Journal of Convex Analysis
Volume 32 (2025), No. 1, 107–118

Lipschitz Differentiability of the Supporting Function from a Geometric Viewpoint, the Nearest and the Farthest Points

Maxim V. Balashov, Kamil Z. Biglov
V. A. Trapeznikov Institute of Control Sciences, Moscow, Russia
balashov73@mail.ru, biglov.kz@phystech.su

Dedicated to Ralph Tyrrell Rockafellar on the occasion of his 90th birthday.

Received: July 25, 2024
Accepted: October 14, 2024

We consider a supporting condition for a convex compact set in $\mathbb{R}^n$. This condition is typical and it is closely related to the local Lipschitz differentiability of the supporting function. With the help of the considered condition we prove a local Lipschitz condition for supporting elements and for nearest and farthest points of a convex compact set. Some applications for numerical algorithms are discussed.

Keywords: Supporting function, subdifferential, supporting condition, nonsmooth analysis.

2010 Mathematics Subject Classification: Primary: 49J52, 49J53, 52A20; secondary: 47H09.

1. Introduction and main notations

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function. Let (p, q) be the inner product of vectors $p, q \in \mathbb{R}^n$ and $\|p\| = \sqrt{(p,p)}$. The *subdifferential* $\partial f(p)$ at a point $p \in \mathbb{R}^n$ is the set of all *subgradients* $x \in \mathbb{R}^n$, i.e.

$$\partial f(p) = \{x \in \mathbb{R}^n \ : \ f(q) \geq f(p) + (x, q-p), \quad \forall q \in \mathbb{R}^n\}.$$

If $\partial f(p)$ is a singleton, then $\partial f(p) = \{f'(p)\}$, where $f'(p)$ is the Frechet gradient of the function f at the point p. The main creators of the subdifferential theory are R. T. Rockafellar and J. J. Moreau. The theory of subdifferential calculus can be found in [23], see also [24].

We want to focus our attention on a very special case of the supporting function $f(p) = s(p, \mathcal{M}) = \sup_{x\in\mathcal{M}}(p, x)$ for a convex compact set $\mathcal{M} \subset \mathbb{R}^n$. The function $s(p, \mathcal{M})$ is homogeneous, convex and Lipschitz continuous with the Lipschitz constant $L = \max\limits_{x\in\mathcal{M}} \|x\|$.

It is well known that $\partial s(p, \mathcal{M}) = \{x \in \mathcal{M} : (p, x) = s(p, \mathcal{M})\}$. From the differentiability almost everywhere of a convex function there is a set $\mathcal{S} \subset \{p\in\mathbb{R}^n \ : \ \|p\| = 1\}$ of full measure such that $\partial s(p, \mathcal{M}) = \{x_p\}$ for all $p \in \mathcal{S}$. We denote such an element x_p as $\mathcal{M}(p)$. If the set $\partial s(p, \mathcal{M})$ consists of more than one point then we also denote its arbitrary element by $\mathcal{M}(p)$. We shall call a point $\mathcal{M}(p)$ a *supporting element*.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

The behavior of supporting elements of a compact convex set $\mathcal{M} \subset \mathbb{R}^n$ is very important in different problems, both in theory and applications.

Denote by $\mathcal{B}_r(x)$ the closed Euclidean ball with center $x \in \mathbb{R}^n$ and radius $r > 0$. In [11], the authors proved the following proposition.

Proposition 1.1. *For a compact convex set $\mathcal{M} \subset \mathbb{R}^n$ and a number $R > 0$ the following three points are equivalent:*

(1) $\mathcal{M} = \bigcap_{x\in\mathcal{X}} \mathcal{B}_R(x)$ *for some set of centers* $\mathcal{X} \subset \mathbb{R}^n$,

(2) $\mathcal{M} \subset \mathcal{B}_R(\mathcal{M}(p) - Rp)$ *for all* $\|p\| = 1$,

(3) $\|\mathcal{M}(p) - \mathcal{M}(q)\| \leq R\|p - q\|$ *for all* $\|p\| = \|q\| = 1$. □

A set of the form $\mathcal{M} = \bigcap_{x\in\mathcal{X}} \mathcal{B}_R(x)$ is called R-strongly convex. The equivalence of (1) and (3) shows that the class of R-strongly convex sets and the class of convex compact sets $\mathcal{M}$ with R-Lipschitz continuous gradient of the function $s(p, \mathcal{M})$ on the unit sphere coincide.

The work [11] and subsequent works [12, 21, 26] show that geometric approach, e. g. items (1), (2) of Proposition 1.1, is productive when studying R-strongly convex sets. Before [11], similar settings were also considered in works [15, 20]. An equivalent characterization of R-convex sets was also considered in earlier works [16, 17].

There are a lot of applications of R-strongly convex sets in different algorithms [2, 14, 25]. There is also an important characterization of strong convexity in terms of nearest and farthest points of a set.

For a compact set $\mathcal{M} \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$ we define the *distance* $\varrho_{\mathcal{M}}(x) = \inf_{y\in\mathcal{M}} \|x - y\|$ and the *antidistance* $f_{\mathcal{M}}(x) = \sup_{y\in\mathcal{M}} \|x - y\|$. The distance function and its properties are well known, while the antidistance function is not so familiar to the general public [1, 18].

Denote by $P_{\mathcal{M}}x = \{y \in \mathcal{M} : \|x - y\| = \varrho_{\mathcal{M}}(x)\}$ the metric projection of a point x onto a set $\mathcal{M}$ and denote by $F_{\mathcal{M}}x = \{y \in \mathcal{M} : \|x - y\| = f_{\mathcal{M}}(x)\}$ the metric antiprojection.

Proposition 1.2. ([6, 7]) *For a compact convex set $\mathcal{M} \subset \mathbb{R}^n$ and a number $R > 0$ the next points are equivalent:*

(1) $\mathcal{M} = \bigcap_{x\in\mathcal{X}} \mathcal{B}_R(x)$ *for some set of centers* $\mathcal{X} \subset \mathbb{R}^n$,

(2) *For any $r > 0$, points $x_0, x_1 \in \mathbb{R}^n$ with $\varrho_{\mathcal{M}}(x_i) \geq r$ and projections $a_i = P_{\mathcal{M}}x_i$, $i = 0, 1$, the inequality $\|a_0 - a_1\| \leq \frac{R}{R+r}\|x_0 - x_1\|$ is satisfied.*

Moreover, if the radius R is minimal then the Lipschitz constant $\frac{R}{R+r}$ is also minimal.

Proposition 1.3. ([1]) *For a compact convex set $\mathcal{M} \subset \mathbb{R}^n$ and a number $R > 0$ the next points are equivalent:*

(1) $\mathcal{M} = \bigcap\limits_{x\in\mathcal{X}} \mathcal{B}_R(x)$ *for some set of centers* $\mathcal{X} \subset \mathbb{R}^n$,

(2) *For any $r > R$, points $x_0, x_1 \in \mathbb{R}^n$ with $f_{\mathcal{M}}(x_i) \geq r$ and antiprojections $a_i = F_{\mathcal{M}}x_i$, $i = 0, 1$, the inequality $\|a_0 - a_1\| \leq \frac{R}{r-R}\|x_0 - x_1\|$ holds.*

If the radius R is minimal then the Lipschitz constant $\frac{R}{r-R}$ is also minimal.

We also want to mention a survey [12], where different properties of R-strongly convex sets and metric projections/antiprojections onto such sets are discussed.

The properties mentioned above, in particular Proposition 1.2, have found different applications, e. g. in gradient methods of minimization.

An important technical tool for proving Propositions 1.2 and 1.3 is the supporting principle (2) from Proposition 1.1. It is this property that allows us to prove the Lipschitz behavior of the metric projection or antiprojection. We want to note that we often don't need a global Lipschitz condition in the spirit of item (3) from Proposition 1.1. For example, when we prove the convergence of minimization algorithms (e.g. the gradient projection algorithm or the method of conditional gradient (Frank-Wolfe method)) we often estimate the distance $\|x_0 - x_k\|$ between the solution x_0 and a current point x_k. Thus, one point x_0 is fixed and another point x_k can "travel" through the set $\mathcal{M}$ during calculations. In the same way we can fix p in item 3) of Proposition 1.1 and change only vector q. Note that in this case, we no longer need strong convexity of the set $\mathcal{M}$. This and similar observations lead us to the following definitions.

Definition 1.4. ([3]) Let $\mathcal{M} \subset \mathbb{R}^n$ be a (convex) compact set, $R > 0$ and $p \in \mathbb{R}^n$ be a unit vector. We shall say that the *supporting condition* $sc(p, R)$ holds for the set $\mathcal{M}$ and write $\mathcal{M} \in sc(p, R)$

$$\text{if} \;\; \partial\, s(p, \mathcal{M}) = \{\mathcal{M}(p)\} \;\; \text{and} \;\; \mathcal{M} \subset \mathcal{B}_R(\mathcal{M}(p) - Rp). \tag{1}$$

Definition 1.5. ([4]) Let $\mathcal{M} \subset \mathbb{R}^n$ be a (convex) compact set, $L > 0$ and $p \in \mathbb{R}^n$ be a unit vector. We shall say that the *Lipschitz condition* $lip(p, L)$ holds for the set $\mathcal{M}$ and write $\mathcal{M} \in lip(p, L)$

$$\text{if} \;\; \partial\, s(p, \mathcal{M}) = \{\mathcal{M}(p)\} \;\; \text{and} \;\; \|\mathcal{M}(p) - \mathcal{M}(q)\| \le L\|p - q\|. \tag{2}$$

Here $\|q\| = 1$ is any unit vector and $\mathcal{M}(q) \in \partial\, s(q, \mathcal{M})$ is any element.

Below for simplicity we will assume that in Definitions 1.4 and 1.5 the set $\mathcal{M}$ is convex, but it is not necessary. We want to note that Definition 1.4 is ideologically very close to the quadratic growth condition for a function. The quadratic growth condition allows one to prove the convergence of some gradient minimization methods in convex optimization [19].

In [4], the implications $lip(p, L) \Rightarrow sc(p, L)$ and $sc(p, R) \Rightarrow lip(p, 2R)$ were proven.

We note that the second implication can be improved by letting q in (2) have arbitrary norm. Indeed, Definition 1.4 means that for any point $x \in \mathcal{M}$ we have $\|\mathcal{M}(p) - Rp - x\| \le R$ or, equivalently, $\frac{1}{2R}\|\mathcal{M}(p) - x\|^2 \le (p, \mathcal{M}(p) - x)$. Suppose that $q \in \mathbb{R}^n$, $x \in \partial\, s(q, M)$, i.e. $x = \mathcal{M}(q)$. Then by the definition of $\partial\, s(p, \mathcal{M})$ we have $(q, \mathcal{M}(q) - \mathcal{M}(p)) \ge 0$ and adding this inequality with the last inequality we get

$$\frac{1}{2R}\|\mathcal{M}(p) - \mathcal{M}(q)\|^2 \le (p - q, \mathcal{M}(p) - \mathcal{M}(q)) \le \|p - q\| \cdot \|\mathcal{M}(p) - \mathcal{M}(q)\|.$$

Thus $\|\mathcal{M}(p) - \mathcal{M}(q)\| \le 2R\|p - q\|$ for arbitrary q.

We want to point out that the constant $2R$ is minimal, see Figure 1. Fix $R > 0$. In $\mathbb{R}^2$, put $p = (0, 1) = e_2$, $q = (\sin\varphi, \cos\varphi)$ for small $\varphi > 0$. Now we define the set $\mathcal{M} = \mathcal{B}_R(0) \cap \{x \in \mathbb{R}^2 : (q, x) \leq R(q, e_2)\}$, i.e. a 2D ball without a circular segment. From elementary planimetry we see that the length of the base of the removed segment is $2R \sin\varphi$. Moreover, $\mathcal{M}(p) = Re_2$. Add to the set $\mathcal{M}$ a point z which is contained in the removed segment strictly higher than the line $\{x \in \mathbb{R}^2 : (q, x) = R(q, e_2)\}$ and close to the point $R(\sin 2\varphi, \cos 2\varphi)$, namely the distance between these points is of the order φ^2. Denote the convex hull of the set $\mathcal{M}$ and the point z by $\widetilde{\mathcal{M}}$. By construction $\widetilde{\mathcal{M}}(q)$ is the added point. Then $\|p - q\| \sim \varphi$, and $\|\widetilde{\mathcal{M}}(p) - \widetilde{\mathcal{M}}(q)\| \sim 2R\varphi$ as $\varphi \to +0$.

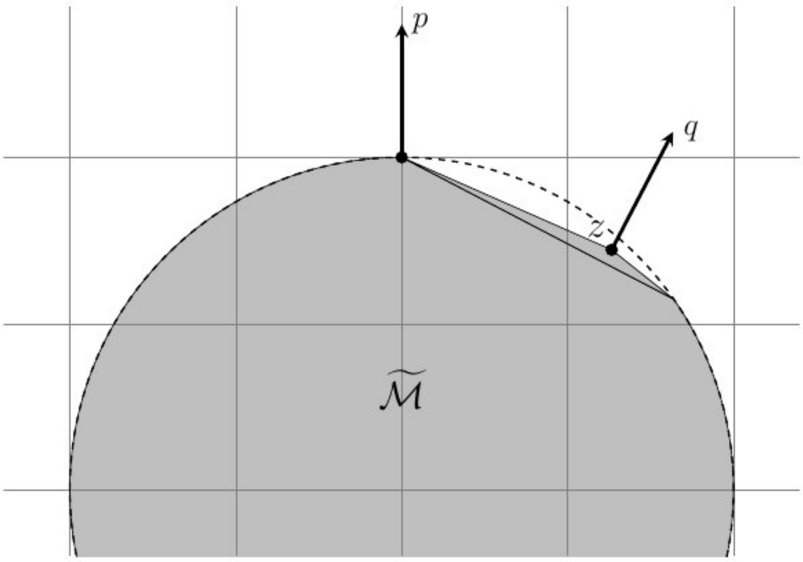

Figure 1: The constant $2R$ is optimal.

A very useful property of *sc* is that it is preserved in cartesian products, while strong convexity is not.

Lemma 1.6. *Let $\mathcal{M}_1 \subset \mathbb{R}^{n_1}$, $\mathcal{M}_2 \subset \mathbb{R}^{n_2}$ be convex compact sets. Assume that $\mathcal{M}_i \in sc(p_i, R_i)$, $i = 1, 2$. Then for any $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_1^2 + \alpha_2^2 = 1$ we have*

$$\mathcal{M}_1 \times \mathcal{M}_2 \in sc\left((\alpha_1 p_1, \alpha_2 p_2), \max\{R_1/\alpha_1, R_2/\alpha_2\}\right)$$

Proof. Take any $x_i \in \mathcal{M}_i$. From the definition of *sc* the following inequalities hold:

$$\|\mathcal{M}_i(p_i) - R_i p_i - x_i\| \leq R_i \Leftrightarrow \frac{1}{2R_i}\|\mathcal{M}_i(p_i) - x_i\|^2 \leq (p_i, \mathcal{M}_i(p_i) - x_i).$$

Adding them with appropriate coefficients and using the fact that supporting elements are the same for collinear vectors, we get

$$\begin{aligned}
&((\alpha_1 p_1, \alpha_2 p_2), (\mathcal{M}_1(\alpha_1 p_1), \mathcal{M}_2(\alpha_2 p_2)) - (x_1, x_2)) \\
&\geq \frac{1}{2R_1/\alpha_1}\|\mathcal{M}_1(\alpha_1 p_1) - x_1\|^2 + \frac{1}{2R_2/\alpha_2}\|\mathcal{M}_2(\alpha_2 p_2) - x_2\|^2 \\
&\geq \frac{1}{2\max\{R_1/\alpha_1, R_2/\alpha_2\}}\|(\mathcal{M}_1(\alpha_1 p_1), \mathcal{M}_2(\alpha_2 p_2)) - (x_1, x_2)\|^2.
\end{aligned}$$

As noted above, this inequality is equivalent to

$$\|(\mathcal{M}_1\times\mathcal{M}_2)(\alpha_1p_1,\alpha_2p_2)-R(\alpha_1p_1,\alpha_2p_2)-(x_1,x_2)\| \le R,\ R=\max\{R_1/\alpha_1,R_2/\alpha_2\}.$$

This concludes the proof. □

Here we also point out that the tangent cone to the set $\mathcal{M}_1 \times \mathcal{M}_2$ at the point $(\mathcal{M}_1 \times \mathcal{M}_2)(\alpha_1p_1,\alpha_2p_2)$ is not sharp.

The condition *sc* for a convex compact set has a typical nature in the following sense.

Proposition 1.7. ([4]) *Let $\mathcal{M} \subset \mathbb{R}^n$ be a convex compact set. Then the set of unit vectors p, for which we have the inclusion $\mathcal{M} \in sc(p, R(p))$ for some $R(p) > 0$ has a full spherical measure* [8, 3.10.82, p. 231] *on the unit sphere in $\mathbb{R}^n$.*

We normalize the measure so that the unit sphere in $\mathbb{R}^n$ has measure 1. In [4], we proved that the set $\mathcal{F}_R = \{p \in \mathbb{R}^n \ :\ \|p\| = 1,\ R(p) > R\}$ has measure

$$\mu\mathcal{F}_R \le 1-\left(1-\frac{R_{\mathcal{M}}}{R}\right)^{n-1}$$

where $R_{\mathcal{M}} = \min\limits_{x\in\mathbb{R}^n}\max\limits_{y\in\mathcal{M}}\|x-y\|$ is the Chebyshev radius of the set $\mathcal{M}$ and $R > R_{\mathcal{M}}$.

2. The nearest and the farthest points

For points $x, y \in \mathbb{R}^n$ denote $[x,y] = \{(1-\lambda)x+\lambda y \ :\ \lambda\in[0,1]\}$.

Theorem 2.1. *Let $\varrho_0 \ge r > 0$, $\|p_0\| = 1$, $R > 0$ and $\mathcal{M} \subset \mathbb{R}^n$ be a convex compact set with $\mathcal{M} \in sc(p_0, R)$. Suppose that $x_0 = \mathcal{M}(p_0)+\varrho_0p_0$ and $x \in \mathbb{R}^n$ is such a point that $\varrho = \varrho_{\mathcal{M}}(x) \ge r$. Then for $a = P_{\mathcal{M}}x$ and $a_0 = P_{\mathcal{M}}x_0 = \mathcal{M}(p_0)$ the following inequality holds*

$$\|a_0-a\| \le \frac{2R}{2R+r}\sqrt{\|x_0-x\|^2-(\varrho_0-\varrho)^2}.$$

Proof. Put $p = (x-a)/\varrho$. The vector p is a unit normal to the set $\mathcal{M}$ at the point $a = P_{\mathcal{M}}x$. By the implication $sc(p_0,R) \Rightarrow lip(p_0,2R)$

$$\|a_0-a\| \le 2R\left\|\frac{x_0-a_0}{\varrho_0}-\frac{x-a}{\varrho}\right\|.$$

Squaring and grouping the terms we get

$$\|a_0-a\|^2 \le 4R^2\left(2+\frac{\|a_0-a\|^2+\|x_0-x\|^2-\|x-a_0\|^2-\|x_0-a\|^2}{\varrho_0\varrho}\right). \tag{3}$$

Let $\alpha = \|a_0-a\|$, $z = a_0-Rp_0$.

Consider triangles x_0a_0a and za_0a. From the inclusion $a_0 \in [z,x_0]$ we have the equality $\angle x_0a_0a = \pi-\angle za_0a$ and

$$\cos\angle x_0a_0a = \cos(\pi-\angle za_0a) = -\frac{R^2+\alpha^2-\|z-a\|^2}{2R\alpha} \le -\frac{\alpha}{2R},$$

hence

$$\|x_0-a\|^2 = \|x_0-a_0\|^2 + \|a_0-a\|^2 - 2\|x_0-a_0\|\cdot\|a_0-a\|\cos\angle x_0a_0a$$
$$\geq \varrho_0^2+\alpha^2+2\varrho_0\alpha\frac{\alpha}{2R} = \varrho_0^2+\alpha^2+\frac{\alpha^2\varrho_0}{R}. \tag{4}$$

The angle $\angle xaa_0$ in the triangle xa_0a is greater or equal to $\pi/2$. Thus

$$\|x-a_0\|^2 \geq \|a_0-a\|^2+\|x-a\|^2 = \alpha^2+\varrho^2. \tag{5}$$

Substituting (4) and (5) in (3) and multiplying both sides on $\varrho_0\varrho$, we obtain that

$$\alpha^2\left(\varrho_0\varrho+4R\varrho_0+4R^2\right) \leq 4R^2\left(\|x_0-x\|^2-(\varrho_0-\varrho)^2\right).$$

By the inequalities $\varrho_0\geq r$, $\varrho\geq r$, the last formula proves the theorem. $\square$

Note that under conditions of Theorem 2.1 we have the estimate

$$\|a_0-a\| \leq \frac{2R}{2R+r}\|x_0-x\|. \tag{6}$$

Compared to Proposition 1.2, the Lipschitz constant is $\frac{2R}{2R+r}$ instead of $\frac{R}{R+r}$.

Theorem 2.2. *Let $\mathcal{M}\subset\mathbb{R}^n$ be a convex compact set, $x_0\in\mathbb{R}^n$, $\varrho_{\mathcal{M}}(x_0)=r>0$, $a_0=P_{\mathcal{M}}x_0$. Suppose that there exists such a constant $\lambda\in(0,1)$, that for any point $x\in\mathbb{R}^n$, $\varrho_{\mathcal{M}}(x)\geq r$, and $a=P_{\mathcal{M}}x$ the following inequality holds*

$$\|a_0-a\|\leq\lambda\|x_0-x\|.$$

Then $\mathcal{M}\in sc(p_0,R)$, where $p_0=\frac{x_0-a_0}{\|x_0-a_0\|}$ and $R=\frac{\lambda r}{1-\lambda}$.

Proof. Fix a unit vector p and a supporting element $a=\mathcal{M}(p)\in\mathcal{M}$ and put $x=a+rp$. Then $\varrho_{\mathcal{M}}(x)=r$ and $P_{\mathcal{M}}x=a$. By the condition of Theorem and equality $x_0=a_0+rp_0$ we have

$$\|a_0-a\|\leq\lambda\|a_0+rp_0-a-rp\|\leq\lambda\|a_0-a\|+\lambda r\|p-p_0\|,$$

i.e. $\|a_0-a\|\leq L\|p_0-p\|$ for $L=\frac{\lambda r}{1-\lambda}$. From the implication $lip(p_0,L)\Rightarrow sc(p_0,L)$ we get $\mathcal{M}\in sc(p_0,L)$. $\square$

Theorem 2.3. *Let $\mathcal{M}\subset\mathbb{R}^n$ be a convex compact set, $x_0\in\mathbb{R}^n$, $a_0=F_{\mathcal{M}}x_0$, $p_0=(a_0-x_0)/\|a_0-x_0\|$. Assume that $\mathcal{M}\in sc(p_0,R)$ and for some $R>0$, $r>0$ we have $\|a_0-x_0\|\geq r>2R$. Then for any point $x\in\mathbb{R}^n$ with $f_{\mathcal{M}}(x)\geq r$, and for any point $a\in F_{\mathcal{M}}x$ we have the estimate*

$$\|a_0-a\|\leq\frac{2R}{r-2R}\|x_0-x\|. \tag{7}$$

Proof. Recall that $\mathcal{M}(p_0)=\{a_0\}$. Fix $a\in F_{\mathcal{M}}x$. Put $p=(a-x)/\|a-x\|$.

By the implication $sc(p_0,R) \Rightarrow lip(p_0,2R)$

$$\|a_0-a\| \le 2R\|p_0-p\| = 2R\left\|\frac{a_0-x_0}{\|a_0-x_0\|}-\frac{a-x}{\|a-x\|}\right\|$$
$$\le 2R\left\|\frac{a_0-x_0}{r}-\frac{a-x}{r}\right\| \le \frac{2R}{r}\left(\|x_0-x\|+\|a_0-a\|\right).$$

From the last estimate and by the inequality $r>2R$ we get (7). □

Compared to Proposition 1.3, the Lipschitz constant is $\frac{2R}{r-2R}$ instead of $\frac{R}{r-R}$.

Below we consider some examples of applying the above results to prove the convergence of numerical algorithms. These examples under more stringent assumptions were considered in [13], where the sharp sufficient conditions for each of the problems to have a unique solution and, moreover, to be Tykhonov well-posed, were obtained.

Theorem 2.4. *Let $\mathcal{M},\mathcal{N}\subset\mathbb{R}^n$ be compact convex sets and $a_0\in\mathcal{M}$, $b_0\in\mathcal{N}$ be such points that*

$$\max_{x\in\mathcal{M}}\max_{y\in\mathcal{N}}\|x-y\| = \|b_0-a_0\|. \tag{8}$$

Assume that $\mathcal{M}\in sc(p_0,R)$ and $\mathcal{N}\in sc(-p_0.R)$ for $p_0=(a_0-b_0)/\|a_0-b_0\|$ and $R>0$. Fix $a_1\in\mathcal{M}$, $b_1\in\mathcal{N}$. For any natural $k\ge 1$ for given $a_k\in\mathcal{M}_0$, $b_k\in\mathcal{M}$ define the vector $p_k=(a_k-b_k)/\|a_k-b_k\|$ and points $a_{k+1}\in\mathcal{M}(p_k)$, $b_{k+1}\in\mathcal{N}(-p_k)$. Suppose that there exists a constant $r>0$ such that $\|a_1-b_1\|\ge r>4R$.

Then the iterations $\{a_k,b_k\}$ converge to a point $\{a_0,b_0\}$ with a linear rate:

$$\|a_{k+1}-a_0\|+\|b_{k+1}-b_0\| \le \frac{4R}{r}\left(\|a_k-a_0\|+\|b_k-b_0\|\right). \tag{9}$$

The point $\{a_0,b_0\}$ is the unique solution of (8).

Proof. By the inequality $r>4R$, the sets $\mathcal{B}_R(b_0+Rp_0)$, $\mathcal{B}_R(a_0-Rp_0)$ do not intersect and have a_0,b_0 as their most distant points. Then uniqueness follows from the inclusions $\mathcal{M}\subset\mathcal{B}_R(a_0-Rp_0)$, $\mathcal{N}\subset\mathcal{B}_R(b_0+Rp_0)$.

Also note that the norm $\|a_k-b_k\|$ is increasing.

$$\|a_{k+1}-b_{k+1}\| \ge (p_k,a_{k+1}-b_{k+1}) \ge (p_k,a_k-b_k) = \|a_k-b_k\|$$

Then by the inequality $r\le\|a_k-b_k\|$ we have

$$\|a_{k+1}-a_0\| \le 2R\|p_k-p_0\| = 2R\left\|\frac{a_k-b_k}{\|a_k-b_k\|}-\frac{a_0-b_0}{\|a_0-b_0\|}\right\| \le \frac{2R}{r}(\|a_k-a_0\|+\|b_k-b_0\|).$$

Similarly $$\|b_{k+1}-b_0\| \le \frac{2R}{r}(\|a_k-a_0\|+\|b_k-b_0\|).$$

The sum of two last inequalities finishes the proof. □

Note that if the set $\mathcal{N}=\{b_0\}$ in Theorem 2.4 is a singleton, then $b_k=b_0$ for all k and under the assumptions $\|a_1-b_0\|>r$ and $r>2R$ we get

$$\|a_k-a_0\| \le \frac{2R}{r}\|a_k-a_0\|.$$

Theorem 2.5. *Let $\mathcal{M}_0 \subset \mathbb{R}^n$ be a convex compact set and $\mathcal{M} \subset \mathbb{R}^n$ be a closed set, $a_0 \in \mathcal{M}_0$, $b_0 \in \mathcal{M}$ such points that*

$$\min_{x\in\mathcal{M}} \max_{y\in\mathcal{M}_0} \|x-y\| = \|b_0 - a_0\|. \tag{10}$$

Suppose that $\mathcal{M}_0 \in sc(p_0, R)$ for a vector $p_0 = (a_0 - b_0)/\|a_0 - b_0\|$ and a number $R > 0$ and there exists such a constant $C > 1$ that for any points $x, y \in \mathcal{M}_0$ we have the estimate $\|P_{\mathcal{M}}x - P_{\mathcal{M}}y\| \le C\|x-y\|$. Put $a_1 \in \mathcal{M}_0$, $b_1 = P_{\mathcal{M}}a_1 \in \mathcal{M}$. For any natural number $k \ge 1$ for given $a_k \in \mathcal{M}_0$, $b_k \in \mathcal{M}$ define $p_k = (a_k - b_k)/\|a_k - b_k\|$ and points $a_{k+1} \in \mathcal{M}_0(p_k)$, $b_{k+1} = P_{\mathcal{M}}a_{k+1}$. Then under the assumption $\frac{2R}{r}(1+C) < 1$, where $r = \min_{x\in\mathcal{M}_0, y\in\mathcal{M}} \|x-y\|$, iterations $\{a_k, b_k\}$ converge to the point $\{a_0, b_0\}$ with a linear rate:

$$\|a_{k+1} - a_0\| \le \frac{2R}{r}(1+C)\|a_k - a_0\|. \tag{11}$$

Moreover, $\|b_k - b_0\| \le C\|a_k - a_0\|$ and the solution $\{a_0, b_0\}$ of the problem (10) *is unique.*

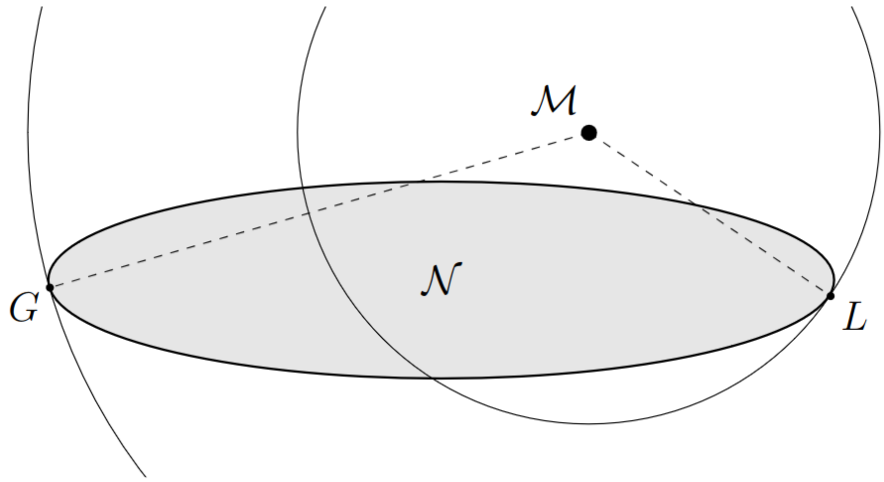

Figure 2: $r < 2R$ and $\mathcal{M}$ is a singleton. In this case the algorithm from Theorem 2.4 can converge to a local maximum L.

Proof. From the Lipschitz condition for $P_{\mathcal{M}}$ with Lipschitz constant $C > 0$ we have

$$\|b_k - b_0\| = \|P_{\mathcal{M}}a_k - P_{\mathcal{M}}a_0\| \le C\|a_k - a_0\|.$$

Thus if (11) holds then $\{b_k\}$ converges to b_0 linearly.

Now prove (11). We have

$$\begin{aligned}\|a_{k+1} - a_0\| &\le 2R\|p_k - p_0\| \le \frac{2R}{r}\|a_k - a_0 - (b_k - b_0)\| \\ &\le \frac{2R}{r}(\|a_k - a_0\| + \|b_k - b_0\|) \le \frac{2R}{r}(1+C)\|a_k - a_0\|.\end{aligned}$$

Assume that we have another solution $(\widetilde{a}_0, \widetilde{b}_0)$ of the problem (10). Take $a_1 = \widetilde{a}_0$, $b_1 = \widetilde{b}_0$, by formula (11) we obtain that the vector $p_1 = \frac{a_1 - b_1}{\|a_1 - b_1\|}$ is not normal to the set $\mathcal{M}_0$ at the point a_1. Thus the necessary condition of extremum is false. $\square$

Notice that the Lipschitz condition for the projector $P_{\mathcal{M}}$ in Theorem 2.5 takes place if the set $\mathcal{M}$ is proximally smooth [9, 10, 22, 26]. If the set $\mathcal{M}$ is proximally smooth with constant $K > 0$ then in any d-neighborhood $U_{\mathcal{M}}(d) = \{x \in \mathbb{R}^n : \varrho_{\mathcal{M}}(x) < d\}$, $d \in (0, K)$, of the set $\mathcal{M}$ for any points $x, y \in U_{\mathcal{M}}(d)$ we have

$$\|P_{\mathcal{M}}x - P_{\mathcal{M}}y\| \leq \frac{K}{K-d}\|x - y\|.$$

Also note that if the sets are R-convex, then we can use Lipschitz condition with constant R instead of $2R$ which improves the convergence speed estimates (and the corresponding conditions) by a factor of two. While the conditions imposed in the theorems are not necessary, there are counterexamples in case they fail, see Figure 2.

3. Some optimization problems with sets

Below we shall give some illustrative examples of using Theorems 2.4 and 2.5.

Consider the sets

$$\mathcal{E}_1 = \{x \mid x^T A_1 x \leq 1\}, \quad A_1 = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 9 \end{pmatrix};$$

$$\mathcal{E}_2 = \{x \mid x^T A_2 x \leq 1\}, \quad A_2 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 12 \end{pmatrix};$$

$$\mathcal{R}_t = \int_0^t e^{As}\,\mathcal{U}\,ds, \quad A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}, \quad \mathcal{U} = \mathrm{co}\{\pm e_3\}.$$

The integral in the definition of $\mathcal{R}_t$ above is considered in the Aumann sense.

We calculate the supporting element $\mathcal{R}_t(p)$ for a unit vector $p \in \mathbb{R}^n$ by the formula

$$\mathcal{R}_t(p) = \int_0^t e^{As}\,\mathcal{U}\left(e^{A^T s}p\right)\,ds,$$

the last integral is calculated with high accuracy using standard Python tools.

Note that $\mathcal{R}_t$ is the reachable set of the differential inclusion $\dot{x} \in Ax + \mathcal{U}$ with zero initial condition. This set is not strongly convex [5].

Example 3.1. Consider problem (8) with the following sets in $\mathbb{R}^3$ (see Figure 3 on the next page):

$$\mathcal{M} = \mathcal{R}_2, \quad \mathcal{N} = \mathcal{E}_1 + \mathcal{E}_2 + \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix},$$

Example 3.2. Consider problem (8) with following sets in $\mathbb{R}^4$

$$\mathcal{M} = \mathcal{B}_{1/2}\left(\begin{smallmatrix} 4 \\ -5 \end{smallmatrix}\right) \times \mathcal{B}_{1/2}\left(\begin{smallmatrix} 0 \\ -6 \end{smallmatrix}\right), \quad \mathcal{N} = [-6, -4] \times \mathcal{R}_2.$$

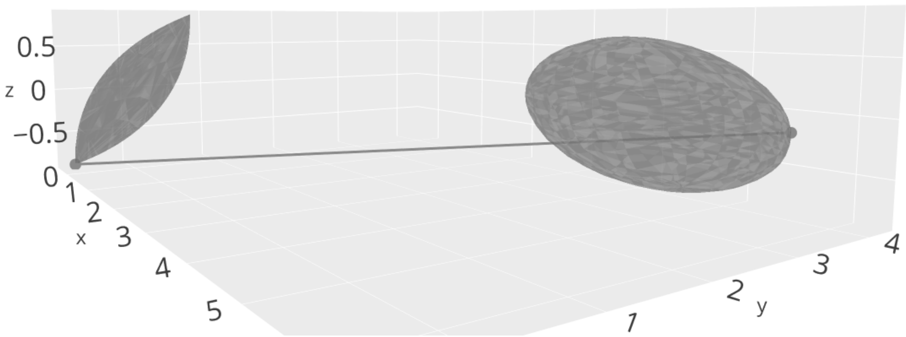

Figure 3: First example, $\mathcal{M}$ on the left, $\mathcal{N}$ on the right

Example 3.3. Consider problem (10) with the following sets in $\mathbb{R}^3$:

$$\mathcal{M}_0 = \mathcal{E}_1 + \mathcal{E}_2;$$
$$e = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$
$$\mathcal{M} = \left\{ x \in \mathbb{R}^3 \mid \|x - x_0\| = 30,\ \frac{(e, x - x_0)}{\|x - x_0\|} \geq \cos(\pi/15) \right\},$$
$$x_0 = -15e.$$

The algorithms stopped when the total change of both a_k and b_k per iteration was less than 10^{-8}. In all cases the convergence was linear as shown in Figure 4 (note the logarithmic scale).

For the Examples 3.1 and 3.2 the plots of $\|a_k - \widehat{a}_0\| + \|b_k - \widehat{b}_0\|$ are depicted (where $\widehat{a}_0$, $\widehat{b}_0$ is the found approximate solution), while for example 3 we plot $\|a_k - \widehat{a}_0\|$. The number of steps N, the estimated rate of linear convergence q, found approximate solution $(\widehat{a}_0, \widehat{b}_0)$, as well the numeric estimates of constants R, r and C from Theorems 2.4, 2.5 are shown in Table 1. In Example 3.2, Lemma 1.6 was also used to estimate R.

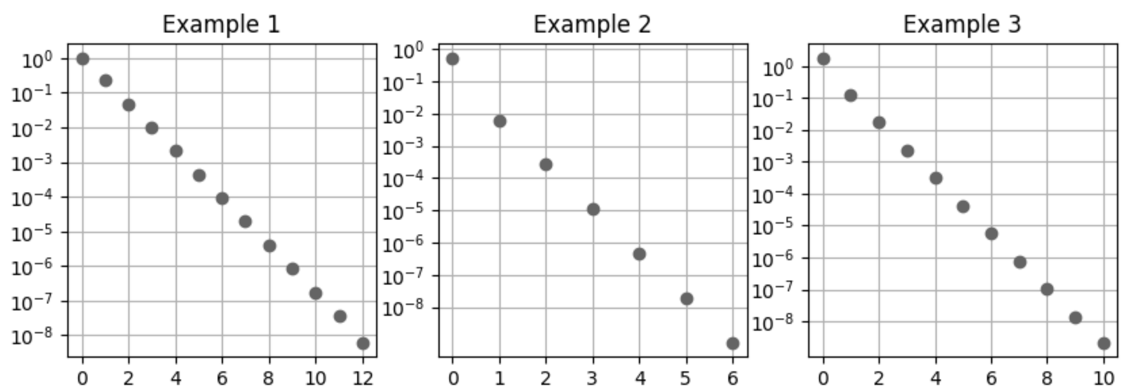

Figure 4: Graphs of distance to the found solution in logarithmic scale

Example	N	q	R	r	C	$(\widehat{a}_0, \widehat{b}_0)$
1	14	0.21	1.8	7.3	-	$\left(\begin{pmatrix}-0.32\\-0.59\\-0.86\end{pmatrix}, \begin{pmatrix}5.81\\3.68\\-0.11\end{pmatrix}\right)$
2	7	0.04	3.0	13.7	-	$\left(\begin{pmatrix}4.44\\-5.23\\-0.04\\-6.50\end{pmatrix}, \begin{pmatrix}-6.00\\0.32\\0.59\\0.86\end{pmatrix}\right)$
3	12	0.13	2.2	14.5	2.1	$\left(\begin{pmatrix}-0.64\\-0.74\\-0.13\end{pmatrix}, \begin{pmatrix}8.37\\8.15\\9.43\end{pmatrix}\right)$

Table 1: Numeric results for the Examples

References

[1] M. V. Balashov: *Antidistance and antiprojection in the Hilbert space*, J. Convex Analysis 22/2 (2015) 521–536.

[2] M. V. Balashov: *About the gradient projection algorithm for a strongly convex function and a proximally smooth set*, J. Convex Analysis 24/2 (2017) 493–500.

[3] M. V. Balashov: *Sufficient conditions for the linear convergence of an algorithm for finding the metric projection of a point onto a convex compact set*, Math. Notes 113/5 (2023) 632–641.

[4] M. V. Balashov, K. Z. Biglov: *The strong convexity supporting condition and the Lipschitz condition for the metric projection*, Math. Notes 115/2 (2024) 164–172.

[5] M. V. Balashov, K. Z. Biglov, A. A. Tremba: *On some problems with multivalued mappings*, Automation Remote Control 85/5 (2024) 491–511.

[6] M. V. Balashov, M. O. Golubev: *About the Lipschitz property of the metric projection in the Hilbert space*, J. Math. Analysis Appl. 394/2 (2012) 545–551.

[7] M. V. Balashov, G. E. Ivanov: *On farthest points of sets*, Math. Notes 80/2 (2006) 159–166.

[8] V. I. Bogachev: *Measure Theory Vol. 1*, Springer, Berlin (2007).

[9] F. H. Clarke, J. R. Stern, P. R. Wolenski: *Proximal smoothness and the lower-C^2 property*, J. Convex Analysis 2/1-2 (1995) 117–144.

[10] H. Federer: *Curvature measures*, Trans. Amer. Math. Soc. 93/3 (1959) 418–491.

[11] H. Frankowska, C. Olech: *R-convexity of the integral of the set-valued functions*, Contributions to Analysis and Geometry, Suppl. Amer. J. Math. (1981) 117–129.

[12] V. V. Goncharov, G. E. Ivanov: *Strong and weak convexity of closed sets in a Hilbert space*, in: *Operations Research, Engineering, and Cyber Security. Trends in Applied Mathematics and Technology*, N. J. Daras et al. (eds.), Nea Peramos, Greece, May 2015, Springer Optimization and Its Applications Vol. 113, Springer, Cham (2017) 259–297.

[13] G. E. Ivanov: *Extremal problems for the distance between elements of two sets*, J. Convex Analysis 29/3 (2022) 767–788.

[14] G. E. Ivanov, E. S. Polovinkin: *On strongly convex differential games*, Diff. Equations 31/10 (1995) 1603–1612.

[15] S. Lojaciewicz: *Some properties of accessible sets in nonlinear control systems*, Ann. Polon. Math. 36 (1979) 123–137.

[16] A. E. Mayer: *Der Inhalt der Gleichdicke*, Math. Ann. 110 (1935) 97–127.

[17] A. E. Mayer: *Eine Überkonvexität*, Math. Zeitschrift 39 (1935) 511–531.

[18] F. Nacry, N. V. A. Thuong, L. Thibault: *Farthest distance function to strongly convex sets*, J. Convex Analysis 30/4 (2023) 1217–1240.

[19] I. Necoara, Yu. Nesterov, F. Glineur: *Linear convergence of first order methods for non-strongly convex optimization*, Math. Programming Ser. A 175/1-2 (2019) 69–107.

[20] A. Pliś: *Uniqueness of optimal trajectories for non-linear control systems*, Ann. Polon. Math. 29 (1975) 397–401.

[21] E. S. Polovinkin: *Strongly convex analysis*, Sb. Math. 187/2 (1996) 259–286.

[22] Yu. G. Reshetnyak: *On a generalization of convex surfaces*, Mat. Sb. (N.S.) 40(82)/3 (1956) 381–398.

[23] R. T. Rockafellar: *Convex Analysis*, Princeton University Press, Princeton (1970).

[24] L. Thibault: *Jean Jacques Moreau. A selected review of his mathematical works*, J. Convex Analysis 23/1 (2016) 005–022.

[25] V. M. Veliov: *On the convexity of integrals of multivalued mappings: application in control theory*, J. Optimization Theory Appl. 54/3 (1987) 541–563.

[26] J.-Ph. Vial: *Strong and weak convexity of sets and functions*, Math. Oper. Res. 8/2 (1983) 231–259.