ANNALES SCIENTIFIQUES
DE L'UNIVERSITÉ DE CLERMONT-FERRAND 2
Série Mathématiques

B. VAN CUTSEM

Convergence des suites d'estimateurs ensemblistes du maximum de vraisemblance

Annales scientifiques de l'Université de Clermont-Ferrand 2, tome 58, série *Mathématiques*, nº 12 (1976), p. 177-182

<http://www.numdam.org/item?id=ASCFM_1976__58_12_177_0>

CONVERGENCE DES SUITES D'ESTIMATEURS ENSEMBLISTES DU MAXIMUM DE VRAISEMBLANCE

B. VAN CUTSEM , Université de Grenoble

Abstract : A theorem on the convergence of sequences of set valued maximum likelihood estimators is given for statistical spaces with convex properties relative to the parameter.

Résumé : On donne un théorème de convergence pour les suites d'estimateurs ensemblistes du maximum de vraisemblance dans le cas de structures statistiques possédant des propriétés de convexité par rapport aux paramètre.

A. WALD (1949) donne des conditions suffisantes de convergence des suites d'estimateurs du maximum de vraisemblance, sous réserve que, pour tout $n \in \mathbb{N}$, il existe un unique estimateur du maximum de vraisemblance (m.v.) pour la structure statistique associée à l'échantillon de taille n. Notre propos est de montrer que pour certaines structures statistiques, pour lesquelles l'unicité de l'estimateur du m.v. n'est pas assurée, les hypothèses de A. WALD entrainent néanmoins la convergence - en un sens à préciser - de la suite des estimateurs ensemblistes du m.v.

I - ESTIMATEURS ENSEMBLISTES DU MAXIMUM DE VRAISEMBLANCE

I.1 - Définition

La structure statistique $(\Omega, \mathcal{A}, \{P_\theta \mid \theta \in \Theta\})$ où Θ est une partie convexe de $\mathbb{R}^k$, est dite structure Log-concave normale si la famille $\{P_\theta \mid \theta \in \Theta\}$ est dominée par une mesure σ-finie μ sur $(\mathbb{R}^k, \mathcal{B}_{\mathbb{R}^k})$, et si il existe, pour tout $\theta \in \Theta$, une densité de probabilité $\varphi(.\ ,\theta)$ de P_θ par rapport à μ, telle que, si on pose

$$\forall \omega \in \Omega\ , \forall \theta \in \Theta,\ f(\omega,\theta) = -\operatorname{Log} \varphi(\omega,\theta)$$

$$\forall \omega \in \Omega\ , \forall \theta \in \mathbb{R}^k \setminus \Theta \qquad = +\infty,$$

la fonction f soit un intégrande convexe normal sur $\Omega \times \mathbb{R}^k$. (R.T. ROCKAFELLAR,1968).

I.2 - Exemple 1 : Soit $(\Omega, \mathcal{A}) = (\mathbb{R}^k, \mathcal{B}_{\mathbb{R}^k})$, μ la mesure de Lebesgue sur $\mathbb{R}^k$, C un convexe compact de $\mathbb{R}^k$, $\gamma = \mu(C)$, $\Theta = \mathbb{R}^k$. Posons

$$f(\omega, \theta) = \delta(\omega, C+\theta) + \operatorname{Log} \gamma ,$$

où $\delta(., C+\theta)$ est l'indicatrice convexe de l'ensemble convexe $C + \theta$, translaté de C par θ. La probabilité P_θ est la loi uniforme sur $C + \theta$.

Exemple 2 : Soit $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, $\Theta = \mathbb{R}$, μ la mesure de Lebesgue sur $\mathbb{R}$.

Posons :

$$f(\omega, \theta) = |\omega-\theta| + \operatorname{Log} 2.$$

La probabilité P_θ est la loi de Laplace centrée en θ .

Exemple 3 : Toute structure exponentielle canonique.

I.3 - L'exemple 1 précédent montre que les structures statistiques Log-concaves normales peuvent faire intervenir des fonctions $\theta \to f(\omega, \theta)$ dont l'infimum est atteint sur un ensemble $\Gamma(\omega)$ non réduit à un point. Dans le cas général, nous poserons

$$\Gamma(\omega) = \{\theta \in \Theta \mid \operatorname{Inf} \{f(\omega, \xi) \mid \xi \in \Theta\} = f(\omega, \theta)\}$$

et la mesurabilité de la multiapplication Γ résulte du corollaire 4.3 de (R.T. ROCKAFELLAR, 1969). Nous définirons donc :

Définition : Soit $(\Omega, \mathcal{A}, \{P_\theta \mid \in \Theta\})$ une structure statistique Log-concave normale telle que, pour tout $\omega \in \Omega$, l'infimum de $f(\omega, .)$ soit atteint. On appelle estimateur ensembliste du m.v. la multiapplication mesurable Γ définie sur $(\Omega , \mathcal{A})$ et à valeurs convexes fermées non vides dans Θ :

$$\omega \to \Gamma(\omega) = \{\theta \in \Theta \mid \operatorname{Inf} \{f(\omega, \xi) \mid \xi \in \Theta\} = f(\omega, \theta)\} .$$

I.4 - Remarque : La convexité des fonctions $f(\omega, .)$ n'est pas indispensable pour obtenir la mesurabilité de la multiapplication Γ et écrire la définition précédente. Nous nous limiterons néanmoins à ce cas, car les propriétés des multiapplications à valeurs convexes sont mieux connues.

II - CONVERGENCE DES MULTIAPPLICATIONS MESURABLES

Nous donnons ici brièvement la définition de la convergence de suites de multiapplications et énonçons la propriété dont nous aurons besoin.

II.1 - <u>Définition</u> : Soit $\{F_n\}$ une suite de sous-ensembles de $\mathbb{R}^k$. On pose :

$$\underline{\lim}\, F_n = \{x \in \mathbb{R}^k \mid \exists\ \{x_n\},\ \lim x_n = x,\ \forall n \in \mathbb{N},\quad x_n \in F_n\}$$

$$\overline{\lim}\, F_n = \left\{x \in \mathbb{R}^k \mid \exists\ \{x_n\},\ \lim x_n = x,\ \exists \Psi : \mathbb{N} \to \mathbb{N} \text{ stric. croiss.} \atop \forall n \in \mathbb{N},\quad x_n \in F_{\Psi(n)}\right\}$$

et on dit que F_n <u>converge vers</u> F si $F = \underline{\lim}\, F_n = \overline{\lim}\, F_n$.

On notera $\lim F_n$ la valeur commune de $\underline{\lim}\, F_n$ et $\overline{\lim}\, F_n$, quand elle existe. On peut, en outre, démontrer que si $\{F_n\}$ est une suite de convexes fermés non vides de $\mathbb{R}^k$, alors les ensembles $\underline{\lim}\, F_n$ et $\overline{\lim}\, F_n$ sont convexes fermés.

II.2 - <u>Définition</u> : Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $\{\Gamma_n\}$ une suite de multiapplications mesurables définies sur $(\Omega, \mathcal{A})$ et à valeurs convexes fermées non vides dans $\mathbb{R}^k$. On dit que la suite $\{\Gamma_n\}$ <u>converge P-presque sûrement</u> si l'ensemble $\{\omega \in \Omega \mid \underline{\lim}\, \Gamma_n(\omega) \neq \overline{\lim}\, \Gamma_n(\omega)\}$ est P-négligeable.

On sait de plus que $\underline{\lim}\, \Gamma_n(\omega) = \bigcap_m \bigcup_{n \geq m} \Gamma_n(\omega)$, et par suite que la multiapplication $\underline{\lim}\, \Gamma_n(\omega)$ est mesurable. Il en résulte alors que si la suite $\{\Gamma_n\}$ est P-presque sûrement convergente, la multiapplication limite Γ, définie sauf sur un négligeable, coïncide avec la multiapplication $\underline{\lim}\, \Gamma_n$ et est mesurable (en prolongeant sa définition par l'égalité $\Gamma = \underline{\lim}\, \Gamma_n$ là où Γ n'est pas définie).

II.3 - Nous utiliserons le théorème suivant :

<u>Théorème</u> : Soit $\{\Gamma_n\}$ une suite de multiapplications mesurables définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et à valeurs convexes fermées non vides dans $\mathbb{R}^k$. Supposons que toute suite $\{f_n\}$ où, pour tout $n \in \mathbb{N}$, f_n est une section mesurable de Γ_n, converge P-presque sûrement vers une fonction mesurable f_o. Alors la suite $\{\Gamma_n\}$ converge P-presque sûrement vers la multiapplication $\{f_o\}$.

La seule difficulté de cette démonstration résulte du fait que l'ensemble d'exception de la convergence de la suite de sections $\{f_n\}$ dépend de cette suite. On utilise pour éviter cet écueil, un théorème d'existence d'une famille dénombrable de sections mesurables possédant une propriété de densité. On trouvera un énoncé de ce résultat par exemple dans VALADIER (1970).

III - CONVERGENCE DES SUITES D'ESTIMATEURS ENSEMBLISTES DU M.V.

III.1 - Introduisons les notations suivantes :

Soit g une application de $\mathbb{R}^k$ dans $\mathbb{R} \cup \{+\infty\}$. Nous poserons pour tout $\rho > 0$ et tout $x \in \mathbb{R}^k$

$$f_\rho(x) = \text{Inf } \{f(y) \mid \|y-x\| \leq \rho\}$$

$$a_\rho = \text{Inf } \{f(y) \mid \rho \leq \|y\|\}$$

Nous utiliseraons aussi la notation $g^- = \text{Sup } \{0, -g\}$

III.2 - Nous pouvons alors énoncer le résultat suivant analogue à celui de WALD (1949).

<u>Lemme</u> : Soit $(\Omega, \mathcal{A}, \{P_\theta \mid \theta \in \Theta\})$ une structure statistique Log-concave normale et soit $\theta_0 \in \Theta$. Supposons que les hypothèses suivantes soient vérifiées :

Hyp. 1) $|f(.,\theta_0)|$ est P_{θ_0} - intégrable

Hyp. 2) $\forall \theta \in \Theta, \theta \neq \theta_0, P_{\theta_0} \{f(.,\theta) \neq f(., \theta_0)\} > 0$

Hyp. 3) $\forall \theta \in \Theta, \exists \rho > 0$, $f_\rho^-(.,\theta)$ est P_{θ_0} - intégrable

Hyp. 4) $\exists r_0 > 0$, $a_{r_0}^-$ est P_{θ_0} - intégrable

Hyp. 5) $\forall \omega \in \Omega$, $\lim\limits_{\|\theta\| \to \infty} f(\omega,\theta) = +\infty$.

Alors, pour tout fermé $F \subset \Theta$, tel que $\theta_0 \notin F$, il existe $B \in \mathcal{A}^{\otimes \infty}$, $P_{\theta_0}^{\otimes \infty}(B) = 1$, tel que pour tout $\{\omega_n\} \in B$

$$\lim_n \text{ Inf } f(\omega_1,\theta) + \ldots + f(\omega_n,\theta) - f(\omega_1,\theta_0) - \ldots - f(\omega_n,\theta_0) = +\infty$$

La démonstration de ce lemme est exactement la traduction dans nos notations de la démonstration de WALD.

III. 3 - Nous en déduirons le théorème :

Théorème : Soit $(\Omega, \mathcal{A}, \{P_\theta \mid \theta \in \Theta\})$ une structure statistique Log-concave normale, Soit $\theta_0 \in \Theta$. Supposons que les hypothèses 1 à 5 du lemme III.2 soient vérifiées. Supposons de plus, que, pour tout $n \in \mathbb{N}$, l'estimateur ensembliste Γ_n du maximum de vraisemblance de la structure statistique $(\Omega, \mathcal{A}, \{P_\theta \mid \theta \in \Theta\})^n$ soit P_{θ_0} - presque sûrement à valeurs non vides. Alors il existe $B \in \mathcal{A}^{\otimes \infty}$ $P_{\theta_0}^{\otimes \infty}(B) = 1$, tel que pour tout $\{\omega_n\} \in B$,

$$\lim \Gamma_n(\omega_1, \dots, \omega_n) = \{\theta_0\}$$

Démonstration : Il suffit d'écrire que toute suite $\{u_n\}$ où u_n est une section mesurable de Γ_n est une suite d'estimateurs ponctuels du m.v. dont on démontre la convergence presque sûrement θ_0 exactement comme le fait WALD. Le théorème II.3 prouve alors le résultat.

III.4 - Exemple 1 : Prenons le cas particulier suivant de l'exemple 1 du paragraphe I.2 :

$$\Omega = \mathbb{R}, \Theta = \mathbb{R}, \quad \text{et } f(\omega,\theta) = \delta(\omega, [\theta, \theta+1]).$$

On vérifie alors que

$$\Gamma_n(\omega_1, \dots, \omega_n) = \left[\sup_{1 \leq i \leq n} \omega_i - 1 \ , \ \inf_{1 \leq i \leq n} \omega_i \right]$$

et il est facile de voir que Γ_n converge presque sûrement vers θ_0.

Exemple 2 : Reprenons maintenant le cas général de l'exemple I.2.

Nous avons $\Omega = \mathbb{R}^k$, $\Theta = \mathbb{R}^k$ et $f(\omega,\theta) = \delta(\omega, C+\theta) + \text{Log}\,\gamma$.

On vérifie alors que

$$\Gamma_n(\omega_1, \dots, \omega_n) = \bigcap_{i=1}^{n} \{\omega_i - C\}$$

$$\text{où} \quad \{\omega_i - C\} = \{\theta \in \mathbb{R}^k \mid \theta = \omega_i - u, \ u \in C\}$$

et on peut s'assurer que la suite décroissante $\{\Gamma_n(\omega_1, \dots, \omega_n)\}$ converge presque sûrement vers $\{\theta_0\}$.

Exemple 3 : Prenons le cas de l'exemple 2 du paragraphe I.2.

On voit cette fois que

$$\Gamma_n(\omega_1,\ldots,\omega_n) = \begin{cases} \left\{\omega_{(\frac{n-1}{2}+1)}\right\} & \text{si n est impair} \\ \left[\omega_{\frac{n}{2}}, \omega_{(\frac{n}{2}+1)}\right] & \text{si n est pair} \end{cases}$$

où $\omega_{(p)}$ désigne le $p^{\text{ième}}$ terme de l'échantillon $\omega_1,\ldots,\omega_n$ ordonné par ordre croissant. Ici encore la suite $\{\Gamma_n\}$ converge presque sûrement vers $\{\theta_o\}$.

B I B L I O G R A P H I E

ROCKAFELLAR, R.T. (1968) Integrals which are convex functionals, Pacific J. Math. 24, p. 525-539.

(1969) Measurable dependance of convex sets and functions on parameters.
J. Math. Anal. Appl. 28, p. 4-25.

VALADIER, M. (1970) Contribution à l'analyse convexe.
Thèse Paris (1970).

VAN CUTSEM, B. (1971) Eléments aléatoires à valeurs convexes compactes
Thèse Grenoble (1971).

WALD, A. (1949) on the consistency of the maximum likelihood estimate.
Ann. Math. Stat. 20, p. 595-601