MIT Sloan School of Management

MIT Sloan Working Paper 4643-07
April 2007

A Geometric Analysis of Renegar's Condition Number,
and its Interplay with Conic Curvature

Alexandre Belloni and Robert M. Freund

A Geometric Analysis of Renegar's Condition Number, and its interplay with Conic Curvature

Alexandre Belloni* and Robert M. Freund†

April 4, 2007

Abstract

For a conic linear system of the form $Ax \in K$, K a convex cone, several condition measures have been extensively studied in the last dozen years. Among these, Renegar's condition number $\mathcal{C}(A)$ is arguably the most prominent for its relation to data perturbation, error bounds, problem geometry, and computational complexity of algorithms. Nonetheless, $\mathcal{C}(A)$ is a representation-dependent measure which is usually difficult to interpret and may lead to overly-conservative bounds of computational complexity and/or geometric quantities associated with the set of feasible solutions.

Herein we show that Renegar's condition number is bounded from above and below by certain purely geometric quantities associated with A and K, and highlights the role of the singular values of A and their relationship with the condition number. Moreover, by using the notion of conic curvature, we show how Renegar's condition number can be used to provide both lower and upper bounds on the width of the set of feasible solutions. This complements the literature where only lower bounds have heretofore been developed.

1 Introduction

We consider the problem of computing a solution of the following conic linear system:

$$\left\{ \begin{array}{rcl} Ax & \in & \textbf{int}\, K \\ x & \in & X \end{array} \right. \tag{1}$$

where X and Y are n- and m-dimensional Euclidean subspaces, respectively, $A : X \to Y$ is a linear operator and $K \subset Y$ is a regular closed convex cone. It is still an open question whether there exists a strongly-polynomial-time algorithm to solve (1) even for the case when $K = \mathbb{R}^m_+$. Indeed, computational complexity analysis of (1) always relies on some condition measure associated with the set of feasible solutions, the linear operator A, and/or the cone K, or combinations thereof.

Herein we focus on the data-perturbation condition number $\mathcal{C}(A)$ of Renegar, see [8]. It is well-known that $\ln(\mathcal{C}(A))$ is tied to the complexity of interior-point methods, the ellipsoid method, and re-scaled perceptron algorithms for computing a feasible solution of (1), see respectively [9], [3], and [1]. Nonetheless, Renegar's condition number depends on the particular norms being used and on the particular choice of A. As such, $\mathcal{C}(A)$ is not an intrinsic geometric measure associated with the feasible solution set of (1), and can provide misleading information about certain geometric aspects of the feasible solution set. Herein we show that $\mathcal{C}(A)$ is bounded from below and above by certain purely geometric quantities, which enable us to see better how $\mathcal{C}(A)$ behaves relative to the underlying problem geometry.

*IBM T. J. Watson Research Center and MIT, 32-221, 1101 Kitchawan Road, Yorktown Heights, New York 10598, email: belloni@mit.edu

†MIT Sloan School of Management, 50 Memorial Drive, Cambridge, MA 02142, USA, email: rfreund@mit.edu

A geometric condition measure associated with (1) is the *width* $\tau_{\mathcal{F}}$ of the cone of feasible solutions of (1). It is well understood that Renegar's condition number - combined with the width of K itself - can be used to bound the width of the feasible set from below. However, there is no reverse relation in general: $\mathcal{C}(A)$ by itself carries no upper bound information on the width $\tau_{\mathcal{F}}$ in general. In fact, one can observe this shortcoming of $\mathcal{C}(A)$ even for the linear inequality case, i.e., when $K = \mathbb{R}^m_+$.

By introducing a simple notion of *conic curvature*, we develop upper-bound information on the width of the feasible solution set as a function of the condition number $\mathcal{C}(A)$ and the conic curvature of the fixed cone K (or the components of K if K is a cartesian product of basic cones). These bounds pertain only to the case when K or its components have strictly positive conic curvature, and so are relevant when K is the cartesian product of second-order cones Q^{k_i} (whose conic curvature is 1) and products of $\mathbb{R}^2_+ \times \cdots \times \mathbb{R}^2_+$ (whose component conic curvature is also 1), but does not include the case when K is the cone of symmetric positive semi-definite matrices.

The paper is organized as follows. After preliminaries and notation in Section 2, in Section 3 we develop an analysis of Renegar's condition number in terms of purely geometric quantities that bound $\mathcal{C}(A)$ from above and below. In Section 4 we present a lower bound on the width of the feasible solution set of (1) in terms of $\mathcal{C}(A)$. In Section 5 we introduce the concept of *conic curvature*, and we present upper bounds on the width of the feasible solution set using $\mathcal{C}(A)$ and the conic curvature K. Section 6 contains comments about three themes underlying conic linear systems (1): conditioning, geometry, and complexity, and relations between these themes.

2 Preliminaries

Let $C \subset Y$ denote a convex cone and $C^* := \{w \in Y^* : \langle w, v\rangle \geq 0, \text{ for all } v \in C\}$ denote the dual cone associated with C, where Y^* is the dual space of Y. C is said to be regular if it is a pointed cone (it contains no lines) and has non-empty interior. Moreover, the *width* of C is defined as

$$\tau_C := \max_{r,x}\{r : \|x\| \leq 1, B(x,r) \subset C\}$$

where $B(x,r)$ denotes a Euclidean ball centered at x with radius r. x is the *center* of C if $\|x\| = 1$ and $B(x, \tau_C) \subset C$. Let ∂C denote the boundary of C and $\partial B(0,1) = S^{m-1}$ denote the $(m-1)$-dimensional unit sphere.

The cone of feasible solutions of (1) is denoted by $\mathcal{F} := \{x \in X : Ax \in K\}$ and its width is $\tau_{\mathcal{F}}$. Considering (1) as a system with fixed cone K and fixed spaces X and Y, let $\mathcal{M}$ denote those operators $A : X \to Y$ for which (1) has a solution. For $A \in \mathcal{M}$, let $\rho(A)$ denote the *distance to infeasibility* for (1), namely:

$$\rho(A) := \min_{\Delta A}\left\{\|\Delta A\| : A + \Delta A \notin \mathcal{M}\right\},$$

where for a linear operator M, $\|M\|$ denotes the operator norm, $\|M\| := \max\{\|Mx\| : \|x\| \leq 1\}$. $\rho(A)$ is the smallest perturbation of our given operator A which would render the system (1) infeasible. Let $\mathcal{C}(A) := \|A\|/\rho(A)$ denote Renegar's *condition number*, which is a scale-invariant reciprocal of the distance to infeasibility.

Let $A^* : Y \to X$ denote the adjoint operator associated with A, and let P_L denote the orthogonal projection on the subspace $L := \textbf{span}(A)$. It follows from basic linear algebra that $A^*w = A^*P_Lw$, and $\textbf{dist}(w, L^\perp) = \|P_Lw\|$ for any $w \in Y$. Let $\lambda_{\min}(A), \lambda_{\max}(A)$ denote the smallest and largest singular values of A, and note that $\lambda_{\max}(A) = \|A\| = \|A^*\|$. Furthermore, $\lambda_{\min}(A)\|w\| \leq \|A^*w\| \leq \lambda_{\max}(A)\|w\|$ for all $w \in L$.

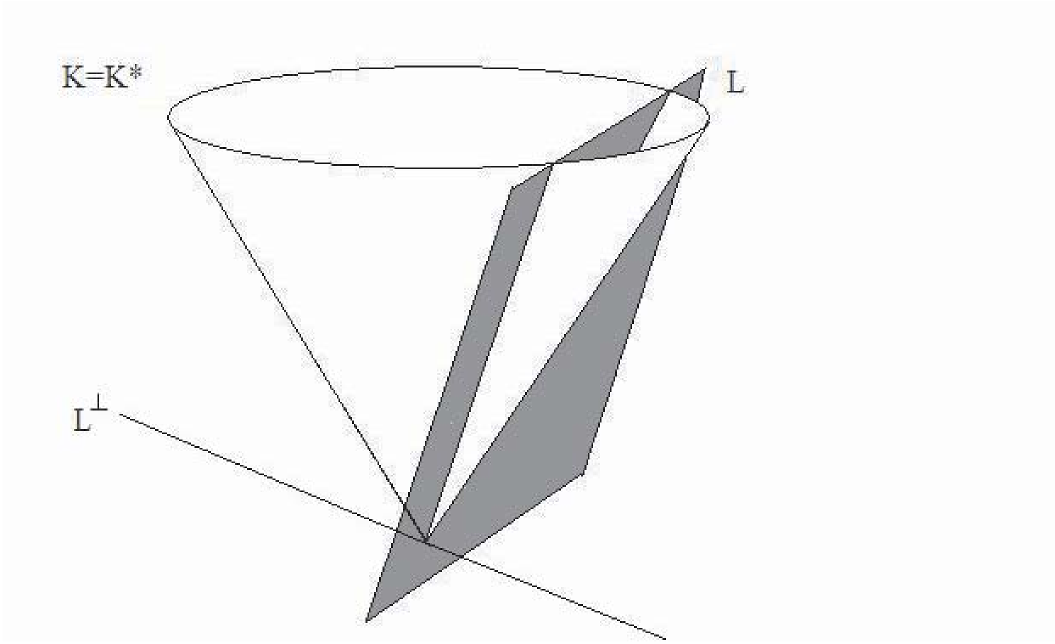

Figure 1: When $K = K^*$ and the original system is feasible, $L := \mathbf{span}(A)$ must have a nontrivial intersection with K. Equivalently, K^* can only intersect $L^\perp$ at the origin. This illustrates a conic theorem of the alternative.

3 Some Intrinsic Geometry related to Renegar's Condition Number

Our motivating geometry is Figure 1 which presumes for visualization that K is self-dual, i.e., $K = K^*$. Under the assumption that (1) is feasible, we have

$$K \cap L \neq \{0\} \quad \text{and} \quad K^* \cap L^\perp = \{0\}. \tag{2}$$

Although equivalent to (1), relation (2) is a geometric statement about the intersection of cones and subspaces. The dependence of the representation of A is embedded into $L^\perp$, the null space of the operator A^*.

The goal of this section is to relate Renegar's condition number to another (more geometric) quantity inspired by (2). In order to properly "quantify" the relations in (2), consider the following defined quantities:

$$\mu_A := \min_w \begin{array}{l} \|A^* w\| \\ \|w\| = 1 \\ w \in K^* \end{array} \quad \text{and} \quad \mu_L = \min_w \begin{array}{l} \|P_L w\| \\ \|w\| = 1 \\ w \in K^* \end{array} = \min_w \begin{array}{l} \mathbf{dist}(w, L^\perp) \\ \|w\| = 1 \\ w \in K^* \, . \end{array} \tag{3}$$

Whereas μ_A turns out to be the distance to infeasibility $\rho(A)$ (see Lemma 3.2), μ_L is a simpler object whose dependence on the data A is only through $L = \mathbf{span}(A)$. By its definition, μ_L captures important aspects of the underlying geometry of (1). The next theorem, which is the main result of this subsection, shows how the condition number is related to μ_L.

Theorem 3.1 *If (1) is feasible, then*

$$\frac{1}{\mu_L} \leq \mathcal{C}(A) \leq \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \frac{1}{\mu_L}.$$

Proof. The proof follows directly from Lemmas 3.1 and 3.2 below. ■

The following elementary lemma illustrates the "gap" between working with the particular representation of A and working with the subspace L.

Lemma 3.1 *If (1) is feasible, then*

$$\lambda_{\min}(A^*)\mu_L \leq \mu_A \leq \lambda_{\max}(A^*)\mu_L \ .$$

Proof. This follows since $\|A^*w\| = \|A^*P_L w\| \leq \lambda_{\max}(A)\|P_L w\|$ and $\|A^*w\| = \|A^*P_L w\| \geq \lambda_{\min}(A)\|P_L w\|$. ■

It turns out that the distance to infeasibility $\rho(A)$ is exactly the quantity μ_A. A version of this statement was initially proved as Theorem 2 of [4] using a more general non-homogeneous framework and arbitrary norms. For the sake of completeness we include a simple proof specialized to our problem setting.

Lemma 3.2 *If (1) is feasible, then $\rho(A) = \mu_A$.*

Proof. Let ΔA denote any perturbation for which the system $(A + \Delta A)x \in \mathbf{int}\ K$ is not feasible. By a theorem of the alternative, there exists $w \in K^*,\ \|w\| = 1$ satisfying $(A + \Delta A)^* w = 0$. Thus we have

$$\|\Delta A\| = \|\Delta A^*\| \geq \|\Delta A^* w\| = \|A^* w\| \geq \mu_A.$$

Taking the infimum on the left-hand side over all such ΔA we obtain $\rho(A) \geq \mu_A$.

On the other hand, let $\hat{w}$ be the point in K^*, $\|\hat{w}\| = 1$, such that $\mu_A = \|A^*\hat{w}\|$. Consider the linear operator $\Delta A_{\hat{w}}(\cdot) := -\hat{w}\,\langle \hat{w}, A(\cdot)\rangle$, whereby $\Delta A^*_{\hat{w}}\hat{w} = -A^*\hat{w}$. Therefore $(A + \Delta A_{\hat{w}})^*\hat{w} = 0$ which in turn implies that $\rho(A) \leq \|\Delta A_{\hat{w}}\| = \|A^*\hat{w}\| = \mu_A$. ■

The following example illustrates these concepts in the case of a small system of linear inequalities. This same example will be revisited in Sections 4 and 5.

Example 3.1 *Given $\varepsilon \in (0, 1/4)$, define $A_\varepsilon := \begin{bmatrix} 2\varepsilon & 0 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$ and let $K := \mathbb{R}^3_+$. Since the columns of A_ε are orthogonal, it is easy to verify that $L_\varepsilon^\perp = \{t(1, -\varepsilon, -\varepsilon)^T \ : \ t \in \mathbb{R}\}$. Note that $\mu_L = \mathbf{dist}(L_\varepsilon^\perp, K^* \cap S^2) = \sqrt{2}\varepsilon/\sqrt{1+2\varepsilon^2}$, $\lambda_{\max}(A_\varepsilon) = \|A_\varepsilon\| = \sqrt{2+4\varepsilon^2}$ and $\lambda_{\min}(A_\varepsilon) = \sqrt{2}$. Therefore, we have $\frac{1}{\sqrt{2}\varepsilon} \leq \frac{\sqrt{1+2\varepsilon^2}}{\sqrt{2}\varepsilon} = 1/\mu_L \leq \mathcal{C}(A_\varepsilon) \leq \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}\frac{1}{\mu_L} = \frac{1+2\varepsilon^2}{\sqrt{2}\varepsilon}$. In this example $1/\mu_L$ is a relatively tight bound on $\mathcal{C}(A)$ for ε small, since $\lambda_{\max}(A)/\lambda_{\min}(A) = \sqrt{1+2\varepsilon^2} \approx 1$ for ε small.*

4 Lower Bounds on the Width of the Cone of Feasible Solutions based on the Condition Number

The cone of feasible solutions of (1) is denoted by $\mathcal{F} := \{x \in X : Ax \in K\}$ and its width is $\tau_{\mathcal{F}}$:

$$\tau_{\mathcal{F}} = \max_{x,r}\{r : \|x\| \leq 1,\ AB(x,r) \subset K\} \ .$$

The following theorem presents lower bounds on $\tau_{\mathcal{F}}$ based on the condition number and/or on more geometric quantity μ_L. This is of interest from a computational complexity perspective since the iteration complexity of many algorithms for (1) involve $1/\tau_{\mathcal{F}}$ either polynomially (conditional gradient, perceptron) or logarithmically (interior-point methods, ellipsoid method, re-scaled perceptron). In conjunction with Theorem 4.1, $\mathcal{C}(A)$ and μ_L can be used to upper bound the computational complexity of these methods directly.

Theorem 4.1 *If (1) is feasible, then $\tau_{\mathcal{F}} \geq \tau_K\left(\frac{1}{\mathcal{C}(A)}\right) \geq \tau_K\left(\mu_L \frac{\lambda_{\min}(A)}{\lambda_{\max}(A)}\right)$.*

The middle quantity in Theorem 4.1 bounds the width of the feasible cone from below using the width of K and the condition number $\mathcal{C}(A)$, whereas the right-most quantity in the theorem bounds the width using quantities that each have a purely geometric interpretation. A version of the first inequality of the theorem was proved in slightly weaker form for non-homogeneous systems and arbitrary norms as Theorem 19 of [4]. Before proving the theorem we illustrate it in the continuation of Example 3.1.

Example 4.1 (Continuation Example 3.1) *Let $\mathcal{F}_\varepsilon := \{x \in \mathbb{R}^2 : A_\varepsilon x \in K\}$ where $K = \mathbb{R}^3_+$. Then $\mathcal{F}_\varepsilon = \{x \in \mathbb{R}^2 : x_1 \geq x_2,\ x_1 \geq -x_2\}$, and $\tau_{\mathcal{F}_\varepsilon} = \sqrt{1/2} \geq \frac{\sqrt{\frac{1}{3}}\sqrt{2}\varepsilon}{\sqrt{1+2\varepsilon^2}} = \tau_K \mu_L \geq \tau_K / \mathcal{C}(A_\varepsilon)$. Note that this bound is very weak when ε is close to zero.*

The proof of Theorem 4.1 is based on the following elementary proposition and corollary, and a technical lemma, Lemma 4.1, whose results will be used here as well as in Section 5.

Proposition 4.1 *Let C be closed convex cone. Then $B(y,r) \subseteq C$ if and only if $\langle d, y\rangle \geq r\|d\|$ for all $d \in C^*$.*

Proof. Suppose $B(y,r) \subset C$. Let $d \in C^*$. Then, $y - r\frac{d}{\|d\|} \in C$ and since $d \in C^*$, $\left\langle d, y - r\frac{d}{\|d\|}\right\rangle \geq 0$. Thus, $\langle d, y\rangle \geq r\frac{\langle d,d\rangle}{\|d\|} = r\|d\|$. Conversely, suppose $\langle d, y\rangle \geq r\|d\|$ for every $d \in C^*$. Let v satisfy $\|v\| \leq r$. Assume $y + v \notin C$, then there exists $d \in C^*$, $\langle d, y+v\rangle < 0$. Therefore $\langle d, y\rangle < -\langle d, v\rangle \leq r\|d\|$, which contradicts $\langle d, y\rangle \geq r\|d\|$. ■

Corollary 4.1 *Let C be closed convex cone. Then for $y \in C$ we have $\max_r\{r : B(y,r) \subseteq C\} = \min_d\{\langle d, y\rangle : d \in C^*,\ \|d\| = 1\}$.*■

The last ingredient of the proof of Theorem 4.1 is Lemma 4.1 below, which is based on the following definition.

Definition 4.1 *Let $K \subset Y$ be a fixed convex cone, $A : X \to Y$ be a linear operator, and $M \subset Y$ be a subspace. Define the deepness of A (in K) with respect to M as*

$$\begin{array}{rll} \rho(A,M) := \max_{x,r} & r \\ \text{s.t.} & Ax + \Big(B(0,r) \cap M\Big) \subset K \\ & \|x\| \leq 1\ . \end{array}$$

Definition 4.1 is a geometric counterpart/generalization of concepts developed in [9] and [4] for $M = Y$. The following technical lemma shows some ways that this deepness measure is related to the distance to infeasibility $\rho(A)$.

Lemma 4.1 *Let $L = \mathbf{span}(A)$ and K be a pointed convex cone. Then:*
(i) $\rho(A, L) \leq \|A\|$;
(ii) $\rho(A) \geq \rho(A, Y) \geq \tau_K \rho(A)$;
(iii) $\rho(A, Y) \geq \frac{1}{2\sqrt{2}} \min\{\rho(A, L), \rho(A, L^\perp)\}$.

Proof. (i) If x, r is feasible for the problem defining $\rho(A, L)$ and $r > \|A\|$, then $Ax + B(0,r) \cap L$ contains a neighborhood of the origin restricted to the subspace L. If this neighborhood is contained in K, then K contains L and is not pointed, thus proving (i) by contradiction.The proof of the first inequality of (ii) is a straightforward consequence of minimax weak duality:

$$\begin{array}{rlcll} \rho(A) = \min_w & \|A^*w\| & = & \min_w & \max_x \quad \langle A^*w, x\rangle \\ \text{s.t.} & w \in K^* & & w \in K^* & \|x\| \leq 1 \\ & \|w\| = 1 & & \|w\| = 1 & \end{array}$$

$$\geq \begin{array}{c}\max_x\\ \|x\|\leq 1\end{array} \begin{array}{c}\min_w\\ w\in K^*\\ \|w\|=1\end{array} \langle w, Ax\rangle = \begin{array}{rl}\max_{x,r} & r\\ \text{s.t.} & B(Ax,r)\subset K\\ & \|x\|\leq 1\ .\end{array} = \rho(A,Y)$$

Here the first equality is from Lemma 3.2, the inequality is from the observation that minimax $\geq$ maximin, and the second-to-last equality above follows from Corollary 4.1. To prove the second inequality of (ii), let z be the center of K, whereby $\|z\| = 1$ and $B(z, \tau_K) \subset K$. It then follows from Proposition 4.1 that $\tau_K\|w\| \leq \langle z, w\rangle \leq \|w\|$ for any $w \in K$. We have:

$$\rho(A) = \begin{array}{rl}\min_w & \|A^*w\|\\ \text{s.t.} & w\in K^*\\ & \|w\|\geq 1\end{array} \leq \begin{array}{rl}\min_w & \|A^*w\|\\ \text{s.t.} & w\in K^*\\ & \langle z,w\rangle = 1\end{array} = \begin{array}{c}\min_w\\ w\in K^*\\ \langle z,w\rangle=1\end{array} \begin{array}{rl}\max_x & \langle A^*w, x\rangle\\ \text{s.t.} & \|x\|\leq 1\end{array}$$

$$= \begin{array}{c}\max_x\\ \|x\|\leq 1\end{array} \begin{array}{rl}\min_w & \langle Ax, w\rangle\\ \text{s.t.} & w\in K^*\\ & \langle z,w\rangle=1\end{array} \leq \begin{array}{c}\max_x\\ \|x\|\leq 1\end{array} \begin{array}{rl}\min_w & (1/\tau_K)\langle Ax, w\rangle\\ \text{s.t.} & w\in K^*\\ & \|w\|\geq 1\end{array}$$

$$= (1/\tau_K)\begin{array}{rl}\max_{x,r} & r\\ \text{s.t.} & B(Ax,r)\subset K\\ & \|x\|\leq 1\end{array} = \rho(A,Y)/\tau_K\ .$$

Here the first equality is from Lemma 3.2 and the third equality follows from a standard minimax theorem, see Bertsekas [2] Proposition 5.4.4, for example. The second-to-last inequality uses $\tau_K\|w\| \leq \langle z, w\rangle$ for any $w \in K$ from Proposition 4.1, and the second-to-last equality follows from Corollary 4.1.

To prove (iii), let x_L and $x_{L^\perp} \in K$ be points that achieve respectively $\rho(A, L)$ and $\rho(A, L^\perp)$. By convexity of K, for $w := A\left(\frac{x_L+x_{L^\perp}}{2}\right)$, we have

$$w + B\left(0, \frac{1}{2}\rho(A,L)\right) \cap L \subset K \text{ and } w + B\left(0, \frac{1}{2}\rho(A,L^\perp)\right) \cap L^\perp \subset K\ .$$

This implies that $B\left(w, \frac{1}{2\sqrt{2}}\min\{\rho(A,L), \rho(A,L^\perp)\}\right) \subset K$, whereby

$$\rho(A,Y) \geq \frac{1}{2\sqrt{2}}\min\{\rho(A,L), \rho(A,L^\perp)\}\ .$$

■

Proof of Theorem 4.1: From Definition 4.1 and Lemma 4.1 there exists x satisfying $\|x\| \leq 1$ and $Ax + B(0, \tau_K\rho(A)) \subset Ax + B(0, \rho(A, Y)) \subset K$, which implies that $AB(x, \tau_K\rho(A)/\|A\|) \subset K$, from which it follows that $\tau_{\mathcal{F}} \geq \tau_K\left(\frac{1}{\mathcal{C}(A)}\right)$. The second inequality follows directly from Theorem 3.1.■

5 Upper Bounds on the Width of the Cone of Feasible Solutions based on the Condition Number and Conic Curvature

As illustrated in Example 4.1 when ε is small, the condition number $\mathcal{C}(A)$ by itself might contain no upper-bound information on the width of the feasibility cone. Herein we show that if one has information about the conic curvature of the cone K, then such upper-bound information is available. Furthermore, as we will see, for standard cross-products of cones such conic curvature information is known *a priori*.

5.1 Conic Curvature

The term "conic" in conic curvature is intended to contrast our concept with other notions of curvature that can be found in the literature. The origins of this concept can be traced to the notion of strongly convex sets which was first defined by Levitin and Poljak in [6]. A systematic study of such sets can be found in [7, 10, 11].

Definition 5.1 *A set S is strongly convex with parameter $\delta \geq 0$ if S is bounded, and for every pair of points $x,\ y \in S$, and every $\lambda \in [0,1]$, it holds that*

$$\lambda x + (1-\lambda)y + B\left(0,\ \frac{\delta}{2}\lambda(1-\lambda)\|x-y\|^2\right) \subset S\ .$$

According to Definition 5.1, even if we relax the boundedness assumption, no convex cone (other than $\mathbb{R}^n$ or $\{0\}$) is strongly convex with $\delta > 0$. To obtain a meaningful concept of strong convexity and curvature for cones, we simply limit attention to points with unit norm:

Definition 5.2 *A cone C has conic curvature δ if δ is the largest nonnegative scalar such that for every pair of points $x,\ y \in C$ satisfying $\|x\| = \|y\| = 1$, and every $\lambda \in [0,1]$, it holds that*

$$\lambda x + (1-\lambda)y + B\left(0,\ \frac{\delta}{2}\lambda(1-\lambda)\|x-y\|^2\right) \subset C\ .$$

Although technically valid, this definition is meaningless for dimension $n = 1$, since in this case all regular cones are half-lines and trivially have infinite curvature.

Example 5.1 ($\mathbb{R}^2_+$) *The nonnegative orthant in $\mathbb{R}^2$ has curvature $\delta = 1$. To see this let e_1 and e_2 be the extreme rays of $\mathbb{R}^2_+$ with unit length, in particular $\|e_1 - e_2\|^2 = 2$. Moreover, for $\lambda \in [0,1]$ we have* $\mathbf{dist}(\lambda e_1 + (1-\lambda)e_2, \partial\mathbb{R}^2_+) = \min\{\lambda, 1-\lambda\}$*. The curvature δ is the largest scalar that satisfies*

$$\min\{\lambda, (1-\lambda)\} \geq \delta\frac{\lambda(1-\lambda)}{2}\|e_1 - e_2\|^2 = \delta\lambda(1-\lambda),$$

which implies that $\delta = 1$.

The nonnegative orthant $\mathbb{R}^n_+$ for $n \geq 3$ and the cone of positive semi-definite symmetric matrices S^n_+ for $n \geq 3$ have zero conic curvature. On the other hand, the second-order cone has conic curvature $\delta = 1$ for $n \geq 2$ as shown in the next lemma.

Lemma 5.1 *For $n \geq 2$, the second-order cone $Q^n := \{(x,t) \in \mathbb{R}^n : \|x\| \leq t\}$ has conic curvature $\delta = 1$.*

Proof. Let $(x,t),(y,s) \in \partial Q^n$, $\|(x,t)\| = \|(y,s)\| = 1$. Therefore we have $\|x\|^2 + t^2 = 1$, $\|x\| = t = 1/\sqrt{2}$, likewise we have $\|y\| = s = 1/\sqrt{2}$.

For $\lambda \in [0,1]$ let $(z,u) = \lambda(x,t) + (1-\lambda)(y,s) = (\lambda x + (1-\lambda)y, 1/\sqrt{2}) \in Q^n$. Define r such that $r + \|z\| = 1/\sqrt{2}$. Note that $\mathbf{dist}((z,u),\partial Q^n) = \frac{r}{\sqrt{2}}$ and the result follows if we show that $r \geq \frac{1}{\sqrt{2}}\lambda(1-\lambda)\|x-y\|^2$, and that equality holds in the limit for some values of λ, x, and y.

By the Pythagorean Theorem noting that $z = (1/2)(x+y) + (\lambda - 1/2)(x-y)$ and $(x+y)^T(x-y) = 0$, we have

$$\|z\|^2 = \left\|\frac{x+y}{2}\right\|^2 + \left(\frac{1}{2} - \lambda\right)^2\|x-y\|^2.$$

Using the definition of r in the relation above we have

$$\begin{aligned} r &= \tfrac{1}{\sqrt{2}} - \sqrt{\left\|\tfrac{x+y}{2}\right\|^2 + \left(\tfrac{1}{2}-\lambda\right)^2\|x-y\|^2} = \tfrac{1}{\sqrt{2}} - \sqrt{\tfrac{1}{2} - \lambda(1-\lambda)\|x-y\|^2} \\ &\geq \tfrac{1}{\sqrt{2}}\lambda(1-\lambda)\|x-y\|^2 \end{aligned}$$

where we used $\left\|\frac{x+y}{2}\right\|^2 + \left\|\frac{x-y}{2}\right\|^2 = \frac{1}{2}$ and $\frac{1}{\sqrt{2}} - \sqrt{\frac{1}{2} - a} \geq \frac{1}{\sqrt{2}} a$ for $a \in [0, 1/2]$.

Equality follows by considering the case of $\lambda = 1/2$, and x arbitrary close to y. ■

Finally, it follows from the definition that the conic curvature of the intersection of two convex cones is at least the minimum of the conic curvature of the two cones. Moreover, except for the case of $\mathbb{R}^2_+$ considered as the the cartesian product $\mathbb{R}^2_+ = \mathbb{R}_+ \times \mathbb{R}_+$, the cartesian product of two cones must have zero conic curvature. This last comment notwithstanding, we show in Corollary 5.1 how to construct useful upper bounds when K is the cartesian product of simpler cones.

5.2 Using Conic Curvature to Upper-Bound the Width of the Feasibility Cone

We will prove the following theorem which bounds the width of the feasible region of (1) from above using $\mathcal{C}(A)$ in conjunction with the conic curvature of K.

Theorem 5.1 *Let K be a regular convex cone with conic curvature δ, and let A be a linear operator satisfying* $\mathbf{rank} A \geq 2$. *If (1) is feasible, then*

$$\tau_{\mathcal{F}} \leq \frac{1}{\sqrt{\mathcal{C}(A)}} \frac{\sqrt[4]{8}}{\sqrt{\delta}} \left(\frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \right) \leq \sqrt{\mu_L} \frac{\sqrt[4]{8}}{\sqrt{\delta}} \left(\frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \right) .$$

Similar to Theorem 4.1, the middle quantity in Theorem 5.1 bounds the width of the feasible cone from above using geometric quantities and the condition number $\mathcal{C}(A)$, whereas the right-most quantity in the theorem bounds the width using quantities that each have a purely geometric interpretation.

Before proving Theorem 5.1, we present a corollary that pertains to the case when K is the cartesian product of basic cones., and the system (1) has the form

$$A_i x \in \mathbf{int}\ K_i, \quad \text{for } i = 1, \ldots, r \ , \tag{4}$$

where $K = K_1 \times K_2 \times \cdots \times K_r \subset Y_1 \times Y_2 \times \cdots \times Y_r = Y$. Let $\rho(A_i)$ and $\mathcal{C}(A_i)$ denote the distance to infeasibility and the condition number of the system $A_i x \in K_i$ for $i = 1, \ldots, r$. Although in this case the conic curvature of K is zero under Definition 5.2 (except for the special case when $K = \mathbb{R}^2_+$), it is sufficient to look at the curvatures δ_i of the basic cones K_i, $i = 1, \ldots, r$, as the following corollary demonstrates.

Corollary 5.1 *Consider the system (4), let δ_i be the conic curvature of K_i, $i = 1, \ldots, r$, and suppose that* $\mathbf{rank} A_i \geq 2$ *for $i = 1, \ldots, r$. Then*

$$\tau_{\mathcal{F}} \leq \sqrt[4]{8} \min_i \left\{ \frac{\sqrt{\tau_{K_i}}}{\sqrt{\delta_i}} \frac{\lambda_{\max}(A_i)}{\lambda_{\min}(A_i)} \right\} ,$$

$$\tau_{\mathcal{F}} \leq \sqrt[4]{8} \min_i \left\{ \frac{1}{\sqrt{\delta_i}} \frac{1}{\sqrt{\mathcal{C}(A_i)}} \frac{\lambda_{\max}(A_i)}{\lambda_{\min}(A_i)} \right\} ,$$

and

$$\tau_{\mathcal{F}} \leq \frac{\sqrt[4]{8}}{\sqrt{\mathcal{C}(A)}} \max_i \left\{ \frac{1}{\sqrt{\delta_i}} \frac{\lambda_{\max}(A_i)}{\lambda_{\min}(A_i)} \sqrt{\frac{\lambda_{\max}(A)}{\lambda_{\max}(A_i)}} \right\} .$$

Remark 5.1 *The requirement that* $\mathbf{rank} A_i \geq 2$ *cannot be relaxed. To see this, suppose $K = \mathbb{R}^m_+$ and we consider K as the cartesian product of m half-lines $\mathbb{R}_+$, in which case* $\mathbf{rank} A_i = 1,\ i = 1, \ldots, m$. *In this case $\delta_i = \infty,\ i = 1, \ldots, m$, and the resulting upper bound on $\tau_{\mathcal{F}}$ would (falsely) be zero.*

Remark 5.2 *Although the conic curvature of* $\mathbb{R}^m_+$ *is zero for* $m \geq 3$*, the conic curvature of* $\mathbb{R}^2_+$ *is equal to one. By adding one redundant inequality if necessary to ensure that m is even, without loss of generality we can write a linear inequality system* $Ax \in \mathbb{R}^m_+$ *as*

$$A_i x \in \mathbb{R}^2, \quad i = 1, \ldots, m/2 \ ,$$

and through Corollary 5.1 with $\delta_i = 1$, $i = 1, \ldots, m/2$*, we obtain upper bounds on the width of the feasible region in terms of the condition number* $\mathcal{C}(A)$*. Notice from the last bound in Corollary 5.1 that the more natural quantity* $\lambda_{\min}(A)$ *is missing in the bound, and cannot in general be inserted since* $\lambda_{\min}(A) \geq \min_{i=1,\ldots,m}\{\lambda_{\min}(A_i)\}$.

A close inspection of the second bound in Corollary 5.1 leads us to conclude that if the original problem has good geometry ($\tau_{\mathcal{F}}$ is large) but has a badly conditioned subsystem ($\mathcal{C}(A_i)$ is large for some i), the ratio between the largest and smallest singular value of A_i must be large to accommodate the discrepancy. Put another way, for instances in which the ratio between the largest and smallest singular values of A_i is small, Renegar's condition number yields relevant upper bounds on the width of the feasible region in the presence of finite conic curvature.

We now proceed to prove Theorem 5.1 and Corollary 5.1. We first prove two intermediary results.

Proposition 5.1 $\rho(A, L) \geq \lambda_{\min}(A)\tau_{\mathcal{F}}$.

Proof. By definition there exists $z \in \mathbb{R}^n$, $\|z\| = 1$, for which $B(z, \tau_{\mathcal{F}}) \subset \mathcal{F}$, i.e., $AB(z, \tau_{\mathcal{F}}) \subset K$. Therefore $Az + B(0, \lambda_{\min}(A)\tau_{\mathcal{F}}) \cap L \subset K$, which implies that $\rho(A, L) \geq \lambda_{\min}(A)\tau_{\mathcal{F}}$. ■

Lemma 5.2 *Let K be a regular convex cone with conic curvature* δ*, and suppose that* **rank**$A \geq 2$ *and (1) is feasible. Then* $\rho(A) \geq \frac{\delta}{2\sqrt{2}} \frac{\lambda^2_{\min}(A)\tau^2_{\mathcal{F}}}{\lambda_{\max}(A)}$.

Proof. From Definition 4.1, there exists x_L satisfying $\|x_L\| \leq 1$ and $Ax_L + B(0, \rho(A, L)) \cap L \subset K$. Since L has dimension at least two, there exists $v \in L, \|v\| = 1, \langle v, Ax_L\rangle = 0$, such that $u_{\pm} := Ax_L \pm \rho(A, L)v \in K$. Therefore we have $\|u_{\pm}\| = \sqrt{\|Ax_L\|^2 + \rho(A, L)^2}$ and $\|u_+ - u_-\| = 2\rho(A, L)$. Since K has conic curvature δ, using $\lambda = 1/2$ we have

$$B\left(\frac{Ax_L}{\sqrt{\|Ax_L\|^2 + \rho(A, L)^2}}, \frac{\delta}{2}\frac{1}{4}\frac{4\rho(A, L)^2}{\|Ax_L\|^2 + \rho(A, L)^2}\right) \subset K.$$

After re-normalizing and applying Lemma 4.1 and Proposition 5.1, we obtain

$$\rho(A) \geq \rho(A, Y) \geq \frac{\delta}{2}\frac{\rho(A, L)^2}{\sqrt{\|Ax_L\|^2 + \rho(A, L)^2}} \geq \frac{\delta}{2\sqrt{2}}\frac{\lambda^2_{\min}(A)\tau^2_{\mathcal{F}}}{\lambda_{\max}(A)} \ .$$

■

Proof of Theorem 5.1: The first inequality follows directly from Lemma 5.2, and the second inequality follows from Theorem 3.1. ■

Proof of Corollary 5.1: By definition, there exist $\bar{x}$, $\|\bar{x}\| = 1$, such that $A_i B(\bar{x}, \tau_{\mathcal{F}}) \subset K_i$ for $i = 1, \ldots, r$. Using a similar construction to that used in the proofs of Proposition 5.1 and Lemma 5.2, for each basic cone K_i we obtain

$$B\left(A_i\bar{x}, \frac{\delta_i}{2\sqrt{2}}\frac{\lambda^2_{\min}(A_i)\tau^2_{\mathcal{F}}}{\lambda_{\max}(A_i)}\right) \subset K_i \ . \tag{5}$$

From this it follows that $\tau_{K_i} \geq \frac{\delta_i \tau^2_{\mathcal{F}}\lambda^2_{\min}(A_i)}{2\sqrt{2}\lambda^2_{\max}(A_i)}$ and rearranging yields the first inequality of the corollary. Applying Lemma 4.1 to the system $A_i x \in K_i$ and using (5) we obtain $\rho(A_i) \geq \rho(A_i, Y_i) \geq \frac{\delta_i \tau^2_{\mathcal{F}}\lambda^2_{\min}(A_i)}{2\sqrt{2}\lambda_{\max}(A_i)}$, and rearranging yields the second inequality of the corollary. Note that (5) also implies

$$B\left(A\bar{x}, \frac{\tau^2_{\mathcal{F}}}{2\sqrt{2}}\min_{i=1,\ldots,r}\frac{\delta_i\lambda^2_{\min}(A_i)}{\lambda_{\max}(A_i)}\right) \subset K$$

and applying Lemma 4.1 to this inclusion yields $\rho(A) \ge \rho(A, Y) \ge \frac{\tau_{\mathcal{F}}^2}{2\sqrt{2}} \min_{i=1,\ldots,r} \frac{\delta_i \lambda_{\min}^2(A_i)}{\lambda_{\max}(A_i)}$, and rearranging yields the third inequality of the corollary. ■

6 Conditioning, Geometry, and Complexity

Renegar's condition number $\mathcal{C}(A)$ is a data-perturbation condition measure that aims to capture how close the data instance A of (1) is to the set of infeasible instances of (1), where A is replaced by $A + \Delta A$ and K is kept fixed. One can easily conjure up other ways to define perturbations of homogeneous or non-homogeneous conic linear systems (perturbing K, perturbing **span**(A), introducing non-homogeneity, etc.) and other metrics of closeness to infeasible instances. Nevertheless, $\mathcal{C}(A)$ is sufficiently robust a measure to be connected to two other important themes underlying conic linear systems (1), namely (i) the *geometry* of the set of feasible solutions (as measured canonically with the width $\tau_{\mathcal{F}}$ of the set of feasible solutions), and (ii) the worst-case *complexity* of efficient algorithms (such as interior-point methods and the ellipsoid method). Indeed, $\mathcal{C}(A)$ yields lower bounds on the width $\tau_{\mathcal{F}}$ (Theorem 4.1) and upper bounds on the complexity of interior-point methods [9] and the ellipsoid method [3], thus implying in vernacular that "good conditioning implies good geometry and good complexity." Concerning other implications among these three notions, results from [3] and [5] show that "good geometry implies good complexity," for example. And the contribution of Theorem 5.1 here is that under positive conic curvature of K or its components, that "good geometry implies good conditioning." However, the plausible statement "good complexity implies good conditioning and/or good geometry" lacks a formal mathematical argument. Furthermore, we believe that the truthfulness of this statement seems reasonably important to ascertain.

Acknowledgement. We are grateful to Santosh Vempala for research discussions leading to the work contained herein.

References

[1] A. Belloni, R. M. Freund, and S. Vempala, *An efficient re-scaled perceptron algorithm for conic systems*, MIT Operations Research Center Working Paper (2006), no. OR-379-06.

[2] D. Bertsekas, *Nonlinear programming*, Athena Scientific, 1999.

[3] R. M. Freund and J. R. Vera, *Condition-based complexity of convex optimization in conic linear form via the ellipsoid algorithm*, SIAM Journal on Optimization **10** (1999), no. 1, 155–176.

[4] ———, *Some characterizations and properties of the "distance to ill-posedness" and the condition measure of a conic linear system*, Mathematical Programming **86** (1999), no. 2, 225–260.

[5] Robert M. Freund, *Complexity of convex optimization using geometry-based measures and a reference point*, Mathematical Programming **99** (2004), 197–221.

[6] E. Levitin and B. J. Poljak, *Constrained minimization methods*, U.S.S.R. Computational Math. and Math Phys. **6** (1963), 787–823.

[7] B. J. Poljak, *Existence theorems and convergence of minimizing sequences in extremum problems with restrictions*, Soviet Math. **7** (1966), 72–75.

[8] J. Renegar, *Some perturbation theory for linear programming*, Mathematical Programming **65** (1994), no. 1, 73–91.

[9] ———, *Linear programming, complexity theory, and elementary functional analysis*, Mathematical Programming **70** (1995), no. 3, 279–351.

[10] J.-P. Vial, *Strong convexity of sets and functions*, Journal of Mathematical Economics **9** (1982), 187–205.

[11] ———, *Strong and weak convexity of sets and functions*, Mathematics of Operaitons Research **8** (1983), no. 2, 231–259.