C. R. Acad. Sci. Paris, t. 322, Série I, p. 695-698, 1996
Probabilités/*Probability Theory*

Grandes déviations pour des formes quadratiques de processus gaussiens stationnaires

Bernard BERCU, Fabrice GAMBOA et Alain ROUAULT

B. B. : Laboratoire de Statistiques, Université Paris-Sud, 91405 Orsay, France ;

F. G. : Laboratoire de Statistiques, Université Paris-Sud, 91405 Orsay, France et Université Paris-Nord, Institut Galilée, 93430 Villetaneuse, France ;

A. R. : Université de Versailles, 78035 Versailles, France.

Résumé. On montre un principe de grandes déviations pour des formes quadratiques de Toeplitz de processus gaussiens stationnaires centrés. La fonction de taux s'obtient par la méthode MEM (maximum d'entropie sur la moyenne) et par une étude fine du comportement des valeurs propres d'un produit de deux matrices de Toeplitz.

Large deviations for quadratic forms of Gaussian stationary processes

Abstract. *A large deviation principle is proved for Toeplitz quadratic forms of centered stationary gaussian processes. The rate functional is obtained by MEM method (maximum entropy on the mean) and by a sharp study of the behaviour of eigenvalues of a product of two Toeplitz matrices.*

1. Introduction

Soit $(X_n)_{n\in\mathbb{Z}}$ un processus gaussien réel stationnaire centré de densité spectrale g continue sur le tore $\mathbb{T}$. Son périodogramme empirique est la mesure aléatoire $\mathcal{W}_n$ définie sur les fonctions continues f par

$$\mathcal{W}_n(f) = \frac{1}{2\pi n}\int_{\mathbf{T}} f(x)\left|\sum_{j=1}^{n} X_j \exp(ijx)\right|^2 dx.$$

Soit $T_n(f)$ la n-ième matrice de Toeplitz associée à f

$$T_n(f) = \left(\frac{1}{2\pi}\int_{\mathbf{T}} \exp(i(j-k)x) f(x)dx\right)_{1\le j,\, k\le n}.$$

On peut écrire

$$\mathcal{W}_n(f) = \frac{1}{n} X^{(n)*} T_n(f) X^{(n)}$$

où $X^{(n)*} = (X_1, \dots, X_n)$. Si l'on pose $X^{(n)} = T_n(g)^{\frac{1}{2}} Y^{(n)}$

$$\mathcal{W}_n(f) = \frac{1}{n} Y^{(n)*} T_n(g)^{\frac{1}{2}} T_n(f) T_n(g)^{\frac{1}{2}} Y^{(n)}$$

Note présentée par Jean-Pierre KAHANE.

avec $Y^{(n)}$ de loi $N(O, I_n)$. Après un changement de base orthonormée, on a

$$\mathcal{W}_n(f) = \frac{1}{n}\sum_{j=1}^{n} \lambda_j^{(n)} Z_j^{(n)}$$

où $(\lambda_1^{(n)}, \ldots, \lambda_n^{(n)})$ sont les valeurs propres de $T_n(g)^{\frac{1}{2}} T_n(f) T_n(g)^{\frac{1}{2}}$ (ou de $T_n(f)T_n(g)$) et $(Z_1^{(n)}, \ldots, Z_n^{(n)})$ est un n-échantillon de loi $\chi^2(1)$. D'après le théorème de Grenander-Szegö [7]

$$\frac{1}{n}\sum_{j=1}^{n} \delta_{\lambda_j^{(n)}} \xrightarrow{\mathcal{L}} P_{fg}$$

où P_{fg} est la loi image de la loi uniforme sur $\mathbb{T}$ par l'application fg. Comme $T_n(f)T_n(g)$ n'est pas une matrice de Toeplitz, la plus petite et la plus grande des valeurs propres de $T_n(f)T_n(g)$ n'ont pas en général pour limites respectives $m(fg) = \min_{x\in\mathbf{T}} f(x)g(x)$ et $M(fg) = \max_{x\in\mathbf{T}} f(x)g(x)$. Si $I_{fg} = [m(fg), M(fg)]$, les valeurs propres de $T_n(f)T_n(g)$ ne sont donc pas toutes à l'intérieur de cet intervalle mais elles sont majorées en valeur absolue par $\|f\|_\infty \|g\|_\infty$. On décompose $\mathcal{W}_n(f) = \widehat{\mathcal{W}}_n(f) + \tilde{\mathcal{W}}_n(f)$ avec

$$\widehat{\mathcal{W}}_n(f) = \frac{1}{n}\sum_{j=1}^{n} \lambda_j^{(n)} Z_j^{(n)} \mathbf{1}_{\lambda_j^{(n)} \in I_{fg}}$$

et $\tilde{\mathcal{W}}_n(f)$ pour les valeurs propres à l'extérieur de I_{fg}. Les grandes déviations de $\widehat{\mathcal{W}}_n(f)$ découlent d'un résultat général obtenu grâce à la méthode MEM [6]. Les grandes déviations de $\tilde{\mathcal{W}}_n(f)$ proviennent de l'étude fine du comportement des valeurs propres situées à l'extérieur de I_{fg}. Le théorème classique de Gärtner-Ellis (*voir* le théorème 2.3.6 de [5]) ne s'applique pas ici pour des raisons de bord. Le passage par MEM permet d'éviter le recours à un changement de probabilité dépendant du temps, utilisé pour des fonctions f particulières par Bryc et Dembo [3].

2. Maximum d'entropie sur la moyenne

Soit S un compact de $\mathbb{R}$ et P une mesure de probabilité de support S. Soit $\mathcal{M}^+(S)$ l'ensemble des mesures positives sur S. Pour une suite triangulaire $(t_1^{(n)}, \ldots, t_n^{(n)})$ d'éléments de S, on suppose que

$$\frac{1}{n}\sum_{j=1}^{n} \delta_{t_j^{(n)}} \xrightarrow{\mathcal{L}} P.$$

On définit la suite de mesures aléatoires $(\nu_n)_{n\geq 1}$ par

$$\nu_n = \frac{1}{n}\sum_{j=1}^{n} Z_j^{(n)} \delta_{t_j^{(n)}}$$

où $(Z_1^{(n)}, \ldots, Z_n^{(n)})$ est un n-échantillon de loi $\chi^2(1)$. Le théorème suivant et son corollaire précisent les grandes déviations de la suite $(\nu_n)_{n\geq 1}$.

THÉORÈME 1. – *Pour toute mesure positive $\mu \in \mathcal{M}^+(S)$ de décomposition de Lebesgue par rapport à P, $\mu = h_\mu P + \eta$, soit*

$$\begin{aligned} I_P(\mu) &= \frac{1}{2}\left(-\int_S \log h_\mu(t) P(dt) + \mu(S) - 1\right) \qquad si \quad \log h_\mu \in L^1(P), \\ &= +\infty \quad sinon. \end{aligned}$$

Alors, $(\nu_n)_{n\geq 1}$ satisfait un principe de grandes déviations (PGD) avec bonne fonction de taux I_P.

COROLLAIRE 2. – *Soit φ une fonction continue définie sur S, à valeurs dans $\mathbb{R}^d$ avec $d \geq 1$. Alors, la suite des vecteurs aléatoires $(\nu_n(\varphi))_{n\geq 1}$ satisfait un PGD avec bonne fonction de taux*

$$K_\varphi(x) = \inf_{\mu(\varphi)=x} I_P(\mu) = \inf_{h>0,\, [hP](\varphi)=x} I_P(hP) \qquad (x \in \mathbb{R}^d).$$

Une forme duale de K_φ est donnée par

$$K_\varphi(x) = \sup_{y\in\mathbb{R}^d} \left(\langle x,\, y\rangle + \frac{1}{2}\int_S \log(1 - 2\langle y,\, \varphi(t)\rangle)P(dt) \right) \qquad (x \in \mathbb{R}^d).$$

Preuve. – Le théorème 1 se montre à partir d'un théorème de Baldi [1] (*voir* aussi le théorème 4.5.20 de [5]). Une partie de sa démonstration figure dans [6]. On peut en déduire le corollaire 2 grâce au principe de contraction. La forme duale de la fonction de taux est établie dans le théorème 3 de [8] (*voir* également le théorème 2.4 de [6]).

3. Résultats sur les formes quadratiques

Soient $\underline{a}_n(f,\, g)$ et $\overline{a}_n(f,\, g)$ la plus petite et la plus grande valeurs propres de $T_n(f)T_n(g)$ respectivement. Les suites $(\underline{a}_n(f,\, g))_{n\geq 1}$ et $(\overline{a}_n(f,\, g))_{n\geq 1}$ sont compactes. On peut donc définir l'ensemble $\mathcal{A}(f,\, g)$ des couples limites des suites $(\underline{a}_n(f,\, g), \overline{a}_n(f,\, g))_{n\geq 1}$. Pour tout $(\underline{a}, \overline{a}) \in \mathcal{A}(f,\, g)$, on a $\max(|\underline{a}|,\, |\overline{a}|) \leq \|f\|_\infty \|g\|_\infty$. Soit $\mathcal{N}(\underline{a},\, \overline{a})$ l'ensemble des suites croissantes d'entiers $(n_j)_{j\geq 1}$ telles que les sous-suites $(\underline{a}_{n_j}(f,\, g), \overline{a}_{n_j}(f,\, g))_{j\geq 1}$ convergent vers $(\underline{a}, \overline{a}) \in \mathcal{A}(f,\, g)$ quand j tend vers l'infini. Pour $\widehat{\mathcal{W}}_n(f)$, on utilise la partie précédente avec $S = I_{fg}$, $P = P_{fg}$ et φ la fonction identité sur $\mathbb{R}$. Pour $\widetilde{\mathcal{W}}_n(f)$, on remarque que les valeurs propres concernées sont en proportion négligeable. Seules les valeurs propres extrêmes contribuent à la détermination de la fonction de taux. On pose

$$K_{fg}(x) = \sup_{y\in\mathbb{R}} \left(xy + \frac{1}{4\pi}\int_{\mathbf{T}} \log\left(1 - 2yf(t)g(t)\right)dt \right) \qquad (x \in \mathbb{R})$$

avec pour convention $\log z = -\infty$ si $z \leq 0$.

$$J_{fg}^{(\underline{a},\overline{a})}(x) = \begin{cases} K_{fg}(x) & \text{si} \quad x \in]x_1,\, x_2[\\ K_{fg}(x_1) + \dfrac{1}{2\underline{a}}(x - x_1) & \text{si} \quad x \in]-\infty,\, x_1] \\ K_{fg}(x_2) + \dfrac{1}{2\overline{a}}(x - x_2) & \text{si} \quad x \in [x_2,\, +\infty[\end{cases}$$

où $x_1,\, x_2$ sont implicitement données par

$$K'_{fg}(x_1) = \frac{1}{2\underline{a}} \quad \text{si } \underline{a} < 0 \quad \text{et} \quad \underline{a} \notin I_{fg}\,; \qquad x_1 = -\infty \quad \text{sinon}$$

$$K'_{fg}(x_2) = \frac{1}{2\overline{a}} \quad \text{si } \overline{a} > 0 \quad \text{et} \quad \overline{a} \notin I_{fg}\,; \qquad x_2 = +\infty \quad \text{sinon.}$$

THÉORÈME 3. – *On suppose que $P_{fg}(\{m_{fg},\, M_{fg}\}) = 0$. Pour $(n_j) \in \mathcal{N}(\underline{a}, \overline{a})$ avec $(\underline{a}, \overline{a}) \in \mathcal{A}(f,\, g)$, la sous-suite $(\mathcal{W}_{n_j}(f))_{j\geq 1}$ satisfait un PGD avec bonne fonction de taux $J_{fg}^{(\underline{a},\overline{a})}$.*

COROLLAIRE 4. – *On suppose que $P_{fg}(\{m_{fg},\, M_{fg}\}) = 0$. Soit $\mathcal{R}(g)$ l'ensemble des fonctions positives f telles que $\|fg\|_\infty = \|f\|_\infty \|g\|_\infty$.*

a) $(\mathcal{W}_n(f))_{n\geq 1}$ satisfait un PGD si et seulement si ou $\mathcal{A}(f,\, g) \subset I_{fg} \times I_{fg}$ ou bien $\mathcal{A}(f,\, g)$ est réduit à un singleton.

b) Si f appartient à $\mathcal{R}(g)$ alors $(\mathcal{W}_n(f))_{n\geq 1}$ satisfait un PGD avec bonne fonction de taux K_{fg}.

Commentaire. – Si la fonction f est constante sur le tore $\mathbb{T}$, $f \in \mathcal{R}(g)$. Le corollaire 4 entraîne alors immédiatement que la suite $\left(\frac{1}{n}\sum_{j=1}^{n} X_j^2\right)_{n\geq 1}$ satisfait un PGD avec bonne fonction de taux K_{fg}. On retrouve ainsi le corollaire 2.1 de [3].

4. Applications

a) Cas du bruit blanc

On suppose que $(X_n)_{n\in\mathbb{Z}}$ est une suite de variables aléatoires indépendantes. La densité spectrale g est donc constante sur le tore $\mathbb{T}$ et l'on choisira sans perte de généralité $g = 1$. P est la probabilité uniforme sur $\mathbb{T}$.

THÉORÈME 5. – *La suite des périodogrammes empiriques* $(\mathcal{W}_n)_{n\geq 1}$ *à valeurs dans* $\mathcal{M}^+(\mathbb{T})$ *satisfait un PGD avec bonne fonction de taux* I_P. *De plus, soit* φ *une fonction continue définie sur* $\mathbb{T}$, *à valeurs dans* $\mathbb{R}^d$ *avec* $d \geq 1$. *Alors, la suite des vecteurs aléatoires* $(\mathcal{W}_n(\varphi))_{n\geq 1}$ *satisfait un PGD avec bonne fonction de taux* K_φ.

b) Cas du processus autorégressif

On considère le processus défini par

$$X_{n+1} = \theta X_n + \varepsilon_{n+1}, \qquad |\theta| < 1$$

où $(\varepsilon_n)_{n\geq 1}$ est une suite i.i.d. de loi $\mathcal{N}(0,\,1)$, X_0 est indépendante de $(\varepsilon_n)_{n\geq 1}$ et de loi $\mathcal{N}(0,\,1/(1-\theta^2))$. $(X_n)_{n\geq 0}$ est un processus gaussien stationnaire centré de densité spectrale $g(x) = (1+\theta^2 - 2\theta\cos x)^{-1}$ avec $x \in \mathbb{T}$. On peut obtenir les grandes déviations de l'estimateur des moindres carrés de θ via le corollaire 4 [2]. On préfère illustrer ici le comportement des valeurs propres de $T_n(f)T_n(g)$ sur la fonction $f(x) = 1 + 2\cos x$ où $x \in \mathbb{T}$. On a $m(fg) = -(1+\theta)^{-2}$, $M(fg) = 3(1-\theta)^{-2}$ et

$$\mathcal{W}_n(f) = \frac{1}{n}\left(\sum_{j=1}^{n} X_j^2 + 2\sum_{j=2}^{n} X_j X_{j-1}\right).$$

Si $-1/2 < \theta \leq 1$, $\underline{a} = m(fg)$ et $\overline{a} = M(fg)$. Par contre, si $-1 < \theta \leq -1/2$, $\underline{a} = m(fg)$ et $\overline{a} = -\theta^{-1}(2+\theta)^{-1}$.

c) Vers d'autres applications

Dans [2], on établit des PGD pour des formes quadratiques qui ne sont pas nécessairement Toeplitz. En particulier, on précise l'étude sur le rapport de vraisemblance de deux processus gaussiens stationnaires menée par Coursol et Dacunha-Castelle [4].

Note remise le 6 février 1996, acceptée le 12 février 1996.

Références bibliographiques

[1] **Baldi P., 1988.** Large deviations and stochastic homogenization, *Ann. Mat. Pura Appl.*, 151, p. 161-177.

[2] **Bercu B., Gamboa F. et Rouault A., 1996.** Large deviations for quadratic forms of gaussian stationary processes, *preprint*, Orsay.

[3] **Bryc W. et Dembo A., 1995.** Large deviations for quadratic functionals of gaussian processes, *preprint*.

[4] **Coursol J. et Dacunha-Castelle D., 1979.** Sur la formule de Chernoff pour deux processus gaussiens stationnaires, *C. R. Acad. Sci. Paris*, 288, série A, p. 769-770.

[5] **Dembo A. et Zeitouni O., 1993.** *Large deviations techniques and applications*, Jones and Barblett pub. Boston.

[6] **Gamboa F. et Gassiat E., 1996.** Bayesian methods for ill posed problems, *Ann. Statist.* (à paraître).

[7] **Grenander V. et Szegö G., 1958.** *Toeplitz forms and their applications*, University of California Press.

[8] **Rockafellar R. T., 1971.** Integrals which are convex functionals, *Pacific J. Math.*, 39, p. 439-469.