GOPT 919499 8-5-2014 Initial CE:YS QA:EK

Optimization, 2014
http://dx.doi.org/10.1080/02331934.2014.919499

Taylor & Francis
Taylor & Francis Group

A Robbins–Monro-type algorithm for computing global minimizer of generalized conic functions

AQ1

Mátyás Barczy[a]*, Ábris Nagy[b], Csaba Noszály[a] and Csaba Vincze[c]

[a]*Faculty of Informatics, Department of Applied Mathematics and Probability, University of Debrecen, Debrecen, Hungary;* [b]*MTA-DE Research Group 'Equations Functions and Curves', Institute of Mathematics, Hungarian Academy of Sciences and University of Debrecen, Debrecen, Hungary;* [c]*Institute of Mathematics, University of Debrecen, Debrecen, Hungary*

AQ2

(*Received 20 February 2013; accepted 17 April 2014*)

AQ3

We generalize the notion and some properties of the conic function introduced by Vincze and Nagy (2012). We provide a stochastic algorithm for computing the global minimizer of generalized conic functions, we prove almost sure and L^q-convergence of this algorithm.

Keywords: global optimization; Markov process; conic function; stochastic algorithm; Robbins–Monro algorithm

AMS Subject Classifications: 90C25; 60D05

1. Introduction

Let K be a compact body in $\mathbb{R}^2$ (a non-empty compact set coinciding with the closure of its interior) and consider the distance function induced by the taxicab norm. The so-called conic function F_K associated to K (introduced by Vincze and Nagy [1, Definition 6], see also Definition 2.1) measures the average taxicab distance of the points from K via integration with respect to the Lebesgue measure, or explaining in another way: the conic function F_K at some point $(x, y) \in \mathbb{R}^2$ can be interpreted as the expectation of the random variable defined as the taxicab distance of (x, y) and (ξ, η), where (ξ, η) is a uniformly distributed random variable on K, for more details see part (ii) of Remark 1. Conic functions are extensively used in geometric tomography since they contain a lot of information about unknown bodies, for a more detailed discussion see Gardner [2] and Vincze and Nagy [1]. We call the attention that in the literature one can find other definitions of 'conic functions' that are completely different from ours. For example, in optimization, a conic function is usually defined to be the ratio of a quadratic function and the square of a linear function on the open halfspace, where the linear function is positive, see, e.g. Luksan [3, formula (2.1)]. Wang et al. [4] introduced another definition of conic functions in metric spaces and obtained a new condition for metric spaces being compact in terms of conic functions.

AQ4

We recall that one of the striking features of the conic function F_K is that a point in $\mathbb{R}^2$ is a global minimizer of F_K if and only if it bisects the area of K, i.e. the vertical

*Corresponding author. Email: barczy.matyas@inf.unideb.hu

and horizontal lines through this point cut the compact body K into two parts with equal areas, see Vincze and Nagy [1, Corollary 1]. We call the attention that points with similar properties are important and well studied in geometry. For instance, we mention that if S is a convex set in $\mathbb{R}^2$, then there exist two perpendicular lines that divide S into four parts with equal areas, see Yaglom and Boltyanskii [5, Section 3].

In Section 2 of the present paper, we generalize the conic function F_K introduced by Vincze and Nagy [1] in a way that it measures the average taxicab distance of the points from K via integration with respect to some measure μ on K with $\mu(K) < \infty$, see Definition 2.5. From geometric point of view, the body K associated with some measure μ can be considered as a mathematical model of a non-homogeneous body and hence our generalization of conic functions may find applications in (geometric) tomography where typically non-homogeneous bodies occur. We generalize Theorems 3, 4, 5, Lemmas 6, 7 and Corollary 1 in Vincze and Nagy [1] for conic functions $F_{K,\mu}$ associated with a compact body K and a measure μ with $\mu(K) < \infty$. We only mention that it turns out that a point in $\mathbb{R}^2$ is a global minimizer of $F_{K,\mu}$ if and only if it bisects the μ-area of K, see Corollary 2.9.

In Section 3, we give a stochastic algorithm for the global minimizer of the convex function $F_{K,\mu}$. In the heart of our algorithm, the well-known Robbins–Monro algorithm (see [6]) lies, and we prove almost sure and L^q-convergence of our algorithm. More precisely, we define recursively a sequence $(X_k)_{k\in\mathbb{Z}_+}$ of random variables (see (3.1)) which forms an inhomogeneous Markov chain and we prove almost sure and L^q-convergence of this Markov chain via Robbins–Monro algorithm, see Theorem 3.3. We also prove almost sure and L^q-convergence of the sequence $(F_{K,\mu}(X_k))_{k\in\mathbb{N}}$, see Theorem 3.6. In general, stochastic algorithms for finding a minimum of a convex function have a vast literature, see, e.g. Robert and Casella [7] and Bouleau and Lépingle [8]. Without giving an introduction of the newest results in the field we only mention the paper [9] of Arnaudon et al., which in some sense motivated our study. They gave a stochastic algorithm which converges almost surely and in L^2 to the so-called p-mean of a probability measure supported by a regular geodesic ball in a manifold.

2. Generalized conic functions

Let $\mathbb{Z}_+$, $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the set of non-negative integers, positive integers, real numbers and non-negative real numbers, respectively. For an $x \in \mathbb{R}^2$, we will denote its Euclidean norm by $\|x\|$. Let $K \subset \mathbb{R}^2$ be a non-empty compact set such that it coincides with the closure of its interior. In geometry, K is called a compact body. By $\mathcal{B}(\mathbb{R}^d)$ and $\mathcal{B}(K)$, we denote the Borel σ-algebra on $\mathbb{R}^d$ and on K, respectively, where $d \in \mathbb{N}$. For all $x, y \in \mathbb{R}$ let us introduce the following notations

$$\begin{aligned}
\{K <_1 x\} &:= \{(\alpha, \beta) \in K : \alpha < x\}, & \{x <_1 K\} &:= \{(\alpha, \beta) \in K : x < \alpha\},\\
\{K <_2 y\} &:= \{(\alpha, \beta) \in K : \beta < y\}, & \{y <_2 K\} &:= \{(\alpha, \beta) \in K : y < \beta\},\\
\{K =_1 x\} &:= \{(\alpha, \beta) \in K : \alpha = x\}, & \{K =_2 y\} &:= \{(\alpha, \beta) \in K : \beta = y\}.
\end{aligned}$$

The notations $\{K \le_1 x\}$, $\{x \le_1 K\}$, $\{K \le_2 y\}$ and $\{y \le_2 K\}$ are defined in the same way. For a function $f : \mathbb{R}^2 \to \mathbb{R}$, we will denote by $D_1 f$ and $D_2 f$ the partial derivatives of f.

Next, we recall the notion of a generalized conic function associated with K due to Vincze and Nagy [1].

Definition 2.1 (Vincze and Nagy [1, Definition 6]) The generalized conic function $F_K : \mathbb{R}^2 \to \mathbb{R}$ associated to K is defined by

$$F_K(x, y) := \frac{1}{A(K)} \int_K d_1((x, y), (\alpha, \beta))\, \mathrm{d}\alpha \mathrm{d}\beta, \qquad (x, y) \in \mathbb{R}^2,$$

where $A(K)$ is the two-dimensional Lebesgue measure (area) of K, and the distance function d_1 is given by $d_1((x, y), (\alpha, \beta)) := |x - \alpha| + |y - \beta|$, $(x, y), (\alpha, \beta) \in \mathbb{R}^2$ (d_1 is known to be the metric induced by the taxicab norm).

The next result is about the global minimizer of F_K.

Proposition 2.2 (Vincze and Nagy [1, Corollary 1]) *A point in $\mathbb{R}^2$ is a global minimizer of the generalized conic function F_K if and only if it bisects the area of K, i.e. the vertical and the horizontal lines through this point cut the compact body K into two parts with equal area.*

We note that the global minimizer of the generalized conic function F_K is not unique in general. In Proposition 2.3, we give a sufficient condition for its uniqueness.

In what follows we will frequently use the following conditions

(C.1) K is connected,
(C.2) $\mu(B(p, \varepsilon) \cap K) > 0$ for all $p \in K$, $\varepsilon > 0$ and $B(p, \varepsilon)$,
where μ is a measure on the measurable space $(K, \mathcal{B}(K))$ and $B(p, \varepsilon)$ denotes the open ball around p with radius ε, and
(C.3) $\mu(\{K =_1 x\}) = \mu(\{K =_2 y\}) = 0$ for all $x, y \in \mathbb{R}$.
We call the attention that Condition (C.3) does not hold for a measure in general. For example, if μ is the distribution of a discrete random variable having values in K, then Condition (C.3) does not hold. However, if μ is the two-dimensional Lebesgue measure on K, then Conditions (C.2) and (C.3) hold automatically.

Proposition 2.3 *If Condition* (C.1) *holds, then the convex function F_K has a unique global minimizer $(x^*, y^*) \in \mathbb{R}^2$, that is, $F_K(x, y) > F_K(x^*, y^*)$ for $(x, y) \neq (x^*, y^*)$, $(x, y) \in \mathbb{R}^2$.*

Proof The existence of a global minimizer of F_K can be checked as follows. By Theorem 3 in Vincze and Nagy [1], F_K is a finite-valued convex function defined on $\mathbb{R}^2$ and its level sets are compact subsets of $\mathbb{R}^2$. Hence, F_K is continuous and consequently it reaches its minimum on every compact set.

Now we turn to prove the uniqueness of (x^*, y^*). Let us suppose that $(x^*, y^*) \in \mathbb{R}^2$ and $(\widetilde{x^*}, \widetilde{y^*}) \in \mathbb{R}^2$ are global minimizers of F_K such that $(x^*, y^*) \neq (\widetilde{x^*}, \widetilde{y^*})$. Then $x^* \neq \widetilde{x^*}$ or $y^* \neq \widetilde{y^*}$. We may assume that $\widetilde{x^*} < x^*$. Then both of the vertical lines $\mathbb{R}^2 =_1 x^*$ and $\mathbb{R}^2 =_1 \widetilde{x^*}$ bisect the area of K. Note that since Condition (C.3) holds automatically for the two-dimensional Lebesgue measure, the bisection of the area of K is well defined. Let us consider the open half-planes

$$H^* := \mathbb{R}^2 <_1 x^* \qquad \text{and} \qquad \widetilde{H^*} := \mathbb{R}^2 >_1 \widetilde{x^*}.$$

Note that $(\widetilde{x^*}, \widetilde{y^*}) \in H^*$ and $(x^*, y^*) \in \widetilde{H^*}$. We show that $K \cap (H^* \cap \widetilde{H^*}) = \emptyset$. On the contrary, let us suppose that there exists $p \in \mathbb{R}^2$ such that $p \in K \cap (H^* \cap \widetilde{H^*})$. Since K is a non-empty compact body, there exist

$$0 < \varepsilon < \min\{d_2(p, \mathbb{R}^2 =_1 x^*), d_2(p, \mathbb{R}^2 =_1 \widetilde{x^*})\}$$

and $q \in B(p, \varepsilon)$ such that q is an interior point of K, where d_2 denotes the standard Euclidean distance on $\mathbb{R}^2$. Hence, there exists

$$0 < \delta < \min\{d_2(p, \mathbb{R}^2 =_1 x^*), d_2(p, \mathbb{R}^2 =_1 \widetilde{x^*})\}$$

such that $B(q, \delta) \subset K \cap (H^* \cap \widetilde{H^*})$. Then

$$\begin{aligned} A(K <_1 \widetilde{x^*}) = A(\widetilde{x^*} <_1 K) \geq A(B(q, \delta)) + A(x^* <_1 K), \\ A(x^* <_1 K) = A(K <_1 x^*) \geq A(B(q, \delta)) + A(K <_1 \widetilde{x^*}), \end{aligned} \tag{2.1}$$

and hence

$$A(K <_1 x^*) \geq 2A(B(q, \delta)) + A(K <_1 x^*),$$

i.e. $0 \geq A(B(q, \delta))$, which yields us to a contradiction. At this point, we implicitly used that Condition (C.2) holds automatically for the two-dimensional Lebesgue measure. Hence $K \cap (H^* \cap \widetilde{H^*}) = \emptyset$. Let $0 < \eta < (x^* - \widetilde{x^*})/2$, and let us consider the open half-planes

$$I^* := \mathbb{R}^2 >_1 x^* - \eta \quad \text{and} \quad \widetilde{I^*} := \mathbb{R}^2 <_1 \widetilde{x^*} + \eta.$$

Then I^* and $\widetilde{I^*}$ are open sets of $\mathbb{R}^2$, $I^* \cap \widetilde{I^*} = \emptyset$, and, since $K \cap (H^* \cap \widetilde{H^*}) = \emptyset$, we have $K \subset I^* \cup \widetilde{I^*}$. Further, $I^* \cap K$ and $\widetilde{I^*} \cap K$ are separated sets such that their union equals K. This is a contradiction due to the connectedness of K. Hence $x^* = \widetilde{x^*}$, and in a similar way we have $y^* = \widetilde{y^*}$. □

We call the attention that Condition (C.1) is sufficient but not necessary in order that the generalized conic function F_K should have a uniquely determined global minimizer. Figure 1 shows three different cases where Condition (C.1) is not satisfied but F_K has a unique global minimizer.

On the subfigure (c) of Figure 1, the circles have centres $(-1/\sqrt{12}, 0)$ and $(1/2^n, 0)$ with radii $1/\sqrt{12}$ and $1/2^{n+2}$, respectively, where $n \in \mathbb{Z}_+$.

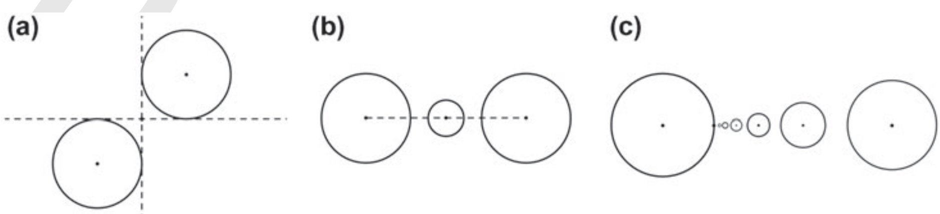

Figure 1. Examples for K such that Condition (C.1) does not hold but F_K has a unique global minimizer.

Example 2.4

(i) If K is the square with vertexes $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, then

$$F_K(x, y) = \left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2 + \frac{1}{2}, \qquad (x, y) \in K,$$

see, e.g. Vincze and Nagy [1, Example 3]. Using that K is connected, by Propositions 2.2 and 2.3, the global minimizer of F_K is $(x, y) = (\frac{1}{2}, \frac{1}{2})$.

(ii) If K is the triangle with vertexes $(0, 0)$, $(0, 1)$, $(1, 0)$, then

$$F_K(x, y) = -\frac{2}{3}(x^3 + y^3) + 2(x^2 + y^2) - (x + y) + \frac{2}{3}, \qquad (x, y) \in K.$$

Indeed, $F_K(x, y) = \mathbb{E}(|\xi - x|) + \mathbb{E}(|\eta - y|)$ for all $(x, y) \in \mathbb{R}^2$, where (ξ, η) is a uniformly distributed random variable on K. Then the joint density function of (ξ, η), and the density functions of the marginals of (ξ, η) take the forms

$$f_{(\xi,\eta)}(\alpha, \beta) = \begin{cases} 2 & \text{if } (\alpha, \beta) \in K, \\ 0 & \text{if} (\alpha, \beta) \notin K, \end{cases}$$

and

$$f_\xi(\alpha) = \begin{cases} -2\alpha + 2 & \text{if } \alpha \in [0, 1], \\ 0 & \text{if } \alpha \notin [0, 1], \end{cases} \qquad f_\eta(\beta) = \begin{cases} -2\beta + 2 & \text{if } \beta \in [0, 1], \\ 0 & \text{if } \beta \notin [0, 1], \end{cases}$$

respectively. Hence for all $(x, y) \in K$,

$$\begin{aligned} \mathbb{E}(|\xi - x|) &= \int_0^1 |\alpha - x|(-2\alpha + 2)\,\mathrm{d}\alpha \\ &= \int_0^x (x - \alpha)(-2\alpha + 2)\,\mathrm{d}\alpha + \int_x^1 (\alpha - x)(-2\alpha + 2)\,\mathrm{d}\alpha \\ &= -\frac{2}{3}x^3 + 2x^2 - x + \frac{1}{3}, \end{aligned}$$

and similarly $\mathbb{E}(|\eta - y|) = -\frac{2}{3}y^3 + 2y^2 - y + \frac{1}{3}$ for all $(x, y) \in K$. Hence, the global minimizer of F_K is $(1 - \sqrt{2}/2, 1 - \sqrt{2}/2)$. Indeed, the solution in K of the system of equations

$$D_1F_K(x, y) = -2x^2 + 4x - 1 = 0 \quad \text{and} \quad D_2F_K(x, y) = -2y^2 + 4y - 1 = 0,$$

is $(1 - \sqrt{2}/2, 1 - \sqrt{2}/2)$. Using that K is connected, by Propositions 2.2 and 2.3, the global minimizer of F_K is $(1 - \sqrt{2}/2, 1 - \sqrt{2}/2)$.

In what follows, we generalize the notion of the conic function introduced by Vincze and Nagy [1, Definition 6], see also Definition 2.1.

Definition 2.5 Let μ be a measure on the measurable space $(K, \mathcal{B}(K))$ such that $\mu(K) < \infty$. The generalized conic function $F_{K,\mu} : \mathbb{R}^2 \to \mathbb{R}$ associated to K and μ is defined by

$$F_{K,\mu}(x, y) := \int_K d_1((x, y), (\alpha, \beta))\,\mu(\mathrm{d}\alpha, \mathrm{d}\beta), \qquad (x, y) \in \mathbb{R}^2.$$

Remark 1

(i) Note that under the conditions of Definition 2.5, we have $F_{K,\mu}(x, y)$ is well defined for all $(x, y) \in \mathbb{R}^2$, since for fixed $(x, y) \in \mathbb{R}^2$, the function $K \ni (\alpha, \beta) \mapsto d_1((x, y), (\alpha, \beta))$ is bounded and $\mu(K) < \infty$.

(ii) If μ is a measure on K such that $\mu(K) < \infty$ and it is absolutely continuous with respect to the Lebesgue measure on K with Radon-Nikodym derivative h_μ, then

$$F_{K,\mu}(x, y) = \int_K d_1((x, y), (\alpha, \beta)) h_\mu(\alpha, \beta)\, \mathrm{d}\alpha \mathrm{d}\beta, \qquad (x, y) \in \mathbb{R}^2.$$

With

$$h_\mu(\alpha, \beta) := \begin{cases} \frac{1}{A(K)} & \text{if } (\alpha, \beta) \in K, \\ 0 & \text{if } (\alpha, \beta) \notin K, \end{cases}$$

we have $F_{K,\mu}$ coincides with F_K given in Definition 2.1. Note also that the conic function F_K can be interpreted as the expectation of an appropriate random variable. Namely, $F_K(x, y) = \mathbb{E}[d_1((x, y), (\xi, \eta))]$, $(x, y) \in \mathbb{R}^2$, where (ξ, η) is a uniformly distributed random variable on K.

Next, we generalize Theorems 3, 4 and 5, Lemmas 6 and 7 and Corollary 1 in Vincze and Nagy [1] for the generalized conic function $F_{K,\mu}$.

THEOREM 2.6 *The generalized conic function* $F_{K,\mu} : \mathbb{R}^2 \to \mathbb{R}_+$ *is a convex function which satisfies the growth condition*

$$\liminf_{\|(x,y)\| \to \infty} \frac{F_{K,\mu}(x, y)}{\sqrt{x^2 + y^2}} \geq \mu(K) > 0.$$

Consequently, the level sets of the function $F_{K,\mu}$ *are bounded and hence compact subsets of* $\mathbb{R}^2$.

Proof Recall that

$$F_{K,\mu}(x, y) = \int_K d_1((x, y), (\alpha, \beta))\, \mu(\mathrm{d}\alpha, \mathrm{d}\beta), \qquad (x, y) \in \mathbb{R}^2.$$

The convexity of $F_{K,\mu}$ is clear, since the integrand is a convex function for any fixed element $(\alpha, \beta) \in K$, and the Lebesgue integral with respect to the measure μ is monotone. Further, since $d_2((x, y), (\alpha, \beta)) \leq d_1((x, y), (\alpha, \beta))$, $(x, y), (\alpha, \beta) \in \mathbb{R}^2$, where d_2 is the standard Euclidean distance on $\mathbb{R}^2$, we have

$$F_{K,\mu}(x, y) \geq \int_K d_2((x, y), (\alpha, \beta))\, \mu(\mathrm{d}\alpha, \mathrm{d}\beta), \qquad (x, y) \in \mathbb{R}^2,$$

and then

$$\frac{F_{K,\mu}(x, y)}{\sqrt{x^2 + y^2}} \geq \int_K \left(\frac{d_2((x, y), (\alpha, \beta)) - \sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}} + 1 \right) \mu(\mathrm{d}\alpha, \mathrm{d}\beta)$$

for $(x, y) \in \mathbb{R}^2$, $(x, y) \neq (0, 0)$. The triangle inequality shows that

$$\sqrt{x^2 + y^2} = d_2((x, y), (0, 0)) \leq d_2((x, y), (\alpha, \beta)) + d_2((\alpha, \beta), (0, 0)) = d_2((x, y), (\alpha, \beta)) + \sqrt{\alpha^2 + \beta^2},$$

and then

$$\frac{F_{K,\mu}(x, y)}{\sqrt{x^2 + y^2}} \geq \int_K \left(1 - \frac{\sqrt{\alpha^2 + \beta^2}}{\sqrt{x^2 + y^2}}\right) \mu(\mathrm{d}\alpha, \mathrm{d}\beta), \qquad (x, y) \in \mathbb{R}^2,\ (x, y) \neq (0, 0).$$

By Fatou's lemma,

$$\liminf_{\|(x,y)\| \to \infty} \frac{F_{K,\mu}(x, y)}{\sqrt{x^2 + y^2}} \geq \liminf_{\|(x,y)\| \to \infty} \int_K \left(1 - \frac{\sqrt{\alpha^2 + \beta^2}}{\sqrt{x^2 + y^2}}\right) \mu(\mathrm{d}\alpha, \mathrm{d}\beta) \geq \int_K \liminf_{\|(x,y)\| \to \infty} \left(1 - \frac{\sqrt{\alpha^2 + \beta^2}}{\sqrt{x^2 + y^2}}\right) \mu(\mathrm{d}\alpha, \mathrm{d}\beta) = \mu(K) > 0.$$

Here for completeness, we note that one can use Fatou's lemma, since for all $c > 0$,

$$\int_K \inf \left\{1 - \frac{\sqrt{\alpha^2 + \beta^2}}{\sqrt{x^2 + y^2}} : \|(x, y)\| \geq c\right\} \mu(\mathrm{d}\alpha, \mathrm{d}\beta) = \int_K \left(1 - \frac{\sqrt{\alpha^2 + \beta^2}}{c}\right) \mu(\mathrm{d}\alpha, \mathrm{d}\beta) > -\infty,$$

where the last inequality follows by that K is compact (hence bounded) and $\mu(K) < \infty$.

Let $d \in \mathbb{R}_+$ and let us suppose that the level set $\{(x, y) \in \mathbb{R}^2 : F_{K,\mu}(x, y) \leq d\}$ is unbounded. Then one can choose a sequence (x_n, y_n), $n \in \mathbb{N}$, such that $F_{K,\mu}(x_n, y_n) \leq d$, $n \in \mathbb{N}$, and $\lim_{n\to\infty} \|(x_n, y_n)\| = \infty$. This would imply that

$$\lim_{n\to\infty} \frac{F_{K,\mu}(x_n, y_n)}{\sqrt{x_n^2 + y_n^2}} = 0,$$

which contradicts to the growth condition. □

Lemma 2.7 *Let us suppose that Condition* (C.3) *holds. For the generalized conic function* $F_{K,\mu}$*, we have*

$$F_{K,\mu}(x, y) = x\big(\mu(\{K <_1 x\}) - \mu(\{x <_1 K\})\big) - \int_K \alpha(\mathbf{1}_{\{\alpha<x\}} - \mathbf{1}_{\{x<\alpha\}})\, \mu(\mathrm{d}\alpha, \mathrm{d}\beta) + y\big(\mu(\{K <_2 y\}) - \mu(\{y <_2 K\})\big) - \int_K \beta(\mathbf{1}_{\{\beta<y\}} - \mathbf{1}_{\{y<\beta\}})\, \mu(\mathrm{d}\alpha, \mathrm{d}\beta)$$

for all $(x, y) \in \mathbb{R}^2$.

Proof By definition,

$$F_{K,\mu}(x, y) = \int_K (|x - \alpha| + |y - \beta|)\, \mu(\mathrm{d}\alpha, \mathrm{d}\beta), \qquad (x, y) \in \mathbb{R}^2.$$

Here,

$$\begin{aligned}\int_K |x-\alpha|\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta) &= \int_{K<_1 x} |x-\alpha|\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta) + \int_{x\leq_1 K} |x-\alpha|\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &= \int_{K<_1 x} (x-\alpha)\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta) + \int_{x\leq_1 K} (\alpha-x)\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &= x\big(\mu(\{K <_1 x\}) - \mu(\{x\leq_1 K\})\big) - \int_{K<_1 x} \alpha\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &\quad + \int_{x\leq_1 K} \alpha\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta),\end{aligned}$$

and the integral $\int_K |y-\beta|\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)$ can be handled similarly. The assertion follows by taking into account Condition (C.3). □

Lemma 2.8 *Let us suppose that Condition* (C.3) *holds. For the generalized conic function* $F_{K,\mu}$*, we have*

$$\begin{aligned}D_1F_{K,\mu}(x,y) &= \mu(\{K <_1 x\}) - \mu(\{x <_1 K\}), \qquad (x,y)\in\mathbb{R}^2,\\ D_2F_{K,\mu}(x,y) &= \mu(\{K <_2 y\}) - \mu(\{y <_2 K\}), \qquad (x,y)\in\mathbb{R}^2.\end{aligned}$$

Proof Let $h>0$. Then for all $(x,y)\in\mathbb{R}^2$,

$$\begin{aligned}&\frac{F_{K,\mu}(x+h,y)-F_{K,\mu}(x,y)}{h}\\ &= \int_K \frac{|x+h-\alpha|-|x-\alpha|}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &= \int_{K<_1 x} \frac{|x+h-\alpha|-|x-\alpha|}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &\quad + \int_{x\leq_1 K\leq_1 x+h} \frac{|x+h-\alpha|-|x-\alpha|}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &\quad + \int_{x+h<_1 K} \frac{|x+h-\alpha|-|x-\alpha|}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &= \int_{K<_1 x} \frac{x+h-\alpha-(x-\alpha)}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &\quad + \int_{x\leq_1 K\leq_1 x+h} \frac{x+h-\alpha-(\alpha-x)}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &\quad + \int_{x+h<_1 K} \frac{\alpha-x-h-(\alpha-x)}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\\ &= \mu(\{K <_1 x\}) - \mu(\{x+h <_1 K\})\\ &\quad + \int_{x\leq_1 K\leq_1 x+h} \frac{|x+h-\alpha|-|x-\alpha|}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta).\end{aligned}$$

Using that $||a|-|b||\leq|a-b|$, $a,b\in\mathbb{R}$, for the integrand, we have

$$\left|\frac{|x+h-\alpha|-|x-\alpha|}{h}\right| \leq \frac{1}{h}|x+h-\alpha-(x-\alpha)| = \frac{|h|}{h} = 1, \qquad x,\alpha\in\mathbb{R},\ h>0,$$

and hence, by dominated convergence theorem,

$$\left|\int_{x\le_1 K\le_1 x+h}\frac{|x+h-\alpha|-|x-\alpha|}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)\right|$$
$$\le \int_{x\le_1 K\le_1 x+h}\left|\frac{|x+h-\alpha|-|x-\alpha|}{h}\right|\mu(\mathrm{d}\alpha,\mathrm{d}\beta)$$
$$\le \mu(\{x\le_1 K\le_1 x+h\})\to\mu(\{K=_1 x\})=0$$

as $h\downarrow 0$. Then, for all $(x,y)\in\mathbb{R}^2$,

$$\lim_{h\downarrow 0}\frac{F_{K,\mu}(x+h,y)-F_{K,\mu}(x,y)}{h}=\mu(\{K<_1 x\})-\mu(\{x\le_1 K\})$$
$$=\mu(\{K<_1 x\})-\mu(\{x<_1 K\}). \qquad (2.2)$$

Similarly, if $h<0$, then

$$\frac{F_{K,\mu}(x+h,y)-F_{K,\mu}(x,y)}{h}=\mu(\{K<_1 x+h\})-\mu(\{x<_1 K\})$$
$$+\int_{x+h\le_1 K\le_1 x}\frac{|x+h-\alpha|-|x-\alpha|}{h}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta)$$

for all $(x,y)\in\mathbb{R}^2$, and hence, using again Condition (C.3),

$$\lim_{h\uparrow 0}\frac{F_{K,\mu}(x+h,y)-F_{K,\mu}(x,y)}{h}=\mu(\{K\le_1 x\})-\mu(\{x<_1 K\})$$
$$=\mu(\{K<_1 x\})-\mu(\{x<_1 K\}) \qquad (2.3)$$

for all $(x,y)\in\mathbb{R}^2$. Then (2.2) and (2.3) yield that $D_1F_{K,\mu}(x,y)=\mu(\{K<_1 x\})-\mu(\{x<_1 K\})$, $(x,y)\in\mathbb{R}^2$.

In a similar way, we have $D_2F_{K,\mu}(x,y)=\mu(\{K<_2 y\})-\mu(\{y<_2 K\})$, $(x,y)\in\mathbb{R}^2$. □

If μ is a measure on $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$, then by the μ-area of a Borel measurable set $S\in\mathcal{B}(\mathbb{R}^d)$, we mean $\mu(S)$.

COROLLARY 2.9 *Let us suppose that Condition* (C.3) *holds. A point in* $\mathbb{R}^2$ *is a global minimizer of the generalized conic function* $F_{K,\mu}$ *if and only if it bisects the* μ*-area of* K*, i.e. the vertical and the horizontal lines through this point cut the body* K *into two parts with equal* μ*-areas. Moreover, if Conditions* (C.1) *and* (C.2) *hold too, then the convex function* $F_{K,\mu}$ *has a unique global minimizer* $(x^*,y^*)\in\mathbb{R}^2$*, that is,* $F_{K,\mu}(x,y)>F_{K,\mu}(x^*,y^*)$ *for* $(x,y)\ne(x^*,y^*)$*,* $(x,y)\in\mathbb{R}^2$.

Proof First note that under Condition (C.3), the concept of bisection of the μ-area of K is well defined. The first part of the corollary is a consequence of Lemma 2.8 using that a local minimum of a convex function defined on $\mathbb{R}^2$ is a global minimum, too. Under Conditions (C.1), (C.2) and (C.3), the existence of a global minimizer (x^*,y^*) of $F_{K,\mu}$ follows by that $F_{K,\mu}$ is a convex function defined on $\mathbb{R}^2$ and its level sets are compact subsets of $\mathbb{R}^2$ (see Theorem 2.6). Indeed, a finite-valued convex function defined on $\mathbb{R}^2$ is continuous and it reaches its minimum on every compact set. Now, we turn to prove the uniqueness of

(x^*, y^*). The proof goes along the very same lines as in the proof of Proposition 2.3. Indeed, the area A (two-dimensional Lebesgue measure) has to be replaced by the measure μ. □

Before we generalize Theorem 4 in Vincze and Nagy [1], we need to introduce some notations and to recall the Cavalieri principle for product measures.

Definition 2.10 Let μ_1 and μ_2 be σ-finite measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and let $\mu := \mu_1 \times \mu_2$ be their product measure on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. Given a measurable set $S \in \mathcal{B}(\mathbb{R}^2)$, the generalized X-ray functions of S with respect to μ into the coordinate directions are defined by

$$X_{S,\mu}(y) := \mu_1(S_y), \quad y \in \mathbb{R}, \quad \text{and} \quad Y_{S,\mu}(x) := \mu_2(S_x), \quad x \in \mathbb{R},$$

where $S_x := \{y \in \mathbb{R} : (x, y) \in S\}$ and $S_y := \{x \in \mathbb{R} : (x, y) \in S\}$. (Note that $S_x, S_y \in \mathcal{B}(\mathbb{R})$ for all $x, y \in \mathbb{R}$, see, e.g. Lemma 5.1.1 in Cohn [10].)

For the product measure μ defined in Definition 2.10, we have $\mu(K) < \infty$.

THEOREM 2.11 (The Cavalieri principle, see, e.g. Cohn [10, Theorem 5.1.3]) *Let μ_1 and μ_2 be σ-finite measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and let $\mu := \mu_1 \times \mu_2$ be their product measure on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. If $S \in \mathcal{B}(\mathbb{R}^2)$, then the functions $X_{S,\mu}, Y_{S,\mu} : \mathbb{R} \to \mathbb{R}_+$ are Borel measurable, and*

$$\mu(S) = (\mu_1 \times \mu_2)(S) = \int_{\mathbb{R}} Y_{S,\mu}(x)\mu_1(\mathrm{d}x) = \int_{\mathbb{R}} X_{S,\mu}(y)\mu_2(\mathrm{d}y).$$

THEOREM 2.12 *Let $K, K^* \subset \mathbb{R}^2$ be compact bodies, let μ_i, μ_i^*, $i = 1, 2$, be σ-finite measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ that are absolutely continuous with respect to the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with Radon-Nikodym derivatives f_i, f_i^*, $i = 1, 2$. Let $\mu := \mu_1 \times \mu_2$ and $\mu^* := \mu_1^* \times \mu_2^*$ be their product measures on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ and we assume that μ and μ^* are supported by K and K^*, respectively. Let us suppose that Condition (C.3) holds for K and μ, and K^* and μ^*, respectively. Then $F_{K,\mu} = F_{K^*,\mu^*}$ if and only if $f_2(y)X_{K,\mu}(y) = f_2^*(y)X_{K^*,\mu^*}(y)$ for (Lebesgue) almost every $y \in \mathbb{R}$, and $f_1(x)Y_{K,\mu}(x) = f_1^*(x)Y_{K^*,\mu^*}(x)$ for (Lebesgue) almost every $x \in \mathbb{R}$.*

Proof By Theorem 2.11 (the Cavalieri principle), for all $x, y \in \mathbb{R}$,

$$\begin{aligned}
\mu(K <_1 x) &= \int_{\mathbb{R}} Y_{K <_1 x,\mu}(s)\,\mu_1(\mathrm{d}s) = \int_{-\infty}^{x} Y_{K,\mu}(s)\,\mu_1(\mathrm{d}s) = \int_{-\infty}^{x} Y_{K,\mu}(s) f_1(s)\,\mathrm{d}s,\\
\mu(x <_1 K) &= \int_{\mathbb{R}} Y_{x <_1 K,\mu}(s)\,\mu_1(\mathrm{d}s) = \int_{x}^{\infty} Y_{K,\mu}(s)\,\mu_1(\mathrm{d}s) = \int_{x}^{\infty} Y_{K,\mu}(s) f_1(s)\,\mathrm{d}s,\\
\mu(K <_2 y) &= \int_{\mathbb{R}} X_{K <_2 y,\mu}(t)\,\mu_2(\mathrm{d}t) = \int_{-\infty}^{y} X_{K,\mu}(t)\,\mu_2(\mathrm{d}t) = \int_{-\infty}^{y} X_{K,\mu}(t) f_2(t)\,\mathrm{d}t,\\
\mu(y <_2 K) &= \int_{\mathbb{R}} X_{y <_2 K,\mu}(t)\,\mu_2(\mathrm{d}t) = \int_{y}^{\infty} X_{K,\mu}(t)\,\mu_2(\mathrm{d}t) = \int_{y}^{\infty} X_{K,\mu}(t) f_2(t)\,\mathrm{d}t,
\end{aligned} \tag{2.4}$$

and, by Fubini's theorem, for all $x, y \in \mathbb{R}$,

$$\begin{aligned}
\int_K \alpha \mathbf{1}_{\{\alpha<x\}}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta) &= \int_{-\infty}^{x} sY_{K,\mu}(s)\,\mu_1(\mathrm{d}s) = \int_{-\infty}^{x} sY_{K,\mu}(s)f_1(s)\,\mathrm{d}s,\\
\int_K \alpha \mathbf{1}_{\{x<\alpha\}}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta) &= \int_{x}^{\infty} sY_{K,\mu}(s)\,\mu_1(\mathrm{d}s) = \int_{x}^{\infty} sY_{K,\mu}(s)f_1(s)\,\mathrm{d}s,\\
\int_K \beta \mathbf{1}_{\{\beta<y\}}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta) &= \int_{-\infty}^{y} tX_{K,\mu}(t)\,\mu_2(\mathrm{d}t) = \int_{-\infty}^{y} tX_{K,\mu}(t)f_2(t)\,\mathrm{d}t,\\
\int_K \beta \mathbf{1}_{\{y<\beta\}}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta) &= \int_{y}^{\infty} tX_{K,\mu}(t)\,\mu_2(\mathrm{d}t) = \int_{y}^{\infty} tX_{K,\mu}(t)f_2(t)\,\mathrm{d}t. \qquad (2.5)
\end{aligned}$$

Indeed, for example, the first statement of (2.5) holds since, by Fubini's theorem for non-rectangular regions,

$$\begin{aligned}
\int_K \alpha \mathbf{1}_{\{\alpha<x\}}\,\mu(\mathrm{d}\alpha,\mathrm{d}\beta) &= \int_{\alpha_b}^{\alpha_u}\left(\int_{K_\alpha} \alpha \mathbf{1}_{\{\alpha<x\}}\,\mu_2(\mathrm{d}\beta)\right)\mu_1(\mathrm{d}\alpha)\\
&= \int_{\alpha_b}^{\alpha_u} \alpha \mathbf{1}_{\{\alpha<x\}}\mu_2(K_\alpha)\,\mu_1(\mathrm{d}\alpha)\\
&= \int_{\alpha_b}^{\alpha_u} \alpha \mathbf{1}_{\{\alpha<x\}}Y_{K,\mu}(\alpha)\,\mu_1(\mathrm{d}\alpha)\\
&= \int_{-\infty}^{x} sY_{K,\mu}(s)\,\mu_1(\mathrm{d}s),
\end{aligned}$$

where $K_\alpha = \{\beta \in \mathbb{R} \mid (\alpha,\beta) \in K\}$ and

$$\alpha_b := \inf\left\{\alpha \mid \exists\, \beta \in \mathbb{R} : (\alpha,\beta) \in K\right\}, \quad \alpha_u := \sup\left\{\alpha \mid \exists\, \beta \in \mathbb{R} : (\alpha,\beta) \in K\right\}.$$

Further, by (2.4), Lemma 2.8 and Lebesgue differentiation theorem,

$$\begin{aligned}
D_1D_1F_{K,\mu}(x,y) &= D_1\big(\mu(\{K <_1 x\}) - \mu(\{x <_1 K\})\big)\\
&= D_1\left(\int_{-\infty}^{x} Y_{K,\mu}(s)f_1(s)\,\mathrm{d}s - \int_{x}^{\infty} Y_{K,\mu}(s)f_1(s)\,\mathrm{d}s\right)\\
&= 2Y_{K,\mu}(x)f_1(x) \quad \text{for all } y \in \mathbb{R} \text{ and almost every } x \in \mathbb{R}, \qquad (2.6)
\end{aligned}$$

and similarly,

$$\begin{aligned}
&D_1D_2F_{K,\mu}(x,y) = D_2D_1F_{K,\mu}(x,y) = 0 \quad \text{for all } (x,y) \in \mathbb{R}^2,\\
&D_2D_2F_{K,\mu}(x,y) = 2X_{K,\mu}(y)f_2(y) \quad \text{for all } x \in \mathbb{R} \text{ and almost every } y \in \mathbb{R}. \qquad (2.7)
\end{aligned}$$

Let us suppose that $F_{K,\mu} = F_{K^*,\mu^*}$. By (2.6) and (2.7), we have $f_1(x)Y_{K,\mu}(x) = f_1^*(x)Y_{K^*,\mu^*}(x)$ for almost every $x \in \mathbb{R}$, and $f_2(y)X_{K,\mu}(y) = f_2^*(y)X_{K^*,\mu^*}(y)$ for almost every $y \in \mathbb{R}$, as desired.

Conversely, let us suppose that $f_2(y)X_{K,\mu}(y) = f_2^*(y)X_{K^*,\mu^*}(y)$ for almost every $y \in \mathbb{R}$, and $f_1(x)Y_{K,\mu}(x) = f_1^*(x)Y_{K^*,\mu^*}(x)$ for almost every $x \in \mathbb{R}$. Then, by Lemma 2.7, (2.4) and (2.5), we get $F_{K,\mu} = F_{K^*,\mu^*}$. □

Remark 2 Note that, under the conditions of Theorem 2.12, for almost every $(x, y) \in \mathbb{R}^2$, the matrix consisting of the second-order partial derivatives of $F_{K,\mu}$ takes the form

$$\begin{bmatrix} 2f_1(x)Y_{K,\mu}(x) & 0 \\ 0 & 2f_2(y)X_{K,\mu}(y) \end{bmatrix},$$

which is a positive semidefinite matrix, since the Radon-Nikodym derivatives f_i and f_i^*, $i = 1, 2$ are non-negative almost everywhere. Note also that this is in accordance with the fact that $F_{K,\mu}$ is a convex function due to Theorem 2.6.

Before we generalize Theorem 5 in Vincze and Nagy [1], we need to recall some notions.

Definition 2.13 Let K be a compact body in $\mathbb{R}^2$. For all $\varepsilon > 0$, the outer parallel body K^ε is the union of closed Euclidean balls centred at the points of K with radius $\varepsilon > 0$.

Definition 2.14 The Hausdorff distance between two compact bodies K and L is given by

$$H(K, L) := \inf\left\{\varepsilon > 0 : K \subset L^\varepsilon \text{ and } L \subset K^\varepsilon\right\}.$$

The collection of compact bodies in $\mathbb{R}^2$ furnished with the Hausdorff distance H is a metric space, see, e.g. Beer [11].

LEMMA 2.15 *Let K_n, $n \in \mathbb{N}$, K be compact bodies, and let μ be a Radon measure on* $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.

(i) *We have* $\lim_{\varepsilon\downarrow 0} \mu(K^\varepsilon) = \mu(K)$.
(ii) *If $K_n \to K$ as $n \to \infty$ with respect to the Hausdorff metric H, then the following regularity properties are equivalent:*

  (a) $\lim_{n\to\infty} \mu((K \setminus K_n) \cup (K_n \setminus K)) = 0$,
  (b) $\lim_{n\to\infty} \mu(K_n) = \mu(K)$.

Proof The proofs go along the very same lines as those of Lemmas 1 and 2 in Vincze and Nagy [1] by replacing the area A (two-dimensional Lebesgue measure) by the measure μ in the proofs and referring to that $\mu(L) < \infty$ for all compact sets $L \subset \mathbb{R}^2$ (due to that μ is a Radon measure). □

Definition 2.16 Let K_n, $n \in \mathbb{N}$, and K be compact bodies, and let μ be a Radon measure on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. The convergence $K_n \to K$ as $n \to \infty$ with respect to the Hausdorff metric is called regular if one of the conditions (a) and (b) of part (ii) of Lemma 2.15 holds.

THEOREM 2.17 *Let K_n, $n \in \mathbb{N}$, and K be compact bodies, and let μ be a Radon measure on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ supported by K^ε for some $\varepsilon > 0$. Let us suppose that the convergence $K_n \to K$ as $n \to \infty$ with respect to the Hausdorff metric is regular. Then*

$$\lim_{n\to\infty} F_{K_n,\mu}(x, y) = F_{K,\mu}(x, y), \qquad (x, y) \in \mathbb{R}^2.$$

Proof The proof goes along the very same lines as that of Theorem 5 in Vincze and Nagy [1], but replacing the integration with respect to the two-dimensional Lebesgue measure by the integration with respect to the measure μ. □

For the remaining sections of the paper, we will need some further properties of the convex function $F_{K,\mu}$. Next, we recall some general facts from the theory of convex functions, see, e.g. Polyak [12, Lemma 3, Section 1.1.4].

LEMMA 2.18 *Let* $F : \mathbb{R}^d \to \mathbb{R}$ *be a differentiable and convex function such that its gradient is Lipschitz continuous with constant* $L > 0$*, i.e.*

$$\|\text{grad } F(p) - \text{grad } F(q)\| \leq L\|p - q\|, \qquad p, q \in \mathbb{R}^d, \tag{2.8}$$

where $\text{grad } F(p) := (D_1 F(p), D_2 F(p))^\top$, $p \in \mathbb{R}^d$. *Then we have an affine lower bound*

$$F(q) \geq F(p) + \langle \text{grad } F(p), q - p \rangle, \qquad p, q \in \mathbb{R}^d.$$

LEMMA 2.19 *Let* μ_1 *and* μ_2 *be* σ*-finite measures on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *that are absolutely continuous with respect to the Lebesgue measure on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *with bounded Radon-Nikodym derivatives. Let* $\mu := \mu_1 \times \mu_2$ *be their product measure on* $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ *and we assume that* μ *is supported by* K*. Further, let us suppose that Condition* (C.3) *holds. Then the generalized conic function* $F_{K,\mu} : \mathbb{R}^2 \to \mathbb{R}$ *associated with* K *and* μ *satisfies the conditions of Lemma 2.18, and consequently, we have an affine lower bound for* $F_{K,\mu}$.

Proof By Theorem 2.6, $F_{K,\mu}$ is convex. Under Condition (C.3), by Lemma 2.8 and (2.4),

$$\begin{aligned} D_1 F_{K,\mu}(x, y) &= \int_{-\infty}^{x} Y_{K,\mu}(s)\, \mu_1(\mathrm{d}s) - \int_{x}^{\infty} Y_{K,\mu}(s)\, \mu_1(\mathrm{d}s) \\ &= \int_{-\infty}^{x} Y_{K,\mu}(s) f_1(s)\, \mu_1(\mathrm{d}s) - \int_{x}^{\infty} Y_{K,\mu}(s) f_1(s)\, \mu_1(\mathrm{d}s) \end{aligned}$$

for $(x, y) \in \mathbb{R}^2$, where f_1 denotes the (bounded) Radon-Nikodym derivative of μ_1 with respect to the Lebesgue measure on $\mathbb{R}$. Using that the integral as a function of the upper limit of the integration is continuous, we have $D_1 F_{K,\mu}$ is continuous on $\mathbb{R}^2$. Similarly, one can check that $D_2 F_{K,\mu}$ is also continuous on $\mathbb{R}^2$. This implies that $F_{K,\mu}$ is differentiable on $\mathbb{R}^2$.

Condition (2.8) for $F_{K,\mu}$ can be checked as follows. Let us start with the difference of the partial derivatives with respect to the first variable

$$\begin{aligned} &D_1 F_{K,\mu}(q) - D_1 F_{K,\mu}(p) \\ &\quad = \mu(K <_1 q^{(1)}) - \mu(q^{(1)} <_1 K) - (\mu(K <_1 p^{(1)}) - \mu(p^{(1)} <_1 K)) \end{aligned}$$

for all $p = (p^{(1)}, p^{(2)})$, $q = (q^{(1)}, q^{(2)}) \in \mathbb{R}^2$, where the equality follows by Lemma 2.8. We have

$$\mu(K <_1 q^{(1)}) = \mu(K <_1 \min\{p^{(1)}, q^{(1)}\}) + \mu(\min\{p^{(1)}, q^{(1)}\} <_1 K <_1 q^{(1)})$$

and

$$\mu(q^{(1)} <_1 K) = \mu(\max\{p^{(1)}, q^{(1)}\} <_1 K) + \mu(q^{(1)} <_1 K <_1 \max\{p^{(1)}, q^{(1)}\}).$$

Of course we can change the role of q and p to express $\mu(K <_1 p^{(1)})$ and $\mu(p^{(1)} <_1 K)$ in a similar way. Then

$$\begin{aligned} & D_1 F_{K,\mu}(q) - D_1 F_{K,\mu}(p) \\ & \quad = \mu(\min\{p^{(1)}, q^{(1)}\} <_1 K <_1 q^{(1)}) - \mu(q^{(1)} <_1 K <_1 \max\{p^{(1)}, q^{(1)}\}) \\ & \qquad - \mu(\min\{p^{(1)}, q^{(1)}\} <_1 K <_1 p^{(1)}) + \mu(p^{(1)} <_1 K <_1 \max\{p^{(1)}, q^{(1)}\}). \end{aligned}$$

Hence, we can see that if $p^{(1)} = \min\{p^{(1)}, q^{(1)}\}$ and consequently, $q^{(1)} = \max\{p^{(1)}, q^{(1)}\}$, then

$$D_1 F_{K,\mu}(q) - D_1 F_{K,\mu}(p) = 2\mu(p^{(1)} <_1 K <_1 q^{(1)}).$$

If $q^{(1)} = \min\{p^{(1)}, q^{(1)}\}$ and $p^{(1)} = \max\{p^{(1)}, q^{(1)}\}$, then

$$D_1 F_{K,\mu}(q) - D_1 F_{K,\mu}(p) = -2\mu(q^{(1)} <_1 K <_1 p^{(1)}).$$

In general,

$$|D_1 F_{K,\mu}(q) - D_1 F_{K,\mu}(p)| = 2\mu(\min\{p^{(1)}, q^{(1)}\} <_1 K <_1 \max\{p^{(1)}, q^{(1)}\}).$$

Therefore, using Theorem 2.11 (the Cavalieri principle), we can estimate the difference of the absolute value of the first-order partial derivatives of $F_{K,\mu}$ as follows:

$$\begin{aligned} |D_1 F_{K,\mu}(q) - D_1 F_{K,\mu}(p)| & \leq 2 \int_{\min\{p^{(1)},q^{(1)}\}}^{\max\{p^{(1)},q^{(1)}\}} Y_{K,\mu}(s)\, \mu_1(\mathrm{d}s) \\ & \leq 2 \left(\sup_{s\in\mathbb{R}} Y_{K,\mu}(s) \right) \mu_1\big((\min\{p^{(1)}, q^{(1)}\}, \max\{p^{(1)}, q^{(1)}\}) \big) \\ & = 2 \left(\sup_{s\in\mathbb{R}} Y_{K,\mu}(s) \right) \int_{\min\{p^{(1)},q^{(1)}\}}^{\max\{p^{(1)},q^{(1)}\}} f_1(s)\, \mathrm{d}s \\ & \leq 2C_1 \left(\sup_{s\in\mathbb{R}} Y_{K,\mu}(s) \right) |p^{(1)} - q^{(1)}| \end{aligned}$$

with some constant $C_1 > 0$, where $\sup_{s\in\mathbb{R}} Y_{K,\mu}(s) < \infty$ (since $\mu(K) < \infty$), and f_1 denotes the bounded Radon-Nikodym derivative of μ_1 with respect to the Lebesgue measure on $\mathbb{R}$. Similarly,

$$|D_2 F_{K,\mu}(q) - D_2 F_{K,\mu}(p)| \leq 2C_2 \left(\sup_{t\in\mathbb{R}} X_{K,\mu}(t) \right) |p^{(2)} - q^{(2)}|$$

with some constant $C_2 > 0$. Therefore,

$$\begin{aligned} & \|\mathrm{grad}\ F_{K,\mu}(p) - \mathrm{grad}\ F_{K,\mu}(q)\| \\ & \quad = \sqrt{(D_1 F_{K,\mu}(p) - D_1 F_{K,\mu}(q))^2 + (D_2 F_{K,\mu}(p) - D_2 F_{K,\mu}(q))^2} \\ & \quad \leq L\|p - q\|, \qquad p, q \in \mathbb{R}^2, \end{aligned}$$

where

$$L := 2 \max \left\{ C_1 \sup_{s\in\mathbb{R}} Y_{K,\mu}(s),\, C_2 \sup_{t\in\mathbb{R}} X_{K,\mu}(t) \right\},$$

i.e. condition (2.8) for $F_{K,\mu}$ is satisfied with $d = 2$ and with the Lipschitz constant L given above. □

3. A stochastic algorithm for the global minimizer of $F_{K,\mu}$

We provide a stochastic algorithm for computing the global minimizer of generalized conic function $F_{K,\mu}$ introduced in Definition 2.5, and we prove almost sure and L^q-convergence of this algorithm.

In this section, we assume that

(C.4) μ is a probability measure on K.

Let $(t_k)_{k\in\mathbb{N}}$ be a decreasing sequence of positive numbers such that $\sum_{k=1}^{\infty} t_k = \infty$ and $\sum_{k=1}^{\infty} t_k^2 < \infty$.

Let $(P_k)_{k\in\mathbb{N}}$ be a sequence of independent identically distributed (two-dimensional) random variables such that their common distribution on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ is given by μ. Let $x_0 \in K$ be arbitrarily chosen. We define recursively a Markov chain $(X_k)_{k\in\mathbb{Z}_+}$ by

$$X_0 := x_0, \quad \text{and} \quad X_{k+1} := X_k - t_{k+1}Q_{k+1}, \quad k \in \mathbb{Z}_+, \tag{3.1}$$

where

$$Q_{k+1} := \begin{cases} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \text{if } X_k^{(1)} \geq P_{k+1}^{(1)} \text{ and } X_k^{(2)} \geq P_{k+1}^{(2)}, \\ \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \text{if } X_k^{(1)} \geq P_{k+1}^{(1)} \text{ and } X_k^{(2)} < P_{k+1}^{(2)}, \\ \begin{pmatrix} -1 \\ 1 \end{pmatrix} & \text{if } X_k^{(1)} < P_{k+1}^{(1)} \text{ and } X_k^{(2)} \geq P_{k+1}^{(2)}, \\ \begin{pmatrix} -1 \\ -1 \end{pmatrix} & \text{if } X_k^{(1)} < P_{k+1}^{(1)} \text{ and } X_k^{(2)} < P_{k+1}^{(2)}, \end{cases}$$

with the notations $X_k := (X_k^{(1)}, X_k^{(2)})$, $P_k := (P_k^{(1)}, P_k^{(2)})$, $k \in \mathbb{N}$.

Remark 1 Note that if μ is a probability measure on K such that it is absolutely continuous with respect to the Lebesgue measure on K with Radon-Nikodym derivative (density function) h_μ given by

$$h_\mu(x, y) = \begin{cases} \frac{1}{A(K)} & \text{if } (x, y) \in K, \\ 0 & \text{if } (x, y) \notin K, \end{cases}$$

i.e. μ is the uniform distribution on K, then $(P_k)_{k\in\mathbb{N}}$ is a sequence of independent identically distributed (two-dimensional) random variables such that their common distribution is the uniform distribution on K.

3.1. *Almost sure and L^q-convergence of $(X_k)_{k\in\mathbb{Z}_+}$*

First, we recall the so-called Robbins–Monro algorithm based on Bouleau and Lépingle [8, Theorem B.5.1, Chapter 2]. This algorithm (in dimension 1) was originally invented by Robbins and Monro [6].

Let $d \in \mathbb{N}$ and $(t_n)_{n\in\mathbb{Z}_+}$ be a decreasing sequence of positive real numbers. Let us suppose that all the random variables introduced below are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The Robbins–Monro algorithm generates a sequence of $\mathbb{R}^d$-valued random variables $(\theta_n)_{n\in\mathbb{Z}_+}$ given by the recursion

$$\theta_{n+1} := \theta_n + t_{n+1}(\beta - \xi_{n+1}), \quad n \in \mathbb{Z}_+,$$

where $\beta \in \mathbb{R}^d$, θ_0 is a given $\mathbb{R}^d$-valued random variable and $(\xi_n)_{n\in\mathbb{Z}_+}$ is a sequence of d-dimensional random variables such that there exists a Borel measurable function $M : \mathbb{R}^d \to \mathbb{R}^d$ satisfying

$$\mathbb{E}(\xi_{n+1} \mid \mathcal{F}_n) = M(\theta_n) \qquad \mathbb{P}\text{-almost surely for all } n \in \mathbb{N},$$

where the filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}_+}$ is defined by $\mathcal{F}_0 := \sigma(\theta_0)$ (the sigma-algebra generated by θ_0) and $\mathcal{F}_n := \sigma(\theta_0, \theta_1, \dots, \theta_n, \xi_1, \dots, \xi_n)$, $n \in \mathbb{N}$ (the sigma-algebra generated by $\theta_0, \theta_1, \dots, \theta_n, \xi_1, \dots, \xi_n$).

The following assumptions will be used.

Assumption (A.1) The $\mathbb{R}^d$-valued random variable θ_0 belongs to $L^q(\Omega, \mathcal{F}, \mathbb{P})$, where $q \in \mathbb{N}$.

Assumption (A.2) There exists some $B > 0$ such that $\|\xi_n\| \le B$ for all $n \in \mathbb{N}$.

Assumption (A.3) There exists some $\theta^* \in \mathbb{R}^d$ such that for each $\varepsilon \in (0, 1)$,

$$\inf_{\varepsilon \le \|\theta-\theta^*\| \le 1/\varepsilon} \langle \theta - \theta^*, M(\theta) - \beta \rangle > 0,$$

where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $\mathbb{R}^d$. Here, Assumption (A.3) could be interpreted as a 'half-space' assumption: roughly speaking, given the value of θ_n, the expected value of θ_{n+1} will be on that side of the hyperplane through θ_n having normal vector $\theta^* - \theta_n$ which contains θ^*.

THEOREM 3.1 [Almost sure and L^q-convergence of Robbins–Monro algorithm] *Let us suppose that Assumptions* (A.1), (A.2) *and* (A.3) *hold and that the decreasing sequence* $(t_n)_{n\in\mathbb{Z}_+}$ *of positive numbers satisfies*

$$\sum_{n=0}^{\infty} t_n = \infty \quad \text{and} \quad \sum_{n=0}^{\infty} t_n^2 < \infty.$$

Then $\mathbb{P}(\lim_{n\to\infty} \theta_n = \theta^*) = 1$ *and* $\lim_{n\to\infty} \mathbb{E}\,\|\theta_n - \theta^*\|^q = 0$ *for all* $q \in \mathbb{N}$.

Note that under the conditions of Theorem 3.1, the point $\theta^* \in \mathbb{R}^d$ exists uniquely due to that, by Theorem 3.1, $\mathbb{P}(\lim_{n\to\infty} \theta_n = \theta^*) = 1$ and the limit of an almost surely convergent sequence of random variables is unique (up to probability one). We also mention that, from a technical point of view, Assumption (A.3) is used for defining an appropriate non-negative supermartingale in order to prove the almost sure convergence of the sequence $(\theta_n)_{n\in\mathbb{Z}_+}$, see, e.g. Bouleau and Lépingle [8, proof of Theorem B.5.1, Chapter 2].

We will prove almost sure and L^q-convergence of the recursion given in (3.1). But, first we present an auxiliary lemma.

Lemma 3.2 *Let us consider the sequence* $(X_k)_{k\in\mathbb{Z}_+}$ *defined by* (3.1). *Let us suppose that Conditions* (C.3) *and* (C.4) *hold. Then*

$$\mathbb{E}(Q_i \mid X_{i-1}) = \operatorname{grad} F_{K,\mu}(X_{i-1}), \qquad i \in \mathbb{N}, \tag{3.2}$$

and

$$\mathbb{E}(X_k) = x_0 - \sum_{i=1}^{k} t_i\, \mathbb{E}(\operatorname{grad} F_{K,\mu}(X_{i-1})), \quad k \in \mathbb{N}.$$

Proof First note that $X_k = x_0 - \sum_{i=1}^{k} t_i Q_i$, $k \in \mathbb{N}$, where the sequence $(Q_i)_{i\in\mathbb{N}}$ is such that the conditional distribution of Q_i with respect to X_{i-1} is given by

$$Q_i = \begin{cases} \begin{pmatrix}1\\1\end{pmatrix} & \text{with probability } \mu\left(\left\{(x,y)\in K : X_{i-1}^{(1)} \ge x,\ X_{i-1}^{(2)} \ge y\right\}\right), \\ \begin{pmatrix}1\\-1\end{pmatrix} & \text{with probability } \mu\left(\left\{(x,y)\in K : X_{i-1}^{(1)} \ge x,\ X_{i-1}^{(2)} < y\right\}\right), \\ \begin{pmatrix}-1\\1\end{pmatrix} & \text{with probability } \mu\left(\left\{(x,y)\in K : X_{i-1}^{(1)} < x,\ X_{i-1}^{(2)} \ge y\right\}\right), \\ \begin{pmatrix}-1\\-1\end{pmatrix} & \text{with probability } \mu\left(\left\{(x,y)\in K : X_{i-1}^{(1)} < x,\ X_{i-1}^{(2)} < y\right\}\right). \end{cases} \tag{3.3}$$

Then

$$\begin{aligned} \mathbb{E}(Q_i \mid X_{i-1}) &= \begin{pmatrix}1\\1\end{pmatrix} \mu\left(\left\{(x,y)\in K : X_{i-1}^{(1)} \ge x,\ X_{i-1}^{(2)} \ge y\right\}\right) \\ &\quad + \begin{pmatrix}1\\-1\end{pmatrix} \mu\left(\left\{(x,y)\in K : X_{i-1}^{(1)} \ge x,\ X_{i-1}^{(2)} < y\right\}\right) \\ &\quad + \begin{pmatrix}-1\\1\end{pmatrix} \mu\left(\left\{(x,y)\in K : X_{i-1}^{(1)} < x,\ X_{i-1}^{(2)} \ge y\right\}\right) \\ &\quad + \begin{pmatrix}-1\\-1\end{pmatrix} \mu\left(\left\{(x,y)\in K : X_{i-1}^{(1)} < x,\ X_{i-1}^{(2)} < y\right\}\right) \\ &= \begin{pmatrix} \mu(\{(x,y)\in K : X_{i-1}^{(1)} \ge x\}) - \mu(\{(x,y)\in K : X_{i-1}^{(1)} < x\}) \\ \mu(\{(x,y)\in K : X_{i-1}^{(2)} \ge y\}) - \mu(\{(x,y)\in K : X_{i-1}^{(2)} < y\}) \end{pmatrix} \end{aligned}$$

for $i \in \mathbb{N}$. Note that by Condition (C.3) and Lemma 2.8, we also have

$$\mathbb{E}(Q_i \mid X_{i-1}) = \begin{pmatrix} D_1 F_{K,\mu}(X_{i-1}^{(1)}, X_{i-1}^{(2)}) \\ D_2 F_{K,\mu}(X_{i-1}^{(1)}, X_{i-1}^{(2)}) \end{pmatrix} = \operatorname{grad} F_{K,\mu}(X_{i-1}), \qquad i \in \mathbb{N}.$$

Hence, by the tower rule, the expectation of X_k takes the form

$$\begin{aligned} \mathbb{E}(X_k) &= x_0 - \sum_{i=1}^{k} t_i\, \mathbb{E}(Q_i) = x_0 - \sum_{i=1}^{k} t_i\, \mathbb{E}(\mathbb{E}(Q_i \mid X_{i-1})) \\ &= x_0 - \sum_{i=1}^{k} t_i\, \mathbb{E}(\operatorname{grad} F_{K,\mu}(X_{i-1})), \qquad k \in \mathbb{N}. \end{aligned}$$

□

THEOREM 3.3 *Let us suppose that Conditions* (C.1)–(C.4) *hold. Then the sequence of 2-dimensional random variables defined in* (3.1) *converges almost surely and in* L^q $(q \in \mathbb{N})$ *to the unique global minimizer* X^* *of the generalized conic function* $F_{K,\mu}$, *i.e.* $\mathbb{P}(\lim_{n\to\infty} X_n = X^*) = 1$ *and* $\lim_{n\to\infty} \mathbb{E}\,\|X_n - X^*\|^q = 0$.

Proof First note that under Conditions (C.1)–(C.3), there exists a unique global minimizer θ^* of $F_{K,\mu}$, that is $F_{K,\mu}(\theta) > F_{K,\mu}(\theta^*)$ for all $\theta \neq \theta^*$, $\theta \in \mathbb{R}^2$, see, Corollary 2.9. Let us apply Theorem 3.1 with the following choices:

- $d := 2$, $\beta := 0 \in \mathbb{R}^2$ and $\xi_{n+1} := Q_{n+1}$, $n \in \mathbb{Z}_+$.
- $\theta^* \in \mathbb{R}^2$ is such that grad $F_{K,\mu}(\theta^*) = 0 \in \mathbb{R}^2$. Note that under the Conditions (C.1)–(C.3), by Corollary 2.9, θ^* is unique, and it is nothing else but the unique global minimizer of $F_{K,\mu}$.

In what follows we check that Assumptions (A.1)–(A.3) hold. Assumption (A.1) holds trivially. Assumption (A.2) holds with $B := \sqrt{2}$, since

$$\left\|\begin{pmatrix}1\\1\end{pmatrix}\right\| = \left\|\begin{pmatrix}1\\-1\end{pmatrix}\right\| = \left\|\begin{pmatrix}-1\\1\end{pmatrix}\right\| = \left\|\begin{pmatrix}-1\\-1\end{pmatrix}\right\| = \sqrt{2}.$$

Since $\mathbb{E}(Q_i \mid X_0, X_1, \dots, X_{i-1}, Q_1, \dots, Q_{i-1}) = \mathbb{E}(Q_i \mid X_{i-1})$, by (3.2), we have $M : \mathbb{R}^2 \to \mathbb{R}^2$, $M(\theta) = \operatorname{grad} F_{K,\mu}(\theta)$, $\theta \in \mathbb{R}^2$, and by Corollary 2.9,

$$M(\theta^*) = \operatorname{grad} F_{K,\mu}(\theta^*) = 0 \in \mathbb{R}^2.$$

Finally, for Assumption (A.3), we have to check that for all $\varepsilon \in (0, 1)$,

$$\inf_{\varepsilon \le \|\theta-\theta^*\| \le 1/\varepsilon} \langle \theta - \theta^*, \operatorname{grad} F_{K,\mu}(\theta)\rangle > 0.$$

Since $F_{K,\mu}$ is a convex and differentiable function defined on $\mathbb{R}^2$ (see, Theorem 2.6 and the proof of Lemma 2.19), we have

$$\langle \operatorname{grad} F_{K,\mu}(\theta), \theta^* - \theta\rangle \le F_{K,\mu}(\theta^*) - F_{K,\mu}(\theta) \le 0, \qquad \forall\, \theta \in \mathbb{R}^2, \tag{3.4}$$

where the last inequality follows by that θ^* is the global minimizer of $F_{K,\mu}$, see also Lemma 2.18. Since θ^* is strict global minimizer of $F_{K,\mu}$, i.e. $F_{K,\mu}(\theta) > F_{K,\mu}(\theta^*)$ for all $\theta \neq \theta^*$, $\theta \in \mathbb{R}^2$ (see Corollary 2.9) and $\{\theta \in \mathbb{R}^2 : \varepsilon \le \|\theta - \theta^*\| \le 1/\varepsilon\}$ is a compact set, by (3.4), we get Assumption (A.3) holds in our case. □

Example 3.4 Let K be the square with vertexes $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$ as in part (i) of Example 2.4. Let us assume that μ is the probability measure on K with Radon-Nikodym derivative with respect to the Lebesgue measure given by

$$h_\mu(x, y) = \begin{cases} 1 & \text{if } (x, y) \in K, \\ 0 & \text{if } (x, y) \notin K. \end{cases}$$

Further, let $x_0 := (0,0)^\top$ and $t_k := \frac{1}{k}$, $k \in \mathbb{N}$. Then

$$X_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad X_k = -\sum_{i=1}^{k} t_i Q_i = -\sum_{i=1}^{k} \frac{1}{i} Q_i, \quad k \in \mathbb{N},$$

where the sequence $(Q_i)_{i \in \mathbb{N}}$ is such that the conditional distribution of Q_i with respect to X_{i-1} is given by (3.3). By Theorem 3.3 and part (i) of Example 2.4, we have $\mathbb{P}(\lim_{k\to\infty} X_k = X^*) = 1$ and $\lim_{k\to\infty} \mathbb{E}\,\|X_k - X^*\|^q = 0$ for all $q \in \mathbb{N}$, where $X^* = (1/2, 1/2)^\top$. Note also that if $X_{i-1} \in K$, then the conditional distribution of Q_i with respect to X_{i-1} takes the form

$$Q_i = \begin{cases} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \text{with probability } X_{i-1}^{(1)} X_{i-1}^{(2)}, \\ \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \text{with probability } X_{i-1}^{(1)} \left(1 - X_{i-1}^{(2)}\right), \\ \begin{pmatrix} -1 \\ 1 \end{pmatrix} & \text{with probability } \left(1 - X_{i-1}^{(1)}\right) X_{i-1}^{(2)}, \\ \begin{pmatrix} -1 \\ -1 \end{pmatrix} & \text{with probability } \left(1 - X_{i-1}^{(1)}\right)\left(1 - X_{i-1}^{(2)}\right). \end{cases}$$

Finally, we remark that $X_1 = (1,1)^\top$ and $X_2 = (1/2, 1/2)^\top$.

3.2. *Almost sure and L^q-convergence of $(F_{K,\mu}(X_k))_{k\in\mathbb{Z}_+}$*

First we recall an equivalent reformulation of L^q-convergence, where $q \in \mathbb{N}$, see, e.g. Chow and Teicher [13, Theorem 4.2.3].

Lemma 3.5 *Let $d, q \in \mathbb{N}$, $\xi : \Omega \to \mathbb{R}^d$ and $\xi_n : \Omega \to \mathbb{R}^d$, $n \in \mathbb{N}$, be $\mathbb{R}^d$-valued random variables such that $\mathbb{E}(\|\xi\|^q) < \infty$ and $\mathbb{E}(\|\xi_n\|^q) < \infty$, $n \in \mathbb{N}$. Then ξ_n converges to ξ in L^q as $n \to \infty$ (i.e. $\lim_{n\to\infty} \mathbb{E}(\|\xi_n - \xi\|^q) = 0$) if and only if ξ_n converges in probability to ξ as $n \to \infty$ and the set of random variables $\{\|\xi_n\|^q : n \in \mathbb{N}\}$ is uniformly integrable, i.e.*

$$\lim_{m\to\infty} \sup_{n\in\mathbb{N}} \mathbb{E}\left(\|\xi_n\|^q \mathbf{1}_{\{\|\xi_n\|^q > m\}}\right) = 0.$$

Theorem 3.6 *Let us suppose that Conditions* (C.1)–(C.4) *hold. Then the sequence of one-dimensional random variables $(F_{K,\mu}(X_k))_{k\in\mathbb{N}}$ converges almost surely and in L^q ($q \in \mathbb{N}$) to $F_{K,\mu}(X^*)$ as $k \to \infty$, where X^* denotes the unique global minimizer of $F_{K,\mu}$.*

Proof By Theorem 3.3, $\mathbb{P}(\lim_{k\to\infty} X_k = X^*) = 1$, and hence to prove that $\mathbb{P}(\lim_{k\to\infty} F_{K,\mu}(X_k) = F_{K,\mu}(X^*)) = 1$, it is enough to check that $F_{K,\mu}$ is continuous. This follows by that $F_{K,\mu}$ is a convex function defined on $\mathbb{R}^2$ (see Theorem 2.6). We give an alternative argument, too. Let $(x_n, y_n)^\top \in \mathbb{R}^2$, $n \in \mathbb{N}$, be such that $\lim_{n\to\infty}(x_n, y_n) = (x, y)$, where $(x, y)^\top \in \mathbb{R}^2$. Then for all $(\alpha, \beta)^\top \in \mathbb{R}^2$, $\lim_{n\to\infty} d_1((x_n, y_n), (\alpha, \beta)) = d_1((x, y), (\alpha, \beta))$, and using that K is bounded,

$$\sup_{n\in\mathbb{N}} \sup_{(\alpha,\beta)\in K} d_1((x_n, y_n), (\alpha, \beta)) < \infty.$$

GOPT 919499 8-5-2014 Initial CE:YS QA:EK

By Lebesgue dominated convergence theorem (which can be used since $\mu(K) < \infty$)

$$\lim_{n\to\infty} F_{K,\mu}(x_n, y_n) = \int_K \lim_{n\to\infty} d_1((x_n, y_n), (\alpha, \beta))\, \mu(\mathrm{d}\alpha, \mathrm{d}\beta)$$
$$= \int_K d_1((x, y), (\alpha, \beta))\, \mu(\mathrm{d}\alpha, \mathrm{d}\beta) = F_{K,\mu}(x, y),$$

yielding that $F_{K,\mu}$ is continuous.

Further, using Lemma 3.5 and that almost sure convergence yields convergence in probability, in order to prove L^q-convergence of $(F_{K,\mu}(X_k))_{k\in\mathbb{N}}$, it is enough (and actually necessary) to check that

$$\lim_{m\to\infty} \sup_{k\in\mathbb{N}} \mathbb{E}\left(\|X_k\|^q \mathbf{1}_{\{\|X_k\|^q > m\}}\right) = 0. \tag{3.5}$$

We show that the sequence $(\|X_k\|^q)_{k\in\mathbb{N}}$ is bounded, and then (3.5) readily follows. Let $D := \sup_{k\in\mathbb{N}}\{t_k\} = t_1 > 0$ (indeed, $(t_k)_{k\in\mathbb{N}}$ is a decreasing sequence of positive numbers). Let us consider the rectangle R with vertexes

$$\left(\inf\{x : (x, y) \in K\} - D\sqrt{2},\ \inf\{y : (x, y) \in K\} - D\sqrt{2}\right),$$
$$\left(\inf\{x : (x, y) \in K\} - D\sqrt{2},\ \sup\{y : (x, y) \in K\} + D\sqrt{2}\right),$$
$$\left(\sup\{x : (x, y) \in K\} + D\sqrt{2},\ \inf\{y : (x, y) \in K\} - D\sqrt{2}\right),$$
$$\left(\sup\{x : (x, y) \in K\} + D\sqrt{2},\ \sup\{y : (x, y) \in K\} + D\sqrt{2}\right).$$

Since $\|Q_k\| = \sqrt{2}$, $k \in \mathbb{N}$, if $X_n \in K$ with some $n \in \mathbb{N}$, then $X_{n+1} \in R$, i.e. the recursion (3.1) cannot leave the rectangle R starting from K by one step. Next we check that if $X_n \in R$ with some $n \in \mathbb{N}$, then $X_{n+1} \in R$, which yields that the recursion (3.1) cannot leave the rectangle R. We distinguish eight cases according to the Figure 2.

If X_n is in the rectangle numbered 1, then $Q_{n+1} = (-1, 1)^\top$ and hence, by the choice of D,

$$X_{n+1} = X_n + t_{n+1}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \in R.$$

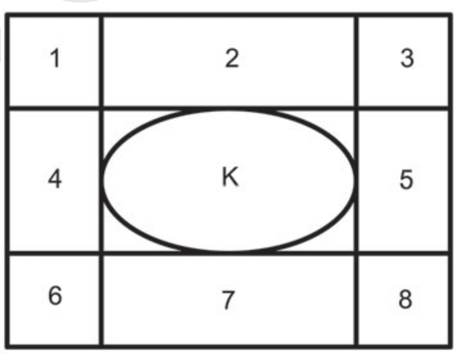

Figure 2. The eight cases.

If X_n is in the rectangle numbered 2, then $Q_{n+1} = (1, 1)^\top$ or $Q_{n+1} = (-1, 1)^\top$ according to the cases $X_n^{(1)} \geq P_{n+1}^{(1)}$ and $X_n^{(1)} < P_{n+1}^{(1)}$, and hence

$$X_{n+1} = X_n + t_{n+1} \begin{pmatrix} -1 \\ -1 \end{pmatrix} \in R \qquad \text{or} \qquad X_{n+1} = X_n + t_{n+1} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in R.$$

If X_n is in the rectangle numbered 3, then $Q_{n+1} = (1, 1)^\top$ and hence

$$X_{n+1} = X_n + t_{n+1} \begin{pmatrix} -1 \\ -1 \end{pmatrix} \in R.$$

The other cases can be handled similarly. □

Acknowledgements

We are grateful for the referee for his/her several valuable comments that have led to an improvement of the manuscript. The research of Mátyás Barczy was realized in the frames of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program – Elaborating and operating an inland student and researcher personal support system'.

Funding

The project was subsidized by the European Union and co-financed by the European Social Fund. Ábris Nagy has been supported, in part, by the Hungarian Academy of Sciences, the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program'. Csaba Vincze was partially supported by the European Union and the European Social Fund through the project Supercomputer, the national virtual lab [grant number TÁMOP-4.2.2.C-11/1/KONV-2012-0010]. Csaba Vincze is supported by the University of Debrecen's internal research project RH/885/2013.

References

[1] Vincze Cs, Nagy Á. On the theory of generalized conics with applications in geometric tomography. J. Approx. Theory. 2012;164:371–390.

[2] Gardner RJ. Geometric tomography. 2nd ed. New York (NY): Cambridge University Press; 2006.

[3] Lukšan L. Conjugate gradient algorithms for conic functions. Aplikace Matematiky. 1986;31:427–440.

[4] Wang G-J, Wang W, Ren F. Conic functions on metric spaces and their application. Appl. Math. Lett. 2005;18:451–454.

[5] Yaglom IM, Boltyanskii VG. Convex figures. Holt, Rinehart and Winston; 1961.

[6] Robbins H, Monro S. A stochastic approximation method. Ann. Math. Stat. 1951;22:400–407.

[7] Robert C, Casella G. Monte Carlo statistical methods. 2nd ed. Springer Verlag; 2004.

[8] Bouleau N, Lépingle D. Numerical methods for stochastic processes. Wiley; 1994.

[9] Arnaudon M, Dombry C, Phan A, Yang L. Stochastic algorithms for computing means of probability measures. Stoch. Process. Appl. 2012;122:1437–1455.

[10] Cohn DL. Measure theory. Boston: Birkhäuser; 1980.

[11] Beer GA. The Hausdorff metric and convergence in measure. Mich. Math. J. 1974;21:63–64.

[12] Polyak BT. Introduction to optimization. New York (NY): Optimization Software Inc., Publications Division; 1987.

[13] Chow YS, Teicher H. Probability theory. 3rd ed. New York (NY): Springer-Verlag; 1997.