Annales scientifiques de l'Université de Clermont-Ferrand 2
Série Mathématiques

Xavier Milhaud

Structures statistiques asymptotiquement différentiables

Annales scientifiques de l'Université de Clermont-Ferrand 2, tome 65, série *Mathématiques*, nº 15 (1977), p. 103-107

<http://www.numdam.org/item?id=ASCFM_1977__65_15_103_0>

STRUCTURES STATISTIQUES ASYMPTOTIQUEMENT DIFFERENTIABLES

Xavier MILHAUD

Université Paul Sabatier, Toulouse

INTRODUCTION

Partant de "Théorie asymptotique de la décision statistique" de Lucien Lecam [1] et de "Contiguity of probability measures : somes applications in statistics" de George, G. Roussa [3] l'auteur en [2] expose la théorie des structures statistiques asymptotiquement différentiables dans un cadre suffisamment général et applique cette théorie aux chaînes de Markov stationnaires et non stationnaires. Un des outils de base est la contiguïté. Ici nous n'aborderons pas les chaînes de Markov.

CONTIGUITE

Soit $(\Omega_n, \mathcal{A}_n)_{n \in \mathbb{N}}$ une suite d'espaces probabilisables. Pour chaque n entier naturel P_n et P'_n désigneront deux probabilités sur $(\Omega_n, \mathcal{A}_n)$. Les suites $(P_n)_{n \in \mathbb{N}}$ et $(P'_n)_{n \in \mathbb{N}}$ sont dites contiguës si :

"$\lim_n P_n(A_n) = 0$ équivaut à $\lim_n P'_n(A_n) = 0$"

Remarque : Pour voir la force de cette notion remarquons que si de plus pour tout $n \in \mathbb{N}$ P_n est équivalente à P'_n alors la contiguïté est équivalente à l'uniforme intégrabilité des suites :

$$\left(\frac{d\,P_n}{d\,P'_n}, P'_n\right)_{n \in \mathbb{N}} \quad \text{et} \quad \left(\frac{d\,P'_n}{d\,P_n}, P_n\right)_{n \in \mathbb{N}}.$$

Dans [3] de nombreux théorèmes peuvent se démontrer par des arguments d'uniforme intégrabilité.

SUITES DE STRUCTURES STATISTIQUES ASYMPTOTIQUEMENT DIFFERENTIABLES

Soit k un entier naturel. Θ un ouvert de $\mathbb{R}^k$. $<.,.>$ désigne le produit scalaire dans $\mathbb{R}^k$.

$\mathcal{S} = (\Omega_n, \mathcal{A}_n, P_{n,\theta}\ \theta \in \Theta)_{n \in \mathbb{N}}$ est une suite de structures statistiques indicées par le même espace paramétrique Θ.

On suppose que pour chaque $n \in \mathbb{N}$ et tout couple θ et $\theta' \in \Theta$ $P_{n,\theta}$ et $P_{n,\theta'}$ sont équivalentes.

$\mathcal{C}$ désigne la tribu des boréliens de Θ.

On suppose que pour chaque $n \in \mathbb{N}$ et chaque $B \in \mathcal{Q}_n$ la fonction de θ $P_{n,\theta}(B)$ est mesurable de $(\Theta, \mathcal{C})$ dans $(\mathbb{R}, \mathcal{B})$. On note :

$$\Lambda_n(\theta,\theta') = \text{Log} \frac{d\, P_{n,\theta'}}{d\, P_{n,\theta}} \ .$$

Soit $(\delta_n,\ n \in \mathbb{N})$ une suite de réels de limite nulle.

Nous dirons qu'un couple $(\theta,\ t)$ d'éléments de $\Theta \times \mathbb{R}^k$ vérifie la condition C si pour tout $n \in \mathbb{N}$ $\theta + \delta_n t \in \Theta$.

Définition :

La suite $\mathcal{S} = (\Omega_n, \mathcal{a}_n, P_{n,\theta}\ \ \theta \in \Theta)_{n \in \mathbb{N}}$ est dite asymptotiquement différentiable sur Θ si :

a) pour tout couple $(\theta,\ t)$ vérifiant la condition C les suites $(P_{n,\theta})_{n \in \mathbb{N}}$ $(P_{n,\theta+\delta_n t})_{n \in \mathbb{N}}$ sont contiguës.

b) pour tout couple $(\theta,\ t)$ vérifiant la condition C il existe une fonction mesurable de $(\Omega_n, \mathcal{a}_n)$ dans $(\mathbb{R}^k, \mathcal{B}_k)$ notée $\Delta_n(\theta)$ et une fonction $A(\theta,.)$ de $\mathbb{R}^k$ dans $\mathbb{R}_+$ telles que on ait :

$$\Lambda_n(\theta,\theta+\delta_n t) - \langle t,\ \Delta_n(\theta)\rangle + A(\theta,\ t) \to 0 \qquad P_{n,\theta}$$

c) pour tout couple $(\theta,\ t)$ vérifiant la condition C et toute suite $(t_n,\ n \in \mathbb{N})$ d'éléments de $\mathbb{R}^k$ telle que pour tout $n \in \mathbb{N}$ $\theta + \delta_n t_n$ appartienne à Θ

$$\Lambda_n(\theta,\theta+\delta_n t_n) - \Lambda_n(\theta,\theta+\delta_n t) \to 0 \qquad P_{n,\theta} \ .$$

d) pour tout $\theta \in \Theta$ la fonction de t $A(\theta,t)$ est une forme quadratique définie positive sur $\mathbb{R}^k$. On pose :

$$A(\theta,\ t) = \frac{1}{2} \langle \Gamma(\theta)\, t, t\rangle$$

e) pour tout $\theta \in \Theta$ la suite des lois $(\mathcal{L}(\Delta_n(\theta) \mid P_{n,\theta}),\ n \in \mathbb{N})$ converge étroitement vers la loi normale centrée $N(0,\Gamma(\theta))$ de covariance $\Gamma(\theta)$.

Remarque : En [1] L. Lecam propose une définition plus générale. La définition que nous donnons ici s'applique bien aux chaînes de Markov.

Conséquences :

1° La condition b) implique :

$$\lim_n \mathcal{L}(\Lambda_n(\theta,\theta+\delta_n t) \mid P_{n,\theta}) = N(-\frac{1}{2} \langle\Gamma(\theta)\, t,t\rangle\ ,\ \langle\Gamma(\theta)\, t,t\rangle)$$

2° La condition a), la contiguité des suites $(P_{n,\theta})_{n \in \mathbb{N}}$ et $(P_{n,\theta + \delta_n t})_{n \in \mathbb{N}}$ implique :

$$\lim_n \mathcal{L}(\Lambda_n(\theta, \theta+\delta_n t) \mid P_{n,\theta+\delta_n t}) = N(\tfrac{1}{2} <\Gamma(\theta)t,t> , <\Gamma(\theta)\, t,t>)$$

$$\lim_n \ (\Delta_n(\theta) \mid P_{n,\theta + \delta_n t}) = N(\Gamma(\theta)\, t, \Gamma(\theta)).$$

APPROXIMATION PAR UNE FAMILLE EXPONENTIELLE

Gardons les notations des paragraphes précédents. Si B est une partie compacte de $\mathbb{R}^k$, θ un point de Θ, δ un scalaire, désignons par $\theta + \delta B$ l'ensemble des points $\theta + \delta t$ de $\mathbb{R}^k$ tels que t appartienne à B.

θ étant maintenant fixé dans $\mathbb{R}^k$ choisissons B compact tel que pour tout entier n $\theta + \delta_n B$ soit inclu dans Θ.

Les conditions de différentiabilité asymptotiques nous permettent de construire pour chaque entier n une famille exponentielle de probabilité $(R_{n,t}, t \in \mathbb{R}^k)$ sur $(\Omega_n, \mathcal{Q}_n)$ indicée par $\mathbb{R}^k$ telle que :

$$\lim_n \ \sup_{t \in B} |||P_{n, \theta + \delta_n t} - R_{n,t}||| = 0$$

où $|||\mu|||$ désigne la variation totale de la mesure signée μ .

Idée de la construction de $R_{n,t}$

La différentiabilité asymptotique de la suite de structure statistique $\mathcal{J}$ implique :

$$\Lambda_n(\theta, \theta + \delta_n t) - t\, \Delta_n(\theta) + \frac{1}{2} <\Gamma(\theta)\, t,t> \to 0 \quad P_{n,\theta}$$

En remarquant que :

$$\frac{d P_{n,\theta + \delta_n t}}{d P_{n,\theta}} = \exp \Lambda_n(\theta, \theta + \delta_n t)$$

nous aimerions que pour tout $n \in \mathbb{N}$ $t \in \mathbb{R}^k$

$$\exp \left[<t, \Delta_n(\theta)> - \frac{1}{2} <\Gamma(\theta)\, t, t> \right]$$

soit une suite de densité de probabilité par rapport à $P_{n,\theta}$ et que :

$$\lim_{n\to\infty} \sup_{t \in B} \int_{\Omega_n} \left| \exp \Lambda_n(\theta, \theta+ \delta_n t) - \exp <t,\Delta_n(\theta)> - \frac{1}{2} <\Gamma(\theta)t,t> \right| dP_{n,\theta}$$

soit nulle. En effet nous pourrions alors poser :

$$d R_{n,t} = \exp <t, \Delta_n(\theta)> - \frac{1}{2} <\Gamma(\theta)t,t> \quad d P_{n,\theta}$$

Malheureusement il n'est pas toujours vrai que exp $<t, \Delta_n(\theta)>$ soit intégrable quelque soit t pour $P_{n,\theta}$.

On tourne la difficulté en montrant que l'on peut prendre pour tout $n \in \mathbb{N}$ et tout $t \in \mathbb{R}^k$

$$d\, R_{n,t} = \exp \{<t, \Delta_n^*(\theta)> - B_n(t)\} \; d\, P_{n,\theta}$$

où

$$\lim_n \; P_{n,\theta} \; (\Delta_n^*(\theta) \neq \Delta_n(\theta)) = 0$$

$$\exp\,(B_n(t)) = \int_{\Omega_n} \exp \; <t, \Delta_n^*(\theta)> \; dP_{n,\theta}$$

Pour la construction voir [1] p.70 et [3] p.67.

EXHAUSTIVITE ASYMPTOTIQUE

Pour chaque entier n la statistique $\Delta_n^*(\theta)$ est exhaustive pour la famille exponentielle de probabilités $(R_{n,t}, \; t \in \mathbb{R}^k)$. On démontre alors que $\Delta_n(\theta)$ est asymptotiquement localement exhaustive en θ.

On montre aussi que quand il existe une estimation préliminaire $(T_n)_{n \in \mathbb{N}}$ telle que pour tout ε réel positif tout n entier il existe $b(\varepsilon)$ tel que :

$$P_{n,\theta} \; (||\delta_n, \; T_n-\theta)|| \; > \; b(\varepsilon)) \; < \varepsilon$$

on peut construire une tribu asymptotiquement exhaustive.

PROBLEMES DE TESTS

Nous supposons ici que l'espace paramétrique Θ est inclu dans $\mathbb{R}$. Si la suite de structures statistiques est asymptotiquement différentiable on saura construire des suites de tests asymptotiquement optimaux grace à l'approximation exponentielle.

ESTIMATION

Supposons que pour chaque $n \in \mathbb{N}$ T_n soit un estimateur du paramètre θ et supposons que la suite de lois

$$(\mathcal{L}(\frac{T_n-\theta}{\delta_n} \;|\; P_{n,\theta}), \quad n \in \mathbb{N})$$

admette une limite pour la convergence en loi : $\mathcal{L}(\theta)$ et cela pour tout $\theta \in \Theta$. Alors pour presque tout θ dans Θ $\mathcal{L}(\theta)$ se factorise au sens de la convolution :

$$\mathcal{L}(\theta) = \mathcal{L}_1(\theta) * \mathcal{L}_2(\theta)$$

$$\text{où} \quad \mathcal{L}_1(\theta) = N(0, \Gamma^{-1}(\theta)).$$

BIBILIOGRAPHIE

[1] Lucien LECAM

Théorie asymptotique de la décision statistique
Presses de l'Université de Montréal 1968.

[2] Xavier MILHAUD

Structures statistiques asymptotiquement différentiables,application aux chaînes de Markov (à paraître).

[3] George. G. ROUSSAS

Contiguity of probability measures : some applications in statistics

Cambridge Tracts in Mathematics and Mathematical Physics, N° 63
Cambridge University Press, London-New York, 1972