Long-Range Dependence

Lance Galletti

MA 780
Boston University
Department of Mathematics & Statistics

April, 2016

Table of Content

Time Series Overview

A stochastic process $\{X(t)\}_{t \in \mathrm{T}}$ is **stationary** if

$$\{X(t)\}_{t \in \mathrm{T}} \stackrel{d}{=} \{X(t+h)\}_{t \in \mathrm{T}}$$

But this assumption is often too strong in practice. A way to relax this assumption is to limit the constraints to mean and covariance. We define a stochastic process $\{X(t)\}_{t \in \mathrm{T}}$ to be **weak stationary** (or second-order stationary) if for any $s < t \in \mathrm{T}$ the two following hold:

$$\mathbb{E}X(t) = \mathbb{E}X(0)$$

and

$$\begin{aligned} \mathrm{Cov}(X(t), X(s)) &= \mathrm{Cov}(X(t-s), X(0)) \\ &= \gamma_X(t-s) \end{aligned}$$

Time Series Overview

Let $h = t - s$, we call $\gamma_X(h)$ the **autocovariance function** and $\rho(h) = \gamma_X(h)/\gamma_X(0)$ the **autocorrelation function** of the time series $\{X(t)\}_{t \in \mathbb{Z}}$. This is enough to fully characterize the dependence structure and therefore the process itself. Sample counterparts of the ACF and ACVF are good measures of the dependence structure of the time series. They can also allow us to uncover simpler processes that together model our process.

Time Series Overview

A stochastic process $\{X(n)\}_{n\in\mathbb{Z}}$ is called a **Moving Average process of order q** if

$$X(n) = X_n = Z_n + \sum_{k=1}^{q} \theta_k Z_{n-k} \tag{1}$$

where $Z_n \sim WN(0, \sigma_Z^2)$.
Using (1), let X_n be MA(1):

$$X_n = Z_n + \theta Z_{n-1}$$

Hence we can find the ACVF and ACF of X_n:

$$\gamma_X(h) = \mathbb{E}X_h X_0 = \begin{cases} \sigma_Z^2(1+\theta^2),\ h=0 \\ \sigma_Z^2\theta,\ h=1 \\ 0,\ h\geq 2 \end{cases}$$

Time Series Overview

A weak stationary time series $\{X(n)\}_{n\in\mathbb{Z}}$ is called **autoregressive of order p** if

$$X_n = \sum_{k=1}^{p} \varphi_k X_{n-k} + Z_n \tag{2}$$

with $Z_n \sim WN(0, \sigma_Z^2)$.
Let $p = 1$ and notice that because (2) is equivalent to

$$X_n = \varphi(\varphi X_{n-2} + Z_{n-1}) + Z_n$$

We have that

$$X_n = \sum_{k=0}^{\infty} \varphi^k Z_{n-k} \tag{3}$$

Since WLOG $X(0) = 0$. Which by geometric series properties only converges in L^2 for $|\varphi| < 1$

Time Series Overview

When this is the case, it is easy to show that $\rho_X(h) = \varphi^{|h|}$. That is, (3) is well-defined when

$$\sum_{h=0}^{\infty} |\varphi^h| = \sum_{h=0}^{\infty} |\rho_X(h)| < \infty \tag{4}$$

Let $d \in \mathbb{Z}$ be positive. In an AR process of order d, if the first $d-1$ φ_j's are zero and $\varphi_d = 1$ then X_n becomes

$$X_n - X_{n-d} = Z_n \tag{5}$$

Such a process is called **Integrated of order d**.

Combining (1) and (2) we get the ARMA(p,q):

$$X_n - \sum_{k=1}^{p} \varphi_k X_{n-k} = Z_n + \sum_{k=1}^{q} \theta_k Z_{n-k} \tag{6}$$

which exists if the characteristic polynomial $1 - \varphi_1 z - ...\varphi_p z^p = 0$ does not have roots on the unit circle. This is consistent with (3).
At this point it becomes more convenient to express processes in terms of an operator B, called the **backshift operator**, defined as:

$$B^k X_n = X_{n-k},\ B^0 = I,\ k \in \mathbb{Z}$$

We write:

$$\varphi(B)X_n = \theta(B)Z_n \tag{7}$$

where $\varphi(z) = 1 - \varphi_1 z - ... - \varphi_p z^p$
and $\theta(z) = 1 + \theta_1 z + ... + \theta_p z^p$

Time Series Overview

Adding an integrated component to the ARMA(p,q) process, we get the following process we call ARIMA(p,d,q)

$$\varphi(B)(1-B)^d X_n = \theta(B) Z_n \tag{8}$$

Long-Range Dependence (ACF Approach)

Attempting to extend the definition of ARIMA where d is only integer valued, we define a FARIMA(0, d, 0) process X_n as

$$X_n = (I - B)^{-d} Z_n$$

For $d \in I$ where I is some interval to be determined later. Notice that we can also write this as

$$X_n = \sum_{j=0}^{\infty} b_j^{(-d)} B^j Z_n = \sum_{j=0}^{\infty} b_j^{(-d)} Z_{n-j} \tag{9}$$

where

$$b_j^{(-d)} = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} \tag{10}$$

are the coefficients of the Taylor expansion of the above $(I - B)^{-d}$

Long-Range Dependence (ACF Approach)

As we have seen before in I.2, (9) is well defined if

$$\sum_{j=0}^{\infty} b_j^2 < \infty \tag{11}$$

Using Stirling's approximation formula, we find that (9) is well defined for $d < \frac{1}{2}$. So $I = (-\infty, 1/2)$
The F in FARIMA stands for Fractional.

Long-Range Dependence (ACF Approach)

If $d < 0$ but not a negative integer,

$$\sum_{h=-\infty}^{\infty} |\gamma_X(h)| \leq \sigma^2 \left(\sum_j |b_j| \right)^2 < \infty$$

If $d = -1$ then $X_n = Z_n - Z_{n-1} = \Delta Z_n$ and hence for $d = -1, -2, \ldots,$ $X_n = \Delta^{-d} Z_n$. Hence in these cases, $\gamma_X(h) = 0$ for $|h| > |d|$ by Remark 1. Hence, for $d < 0$ the autocovariances of the times series are absolutely summable.

If $d = 0$, then $X_n = Z_n$ is just white noise, so the autocovariances are absolutely summable.

If $0 < d < 1/2$ then the autocovariances are not absolutely summable.

Long-Range Dependence (ACF Approach)

A time series X_n is **short range dependent** (SRD) if

$$\sum_{h=-\infty}^{\infty} |\gamma_X(h)| < \infty$$

That is, the so called **long-run variance** of the times series is constant.
This corresponds to $d \leq 0$ above.
A special case of SRD time series is when the long-run variance is 0. Such a process is called **antipersistent**.

Long-Range Dependence (ACF Approach)

Condition I. The times series X_n has a linear representation

$$X_n = \mu_X + \sum_{k=0}^{\infty} \psi_k Z_{n-k} = \mu_X + \sum_{k=-\infty}^{\infty} \psi_{n-j} Z_j$$

with $\{\psi_k\}_{k \geq 0}$ satisfying

$$\psi_k = L_1(k) k^{d-1}$$

where $d \in (0, 1/2)$ and $L_1(k)$ is positive on $[c, \infty]$ for some $c \geq 0$ satisfies, for any $a > 0$

$$\lim_{k \to \infty} \frac{L(ak)}{L(k)} = 1$$

we call such a function slowly varying at infinity.

Condition II. The autocovariance function of X_n satisfies

$$\gamma_X(k) = L_2(k) k^{2d-1}, \ k = 0, 1, 2, ...$$

with d as before and L_2 slowly varying at infinity.

Long-Range Dependence (ACF Approach)

Condition III. The autocovariances of the time series of X_n are not absolutely summable, that is,

$$\sum_{h=-\infty}^{\infty} |\gamma_X(h)| = \infty$$

Condition V. The time series X_n satisfies

$$\text{Var}(X_1 + \ldots + X_N) = L_5(N)N^{2d+1}$$

Long-Range Dependence (ACF Approach)

Definition A weak stationary time series X_n is called **long-range dependent** (LRD) if one of the non-equivalent conditions above are satisfied. The parameter $d \in (0, 1/2)$ is called the **long-range dependence parameter**.

Note: It can be shown that

$$\text{I} \implies \text{II} \implies \text{III, V}$$

Long-Range Dependence (ACF Approach)

Example 1: Let $X = X_n$ be a Markov Chain with finite state space. Let P be the transition matrix. Assume X is is irreducible (all states are connected) and aperiodic. Then, let $\pi_i > 0,\ i = 1, ..., m$ be the invariant probabilities.

A Markov Chain X with initial distribution $\mathbb{P}(X_0 = i) = \pi_i$ is strictly stationary

It can be shown that if P has distinct eigenvalues, then

$$|\gamma_{g(X)}(k)| \leq Cs^k, \text{ some } 0 < s < 1$$

for some function g and some s that depends on the eigenvalues of P. That is, $g(X_n)$ has at least exponentially decreasing autovariances and hence is SRD.

Long-Range Dependence (ACF Approach)

Example 2: The series X_n is called FARIMA(p, d, q) if it satisfies

$$X_n = \varphi^{-1}(B)\theta(B)(I - B)^{-d} Z_n$$

with $d < 1/2$

Remark The operator $(1 - B)^{-d}$ is responsible for LRD features of X_n while $\varphi^{-1}(B)\theta(B)$ is responsible for SRD behavior.
If the $\varphi(z)$ has roots outside of the unit circle, then the FARIMA(p, d, q) series with $0 < d < 1/2$ is LRD in the sense of conditions I, II, III, and V.

Long-Range Dependence (Self-Similar Approach)

Definition A process $\{X(t)\}_{t\in\mathbb{R}}$ is self-similar if there is an $H > 0$, called the self-similarity parameter such that for all $c > 0$,

$$\{X(ct)\}_{t\in\mathbb{R}} \stackrel{d}{=} \{c^H X(t)\}_{t\in\mathbb{R}} \tag{12}$$

Theorem If $X = X(t)$ is H-SS, then $Y(t) = e^{-tH} X(e^t)$ is stationary. Conversely, If $Y = Y(t)$ is stationary, then $X(t) = t^H X(\log(t)))$ is H-SS.

Definition A process $\{X(t)\}_{t\in\mathbb{R}}$ has stationary increments if, for any $h \in \mathbb{R}$

$$\{X(t+h) - X(t)\}_{t\in\mathbb{R}} \stackrel{d}{=} \{X(t) - X(0)\}_{t\in\mathbb{R}}$$

Long-Range Dependence (Self-Similar Approach)

Proposition Let $\{X(t)\}_{t\in\mathbb{R}}$ be an H-SSSI process. Then, if $\mathbb{E}|X(1)|^2 < \infty$, then for $s, t \in \mathbb{R}$

$$\mathbb{E}X(t)X(s) = \frac{\mathbb{E}X(1)^2}{2}\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\}$$

Definition A Gaussian H-SSSI process $\{B_H(t)\}_{t\in\mathbb{R}}$ with $0 < H \leq 1$ is called Fractional Brownian Motion.

Corollary If $\{X(t)\}_{t\in\mathbb{R}}$ satisfies the following conditions:
(i) It is a Gaussian process with zero mean and X(0) = 0
(ii) $\mathbb{E}(X(t))^2 = \sigma^2|t|^{2H}$ for all $t \in \mathbb{R}$, $\sigma > 0$, and $0 < H \leq 1$
(iii) $\mathbb{E}(X(t) - X(s))^2 = \mathbb{E}(X(t-s))^2$ for all $s, t \in \mathbb{R}$
Then, $\{X(t)\}_{t\in\mathbb{R}}$ is fractional Brownian Motion.

Long-Range Dependence (Self-Similar Approach)

Proposition For $H \in (0,1)$, standard FBM admits the following integral representation

$$\{B_H(t)\}_{t\in\mathbb{R}} =$$
$$\{\frac{1}{c_1(H)}\int_{\mathbb{R}}\left((t-u)_+^{H-1/2}-(-u)_-^{H-1/2}\right)B(du)\}_{t\in\mathbb{R}} \tag{13}$$

Proof Must show that RHS is well defined (i.e. the integrand is square integrable), and show that the three conditions of the Corollary are satisfied.

Values of H If $1/2 < H < 1$, then we can express (13) as

$$\{\frac{H-1/2}{c_1(H)}\int_{\mathbb{R}}\left(\int_0^t (s-u)_+^{H-3/2}ds\right)B(du)\}_{t\in\mathbb{R}}$$

If $H = 1/2$ then FBM reduces to Brownian Motion.

Long-Range Dependence (Self-Similar Approach)

Definition Bifractional Brownian Motion is a natural extension to FBM and is defined as a zero mean Gaussian process with covariance function

$$\mathbb{E}B_{H,K}(t)B_{H,K}(s) = \frac{1}{2^K}\left((t^{2H}+s^{2H})^K - |t-s|^{2HK}\right)$$

where $H, K \in (0,1)$

Note When $K = 1$, or $H = 1/2$, biFBM reduces to FBM.

Proposition Let $B_{H,K}$ be biFBM, then:
(i) It is (HK)-self-similar
(ii) It has stationary increments iff $K = 1$
(iii) For all $s, t \geq 0$,

$$2^{-K}|t-s|^{2HK} \leq \mathbb{E}(B_{H,K}(t) - B_{H,K}(s))^2 \leq 2^{1-K}|t-s|^{2HK}$$

Long-Range Dependence (Self-Similar Approach)

Let $Y = Y(t)$ be an H-SSSI process with $0 < H < 1$ and consider the stationary process

$$X_n = Y(n) - Y(n-1),\ n \in \mathbb{Z}$$

Notice that X_n has zero mean, $\mathbb{E}X_n^2 = \mathbb{E}Y(1)^2$, and

$$\gamma_X(k) = \frac{\mathbb{E}X(1)^2}{2}\{|k+1|^{2H} + 2|k|^{2H} - |k-1|^{2H}\}$$

$$\sim \mathbb{E}Y(1)^2 H(2H-1)k^{2H-2}$$

Long-Range Dependence (Self-Similar Approach)

By the above observation, if $1/2 < H < 1$ then the series X is LRD in the sense of condition II (and hence III) with

$$d = H - \frac{1}{2} \in \left(0, \frac{1}{2}\right) \tag{14}$$

So d and H are closely related.
Note When Y is FBM, X is called fractional Gaussian Noise.

Bibliography

Long-Range Dependence and Self-Similarity (Cambridge Series in Statistical and Probabilistic Mathematics) **1st Edition, 2017**, Murad Taqqu, Vladas Pipiras