Scaling Processes

—

A Wavelet Perspective

Patrick Flandrin

CNRS,
Ecole Normale Supérieure de Lyon (France)

Self-similarity and $1/f$ spectra

Ubiquity — Self-similar processes and empirical "$1/f$" spectra are observed in many areas (communications, solid-state physics, biology, turbulence, . . .)

Variations — This however may correspond to different features, depending on the frequency range :

- $A \leq f \leq B$: *scaling* (e.g., turbulence in the inertial range)
- $0 \leq f \leq A$ (large scales) : *long-range dependence*, slowly-decaying correlation
- $B \leq f \leq \infty$ (small scales) : *fractality*, Hölder regularity

Scaling processes

Common property is *scale invariance*

In each case, no characteristic scale exists (in a given range), the important feature being rather the existence of some invariant relation *between* scales.

Non-standard situations in signal processing or time series analysis (nonstationarity, long-range dependence, . . .) $\Rightarrow$ challenging problems in term of analysis, synthesis, and processing (filtering, prediction, . . .).

Specific tools have been developed over the years and, in a recent past, a natural approach has been to consider scaling processes from the perspective of the multiresolution tools introduced around the concept of *wavelet*

Continuous wavelet transform

The *wavelet transform* of a signal $X(t) \in L^2(\mathbb{R})$ is defined in continuous time by

$$T_X(a,t) := \frac{1}{\sqrt{a}} \int X(s)\, \psi\left(\frac{s-t}{a}\right) ds,$$

where :

- t is *time*
- a is *scale*
- $\psi(.)$ is the *analyzing wavelet*, i.e., some zero-mean function, well-localized in both time and frequency

It is a *time-scale* transform, that can be viewed as a "mathematical microscope"

Multiresolution analysis

Slogan — "signal (at any resolution = approximation (at a coarser resolution) + detail"

Theory — More precisely, a *MultiResolution Analysis* of $L^2(\mathbb{R})$ is given by

1. a collection of *nested approximation spaces*

$$\ldots V_1 \subset V_0 \subset V_{-1} \ldots$$

   such that their intersection is zero and their closure dense in $L^2(\mathbb{R})$;

2. a *dyadic scaling relation* between the different approximation spaces :

$$X(t) \in V_j \Leftrightarrow X(2t) \in V_{j-1} \, ;$$

3. a *scaling function* $\varphi(t)$ such its integer translates $\{\varphi(t-n), n \in \mathbb{Z}\}$ form a basis of V_0.

Discrete wavelet transform 1.

It follows that, given a resolution depth J, any signal $X(t) \in V_0$ admits the decomposition :

$$X(t) = \underbrace{\sum_k a_X(J,k)\,\varphi_{J,k}(t)}_{\text{approximation}} +$$

$$\underbrace{\sum_{j=1}^{J}}_{J \text{ scales}} \underbrace{\sum_k d_X(j,k)\,\psi_{j,k}(t)}_{\text{details}},$$

with $\{\varphi_{j,k}(t) := 2^{-j/2}\varphi(2^{-j}t - k), j \text{ and } k \in \mathbb{Z}\}$ and $\{\psi_{j,k}(t) := 2^{-j/2}\psi(2^{-j}t - k), j \text{ and } k \in \mathbb{Z}\}$, the *wavelet* $\psi(.)$ being such that its integer translates are a basis of W_0, the complement of V_0 in V_{-1}.

The a_X's and d_X's stand, respectively, for the *approximation* and *detail* (or wavelet) coefficients of $X(t)$.

Discrete wavelet transform 2.

- The detail coefficients $d_X(j,k)$ measure a *difference in information* between two successive approximations, and are obtained as

$$d_X(j,k) := 2^{-j/2} \int X(t)\, \psi\left(2^{-j}t - k\right)\, dt\,.$$

- From a practical point of view, they can be computed *recursively* with efficient *pyramidal algorithms* (faster than FFT).

- An important property of a wavelet is its *number of vanishing moments*, i.e., the number $N \geq 1$ such that

$$\int t^k\, \psi(t)\, dt \equiv 0, \quad k = 0, 1, \ldots N-1\,.$$

Self-similarity 1.

Definition — A process $X = \{X(t), t \in \mathbb{R}\}$ is said to be *self-similar* with self-similarity parameter $H > 0$ (or, "H-ss") if and only if the processes $X_1 := X$ and $X_c := \left\{c^{-H} X(ct), t \in \mathbb{R}\right\}$ have the same finite dimensional distributions for any $c > 0$ (statistical scale-invariance).

(Remark : self-similarity $\Rightarrow$ nonstationarity.)

Wavelet characterization — The wavelet coefficients of an H-ss process X exactly reproduce its self-similarity through :

$$(d_X(j,0), \ldots, d_X(j, N_j - 1)) \stackrel{d}{=}$$

$$2^{j(H+1/2)}(d_X(0,0), \ldots, d_X(0, N_j - 1)) \,.$$

Self-similarity 2.

Two consequences of the wavelet characterization of self-similarity :

1. For processes whose wavelet coefficients have *finite* second-order statistics, one has :

   $$\mathbb{E} d_X^2(j,k) = 2^{j(2H+1)} \mathbb{E} d_X^2(0,k),$$

   and, thus,

   $$\log_2 \mathbb{E} d_X^2(j,k) = j(2H+1) + \log_2 \mathbb{E} d_X^2(0,k).$$

2. For processes whose wavelet coefficients may have *infinite* second-order statistics, but for which the quantity $\mathbb{E} \log_2 |d_X(j,k)|$ exists, one has :

   $$\mathbb{E} \log_2 |d_X(j,k)| =$$

   $$j(H+1/2) + \mathbb{E} \log_2 |d_X(0,k)|.$$

Stationary increments 1.

Definition — A process $X = \{X(t), t \in \mathbb{R}\}$ is said to have *stationary increments* (or, to be "si") if and only if, for any $h \in \mathbb{R}$, the finite-dimensional distributions of the increment processes

$$X^{(h)} := \left\{X^{(h)}(t) := X(t+h) - X(t), t \in \mathbb{R}\right\}$$

do not depend on t.

Wavelet characterization — In the case of the wavelet transform, this results in a *stationarization property*, according to which the details $\{d_X(j,k), k \in \mathbb{Z}\}$ of a si process form, at each octave j, a stationary sequence.

Stationary increments 2.

Some remarks :

1. Stationarization can be extended to processes which have increments of order p, under the condition that $N \geq p$.

2. Stationarization applies as well to the continuous wavelet transform.

3. The increments of a process X can be read as a specific example of wavelet coefficients, since we have

   $$X^{(ah_0)}(t) := X(t+ah_0)-X(t) \equiv \frac{1}{\sqrt{a}} T_X(a,t),$$

   with $\psi(t) = \delta(t+h_0)-\delta(t)$ (the “poorman’s wavelet”).

Finite variance and Gaussian processes

Let X denote a zero-mean H-sssi process with finite variance. Assuming that $X(0) \equiv 0$, one has necessarily :

$$\mathbb{E}X(t)X(s) = \frac{\sigma^2}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right),$$

with the two consequences that wavelet coefficients

1. may be asymptotically decorrelated when $|2^{-j}k - 2^{-j'}k'| \to \infty$:

   $\mathbb{E}d_X(j,k)d_X(j',k') \sim |2^{-j}k - 2^{-j'}k'|^{2(H-N)}$;

2. have a variance reproducing the scaling law :

   $$\text{Var } d_X(j,k) = 2^{j(2H+1)}\sigma^2 C(\psi, H),$$

   with

   $$C(\psi, H) := \int |u|^{2H}\left(\int \psi(v)\psi(v+u)dv\right)du\,.$$

(Remark : Gaussianity $\Rightarrow$ FBM.)

Long-range dependence

Definition 1 — A second-order stationary process X is said to be *long-range dependent* (LRD) if its stationary covariance function c_X satisfies :

$$c_X(\tau) \sim c_r\, \tau^{-\beta},\, \tau \to +\infty,$$

with $0 < \beta < 1$.

Definition 2 — An equivalent definition amounts to saying that the spectrum Γ_X of a LRD process satisfies :

$$\Gamma_X(f) \sim c_f\ |f|^{-\gamma},\, f \to 0\,,$$

with $0 < \gamma < 1$.

Interpretation — Both definitions imply that

$$\int c_X(\tau)\, d\tau = \infty\,,$$

i.e., that the asymptotic decay of the stationary covariance function is so slow that it is not summable ("long memory").

LRD processes and wavelets

Scaling — The variance of the wavelet coefficients of a LRD process reproduces the underlying power-law :

$$\text{Var } d_X(j,k) \sim 2^{j\gamma} c_f C(\psi,\gamma)\,, j \to \infty,$$

with

$$C(\psi,\gamma) := \int |f|^{-\gamma} |\Psi(f)|^2 \, df.$$

Weak dependence — Moreover, the covariance of wavelet coefficients has as asymptotic behaviour :

$$\mathbb{E} d_X(j,k) d_X(j',k') \sim |2^{-j}k - 2^{-j'}k'|^{\gamma-1-2N}\,,$$

when $|2^{-j}k - 2^{-j'}k'| \to \infty$: long-range correlation within X can be turned to short-range correlation within its wavelet coefficients, under the condition that $N > \gamma/2$.

Wavelets for scaling exponent estimation

Wavelet analysis offers a unifying framework for the scaling processes discussed so far, since wavelet coefficients

1. form *stationary sequences*, at any scale ;

2. *reproduce the scale invariance*, through either

   $$\log_2 \text{Var } d_X(j,k) = j\gamma + C$$

   or

   $$\mathbb{E} \log_2 |d_X(j,k)| = j(H + 1/2) + C \,;$$

3. are *weakly correlated* if the number of vanishing moments is high enough.

Efficient estimation of scaling exponent is therefore possible on the basis of a simple linear regression in a log-log plot (*logscale diagram*).

Estimation for finite variance and Gaussian processes 1. Definition

Variance — At each scale j, Var $d_X(j,k)$ can be efficiently estimated by

$$\mu_j = \frac{1}{n_j} \sum_{k=1}^{n_j} d_X^2(j,k)\,,$$

and there exists a (known) correction number g_j such that $y_j := \log_2 \mu_j - g_j$ verifies

$$\mathbb{E} y_j = \gamma j + C.$$

Exponent — Estimation is achieved through the weighted linear regression

$$\widehat{\gamma} = \sum_j \left(\frac{1}{a_j} \frac{S_0 j - S_1}{S_0 S_2 - S_1^2} \right) y_j\,,$$

with $S_m := \sum_j j^m / \text{Var } y_j$ for $m = 0, 1, 2$.

Estimation for finite variance and Gaussian processes 2. Properties

1. By construction, the estimator is unbiased :

$$\mathbb{E}\widehat{\gamma} = \gamma.$$

2. Assuming no correlation between wavelet coefficients, Var $\widehat{\gamma}$ decreases as $1/n$ for data of size n (in the limit of large n_j at each scale j under consideration).

3. The Cramér-Rao lower bound is attained in this case, and the estimate $\widehat{\gamma}$ is (asymptotically) normally distributed.

Wavelet-based estimations — Additional benefits

1. *Robustness to non-Gaussianity.*

2. *Insensitivity to polynomial trends.* Imposing $N > 1$ vanishing moments results in a wavelet that is blind to polynomials up to orders $p \leq N - 1$.

3. *Computational efficiency.* Because of their multiresolution structure and their pyramidal implementation, wavelet-based methods are associated with fast algorithms over-performing FFT-based algorithms (complexity O(n) *vs.* O($n \log n$), for n data points).

Conclusion

- Wavelet analysis offers a *unified framework* for the characterization of scaling processes.

- It allows for a unique treatment of a large variety of processes, be they self-similar, fractal, long-range dependent, Gaussian or not, . . .

- Further extensions can be given to the results presented so far (*point processes* with fractal-type characteristics, *multifractional* or *multifractal* processes, . . .).

Wedding wavelets with scaling processes is in some sense “natural,” in terms of a structural adequacy between the *mathematical* framework they offer (multiresolution) and the *physical* nature of the processes under study (scaling).