Set-Valued Analysis **9:** 375–381, 2001.

Ill-Conditioned Inclusions

A. S. LEWIS
Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. e-mail: aslewis@orion.uwaterloo.ca

(Received: 30 January 2000; in final form: 3 November 2000)

Abstract. A square system of linear equations is 'ill-conditioned' when the norm of the corresponding inverse matrix is large. This norm bounds the size of the solution, and measures how close the system is to being inconsistent: it is thus of fundamental computational significance. We generalize this idea from linear equations to inclusions governed by closed convex processes, and hence to 'conic linear systems'.

Mathematics Subject Classifications (2000): Primary: 15A12, 90C31; Secondary: 65F35, 46A30.

Key words: condition number, conic linear system, distance to inconsistency, convex process, surjectivity, homogenization.

1. Introduction

Given an invertible $n \times n$ real matrix F and a vector b in $\mathbf{R}^n$, consider the linear system

$$Fx = b, \quad x \in \mathbf{R}^n.$$

It is a simple consequence of the Eckart–Young theorem (see [7], for example) that the smallest operator norm of a matrix G making the perturbed matrix $F + G$ singular is just $\|F^{-1}\|^{-1}$. This quantity is a fundamental measure of the 'conditioning' of the system: for example, the solution of the original system clearly has norm no larger than $\|F^{-1}\|\|b\|$.

In 1995 Renegar extended this notion of conditioning to 'conic linear systems' (systems of equations and inequalities defined by convex cones). He was able to relate this idea to the complexity of solving such systems by interior point methods, thus establishing its importance for linear and semidefinite programming [4]. Our aim here is to give a simple and concise extension of Renegar's work, using the elegant language of 'convex processes'. This language, originating with [6], is perfectly suited to a unified study of convex cones and inequality systems.

Given real vector spaces X and Y, we call a set-valued map Φ: $X \to Y$ a *convex process* if its *graph* $G(\Phi) = \{(x, y) : y \in \Phi(X)\}$ is a convex cone [6]. We are interested in the conditioning of the inclusion

$$b \in \Phi(x), \quad x \in X, \tag{1.1}$$

for a given vector b in Y. Such inclusions, as we shall see, include conic linear systems as a special case (and hence both linear and semidefinite programming), but are both concise in notation and broader in scope.

We define the *inverse process* $\Phi^{-1}\colon\ Y \to X$ by

$$x \in \Phi^{-1}(y) \Leftrightarrow y \in \Phi(x), \quad \text{for } x \in X,\ y \in Y.$$

In general, system (1.1) may be *inconsistent* (that is, $\Phi^{-1}(b) = \emptyset$) unless Φ is *surjective* (by which we mean the *range* $\Phi(X)$ equals Y). Our aim in this paper is to relate the ideas of conditioning, surjectivity, and solution size.

Since the map Φ is positively homogeneous, solving the inclusion (1.1) is structurally no different than solving $x \in \Phi^{-1}(tb)$ for some real $t > 0$. This motivates considering a 'homogenization' construction, which we discuss next.

2. Homogenization

We can *homogenize* the inclusion $b \in \Phi(x)$ by considering a new convex process $\Phi_b\colon\ X \times \mathbf{R} \to Y$ defined by

$$\Phi_b(x,t) = \begin{cases} \Phi(x) - tb & \text{if } t \geqslant 0, \\ \emptyset & \text{if } t < 0. \end{cases} \tag{2.1}$$

Our fundamental tool, a purely algebraic result following an idea from [1], relates the 'strong' consistency of system $b \in \Phi(x)$ with the surjectivity of Φ_b. The *core* of a subset S of Y, written core S, is the set of those vectors b in S such that for any vector d in Y we have $b + rd \in S$ for all small real $r \geqslant 0$.

PROPOSITION 2.2 (Surjectivity and stability). *The homogenized process Φ_b is surjective if and only if b lies in the core of the range of Φ.*

Proof. Suppose Φ_b is surjective. Since there is a vector x and a real $t \geqslant 0$ such that $b \in \Phi(x) - tb$, certainly we have $b \in \Phi(X)$. On the other hand, for any vector d in Y there is a vector $z \in X$ and a real $s \geqslant 0$ such that $d \in \Phi(z) - sb$, so $d + sb \in \Phi(X)$. Now for any real $r \geqslant 0$ satisfying $rs \leqslant 1$ we have $b + rd = (1 - rs)b + r(d + sb) \in \Phi(X)$, so we deduce $b \in \mathrm{core}(\Phi(X))$.

Conversely, if Φ_b is not surjective, there is a vector d in Y satisfying $\Phi_b^{-1}(d) = \emptyset$, so $d + tb \notin \Phi(X)$ for all real $t \geqslant 0$. Hence $b + t^{-1}d \notin \Phi(X)$ for all real $t > 0$, so $b \notin \mathrm{core}(\Phi(X))$. □

The core is an algebraic notion of interior: to obtain more quantifiable results we need to introduce some norms.

3. Small Solutions

When does the inclusion $b \in \Phi(x)$ have a small solution $x \in X$? To quantify this question, we suppose now that X and Y are normed spaces (with closed unit balls

B_X and B_Y respectively) and that the convex process Φ is *closed* (which is to say Φ has closed graph). Following [5], we generalize the operator norm to processes by defining

$$\|\Phi\| = \inf\{0 < r \in \mathbf{R} : B_X \subset r\Phi^{-1}(B_Y)\}.$$

Notice $\|\Phi\|$ may take any value in the interval $[0, +\infty]$, and Φ is surjective exactly when $\|\Phi^{-1}\|$ is finite. The following simple result amounts to an equivalent definition for the norm.

PROPOSITION 3.1 (Process norm).

$$\|\Phi^{-1}\| = \sup_{\|b\|\leqslant 1} \inf_{x\in\Phi^{-1}(b)} \|x\|.$$

Proof. A real number s is no less than the right-hand side if and only if, for all vectors $b \in B_Y$ and all real $r > s$, there is a vector x in $\Phi^{-1}(b) \cap rB_X$. Equivalently, $B_Y \subset \Phi(rB_X)$ for all $r > s$, so by definition, $\|\Phi^{-1}\| \leqslant s$. The result follows. □

We denote the quantity $\inf\{\|x\| : x \in \Phi^{-1}(b)\}$ by $\inf \|\Phi^{-1}(b)\|$. Throughout, we use the convention $+\infty \cdot 0 = 0$. For a given process Φ, the above result shows, for any vector b in Y, the inequality

$$\inf \|\Phi^{-1}(b)\| \leqslant \|\Phi^{-1}\| \|b\|, \tag{3.2}$$

and this bound is in some sense tight. However, for any particular vector b the bound may not be helpful: for example, it gives information only when Φ is surjective. Our goal is therefore to refine it.

We can accomplish this with the simple trick of considering the homogenized process Φ_b instead of Φ. We consider the product space $X \times \mathbf{R}$ with the norm

$$\|(x, t)\| = \|x\| + |t|. \tag{3.3}$$

LEMMA 3.4 $\|\Phi_b^{-1}\| \leqslant \|\Phi^{-1}\|$.

Proof. This follows from the fact $\Phi(B_X) \subset \Phi_b(B_{X\times\mathbf{R}})$. □

PROPOSITION 3.5 (Small solutions).

$$\inf \|\Phi^{-1}(b)\| \leqslant \|\Phi_b^{-1}\| \|b\| \leqslant \|\Phi^{-1}\| \|b\|.$$

Proof. Given any point (x, t) in $\Phi_b^{-1}(b)$, the point $(1+t)^{-1}x$ lies in $\Phi^{-1}(b)$ and has smaller norm. We deduce the inequality $\inf \|\Phi^{-1}(b)\| \leqslant \inf \|\Phi_b^{-1}(b)\|$, and the result now follows from inequality (3.2) and the previous lemma. □

For varying b, the above result is again tight, by Proposition 3.1. Furthermore, it really is a refinement of inequality (3.2). If, for example, the process Φ: $\mathbf{R} \to \mathbf{R}$

is defined by $\Phi(x) = \{x\}$ when $x \geqslant 0$ and is empty otherwise, then $\|\Phi^{-1}\| = +\infty$, whereas if $b = 1$ we obtain $\|\Phi_b^{-1}\| = 1$.

It is interesting to express the important quantity $\|\Phi_b^{-1}\|$ more explicitly. Consider, for a given vector y in Y, the system (in the unknowns t and x)

$$tb + y \in \Phi(x), \qquad \|x\| + t \leqslant 1, \quad 0 \leqslant t \in \mathbf{R}, \quad x \in X. \qquad P(y)$$

We adopt the conventions $1/0 = +\infty$ and $1/+\infty = 0$. The following result is a direct calculation from the definition of the process norm.

PROPOSITION 3.6. $\|\Phi_b^{-1}\|^{-1} = \inf\{\|y\| : \textit{system } P(y) \textit{ inconsistent}\}$.

In the next section we investigate this quantity further.

4. Distance to Inconsistency

Assume henceforth X and Y are Banach spaces. In this case the quantity $\|\Phi_b^{-1}\|^{-1}$ that appeared in the last section has an important alternative description using the following generalization of the Eckart–Young theorem [3, Thm 2.8]. We denote the space of continuous linear maps from X to Y by $L(X, Y)$.

THEOREM 4.1 (Distance to nonsurjectivity). *For any closed convex precess* Φ: $X \to Y$,

$$\|\Phi^{-1}\|^{-1} = \inf_{G \in L(X,Y)} \{\|G\| : \Phi + G \textit{ not surjective}\}.$$

In particular, if Φ is surjective then so is $\Phi + G$ for all small perturbations G in $L(X, Y)$.

To apply this result to the process Φ_b we consider maps G in $L(X \times \mathbf{R}, Y)$. Any such map has the form

$$G(x, t) = Ax - ty \quad (x \in X, t \in \mathbf{R})$$

for some map A in $L(X, Y)$ and vector y in Y, and a standard calculation shows the norm of this map is given by $\|G\| = \|A\| \vee \|y\|$ (where $\vee$ denotes max). In other words, $L(X \times \mathbf{R}, Y)$ is isomorphic to $L(X, Y) \times Y$ (with this norm). Hence we deduce the relationship

$$\|\Phi_b^{-1}\|^{-1} = \inf\{\|A\| \vee \|y\| : (\Phi + A)_{b+y} \text{ not surjective}\}. \qquad (4.2)$$

Our aim is now to relate this to the original inclusion $b \in \Phi(x)$ using Proposition 2.2 (Surjectivity and stability). It is therefore natural to consider the following quantity, measuring how close the inclusion is to inconsistency.

DEFINITION 4.3 (Distance to inconsistency).

$$\rho(\Phi, b) = \inf\{\|A\| \vee \|y\| : (\Phi + A)^{-1}(b + y) = \emptyset\}.$$

To connect this quantity with Equation (4.2) we need the following result.

PROPOSITION 4.4 *In the space* $L(X, Y) \times Y$ *we have*

$$\{(A, y) : (\Phi + A)_{b+y} \text{ not surjective}\} = \mathrm{cl}\{(A, y) : (\Phi + A)^{-1}(b + y) = \emptyset\}.$$

Proof. Theorem 4.1 (Distance to nonsurjectivity) shows that the surjectivity of a closed convex process is stable under small continuous linear perturbations. Hence the left-hand side is a closed set, which therefore contains the right-hand side by Proposition 2.2 (Surjectivity and stability).

On the other hand, for any pair (A, y) in the left-hand side, we know $b + y \notin \mathrm{core}((\Phi+A)(X))$, by Proposition 2.2. Hence there is a sequence $y^r \to y$ satisfying $(\Phi + A)^{-1}(b + y^r) = \emptyset$ for all indices r, which shows (A, y) belongs to the right-hand side. □

We can now state our main result.

THEOREM 4.5 (Surjectivity and consistency). *Suppose X and Y are Banach spaces and consider a given vector b in Y and a closed convex process* Φ: $X \to Y$. *The following four quantities are equal:*

(i) *the distance to inconsistency of the inclusion* $b \in \Phi(x)$ (*namely* $\rho(\Phi, b)$)*;*
(ii) *the distance to nonsurjectivity of the homogenized process* Φ_b*;*
(iii) $\|\Phi_b^{-1}\|^{-1}$*;*
(iv) $\inf\{\|y\|$: *system* $P(y)$ *is inconsistent*$\}$.

Proof. The equality of quantities (i) and (ii) follows from the previous result, that of (ii) and (iii) is Equation (4.2), and that of (iii) and (iv) is Proposition 3.6. □

Notice that whereas the distance to inconsistency (quantity (i) in the above result) involves perturbations to the process Φ, quantities (iii) and (iv) involve only the original process.

By comparing with Proposition 3.5 (Small solutions), we can now relate the distance to inconsistency to our original question of determining the minimal norm of solutions to inclusions.

COROLLARY 4.6 (Condition measure).

$$\inf \|\Phi^{-1}(b)\| \leqslant \frac{\|b\|}{\rho(\Phi, b)}.$$

Thus the quantity $\|b\|/\rho(\Phi, b)$ is a natural condition measure for the inclusion $b \in \Phi(b)$.

We end with an example.

EXAMPLE 4.7 (Conic linear systems). A *conic linear system* is a system of the form

$$b - Ax \in C_Y, \quad x \in C_X,$$

where $A\colon X \to Y$ is a continuous linear map and $C_X \subset X$ and $C_Y \subset Y$ are closed convex cones (see for example [2]). We can model any such system as a process inclusion $b \in \Phi(x)$ by defining

$$\Phi(x) = \begin{cases} Ax + C_Y & \text{if } x \in C_X, \\ \emptyset & \text{otherwise.} \end{cases} \tag{4.8}$$

In this case the system $P(y)$ becomes

$$tb - Ax + y \in C_Y, \qquad \|x\| + t \leqslant 1, \quad 0 \leqslant t \in \mathbf{R},\ x \in C_X.$$

Theorem 4.5 now shows that the infimum of $\|y\|$ such that the above system is inconsistent equals the distance to inconsistency of the original conic linear system, namely the infimum of $\|G\| \vee \|y\|$ over continuous linear maps $G\colon X \to Y$ and vectors y in Y such that the system

$$b + y - (A + G)x \in C_Y, \quad x \in C_X,$$

is inconsistent. This recaptures [4, Thm 1.3], without the assumption made there of reflexivity.

It is worth remarking that the processes we consider in this paper are considerably more general than those arising immediately from conic linear systems. To take a simple example, the closed convex process

$$x \in \mathbf{R} \mapsto \{z \in \mathbf{R} : 0 \leqslant z \leqslant x\}$$

cannot be written in the conic linear form (4.8).

Acknowledgements

Research partially supported by the Natural Sciences and Engineering Research Council of Canada. The author thanks the University of Montpellier II for their hospitality while this work was completed.

References

1. Borwein, J. M.: Norm duality for convex processes and applications, *J. Optim. Theory Appl.* **48** (1986), 53–64.
2. Freund, R. M. and Vera, J. R.: Some characterizations and properties of the ‘distance to ill-posedness’ and the condition measure of a conic linear system, *Math. Programming* **86** (1999), 225–260.

3. Lewis, A. S.: Ill-conditioned convex processes and conic linear systems, *Math. Oper. Res.* **24** (1999), 829–834.
4. Renegar, J.: Linear programming, complexity theory and elementary functional analysis, *Math. Programming* **70** (1995), 279–351.
5. Robinson, S. M.: Normed convex processes, *Trans. Amer. Math. Soc.* **174** (1972), 127–140.
6. Rockafellar, R. T.: *Convex Analysis*, Princeton Univ. Press, Princeton, NJ, 1970.
7. Stewart, G. W. and Sun, J.-G.: *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.