - N_d$ zeros behind $p_0,\ldots,p_{N_d-1}$.

First, we decompose the 2^M-point sequences $\{p_{\nu_d}\}$ into two $2^{M-1}-$point sequences:

$$y = (p_0, p_2, \dots, p_{2^M-2}) =: (y_0, y_1, \dots, y_{2^{M-1}-1}),$$

$$z = (p_1, p_3, \dots, p_{2^M-1}) =: (z_0, z_1, \dots, z_{2^{M-1}-1}).$$

The first half P_{μ_d} ($\mu_d = 0, 1, \dots, 2^{M-1} - 1$) is computed as follows:

$$P_{\mu_d} = \tfrac{1}{2^M} \sum_{\nu=0}^{2^{M-1}-1} y_\nu w_{2^M}^{-2\nu\mu_d} + \tfrac{1}{2^M} \sum_{\nu=0}^{2^{M-1}-1} z_\nu w_{2^M}^{-(2\nu+1)\mu_d}$$

$$= \tfrac{1}{2}(Y_{\mu_d} + w_{2^M}^{\mu_d} Z_{\mu_d}) \qquad (\mu_d = 0, \dots, 2^{M-1} - 1);$$

the second half $P_{\mu_d}(\mu_d = 2^{M-1}, \dots, 2^M)$ is similarly computed as follows:

$$P_{\mu_d+2^{M-1}} = \frac{1}{2}(Y_{\mu_d} - w_{2^M}^{\mu_d} Z_{\mu_d}) \qquad (\mu_d = 0, \dots, 2^{M-1} - 1),$$

where

$$Y_{\mu_d} = \tfrac{1}{2^{M-1}} \sum_{\nu=0}^{2^{M-1}-1} y_\nu w_{2^{M-1}}^{\nu\mu_d},$$

$$Z_{\mu_d} = \tfrac{1}{2^{M-1}} \sum_{\nu=0}^{2^{M-1}-1} z_\nu w_{2^{M-1}}^{\nu\mu_d}$$

are the DFT of y and z, respectively. It means that the computation of P_{μ_d} or R_1 can be reduced to that of Y_{μ_d} and Z_{μ_d}.

Next, we rewrite $Y_{\mu_d}(\mu_d = 0, \dots, 2^{M-1} - 1)$ and Z_{μ_d} ($\mu_d = 0, \dots, 2^{M-1} - 1$), respectively, as follows:

$$Y_{\mu_d} = \tfrac{1}{2}(Y'_{\mu_d} + w_{2^{M-1}}^{\mu_d} Y''_{\mu_d}),$$

$$Y_{\mu_d+2^{M-2}} = \tfrac{1}{2}(Y'_{\mu_d} - w_{2^{M-1}}^{\mu_d} Y''_{\mu_d}) \qquad (\mu_d = 0, \dots, 2^{M-2} - 1)$$

and

$$Z_{\mu_d} = \tfrac{1}{2}(Z'_{\mu_d} + w_{2^{M-1}}^{\mu_d} Z''_{\mu_d}),$$

$$Z_{\mu_d+2^{M-2}} = \tfrac{1}{2}(Z'_{\mu_d} - w_{2^{M-1}}^{\mu_d} Z''_{\mu_d}) \qquad (\mu_d = 0, \dots, 2^{M-2} - 1),$$

where Y'_{μ_d} and Y''_{μ_d} are DFTs of $2^{M-2}-$even samples and $2^{M-2}-$odd samples of $\{Y_{\mu_d}\}$, respectively, and Z'_{μ_d} and Z''_{μ_d} are DFTs of $2^{M-2}-$even samples and $2^{M-2}-$odd samples of $\{Z_{\mu_d}\}$, respectively. It means that the computation of P_{μ_d} or R_1 can be reduced to that of Y'_{μ_d}, Y''_{μ_d}, Z'_{μ_d}, and Z''_{μ_d}.

Finally, by repeating the above procedure again and again, it terminates at the computation of DFT of one sample. This procedure gives a fast algorithm to compute $R_1(\nu_1, \ldots, \nu_{d-1}, \mu_d)$.

Let

$$R_2(\nu_1, \ldots, \nu_{d-2}, \mu_{d-1}, \mu_d) := \frac{1}{N_{d-1}} \sum_{\nu_{d-1}=0}^{N_{d-1}-1} R_1(\nu_1, \ldots, \nu_{d-1}, \mu_d)\, e^{-2\pi i \frac{\mu_{d-1}\nu_{d-1}}{N_{d-1}}}.$$

Using the same algorithm as in R_1, the $R_2(\nu_1, \ldots, \nu_{d-2}, \mu_{d-1}, \mu_d)$ can be computed fast. Continuing this procedure, we finally fast compute

$$X_{\mu_1,\ldots,\mu_d} = \frac{1}{N_1} \sum_{\nu_1=0}^{N_1-1} R_{d-1}(\nu_1, \mu_2, \ldots, \mu_d)\, e^{-2\pi i \frac{\mu_1\nu_1}{N_1}}.$$

The total number of operations in the whole procedure is equal approximately to

$$\left(\prod_{k=1}^{d} N_k\right) \log_2 \left(\prod_{k=1}^{d} N_k\right).$$

2.3 Discrete Cosine/Sine Transform

The DFT is a linear transform from a sequence $x_0, \ldots, x_{N-1}$ of real numbers to a sequence $X_0, \ldots, X_{N-1}$ of complex numbers. In order to reduce boundary effects, one often takes DFTs after even extension/odd extension of the original sequence $x_0, \ldots, x_{N-1}$ around end points. As a result, various discrete cosine/sine transforms (DCT/DST) are introduced. The DCT/DST is similar to the DFT: They are also the transform connecting between the time/spatial domain to the frequency domain.

2.3.1 Four Forms of DCTs

DCTs transform real numbers $x_0, \ldots, x_{N-1}$ into real numbers $X_0, \ldots, X_{N-1}$.

Define DCT-1 as

$$X_k = \sum_{n=0}^{N-1} \beta_n x_n \cos \frac{\pi n k}{N-1} \qquad (k = 0, \ldots, N-1), \tag{2.3.1}$$

where $\beta_0 = \beta_{N-1} = \frac{1}{2}$ and $\beta_n = 1$ $(n = 1, \ldots, N-2)$. Its inverse transform is

$$x_k = \frac{2}{N-1} \sum_{n=0}^{N-1} \beta_n X_n \cos \frac{\pi n k}{N-1} \qquad (k = 0, \ldots, N-1).$$

In fact, substituting this into the right-hand side of (2.3.1), by Euler's formula, we get

$$\begin{aligned} J_k :&= \sum_{n=0}^{N-1} \beta_n x_n \cos \tfrac{\pi n k}{N-1} \\ &= \tfrac{2}{N-1} \sum_{n=0}^{N-1} \sum_{l=0}^{N-1} \beta_n \beta_l X_l \cos \tfrac{\pi l n}{N-1} \cos \tfrac{\pi n k}{N-1} \\ &= \tfrac{1}{N-1} \sum_{l=0}^{N-1} \beta_l X_l \operatorname{Re} \left\{ \sum_{n=0}^{N-1} \beta_n \left(e^{i \frac{\pi n (l+k)}{N-1}} + e^{i \frac{\pi n (l-k)}{N-1}} \right) \right\}. \end{aligned}$$

By the summation formula of geometric series,

$$\operatorname{Re} \left\{ \sum_{n=0}^{N-1} \beta_n \left(e^{i \frac{\pi n (l+k)}{N-1}} + e^{i \frac{\pi n (l-k)}{N-1}} \right) \right\} = \begin{cases} 0, & l \neq k \ (k \neq 0, \ N-1), \\ N-1, & l = k \ (k \neq 0, \ N-1), \\ 2(N-1), & l = k \ (k = 0, N-1). \end{cases}$$

So $J_k = X_k$.

The even extension of the sequence $x_0, \ldots, x_{N-1}$ around x_{N-1} is $x_0, \ldots, x_{N-2}$, $x_{N-1}, x_{N-2}, \ldots, x_1$, and then the DCT-1 of $x_0, \ldots, x_{N-1}$ is equivalent to the DFT of its even extension around x_{N-1}.

In fact, let

$$y_n = \begin{cases} x_n, & n = 0, \ldots, N-1, \\ x_{2N-2-n}, & n = N, \ldots, 2N-3. \end{cases}$$

The DFT of $y_n (n = 0, \ldots, 2N-3)$ is equal to

$$\begin{aligned} Y_k &= \tfrac{1}{2N-2} \sum_{n=0}^{2N-3} y_n \, e^{-i \frac{2\pi n k}{2N-2}} \\ &= \tfrac{1}{2N-2} \left(y_0 + (-1)^k y_{N-1} + \sum_{n=1}^{N-2} y_n \left(e^{-i \frac{\pi n k}{N-1}} + e^{i \frac{\pi n k}{N-1}} \right) \right) \\ &= \tfrac{1}{N-1} \left(\tfrac{1}{2} (x_0 + (-1)^k x_{N-1}) + \sum_{n=1}^{N-2} x_n \cos \tfrac{\pi n k}{N-1} \right) = \tfrac{1}{N-1} X_k \quad (k = 0, \ldots, N-1). \end{aligned}$$

The *high-dimensional form of DCT-1* is

$$X_{k_1,\ldots,k_d} = \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} x_{n_1,\ldots,n_d} \prod_{j=1}^{d} \beta_{n_j} \cos \frac{\pi n_j k_j}{N_j - 1} \qquad (k_l = 1, \ldots, N_l - 1).$$

Its inverse transform is

$$x_{k_1,\ldots,k_d} = \frac{2^d}{(N_1-1)\cdots(N_d-1)} \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} X_{n_1,\ldots,n_d} \prod_{j=1}^{d} \beta_{n_j} \cos \frac{\pi n_j k_j}{N_j - 1},$$

where $\beta_{n_j} = \frac{1}{2}$ $(n_j = 0, N_j - 1)$ and $\beta_{n_j} = 1$ $(n_j = 1, \ldots, N_j - 2)$.

DCT-2 is the most commonly used form of discrete cosine transform. Define DCT-2 as

$$X_k = \sum_{n=0}^{N-1} x_n \cos \frac{\pi(n+\frac{1}{2})k}{N} \qquad (k = 0, \ldots, N-1). \tag{2.3.2}$$

The inverse formula is

$$x_k = \frac{2}{N} \sum_{n=0}^{N-1} \alpha_n X_n \cos \frac{\pi n(k+\frac{1}{2})}{N} \qquad (k = 0, \ldots, N-1), \tag{2.3.3}$$

where $\alpha_0 = \frac{1}{2}$, $\alpha_n = 1$ $(n = 1, \ldots, N-1)$.

In fact, substituting (2.3.3) into the right-hand side of (2.3.2),

$$\frac{1}{N} \sum_{l=0}^{N-1} \alpha_l X_l \sum_{n=0}^{N-1} 2 \cos \frac{\pi(n+\frac{1}{2})l}{N} \cos \frac{\pi(n+\frac{1}{2})k}{N} =: J_k.$$

Note that

$$\sum_{n=0}^{N-1} 2 \cos \frac{\pi(n+\frac{1}{2})l}{N} \cos \frac{\pi(n+\frac{1}{2})k}{N} = \begin{cases} 0, & k \neq l, \\ N, & k = l \neq 0, \\ 2N, & k = l = 0 \end{cases}$$

So $J_k = X_k$ $(k = 0, \ldots, N-1)$, i.e., (2.3.3) holds.

Given a sequence $x_0, \ldots, x_{N-1}$, one constructs a new sequence $y_0, \ldots, y_{4N-1}$ satisfying the following:

$$y_{2n} = 0 \qquad (n = 0, \ldots, N-1),$$

$$y_{2n+1} = x_n \qquad (n = 0, \ldots, N-1),$$

$$y_{4N-n} = y_n \qquad (n = 1, \ldots, 2N-1).$$

The DFT of the sequence $y_0, \ldots, y_{4N-1}$ is

$$\begin{aligned} Y_k &= \sum_{n=0}^{4N-1} y_n \, e^{-i\frac{2\pi nk}{4N}} \\ &= \sum_{n=0}^{2N-1} y_{2n+1} \, e^{-i\frac{2\pi(2n+1)k}{4N}} \\ &= 2\sum_{n=0}^{N-1} x_n \, \cos\frac{\pi(n+\frac{1}{2})k}{N} = 2X_k. \end{aligned}$$

So the DCT-2 of a sequence of N real numbers $x_0, \ldots, x_{N-1}$ is equivalent to the DFT of the sequence of $4N$ real numbers $y_0, \ldots, y_{4N-1}$ of even symmetry whose even-indexed elements are zero.

The *high-dimensional form of DCT-2* of a sequence $\{x_{n_1,\ldots,n_d}\}_{n_i=0,\ldots,N_i-1(i=1,\ldots,d)}$ is

$$X_{k_1,\ldots,k_d} = \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} x_{n_1,\ldots,n_d} \prod_{l=1}^{d} \cos\frac{\pi(n_l+\frac{1}{2})k_l}{N_l}.$$

Its inverse formula is

$$x_{k_1,\ldots,k_d} = \frac{2^d}{N_1\cdots N_d} \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} X_{n_1,\ldots,n_d} \prod_{l=1}^{d} \alpha_{n_l} \cos\frac{\pi n_l(k_l+\frac{1}{2})}{N_l},$$

where $\alpha_{n_l} = \frac{1}{2}$ $(n_l = 0)$ and $\alpha_{n_l} = 1$ $(n_l = 1, \ldots, N_l - 1)$ for $l = 1, \ldots, d$.

The DCT-3 is defined as the inverse transform of the DCT-2.

Define DCT-4 as

$$X_k = \sum_{n=0}^{N-1} x_n \cos\frac{\pi(n+\frac{1}{2})(k+\frac{1}{2})}{N} \qquad (k = 0, \ldots, N-1). \tag{2.3.4}$$

Its inverse formula is

$$x_k = \frac{2}{N}\left(\sum_{n=0}^{N-1} X_n \cos\frac{\pi(n+\frac{1}{2})(k+\frac{1}{2})}{N}\right) \qquad (k = 0, \ldots, N-1).$$

In fact, substituting this into the right-hand side of (2.3.4),

$$\begin{aligned} J_k : &= \sum_{n=0}^{N-1} x_n \cos \tfrac{\pi(n+\frac{1}{2})(k+\frac{1}{2})}{N} \\ &= \tfrac{2}{N} \sum_{n=0}^{N-1} \sum_{l=0}^{N-1} X_l \cos \tfrac{\pi(l+\frac{1}{2})(n+\frac{1}{2})}{N} \cos \tfrac{\pi(n+\frac{1}{2})(k+\frac{1}{2})}{N} \\ &= \tfrac{1}{N} \sum_{l=0}^{N-1} X_l \mathrm{Re} \left(e^{i\frac{\pi(l+k+1)}{2N}} \sum_{n=0}^{N-1} e^{i\frac{\pi n(l+k+1)}{N}} + e^{i\frac{\pi(l-k)}{2N}} \sum_{n=0}^{N-1} e^{i\frac{\pi n(l-k)}{N}} \right). \end{aligned} \tag{2.3.5}$$

Since

$$e^{i\frac{\pi(l+k+1)}{2N}} \sum_{n=0}^{N-1} e^{i\frac{\pi n(l+k+1)}{N}} = i \tfrac{1-(-1)^{l+k+1}}{2\sin\frac{\pi}{2N}(l+k+1)},$$

$$e^{i\frac{\pi(l-k)}{2N}} \sum_{n=0}^{N-1} e^{i\frac{\pi n(l-k)}{N}} = \begin{cases} i \frac{1-(-1)^{l-k}}{2\sin\frac{\pi}{2N}(l-k)}, & l \neq k, \\ N, & l = k, \end{cases}$$

the representative in large parentheses in (2.3.5) is equal to a pure imaginary number for $l \neq k$ and to N for $l = k$. So $J_k = \frac{1}{N} X_k$ $(k = 0, \ldots, N-1)$.

The *high-dimensional form of DCT-4* of a real sequence $\{x_{n_1,\ldots,n_d}\}_{n_i=0,\ldots,N_i-1(i=1,\ldots,d)}$ is defined as

$$X_{k_1,\ldots,k_d} = \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} x_{n_1,\ldots,n_d} \prod_{j=1}^{d} \cos \frac{\pi(n_j+\frac{1}{2})(k_j+\frac{1}{2})}{N_j} \quad (k_l = 0, \ldots, N_l - 1). \tag{2.3.6}$$

Using the one-dimensional DCT successively, the high-dimensional form of DCT-4 becomes

$$X_{k_1,\ldots,k_d} = \sum_{n_1=0}^{N_1-1} \cdots \left(\sum_{n_d=0}^{N_d-1} x_{n_1,\ldots,n_d} \cos \frac{\pi(n_d+\frac{1}{2})(k_d+\frac{1}{2})}{N_d} \right) \cdots \cos \frac{\pi(n_1+\frac{1}{2})(k_1+\frac{1}{2})}{N_1}.$$

Its inverse transform is

$$x_{k_1,\ldots,k_d} = \frac{2^d}{N_1 \cdots N_d} \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} X_{n_1,\ldots,n_d} \prod_{j=1}^{d} \cos \frac{\pi(n_j+\frac{1}{2})(k_j+\frac{1}{2})}{N_j}.$$

2.3.2 Four Forms of DSTs

Similar to DCTs, there are four forms of DSTs. The DST-1 of a sequence $x_0, \ldots, x_{N-1}$ is defined as

$$X_k = \sum_{n=0}^{N-1} x_n \sin \frac{\pi(n+1)(k+1)}{N+1} \qquad (k = 0, \ldots, N-1). \tag{2.3.7}$$

Its inverse formula is

$$x_k = \frac{2}{N+1} \sum_{n=0}^{N-1} X_n \sin \frac{\pi(n+1)(k+1)}{N+1} \qquad (k = 0, \ldots, N-1). \tag{2.3.8}$$

In fact, substituting (2.3.8) into the right-hand side of (2.3.7),

$$\begin{aligned} J_k :&= \frac{2}{N+1} \sum_{n=0}^{N-1} \sum_{j=0}^{N-1} X_l \sin \frac{\pi(l+1)(n+1)}{N+1} \sin \frac{\pi(n+1)(k+1)}{N+1} \\ &= \frac{1}{N+1} \sum_{l=0}^{N-1} X_l (A_{kl} - B_{kl}), \end{aligned}$$

where

$$A_{kl} = \mathrm{Re}\left(\sum_{n=0}^{N-1} e^{i\,\frac{\pi(n+1)(l-k)}{N+1}} \right) = \begin{cases} N, & l = k, \\ -1, & l \neq k,\ l-k \text{ is even}, \\ 0, & l-k \text{ is odd}, \end{cases}$$

$$B_{kl} = \mathrm{Re}\left(\sum_{n=0}^{N-1} e^{i\,\frac{\pi(n+1)(l+k+2)}{N+1}} \right) = \begin{cases} -1, & l = k, \\ -1, & l \neq k,\ l-k \text{ is even}, \\ 0, & l-k \text{ is odd}. \end{cases}$$

So $J_k = X_k$.

The *high-dimensional form of DST-1* is

$$X_{k_1,\ldots,k_d} = \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} x_{n_1,\ldots,n_d} \prod_{l=1}^{d} \sin\left(\frac{\pi}{N} (n_l+1)(k_l+1) \right).$$

The *high-dimensional form of DST-2* is

$$X_{k_1,\ldots,k_d} = \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} x_{n_1,\ldots,n_d} \prod_{l=1}^{d} \sin \frac{\pi(n_l+\frac{1}{2})(k_l+1)}{N_l}.$$

The *high-dimensional form of DST-3* is

$$X_{k_1,\ldots,k_d} = \left(\prod_{l=1}^{d} \frac{2}{N_l+1}\right)\left(\sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} x_{n_1,\ldots,n_d} \prod_{l=1}^{d} \gamma_{n_l} \sin \frac{\pi(n_l+1)(k_l+\frac{1}{2})}{N_l}\right),$$

where $\gamma_{n_l} = 1$ $(n_l = 0, \ldots, N_l - 2)$ and $\gamma_{n_l} = \frac{1}{2}$ $(n_l = N_l - 1)$.

The *high-dimensional form of DST-4* is

$$X_{k_1,\ldots,k_d} = \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_d=0}^{N_d-1} x_{n_1,\ldots,n_d} \prod_{l=1}^{d} \sin \frac{\pi(n_l+\frac{1}{2})(k_l+\frac{1}{2})}{N_l}.$$

2.4 Filtering

Climatologists and environmentalists might be interested in some phenomena that would be cyclic over a certain period, or exhibit slow changes, so filtering becomes a key preprocessing step of climate and environmental data. The design of filters is closely related to the Fourier transform of the $d-$dimensional sequence which is defined as

$$F(x_{n_1,\ldots,n_d}) = X(e^{i\omega_1}, \ldots, e^{i\omega_d}) = \sum_{n_1\in\mathbb{Z}} \cdots \sum_{n_d\in\mathbb{Z}} x_{n_1,\ldots,n_d}\, e^{-in_1\omega_1} \cdots e^{-in_d\omega_d},$$

where $X(e^{i\omega_1}, \ldots, e^{i\omega_d})$ is called the *frequency characteristic* of $h_{n_1,\ldots,n_d}$. Its inverse transform is defined as

$$\begin{aligned} &F^{-1}(X(e^{i\omega_1}, \ldots, e^{i\omega_d})) \\ &= \frac{1}{(2\pi)^d} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} X(e^{i\omega_1}, \ldots, e^{i\omega_d})\, e^{in_1\omega_1} \cdots e^{in_d\omega_d} \mathrm{d}\omega_1 \cdots \mathrm{d}\omega_d. \end{aligned}$$

With the help of Poisson summation formula in Sect. 2.1, it is clear that the Fourier transform of the $d-$dimensional sequence is a transform from time/spatial domain to frequency domain.

(i) Low-Pass Filter

For a filter $h_{n_1,\ldots,n_d}$, if its frequency characteristic $H(e^{i\theta_1}, \ldots, e^{i\theta_d})$ satisfies the condition:

$$H(e^{i\theta_1}, \ldots, e^{i\theta_d}) = \begin{cases} 1, \ \sqrt{\theta_1^2 + \cdots + \theta_d^2} \le R < \pi, \\ \\ 0, \ \text{otherwise}, \end{cases}$$

the filter $h_{n_1,\ldots,n_d}$ is called a *spherically symmetric low-pass filter*.

In the case $d = 1$. If h_n is a low-pass filter and its frequency characteristic is

$$H(e^{i\theta}) = \begin{cases} 1, \ |\theta| \le R < \pi, \\ \\ 0, \ \text{otherwise}, \end{cases}$$

then the low-pass filter is

$$h_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(e^{i\theta})\, e^{in\theta}\mathrm{d}\theta = \frac{1}{2\pi}\int_{-R}^{R} e^{in\theta}\mathrm{d}\theta = \frac{1}{2\pi}\int_{-R}^{R} e^{i\theta}\mathrm{d}\theta = \begin{cases} \frac{\sin(nR)}{\pi n}, & n \ne 0, \\ \frac{R}{\pi}, & n = 0, \end{cases}$$

In the case $d = 2$. If h_{n_1,n_2} is a low-pass filter and its frequency characteristic is

$$H(e^{i\theta_1}, e^{i\theta_2}) = \begin{cases} 1, \ \sqrt{\theta_1^2 + \theta_2^2} \le R < \pi, \\ \\ 0, \ \text{otherwise}, \end{cases}$$

then the low-pass filter is

$$\begin{aligned} h_{n_1,n_2} &= \tfrac{1}{(2\pi)^2} \int\int_{\sqrt{\theta_1^2+\theta_2^2} \le R} e^{in_1\theta_1}\, e^{in_2\theta_2}\, \mathrm{d}\theta_1 \mathrm{d}\theta_2 \\ &= \tfrac{1}{4\pi^2} \int_0^R r \left(\int_0^{2\pi} e^{ir(n_1\cos\theta + n_2\sin\theta)}\mathrm{d}\theta \right) \mathrm{d}r, \end{aligned} \tag{2.4.1}$$

where $\theta_1 = r\cos\theta$ and $\theta_2 = r\sin\theta$. Note that

$$n_1\cos\theta + n_2\sin\theta = \sqrt{n_1^2 + n_2^2}(\sin\varphi\cos\theta + \cos\varphi\sin\theta) = \sqrt{n_1^2 + n_2^2}\,\sin(\theta + \varphi),$$

where

$$\sin\varphi = \frac{n_1}{\sqrt{n_1^2 + n_2^2}}, \qquad \cos\varphi = \frac{n_2}{\sqrt{n_1^2 + n_2^2}},$$

and the integral representation of Bessel function of degree 0 is

$$J_0(t) = \frac{1}{2\pi}\int_0^{2\pi} e^{it\sin\theta}\mathrm{d}\theta.$$

By the periodicity of sine function, the inner integral of (2.4.1) is

$$\begin{aligned} \int_0^{2\pi} e^{ir(n_1\cos\theta + n_2\sin\theta)}\mathrm{d}\theta &= * \int_0^{2\pi} e^{ir\sqrt{n_1^2+n_2^2}\sin(\theta+\varphi)}\mathrm{d}\theta = \int_0^{2\pi} e^{ir(\sqrt{n_1^2+n_2^2})\sin\theta}\mathrm{d}\theta \\ &= J_0\left(r\sqrt{n_1^2 + n_2^2}\right). \end{aligned}$$

By (2.4.1), this follows that the low-pass filter is

$$h_{n_1,n_2} = \frac{1}{4\pi^2} \int_0^R r J_0 \left(r\sqrt{n_1^2 + n_2^2} \right) \mathrm{d}r,$$

where J_0 is the Bessel function of degree 0.

(ii) Band-Pass Filter
For a filter $h_{n_1,\ldots,n_d}$, if its frequency characteristic $H(e^{i\theta_1}, \ldots, e^{i\theta_d})$ satisfies the condition:

$$H(e^{i\theta_1}, \ldots, e^{i\theta_d}) = \begin{cases} 1, \ 0 < R_2 \leq \sqrt{\theta_1^2 + \cdots + \theta_d^2} \leq R_1 < \pi, \\ \\ 0, \ \text{otherwise}, \end{cases}$$

then the filter $h_{n_1,\ldots,n_d}$ is called a *spherically symmetric band-pass filter*.
In the case $d = 1$. The band-pass filter is

$$h_n = \begin{cases} \frac{1}{\pi n}(\sin(nR_1) - \sin(nR_2)), \ n \neq 0, \\ \\ \frac{1}{\pi}(R_1 - R_2), \qquad\qquad n = 0. \end{cases}$$

In the case $d = 2$. The band-pass filter is

$$h_{n_1,n_2} = \frac{1}{4\pi^2} \int_{R_2}^{R_1} r J_0 \left(r\sqrt{n_1^2 + n_2^2} \right) \mathrm{d}r,$$

where J_0 is the Bessel function of degree 0.

(iii) High-Pass Filter
For a filter $h_{n_1,\ldots,n_d}$, if its frequency characteristic $H(e^{i\theta_1}, \ldots, e^{i\theta_d})$ satisfies the condition:

$$H(e^{i\theta_1}, \ldots, e^{i\theta_d}) = \begin{cases} 1, \ 0 < R \leq \sqrt{\theta_1^2 + \cdots + \theta_d^2} < \pi, \\ \\ 0, \ \text{otherwise}, \end{cases}$$

the filter $h_{n_1,\ldots,n_d}$ is called a *spherically symmetric high-pass filter*.
In the case $d = 1$. The high-pass filter is

$$h_n = \begin{cases} -\frac{\sin(nR)}{\pi n}, \ n \neq 0 \\ \\ 1 - \frac{R}{\pi}, \quad n = 0. \end{cases}$$

In the case $d = 2$. The high-pass filter is

$$h_{n_1,n_2} = \begin{cases} -\frac{1}{4\pi^2}\int_0^R r\,J_0\left(r\sqrt{n_1^2+n_2^2}\right)\,\mathrm{d}r, & n_1 \neq 0 \text{ or } n_2 \neq 0, \\ 1 - \frac{R^2}{4\pi}, & n_1 = n_2 = 0, \end{cases}$$

where J_0 is the Bessel function of degree 0.

2.5 Fractional Fourier Transform

Fractional Fourier transform (FRFT) is an extension of Fourier transform. It is thought of as Fourier transform to αth power. FRFT can transform a time series into the domain between time and frequency, so it demonstrates obvious advantages over Fourier transform.

2.5.1 Continuous FRFT

Starting from successive Fourier transform operator F, we study its nth iterated F^n given by

$$F^n(f) = F(f^{n-1}(f)), \qquad F^{-n} = (F^{-1})^n,$$

where n is a nonnegative integer and $F^0(f) = f$. From Fourier transform operator F and its inverse operator F^{-1}, it follows that

$$F^2(f)(t) = F \circ F(f)(t) = f(-t),$$

$$F^3(f)(t) = F^{-1}(f)(t),$$

$$F^4(f))(t) = f(t).$$

FRFT provides a family of linear transforms that further extend Fourier transform to handle non-integer power $n = \frac{2\alpha}{\pi}$ of the Fourier transform. For a $\alpha \in \mathbb{R}$, the $\alpha-$angle FRFT of a function f is defined as

$$F_\alpha(f)(\omega) = \int_{\mathbb{R}} K_\alpha(t,\omega) f(t)\mathrm{d}t$$

and $K_\alpha(t,\omega)$ is called *kernel function* and

$$K_\alpha(t,\omega) = \sqrt{1 - i\cot\alpha}\, e^{\pi i(\omega^2+t^2)\cot\alpha - 2\pi i\omega t\csc\alpha}. \tag{2.5.1}$$

It is clear that $F_{\frac{\pi}{2}}(f)(\omega)$ is the Fourier transform and $F_{-\frac{\pi}{2}}(f)(\omega)$ is the inverse Fourier transform. When α is an integer multiple of π,

$$(F_\alpha f)(\omega) = \begin{cases} f(\omega), & \text{if } \alpha \text{ is a multiple of } 2\pi, \\ f(-\omega), & \text{if } \alpha + \pi \text{ is a multiple of } 2\pi. \end{cases}$$

In FRFT, the variable ω is neither a time nor a frequency, and it is an interpolation between time and frequency. Note that $F_0 f$ is f itself, $F_{\frac{\pi}{2}} f$ is Fourier transform of f, $F_\pi f$ results in an inversion of the time axis of f, and $F_{2\pi} f$ is f itself. So $F_\alpha f$ is regarded as a counterclockwise rotation of the axis by an angle α of f in the time–frequency domain.

(i) FRFT and Eigenfunctions of Fourier Transform
Hermite–Gaussian functions are defined as

$$\psi_k(t) = \frac{2^{\frac{1}{4}}}{\sqrt{2^k k!}} H_k(\sqrt{2\pi}\, t)\, e^{-\pi t^2} \qquad (k = 0, 1, \ldots), \tag{2.5.2}$$

where H_k is the kth Hermite polynomial:

$$H_k(t) = e^{t^2} \frac{d^n (e^{-t^2})}{dt^n}.$$

Hermite–Gaussian functions are the unique finite energy eigensolutions of the Hermite–Gaussian equation:

$$\frac{d^2 f(t)}{dt^2} - 4\pi^2 t^2 f(t) = \lambda f(t). \tag{2.5.3}$$

Let $D = \frac{d}{dt}$, F denote the Fourier transform operator, and $S = D^2 + F D^2 F^{-1}$. Then, the Eq. (2.5.3) can be rewritten in the form $Sf(t) = \lambda f(t)$, and $\psi_k (k = 0, 1, \ldots)$ are the eigenfunctions of the operator S. Since the operators S and F are exchangeable and two exchangeable operators must have a common eigenvector set, the Fourier transform operator F has eigenfunctions $\psi_k(t) (k = 0, 1, \ldots)$, and the eigenvalue corresponding to $\psi_k(t)$ is $e^{-i\frac{\pi k}{2}}$, i.e.,

$$F(\psi_k)(\omega) = e^{-i\frac{\pi k}{2}} \psi_k(\omega). \tag{2.5.4}$$

Since $\{\psi_k(t)\}_{k=0,1,\ldots}$ can form a normal orthogonal basis for $L^2(\mathbb{R})$, any $f \in L^2(\mathbb{R})$ can be expanded into a series

$$f(t) = \sum_{k=0}^{\infty} c_k \psi_k(t),$$

where the coefficients $c_k = \int_{\mathbb{R}} f(t) \psi_k(t) \mathrm{d}t$. Taking Fourier transform on both sides gives

$$F(f)(\omega)=\sum_{k=0}^{\infty}c_kF(\psi_k)(\omega)=\sum_{k=0}^{\infty}c_ke^{-i\frac{\pi k}{2}}\psi_k(\omega)=\int_{\mathbb{R}}K(t,\omega)f(t)\mathrm{d}t, \quad (2.5.5)$$

where $K(t,\omega)=\sum_{k=0}^{\infty}\psi_k(\omega)\psi_k(t)\,e^{-i\frac{\pi k}{2}}$. For FRFT, its kernel function $K_\alpha(t,\omega)$ can be expanded into

$$K_\alpha(t,\omega)=\sum_{k=0}^{\infty}\psi_k(\omega)\psi_k(t)\,e^{-i\alpha\frac{\pi k}{2}}. \quad (2.5.6)$$

This means that FRFT can be derived by the eigenfunctions of Fourier transform.

(ii) Index Additivity
Successive applications of FRFT are equivalent to a single transform whose order is the sum of individual orders, i.e., $F_\alpha\circ F_\beta=F_\alpha(F_\beta)=F_{\alpha+\beta}$.

In fact,

$$F_\alpha\circ F_\beta(u)=\int_{\mathbb{R}}\left(\int_{\mathbb{R}}K_\alpha(u,\omega)K_\beta(t,\omega)\mathrm{d}\omega\right)f(t)\mathrm{d}t.$$

Since $\int_{\mathbb{R}}\psi_k(\omega)\psi_l(\omega)\mathrm{d}\omega=\delta_{k,l}$, by the expansion (2.5.6) of the kernel function $K_\alpha(t,\omega)$, it follows that

$$\int_{\mathbb{R}}K_\alpha(u,\omega)K_\beta(t,\omega)\mathrm{d}\omega=\sum_{k=0}^{\infty}e^{-i\frac{\pi k}{2}(\alpha+\beta)}\psi_k(u)\psi_k(t)=K_{\alpha+\beta}(u,t).$$

So

$$F_\alpha\circ F_\beta(u)=\int_{\mathbb{R}}K_{\alpha+\beta}(u,t)f(t)\mathrm{d}t=F_{\alpha+\beta}(u).$$

According to the index additivity, FRFT has the following properties:

(a) Inverse. $f(t)=\int_{\mathbb{R}}F_\alpha(f)(\omega)K_{-\alpha}(\omega,t)\mathrm{d}\omega$.
(b) Parseval Identity. $\int_{\mathbb{R}}f(t)\overline{g(t)}\mathrm{d}t=\int_{\mathbb{R}}F_\alpha(f)(\omega)\overline{F_\alpha(g)(\omega)}\mathrm{d}\omega$.
(c) Shift. Let $g(t)=f(t-\tau)$. Then,

$$F_\alpha(g)(\omega)=e^{i\pi\tau^2\sin\alpha\cos\alpha}\,e^{2\pi i\omega\tau\sin\alpha}F_\alpha(f)(\omega-\tau\cos\alpha).$$

(d) Modulation. Let $g(t)=f(t)\,e^{i\eta t}$. Then,

$$F_\alpha(g)(\omega)=e^{-i\pi\eta^2\sin\alpha\cos\alpha}e^{2\pi i\omega\eta\cos\alpha}F_\alpha(f)(\omega-\eta\sin\alpha).$$

(e) Dilation. Let $g(t)=f(ct)$. Then,

$$F_\alpha(g)(\omega)=\sqrt{\frac{1-i\cot\alpha}{c^2-i\cot\alpha}}\,e^{i\pi\cot\alpha\left(1-\frac{\cos^2\beta}{\cos^2\alpha}\right)}F_\beta(f)\left(\frac{\omega\sin\beta}{c\sin\alpha}\right),$$

where $\cot\beta=c^{-2}\cot\alpha$.

2.5.2 Discrete FRFT

The discrete FRFT is a discrete version of the continuous FRFT. It can also be thought as DFT to αth power. The algorithm of discrete FRFT is based on a special set of eigenvectors of DFT matrix.

A normalized discrete Fourier transform

$$X_\mu = N^{-\frac{1}{2}} \sum_{\nu=0}^{N-1} x_\nu \, e^{-2\pi i \frac{\mu\nu}{N}} \qquad (\mu = 0, 1, \ldots, N-1) \tag{2.5.7}$$

can be rewritten into the matrix form $\mathbf{X} = F\mathbf{x}$, where

$$\mathbf{X} = (X_0, \ldots, X_{N-1})^T, \quad \mathbf{x} = (x_0, \ldots, x_{N-1})^T, \quad F = \left(N^{-\frac{1}{2}} \, e^{-2\pi i \frac{\mu\nu}{N}}\right)_{\mu,\nu=0,\ldots,N-1}.$$

Let $\{\lambda_l\}_{l=0,\ldots,N-1}$ and $\{\mathbf{p}_l\}_{l=0,\ldots,N-1}$ be the eigenvalues and eigenvectors of the DFT matrix F, respectively, where the eigenvectors $\{\mathbf{p}_l\}_{l=0,\ldots,N-1}$ form a normal orthogonal basis for $\mathbb{R}^N$. For any $\mathbf{x} \in \mathbb{R}^N$,

$$\mathbf{x} = \sum_{l=0}^{N-1} c_l \mathbf{p}_l,$$

where $c_l = (\mathbf{p}_l, \mathbf{x}) = \mathbf{p}_l^T \mathbf{x}$. Taking DFT on both sides,

$$F\mathbf{x} = \sum_{l=0}^{N-1} c_l F \mathbf{p}_l = \sum_{l=0}^{N-1} c_l \lambda_l \mathbf{p}_l = \sum_{l=0}^{N-1} \mathbf{p}_l \lambda_l \mathbf{p}_l^T \mathbf{x}.$$

Therefore, the *spectral decomposition* of DFT matrix F is

$$F = \sum_{l=0}^{N-1} \mathbf{p}_l \lambda_l \mathbf{p}_l^T \qquad \text{or} \qquad F(\mu, \nu) = \sum_{l=0}^{N-1} \mathbf{p}_l(\mu) \lambda_l \mathbf{p}_l(\nu),$$

where $F = (F(\mu, \nu))_{\mu,\nu=0,\ldots,N-1}$ and $\mathbf{p}_l = (p_l(0), \ldots, p_l(N-1))^T$.

The *discrete FRTF matrix* F^α is defined as

$$F^\alpha = \sum_{l=0}^{N-1} \mathbf{p}_l \lambda_l^\alpha \mathbf{p}_l^T \quad \text{or} \quad F^\alpha(\mu, \nu) = \sum_{l=0}^{N-1} \mathbf{p}_l(\mu) \lambda_l^\alpha \mathbf{p}_l(\nu) \quad (\alpha \in \mathbb{R}), \tag{2.5.8}$$

where $\lambda_l^\alpha = e^{-i \frac{\pi l \alpha}{2}}$ and $F^\alpha = (F^\alpha(\mu, \nu))_{\mu,\nu=0,\ldots,N-1}$ and $\mathbf{p}_l = (p_l(0), \ldots, p_l(N-1))^T$. It is clear that F^1 is the DFT matrix F. The *discrete FRFT of order* α is defined as

$$X^\alpha(\mu) = \sum_{\nu=0}^{N-1} x_\nu F^\alpha(\mu, \nu).$$

(i) Index Additivity
For the discrete FRFT, the index additivity is $F^\alpha \circ F^\beta(f) = F^{\alpha+\beta}(f)$ for any $f \in \mathbb{R}^N$.

In fact, by (2.5.8), since $\{p_k\}_{k=0,\dots,N-1}$ are a normal orthogonal basis for $\mathbb{R}^N$, we get

$$\begin{aligned}(F^\alpha \circ F^\beta)(\mu, \nu) &= \sum_{n=0}^{N-1} F^\alpha(\mu, n) F^\beta(n, \nu)\\ &= \sum_{l=0}^{N-1} \sum_{s=0}^{N-1} \mathbf{p}_l(\mu)\lambda_l^\alpha \lambda_s^\beta \mathbf{p}_s(\nu) \left(\sum_{n=0}^{N-1} \mathbf{p}_l(n)\mathbf{p}_s(n) \right)\\ &= \sum_{l=0}^{N-1} \mathbf{p}_l(\mu)\lambda_l^{\alpha+\beta} \mathbf{p}_l(\nu) = F^{\alpha+\beta}(\mu, \nu).\end{aligned}$$

So $F^{\alpha+\beta} f = F^\alpha \circ F^\beta(f)$ for any $f \in \mathbb{R}^N$. Since $(p_l(\mu))_{l,\mu}$ is an orthogonal matrix, by (2.5.8), $F^0(\mu, \nu) = \delta_{\mu,\nu}$, and so $F^0 = I$.

According to the index additivity, the inversion of F^α is equal to $F^{-\alpha}$, i.e., $(F^\alpha)^{-1} = F^{-\alpha}$.

(ii) Discrete Hermite–Gaussian Functions
Note that the set of eigenvectors for DFT matrix F is not unique. In order to make the discrete FRFT completely analogous to continuous FRFT, we give the discrete form of the Hermite–Gaussian equation (2.5.3).

The first term of (2.5.3) is equal approximately to

$$\frac{d^2 f}{dt^2} \approx \widetilde{D}^2 f(t) = \frac{f(t+h) - 2f(t) + f(t-h)}{h^2}.$$

The second term of (2.5.3) is equal approximately to $(-4\pi^2 t^2) f(t) \approx (F \widetilde{D}^2 F^{-1}) f(t)$. So the discrete form of the Hermite–Gaussian equation is

$$(\widetilde{D}^2 + F \widetilde{D}^2 F^{-1}) f(t) = \lambda f(t). \tag{2.5.9}$$

Using the Taylor formula,

$$\widetilde{D}^2 f(t) = \frac{f(t+h) - 2f(t) + f(t-h)}{h^2} = \frac{2}{h^2} \sum_{n=1}^{\infty} \frac{d^{2n} f(t)}{dt^{2n}} \frac{h^{2n}}{(2n)!}. \tag{2.5.10}$$

Substituting $F^{-1} f$ into f in (2.5.10) and then taking Fourier transform on both sides, we get

$$(F\widetilde{D}^2F^{-1})f(t) = F\left(\frac{2}{h^2}\sum_{n=1}^{\infty}\frac{d^{2n}(F^{-1}f)(t)}{dt^{2n}}\frac{h^{2n}}{(2n)!}\right) = \frac{2}{h^2}\sum_{n=1}^{\infty}F\left(\frac{d^{2n}(F^{-1}f)(t)}{dt^{2n}}\right)\frac{h^{2n}}{(2n)!}.$$

Using the derivative formula of Fourier transform:

$$F\left(\frac{d^{2n}(F^{-1}f)(t)}{dt^{2n}}\right) = (2\pi it)^{2n}F(F^{-1}f)(t) = (2\pi it)^{2n}f(t)$$

and the known formula: $\cos(2\pi ht) = 1 + \sum_{n=1}^{\infty}(-1)^n(2\pi t)^{2n}\frac{h^{2n}}{(2n)!}$, we get

$$(F\widetilde{D}^2F^{-1})f(t) = \frac{2}{h^2}\left(\sum_{n=1}^{\infty}(-1)^n(2\pi t)^{2n}\frac{h^{2n}}{(2n)!}\right)f(t) = \frac{2}{h^2}(\cos(2\pi ht)-1)f(t).$$

So (2.5.9) is equivalent to

$$f(t+h) - 2f(t) + f(t-h) + 2(\cos(2\pi ht)-1)f(t) = h^2\lambda f(t). \tag{2.5.11}$$

Now, we find the solution of (2.5.11).

Let $t = nh$, $g(n) = f(nh)$, and $h = \frac{1}{\sqrt{N}}$. Then, (2.5.11) becomes:

$$g(n+1) + g(n-1) + 2\left(\cos\frac{2\pi n}{N} - 2\right)g(n) = \lambda g(n). \tag{2.5.12}$$

Since coefficients of (2.5.12) are periodic, its solutions are periodic. Let $n = 0, \ldots, N-1$. Then, for any $\mathbf{g} \in \mathbb{R}^N$,

$$M\mathbf{g} = \lambda\mathbf{g}, \tag{2.5.13}$$

where $\mathbf{g} = (g(0), \ldots, g(N-1))^T$ and $M = (\gamma_{kl})_{k,l=0,\ldots,N-1}$, and

$$\begin{cases} \gamma_{kk} = 2\cos\frac{2\pi k}{N} - 4 \qquad (k = 0, \ldots, N-1), \\ \gamma_{k,k+1} = \gamma_{k+1,k} = 1 \qquad (k = 0, \ldots, N-1), \\ \gamma_{0,N-1} = \gamma_{N-1,0} = 1, \\ \gamma_{k,l} = 0 \qquad (\text{otherwise}). \end{cases}$$

Namely,

$$M = \begin{pmatrix} 2\cos\frac{2\pi 0}{N}-4 & 1 & 0 & \cdots & \cdots & 0 & 1 \\ 1 & 2\cos\frac{2\pi}{N}-4 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 2\cos\frac{4\pi}{N}-4 & 1 & 0 & \cdots & 0 \\ & \vdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \cdots & 1 & 0 \\ 0 & \vdots & \vdots & 0 & 1 & 2\cos\frac{2(N-2)\pi}{N}-4 & 1 \\ 1 & 0 & \cdots & \cdots & 0 & 1 & 2\cos\frac{2(N-1)\pi}{N}-4 \end{pmatrix}$$

It is easy to prove that the matrices M and F are exchangeable, i.e., $MF = FM$. So the matrices M and F have a common eigenvector set $\{u_k(n)\}_{k=0,\ldots,N-1}$ which is called the *discrete Hermite–Gaussians*. This eigenvector set forms a normal orthogonal basis for $\mathbb{R}^N$. The corresponding discrete FRFT becomes

$$F^{\alpha}(\mu,\nu) = \sum_{k=0}^{N-1} u_k(\mu)\, e^{-i\pi\frac{k\alpha}{2}} u_k(\nu).$$

(iii) Algorithm

We introduce a matrix P that maps the even part of $\{g(n)\}_{n=0,\ldots,N-1}$ in (2.5.13) to the first $\left[\frac{N}{2}\right]+1$ components and its odd part to the remaining components. Since the matrix P is orthogonal symmetric, the similarity transform of the matrix M in (2.5.13) has a block diagonal form as follows:

$$PMP^{-1} = PMP = \begin{pmatrix} E_v & 0 \\ 0 & O_d \end{pmatrix}$$

The eigenvectors of PMP^{-1} can be determined separately from E_v and O_d matrices. Sort the eigenvectors of the matrices E_v and O_d in the descending order of eigenvalues. Denote by $\mathbf{e}_k$ $(k = 0, \ldots, \left[\frac{N}{2}\right])$ and $\mathbf{o}_k$ $(k = \left[\frac{N}{2}\right]+1, \ldots, N-1)$ the corresponding eigenvectors, respectively. Then, the even and odd eigenvectors of M can be formed, respectively, by

$$\begin{aligned} u_{2k} &= P(\mathbf{e}_k^T, 0, \ldots, 0)^T, \\ u_{2k+1} &= P(0, \ldots, 0, \mathbf{o}_k^T)^T. \end{aligned} \tag{2.5.14}$$

Finally, the discrete FRFT matrix is given by

$$F^{\alpha}(\mu,\nu)=\begin{cases}\sum\limits_{k=0}^{N-1}u_k(\mu)\,e^{-i\frac{\pi}{2}k\alpha}u_k(\nu) & (N \text{ is odd}),\\ \left(\sum\limits_{k=0}^{N-2}u_k(\mu)\,e^{-i\frac{\pi}{2}k\alpha}u_k(\nu)\right)+u_N(\mu)\,e^{-i\frac{\pi}{2}N\alpha}u_N(\nu) & (N \text{ is even}).\end{cases} \tag{2.5.15}$$

The algorithm of discrete FRFT:
Step 1. For given N, write the matrices M and P [see (2.5.13) and (2.5.14)].
Step 2. Calculate the matrices E_v and O_d defined by

$$PMP=\begin{pmatrix}E_v & 0\\ 0 & O_d\end{pmatrix}$$

and sort the eigenvectors of the matrices E_v and O_d in the descending order of eigenvalues and denote the sorted eigenvectors by $\mathbf{e}_k$ and $\mathbf{o}_k$, respectively.
Step 3. Calculate the discrete Hermite–Gaussian functions as u_k by (2.5.14).
Step 4. The $\alpha-$order discrete FRFT is

$$(F^{\alpha}\mathbf{f})(\mu)=\sum_{\nu=0}^{N-1}F^{\alpha}(\mu,\nu)f(\nu),$$

where $F^{\alpha}(\mu,\nu)$ is stated in (2.5.15) and $\mathbf{f}=(f(0),\ldots,f(N-1))^T$.

2.5.3 Multivariate FRFT

For $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_d)\in\mathbb{R}^d$, the *multivariate FRFT* of a $d-$variate function $f(\mathbf{t})$ is defined as

$$F_{\boldsymbol{\alpha}}(\boldsymbol{\omega})=\int_{\mathbb{R}^d}f(\mathbf{t})K_{\boldsymbol{\alpha}}(\mathbf{t},\boldsymbol{\omega})\mathrm{d}\mathbf{t}\qquad (\mathbf{t}=(t_1,\ldots,t_d),\ \boldsymbol{\omega}=(\omega_1,\ldots,\omega_d)),$$

where the kernel is

$$K_{\boldsymbol{\alpha}}(\mathbf{t},\boldsymbol{\omega})=\prod_{k=1}^{d}\left(\sqrt{1-i\cot\alpha_k}\,e^{\pi i(\omega_k^2+t_k^2)\cot\alpha_k-2\pi i\omega_k t_k\csc\alpha_k}\right).$$

Its inverse formula is $f(\mathbf{t})=\int_{\mathbb{R}^d}F_{\boldsymbol{\alpha}}(\boldsymbol{\omega})\,K_{-\boldsymbol{\alpha}}(\boldsymbol{\omega},\mathbf{t})\mathrm{d}\boldsymbol{\omega}$.

For $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_d)\in\mathbb{R}^d$, the multivariate DFRFT of a $d-$dimensional sequence $\{x_{\nu}\}_{\nu}$ is defined as

$$X^{\alpha_1,\ldots,\alpha_d}(\mu_1,\ldots,\mu_d)=\sum_{\nu_d=0}^{N-1}\cdots\sum_{\nu_1=0}^{N-1}x_{\nu_1,\ldots,\nu_d}\prod_{j=1}^{d}F^{\alpha_j}(\mu_j,\nu_j),$$

where $F^{\alpha_j}(\mu_j,\nu_j)$ is stated in (2.5.15), $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_d)$, $\boldsymbol{\nu}=(\nu_1,\ldots,\nu_d)$, and $\boldsymbol{\mu},\ \boldsymbol{\nu}\in([0,N-1]^d\bigcap\mathbb{Z}^d)$. The inverse transform of multivariate DFRFT is

$$x_{\nu_1,\ldots,\nu_d}=\sum_{\mu_d=0}^{N-1}\cdots\sum_{\mu_1=0}^{N-1}X^{\alpha_1,\ldots,\alpha_d}(\mu_1,\ldots,\mu_d)\prod_{j=1}^{d}F^{-\alpha_j}(\mu_j,\nu_j).$$

2.6 Space–Frequency Distribution

In multivariate time series analysis, one important topic is to research how the dominant frequencies of the variations of time series analysis change with time/spatial scales. In this section, we will introduce various transforms which can give information about time series simultaneously in the time/spatial domain and the frequency domain, including windowed Fourier transform, Wigner–Ville distribution, page distribution, Levin distribution.

2.6.1 Multivariate Windowed Fourier Transform

The multivariate Fourier transform can provide only global frequency information. In order to overcome this disadvantage, multivariate windows are used to modify Fourier transform. One always assumes that these multivariate windows are even, real-valued. For example, Gaussian window function:

$$W(\mathbf{t})=\left(\frac{1}{2\sqrt{\pi\alpha}}\right)^d e^{-\frac{\|\mathbf{t}\|^2}{4\alpha}}\qquad(\alpha>0),$$

is the most used multivariate window. Other multivariate windows include:

(a) Rectangular window: $\prod_{k=1}^{N}\chi_{-\frac{1}{2},\frac{1}{2}}(t_k)$;

(b) Hamming window: $\prod_{k=1}^{N}(0.54+0.46\cos(2\pi t_k))\chi_{[-\frac{1}{2},\frac{1}{2}]}(t_k)$;

(c) Hanning window: $\prod_{k=1}^{N}\cos^2(\pi t_k)\chi_{-\frac{1}{2},\frac{1}{2}}(t_k)$;

(d) Blackman window: $\prod_{k=1}^{N}(0.42+0.5\cos(2\pi t_k)+0.08\cos(4\pi t_k))\chi_{[-\frac{1}{2},\frac{1}{2}]}(t_k)$,

where $\chi_{[-\frac{1}{2},\frac{1}{2}]}(t)$ is the characteristic function of $[-\frac{1}{2},\frac{1}{2}]$.

The *windowed Fourier transform* is defined as

$$S_f^W(\mathbf{t}, \boldsymbol{\omega}) = \int_{\mathbb{R}^d} e^{-2\pi i \boldsymbol{\tau}\cdot\boldsymbol{\omega}} f(\boldsymbol{\tau}) W(\boldsymbol{\tau} - \mathbf{t}) \mathrm{d}\boldsymbol{\tau},$$

where $f \in L(\mathbb{R}^d)$ and W is a window function. When the window function is Gaussian window, the corresponding windowed Fourier transform is also called a *Gabor transform* or *short-time Fourier transform*. If the window function W satisfies $\int_{\mathbb{R}^d} W(\mathbf{t})\mathrm{d}\mathbf{t} = 1$, the *inversion formula* is

$$f(t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{2\pi i \mathbf{t}\cdot\boldsymbol{\omega}} S_f^W(\boldsymbol{\tau}, \boldsymbol{\omega}) W(\mathbf{t} - \boldsymbol{\tau}) \mathrm{d}\boldsymbol{\tau}\mathrm{d}\boldsymbol{\omega}.$$

Denote by $\rho_f^W(\mathbf{t}, \boldsymbol{\omega})$ the squared magnitude of the windowed Fourier transform, i.e.,

$$\rho_f^W(\mathbf{t}, \boldsymbol{\omega}) = |S_f^W(\mathbf{t}, \boldsymbol{\omega})|^2.$$

The $\rho_f^W(\mathbf{t}, \boldsymbol{\omega})$ is called the *spectrogram*. By the definition of the windowed Fourier transform,

$$\begin{aligned}\rho_f^W(\mathbf{t}, \boldsymbol{\omega}) &= \left|\int_{\mathbb{R}^d} e^{-2\pi i \boldsymbol{\tau}\cdot\boldsymbol{\omega}} f(\boldsymbol{\tau}) W(\boldsymbol{\tau} - \mathbf{t}) \mathrm{d}\boldsymbol{\tau}\right|^2 \\ &= \int_{\mathbb{R}^d} e^{-2\pi i \boldsymbol{\tau}\cdot\boldsymbol{\omega}} f(\boldsymbol{\tau}) W(\boldsymbol{\tau} - \mathbf{t}) \mathrm{d}\boldsymbol{\tau}\ \overline{\int_{\mathbb{R}^d} e^{-2\pi i \boldsymbol{\tau}\cdot\boldsymbol{\omega}} f(\boldsymbol{\tau}) W(\boldsymbol{\tau} - \mathbf{t}) \mathrm{d}\boldsymbol{\tau}}\end{aligned} \tag{2.6.1}$$

Let $\lambda_f(\mathbf{t}, \boldsymbol{\tau})$ be the inverse Fourier transform of $\rho_f^W(\mathbf{t}, \boldsymbol{\omega})$, i.e.,

$$\lambda_f(\mathbf{t}, \boldsymbol{\tau}) = \int_{\mathbb{R}^d} \rho_f^W(\mathbf{t}, \boldsymbol{\omega})\, e^{2\pi i \boldsymbol{\tau}\cdot\boldsymbol{\omega}} \mathrm{d}\boldsymbol{\omega}.$$

Taking the inverse Fourier transform on both sides of (2.6.1) and using properties of Fourier transform, we get

$$\begin{aligned}\lambda_f(\mathbf{t}, \boldsymbol{\tau}) &= (f(\boldsymbol{\tau}) W(\boldsymbol{\tau} - \mathbf{t})) * (\overline{f}(-\boldsymbol{\tau}) \overline{W}(-\boldsymbol{\tau} - \mathbf{t})) \\ &= \int_{\mathbb{R}^d} f(\boldsymbol{\alpha}) W(\boldsymbol{\alpha} - \mathbf{t}) \overline{f}(\boldsymbol{\alpha} - \boldsymbol{\tau}) \overline{W}(\boldsymbol{\alpha} - \boldsymbol{\tau} - \mathbf{t}) \mathrm{d}\boldsymbol{\alpha} \\ &= \int_{\mathbb{R}^d} f\left(\mathbf{u} + \tfrac{\boldsymbol{\tau}}{2}\right) W\left(\mathbf{u} - \mathbf{t} + \tfrac{\boldsymbol{\tau}}{2}\right) \overline{f}\left(\mathbf{u} - \tfrac{\boldsymbol{\tau}}{2}\right) \overline{W}\left(\mathbf{u} - \mathbf{t} - \tfrac{\boldsymbol{\tau}}{2}\right) \mathrm{d}\mathbf{u},\end{aligned}$$

and the last equality uses the substitution $\boldsymbol{\alpha} = \mathbf{u} + \frac{\boldsymbol{\tau}}{2}$. Let

$$\Gamma(\mathbf{t}, \boldsymbol{\tau}) = \overline{W}\left(\mathbf{t} + \frac{\boldsymbol{\tau}}{2}\right) W\left(\mathbf{t} - \frac{\boldsymbol{\tau}}{2}\right).$$

Since the window function $W(\mathbf{t})$ is even, real-valued, then $\Gamma(\mathbf{t}, \boldsymbol{\tau})$ is the instantaneous autocorrelation function of the window function, and $\Gamma(\mathbf{t} - \mathbf{u}, \boldsymbol{\tau}) = W\left(\mathbf{u} - \mathbf{t} + \frac{\boldsymbol{\tau}}{2}\right) \overline{W}\left(\mathbf{u} - \mathbf{t} - \frac{\boldsymbol{\tau}}{2}\right)$. This implies that

$$\lambda_f(\mathbf{t}, \boldsymbol{\tau}) = \int_{\mathbb{R}^d} \Gamma(\mathbf{t} - \mathbf{u}, \boldsymbol{\tau}) f\left(\mathbf{u} + \frac{\boldsymbol{\tau}}{2}\right) \overline{f}\left(\mathbf{u} - \frac{\boldsymbol{\tau}}{2}\right) \mathrm{d}\mathbf{u} = \Gamma(\mathbf{t}, \boldsymbol{\tau}) * K_f(\mathbf{t}, \boldsymbol{\tau}),$$

where $K_f(\mathbf{t}, \boldsymbol{\tau}) = f\left(\mathbf{t} + \frac{\boldsymbol{\tau}}{2}\right) \overline{f}\left(\mathbf{t} - \frac{\boldsymbol{\tau}}{2}\right)$. Taking Fourier transforms on both sides, the spectrogram is

$$\rho_f^W(\mathbf{t}, \boldsymbol{\omega}) = \int_{\mathbb{R}^d} (\Gamma(\mathbf{t}, \boldsymbol{\tau}) * K_f(\mathbf{t}, \boldsymbol{\tau}))\, e^{-2\pi i \boldsymbol{\tau} \cdot \boldsymbol{\omega}} \mathrm{d}\boldsymbol{\tau},$$

where $K_f(\mathbf{t}, \boldsymbol{\tau})$ is stated as above and is called the *signal kernel*.

2.6.2 General Form

For a function $f(t) (t \in \mathbb{R})$, if the Cauchy's principal value:

$$\widetilde{f}(t) = \frac{1}{\pi} \lim_{\epsilon \to 0} \int_{|t-\tau|>\epsilon} \frac{f(\tau)}{t - \tau} \mathrm{d}\tau$$

exists, then $\widetilde{f}(t)$ is called the *Hilbert transform* of $f(t)$. Let f be a signal and $\widetilde{f}$ be its Hilbert transform. The $z(t) = f(t) + i\widetilde{f}(t)$ is called *analytic associate* of f. Fourier transform of z vanishes for the negative frequency, i.e.,

$$\widehat{z}(\omega) = \begin{cases} 2\widehat{f}(\omega), & \omega \geq 0, \\ 0, & \omega < 0. \end{cases}$$

Let $\mathbf{f}(t) = (f_1(t), \ldots, f_m(t))^T$ be a real vector and $\mathbf{z}(t) = (z_1(t), \ldots, z_m(t))^T$ be the analytic associate vector of $\mathbf{f}$. The *general form of space–frequency distribution* is defined as

$$P_{\mathbf{zz}}(t, \omega) = \int_{\mathbb{R}^2} G(t - u, \tau) K_{\mathbf{zz}}(u, \tau)\, e^{-2\pi i \omega \tau} \mathrm{d}u \mathrm{d}\tau,$$

where $G(t, \tau)$ is the *time-lag kernel* and the signal kernel matrix is

$$K_{\mathbf{zz}}(u,\tau)=(K_{z_iz_j}(u,\tau))_{m\times m}=\begin{pmatrix} K_{z_1z_1}(u,\tau) & K_{z_1z_2}(u,\tau) & \cdots & K_{z_1z_m}(u,\tau) \\ K_{z_2z_1}(u,\tau) & K_{z_2z_2}(u,\tau) & \cdots & K_{z_2z_m}(u,\tau) \\ \vdots & \vdots & \ddots & \vdots \\ K_{z_mz_1}(u,\tau) & K_{z_mz_2}(u,\tau) & \cdots & K_{z_mz_m}(u,\tau) \end{pmatrix}$$

$$=\begin{pmatrix} z_1\left(u+\frac{\tau}{2}\right)\overline{z}_1\left(u-\frac{\tau}{2}\right) & z_1\left(u+\frac{\tau}{2}\right)\overline{z}_2\left(u-\frac{\tau}{2}\right) & \cdots & z_1\left(u+\frac{\tau}{2}\right)\overline{z}_m\left(u-\frac{\tau}{2}\right) \\ z_2\left(u+\frac{\tau}{2}\right)\overline{z}_1\left(u-\frac{\tau}{2}\right) & z_2\left(u+\frac{\tau}{2}\right)\overline{z}_2\left(u-\frac{\tau}{2}\right) & \cdots & z_2\left(u+\frac{\tau}{2}\right)\overline{z}_m\left(u-\frac{\tau}{2}\right) \\ \vdots & \vdots & \ddots & \vdots \\ z_m\left(u+\frac{\tau}{2}\right)\overline{z}_1\left(u-\frac{\tau}{2}\right) & z_m\left(u+\frac{\tau}{2}\right)\overline{z}_2\left(u-\frac{\tau}{2}\right) & \cdots & z_m\left(u+\frac{\tau}{2}\right)\overline{z}_m\left(u-\frac{\tau}{2}\right) \end{pmatrix},$$

i.e., $K_{\mathbf{zz}}(u,\tau)=\mathbf{z}\left(u+\frac{\tau}{2}\right)\overline{\mathbf{z}}^T\left(u-\frac{\tau}{2}\right)$. So the class of space–frequency distribution is

$$P_{\mathbf{zz}}(t,\omega)=\int_{\mathbb{R}^2}G(t-u,\tau)\mathbf{z}\left(u+\frac{\tau}{2}\right)\overline{\mathbf{z}}^T\left(u-\frac{\tau}{2}\right)e^{-2\pi i\omega\tau}\mathrm{d}u\mathrm{d}\tau, \tag{2.6.2}$$

2.6.3 Popular Distributions

Various popular space–frequency distributions can be derived directly from the general form (2.6.2).

(a) Let $G(t,\tau)=W\left(t+\frac{\tau}{2}\right)\overline{W}\left(t-\frac{\tau}{2}\right)$, where W is a univariate window function. By (2.6.2),

$$P_{\mathbf{zz}}(t,\omega)=\int_{\mathbb{R}^2}W\left(t-u+\frac{\tau}{2}\right)\overline{W}\left(t-u-\frac{\tau}{2}\right)\mathbf{z}\left(u+\frac{\tau}{2}\right)\overline{\mathbf{z}}^T\left(u-\frac{\tau}{2}\right)e^{-2\pi i\omega\tau}\mathrm{d}u\mathrm{d}\tau$$

which is called a *space spectrogram*. It can be rewritten as $P_{\mathbf{zz}}(t,\omega)=(P_{z_iz_j}(t,\omega))_{m\times m}$, where

$$P_{z_iz_j}(t,\omega)=\int_{\mathbb{R}^2}W\left(t-u+\frac{\tau}{2}\right)\overline{W}\left(t-u-\frac{\tau}{2}\right)z_i\left(u+\frac{\tau}{2}\right)\overline{z}_j\left(u-\frac{\tau}{2}\right)e^{-2\pi i\omega\tau}\mathrm{d}u\mathrm{d}\tau.$$

If $i=j$, then

$$P_{z_iz_j}(t,\omega)=\left|\int_{\mathbb{R}}e^{-2\pi i\omega\tau}z_i(\tau)W(\tau-t)\mathrm{d}\tau\right|^2=|S_{z_i}^W(t,\omega)|^2,$$

where $S_{z_i}^W(t,\omega)$ is the windowed Fourier transform of $z_i(\tau)$.

(b) Let $G(t,\tau)=\delta(t)$, where δ is Dirac function. Note that $\delta(-t)=\delta(t)$. Since

$$\int_{\mathbb{R}}\delta(t-u)\mathbf{z}\left(u+\frac{\tau}{2}\right)\overline{\mathbf{z}}^T\left(u-\frac{\tau}{2}\right)\mathrm{d}u=\mathbf{z}\left(t+\frac{\tau}{2}\right)\overline{\mathbf{z}}^T\left(t-\frac{\tau}{2}\right),$$

we get

$$P_{\mathbf{zz}}(t,\omega)=\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\delta(t-u)\mathbf{z}\left(u+\tfrac{\tau}{2}\right)\overline{\mathbf{z}}^T\left(u-\tfrac{\tau}{2}\right)\mathrm{d}u\right)\,e^{-2\pi i\omega\tau}\mathrm{d}\tau$$
$$=\int_{\mathbb{R}}\mathbf{z}\left(t+\tfrac{\tau}{2}\right)\overline{\mathbf{z}}^T\left(t-\tfrac{\tau}{2}\right)\,e^{-2\pi i\tau\omega}\mathrm{d}\tau$$

which is called a *space Wigner–Ville distribution*.
(c) Let $G(t,\tau)=\delta(t-\frac{\tau}{2})$. Note that $\delta(t-u-\frac{\tau}{2})=\delta(u-t+\frac{\tau}{2})$. Then,

$$P_{\mathbf{zz}}(t,\omega)=z(t)\overline{Z}^T(\omega)\,e^{-2\pi i\tau\omega}$$

is called the *space Rihaczek distribution*, where $Z(\omega)$ is the Fourier transform of $z(t)$.
(d) Let $G(t,\tau)=\delta(t-\frac{\tau}{2})W(t)$, where $W(t)$ is a window function. The function:

$$P_{\mathbf{zz}}(t,\omega)=\mathbf{z}(t)\int_{\mathbb{R}}\overline{\mathbf{z}}^T(t-\tau)W(\tau)\,e^{-2\pi i\tau\omega}\mathrm{d}\tau$$
$$=\mathbf{z}(t)\,e^{-2\pi it\omega}\left(\int_{\mathbb{R}}\overline{\mathbf{z}}^T(\tau)W(t-\tau)\,e^{-2\pi i\tau\omega}\mathrm{d}\tau\right)$$

is called a *space $W-$Rihaczek distribution*.
(e) Let $G(t,\tau)=\frac{1}{2}\left(\delta\left(t+\frac{\tau}{2}\right)+\delta\left(t-\frac{\tau}{2}\right)\right)$. Then,

$$P_{\mathbf{zz}}(t,\omega)=\mathrm{Re}\left(\mathbf{z}(t)\overline{\widehat{\mathbf{z}}}^T(\omega)\,e^{-2\pi it\omega}\right)$$

which is called a *space Levin distribution*.
(f) Let $G(t,\tau)=\frac{W(\tau)}{2}\left(\delta\left(t+\frac{\tau}{2}\right)+\delta\left(t-\frac{\tau}{2}\right)\right)$, where $W(\tau)$ is a window function. Then,

$$P_{\mathbf{zz}}(t,\omega)=\mathrm{Re}\left(\int_{\mathbb{R}}\overline{\mathbf{z}}^T(\tau)W(t-\tau)\,e^{-2\pi i\tau\omega}\mathrm{d}\tau\right)e^{-2\pi it\omega}$$

which is called a *space $W-$Levin distribution*.
(g) Let $G(t,\tau)=v(\tau)\delta\left(t-\frac{\tau}{2}\right)+v(-\tau)\delta\left(t+\frac{\tau}{2}\right)$, where $v(\tau)$ is the unit step function: $v(\tau)=1$ $(\tau\geq 0)$ and $v(\tau)=0$ $(\tau<0)$. Then,

$$P_{\mathbf{zz}}(t,\omega)=\frac{\partial}{\partial t}\left|\int_{-\infty}^{t}\mathbf{z}(\tau)\,e^{-2\pi i\omega\tau}\mathrm{d}\tau\right|^2,$$

which is called a *space Page distribution*.
(h) Let $G(t,\tau)=W(\tau)\,\mathrm{rect}\left(\frac{\alpha}{2\tau}\right)$. The $P_{\mathbf{zz}}(t,\omega)$ is called a *Zhao-Atlas-Marks distribution*.
(i) Let $G(t,\tau)=|\tau|^{\beta}\cosh^{-2\beta}t$. The $P_{\mathbf{zz}}^{W}(t,\omega)$ is called a *B-distribution*.
(j) Let $G(t,\tau)=\frac{\sqrt{\pi\sigma}}{|\tau|}\,e^{-\frac{\pi^2\sigma t^2}{\tau^2}}$. The $P_{\mathbf{zz}}(t,\omega)$ is called an *E-distribution*.

2.7 Multivariate Interpolation

Long-term multivariate time series always contain missing data or data with different sampling intervals. In this section, we introduce new advances in multivariate interpolation, including polynomial interpolation, positive definite function interpolation, radial function interpolation, and interpolation on sphere. These interpolation methods are widely applied in climate and environmental time series.

2.7.1 Multivariate Polynomial Interpolation

Let $\mathbf{x} = (x_1, \ldots, x_d)$ be a $d-$dimensional real-valued vector and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ $(\alpha_i \in \mathbb{Z}_+ \ (i = 1, \ldots, d))$, denoted by $\mathbf{x} \in \mathbb{R}^d$ and $\boldsymbol{\alpha} \in \mathbb{Z}_+^d$. Let $|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_d$. Define $\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}$. The polynomial $P_k(\mathbf{x}) = \sum_{|\boldsymbol{\alpha}| \leq k} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ is called a *polynomial of degree* k.

The multivariate interpolation of polynomials is different from the univariate interpolation of polynomials. In the one-dimensional case, a polynomial $\sum_{k=0}^n c_k x^k$ $(x \in \mathbb{R})$ of degree $\leq n$ is a linear combinations of $1, x, \ldots, x^n$. It can interpolate arbitrary data $\lambda_0, \ldots, \lambda_n$ on any set of $n+1$ distinct nodes $t_0, \ldots, t_n \in \mathbb{R}$. Namely, the coefficients $c_0, \ldots, c_n$ can be chosen so that $\lambda_i = \sum_{k=0}^n c_k t_i^k$ $(i = 0, \ldots, n)$ since the Vandermonde determinant:

$$\begin{vmatrix} 1 & t_1 & \cdots & t_1^{n-1} \\ 1 & t_2 & \cdots & t_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & \cdots & t_n^{n-1} \end{vmatrix} = \prod_{1 \leq j < i \leq n} (t_i - t_j) > 0.$$

This results in the well-known Lagrange interpolation formula. However, in the high-dimensional case, there do not exist $n+1$ multivariate continuous functions $g_k(\mathbf{x})(k = 0, \ldots, n)$ (not only polynomials) such that their linear combination $\sum_{k=0}^n c_k g_k(\mathbf{x})$ $(\mathbf{x} \in \mathbb{R}^d)$ can interpolate arbitrary data $\lambda_0, \ldots, \lambda_n$ on any set of $n+1$ distinct nodes $x_0, \ldots, x_n$ in $\mathbb{R}^d$. In fact, if there exist such $\{g_k\}_{k=0,\ldots,n}$, the determinant $\det(g_k(\mathbf{t}_i)) \neq 0$ for any set of $n+1$ distinct nodes $\{\mathbf{t}_k\}_{k=0,\ldots,n}$ in $\mathbb{R}^d$. However, this is not possible. Select a closed path containing $\mathbf{t}_1$ and $\mathbf{t}_2$ in $\mathbb{R}^d$ but no other nodes. Continuously moving $\mathbf{t}_1$ and $\mathbf{t}_2$ in the same direction in the closed path can exchange positions of $\mathbf{t}_1$ and $\mathbf{t}_2$, so the determinant will change sign. Since the determinant is a continuous function, it will vanish at some stage. This is a contradiction.

An efficient method is on a Cartesian grid. For example, for $d = 3$, we interpolate $F(x, y, z)$ in the grid:

$$\{(x_i, y_j, z_k): \quad 1 \le i \le n, \quad 1 \le j \le m, \quad 1 \le k \le l\}.$$

The interpolation formula is

$$L(F(x, y, z)) = \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{k=1}^{l} F(x_i, y_j, z_k) u_i(x) v_j(y) w_k(z),$$

where

$$u_i(x) = \prod_{\substack{\nu=1 \\ \nu \ne i}}^{n} \frac{x - x_i}{x_\nu - x_i}, \qquad v_j(y) = \prod_{\substack{\nu=1 \\ \nu \ne j}}^{m} \frac{y - y_j}{y_\nu - y_j}, \qquad w_k(z) = \prod_{\substack{k=1 \\ \nu \ne k}}^{l} \frac{z - z_k}{z_\nu - z_k}.$$

This interpolation formula is a direct generalization of the univariate case.

Consider a set N consists of $\frac{1}{2}(k+1)(k+2)$ nodes in $\mathbb{R}^d$ and there are k hyperplanes $H_0, \ldots, H_k \subset \mathbb{R}^d$ such that $N \subset \bigcup_{i=0}^{k} H_i$ and just $i+1$ nodes lie in $H_i (0 \le i \le k)$, then arbitrary data on the set N can be interpolated by a $d-$variate polynomial of degree k. For example, if the set N consists of six nodes $\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_5$ in $\mathbb{R}^2$ and there are three straight lines $L_0, L_1, L_2 \subset \mathbb{R}^2$ such that $\mathbf{t}_0 \in L_0, \mathbf{t}_1, \mathbf{t}_2 \in L_1$, and the remaining $\mathbf{t}_3, \mathbf{t}_4, \mathbf{t}_5 \in L_2$, then arbitrary data on the set N can be interpolated by a bivariate polynomial of degree 2.

2.7.2 Schoenberg Interpolation

A complex-valued function f in $\mathbb{R}^d$ is said to be *strictly positive definite* if, for any finite set $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$, the $n \times n$ matrix $M_{jk} = f(\mathbf{x}_j - \mathbf{x}_k)$ is positive definite, i.e., for all $\mathbf{v} = (v_1, \ldots, v_n)^T$,

$$\overline{\mathbf{v}}^T M \mathbf{v} = \sum_{j=1}^{n} \sum_{k=1}^{n} \overline{v}_j v_k M_{jk} > 0,$$

where $M = (M_{jk})_{n \times n}$. A linear combination of translation of a strictly positive definite function, $\sum_{k=1}^{n} a_k f(\mathbf{x} - \mathbf{x}_k)$, can interpolate arbitrary data on any distinct nodes $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$.

A function $f(t)$ is said to be *completely monotone* on $[0, \infty)$ if $f(t)$ is infinitely differentiable on $(0, \infty)$ and is continuous at $t = 0$ and $(-1)^k f^{(k)}(t) \ge 0\ (t > 0,\ k = 0, 1, \ldots)$. For example, the following three functions:

$$f(t) = (\alpha + t)^{-\beta} \quad (\alpha > 0,\ \beta > 0),$$

$$f(t) = e^{-\alpha t} \qquad (\alpha > 0),$$

$$f(t) = t^{-1}(1 - e^{t})$$

are all completely monotone.

Schoenberg Interpolation Theorem. *Schoenberg Interpolation Theorem. If f is completely monotone but not constant on $[0, \infty)$, then, for any distinct nodes $\{\mathbf{x}_k\}_{k=1,\dots,n} \in \mathbb{R}^d$, the linear combination:*

$$\sum_{k=1}^{n} c_k f(\| \mathbf{x} - \mathbf{x}_k \|^2) = \sum_{k=1}^{n} c_k f\left((x_1 - x_{k1})^2 + \cdots + (x_d - x_{kd})^2\right),$$

where $\mathbf{x} = (x_1, \dots, x_d)$ and $\mathbf{x}_k = (x_{k1}, \dots, x_{kd})$, can interpolate arbitrary data on these nodes.

For any distinct $n + 1$ nodes $\{\mathbf{x}_k\}_{k=1,\dots,n} \in \mathbb{R}^d$, Schoenberg interpolation theorem shows that the following linear combinations:

$$\sum_{k=1}^{n} c_k(\alpha + \| \mathbf{x} - \mathbf{x}_k \|^2)^{-\beta} \qquad (\alpha > 0,\ \beta > 0),$$

$$\sum_{k=1}^{n} c_k\, e^{-\alpha \|\mathbf{x} - \mathbf{x}_k\|^2} \qquad (\alpha > 0),$$

$$\sum_{k=1}^{n} c_k \| \mathbf{x} - \mathbf{x}_k \|^{-2} \left(1 - e^{-\|\mathbf{x} - \mathbf{x}_k\|^2}\right)$$

can interpolate arbitrary data $\lambda_1, \dots, \lambda_n$ on nodes $\mathbf{x}_1, \dots, \mathbf{x}_n$.

2.7.3 Micchelli Interpolation

There exist more functions such that the following form of interpolation holds

$$\sum_{k=0}^{\infty} c_k f(\| \mathbf{x} - \mathbf{x}_k \|^2).$$

Micchelli Interpolation Theorem. Let f be a positive-valued continuous function on $[0, \infty)$ and the derivative f' be completely monotone but not constant on $[0, \infty)$. Then, the combination $\sum_{k=1}^{n} c_k f(\| \mathbf{x} - \mathbf{x}_k \|^2)$ can interpolate arbitrary data $\lambda_1, \dots, \lambda_n$ on any distinct nodes $\mathbf{x}_1, \dots, \mathbf{x}_n$ in $\mathbb{R}^d$.

Proof Let $\mathbf{v} = (v_1, \dots, v_n)^T$ be a nonzero real-valued vector, where $\sum_{k=1}^n v_k = 0$, and let $M = (M_{jk})_{n\times n}$, where $M_{jk} = f(\| \mathbf{x}_j - \mathbf{x}_k \|^2)$. Note that

$$\mathbf{v}^T M\mathbf{v} = \sum_{j=1}^n \sum_{k=1}^n v_j v_k f(\| \mathbf{x}_j - \mathbf{x}_k \|^2).$$

Since $f(t)$ is infinitely differentiable in $(0, \infty)$ and continuous at $t = 0$, $f(t)$ can be expressed by a Stieltjes integral:

$$f(t) = f(0) + \int_0^\infty \frac{1 - e^{-st}}{s} \mathrm{d}\alpha(s),$$

where $\alpha(s)$ is a non-decreasing bounded function and $\int_0^\infty \frac{\mathrm{d}\alpha(s)}{s} < \infty$. So

$$\begin{aligned}\mathbf{v}^T M\mathbf{v} &= \sum_{j=1}^n \sum_{k=1}^n v_j v_k \left(f(0) + \int_0^\infty \frac{1-e^{-s\|\mathbf{x}_j - \mathbf{x}_k\|^2}}{s} \mathrm{d}\alpha(s)\right)\\ &= -\sum_{j=1}^n \sum_{k=1}^n v_j v_k \int_0^\infty \frac{e^{-s\|\mathbf{x}_j - \mathbf{x}_k\|^2}}{s} \mathrm{d}\alpha(s).\end{aligned}$$

Since the matrix $(e^{-s\|\mathbf{x}_j - \mathbf{x}_k\|^2})_{n\times n}$ is positive definite, $\mathbf{v}^T(-M)\mathbf{v} > 0$, and then M is non-singular. From this, the desired result is obtained.

The functions $\sqrt{t}$ and $\log(1 + t)$ satisfy the conditions of Micchelli interpolation theorem. The function $\sqrt{1 + t}$ used often in geophysics also satisfies the conditions of Micchelli interpolation theorem. Therefore, the following three linear combinations:

$$\sum_{k=1}^n c_k \| \mathbf{x} - \mathbf{x}_k \|,$$

$$\sum_{k=1}^n c_k \log(1+ \| \mathbf{x} - \mathbf{x}_k \|^2),$$

$$\sum_{k=1}^n c_k \sqrt{1+ \| \mathbf{x} - \mathbf{x}_k \|^2}$$

can interpolate arbitrary data $\lambda_1, \dots, \lambda_n$ on any distinct nodes $\mathbf{x}_1, \dots, \mathbf{x}_n$ in $\mathbb{R}^d$.

2.7.4 Interpolation on Spheres

Interpolation on spheres is often used in geography, climate, and environment, whose information is gathered over the Earth's surface. Denote by $\mathbb{S}^d$ the $d-$dimensional unit sphere in $\mathbb{R}^{d+1}$, i.e.,

$$\mathbb{S}^d = \{\mathbf{x} \in \mathbb{R}^{d+1} : \ \| \mathbf{x} \| = 1\}.$$

Ultraspherical polynomials are introduced to establish the interpolation formula on spheres. The ultraspherical polynomial $P_n^{(\lambda)}(x)$ is a polynomial of degree n which is determined by the equation:

$$(1 - 2rx + r^2)^{-\lambda} = \sum_{n=0}^{\infty} r^n P_n^{(\lambda)}(x) \qquad (\lambda > 0).$$

For any $\lambda > -\frac{1}{2}$, the ultraspherical polynomial family $\{P_n^{(\lambda)}\}_{n=0,1,\ldots}$ satisfies the conditions:

$$\int_{-1}^{1} P_j^{(\lambda)}(x) P_k^{(\lambda)}(x)(1 - x^2)^{\lambda-\frac{1}{2}} \mathrm{d}x = 0 \ (j \neq k),$$

$$P_n^{(\lambda)}(1) = C_{n+2\lambda-1}^{n} = \frac{(n+2\lambda-1)!}{n!(2\lambda-1)!}.$$

Interpolation Theorem on Spheres. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be distinct points on the unit sphere $\mathbb{S}^d$. For the ultraspherical polynomial series of $f(t)$ on $[-1,1]$,

$$f(t) = \sum_{j=0}^{\infty} \alpha_j P_j^{(\frac{d-1}{2})}(t),$$

if $\alpha_j \geq 0$, $\alpha_k > 0 \, (0 \leq k < n)$, and $\sum_{j=0}^{\infty} \alpha_j P_j^{(\frac{d-1}{2})}(1) < \infty$, then the linear combination $\sum_{k=1}^{n} c_k f((\mathbf{x}, \mathbf{x}_k))$ can interpolate arbitrary values on nodes $\mathbf{x}_1, \ldots, \mathbf{x}_n$, where $(\mathbf{x}, \mathbf{x}_k)$ is the inner product of $\mathbf{x}$ and $\mathbf{x}_k$.

2.8 Sparse Approximation

Sparse approximation has attracted considerable interest recently because of its application in many aspects of data analysis. The crucial rule in sparse approximation is to seek a good approximation using as few elementary signals as possible, so the amount of space needed to store large multivariate spatial datasets would be reduced to a fraction of what was originally needed.

2.8.1 Approximation Kernels

Let $H_n(x)$ $(n \in \mathbb{Z}_+)$ be kernel functions on $\mathbb{R}^d$ satisfying

$$\int_{\mathbb{R}^d} H_n(\mathbf{x})\mathrm{d}\mathbf{x} = 1,$$

$$\sup_{n\in\mathbb{Z}_+} \int_{\mathbb{R}^d} |H_n(\mathbf{x})|\mathrm{d}\mathbf{x} < \infty,$$

$$\lim_{n\to\infty} \int_{\|\mathbf{x}\|\geq\delta} |H_n(\mathbf{x})|\mathrm{d}\mathbf{x} = 0 \quad \text{for any } \delta > 0,$$

and let f be any bonded continuous function on $\mathbb{R}^d$. Then,

$$\lim_{n\to\infty} (H_n * f)(\mathbf{x}) = f(\mathbf{x})\ (\mathbf{x} \in \mathbb{R}^d),$$

where $H_n * f$ represents the convolution of H_n and f on $\mathbb{R}^d$.

Main kernel functions used often are as follows:

(a) Polynomial Kernel:

$$H_n(\mathbf{x}) = \begin{cases} c_n(1- \parallel \mathbf{x} \parallel^2)^n, & \parallel \mathbf{x} \parallel \leq 1, \\ 0, & \parallel \mathbf{x} \parallel > 1, \end{cases}$$

where $\parallel \mathbf{x} \parallel$ is the Euclidean norm of $\mathbf{x} \in \mathbb{R}^d$ and c_n is determined by $\int_{\mathbb{R}^d} H_n(\mathbf{x})\mathrm{d}\mathbf{x} = 1$.

(b) Dilation Kernels: $H_n(\mathbf{x}) = n^d H(n\mathbf{x})$ $(\mathbf{x} \in \mathbb{R}^d,\ n \in \mathbb{Z}_+)$, where $\int_{\mathbb{R}^d} H(\mathbf{x})\mathrm{d}\mathbf{x} = 1$.

(c) Weierstrass Kernel: $W(\mathbf{x}) = \pi^{-\frac{d}{2}} e^{-\|\mathbf{x}\|^2}$ $(\mathbf{x} \in \mathbb{R}^d)$.

(d) Cauchy Kernel: $C(\mathbf{x}) = \frac{\tau}{1+\|\mathbf{x}\|^2}$ $(\mathbf{x} \in \mathbb{R}^d)$, where $\tau > 0$ is determined by $\int_{\mathbb{R}^d} C(\mathbf{x})\mathrm{d}\mathbf{x} = 1$.

If a function $f(\mathbf{x})$ on $\mathbb{R}^d$ has the form:

$$f(\mathbf{x}) = \sum_{\substack{\mathbf{k}\in\mathbb{Z}^d \\ |k_1|+\cdots+|k_d|\leq n}} c_{\mathbf{k}}\, e^{2\pi i(\mathbf{k},\mathbf{x})},$$

then f is called an *multivariate trigonometric polynomial* of degree at most n. If

$$\int_{\mathbb{R}^d} |H(\mathbf{x})|\mathrm{d}\mathbf{x} = 1,$$

$$\widehat{H}(\mathbf{x}) = 0 \qquad (|x_1| + \cdots + |x_d| \geq 1),$$

and if f is an integer-periodic continuous function on $\mathbb{R}^d$, it can be proved that $H_\lambda * f$ is an multivariate trigonometric polynomial of degree at most $\lambda - 1$ and $\lim_{k\to\infty}(H_\lambda * f) = f$, where $H_\lambda(\mathbf{x}) = \lambda^d H(\lambda\mathbf{x})$.

Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$. Define

$$(D^{\alpha} f)(\mathbf{x}) = \frac{\partial^{|\alpha|} f(\mathbf{x})}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \qquad (|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_d).$$

If $D^{\alpha} f$ is continuous and bounded on $\mathbb{R}^d$ for all $|\boldsymbol{\alpha}| \le k$, we say $f \in C^k(\mathbb{R}^d)$. If a kernel function H satisfies $\int_{\mathbb{R}^d} |H(\mathbf{x})\mathbf{x}^{\alpha}|d\mathbf{x} < \infty$ and its dilation kernel $H_\lambda(\mathbf{x})$ satisfies $H_\lambda(\mathbf{x}) * \frac{\mathbf{x}^\alpha}{\alpha!} = \frac{\mathbf{x}^\alpha}{\alpha!}$ $(|\boldsymbol{\alpha}| < k)$, it can be proved that for $f \in C^k(\mathbb{R}^d)$,

$$|(H_\lambda * f)(\mathbf{x}) - f(\mathbf{x})| \le M|f|_k \lambda^{-k} \qquad (\mathbf{x} \in \mathbb{R}^d),$$

where $M = \sum_{|\alpha|=k} \frac{1}{\alpha!} \int_{\mathbb{R}^d} |H(\mathbf{y})||\mathbf{y}^\alpha|d\mathbf{y}$ and $|f|_k = \max_{|\alpha|=k,\, \mathbf{x}\in\mathbb{R}^d} |D^\alpha f(\mathbf{x})|$. It is not difficult to construct this kind of kernel functions, for example,

$$H(\mathbf{x}) = \sum_{i=1}^{k} a_i h\left(\frac{\mathbf{x}}{t_i}\right) t_i^{-d} \qquad (\mathbf{x} \in \mathbb{R}^d),$$

where $t_1, \ldots, t_k \in \mathbb{R}$ are different nonzero points, $a_i = \prod_{j=1,\ldots,k;\, j\ne i} \frac{-t_j}{t_i - t_j}$ $(i = 1, \ldots, k)$, and h satisfies

$$\int_{\mathbb{R}^d} h(\mathbf{y})d\mathbf{y} = 1, \qquad \int_{\mathbb{R}^d} |h(\mathbf{y})\mathbf{y}^\alpha|d\mathbf{y} < \infty \quad (|\boldsymbol{\alpha}| < k).$$

2.8.2 Sparse Schemes

For a smooth function f on the cube $[0, 1]^d$, we extend it to an integer-periodic function on $\mathbb{R}^d$ and then expand it into Fourier series. Its Fourier coefficients decay slowly due to the discontinuity on the boundary, so we need a lot of Fourier coefficients to reconstruct it. In this subsection, we give a sparse scheme to reconstruct the smooth functions on the cube $[0, 1]^d$.

Let $f(x_1, \ldots, x_d)$ be defined on $[0, 1]^d$ and $\frac{\partial^{i_1+\cdots+i_d} f}{\partial x_1^{i_1} \cdots \partial x_d^{i_d}}$ $(i_1, \ldots, i_d = 0, 1, 2)$ be continuous on $[0, 1]^d$. The *reconstruction scheme* consists of five steps.

Step 1. Start from the $(d-1)-$variate boundary functions of $f(x_1, \ldots, x_d)$

$$f(0, x_2, \ldots, x_d), \quad f(x_1, 0, \ldots, x_d), \quad \ldots, \quad f(x_1, x_2, \ldots, 0),$$

$$f(1, x_2, \ldots, x_d), \quad f(x_1, 1, \ldots, x_d), \quad \ldots, \quad f(x_1, x_2, \ldots, 1).$$

Step 2. Construct a $d-$variate function $g(x_1, \ldots, x_d)$, which is a combination of these $(d-1)-$variate boundary functions and the factors $x_1, \ldots, x_d$, $(1-x_1), \ldots, (1-x_d)$, to satisfy a condition:

$$g(x_1, \ldots, x_d) = f(x_1, \ldots, x_d) \quad \text{on} \quad \partial([0, 1]^d).$$

Step 3. Let $h(x_1, \ldots, x_d) = f(x_1, \ldots, x_d) - g(x_1, \ldots, x_d)$. Compute the Fourier coefficients $c_{n_1,\ldots,n_d}(h)$ of h, where $|n_1 n_2 \cdots n_d| \le N$ and $|n_1|, |n_2|, \ldots, |n_d| \le N$.
Step 4. Compute the hyperbolic cross truncations of Fourier series of $f(x_1, \ldots, x_d)$ by the formula:

$$s_N^{(h,c)}(f; x_1, \ldots, x_d) = \sum_{\substack{|n_1 \cdots n_d| \le N \\ |n_1|, |n_2|, \ldots, |n_d| \le N}} c_{n_1,\ldots,n_d}(h)\, e^{2\pi i (n_1 x_1 + \cdots + n_d x_d)}.$$

Step 5. The reconstruction formula of $f(x_1, \ldots, x_d)$ is

$$f(x_1, \ldots, x_d) \approx g(x_1, \ldots, x_d) + s_N^{(h,c)}(x_1, \ldots, x_d). \tag{2.8.1}$$

The square error $e_N^{(d)}$ is

$$(e_N^{(d)})^2 = O\left(\frac{\log^{d-1} N}{N^3}\right),$$

and the total number of Fourier coefficients used is equivalent to $N \log^{d-1} N$.

In the reconstruction formula (2.8.1), the term $g(x_1, \ldots, x_d)$ is a combination of these $(d-1)-$variate boundary functions:

$$f(0, x_2, \ldots, x_d), \quad f(x_1, 0, \ldots, x_d), \quad \ldots, \quad f(x_1, \ldots, x_{d-1}, 0),$$

$$f(1, x_2, \ldots, x_d), \quad f(x_1, 1, \ldots, x_d), \quad \ldots, \quad f(x_1, \ldots, x_{d-1}, 1)$$

and the factors $x_1, \ldots, x_d,\ (1-x_1), \ldots, (1-x_d)$. However, each of the above $(d-1)-$variate boundary functions can be processed again and again by Steps 1–5 until the involved boundary functions are univariate functions. Finally, $f(x_1, \ldots, x_d)$ on $[0, 1]^d$ can be reconstructed by few Fourier coefficients with small approximation error and its values at the vertexes of $[0, 1]^d$. Such approximation scheme by using few coefficients is just *sparse approximation*.

Compared with this approximation scheme, if the function $f(x_1, \ldots, x_d)$ is reconstructed by the partial sum of its Fourier series:

$$s_N(f; x_1, \ldots, x_d) = \sum_{|n_1| \le N} \cdots \sum_{|n_d| \le N} c_{n_1,\ldots,n_d}(f)\, e^{2\pi i (n_1 x_1 + \cdots + n_d x_d)},$$

then the corresponding square error $\widetilde{e}_N^{(d)}$ decays slowly,

$$(\widetilde{e}_N^{(d)})^2 = O(\widetilde{N}^{-\frac{1}{d}}),$$

where $\widetilde{N}$ is the total number of the Fourier coefficients and $\widetilde{N} \approx N^d$.

Now, we give the detail algorithm for the univariate, bivariate, and triple functions used often.

(i) Let $f(x)$ be a univariate function defined on [0, 1] and $f''(x)$ be continuous on [0, 1]. We reconstruct $f(x)$ as follows. Clearly, two boundary values are $f(0)$ and $f(1)$. Construct a univariate function:

$$g(x) = f(x) - f(0)(1-x) - f(1)x \qquad (0 \le x \le 1).$$

Clearly, $g(0) = f(0)$ and $g(1) = f(1)$. Find the Fourier coefficients of $g(x)$ by the formula:

$$c_n(g) = \int_0^1 g(x)\, e^{-2\pi i n x} \mathrm{d}x.$$

The reconstruction formula of $f(x)$ is

$$f(x) \approx f(0)(1-x) + f(1)x + \sum_{n=-N}^{N} c_n(g)\, e^{2\pi i n x} \qquad (0 \le x \le 1)$$

and the square errors $e_N^{(1)}$ satisfy

$$(e_N^{(1)})^2 = O\left(\frac{1}{N^3}\right).$$

In the reconstruction formula, $f(x)$ can be reconstructed by $f(0)$, $f(1)$, and coefficients $\{c_n(g)\}(|n| \le N)$, and the total number of Fourier coefficients is equivalent to N.

(ii) Let $f(x, y)$ be a bivariate function defined on $[0, 1]^2$ and $\frac{\partial^{i+j} f}{\partial x^i \partial y^j}$ $(i, j = 0, 1, 2)$ be continuous on $[0, 1]^2$. We reconstruct $f(x, y)$ as follows. Clearly, four univariate boundary functions are

$$f(x, 0), \quad f(x, 1) \qquad (0 \le x \le 1),$$

$$f(0, y), \quad f(1, y) \qquad (0 \le y \le 1).$$

where each boundary function can be reconstructed by values at the end points and $2N + 1$ univariate Fourier coefficients. Based on these four boundary functions, we construct a bivariate function:

$$\begin{aligned} g(x, y) = {} & f(x, 0)(1-y) + f(x, 1)y + f(0, y)(1-x) + f(1, y)x \\ & - f(0, 0)(1-x)(1-y) - f(0, 1)(1-x)y - f(1, 0)x(1-y) - f(1, 1)xy. \end{aligned}$$

Clearly, $g(x, y)$ is a combination of four boundary functions and four factors $x, y, (1-x), (1-y)$ and satisfies

$$g(x,0) = f(x,0), \qquad g(x,1) = f(x,1) \qquad (0 \le x \le 1),$$

$$g(0,y) = f(0,y), \qquad g(1,y) = f(1,y) \qquad (0 \le y \le 1).$$

Let

$$h(x, y) = f(x, y) - g(x, y).$$

Find Fourier coefficients of $h(x, y)$: $\{ c_{mn}(h) \} \, (|mn| \le N, \ \ |m|, |n| \le N)$ and then find the hyperbolic cross truncations of Fourier series of $h(x, y)$ by the formula:

$$s_N^{(h,c)}(x, y) = \sum_{\substack{|mn| \le N \\ |m|, |n| \le N}} c_{mn}(h) \, e^{2\pi i (mx+ny)}.$$

The reconstruction formula is

$$f(x, y) \approx g(x, y) + s_N^{(h,c)}(x, y) \qquad ((x, y) \in [0, 1]^2)$$

Since four univariate boundary functions, which are used to construct $g(x, y)$, can be reconstructed as those in (i), the total number of the Fourier coefficients used in the reconstruction of $f(x, y)$ is equivalent to $N \log N$, and the square errors $e_N^{(2)}$ satisfy

$$(e_N^{(2)})^2 = O\left(\frac{\log N}{N^3} \right).$$

(iii) Let $f(x, y, z)$ be a triple function on $[0, 1]^3$ and $\frac{\partial^{i+j+k} f}{\partial x^i \partial y^j \partial z^k} (i, j, k = 0, 1, 2)$ be continuous on $[0, 1]^3$. We reconstruct f as follows. Clearly, six bivariate boundary functions are

$$f(0, y, z), \quad f(1, y, z) \qquad (0 \le y \le 1, \ \ 0 \le z \le 1),$$

$$f(x, 0, z), \quad f(x, 1, z) \qquad (0 \le x \le 1, \ \ 0 \le z \le 1),$$

$$f(x, y, 0), \quad f(x, y, 1) \qquad (0 \le x \le 1, \ \ 0 \le y \le 1).$$

Based on these six bivariate boundary functions, construct a triple function as follows:

$$g(x, y, z) = g_1(x, y, z) + g_2(x, y, z) + g_3(x, y, z),$$

where

$$\begin{aligned}
g_1(x,y,z) = {} & f(x,y,0)(1-z) + f(x,y,1)z + f(x,0,z)(1-y) + f(x,1,z)y \\
& + f(0,y,z)(1-x) + f(1,y,z)x, \\
g_2(x,y,z) = {} & -f(x,0,0)(1-y)(1-z) - f(x,0,1)(1-y)z \\
& - f(x,1,0)y(1-z) - f(x,1,1)yz - f(0,y,0)(1-x)(1-z) \\
& - f(0,y,1)(1-x)z - f(1,y,0)x(1-z) - f(1,y,1)xz \\
& - f(0,0,z)(1-x)(1-y) - f(0,1,z)(1-x)z \\
& - f(1,0,z)x(1-y) - f(1,1,z)xy, \\
g_3(x,y,z) = {} & f(0,0,0)(1-x)(1-y)(1-z) + f(1,0,0)x(1-y)(1-z) \\
& + f(0,1,0)(1-x)y(1-z) + f(0,0,1)(1-x)(1-y)z \\
& + f(1,1,0)xy(1-z) + f(1,0,1)x(1-y)z \\
& + f(0,0,1)(1-x)yz + f(1,1,1)xyz.
\end{aligned}$$

Clearly, $g(x,y,z) = f(x,y,z)$ on $\partial([0,1]^3)$. Let $h(x,y,z) = f(x,y,z) - g(x,y,z)$. Find Fourier coefficients $\{c_{n_1n_2n_3}(h)\}\,(|n_1n_2n_3| \le N,\ |n_1|, |n_2|, |n_3| \le N)$ and find the hyperbolic cross truncations of Fourier series of $h(x,y,z)$ by the formula:

$$s_N^{(h,c)}(x,y,z) = \sum_{\substack{|n_1n_2n_3|\le N \\ |n_1|,|n_2|,|n_3|\le N}} c_{n_1n_2n_3}(h)\, e^{2\pi i(n_1x+n_2y+n_3z)}.$$

The reconstruction formula is

$$f(x,y,z) \approx g(x,y,z) + s_N^{(h,c)}(x,y,z), \qquad (x,y,z) \in [0,1]^2$$

Since six bivariate boundary functions, which are used to construct $g(x,y,z)$, can be reconstructed as those in (ii), the total number of the Fourier coefficients in the reconstruction of $f(x,y,z)$ is equivalent to $N\log^2 N$, and the square errors $e_N^{(3)}$ satisfy $(e_N^{(3)})^2 = O(\frac{\log^2 N}{N^3})$.

2.8.3 Greedy Algorithm

An arbitrary subset of $L^2(\mathbb{R}^d)$ is called a *dictionary*. One approximates to f by a linear combination of N functions in a dictionary D. For example, the translations of Gabor functions $g_{\alpha,\beta}(t) = e^{-i\alpha t}\, e^{-\beta t^2}$ $(\alpha, \beta \in \mathbb{R})$ can generate a dictionary $\{g_{\alpha,\beta}(t-\gamma) : \alpha,\ \beta,\ \gamma \in \mathbb{R}\}$ in $L^2(\mathbb{R})$.

(a) Pure Greedy Algorithm

Let D be a dictionary in $L^2(\mathbb{R}^d)$. For $f \in L^2(\mathbb{R}^d)$, let $h = h(f) \in D$ be such that $(f, h(f)) = \sup_{h\in D}(f,h)$, where $(f,g) = \int_{\mathbb{R}^d} f(\mathbf{t})\overline{g}(\mathbf{t})\mathrm{d}\mathbf{t}$. Define

$$G_1(f) = (f, h(f))h(f), \qquad R_1(f) = f - G_1(f),$$

$$G_2(f) = G_1(f) + G_1(R_1(f)), \qquad R_2(f) = f - G_2(f).$$

For $m \geq 2$, inductively define

$$G_m(f) = G_{m-1}(f) + G_1(R_{m-1}(f)), \qquad R_m(f) = f - G_m(f).$$

For the dictionary D generated by an orthogonal basis, $G_m(f)$ is the best $m-$term approximation of f from D. However, in general, this algorithm is not ideal.
(b) Relaxed Greedy Algorithm
Let $G_1(f)$ and $R_1(f)$ be stated in (a). Define

$$G_0^{(r)}(f) = 0, \qquad R_0^{(r)}(f) = f,$$

$$G_1^{(r)}(f) = G_1(f), \; R_1^{(r)}(f) = R_1(f).$$

For a function g, let $h = h(g) \in D$ be such that $(g, h) = \sup_{\tau \in D}(g, \tau)$. Inductively define

$$G_m^{(r)}(f) = \left(1 - \tfrac{1}{m}\right) G_{m-1}^{(r)}(f) + \tfrac{1}{m} h(R_{m-1}^{(r)}(f)),$$

$$R_m^{(r)}(f) = f - G_m^{(r)}(f).$$

In general, $G_m^{(r)}(f)$ can approximate to f better than $G_m(f)$.

2.9 Spherical Harmonics

Spherical harmonics are a frequency–space basis for representing multivariate time series defined over the sphere. Based on the rotation and spherical harmonics, a square integrable function can be expanded into an orthogonal series whose term is a product of a radial function and a solid spherical harmonic function of degree k. The structure of each term is invariant under Fourier transform.

2.9.1 Spherical Harmonic Functions

A function u defined on a domain D on $\mathbb{R}^d$ is called a *harmonic function* if it satisfies the Laplace equation $\Delta u = \frac{\partial^2 u}{\partial t_1^2} + \cdots + \frac{\partial^2 u}{\partial t_d^2} = 0$ on D.

A polynomial of the form $p(\mathbf{t}) = \sum_{|\boldsymbol{\alpha}|=k} C_{\boldsymbol{\alpha}} \mathbf{t}^{\boldsymbol{\alpha}}$ $(k \in \mathbb{Z}_+)$, where

$$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d), \quad \text{each } \alpha_k = 0, 1, \ldots,$$

$$|\boldsymbol{\alpha}| = \alpha_1 + \alpha_2 + \cdots + \alpha_d,$$

$$\mathbf{t}^{\boldsymbol{\alpha}} = t_1^{\alpha_1} t_2^{\alpha_2} \cdots t_d^{\alpha_d},$$

is called a *homogeneous polynomial* of degree k on $\mathbb{R}^d$.

A homogeneous harmonic polynomial of degree k is called a *solid spherical harmonics* of degree k, and the restriction of such function on the surface of the unit sphere is called a *surface spherical harmonics* of degree k. For example, the polynomial $p(x, y) = x^3 - 3xy^2$ is a solid spherical harmonics of degree 3 on $\mathbb{R}^2$. Let $x = r\cos\theta,\ y = r\sin\theta$. The restriction of $p(x, y)$ on the unit circle is $\cos(3\theta)$. So $\cos(3\theta)$ is a surface spherical harmonics of degree 3 on $\mathbb{R}^2$.

2.9.2 *Invariant Subspace under Fourier Transform*

(i) Radial Function

Let f be a function on $\mathbb{R}^d$. If

$$f(\rho\mathbf{t}) = f(\mathbf{t}), \tag{2.9.1}$$

where ρ is any orthogonal transform on $\mathbb{R}^d$, then $f(\mathbf{t})$ is called a *radial function*. If $f(\mathbf{t}) = \xi(\| \mathbf{t} \|)$, i.e., $f(\mathbf{t})$ on $\mathbb{R}^d$ only depends on the norm of $\mathbf{t}$, then $f(\mathbf{t})$ must be a radial function.

(a) Fourier transform and orthogonal transform are exchangeable

Let $f(\mathbf{x})$ be an integrable function and ρ be a $d \times d$ orthogonal matrix. Fourier transform of $f(\rho\mathbf{t})$ is

$$(f(\rho\mathbf{t}))^\wedge(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\rho\mathbf{t})\, e^{-2\pi i(\boldsymbol{\omega},\, \mathbf{t})} \mathrm{d}\mathbf{t}.$$

Since ρ is an orthogonal matrix, $\rho^{-1} = \rho^T$ and $\det(\rho) = \pm 1$. Making the substitution $\mathbf{x} = \rho\mathbf{t}$, we get

$$(f(\rho\mathbf{t}))^\wedge(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\mathbf{x})\, e^{-2\pi i(\boldsymbol{\omega},\, \rho^T\mathbf{x})} \mathrm{d}\mathbf{x}. \tag{2.9.2}$$

The inner product can be rewritten into the product of matrices $(\boldsymbol{\omega},\ \rho^T\mathbf{x}) = \boldsymbol{\omega}^T\rho^T\mathbf{x}$. The transpose of 1×1 matrix is equal to itself and $(\boldsymbol{\omega},\ \rho^T\mathbf{x})$ is a 1×1 matrix, so $(\boldsymbol{\omega},\ \rho^T\mathbf{x}) = \mathbf{x}^T\rho\boldsymbol{\omega} = (\mathbf{x},\ \rho\boldsymbol{\omega}) = (\rho\boldsymbol{\omega},\ \mathbf{x})$. From this and (2.9.2), we get

$$(f(\rho\mathbf{t}))^\wedge(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\mathbf{x})\, e^{-2\pi i(\rho\boldsymbol{\omega},\, \mathbf{x})} \mathrm{d}\mathbf{x} = \widehat{f}(\rho\boldsymbol{\omega}). \tag{2.9.3}$$

In detail, $f(\mathbf{t}) \overset{\rho}{\to} f(\rho\mathbf{t}) \overset{F}{\to} (f(\rho\mathbf{t}))^\wedge(\boldsymbol{\omega})$ is equivalent to $f(\mathbf{t}) \overset{F}{\to} \widehat{f}(\boldsymbol{\omega}) \overset{\rho}{\to} \widehat{f}(\rho\boldsymbol{\omega})$, where F is the Fourier transform.

(b) Fourier transforms of radial functions are still radial functions

If $f(\mathbf{t})$ is a radial function on $\mathbb{R}^d$, ρ is an orthogonal transform on $\mathbb{R}^d$, by (2.9.1) and (2.9.3), we get

$$\widehat{f}(\boldsymbol{\omega}) = (f(\rho\mathbf{t}))^\wedge(\boldsymbol{\omega}) = \widehat{f}(\rho\boldsymbol{\omega}).$$

So $\widehat{f}(\boldsymbol{\omega})$ is also a radial function on $\mathbb{R}^d$.

(ii) 2−Dimensional Invariant Subspace

The polar coordinate representation of Fourier transform on $\mathbb{R}^2$:

$$\widehat{f}(\boldsymbol{\omega}) = \int_0^\infty \int_0^{2\pi} f(r\, e^{i\theta})\, e^{-2\pi i Rr\cos(\varphi-\theta)}\, r\,\mathrm{d}r\,\mathrm{d}\theta \qquad (\boldsymbol{\omega} = R\, e^{i\varphi}). \tag{2.9.4}$$

The subspace $D_k^{(2)} (k \in \mathbb{Z})$ is the set of functions f on $\mathbb{R}^2$ satisfying $f(\mathbf{z}) = g(r)\, e^{ik\theta}$, where $\mathbf{z} = r\, e^{i\theta}$ and $g(r)$ only depends on r. If $f \in D_k^{(2)}$, then $f(\mathbf{z}) = g(r)\, e^{ik\theta} (\mathbf{z} = re^{i\theta})$. Let $h(\mathbf{z}) = f(e^{i\varphi}\mathbf{z})$ for a fixed φ. Then,

$$h(\mathbf{z}) = f(e^{i\varphi}\mathbf{z}) = g(r)\, e^{ik(\theta+\varphi)} = e^{ik\varphi} f(\mathbf{z}).$$

This implies that

$$\widehat{h}(\boldsymbol{!}) = e^{ik\varphi}\widehat{f}(\boldsymbol{!}). \tag{2.9.5}$$

On the other hand, since Fourier transform and orthogonal transform are exchangeable and $e^{i\varphi}\mathbf{z}$ is a rotation, it follows from $h(\mathbf{z}) = f(e^{i\varphi}\mathbf{z})$ that $\widehat{h}(\boldsymbol{\omega}) = \widehat{f}(e^{i\varphi}\boldsymbol{\omega})$. Comparing this with (2.9.5) gives $\widehat{f}(e^{i\varphi}\boldsymbol{\omega}) = e^{ik\varphi}\widehat{f}(\boldsymbol{\omega})$. Let $\boldsymbol{\omega} = R$. Then,

$$\widehat{f}(R\, e^{i\varphi}) = \widehat{f}(R)\, e^{ik\varphi}, \tag{2.9.6}$$

i.e., $\widehat{f} \in D_k^{(2)}$. Hence, $D_k^{(2)}$ is a invariant subspace under Fourier transform.

In the invariant subspace $D_k^{(2)}$, if $k \neq l$, let $\xi \in D_k^{(2)}$ and $\eta \in D_l^{(2)}$. Then, $\xi(t) = \mu(r)\, e^{ik\theta}$, $\eta(t) = \tau(r)\, e^{il\theta}$. Since $\int_0^{2\pi} e^{i(k-l)\theta}\mathrm{d}\theta = 0\, (k \neq l)$,

$$(\xi,\, \eta) = \int_{\mathbb{R}} \mu(r)\tau(r) \left(\int_0^{2\pi} e^{i(k-l)}\mathrm{d}\theta\right) r\mathrm{d}r = 0,$$

So $D_k^{(2)} \perp D_l^{(2)}$, i.e., $D_k^{(2)}$ and $D_l^{(2)}$ are orthogonal.

This means that the sequence of subspaces $\{D_k^{(2)}\}_{k\in\mathbb{Z}}$ is an orthogonal sequence. Any $f(\mathbf{z}) = f(r\, e^{i\theta})$ is a $2\pi-$periodic function of θ. Expand it into Fourier series:

$$f(\mathbf{z}) = f(r\, e^{i\theta}) = \sum_{k\in\mathbb{Z}} f_k(r)\, e^{ik\theta},$$

where Fourier coefficient $f_k(r) = \frac{1}{2\pi}\int_0^{2\pi} f(r\,e^{i\theta})\,e^{-ik\theta}\mathrm{d}\theta$. Clearly, each $f_k(r)\,e^{ik\theta}$ is a function in $D_k^{(2)}$. Hence, $L^2(\mathbb{R}^2)$ is an orthogonal sum of $D_k^{(2)}$ ($k \in \mathbb{Z}$),

$$L^2(\mathbb{R}^2) = \bigoplus_{k=0}^{\infty} D_k^{(2)}$$

and each subspace $D_k^{(2)}$ is invariant under the Fourier transform. So any function $f \in L^2(\mathbb{R}^2)$ can be decomposed into $f(\mathbf{z}) = \sum_k F_k$, where $F_k \in D_k^{(2)}$.

If $f \in D_k^{(2)}$ and $f(\mathbf{z}) = g(r)\,e^{ik\theta}$, then, by (2.9.4),

$$\widehat{f}(\boldsymbol{\omega}) = \int_0^{\infty}\int_0^{2\pi} g(r)\,e^{ik\theta}\,e^{-2\pi i Rr\cos(\varphi-\theta)} r\mathrm{d}r\mathrm{d}\theta \qquad (\boldsymbol{\omega} = R\,e^{i\varphi}).$$

Let $\varphi = 0$ and then let $\theta = \frac{\pi}{2} - \theta^*$. Then,

$$\begin{aligned}\widehat{f}(R) &= \textstyle\int_0^{\infty} g(r)\left(\int_0^{2\pi} e^{-2\pi i Rr\cos\theta}\,e^{ik\theta}\,\mathrm{d}\theta\right) r\mathrm{d}r \\ &= 2\pi(-i)^k \textstyle\int_0^{\infty} g(r)\left(\frac{1}{2\pi}\int_0^{2\pi} e^{-2\pi i Rr\sin\theta^*}\,e^{-ik\theta^*}\mathrm{d}\theta^*\right) r\mathrm{d}r.\end{aligned}$$

Note that Bessel function $J_k(t)$ has the integral representation: $J_k(t) = \frac{1}{2\pi}\int_0^{2\pi} e^{it\sin\theta}$ $e^{-ik\theta}\mathrm{d}\theta$ ($k \in \mathbb{Z}$). So

$$\widehat{f}(R) = 2\pi(-i)^k \int_0^{\infty} g(r) J_k(2\pi Rr) r\mathrm{d}r.$$

From this and (2.9.6),

$$\widehat{f}(R\,e^{i\varphi}) = 2\pi(-i)^k \left(\int_0^{\infty} g(r) J_k(2\pi Rr) r\mathrm{d}r\right) e^{ik\varphi}.$$

(iii) $d-$Dimensional Invariant Subspace

The *subspace* $D_k^{(d)}$ is the set of all linear combinations of functions of the form $\tau(\|\,\mathbf{t}\,\|)p(\mathbf{t})$, where $\tau(\|\,\mathbf{t}\,\|)$ ranges over all radial functions and $p(\mathbf{t})$ over all solid spherical harmonic functions of degree k.

If $f(\mathbf{t}) = \tau(\|\,\mathbf{t}\,\|)$, where $\tau(\|\,\mathbf{t}\,\|)$ is a radial function, then its Fourier transform $\widehat{f}(\boldsymbol{\omega}) = F_0(|\boldsymbol{\omega}|)$, where

$$F_0(|\boldsymbol{\omega}|) = F_0(R) = 2\pi R^{-\frac{d-2}{2}} \int_0^{\infty} \tau(s) J_{\frac{d-2}{2}}(2\pi Rs) s^{\frac{d}{2}}\mathrm{d}s \qquad (|\omega| = R)$$

is also a radial function, where $J_\nu(x)$ is the $\nu-$order Bessel function and

$$J_\nu(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(\nu\theta - x\sin\theta)}\mathrm{d}\theta.$$

If $f(\mathbf{t}) = \tau(\|\mathbf{t}\|)\, p(\mathbf{t})$, where $p(\mathbf{t})$ is a (solid) spherical harmonic function of degree k, then its Fourier transform $\widehat{f}(\boldsymbol{\omega}) = F_k(|\boldsymbol{\omega}|)p(\boldsymbol{\omega})$, where

$$F_k(|\boldsymbol{\omega}|) = F_k(R) = 2\pi(i)^{-k}R^{-\frac{d-2}{2}-k}\int_0^\infty \tau(s)J_{\frac{d-2}{2}+k}(2\pi Rs)s^{\frac{d}{2}+k}\mathrm{d}s \quad (|\omega| = R).$$

This implies that each $D_k^{(d)}$ is an invariant subspace in $L^2(\mathbb{R}^d)$ and the Fourier transform in $D_k^{(d)}$ corresponds to a Bessel transform.

Similarly, if $k \neq l$, then $D_k^{(d)} \perp D_l^{(d)}$, i.e., $D_k^{(d)}$ and $D_l^{(d)}$ are orthogonal. Hence, $L^2(\mathbb{R}^2)$ is an orthogonal sum of $D_k^{(d)}$ $(k = 0, 1, \ldots)$, i.e.,

$$L^2(\mathbb{R}^d) = \bigoplus_{k=0}^{\infty} D_k^{(d)}.$$

For $f \in L^2(\mathbb{R}^d)$, there exists a sequence $\{f_k\}_{k=0,1,\ldots}$ of functions such that

$$f(\mathbf{t}) = \sum_{k=0}^{\infty} f_k(\mathbf{t}),$$

where $f_k \in D_k^{(d)}$ and $f_k(\mathbf{t}) = \tau_k(\|\mathbf{t}\|)p_k(\mathbf{t})$, i.e., f can be expanded into an orthogonal series:

$$f(\mathbf{t}) = \sum_{k=0}^{\infty} \tau_k(\|\mathbf{t}\|)p_k(\mathbf{t}),$$

where $\tau_k(\|\mathbf{t}\|)$ is a radial function and $p_k(\mathbf{t})$ is a solid spherical harmonic function of degree k.

2.10 Harmonic Analysis on General Domains

More and more data are collected in a distributed and irregular manner. In this section, we will analyze the spatial frequency information of multivariate time series on general domains.

2.10.1 Symmetric Kernels

Let $k(\mathbf{x}, \mathbf{t})$ be a continuous function on a domain $\Omega \times \Omega$, where $\mathbf{x} = (x_1, \ldots, x_d) \in \Omega$ and $\mathbf{t} = (t_1, \ldots, t_d) \in \Omega$. A continuous functional F on Ω is called a functional expressed by the kernel $k(\mathbf{x}, \mathbf{t})$ if it maps a continuous function $h(\mathbf{t})$ on Ω into

$$\int_\Omega k(\mathbf{x}, \mathbf{t}) h(\mathbf{t}) \mathrm{d}\mathbf{t} \ (\mathbf{x} \in \Omega).$$

The kernel $k(\mathbf{x}, \mathbf{t})$ is said to be a *symmetric kernel* if it satisfies the condition $k(\mathbf{x}, \mathbf{t}) = k(\mathbf{t}, \mathbf{x})$ $(\mathbf{x},\ \mathbf{t} \in \Omega)$. When $\lambda = \lambda_k$, if the integral equation:

$$\varphi(\mathbf{x}) = \lambda \int_\Omega k(\mathbf{x}, \mathbf{t}) \varphi(\mathbf{t}) \mathrm{d}\mathbf{t}$$

has nonzero solutions, then λ_k is said to be an *eigenvalue of the kernel function* $k(\mathbf{x}, \mathbf{t})$ and the nonzero solutions, denoted by φ_k, of the integral equation are said to be *eigenfunctions corresponding to the eigenvalue* λ_k, i.e.,

$$\varphi_k(\mathbf{x}) = \lambda_k \int_\Omega k(\mathbf{x}, \mathbf{t}) \varphi_k(\mathbf{t}) \mathrm{d}\mathbf{t} \qquad (\mathbf{x} \in \Omega).$$

Any function expressed by the symmetric kernel $k(\mathbf{x}, \mathbf{t})$ has infinite eigenvalues. Denote all eigenvalues by $|\lambda_1| \le |\lambda_2| \le |\lambda_3| \le \cdots$ and $|\lambda_k| \to \infty$ $(k \to \infty)$. The corresponding eigenfunctions $\{\varphi_k\}$ are normal orthogonal, i.e., $\int_\Omega \varphi_k(\mathbf{t}) \varphi_l(\mathbf{t}) \mathrm{d}\mathbf{t} = \delta_{k,l}$, where $\delta_{k,l}$ is the Kronecker delta.

Let $f(\mathbf{x})$ be a function on the general domain Ω. Then, the Fourier series of f:

$$f(\mathbf{x}) = \sum_{k \in \mathbb{Z}_+} c_k \varphi_k(\mathbf{x}),$$

where $c_k = \int_\Omega f(\mathbf{x}) \varphi_k(\mathbf{x}) \mathrm{d}\mathbf{x}$.

2.10.2 Smooth Extensions and Approximation

Smooth extensions of the higher-dimensional function are similar to those of the two-dimensional functions. However, the representation of the higher-dimensional case is too complicated. Therefore, we only consider the two-dimensional case.

(i) Smooth Extension

Let $\Omega \subset [0, 1]^2$ be a planer domain and its boundary $\partial\Omega$ be a piecewise very smooth curve, and let f be a function on Ω and its all partial derivatives are continuous on Ω. Then, for $r \in \mathbb{Z}_+$, there exists a smooth function $F(x, y)$ on $[0, 1]^2$ such that

(a) $F(x, y) = f(x, y)$ $((x, y) \in \Omega)$;
(b) $\frac{\partial^{i+j} F}{\partial x^i \partial y^j}(x, y)$ $(i + j \leq r)$ is continuous on $[0, 1]^2$;
(c) $\frac{\partial^{i+j} F}{\partial x^i \partial y^j}(x, y) = 0$ $(i + j \leq r)$ on the boundary $\partial[0, 1]^2$;
(d) $F(x, y)$ can be expressed locally in the form:

$$\sum_{j=0}^{L} \xi_j(x) y^j \quad \text{or} \quad \sum_{j=0}^{L} \eta_j(y) x^j \quad \text{or} \quad \sum_{i,j=0}^{L} c_{ij} x^i y^j$$

on the complement $[0, 1]^2 \setminus \Omega$, where $L \in \mathbb{Z}_+$ and each c_{ij} is constant.

The construction of the smooth extension $F(x, y)$ of $f(x, y)$ from Ω to $[0, 1]^2$ is stated as follows. Without loss of generality, the complement $[0, 1]^2 \setminus \Omega$ is divided into some trapezoids with a curved side and some rectangles. For convenience of representation, assume that there are four points $(x_\nu, y_\nu) \in \partial\Omega$ $(\nu = 1, 2, 3, 4)$ such that $[0, 1]^2 \setminus \Omega$ is divided into the following four rectangles:

$$H_1 = [0, x_1] \times [0, y_1], \qquad H_1 = [x_2, 1] \times [0, y_2],$$

$$H_3 = [x_3, 1] \times [y_3, 1], \qquad H_4 = [0, x_4] \times [y_4, 1]$$

and four trapezoids with a curved side:

$$E_1 = \{(x, y) : \quad x_1 \leq x \leq x_2, \quad 0 \leq y \leq g(x)\},$$

$$E_2 = \{(x, y) : \quad h(y) \leq x \leq 1, \quad y_2 \leq y \leq y_1\},$$

$$E_3 = \{(x, y) : \quad x_4 \leq x \leq x_3, \quad g^*(x) \leq y \leq 1\},$$

$$E_4 = \{(x, y) : \quad 0 \leq x \leq h^*(y), \quad y_1 \leq y \leq y_4\}.$$

The curve side of each trapezoid E_ν is on $\partial\Omega$, and the bottom side is on $\partial([0, 1]^2)$.

First, we extend f smoothly to each trapezoid E_ν.

By similarity, we only state how to extend f from Ω to $\Omega \bigcup E_1$. Using two sequences $\{a_{k,1}(x, y)\}_{k=0,1,\ldots}$ and $\{b_{k,1}(x, y)\}_{k=0,1,\ldots}$, where

$$a_{0,1}(x, y) = \tfrac{y}{g(x)}, \qquad a_{k,1}(x, y) = \tfrac{(y-g(x))^k}{k!} \left(\tfrac{y}{g(x)}\right)^{k+1} \quad (k \in \mathbb{Z}_+),$$

$$b_{0,1}(x, y) = \tfrac{y-g(x)}{-g(x)}, \qquad b_{k,1}(x, y) = \tfrac{y^k}{k!} \left(\tfrac{y-g(x)}{-g(x)}\right)^{k+1} \quad (k \in \mathbb{Z}_+),$$

by induction, we define a sequence of functions $\{S_1^{(k)}(x, y)\}_{k=0,1,\ldots}$ on E_1 as follows:

$$S_1^{(0)}(x,y) = f(x,g(x))\,a_{0,1}(x,y) \qquad (x_1 \le x \le x_2,\ \ 0 \le y \le g(x)).$$

In general, for $k \in \mathbb{Z}_+$,

$$S_1^{(k)}(x,y) = S_1^{(k-1)}(x,y) + a_{k,1}(x,y)\left(\tfrac{\partial^k f}{\partial y^k}(x,g(x)) - \tfrac{\partial^k S_1^{(k-1)}}{\partial y^k}(x,g(x))\right)$$

$$-b_{k,1}(x,y)\tfrac{\partial^k S_1^{(k-1)}}{\partial y^k}(x,0) \qquad (x_1 \le x \le x_2,\ \ 0 \le y \le g(x)).$$

Similarly, we construct $S_2^{(k)}(x,y)$, $S_3^{(k)}(x,y)$, and $S_4^{(k)}(x,y)$ on E_1, E_2, and E_3, respectively.

Next, we smoothly extend f to each rectangle $H_\nu(\nu = 1,2,3,4)$. By similarity, we only state the smooth extension of f on H_1. Let

$$\alpha_{k,11}(y) = \tfrac{(y-y_1)^k}{k!}\left(\tfrac{y}{y_1}\right)^{k+1},$$

$$\beta_{k,11}(y) = \tfrac{y^k}{k!}\left(\tfrac{y-y_1}{-y_1}\right)^{k+1} \qquad (k = 0,1,\ldots),$$

and let

$$M_1^{(0)}(x,y) = S_4^{(l_r)}(x,y_1)\,\alpha_{0,11}(y) \qquad (l_r = r(r+1)(r+2)(r+3)),$$

$$M_1^{(k)}(x,y) = M_1^{(k-1)}(x,y) + \alpha_{k,11}(y)\left(\tfrac{\partial^k S_4^{(l_r)}}{\partial y^k}(x,y_1) - \tfrac{\partial^k M_1^{(k-1)}}{\partial y^k}(x,y_1)\right)$$

$$-\beta_{k,11}(y)\tfrac{\partial^k M_1^{(k-1)}}{\partial y^k}(x,0)(k = 1,\ldots,\tau_r,\ (x,y) \in H_1),$$

where $\tau_r = r(r+2)$, and let

$$\alpha_{k,14}(x) = \tfrac{(x-x_1)^k}{k!}\left(\tfrac{x}{x_1}\right)^{k+1},$$

$$\beta_{k,14}(x) = \tfrac{x^k}{k!}\left(\tfrac{x-x_1}{-x_1}\right)^{k+1} \qquad (k = 0,1,\ldots),$$

and let

$$N_1^{(0)}(x,y) = (S_1^{(l_r)}(x_1,y) - M_1^{(\tau_r)}(x_1,y))\,\alpha_{0,14}(x),$$

$$N_1^{(k)}(x,y) = N_1^{(k-1)}(x,y) + \alpha_{k,14}(x)\left(\tfrac{\partial^k (S_1^{(l_r)} - M_1^{(\tau_r)})}{\partial x^k}(x_1,y) - \tfrac{\partial^k N_1^{(k-1)}}{\partial x^k}(x_1,y)\right)$$

$$-\beta_{k,14}(x)\tfrac{\partial^k N_1^{(k-1)}}{\partial x^k}(0,y)(k = 1,\ldots,r,\ (x,y) \in H_1).$$

Similarly, we construct $M_\nu^{(\tau_r)}(x, y)$ and $N_\nu^{(r)}(x, y)$ on $H_\nu (\nu = 2, 3, 4)$.

Finally, let

$$F(x, y) = \begin{cases} f(x, y), & (x, y) \in \Omega, \\ S_\nu^{(l_r)}(x, y), & (x, y) \in E_\nu \ \ (\nu = 1, 2, 3, 4), \\ M_\nu^{(\tau_r)}(x, y) + N_\nu^{(r)}(x, y), & (x, y) \in H_\nu \ \ (\nu = 1, 2, 3, 4). \end{cases} \tag{2.10.1}$$

Then, the function $F(x, y)$ is the desired smooth extension of f from Ω to $[0, 1]^2$.
(ii) Fourier Approximation
Let $f(x, y)$ and its all partial derivatives be continuous on $\Omega \subset [0, 1]^2$, and let $F(x, y)$ [see (2.10.1)] be its smooth extension from Ω to $[0, 1]^2$. We extend $F(x, y)$ to an one-periodic function on $\mathbb{R}^2$, still denoted by $F(x, y)$, and then expand $F(x, y)$ into a Fourier series. Note that $F(x, y) = f(x, y)$ on Ω. Then,

$$f(x, y) = \sum_{m,n\in\mathbb{Z}} c_{mn}(F)\, e^{2\pi i(mx+ny)} \qquad ((x, y) \in \Omega), \tag{2.10.2}$$

and the Fourier coefficients $c_{mn}(F)$ decay fast, where $c_{mn}(F) = \int_0^1 \int_0^1 F(x, y)$ $e^{-2\pi i(mx+ny)}\mathrm{d}x\mathrm{d}y$.

Take the hyperbolic cross truncations of the Fourier series (2.10.2):

$$s_N^{(h,c)}(f; x, y) = \sum_{\substack{|n_1 n_2|\le N \\ |n_1|,|n_2|\le N}} c_{mn}(F)\, e^{2\pi i(mx+ny)}.$$

So $f(x, y) \sim s_N^{(h,c)}(f; x, y)$ with a small error, i.e., $f(x, y)$ is reconstructed by few Fourier coefficients $c_{mn}(F)$.

Let $F^{(o)}(x, y)$ be an odd extension of $F(x, y)$ from $[0, 1]^2$ to $[-1, 1]^2$:

$$F^{(o)}(x, y) = \begin{cases} F(x, y), & (x, y) \in [0, 1]^2, \\ -F(-x, y), & (x, y) \in [-1, 0] \times [0, 1], \\ F(-x, -y), & (x, y) \in [-1, 0]^2, \\ -F(x, -y), & (x, y) \in [0, 1] \times [-1, 0]. \end{cases}$$

We extend $F^{(o)}(x, y)$ to a 2-periodic function, still denoted by $F^{(o)}(x, y)$, and then expand it into a Fourier sine series. Note that $F^{(o)}(x, y) = f(x, y)$ on Ω. Then,

$$f(x, y) = \sum_{m\in\mathbb{Z}_+} \sum_{n\in\mathbb{Z}_+} c_{mn}(F^{(o)}) \sin(m\pi x) \sin(n\pi y) \qquad ((x, y) \in \Omega),$$

where the Fourier sine coefficients are

$$c_{mn}(F^{(o)}) = \int_{-1}^{1}\int_{-1}^{1} F^{(o)}(x, y)\sin(m\pi x)\sin(n\pi y)\mathrm{d}x\mathrm{d}y$$

and $c_{mn}(F^{(o)})$ decay fast.

Let $F^{(e)}(x, y)$ be an even extension of $F(x, y)$ from $[0, 1]^2$ to $[-1, 1]^2$:

$$F^{(e)}(x, y) = \begin{cases} F(x, y), & (x, y) \in [0, 1]^2, \\ F(-x, y), & (x, y) \in [-1, 0] \times [0, 1], \\ F(-x, -y), & (x, y) \in [-1, 0]^2, \\ F(x, -y), & (x, y) \in [0, 1] \times [-1, 0]. \end{cases}$$

We extend $F^{(e)}(x, y)$ to a 2-periodic function, still denoted by $F^{(e)}(x, y)$, and then expand it into a Fourier cosine series. Note that $F^{(e)}(x, y) = f(x, y)$ on Ω. Then,

$$f(x, y) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} c_{mn}(F^{(e)})\cos(m\pi x)\cos(n\pi y) \qquad ((x, y) \in \Omega),$$

where the Fourier cosine coefficients $c_{mn}(F^{(o)})$ decay fast.

If we consider the hyperbolic cross truncations of the Fourier sine (or cosine) series, $f(x, y)$ on Ω can be reconstructed by few Fourier sine (or cosine) coefficients $c_{mn}(F^{(o)})$ or $c_{mn}(F^{(e)})$.

2.11 Harmonic Analysis on Graphs

Harmonic analysis tools originally developed for Euclidean spaces and regular lattices are now being transferred to the general settings of graphs in order to analyze geometric and topological structures, as well as data measured on them. A graph G has n vertices $v_0, v_1, \ldots, v_{n-1}$. If there is an edge between two vertices v_i and v_j, we say that v_i and v_j are *adjacent*, denoted by $v_i \sim v_j$. Denote by d_i the number of adjacent vertices of the vertex v_i. The d_i is called the *degree* of the vertex v_i.

2.11.1 The Laplacian of a Graph

The Laplacian matrix $\mathcal{L}$ of a graph G is defined as $\mathcal{L} = (\mathcal{L}(v_i, v_j))_{i,j=0,1,\ldots,n-1}$, where

$$\mathcal{L}(v_i, v_j) = \begin{cases} 1 & \text{if } i = j \text{ and } d_i \neq 0, \\ -\frac{1}{\sqrt{d_i d_j}} & \text{if } v_i \text{ and } v_j \text{ are adjacent}, \\ 0 & \text{otherwise.} \end{cases} \quad (2.11.1)$$

Let g be a function defined at vertices $v_0, v_1, \ldots, v_{n-1}$. Then, $\mathbf{g} = (g(v_0), \ldots, g(v_{n-1}))^T$ is a column vector. By (2.11.1), it follows that $\mathcal{L}\mathbf{g} = (\gamma_0, \ldots, \gamma_{n-1})^T$, where

$$\gamma_i = \frac{1}{\sqrt{d_i}} \sum_{j=0}^{n-1} \left(\frac{g(v_i)}{\sqrt{d_i}} - \frac{g(v_j)}{\sqrt{d_j}} \right) \qquad (i = 0, 1, \ldots, n-1). \quad (2.11.2)$$

Let $\mathbf{g}_0 = (\sqrt{d_0}, \ldots, \sqrt{d_{n-1}})^T$. By (2.11.2), $\mathcal{L}\mathbf{g}_0 = \mathbf{0}$ or $\mathcal{L}\mathbf{g}_0 = \lambda_0 \mathbf{g}_0$, where $\lambda_0 = 0$. This implies that λ_0 is a eigenvalue of the Laplacian matrix $\mathcal{L}$ and $\mathbf{g}_0$ is the corresponding eigenvector. Since the Laplacian matrix $\mathcal{L}$ is symmetric, its all eigenvalues are nonnegative real numbers, denoted by $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1}$. The set of eigenvalues is called the *spectrum* of $\mathcal{L}$ (or the spectrum of the associated graph G).

Let T be a diagonal matrix:

$$T = \begin{pmatrix} d_0 & 0 & \cdots & 0 \\ 0 & d_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{n-1} \end{pmatrix},$$

and let $\mathbf{I}$ be a $n-$dimensional column vector: $\mathbf{I} = (1, 1, \ldots, 1)^T$. Then, $T\mathbf{I} = (d_0, d_1, \ldots, d_{n-1})^T$.

Let $\mathbf{f} = (f(v_0), f(v_1), \ldots, f(v_{n-1}))^T$. The notation $\mathbf{f} \perp T\mathbf{I}$ means that the inner product of two $n-$dimensional vectors $\mathbf{f}$ and $T\mathbf{I}$ vanishes, i.e., $\sum_{i=0}^{n-1} f(v_i) d_i = 0$. The eigenvalue λ_1 is

$$\lambda_1 = \inf_{\mathbf{f} \perp T\mathbf{I}} \frac{\sum\limits_{v_i \sim v_j} (f(v_i) - f(v_j))^2}{\sum\limits_{i=0}^{n-1} f^2(v_i)\, d_i}. \tag{2.11.3}$$

If the function $\mathbf{f}_1$ at vertices attains the infimum, then $\mathbf{f}_1$ is a harmonic eigenfunction of $\mathcal{L}$. The largest eigenvalue is

$$\lambda_{n-1} = \sup_{\mathbf{f}} \frac{\sum\limits_{v_i \sim v_j} (f(v_i) - f(v_j))^2}{\sum\limits_{i=0}^{n-1} f^2(v_i)\, d_i}.$$

Let S_{k-1} be the subspace generated by the harmonic eigenfunctions corresponding to λ_i for $i \le k-1$. Then,

$$\lambda_k = \inf_{\mathbf{f}} \sup_{\mathbf{g} \in S_{k-1}} \frac{\sum\limits_{v_i \sim v_j} (f(v_i) - f(v_j))^2}{\sum\limits_{i=0}^{n-1} (f(v_i) - g(v_i))^2} = \inf_{\mathbf{f} \perp T S_{k-1}} \frac{\sum\limits_{v_i \sim v_j} (f(v_i) - f(v_j))^2}{\sum\limits_{i=0}^{n-1} f^2(v_i)\, d_i} \quad (k = 2, \ldots, n-2).$$

2.11.2 Eigenvalues and Eigenfunctions

From $(f(v_i) - f(v_j))^2 \le 2f^2(v_i) + 2f^2(v_j)$, it follows that

$$\sum_{v_i \sim v_j} (f(v_i) - f(v_j))^2 \le \sum_{i=0}^{n-1} f^2(v_i)\, d_i + \sum_{j=0}^{n-1} f^2(v_j)\, d_j = 2 \sum_{i=0}^{n-1} f^2(v_i)\, d_i,$$

where in the sum of the left-hand side, $f(v_i) - f(v_j)$ occurs only one time. So

$$\frac{\sum\limits_{v_i \sim v_j} (f(v_i) - f(v_j))^2}{\sum\limits_{i=0}^{n-1} f^2(v_i)\, d_i} \le 2.$$

From this and the above expressions of λ_k, we get an estimate $0 \le \lambda_k \le 2\,(k = 0, 1, \ldots, n-1)$.

Now, we estimate the lower bound of λ_1. Define the *distance* between two vertices

v_i and v_j by the number of edges in a shortest path joining v_i and v_j, denoted by $\operatorname{dist}(v_i, v_j)$. Define the *diameter* of a graph G as $D = \max_{i,j=0,1,\ldots,n-1} \operatorname{dist}(v_i, v_j)$. Let G be a connected graph with the diameter D.

Suppose that $\mathbf{f}$ is a harmonic eigenfunction. By (2.11.3), it follows that $\mathbf{f} \perp T\mathbf{I}$ and

$$\lambda_1 = \frac{\sum\limits_{v_i \sim v_j} (f(v_i) - f(v_j))^2}{\sum\limits_{i=0}^{n-1} f^2(v_i) d_i}. \tag{2.11.4}$$

Let v_{k_0} be a vertex with $|f(v_{k_0})| = \max_{i=0,\ldots,n_1} |f(v_i)|$. Since $\mathbf{f} \perp T\mathbf{I}$, the sum $\sum_{i=0}^{n-1} f(v_i) d_i = 0$. So there exists a v_s such that $f(v_{k_0}) f(v_s) < 0$. Let P be the shortest path in G joining v_{k_0} and v_s. By (2.11.4),

$$\lambda_1 \ge \frac{\sum\limits_{(v_i, v_j) \in P} (f(v_i) - f(v_j))^2}{f^2(v_{k_0}) \operatorname{vol} G}, \tag{2.11.5}$$

where $\operatorname{vol} G = \sum\limits_{i=0}^{n-1} d_i$ is called the *volume* of G. Let $l+1$ vertices $v_{k_0}, v_{k_1}, \ldots, v_{k_l} \in P$, where $v_{k_l} = v_s$ and $l = \operatorname{dist}(v_{k_0}, v_s)$. According to the Schwarz inequality and $v_{k_l} = v_s$, $l \le D$, and $f(v_{k_0}) f(v_s) < 0$, it follows that

$$\begin{aligned} \sum_{(v_i, v_j) \in P} (f(v_i) - f(v_j))^2 &\ge \tfrac{1}{l} \sum_{j=1}^{l} (f(v_{k_j}) - f(v_{k_{j-1}}))^2 \\ &\ge \tfrac{1}{l} (f(v_{k_l}) - f(v_{k_0}))^2 = \tfrac{(f(v_s) - f(v_{k_0}))^2}{\operatorname{dist}(v_{k_0}, v_s)} \ge \tfrac{f^2(v_{k_0})}{D}. \end{aligned}$$

From this and (2.11.5),

$$\lambda_1 \ge (D \operatorname{vol} G)^{-1} = \left(D \sum_{i=0}^{n-1} d_i \right)^{-1}.$$

2.11.3 Fourier Expansions

Let G be a graph with vertices $v_0, \ldots, v_{n-1}$ and the set $E(G)$ be edges. The Laplacian matrix $\mathcal{L}$ of G is an $n \times n$ positive semi-definite matrix which is stated in (2.11.1). The eigenvectors $\mathbf{g}_0, \ldots, \mathbf{g}_{n-1}$ of $\mathcal{L}$ form a normal orthogonal basis for the space H_n of functions at vertices $v_0, \ldots, v_{n-1}$, i.e., for $i, j = 0, \ldots, n-1$,

$$(\mathbf{g}_i, \mathbf{g}_j) = \sum_{k=0}^{n-1} g_i(v_k) g_j(v_k) = \begin{cases} 1, \ i = j, \\ \\ 0, \ i \neq j, \end{cases}$$

where $\mathbf{g}_i = (g_i(v_0), \ldots, g_i(v_{n-1}))^T$ and $\mathcal{L}\mathbf{g}_i = \lambda_i \mathbf{g}_i$ $(i = 0, 1, \ldots, n-1)$. So any $\mathbf{f} = (f(v_0), \ldots, f(v_{n-1}))^T \in H_n$ can be expanded into Fourier series with respect to eigenvectors $\{\mathbf{g}_i\}_{i=0,\ldots,n-1}$:

$$f(v) = \sum_{i=0}^{n-1} c_i g_i(v) \qquad (v = v_0, \ldots, v_{n-1}),$$

where Fourier coefficients $c_i = (\mathbf{f}, \mathbf{g}_i) = \sum_{k=0}^{n-1} f(v_k) g_i(v_k)$ $(i = 0, \ldots, n-1)$, and they satisfy the Parseval identity:

$$\sum_{k=0}^{n-1} f^2(v_k) = \sum_{k=0}^{n-1} c_k^2.$$

Consider a *graph G (possibly with loops) with weighted function* $w(v_i, v_j)$ $(i, j = 0, \ldots, n-1)$ satisfying $w(v_i, v_j) = w(v_j, v_i)$ and $w(v_i, v_j) \geq 0$, where $v_0, \ldots, v_{n-1}$ are vertices of the graph G. Unweighted graphs are the special case where all weights are 0 or 1. The *degree* of a vertex v_i is defined as $d_i = \sum_{j=0}^{n-1} w(v_i, v_j)$. The Laplacian matrix $\mathcal{L}_w$ of the graph G is

$$\mathcal{L}_w(v_i, v_j) = \begin{cases} 1 - \frac{w(v_i, v_j)}{d_i} & \text{if } i = j \text{ and } d_i \neq 0, \\ \\ -\frac{w(v_i, v_j)}{\sqrt{d_i d_j}} & \text{if } i \text{ and } j \text{ are adjacent}, \\ \\ 0 & \text{otherwise}. \end{cases}$$

All definitions and results for graphs can be generalized to the weighted graphs. So the eigenvectors $\{\mathbf{g}_i^w\}_{i=0,\ldots,n-1}$ of the Laplacian matrix $\mathcal{L}_w$ form a normal orthogonal basis for H_n. Any function $\mathbf{f} = (f(v_0), \ldots, f(v_{n-1}))^T \in H_n$ at vertices of the graph G can be expanded into Fourier series:

$$f(v) = \sum_{i=0}^{n-1} c_i^w g_i^w(v) \qquad (v = v_0, \ldots, v_{n-1}),$$

where $c_i^w = (\mathbf{f}, \mathbf{g}_i^w) = \sum_{k=0}^{n-1} f(v_k) g_i^w(v_k)$. The *contraction* of a graph means that two distinct vertices v_i and v_j are identified into a vertex v^*. The *weights* of edges associated with v^* are defined as

$$w(v_k, v^*) = w(v_k, v_i) + w(v_k, v_j),$$

$$w(v^*, v^*) = w(v_i, v_i) + w(v_j, v_j) + 2w(v_i, v_j).$$

If a graph G^* is the contraction of the graph G, then the eigenvalue $\lambda_1^G \leq \lambda_1^{G^*}$.

Further Reading

J. Aalto, P. Pirinen, J. Heikkinen et al., Spatial interpolation of monthly climate data for Finland: comparing the performance of kriging and generalized additive models. Theor. Appl. Climatol. **112**, 99–111 (2013)

M.W. Ashiq, C. Zhao, J. Ni et al., GIS-based high-resolution spatial interpolation of precipitation in mountain-plain areas of Upper Pakistan for regional climate change impact studies. Theor. Appl. Climatol. **99**, 239 (2010)

N.M. Atakishiyev, L.E. Vicent, K.B. Wolf, Continuous vs. discrete fractional Fourier transforms. J. Comput. Appl. Math. **107**, 73–95 (1999)

B.G.J. Brooks, Applying Wavelet and Fourier Transform Analysis to Large Geophysical Datasets, in *Computational Science*, vol. 5545, Lecture Notes in Computer Science, ed. by G. Allen, et al. (Springer, Heidelberg, 2009)

C. Candan, M.A. Kutay, H.M. Ozaktas, The discrete fractional Fourier transform. IEEE Trans. Sig. Process. **48**, 1329–1337 (2000)

X. Chen, P. Xing, Y. Luo et al., Surface temperature dataset for North America obtained by application of optimal interpolation algorithm merging tree-ring chronologies and climate model output. Theor. Appl. Climatol. **127**, 533–549 (2017)

W. Cheney, W. Light, *A Course in Approximation Theory* (Brooks/Cole Publishing, Pacific Grove, 2000)

R.A. DeVore, Nonlinear approximation. Acta Numerica **7**, 51–150 (1998)

S. Eghdamirad, F. Johnson, A. Sharma, Using second-order approximation to incorporate GCM uncertainty in climate change impact assessments. Clim. Change **142**, 37–52 (2017)

J. Fan, J. Meng, X. Chen et al., Network approaches to climate science. Sci. China Phys. Mech. Astron. **60**, 010531 (2017)

I. Fountalis, A. Bracco, C. Dovrolis, Spatio-temporal network analysis for studying climate patterns. Clim. Dyn. **42**, 879–899 (2014)

E.D. Giuseppe, G.J. Lasinio, S. Esposito et al., Functional clustering for Italian climate zones identification. Theor. Appl. Climatol. **114**, 39–54 (2013)

S.V. Henriksson, P. Ralsanen, J. Silen et al., Quasiperiodic climate variability with a period of 50–70 years: Fourier analysis of measurements and earth system model simulations. Clim. Dyn. **39**, 1999–2011 (2012)

A.L. Kay, S.M. Crooks, H.N. Davies et al., Probabilistic impacts of climate change on flood frequency using response surfaces I: England and Wales. Reg. Environ. Change **14**, 1215–1227 (2014)

R. Kandel, Understanding and measuring earth's energy budget: from fourier, Humboldt, and Tyndall to CERES and Beyond. Surv. Geophys. **33**, 337–350 (2012)

K. Kikuchi, An introduction to combined Fourier-wavelet transform and its application to convectively coupled equatorial waves. Clim. Dyn. **43**, 1339–1356 (2014)

V.A. Narayanan, K.M.M. Prabhu, The fractional Fourier transform: theory, implementation and error analysis. Microprocess. Microsyst. **27**, 511–521 (2003)

M. Ogurtsov, G. Kocharov, M. Lindholm et al., Evidence of solar variation in tree-ring-based climate reconstructions. Solar Phys. **205**, 403–417 (2002)

S. Samanta, D.K. Pal, D. Lohar et al., Interpolation of climate variables and temperature modeling. Theor. Appl. Climatol. **107**, 35–45 (2012)
E. Sejdic, I. Djurovic, L. Stankovic, Fractional Fourier transform as a signal processing tool: an overview of recent developments. Sig. Process. **91**, 1351–1369 (2011)
E.M. Stein, G. Weiss, *Introduction to Fourier analysis on Euclidean spaces* (Princeton University Press, Princeton, 1971)
Z. Wu, E.K. Schneider, B.P. Kirtman et al., The modulated annual cycle: an alternative reference frame for climate anomalies. Clim. Dyn. **31**, 823–841 (2008)
Z. Zhang, Approximation of bivariate functions via smooth extensions. Sci. World J. **2014**, 102062 (2014)
F. Zwiers, S. Shen, Errors in estimating spherical harmonic coefficients from partially sampled GCM output. Clim. Dyn. **13**, 703–716 (1997)

Springer
http://www.springer.com/978-3-319-67339-4

Multivariate Time Series Analysis in Climate and
Environmental Research
Zhang, Z.
2018, XII, 287 p. 7 illus., 3 illus. in color., Hardcover
ISBN: 978-3-319-67339-4