Differentiability properties of metric projections onto convex sets

Alexander Shapiro[*]

School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA
e-mail: ashapiro@isye.gatech.edu

Abstract

It is known that directional differentiability of metric projection onto a closed convex set in a finite dimensional space is not guaranteed. In this paper we discuss sufficient conditions ensuring directional differentiability of such metric projections. The approach is based on a general theory of sensitivity analysis of parameterized optimization problems.

Keywords: metric projection, directional differentiability, second order regularity, cone reducibility, nondegeneracy

[*]This research was partly supported by the NSF award CMMI 1232623.

1 Introduction

In this paper we discuss directional differentiability properties of metric projections onto convex sets in finite dimensional spaces. Let S be a nonempty convex closed subset of a finite dimensional Euclidean space $\mathcal{X}$ equipped with scalar product $\langle \cdot, \cdot \rangle$ and norm $\|x\| = \sqrt{\langle x, x \rangle}$. Metric projection $P_S : \mathcal{X} \to S$ onto the set S is defined as

$$P_S(x) := \arg\min\{\|x - v\| : v \in S\}. \tag{1.1}$$

That is, $v = P_S(x)$ is the closest point of S to x. Since the set S is convex and closed such point exists and is unique, and hence $P_S : \mathcal{X} \to S$ is well defined. Of course if $x \in S$, then $P_S(x) = x$.

Recall that a mapping $G : \mathcal{X} \to \mathcal{Y}$, from $\mathcal{X}$ to a finite dimensional linear space $\mathcal{Y}$, is said to be directionally differentiable at a point $x \in \mathcal{X}$ if the directional derivative

$$G'(x, h) = \lim_{t \downarrow 0} \frac{G(x + th) - G(x)}{t}.$$

exists for every $h \in \mathcal{X}$. It is known that if $x \in S$, then P_S is directionally differentiable at x and

$$P'_S(x, d) = P_{T_S(x)}(d),\ d \in \mathcal{X}, \tag{1.2}$$

where

$$T_S(x) := \big\{h \in \mathcal{X} : \operatorname{dist}(x + th, S) = o(t),\ t \geq 0\big\}, \tag{1.3}$$

denotes the tangent cone to S at $x \in S$ (cf., Zarantonello [10]).

It is well known that

$$\|P_S(x_1) - P_S(x_2)\| \leq \|x_1 - x_2\|,\ \forall x_1, x_2 \in \mathcal{X}. \tag{1.4}$$

Therefore if P_S is directionally differentiable at a point $x \in \mathcal{X}$, then

$$\|P'_S(x, d_1) - P'_S(x, d_2)\| \leq \|d_1 - d_2\|,\ \forall d_1, d_2 \in \mathcal{X}, \tag{1.5}$$

and P_S is directionally differentiable at x in the sense of Hadamard (e.g., [7]).

When $x \notin S$ directional differentiability of P_S at x is not guaranteed. First example of a convex set with nondirectionally differentiable metric projection was constructed by Kruskal [5]. Kruskal's example is of a convex set in $\mathbb{R}^3$, in fact such sets exist already in $\mathbb{R}^2$, [8]. In this paper we discuss sufficient conditions ensuring directional differentiability of such metric projections. The approach is based on a general theory of sensitivity analysis of parameterized optimization problems. In the next section we briefly survey some basic concepts relevant for the developed theory. Main developments are presented in section 3. Although this is mainly a survey paper, some of the presented results are new.

We use the following notation throughout the paper. By $\operatorname{dist}(x, A) := \inf_{v \in A} \|x - v\|$ we denote the distance from $x \in \mathcal{X}$ to a set $A \subset \mathcal{X}$, and by

$$\sigma(h, A) := \sup_{v \in A} \langle h, v \rangle$$

the support function of set A. For a convex cone C its *lineality space* $\text{lin}(C) := C \cap (-C)$. Let $\mathcal{Y}$ be a finite dimensional Euclidean space. For a mapping $G : \mathcal{X} \to \mathcal{Y}$ its first order derivative, at a point $x \in \mathcal{X}$, is denoted $DG(x)$. For locally Lipschitz mappings and finite dimensional vector spaces $\mathcal{X}$ and $\mathcal{Y}$, all standard concepts of the derivative $DG(x) : \mathcal{X} \to \mathcal{Y}$ do coincide and mean that the directional derivative $G'(x, h)$ exists and is linear in $h \in \mathcal{X}$, and $DG(x)h = G'(x, h)$ for all $h \in \mathcal{X}$. The corresponding second order derivative is denoted by $D^2G(x)$. If G is twice continuously differentiable at x, then we have the following second order Taylor expansion

$$G(x + h) = G(x) + DG(x)h + \tfrac{1}{2}D^2G(x)(h, h) + o(\|h\|^2),$$

with $D^2G(x) : \mathcal{X} \times \mathcal{X} \to \mathcal{Y}$ being a symmetric bilinear mapping.

By $\text{sp}(h)$ we denote the linear space generated by $h \in \mathcal{X}$. Of course, the space $\text{sp}(h)$ is one dimensional unless $h = 0$. For a set $C \subset \mathcal{X}$ we denote by $\text{cl}(C)$ its topological closure, and by $\text{int}(C)$ its interior. By $\text{Tr}(A)$ we denote trace of a square matrix A, and write $A \preceq 0$ to denote that (symmetric) matrix A is negative semidefinite. For a linear mapping $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$ we denote by $\mathcal{A}^* : \mathcal{Y} \to \mathcal{X}$ its conjugate mapping defined by $\langle \mathcal{A}x, y\rangle = \langle x, \mathcal{A}^*y\rangle$. By

$$K^- := \{z \in \mathcal{Y} : \langle z, y\rangle \le 0,\ \forall y \in K\},$$

we denote the (negative) dual of cone $K \subset \mathcal{Y}$.

2 Basic concepts

In this section we discuss some concepts which will be needed for the subsequent analysis. The outer and inner second order tangent sets to the set S at $x \in S$ in direction h are defined as

$$\mathcal{T}^2_S(x, h) := \big\{w \in \mathcal{X} : \exists t_n \downarrow 0,\ \text{dist}(x + t_n h + \tfrac{1}{2}t_n^2 w, S) = o(t_n^2)\big\} \tag{2.1}$$

and

$$T^2_S(x, h) := \big\{w \in \mathcal{X} : \text{dist}(x + th + \tfrac{1}{2}t^2 w, S) = o(t^2),\ t \ge 0\big\}, \tag{2.2}$$

respectively. Clearly $T^2_S(x, h) \subset \mathcal{T}^2_S(x, h)$.

For convex sets similar concepts of *first* order tangent sets (cones) do coincide. That is, the “inner” tangent cone $T_S(x)$, defined in (1.3), coincides with the respective “outer” (also called contingent) tangent cone. On the other hand, the second order tangent sets $T^2_S(x, h)$ and $\mathcal{T}^2_S(x, h)$ can be different even if the set S is convex (e.g., [3, Example 3.31]). It follows from the definition that $T^2_S(x, h)$ can be nonempty only if $h \in T_S(x)$. Even if $h \in T_S(x)$ the set $T^2_S(x, h)$ can be empty (e.g., [3, Example 3.29]). The inner second order tangent set $T^2_S(x, h)$ is closed and convex. On the other hand, the outer second order tangent set $\mathcal{T}^2_S(x, h)$ can be nonconvex even if the set S is convex (cf., [3, Example 3.35]).

We can formulate now the basic concept of second order regularity.

Definition 2.1 *It is said that the set S is* second order regular *at a point $\bar{x} \in S$ if for any sequence $x_n \in S$ of the form $x_n = \bar{x} + t_n h + \frac{1}{2} t_n^2 r_n$, where $t_n \downarrow 0$ and $t_n r_n \to 0$, it follows that*

$$\lim_{n \to \infty} \operatorname{dist}\left(r_n, T_S^2(\bar{x}, h)\right) = 0. \tag{2.3}$$

We say that the set S is second order regular if it is second order regular at its every point.

The concept of second order regularity was introduced in Bonnans, Cominetti and Shapiro [2] and discussed in details in [3, sction 3.3.3]. In the above definition the term $t_n^2 r_n = o(t_n)$ and since $x_n \in S$, it follows that $h \in T_S(\bar{x})$. If the sequence r_n is *bounded*, then condition (2.3) holds by the definition of the set $T_S^2(\bar{x}, h)$.

Second order regularity of S implies the following two properties at $\bar{x} \in S$. For any $h \in \mathcal{X}$ the outer and inner second order tangent sets $\mathcal{T}_S^2(\bar{x}, h)$ and $T_S^2(\bar{x}, h)$ do coincide. Indeed, this should be verified only for $h \in T_S(\bar{x})$, since otherwise both sets are empty. Now by the definition, for $w \in \mathcal{T}_S^2(\bar{x}, h)$ there exists a sequence $t_n \downarrow 0$ and $r_n \in \mathcal{X}$ such that r_n tends to w and $x + t_n h + \frac{1}{2} t_n^2 r_n \in S$. By (2.3) it follows that $w \in T_S^2(\bar{x}, h)$. This shows that $\mathcal{T}_S^2(\bar{x}, h) \subset T_S^2(\bar{x}, h)$. Since the opposite inclusion always holds, it follows that these two sets are equal to each other. Second order regularity of S also implies that for any $h \in T_S(\bar{x})$ the set $T_S^2(\bar{x}, h)$ is nonempty. Indeed, for $h \in T_S(\bar{x})$ we have that there exists $v_n \in S$ such that $v_n - \bar{x} - t_n h = o(t_n)$. Then take $r_n := 2t_n^{-2}(v_n - \bar{x} - t_n h)$. For this r_n the distance in (2.3) tends to zero and hence the set $T_S^2(\bar{x}, h)$ cannot be empty.

Although necessary, the above two properties are not sufficient for the second order regularity of S at $\bar{x}$ (cf., [3, Example 3.87]). Nevertheless, it turns out that many interesting convex sets are second order regular. Let us show first that the second order regularity holds in the following case. Suppose for the moment that S is a convex cone and $\bar{x} = 0$. Then (e.g., [3, p.168])

$$T_S^2(\bar{x}, h) = \mathcal{T}_S^2(\bar{x}, h) = T_S(h).$$

Moreover, since S is a cone and $\bar{x} = 0$, the condition $\bar{x} + t_n h + \frac{1}{2} t_n^2 r_n \in S$ means that $h + \frac{1}{2} t_n r_n \in S$. Since $S - h \subset T_S(h)$ it follows that $r_n \in T_S(h)$, and hence $\operatorname{dist}\left(r_n, T_S^2(\bar{x}, h)\right) = 0$. It follows that the convex cone S is second order regular at $\bar{x} = 0$. This shows that if the set S coincides with $\bar{x} + T_S(\bar{x})$ in vicinity of the point $\bar{x}$, then S is second order regular at $\bar{x}$. In particular polyhedral sets are second order regular.

Let us consider now sets of the form $S = G^{-1}(K)$, i.e.,

$$S := \{x \in \mathcal{X} : G(x) \in K\}, \tag{2.4}$$

where $K \subset \mathcal{Y}$ is a closed convex cone in a finite dimensional space $\mathcal{Y}$, and $G : \mathcal{X} \to \mathcal{Y}$ is a twice continuously differentiable mapping. Of course the mapping G should satisfy some conditions in order for the set S to be convex, we will discuss this later. Consider a point $\bar{x} \in S$ and $\bar{y} := G(\bar{x})$. Recall that Robinson's constraint qualification holds at $\bar{x}$ if

$$DG(\bar{x})\mathcal{X} + T_K(\bar{y}) = \mathcal{Y}. \tag{2.5}$$

We have the following result ([2, Proposition 4.2], [3, Proposition 3.88]).

Proposition 2.1 *Suppose that the mapping $G : \mathcal{X} \to \mathcal{Y}$ is twice continuously differentiable, Robinson's constraint qualification holds at a point $\bar{x} \in S$ and the cone K is second order regular at $G(\bar{x}) \in K$. Then the set S is second order regular at $\bar{x}$.*

It follows by the above discussion that if the cone K is polyhedral and Robinson's constraint qualification holds, then the set $S = G^{-1}(K)$ is second order regular. In fact Proposition 2.1 implies second order regularity for a larger class of sets. Consider the following concept ([3, Definition 3.135]). Recall that a convex cone C is said to be *pointed* if its *lineality space* $\text{lin}(C)$ is $\{0\}$.

Definition 2.2 *It is said that a set $S \subset \mathcal{X}$ is* cone reducible *at a point $\bar{x} \in S$ if there exists a neighborhood N of $\bar{x}$, a closed pointed convex cone C in a finite dimensional space $\mathcal{Z}$ and a twice continuously differentiable mapping $\Xi : N \to \mathcal{Z}$ such that:* (i) $\Xi(\bar{x}) = 0$, (ii) $D\Xi(\bar{x}) : N \to \mathcal{Z}$ *is onto, and* (iii) $S \cap N = \{x \in N : \Xi(x) \in C\}$.

If S is cone reducible at its every point, then it is said that S is cone reducible.

Note that we require here for the cone C to be *pointed*. In [3] such cone reduction was called pointed. Condition (iii) of the above definition means that locally, in vicinity of the point $\bar{x}$, the set S can be defined by the constraint $\Xi(x) \in C$. Since the cone C is second order regular at $\Xi(\bar{x}) = 0$, it follows by Proposition 2.1, that the corresponding set S is second order regular at $\bar{x}$. That is, cone reducibility implies second order regularity. It is not difficult to see that polyhedral sets and spherical (also called ice-cream or Lorentz) cones are cone reducible. It is also possible to show that the sets (cones) of positive semi-definite symmetric matrices are cone reducible (cf., [1, Example 4], [3, Example 3.140]).

Condition stronger than Robinsons constraint qualification (2.5) is the following condition of nondegeneracy.

Definition 2.3 *Consider set S of the form* (2.4)*, a point $\bar{x} \in S$ and $\bar{y} := G(\bar{x})$. It is said that $\bar{x}$ is a* nondegenerate *point of G, with respect to K, if*

$$DG(\bar{x})\mathcal{X} + \text{lin}\,(T_K(\bar{y})) = \mathcal{Y}. \tag{2.6}$$

The above concept of nondegeneracy was discussed in Roninson [6] for polyhedral sets K, and in Bonnans and Shapiro [1] (see also [3, section 4.6.1]) for cone reducible sets. That is, suppose that the cone K is cone reducible at $\bar{y} = G(\bar{x})$ to a (pointed) cone C. Then it follows from conditions (i)-(iii) of Definition 2.2 that the set

$$W := \{y \in N : \Xi(y) = 0\}$$

forms a smooth manifold in a neighborhood of the point $\bar{y}$, and the tangent space $T_W(\bar{y})$, to this manifold, coincides with $\text{lin}\,(T_K(\bar{y}))$ (cf., [3, Proposition 4.73]). In that case the nondegeneracy condition (2.6) coincides with the transversality condition used in differential geometry. Transversality is stable under small perturbations (see, e.g., discussion in [3, p.475]). As the following example shows, without the cone reducibility the nondegeneracy can be unstable.

Example 2.1 Let us construct the following set $K \subset \mathbb{R}^2$. Consider a sequence $t_n \downarrow 0$ (e.g., take $t_n = 1/n$) and the following sequence of points $y_0 = (0,0)$, $y_1 = (t_1, t_1^2)$, $y_2 = (-t_1, t_1^2)$, $y_3 = (t_2, t_2^2)$, $y_4 = (-t_2, t_2^2)$, ..., in $\mathbb{R}^2$. Let K be the convex hull of these points $y_0, y_1, \ldots$. Define mapping $G : \mathbb{R} \to \mathbb{R}^2$ as $G(x) = (0, x)$, and let $\bar{x} = 0$, and hence $\bar{y} = G(\bar{x}) = (0,0)$. We have that $T_K(\bar{y}) = \{y \in \mathbb{R}^2 : y_2 \geq 0\}$ and hence $\text{lin}\,(T_K(\bar{y})) = \{y : y_2 = 0\}$. It follows that $\bar{x} = 0$ is a nondegenerate point of G with respect to K. On the other hand consider slightly perturbed mappings $G_n(x) := (t_n, x)$. Note that $G_n(x_n) = y_{2n-1}$, where $x_n := t_n^2$. It is not difficult to see that $\text{lin}\,(T_K(y_n)) = (0,0)$ for any $n \geq 1$. It follows that x_n is not a nondegenerate point of G_n for $n \geq 1$. ∎

We also have the following result (cf., [9, Proposition 3.1]).

Proposition 2.2 *Suppose that the set K is cone reducible at the point $\bar{y} = G(\bar{x})$, and that the point $\bar{x} \in S$ is nondegenerate with respect to G and K. Then the set S is cone reducible at $\bar{x}$.*

3 Main results

In this section we discuss differentiability properties of the metric projection $P_S : \mathcal{X} \to S$. Let us recall the following basic result from [2, Theorem 7.2].

Theorem 3.1 *Let $S \subset \mathcal{X}$ be a closed convex set, a point $\bar{x} \in \mathcal{X}$ and $\bar{v} := P_S(\bar{x})$. Suppose that the set S is second order regular at $\bar{v}$. Then P_S is directionally differentiable at $\bar{x}$ and the directional derivative $P'_S(\bar{x}, d)$ is given by the optimal solution of the problem*

$$\operatorname*{Min}_{h \in C(\bar{v})} \left\{ \|d - h\|^2 - \sigma\big(\bar{x} - \bar{v}, T_S^2(\bar{v}, h)\big) \right\}, \tag{3.1}$$

where $C(\bar{v})$ is the so-called critical cone

$$C(\bar{v}) := \{h \in T_S(\bar{v}) : \langle \bar{x} - \bar{v}, h \rangle = 0\}. \tag{3.2}$$

Let us discuss this result. In a sense the term (called *sigma term*)

$$\mathfrak{s}_{\bar{x}}(h) := \sigma\big(\bar{x} - \bar{v}, T_S^2(\bar{v}, h)\big), \tag{3.3}$$

in formula (3.1), represents the curvature of the set S at the point $\bar{v} = P_S(\bar{x})$. Suppose for the moment that $\bar{x} \in S$. Then $\bar{v} = P_S(\bar{x})$ coincides with $\bar{x}$ and hence $C(\bar{v}) = T_S(\bar{v})$ and the sigma term $\mathfrak{s}_{\bar{x}}(h)$ vanishes. In that case optimal solution of problem (3.1) is given by $P_{T_S(\bar{x})}(d)$, and hence the above theorem gives the same formula as in (1.2). Also in that case there is no need for the second order regularity condition.

If $\bar{x} \notin S$, and hence $P_S(\bar{x}) \neq \bar{x}$, then the directional differentiability of P_S at $\bar{x}$ is ensured by the second order regularity condition. Since the set S is second order regular at $\bar{v} = P_S(\bar{x})$, it follows that the second order tangent set $T_S^2(\bar{v}, h)$ is nonempty for every

$h \in T_S(\bar{v})$. Therefore $\mathfrak{s}_{\bar{x}}(h) > -\infty$ for every $h \in C(\bar{v})$. Moreover, we have that if $h \in T_S(\bar{v})$ and $w \in T_S^2(\bar{v}, h)$, then (cf., [4])

$$w + T_S(\bar{v}) + \mathrm{sp}(h) \subset T_S^2(\bar{v}, h) \subset \mathrm{cl}\{T_S(\bar{v}) + \mathrm{sp}(h)\}. \tag{3.4}$$

Also by the first order optimality conditions we have that

$$\langle \bar{x} - \bar{v}, h \rangle \leq 0, \ \forall h \in T_S(\bar{v}). \tag{3.5}$$

It follows that $\mathfrak{s}_{\bar{x}}(h) \leq 0$ for every $h \in C(\bar{z})$. If the set S is *polyhedral*, then this sigma term vanishes, and the directional derivative $P'_S(\bar{x}, d)$ is given by the metric projection of d onto the corresponding critical cone $C(\bar{z})$. Note that the function $\mathfrak{s}_{\bar{x}}(\cdot)$ is concave and hence $-\mathfrak{s}_{\bar{x}}(\cdot)$ is convex (cf., [2, Lemma 4.1]). Therefore (3.1) is a convex problem.

Suppose now that the set S is cone reducible, to a cone $\mathcal{C}$ by mapping Ξ, at the point $\bar{v} = P_S(\bar{x})$. Then (e.g., [3, Proposition 3.136])

$$T_S^2(\bar{v}, h) = D\Xi(\bar{v})^{-1} \left\{ T_{\mathcal{C}}^2(0, D\Xi(\bar{v})h) - D^2\Xi(\bar{v})(h, h) \right\}, \tag{3.6}$$

and since $\mathcal{C}$ is a convex cone,

$$T_{\mathcal{C}}^2(0, D\Xi(\bar{v})h)) = T_{\mathcal{C}}(D\Xi(\bar{v})h) = \mathrm{cl}\big\{\mathcal{C} + \mathrm{sp}(D\Xi(\bar{v})h)\big\}. \tag{3.7}$$

Moreover, $T_S(\bar{v}) = D\Xi(\bar{v})^{-1}\mathcal{C}$. Together with (3.2) and (3.5) this implies that for $h \in C(\bar{v})$,

$$\langle \bar{x} - \bar{v}, w \rangle \leq 0, \ \forall w \in A := D\Xi(\bar{v})^{-1} \left\{ T_{\mathcal{C}}^2(0, D\Xi(\bar{v})h)) \right\}, \tag{3.8}$$

and hence $\sigma(\bar{x} - \bar{v}, A) = 0$. Thus for $h \in C(\bar{v})$ we have by (3.3) that

$$\mathfrak{s}_{\bar{x}}(h) = \langle \bar{x} - \bar{v}, D\Xi(\bar{v})^{-1}[D^2\Xi(\bar{v})(h, h)] \rangle. \tag{3.9}$$

It follows that in the cone reducible case the sigma term $\mathfrak{s}_{\bar{x}}(\cdot)$ is *quadratic* on $C(\bar{v})$. This leads to the following result.

Proposition 3.1 *Suppose that the set S is cone reducible at the point $\bar{v} = P_S(\bar{x})$. Then P_S is directionally differentiable at $\bar{x}$. Moreover, P_S is differentiable at $\bar{x}$ iff the critical cone $C(\bar{v})$ is a linear space.*

Proof. Since the cone reducibility implies the second order regularity, it follows by Theorem 3.1 that P_S is directionally differentiable at $\bar{x}$.

In order to verify differentiability of P_S at $\bar{x}$ we only need to verify that $P'_S(\bar{x}, d)$ is linear in d. By the above discussion the sigma term $-\mathfrak{s}_{\bar{x}}(\cdot)$ is quadratic on $C(\bar{v})$. Since $-\mathfrak{s}_{\bar{x}}(\cdot)$ is convex, the corresponding quadratic function is positive semidefinite on the linear space generated by $C(\bar{v})$. Therefore we can write the objective function of (3.1) as $\|d-h\|^2 + \langle h, Qh \rangle$ for some positive semidefinite matrix Q. It follows that the minimizer of $\|d - h\|^2 + \langle h, Qh \rangle$ over $h \in C(\bar{v})$ is a linear function of d iff the convex cone $C(\bar{v})$ is a linear space. ∎

Example 3.1 (semidefinite cone) Let $\mathcal{X} := \mathcal{S}^n$ be the space of $n \times n$ symmetric matrices, equipped with the scalar product $\langle X, Y \rangle := \mathrm{Tr}(XY)$, and $S := \mathcal{S}^n_+$ be the cone of positive semidefinite matrices. Since the set (cone) $\mathcal{S}^n_+$ is cone reducible, and hence is second order regular, we have that P_S is directionally differentiable. Consider $\bar{X} \in \mathcal{S}^n$ and let $\bar{V} := P_S(\bar{X})$ and $\Omega := \bar{X} - \bar{V}$. Since S is a convex cone, Ω belongs to the dual of the cone S, and hence $\Omega \preceq 0$. Also we have that $\langle \Omega, \bar{V} \rangle = 0$. This implies that if $\bar{X} \notin S$ and hence $\Omega \neq 0$, then rank $r = \mathrm{rank}(\bar{V})$ is less than n.

The tangent cone to $S = \mathcal{S}^n_+$ at $\bar{V} \in S$ can be written as

$$T_{\mathcal{S}^n_+}(\bar{V}) = \{H \in \mathcal{S}^n : E^T H E \succeq 0\}, \tag{3.10}$$

where E is an $n \times r$ matrix of full column rank $r = \mathrm{rank}(\bar{V})$ such that $\bar{V} E = 0$. Moreover, if $H \in T_{\mathcal{S}^n_+}(\bar{V})$, then (cf., [3, p.487])

$$\sigma\left(\Omega, T^2_{\mathcal{S}^n_+}(\bar{V}, H)\right) = \begin{cases} 2\mathrm{Tr}(\Omega H \bar{V}^\dagger H), & \text{if } \Omega \preceq 0,\ \mathrm{Tr}(\Omega \bar{V}) = 0,\ \mathrm{Tr}(\Omega H) = 0, \\ -\infty, & \text{otherwise}, \end{cases} \tag{3.11}$$

where $\bar{V}^\dagger$ denotes the Moore-Penrose pseudoinverse of matrix $\bar{V}$.

Here the sigma term $\sigma\left(\Omega, T^2_{\mathcal{S}^n_+}(\bar{V}, H)\right)$ is quadratic in H. This should be not surprising in view that the set S is cone reducible. ■

3.1 Sets defined by constraints

Let us consider now convex sets defined by constraints. Specifically we assume that the set $S = G^{-1}(K)$ is defined as in (2.4) with $K \subset \mathcal{Y}$ being a closed convex cone and the mapping $G : \mathcal{X} \to \mathcal{Y}$ being twice continuously differentiable. In order for the set S to be convex, we need to impose some conditions on the mapping G. Of course, the set S is convex if G is an affine mapping, i.e.,

$$G(x) = a + \mathcal{A}x, \tag{3.12}$$

with $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$ being a linear mapping. More generally we assume that the mapping G is convex with respect to the cone $-K$. That is, for any $x_1, x_2 \in \mathcal{X}$ and $t \in [0,1]$ it holds that

$$G(tx_1 + (1-t)x_2) \succeq_K tG(x_1) + (1-t)G(x_2), \tag{3.13}$$

where $a \succeq_K b$ means that $a - b \in K$. For example let $K := -\mathbb{R}^n_+$ and $G(x) = (g_1(x), ..., g_n(x))$, and hence the constraint $G(x) \in K$ means that $g_i(x) \leq 0$, $i = 1, ..., n$. Then condition (3.13) means that the functions $g_i : \mathcal{X} \to \mathbb{R}$ are convex.

We also make the following assumptions. Let $\bar{x} \in \mathcal{X}$ be the considered point, $\bar{v} := P_S(\bar{x})$ and $\bar{y} := G(\bar{v})$. By the definition we have that $\bar{v} \in S$ and $\bar{y} \in K$.

(A1) The mapping $G : \mathcal{X} \to \mathcal{Y}$ is twice continuously differentiable and convex with respect to the cone $-K$.

(A2) Robinson's constraint qualification holds:

$$DG(\bar{v})\mathcal{X} + T_K(\bar{y}) = \mathcal{Y}. \tag{3.14}$$

(A3) The cone K is second order regular at $\bar{y}$.

Recall that, provided the cone K has a nonempty interior, for convex constraints Robinson's constraint qualification is equivalent to the Slater condition: there exists a point $\hat{x} \in \mathcal{X}$ such that $G(\hat{x}) \in \text{int}(K)$.

By Proposition 2.1 it follows that the set S is also second order regular at $\bar{v}$. Let us calculate the corresponding sigma term $\mathfrak{s}_{\bar{x}}(h)$. Denote $\mathcal{A} := DG(\bar{v})$ (of course, if $G(x) = a + \mathcal{A}x$ is affine as in (3.12), then this holds). We have that (e.g., [3, p.167])

$$T_S^2(\bar{v}, h) = \left\{ w : DG(\bar{v})w + D^2G(\bar{v})(h,h) \in T_K^2(\bar{y}, \mathcal{A}h) \right\}, \tag{3.15}$$

and hence $-\sigma(\bar{x} - \bar{v}, T_S^2(\bar{v}, h))$ is equal to the optimal value of the problem

$$\operatorname*{Min}_{w \in \mathcal{X}} \langle \bar{v} - \bar{x}, w \rangle \ \text{ s.t. } \mathcal{A}w + D^2G(\bar{v})(h,h) \in T_K^2(\bar{y}, \mathcal{A}h). \tag{3.16}$$

The (Lagrangian) dual of problem (3.16) is (e.g., [3, eq. (2.305)])

$$\operatorname*{Max}_{\lambda \in \mathcal{X}} \left\{ \inf_{w \in \mathcal{X}} \langle \bar{v} - \bar{x}, w \rangle + \langle \lambda, \mathcal{A}w + D^2G(\bar{v})(h,h) \rangle - \sigma\big(\lambda, T_K^2(\bar{y}, \mathcal{A}h)\big) \right\}. \tag{3.17}$$

Because of Robinson's constraint qualification (assumption (A2)) optimal values of problems (3.16) and (3.17) are equal to each other.

Equivalently the dual problem (3.17) can be written as

$$\begin{array}{ll} \operatorname*{Max}_{\lambda \in \mathcal{Y}} & \langle \lambda, D^2G(\bar{v})(h,h) \rangle - \sigma\big(\lambda, T_K^2(\bar{y}, \mathcal{A}h)\big) \\ \text{s.t.} & \mathcal{A}^*\lambda = \bar{x} - \bar{v}. \end{array} \tag{3.18}$$

We have (compare with (3.4)) that for $\mathcal{A}h \in T_K(\bar{y})$,

$$T_K^2(\bar{y}, \mathcal{A}h) + T_K(\bar{y}) + \text{sp}(\mathcal{A}h) \subset T_K^2(\bar{y}, \mathcal{A}h) \subset \text{cl}\{T_K(\bar{y}) + \text{sp}(\mathcal{A}h)\},$$

and $T_K(\bar{y}) = \text{cl}\{K + \text{sp}(\bar{y})\}$. Therefore $\sigma\left(\lambda, T_K^2(\bar{y}, \mathcal{A}h)\right) < \infty$ iff $\langle \lambda, \mathcal{A}h \rangle = 0$, $\langle \lambda, \bar{y} \rangle = 0$, $\lambda \in K^-$. Note that if $\mathcal{A}^*\lambda = \bar{x} - \bar{v}$ and $h \in C(\bar{v})$, then $\langle \mathcal{A}^*\lambda, h \rangle = \langle \lambda, \mathcal{A}h \rangle = 0$. It follows that it suffices to maximize in (3.18) over $\lambda \in \Lambda(\bar{x})$, where

$$\Lambda(\bar{x}) := \left\{ \lambda : [DG(\bar{v})]^*\lambda = \bar{x} - \bar{v}, \ \lambda \in K^-, \ \langle \lambda, G(\bar{v}) \rangle = 0 \right\}. \tag{3.19}$$

Note that because of Robinson's constraint qualification the set $\Lambda(\bar{x})$, of Lagrange multipliers, is nonempty and bounded, and hence is compact.

Consequently by Theorem 3.1 we obtain the following result (this can be also derived from the general theory of sensitivity analysis, cf., [2, section 4],[3, section 4.7]).

Theorem 3.2 *Let $S := G^{-1}(K)$, where $K \subset \mathcal{Y}$ is a closed convex cone and suppose that the assumptions* (A1) – (A3) *hold. Then P_S is directionally differentiable at $\bar{x}$ and the directional derivative $P'_S(\bar{x}, d)$ is given by the optimal solution of the problem*

$$\operatorname*{Min}_{h \in C(\bar{v})} \sup_{\lambda \in \Lambda(\bar{x})} \left\{ \|d - h\|^2 + \langle \lambda, D^2 G(\bar{v})(h, h) \rangle - \sigma\big(\lambda, T_K^2(\bar{y}, DG(\bar{v})h)\big) \right\}. \tag{3.20}$$

If the cone K is polyhedral, then the sigma term $\sigma\big(\lambda, T_K^2(\bar{y}, DG(\bar{v})h)\big)$ in (3.20) vanishes. Also if the nondegeneracy condition

$$DG(\bar{v})\mathcal{X} + \operatorname{lin}(T_K(\bar{y})) = \mathcal{Y}. \tag{3.21}$$

(rather than Robinsons constraint qualification (3.14)) holds, then the set of Lagrange multipliers $\Lambda(\bar{x}) = \{\bar{\lambda}\}$ is a singleton. Moreover, if K is cone reducible at $\bar{y}$, then the sigma term is quadratic in h and the set S is cone reducible at $\bar{v}$ (see Proposition 2.2). In that case P_S is differentiable at $\bar{x}$ iff the critical cone $C(\bar{v})$ is a linear space (see Proposition 3.1). If the mapping G is affine, as defined in (3.12), then $D^2G(\bar{v}) = 0$ and hence the term $\langle \lambda, D^2 G(\bar{v})(h, h) \rangle$ in (3.20) vanishes.

For affine mappings we can also formulate the result of Theorem 3.2 in the following framework. Let

$$S := K \cap (L + b), \tag{3.22}$$

where $b \in \mathcal{Y}$ and L is a linear subspace of $\mathcal{Y}$. It can be defined as $L = \{w \in \mathcal{Y} : \mathcal{A}w = 0\}$, where $\mathcal{A} : \mathcal{Y} \to \mathcal{X}$ is a linear mapping. Without loss of generality we can assume that $\mathcal{A}$ is onto and hence its conjugate $\mathcal{A}^* : \mathcal{X} \to \mathcal{Y}$ is one-to-one. Note that $\mathcal{A}^*\mathcal{X} = L^\perp$, where

$$L^\perp := \{y \in \mathcal{Y} : \langle y, w \rangle = 0, \ \forall w \in L\}.$$

Assume that:

(A′1) the Slater condition holds, i.e., the intersection $\operatorname{int}(K) \cap (L + b)$ is nonempty.

(A′2) the cone K is second order regular.

Then the set S is second order regular (cf., [3, Proposition 3.90]). Consider a point $\bar{x} \in \mathcal{Y}$ and let $\bar{y} := P_S(\bar{x})$. Since S is second order regular, it follows that P_S is directionally differentiable at $\bar{x}$ and formula (3.1) holds. The corresponding sigma term

$$\mathfrak{s}_{\bar{x}}(h) := \sigma\big(\bar{x} - \bar{y}, T_S^2(\bar{y}, h)\big), \tag{3.23}$$

can be calculated as follows. We have that $T_S(\bar{y}) = T_K(\bar{y}) \cap L$, and for $h \in T_S(\bar{y})$,

$$T_S^2(\bar{y}, h) = T_K^2(\bar{y}, h) \cap L. \tag{3.24}$$

It follows that $\sigma(\bar{x} - \bar{y}, T_S^2(\bar{y}, h))$ is equal to the optimal value of the problem

$$\operatorname*{Max}_{w \in T_K^2(\bar{y}, h)} \langle \bar{x} - \bar{y}, w \rangle \ \text{ s.t. } \mathcal{A}w = 0. \tag{3.25}$$

The dual of that problem is

$$\operatorname*{Min}_{\lambda\in L^\perp}\ \sup_{w\in T^2_K(\bar y,h)} \langle \bar x-\bar y+\lambda, w\rangle. \tag{3.26}$$

By Slater condition optimal values of problems (3.25) and (3.26) are equal to each other. By (3.4) we have that the minimum in (3.26) is finite iff

$$\bar x-\bar y+\lambda\in K^-,\ \langle \bar x-\bar y+\lambda,\bar y\rangle=0,\ \langle \bar x-\bar y+\lambda,h\rangle=0. \tag{3.27}$$

Hence the sigma term $\mathfrak{s}_{\bar x}(h)$ is given by the optimal value of the problem

$$\begin{array}{ll}\operatorname*{Min}_{\lambda\in L^\perp} & \sigma\big(\bar x-\bar y+\lambda, T^2_K(\bar y,h)\big)\\ \text{s.t.} & \bar x-\bar y+\lambda\in K^-,\ \langle \bar x-\bar y+\lambda,\bar y\rangle=0,\ \langle \bar x-\bar y+\lambda,h\rangle=0.\end{array} \tag{3.28}$$

Example 3.2 Let $\mathcal{Y}:=\mathcal{S}^n$ be the space of $n\times n$ symmetric matrices, $K:=\mathcal{S}^n_+$ be the cone of positive semidefinite matrices and

$$S:=\mathcal{S}^n_+\cap(L+b), \tag{3.29}$$

where L is a linear subspace of $\mathcal{S}^n$. Assuming Slater condition we have that metric projection P_S is directionally differentiable at any $\bar X\in\mathcal{S}^n$ and $P'(\bar X,D)=\bar H$, where $\bar H$ is the optimal solution of the problem

$$\begin{array}{lll}\operatorname*{Min}_{H\in T_{\mathcal{S}^n_+}(\bar Y)\cap L} & \sup_{\Lambda\in L^\perp} & \left\{\|D-H\|^2-2\mathrm{Tr}[(\bar X-\bar Y+\Lambda)H\bar Y^\dagger H]\right\}\\ & \text{s.t.} & \bar X-\bar Y+\Lambda\preceq 0, \mathrm{Tr}[(\bar X-\bar Y+\Lambda)\bar Y]=0,\\ & & \mathrm{Tr}[(\bar X-\bar Y+\Lambda)H]=0,\end{array} \tag{3.30}$$

where $\bar Y:=P_S(\bar X)$.

References

[1] Bonnans, J.F. and Shapiro, A., Nondegeneracy and quantitave stability of parameterized optimization problems with multiple solutions, *SIAM J. Optimization*, 8, 940-946, 1998.

[2] Bonnans, J.F., Cominetti, R. and Shapiro, A., Sensitivity analysis of optimization problems under second order regular constraints, *Mathematics of Operations Research*, 23, 806-831, 1998.

[3] Bonnans, J.F. and Shapiro, A., *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[4] Cominetti, R. Metric regularity, tangent sets and second order optimality conditions, *Applied Math. & Opt.*, 21, 265–287, 1990.

[5] Kruskal, J., Two convex counterexamples: a discontinuous envelope function and a non-differentiable nearest point mapping, *Proc. Amer. Math. Soc.*, 23, 697-703, 1969.

[6] Robinson, S.M., Local structure of feasible sets in nonlinear programming. II. Nondegeneracy, *Math. Programming Stud.*, 22, 217-230, 1984.

[7] Shapiro, A., On Concepts of Directional Differentiability, *Journal of Optimization Theory and Applications*, 66, 447-487, 1990.

[8] Shapiro, A., Directionally Nondifferentiable Metric Projection, *Journal of Optimization Theory and Applications*, 81, 203-204, 1994.

[9] Shapiro, A., Sensitivity analysis of generalized equations, *Journal of Mathematical Sciences*, 115, 2554-2565, 2003.

[10] Zarantonello, E.H., Projections on convex sets in Hilbert space and spectral theory, In contributions to *Nonlinear Functional Analysis*, Academic Press, New York, 1971, pp. 237-424.