



Statistique

Sur l'approximation des fonctions convexes par les fonctions génératrices des cumulants [☆]

Roland Hildebrand

LMC, université Joseph-Fourier, tour IRMA, 51, rue des mathématiques, 38400 St. Martin d'Hères, France

Reçu le 7 octobre 2004 ; accepté après révision le 5 septembre 2006

Disponible sur Internet le 12 octobre 2006

Présenté par Paul Deheuvels

Résumé

Nous montrons que chaque fonction convexe f définie sur la droite réelle ou un intervalle réel peut être approximée dans la norme C^0 par la fonction génératrice des cumulants d'une mesure non-négative avec une erreur bornée par une constante absolue, qui ne dépend pas de f . Nous fournissons des bornes supérieures et inférieures sur la meilleure de telles constantes, notamment $\ln 2$ et $\frac{\ln 2}{2}$. La déduction de ces bornes est constructive. Nous montrons également que dans le cas multi-dimensionnel l'erreur de l'approximation n'est pas bornée. *Pour citer cet article : R. Hildebrand, C. R. Acad. Sci. Paris, Ser. I 343 (2006).*

© 2006 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

Abstract

On the approximation of convex functions with cumulant generating functions. In this Note we show that any convex function f on the real line or an interval thereof can be approximated in the C^0 norm by the cumulant generating function of a non-negative measure with an error bounded by an absolute constant which does not depend on f . We give upper and lower bounds on the best of such constants, which equal $\ln 2$ and $\frac{\ln 2}{2}$, respectively. The proofs for these bounds are constructive. We also show that the approximation error in the multi-dimensional case is not bounded. *To cite this article: R. Hildebrand, C. R. Acad. Sci. Paris, Ser. I 343 (2006).*

© 2006 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

Abridged English version

Let $f(t)$ be a real-valued measurable function defined on the real line. The *two-sided Laplace transform* of f is defined as $F(s) = \int_{-\infty}^{+\infty} e^{-st} f(t) dt$. It is well known that the region of convergence of this integral is a vertical strip in the complex plane, which may be unbounded to the left or to the right or to both sides, and in this region the function $F(s)$ is analytic [6].

[☆] This paper presents research results of the Belgian Programme on Interuniversity Poles of Attraction, Phase V, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, and of the Action Concertée Incitative 'Masses de données' of CNRS, France. The scientific responsibility rests with its author.

Adresse e-mail : roland.hildebrand@imag.fr (R. Hildebrand).

Let $\mu(t)$ be a probability measure and $M(s)$ its two-sided Laplace transform. Then the function $M(-s) = \int_{-\infty}^{\infty} e^{st} \mu(t) dt$ generates the *moments* of this measure. Its logarithm $\varphi(s) = \ln M(-s)$ generates the *cumulants* of the probability measure and is thus called *cumulant generating function* (CGF) [5]. Since μ is a probability measure, the cumulant generating function satisfies the relation $\varphi(0) = 0$. To avoid this restriction, we further consider CGFs of unnormalized non-negative measures, which not necessarily sum up to 1. In the sequel, if we speak of a CGF we assume the CGF of an unnormalized non-negative measure. CGFs have many applications in statistics, e.g. they describe the asymptotics of large deviations [4] and play a central role in the apparatus of the exponential family of distributions [3].

A CGF $\varphi(s)$ of a non-negative measure $\mu(t)$ is an analytic function whose derivatives satisfy certain polynomial inequalities. To obtain these inequalities, recall that a sequence of real numbers $\{m_0, m_1, \dots, m_n, \dots\}$ can be represented as the sequence of moments of a non-negative measure if and only if all Hankel matrices of the form

$$H(m_0, m_1, \dots, m_{2n}) = \begin{pmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & m_2 & \cdots & m_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_n & m_{n+1} & \cdots & m_{2n} \end{pmatrix}$$

are positive semidefinite [2]. This yields the non-negativity of the principal minors of these matrices. These minors are polynomials in the m_k . But m_k is the k -th order derivative at zero of the exponent $e^{\varphi(s)}$ of the CGF of the measure and hence can be expressed through the derivatives of $\varphi(s)$ at zero. In this way one obtains a sequence of inequalities on the derivatives of $\varphi(s)$, which can easily be verified to be also polynomial. By multiplication of the measure with $e^{s_0 t}$ one obtains the same inequalities for the derivatives at an arbitrary point s_0 .

The first two non-trivial inequalities are $m_0 m_2 - m_1^2 \geq 0$, $m_0 m_2 m_4 + 2m_1 m_2 m_3 - m_0 m_3^2 - m_4 m_1^2 - m_2^3 \geq 0$, which lead to $\varphi'' \geq 0$, $\varphi''(\varphi^{(IV)} + 2(\varphi'')^2) - (\varphi''')^2 \geq 0$. While the first inequality involves only moments up to order 2 of the corresponding measure, the following inequalities involve higher-order moments. Note that the first inequality expresses the condition of convexity of the CGF.

Obviously not every convex function is a CGF of some non-negative measure, already because CGFs are analytic functions. One can then ask how well an arbitrary convex function can be approximated by a CGF. This question constitutes the subject of the present note.

The surprising answer is that a convex function can be approximated by a CGF with an error in the C^0 norm that is bounded above by an absolute constant, no matter whether the convex function is defined on a finite interval or an infinite one or whether it is bounded. We prove an upper and a lower bound to the best such constant. These bounds equal $\ln 2$ and $\frac{\ln 2}{2}$, respectively. Namely, we have the following results:

Theorem 0.1. *Let $f(x)$ be a continuous convex function, defined on the real line or an interval of the real line. Then there exists a non-negative measure whose CGF $\varphi(x)$ is defined on the same interval and $\|f - \varphi\|_{C^0} \leq \ln 2$.*

Theorem 0.2. *Let $\varepsilon > 0$. Then there exists a continuous convex function $f(x)$ on the interval $[-1, 1]$ such that for any CGF $\varphi(x)$ of a non-negative measure we have $\|f - \varphi\|_{C^0([-1, 1])} \geq \frac{\ln 2}{2} - \varepsilon$.*

We also consider the approximation of convex functions in real spaces of higher dimensions by multivariate CGFs. The CGF of a non-negative measure $\mu(t)$ on \mathbb{R}^n is defined to be the function $\varphi(s) = \ln \int_{\mathbb{R}^n} e^{(s,t)} \mu(t) dt$, where $\langle \cdot, \cdot \rangle$ is the usual scalar product in \mathbb{R}^n . This problem is the straightforward generalization of the preceding approximation problem from the case of one to the case of several independent variables. We show that the above results cannot be extended to the multivariable case, and there exist convex functions on \mathbb{R}^n , $n \geq 2$, for which the approximation by CGFs is arbitrarily bad. Namely, for $n \geq 2$ there exist convex functions on convex domains of \mathbb{R}^n which are approximated by multivariate CGFs with an arbitrarily large error in the C^0 norm. We have the following result:

Theorem 0.3. *Let $n \geq 2$ be an integer and $R > 0$ a real number. Then there exists a continuous convex function $f(x)$ on the closed ball $B_1 \subset \mathbb{R}^n$ with unit radius such that for any n -variate CGF $\varphi(x)$ of a non-negative measure on B_1 we have $\|f - \varphi\|_{C^0(B_1)} \geq R$.*

Theorems 0.1–0.3 will be proven in the next three sections, respectively.

1. Borne supérieure

Théorème 1.1. Soit $f(x)$ une fonction convexe continue, définie sur la droite réelle ou un intervalle réel. Alors il existe une mesure non-négative dont la fonction génératrice des cumulants $\varphi(x)$ est définie sur le même intervalle et $\|f - \varphi\|_{C^0} \leq \ln 2$.

Preuve. Soit x_0 un point dans le domaine D de définition de f et soit p_0 un sous-gradient de f à x_0 . Définissons la fonction $t_{x_0, p_0}(x) = f(x) - f(x_0) - p_0(x - x_0)$. Elle est non-négative, convexe (par rapport à x) et possède un zéro à $x = x_0$. De plus, si $x_0 \leq x_1 \leq x_2$ sont des points dans D et $p_0 \leq p_1 \leq p_2$ sont des sous-gradients de f dans ces points, alors on a les inégalités suivantes :

$$t_{x_0, p_0}(x_2) \geq t_{x_0, p_0}(x_1) + t_{x_1, p_1}(x_2), \quad t_{x_2, p_2}(x_0) \geq t_{x_2, p_2}(x_1) + t_{x_1, p_1}(x_0). \tag{1}$$

Nous allons construire d'une manière itérative une séquence de fonctions linéaires $\dots, l_{-2}(x), l_{-1}(x), l_0(x), l_1(x), l_2(x), \dots$ sur D , où l'index k parcourt un ensemble $K \subset \mathbb{Z}$ des entiers consécutifs. L'ensemble K peut être infini sur chaque côté et contient le zéro. Ces fonctions obéissent l'inégalité $l_k(x) \leq f(x)$ pour tout $x \in D$. Définissons la fonction $l_0(x)$ en fixant un point $x_0 \in D$ et un sous-gradient p_0 de f à x_0 et en mettant $l_0(x) = f(x_0) + p_0(x - x_0)$. Les fonctions $l_{\pm(k+1)}(x)$ seront construites à partir des fonctions $l_{\pm k}(x)$. D'ailleurs, une de ces fonctions ou toutes les deux peuvent faillir d'exister. Dans ce cas-là la séquence se termine dans la direction correspondante. La construction pour la direction positive et la direction négative est similaire, et nous ne décrivons que celle pour les indices non-négatifs.

Pour $k = 0$ l'inégalité $l_0(x) \leq f(x)$ est satisfaite par la convexité de f . Soit $k \geq 0$ et $\alpha = \ln 2$. Supposons que $l_k(x) \leq f(x)$ est déjà construite et montrons la construction de la fonction linéaire $l_{k+1}(x)$. Soit p_k la dérivée de $l_k(x)$. S'il n'y a aucun point $x \in D$ tel que $l_k(x) = f(x)$, alors l_{k+1} n'existe pas et k est l'index maximal de l'ensemble K . S'il existe un tel point $x \in D$, on en choisit un, disons x_k . Alors il suit de l'inégalité $l_k(x) \leq f(x)$ que p_k est un sous-gradient de f à x_k . Si on a $t_{x_k, p_k}(x) \leq \alpha$ pour tout $x > x_k, x \in D$, alors l_{k+1} n'existe pas non plus et k est l'index maximal de K . S'il existe $x \in D, x > x_k$ tel que $t_{x_k, p_k}(x) > \alpha$, alors il existe un point unique $x^* > x_k$ tel que $t_{x_k, p_k}(x^*) = \alpha$. Posons $p_{k+1} = \sup\{p \mid f(x^*) - \alpha + p(x - x^*) \leq f(x) \forall x \in D\}$. L'ensemble de tels points p n'est pas vide et son supremum est supérieur à p_k parce qu'il contient tous les sous-gradients de f à x^* . Le supremum est aussi fini, parce qu'il existe $x \in D, x > x^*$, et est alors atteint. Définissons $l_{k+1}(x) = f(x^*) - \alpha + p_{k+1}(x - x^*)$. Par la construction nous avons $l_{k+1}(x) \leq f(x)$ pour tout $x \in D$.

La séquence des fonctions linéaires construites en-dessus possède les propriétés suivantes : (i) Les dérivées p_k des fonctions $l_k(x)$ forment une séquence strictement croissante. (ii) Pour chaque $x \in D$ il existe un index $k \in K$ tel que $l_k(x) \leq f(x) \leq l_k(x) + \alpha$.

Notons \bar{D} pour la clôture de D dans la compactification à deux points $\mathbb{R} \cup \{\pm\infty\}$ de \mathbb{R} . Pour chaque index $k \in K$ nous définissons un nombre (possiblement infini) $\Delta_k \geq \alpha$ dans $\mathbb{R} \cup \{\pm\infty\}$ et un point $\hat{x}_k \in \bar{D}$. Notamment, si $k - 1, k + 1 \in K$, alors il existe un point unique x^* tel que $l_{k-1}(x^*) = l_{k+1}(x^*)$. Dans ce cas-là on pose $\Delta_k = f(x^*) - l_{k-1}(x^*)$ et $\hat{x}_k = x^*$. Si k est l'index maximal de K et $k - 1 \in K$, alors on définit x^* comme le point unique tel que $l_{k-1}(x^*) = l_k(x^*)$ et on pose $\Delta_k = \sup_{x > x^*, x \in D} (f(x) - l_{k-1}(x))$, $\hat{x}_k = \sup_{x \in D} x$. On définit Δ_k d'une manière similaire si k est l'index minimal de K et $k + 1 \in K$. Enfin, on pose $\Delta_k = \infty$ et $\hat{x}_k \in D$ arbitraire si $K = \{k\}$. De plus, définissons $\delta_k = 1 - e^{-\max(\alpha, \Delta_k - \alpha)}$ (avec $\delta_k = 1$ si $\Delta_k = \infty$).

Définissons la fonction $\varphi(x) = \ln \frac{4}{3} + \ln \sum_{k \in K} e^{l_k(x) + \delta_k}$. Dans le domaine de convergence de cette somme la fonction φ est la fonction génératrice des cumulants d'une mesure non-négative discrète. Nous allons sortir des bornes supérieures et inférieures sur la différence $\varphi(x) - f(x)$ pour tout $x \in D$ et alors prouver la convergence sur D .

Pour tout index $k \in K$ nous avons $l_k(\hat{x}_k) - l_{k-1}(\hat{x}_k) \geq \Delta_k - \alpha, l_k(\hat{x}_k) - l_{k+1}(\hat{x}_k) \geq \Delta_k - \alpha$, si les indices concernés sont contenus dans K . De plus, nous avons $l_k(\hat{x}_k) - l_{k-1}(\hat{x}_k) \geq \alpha, l_k(\hat{x}_k) - l_{k+1}(\hat{x}_k) \geq \alpha$. C'est parce que les lignes en pointillés longs en Fig. 1 (qui sont parallèles aux graphes de l_{k-1}, l_{k+1}) ne peuvent pas s'intersecter en-dessus du graphe de l_k . Par conséquence, $l_k(\hat{x}_k) - l_{k-1}(\hat{x}_k) \geq \max(\alpha, \Delta_k - \alpha), l_k(\hat{x}_k) - l_{k+1}(\hat{x}_k) \geq \max(\alpha, \Delta_k - \alpha)$. La propriété (1) nous donne alors pour tout $n > 0$

$$l_k(\hat{x}_k) - l_{k-n}(\hat{x}_k) \geq \sum_{j=0}^{n-1} \max(\alpha, \Delta_{k-j} - \alpha), \quad l_k(\hat{x}_k) - l_{k+n}(\hat{x}_k) \geq \sum_{j=0}^{n-1} \max(\alpha, \Delta_{k+j} - \alpha), \tag{2}$$

si les indices concernés sont contenus dans K .

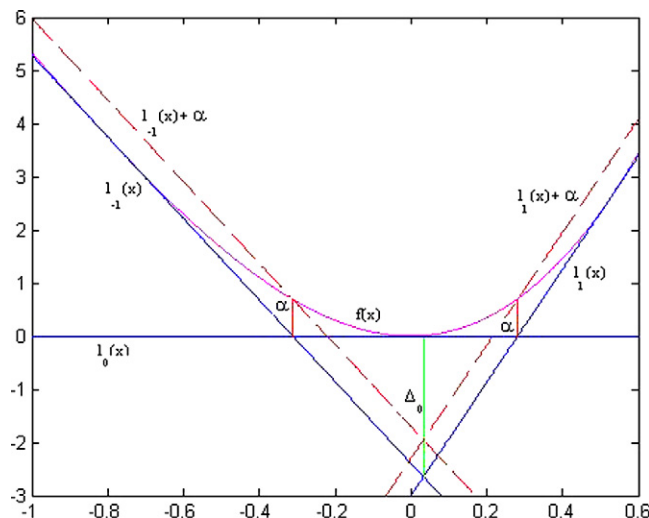


Fig. 1. Fonctions l_{-1}, l_0, l_1 supportant f .

Considérons maintenant un point $x \in D$. Alors il existe $k \in K$ tel que $l_k(x) \leq f(x) \leq l_k(x) + \alpha$. Supposons sans restreindre la généralité que $x \geq \hat{x}_k$ (sinon on a à inverser la séquence des indices autour de k dans les inégalités en-dessous). Alors on obtient $e^{\varphi(x)-f(x)} = \frac{4}{3} \sum_{j \in K} e^{l_j(x)+\delta_j-f(x)}$. On peut maintenant déduire les bornes suivantes sur $l_j(x) - f(x)$. Par (2) et les propriétés de monotonie de $l_{k \pm n}(x) - f(x)$ nous avons

$$l_k(x) - f(x) \leq 0, \quad l_{k+1}(x) - f(x) \leq -\alpha, \quad l_{k+n}(x) - f(x) \leq -\alpha - \sum_{j=1}^{n-1} \max(\alpha, \Delta_{k+j} - \alpha),$$

$$l_{k-n}(x) - f(x) \leq -\sum_{j=0}^{n-1} \max(\alpha, \Delta_{k-j} - \alpha), \quad l_k(x) - f(x) \geq -\alpha, \quad l_{k+1}(x) - f(x) \geq -\Delta_k$$

pour tout $n > 0$, sous condition que les indices concernés sont contenus dans K . Cela nous donne les bornes

$$e^{\varphi(x)-f(x)} \leq \frac{4}{3} (e^{\delta_k} + e^{-\max(\alpha, \Delta_k - \alpha)} e^{\delta_{k-1}} + e^{-\max(\alpha, \Delta_k - \alpha)} e^{-\max(\alpha, \Delta_{k-1} - \alpha)} e^{\delta_{k-2}} + \dots$$

$$+ e^{-\alpha} e^{\delta_{k+1}} + e^{-\alpha} e^{-\max(\alpha, \Delta_{k+1} - \alpha)} e^{\delta_{k+2}} + \dots),$$

$$e^{\varphi(x)-f(x)} \geq \frac{4}{3} (e^{\delta_k} e^{-\alpha} + e^{\delta_{k+1}} e^{-\Delta_k}).$$

Pour évaluer les expressions en-dessus, on aura besoin du lemme suivant :

Lemme 1.1. Soit $c \in (0, 1)$ et soit $\kappa_0, \kappa_1, \kappa_2, \dots$ une séquence finie ou infinie de nombres réels non-négatifs inférieurs ou égaux à c . Alors la somme finie ou infinie $S = (1 - \kappa_0) + \kappa_0(1 - \kappa_1) + \kappa_0\kappa_1(1 - \kappa_2) + \kappa_0\kappa_1\kappa_2(1 - \kappa_3) + \dots$ converge et est inférieure ou égale à 1.

Preuve. Le produit $\prod_{j=0}^k \kappa_j$ est borné par c^{k+1} . Par conséquent, la somme $1 + 2\kappa_0 + 2\kappa_0\kappa_1 + \dots$ est bornée par une progression géométrique décroissante et S converge absolument. Il suit que $S = 1 - \prod_{j=0}^k \kappa_j \leq 1$ si la séquence $\kappa_0, \kappa_1, \dots$ est finie et k est l'index maximal de la séquence, ou $S = 1$ si la séquence est infinie. \square

En utilisant la définition de δ_k et le fait que $e^{-\max(\alpha, \Delta_k - \alpha)} \leq e^{-\alpha} = \frac{1}{2} < 1$, nous pouvons simplifier les bornes sur $\varphi(x) - f(x)$ par ce lemme. Notamment, nous avons $e^{\varphi(x)-f(x)} \leq \frac{4}{3}(1 + e^{-\alpha})$, $e^{\varphi(x)-f(x)} \geq \frac{4}{3}((1 - e^{-\max(\alpha, \Delta_k - \alpha)})e^{-\alpha} + (1 - e^{-\alpha})e^{-\Delta_k})$. On vérifie facilement que $\min_{\Delta_k > 0} [(1 - e^{-\max(\alpha, \Delta_k - \alpha)})e^{-\alpha} + (1 - e^{-\alpha})e^{-\Delta_k}] = e^{-\alpha}(1 + e^{-\alpha})(1 - e^{-\alpha})$. L'insertion de la valeur $\alpha = \ln 2$ nous donne finalement $\varphi(x) - f(x) \leq \ln 2$, $\varphi(x) - f(x) \geq -\ln 2$. Il suit que φ est définie sur D et $\|\varphi - f\|_{C^0(D)} \leq \ln 2$. \square

2. Borne inférieure

Théorème 2.1. Soit $\varepsilon > 0$. Alors il existe une fonction convexe continue $f(x)$ définie sur l'intervalle $[-1, 1]$ tel que pour toute fonction génératrice des cumulants $\varphi(x)$ d'une mesure non-négative on a $\|f - \varphi\|_{C^0([-1,1])} \geq \frac{\ln 2}{2} - \varepsilon$.

Preuve. Notons Φ l'ensemble de toutes les fonctions $\varphi(x)$ qui peuvent être représentées comme $\varphi(x) = \ln \sum_{k=1}^n \alpha_k e^{s_k x}$ avec s_k des nombres réels, α_k des nombres positifs et n fini. Soit Φ' l'ensemble de toutes les fonctions $\varphi(x)$ qui peuvent être représentées comme $\varphi(x) = \ln \sum_{k=1}^n \alpha_k \frac{e^{s_k x} + e^{-s_k x}}{2}$ avec s_k des nombres non-négatifs et α_k des nombres positifs. Alors l'ensemble de toutes les fonctions génératrices des cumulants qui sont définies sur $[-1, 1]$ est contenu dans la clôture de Φ dans la norme $C^0([-1, 1])$. Il est alors suffisant de montrer que pour tout $\varepsilon > 0$ il existe une fonction convexe continue $f(x)$ sur $[-1, 1]$ tel que $\inf_{\varphi \in \Phi} \|f - \varphi\|_{C^0([-1,1])} \geq \frac{\ln 2}{2} - \varepsilon$.

Choisissons $t \in (0, 1)$ et notons $S_t = \{-1, -t, 0, t, 1\} \subset [-1, 1]$, $S'_t = \{0, t, 1\} \subset [-1, 1]$. Pour deux fonctions $f(x), \varphi(x)$ définies sur $[-1, 1]$ nous avons $\max_{x \in S_t} |f(x) - \varphi(x)| \leq \|f - \varphi\|_{C^0([-1,1])}$. Alors $\inf_{\varphi \in \Phi} \|f - \varphi\|_{C^0([-1,1])} \geq \inf_{\varphi \in \Phi} \max_{x \in S_t} |f(x) - \varphi(x)|$. Si $f(x)$ est une fonction paire, $f(x) = f(-x)$, alors pour chaque fonction $\varphi(x)$ définie sur $[-1, 1]$ nous avons que $\max_{x \in S_t} |f(x) - \varphi(x)| \geq \max_{x \in S'_t} |f(x) - \ln \frac{e^{\varphi(x)} + e^{\varphi(-x)}}{2}|$. Alors si la fonction f est paire, nous avons $\inf_{\varphi \in \Phi} \max_{x \in S_t} |f(x) - \varphi(x)| \geq \inf_{\varphi \in \Phi'} \max_{x \in S'_t} |f(x) - \varphi(x)|$. Appelons l'infimum sur le côté droit de cette inégalité $\Delta_t(f)$.

$\Delta_t(f)$ ne dépend que des valeurs de $f(x)$ aux points $x \in S'_t$, c'est-à-dire du vecteur 3-dimensionnel $v(f) = (f(0), f(t), f(1))^T$. Définissons également le vecteur $w(f) = (e^{f(0)}, e^{f(t)}, e^{f(1)})^T$. Alors $\Delta_t(f)$ peut être exprimé comme $\inf_{\varphi \in \Phi'} \|v(f) - v(\varphi)\|_\infty$.

Considérons la famille de fonctions paires convexes continues $f_d(x)$ définies sur $[-1, 1]$ par $f_d(x) = d|x|$, où $d > 0$ est le paramètre de la famille. Nous avons $w(f_d) = (1, e^{dt}, e^d)^T$. D'autre côté, l'ensemble $\{w(\varphi) \mid \varphi \in \Phi'\}$ est un cône convexe, notamment l'enveloppe convexe conique de la courbe $\Gamma = \{(1, \frac{e^{st} + e^{-st}}{2}, \frac{e^s + e^{-s}}{2})^T \mid s \geq 0\}$.

Il n'est pas difficile à vérifier que $w(f_d)$ est situé en dehors de l'enveloppe convexe conique de Γ et que chaque segment de droite entre $w(f_d)$ et un point $w(\varphi)$, $\varphi \in \Phi'$, contient un point représentable comme $\alpha\gamma$, où $\alpha \geq 0$ et $\gamma = (\gamma_1, \gamma_2, \gamma_3)^T \in \Gamma$. Mais cela implique l'existence d'un $\beta \in [0, 1]$ tel que $w_k(f_d) - \alpha\gamma_k = \beta(w_k(f_d) - w_k(\varphi))$ pour $k = 1, 2, 3$, où w_k sont les éléments du vecteur w . Par la monotonie de la fonction exponentielle nous avons alors $|v_k(f_d) - v_k(\varphi)| \geq |v_k(f_d) - \ln \alpha - \ln \gamma_k|$ pour tout $k = 1, 2, 3$, où v_k sont les éléments du vecteur v . Par conséquent, $\|v(f_d) - v(\varphi)\|_\infty \geq \|v(f_d) - (\ln \alpha + \ln \gamma_1, \ln \alpha + \ln \gamma_2, \ln \alpha + \ln \gamma_3)^T\|_\infty$. Il suit que l'infimum de $\|v(f_d) - v(\varphi)\|_\infty$ sur toutes fonctions $\varphi \in \Phi'$ est égal à l'infimum sur les fonctions φ qui satisfont en plus $w(\varphi) = \alpha\gamma$ pour quelques $\alpha \geq 0, \gamma \in \Gamma$.

On peut réécrire cela comme $\Delta_t(f_d) = \inf_{\varphi: w(\varphi) \in \mathbb{R}_+ \Gamma} \|v(f_d) - v(\varphi)\|_\infty$, où $\mathbb{R}_+ \Gamma$ est l'ensemble $\{\alpha\gamma \mid \alpha \geq 0, \gamma \in \Gamma\} \subset \mathbb{R}^3$. En utilisant les descriptions explicites de Γ et de la fonction f_d , nous avons $\Delta_t(f_d) = \inf_{s,c \in \mathbb{R}} \max\{|c|, |c + \ln \frac{e^{st} + e^{-st}}{2} - td|, |c + \ln \frac{e^s + e^{-s}}{2} - d|\}$. Un calcul direct montre que l'infimum est atteint à s donné par $e^s + e^{-s} = 2e^d$ et à c donné par $2c + \ln \frac{e^{st} + e^{-st}}{2} - td = 0$. Si on insère ces valeurs de c et s , on obtient

$$\begin{aligned} \Delta_t(f_d) &= c = \frac{1}{2} \left(td - \ln \frac{e^{td} (1 + \sqrt{1 - e^{-2d}})^t + e^{-td} (1 + \sqrt{1 - e^{-2d}})^{-t}}{2} \right) \\ &= \frac{1}{2} \ln \frac{2}{(1 + \sqrt{1 - e^{-2d}})^t + e^{-2td} (1 + \sqrt{1 - e^{-2d}})^{-t}}. \end{aligned}$$

Posons $t = t(d) = \frac{\log d}{2d}$ et laissons tendre $d \rightarrow \infty$. Alors on obtient $e^{2td} = d, 1 + \sqrt{1 - e^{-2d}} = 2 + O(e^{-2d}), t \rightarrow 0$ et par conséquent $\Delta_{t(d)}(f_d) \rightarrow \frac{\ln 2}{2}$. Il suit que $\Delta_{t(d)}(f_d) > \frac{\ln 2}{2} - \varepsilon$ si d est suffisamment grand, mais $\inf_{\varphi \in \Phi} \|f_d - \varphi\|_{C^0([-1,1])} \geq \Delta_{t(d)}(f_d)$. \square

3. Le cas multi-dimensionnel

Théorème 3.1. Soit $n \geq 2$ un entier et $R > 0$ un nombre réel. Alors il existe une fonction convexe continue $f(x)$ définie sur la boule fermée $B_1 \subset \mathbb{R}^n$ de rayon 1 tel que pour chaque fonction $\varphi(x)$ génératrice des cumulants d'une mesure non-négative n -variée, définie sur B_1 , on a $\|f - \varphi\|_{C^0(B_1)} \geq R$.

Preuve. La preuve est similaire à celle du Théorème 2.1.

Soit $n \geq 2$ et $R > 0$. Notons Φ_n l'ensemble de toutes les fonctions $\varphi(x)$ sur \mathbb{R}^n qui peuvent être représentées comme $\varphi(x) = \ln \sum_{k=1}^K \alpha_k e^{(s_k, x)}$ où $s_k \in \mathbb{R}^n$ sont des vecteurs, α_k des nombres positifs et K est fini. Soit Φ'_n l'ensemble de toutes les fonctions $\varphi(x)$ qui peuvent être représentées comme $\varphi(x) = \ln \sum_{k=1}^K \alpha_k \left(\frac{1}{A(n)} \int_{S^{n-1}} e^{(r_k \sigma, x)} d\sigma \right)$ où r_k sont des nombres non-négatifs et α_k des nombres positifs. Ici S^{n-1} est la sphère normée dans \mathbb{R}^n et $A(n) = \frac{2\pi^{n/2}}{\Gamma(n/2)}$ sa surface. Alors l'ensemble des fonctions génératrices des cumulants qui sont définies sur B_1 est contenu dans la clôture de Φ_n dans la norme $C^0(B_1)$. Alors il suffit de montrer qu'il existe une fonction convexe continue $f(x)$ définie sur B_1 tel que $\inf_{\varphi \in \Phi_n} \|f - \varphi\|_{C^0(B_1)} \geq R$.

Choisissons $t \in (0, 1)$ et posons $S_t = \{0\} \cup tS^{n-1} \cup S^{n-1} \subset B_1$, $S'_t = \{0, te_1, e_1\} \subset B_1$, où $e_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^n$ est le premier vecteur de la base orthonormale canonique. Des arguments similaires à ceux dans la preuve du Théorème 2.1 permettent de conclure que pour chaque deux fonctions $f(x), \varphi(x)$ définies sur B_1 on a $\inf_{\varphi \in \Phi_n} \|f - \varphi\|_{C^0(B_1)} \geq \inf_{\varphi \in \Phi_n} \sup_{x \in S_t} |f(x) - \varphi(x)|$. Si $f(x)$ possède une symétrie sphérique, nous avons d'une manière similaire $\inf_{\varphi \in \Phi_n} \sup_{x \in S_t} |f(x) - \varphi(x)| \geq \inf_{\varphi \in \Phi'_n} \max_{x \in S'_t} |f(x) - \varphi(x)|$. Appelons l'infimum sur le côté droit de cette inégalité $\Delta_t(f)$.

$\Delta_t(f)$ ne dépend que des valeurs de $f(x)$ aux points $x \in S'_t$, c'est-à-dire du vecteur 3-dimensionnel $v(f) = (f(0), f(te_1), f(e_1))^T$. Définissons également le vecteur $w(f) = (e^{f(0)}, e^{f(te_1)}, e^{f(e_1)})^T$. Avec cette notation $\Delta_t(f)$ peut être exprimé comme $\inf_{\varphi \in \Phi'_n} \|v(f) - v(\varphi)\|_\infty$.

Considérons la famille de fonctions convexes continues $f_d(x)$, possédant une symétrie sphérique et définies sur B_1 , donnée par $f_d(x) = d|x|$, $d > 0$. Nous avons $w(f_d) = (1, e^{dt}, e^d)^T$. D'autre côté, l'ensemble $\{w(\varphi) \mid \varphi \in \Phi'_n\} \subset \mathbb{R}^3$ est un cône convexe, notamment l'enveloppe convexe conique de la courbe $\Gamma_n = \{(1, \frac{1}{A(n)} \int_{S^{n-1}} e^{(t e_1, s \sigma)} d\sigma, \frac{1}{A(n)} \int_{S^{n-1}} e^{(e_1, s \sigma)} d\sigma)^T \mid s \geq 0\}$. Le calcul des intégraux nous donne la représentation suivante explicite de la courbe Γ_n . $\Gamma_n = \{(1, g(st), g(s))^T \mid s \geq 0\}$, où $g(s) = {}_0F_1(\frac{n}{2}; \frac{s^2}{4})$, et ${}_0F_1$ est la fonction hypergéométrique confluyente de limite [1]. Pour $n = 1$ ${}_0F_1(\frac{n}{2}; \frac{s^2}{4})$ se simplifie à la fonction $\cosh(s)$ qu'on a rencontré dans la section précédente.

Par de raisons similaires à celles dans la preuve du Théorème 2.1, on obtient que l'infimum de $\|v(f_d) - v(\varphi)\|_\infty$ sur tous $\varphi \in \Phi'_n$ est égal à l'infimum sur des fonctions φ qui satisfont en plus $w(\varphi) = \alpha\gamma$ pour quelques $\alpha \geq 0$, $\gamma \in \Gamma_n$, et $\Delta_t(f_d) = \inf_{\varphi \in \Phi'_n} \|v(f_d) - v(\varphi)\|_\infty$. En utilisant les descriptions explicites de Γ_n et de la fonction f_d , on obtient $\Delta_t(f_d) = \inf_{s \geq 0, c \in \mathbb{R}} \max\{|c|, |c + \ln g(st) - td|, |c + \ln g(s) - d|\}$. Un calcul direct montre que l'infimum est atteint à $s = s(d)$ donné par $\ln g(s) = d$ et à c donné par $2c + \ln g(st) = td$. Si on insère ces valeurs de c et s , on obtient $\Delta_t(f_d) = \frac{1}{2}(t \ln g(s) - \ln g(st))$.

Posons $t = t(s, d) = \frac{1}{\ln s}$ et laissons tendre $d \rightarrow \infty$. Alors on a aussi $s \rightarrow \infty$ et $\Delta_{t(d)}(f_d) = \frac{1}{2}(\frac{\ln g(s)}{\ln s} - \ln g(\frac{s}{\ln s}))$. Pour $s \rightarrow \infty$ on obtient l'expansion asymptotique $\Delta_{t(d)}(f_d) \sim \frac{n-1}{4}(\ln s - \ln \ln s) + O(1)$ [1], et $\Delta_{t(d)}(f_d)$ est croissant vers l'infini. Par conséquent $\Delta_{t(d)}(f_d) > R$ si d est suffisamment grand, mais $\inf_{\varphi \in \Phi_n} \|f_d - \varphi\|_{C^0(B_1)} \geq \Delta_{t(d)}(f_d)$. \square

Références

- [1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, ninth Dover printing, tenth GPO printing edition, Dover, New York, 1964.
- [2] N.I. Akhiezer, The Classical Moment Problem and Some Related Questions in Analysis, FizMatLit, Moscow, 1961.
- [3] L.D. Brown, Fundamentals of Statistical Exponential Families, IMS Lecture Notes, vol. 9, Institute of Mathematical Statistics, Hayward, CA, 1986.
- [4] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, second ed., Springer, New York, 1998.
- [5] E.W. Weisstein, Cumulant-generating function, From MathWorld – A Wolfram Web Resource.
- [6] D.V. Widder, The Laplace Transform, Princeton Mathematical Series, vol. 6, Princeton University Press, Princeton, 1946.