The Maximum Likelihood Estimation of Uniform Distribution Support

Tomasz Schreiber,
Faculty of Mathematics and Computer Science,
Nicholas Copernicus University,
ul. Chopina 12 /18, 87-100 Toruń, Poland,
e-mail: tomeks@mat.uni.torun.pl

September 8, 2000

Abstract: *Let $Z_1, Z_2, ...$ be a sequence of i.i.d. $\mathbb{R}^k$-valued random vectors distributed uniformly on a star-shaped set A_0 belonging to a certain known class $\mathcal{A}$ of star-shaped sets satisfying some regularity assumptions. For A_0 we consider the maximum likelihood estimator $\hat{A}_n$ which minimizes the volume among the sets from $\mathcal{A}$ covering the whole sample $Z_1, Z_2, ...$ The purpose of the paper is to investigate the asymptotic properties of $\hat{A}_n$. We prove a functional limit theorem for the radius-vector functions of $\hat{A}_n$. In particular we show that $n\lambda(\hat{A}_n \triangle A_0)$ converges in distribution and we find the limit.*

Keywords: *support estimation, radius-vector function, point processes*

1 Introduction and main results

Let $Z_1, Z_2, ...$ be a sequence of i.i.d. $\mathbb{R}^k$-valued random vectors defined on a probability space $(\Omega, \Im, P)$. Suppose that the distribution of Z_i is uniform and is supported by a compact set $A_0 \subset \mathbb{R}^k$ with a positive volume. The problem of the estimation of A_0 has been widely considered in the literature. For A_0 convex the estimator which was particularly popular since the pioneering works of Rényi & Sulanke [16] and Efron [5] was the convex hull $\mathrm{conv}\{Z_1, ..., Z_n\}$ of the sample $Z_1, ..., Z_n$. A natural functional which characterises the quality of this approximation is the defect of volume of the convex hull

$$D_n = \lambda(A_0) - \lambda(\mathrm{conv}\{Z_1, ..., Z_n\}) = \lambda(A_0 \triangle \mathrm{conv}\{Z_1, ..., Z_n\}),$$

where λ denotes the Lebesgue measure. Rényi and Sulanke proved in their paper that for $A_0 \subset \mathbb{R}^2$ with smooth boundary D_n satisfies the following asymptotic relation

$$ED_n \sim cn^{-\frac{2}{3}},$$

where c is a constant depending on A_0. A survey of these and related results can be found in Schneider [18].

Very strong results concerning the asymptotic behaviour of the convex hull and its functionals have been obtained (in choronological order) by Groeneboom [6], Khamdamov & Nagaev [10], Cabo & Groeneboom [2], Hsing [7] and Braeker & Hsing [1].

For results in higher-dimensional case see e.g. Kuefer [14] and references therein. It has been proven that for $A_0 \subset \mathbb{R}^k$

$$ED_n \sim c_k n^{-\frac{2}{k+1}}$$

with some constants c_k. This means that the convergence is extremely slow in high dimensions.

In order to construct a better estimator of A_0 we have to assume some additional a priori knowledge about this set. We assume that A_0 belongs to a certain known class $\mathcal{A}$ of compact sets in $\mathbb{R}^k$. Leaving for a while existence and measurability questions we can consider for A_0 the maximum likelihood estimator $\hat{A}_n$ which minimizes the Lebesgue measure among the sets belonging to $\mathcal{A}$ and covering the whole sample. It is natural to expect that for $\mathcal{A}$ properly chosen such an estimator is consistent and has a good asymptotic performance. For example if we take $\mathcal{A}$ to be the class of convex polygons on the plane, $\hat{A}_n$ is once more the convex hull of the sample, but we get here for A_0 with r vertices

$$ED_n \sim \frac{2}{3} r \frac{\log n}{n},$$

which was originally proved by Rényi and Sulanke. There is however a trade-off between the richness of the class $\mathcal{A}$ and the asymptotic performance of the estimator, so the question arises how to find a reasonable compromise.

Mammen and Tsybakov [12] considered $\mathcal{A}$ satisfying certain smoothness conditions (they did not assume convexity) and suggested a certain modification of the maximum likelihood principle to construct the estimator of A_0. If we denote by D_n the volume of the symmetric difference of A_0 and the estimator, they prove in particular that under their assumptions

$$ED_n = O(n^{-\frac{\gamma}{\gamma+k-1}}),$$

where γ is the so called "smoothness parameter" for $\mathcal{A}$ in the sense of Dudley [4].

In this paper we propose another regularity assumptions to be imposed on class $\mathcal{A}$. Under these assumptions we prove that for the ML-estimator $\hat{A}_n$, defined as above, sequences

$$n(\lambda(A_0) - \lambda(\hat{A}_n))$$

and

$$n\lambda(A_0 \triangle \hat{A}_n)$$

converge in distribution and we find their limits. This means that the convergence is extremely fast in this case. It is also worth observing that the rate of the convergence does not depend on the dimension.

One of the assumptions that we make about the class $\mathcal{A}$ is that all the sets belonging to this class are star-shaped with respect to 0, so that they can be represented by the corresponding radius-vector functions.

The emphasis is on establishing the functional limit theorem for the radius-vector function of the estimator $\hat{A}_n$. The theorem allows us to prove the aforementioned results as well as to find the limiting distributions for certain functionals of $\hat{A}_n$.

It is worth noting that the ideas underlying the methods used to prove this theorem turn out to have a lot in common with those presented in the paper of Molchanov [13].

We assume that

(R0) Each set $A \in \mathcal{A}$ is compact and star-shaped with respect to 0 (i.e. $\alpha A \subset A$ for $0 \leq \alpha \leq 1$). Moreover, $0 \in \mathrm{Int}(A)$.

(R1) $\mathcal{A}$ is closed in Hausdorff metric,

(R2) $\mathcal{A}$ is invariant with respect to scalar multiplication, i.e. $\alpha A \in \mathcal{A}$ for $A \in \mathcal{A}$ and $\alpha > 1$.

Let $\mathcal{S}_{k-1}$ be the unit sphere in $\mathbb{R}^k$. For each compact set $A \subset \mathbb{R}^k$ star-shaped with respect to 0 by $R_A : \mathcal{S}_{k-1} \to \mathbb{R}$ we denote its radius-vector function, i.e.

$$R_A(x) = \max\{\alpha \mid \alpha x \in A\}.$$

Define also $F_A : \mathcal{S}_{k-1} \to \mathbb{R}$ by

$$F_A(x) = \frac{R_{A_0}(x) - R_A(x)}{R_{A_0}(x)}.$$

It is not difficult to observe that the set A is uniquely determined by F_A. Note also that F_A is bounded because A is compact.

Having established the equivalence between A and F_A for each compact star-shaped set A we can now represent the class $\mathcal{A}$ by the corresponding functional class

$$\mathcal{F}_1 = \{F_A \mid A \in \mathcal{A}\}.$$

We impose on class $\mathcal{A}$ one more regularity assumption expressed in terms of $\mathcal{F}_1$

(REG) There exists a positive constant L_0 such that for any $f \in \mathcal{F}_1$ and $x, y \in \mathcal{S}_{k-1}$ we have $|f(x) - f(y)| \leq L_0 d(x, y) \sup_{z \in \mathcal{S}_{k-1}} |f(z)|$, where d denotes the Euclidean distance.

It is easily seen that condition (**REG**) implies in particular that F_A is continuous for $A \in \mathcal{A}$.

To proceed, we put for $n \in \mathbb{N}$

$$\mathcal{F}_n = n\mathcal{F}_1 = \{nF_A \mid A \in \mathcal{A}\}.$$

It is natural to define the limiting class

$$\mathcal{F}_\infty = \{\lim_{i\to\infty} f_i \mid f_i \in \mathcal{F}_{n_i}\ ,\ n_1 < n_2 < n_3 < ...\},$$

where the limits are taken in the sense of the supremum-norm $||.||$ on the space $C(\mathcal{S}_{k-1};\mathbb{R})$ of all the real-valued continuous functions defined on $\mathcal{S}_{k-1}$. $\mathcal{F}_\infty$ may be regarded to some extent as a class tangent to $\mathcal{F}_1$ at 0. Note also that by condition (**R1**) the classes $\mathcal{F}_n$ are closed in $||.||$ for $n \in \mathbb{N} \cup \{\infty\}$.

Condition (**REG**) can be now expressed in the following equivalent form

(REG') There exists a positive constant L_0 such that for any $n \in \mathbb{N} \cup \{+\infty\}$, $f \in \mathcal{F}_n$ and $x, y \in \mathcal{S}_{k-1}$ we have $|f(x) - f(y)| \leq L_0||f||d(x,y)$.

The conditions (**REG**) and (**REG**)$'$ turn out to be rather restrictive. Indeed, combined with Arzelá-Ascoli lemma they imply the relative compactness of bounded subsets of $\mathcal{F}_\infty$. In particular, in many natural cases where $\mathcal{F}_\infty$ is a linear space (see Section 4) we conclude that dim$\mathcal{F}_\infty < \infty$ which is not far from the parametric setting.

Let $\mathcal{B}_\mathcal{A}$ denote the Borel σ-field on $\mathcal{A}$ corresponding to the Hausdorff distance ρ_H. As an estimator of A_0 we suggest the random element $\hat{A}_n : (\Omega, \Im) \to (\mathcal{A}, \mathcal{B}_\mathcal{A})$ such that

$$\lambda(\hat{A}_n) = \min\{\lambda(A) \mid A \in \mathcal{A},\ \{Z_1, ..., Z_n\} \subset A\},$$

i.e. $\hat{A}_n$ minimizes the Lebesgue measure among the sets from $\mathcal{A}$ covering the whole sample. From condition (**REG**) and the Arzelá-Ascoli lemma it follows that bounded subsets of $\mathcal{F}_1$ are relatively compact. Hence condition (**R1**) implies that that the set argmin$\{\lambda(A) \mid A \in \mathcal{A},\ \{Z_1, ..., Z_n\} \subset A\}$ is nonempty with probability 1. If this set contains more than one element, we choose $\hat{A}_n$ in an arbitrary way. In order to avoid technical measurability questions we simply assume that σ-field $\Im$ is rich enough so that $\hat{A}_n$ can be chosen to be measurable.

For any real-valued function $f : \mathcal{S}_{k-1} \to \mathbb{R}$ we consider its epigraph and hypograph defined by

$$\text{epi}(f) = \{(x,y) \in \mathcal{S}_{k-1} \times \mathbb{R} \mid y \geq f(x)\}$$

and

$$\text{hypo}(f) = \{(x,y) \in \mathcal{S}_{k-1} \times \mathbb{R} \mid y < f(x)\}.$$

Put also

$$\text{epi}(0) = \{(x,y) \in \mathcal{S}_{k-1} \times \mathbb{R} \mid y \geq 0\}$$

and

$$\text{hypo}(0) = \{(x, y) \in \mathcal{S}_{k-1} \times \mathbb{R} \mid y < 0\}.$$

On $\mathcal{S}_{k-1}$ we define the probability measure τ to be the distribution of the random vector $\frac{Z_1}{|Z_1|}$. Note that

$$\frac{d\tau}{ds}(x) = \frac{R^k_{A_0}(x)}{k\lambda(A_0)}$$

where s denotes the surface measure on $\mathcal{S}_{k-1}$. Further, let τ^+ be the measure on $\mathcal{S}_{k-1} \times \mathbb{R}$ defined as the product

$$\tau^+ = k\tau \times \lambda_1((.) \cap \mathbb{R}_+),$$

where λ_1 is the one-dimensional Lebesgue measure. Note that τ^+ concentrates on epi(0). Let Π_{τ^+} be the Poisson point process on $\mathcal{S}_{k-1} \times \mathbb{R}$ with intensity measure τ^+.

The main result of the paper is

Theorem 1 *Let*

$$\mathcal{T} = \sup\{\int_{\mathcal{S}_{k-1}} f(x)\tau(dx) \mid f \in \mathcal{F}_\infty,\ \Pi_{\tau^+}(hypo(f)) = 0\}.$$

Then $\mathcal{T}$ is a random variable, $P(\mathcal{T} < \infty) = 1$ and the sequence

$$n(\lambda(A_0) - \lambda(\hat{A}_n))$$

converges in distribution to $\lambda(A_0)k\mathcal{T}$. Moreover, there exists a probability space $(\hat{\Omega}, \hat{\Im}, \hat{P})$ carrying versions of $Z_1, Z_2, \ldots$ and such that for $\hat{P}$-almost all $\omega \in \hat{\Omega}$ from any subsequence of $(nF_{\hat{A}_n}(\omega))_{n=1}^\infty$ we may extract a further subsequence converging in $||.||$ to some $f \in \mathcal{F}_\infty$ such that $\int_{\mathcal{S}_{k-1}} f(x)\tau(dx) = \mathcal{T}(\omega)$.

It is worthwile observing that the random variable $\mathcal{T}$ is the solution of the extreme problem which may be regarded to some extent as an asymptotic equivalent of the original problem of minimizing the Lebesgue measure among the sets from $\mathcal{A}$ covering the whole sample, whose solution is $\hat{A}_n$. If we assume in addition that the solution of the limiting problem is unique, we obtain immediately

Corollary 1 *Assume that the set*

$$argmax_f\{\int_{\mathcal{S}_{k-1}} f(x)\tau(dx) \mid f \in \mathcal{F}_\infty,\ \Pi_{\tau_+}(hypo(f)) = 0\ \}$$

is a singleton with probability 1. *Then*

$$\Psi = argmax_f\{\int_{\mathcal{S}_{k-1}} f(x)\tau(dx) \mid f \in \mathcal{F}_\infty,\ \Pi_{\tau_+}(hypo(f)) = 0\ \}$$

is a $C(\mathcal{S}_{k-1};\ \mathbb{R})$*-valued random element,*

$$\mathcal{T} = \int_{\mathcal{S}_{k-1}} \Psi(x)\tau(dx)$$

and the sequence $(nF_{\hat{A}_n})$ *converges in distribution to* Ψ *on* $C(\mathcal{S}_{k-1};\ \mathbb{R})$*.*

Using Lemma 2, which expresses the voulume of a star-shaped set A in terms of F_A, we get also

Corollary 2 *Under assumptions of Corollary 1 the sequence*

$$n\lambda(A_0 \triangle \hat{A}_n)$$

converges in distribution to $\lambda(A_0)k\int_{\mathcal{S}_{k-1}} |\Psi(x)|\tau(dx)$*.*

The paper is organized as follows. In the following section we give the proof of the main theorem. Further, we consider a special case where class $\mathcal{A}$ consists of convex sets only. This allows us to obtain some results about the asymptotic properties of the surface measure of $\hat{A}_n$. Finally in the last section, as an example, we find an explicit representation of $\mathcal{F}_\infty$ for certain classes of ellipsoids.

2 Proof of Theorem 1

To proceed we need some notation. We denote by $M(\mathcal{S}_{k-1} \times \mathbb{R})$ the space of all the boundedly finite integer-valued Borel measures on $\mathcal{S}_{k-1} \times \mathbb{R}$. In other words a Borel measure μ on $\mathcal{S}_{k-1} \times \mathbb{R}$ belongs to $M(\mathcal{S}_{k-1} \times \mathbb{R})$ iff $\mu(A) \in \mathbb{N} \cup \{+\infty\}$ for each Borel set A and, in addition, $\mu(A) < \infty$ for A bounded. The elements of $M(\mathcal{S}_{k-1} \times \mathbb{R})$ will be referred to as counting measures on $\mathcal{S}_{k-1} \times \mathbb{R}$.

We endow $M(\mathcal{S}_{k-1} \times \mathbb{R})$ with the σ-field $\mathcal{M}(\mathcal{S}_{k-1} \times \mathbb{R})$ defined to be the least σ-field with respect to which the functions $\varsigma_A : M(\mathcal{S}_{k-1} \times \mathbb{R}) \to \mathbb{N} \cup \{+\infty\}$,

$$\varsigma_A(\mu) = \mu(A),$$

are measurable. By a point process on $\mathbb{S}_{k-1} \times \mathbb{R}$ we shall mean any random element taking values in $(M(\mathcal{S}_{k-1} \times \mathbb{R}), \mathcal{M}(\mathcal{S}_{k-1} \times \mathbb{R}))$.

We say that a sequence μ_n of counting measures converges vaguely to a counting measure μ iff

$$\lim_{n\to\infty} \int f(x)\mu_n(dx) = \int f(x)\mu(dx)$$

for each f continuous with compact support. It is worth noting that, the bounded subsets of $\mathcal{S}_{k-1} \times \mathbb{R}$ being relatively compact, the vague convergence coincides in this case with the $\hat{w}$-convergence (see e.g. Section A2.6 in Daley & Vere-Jones [3]).

Consider a sequence Ψ_n of point processes on $\mathcal{S}_{k-1} \times \mathbb{R}$. We say that Ψ_n converges weakly to a point process Ψ and we write $\Psi_n \rightharpoonup \Psi$ iff

$$\lim_{n\to\infty} E(g(\Psi_n)) = E(g(\Psi))$$

for any $g : M(\mathrm{S}_{k-1} \times \mathbb{R}) \to \mathbb{R}$ bounded and continuous with respect to the vague topology on $M(\mathcal{S}_{k-1} \times \mathbb{R})$ (see Section 1.3 in Karr [8], Chapter 4 in Kerstan, Matthes & Mecke [9] or Section 9.1 in Daley & Vere-Jones [3]).

For any σ-finite Borel measure ν we denote by Π_ν the Poisson point process with intensity measure ν. Further, let $\mathcal{B}_n$ for $n \in \mathbb{N}$ be the binomial point process corresponding to the sample $\{Z_1, ..., Z_n\}$, i.e.

$$\mathcal{B}_n(A) = \sum_{i=1}^{n} \mathbf{1}_A(Z_i).$$

It turns out that the Poisson point process Π_{τ^+}, where τ^+ is defined as in the previous section, describes the local behaviour of the sample in the neighbourhood of the boundary ∂A_0. To make it more precise, we need first to introduce some additional notations. For each $\alpha \in \mathbb{R}$ consider a function $\gamma_\alpha : \mathbb{R}^k \setminus \{0\} \to \mathcal{S}_{k-1} \times \mathbb{R}$ such that

$$\gamma_\alpha(z) = (\frac{z}{|z|}, r_{A_0}(z)),$$

where

$$r_{A_0}(z) = \frac{R_{A_0}(\frac{z}{|z|}) - |z|}{R_{A_0}(\frac{z}{|z|})}.$$

It is worth noting that for $\alpha = 1$ the graph Γ_{F_A} of function F_A coincides with $\gamma_1(\partial A)$, since for $z \in \partial A$ we have $|z| = R_A(\frac{z}{|z|})$. More generally, $\Gamma_{(\alpha F_A)}$ coincides with $\gamma_\alpha(\partial A)$. Observe also that ∂A_0 is mapped by γ_α onto the unit sphere $\mathcal{S}_{k-1} \times \{0\}$ for each α. In other words, the functions γ_α allow us to represent the neighbourhoods of ∂A_0 by the neighbourhoods of $\mathcal{S}_{k-1} \times \{0\}$ on the surface of the cylinder $\mathcal{S}_{k-1} \times \mathbb{R}$. We are now ready to state our lemma, which follows easily from classical results.

Lemma 1

$$\gamma_n(\mathcal{B}_n) \rightharpoonup_{n\to\infty} \Pi_{\tau^+}.$$

Proof According to proposition 3.21 in Resnick [17] the lemma will be proved if we show that the sequence of measures

$$nP(\gamma_n(Z_1) \in (.))$$

converges weakly to τ^+. In order to prove it we need a formula expressing $\lambda(A)$ in terms of F_A. It is easy to verify that

Lemma 2 *For each set A star-shaped with respect to* 0 *we have*

$$\lambda(A) = \lambda(A_0) \int_{\mathcal{S}_{k-1}} (1 - F_A(x))^k \tau(dx).$$

Now let f be a bounded measurable function from $\mathcal{S}_{k-1}$ to $\mathbb{R}_+$ and let $A \subset A_0$ be a star-shaped set such that $F_A = n^{-1} f$. We get then

$$nP(\gamma_n(Z_1) \in \text{hypo}(f)) = n \frac{\lambda(A_0) - \lambda(A)}{\lambda(A_0)} = n(1 - \int_{\mathcal{S}_{k-1}} (1 - \frac{f(x)}{n})^k \tau(dx)) =$$

$$= k \int_{\mathcal{S}^{k-1}} f(x)\tau(dx) + O(n^{-1}) = \tau^+(\text{hypo}(f)) + O(n^{-1}).$$

The desired weak convergence now follows easily. Lemma 1 is proved. □

Since $\mathcal{S}_{k-1} \times \mathbb{R}$ is separable and complete, the space $M(\mathcal{S}_{k-1} \times \mathbb{R})$ endowed with vague topology may be metrised in the way that makes it complete and separable. Moreover, the Borel σ-field generated on $M(\mathcal{S}_{k-1} \times \mathbb{R})$ by the vague topology coincides with $\mathcal{M}(\mathcal{S}_{k-1} \times \mathbb{R})$ (see e.g. Theorem 1.9.6 in Kerstan, Matthes & Mecke [9] or Theorem A2.6.III in Daley & Vere-Jones [3]). Thus, the Skorohod representation theorem holds and in Lemma 1 we can replace the convergence in distribution by the pointwise one, eventually changing the probability space.

Corollary 3 *There exists a probability space $(\hat{\Omega}, \hat{F}, \hat{P})$ carrying the versions of $\mathcal{B}_n$ and Π_{τ^+} such that for $\hat{P}$-almost every $\omega \in \hat{\Omega}$*

$$\gamma_n(\mathcal{B}_n)(\omega) \Rightarrow_{n\to\infty} \Pi_{\tau^+}(\omega)$$

with $\Rightarrow$ denoting the vague convergence of counting measures.

For a continuous function $f : \mathcal{S}_{k-1} \to \mathbb{R}$ set

$$T_\infty(f) = k \int_{\mathcal{S}_{k-1}} f(x)\tau(dx)$$

and

$$T_n(f) = n \int_{\mathcal{S}_{k-1}} (1 - (1 - \frac{f(x)}{n})^k)\tau(dx).$$

Note that, according to Lemma 2, for $A \in \mathcal{A}$

$$\lambda(A_0)T_n(nF_A) = n(\lambda(A_0) - \lambda(A)).$$

Moreover it is easily seen that

Lemma 3 *The sequence T_n converges uniformly to T_∞ as $n \to \infty$ on bounded subsets of $C(\mathcal{S}_{k-1}; R)$.*

For a counting measure μ on Λ_{A_0} concentrated on epi(0) we define

$$\Theta_\infty(\mu) = \max\{T_\infty(f) \mid f \in \mathcal{F}_\infty,\ \mu(\mathrm{hypo}(f)) = 0\}$$

and

$$\Theta_n(\mu) = \max\{T_n(f) \mid f \in \mathcal{F}_n,\ \mu(\mathrm{hypo}(f)) = 0\}.$$

Note that $\Theta_\infty(\mu), \Theta_n(\mu) \geq 0$. Moreover, the separability of $C(\mathcal{S}_{k-1}; R)$ implies that the mappings Θ_n are measurable for $n \in \mathbb{N} \cup \{\infty\}$. It is easily seen that

Lemma 4

$$\frac{n(\lambda(A_0) - \lambda(\hat{A}_n))}{\lambda(A_0)} = T_n(nF_{\hat{A}_n}) = \Theta_n(\gamma_n(\mathcal{B}_n))$$

and

$$k\mathcal{T} = \Theta_\infty(\Pi_{\mathcal{T}^+}).$$

We introduce here one more technical definition.

Definition 1 *We say that a counting measure μ on $\mathcal{S}_{k-1} \times \mathbb{R}$ concentrated on epi(0) generates on $\mathcal{S}_{k-1}$ an (ϵ, K)-net for some $\epsilon > 0$ and $K > 0$ if for any open ball $B \subset \mathcal{S}_{k-1}$ with radius ϵ we have $\mu(B \times [0, K)) > 0$.*

It is not difficult to observe that for almost every $\omega \in \Omega$ for every $\epsilon > 0$ we can find $K := K_{(\epsilon,\omega)}$ such that the counting measure $\Pi_{\mathcal{T}^+}(\omega)$ generates an (ϵ, K)-net. Therefore, in view of Corollary 3 and Lemma 4 the proof of Theorem 1 will be completed if we prove the following crucial lemma showing that Θ_n converges continuously to Θ_∞ as $n \to \infty$ under an additional regularity condition.

Lemma 5 *Let the sequence μ_n of counting measures on $\mathcal{S}_{k-1} \times \mathbb{R}$ concentrated on epi(0) converge vaguely to a counting measure μ_∞ and assume that for each $\epsilon > 0$ there exists $K = K_\epsilon$ such that μ_∞ generates an (ϵ,K)-net. Then*

- **(A1)** $\Theta_\infty(\mu_\infty) < \infty$
- **(A2)** $\Theta_\infty(\mu_\infty) = \lim_{n\to\infty} \Theta_n(\mu_n)$.
- **(A3)** *For each $n \in \mathbf{N} \cup \{\infty\}$ there exists $\phi_n \in \mathcal{F}_n$ such that $T_n(\phi_n) = \Theta_n(\mu_n)$.*
- **(A4)** *if a sequence $\psi_n \in \mathcal{F}_n$ is such that $T_n(\psi_n) = \Theta_n(\mu_n)$ then from any of its subsequences we can extract a further subsequence converging in $||.||$ to some $\psi_\infty \in \mathcal{F}_\infty$ such that $T_\infty(\psi_\infty) = \Theta_\infty(\mu_\infty)$.*

Proof At first we investigate some properties of classes $\mathcal{F}_n$ for $n \in \mathbb{N}$. Let $f \in \mathcal{F}_n$ be such that $f = nF_A$ for certain $A \in \mathcal{A}$ and let $\epsilon > 0$. Then, according to condition (**R2**), $B = (1 + \frac{\epsilon}{n})A$ also belongs to $\mathcal{A}$ and

$$nF_B = (1 + \frac{\epsilon}{n})f - \epsilon$$

is an element of $\mathcal{F}_n$. Thus we have proven

Lemma 6 *For $f \in \mathcal{F}_n$ and $\epsilon > 0$ we have also $((1 + \frac{\epsilon}{n})f - \epsilon) \in \mathcal{F}_n$.*

As a corollary we get immediately

Corollary 4 *For $f \in \mathcal{F}_\infty$ also $(f - \epsilon) \in \mathcal{F}_\infty$ for $\epsilon > 0$.*

Further, as a simple consequence of condition (**REG**$'$) and the Arzelà-Ascoli lemma we obtain

Lemma 7 *The subsets of $\bigcup_{n=1}^{\infty} \mathcal{F}_n \cup \mathcal{F}_\infty$ bounded in $||.||$ are relatively compact.*

The next step will be to prove that $\Theta_\infty(\mu_\infty) < \infty$. Choose an arbitrary $h \in \mathcal{F}_\infty$ such that $\mu_\infty(\text{hypo}(h)) = 0$. Assume that $T_\infty(h) \geq 0$, because otherwise, since $0 \in \mathcal{F}_\infty$, we may ignore h when determining $\Theta_\infty(\mu_\infty)$. Let $h^+ = \max(h, 0)$ and $h^- = -\min(h, 0)$. Condition **(REG')** and the assumptions of the lemma imply that for any $\epsilon > 0$ there exists K such that measure μ_∞ generates an (ϵ, K)-net. Hence

$$||h^+|| = \sup_{x \in \mathcal{S}_{k-1}} h^+(x) \leq K + L_0\epsilon||h||.$$

Let $x_0 \in \mathcal{S}_{k-1}$ be such that $h^-(x_0) = -||h^-||$. Then from condition (**REG**$'$) it follows that

$$h^-(x) \leq -\frac{||h^-||}{2}$$

for $x \in \mathcal{S}_{k-1},\ d(x_0, x) \leq \frac{||h^-||}{2L_0||h||}$. Choose $\sigma > 0$ such that for any closed ball $B \subset \mathcal{S}_{k-1}$

$$\tau(B) \geq \sigma(\text{diam}B)^{k-1}.$$

The existence of such σ follows from $0 \in \text{Int}(A_0)$. We get

$$0 \leq T_\infty(h) = k\int_{\mathcal{S}_{k-1}} h(x)\tau(dx) \leq k||h^+|| - k\sigma\frac{||h^-||}{2}(\frac{||h^-||}{L_0||h||})^{k-1}.$$

Thus

$$||h^-|| \leq (\frac{2L_0^{k-1}||h||^{k-1}||h^+||}{\sigma})^{\frac{1}{k}} \leq$$
$$\leq (\frac{2L_0^{k-1}||h||^{k-1}(K + L_0\epsilon||h||)}{\sigma})^{\frac{1}{k}}.$$

Now, since $||h|| = \max(||h^-||, ||h^+||)$,

$$||h|| \leq \max(K + L_0\epsilon||h||, (\frac{2L_0^{k-1}||h||^{k-1}(K + L_0\epsilon||h||)}{\sigma})^{\frac{1}{k}}).$$

Taking ϵ sufficiently small we obtain a uniform bound for $||h||$. Therefore $\Theta_\infty(\mu_\infty) < \infty$, which proves assertion **(A1)**.

In particular, according to Lemma 7 and taking into account the closedness of $\mathcal{F}_\infty$ in $||.||$ we can infer that there exists $\phi_\infty \in \mathcal{F}_\infty$ such that

$$T_\infty(\phi_\infty) = \Theta_\infty(\mu_\infty)$$

and

$$\mu(\text{hypo}(\phi_\infty)) = 0.$$

Indeed, by definition of Θ_∞ this follows easily from continuity of T_∞ and closedness of the set $\{f \in \mathcal{F}_\infty \mid \mu_\infty(\text{hypo}(f)) = 0\}$.

To proceed, note that since $\mu_n \Rightarrow \mu_\infty$ the fact that μ_∞ generates an (ϵ, K)-net implies that the measure μ_n generates a $(2\epsilon, K)$-net for n large enough. Indeed, covering $\mathcal{S}_{k-1}$ with a finite number of open balls $B_1, ..., B_m$ with radius ϵ, for sufficiently large n we have $\mu_n(B_i \times [0, K)) > 0$ for $i = 1, ..., m$. This proves our assertion.

When investigating $\Theta_n(\mu_n)$ we can, similarly as above, restrict our attention to such functions $h \in \mathcal{F}_n$ such that $T_n(h) \geq 0$. It is however easy to check that $T_\infty(h) \geq T_n(h)$, so we can consider only such $h \in \mathcal{F}_n$ for which $T_\infty(h) \geq 0$.

Thus, we can repeat for μ_n, T_n and Θ_n the arguments which we used above for μ_∞, T_∞ and Θ_∞. In this way we show that for sufficiently large n

$$\Theta_n(\mu_n) < \infty$$

and there exists ϕ_n such that

$$T_n(\phi_n) = \Theta_n(\mu_n).$$

This proves assertion **(A3)** of the lemma.

Now, choose a sequence $\psi_n \in \mathcal{F}_n$ such that

$$T_n(\psi_n) = \Theta_n(\mu_n)$$

and

$$\mu_n(\text{hypo}(\psi_n)) = 0$$

for n large enough. Applying once more the arguments used to prove assertions **(A1)** and **(A3)** we get

$$\sup_n ||\psi_n|| < \infty.$$

Since, according to Lemma 7, the bounded subsets of $\bigcup_{n=1}^\infty \mathcal{F}_n$ are relatively compact in $||.||$, from any subsequence (n') we may choose a further subsequence

(n'') such that $\psi_{n''}$ converges uniformly to some $\psi_\infty \in \mathcal{F}_\infty$. Obviously, by Lemma 3

$$T_\infty(\psi_\infty) = \lim_{n''\to\infty} \Theta_{n''}(\mu_{n''}).$$

It is not difficult to prove that $\mu_\infty(\text{hypo}(\psi_\infty)) = 0$. Indeed, choose $M > \sup_{n''} ||\psi_{n''}||$ such that $\mu_\infty(\mathcal{S}_{k-1} \times \{M\}) = 0$ and let $\Xi_{n''}$ and Ξ_∞ be the sets of atoms respectively of $\mu_{n''}$ and μ_∞ on $\mathcal{S}_{k-1} \times [0, M]$. Obviously then $\lim_{n''\to\infty} \rho_H(\Xi_{n''}, \Xi_\infty) = 0$. Hence

$$\rho_H(\Xi_\infty, \text{epi}(\psi_\infty) \cap (\mathcal{S}_{k-1} \times [0, M])) =$$

$$= \lim_{n''\to\infty} \rho_H(\Xi_{n''}, \text{epi}(\psi_{n''}) \cap (\mathcal{S}_{k-1} \times [0, M])) = 0,$$

which proves our assertion. Therefore

$$\Theta_\infty(\mu_\infty) \geq T_\infty(\psi_\infty) = \lim_{n''\to\infty} \Theta_n(\mu_{n''}).$$

Since (n') was arbitrary,

$$\Theta_\infty(\mu_\infty) \geq \limsup_{n\to\infty} \Theta_n(\mu_n).$$

Thus, to establish assertions **(A2)** and **(A4)** of the lemma, it remains to prove that

$$\Theta_\infty(\mu_\infty) \leq \lim\inf_{n\to\infty} \Theta_n(\mu_n).$$

Let ϕ_∞ be such as in **(A3)**, i.e.

$$T_\infty(\phi_\infty) = \Theta_\infty(\mu_\infty)$$

and

$$\mu_\infty(\text{hypo}(\phi_\infty)) = 0.$$

From Lemma 4 we conclude that for arbitrary $\eta > 0$

$$f_{\infty,\eta} = \phi_\infty - \eta$$

belongs to $\mathcal{F}_\infty$. Note that

$$\Theta_\infty(\mu_\infty) = T_\infty(f_{\infty,\eta}) + k\eta$$

and

$$\mu_\infty(\text{cl}(\text{hypo}(f_{\infty,\eta}))) = 0.$$

This means that $\mu_n(\text{hypo}(f_{\infty,\eta})) = 0$ for sufficiently large n. On the other hand, for sufficiently large n we may choose $f_{n,\eta} \in \mathcal{F}_n$ such that

$$||f_{n,\eta} - f_{\infty,\eta}|| < \eta.$$

Lemma 6 implies that

$$g_{n,\eta} = ((1+\frac{2\eta}{n})f_{n,\eta} - 2\eta) \in \mathcal{F}_n.$$

Obviously, $||g_{n,\eta} - f_{\infty,\eta}|| < 3\eta + \frac{2\eta}{n}||f_{n,\eta}||$ and $g_{n,\eta} < f_{\infty,\eta}$ for n large enough. Thus, taking into account that $\mu_n(\mathrm{hypo}(f_{\infty,\eta})) = 0$ we get $\mu_n(\mathrm{hypo}(g_{n,\eta})) = 0$ for sufficiently large n. To sum up, for large n we have

$$\Theta_\infty(\mu_\infty) \leq T_\infty(f_{\infty,\eta}) + k\eta = T_\infty(g_{n,\eta}) + 5k\eta \leq$$

$$\leq \Theta_n(\mu_n) + 5k\eta + (T_\infty(g_{n,\eta}) - T_n(g_{n,\eta})).$$

Since η was arbitrary and the sequence $||g_{n,\eta}||$ is bounded by $||\phi_\infty|| + 5\eta$, in view of Lemma 3 we obtain finally

$$\Theta_\infty(\mu_\infty) \leq \lim\inf_{n\to\infty} \Theta_n(\mu_n).$$

The proof of the lemma is complete. □

The proof of Theorem 1 is complete too. □

3 The convex case

In this section we assume that all the sets belonging to $\mathcal{A}$ are convex. It is easily seen that under this assumption when constructing the ML-estimator $\hat{A}_n$ we can restrict our attention to the vertices of the convex hull of the sample, ignoring the interior points. Considering an estimator based only on the vertices of the convex hull we do not lose any information, since the convex hull possesses the property of sufficiency in the sense that the conditional distribution of the remaining points of the sample given the convex hull is uniform in the interior of the hull.

The assumption of convexity of sets belonging to $\mathcal{A}$ allows us to investigate the asymptotic behaviour of the surface measure $s(\hat{A}_n)$ of the estimator $\hat{A}_n$.

Theorem 2 *Assume that the conditions of Theorem 1 are satisfied and that in addition all the sets belonging to $\mathcal{A}$ are convex. Then*

$$|s(A_0) - s(\hat{A}_n)| = O_{\mathbb{P}}(n^{-1}).$$

Proof Let $(\hat{\Omega}, \hat{\Im}, \hat{P})$ be as in Theorem 1, i.e. in particular for $\hat{P}$-almost all $\omega \in \hat{\Omega}$ the sequence $nF_{\hat{A}_n}(\omega)$ is relatively compact and hence bounded. Let $\epsilon = \epsilon(\omega) < \infty$ be such that $||nF_{\hat{A}_n}(\omega)|| < \epsilon$ for all $n \in \mathbb{N}$. This means in particular that for $n > \epsilon$

$$(1-\frac{\epsilon}{n})A_0 \subset \hat{A}_n \subset (1+\frac{\epsilon}{n})A_0.$$

Since for compact convex sets $C_1 \subset C_2$ implies $s(C_1) \leq s(C_2)$, we get

$$(1-\frac{\epsilon}{n})^{k-1}s(A_0) = s((1-\frac{\epsilon}{n})A_0) \leq s(\hat{A}_n(\omega)) \leq$$

$$\leq s((1+\frac{\epsilon}{n})A_0) = (1+\frac{\epsilon}{n})^{k-1}s(A_0).$$

This completes the proof. $\Box$

4 The ellipsoids

In this section we require that $\mathcal{A}$ be a certain class of ellipsoids. Under this assumption we obtain an explicit representation for the functional class $\mathcal{F}_\infty$. This is done in Theorem 3. In particular it turns out that the assumptions of Corollary 1 are satisfied in this case, so our functional limit theorem also holds here, which is stated in Corollary 5.

Throughout this section we assume A_0 to be a unit ball in $\mathbb{R}^k$ centered in 0. Consider a Lie group $\mathcal{G} \subset SL(k)$ where $SL(k)$ denotes the group of all the real $k \times k$ matrices with determinant 1. Group $\mathcal{G}$ acts in the natural way on the space of all Borel subsets of $\mathbb{R}^k$. Let $\mathcal{U} \subset \mathcal{G}$ be an open bounded neighbourhood of $\mathbf{1}$. Clearly, $\mathcal{U}$ can be chosen so that $\mathcal{U}^{-1} = \mathcal{U}$. It is natural to consider the class $\mathcal{A}$ of ellipsoids generated by actions of the elements of $\mathcal{U}$ on A_0

$$\mathcal{A} = \{\beta g A_0 \mid g \in \mathrm{cl}(\mathcal{U}),\ \beta \in R_+\}$$

with cl denoting the topological closure. It is easily seen that $\mathcal{A}$ satisfies conditions (**R0**), (**R1**) and (**R2**). The only regularity condition imposed on class $\mathcal{A}$ in Section 1 which requires verification here is condition (**REG**). It turns out that if $\mathcal{U}$ is taken sufficiently small, condition **(REG)** is also satisfied and the class $\mathcal{F}_\infty$ corresponding to $\mathcal{A}$ can be described explicitly in terms of the Lie algebra $\mathfrak{G}$ of the group $\mathcal{G}$. Recall that $\mathfrak{G}$ is defined to be the space tangent to the smooth manifold $\mathcal{G}$ at the identity $\mathbf{1}$.

Theorem 3 *Let $\mathcal{A}$ be defined as above. If $\mathcal{U}$ is chosen sufficiently small then assumptions of Theorem 1 (i.e. conditions* **(R0)**, **(R1)**, **(R2)** *and* **(REG)***) for class $\mathcal{A}$ are satisfied and*

$$\mathcal{F}_\infty = \{f_{\mathfrak{g},\alpha} \mid \alpha \in \mathbb{R},\ \mathfrak{g} \in \mathfrak{G}\},$$

where $f_{\mathfrak{g},\alpha} : \mathcal{S}_{k-1} \to R$ is defined by

$$f_{\mathfrak{g},\alpha}(x) = \langle x, \mathfrak{g}x \rangle + \alpha,$$

with $\langle .,. \rangle$ denoting the scalar product.

It is easily seen that from this theorem it follows that class $\mathcal{A}$ satisfies also the assumptions of Corollary 1. Thus, we obtain

Corollary 5 *Let $\mathcal{A}$ be defined as above and assume that $\mathcal{U}$ is sufficiently small. Then the sequence $nF_{\hat{A}_n}$ converges in distribution to a $\mathcal{C}(\mathcal{S}_{k-1};\mathbb{R})$-valued random element Ψ defined by*

$$\Psi = argmax_{f_{\mathfrak{g},\alpha}}\{\int_{\mathcal{S}_{k-1}} f_{\mathfrak{g},\alpha}(x)U(dx) \mid \alpha \in \mathbb{R},\ \mathfrak{g} \in \mathfrak{G};\ \Pi_{U^+}(hypo(f)) = 0\},$$

where U is the uniform distribution on $\mathcal{S}_{k-1}$, U^+ is defined as the product $U^+ = U \times \lambda((.) \cap \mathbb{R}_+)$ and Π_{U^+} is the Poisson point process with intensity measure U^+.

Proof of Theorem 3 First, we have to determine the class $\mathcal{F}_1$ defined as in previous section by

$$\mathcal{F}_1 = \{F_A \mid A \in \mathcal{A}\}.$$

We will use the following lemma

Lemma 8 *For $g \in \mathcal{G}$ and $x \in \mathcal{S}_{k-1} = \partial A_0$*

$$F_{(gA_0)}(x) = \frac{-r(g^{-1}x)}{1 - r(g^{-1}x)},$$

where

$$r(x) = 1 - |x|$$

for $x \in \mathbb{R}^k \setminus \{0\}$.

Proof For any $x \in \mathcal{S}_{k-1}$

$$F_{(gA_0)}(x) = 1 - |g(\frac{g^{-1}x}{|g^{-1}x|})| = 1 - \frac{1}{|g^{-1}x|} = 1 - \frac{1}{1 - r(g^{-1}x)} =$$

$$= \frac{-r(g^{-1}x)}{1 - r(g^{-1}x)}.$$

□

As a conclusion we get

Corollary 6 *The class $\mathcal{F}_1$ admits the following representation*

$$\mathcal{F}_1 = \{\frac{-(1+\epsilon)r(g^{-1}x)}{1 - r(g^{-1}x)} - \epsilon \mid \epsilon \in (-1, \infty)\}.$$

More generally

$$\mathcal{F}_n = \{\frac{-(1+\epsilon)(nr(g^{-1}x))}{1 - r(g^{-1}x)} - n\epsilon \mid \epsilon \in (-1, \infty)\}.$$

Proof Recall that for any $A \in \mathcal{A}$ and $\epsilon \in (-1, \infty)$

$$F_{(1+\epsilon)A} = (1+\epsilon)F_A - \epsilon$$

and use Lemma 8. $\square$

The next step of the proof of Theorem 3 is to show that $\mathcal{F}_1$ satisfies condition **(REG)**. We will obtain it as an immediate conclusion of the following lemma

Lemma 9 *There exists a constant K such that for any $f \in \mathcal{F}_1$*

$$||f'|| \leq K||f||$$

In order to prove this lemma we need

Lemma 10 *For each $g \in GL(k)$, where $GL(k)$ is the k-dimensional linear group, define $\hat{g} := (gg^*)^{\frac{1}{2}}$. Then operator $\hat{g}$ is self-adjoint and there exists an isometry ϕ_g such that $g = \hat{g} \circ \phi_g$. In particular, $gA_0 = \hat{g}A_0$. Further, if $g \in SL(k)$ then also $\hat{g} \in SL(k)$.*

Proof Take $\phi_g = ((gg^*)^{-\frac{1}{2}}g)$. Obviously $\phi_g^* = g^*(gg^*)^{-\frac{1}{2}}$, so $\phi_g^*\phi_g = \mathbf{1}$, which completes the proof. $\square$

Proof of Lemma 9 Actually we are going to prove a slightly stronger thesis. Namely we show that there exists a constant K such that

$$||f'|| \leq K(\sup_{x \in \mathcal{S}_{k-1}} f(x) - \inf_{x \in \mathcal{S}_{k-1}} f(x))$$

(note that $\sup_{x \in \mathcal{S}_{k-1}} f(x) - \inf_{x \in \mathcal{S}_{k-1}} f(x) \leq 2||f||$). It is not difficult to observe that whenever this inequality holds for some function f, it holds also for $\beta f + \alpha$ for any $\alpha, \beta \in \mathbb{R}$. Therefore in view of the Corollary 6 it suffices to prove our thesis for f of the form

$$f(x) = F_{gA_0}(x) = \frac{-r(g^{-1}x)}{1 - r(g^{-1}x)}$$

with $g \in \mathcal{U}$. According to Lemma 10 we may write

$$f(x) = F_{\hat{g}A_0}(x) = \frac{-r(\hat{g}^{-1}x)}{1 - r(\hat{g}^{-1}x)},$$

where $\hat{g}$ is self-adjoint. Taking the derivative we get

$$f'(x)[h] = -\frac{(r \circ \hat{g}^{-1})'(x)[h]}{1 - r(\hat{g}^{-1}x)} - \frac{r(\hat{g}^{-1}x)(r \circ \hat{g}^{-1})'(x)[h]}{(1 - r(\hat{g}^{-1}x))^2}.$$

Taking $\mathcal{U}$ small enough we may make $r(\hat{g}^{-1}x)$ arbitrarily close to 0 uniformly on x. Therefore the proof will be terminated if we prove the existence of a constant K_1 independent on $\hat{g}$ and such that for all $x \in \mathcal{S}_{k-1}$

$$|(r \circ \hat{g}^{-1})'(x)[h]| \leq K_1 (\sup_{x \in \partial A_0} r(\hat{g}^{-1}x) - \inf_{x \in \partial A_0} r(\hat{g}^{-1}x))$$

for all $h \in T_x$ such that $|h| \leq 1$, where T_x is the space tangent to $\mathcal{S}_{k-1}$ in x. Observe that

$$|(r \circ \hat{g}^{-1})'(x)[h]| = | - \frac{\langle \hat{g}^{-1}x, \hat{g}^{-1}h \rangle}{|x|} | = |\langle \hat{g}^{-1}x, \hat{g}^{-1}h \rangle| =$$

$$= |\langle \hat{g}^{-2}x, h \rangle| = |\langle \hat{g}^{-2}x - x, h \rangle| = |\langle (\hat{g}^{-2} - \mathbf{1})x, h \rangle| \leq ||\hat{g}^{-2} - \mathbf{1}|| \leq$$

$$\leq ||\hat{g}^{-2} - \hat{g}^{-1}|| + ||\hat{g}^{-1} - \mathbf{1}|| \leq ||\hat{g}^{-1} - \mathbf{1}|| ||\hat{g}^{-1}|| + ||\hat{g}^{-1} - \mathbf{1}|| =$$

$$= (||\hat{g}^{-1}|| + 1)||\hat{g}^{-1} - \mathbf{1}|| = (||g^{-1}|| + 1)||\hat{g}^{-1} - \mathbf{1}||.$$

Set

$$K_1 := \sup_{g \in \mathcal{U}} ||g^{-1}|| + 1.$$

Obviously $K_1 < \infty$ if $\mathcal{U}$ is sufficiently small. Since $\hat{g}^{-1} - \mathbf{1}$ is self-adjoint, there exists $x_0 \in \mathcal{S}_{k-1}$ such that $||\hat{g}^{-1} - \mathbf{1}|| = |(\hat{g}^{-1} - \mathbf{1})(x_0)| = |\langle (\hat{g}^{-1} - \mathbf{1})(x_0), x_0 \rangle| = |r(\hat{g}^{-1}x_0)|$. Therefore we get

$$|(r \circ \hat{g}^{-1})'(x)[h]| \leq K_1 \sup_{x \in \partial A_0} |r(\hat{g}^{-1}x)|.$$

To complete the proof it suffices to show that

$$\sup_{x \in \mathcal{S}_{k-1}} |r(\hat{g}^{-1})| \leq \sup_{x \in \mathcal{S}_{k-1}} r(\hat{g}^{-1}x) - \inf_{x \in \mathcal{S}_{k-1}} r(\hat{g}^{-1}x).$$

This is however an immediate conclusion from the following lemma

Lemma 11 *For each $g \in \mathcal{G}$ there exists $\overline{x}$ such that*

$$r(g^{-1}\overline{x}) = 0.$$

Proof of Lemma 11 Since $\mathcal{G} \subset SL(k)$, g preserves the Lebesgue measure. In particular $\lambda(gA_0) = \lambda(A_0)$, so ∂gA_0 cannot lie neither totally inside nor totally outside $\partial A_0 = \mathcal{S}_{k-1}$. It suffices now to take $\overline{x} \in \partial A_0 \cap \partial gA_0$. This proves the lemma. □

The proof of Lemma 9 is now also complete. □

To proceed, we expand $r(x) = 1 - |x|$ in a Taylor series obtaining

Lemma 12 *Let $x \in \mathcal{S}_{k-1}$. Then for any $y \in \mathbb{R}^k \setminus \{0\}$ there exists $\theta \in [0,1]$ such that*

$$r(y) = -\langle x, y - x \rangle - \frac{1}{2} \frac{|y - x|^2}{|x + \theta(y - x)|} + \frac{1}{2} \frac{\langle x + \theta(y - x), y - x \rangle^2}{|x + \theta(y - x)|^3}.$$

Proof Note that

$$r'(x)[h] = -\frac{\langle x, h\rangle}{|x|}$$

and

$$r''(x)[h_1][h_2] = -\frac{\langle h_1, h_2\rangle}{|x|} + \frac{\langle h_1, x\rangle\langle h_2, x\rangle}{|x|^3}.$$

To complete the proof it suffices to observe that $r(x) = 0$ and $|x| = 1$. □

The above expansion together with Corollary 6 allows us to prove

Lemma 13 *Let* $\mathfrak{g} \in \mathfrak{G}$, $\alpha \in R$ *and let* $f_\infty : \partial A_0 \to R$ *be defined by*

$$f(x) = \alpha + \langle x, \mathfrak{g}\mathfrak{x}\rangle.$$

Then $f_\infty \in \mathcal{F}_\infty$.

Proof Choose a sequence $(g_n)_{n=1}^{\infty} \subset \mathcal{U}$ such that

$$\lim_{n\to\infty} ||n(g_n^{-1} - 1) + \mathfrak{g}|| = 0.$$

Take also $\epsilon_n = -\frac{\alpha}{n}$. According to the Corollary 6 for any $n \in \mathbb{N}$ the function

$$f_n(x) = \frac{-(1+\epsilon_n)(nr(g_n^{-1}x))}{1 - r(g_n^{-1}x)} - n\epsilon_n$$

belongs to $\mathcal{F}_n$. It suffices now to prove that

$$\sup_{x\in\partial A_0} |nr(g_n^{-1}x) + \langle x, \mathfrak{g}x\rangle| = 0.$$

But from Lemma 12 it follows that

$$nr(g_n^{-1}x) = -\langle x, n(g_n^{-1} - \mathbf{1})(x)\rangle + O(n||g_n^{-1} - \mathbf{1}||^2).$$

The lemma is proven. □

Now let us assume that n' is a strictly inreasing sequence of natural numbers and that a sequence $f_{n'} \in \mathcal{F}_{n'}$ converges uniformly to some f_∞. We will show that there exists $\mathfrak{g} \in \mathfrak{G}$ and $\alpha \in R$ such that $f_\infty(x) = \alpha + \langle x, \mathfrak{g}x\rangle$. In view of Lemmas 9 and 13 this will complete the proof of Theorem 3.

By the definition of $\mathcal{A}$ and $\mathcal{F}_{n'}$ there exist sequences $g_{n'} \in \mathcal{U}$ and $\epsilon_{n'} \in (-1, \infty)$ such that

$$f_{n'} = F_{(1+\epsilon_{n'})g_{n'}A_0}.$$

From Lemma 10 it follows that there exists a sequence $\hat{g}_{n'}$ of self-adjoint operators belonging to $SL(k)$ such that

$$g_{n'}A_0 = \hat{g}_{n'}A_0.$$

Thus, according to Lemma 8 and Corollary 6 we obtain

$$f_{n'}(x) = \frac{-(1+\epsilon_{n'})(n' r(\hat{g}_{n'}^{-1}x))}{1-r(\hat{g}_{n'}^{-1}x)} - n'\epsilon_{n'}.$$

According to Lemma 11 there exists a sequence $x_{n'} \in \partial A_0$ such that $r(\hat{g}_{n'}^{-1}x_{n'}) = 0$. Since ∂A_0 is compact we may assume without loss of generality that $x_{n'}$ converges to some $x_\infty \in \partial A_0$. Therefore there exists the limit

$$\alpha_\infty = -\lim_{n'\to\infty} n'\epsilon_{n'} = f_\infty(x_\infty).$$

It follows also that $\lim_{n'\to\infty}\epsilon_{n'} = 0$.

Henceforth we assume without loss of generality that $\epsilon_{n'} = 0$ and

$$f_{n'}(x) = \frac{-n' r(\hat{g}_{n'}^{-1}x)}{1-r(\hat{g}_{n'}^{-1}x)}.$$

We will prove that if $\mathcal{U}$ is sufficiently small

Lemma 14

$$\lim_{n'\to\infty} ||n' r\circ \hat{g}_{n'}^{-1} + f_\infty|| = 0.$$

Proof If $\mathcal{U}$ is chosen small enough, there exists a constant $\eta > 0$ such that

$$\forall_{n'}\forall_{x\in\mathcal{S}_{k-1}}\ \ 0 < 1 - r(\hat{g}_{n'}^{-1}x) < \eta.$$

Moreover there exists $M > 0$ such that

$$\forall_{n'}\forall_{x\in\mathcal{S}_{k-1}}\ \ |f_{n'}(x)| < M,$$

because $f_{n'}$ converges uniformly. Hence we obtain

$$\forall_{n'}\forall_{x\in\partial A_0}\ \ |r(\hat{g}_{n'}^{-1}x)| < \frac{\eta M}{n'}.$$

In particular it means that $|r(\hat{g}_{n'}^{-1}x)|$ converges to 0 uniformly on x. This proves our assertion because $f_{n'}$ converges uniformly to f_∞ and $f_{n'} = -\frac{n' r(\hat{g}_{n'}^{-1}x)}{1-r(\hat{g}_{n'}^{-1}x)}$. □

From Lemma 12 we conclude that for each $x \in \mathcal{S}_{k-1}$

$$n' r(\hat{g}_{n'}^{-1}x) = -n'\langle \hat{g}_{n'}^{-1}x - x, x\rangle - \frac{n'}{2}\frac{|\hat{g}_{n'}^{-1}x - x|^2}{|x+\theta_x(\hat{g}_{n'}^{-1}x - x)|} +$$

$$+\frac{n'}{2}\frac{\langle \hat{g}_{n'}^{-1}x - x, x+\theta_x(\hat{g}_{n'}^{-1}x - x)\rangle^2}{|x+\theta_x(\hat{g}_{n'}^{-1}x - x)|^3}$$

for some $\theta \in [0,1]$. If $\mathcal{U}$ is taken small enough two last terms in this expansion can be made arbitrarily small with respect to the first term, because the operators $\hat{g}_{n'}^{-1} - \mathbf{1}$ are self-adjoint and hence $||\hat{g}_{n'}^{-1} - \mathbf{1}|| = \max_{x \in \partial A_0} \langle \hat{g}_{n'}^{-1} x - x, x \rangle$. Thus, since by Lemma 14 the sequence $n' r(\hat{g}_{n'}^{-1} x)$ converges uniformly on x, so does the first term of the expansion equal to $-n' \langle \hat{g}_{n'}^{-1} x - x, x \rangle$. In particular we conclude that

$$\limsup_{n' \to \infty} ||n'(\hat{g}_{n'}^{-1} - \mathbf{1})|| < \infty.$$

Hence, without loss of generality we can assume that the sequence $n'(\hat{g}_{n'}^{-1} - \mathbf{1})$ converges to some $k \times k$-matrix $\mathfrak{h}$ for $n' \to \infty$. Clearly, $\mathfrak{h}$ is symmetric because so are $\hat{g}_{n'}^{-1} - \mathbf{1}$. Now, applying once more Lemma 14 and the considered expansion of $n' r(\hat{g}_{n'}^{-1} x)$ we obtain the uniform convergence

$$\lim_{n' \to \infty} ||f_{n'}(x) - \langle \mathfrak{h} x, x \rangle|| = 0,$$

so

$$f_\infty(x) = \langle \mathfrak{h} x, x \rangle.$$

The proof of Theorem 3 will be complete if we succeed to show that there exists $\mathfrak{g} \in \mathfrak{G}$ such that $\mathfrak{h} = (\mathfrak{g})_{Sym}$, where $(\mathfrak{g})_{Sym}$ denotes the symmetric part of matrix $\mathfrak{g}$, i.e. $(\mathfrak{g})_{Sym} = \frac{1}{2}(\mathfrak{g} + \mathfrak{g}^*)$. Obviously then

$$f_\infty(x) = \langle \mathfrak{h} x, x \rangle = \langle \mathfrak{g} x, x \rangle.$$

To proceed we need an auxiliary lemma

Lemma 15 *Let Φ be a mapping defined on the space of all $k \times k$-matrices by*

$$\Phi(g) = \hat{g} = (gg^*)^{\frac{1}{2}}.$$

Then the rank of the derivative Φ' is constant in the neighbourhood of $\mathbf{1}$.

Proof We have $\Phi = \Psi \circ \Sigma$, where $\Psi(a) = a^{\frac{1}{2}}$ (for a symmetric and positive-definite) and $\Sigma(g) = gg^*$. Obviously, for $g \in GL(k)$ and for any $k \times k-$ matrix h,

$$\Phi'(g)[h] = \Psi'(\Sigma(g))[\Sigma'(g)[h]].$$

Note however that for g sufficiently close to $\mathbf{1}$ the operator $\Psi'(\Sigma(g))$ is a linear automorphism of the space of symmetric matrices. Further,

$$\Sigma'(g)[h] = hg^* + gh^* = 2(gh^*)_{Sym},$$

so

$$\dim(\mathrm{Im}\Sigma'(g)) = \frac{n(n+1)}{2}$$

and hence this dimension is constant for g close enough to $\mathbf{1}$. The proof is complete. $\square$

Applying Lemma 15 we prove that the space tangent to $\Phi(\mathcal{G})$ in $\mathbf{1}$ coincides with $\Phi'(\mathbf{1})(\mathfrak{G})$. In particular, there exists $\mathfrak{g} \in \mathfrak{G}$ such that $\mathfrak{h} = \Phi'(\mathbf{1})[\mathfrak{g}]$. But $\Phi'(\mathbf{1})[\mathfrak{g}] = (\mathfrak{g})_{Sym}$, what completes the proof of Theorem 3. $\square$

5 Concluding remarks

It is worth noting that the ML-principle which in classical statistics gives optimal rates of convergence works also very well when applied to the problems of support estimation and gives very fast convergence.

The important problem that arises here is the serious computational complexity of ML-estimators. However within our setting the additional assumption of the convexity of the sets belonging to $\mathcal{A}$ can facilitate the problem considerably. It is easily seen that in this situation when constructing the ML-estimator we can restrict our attention to the vertices of the convex hull of the sample, ignoring the interior points. These observations together with appropriate formulas for the expected number of vertices of the convex hull of a uniform sample (see e.g. Rényi & Sulanke [16] and Efron [5]) prove that for example in case where the class $\mathcal{A}$ is a certain class of ellipsoids (as considered in Section 4) the expected complexity of the algorithm constructing the estimator is polynomial.

Acknowledgments Special thanks are due to Prof. A. V. Nagaev for suggesting me the topic and for carefully revising this text. I would also like to express my gratitude to Prof. I. Molchanov for his valuable comments and suggestions.

References

[1] H. Braeker and T. Hsing, On the area and perimeter of a random convex hull in a bounded convex set. *Probab. Theory Related Fields,* **111** (1998), pp. 517-550.

[2] A. J. Cabo and P. Groeneboom, Limit theorems for functionals of convex hulls. *Probab. Theory Related Fields,* **100** (1994), pp. 31-55.

[3] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes.* Springer-Verlag, New York Inc., 1988.

[4] R. M. Dudley, Metric entropy of some classes of sets with differentiable boundaries, *J. Approx. Theory,* **10** (1974), pp. 227-236.

[5] B. Efron, The convex hull of a random set of points. *Biometrika,* **52** (1965), pp. 331-343.

[6] P. Groeneboom, Limit theorems for convex hulls. *Probability Theory and Related Fields,* **79** (1988), pp. 327-368.

[7] T. Hsing, On the asymptotic distribution of the area outside a random convex hull in a disk. *Annals Appl. Probab.,* Vol. **4** (1994), No. 2, pp. 478-493.

[8] A. F. Karr, *Point Processes and Their Statistical Inference,* Probability: pure and applied; Marcel Dekker, Inc., New York, 1991.

[9] J. Kerstan, K. Matthes and J. Mecke, *Infinitely Divisible Point Processes.* (Russian edition), Nauka, Moscow, 1982.

[10] I. M. Khamdamov and A. V. Nagaev, Limiting distributions for functionals of the convex hull generated by uniformly distributed variables. *Dokl. Akad. Nauk UzSSR,* **7** (1991), pp. 8-9 (in Russian).

[11] A. P. Korostelev and A. B. Tsybakov, *Minimax Theory of Image Reconstruction.* Lecture Notes in Statistics, **82** (1993), Springer-Verlag, New York.

[12] E. Mammen and A. B. Tsybakov, Asymptotical minimax recovery of sets with smooth boundaries. *Ann. Statist.,* **23** (1995), pp. 502-524.

[13] I. S. Molchanov, On the convergence of random processes generated by polyhedral approximation of convex compacts. *Theory Probab. Appl.,* **40** (1995), pp. 383-390.

[14] K. H. Kuefer, On the approximation of a ball by random polytopes. *Adv. Appl. Probab.,* **26** (1994), pp. 876-892.

[15] J. P. Rasson and B. D. Ripley, Finding the edge of a Poisson forest, *J. Appl. Probab.,* **14** (1977), pp. 483-491.

[16] A. Rényi and R. Sulanke, Ueber die konvexe Huelle von n zufaellig gewaehlten Punkten (I; II) *Z. Wahrscheinlichkeitsth. verw. Geb.* **2** (1963), pp. 75-84; **3** (1964), pp. 138-147.

[17] S. I. Resnick, *Extreme Values, Regular Variation and Point Processes* Springer-Verlag, New York, 1987.

[18] R. Schneider, Random approximation of convex sets. *J. Microscopy,* **151** (1988), pp. 211-227