Machine learning under physical constraints
Fourier and wavelet representations

Sixin Zhang
(sixin.zhang@toulouse-inp.fr)

Outline

Outline

Fourier representation

Wavelet representation

Discrete Fourier transform (DFT)

- Discrete Fourier transform of $x \in \mathbb{R}^N$: Let $\omega_k = \frac{2\pi k}{N}$f

$$\widehat{x}(\omega_k) = \sum_{u=0}^{N-1} x(u) e^{-i\omega_k u}, \quad 0 \leq k < N$$

- It transforms x to $\widehat{x}$ using the orthonormal basis $\{\phi_k(u) = e^{i\omega_k u}\}_{k<N}$, i.e. $\widehat{x}(\omega_k) = \langle x, \phi_k \rangle$.
- Inverse DFT: As $\|\phi_k\|^2 = N$,

$$x(u) = \sum_{k=0}^{N-1} \frac{\widehat{x}(\omega_k)}{N} e^{i\omega_k u}$$

Discrete Fourier transform (DFT)

- DFT assumes a periodic extension of $x \in \mathbb{R}^N$ using circular translation.
- This means that when x is translated by g_τ:

$$\widehat{g_\tau \cdot x}(\omega_k) = \sum_{u=0}^{N-1} x(u-\tau) e^{-i\omega_k u} = \widehat{x}(\omega_k) e^{i\omega_k \tau}.$$

- When $N = 2^n$, Fast Fourier transform (FFT) computes $\widehat{x}$ in $O(N \log_2(N))$ time complexity $\Rightarrow$ Digital revolution.

Parseval identity and Convolution theorem

- Parseval identity (energy conservation): $\forall x \in \mathbb{R}^N$,

$$\|x\|^2 = \|\widehat{x}\|^2 / N$$

- Convolution theorem: consider circular convolution of $x \in \mathbb{R}^N$ and $h \in \mathbb{R}^N$,

$$x \star h(u) = \sum_v x(u-v) h(v)$$

Then for any $\omega_k = \frac{2\pi k}{N}$,

$$\widehat{x \star h}(\omega_k) = \widehat{x}(\omega_k) \widehat{h}(\omega_k)$$

  - This gives a fast way to compute the convolution using FFT.

Periodogram: Fourier invariant representations

- To build a representation $\Phi(x)$ which is translation invariant, we can take the modulus of Fourier coefficients:

$$\Phi(x) = \{|\widehat{x}(\omega_k)|^p\}_k, \quad \text{with} \quad p = 1, 2, \ldots$$

- Recall

$$\widehat{g_\tau \cdot x}(\omega_k) = \widehat{x}(\omega_k) e^{i\omega_k \tau}.$$

- Since $\widehat{g_\tau \cdot x}$ changes only the phase of $\widehat{x}$, therefore

$$\Phi(g_\tau \cdot x) = \Phi(x), \forall \tau \in \{0, \cdots, N-1\}$$

- When $p = 2$, $\Phi(X_N)/N$ is the **periodogram** of a stationary process $X_N(u)$ restricted on the interval $u \in \{0, \cdots, N-1\}$.

Power spectrum of stationary process

- For a (zero-mean) stationary process $X(u)$ on $u \in \mathbb{Z}$, its periodogram is related to its power spectrum. Let X_N be the restriction of X on the interval $u \in \{0, \cdots, N-1\}$.
- Taking the limit of N to infinity (as the domain size grows), the expected periodogram of X_N converges (under suitable assumptions) to the **power spectrum** of X:

$$R_X(\omega) = \lim_{N\to\infty} \frac{\mathbb{E}(|\widehat{X_N}(\omega)|^2)}{N}, \quad \omega \in [0, 2\pi]$$

- The power spectrum R_X is a property of the distribution of X (vs. periodogram)

Outline

Motivation

- Fourier transform is localized in frequency, but not in time or space.
- Wavelet transform is localized in both time and frequency, as music notes.
- Uncertainty principle: there is a trade-off to achieve both localization in both time and frequency.
  $\Rightarrow$ To study how to choose a proper transform to extract localized structures.

Construct wavelet family in 1d

- A wavelet is a localized function $\psi \in L^2(\mathbb{R})$ (in time and frequency) such that

$$\int_{\mathbb{R}} \psi(u) du = 0$$

- One can construct a wavelet family based on 2 groups:
  - Translation ($b \in \mathbb{R}$)
  - Dilation ($s > 0$)

$$\psi_{s,b}(u) = \frac{1}{\sqrt{s}} \psi(\frac{u-b}{s})$$

- The dilation s changes the scale of ψ.

Time localization of ψ

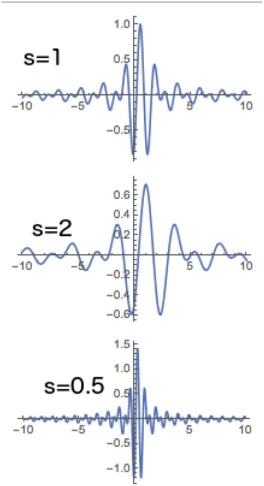


Figure: Figure 22 in Campagne et Mallat 18'

Wavelet transform for scale analysis

- Given a wavelet family, the wavelet transform of x is

$$Wx(s,u) = \langle x, \psi_{s,u} \rangle = \int x(v) \frac{1}{\sqrt{s}} \psi^*(\frac{v-u}{s}) dv$$

- Wavelet transform is a convolution,

$$Wx(s,u) = x \star \tilde{\psi}_s(u)$$

where $\tilde{\psi}_s(u) = \frac{1}{\sqrt{s}} \psi^*(-u/s)$.

- Wx filters structures of x using multiple scales $s > 0$: scale analysis.

Frequency localization of ψ

- Properties of the Fourier transform of wavelet: $\widehat{\psi}(\omega) = \int \psi(u) e^{-i\omega u} du$
- Zero-DC component: $\widehat{\psi}(0) = 0$.
- $\widehat{\psi}_s(\omega) = \sqrt{s}\widehat{\psi}(s\omega)$
  - $s > 1$: lower-frequency (smaller support, larger scale)
  - $s < 1$: higher frequency (larger support, smaller scale)

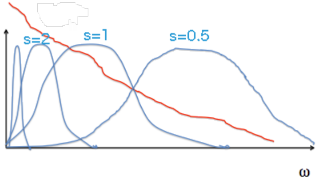


Figure: Support of $\widehat{\psi_s}(\omega)$ from Figure 27 in Campagne et Mallat 18'

From complex wavelets to orthogonal wavelets

- Gabor (1946')/Morlet (1984') wavelet is complex-valued and it is not an orthogonal basis in $L^2(\mathbb{R})$:

$$\psi^{Gabor}(u) \propto e^{-u^2/2} e^{i\nu u}$$

- Y. Meyer (1985') constructed an orthogonal basis by discrete scales $s = 2^j, j \in \mathbb{Z}$.

$$\psi_j(u) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{u}{2^j}\right), \quad Wx(j,u) = x \star \bar{\psi}_j(u)$$

- To construct an orthogonal basis, one key question is: when it is possible to recover x from Wx? (Haar 1910', Mallat 1989')

Back to complex wavelet transform

- Complex wavelet transform provides phase information in Wx which is sometimes crucial.
- In 2d, we can construct complex wavelet transform by dilation and rotation groups. For $0 \leq \ell < L, 0 \leq j < J$,

$$\psi_{j,\ell}(u) = 2^{-2j}\psi(2^{-j} r_{\theta_\ell} u), \quad \theta_\ell = \frac{\pi \ell}{L}$$

- Example: Morlet wavelet (1984') in 2d with $J = 5, L = 8$,

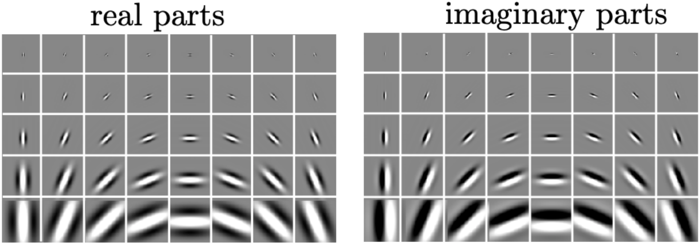
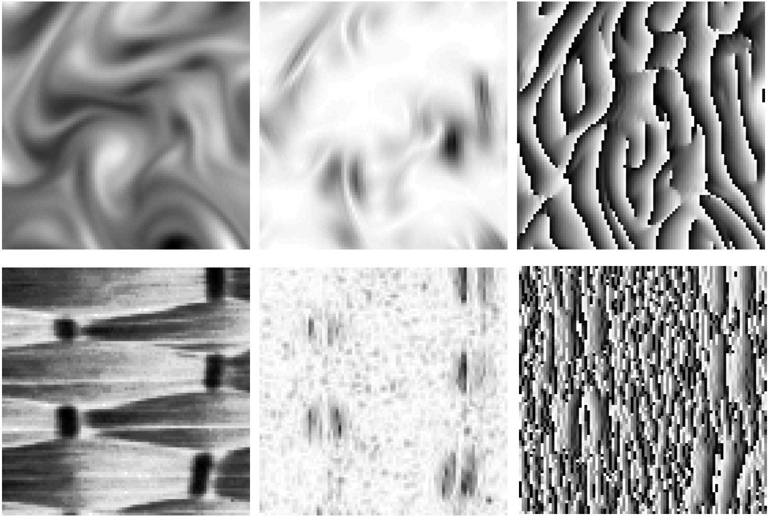


Figure: Real and imaginary part of $\psi_{j,\ell}$. Top to bottom: increasing j. Left to right: increasing ℓ.

How to read phase information?

$x(u)$ | Modulus of $x \star \psi$ | Phase of $x \star \psi$



Properties of Morlet wavelet in 2d

- Basic form: $\psi(u) = \alpha(e^{i\langle u,\xi\rangle} - \beta)e^{-\|u\|^2/2\sigma^2}, u \in \mathbb{R}^2$
  - Choose $\beta \neq 0$ so that $\int \psi(u)du = 0$
  - ψ is nearly analytic: $\hat{\psi}$ is supported on a half plane of $\mathbb{R}^2$.
- Wavelet family: $\psi_{j,\ell}(u) = 2^{-2j}\psi(2^{-j}r_{\theta_\ell}u)$
  - Rotation matrix A_θ in 2d: $r_\theta u = A_\theta u$.
  - Restrict $0 \leq j \leq J-1$ to discretize ψ_j into pixels.
  - Use a low-pass Gaussian filter ϕ_J to capture large scales

$$\phi_J(u) = e^{-|u|^2/2\sigma_J^2}\frac{1}{2\pi\sigma_J^2}$$

See Fig. 46-48 in Campagne et Mallat 20'.

- The Morlet wavelet transform is

$$Wx = \{x \star \phi_J, x \star \psi_{j,\ell}\}_{j<J,\ell<L}.$$

Invariant representation from wavelet coefficients

- Wavelet coefficients: $Wx(\cdot, u)$
- Order p invariant coefficients:

$$\Phi(x) = \frac{1}{N^d} \sum_u |Wx(\cdot, u)|^p$$

  - Case $p = 2$: $\frac{1}{N^d} \sum_u |x \star \psi_{j,\ell}(u)|^2$, $\frac{1}{N^d} \sum_u |x \star \phi_J(u)|^2$.
  - Case $p = 1$: $\frac{1}{N^d} \sum_u |x \star \psi_{j,\ell}(u)|$, $\frac{1}{N^d} \sum_u |x \star \phi_J(u)|$.
- Order $p = 2$ invariant coefficients are related to the *power spectrum* of a stationary process X.
- The order 1 and order 2 coefficients can distinguish if X is from *white noise* or a *Dirac* function.

Relation with power spectrum

- Consider X_N defined on $u \in \{0, \cdots, N-1\}$. From Parseval identity, we have

$$\mathbb{E}(\|X_N \star \psi\|^2) = \sum_{\omega} \mathbb{E}(|\widehat{X}_N(\omega)|^2)|\widehat{\psi}(\omega)|^2/N$$

- Take the limit of $N \to \infty$:

$$\lim_{N\to\infty} \mathbb{E}(\|X_N \star \psi\|^2) = \frac{1}{2\pi}\int_0^{2\pi} R_X(\omega)|\widehat{\psi}(\omega)|^2 d\omega$$

- As the wavelet transform covers the whole frequency range, the order $p = 2$ coefficients capture average information of $R_X(\omega)$ over the support of $\widehat{\phi}_J$ and $\widehat{\psi}_{j,\ell}$.