Probabilités/*Probability Theory*, **Statistique**/*Statistics*

Inégalités de covariance

Covariance inequalities

Jérôme Dedecker [a]

[a] *Laboratoire de Statistique Théorique et Appliquée, Université Paris 6, Site Chevaleret, 13 rue Clisson, 75013 Paris.*

Abstract

In this note, we show that some well known covariance inequalities expressed in terms of mixing coefficients remain true for weaker coefficients. Next, we give some examples of non-mixing processes for which these weaker coefficients can be easily bounded. *To cite this article: J. Dedecker, C. R. Acad. Sci. Paris, Ser. I ...*

Résumé

Dans cette note, nous montrons que certaines inégalités de covariance écrites en terme de coefficients de mélange restent vraies pour des versions faibles de ces coefficients. Nous donnons ensuite quelques exemples de processus non-mélangeants pour lesquels nous pouvons obtenir sans peine des bornes pour les coefficients faibles. *Pour citer cet article : J. Dedecker, C. R. Acad. Sci. Paris, Ser. I ...*

1. Les coefficients

Soient X et Y deux variables aléatoires réelles sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Rappelons d'abord les définitions classiques des coefficients de β-mélange et de ϕ-mélange. Soit $P_{X|Y}$ une distribution régulière de X sachant Y et P_X la distribution de X. On définit la variable aléatoire

$$b(\sigma(Y), \sigma(X)) = \sup_{A \in \mathcal{B}(\mathbb{R})} |P_{X|Y}(A) - P_X(A)|,$$

et les coefficients $\beta(\sigma(Y), \sigma(X)) = \|b(\sigma(Y), \sigma(X))\|_1$ et $\phi(\sigma(Y), \sigma(X)) = \|b(\sigma(Y), \sigma(X))\|_\infty$.

Pour sortir du cadre mélangeant, et considérer une classe de modèles plus vaste que celle des processus mélangeants, nous avons proposé dans [2] une version faible des coefficients de mélange ci-dessus. Soit

Email address: `dedecker@ccr.jussieu.fr` (Jérôme Dedecker).

$F_{X|Y} = \{t \to P_{X|Y}(]-\infty, t])\}$ une fonction de répartition de X sachant Y et F_X la fonction de répartition de X. On définit la variable aléatoire

$$b(\sigma(Y), X) = \sup_{x\in\mathbb{R}} |F_{X|Y}(x) - F_X(x)| ,$$

et les coefficients $\beta(\sigma(Y), X) = \|b(\sigma(Y), X)\|_1$ et $\phi(\sigma(Y), X) = \|b(\sigma(Y), X)\|_\infty$. Remarquons que la valeur de $b(\sigma(Y), X)$ ne change pas si l'on remplace $F_{X|Y}(x)$ et $F_X(x)$ par $P_{X|Y}(] - \infty, x[)$ et $P_X(] - \infty, x[)$, et par conséquent $b(\sigma(Y), X) = b(\sigma(Y), -X)$ presque sûrement.

On peut aussi définir le coefficient de dépendance $\phi(X, Y)$ entre X et Y par

$$\phi(Y, X) = \sup_{x,y\in\mathbb{R}} \|\mathbb{E}(\mathbf{1}_{X\leq x}|\mathbf{1}_{Y\leq y}) - F_X(x)\|_\infty .$$

La valeur de ce coefficient ne change pas si l'on change les inégalités larges en inégalités strictes, et par conséquent $\phi(Y, -X) = \phi(Y, X) = \phi(-Y, X) = \phi(-X, -Y)$.

On voit immédiatement, en partant de ces définitions que

$$\beta(\sigma(Y), X) \leq \beta(\sigma(Y), \sigma(X)) \quad \text{et} \quad \phi(Y, X) \leq \phi(\sigma(Y), X) \leq \phi(\sigma(Y), \sigma(X)) .$$

Dans [2] nous avons donné de nombreux exemples de processus $(X_i)_{i\geq 0}$ non mélangeants pour lesquels $\beta(\sigma(X_0), X_n)$ ou même $\phi(\sigma(X_0), X_n)$ tend vers 0, avec une vitesse que l'on peut évaluer. On rappelera certains de ces exemples dans la section 3.

2. Les inégalités

Proposition 2.1 *Pour tous les exposants conjugués p et q, on a les inégalités (qui restent vraies pour $p = 1$ et $p = \infty$ avec les énoncés adéquats)*

$$|\mathrm{Cov}(Y, X)| \leq 2\{\mathbb{E}(|X|^p b(\sigma(X), Y))\}^{1/p}\{\mathbb{E}(|Y|^q b(\sigma(Y), X))\}^{1/q}, \quad et \tag{1}$$

$$|\mathrm{Cov}(Y, X)| \leq 2\phi(X, Y)^{1/p}\phi(Y, X)^{1/q}\|X\|_p\|Y\|_q . \tag{2}$$

Remarque 1 L'inégalité (1) est celle donnée dans [3] (voir aussi [6]), mais avec les versions faibles $b(\sigma(Y), X)$ et $b(\sigma(X), Y)$ au lieu de $b(\sigma(Y), \sigma(X))$ et $b(\sigma(X), \sigma(Y))$. L'inégalité (2) est celle donnée dans [4], mais avec $\phi(Y, X)$ et $\phi(X, Y)$ au lieu de $\phi(\sigma(Y), \sigma(X))$ et $\phi(\sigma(X), \sigma(Y))$. En partant de (1) on obtient l'inégalité $|\mathrm{Cov}(X, Y)| \leq 2\phi(\sigma(X), Y)^{1/p}\phi(\sigma(Y), X)^{1/q}\|X\|_p\|Y\|_q$ qui est moins précise que (2).

Preuve de la proposition 2.1. On part de l'égalité

$$\mathrm{Cov}(Y, X) = \int_0^\infty \int_0^\infty \mathrm{Cov}(\mathbf{1}_{X>x} - \mathbf{1}_{X<-x}, \mathbf{1}_{Y>y} - \mathbf{1}_{Y<-y})\, dx\, dy . \tag{3}$$

Montrons d'abord (1). On a les quatre inégalités

$$\begin{aligned}
|\mathbb{E}(\mathbf{1}_{X>x}\mathbf{1}_{Y>y} - \mathbb{P}(X > x)\mathbb{P}(Y > y))| &\leq \mathbb{E}(\mathbf{1}_{Y>y}b(\sigma(Y), X)) \wedge \mathbb{E}(\mathbf{1}_{X>x}b(\sigma(X), Y))\\
|\mathbb{E}(\mathbf{1}_{X<-x}\mathbf{1}_{Y>y} - \mathbb{P}(X < -x)\mathbb{P}(Y > y))| &\leq \mathbb{E}(\mathbf{1}_{Y>y}b(\sigma(Y), X)) \wedge \mathbb{E}(\mathbf{1}_{X<-x}b(\sigma(X), Y))\\
|\mathbb{E}(\mathbf{1}_{X>x}\mathbf{1}_{Y<-y} - \mathbb{P}(X > x)\mathbb{P}(Y < -y))| &\leq \mathbb{E}(\mathbf{1}_{Y<-y}b(\sigma(Y), X)) \wedge \mathbb{E}(\mathbf{1}_{X>x}b(\sigma(X), Y))\\
|\mathbb{E}(\mathbf{1}_{X<-x}\mathbf{1}_{Y<-y} - \mathbb{P}(X < -x)\mathbb{P}(Y < -y))| &\leq \mathbb{E}(\mathbf{1}_{Y<-y}b(\sigma(Y), X)) \wedge \mathbb{E}(\mathbf{1}_{X<-x}b(\sigma(X), Y)) .
\end{aligned}$$

Puisque $a_1 \wedge b_1 + a_1 \wedge b_2 + a_2 \wedge b_1 + a_2 \wedge b_2 \leq 2(a_1 + a_2) \wedge (b_1 + b_2)$, on obtient d'après (3) que

$$|\mathrm{Cov}(Y, X)| \leq 2\int_0^\infty \int_0^\infty \mathbb{E}(\mathbf{1}_{|X|>x}b(\sigma(X), Y)) \wedge \mathbb{E}(\mathbf{1}_{|Y|>y}b(\sigma(Y), X))\, dx\, dy . \tag{4}$$

Notons $\beta_1 = \beta(\sigma(X), Y)$ et $\beta_2 = \beta(\sigma(Y), X)$. Remarquons que β_1 (ou β_2) est nul si et seulement si X est indépendante de Y, et par conséquent (1) est vraie dans ce cas. Dans le cas contraire, notons $\mathbb{P}_{\beta_1}$, $\mathbb{P}_{\beta_2}$ les probabilités de densités $b(\sigma(X), Y)/\beta_1$ et $b(\sigma(Y), X)/\beta_2$ par rapport à $\mathbb{P}$. L'inégalité (4) s'écrit donc

$$|\mathrm{Cov}(Y, X)| \leq 2 \int_0^\infty \int_0^\infty \beta_1 \mathbb{P}_{\beta_1}(|X| > x) \wedge \beta_2 \mathbb{P}_{\beta_2}(|Y| > y)\, dx\, dy\,. \tag{5}$$

Notons $Q_{\beta_1,X}$ et $Q_{\beta_2,Y}$ les inverses càdlàg de $x \to \mathbb{P}_{\beta_1}(|X| > x)$ et $y \to \mathbb{P}_{\beta_2}(|Y| > y)$. En partant de (5), on obtient successivement

$$\begin{aligned}
|\mathrm{Cov}(Y, X)| &\leq 2 \int_0^\infty \int_0^\infty \int_0^{\beta_1 \wedge \beta_2} \mathbf{I}_{u < \beta_1 \mathbb{P}_{\beta_1}(|X| > x)} \mathbf{I}_{u < \beta_2 \mathbb{P}_{\beta_2}(|Y| > y)}\, du\, dx\, dy \\
&\leq 2 \int_0^\infty \int_0^\infty \int_0^{\beta_1 \wedge \beta_2} \mathbf{I}_{Q_{\beta_1,X}(u/\beta_1) > x} \mathbf{I}_{Q_{\beta_2,Y}(u/\beta_2) > y}\, du\, dx\, dy \\
&\leq 2 \int_0^{\beta_1 \wedge \beta_2} Q_{\beta_1,X}(u/\beta_1) Q_{\beta_2,Y}(u/\beta_2)\, du\,.
\end{aligned}$$

En appliquant l'inégalité de Hölder et en effectuant les changements $s = u/\beta_1$ et $t = u/\beta_2$, on obtient

$$|\mathrm{Cov}(Y, X)| \leq 2 \left(\int_0^1 \beta_1 Q^p_{\beta_1,X}(s)\, ds \right)^{1/p} \left(\int_0^1 \beta_2 Q^q_{\beta_2,Y}(t)\, dt \right)^{1/q}.$$

On termine la preuve de (1) en notant que, grâce aux définitions de $Q_{\beta_1,X}$ et $Q_{\beta_2,Y}$,

$$\int_0^1 \beta_1 Q^p_{\beta_1,X}(s)\, ds = \mathbb{E}(|X|^p b(\sigma(X), Y)) \quad \text{et} \quad \int_0^1 \beta_2 Q^q_{\beta_2,Y}(t)\, dt = \mathbb{E}(|Y|^q b(\sigma(Y), X))\,.$$

Pour montrer (2), on part de (3) et on transforme les quatre inégalités qui suivent ainsi :

$$|\mathbb{E}(\mathbf{I}_{X>x} \mathbf{I}_{Y>y} - \mathbb{P}(X > x)\mathbb{P}(Y > y))| \leq \mathbb{P}(Y > y)\phi(Y, X) \wedge \mathbb{P}(X > x)\phi(X, Y)$$

(de même pour les trois autres). On termine la preuve de la même manière que pour (1).

Corollaire 2.2 *Soient f_1, f_2, g_1, g_2 quatre fonctions croissantes, $f = f_1 - f_2$ et $g = g_1 - g_2$. Pour toute variable aléatoire Z, notons $\Delta_p(Z) = \inf_{a\in\mathbb{R}} \|Z - a\|_p$ et $\Delta_{p,\sigma(X),Y}(Z) = \inf_{a\in\mathbb{R}}(\mathbb{E}(|Z-a|^p b(\sigma(X), Y)))^{1/p}$. Pour tous les exposants conjugués p et q, on a les inégalités (qui restent vraies pour $p = 1$ et $p = \infty$)*

$$|\mathrm{Cov}(g(Y), f(X))| \leq \; 2\big\{\Delta_{p,\sigma(X),Y}(f_1(X)) + \Delta_{p,\sigma(X),Y}(f_2(X))\big\}\big\{\Delta_{q,\sigma(Y),X}(g_1(Y)) + \Delta_{q,\sigma(Y),X}(g_2(Y))\big\}$$

$$|\mathrm{Cov}(g(Y), f(X))| \leq \; 2\,\phi(X, Y)^{1/p}\, \phi(Y, X)^{1/q} \,\big\{\, \Delta_p(f_1(X)) \,+\, \Delta_p(f_2(X)) \,\big\} \big\{\, \Delta_q(g_1(Y)) \,+\, \Delta_q(g_2(Y)) \,\big\}$$

En particulier, si μ est une mesure signée de variation totale $\|\mu\|$ et $f(x) = \mu(]-\infty, x])$, on a

$$|\mathrm{Cov}(Y, f(X))| \leq \|\mu\| \mathbb{E}(|Y| b(\sigma(Y), X)) \quad \textit{et} \quad |\mathrm{Cov}(Y, f(X))| \leq \phi(Y, X)\|\mu\|\, \|Y\|_1\,. \tag{6}$$

Remarque 2 Dans [2], nous avons obtenu (6) par une autre méthode. Ces inégalités permettent d'évaluer le risque en moyenne quadratique intégrée des estimateurs de la densité marginale d'une suite stationnaire.

Preuve du corollaire 2.2. Pour montrer les deux premières inégalités, notons que, pour tout a_1, a_2, b_1, b_2,

$$\begin{aligned}
|\mathrm{Cov}(g(Y), f(X))| \leq\; &|\mathrm{Cov}(g_1(Y) - b_1, f_1(X) - a_1)| + |\mathrm{Cov}(g_1(Y) - b_1, f_2(X) - a_2)| \\
&+ |\mathrm{Cov}(g_2(Y) - b_2, f_1(X) - a_1)| + |\mathrm{Cov}(g_2(Y) - b_2, f_2(X) - a_2)|\,.
\end{aligned}$$

Les fonctions $f_1 - a_1, f_2 - a_2, g_1 - b_1, g_2 - b_2$ étant croissantes, on a $\phi(f_i(X) - a_i, g_j(Y) - b_j) \leq \phi(X, Y)$ et $b(\sigma(f_i(X)), g_j(Y) - b_j) \leq \mathbb{E}(b(\sigma(X), Y)|\sigma(f_i(X)))$ presque sûrement. On applique les inégalités (1) et (2), et on optimise en a_1, b_1, a_2, b_2. Pour montrer les inégalités de (6), on prend $q = 1$ et $p = \infty$. Soit $\mu = \mu_+ - \mu_-$ la décomposition de Jordan de μ. On a alors $f(x) = f_1(x) - f_2(x)$, avec $f_1(x) = \mu_+(]-\infty, x])$ et $f_2(x) = \mu_-(]-\infty, x])$. On applique les inégalités précédentes en notant que $\|\mu\| = \mu_+(\mathbb{R}) + \mu_-(\mathbb{R})$ et que $\Delta_{1,\sigma(Y),X}(Y) \leq \mathbb{E}(|Y| b(\sigma(Y), X))$, $2\Delta_\infty(f_1(X)) \leq \mu_+(\mathbb{R})$, et $2\Delta_\infty(g_1(X)) \leq \mu_-(\mathbb{R})$.

3. Exemples

Dans cette section, nous reprenons certains des exemples développés dans [2]. Nous renvoyons à [2] pour plus de détails, ainsi que l'étude d'autres modèles (notamment les transformations dilatantes de $[0,1]$).

Processus linéaires. Soit $(\xi_i)_{i\in\mathbb{Z}}$ une suite de variables i.i.d. ayant un moment d'ordre $p \geq 1$, et $(a_j)_{j\geq 0}$ une suite de réels dans ℓ^1. Soit $(X_i)_{i\in\mathbb{Z}}$ la suite stationnaire définie par $X_n = \sum_{j\geq 0} a_j \xi_{n-j}$. Si X_0 possède une densité bornée par K, on a les majorations

$$\beta(\sigma(X_0), X_n) \leq 2\Big(K\|\xi_1 - \xi_0\|_p \sum_{j\geq n} |a_j|\Big)^{p/(p+1)} \quad \text{et} \quad \phi(\sigma(X_0), X_n) \leq K\|\xi_1 - \xi_0\|_\infty \sum_{j\geq n} |a_j| \,.$$

En particulier, si $a_i = 2^{-i-1}$ et $\xi_0 \sim \mathcal{B}(1/2)$, alors X_0 est uniformément distribuée sur $[0,1]$ et donc $\phi(\sigma(X_0), X_n) \leq 2^{-n}$. Pourtant $(X_i)_{i\in\mathbb{Z}}$ n'est pas mélangeante car $\beta(\sigma(X_0), \sigma(X_n)) \geq 1/2$ pour tout n (voir par exemple [1], page 180). En fait on peut même montrer que $\beta(\sigma(X_0), \sigma(X_n)) = 1$ pour tout n, comme me l'a fait remarquer un rapporteur.

Les propriétés de β-mélange des processus linéaires sont étudiées dans [5]. Dans cet article, les auteurs font des hypothèses sur la loi de ξ_0 qui ne sont pas nécessaires pour les coefficients faibles.

Chaînes de Markov. Soit $(X_i)_{i\geq 0}$ une solution stationnaire de l'équation $X_n = F(X_{n-1}, \xi_n)$, où $(\xi)_{i\geq 1}$ est une suite i.i.d. indépendante de X_0. Supposons que la fonction F satisfasse la condition de lipschitz faible $\|F(x,\xi_1) - F(y,\xi_1)\|_p \leq \kappa|x-y|$ pour un $\kappa < 1$ et un $p \geq 1$. Soit X'_0 une variable aléatoire indépendante et de même loi que X_0. Si X_0 possède une densité bornée par K, on a les majorations

$$\beta(\sigma(X_0), X_n) \leq 2(K\|X_0 - X'_0\|_p \kappa^n)^{p/(p+1)} \quad \text{et} \quad \phi(\sigma(X_0), X_n) \leq 2(K\|X_0 - X'_0\|_\infty \kappa^n)^{p/(p+1)} \,.$$

Si la condition de lipschitz est vraie uniformément ($p = \infty$), on a $\phi(\sigma(X_0), X_n) \leq K\|X_0 - X'_0\|_\infty \kappa^n$.

Par exemple, dans le cas du modèle autorégressif $X_n = f(X_{n-1}) + \xi_n$, la condition de lipshitz est vraie uniformément dès que f est κ-lipschitzienne.

Remerciements

Je remercie Magda Peligrad d'avoir passé une journée à discuter de ces problèmes avec moi. Les exemples de la section 3 sont tirés de l'article [2] écrit avec Clémentine Prieur. Merci également aux deux rapporteurs pour les améliorations qu'ils m'ont suggérées.

Références

[1] R. C. Bradley, Basic properties of strong mixing conditions, Dependence in probability and statistics. A survey of recent results. Oberwolfach, 1985. E. Eberlein and M. S. Taquu editors, Birkhäuser. (1986) 165-192.

[2] J. Dedecker et C. Prieur, New dependence coefficients. Examples and applications to statistics, (2003) Prépublication. www.ccr.jussieu.fr/lsta/prepublications.html

[3] B. Delyon, Limit theorems for mixing processes, (1990) Tech. Report 546 IRISA, Rennes I.

[4] M. Peligrad, A note on two measures of dependence and mixing sequences, Adv. Appl. Probab. 15 (1983) 461-464.

[5] T. D. Phan et L. T. Tran, Some mixing properties of time series models, Stochastic Process. Appl. 19 (1985) 297-303.

[6] G. Viennet, Inequalities for absolutely regular sequences : application to density estimation. Probab. Theory Relat. Fields 107 (1997) 467-492.