ANNALES DE L'I. H. P., SECTION B

DIDIER DACUNHA-CASTELLE
FABRICE GAMBOA

Maximum d'entropie et problème des moments

Annales de l'I. H. P., section B, tome 26, nº 4 (1990), p. 567-596
<http://www.numdam.org/item?id=AIHPB_1990__26_4_567_0>

Ann. Inst. Henri Poincaré,

Vol. 26, n° 4, 1990, p. 567-596. *Probabilités et Statistiques*

Maximum d'entropie et problème des moments

par

Didier DACUNHA-CASTELLE et Fabrice GAMBOA

Université de Paris-Sud, U.R.A. D0743, C.N.R.S.,
Statistique appliquée, Mathématiques, bât. n° 425, 91405 Orsay Cedex, France

RÉSUMÉ. — Nous présentons une extension du principe d'entropie maximum (principe d'entropie maximum pour la moyenne). Ceci permet de prendre en compte des contraintes non linéaires dans la problématique, donne de nouveaux résultats pour les problèmes de moments et une vue plus unifiée des différents principes d'optimisation (moindres carrés, entropie usuelle, entropie de la prédiction, etc.).

Mots clés : Maximum d'entropie, discrétisation, moments, espaces fonctionnels, analyse convexe, procédés sommatoires.

ABSTRACT. — We give an extension of the so-called maximum entropy principle in order to extend its application to non linear situation. We call the new principle maximum entropy for the mean. It gives also new results for classical moments problem and also a more unified insight towards the different principles of optimization as least squares, usual entropy, entropy of prediction, and so on.

Classification A.M.S. : 62 A 99, 60 A 99, 41 00.

I. INTRODUCTION

La méthode du maximum d'entropie (ME) est utilisée pour reconstruire une densité g à partir d'informations incomplètes, essentiellement de moments. Elle est fondée sur le choix d'une mesure μ dite de référence (choix fait en physique à partir de principes d'invariance). L'entropie de g est $S(g) = -\int_E \log g(x) g(x)\, dx$ lorsque E est $\mathbb{R}^k$ ou $\mathbb{T}^k$ et μ est égale à dx. Si P est une probabilité et μ une mesure positive quelconque sur $(\Omega, \mathscr{A})$ espace mesurable, on pose $S_\mu(P) = -\int_\Omega \log \frac{dP}{d\mu} dP$ et $K(P, \mu) = -S_\mu(P)$; lorsque μ est une probabilité, K est l'information de Kullback. Rappelons que (ME) reconstruit g par le principe métamathématique,

$$g^{ME} = \arg\max_{g \in \mathscr{C}} S_\mu(g), \tag{0}$$

(arg signifiant: le point où le maximum supposé unique est atteint).

Notre but est d'étendre ME lorsque l'on impose à g d'être dans un ensemble $\mathscr{C}$ (qui sera précisé plus tard), $\mathscr{C}$ étant défini par contraintes ne relevant pas de ME (*i.e.* ME ne donne en général ni méthode, ni solution).

Examinons de plus près ces contraintes. La littérature ([1], [2]) traite des contraintes de moment du type

$$\int_\Omega g \Phi\, d\mu = c \qquad \text{où} \quad \Phi : \Omega \to \mathbb{R}^k \qquad \text{et} \qquad c \in \mathbb{R}^k \text{ fixé.} \tag{1}$$

Lorsque ce problème est remplacé par

$$\int_\Omega g \Phi\, d\mu \in K, \qquad K \text{ compact de } \mathbb{R}^k \tag{2}$$

nous parlerons de contraintes relaxées. Outre cette généralisation qui relève d'une technique minimax [6], on peut traiter de nouvelles contraintes que nous situons maintenant sur des exemples.

Plaçons-nous dans le cas où $\Omega = [0, 1]$ et considérons d'abord des contraintes d'interpolation du type $g(x_i) = d_i$, $i = 1 \ldots k$ avec la contrainte supplémentaire g continue. La forme linéaire $\delta_{x_i}(g)$ est trop « locale » pour être traitée par (ME) pour la mesure de Lebesgue. Une méthode naturelle est de remplacer la mesure de Lebesgue par la loi Q d'un processus $[X_x, x \in (0,1)]$. Les trajectoires ne satisferont pas p. s. la contrainte mais la réaliseront en moyenne pour une probabilité Q^{ME}, choisie suivant le principe d'entropie, la plus voisine possible de Q. Si par le choix de Q, l'on impose aux trajectoires d'autres conditions, on pourra traiter de nouvelles contraintes.

On peut aussi chercher g comme une dérivée, Q étant la loi d'une mesure aléatoire dX_x (dérivée de X_x au sens des distributions) et la contrainte sera alors réalisée par la dérivée de la moyenne de X_x.

Par exemple, pour un processus de loi Q, les contraintes s'écriront $E_{Q^{ME}}\left[\int \Phi(x) X_x\,dx\right]=c$, $E_{Q^{ME}} X_{x_i}=d_i$ et la fonction cherchée sera $E_{Q^{ME}} X_x$; pour une mesure aléatoire, elles s'écriront $E_{Q^{ME}}\left[\int \Phi(x)\,dX_x\right]$ et la fonction $g^{MEM}(x)$ cherchée sera définie par

$$E_{Q^{ME}}[dX_x]=g^{MEM}(x)\,dx.$$

Nous appellerons méthode du maximum d'entropie pour la moyenne et noterons MEM cette construction, dont ME est un cas particulier.

Elle se formule ainsi généralement. Soit B un E.V.T.L.C., $\mathscr{C}$ un convexe de B, X une v. a. à valeurs dans B de loi Q, ν une probabilité de référence sur B, $\Phi\in(B^*)^k$, $c\in\mathbb{R}^k$. MEM consiste à minimiser $K(Q,\nu)$ sous la contrainte $E_Q(\langle\Phi, X\rangle|=c$.

Soit μ une mesure de référence, $\mathscr{G}$ le groupe des automorphismes laissant μ invariante et $\mathscr{G}_1$ un sous groupe de $\mathscr{G}$. On s'intéresse aux constructions invariantes par $\mathscr{G}$ ou $\mathscr{G}_1$. Par exemple, pour μ mesure de Lebesgue sur [0, 1], la $\mathscr{G}$-invariance amène à travailler avec des processus à accroissements échangeables et leurs dérivées.

Dans le cas de $\mathbb{R}$, les bruits blancs stationnaires (BBS) joueront le rôle central; la partie II est consacrée à ces situations.

Les parties III et IV sont consacrées à l'introduction de contraintes non linéaires. Prenons un exemple important qui permet des calculs explicites.

$$g \text{ continue}, \qquad a<g(x)<b \qquad \text{pour} \quad x\in[0,1] \tag{3}$$

et

$$\int_0^1 \varphi(x)g(x)\,dx=c\in\mathbb{R}^k \quad \text{avec } \varphi \text{ continue de } [0,1] \text{ dans } \mathbb{R}^k.$$

Là encore, (ME) ne peut prendre en compte naturellement le non linéaire. Heuristiquement, on cherche une construction MEM $\mathscr{G}$-invariante, c'est-à-dire une probabilité $\mathscr{G}$-invariante portée par le convexe de C([0,1]) des fonctions vérifiant (3). Les BBS ne sont pas de bons candidats du fait qu'ils sont à trajectoires non bornées.

Nous proposons donc de définir (MEM) par une discrétisation (*i.e.* grossièrement par une méthode projective), que l'on peut décrire ainsi sur l'exemple choisi. Soit F une mesure positive dont l'enveloppe convexe du support contient $[a, b]$. Soit $\left\{\frac{i}{n}, i=0 \ldots n-1\right\}$ la discrétisation de $[0, 1]$. Soit $(1.n)$ $\frac{1}{n}\sum_{i=0}^{n-1} \varphi\left(\frac{i}{n}\right) g\left(\frac{i}{n}\right)=c$, la version discrétisée de (1) (avec ici φ pour Φ).

A l'étape n, soit $\tilde{\mathbb{P}}_n=\tilde{\mathbb{P}}_n(\varphi, c)$ l'ensemble des probabilités P_n sur $\mathbb{R}^n$ telles que $\frac{1}{n}\sum_{i=0}^{n-1} \varphi\left(\frac{i}{n}\right) \mathrm{E}_{\mathrm{P}_n}(\mathrm{X}_i^n)=c$ où $(\mathrm{X}_i^n)_{i=0, \ldots, n-1 ; n \in \mathbb{N}}$ est une famille de variables aléatoires (v.a.) de loi P_n. Définissons alors $\mathrm{Q}_n^{\mathrm{MEM}}$ par $\arg\max_{\tilde{\mathbb{P}}_n} \mathrm{S}_{\mathrm{F}^{\otimes n}}(\mathrm{P}_n)$ et enfin $g_n^{\mathrm{MEM}}(x)=\frac{1}{n}\sum_{i=0}^{n-1} \mathrm{E}_{\mathrm{Q}_n^{\mathrm{MEM}}} \mathrm{X}_i^n \; \mathbf{1}_{[i/n, i+1/n[}(x)$. Nous démontrerons que, sous des hypothèses raisonnables, $g_n^{\mathrm{MEM}}(x)$ tend vers $g^{\mathrm{MEM}}(x)$ solution de (MEM). En fait pour tout x

$$g^{\mathrm{MEM}}(x)=\lim_{n \to \infty}\left\{\mathrm{E}_{\mathrm{Q}_n^{\mathrm{MEM}}} \mathrm{X}_i^n, \text{ avec } \frac{i}{n} \leqq x<\frac{i+1}{n}\right\}.$$

Les parties suivantes seront plus analytiques. Le problème MEM pour la bande $a<g(x)<b$ est lié à un problème d'optimisation d'une fonctionnelle concave. Ce qui permet de lier MEM et des problèmes usuels type moindres carrés, entropie de Burg, etc.

Les techniques simples de grandes déviations donnent un éclairage nouveau sur le problème des moments dit de Markov (traité surtout par Krein-Nudelman [11]).

Enfin, le problème d'une suite infinie de moments est traité sous l'angle des méthodes sommatoires par (ME) et (MEM), méthodes qui peuvent avoir un grand intérêt numérique (Gassiat [8]).

II. NOTATIONS, LE PROBLÈME LINÉAIRE

II.1. Notations

$(\Omega, \mathscr{A})$ est un espace mesurable, $\mathbb{M}(\Omega)$ [resp. $\mathbb{P}(\Omega)$] est l'ensemble des mesures positives (resp. des probabilités) sur $(\Omega, \mathscr{A})$.

Pour

$$\mathrm{P}, \mathrm{Q} \in \mathbb{P}(\Omega), \quad \mathrm{K}(\mathrm{P}, \mathrm{Q}) = \mathrm{E}_{\mathrm{P}}\left(\operatorname{Log}\frac{d\mathrm{P}}{d\mathrm{Q}}\right) \quad \text{si } \mathrm{P} \ll \mathrm{Q} \text{ et } \operatorname{Log}\frac{d\mathrm{P}}{d\mathrm{Q}} \in \mathrm{L}^1(\mathrm{P})$$
$$= +\infty \text{ sinon}$$

est l'information de Kullback de P sur Q $(0 \leqq \mathrm{K} \leqq +\infty)$.

Pour des mesures positives bornées μ et ν,

$$\mathrm{K}(\nu, \mu) = \int \operatorname{Log}\frac{d\nu}{d\mu}\, d\nu + \mu(\Omega) - \nu(\Omega).$$

Pour

$$\mu \in \mathbb{M}(\Omega), \qquad \mathrm{P} \in \mathbb{P}(\Omega), \qquad \mathrm{S}_\mu(\mathrm{P}) = -\int \operatorname{Log}\frac{d\mathrm{P}}{d\mu}\, d\mathrm{P}$$

est la μ-entropie de P lorsqu'elle est définie.

Si μ est une probabilité, $\mathrm{S}_\mu(\mathrm{P}) = -\mathrm{K}(\mathrm{P}, \mu)$.

Soit Φ une application mesurable (fixée) de Ω dans $\mathbb{R}^k$.

On notera

$$\mathrm{Z}_{\mu, \Phi}(u) = \int_\Omega \exp(\langle u, \Phi(\omega) \rangle)\, d\mu(\omega), \qquad \mu \in \mathbb{M}(\Omega),$$
$$\mathrm{D}^\Phi(\mu) = \{ u \in \mathbb{R}^k, \mathrm{Z}_{\mu, \Phi}(u) < +\infty \},$$
$$\mathrm{H}_{\mu, \Phi}(u, c) = \operatorname{Log}[\mathrm{Z}_{\mu, \Phi}(u)] - \langle u, c \rangle,$$

et par Φ-arc d'Hellinger de μ, on entendra

$$d\mu_u^\Phi(\omega) = (\mathrm{Z}_{\mu, \Phi}(u))^{-1} \exp(\langle u, \Phi(\omega) \rangle)\, d\mu(\omega).$$

Si $c \in \mathbb{R}^k$ est fixé, on note

$$\mathbb{P}_{c, \Phi} = \tilde{\mathbb{P}} = \{ \mathrm{P} \in \mathbb{P}(\Omega), \mathrm{E}_{\mathrm{P}}[\Phi(\omega)] = c \}.$$

II.2. Le problème linéaire

Les données d'un problème de maximum d'entropie (ME) tel qu'il va être rappelé ci-dessous seront le triplet $(\Phi, c, \mu\,(\text{ou}\,\mathrm{Q}))$. Par problème de ME on entend: trouver P dans $\tilde{\mathbb{P}}$ tel que $\mathrm{S}_\mu(\mathrm{P})$ soit maximum. Le problème est largement résolu par le théorème suivant dû à Csiszar (*cf.* [1]).

THÉORÈME 1. — *Soit* $\mathrm{Q} \in \mathbb{P}(\Omega)$. *Si* $\mathrm{D}^\Phi(\mathrm{Q})$ *est ouvert, et si* $\tilde{\mathbb{P}}$ *contient au moins une probabilité* P *équivalente à* Q, *alors il existe une unique probabilité* $\mathrm{Q}_{u_c}^\Phi$ *de* $\tilde{\mathbb{P}}$ *telle que*:

$$\mathrm{Q}_{u_c}^\Phi = \arg \inf_{\mathrm{P} \in \tilde{\mathbb{P}}} \mathrm{K}(\mathrm{P}, \mathrm{Q}),$$

$Q^{\Phi}_{u_c}$ *est l'élément de l'arc d'Hellinger de* Q *défini par*

$$u_c = [\operatorname{grad}(\operatorname{Log}[Z_{Q,\Phi}])]^{-1}(c).$$

De plus

$$K(Q^{\Phi}_{u_c}, Q) = -\inf_u H_{\mu,\Phi}(u, c) = -H_{\mu,\Phi}(u_c, c).$$

COROLLAIRE 1. – *Soit* $\mu \in \mathbb{M}(\Omega)$, *on suppose que* $D^{\Phi}(\mu)$ *est ouvert non vide, alors on a un énoncé identique au théorème* 1 *en remplaçant* Q *par* μ *et* $K(P, Q)$ *par* $-S_{\mu}(P)$.

DÉFINITION 1. – On dira que c est Φ-réalisable par Q si il existe $P \in \tilde{\mathbb{P}}$ avec P équivalente à Q.

Notation. – Dans la suite lorsque $Q^{\Phi}_{u_c}$ sera définie on la notera plus simplement Q^{ME}.

Considérons d'abord la technique usuelle ME pour reconstruire g.

Considérons les contraintes suivantes pour g définie sur $[0, 1]$:

(1) g continue,

(2) $g > 0$.

(3) $\int_0^1 g(x)\,dx = 1$.

(4) $\int_0^1 \varphi(x) g(x)\,dx = c$ où $\varphi \in C[[0, 1], \mathbb{R}^k]$ et $c \in \mathbb{R}^k$.

(4') $g(x_i) = d_i$, $i = 1 \ldots k$ avec $x_i \in [0, 1]$ et $d_i \in \mathbb{R}$.

On s'intéressera particulièrement aux reconstructions invariantes par $\mathcal{G}$, groupe des réarrangements de $[0, 1]$ laissant la mesure de Lebesgue invariante. Sous les contraintes 1, 2, 3, 4, si c est φ-réalisable pour dx le théorème 1 donne la solution ME

$$g^{ME}(x) = \frac{\exp\langle u_c, \varphi(x)\rangle}{Z_{dx,\varphi}(u_c)}$$

et

$$\int_0^1 \varphi(x) g^{ME}(x)\,dx = c.$$

Si l'on remplace (4) par (4'), le problème ME est sans solution dans $C(0, 1)$.

Nous allons introduire dans le cas linéaire la méthode MEM, à partir d'exemples.

Exemple 1. – Soit $(X_x)_{x \in [0, 1]}$ un processus à accroissements échangeables [10], c'est-à-dire pour tout $k \in \mathbb{N}$, tout choix $I_1, \ldots, I_k$ d'intervalles disjoints de même mesure, si $X_I = X_b - X_a$ pour $I = [a, b[$, alors $(X_{I_1}, \ldots, X_{I_k})$ sont des v. a. échangeables.

Les processus positifs échangeables (distincts de la translation) forment un convexe ayant pour points extrémaux, les processus dont les dérivées (au sens des distributions) sont les mesures aléatoires

$$dU = \sum_{n=0}^{\infty} \alpha_n \delta_{U_n},$$

où (α_n) est une suite positive de somme 1 et (U_n) une suite de v. a. uniformes sur [0, 1] et indépendantes [10]. Soit Q la loi d'un tel processus. Parmi les processus de loi absolument continue par rapport à Q on a les processus de dérivée

$$dV = \sum_{n=0}^{\infty} \alpha_n \delta_{V_n}, \qquad V_n \text{ indépendantes de loi } f(x)\,dx.$$

Soit $\Phi(\omega) = \int \varphi(x)\, dU_x(\omega)$ et $\tilde{P} = \mathbb{P}_{c, \Phi} = \{P, P \ll Q, E_Q(\Phi) = c\}$.

Commençons par le cas $\alpha_0 = 1$, $\alpha_i = 0$ pour $i > 0$. Supposons c réalisable par la densité g pour dx, g continue. Soit V_0 de densité $g(x)\,dx$. Alors la mesure aléatoire δ_{V_0} est équivalente à δ_{U_0}. Le théorème de Csiszar donne Q^{ME} avec $E_{Q^{ME}}[\Phi] = c$. Et donc on obtient comme réalisation MEM.

$$g^{MEM}(x)\,dx = E_{Q^{ME}}(\delta_{V_0}).$$

Mais $\dfrac{dQ^{ME}}{dQ}(x) = Z_{dx, \varphi}^{-1}(u_c)\, e^{\langle u_c, \varphi(x) \rangle}$, et donc ME et MEM ont la même solution.

Plus généralement si

$$dU = \sum_{n=0}^{\infty} \alpha_n \delta_{U_n}, \qquad \tilde{x} = \{(x_n)_{n \in \mathbb{N}}\}, \qquad x_n \in [0, 1].$$

Alors

$$\frac{dQ^{ME}(\tilde{x})}{dQ} = \frac{e^{\langle u_c, \sum_0^{\infty} \alpha_n \varphi(x_n) \rangle}}{\prod_{n=0}^{\infty} Z_{dx, \varphi}(u_c \alpha_n)}$$

et de

$$c=\int_{[0,1]^{\mathbb{N}}}\sum_{n=0}^{\infty}\alpha_n\varphi(x_n)\frac{e^{\langle u_c,\sum_0^\infty \alpha_n\varphi(x_n)\rangle}}{\prod_{n=0}^{\infty}Z_{dx,\varphi}(u_c,\alpha_n)}d\tilde{x}.$$

(on donne immédiatement un sens à cette notation).

On tire

$$g^{\mathrm{MEM}}(x)=\sum_{n=0}^{\infty}\alpha_n\frac{e^{\langle u_c,\alpha_n\varphi(x)\rangle}}{Z_{dx,\varphi}(u_c,\alpha_n)}.$$

– Soit $dU_{(n)}$ correspondant à $\alpha_0=\alpha_1=\ldots=\alpha_n=\dfrac{1}{n+1}$ et $\alpha_i=0$, $i>n$. On trouve encore la solution ME pour Q associée à $dU_{(n)}$.

– Soit dN la mesure de Poisson (dérivée du processus de Poisson). Alors

$$dN=\sum_{n=0}^{\infty}e^{-1}\frac{dU_{(n)}}{n!}$$

et donc dN pris comme mesure de référence Q donne encore la solution ME dans le problème MEM.

Dans le cas où $[0,1]$ est remplacé par $\mathbb{R}^+$ (ou $\mathbb{R}$), dN garde cette propriété (*voir* ci-dessous).

Exemple 2. – Considérons un bruit $\mathbb{B}$ blanc stationnaire sur $\mathbb{R}$, *i.e.* une application $A\to\mathbb{B}(A)$ des boréliens dans l'ensemble des v.a. telle que $\mathbb{B}(A_1)\ldots\mathbb{B}(A_n)$ soient indépendantes si $A_1\ldots A_n$ sont disjoints et que la loi Q de $\mathbb{B}$ soit invariante sous $\mathscr{G}$. Donnons une introduction formelle, sans étudier les problèmes d'existence. Remarquons que les bruits blancs sont associés aux points extrémaux des processus échangeables sur $\mathbb{R}^+$ et généralisent en ce sens l'exemple 1 [10].

Soit

$$\begin{aligned}\tilde{\mathbb{P}}&=\mathbb{P}_{c,\Phi}\\&=\{P\ll Q, E_P(\Phi)=c\}\end{aligned}$$

où $\Phi(\omega)=\int\varphi(t)\,d\mathbb{B}_t(\omega)$. Supposons c Φ-réalisable pour Q.

Le théorème 1 donne

$$Q^{\mathrm{ME}}=\arg\min_{\tilde{\mathbb{P}}}K(P,Q)$$

$$\frac{dQ^{ME}}{dQ} = Z_{Q,\Phi}^{-1}(u_c) \exp \langle u_c, \Phi \rangle$$

où u_c est calculée par $E_{Q^{ME}}(\Phi) = c$. Toujours formellement

$$E_{Q^{ME}}\left(\int \varphi(t)\, d\mathbb{B}_t\right) = \int \varphi(t)\, E_{Q^{ME}}\, d\mathbb{B}_t.$$

Posant $E_{Q^{ME}}\, d\mathbb{B}(t) = g^{MEM}(t)\, dt$, on voit que g^{MEM} est solution du problème. Nous verrons (problème d'interpolation) que, si Φ n'est pas continue, en général g^{MEM} ne vérifie pas (1) et que, si $d\mathbb{B}$ n'est pas p.s. une mesure positive, g^{MEM} ne vérifie pas (2).

Justifions ce formalisme sur les exemples suivants :

Exemple A. – $d\mathbb{B}$ est le bruit poissonnien dN.

$$\frac{dQ^{ME}}{dN} = Z_{N,\Phi}^{-1}(u_c) \exp\left[\int_0^1 \langle u_c, \varphi(t) \rangle\, dN_t\right].$$

La densité d'un bruit poissonnien d'intensité g par rapport au bruit homogène [4] est donnée par $\exp\left[\int_0^1 \log g(u)\, dN(u) + \int_0^1 (g(u) - 1)\, du\right]$ donc Q^{ME} est le bruit poissonnien d'intensité

$$g^{MEM}(t) = Z_{dx,\varphi}^{-1}(u_c) \exp \langle u_c, \varphi(t) \rangle.$$

Le problème MEM *de référence un bruit poissonnien donne la même solution que le problème* ME *de référence la mesure de Lebesgue.*

Exemple B. – $d\mathbb{B}$ est le bruit brownien dW:

$$\frac{dQ^{ME}}{dW} = Z_{W,\Phi}^{-1}(u_c) \exp\left\{\int_0^1 \langle u_c, \varphi(t) \rangle\, dW(t)\right\}.$$

Le théorème de Girsanov [4] montre que Q^{ME} est la loi de la dérivée du mouvement brownien de dérive $\langle u_c, \varphi(t) \rangle$, *i.e.* sous Q^{ME} le processus X_t canonique définie sur $C[0, 1]$ vérifie

$$dX_t = \langle u_c, \varphi(t) \rangle\, dt + dW_t$$

donc $g^{MEM}(t) = \langle u_c, \varphi(t) \rangle$ avec $\int_0^1 \varphi(t) \langle u_c, \varphi(t) \rangle\, dt = c$.

Pour W, MEM *donne la même solution que la méthode des moindres carrés.*

Remarque. – Pour un bruit blanc de loi $\gamma(\alpha, p)$. si on impose (3), on trouve que MEM est équivalent à une entropie dite de Burg, pour $\alpha=1$, $\beta=0$, par exemple $-\int \log g(x)\,dx$ doit être minimisée sous contrainte.

On reverra en détail ces aspects dans la partie IV.

Revenons par exemple au cas de W mais pour la contrainte (4′), on souhaite écrire

$$E_P W_{t_i}=d_i, \qquad i=1\ldots k.$$

Soit

$$E_P\left[\int_0^1 \mathbf{1}_{(0,\,t_i)}(t)\,dW_t\right]=d_i.$$

La solution de MEM s'écrit

$$\int_0^t \sum_{i=1}^{k} \lambda_i \mathbf{1}_{(0,\,t_i)}(u)\,du.$$

Si l'on veut une solution m fois différentiable, il faut prendre pour référence le processus $X_t=\int_0^t (t-u)^m\,dW_u$, et on obtient alors *la méthode d'interpolation par fonctions splines.*

III. LE CAS NON LINÉAIRE

Nous nous limiterons à la situation suivante :

Les contraintes non linéaires sur la fonction à reconstruire g seront du type $g\in\mathscr{C}$, où $\mathscr{C}$ est un convexe ouvert de l'espace C([0, 1]) des fonctions continues. Nous traiterons en détail deux cas particuliers :

$$\mathscr{C}_T=\{g,\ -\infty\leqq a<g(x)<b\leqq+\infty\},$$
$$\mathscr{C}_\theta=\left\{g,\ \int g^2(x)\,dx<1\right\}.$$

La contrainte linéaire sera du type

$$\int_0^1 \varphi(x)\,g(x)\,dx=c.$$

avec φ application continue de [0, 1] vers $\mathbb{R}^k$.

Comme nous l'avons vu en introduction, les méthodes d'entropie sont fondées sur une mesure de référence μ sur $\mathscr{C}$.

S'il y a des contraintes non linéaires, si l'on exige de μ une propriété d'invariance du type de celles évoquées au paragraphe précédent, on arrive à une impossibilité: il n'existe pas sur $\mathscr{C}$ de telle mesure. C'est le cas, par exemple du tore muni du groupe des translations, et du groupe des transformations orthogonales pour $\mathscr{C}_\theta$. Pour respecter cette exigence d'invariance, nous allons employer une procédure d'approximation par des problèmes de dimension finie qui ne définira pas en général une mesure sur $C([0, 1])$ ou sur un espace fonctionnel utile. Précisément, soit $\{i/n, i=0\ldots n-1\}$ la discrétisation de $[0,1]$. Soient h_n la trace d'une fonction h, $h_n=\{h(i/n),\ i=0\ldots n-1\}$ et $\mathscr{C}_n$ la trace de $\mathscr{C}$. Pour chaque n, on va définir une mesure de référence μ_n, portée par $\mathscr{C}_n$ et satisfaisant aux propriétés d'invariance pour les traces des groupes considérés [par exemple $(\mathscr{C}_T)_n$ sera le cube $]a, b[^n$ et $\mu_n=F^{\otimes n}$, F une mesure positive sur $]a, b[$: $(\mathscr{C}_\theta)_n$ sera B_n la boule unité de $\mathbb{R}^n$ et μ_n sera la loi uniforme sur la boule].

Les contraintes non linéaires seront donc satisfaites avec les invariances souhaitables par les suites $(\mathscr{C}_n, \mu_n)$, $\mathscr{C}_n$ convexe ouvert de $\mathbb{R}^n$. La contrainte linéaire sera discrétisée en $1/n\ E_{P_n}\left[\sum_{i=1}^{n} \varphi(i/n)X_i^n\right]=c$, où $(X_1^n, \ldots, X_n^n)$ sont les n coordonnées; et Q_n^{ME}, si elle existe sera la solution du problème d'entropie associé à μ_n, pour la contrainte discrétisée.

DÉFINITION 2. — Soit $(\mathscr{C}, (\mathscr{C}_n)_{n\in\mathbb{N}}, (\mu_n)_{n\in\mathbb{N}}, \varphi, c)$, la donnée d'une suite de problèmes d'entropie sur la moyenne obtenus par discrétisation à partir d'un convexe ouvert de $C([0, 1])$. On dira que la suite est convergente (pour $n\to\infty$) si:

(1) pour tout n assez grand, le problème (MEM) de taille n admet une solution unique Q_n^{ME};

(2) pour tout x, $E_{Q_n^{ME}} X_i^n \to g^{MEM}(x)$ lorsque $n\to\infty$, $i(n)/n\to x$, X_i^n étant la i-ème coordonnée dans le n-ième problème;

(3) $\int_0^1 g^{MEM}(x)\varphi(x)\,dx=c$.

LEMME 1. — *Soit $\mathscr{C}$ un convexe ouvert de* $C([0, 1])$. *On suppose que la contrainte est réalisable c'est-à-dire qu'il existe $g\in\mathscr{C}$ tel que $\int_0^1 g(x)\varphi(x)\,dx=c$. Si de plus φ est de rang k (i.e.; si $\langle\alpha, \varphi(x)\rangle=0$, pour tout $x\in[0, 1]$, alors $\alpha=0$), alors pour n assez grand, la contrainte* *est réalisable dans* $\mathscr{C}_n$, *c'est-à-dire qu'il existe* $\sigma_n \in \mathscr{C}_n$ *telle que* $\frac{1}{n} \sum_{i=1}^{n} \varphi(i/n)\, \sigma_n^i = c.$

Démonstration. — Remarquons que le résultat est presque évident si $k=1$.

Soit $g \in \mathscr{C}$ tel que $\int_0^1 \varphi(x)\, g(x)\, dx = c$ et soit $\mathrm{B}(g, \varepsilon)$ la boule ouverte de centre g et de rayon ε, où ε est tel que $\mathrm{B}(g, \varepsilon) \subset \mathscr{C}$.

Soit $\mathrm{B}_n(g_n, \varepsilon)$ la boule ouverte de $\mathbb{R}^n$ (pour la norme du sup) centrée en g_n et de rayon ε.

Notons pour $\sigma_n \in \mathbb{R}^n$, $\mathrm{T}_n(\sigma_n) = 1/n \sum_{i=1}^{n} \sigma_n^i\, \varphi(i/n)$.

Supposons que pour tout n de $\mathbb{N}$, $c \notin \mathrm{T}_n(\mathrm{B}_n(g_n, \varepsilon))$.

Choisissons σ_n de la forme $\sigma_n^i = g(i/n) + \varepsilon\, \Delta_n^i$, $|\Delta_n^i| \leqq 1$.

Comme $\mathrm{B}_n(g_n, \varepsilon)$ est convexe, il existe un hyperplan

$$\alpha_n = (\alpha_n^1, \ldots, \alpha_n^k), \qquad \sum_{i=1}^{k} (\alpha_n^i)^2 = 1,$$

qui sépare c de $\mathrm{T}_n(\mathrm{B}_n(g_n, \varepsilon))$.

Donc, pour tout choix de Δ_n^i, on a

$$\langle \alpha_n, c \rangle = \int_0^1 \langle \alpha_n, \varphi(x) \rangle\, g(x)\, dx > 1/n \sum_{i=1}^{n} \langle \alpha_n, \varphi(i/n) \rangle\, (g(i/n) + \varepsilon\, \Delta_n^i),$$

et donc

$$\int_0^1 \langle \alpha_n, \varphi(x) \rangle\, g(x)\, dx - 1/n \sum_{i=1}^{n} \langle \alpha_n, \varphi(i/n) \rangle\, g(i/n)$$
$$> 1/n \sum_{i=1}^{n} \langle \alpha_n, \varphi(i/n) \rangle\, \varepsilon\, \Delta_n^i. \qquad (\mathrm{A})$$

Soient $\omega_{\varphi, g}$ le module d'uniforme continuité des composantes de φ et de g et K le sup des normes de ces mêmes fonctions dans $\mathrm{C}([0, 1])$. Le terme de gauche de l'inéquation (A) est majoré par $2\sqrt{k}\, \mathrm{K}\, \omega_{\varphi, g}(1/n)$. Comme le choix des Δ_n^i est arbitraire dans $[-1, 1]$, on doit avoir

$$2\sqrt{k}\, \mathrm{K}\, \omega_{\varphi, g}(1/n) > \varepsilon/n \sum_{i=1}^{n} |\langle \alpha_n, \varphi(i/n) \rangle|.$$

La suite α_n est relativement compacte (sur la sphère unité de $\mathbb{R}^k$). Soit α_∞ un de ses points d'adhérence. On aurait donc

$$0 > \varepsilon \int_0^1 |\langle \alpha_\infty, \varphi(x) \rangle| \, dx, \qquad \text{soit} \langle \alpha_\infty, \varphi \rangle = 0$$

ce qui contredit l'hypothèse faite sur le rang de φ.

LEMME 2. — *Si les hypothèses du lemme* 1 *sont vérifiées, si l'enveloppe convexe du support de* μ_n *contient* $\mathscr{C}_n$, *et si*

$$D(\mu_n) = \left\{ u \in \mathbb{R}^k, \int \exp\left[1/n \sum_{i=1}^{n} \langle u, \varphi(i/n) \rangle x_i^n \right] d\mu_n(x^n) < +\infty \right\}$$

est un ouvert non vide alors le problème d'entropie de taille n admet une solution unique définie par

$$dQ^{ME}(x^n) = [Z_n(u_n)]^{-1} \exp\left[1/n \sum_{i=1}^{n} \langle u_n, \varphi(i/n) \rangle x_i^n \right] d\mu_n(x^n),$$

où

$$x^n = (x_1^n, \ldots, x_n^n), \qquad Z_n(u) = \int_{\mathscr{C}_n} \exp\left[1/n \sum_{i=1}^{n} \langle u, \varphi(i/n) \rangle x_i^n \right] d\mu_n(x^n),$$

et u_n *est défini par* $\operatorname{grad}[\operatorname{Log} Z_n(u_n)] = c$.

Démonstration. — D'après le lemme précédent, il existe σ_n de $\mathscr{C}_n$ tel que, $\frac{1}{n} \sum_{i=1}^{n} \sigma_n^i \varphi(i/n) = c$. Pour pouvoir appliquer le corollaire 1, il suffit de trouver ν_n équivalente à μ_n, de moyenne $E_{\nu_n}[X_i^n] = \sigma_n^i$.

Sous les hypothèses du lemme, il existe pour tout σ_n de $\mathscr{C}_n$ une loi équivalente à μ_n d'espérance σ_n [7], d'où le résultat.

Remarques. — L'hypothèse sur le support de μ_n peut être remplacé par: il existe ν_n équivalente à μ_n et $E_{\nu_n} X_i^n = \sigma_n^i$, où (σ_n^i) est une suite quelconque satisfaisant la contrainte (il est des situations où (σ_n^i) peut être construite indépendamment du lemme 1).

Pour un convexe $\mathscr{C}$ quelconque, il n'y a pas d'approximation privilégiée par des espaces de dimension finie. La formulation la plus naturelle est de choisir pour μ_n la loi uniforme sur $\mathscr{C}_n$. Nous allons étudier les deux cas particuliers $\mathscr{C}_T$ et $\mathscr{C}_\theta$ introduits plus haut. La méthode définie pour $\mathscr{C}_T$ vaut pour toute bande, $-\infty \leqq a < g(x) < b \leqq +\infty$, et non pas seulement pour a et b bornés. De plus on pourrait transformer facilement les hypothèses, pour remplacer a par $a(x)$ et b par $b(x)$ (avec $a(x)$ et $b(x)$ des fonctions continues) et remplacer $C([0, 1])$ par un espace L^∞ (*voir* [7]). La méthode introduite pour $\mathscr{C}_\theta$ est aussi valable pour un ellipsoïde de $L^2([0, 1]) \cap C([0, 1])$.

IV. LE CAS DES CONTRAINTES

$$-\infty \leqq a < g(x) < b \leqq +\infty \qquad \text{et} \qquad \int g^2(x)\,dx \leqq 1.$$

Soit F *une mesure positive sur* $[a, b]$ *qui vérifie les hypothèses :*
(H1) $]a, b[$ *est inclus dans l'enveloppe convexe du support de* F.

(H2) $D(F) = \left\{ t \in \mathbb{R}, \int [\exp(ty)]\, dF(y) < +\infty \right\}$ *est un ouvert non vide.*

(H3) $V = \{ \Lambda \in \mathbb{R}^k, \forall x \in [0, 1], \langle \Lambda, \varphi(x) \rangle \in D(F) \}$ *est non vide, et coïncide avec* $V' = \{ \Lambda \in \mathbb{R}^k, \psi(\langle \Lambda, \varphi(x) \rangle)$ *est dx intégrable sur* $[0, 1] \}$, *où l'on a posé*, $\exp[\psi(t)] = \int [\exp ty]\, dF(y)$.

Ces hypothèses ne sont pas très restrictives. Lorsque a et b sont finis (H1), (H2) et (H3) sont vérifiées si F est une mesure de masse totale finie dont l'enveloppe convexe du support est $[a, b]$. Dans le cas non compact les hypothèses (H2) et (H3) sont vérifiées par exemple lorsque la transformée de Laplace de F est définie sur $\mathbb{R}$ tout entier. H1 entraîne l'existence d'une mesure équivalente à F et d'espérance arbitraire dans $]a, b[$. Remarquons que l'on a toujours V′ inclus dans l'adhérence de V, et que sous (H2), (H3) V′ est un ensemble ouvert et non vide.

$\Lambda_n = n^{-1} u_n$ est défini (lorsqu'il existe et est unique) par

$$\frac{1}{n} \sum_{i=1}^{n} \varphi(i/n)\, \psi'(\langle \Lambda_n, \varphi(i/n) \rangle) = c.$$

Si Q_n^{ME} est la solution du n-ième problème ME associé à la contrainte

$$E_{Q_n^{ME}}\left(n^{-1} \sum_{i=1}^{n} X_i^n \varphi(i/n) \right) = c$$

alors

$$E_{Q_n^{ME}}(X_i^n) = \psi'(\langle \Lambda_n, \varphi(i/n) \rangle).$$

On a

$$\log Z_n(n\Lambda) = \sum_{i=1}^{n} \psi(\langle \Lambda, \varphi(i/n) \rangle).$$

LEMME 2 *bis*. — *Sous les hypothèses* (H1), (H2), (H3), Λ_n *est défini de manière unique, pour n assez grand, par*

$$\frac{1}{n}\sum_{i=1}^{n} \varphi(i/n)\,\psi'(\langle \Lambda_n, \varphi(i/n) \rangle) = c$$

et le n-ième problème ME *associé à la contrainte*

$$E_{P_n}(n^{-1}\sum_{i=1}^{n} X_i^n\,\varphi(i/n)) = c$$

de référence $F^{\otimes n}$, *admet pour solution* Q_n^{ME} *définie par*

$$dQ_n^{ME}(x^n) = Z_n^{-1}(n\Lambda_n)\prod_{1}^{n} \exp\{\langle \Lambda_n, \varphi(i/n) \rangle x_i^n\}\, dF^{\otimes n}.$$

D'autre part

$$\lim_{n\to\infty} 1/n \sum_{i=1}^{n} \varphi(i/n)\,\psi'(\langle \Lambda, \varphi(i/n) \rangle) = \int_0^1 \varphi(x)\,\psi'(\langle \Lambda, \varphi(x) \rangle)\,dx.$$

Il résulte alors du lemme ci-dessous que $\lim_{n\to\infty} \Lambda_n = \Lambda_\infty$.

En résumé. — *Le problème du* (MEM) *pour une contrainte* $a < g(x) < b$ *admet dans ce cadre la solution* $g^{MEM}(x) = \psi'(\langle \Lambda_\infty, \varphi(x) \rangle)$, *où* $\psi(t) = \mathrm{Log}\int \exp(ty)\,dF(y)$, $F^{\otimes n}$ *est la suite des mesures de référence et* Λ_∞ *est donné par*

$$\int_0^1 \varphi(x)\,\psi'(\langle \Lambda_\infty, \varphi(x) \rangle)\,dx = c.$$

LEMME 3. — *Sous les hypothèses* H1, H2, H3 $\lim_{n\to\infty} \Lambda_n = \Lambda_\infty$, *avec* Λ_∞ *donné par*

$$\int_0^1 \varphi(x)\,\psi'(\langle \Lambda_\infty, \varphi(x) \rangle)\,dx = c.$$

Pour prouver le lemme 3 nous avons besoin d'un lemme d'analyse convexe. Soit U un ouvert de $\mathbb{R}^k$ et soit U^* la frontière de U dans

$\mathbb{R}^k \cup \{\infty\}$, nous dirons qu'une suite (Λ_n) $\Lambda_n \in \mathrm{U}$ tend vers U^* et nous noterons $\Lambda_n \to \mathrm{U}^*$, si les points d'accumulation de (Λ_n) sont tous inclus dans U^*.

LEMME 4. — *Soit* $(\Lambda, c) \to \mathrm{K}(\Lambda, c)$ *une fonction de deux variables définie sur* $\mathrm{U} \times \mathrm{W}$, U (resp. W) *est un convexe ouvert de* $\mathbb{R}^k$ (resp. $\mathbb{R}^m$). *On suppose que* $\mathrm{K}(\Lambda, c)$ *est strictement convexe en* Λ, *concave en* c *et que* :

(*a*) *Il existe* c^* *de* W *tel que* $\mathrm{K}(\Lambda, c^*)$ *atteigne son minimum en un point* Λ^* *de* U.

(*b*) *Il existe une fonction* L *à valeur finie sur* W *telle que*

$$\forall \Lambda \in \mathrm{U}, \qquad \mathrm{K}(\Lambda, c) \geqq \mathrm{L}(c).$$

(*c*) *En convenant que* $\mathrm{K}(\Lambda, c) = +\infty$ *si* $\Lambda \notin \mathrm{U}$, *pour* c *de* W *fixé, la fonction* $\mathrm{K}(\,.\,, c)$ *est semi-continue inférieurement.*

Alors $\forall c \in \mathrm{W}$, $\mathrm{K}(\Lambda_n, c) \to +\infty$ *si* $\Lambda_n \to \mathrm{U}^*$.

Preuve du lemme 4. — Soit $c \in \mathrm{W}$, il existe c^{**} de W et un réel λ de $]0, 1[$ tels que $c = \lambda c^* + (1-\lambda) c^{**}$ (car W est un convexe ouvert).

Utilisons maintenant la concavité de K

$$\mathrm{K}(\Lambda, c) \geqq \lambda \mathrm{K}(\Lambda, c^*) + (1-\lambda) \mathrm{K}(\Lambda, c^{**}) \geqq \lambda \mathrm{K}(\Lambda, c^*) + (1-\lambda) \mathrm{L}(c^{**}).$$

On conclut en remarquant que $\mathrm{K}(\Lambda_n, c^*) \to +\infty$ lorsque $\Lambda_n \to \mathrm{U}^*$.

Preuve du lemme 3. — Posons

$$\mathrm{V}_n = \{\Lambda \in \mathbb{R}^k, \langle \Lambda, \varphi(x) \rangle \in \mathrm{D}(\mathrm{F}),\ x = 1/n,\ 2/n, \ldots, 1\},$$

pour c élément de

$$\varphi(\mathscr{C}_\mathrm{T}) = \left\{ c \in \mathbb{R}^k, \int_0^1 \varphi(x) g(x)\, dx, g \in \mathscr{C}_\mathrm{T} \right\}$$

(qui un convexe ouvert de $\mathbb{R}^k$ par application du théorème de l'image ouverte) on pose

$$\begin{aligned} \mathrm{H}_n(\Lambda, c) &= 1/n \sum_{i=1}^{n} \psi(\langle \Lambda, \varphi(i/n) \rangle) - \langle \Lambda, c \rangle, \qquad \text{si} \quad \Lambda \in \mathrm{V}_n \\ &= +\infty, \qquad \text{si} \quad \Lambda \notin \mathrm{V}_n, \end{aligned}$$

et

$$\begin{aligned} \mathrm{H}(\Lambda, c) &= \int_0^1 \psi(\langle \Lambda, \varphi(x) \rangle)\, dx - \langle \Lambda, c \rangle, \qquad \text{si} \quad \Lambda \in \mathrm{V} \\ &= +\infty, \qquad \text{si} \quad \Lambda \notin \mathrm{V}. \end{aligned}$$

On a $\lim_{n \to \infty} \mathrm{H}_n(\Lambda, c) = \mathrm{H}(\Lambda, c)$, la convergence est uniforme sur tout compact de V, donc si l'on montre que $\mathrm{H}(\Lambda, c)$ possède un minimum en un point Λ_∞ de V le lemme sera établi. Pour cela, montrons que les hypothèses du lemme 4 sont vérifiées pour la fonction H.

Soit $\Lambda^* \in V$: posons $c^* = \int_0^1 \varphi(x)\psi'(\langle \Lambda^*, \varphi(x)\rangle)\,dx$. Alors Λ^* est le minimum de $H(\Lambda, c^*)$, et l'hypothèse (*a*) du lemme 4 est vérifiée par H.

Soit $g \in \mathscr{C}_T$ vérifiant $c = \int_0^1 \varphi(x) g(x)\,dx$.

Posons $c^{(n)} = 1/n \sum_{i=1}^n \varphi(i/n) g(i/n)$, le n-ième problème d'entropie sur la moyenne associé à $\mu_n = F^{\otimes n}$, φ, $c^{(n)}$ a pour solution $d\mu_{\Lambda_n}$ avec

$$c^{(n)} = 1/n \sum_{i=1}^n \varphi(i/n)\psi'(\langle \Lambda_n, \varphi(i/n)\rangle)$$

et

$$1/n\, S_{\mu_n}(\mu_{\Lambda_n}) = H_n(\Lambda_n, c^{(n)}).$$

Soient χ_n la loi de densité (par rapport à μ_n)

$$\prod_{i=1}^n \exp(x_i^n (\psi')^{-1}[g(i/n)] - \psi \circ (\psi')^{-1}(g(i/n))),$$

et pour y de $]a, b[$,

$$\gamma_F(y) = y[\psi']^{-1}(y) - [\psi \circ (\psi')^{-1}](y).$$

la transformée de Cramer de la mesure F.

On a alors

$$E_{\chi_n}[X_i^n] = \psi' \circ (\psi')^{-1}[g(i/n)] = g(i/n).$$

et

$$S_{\mu_n}(\chi_n) = -\sum_{i=1}^n \gamma_F[g(i/n)].$$

La loi χ_n vérifie les contraintes discrétisées pour $c^{(n)}$ ce qui implique l'inégalité

$$-1/n \sum_{i=1}^n \gamma_F[g(i/n)] \leqq \frac{1}{n} S_{\mu_n}(\mu_{\Lambda_n}) = H_n(\Lambda_n, c^{(n)}) \leqq H_n(\Lambda, c^{(n)}) \qquad (\Lambda \in \mathbb{R}^k).$$

Par passage à la limite on en déduit que

$$\Gamma_F(g) = -\int_0^1 \gamma_F[g(x)]\,dx \leqq H(\Lambda, c) \qquad (\Lambda \in \mathbb{R}^k).$$

Pour g de $\mathscr{C}_T$, $\Gamma_F(g)$ est toujours fini car la fonction γ_F est continue sur $]a, b[$. Ainsi l'hypothèse (b) du lemme 4 est vérifiée. Le lemme de Fatou donne l'hypothèse (c) et le lemme 3 est démontré.

Remarquons que la loi χ_n est la solution du problème de maximisation de l'entropie $S_{\mu_n}(P_n)$ sous les contraintes

$$E_{P_n}[X_i^n] = g(i/n), \qquad i = 1, \ldots, n.$$

On peut rassembler les résultats précédents dans le

THÉORÈME 2. — *Soient* $\varphi : [0, 1] \to \mathbb{R}^k$ *une application continue,* $c \in \mathbb{R}^k$ *donnés; soit* (P) *le problème*: *trouver* g *continue* $-\infty \leqq a < g(x) < b \leqq \infty$, telle que

$$\int_0^1 \varphi(x) g(x) \, dx = c.$$

Alors les conditions suivantes sont équivalentes:

(1) (P) *a une solution.*

(2) *Il existe* ψ, *tel que* $\exp \psi$ *soit la transformée de Laplace d'une mesure positive vérifiant* (H1), (H2) *et* (H3), *telle que* (P) ait pour solution $\psi'(\langle \Lambda_\infty, \varphi(x) \rangle)$ *où* Λ_∞ *est défini par*

$$\int_0^1 \psi'(\langle \Lambda_\infty, \varphi(x) \rangle) \varphi(x) \, dx = c.$$

(3) *Pour tout* ψ *tel que* $\exp \psi$ *soit la transformée de Laplace d'une mesure positive vérifiant* H1, H2 *et* H3, (P) *a pour solution* $\psi'(\langle \Lambda_\infty, \varphi(x) \rangle)$ *où* Λ_∞ *est défini par*

$$\int_0^1 \psi'(\langle \Lambda_\infty, \varphi(x) \rangle) \varphi(x) \, dx = c.$$

Le cas $\{g \in C([0, 1]), \|g\|_{L^2} < 1\}$. — Remarquons d'abord que si la mesure de référence est la densité uniforme sur la boule unité B_n de $\mathbb{R}^n$ de volume v_n alors les hypothèses du lemme 2 sont vérifiées, et l'on a pour u dans $\mathbb{R}^k$

$$Z_n(u) = (v_n)^{-1} \int_{B_n} \exp[1/n \langle u, \varphi_n \rangle . x^n] \, dx_1^n \ldots dx_n^n,$$

où

$$x^n = (x_1^n, \ldots, x_n^n), \qquad \langle u, \varphi_n \rangle = (\langle u, \varphi(i/n) \rangle, \, i = 1 \ldots n).$$

$Z_n(u)$ ne dépend en fait que de la norme du vecteur $1/n \langle u, \varphi_n \rangle$, En effet pour t de $\mathbb{R}^n$,

$$\int_{B_n} \exp[t.x]\,dx_1^n\ldots dx_n^n = \int_{-1}^{1} \exp[\|t\|y](1-y^2)^{(n-1)/2}\,v_{n-1}\,dy.$$

où $\|t\|$ est la norme euclidienne sur $\mathbb{R}^n$, dans la suite on notera $\mathrm{I}(t)$ cette intégrale.

On a donc

$$Z_n(u) = \int_{-1}^{1} \exp[1/n\,\|\langle u, \varphi_n\rangle\|y](1-y^2)^{(n-1)/2}\,\frac{v_{n-1}}{v_n}\,dy.$$

De

$$\operatorname{grad}[\operatorname{Log}\mathrm{I}(t)] = t\,[\|t\|\,\mathrm{I}(t)]^{-1}\,\mathrm{J}(t),$$

où

$$\mathrm{J}(t) = \int_{-1}^{1} y\exp[y\,\|t\|](1-y^2)^{(n-1)/2}\,v_{n-1}\,dy,$$

on déduit que

$$\mathrm{E}_{Q_n^{\mathrm{ME}}}[X_i^n] = \langle u_n, \varphi(i/n)\rangle(\|\langle u_n, \varphi_n\rangle\|)^{-1}\,\mathrm{G}_n(u_n).$$

où u_n est défini par

$$1/n(\|\langle u_n, \varphi_n\rangle\|)^{-1}\,\mathrm{G}_n(u_n)\sum_{i=1}^{n}\langle u_n, \varphi(i/n)\rangle\,\varphi(i/n) = c,$$

et $\mathrm{G}_n(u)$ par

$$(\mathrm{I}(n^{-1}\langle u, \varphi_n\rangle))^{-1}\,\mathrm{J}(n^{-1}\langle t, \varphi_n\rangle).$$

Posons alors $\Lambda_n = u_n(\|\langle u_n, \varphi_n\rangle\|)^{-1}\,\mathrm{G}_n(u_n)$; on obtient

$$g_n^{\mathrm{MEM}}(x) = \sum_{i=1}^{n}\langle \Lambda_n, \varphi(i/n)\rangle\,1|_{(i-1)/n \leqq x < i/n}(x)$$

où Λ_n est l'unique minimum de

$$\mathrm{H}_n(\Lambda, c) = (2n)^{-1}\sum_{i=1}^{n}\langle \Lambda, \varphi(i/n)\rangle^2 - \langle \Lambda, c\rangle.$$

La suite H_n converge (uniformément sur tout compact) vers

$$\mathrm{H}(\Lambda, c) = \int_0^1 1/2\,\langle \Lambda, \varphi(x)\rangle^2\,dx - \langle \Lambda, c\rangle,$$

qui possède un unique minimum en $[\mathrm{M}_\varphi]^{-1}c$ (où M_φ est la matrice des produits scalaires des composantes de φ dans L^2).

La suite $(g_n^{\mathrm{MEM}}(x))$ est donc convergente de limite

$$g^{\mathrm{MEM}}(x)=\langle [\mathrm{M}_\varphi^{-1}]\,c, \varphi(x)\rangle, \qquad (\mathrm{M}_\varphi)_{ij}=\int_0^1 \varphi_i(x)\,\varphi_j(x)\,dx,$$

qui est la reconstruction des moindres carrés.

Remarquons que $\|g^{\mathrm{MEM}}\|_{\mathrm{L}^2}$ est bien plus petite que 1, en effet la contrainte est par hypothèse réalisée par une fonction de norme L^2 inférieure à 1, et g^{MEM} minimise $\|g\|_{\mathrm{L}^2}$ sous la contrainte

$$\int_0^1 \varphi(x)\,g(x)\,dx=c.$$

V. MAXIMISATION DES FONCTIONNELLES CONCAVES ET MAXIMUM D'ENTROPIE SUR LA MOYENNE

Comme cadre général à la minimisation de normes, à la maximisation de l'entropie usuelle ou de celle de Burg, on peut proposer la situation suivante:

Soit $\exp[\psi_{\mathrm{F}}(u)]$ la transformée de Laplace d'une mesure F (positive) sur $\mathbb{R}$; on suppose que F vérifie les hypothèses H1, H2, H3. Le problème (MEM) a comme solution ainsi que nous venons de le voir

$$g^{\mathrm{MEM}}(x)=\psi'(\langle \Lambda_\infty, \varphi(x)\rangle),$$

où

$$\int_0^1 \psi'(\langle \Lambda_\infty, \varphi(x)\rangle)\,\varphi(x)\,dx=c.$$

Les fonctions de type négatif jouent donc un rôle très particulier. Si l'on considère la fonctionnelle de Legendre

$$\Gamma(h)=-\int_0^1 \gamma[h(y)]\,dx,$$
$$\gamma(y)=y\,(\psi')^{-1}(y)-[\psi\circ(\psi')^{-1}](y),$$

définie sur $\mathscr{C}_{\mathrm{T}}$, alors

THÉORÈME 3. — g^{MEM} *est l'unique solution de* $g = \arg\max_{h \in \mathbb{L}(c)} \Gamma(h)$, *où*

$$\mathbb{L}(c) = \left\{ h \in \mathscr{C}_{\mathrm{T}},\ \int_0^1 h(x)\,\varphi(x)\,dx = c \right\}.$$

Démonstration. — D'après la démonstration du lemme 3 on a

$$\forall h \in \mathbb{L}(c), \qquad \Gamma(h) \leqq \mathrm{H}(\Lambda_\infty, c).$$

Un simple calcul donne d'autre part $\mathrm{H}(\Lambda_\infty, c) = \Gamma(g^{\mathrm{MEM}})$, ce qui entraîne que g^{MEM} est une solution du problème d'optimisation. On conclut en remarquant que γ est une fonction strictement convexe ce qui assure l'unicité de la solution.

Remarquons que l'on aurait pu conclure directement par un calcul standard de variation (par les équations d'Euler).

Exemples :

(1) $\psi(t) = t^2/2$, F la loi normale centrée réduite, $\psi'(t) = (\psi')^{-1}(t) = t$, dans ce cas le critère est

$$\Gamma(g) = \int_0^1 -g^2(t)\,dt \qquad \text{(moindres carrés).}$$

(MEM) est ici la discrétisation du problème usuel, pour la mesure de Wiener sur C([0, 1]) (*cf.* exemple B dans II).

(2) $\psi(t) = \exp t - 1$, F la loi de Poisson de paramètre 1, $\psi'(t) = \exp t$, $(\psi')^{-1}(t) = \operatorname{Log} t$, le critère est ici l'entropie usuelle

$$\Gamma(g) = \int_0^1 (-g(x)\operatorname{Log}[g(x)] + g(x) - 1)\,dx.$$

En général l'entropie est utilisée pour reconstruire une densité de probabilité ($\varphi_1 = 1$ et $c_1 = 1$), on retrouve dans ce cas l'expression habituelle

$$\Gamma(g) = -\int_0^1 g(x)\operatorname{Log}[g(x)]\,dx.$$

(MEM) est encore la discrétisation d'un problème usuel (mesure de Poisson, *cf.* exemple A de II).

(3) F est la mesure de Lebesgue sur $]0, +\infty[$, $\psi(t) = -\operatorname{Log}(-t)$, $\psi'(t) = (\psi')^{-1}(t) = -1/t$; $(t < 0)$.

Le critère associé est l'entropie de Burg

$$\Gamma(g)=\int_0^1 [-g(x)(g(x))^{-1}-\operatorname{Log}([g(x)]^{-1})]\,dx$$
$$=\int_0^1 \operatorname{Log}[g(x)]\,dx-1.$$

Dans le cas où $\varphi_1=1$, on peut aussi interpréter l'entropie de Burg comme un problème usuel avec la mesure *a priori* G associée à la mesure aléatoire définie par le semi-groupe $\gamma(t, 1)$ de lois indéfiniment divisibles.

(4) Navaza [12], [13] propose pour reconstruire une densité électronique bornée (on se ramène au cas où elle est comprise entre 0 et 1) de prendre pour F la loi uniforme sur]0, 1[, la solution est alors

$$g^{\mathrm{MEM}}(x)=(\exp[\langle \Lambda_\infty, \varphi(x)\rangle]-1)/\langle \Lambda_\infty, \varphi(x)\rangle.$$

Le critère concave associé n'est pas ici explicitement calculable.

Remarque sur le groupe naturel associé à une mesure. — Considérons la mesure de Lebesgue sur [0, 1]. On peut demander que toute suite (μ_n) de mesures sur $\mathbb{R}^n$ associée à la discrétisation soit invariante par le groupe des permutations $\mathscr{G}_n$ qui est la trace sur $\mathbb{R}^n$ du groupe des transformations conservant la mesure. La recherche de densité par rapport à la mesure de Lebesgue se fera alors à l'aide de suites de probabilités (μ_n) échangeables [mais non nécessairement cohérentes dans C([0, 1])].

On s'intéresse particulièrement à des contraintes non linéaires invariantes par réarrangement; les cas simples sont les normes L^p, les normes d'Orlicz ou de Lorentz, ou des combinaisons de ces normes. Ainsi $\mathscr{C}_T$ correspond à $\|g \vee 0\|_\infty<b$, $\|g \wedge 0\|_\infty<a$, $\mathscr{C}_\theta$ à $\|g\|_2<1$.

VI. EXISTENCE D'UNE SOLUTION AU PROBLÈME DES MOMENTS DANS $\mathscr{C}_T$

Nous allons maintenant étudier l'existence d'une fonction $g\in\mathscr{C}_T$ telle que

$$\int_0^1 \varphi(x)\,g(x)\,dx=c,$$

où

$$\mathscr{C}_T=\{g, g\in C([0, 1]), a<g(x)<b\}.$$

Un des cas bien connu est celui où $]a, b[=\mathbb{R}^{+*}$ et où φ est un système d'exponentielles $\{\exp(i 2\pi \alpha x), |\alpha| \leqq k\}$.

Ce cas peut être entièrement résolu par des méthodes hilbertiennes et de variable complexe [11]. Le problème a alors une solution si et seulement si c est de type positif. La solution continue, classiquement introduite, est de la forme $(|P_k(e^{-i2\pi x})|^{-2})$, où P_k est le polynôme de degré k défini par, $T_k(c) P_k=(0, 0, \ldots, 1)$, $T_k(c)$ est la matrice de Toeplitz associée à c. Cette fonction est la densité spectrale du processus d'entropie maximale dont la covariance est une extension de c. Remarquons que le théorème 2 nous donne, quand c est de type positif, une solution $\psi'[P_k(\exp(-i2\pi x))]$ avec ψ de type négatif. C'est une famille très intéressante pour la réalisation pratique d'une covariance.

Avant d'énoncer le théorème, faisons une remarque très importante: la construction par discrétisation qui est à la base des sections IV et V n'utilise pas les propriétés particulières de l'ensemble $[0, 1]$ sur lequel φ et g sont définis. En fait $[0, 1]$ pourrait être remplacé par un ensemble E compact, métrique, séparable.

En particulier dans le cas des systèmes d'exponentielles on peut remplacer $[0, 1]$ par $[0, 1]^d=\mathbb{T}^d$, et prendre $\alpha \in \mathbb{Z}^d$.

Ce type d'extensions est beaucoup plus difficile à mettre en œuvre avec les méthodes hilbertiennes (dans le cas $]a, b[=\mathbb{R}^{+*}$).

Nous noterons,

$$V_1=\left\{c \in \mathbb{R}^k, \text{ tel qu'il existe } g \text{ de } \mathscr{C}_T \quad \text{avec} \quad \int_0^1 \varphi(x) g(x)\, dx=c\right\}$$
$$=\left\{\int_0^1 \varphi(x) g(x)\, dx, g \in \mathscr{C}_T\right\}$$

et $V_2=\mathbb{R}^k/V_1$, d'intérieur $\mathring{V}_2$.

On dira que c est réalisable dans $\mathscr{C}_T$ si $c \in V_1$.

THÉORÈME 4. — *Soient* $[a, b]=[0, 1]$, $\psi(t)=\operatorname{Log}(e^t+1)$,

$$h(\Lambda)=\int_0^1 \psi(\langle \Lambda, \varphi(x) \rangle)\, dx,$$
$$H(\Lambda, c)=h(\Lambda)-\langle \Lambda, c \rangle,$$
$$h^*(c)=-\inf_{\Lambda \in \mathbb{R}^k} H(\Lambda, c).$$

On suppose que $\varphi(x)$ *est de rang plein pour presque tout* x. *Alors* c *est réalisable dans* $\mathscr{C}_T$ *si et seulement si* $h^*(c)<0$

$$c \in \mathring{V}_2 \text{ équivaut à } h^*(c)=+\infty$$

et

$$c \in V_2/\mathring{V}_2 \text{ équivaut à } h^*(c)=0.$$

Remarques. — En renormalisant ψ, on a un résultat identique pour a et b quelconques.

— Ce résultat est à rapprocher de celui de Krein Nudelman [11]: $c \in V_1$ si et seulement si pour tout Λ de $\mathbb{R}^k$, on a

$$2^{-1}\int_0^1 (|\langle \Lambda, \varphi(x)\rangle| + \langle \Lambda, \varphi(x)\rangle)\,dx > \langle \Lambda, c\rangle.$$

COROLLAIRE 2. — *Soit* $\varphi(x)$ *défini par un système d'exponentielles*

$$\varphi(x)=\{\exp(i2\pi\alpha x), |\alpha| \leqq N\},$$

soit $P_k(e^{i2\pi x})$ *un polynôme trigonométrique réel de la forme*

$$\sum_{j=1}^{k} (\Lambda_j \exp(i2\pi jx) + \bar{\Lambda}_j \exp(-i2\pi jx)) + \Lambda_0.$$

$$\Lambda=(\Lambda_0, \Lambda_1, \ldots, \Lambda_k) \in \mathbb{R} \times \mathbb{C}^k.$$

Alors c *est réalisable dans* $\mathscr{C}_T$ *si et seulement si pour tout polynôme* P_k, *on a*

$$P_k(c) < \int_0^1 \operatorname{Log}(1+\exp(P_k(e^{i2\pi x}))\,dx.$$

Il existe alors une réalisation de c *dans* $\mathscr{C}_T$ *de la forme*

$$P_k^*(e^{i2\pi x}).(1+P_k^*(e^{i2\pi x}))^{-1}.$$

Pour démontrer le théorème 4 nous avons besoin du lemme suivant:

LEMME 5. — (1) *Soit* $c \in \mathring{V}_2$ *et* $B(c, \varepsilon)$ *une boule ouverte de centre* c, *de rayon* $\varepsilon > 0$ *et d'adhérence incluse dans* $\mathring{V}_2$. *Alors il existe* n_0 *tel que*

$$\forall n \geqq n_0, \qquad \forall c' \in \bar{B}(c, \varepsilon),$$
$$\forall \sigma^n = (\sigma_1^n, \ldots, \sigma_n^n) \in]0, 1[^n,$$
$$c' \neq 1/n \sum_{j=1}^{n} \sigma_j^n \varphi(j/n).$$

(2) *On peut dans* (1) *remplacer* $]0, 1[^n$ *par* $\{0, 1\}^n$.

Démonstration du lemme. — (1) Soient la matrice $k \times n$ $T_n = (\varphi(j/n))_{j=1\ldots n}$ et m un entier fixé tel que T_m soit de rang k (ce qui est possible puisque φ est de rang k).

Soit $\varepsilon_1 > 0$ tel que:

$$\text{pour tout } c' \in \bar{B}(c, \varepsilon), \quad B(c', \varepsilon_1) \subset \mathring{V}_2.$$

Posons $G = \bigcup_{c' \in \bar{B}(c, \varepsilon)} B(c', \varepsilon_1)$.

Supposons qu'il existe une suite $(c^l)_{l>0}$ de $\bar{B}(c, \varepsilon)$ telle que

$$c^l = \frac{1}{n(l)} \sum_{j=1}^{n(l)} \sigma_j^{n(l)} \varphi\left(\frac{j}{n(l)}\right), \qquad \sigma^{n(l)} \in]0, 1[^n,$$

où $n(l)$ est une suite croissante d'entiers.

Il existe un hyperplan α qui sépare G de V_1; on a alors

$$\forall c \in V_1, \qquad \forall \eta \in B(0, \varepsilon_1), \qquad \langle \alpha, c \rangle > \langle \alpha, c^{(l)} + \eta \rangle.$$

Posons $B = \{ \xi \in \mathbb{R}^m, T_m \xi = \eta, \eta \in B(0, \varepsilon_1) \}$; B est un voisinage de 0 et on a, pour tout $g \in \mathscr{C}_T$, pour tout $\xi \in B$.

$$\int_0^1 \langle \alpha, \varphi(x) \rangle g(x)\, dx > \frac{1}{n(l)} \sum_{j=1}^{n(l)} \sigma_j^{n(l)} \left\langle \iota, \varphi\left(\frac{j}{n(l)}\right)\right\rangle + \sum_{j=1}^{m} \xi_j \langle \alpha, \varphi(j/m) \rangle.$$

On en déduit

$$\beta \sum_{j=1}^{m} |\langle \alpha, \varphi(j/m) \rangle| < \varepsilon_{n(l)}$$

où $\beta > 0$, et $(\varepsilon_{n(l)})_l$ est une suite de réels positifs qui tend vers 0.

L'inégalité précédente implique que,

$$|\langle \alpha, \varphi(j/m) \rangle| = 0, \qquad \text{pour} \quad j = 1, \ldots, m.$$

ce qui contredit l'hypothèse sur le rang de T_m.

(2) Il suffit de considérer $B(\lambda c + (1-\lambda) c^*, \varepsilon)$ avec $c^* \in V_1$, et $\lambda \in]0, 1[$, tels que l'adhérence de cette boule soit incluse dans $\mathring{V}_2$ et d'utiliser (1).

Démonstration du théorème 4. — (*a*) Soit $c \in V_1$, d'après le théorème 2

$$\left(V_1 = \left\{ \int_0^1 \varphi(x) g(x)\, dx, g \in \mathscr{C}_T \right\}\right),$$

la fonction $g^{\text{MEM}}(x) = \psi'(\langle \Lambda_\infty, \varphi(x) \rangle)$ vérifie

$$\int_0^1 g^{\text{MEM}}(x) \varphi(x)\, dx = c \qquad \text{avec} \quad \psi'(t) = e^t (1 + e^t)^{-1}.$$

Si γ désigne comme précédemment la conjuguée de Young-Legendre de ψ et si

$$\Gamma(g) = -\int_0^1 \gamma(g(x))\, dx,$$

on a $\Gamma(g^{\text{MEM}}) = H(\Lambda_\infty, c)$ (*cf.* démonstration du théorème 3) donc $-\text{Log}\, 2 \leqq h^*(c) < 0$, car $-\text{Log}\, 2 \leqq \gamma(y) < 0$ pour $y \in]0, 1[$.

(*b*) Soit $c \in \mathring{V}_2$ et $B(c, \varepsilon)$ d'adhérence incluse dans $\mathring{V}_2$.

Notons $\Pi_n = (1/2\,\delta_0 + 1/2\,\delta_1)^{\otimes n}$ et X_j^n, $j = 1 \ldots n$, les applications coordonnées. D'après le lemme 5 pour $n \geqq n_0$ on a:

$$\Pi_n (1/n \sum_{j=1}^{n} X_j^n \varphi (j/n) \in B(c, \varepsilon)) = 0.$$

Le théorème de grandes déviations ([5], p. 47) donne,

$$\lim_{n \to \infty} \inf 1/n \operatorname{Log} \Pi_n \left(1/n \sum_{i=1}^{n} X_j^n \varphi (j/n) \in \mathring{B}(c, \varepsilon) \right) \geqq - \inf_{c' \in \mathring{B}(c, \varepsilon)} h^*(c') - \operatorname{Log} 2$$

et donc $h^*(c) = +\infty$.

(*c*) Soit $c \in V_2 / \mathring{V}_2$, comme h^* est convexe fermée,

$$-\operatorname{Log} 2 \leqq \operatorname*{Lim}_{\substack{c' \to c \\ c' \in V_1}} h^*(c') = h^*(c) \leqq 0.$$

Il existe une suite (Λ_n) avec $\operatorname*{Lim}_{n \to \infty} \| \Lambda_n \| = +\infty$,

$$\operatorname*{Lim}_{n \to \infty} - H(\Lambda_n, c) = h^*(c) \qquad \text{et} \qquad \operatorname*{Lim}_{n \to \infty} \Lambda_n (\| \Lambda_n \|)^{-1} = \alpha$$

(d'après le paragraphe IV, si $\Lambda_n^* \to \Lambda$ alors $c \in V_1$).

On a

$$-H(\Lambda_n, c) = \langle \Lambda_n, c \rangle - \int_0^1 \operatorname{Log} (\exp (\langle \Lambda_n, \varphi (x) \rangle) + 1)\, dx,$$

$$\int_{\{ \langle \Lambda_n, \varphi (x) \rangle < 0 \}} \operatorname{Log} (\exp (\langle \Lambda_n, \varphi (x) \rangle) + 1)\, dx \to 0 \qquad \text{si} \quad n \to \infty,$$

d'après le théorème de Lebesgue, de même

$$\int_{\{ \langle \Lambda_n, \varphi (x) \rangle > 0 \}} \operatorname{Log} (\exp (\langle \Lambda_n, \varphi (x) \rangle + 1)\, dx - \int_{\{ \langle \Lambda_n, \varphi (x) \rangle > 0 \}} \langle \Lambda_n, \varphi (x) \rangle\, dx \to 0$$

donc

$$h^*(c) = \operatorname*{Lim}_{n \to \infty} \langle \Lambda_n, c - \int_{\{ \langle \Lambda_n, \varphi (x) \rangle > 0 \}} \varphi (x)\, dx \rangle$$

et puisque $|h^*(c)| < \infty$; on a

$$\langle \alpha, c - \int_{\{\langle \alpha, \varphi(x) \rangle > 0\}} \varphi(x)\, dx \rangle = 0.$$

en prenant $\Lambda_n = n\alpha$, on a d'après le raisonnement précédent

$$\lim_{n \to \infty} -\mathrm{H}(\Lambda_n, c) = 0 \qquad \text{donc} \quad h^*(c) = 0.$$

VII. LA MÉTHODE DE MAXIMISATION DE L'ENTROPIE SUR LA MOYENNE COMME MÉTHODE SOMMATOIRE DE FOURIER

Supposons que $\varphi(x) = (\exp(ijx))_{j=0-k}$. On se donne donc pour une fonction réelle g les coefficients Fourier $c = (c_j)_{j=0\ldots k}$ $\left(\text{par la contrainte } \int_0^{2\pi} g(x)\, \varphi(x)\, dx = c\right)$. Une méthode sommatoire est la donnée pour chaque k d'une application, $\mathbb{R}^k \to \mathrm{C}([0, 1])$

$$\{c_j(g), j = 0 \ldots k\} \to g_k,$$

telle que (au moins): $\lim_{k \to \infty} g_k = g$.

Les qualités de la méthode sommatoire par maximum d'entropie ME pour une fonction positive ont été signalées dans [3] et [8]. (ME) consiste à poser: $g_k^{\mathrm{MEM}}(x)$ est la densité de la mesure d'entropie maximum (par rapport à la mesure de Lebesgue) vérifiant la contrainte $\int_0^{2\pi} g(x)\, \varphi(x)\, dx = c$. (ME) est très supérieure aux méthodes usuelles pour k « moyen » et g de forte concentration. Elle consiste d'après ce qui précède à approcher g par des exponentielles de polynômes trigonométriques.

Si l'on utilise (MEM), en imposant à g_k d'appartenir à un convexe donné *a priori*, alors on définit pour chaque convexe $\mathscr{C}$ et chaque suite de mesures de référence (μ_n) une méthode sommatoire où $g_k \in \mathscr{C}$. Ces méthodes sont convergentes aux sens des distances en mesures (L^p, etc.) et ne donnent évidemment que des interpolées. On peut aussi regarder les fonctionnelles concaves introduites précédemment et associées à chaque méthode (MEM), comme méthode sommatoire. Chacune d'entre elles est convergente et l'on peut obtenir des inégalités intéressantes, pour l'entropie de Burg par exemple. Nous donnons ici un exemple à titre d'illustration.

On suppose à partir de maintenant que $a=0$ et $b=1$ ($\mathscr{C}_{\mathrm{T}}$ est l'ensemble des fonctions continues comprises strictement entre 0 et 1). Les notations de IV sont conservées.

Pour $\Lambda=(\Lambda^0, \Lambda^1, \ldots, \Lambda^k)$ de $\mathbb{R}\times\mathbb{C}^k$ on pose

$$\mathrm{P}_\Lambda^k(t)=\langle \Lambda, \varphi(t)\rangle=\sum_{j=1}^{k}(\Lambda^j e^{ijt}+\bar{\Lambda}^j e^{-ijt})+\Lambda^0.$$

THÉORÈME 5. — *Soit* F *une mesure positive sur* [0, 1]. *Si* H_1 *est vérifiée, et si* F *est de masse totale finie alors on a l'inégalité sommatoire:* $\forall k$, $\forall \Lambda\in\mathbb{R}\times\mathbb{C}^k$,

$$\|g-g_k^{\mathrm{MEM}}\|_{\mathrm{L}^1}\leqq[2\|\mathrm{P}_\Lambda^k-(\psi')^{-1}(g)\|_{\mathrm{L}^1}+2\|\mathrm{P}_\Lambda^k-(\psi')^{-1}(g)\|_{\mathrm{L}^p}\|g\|_{\mathrm{L}^q}]^{1/2},$$

avec $1/p+1/q=1$ *et* g *une fonction continue sur* [0, 1] *qui vérifie*, $0<g(x)<1$, *pour tout* $x\in[0, 1]$ *et* $\int_0^1 g(x)\varphi(x)\,dx=c$.

Démonstration du théorème 8. — On reprend ici les notations de la démonstration du lemme 3.

Soient $\chi_n^i(x_i^n)$ la i-ième marginale de χ_n

$$d\chi_n^i(x_i^n)=\exp(x_i^n(\psi')^{-1}[g(i/n)]-\psi\circ(\psi')^{-1}[g(i/n)])\,d\mathrm{F}(x_i^n),$$

et $\mu_{\Lambda_n}^i$ la i-ième marginale de μ_{Λ_n} solution du n-ième problème (MEM) avec les coefficients approchés $c^{(n)}$

$$d\mu_{\Lambda_n}^i(x_i^n)=\exp(x_i^n\langle \Lambda_n, \varphi(i/n)\rangle-\psi[\langle \Lambda_n, \varphi(i/n)\rangle])\,d\mathrm{F}(x_i^n).$$

En utilisant l'inégalité classique entre distance en variation et information de Kullback on obtient l'inéquation

$$\begin{aligned}[2\mathrm{K}(\chi_n^i, \mu_{\Lambda_n}^i)]^{1/2}\geqq\int_0^1\left|\frac{d\mu_{\Lambda_n}^i}{d\mathrm{F}}-\frac{d\chi_n^i}{d\mathrm{F}}\right|d\mathrm{F}(x)&\\ \geqq|\mathrm{E}_{\mu_{\Lambda_n}^i}[\mathrm{X}]-\mathrm{E}_{\chi_n^i}[\mathrm{X}]|&\\ =|\psi[\langle \Lambda_n, \varphi(i/n)\rangle]-g(i/n)|.&\end{aligned}$$

$$1/n\sum_{i=1}^{n}[2\mathrm{K}(\chi_n^i, \mu_{\Lambda_n}^i)]^{1/2}\geqq 1/n\sum_{i=1}^{n}|\psi'[\langle \Lambda_n, \varphi(i/n)\rangle]-g(i/n)|.$$

l'information de Kullback est additive donc:

$$1/n\sum_{i=1}^{n}[2\mathrm{K}(\chi_n^i, \mu_{\Lambda_n}^i)]^{1/2}\leqq[2/n\sum_{i=1}^{n}\mathrm{K}(\chi_n^i, \mu_{\Lambda_n}^i)]^{1/2}=[2/n\,\mathrm{K}(\chi_n, \mu_{\Lambda_n})]^{1/2}.$$

D'autre part on a l'identité:

$$\forall \Lambda \in \mathbb{R} \times \mathbb{C}^k, \qquad 1/n\,[\mathrm{K}(\chi_n, \mu_\Lambda) - \mathrm{K}(\chi_n, \mu_{\Lambda_n})] = \mathrm{H}_n(\Lambda, c^{(n)}) - \mathrm{H}_n(\Lambda_n, c^{(n)}),$$

qui entraîne

$$\mathrm{K}(\chi_n, \mu_\Lambda) \geqq \mathrm{K}(\chi_n, \mu_{\Lambda_n}).$$

En résumé

$$1/n \sum_{i=1}^{n} |\psi'(\langle \Lambda_n, \varphi(i/n) \rangle) - g(i/n)| \leqq [2/n\,\mathrm{K}(\chi_n, \mu_\Lambda)]^{1/2}.$$

Par application du théorème de Lebesgue,

$$\lim_{n \to \infty} 1/n \sum_{i=1}^{n} |\psi'(\langle \Lambda_n, \varphi(i/n) \rangle) - g(i/n)| = \| g_k^{\mathrm{MEM}} - g \|_{\mathrm{L}^1}.$$

Il reste donc à majorer le terme de droite de l'inégalité précédente et le lemme sera établi. On a

$$1/n\,\mathrm{K}(\chi_n, \mu_\Lambda) = 1/n \sum_{i=1}^{n} [\psi(\langle \Lambda, \varphi(i/n) \rangle) - \psi \circ (\psi')^{-1}(g(i/n))]$$
$$+ 1/n \sum_{i=1}^{n} g(i/n)\,[(\psi')^{-1}(g(i/n)) - \langle \Lambda, \varphi(i/n) \rangle].$$

La limite du second terme est majorée par

$$\| g \|_{\mathrm{L}^q} \| (\psi')^{-1}(g) - \mathrm{P}_\Lambda^k \|_{\mathrm{L}^p}, \qquad 1/p + 1/q = 1;$$

quant à la limite de la première somme, en appliquant le théorème des accroissements finis et en utilisant le fait que ψ' est uniformément majoré par 1, on la majore par:

$$\| \mathrm{P}_\Lambda^k - (\psi')^{-1}(g) \|_{\mathrm{L}^1}.$$

RÉFÉRENCES

[1] I. CSISZAR, Sanov Property, Generalized I-Projection and a Conditional Limit Theorem, *Ann. Prob.*, vol. **12**, n° **3**, 1984, p. 768-793.

[2] I. CSISZAR, I-Divergence Geometry of Probability Distributions and Minimization Problems, *Ann. Prob.*, vol. **3**, n° **1**, 1975, p. 148-158.

[3] D. DACUNHA-CASTELLE, Congrès de la Société Bernouilli Tachkent 1987, North Holland, 1987.

[4] D. DACUNHA-CASTELLE et M. DUFLO, *Probabilités et statistiques*, tome 2, Problèmes à temps mobile, Masson, 1983.

[5] R. S. ELLIS, *Entropy, Large Deviations and Statistical Mechanics*, Springer-Verlag, 1985.

[6] F. GAMBOA, Minimisation de l'information de Kullback et maximisation de l'entropie sous une contrainte quadratique, *C.R. Acad. Sci. Paris*, t. **306**, série I, 1988, p. 425-427.

[7] F. GAMBOA et E. GASSIAT, Maximum d'entropie et problème des moments cas multidimensionnels, *Prob. and Math. Stat*, 1990, (à paraître).

[8] E. GASSIAT, Problème sommatoire par maximum d'entropie, *C.R. Acad. Sci. Paris*, t. **303**, série I, 1986.

[9] S. F. GULL et J. SKILLING, *The Entropy of an Image. Maximum Entropy and Bayesian Methods in Inverse Problems*, D. Reidel publishing company, 1985, p. 287-301.

[10] O. KALLENBERG, Canonical Representations and Convergence Criteria for Processus with Interchangeable Moments, *Z. Warschein, verw. Gebiete*, Vol. **27**, 1973, p. 23-26.

[11] M. G. KREIN et A. A. NUDELMAN, The Markov Moment Problem and Extremal Problems, vol. **50**, *Am. Math. Soc*, 1977.

[12] J. NAVAZA, On the Maximum Entropy Estimate of the Electron Density Fonction, *Acta-Cryst*, **A41**, 1985, p. 232-241.

[13] J. NAVAZA, Use of Non Local Constraints in Maximum Entropy Electron Density Reconstruction, *Acta-Cryst.*, Vol. **A42**, 1986, p. 212-222.

(*Manuscrit reçu le* 1er *mars* 1988;
révisé le 28 *mai* 1990.)