Optimization
Vol. 61, No. 9, September 2012, 1087–1098

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

Generalized isotone projection cones

O.P. Ferreira[a] and S.Z. Németh[b]*

[a]*IME/UFG, Campus II – Caixa Postal 131, Goiânia, GO 74001-970, Brazil;* [b]*School of Mathematics, The University of Birmingham, The Watson Building, Edgbaston, Birmingham B15 2TT, United Kingdom*

(*Received 9 June 2010; final version received 2 November 2010*)

This article extends the notion of isotone projection cones to generalized isotone projection cones by replacing the usual metric projection with a generalized one. It is shown that all such cones are simplicial.

Keywords: projections; isotone projection cone; convex functions

AMS Subject Classifications: 90C25; 47H99

1. Introduction

A projection onto a convex set is obtained by minimizing the distance function on the considered convex set. More specifically, given a distance d and a nonempty closed convex set C in the Euclidean space $\mathbb{R}^n$ the projection of the point $x \in \mathbb{R}^n$ onto C with respect to d is the set

$$P_C(x) := \{p \in C : d(x, p) \leq d(x, y), \ \forall y \in C\}. \tag{1}$$

Thus, for each fixed distance function we have an associated projection. This view point allows the study of the notion of projection in several different contexts. Works dealing with projections associated to different distances include Carrizosa and Plastria [6], Mangasarian [19,20], Dax [9,10] and Plastria and Carrizosa [28].

The interest in the subject of projection arises in several situations, having a wide range of applications in pure and applied mathematics such as Functional Analysis (see, e.g. [33]), Convex Analysis (see, e.g. [14]), Optimization (see, e.g. [1]), Numerical Linear Algebra (see, e.g. [31]), Statistics (see, e.g. [3,12]), Computer Graphics (see, e.g. [13]) and other fields.

Several distance functions defined in Euclidean spaces are differentiable except at the origin and are convex. For instance, the distance associated to the Euclidean norm $\| \ \|$ is defined by

$$d(x, y) = \|x - y\|, \quad x, y \in \mathbb{R}^n,$$

*Corresponding author. Email: nemeths@for.mat.bham.ac.uk

ISSN 0233–1934 print/ISSN 1029–4945 online
© 2012 Taylor & Francis
http://dx.doi.org/10.1080/02331934.2010.538057
http://www.tandfonline.com

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

and therefore, $d(x,\cdot):\mathbb{R}^n\to\mathbb{R}$, the distance function from the point x, is differentiable except at the origin and its convexity is a consequence of the positive homogeneity and triangle inequality property of the norm. Since the projection of the point $x\in\mathbb{R}^n$ onto a convex set C is the minimizer set of the distance function from the point x to the convex set C (1), it is natural to study the concept of projection as the minimizer set of positive convex functions which are differentiable except at the origin. This extension facilitates obtaining several important properties of the projection associated to different distances in a unified manner.

The main goal of this article is to generalize the notion of the isotone projection cone. The introduction and investigation of the isotone projection cone is originated in complementarity problems which are widely investigated useful models in optimization, economics, physics and engineering.

Let $K\subset\mathbb{R}^n$ be a generating pointed closed convex cone. If K° is the polar of K (see the definitions in Sections 2.1 and 3) and $F:\mathbb{R}^n\to\mathbb{R}^n$ is a mapping, then the complementarity problem defined by F and K consists of getting an element $x\in\mathbb{R}^n$ satisfying

$$(NCP)\quad x\in K,\ F(x)\in -K^\circ,\ \langle x, F(x)\rangle = 0.$$

Denoting by P_K the projection mapping onto K [33], it turns out that x is a solution of (NCP) if and only if it is the solution of the fixed point problem

$$(FIX)\quad x = P_K(x - F(x)).$$

(For a proof see, e.g. Proposition 6.1 in [17].)

The equivalence of problems (NCP) and (FIX) emphasizes the importance of studying the properties of projection mappings onto cones and finding efficient methods of projecting onto cones. This was the main motivation for introducing and analysing the isotone projection cones; that is, the cones K whose projection P_K is isotone with respect to the order relation induced by K. (A cone is isotone projection, if and only if its polar is generated by rays forming non-acute angles [15–17].)

For isotone projection cones K and continuous monotone mappings F several iterative methods were developed for finding the solutions of (FIX) and equivalently of (NCP) [18,24].

The drawback of these methods is that at each step of iteration one needs to project a different point onto the cone. The usual projection methods (see, e.g. the Dykstra algorithm [11,12,30]) are rather slow (see the numerical results in [22] and the remark preceding Section 6.3 in [21]). Therefore, the successive projection onto the cone (by using classical methods) makes the above cited iterative algorithms inefficient.

But the projection of a point onto an isotone projection cone can be reduced to a finite number of projections onto subspaces of decreasing dimension [23]. This recent result completes some earlier results in [18,24] concerning the solution of (NCP) problems for isotone projection cones with a cheep, easily implementable numerical method.

The monotone nonnegative cone ([8], Section 2.13.9.4.2) is a particular isotone projection cone. It is easy to apply the above mentioned algorithm (see its simplified version due to Dattorro in <http://www.convexoptimization.com/wikimization/index.php/Projection_on_Polyhedral_Cone> to compute projections onto monotone nonnegative cones which occur in various reconstruction problems ([8], Section 5.13). This is another application of the isotone projection cones. It shows that the isotone projection cones and the projections onto them are subjects of interest for various applications.

It turns out that isotone projection cones are simplicial cones of a particular form. Their investigation becomes part of the study of latticially ordered Euclidean (and Hilbert) spaces. Besides the theoretical interest in seeking the relation of the ordering and the geometry of the space, there is a real hope that the generalization of the isotone projection cones will be useful in further applications. We mention, in the line of developing a wide theoretical background for these questions, the recent results of the second author [26] on the so-called isotone retraction cones (see, e.g. Theorem 8). However, isotone retraction is merely a topological rather than a geometric notion. Hence a more geometric notion would be in a stronger relation with the earlier theory. A first step towards this is to consider various non-Euclidean projections. This is the subject of this article note which has, at this stage, merely a theoretic character.

Our main result is that a generalized isotone projection cone, that is, a generating closed pointed cone in $\mathbb{R}^n$ which admits an isotone generalized projection, is simplicial (Theorem 9). Moreover, every simplicial cone is a generalized isotone projection cone with respect to an appropriate norm (Theorem 10).

Although, by using results from convex analysis, more general properties of the projection mapping associated to different distances can be obtained (extending some of the results of [33] in the Euclidian case), we present here only the results which are needed to achieve our main task. Although most of our results hold in general Hilbert spaces too, for simplicity of the ideas we considered presenting them in Euclidean spaces only.

It is worth remarking that in the applications several other distance-like notions and projections were considered (see, e.g. proximity mappings, Bregman divergence [5], Kullback–Leibler divergence, Csiszár's f-divergence [7] etc.) Most of them are also generalizations of the Euclidean *distance*. Our approach is different since it focuses on the generalization of the Euclidean *norm* (hence it is translation invariant in contrast with some of the above mentioned extensions). Of course, various questions occur with respect to the behaviour of these projections which constitute topics for the geometry of normed spaces, abstract best approximation theory (in the case of norms the history goes back to the nineteenth century) etc. Our special view consists in the fact that *we are searching for connections between the order relation induced by a cone and the generalized projection onto that cone*.

The structure of this article is as follows. In Section 2 we fix some notations needed later and define the distance and the projection with respect to a convex function which is differentiable except at the origin. We also present several examples of convex functions which generate distances and projection mappings and establish several properties of the distance and the projection mapping onto a convex set and, in particular, onto a convex cone. In Section 3 we introduce the notion of isotone projection cone with respect to a convex function and establish that all such isotone projection cones are simplicial and present some examples of generalized

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

isotone projection cones. We conclude this article by making some final remarks in Section 4.

2. Generalized projection

In this section we define the distance and the projection with respect to a convex function. We will establish the properties of the distance and the projection onto a convex set (in particular, onto a convex cone), necessary for establishing the results in Section 3.

Denote by $\langle , \rangle$ the canonical scalar product in $\mathbb{R}^n$ and by $\| \ \|$ the associated Euclidean norm. A set $C \subset \mathbb{R}^n$ is called *convex* if $\lambda x + (1-\lambda)y \in C$ for all $0 \le \lambda \le 1$ and each $x, y \in C$. A function $f: \mathbb{R}^n \to \mathbb{R}$ is called *convex* if for each $0 \le \lambda \le 1$ and each $x, y \in \mathbb{R}^n$ we have

$$f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y).$$

For a function $f: \mathbb{R}^n \to \mathbb{R}$ differentiable at $x \in \mathbb{R}^n$, the vector $\nabla f(x) \in \mathbb{R}^n$ denotes its gradient at $x \in \mathbb{R}^n$.

Throughout this article we suppose that $\varphi: \mathbb{R}^n \to \mathbb{R}$ is a convex function, differentiable in $\mathbb{R}^n \backslash \{0\}$; and satisfying the following four conditions:

H1 $\varphi(0) = 0$;
H2. $\varphi(x) \ge 0$ for all $x \in \mathbb{R}^n$;
H3. For any $C \subset \mathbb{R}^n$ nonempty closed convex set and any $x \in \mathbb{R}^n$ the function $C \ni y \mapsto \varphi(x-y)$ has exactly one minimizer;
H4. $\varphi(x) = \varphi(-x)$, for all $x \in \mathbb{R}^n$.

We remark that φ is also continuous, because every convex function on $\mathbb{R}^n$ is continuous (see, Theorem 10.1, p. 82 of [29]). In the next two remarks, we will show that the assumptions above are not too restrictive; that is, there exist many functions satisfying that assumptions, including many differentiable norms in $\mathbb{R}^n \backslash \{0\}$.

Remark 1 A convex set C is strictly convex if its boundary ∂C does not contain any line segment. Formally this means that for each $x, y \in \partial C$ with $x \ne y$ there is no $0 < \lambda < 1$ such that $\lambda x + (1-\lambda) y \in \partial C$. *A* function $f: \mathbb{R}^n \to \mathbb{R}$ is called strictly convex if for each $0 < \lambda < 1$ and each x, $y \in \mathbb{R}^n$ with $x \ne$ y we have

$$f(\lambda x + (1-\lambda) y) < \lambda f(x) + (1-\lambda) f(y).$$

A function $f: \mathbb{R}^n \to \mathbb{R}$ is called strongly quasiconvex if for each $0 < \lambda < 1$ and each x, $y \in \mathbb{R}^n$ with $x \ne y$ we have

$$f(\lambda x + (1-\lambda) y) < \max\{f(x), f(y)\}.$$

All strictly convex functions are strongly quasiconvex, but there exist strongly quasiconvex functions that are not strictly convex (for e.g. the Euclidean norm). It follows from the definition that the function φ is strongly quasiconvex if and only if all its nonempty sublevel sets $\{x \in \mathbb{R}^n : \varphi(x) \le L\}$ for $L \in \mathbb{R}$, are strictly convex (remember that φ is continuous). Moreover, for each $x \in \mathbb{R}^n$ the function $C \ni y \mapsto \varphi(x-y)$ has only one minimizer (see Theorem 3.5.9 of [2]), that is, φ satisfy H3.

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

Remark 2 Rockafellar [29] called a convex and positively homogeneous function $\varphi : \mathbb{R}^n \to \mathbb{R}$ satisfying conditions H1, H2 a gauge function. Note that condition H4 together with positive homogeneity is equivalent to $\varphi(\lambda x) = |\lambda|\varphi(x)$, for all $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. All convex functions φ which are positively homogeneous are subadditive, that is, $\varphi(x+y) \le \varphi(x) + \varphi(y)$ for all $x, y \in \mathbb{R}^n$ (see, Theorem 4.7, p. 30 of [29]). Thus, if φ is positive except at the origin, satisfies H1, H2, H4 and is positively homogeneous, then φ is a norm (see, p. 131 of [29]).

Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set. The *distance function* $d^{\varphi}(\cdot, C) : \mathbb{R}^n \to \mathbb{R}$ *to* C *with respect to* φ is defined by

$$d^{\varphi}(x\,, C) := \min\{\varphi(x - y) : y \in C\}, \tag{2}$$

and the *projection mapping* $P_C^{\varphi}(.) : \mathbb{R}^n \to C$ *with respect to* φ *onto the set* C is defined by

$$P_C^{\varphi}(x) := \operatorname{argmin}\{\varphi(x - y) : y \in C\}. \tag{3}$$

Conditions H1 and H2 imply, in particular, that 0 is a minimizer of φ. By using conditions H1, H2 and H3, we also have that $P_C^{\varphi}(z) = z$ if and only if $z \in C$. Moreover, condition H3 implies that the distance function to C and the projection mapping onto C are well defined. Therefore, by using the previous two equalities, it is easy to conclude that

$$d^{\varphi}(x, C) = \varphi(x - P_C^{\varphi}(x)), \quad \forall x \in \mathbb{R}^n. \tag{4}$$

For each $x \in \mathbb{R}^n$ we shall denote by $x_1, \dots, x_n$ the components of x with respect to the canonical basis. Let $p \in (1, +\infty)$ and $\eta_p : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\eta_p(x) = (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$$

It is known that η_p is a norm and a strongly quasiconvex function (see Proposition 7.3.2 on P. 186 of [27]).

Corollary 1 *The function* $\zeta_{p,q} : \mathbb{R}_n \to \mathbb{R}$ *defined by*

$$\zeta_{p,q}(x) = (|x_1|^p + \cdots + |x_n|^p)^{\frac{q}{p}},$$

where $p, q \in (1, +\infty)$ *is strongly quasiconvex.*

Proof First note that the functions $\zeta_{p,q}$ and η_p have the same sub-level sets. Indeed, take $L \in \mathbb{R}$. If $L < 0$, then $\{x \in \mathbb{R}^n : \zeta_{p,q}(x) \le L\} = \{x \in \mathbb{R}^n : \eta_p(x) \le L\} = \varnothing$. Now, if $L \ge 0$ then

$$\{x \in \mathbb{R}^n : \zeta_{p,q}(x) \le L\} = \left\{x \in \mathbb{R}^n : \eta_p(x) \le L^{\frac{1}{q}}\right\}.$$

The definition of strongly quasiconvex functions implies that a continuous function is strongly quasiconvex if and only if its sub-level sets are strictly convex. Since the functions $\zeta_{p,q}$ and η_p have the same sub-level sets, the result follows from the strong quasiconvexity of η_p. ∎

Proposition 2 *Let* Q *be an* $n \times n$ *positive definite symmetric matrix. The function* $\eta_Q : \mathbb{R}^n \to \mathbb{R}$ *defined by*

$$\eta_Q(x) = \sqrt{\langle Qx, x\rangle}$$

is strongly quasiconvex.

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

Proof It is easy to see that the function $\zeta(x) = \eta_Q^2(x)$ is strictly convex and, in particular, strongly quasiconvex. Since ζ is strongly quasiconvex, ζ and η_Q^2 has the same sub-level sets and a continuous function is strongly quasiconvex if and only if its sub-level sets are strictly convex, it follows that η_Q is also strongly quasiconvex. ■

It is easy to see that the function η_Q of Proposition 2 is subadditive, positive except at the origin and positively homogeneous. So, from Remark 2 it follows that it is a norm. In general, if $\| \|'$ is a norm in $\mathbb{R}^n$ then the function $\eta(x) = \|x\|'$ is convex, positive except at the origin and satisfies conditions H1, H2 and H4, but not necessarily H3. The convex function in Corollary 1 is differentiable at nonzero points, satisfies conditions H1–H4, but it is not positively homogeneous (and therefore it is not a norm) and if $p < q$, then it is differentiable everywhere.

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

2.1. *Properties of the generalized projection*

In this section we present some basic properties of the distance and the projection with respect to a convex function necessary for establishing the results in Section 3. The next result is an extension of Proposition 6.1.4 of [27] (p. 161) and has a similar proof. For the sake of completeness we will present this proof here.

LEMMA 3 *Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set. Then, the distance function $d^\varphi(\cdot, C)$ is convex.*

Proof Let $\epsilon > 0$, $\lambda \in [0, 1]$ and $x, y \in \mathbb{R}^n$. The definition of the distance in (2) implies that there exist $\bar{x}, \bar{y} \in C$ such that

$$\varphi(x - \bar{x}) \le d^\varphi(x, C) + \epsilon, \quad \varphi(y - \bar{y}) \le d^\varphi(y, C) + \epsilon.$$

Since $\lambda \in [0, 1]$, the convexity of φ and last two inequalities imply that

$$\varphi(\lambda x + (1 - \lambda) y - (\lambda\bar{x} + (1 - \lambda)\bar{y})) \le \lambda d^\varphi(x, C) + (1 - \lambda) d^\varphi(y, C) + \epsilon.$$

Since C is a convex set and $\bar{x}, \bar{y} \in C$, we get $\lambda\bar{x} + (1 - \lambda)\bar{y} \in C$. Hence, again using the definition of the distance and the last inequality we obtain

$$d^\varphi(\lambda x + (1 - \lambda) y, C) \le \lambda d^\varphi(x, C) + (1 - \lambda) d^\varphi(y, C) + \epsilon.$$

As the above inequality holds for any $\epsilon > 0$ and $\lambda \in [0, 1]$, the result follows. ■

PROPOSITION 4 *Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set. Then, the projection P_C^φ is continuous.*

Proof Let $\{x^k\} \subset \mathbb{R}^n$ such that $\lim_{k \to +\infty} x^k = \bar{x}$. Hence, $\{x^k\} \subset \mathbb{R}^n$ is bounded. Using equation (4) we have

$$d^\varphi(x^k, C) = \varphi(x^k - P_C^\varphi(x^k)) \quad \forall k. \tag{5}$$

Since $d^\varphi(\cdot, C)$ is convex, in particular, it is continuous (see, Theorem 10.1, p. 82 of [29]). Thus,

$$\lim_{k \to +\infty} d^\varphi(x^k, C) = d^\varphi(\bar{x}, C).$$

Now, fix $z \in C$. Thus using (2), (5) we have

$$\{x^k - P_C^{\varphi}(x^k)\} \subset A = \{x \in \mathbb{R}^n : \varphi(x) \leq L\}, \quad L = \sup\{\varphi(x^k - z) : k = 0, 1, \ldots\}.$$

Assumptions H1, H2 and H3 imply that φ has zero as a unique minimizer and, in particular, its minimizer set is bounded. Since φ is convex and has the minimizer set bounded, by using Proposition 2.3.1 of [4] we conclude that the sublevel set A is bounded. As a consequence, the sequence $\{x^k - P_C^{\varphi}(x^k)\}$ is also bounded. By using the triangle inequality we get

$$\|P_C^{\varphi}(x^k)\| \leq \|x^k\| + \|x^k - P_C^{\varphi}(x^k)\|.$$

Hence, the sequence $\{P_C^{\varphi}(x^k)\}$ is bounded because $\{x^k\}$ and $\{x^k - P_C^{\varphi}(x^k)\}$ are bounded. Let $\bar{y}$ be a cluster point of $\{P_C^{\varphi}(x^k)\}$ and let $\{x^{k_j}\}$ be such that $\lim_{k \to +\infty} P_C^{\varphi}(x^{k_j}) = \bar{y}$. Therefore, (5) implies

$$d^{\varphi}(x^{k_j}, C) = \varphi(x^{k_j} - P_C^{\varphi}(x^{k_j})) \quad \forall\, k_j.$$

Tending with j to infinity we have $d^{\varphi}(\bar{x}, C) = \varphi(\bar{x} - \bar{y})$. Since C is closed, it follows that $\bar{y} \in C$, which together with (2), (3) and $d^{\varphi}(\bar{x}, C) = \varphi(\bar{x} - \bar{y})$ imply that $\bar{y} = P_C^{\varphi}(\bar{x})$, because the function $C \ni y \mapsto \varphi(\bar{x} - y)$ has only one minimizer. Therefore, the sequence $\{P_C^{\varphi}(x^k)\}$ has only one cluster point, namely, $P_C^{\varphi}(\bar{x})$. Thus, $\lim_{k \to +\infty} P_C^{\varphi}(x^k) = P_C^{\varphi}(\bar{x})$ and the proof is concluded. ■

By using results of convex analysis, it is possible to prove a converse of the next theorem. However, its result is enough for our purpose.

Proposition 5 *Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set. Then,*

$$\left\langle \nabla\varphi(x - P_C^{\varphi}(x)), y - P_C^{\varphi}(x) \right\rangle \leq 0 \quad \forall y \in C,\ \forall x \in \mathbb{R}^n \backslash C.$$

Proof Let $x \in \mathbb{R}^n \backslash C$ and $y \in C$. Since C is convex we have $ty + (1-t)P_C^{\varphi}(x) \in C$ for all $t \in (0, 1)$. Hence, the definition of the projection $P_C^{\varphi}(x)$ in (3) implies that

$$0 \leq \varphi\big(x - [ty + (1-t)P_C^{\varphi}(x)]\big) - \varphi\big(x - P_C^{\varphi}(x)\big) \quad \forall t \in (0, 1).$$

Simple algebraic manipulation give us $x - [ty + (1-t)P_C^{\varphi}(x)] = x - P_C^{\varphi}(x) - t [y - P_C^{\varphi}(x)]$, which together with the last inequality, imply

$$0 \leq \frac{\varphi\big(x - P_C^{\varphi}(x) - t(y - P_C^{\varphi}(x))\big) - \varphi\big(x - P_C^{\varphi}(x)\big)}{t} \quad \forall t \in (0, 1).$$

As $x \in \mathbb{R}^n \backslash C$ we conclude that $x - P_C^{\varphi}(x) \neq 0$. Hence, letting t goes to 0 in the latter inequality and taking into account that φ is differentiable in $\mathbb{R}^n \backslash \{0\}$ we obtain

$$0 \leq \left\langle \nabla\varphi(x - P_C^{\varphi}(x)),\ -[y - P_C^{\varphi}(x)] \right\rangle,$$

which is equivalent to the statement of the proposition. ■

A closed set $K \subset \mathbb{R}^n$ is called a *closed convex cone* if $\lambda x \in K$ and $x + y \in K$ for all $\lambda > 0$ and all x, $y \in K$. Let $K \subset \mathbb{R}^n$ be a closed convex cone. The *polar cone* of K is the set

$$K^{\circ} := \{x \in \mathbb{R}^n \mid \langle x, y \rangle \leq 0\ \forall\, y \in K\}.$$

Then, we have the following.

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

COROLLARY 6 *Let $K \subset \mathbb{R}^n$ be a closed convex cone and $x \in \mathbb{R}^n \backslash K$. Then,*

$$\nabla\varphi(x - P_K^{\varphi}(x)) \in K^{\circ}, \quad \langle \nabla\varphi(x - P_K^{\varphi}(x)), P_K^{\varphi}(x)\rangle = 0. \tag{6}$$

Proof Let $x \in \mathbb{R}^n \backslash K$. From Proposition 5 it follows that

$$\left\langle \nabla\varphi(x - P_K^{\varphi}(x)), y - P_K^{\varphi}(x)\right\rangle \leq 0 \quad \forall y \in K. \tag{7}$$

Since K is a cone, in the last inequality we can replace y with λy, where $\lambda > 0$ to obtain

$$\left\langle \nabla\varphi(x - P_K^{\varphi}(x)), y - P_K^{\varphi}(x)/\lambda\right\rangle \leq 0 \quad \forall y \in K.$$

Thus, letting $\lambda \to +\infty$ in last inequality we obtain $\left\langle \nabla\varphi(x - P_K^{\varphi}(x)), y\right\rangle \leq 0$, for all $y \in K$, which implies the inclusion in (6). For proving the equality in (6), we use (7) with $y = 0$ and $y = 2P_K^{\varphi}(x)$ together with the assumption that K is a cone. ■

The next lemma is an immediate consequence of H4 and the property of the derivative.

LEMMA 7 *Assume that φ satisfies* H4. *Then, $\nabla\varphi(x) = -\nabla\varphi(-x)$ for all $x \in \mathbb{R}^n \backslash \{0\}$.*

3. Generalized isotone projection cones

In this section we introduce the notion of isotone projection cones with respect to a generalized projection, which extends the definition introduced in [15]. We also establish that all isotone projection cones with respect to a generalized projection are simplicial. We begin with some important definitions.

Let $K \subset \mathbb{R}^n$ be a closed convex cone. The cone K is called *generating* if $\mathbb{R}^n = K - K$ and K is called *pointed* if $K \cap (-K) = \{0\}$. Let K be a pointed convex cone. We set $x \leq y$ whenever $y - x \in K$. Thus, $\leq$ is an order relation; that is, a reflexive, antisymmetric and transitive relation. If for any two points x, $y \in \mathbb{R}^n$ there exists $\sup\{x, y\}$ in the ordered vector space $(\mathbb{R}^n, \leq)$, then $(\mathbb{R}^n, \leq)$ will be called a *vector lattice*. Let $e^1, \ldots, e^n$ be n linearly independent vectors in R^n. Then, the cone

$$K = \{\lambda_1 e^1 + \cdots + \lambda_n e^n : \lambda_1 \geq 0, \ldots, \lambda_n \geq 0\} \subset \mathbb{R}^n,$$

is called *simplicial* and $e^1, \ldots, e^n$ are called the *generators* of K. It is known that $(\mathbb{R}^n, \leq)$ is a vector lattice if and only if K is a simplicial cone (see, e.g. Lemma 3 of [32]).

Let $\|\cdot\|$ be the Euclidean norm in $\mathbb{R}^n$, the set $K \subset \mathbb{R}^n$ be a pointed convex cone and $\leq$ be the order induced by K. The cone K is called *normal* if there exists a constant $\gamma > 0$ such that $\gamma\|x\| \leq \|y\|$ for all x, $y \in \mathbb{R}^n$ with $0 \leq x \leq y$.

Let $K \subset \mathbb{R}^n$ be a generating pointed closed convex cone. A mapping $F: \mathbb{R}^n \to \mathbb{R}^n$ is called *isotone* if $F(x) \leq F(y)$ for all $x, y \in K$ with $x \leq y$. The cone K is called a *φ-isotone projection cone* if $P_K^{\varphi}(x) \leq P_K^{\varphi}(y)$ for all $x, y \in \mathbb{R}^n$ with $x \leq y$; that is, the projection mapping onto the set K with respect to φ is isotone. The cone K is called a *generalized isotone projection cone* if there is a φ such that K is a *φ-isotone projection cone*.

Let $K \subset \mathbb{R}^n$ be a generating pointed closed convex cone. A mapping $\Psi : \mathbb{R}^n \to \mathbb{R}^n$ is called a *retraction* onto K if $\Psi(x) \in K$ for all $x \in \mathbb{R}^n$ and $\Psi(u) = u$ for all $u \in K$. A mapping Ψ is called *sharp* if $\Psi(0) = 0$ and $\mathrm{Im}(\Psi) \cap \mathrm{Im}(-\Psi) = \{0\}$.

The following theorem is proved in [25] (see also [26]).

Theorem 8 *Let $K \subset \mathbb{R}^n$ be a pointed closed convex generating normal cone. Then, K is simplicial if and only if there is a continuous isotone retraction $\Psi : \mathbb{R}^n \to K$ such that the complement $I - \Psi$ of Ψ is sharp.*

In [15] it was shown that all isotone projection cones are simplicial. The next theorem extends this result for a projection mapping with respect to a convex function φ which is differentiable at nonzero points and satisfies conditions H1–H4, and therefore is more general than the Euclidian norm. Any norm whose sublevel sets are strictly convex can be considered here, but φ does not have to be a norm. For example, φ can be the function given in Corollary 1 which is not a norm.

Theorem 9 *Let $K \subset \mathbb{R}^n$ be a pointed closed convex generating normal cone. If K is a φ-isotone projection cone, then K is simplicial.*

Proof Let K be a φ-isotone projection cone. It is easy to see that P_K^{φ} is an isotone retraction and, from Proposition 4, is also continuous. By Theorem 8, to conclude the proof it is enough to show that $I - P_K^{\varphi}$ is a sharp mapping. Let x, y such that

$$x - P_K^{\varphi}(x) = P_K^{\varphi}(y) - y. \tag{8}$$

We trivially have $(I - P_K^{\varphi})(0) = 0$. Thus, by the definition of sharp mapping and the last equality, it is sufficient to show that $x - P_K^{\varphi}(x) = 0$. Assume to the contrary that $x - P_K^{\varphi}(x) \neq 0$. So, combining (8) and Lemma 7 we obtain that

$$\nabla\varphi\big(x - P_K^{\varphi}(x)\big) = -\nabla\varphi\big(y - P_K^{\varphi}(y)\big).$$

On the other hand, Corollary 6 implies that $\nabla\varphi(x - P_K^{\varphi}(x))$ and $\nabla\varphi(y - P_K^{\varphi}(y))$ belong to the set K°. Therefore, the latter equality implies that

$$\nabla\varphi(x - P_K^{\varphi}(x)) \in K^{\circ} \cap (-K^{\circ}) = \{0\}.$$

Since $\nabla\varphi(x - P_K^{\varphi}(x)) = 0$ and φ is convex, we conclude that $x - P_K^{\varphi}x$ is a minimizer of φ. But 0 is the unique minimizer of φ (by H3). Thus, $x - P_K^{\varphi}x = 0$ which is a contradiction. Therefore, K is simplicial. ■

The next theorem is a kind of reciprocal of Theorem 9.

Theorem 10 *For every simplicial cone $K \subset \mathbb{R}^n$, there exists a norm η in $\mathbb{R}^n$, that comes from a scalar product, such that K is an η-isotone projection cone.*

Proof Let $e^1, \dots, e^n$ be the generators of K. Any $x \in \mathbb{R}^n$ can be uniquely written as

$$x = x_1^e e^1 + \cdots + x_n^e e^n,$$

where $x_1^e, \dots, x_n^e$ are the coordinates of x with respect to the basis $\{e^1, \dots, e^n\}$. Let $\langle \cdot, \cdot \rangle_e$ be the scalar product defined by

$$\langle x, y \rangle_e := x_1^e y_1^e + \cdots + x_n^e y_n^e,$$

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

where $x = x_1^e e^1 + \cdots + x_n^e e^n$ and $y = y_1^e e^1 + \cdots + y_n^e e^n$, and let η be the norm defined by

$$\eta(x) := \sqrt{\langle x, x \rangle_e}.$$

Hence, the basis $\{e^1, \ldots, e^n\}$ is orthonormal with respect to the scalar product $\langle \cdot, \cdot \rangle_e$. Thus, K is the nonnegative orthant with respect to the scalar product $\langle \cdot, \cdot \rangle_e$. Moreover, it is easy to see that

$$P_K^\eta(x) = \big(\max\{x_1^e, 0\}, \ldots, \max\{x_n^e, 0\}\big),$$

where $x = x_1^e e^1 + \cdots + x_n^e e^n$ and, as a consequence, for $x \le y$ we have $P_K^\eta(x) \le P_K^\eta(y)$. Therefore, K is an η-isotone projection cone. ■

THEOREM 11 *The nonnegative orthant $\mathbb{R}_+^n$ is an η_p-isotone projection cone for all $p \in (1, +\infty)$.*

Proof Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$. The definition of η_p on $\mathbb{R}^n$ gives

$$\eta_p(x-y) := (|x_1 - y_1|^p + \cdots + |x_n - y_n|^p)^{\frac{1}{p}}, \quad p \in (1, +\infty).$$

On the other hand, it is easy to see that $\max\{a, 0\} = \operatorname{argmin}\{|a-t|^p : t \ge 0\}$, for $p \in (1, +\infty)$. Since the function $0 \le t \mapsto t^{\frac{1}{p}}$ is increasing for $p \in (1, +\infty)$, letting $x^+ = (x_1^+, \ldots, x_n^+)$ and $x_i^+ = \max\{x_i, 0\}$, for $i = 1, \ldots, n$, we have

$$\begin{aligned} \eta_p(x-y) &= (|x_1 - y_1|^p + \cdots + |x_n - y_n|^p)^{\frac{1}{p}} \\ &\ge \big(|x_1 - x_1^+|^p + \cdots + |x_n - x_n^+|^p\big)^{\frac{1}{p}} = \eta_p(x - x^+) \end{aligned} \tag{9}$$

for all $y \in \mathbb{R}_+^n$ and $p \in (1, +\infty)$. We have already remarked that η_p is strongly quasiconvex. Therefore, letting $K = \mathbb{R}_+^n$ it follows that $P_K^{\eta_p}$ is a single valued mapping for any $p \in (1, +\infty)$. Hence, by using (9) we get $P_K^{\eta_p}(x) = x^+$. So, for $x \le y$ we have $P_K^{\eta_p}(x) \le P_K^{\eta_p}(y)$, for all $p \in (1, +\infty)$. Therefore, $\mathbb{R}_+^n$ is an η_p-isotone projection cone, for all $p \in (1, +\infty)$. ■

The following corollary follows from Theorem 11, by replacing the canonical basis of $\mathbb{R}^n$ with the basis $\{e^1, \ldots, e^n\}$ and the canonical scalar product $\langle \cdot, \cdot \rangle$ of $\mathbb{R}^n$ with the scalar product $\langle \cdot, \cdot \rangle_e$ defined by

$$\langle x, y \rangle_e = x_1^e y_1^e + \cdots + x_n^e y_n^e,$$

where $x = x_1^e e^1 + \cdots + x_n^e e^n$ and $y = y_1^e e^1 + \cdots + y_n^e e^n$.

COROLLARY 12 *Let $K \subset \mathbb{R}^n$ be a simplicial cone generated by the linearly independent vectors $e^1, \ldots, e^n$. For any $x \in \mathbb{R}^n$ consider the unique decomposition $x = x_1^e e^1 + \cdots + x_n^e e^n$. Let η_p^e be the norm in $\mathbb{R}^n$ defined by*

$$\eta_p^e(x) := \big(|x_1^e|^p + \cdots + |x_n^e|^p\big)^{\frac{1}{p}}, \quad p \in (1, +\infty).$$

Then, K is an η_p^e-isotone projection cone for any $p \in (1, +\infty)$.

It can be similarly shown that Theorem 11 and Corollary 12 hold for the function $\zeta_{p,q}$ defined in Corollary 1 which is not a norm.

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

4. Final remarks

In general, computing the projection of a point onto a convex cone is a difficult and computationally expensive problem. Németh and Németh [23] have shown that the projection of a point onto an isotone projection cone in $\mathbb{R}^n$ can be obtained by projecting recursively at most $n-1$ times into subspaces of decreasing dimensions. In Section 3 we introduced the concept of isotone projection cones with respect to much more general convex functions than the euclidean norm and showed that all such cones are simplicial. So, a natural and logical question is: How to extend the ideas of [23] for this more general context. It would be also interesting to see whether Theorem 9 holds if we remove the differentiability and the uniqueness in the assumption H3 in order to include the 1-norm and $+\infty$-norm (i.e. the p-norm when $p \to +\infty$) too. In this case we remark that Theorem 11 and Corollary 12 can be extended for $p=1$ and $p \to \infty$ too.

Downloaded by [O.P. Ferreira] at 06:25 01 January 2013

Acknowledgements

The authors express their gratitude to the referee for the many helpful comments which led to a substantially improved version of this article. This work was done while Ferreira was visiting Németh at the School of Mathematics in The University of Birmingham, during the period of January 2010 to July 2010. Ferreira was supported in part by FUNAPE/UFG, CNPq Grants 201112/2009-4, 475647/2006-8 and PRONEX–Optimization(FAPERJ/CNPq). Németh was supported in part by the Hungarian Research Grant OTKA 60480.

References

[1] H.H. Bauschke and J.M. Borwein, *On projection algorithms for solving convex feasibility problems*, SIAM Rev. 38 (1996), pp. 367–426.
[2] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty, *Nonlinear Programming, Theory and Algorithms*, John Wiley & Sons, New York–Chichester–Brisbane, 1979.
[3] R. Berk and R. Marcus, *Dual cones, dual norms, and simultaneous inference for partially ordered means*, J. Am. Stat. Assoc. 91 (1996), pp. 318–328.
[4] D.P. Bertsekas, A. Nedić, and A.E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, Belmont, MA, 2003.
[5] D. Butnariu and G. Kassay, *A proximal-projection method for finding zeros of set-valued operators*, SIAM J. Control Optim. 47 (2008), pp. 2096–2136.
[6] E. Carrizosa and F. Plastria, *Optimal expected-distance separating halfspace*, Math. Oper. Res. 33 (2008), pp. 662–677.
[7] I. Csiszar, *Generalized projections for non-negative functions*, Acta. Math. Hungar. 68 (1995), pp. 161–185.
[8] J. Dattorro, *Convex Optimization and Euclidean Distance Geometry*, Meboo, 2005, v2010.08.25.
[9] A. Dax, *The distance between two convex sets*, Linear Algebra Appl. 416 (2006), pp. 184–213.
[10] A. Dax and V.P. Sreedharan, *Theorems of the alternative and duality*, J. Optim. Theory Appl. 94 (1997), pp. 561–590.
[11] F. Deutsch and H. Hundal, *The rate of convergence of Dykstra's cyclic projections algorithm: The polyhedral case*, Numer. Funct. Anal. Optim. 15 (1994), pp. 537–565.
[12] R.L. Dykstra, *An algorithm for restricted least squares regression*, J. Am. Stat. Assoc. 78 (1983), pp. 837–842.