MATHEMATICS OF OPERATIONS RESEARCH
Vol. 24, No. 4, November 1999
Printed in U.S.A.

ILL-CONDITIONED CONVEX PROCESSES AND CONIC LINEAR SYSTEMS

A. S. LEWIS

We prove the smallest possible norm of a linear perturbation making a closed convex process nonsurjective is the inverse of the norm of the inverse process. This generalizes the fundamental property of the condition number of a linear map. We then apply this result to strengthen a theorem of Renegar measuring the size of perturbation necessary to make a conic linear system inconsistent.

1. Introduction. The *condition number* of an invertible linear operator $A: X \to X$ (where X is a Banach space) is the quantity $\|A\| \|A^{-1}\|$, where $\|\cdot\|$ denotes the operator norm:

$$\|A\| = \sup\{\|Ax\|:\ \|x\| = 1\}.$$

The condition number is large exactly when A is relatively close to a singular operator: specifically,

$$\inf\{\|G\|:\ A + G \text{ singular}\} = \|A^{-1}\|^{-1}. \tag{1.1}$$

This identity measures how close to inconsistent a linear equation of the form

$$Ax = b$$

is (for a given element b of X). In the finite-dimensional case, $\|A^{-1}\|^{-1}$ is just the smallest singular value of A.

Recently Renegar (1995b) studied a similar question for "conic linear systems" of the form

$$Ax \le b, \qquad x \ge 0, \tag{1.2}$$

where now A lies in $L(X, Y)$, the space of continuous linear maps from X to another Banach space Y (with b in Y) and the inequalities $\le$ and $\ge$ correspond to closed convex cones in Y and X respectively. He derived an analogous measure of the distance to inconsistency for such systems, and related this measure to the complexity of interior point methods for solving corresponding (generalized) linear programs. This type of analysis has since proved very productive in interior point theory: see Renegar (1995a), Freund and Vera (1998), Renegar (1996), Vera (1996, 1998), Peña (1998), and Nunez and Freund (1998).

An elegant tool for studying linear systems like (1.2) is the idea, due to Rockafellar (1970), of a *closed convex process* $\Phi: X \to Y$. This is a set-valued map whose *graph*

$$\mathrm{Gr}(\Phi) = \{(x, y) \in X \times Y:\ y \in \Phi(x)\}$$

Received June 24, 1998; revised May 26, 1999.
AMS 1991 subject classification. Primary: 15A12, 90C31; Secondary: 65F35, 46A30.
OR/MS subject classification. Primary: Programming/Nonlinear.
Key words. Condition number, conic linear system, distance to inconsistency, convex process, adjoint process, duality, open mapping theorem, surjectivity.

0364-765X/99/2404/0829/$05.00
1526-5471 electronic ISSN, © 1999, INFORMS

is a closed convex cone. The *inverse* $\Phi^{-1}: Y \to X$ is the closed convex process defined by

$$x \in \Phi^{-1}(y) \Leftrightarrow y \in \Phi(x).$$

Given a second closed convex process $\Psi: X \to Y$, the *sum* $\Phi + \Psi: X \to Y$ is the closed convex process defined by $(\Phi + \Psi)(x) = \Phi(x) + \Psi(x)$ for all x in X. Denoting the closed unit balls in X and Y by B_X and B_Y respectively, we follow Robinson (1972) in generalizing the operator norm to processes by defining a function $\Phi \mapsto \|\Phi\| \in [0, +\infty]$:

$$\|\Phi\| = \inf\{0 < r \in \mathbf{R}: B_X \subset r\Phi^{-1}(B_Y)\}. \tag{1.3}$$

This paper concisely proves the following generalization of identity (1.1):

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

This leads quickly to the central characterization in Renegar (1995b) of the distance to inconsistency of system (1.2). The new approach is concise, and dispenses with Renegar's reflexivity assumption. Our proof relies on the norm duality of convex processes developed in Borwein (1983), a homogenization approach pursued in Borwein (1986) and Renegar (1995b), and a rank-one perturbation technique used in Peña (1998).

2. Distance to nonsurjectivity. Throughout this paper, X and Y are Banach spaces and $\Phi: X \to Y$ is a closed convex process.

There is an attractive duality between the "norm" defined by equation (1.3) and the "norm" $n(\cdot)$ defined by

$$n(\Phi) = \inf\{0 < r \in \mathbf{R}: \Phi(B_X) \subset rB_Y\}$$

(see Borwein 1983). It relies on the notion of the *adjoint* $\Phi^*: Y^* \to X^*$, a closed convex process from the dual space to Y to the dual space to X, defined by

$$\mu \in \Phi^*(\nu) \Leftrightarrow \langle \nu, y \rangle \geq \langle \mu, x \rangle \quad \text{whenever } y \in \Phi(x)$$

(see Rockafellar 1970). The key *norm duality* result (Theorem 10.1 in Borwein 1983) on which our argument relies is

$$\|\Phi\| = n(\Phi^*). \tag{2.1}$$

We say Φ is *surjective* if $\Phi(X) = Y$, and Φ is *bounded* if $n(\Phi)$ is finite. These ideas are elegantly related by the *open-mapping/closed-graph theorem* (see Robinson 1976, Ursescu 1975 and Borwein 1986):

$$\Phi \text{ surjective} \Leftrightarrow 0 \in \operatorname{int} \Phi(B_X) \Leftrightarrow \Phi^{*-1} \text{ bounded}. \tag{2.2}$$

Part of our central result consists of the claim that $\|\Phi^{-1}\|^{-1}$ is a lower bound for the norm of a linear perturbation making Φ nonsurjective. This result is in fact a special case of Theorem 5 in Robinson (1976): for completeness we give a concise proof based on norm duality. We begin with an easy result showing a stability property of bounded processes.

LEMMA 2.3 (STABILITY OF BOUNDEDNESS). *If the map $G \in L(X, Y)$ satisfies $\|G\| < (n(\Phi^{-1}))^{-1}$ then the process $(\Phi + G)^{-1}$ is bounded.*

PROOF. If $(\Phi + G)^{-1}$ is not bounded then there is a sequence (y_r) in $\operatorname{int} B_Y$ and corresponding points $x_r \in (\Phi + G)^{-1}(y_r)$ satisfying $\|x_r\| \to \infty$. Since

$$x_r \in \Phi^{-1}(y_r - Gx_r) \subset n(\Phi^{-1})(1 + \|G\|\,\|x_r\|)B_X,$$

we deduce

$$\|x_r\| \leq n(\Phi^{-1})(1 + \|G\| \, \|x_r\|),$$

and letting $r \to \infty$ we see $n(\Phi^{-1})\|G\| \geq 1$. □

We next combine this stability property with the open-mapping/closed-graph theorem to deduce a lower bound on the distance to nonsurjectivity for a closed convex process, specializing Theorem 5 in Robinson (1976).

LEMMA 2.4 (STABILITY OF SURJECTIVENESS). *If the map $G \in L(X, Y)$ satisfies $\|G\| < \|\Phi^{-1}\|^{-1}$ then the process $\Phi + G$ is surjective. In particular, if the process Φ is surjective then so is $\Phi + G$ for all small G in $L(X, Y)$.*

PROOF. We know, by norm duality (2.1), that

$$\|G\| < \|\Phi^{-1}\|^{-1} = (n(\Phi^{-1*}))^{-1}.$$

But direct calculation (see (3.4.5) in Borwein 1983) shows that

$$\Phi^{-1*}(y) = -\Phi^{*-1}(-y), \quad \text{for all } y \text{ in } Y, \tag{2.5}$$

and since $\|G\| = \|G^*\|$ we deduce

$$\|G^*\| < (n(\Phi^{*-1}))^{-1}.$$

Hence by the previous lemma the process $(\Phi^* + G^*)^{-1}$ is bounded. Direct calculation shows $\Phi^* + G^* = (\Phi + G)^*$ (c.f. Theorem 7.4 in Borwein 1983), so $\Phi + G$ is surjective by the open-mapping/closed-graph Theorem (2.2). The same reasoning shows $\|\Phi^{-1}\|$ is finite if Φ is surjective, and the final comment follows. □

We denote the polar of a set $C \subset Y$ by

$$C^\circ = \{v \in Y^* : \langle v, y \rangle \leq 1 \text{ for all } y \text{ in } C\}.$$

Our next step depends on the underlying identity in the proof of norm duality, namely

$$(\Phi(B_X))^\circ = -\Phi^{*-1}(B_{X^*}) \tag{2.6}$$

(see Proposition 1.1 in Borwein 1986). We use this to show how surjective processes can be perturbed to nonsurjectivity by adding rank-one linear maps.

LEMMA 2.7 (RANK-ONE PERTURBATION). *For any point y not in* $\operatorname{int} \Phi(B_X)$ *there is a rank-one map G in $L(X, Y)$ satisfying $\|G\| \leq \|y\|$ with $\Phi + G$ not surjective.*

PROOF. If Φ is not surjective there is nothing to prove, so assume Φ is surjective, and hence $0 \in \operatorname{int} \Phi(B_X)$ by the open-mapping/closed-graph Theorem (2.2). By the Hahn-Banach theorem there is an element v of $(\Phi(B_X))^\circ$ satisfying $\langle v, y \rangle \geq 1$.

The polarity identity (2.6) shows there is an element μ of $B_{X^*} \cap \Phi^*(-v)$. Now define a rank-one map $G: X \to Y$ by

$$Gx = \frac{\langle \mu, x \rangle}{\langle v, y \rangle} y \qquad (x \in X).$$

It is easy to check $\|G\| \leq \|y\|$. Finally, notice for any points x in X and z in $\Phi(x)$ we have

$$\langle v, z + Gx \rangle = \langle v, z \rangle + \langle \mu, x \rangle \leq 0,$$

since μ lies in $\Phi^*(-v)$. Thus $\Phi + G$ is not surjective. □

We can now prove our main result.

THEOREM 2.8 (DISTANCE TO NONSURJECTIVITY). *For any Banach spaces X and Y, if $\Phi: X \to Y$ is a closed convex process then*

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

NOTE. We interpret the right-hand side as 0 or $+\infty$ if $\|\Phi^{-1}\|$ is $+\infty$ or 0 respectively.

PROOF. Call the left-hand side α. If $\|\Phi^{-1}\| = 0$ then $\Phi(B_X) = Y$ so $\alpha = +\infty$. On the other hand, $\|\Phi^{-1}\| = +\infty$ implies $0 \notin \operatorname{int} \Phi(B_X)$, in which case Φ is not surjective (by the open-mapping/closed-graph Theorem (2.2)) so $\alpha = 0$. This covers the two extreme cases, so we can assume $0 < \alpha < +\infty$.

Lemma 2.4 (stability of surjectiveness) proves $\|\Phi^{-1}\|^{-1} \le \alpha$, so suppose $\|\Phi^{-1}\| > \alpha^{-1}$, and choose a number β in $(\|\Phi^{-1}\|^{-1}, \alpha)$. By definition we know $B_Y \not\subset \beta^{-1}\Phi(B_X)$, so there is a point y outside $\Phi(B_X)$ satisfying $\|y\| \le \beta$. By Lemma 2.7 (rank-one perturbation) there is a map G in $L(X,Y)$ satisfying $\|G\| \le \beta$ with $\Phi + G$ not surjective, giving the contradiction $\alpha \le \beta$. □

NOTES. (1). The same results hold when G is a closed convex process with nonempty images $G(x)$ (for all points x in X). The only change necessary is to replace $\|G\|$ by $n(G)$ in Lemma 2.3 (stability of boundedness). Hence we see

$$\begin{aligned} \|\Phi^{-1}\|^{-1} &= \inf\{\|G\| \mid G \text{ has nonempty images, } \Phi + G \text{ not surjective}\} \\ &= \inf\{\|G\| \mid G \text{ linear, } \Phi + G \text{ not surjective}\} \\ &= \inf\{\|G\| \mid G \text{ rank-one, } \Phi + G \text{ not surjective}\}. \end{aligned}$$

(2). When Y is finite-dimensional the infimum in this result is attained (when finite) by a rank-one linear map G. To see this, take the point y in Lemma 2.7 (rank-one perturbation) to be a point on the boundary of $\Phi(B_X)$ of minimal norm.

3. Distance to inconsistency.

Given closed convex cones $K_X \subset X$ and $K_Y \subset Y$, we write $x \le z$ for points x and z in X to mean $z - x \in K_X$, with an analogous definition in the space Y. In this section we investigate how much we need to perturb a map A in $L(X,Y)$ and a point b in Y to render the system

$$(3.1) \qquad Ax \le b, \qquad 0 \le x \in X$$

inconsistent.

Consider the product space $X \times \mathbf{R}$ with the norm

$$\|(x,t)\| = \|x\| + \|t\| \qquad (x \in X,\ t \in \mathbf{R}).$$

Any map in $L(X \times \mathbf{R}, Y)$ has the form

$$(x,t) \mapsto Ax - tb \qquad (x \in X,\ t \in \mathbf{R})$$

for some map A in $L(X,Y)$ and point b in Y, and it is easy to check that this map has norm

$$\|A\| \vee \|b\| \qquad (= \max\{\|A\|, \|b\|\}):$$

in other words, $L(X \times \mathbf{R}, Y)$ is isomorphic to the product space $L(X,Y) \times Y$ (with this norm).

Define the *inconsistent set*

$$I = \{(A, b) \in L(X, Y) \times Y: \text{ system (3.1) inconsistent}\}.$$

Our aim is to calculate the *distance to inconsistency*

$$\text{dist}_I(A, b) = \inf\{\|\tilde{A} - A\| \vee \|\tilde{b} - b\|: (\tilde{A}, \tilde{b}) \in I\}.$$

We can associate the system (3.1) with a certain closed convex process $\Phi_{A,b} : X \times \mathbf{R} \to Y$ defined by

$$\Phi_{A,b}(x, t) = \begin{cases} Ax - tb + K_Y & \text{if } x \in K_X \text{ and } t \geq 0, \\ \emptyset & \text{otherwise,} \end{cases}$$

and there is an associated set

$$I_0 = \{(A, b) \in L(X, Y) \times Y: \Phi_{A,b} \text{ not surjective}\}.$$

LEMMA 3.2 (INCONSISTENCY AND NONSURJECTIVITY). $I_0 = \text{cl}\, I$.

PROOF. If the pair $(A, b) \in L(X, Y) \times Y$ lies outside I_0 then by definition the process $\Phi_{A,b}$ is surjective. Hence the process $\Phi_{\tilde{A},\tilde{b}}$ is also surjective for all pairs $(\tilde{A}, \tilde{b})$ close to (A, b) by Lemma 2.4 (stability of surjectiveness). For any such $(\tilde{A}, \tilde{b})$ there is a point x in K_X and real $t \geq 0$ satisfying

$$\tilde{b} \in \tilde{A}x - t\tilde{b} + K_Y,$$

so we have

$$\tilde{A}(1 + t)^{-1}x \leq \tilde{b}, \qquad 0 \leq (1 + t)^{-1}x \in X,$$

and hence $(\tilde{A}, \tilde{b}) \notin I$. This shows $\text{dist}_I(A, b) > 0$.

Conversely, if (A, b) lies in I_0 then $\Phi_{A,b}$ is not surjective, so there is a point y in Y outside $\Phi_{A,b}(X \times \mathbf{R})$. Consequently the system

$$Ax \leq b + \varepsilon y, \qquad 0 \leq x \in X$$

is inconsistent for all real $\varepsilon > 0$, so $\text{dist}_I(A, b) = 0$. This completes the proof. □

An important "distance to inconsistency" result (Theorem 1.3 in Renegar 1995b) assumes the space X is reflexive and relates the system (3.1) to the system

$$\begin{aligned} Ax &\leq tb + z, & \|x\| + |t| &\leq 1, \\ 0 &\leq x \in X, & 0 &\leq t \in \mathbf{R}, \end{aligned} \tag{3.3}$$

for a given point z in X. We now recapture this result, without the reflexivity assumption.

THEOREM 3.4 (DISTANCE TO INCONSISTENCY). *Perturbed systems*

$$\tilde{A}x \leq \tilde{b}, \qquad 0 \leq x \in X \tag{3.5}$$

have the property

$$\inf\{\|\tilde{A} - A\| \vee \|\tilde{b} - b\|: \textit{system } (3.5) \textit{ inconsistent}\} = \inf\{\|z\|: \textit{system } (3.3) \textit{ inconsistent}\}.$$

PROOF. By Lemma 3.2 (inconsistency and nonsurjectivity), the left-hand side is just $\text{dist}_I(A, b) = \text{dist}_{I_0}(A, b)$. By Theorem 2.8 (distance to nonsurjectivity) this is just $\|\Phi_{A,b}^{-1}\|^{-1}$, and a simple calculation from the definition shows this is exactly the right-hand side. □

Notice how one side of this identity considers perturbations to both the constraint map A and the right-hand side vector b, while the other side of the identity leaves A fixed. The case $b = 0$ and $K_Y = \{0\}$ is particularly simple.

Acknowledgement. Thanks to Jim Renegar and Javier Peña for stimulating my interest in this problem, and to Jon Borwein for helpful suggestions. Research was partially supported by the National Sciences and Engineering Research Council of Canada. The author thanks the School of Operations Research and Industrial Engineering at Cornell University, where this work was completed.

References

Borwein, J. M. 1983. Adjoint process duality. *Math. Oper. Res.* **8** 403–434.

——, 1986. Norm duality for convex processes and applications. *J. Optim. Theory Appl.* **48** 53–64.

Freund, R. M., J. R. Vera. 1998. Some characterizations and properties of the "distance to ill-posedness" and the condition number of a conic linear system. Technical Report 3862-95-MSA, Sloan School of Management, Massachusetts Institute of Technology, Boston, MA.

Nunez, M. A., R. M. Freund. 1998. Condition measures and properties of the central trajectory of a linear program. *Math. Programming* **83** 1–28.

Peña, J. 1998. Understanding the geometry of infeasible perturbations of a conic linear system. Technical Report, Cornell University, Forthcoming in *SIAM J. Optim.*

Renegar, J. 1995a. Incorporating condition measures into the complexity theory of linear programming. *SIAM J. Optim.* **5** 506–524.

——, 1995b. Linear programming, complexity theory and elementary functional analysis. *Math. Programming* **70** 279–351.

——, 1996. Condition numbers, the barrier method and the conjugate gradient method. *SIAM J. Optim.* **6** 879–912.

Robinson, S. M. 1972. Normed convex processes. *Trans. Amer. Math. Soc.* **174** 127–140.

——, 1976. Regularity and stability for convex multivalued functions. *Math. Oper. Res.* **1** 130–143.

Rockafellar, R. T. 1970. *Convex Analysis.* Princeton University Press, Princeton, N.J.

Ursescu, C. 1975. Multifunctions with convex closed graph. *Czech. Math. J.* **7** 438–441.

Vera, J. R. 1996. Ill-posedness and the complexity of deciding existence of solutions to linear programs. *SIAM J. Optim.* **6** 549–569.

——, On the complexity of linear programming under finite precision arithmetic. *Math. Programming* **80** 91–123.

A. S. Lewis: Department of Combinatories and Optimization, University of Waterloo, Waterloo, Ontario, Canada N2L 3Gl; e-mail: aslewis@orion.uwaterloo.ca