Asymptotic spectral theory for nonlinear random fields

Karima Kimouche

Département de Mathematiques
Université 20 août 1955-Skikda, Algeria.
E-mails: k.kimouche@gmail.com

Résumé. Dans ce travail, nous considérons le problème asymptotique dans l'analyse spectrale des champs aléatoires stationnaires. Nous proposons des conditions qui sont facilement vérifiées pour différentes classes de champs aléatoires non linéaires pour obtenir la consistance et la normalité asymptotique de l'estimateur de la densité spectrale. La distribution asymptotique de la déviation maximale de l'estimateur de densité spectrale est aussi dérivée.

Mots-clés. Processus non linéaire spatial, densité spectrale, mesure de dépendance physique.

Abstract. In this paper, we consider the asymptotic problem in spectral analysis of stationary random fields. We impose conditions which are easily verifiable for a variety of nonlinear random fields to obtain the consistency and the asymptotic normality of spectral density estimates. Asymptotic distribution of maximum deviations of the spectral density estimates is also derived.

Keywords. Spatial nonlinear process, spectral density, physical dependence measure.

1 Introduction

Spatial data arise in various area of research, including astronomy, epidemiology, image analysis. Important development in area of spatial statistics are found in Cressie (1991), Guyon (1995) and the references therein. Spectral analysis of stationary processes is a powerful tool for analyzing spatial data sets on a grid see Guyon (1995), Rosenblatt (1985) among other. However, the asymptotic results are certainly needed in the related statistical inference in the frequency domain, such as hypothesis testing and the construction of confidence intervals. So, to analyze the properties of any spectral estimator it is important to quantify the strength of dependence of the random fields. This can be done using cumulant summability conditions, moment and mixing conditions, for example, Rosenblatt(1985) considered strong mixing random field and assumed the summability condition of cumulants up to the eighth order, etc. In the earlier literature, the assumption of linearity of the Random fields is prevalent (see Guyon (1995), Rosenblatt (1985)).

There has been a recent surge of interest in nonlinear random fields see Bibi and Kimouche (2014) and the references therein. It seems that a systematic asymptotic spectral theory for such processes is lacking. The goal of this paper is to establish an asymptotic spectral theory for spatial processes under very mild and natural conditions, thus substantially extending the applicability of spectral analysis to nonlinear and/or non–strong mixing spatial processes. Here we adopt the setting of physical dependence measure (p-stability) introduced in Wu (2005) in dimension 1 and extended by El Machkouri et al. (2013) to general dimension.

The paper is organized as follows. In section 2, we extend the concept of physical dependence measure used in the time series literature to spatial processes. In section 3 we present the estimate of spectral density. The main results are given in Section 4.

The following notations are used throughout this paper. Let $\mathbf{l} = (l_1, ..., l_d)$ and $\mathbf{k} = (k_1, ..., k_d)$ two vectors of non negative integers belonging to $\mathbb{Z}^d$, we have $\mathbf{l}.\mathbf{k} = l_1k_1 + ... + l_dk_d, \mathbf{l} \odot \mathbf{k} = (l_1k_1, ..., l_dk_d)\,, \frac{\mathbf{1}}{\mathbf{k}} = (\frac{l_1}{k_1}, ..., \frac{l_d}{k_d})$ if $k_1, ..., k_d \neq 0$, $\widehat{\mathbf{k}} = \prod_{i=1}^{d} k_i$ and $|\underline{\mathbf{l}}| = (|l_1|, ..., |l_d|)$. Further, for a column vector $\mathbf{a} = (a_1, ..., a_d) \in \mathbb{R}^d$ let $|\mathbf{a}|_p = \left(\sum_{j=1}^{d} a_j^p\right)^{1/p}$ (i.e.$|\mathbf{a}|_2 = |\mathbf{a}|$). Also, let $\mathbf{0} = (0, ..., 0)$ and $\mathbf{1} = (1, ..., 1)$ denote the $d-$dimensional vectors of zeros and ones. Define the partial order $\mathbf{l} \preceq \mathbf{k}$ if $l_i \leq k_i$ (i.e. for $\mathbf{l} \prec \mathbf{k}$ if $l_i < k_i$) for each i, $\Gamma[\mathbf{n}] = \{\mathbf{x} \in \mathbb{Z}^d, \mathbf{0} \preceq \mathbf{x} \preceq \mathbf{n}\}$ and $\mathcal{R}(\mathbf{l}, \mathbf{k}) = \{\mathbf{x} \in \mathbb{Z}^d, \mathbf{l} \preceq \mathbf{x} \preceq \mathbf{k}\}$.

2 The Dependence measure

Let $\mathbb{Z}^d$ denote the integer lattice in $d-$dimensional Euclidean space, where $d \geq 1$ and let $\mathcal{I}_{\mathbf{n}}$ be a subset of $\mathbb{Z}^d$ which is the observation region of the data, i.e. the location at which the data is collected. We assume that the random field $(X(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ has the form

$$X(\mathbf{t}) = G(\varepsilon(\mathbf{t} - \mathbf{h}); \mathbf{h} \in \mathbb{Z}^d), \mathbf{t} \in \mathbb{Z}^d. \tag{1}$$

where $(\varepsilon(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ are independent and identically distributed random fields and G is a measurable function. In the one-dimensional case ($d = 1$), (1) includes linear as well as many widely used nonlinear time series models as special cases in Shao and Wu (2007), Liu and Wu (2010).

The physical dependence measure should be seen as a measure of the dependence of the function G defined in (1). It turns out that, with the dependence measure, the consistency and asymptotic normality can be established in a very elegant and natural way.

Let $\tau : \mathbb{Z} \to \mathbb{Z}^d$ be a bijection and for any $k \in \mathbb{Z}$ and $l \in \mathbb{Z}, \Im_k = \sigma(\varepsilon(\tau(l)); l \leq k)$. Let $(\varepsilon'(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ be an i.i.d copy of $(\varepsilon(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$. For a set $T \subset \mathbb{Z}$, let

$$\varepsilon_T(\mathbf{t}) = \left\{ \begin{array}{c} \varepsilon'(\tau(l)) \text{ if } \mathbf{t} = \tau(l), l \in T \\ \varepsilon(\mathbf{t}) \text{ if } \mathbf{t} \neq \tau(l), l \in T \end{array} \right. .$$

Definition 2.1 *For $p \in]0, +\infty]$ and $T \subset \mathbb{Z}$ we define the physical dependence measure by*

$$\delta_p(T, \mathbf{t}) = \|G(\varepsilon(\mathbf{t}-\mathbf{h})) - G(\varepsilon_T(\mathbf{t}-\mathbf{h}))\|_p \,; \mathbf{h} \in \mathbb{Z}^d \tag{2}$$

Remark 2.1 - *If $T = \{0\}$, then (2) is defined as*

$$\delta_p(\mathbf{t}) = \|G(\varepsilon(\mathbf{t}-\mathbf{h})) - G(\varepsilon^*(\mathbf{t}-\mathbf{h}))\|_p \,; \mathbf{h} \in \mathbb{Z}^d \tag{3}$$

where

$$\varepsilon^*(\mathbf{t}) = \begin{cases} \varepsilon(\mathbf{t}) & \text{if } \mathbf{t} \neq \mathbf{0}, \\ \varepsilon'(\mathbf{0}) & \text{if } \mathbf{t} = \mathbf{0}. \end{cases}$$

- *If $\tau(\min\{l : l \in T\}) \succ \mathbf{t}$ then $\delta_p(T, \mathbf{t}) = 0$.*

Definition 2.2 *We say that the random field $(X(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ defined in (1) is p-stable if*

$$\Delta_p := \sum_{\mathbf{t}\in\mathbb{Z}^d} \delta_p(\mathbf{t}) < \infty$$

As an illustration, we give some examples of p-stable random fields.

Example 2.1 *(Linear random fields) Let $(\varepsilon(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ be i.i.d random field with $\varepsilon(\mathbf{0}) \in \mathcal{L}^p$ for some $p \geq 1$. When $\sum_{\mathbf{i}\in\mathbb{Z}^d} |a_\mathbf{i}| < \infty$, one can define in L^p the centered linear random field $X(\mathbf{t}) = \sum_{\mathbf{i}\succeq\mathbf{0}} a_\mathbf{i} \varepsilon(\mathbf{t}-\mathbf{i}), \mathbf{t} \in \mathbb{Z}^d$. Moreover, for $p > 2$, for any $\mathbf{t} \in \mathbb{Z}^\mathbf{d}$ $\delta_p(\mathbf{t}) = |a_\mathbf{t}| \|\varepsilon(\mathbf{0}) - \varepsilon'(\mathbf{0})\|_p$. So, $X(\mathbf{t})$ is p-stable if and only if*

$$\sum_{\mathbf{i}\in\mathbb{Z}^d} |a_\mathbf{i}| < \infty.$$

Example 2.2 *Let $(\varepsilon(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ be i.i.d random fields and consider the recursion*

$$X(\mathbf{t}) = G(X(\mathbf{t}-\mathbf{1}), \varepsilon(\mathbf{t})) \tag{4}$$

where G is a measurable function. The framework (4) is quite general, and it includes many popular nonlinear random fields such as autoregressive–autoregressive conditionally heteroskedastic random fields, amplitude-dependent exponential autoregressive random fields and signed volatility models. Assume that there exist $p > 0$ and $x_0 \in \mathbb{R}$ such that $G(x_0, \varepsilon(\mathbf{0})) \in \mathcal{L}^p$ and $E\left(L^p_{\varepsilon(\mathbf{0})}\right) < 1$, where $L_{\varepsilon(\mathbf{0})} = \sup_{x\neq x'} \frac{|G(x,\varepsilon(\mathbf{0}))-G(x',\varepsilon(\mathbf{0}))|}{|x-x'|}$. Then $X(\mathbf{t})$ has a unique stationary solution of the form (1), and $\delta_p(\mathbf{n}) = O(\rho^{\widehat{\mathbf{n}}})$ for some $\rho \in (0,1)$.

3 The estimate

Let $(X(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ be a stationary random field with mean 0 and finite covariances $C(\mathbf{h}) := Cov\left\{X(\mathbf{t})X(\mathbf{t}+\mathbf{h})\right\}$. Assume that the covariances are absolutely summable, then the spectral density function

$$f(\lambda) = \frac{1}{(2\pi)^d}\sum_{\mathbf{h}\in\mathbb{Z}^d} C(\mathbf{h})e^{-i\mathbf{h}.\lambda}, \lambda \in \pi \tag{5}$$

exists and is continuous and finite (i.e., $\pi = [-\pi,\pi[\times ... \times [-\pi,\pi[$, $d-$times).

Given the observations $\{X(\mathbf{1}),...,X(\mathbf{n})\}$, the observation region $\mathcal{I}_\mathbf{n}$ mentioned in the previous section will be the rectangular region of $\mathbb{Z}^d$ defined by $\mathcal{I}_\mathbf{n} := \left\{\mathbf{t}\in\mathbb{Z}^d : 1 \leq t_i \leq n_i, i = 1,...,d\right\}$, the number of sites in $\mathcal{I}_\mathbf{n}$ is denoted as $\widehat{\mathbf{n}}= \prod_{i=1}^d n_i$. Consider the lag-window estimate

$$f_\mathbf{n}(\lambda) = \frac{1}{(2\pi)^d}\sum_{\mathbf{h}\in\mathcal{R}(\mathbf{1}-\mathbf{n},\mathbf{n}-\mathbf{1})} \widehat{C}(\mathbf{h})K(\frac{\mathbf{h}}{B_\mathbf{n}})e^{-i\mathbf{h}.\lambda}, \tag{6}$$

where

$$\widehat{C}(\mathbf{h}) = \frac{1}{\widehat{\mathbf{n}}}\sum_{\mathbf{t}\in\mathcal{R}(|\underline{\mathbf{h}}|+\mathbf{1},\mathbf{n})} X(\mathbf{t})\,X(\mathbf{t}-|\underline{\mathbf{h}}|)\,, |\underline{\mathbf{h}}| \prec \mathbf{n}-\mathbf{1}, \tag{7}$$

be the estimated covariances, $b_\mathbf{n} = B_\mathbf{n}^{-1}$ is the bandwidth satisfying $\mathbf{n}\odot b_\mathbf{n} \longrightarrow \infty$, $b_\mathbf{n} \longrightarrow \mathbf{0}$ as $\mathbf{n} \longrightarrow \infty$ and the kernel function K satisfied the following assumption

Assumption 1

K is an even, bounded, continuous and absolutely integrable function with $\lim_{\mathbf{u}\to\mathbf{0}} K(\mathbf{u}) = K(\mathbf{0}) = 1$, $\int_{\mathbb{R}^d} K^2(\mathbf{u})d\mathbf{u} =: \kappa < \infty$.

For asymptotic results in this paper, we consider the following assumption on the Kernel K when all those assumptions are mild, and they are satisfied for Parzen, triangle, Tukey, and many other commonly used windows.

Assumption 2

a) $\lim_{\mathbf{w}\to 0}\widehat{\mathbf{w}}\sum_{\mathbf{h}\in\mathbb{Z}^d} K^2(\mathbf{h}\odot\mathbf{w}) = \int_{\mathbb{R}^d} K^2(\mathbf{u})d\mathbf{u} =: \kappa < \infty$ *and its Fourier transform* $\widehat{K}(\mathbf{x}) = \int_{\mathbb{R}^d} K(\mathbf{u})e^{i\mathbf{x}.\mathbf{u}}d\mathbf{u}$ *satisfies* $\int_{\mathbb{R}^d}\left|\widehat{K}(\mathbf{x})\right| d\mathbf{x} < \infty$.

b) $\sup_{\mathbf{0}\prec\mathbf{w}\preceq\mathbf{1}} \widehat{\mathbf{w}}\sum_{\mathbf{h}\succeq\mathbf{c}/\mathbf{w}} K^2(\mathbf{h}\odot\mathbf{w}) \longrightarrow 0$ *as* $\mathbf{c} \longrightarrow \infty$.

c) *K is bounded function with bounded support* $I = [-1,1]^d, \kappa := \int_I K^2(\mathbf{u})d\mathbf{u} < \infty$ *and* $\sup_{|\underline{\mathbf{s}-\mathbf{h}}|\preceq\mathbf{1}}\sum_{\mathbf{h}\in\mathbb{Z}^d} |K(\mathbf{h}\odot\mathbf{w}) - K(\mathbf{s}\odot\mathbf{w})| = O(1)$ *as* $\mathbf{w} \longrightarrow \infty$.

4 Main results

In this section our consistency and asymptotic normality result require the short-range dependence assumptions $\Delta_p := \sum_{\mathbf{t}\succeq\mathbf{0}} \delta_p(\mathbf{t}) < \infty$; namely, the cumulative dependence of $\varepsilon(\mathbf{0})$ on the future values $X(\mathbf{t})_{\mathbf{t}\succeq\mathbf{0}}$ is finite. Our first main result is the following

Theorem 4.1 *Let Assumptions 1, 2:a) are satisfied. Assume* $E\{X(\mathbf{t})\} = 0$, $X(\mathbf{t}) \in \mathcal{L}^p, p \geq 2$ *and* $\Delta_p < \infty$. *Let* $B_{\mathbf{n}} \longrightarrow \infty$ *and* $\widehat{B}_{\mathbf{n}} = o(\widehat{\mathbf{n}})$ *as* $\mathbf{n} \longrightarrow \infty$. *Then*

$$\sup_{\lambda} \|f_{\mathbf{n}}(\lambda) - f(\lambda)\|_{p/2} \longrightarrow 0. \tag{8}$$

The following theorem establishes the asymptotic normality of the estimator spectral density.

Theorem 4.2 *Assume* $E\{X(\mathbf{t})\} = 0$, $E\{X^4(\mathbf{t})\} < \infty$ *and* $\Delta_4 < \infty$. *Let* $B_{\mathbf{n}} \longrightarrow \infty, \widehat{B}_{\mathbf{n}} = o(\widehat{\mathbf{n}})$ *as* $\mathbf{n} \longrightarrow \infty$. *Then under Assumptions 1, 2:b), we have*

$$(\widehat{\mathbf{n}}\widehat{b_{\mathbf{n}}})^{1/2} \left(f_{\mathbf{n}}(\lambda) - E\{f_{\mathbf{n}}(\lambda)\}\right) \longrightarrow \mathcal{N}\left(0, \sigma^2(\lambda)\right), \tag{9}$$

where

$$\sigma^2(\lambda) = \eta(\lambda) f^2(\lambda)\kappa, \mathbf{0} \preceq \lambda \preceq \pi$$

$$\eta(\lambda) = \begin{cases} 2, & \text{if } \frac{\lambda}{\pi} \in \mathbb{Z}^d. \\ 1, & \text{if } \frac{\lambda}{\pi} \notin \mathbb{Z}^d. \end{cases}$$

To state the maximum deviations result, we need the following assumptions

Assumption 3

There exists $0 < \underline{\delta} < \delta < 1$ *and* $c_1, c_2 > 0$ *such that for all large* $\widehat{\mathbf{n}}, c_1\widehat{\mathbf{n}}^{\underline{\delta}} \leq \widehat{B}_{\mathbf{n}} \leq c_2\widehat{\mathbf{n}}^{\underline{\delta}}$ *holds.*

Assumption 4

a) $\Delta_p = O\left(\widehat{\mathbf{n}}^{-T_1}\right)$ *with* $T_1 > \max\left[\frac{1}{2} - (p-4)/(2p\delta), 2\delta/p\right].$

b) $\Delta_p = O\left(\widehat{\mathbf{n}}^{-T_2}\right)$ *with* $T_2 > \max\left[0, 1 - (p-4)/(2p\delta)\right].$

Theorem 4.3 *Assume* $X(\mathbf{0}) \in \mathcal{L}^p, p > \max(4, 2/(1-\delta))$ *and* $E\{X(\mathbf{0})\} = 0$. *Further assume Assumptions 1, 2:c), 3 and 4. Let* $\lambda_i^* = \pi|i|.\frac{1}{B_{\mathbf{n}}}$. *Then, for all* $x \in \mathbb{R}$,

$$P\left[\max_i \frac{\widehat{\mathbf{n}}}{\widehat{B}_{\mathbf{n}}} \frac{|f_{\mathbf{n}}(\lambda_i^*) - E\{f_{\mathbf{n}}(\lambda_i^*)\}|^2}{f^2(\lambda_i^*)\kappa} - 2\log\widehat{B}_{\mathbf{n}} + \log\left(\pi\log\widehat{B}_{\mathbf{n}}\right) \leq x\right] \longrightarrow e^{-e^{-x/2}}. \tag{10}$$

Bibliographie

[1] Bibi, A., Kimouche, K. (2014), On stationarity and second order properties of bilinear random fields, Stat inference stoch process, 17, 221-244.
[2] Cressie, N.A.C. (1991), *Statistics for Spatial Data*, Wiley, New York.
[3] El Machkouri, M., Voln y, D. Wu, W. B. (2013), A central limit theorem for stationary random fields, Stochastic Process, Appl, 123, 1–14.
[4] Guyon, X. (1995), *Random fields on the network,* Springer-Verglag, New York.
[5] Liu, W., Wu, W. B. (2010), Asymptotics of spectral density estimates, Econometric Theorm, 26, 1218–1245.
[6] Rosenblatt, M. (1985), *Stationary sequences and random fields*, Birkhäuser.
[7] Shao, X., Wu, W. B. (2007), Asymptotic spectral theory for nonlinear time series, The annals of statistics, 35, 1773–1801.
[8] Wu, W. B,. (2005), Nonlinear system theory: another look at dependence, Proc. Natl. Acad. Sci. USA, 102, 14150–14154.