UDC 519.178 Вестник СПбГУ. Прикладная математика. Информатика... 2022. Т. 18. Вып. 4
MSC 52A27, 52A41, 90C25

Smooth approximations of nonsmooth convex functions

L. N. Polyakova

St Petersburg State University, 7–9, Universitetskaya nab., St Petersburg,
199034, Russian Federation

For citation: Polyakova L. N. Smooth approximations of nonsmooth convex functions. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2022, vol. 18, iss. 4, pp. 535–547. https://doi.org/10.21638/11701/spbu10.2022.408

For an arbitrary convex function, using the infimal convolution operation, a family of continuously differentiable convex functions approximating it is constructed. The constructed approximating family of smooth convex functions Kuratowski converges to the function under consideration. If the domain of the considered function is compact, then such smooth convex approximations are uniform in the Chebyshev metric. The approximation of a convex set by a family of smooth convex sets is also considered.

Keywords: set-valued mapping, semicontinuous mapping, conjugate function, Kuratowski converge, infimal convolution operation, smooth approximation.

1. Introduction and preliminaries. The concepts of convex sets and convex functions are fundamental in Convex Analysis (see, e. g., [1–3]). The class of convex functions is one of the most studied among the family of nonsmooth functions. Convex functions are known to be nondifferentiable. Convex sets and convex functions are the main tools in theoretical studies in many subjects of nondifferentiable optimization. In the absence of smoothness, the convexity enables us to use a rich set of analytical tools for the development of the theory of optimality conditions.

The aim of this paper is to construct a family of smooth convex functions, which approximates a given convex function and a family of smooth convex sets, which approximates a given convex set. The need for function approximation arises in many branches of applied mathematics, and in particular in computer science (see, e. g., [4]). For constructing such an approximation family, the operation of taking the infimal convolution is used. As it is known from Convex Analysis [3], if one of the convex functions involved in the infimal convolution operation is essentially smooth, then the resulting function is also smooth. The Moreau — Yoside regularization is the most well-studied among the functions obtained as a result of the infimal convolution. The Moreau envelope also smoothes a nonsmooth convex function. However, these functions approximate well the given function in a neighborhood of an optimal point. Based on such regularization, algorithms, called proximal algorithms, are widely used for solving convex optimization problems. A lot of investigations have been done on the properties of the Moreau envelope, including differentiability, regularization (see, e. g., [5–7]).

A new approach for constructing a family of smooth convex functions uniformly approximating a given convex function on a convex compact set is proposed. If the function is finite on the whole Euclidean space, then it is shown that the epigraphs of the resulting family are Kuratowski continuous.

The article is organized as follows. The most important properties of convex functions and set-valued mappings which are applied in proving the main theorems are collected in Section 1. The main results of this paper are presented in Sections 2 and 3. In Section 2, a family of approximation convex sets is constructed for an arbitrary closed set. Using this family, we form a set-valued mapping and prove that any convex set from this family is smooth and the set-valued mapping is Kuratowski continuous. In Section 3, an algorithm for forming a family of smooth convex functions, which approximates a given convex function is presented. The properties of this family are investigated. Some examples are given.

2. The main theorems.

2.1. Notation. In the paper, the standard notation and terminology of Convex Analysis (see, e. g., [1–3]), are used.

Let $f : \mathbb{R}^n \to \mathbb{R} \bigcup \{+\infty\} \bigcup \{-\infty\}$. A set $\mathrm{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$ is called an effective domain of a convex function f. A set

$$\mathrm{epi} f = \{(x, \mu) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leqslant \mu\}$$

is called an epigraph of f.

A convex function f is said to be proper if its epigraph is nonempty and contains no vertical lines, i. e., if $f(x) < +\infty$ for at least one x and $f(x) > -\infty$ for every x. In the future, we will consider only proper convex functions. For proper convex functions, it is possible to give another definition, which equivalent to the above. A function f is called closed, if its epigraph is a closed set. A proper convex function is called essentially smooth, if it satisfies the following three conditions [3]:

- the set $C = \mathrm{int}\ (\mathrm{dom} f)$ is not empty;
- f is differentiable for each $x \in C$;
- if $x_1, x_2, \ldots$ is a sequence in C converging to a boundary point x of C, then $\lim\limits_{i \to +\infty} ||f'(x_i)|| = +\infty$.

Here and further, we will consider only the Euclidean norm $||x|| = \sqrt{\langle x, x \rangle}$. Note that any smooth convex function on $\mathbb{R}^n$ will be essentially smooth, as the set of sequences satisfying the last condition is empty.

The conjugate function of f is

$$f^*(v) = \sup_{x \in \mathbb{R}^n} \{\langle x, v \rangle - f(x)\}, \quad v \in \mathbb{R}^n.$$

Clearly, the equality

$$f^*(v) = \sup_{x \in \mathrm{dom} f} \{\langle x, v \rangle - f(x)\}, \quad v \in \mathbb{R}^n,$$

is true. Note some of the properties of the conjugate functions [3]:

- f^* is closed and convex (even when f is not convex);
- the Fenchel inequality: the definition implies that

$$f(x) + f^*(v) \geqslant \langle x, v \rangle \quad \forall x \in \mathbb{R}^n, \quad \forall v \in \mathbb{R}^n;$$

- if f is a closed proper convex function, then f^* is also a closed proper convex function and the following equality $f(x) = f^{**}(x)$ is true.

2.2. Distance function and set-valued mappings. Let $C(\mathbb{R}^n)$ be the collection of nonempty closed subsets of $\mathbb{R}^n$. Take a set $X \in C(\mathbb{R}^n)$. In our case the distance function $d(\cdot, X) : \mathbb{R}^n \to [0, +\infty)$ is defined by

$$d(z, X) = \min_{x \in X} ||z - x||.$$

Let $\{X_n\}$ be a sequence of closed sets $X_n \in C(\mathbb{R}^n)$ and $X \in C(\mathbb{R}^n)$. We will define $X_n \to X$, if $d(\cdot, X_n) \to d(\cdot, X)$ pointwise [8].

For any sequence of sets $\{X_n\},\ X_n \in C(\mathbb{R}^n)$ and a set $X \in C(\mathbb{R}^n)$ define [9, 10] the Kuratowski limit inferior (or lower closed limit) of $X_n \to X,\ n \to +\infty$, is

$$\operatorname*{Li}_{n\to+\infty} X_n = \left\{ x \in X \;\middle|\; \limsup_{n\to+\infty} d(x, X_n) = 0 \right\} =$$

$$= \left\{ x \in X \;\middle|\; \begin{array}{c} \text{for all open neighbourhoods } U \text{ of } x, \\ U \cap X_n \neq \varnothing \text{ for large enough } n \end{array} \right\},$$

the Kuratowski limit superior (or upper closed limit) of $X_n \to X,\ n \to +\infty$, is

$$\operatorname*{Ls}_{n\to\infty} X_n = \left\{ x \in X \;\middle|\; \liminf_{n\to+\infty} d(x, X_n) = 0 \right\} =$$

$$= \left\{ x \in X \;\middle|\; \begin{array}{c} \text{for all open neighbourhoods } U \text{ of } x, \\ U \cap X_n \neq \varnothing \text{ for infinitely many } n \end{array} \right\}.$$

If

$$\operatorname*{Li}_{n\to\infty} X_n = \operatorname*{Ls}_{n\to\infty} X_n = X,$$

then we say that $\{X_n\}$ Kuratowski converges to X.

Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be some sets. Denote by 2^Y the set of all nonempty subsets of Y. Let $\psi : X \to 2^Y$ be set-valued mapping. A set-valued mapping ψ is called upper semicontinuous at a point $x \in X$, if from

$$x_n \to X, \quad x_n \in X, \quad y_n \to y, \quad y_n \in \psi(x_n),$$

it follows $y \in \psi(x)$. A set-valued mapping ψ is called upper semicontinuous, if it is upper semicontinuous at each point $x \in X$. A set-valued mapping ψ is called lower semicontinuous at a point $x \in X$, if that for any $y \in \psi(x)$ and any sequence $\{x_n\}$, $x_n \to x$, $x_n \in X$, there is such a sequence $\{y_n\}$, $y_n \in \psi(x_n)$, that $y_n \to y$. A set-valued mapping ψ is called lower semicontinuous, if it is lower semicontinuous at each point $x \in X$. If a set-valued mapping ψ is upper semicontinuous and lower semicontinuous at each point $x \in X$, then ψ is Kuratowski continuous. If a set-valued mapping ψ is upper semicontinuous, then for any $x \in X$ the set $\psi(x)$ is closed. A set-valued mapping ψ is called bounded, if it translats bounded sets into bounded sets.

Denote by $\delta(X, Y) = \sup\limits_{x\in X} \inf\limits_{y\in Y} ||x - y||$. The function

$$\rho_H(X, Y) = \sup\{\delta(X, Y), \delta(Y, X)\}$$

is called the Hausdorff distance between the convex sets X and Y. A set-valued mapping ψ is called Hausdorff continuous at a point $x \in X$, if from $x_n \to x,\ x_n \in X$, it follows

$$\rho_H(\psi(x_n), \psi(x)) \to 0.$$

A set-valued mapping ψ is called Hausdorff continuous, if it is Hausdorff continuous at each point $x \in X$. If a set-valued mapping ψ is Hausdorff continuous on X, then it is Kuratowski continuous. If a bounded set-valued mapping ψ is Kuratowski continuous on X, then it is Hausdorff continuous.

2.3. Infimal convolution of two convex functions. Let $f_1, f_2 : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be proper convex functions. A function

$$f(x) = \inf_{\substack{x_1 + x_2 = x \\ x_1, x_2 \in \mathbb{R}^n}} \{f_1(x_1) + f_2(x_2)\} = \inf_{x_1 \in R^n} \{f_1(x_1) + f_2(x - x_1)\}$$

is called the infimal convolution of functions f_1, f_2 and is denoted by $f(x) = (f_1 \oplus f_2)(x)$. It is known, that

- the function f is convex on $\mathbb{R}^n$;
- the operation of taking the infimal convolution is commutative and associative;
- an infimal convolution is known as epigraphical addition. Because geometrically performing the infimal convolution of the function f_1 using the function f_2, we add the epigraph of f_1 to the epigraph of f_2:

$$(f_1 \oplus f_2)(x) = \inf \left\{ \mu \in \mathbb{R} \;\middle|\; (x, \mu) \in \text{epi } f_1 + \text{epi } f_2 \right\}.$$

The infimal convolution $f_1 \oplus f_2$ is called exact at a point $x = x_1 + x_2$, if

$$f_1(x_1) + f_2(x_2) = \min_{\substack{y_1 + y_2 = x \\ y_1, y_2 \in \mathbb{R}^n}} \{f_1(y_1) + f_2(y_2)\}.$$

Note some properties of convex functions obtained as the result of the infimal convolution operation. Let f_1 and f_2 be convex functions on $\mathbb{R}^n$, then

- dom $(f_1 \oplus f_2) =$ dom $f_1 +$ dom f_2;
- the following equality

$$(f_1 \oplus f_2)^* = f_1^* + f_2^* \tag{1}$$

holds

- if ri (dom f_1) $\cap$ ri (dom f_2) $\neq \varnothing$, then $(f_1 + f_2)^* = f_1^* \oplus f_2^*$;
- if ri (dom f_1) $\cap$ ri (dom f_2) $\neq \varnothing$, and f_1 is essentially smooth, then $f_1 \oplus f_2$ is essentially smooth;
- if the functions f_1 and f_2 are not identically equal $+\infty$ and the infimal convolution $f_1 \oplus f_2$ is exact at a point $x = x_1 + x_2$, then

$$\partial(f_1 \oplus f_2)(x) = \partial f_1(x_1) \cap \partial f_2(x_2).$$

Let f_1 be a continuous convex function on $\mathbb{R}^n$ and $f_2(x) = \frac{1}{2}\langle Mx, x\rangle$, where M is a definite positive matrix. The function

$$f(x) = (f_1 \oplus f_2)(x) = \inf_{y \in \mathbb{R}^n} \left\{ f_1(y) + \frac{1}{2}\langle M(x - y), (x - y)\rangle \right\}$$

is called the Moreau—Yosida regularization.

Example 1. Let $X \subset \mathbb{R}^n$ be a convex set, $f_1(x) = \delta(X, x)$ be the indicator function of this set, $f_2(x) = ||x||,\ x \in \mathbb{R}^n$, then

$$f(x) = (f_1 \oplus f_2)(x) = \inf_{x_1 \in X} ||x - x_1||.$$

Example 2. Fix $\varepsilon > 0$. Denote

$$t_\varepsilon(x) = \begin{cases} -\sqrt{\varepsilon^2 - \langle x, x\rangle}, & ||x|| \leqslant \varepsilon, \\ +\infty, & ||x|| > \varepsilon, \end{cases} \qquad x \in \mathbb{R}^n.$$

The function $t_\varepsilon(x)$ is determined only in a ball of radius ε centered at the zero point and it is essentially smooth, i. e., it is differentiable in each internal point $x \in \text{int dom } t_\varepsilon$, and if $x_1, x_2, \ldots$ is a sequence of elements of int dom t_ε, which converges to the point $x \notin \text{int dom } t_\varepsilon$, then $\lim\limits_{i\to+\infty} ||f'(x_i)|| = +\infty$. It is easy to see that

$$t^*_\varepsilon(v) = \varepsilon\sqrt{1 + \langle v, v\rangle}, \quad v \in R^n, \quad \varepsilon > 0.$$

Therefore, the effective domain of the conjugate function t^*_ε is the whole space $\mathbb{R}^n$.

3. Smooth approximation of convex sets. In this section, we will propose a method for constructing a family of smooth convex sets approximating a given set.

Let $K \subset \mathbb{R}^n$ be a cone. A cone $K^* = \{y \in \mathbb{R}^n \mid \langle y, x\rangle \geqslant 0 \quad \forall x \in K\}$ is called a dual cone to K. Let $X \subset \mathbb{R}^n$ be a closed and convex set. A set $N(X, x) = \Big\{y \in \mathbb{R}^n \;\Big|\; \langle y, z - x\rangle \leqslant 0 \quad \forall z \in X \Big\}$ is called a normal cone to the set X at $x \in X$.

Note some properties of normal cones:

- the normal cone is a closed convex cone;
- let $X \subset \mathbb{R}^n$ be a closed convex set. If $x \in X$, then

$$N(X, x) = -[\text{cone } (X - x)]^* = -\Gamma^*(X, x),$$

where $\Gamma(X, x)$ is the cone of feasible directions at the point x. Here cone A denotes a convex conical hull of a set A.

A closed convex set is called smooth, if for each one of its boundary point there is a unique support hyperplane. Thus, if the normal cone at every boundary point of a closed convex set consists of a single ray, then this set is smooth.

Let a set $X \subset \mathbb{R}^n$ be closed and convex and assume that it does not coincide with $\mathbb{R}^n$. Fix $\varepsilon > 0$ and form a closed convex set

$$X(\varepsilon) = X + \varepsilon B_1(0_n),$$

where $B_r(x_0) = \{x \in \mathbb{R}^n \mid ||x - x_0|| \leqslant r\}$.

Theorem 1. *A normal cone to an arbitrary boundary point $z_0 \in \text{bd } X(\varepsilon)$ of the set $X(\varepsilon)$ consists of a single ray.*

P r o o f. Fix $\varepsilon > 0$. Take a boundary point $z_0 \in \text{bd } X(\varepsilon)$ and project it onto the set X, i. e., we find a point x_0 such that

$$x_0 = \arg \min_{x \in X} ||x - z_0||.$$

The point x_0 is unique and $||x_0 - z_0|| = \varepsilon$. Show that

$$N(X(\varepsilon), z_0) = \Big\{ y \in \mathbb{R}^n \;\Big|\; y = \lambda(z_0 - x_0) \quad \forall \lambda \geqslant 0 \Big\}.$$

First, let us prove that $(z_0 - x_0) \in N(X(\varepsilon), z_0)$, i. e.

$$\langle z - z_0, z_0 - x_0\rangle \leqslant 0 \quad \forall z \in X(\varepsilon).$$

Take a point $z \in X(\varepsilon)$. If $z \in X$, then

$$\langle z - z_0, z_0 - x_0\rangle \;\leqslant -||x_0 - z_0||^2 = -\varepsilon^2 < 0. \tag{2}$$

If $z \notin X$, then there exist points $x \in X$, $p \in \mathbb{R}^n$, $||p|| = 1$, and a number $\varepsilon_1 \in (0, \varepsilon]$ such that $z = x + \varepsilon_1 p$. In this case

$$\langle z - z_0, z_0 - x_0\rangle = \langle x + \varepsilon_1 p - z_0, z_0 - x_0\rangle =$$

$$= \langle x - z_0, z_0 - z_0 \rangle + \varepsilon_1 \langle p, z_0 - x_0 \rangle \leqslant -\varepsilon^2 + \varepsilon_1 \varepsilon \leqslant 0. \quad (3)$$

Thus, from (2) and (3) it follows that for each $z \in X(\varepsilon)$ the inequality

$$\langle z - z_0, z_0 - z_0 \rangle \leqslant 0$$

is satisfied. This means that the ray with a direction vector $y_0 = z_0 - x_0$ belongs to the cone $N(X(\varepsilon), z_0)$.

Let us prove its uniqueness. Note that z_0 is a boundary point not only of the set $X(\varepsilon)$, but it is a boundary point of a closed ball $B_\varepsilon(x_0)$ of radius ε centered at x_0. The vector y_0 is also normal to the tangent plane of the ball at the point z_0. Therefore, if we assume the existence of a vector

$$y_1 \in N(X(\varepsilon), z_0), \quad y_1 \neq \lambda y_0 \quad \forall \lambda \geqslant 0,$$

then it should be normal to set $B_\varepsilon(x_0)$. The obtained contradiction completes the proof of the theorem. ■

In Figure 1 you can see an example of a rectangle smooth approximation.

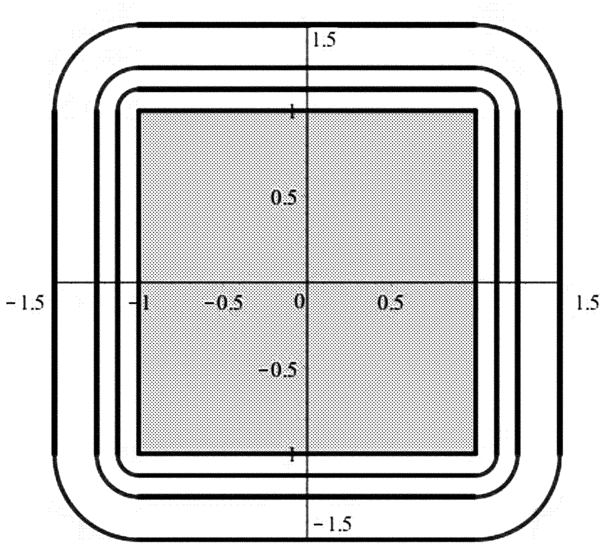

Figure 1. The family X_ε

Corollary 1. Using this theorem, it is not difficult to show the validity of the following statements:

- for points x_0, z_0, from Theorem 1, the next inclusion

$$N(X(\varepsilon), z_0) \subset N(X, x_0)$$

is true;

- let $X \subset \mathbb{R}^n$ be a closed convex set. For every $\varepsilon > 0$, the set $X(\varepsilon)$ is smooth.

Let $X \subset \mathbb{R}^n$ be a closed convex set. Consider a set-valued mapping

$$\psi : X(\cdot) : (0, +\infty) \to 2^{\mathbb{R}^n}.$$

By using the results presented in the paper by G. Beer [9], it is easy to prove the following theorem.

Theorem 2.

• *The set-valued mapping* ψ *is Kuratowski continuous.*

• *Let* $X \subset \mathbb{R}^n$ *be a compact convex set. Then* $\rho_H(X(\varepsilon), X) \to 0$, *if* $\varepsilon \to +0$, *where* $\rho_H(X(\varepsilon), X)$ *is the Hausdorff distance.*

4. Smooth approximation of convex functions. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $f_n : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be real-valued functions. We say that the sequence $\{f_n\}$ epi-converges to a function f if for each $x \in X$:

$$\liminf_{n\to+\infty} f_n(x_n) \geqslant f(x) \text{ for every } x_n \to x,$$

$$\limsup_{n\to+\infty} f_n(x_n) \leqslant f(x) \text{ for some } x_n \to x.$$

A collection Ω of real-valued functions on $\mathbb{R}^n$ is called pointwise equicontinuous [9], if for each $y \in \mathbb{R}^n$ and $\varepsilon > 0$ there exists $\delta > 0$, depending on ε and y, such that whenever $d(x, y) < \delta$ then $|f(x) - f(y)| < \varepsilon$ for all $f \in \Omega$.

The following theorem by G. Beer [9] establishes the relationship between the pointwise convergence of distance functions and the convergence of distance functions of sets in $\mathbb{R}^n \times \mathbb{R}$.

Theorem 3 [9]. *Let* $\{f_n\}$ *be a pointwise equicontinuous sequence of real-valued continuous functions on* $\mathbb{R}^n$, *and let* $f : \mathbb{R}^n \to \mathbb{R}$ *be continuous. The following statements are equivalent:*

• *whenever* $\{x_n\}$ *is a sequence in* $\mathbb{R}^n$ *convergent to* x, *then*

$$\lim_{n\to+\infty} f_n(x_n) = f(x);$$

• $\{f_n\}$ *converges to* f *uniformly on compact subsets of* $\mathbb{R}^n$;

• $\{f_n\}$ *converges pointwise to* f ;

• $\{f_n\}$ *Kuratowski converges to* f ;

• $\{f_n\}$ *epi-converges to* f.

Consider a convex function $f : \mathbb{R}^n \to \mathbb{R}$ and a closed convex set $D \subset \mathbb{R}^n$. Denote

$$X = \mathrm{epi} f = \left\{(x, \mu) \in \mathbb{R}^n \times \mathbb{R} \;\middle|\; \mu \geqslant f(x), \quad x \in D\right\}.$$

Construct families of convex closed sets

$$X(\varepsilon) = X + \varepsilon B_1(0_{n+1}) \subset \mathbb{R}^{n+1}, \quad D(\varepsilon) = D + \varepsilon B_1(0_n) \subset \mathbb{R}^n, \quad \varepsilon > 0,$$

and a family of convex functions

$$f_\varepsilon(x) = \begin{cases} \inf \mu, & (x, \mu) \in X(\varepsilon), \\ +\infty, & \text{at other points.} \end{cases}$$

It is not difficult to see that dom $f_\varepsilon = D(\varepsilon)$, and for each fixed $\varepsilon > 0$, the graph of the function f_ε is the lower envelope of the corresponding set $X(\varepsilon)$.

Fix $\varepsilon > 0$. Let $z \in D$. Consider a family of convex functions $\{\varphi_\varepsilon(x, z)\}$

$$\varphi_\varepsilon(x, z) = f(z) + t_\varepsilon(x, z),$$

where

$$t_\varepsilon(x,z)=\begin{cases} -\sqrt{\varepsilon^2-||x-z||^2}, & x\in B_\varepsilon(z),\\ +\infty, & \text{at other points.}\end{cases}$$

Here $B_\varepsilon(z)\subset D_\varepsilon$. It is obvious that

$$\operatorname{dom}\,\varphi_\varepsilon(\cdot,z)=B_\varepsilon(z),\quad \bigcup_{z\in D}B_\varepsilon(z)=D(\varepsilon).$$

Denote $H_\varepsilon(z)=\operatorname{epi}\,\varphi_\varepsilon(\cdot,z)$. Consider also functions

$$\varphi_\varepsilon(x)=\inf_{z\in D}\varphi_\varepsilon(x,z)$$

and its epigraphs $H_\varepsilon=\operatorname{epi}\,\varphi_\varepsilon$.

From the constructing of functions f_ε and φ_ε it is not difficult to prove that the statements are true [11]:

- for any fixed point x_0, there exists a unique point $z_0\in D$ for which

$$\varphi_\varepsilon(x_0)=f(z_0)+t_\varepsilon(x_0,z_0);$$

- $H_\varepsilon=\operatorname{epi}\,\varphi_\varepsilon=\bigcup\limits_{z\in D}\operatorname{epi}\,\varphi_\varepsilon(\cdot,z)=\bigcup\limits_{z\in D}H_\varepsilon(z)$;
- $H_\varepsilon=X(\varepsilon)$;
- for any fixed $\varepsilon>0$, the following statement $f_\varepsilon(x)=\varphi_\varepsilon(x)$ holds;
- $f_\varepsilon(x)=(f\oplus t_\varepsilon)(x)$, where

$$t_\varepsilon(x)=\begin{cases} -\sqrt{\varepsilon^2-||x||^2}, & ||x||\leqslant\varepsilon,\\ +\infty, & \text{at other points.}\end{cases}$$

Note the fact that t_ε is essentially smooth for every fixed positive ε. Consider the function $f_\varepsilon(x)=(f\oplus t_\varepsilon)(x)$. The function f_ε is convex and

$$f^*_\varepsilon(v)=f^*(v)+t^*_\varepsilon(v),\quad v\in\mathbb{R}^n.$$

Then the next statements are true [11]:

• for the function f_ε, the statements

$$\operatorname{dom}\,f_\varepsilon=\operatorname{dom}\,f_1+B_\varepsilon(0_n),\quad \operatorname{epi}\,f_\varepsilon=\operatorname{epi}\,f_1+B_\varepsilon(0_{n+1})$$

hold, where $B_\varepsilon(0_n)=\{x\in\mathbb{R}^n\mid ||x||\leqslant\varepsilon\}$, $B_\varepsilon(0_{n+1})=\{x\in\mathbb{R}^{n+1}\mid ||x||\leqslant\varepsilon\}$;

• the function f_ε for any fixed $\varepsilon>0$ is continuously differentiable at each interior point of $D(\varepsilon)$;

• the set $\operatorname{epi}\,f_\varepsilon\subset\mathbb{R}^n\times\mathbb{R}$ is smooth for any positive number ε.

As the function t_ε is essentially smooth, then the function f_ε is also essentially smooth [3]. Therefore it is continuously differentiable at any interior point of D_ε.

Theorem 4. *Let a point $x_0\in\operatorname{int}\,D(\varepsilon)$. Then there exists a unique point $z_0\in D$ for which*

$$f'_\varepsilon(x_0)\in\partial f(z_0),$$

where $f'_\varepsilon(x_0)$ is the gradient of the function $f_\varepsilon(x_0)$ at x_0, $\partial f(z_0)$ is the subdifferential of the function f at z_0.

P r o o f. Take a point $x_0 \in \mathrm{int} D_\varepsilon$. Then by using Theorem 1 for any point $\bar{x}_0 = (x_0, f_\varepsilon(x_0))$, the normal cone $N(X(\varepsilon), \bar{x}_0)$ to the set $X(\varepsilon)$ consists of the ray with the direction vector

$$y_0 = \bar{x}_0 - \bar{z}_0 = (x_0 - z_0, f_\varepsilon(x_0) - f(z_0)),$$

where $\bar{z}_0 = \arg\ \min_{\tilde{z} \in X} ||\tilde{z} - \bar{x}_0|| = [z_0, f(x_0)]$, and $f_\varepsilon(x_0) - f(z_0) < 0$, that is,

$$N(X(\varepsilon), \bar{x}_0) = \{y \in \mathbb{R}^{n+1} \bigm| y = \lambda(\bar{x}_0 - \bar{z}_0) \quad \forall \lambda \geqslant 0\}.$$

As the set X is the epigraph of f, then by using one of the properties of the normal cone to the epigraph of f at $\bar{z}_0$, we have

$$(f'_\varepsilon(x_0), -1) \in N(X(\varepsilon), \bar{x}_0) \subset N(X, \bar{z}_0).$$

Thus $f'_\varepsilon(x_0) \in \partial f(z_0)$. ■

Note some properties of functions conjugate to the functions f and f_ε. Let f be a closed proper convex function on $\mathbb{R}^n$. A set

$$\mathrm{dom}\, \partial f = \{x \in \mathbb{R}^n \bigm| \partial f(x) \neq \varnothing\}$$

and

$$\mathrm{range}\, \partial f = \bigcup_{x \in \mathbb{R}^n} \partial f(x)$$

are called, respectively, the effective set and the image of ∂f. It is known [3], that

$$\mathrm{ri}(\mathrm{dom}\ f^*) \subset \mathrm{range}\ \partial f \subset \mathrm{dom}\ f^*.$$

Since the function f_ε is the infimal convolution of the functions f and t_ε, then by using property (1) we have that at each point $v \in \mathrm{range}\ \partial f_\varepsilon$ for every positive $\varepsilon > 0$, the next equality

$$f^*_\varepsilon(x) = f^*(v) + \varepsilon\sqrt{1 + ||v||^2}$$

holds.

Take $v \in \mathrm{range}\ \partial f_\varepsilon \subset \mathrm{dom} f^*_\varepsilon$. Then there exists a point $x \in \mathrm{dom} f_\varepsilon$ for which $v \in \partial f_\varepsilon(x)$, therefore,

$$f_\varepsilon(x) + f^*_\varepsilon(v) = \langle x, v \rangle. \tag{4}$$

Consider the point $\bar{x} = (x, f_\varepsilon(x))$. Find

$$\bar{z} = \arg \min_{\tilde{z} \in X} ||\tilde{z} - \bar{x}|| = (z, f(z)),$$

then $v \in \partial f(z),\ \bar{x} = \bar{z} + \varepsilon\mu(v)[v, -1]$, where $\mu(v) = \frac{1}{\sqrt{1+||v||^2}}$. From the equalities

$$x = z + \varepsilon\mu(v)v, \quad f_\varepsilon(x) = f(z) - \varepsilon\mu(v) \tag{5}$$

hold. Thus, if a point $x \in \mathrm{int}\ (\mathrm{dom}\ f)$, then the function f_ε is differentiable at it. Therefore

$$v = f'_\varepsilon(x), \quad \mu(v) = \frac{1}{\sqrt{1 + ||f'_\varepsilon(x)||^2}}, \quad x = z + \varepsilon\mu(v) f'_\varepsilon(x).$$

Since $v \in \partial f(z)$, then

$$f(z) + f^*(v) = \langle z, v \rangle.$$

From this equality, from equalities (4) and (5) the next formula is true

$$\min_{x\in D} f(x) = \min_{x\in D(\varepsilon)} f_\varepsilon(x) + \varepsilon.$$

Theorem 5. *Let M^* be the set of minimizers of the function f on the set D, and M^*_ε be the set of minimizers of the function f_ε on the set $D(\varepsilon)$. The case when these sets are empty is not excluded. The following statements are true:*

- if the set D is convex compact, then

$$\min_{x\in D} f(x) = \min_{x\in D_\varepsilon} f_\varepsilon(x) + \varepsilon;$$

- the next equality $M^* = M^*_\varepsilon$ holds;
- if M is not an empty set, then

$$f_\varepsilon(z^*) = f(z^*) - \varepsilon \quad \forall z^* \in M^*.$$

P r o o f. First, note that if a point $x_0 \notin D$, but $x_0 \in D(\varepsilon)$, then there exists a point $z_0 \in D$ for which

$$f_\varepsilon(x_0) = f(z_0) + t_\varepsilon(x_0, z_0) > f(z_0) - \varepsilon \geqslant f_\varepsilon(z_0).$$

Therefore $M^*_\varepsilon \subset D \subset \text{ int } D(\varepsilon)$.

Assume that the set M^*_ε is not empty and a point $z^* \in M^*_\varepsilon$. Show that this set belongs to the set M^*. Consider a point $\bar z = (z^*, f_\varepsilon(z^*)) \in X(\varepsilon)$. Then there exists a point $\bar x \in X,\ \bar x = (x, f(x))$, for which

$$(z^* - x, f_\varepsilon(z^*) - f(x)) \in N(X(\varepsilon), \bar z).$$

If a point z^* is a minimizer of f_ε on D_ε, then

$$(z^* - x, f_\varepsilon(z^*) - f(x)) = \varepsilon(0_n, -1) \subset N(X(\varepsilon), \bar z^*),$$

where $\bar z^* = [z^*, f_\varepsilon(z^*)]$. Therefore $z^* = x$ and $f_\varepsilon(x) - f(x) = -\varepsilon$. Hence $z^* \in M^*$. The inclusion of $M^*_\varepsilon \subset M$ is proved.

Show the correctness of the inverse inclusion. Let $z^* \in M^*$. Consider points $\bar z = (z^*, f(z^*))$, $\tilde z = (z^*, f_\varepsilon(z^*))$ and the vector $\bar g = \tilde z - \bar z = (0_n, f_\varepsilon(z^*) - f(z^*))$. By constructing the set X, we have $||\bar g|| \geqslant \varepsilon$ and

$$f(z^*) - f_\varepsilon(z^*) \geqslant \varepsilon.$$

Suppose that $f(z^*) - f_\varepsilon(z^*) > \varepsilon$. Then there exists a point $\bar x = (x, f(x))$, $x \in D$, for which $||\bar x - \tilde z|| = \varepsilon$. Hence $|f(x) - f_\varepsilon(z^*)| \leqslant \varepsilon$. From here we have

$$\varepsilon < f(z^*) - f_\varepsilon(z^*) \leqslant f(z^*) + \varepsilon - f(x).$$

Or $f(z^*) > f(x)$. However, this inequality contradicts with the fact that z^* is a minimizer of the function f on D. ■

Example 3. Let we have

$$f(x) = \max\left\{-2x - 6, -\frac{1}{2}x - 3, 2x - 8\right\}, \quad x \in \mathbb{R},$$

or

$$f(x)=\begin{cases} -2x-6, & x\in(-\infty,-2),\\ -\frac{1}{2}x-3, & x\in[-2,2),\\ 2x-8, & x\in[2,+\infty). \end{cases}$$

Consider two variants.

Variant 1. Let the set D be the Euclidean space $\mathbb{R}$. Then the set of minimizers of this function consists of a single point $x^*=2$ and $f(2)=-4$. Fix an arbitrary positive $\varepsilon>0$. Then

$$f_\varepsilon(x)=\begin{cases} -2x-6-\sqrt{5}\varepsilon, & x\in\left(-\infty,-2-\frac{2\sqrt{5}\varepsilon}{5}\right),\\ -2-\sqrt{\varepsilon^2-(x+2)^2}, & x\in\left[-2-\frac{2\sqrt{5}\varepsilon}{5},-2-\frac{\sqrt{5}\varepsilon}{5}\right),\\ -\frac{1}{2}x-3-\frac{\sqrt{5}\varepsilon}{2}, & x\in\left[-2-\frac{\sqrt{5}\varepsilon}{5},2-\frac{\sqrt{5}\varepsilon}{5}\right),\\ -4-\sqrt{\varepsilon^2-(x-2)^2}, & x\in\left[2-\frac{\sqrt{5}\varepsilon}{5},2+\frac{2\sqrt{5}\varepsilon}{5}\right),\\ 2x-8-\sqrt{5}\varepsilon, & x\in\left[2+\frac{2\sqrt{5}\varepsilon}{5},+\infty\right). \end{cases}$$

The function f_ε is continuously differentiable on $\mathbb{R}$ and

$$f'_\varepsilon(x)=\begin{cases} -2, & x\in\left(-\infty,-2-\frac{2\sqrt{5}\varepsilon}{5}\right),\\ \frac{x+2}{\sqrt{\varepsilon^2-(x+2)^2}}, & x\in\left[-2-\frac{2\sqrt{5}\varepsilon}{5},-2-\frac{\sqrt{5}\varepsilon}{5}\right),\\ -\frac{1}{2}, & x\in\left[-2-\frac{\sqrt{5}\varepsilon}{5},2-\frac{\sqrt{5}\varepsilon}{5}\right),\\ \frac{x-2}{\sqrt{\varepsilon^2-(x-2)^2}}, & x\in\left[2-\frac{\sqrt{5}\varepsilon}{5},2+\frac{2\sqrt{5}\varepsilon}{5}\right),\\ 2, & x\in\left[2+\frac{2\sqrt{5}\varepsilon}{5},+\infty\right). \end{cases}$$

Hence $f'_\varepsilon(2)=0$ and $f_\varepsilon(2)=-4-\varepsilon$. We have

$$f^*(v)=\begin{cases} -2v+2, & v\in\left[-2,-\frac{1}{2}\right),\\ 2v+4, & v\in\left[-\frac{1}{2},2\right],\\ +\infty, & \text{at other points.} \end{cases}$$

Variant 2. Consider the case when the set D is the segment $[-3,0]$ and the functions

$$\tilde{f}(x)=\max\left\{-2x-6,-\frac{1}{2}x-3\right\},\quad x\in[-3,1]\subset\mathbb{R},$$

$$\tilde{f}(x)=\begin{cases} +\infty, & x\in(-\infty,-2),\\ -2x-6, & x\in[-2,-\frac{1}{2}),\\ -\frac{1}{2}x-3, & x\in[-\frac{1}{2},0],\\ +\infty, & x\in(-\frac{1}{2},+\infty). \end{cases}$$

Then $D_\varepsilon=[-3-\varepsilon,\varepsilon]$ (Figure 2) and the function

$$\tilde{f}_\varepsilon(x) = \begin{cases} +\infty, & x \in (-\infty, -3-\varepsilon), \\ -\sqrt{\varepsilon^2 - (x+3)^2}, & x \in \left[-3-\varepsilon, -3-\frac{2\sqrt{5}\varepsilon}{5}\right), \\ -2x - 6 - \sqrt{5}\varepsilon, & x \in \left[-3-\frac{2\sqrt{5}\varepsilon}{5}, -2-\frac{2\sqrt{5}\varepsilon}{5}\right), \\ -2 - \sqrt{\varepsilon^2 - (x+2)^2}, & x \in \left[-2-\frac{2\sqrt{5}\varepsilon}{5}, -2-\frac{\sqrt{5}\varepsilon}{5}\right), \\ -\frac{1}{2}x - 3 - \frac{\sqrt{5}\varepsilon}{2}, & x \in \left[-2-\frac{\sqrt{5}\varepsilon}{5}, -\frac{\sqrt{5}\varepsilon}{4}\right), \\ -3 - \sqrt{\varepsilon^2 - x^2}, & x \in \left[-\frac{\sqrt{5}\varepsilon}{4}, \varepsilon\right], \\ +\infty, & x \in (\varepsilon, +\infty). \end{cases}$$

The function $\tilde{f}_\varepsilon$ is continuously differentiable for all $x \in (-3-\varepsilon, \varepsilon)$ and $\tilde{f}'_\varepsilon(0) = 0$. As

$$\tilde{f}_\varepsilon(0) = -3-\varepsilon, \quad \tilde{f}_\varepsilon(-3-\varepsilon) = 0, \quad f_\varepsilon(\varepsilon) = -3,$$

then

$$\min_{x \in D_\varepsilon} f_\varepsilon(x) = -3 - \varepsilon.$$

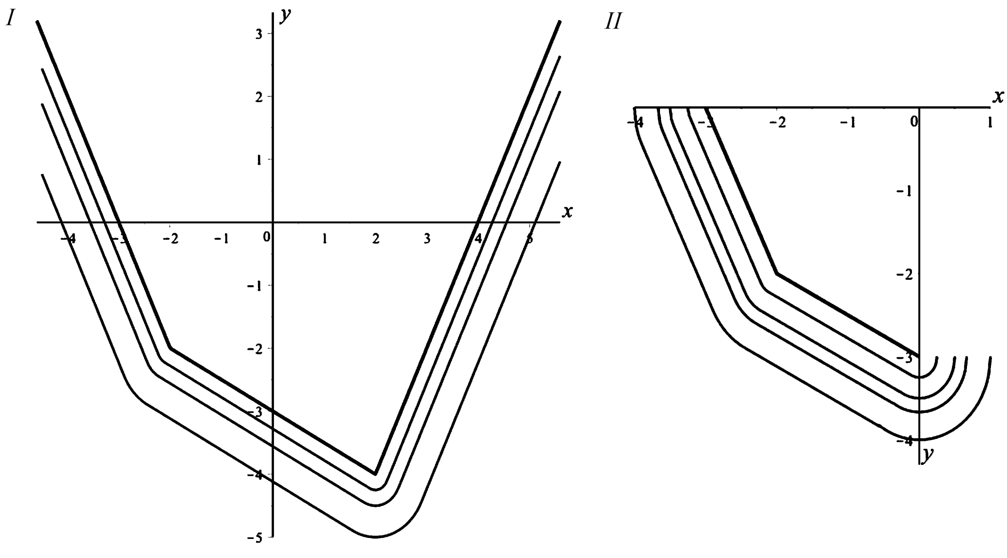

Figure 2. Family $f_\varepsilon(x)$ (I) and $\tilde{f}_\varepsilon(x)$ (II)

Acknowledgements. The author is grateful to the MATRIX research institute for organising the program in algebraic geometry, approximation and optimisation, which provided a fertile research environment that helped this discovery.

References

1. Bertsekas P. *Convex optimization theory.* Belmont, MA, Athena Scientific Publ., 2009, 444 p.
2. Leichtweiß K. *Konvexe mengen.* Berlin, Springer-Verlag Publ., 1980, 330 p.
3. Rockafellar R. T. *Convex analysis.* Princeton, Princeton University Press, 1970, 472 p.

4. Zillober C. Convex approximation methods for practical optimization. *Operations Research Proceedings.* Eds. B. Fleischmann, R. Lasch, U. Derigs, W. Domschke, U. Rieder. Berlin, Heidelberg, Springer Publ., 2001, vol. 2000, pp. 20–25. https://doi.org/10.1007/978-3-642-56656-1_4

5. Moreau J.-J. Propriétés des applications "prox". *C. R. Acad. Sci.* (Paris), 1963, vol. 256, pp. 1069–1071.

6. Moreau J.-J. Proximité et dualité dans un espace Hilbertien. *Bull. Soc. Math. France*, 1965, vol. 93, pp. 273–299.

7. Planiden C., Wang X. Strongly convex functions. Moreau envelopes and the generic nature of convex functions with strong minimizers. *SIAM J. Optim.*, 2016, vol. 26 (2), pp. 1341–1364. hhttps://doi.org/10.1137/15M1035550

8. Wijsman R. A. Convergence of sequences of convex sets, cones, and functions. II. *Trans. Amer. Math. Soc.*, 1966, vol. 123, pp. 32–45.

9. Beer G. On convergence of closed sets in a metric space and distance functions. *Bull. Australian Math. Soc.*, 2009, vol. 31, pp. 421–432. https://doi.org/10.1017/S0004972700009370

10. Kuratowski K. *Topologie I et II.* Warszawa, Panstwowe Wydawnictwa Naukowe Publ., 1961, 528 p. (PWN Monogr. Mat.)

11. Polyakova L. N. O gladkoi approksimatsii vypuklykh funktsii [On a smooth approximation of convex funchions]. *VIII Moscow Internation Conference on Operations Research* (*ORM2016*). Moscow, MAKS Press, 2016, pp. 58–61. (In Russian)

Received: July 21, 2022.
Accepted: September 1, 2022.

Author's information:

Lyudmila N. Polyakova — Dr. Sci. in Physics and Mathematics, Professor; l.polyakova@spbu.ru

Гладкие аппроксимации негладких выпуклых функций

Л. Н. Полякова

Санкт-Петербургский государственный университет, Российская Федерация,
199034, Санкт-Петербург, Университетская наб., 7–9

Для цитирования: *Polyakova L. N.* Smooth approximations of nonsmooth convex functions // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2022. Т. 18. Вып. 4. С. 535–547.
https://doi.org/10.21638/11701/spbu10.2022.408

Используя операцию инфимальной конволюции, для произвольной негладкой выпуклой функции строится аппроксимирующее семейство непрерывно дифференцируемых выпуклых функций. Построенное аппроксимирующее семейство гладких выпуклых функций сходится по Куратовскому к рассматриваемой функции. Если множество определения данной функции компактно, то такие гладкие выпуклые приближения непрерывны в метрике Чебышева. Также рассматривается аппроксимация негладкого выпуклого множества семейством гладких выпуклых множеств.

Ключевые слова: многозначное отображение, полунепрерывное отображение, сопряженная функция, сходимость по Куратовскому, операция инфимальной конволюции, гладкая аппроксимация.

Контактная информация:

Полякова Людмила Николаевна — д-р физ.-мат. наук, проф.; l.polyakova@spbu.ru