Antipodality in Convex Cones and Distance to Unpointedness

Alfredo Iusem [a], Alberto Seeger [b]

[a] *Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, Brazil*

[b] *University of Avignon, Department of Mathematics, 33 rue Louis Pasteur, 84000 Avignon, France*

Received *****; accepted after revision +++++

Abstract

We provide a complete answer to the problem which consists in finding an unpointed convex cone lying at minimal bounded Pompeiu-Hausdorff distance from a pointed one. We give also a simple and useful characterization of the radius of pointedness of a convex cone. A corresponding characterization for the radius of solidity of a convex cone is then derived by using a duality argument.

Keywords: Convex cone, Antipodality, Distance to unpointedness

1. Introduction

How far is a pointed convex cone, say K, from an unpointed one? How to construct an unpointed convex cone that is at minimal distance from K? These questions arise in the theory of convex cones and have a large variety of applications.

To start with, we fix the notation and terminology. The Euclidean space $\mathbb{R}^n$ is equipped with the usual inner product $\langle u, v\rangle = u^T v$ and associated norm $\|\cdot\|$. The symbol $\mathbb{S}_n$ refers to the unit sphere in $\mathbb{R}^n$. We equip the set

$$\mathcal{C}(\mathbb{R}^n) = \{K \subset \mathbb{R}^n : \ K \text{ is a (nontrivial) closed convex cone}\}$$

with the bounded Pompeiu-Hausdorff distance (cf. [6])

$$\delta(K,Q) = \max\left\{ \max_{x\in K\cap\mathbb{S}_n} \operatorname{dist}(x,Q), \max_{x\in Q\cap\mathbb{S}_n} \operatorname{dist}(x,K)\right\}.$$

That K is nontrivial simply means that K is different from $\{0\}$ and different from $\mathbb{R}^n$.

Email addresses: `iusp@impa.br` (Alfredo Iusem), `alberto.seeger@univ-avignon.fr` (Alberto Seeger).

Pointedness is an essential hypothesis in many theorems in which convex cones enter into the picture. One says that $K \in \mathcal{C}(\mathbb{R}^n)$ is *pointed* if $K \cap -K = \{0\}$, that is to say, if K doesn't contain a line. The number

$$\rho(K) = \min_{\substack{Q \in \mathcal{C}(\mathbb{R}^n) \\ Q \text{ unpointed}}} \delta(K, Q), \tag{1}$$

is called the *radius of pointedness* of K and it has been suggested in [3] as tool for measuring the degree of pointedness of K.

In general, the evaluation of (1) is a cumbersome task even for cones having a relatively simple structure. Fortunately, the least-distance problem (1) is related to the angle-maximization problem

$$\theta_{\max}(K) = \max_{u,v \in K \cap \mathbb{S}_n} \arccos\langle u, v\rangle, \tag{2}$$

which, in principle, is easier to solve because the decision variables u, v live in a standard Euclidean space.

Following [1], we say that $(u_0, v_0) \in \mathbb{R}^n \times \mathbb{R}^n$ is an *antipodal pair* of $K \in \mathcal{C}(\mathbb{R}^n)$ if

$$u_0, v_0 \in K \cap \mathbb{S}_n \quad \text{and} \quad \arccos\langle u_0, v_0\rangle = \theta_{\max}(K).$$

By compactness of $K \cap \mathbb{S}_n$, the nonconvex optimization problem (2) admits always a solution, so we don't have to worry about the existence of antipodal pairs.

From now on, the symbol $\langle w\rangle = \{\alpha w : \alpha \in \mathbb{R}\}$ denotes the line generated by a nonzero vector $w \in \mathbb{R}^n$ and $\langle w\rangle^\perp$ refers to the hyperplane which is orthogonal to this line.

Theorem 1 (Main Result) *For any $K \in \mathcal{C}(\mathbb{R}^n)$ one has*

$$\min_{\substack{Q \in \mathcal{C}(\mathbb{R}^n) \\ Q \text{ unpointed}}} \delta(K, Q) = \cos\left[\frac{\theta_{\max}(K)}{2}\right]. \tag{3}$$

Moreover, if K is not a half-line and admits (u_0, v_0) as antipodal pair, then the closed convex cone

$$Q_0 = (K \cap \langle u_0 - v_0\rangle^\perp) + \langle u_0 - v_0\rangle \tag{4}$$

is unpointed and lies at minimal distance from K.

Formula (3) was known to hold, until now, only under the additional (and bothering) hypothesis that $\theta_{\max}(K) \leq 2\pi/3$, see [4, Theorem 1]. The solution (4) to the least-distance problem (1) is given here for the first time.

2. Proof of the Main Result

For the sake of readibility we split the proof of Theorem 1 in five clearly distinguished steps. Throughout the proof we use the notation

$$w_0 = \frac{u_0 - v_0}{\|u_0 - v_0\|}\ .$$

We assume that K is not a half-line, otherwise both sides of (3) are equal to 1 and we are done. If K is unpointed, then both sides (3) are equal to 0 and Q_0 coincides with K as expected. So, there is no loss of generality in assuming that K is pointed.

Step 1. We start with some preliminary words on Q_0. The set Q_0 is clearly a convex cone in $\mathbb{R}^n$. On the other hand, Q_0 is closed because it is expressible as sum of a line $\langle w_0\rangle$ and a closed set contained in $\langle w_0\rangle^\perp$. Finally, Q_0 is unpointed because $Q_0 \cap -Q_0$ contains the nonzero vector w_0.

Step 2. We establish the inequality

$$\operatorname{dist}(x,K) \le \sqrt{\frac{1+\langle u_0,v_0\rangle}{2}} \qquad \forall x \in Q_0 \cap \mathbb{S}_n. \tag{5}$$

Take any $x \in Q_0 \cap \mathbb{S}_n$, so that $x = z + \alpha w_0$, with $z \in K \cap \langle w_0\rangle^\perp$ and $\alpha \in \mathbb{R}$. Clearly $\alpha = \langle x, w_0\rangle$, and therefore $|\alpha| \le 1$ in view of Cauchy-Schwarz inequality. Consider the point y defined as

$$y = \begin{cases} z + \alpha\sqrt{(1-\langle u_0,v_0\rangle)/2}\, u_0 & \text{if } \alpha \ge 0, \\ z - \alpha\sqrt{(1-\langle u_0,v_0\rangle)/2}\, v_0 & \text{if } \alpha \le 0. \end{cases}$$

Note that in both cases y belongs to K, because $z, u_0, v_0 \in K$. We proceed to estimate the distance between x and y. Consider first the case of $\alpha \ge 0$. One has

$$\|x-y\|^2 = \alpha^2 \left\| w_0 - \sqrt{\frac{1-\langle u_0,v_0\rangle}{2}}\; u_0 \right\|^2 = \alpha^2 \left[1 + \frac{1-\langle u_0,v_0\rangle}{2} - 2\sqrt{\frac{1-\langle u_0,v_0\rangle}{2}}\; \langle u_0, w_0\rangle\right].$$

A bit of elementary algebra yields

$$\begin{aligned} \|x-y\|^2 &= \alpha^2 \left[1 + \frac{1-\langle u_0,v_0\rangle}{2} - 2\sqrt{\frac{1-\langle u_0,v_0\rangle}{2}}\; \frac{1-\langle u_0,v_0\rangle}{\|u_0-v_0\|}\right] \\ &= \alpha^2 \left[1 + \frac{1-\langle u_0,v_0\rangle}{2} - 2\sqrt{\frac{1-\langle u_0,v_0\rangle}{2}}\; \frac{1-\langle u_0,v_0\rangle}{\sqrt{2(1-\langle u_0,v_0\rangle)}}\right] \\ &= \alpha^2 \left[1 + \frac{1-\langle u_0,v_0\rangle}{2} - (1-\langle u_0,v_0\rangle)\right] = \alpha^2 \left[\frac{1+\langle u_0,v_0\rangle}{2}\right]. \end{aligned}$$

Hence, $\operatorname{dist}(x,K) \le \|x-y\| \le \sqrt{(1+\langle u_0,v_0\rangle)/2}$. The case of $\alpha \le 0$ is dealt in a similar way.

Step 3. We now prove the inequality

$$d(x,Q_0) \le \sqrt{\frac{1+\langle u_0,v_0\rangle}{2}} \qquad \forall x \in K \cap \mathbb{S}_n. \tag{6}$$

It is at this stage where antipodality enters into action for the first time. Take $x \in K \cap \mathbb{S}_n$ and consider the vector

$$y = x + \frac{|\langle x,w_0\rangle|}{\|u_0-v_0\|}\,(u_0+v_0). \tag{7}$$

Note that y can be decomposed in the form

$$y = \underbrace{x - \langle x,w_0\rangle w_0 + \frac{|\langle x,w_0\rangle|}{\|u_0-v_0\|}(u_0+v_0)}_{\tilde{y}} \;+\; \underbrace{\langle x,w_0\rangle w_0}_{\hat{y}}.$$

Clearly, $\hat{y}$ belongs to $\langle w_0\rangle$. We claim that $\tilde{y} \in K \cap \langle w_0\rangle^\perp$. For checking that $\tilde{y} \in \langle w_0\rangle^\perp$, note that

$$\langle \tilde{y}, w_0\rangle = \langle x,w_0\rangle - \langle x,w_0\rangle\,\|w_0\|^2 + \frac{|\langle x,w_0\rangle|}{\|u_0-v_0\|^2}\langle u_0+v_0, u_0-v_0\rangle = \frac{|\langle x,w_0\rangle|}{\|u_0-v_0\|^2}\langle u_0+v_0, u_0-v_0\rangle = 0,$$

using the fact that $\|w_0\| = \|u_0\| = \|v_0\| = 1$. For checking that $\tilde{y} \in K$, rewrite $\tilde{y}$ as

$$\tilde{y} = x - \frac{\langle x,w_0\rangle}{\|u_0-v_0\|}(u_0-v_0) \;+\; \frac{|\langle x,w_0\rangle|}{\|u_0-v_0\|}(u_0+v_0)$$

$$= \begin{cases} x + 2\,\|u_0 - v_0\|^{-1}\,|\langle x, w_0\rangle|\,v_0 & \text{if } \langle x, w_0\rangle \geq 0, \\ x + 2\,\|u_0 - v_0\|^{-1}\,|\langle x, w_0\rangle|\,u_0 & \text{if } \langle x, w_0\rangle \leq 0. \end{cases}$$

In both cases, $\tilde{y} \in K$ because x, u_0 and v_0 belong to K. We conclude that $y = \tilde{y} + \hat{y}$ belongs to Q_0. We estimate next the distance between x and y. Directly from (7) one gets

$$\|x - y\| = \frac{|\langle x, u_0 - v_0\rangle|}{\|u_0 - v_0\|^2}\,\|u_0 + v_0\| = \frac{|\langle x, u_0 - v_0\rangle|}{2(1 - \langle u_0, v_0\rangle)}\sqrt{2(1 + \langle u_0, v_0\rangle)}.$$

In other words,

$$\|x - y\| = \eta\,\sqrt{\frac{1 + \langle u_0, v_0\rangle}{2}} \quad \text{with} \quad \eta = \frac{|\langle x, u_0 - v_0\rangle|}{1 - \langle u_0, v_0\rangle} \geq 0.$$

We claim that $\eta \leq 1$, which is equivalent to

$$|\langle x, u_0 - v_0\rangle| \leq 1 - \langle u_0, v_0\rangle. \tag{8}$$

If $\langle x, u_0 - v_0\rangle \geq 0$, then (8) is equivalent to

$$\langle x, u_0\rangle + \langle u_0, v_0\rangle \leq 1 + \langle x, v_0\rangle,$$

which holds because $\langle x, u_0\rangle \leq 1$, since both x and u_0 belong to $\mathbb{S}_n$, and also $\langle u_0, v_0\rangle \leq \langle x, v_0\rangle$, because x belongs to $K \cap \mathbb{S}_n$ and (u_0, v_0) is an antipodal pair of K. If $\langle x, u_0 - v_0\rangle \leq 0$, then (8) is equivalent to

$$\langle x, v_0\rangle + \langle u_0, v_0\rangle \leq 1 + \langle x, u_0\rangle,$$

which holds by the same reasons. This confirm that $\eta \leq 1$ as claimed. In this way we have shown that

$$\operatorname{dist}(x, Q_0) \leq \|x - y\| \leq \sqrt{(1 + \langle u_0, v_0\rangle)/2}\,.$$

Step 4. We prove the inequality

$$\rho(K) \leq \sigma(K) := \cos\left[\frac{\theta_{\max}(K)}{2}\right]. \tag{9}$$

Since (u_0, v_0) is an antipodal pair of K, one has

$$\sigma(K) = \sqrt{\frac{1 + \theta_{\max}(K)}{2}} = \sqrt{\frac{1 + \langle u_0, v_0\rangle}{2}}\,.$$

So, the combination of (5) and (6) yields in fact $\delta(K, Q_0) \leq \sigma(K)$. It suffices then to observe that $\rho(K) \leq \delta(K, Q_0)$ because $Q_0 \in \mathcal{C}(\mathbb{R}^n)$ is unpointed.

Step 5. We now prove the reverse inequality

$$\sigma(K) \leq \rho(K). \tag{10}$$

It has been shown in [2, Theorem 3.9] that $\sigma(\cdot)$ is a nonexpansive function over the metric space $(\mathcal{C}(\mathbb{R}^n), \delta)$, that is,

$$|\sigma(K_1) - \sigma(K_2)| \leq \delta(K_1, K_2) \qquad \forall K_1, K_2 \in \mathcal{C}(\mathbb{R}^n).$$

In particular,

$$\sigma(K) \leq \sigma(Q) + \delta(K, Q) \qquad \forall Q \in \mathcal{C}(\mathbb{R}^n).$$

So, one arrives at (10) by taking in the above line the infimum with respect to all unpointed cones in $\mathcal{C}(\mathbb{R}^n)$.

The proof of Theorem 1 is complete. Not only we proved the validity of formula (3), but also the fact that Q_0 solves the least-distance problem (1).

3. Conclusions

The proof of the formula (3) is certainly long and subtle, but we hope that the reader didn't find it excessively complicated. The merit of Theorem 1 is twofold : first of all, one obtains a nice and useful characterization of $\rho(K)$, regardless of whether or not the maximal angle of K falls beyond the critical value $2\pi/3$ that was bothering us so much in [4, Theorem 1]. And, secondly, one obtains an explicit solution to the least-distance problem (1).

Some interesting by-products of Theorem 1 deserve to be properly recorded. By using the very definition of the function $\rho : \mathcal{C}(\mathbb{R}^n) \to \mathbb{R}$, one can prove the following properties (cf.[3]) :

(A_0) *nonexpansiveness* : $|\rho(K_1) - \rho(K_2)| \leq \delta(K_1, K_2) \quad \forall K_1, K_2 \in \mathcal{C}(\mathbb{R}^n)$.

(A_1) *minimal pointedness* : $\rho(K) = 0$ if and only if K is unpointed.

(A_2) *maximal pointedness* : $\rho(K) = 1$ if and only if K is a half-line.

(A_3) *invariance property* : $\rho(U(K)) = \rho(K) \quad \forall K \in \mathcal{C}(\mathbb{R}^n),\ \forall U \in \mathbb{R}^{n\times n}$ orthonormal.

We challenge the reader to obtain a simple and rigorous proof of the property

(A_4) *downward monotonicity* : $K_1 \subset K_2$ implies $\rho(K_1) \geq \rho(K_2)$.

The proof of this monotonicity condition eluded us for a long time! Now we are getting it for free from formula (3). It suffices to observe that $\cos : [0, \pi/2] \to [0, 1]$ is a decreasing function and $\theta_{\max}(K)$ doesn't decrease if we enlarge the cone K.

As a second by-product of Theorem 1 one obtains a simple characterization for the *radius of solidity*

$$\mu(K) = \min_{\substack{R \in \mathcal{C}(\mathbb{R}^n) \\ R \text{ flat}}} \delta(K, R),$$

of a given $K \in \mathcal{C}(\mathbb{R}^n)$. That a cone $R \in \mathcal{C}(\mathbb{R}^n)$ is *flat* simply means that its topological interior is empty, that is to say, flatness is the concept which is opposite to solidity.

In the next corollary, the notation

$$K^+ = \{y \in \mathbb{R}^n : \ \langle y, x\rangle \geq 0,\ \forall x \in K\}$$

refers to the dual cone of K.

Corollary 2 *For any $K \in \mathcal{C}(\mathbb{R}^n)$ one has*

$$\min_{\substack{R \in \mathcal{C}(\mathbb{R}^n) \\ Q \text{ flat}}} \delta(K, R) = \cos\left[\frac{\theta_{\max}(K^+)}{2}\right]. \tag{11}$$

Moreover, if K^+ is not a half-line and admits (y_0, z_0) as antipodal pair, then the closed convex cone

$$R_0 = \left[(K^+ \cap \langle y_0 - z_0\rangle^{\perp}) + \langle y_0 - z_0\rangle\right]^+ \tag{12}$$

is flat and lies at minimal distance from K.

Proof. It is a matter of combining Theorem 1 and a certain duality relationship that exists between the functions ρ and μ (cf. Theorems 4.1 and 4.5 in [3]). □

Our last remark is addressed to the readers that are familiar with the theory of critical angles in convex cones (cf. [1], [5]). As one can see, K^+ plays a prominent role in the formulation of Corollary 2. Strictly speaking, we could have stated everything in terms of the original cone K. The explanation is as follows. We distinguish three disjoint cases :

i) K is a half-space. Then, K^+ is a half-line and both sides in (11) are equal to 1.

ii) K is flat and not a half-space. Then, both sides sides in (11) are equal to 0, and $R_0 = K$ as expected.

iii) K is solid and not a half-space. In this case one can write (11) in the equivalent form

$$\min_{\substack{R \in \mathcal{C}(\mathbb{R}^n) \\ Q \text{ flat}}} \delta(K, R) = \sin\left[\frac{\theta_{\min}(K)}{2}\right]$$

with $\theta_{\min}(K)$ standing for the smallest nonzero critical angle of K. On the other hand, one can take

$$y_0 = \frac{u_0 - \langle u_0, v_0\rangle v_0}{\sqrt{1 - \langle u_0, v_0\rangle^2}}\,, \quad z_0 = \frac{v_0 - \langle u_0, v_0\rangle u_0}{\sqrt{1 - \langle u_0, v_0\rangle^2}}\,,$$

where (u_0, v_0) is a critical pair of K forming the angle $\theta_{\min}(K)$. Such (y_0, z_0) is necessarily an antipodal pair of K^+. Finally, (12) can be written in the more compact form

$$R_0 = P_{\langle y_0 - z_0\rangle^\perp}(K)$$

with the symbol P_L standing for the orthogonal projector onto a subspace L. The last characterization of R_0 is obtained from (12) by applying standard calculus rules on dual cones. In general the projection of a closed convex cone into a subspace may not be closed. In the present situation, however, the closedness of $P_{\langle y_0 - z_0\rangle^\perp}(K)$ is guaranteed by using special arguments.

References

[1] A. Iusem, A. Seeger, On pairs of vectors achieving the maximal angle of a convex cone, Math. Programming, 104 (2005) 501–523.

[2] A. Iusem, A. Seeger, Axiomatization of the index of pointedness for closed convex cones, Computational and Applied Mathematics, 24 (2005) 245–283.

[3] A. Iusem, A. Seeger, Measuring the degree of pointedness of a closed convex cone: a metric approach, Math. Nachrichten, 279 (2006) 599–618.

[4] A. Iusem, A. Seeger, Computing the radius of pointedness of a convex cone, Math. Programming, 2006, in press (temporarily available at the preprint server of IMPA, http://www.preprint.impa.br).

[5] A. Iusem, A. Seeger, Searching for critical angles in a convex cone, Math. Programming, 2006, in press (temporarily available at the preprint server of IMPA, http://www.preprint.impa.br).

[6] D.W. Walkup, R.J.B. Wets, Continuity of some convex-cone-valued mappings, Proc. Amer. Math. Soc., 18 (1967) 229–235.