

A CONDITION NUMBER FOR
DIFFERENTIABLE CONVEX INEQUALITIES

by

O. L. Mangasarian

Computer Sciences Technical Report #491

December 1982

UNIVERSITY OF WISCONSIN-MADISON
COMPUTER SCIENCES DEPARTMENT

A CONDITION NUMBER FOR DIFFERENTIABLE CONVEX INEQUALITIES

O. L. Mangasarian

Technical Report #491

ABSTRACT

For a system of differentiable convex inequalities, a new bound is given for the absolute error in an infeasible point in terms of the absolute residual. By using this bound a condition number is defined for the system of inequalities which gives a bound for the relative error in an infeasible point in terms of the relative residual.

AMS (MOS) Subject Classifications: 90C25, 65F35

Key Words: Convex programming, inequalities, condition number

A CONDITION NUMBER FOR DIFFERENTIABLE CONVEX INEQUALITIES

O. L. Mangasarian

1. Introduction

For a system of linear equalities

$$(1.1) \quad Ax = b$$

where A is a given $n \times n$ real nonsingular matrix and b is a given nonzero vector in the n -dimensional real Euclidean space R^n , the norm of the inverse $\|A^{-1}\|$ and the condition number $\|A^{-1}\| \|A\|$ provide the following useful bounds for the absolute error $\|x - \bar{x}\|$ in terms of the residual $\|Ax - b\|$, and for the relative error $\frac{\|x - \bar{x}\|}{\|\bar{x}\|}$ in terms of the relative residual $\frac{\|Ax - b\|}{\|b\|}$ [1,7]

$$(1.2) \quad \|x - \bar{x}\| \leq \|A^{-1}\| \|Ax - b\|$$

$$(1.3) \quad \frac{\|x - \bar{x}\|}{\|\bar{x}\|} \leq \|A^{-1}\| \|A\| \frac{\|Ax - b\|}{\|b\|}$$

Here, x is any point in R^n , \bar{x} is the exact solution $A^{-1}b$ and $\|\cdot\|$ denotes a vector norm on R^n or its subordinate matrix norm [1,7]. The condition number $\|A^{-1}\| \|A\|$, which depends on the specific norm employed and which is never less than 1, provides a very useful stability measure for the system (1.1). It is the purpose of this work to obtain a corresponding number for the system of inequalities

$$(1.4) \quad g(x) \leq 0$$

where $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is differentiable and convex on \mathbb{R}^n . The key to obtaining a condition number for (1.4) is the definition of a quantity $\mu_{\beta\gamma}$ of (2.4) that plays the same role as that of $\|A^{-1}\|$ for (1.1) and which would provide an error bound (2.3) similar to (1.2). In [3] Hoffman extended the bound (1.2) to a system of linear inequalities, and in [6] a new explicit expression was derived for that bound for a system of linear inequalities and equalities. In [8] Robinson extended the bound (1.2) to a system of convex inequalities that define a bounded feasible region with a nonempty interior. In Section 2 of this paper we shall extend the bound (1.2) to a system of differentiable convex inequalities which satisfy a constraint qualification, but without any boundedness assumption on the feasible region. The diameter of the bounded feasible region which appears linearly in Robinson's bound [8, equation (4)] does not appear in our bound (2.3). In Section 3 of this paper we employ the results of Section 2 to obtain a condition number for (1.4) and thereby extend the relative error bound (1.3) to a system of differentiable convex inequalities satisfying a constraint qualification. We remark that the error bound coefficient $\mu_{\beta\gamma}$ is in general difficult to compute. However its existence and finiteness under the assumptions of Theorem 2.1 is theoretically significant. This parallels the situation of $\|A^{-1}\|$ which is also nontrivial to compute in general but has some very useful properties.

We briefly describe now the notation and some of the basic concepts used in this work. For a vector x in the n -dimensional real Euclidean space \mathbb{R}^n , $|x|$ and x_+ will denote the vectors in \mathbb{R}^n with components $|x|_i = |x_i|$ and $(x_+)_i = \max\{x_i, 0\}$, $i=1,2,\dots,n$, respectively. For

a norm $\|x\|_\beta$ on R^n , $\|x\|_{\beta^*}$ will denote the dual norm on R^n , that is $\|x\|_{\beta^*} = \max_{\|y\|_\beta=1} xy$, where xy denotes the scalar product. The generalized Cauchy-Schwarz inequality $|xy| \leq \|x\|_\beta \|y\|_{\beta^*}$, for x and y in R^n , follows immediately from this definition of the dual norm. For $1 \leq p, q \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ the p -norm $(\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ and the q -norm are dual norms on R^n [4]. If $\|\cdot\|_\beta$ is a norm on R^n , we shall, with a slight abuse of notation, let $\|\cdot\|_\beta$ also denote the corresponding norm on R^m for $m \neq n$. For an $m \times n$ real matrix A , A_i denotes the i th row and $A_{.j}$ denotes the j th column, while $\|A\|_\beta$ denotes the matrix norm [1,7] subordinate to the vector norm $\|\cdot\|_\beta$, that is $\|A\|_\beta = \max_{\|x\|_\beta=1} \|Ax\|_\beta$. The consistency condition $\|Ax\|_\beta \leq \|A\|_\beta \|x\|_\beta$ follows immediately from this definition of a matrix norm. We shall also use $\|\cdot\|$ to denote an arbitrary vector norm and its subordinate matrix norm. A vector norm $\|\cdot\|$ on R^n is said to be monotonic if and only if $|x_i| \leq |y_i|, i=1, \dots, n$, implies that $\|x\| \leq \|y\|$ [7]. The p -norms $1 \leq p \leq \infty$, are all monotonic norms [7]. A vector of ones in any real Euclidean space will be denoted by e . For a differentiable function $g: R^n \rightarrow R^m$, $\nabla g(x)$ will denote the $m \times n$ Jacobian matrix evaluated at the point x in R^n . For a subset $I \subset \{1, 2, \dots, m\}$, $g_I(x)$ or $g_{i \in I}(x)$ will denote those components of $g_i(x)$ such that $i \in I$. Similarly $\nabla g_I(x)$ will denote the rows $(\nabla g(x))_i$ of $\nabla g(x)$ such that $i \in I$.

2. An Absolute Error Bound for Differentiable Convex Inequalities

We shall use the approach of [6] to obtain a bound (2.3) for the absolute error $\|x-p(x)\|_\gamma$, where x is any infeasible point for (1.4) and $p(x)$ is some feasible point for (1.4), in terms of the absolute residual $\|g(x)_+\|_\beta$. The constant $\mu_{\beta\gamma}$ relating these two quantities plays a similar role for the differentiable convex inequalities (1.4) as $\|A^{-1}\|$ does for the linear system (1.1).

2.1 Theorem Let g be a differentiable convex function from R^n into R^m , let S° and S be defined by

$$(2.1) \quad \phi \neq S^\circ := \{x | g_{J_1}(x) < 0, g_{J_2}(x) \leq 0\} \subset S := \{x | g(x) \leq 0, x \in R^n\}$$

where $J_1 \cup J_2$ is a partition of $\{1, \dots, m\}$ such that g_{J_1} is nonlinear and g_{J_2} is linear and let the following asymptotic constraint qualification hold

$$(2.2) \quad \left\{ \begin{array}{l} \text{For each nonempty } I \subset \{1, \dots, m\} \text{ and each sequence of points} \\ \{x^i\} \subset S \text{ such that } g_I(x^i) = 0 \text{ and } \nabla g_{j \in I}(x^i) \text{ are linearly inde-} \\ \text{pendent, each accumulation point } (\overline{\nabla}g_{I_0}, \overline{\nabla}g_{I_1}, \overline{\nabla}g_{I_2}) \text{ of the sequence} \\ \{\nabla g_{j \in I_0}(x^i) / \|\nabla g_{j \in I_0}(x^i)\|, \nabla g_{I_1}(x^i), \nabla g_{I_2}(x^i)\} \text{ satisfies} \\ \overline{\nabla}g_{I_0} z > 0, \overline{\nabla}g_{I_1} z > 0, \overline{\nabla}g_{I_2} z \geq 0 \text{ for some } z \in R^n \\ \text{where } I_0 \cup I_1 \cup I_2 \text{ is a partition of } I \text{ such that the sequence} \\ \{\nabla g_j(x^i)\} \text{ is unbounded for } j \in I_0 \text{ and bounded for } j \in I_1, g_{j \in I_1} \\ \text{is nonlinear and } g_{j \in I_2} \text{ is linear.} \end{array} \right.$$

Then for each x in R^n there exists a point $p(x)$ in S such that

$$(2.3) \quad \|x-p(x)\|_\gamma \leq \mu_{\beta\gamma} \|g(x)_+\|_\beta$$

where $\mu_{\beta\gamma}$ is a constant independent of x and defined by

$$(2.4) \quad \mu_{\beta\gamma} := \alpha_{\gamma\infty} \sup_{w,p,I} \{ \|w_I\|_{\beta^*} \mid p \in S, w_I > 0, g_I(p) = 0, \|\nabla g_I(p)\|_1 = 1 \\ \nabla g_{j \in I}(p) \text{ lin. indep., } I \subset \{1, \dots, m\} \}$$

where $\|\cdot\|_{\beta^*}$ is the dual norm to $\|\cdot\|_{\beta}$ and $\alpha_{\gamma\infty}$ is the positive constant relating the γ -norm and ∞ -norm by

$$\|z\|_{\gamma} \leq \alpha_{\gamma\infty} \|z\|_{\infty} \text{ for all } z \in \mathbb{R}^n$$

Proof We first establish the finiteness of $\mu_{\beta\gamma}$ under the asymptotic constraint qualification (2.2). For if $\mu_{\beta\gamma}$ is infinite then there exists

a fixed nonempty $I \subset \{1, \dots, m\}$ and a sequence $\{(w_I^i, p^i)\}$ such that $\|w_I^i\| \rightarrow \infty$, $p^i \in S$, $w_I^i > 0$, $g_I(p^i) = 0$, $\|w_I^i \nabla g_I(p^i)\|_1 = 1$ and $\nabla g_{j \in I}(p^i)$ are linearly independent. Let the sequence $\{\nabla g_j(p^i)\}$, $j \in I_0$, be unbounded (hence $g_{j \in I_0}$ is nonlinear), let $\{\nabla g_j(p^i)\}$, $j \in I_1$, be bounded and

$g_{j \in I_1}$ nonlinear, and let $g_{j \in I_2}$ be linear. Since

$$1 = \|w_I^i \nabla g_I(p^i)\|_1 = \left\| \sum_{j \in I_0} w_j^i \|\nabla g_j(p^i)\| \cdot \nabla g_j(p^i) / \|\nabla g_j(p^i)\| + \sum_{j \in I_1 \cup I_2} w_j^i \nabla g_j(p^i) \right\|_1$$

then the sequences

$$\left\{ \frac{w_{j \in I_0}^i \|\nabla g_{j \in I_0}(p^i)\|, w_{j \in I_1 \cup I_2}^i}{\|w_{j \in I_0}^i \cdot \|\nabla g_{j \in I_0}(p^i)\|, w_{j \in I_1 \cup I_2}^i\|} \right\}, \left\{ \frac{\nabla g_{j \in I_0}(p^i)}{\|\nabla g_{j \in I_0}(p^i)\|} \right\}, \left\{ \nabla g_{j \in I_1 \cup I_2}(p^i) \right\}$$

have respective accumulation points $(\bar{w}_{I_0}, \bar{w}_{I_1}, \bar{w}_{I_2})$, $\bar{\nabla}g_{I_0}$, and $(\bar{\nabla}g_{I_1}, \bar{\nabla}g_{I_2})$

which, since $\{\|w_{j \in I_0}^i \cdot \|\nabla g_{j \in I_0}(p^i)\|, w_{j \in I_1 \cup I_2}^i\| \} \rightarrow \infty$, satisfy

$$0 = \bar{w}_{I_0} \bar{\nabla}g_{I_0} + \bar{w}_{I_1} \bar{\nabla}g_{I_1} + \bar{w}_{I_2} \bar{\nabla}g_{I_2}, (\bar{w}_{I_0}, \bar{w}_{I_1}, \bar{w}_{I_2}) \geq 0, \|\bar{w}_{I_0}, \bar{w}_{I_1}, \bar{w}_{I_2}\| = 1$$

If $I_0 \cup I_1 = \emptyset$ or $(\bar{w}_{I_0}, \bar{w}_{I_1}) = 0$ then we have a contradiction to the linear independence of $\bar{\nabla}g_{j \in I_2}$. If $I_0 \cup I_1 \neq \emptyset$ and $0 \neq (\bar{w}_{I_0}, \bar{w}_{I_1}) \geq 0$, then we contradict the asymptotic constraint qualification that $\bar{\nabla}g_{I_0} z > 0$, $\bar{\nabla}g_{I_1} z > 0$, $\bar{\nabla}g_{I_2} z \geq 0$ for some $z \in \mathbb{R}^n$. Hence $\mu_{\beta\gamma}$ is finite.

Now for any x in R^n not in S define $p(x)$ as the projection of x on S using the ∞ -norm. Hence $p(x)$ and some $\delta(x) > 0$ solve the following convex programming problem

$$(2.5) \quad \min_{p, \delta} \{ \delta \mid -e\delta \leq p-x \leq e\delta, g(p) \leq 0, p \in R^n, \delta \in R \}$$

Since by (2.1) the inequalities $-e\delta \leq p-x \leq e\delta$, $g_{j_1}(p) < 0$, $g_{j_2}(p) \leq 0$ have a solution it follows that $p(x)$, $\delta(x)$ and some $u(x)$, $v(x)$ and $w(x)$ satisfy the following Karush-Kuhn-Tucker conditions [5] for (2.5), where the explicit dependence of p , δ , u , v and w on x has been dropped for simplicity,

$$(2.6) \quad \begin{cases} 1 - e(u+v) = 0, u - v + w\nabla g(p) = 0, u, v, w \geq 0 \\ -e\delta \leq p - x \leq e\delta, g(p) \leq 0 \\ wg(p) = 0, u_j(p-x-e\delta)_j = 0, v_j(-e\delta-p+x)_j = 0, j=1, \dots, n \end{cases}$$

From the last two equalities it follows that $u_j v_j = 0$ for $j=1, \dots, n$, because

$$(p-x-e\delta)_j + (-e\delta-p+x)_j = -2\delta < 0$$

Consequently

$$p \in S, wg(p) = 0, w \geq 0, \|w\nabla g(p)\|_1 = 1$$

It follows by the fundamental theorem of basic feasible solutions [2, Theorem 2.11] that there exists a nonempty $I \subset \{1, \dots, m\}$ such that (2.6) is satisfied and

$$(2.7) \quad p \in S, w_I > 0, g_I(p) = 0, \|w_I \nabla g_I(p)\|_1 = 1, \nabla g_{j \in I}(p) \text{ lin. indep.}, w_{i \notin I} = 0$$

Hence by (2.5) and (2.6) it follows that

$$\begin{aligned} 0 &< \|x-p\|_\infty = \delta \\ &= \delta + u(p-x-e\delta) + v(-e\delta-p+x) + wg(p) \\ &= x(v-u) + wg(p) - w\nabla g(p)p \\ &= w\nabla g(p)x + wg(p) - w\nabla g(p)p \\ &\leq wg(x) \quad (\text{By convexity of } g \text{ and } w \geq 0) \\ &\leq wg(x)_+ \end{aligned}$$

$$\begin{aligned} &\leq \|w\|_{\beta^*} \|g(x)_+\|_{\beta} \quad (\text{By Cauchy-Schwarz inequality}) \\ &\leq \frac{\mu_{\beta\gamma}}{\alpha_{\gamma\infty}} \|g(x)_+\|_{\beta} \quad (\text{By (2.4) and (2.7)}) \end{aligned}$$

Hence

$$\|x-p\|_{\gamma} \leq \alpha_{\gamma\infty} \|x-p\|_{\infty} \leq \mu_{\beta\gamma} \|g(x)_+\|_{\beta} \quad \square$$

2.2 Remark The asymptotic constraint qualification (2.2) is needed, only when J_1 is nonempty, as a sufficient condition for the finiteness of $\mu_{\beta\gamma}$ as defined by (2.4). Hence Theorem 2.1 can be stated for the case when J_1 is nonempty with the constraint qualification (2.2) replaced by the assumption that the supremum of (2.4) defining $\mu_{\beta\gamma}$ is finite.

When g is linear, that is J_1 is empty, $\mu_{\beta\gamma}$ is automatically finite under no additional assumption as was the case in [6, Theorem 1], because (2.2) holds trivially with $z = 0$.

We give now a simple example illustrating the above theorem which is not covered by Robinson's result [8, equation (4)] because the feasible region is unbounded.

2.3 Example $S := \{x \mid x \in \mathbb{R}^2, x_2 \geq e^{x_1}, x_1 \geq 0\}.$

It is easy to verify that the assumptions of Theorem 2.1 are satisfied with $g_1(x) := e^{x_1} - x_2$, $g_2(x) = -x_1$ and that

$$\mu_{\infty\infty} := \sup_{w, p, I} \{ \|w_I\|_1 \mid p \in S, w_I > 0, g_I(p) = 0, \|w_I \nabla g_I(p)\|_1 = 1, I \subset \{1, 2\} \} = 2$$

Consequently for each x in \mathbb{R}^2 there exists a $p(x)$ in S such that

$$\|x-p(x)\|_{\infty} \leq 2 \left\| \begin{pmatrix} (e^{x_1} - x_2)_+ \\ (-x_1)_+ \end{pmatrix} \right\|_{\infty}$$

The bound $\mu_{\infty\infty} = 2$ is sharp here, for take the sequence of points $x_1 = -t$, $x_2 = 1 - 2t$ with t a nonnegative number converging to zero.

Then

$$\frac{\|x-p(x)\|_{\infty}}{\|g(x)_{+}\|_{\infty}} = \frac{\|-t, -2t\|_{\infty}}{e^{-t}-1+2t} = \frac{2t}{e^{-t}-1+2t}$$

which approaches 2 as t approaches 0.

3. A Relative Error Bound for Differentiable Convex Inequalities

We extend now the relative error bound (1.3) to convex differentiable inequalities by using the error bound (2.3). We will again use the approach of [6] and will need the following simple lemma established there.

3.1 Lemma [6, Lemma 2] Let $\|\cdot\|_\beta$ be a monotonic norm on \mathbb{R}^m and let a, b be in \mathbb{R}^m . Then $a \leq b$ implies that $\|(a)_+\|_\beta \leq \|b\|_\beta$.

The following is a direct consequence of the above lemma.

3.2 Lemma Let $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be differentiable and convex on \mathbb{R}^n and let $\|\cdot\|_\beta$ be a monotonic norm on \mathbb{R}^m . Then

$$\|g(0)_+\|_\beta \leq \|\nabla g(0)p\|_\beta \quad \text{for } g(p) \leq 0$$

Proof By the convexity of g , $0 \geq g(p) \geq g(0) + \nabla g(0)p$. Hence applying Lemma 3.1 to $g(0) \leq -\nabla g(0)p$ we obtain the desired inequality. \square

3.3 Theorem (Condition number bound) Let the assumptions of Theorem 2.1 hold, let $g(0)_+ \neq 0$ and let $\|\cdot\|_\beta$ be a monotonic norm on \mathbb{R}^m . Then for each x in \mathbb{R}^n there exists a $p(x)$ in S such that

$$(3.1) \quad \frac{\|x-p(x)\|_\beta}{\|p(x)\|_\beta} \leq \mu_{\beta\beta} \|\nabla g(0)\|_\beta \frac{\|g(x)_+\|_\beta}{\|g(0)_+\|_\beta}$$

where $\mu_{\beta\beta}$ is defined by (2.4), and $\mu_{\beta\beta} \|\nabla g(0)\|_\beta$ defines the condition number of (1.4).

Proof For each x in \mathbb{R}^n there exists a $p(x)$ in S such that

$$\frac{\|x-p(x)\|_{\beta}}{\|p(x)\|_{\beta}} \leq \mu_{\beta\beta} \frac{\|g(0)_+\|_{\beta}}{\|p(x)\|_{\beta}} \frac{\|g(x)_+\|_{\beta}}{\|g(0)_+\|_{\beta}} \quad (\text{By Theorem 2.1})$$

$$\leq \mu_{\beta\beta} \frac{\|\nabla g(0)p(x)\|_{\beta}}{\|p(x)\|_{\beta}} \frac{\|g(x)_+\|_{\beta}}{\|g(0)_+\|_{\beta}} \quad (\text{By Lemma 3.2})$$

$$\leq \mu_{\beta\beta} \|\nabla g(0)\|_{\beta} \frac{\|g(x)_+\|_{\beta}}{\|g(0)_+\|_{\beta}} \quad \square$$

For Example 2.3, it is easy to verify that $\|\nabla g(0)\|_{\infty} = 2$ and hence the condition number for the example, using the ∞ -norm, is: $\mu_{\infty\infty} \|\nabla g(0)\|_{\infty} = 4$.

Acknowledgements

I am indebted to Stephen M. Robinson for referring me to his paper [8] and to my student T.-H. Shiau for reading the paper and suggesting improvements. I am also grateful to the referees for a correction in the finiteness argument for $\mu_{\beta\gamma}$.

REFERENCES

1. S. D. Conte and C. de Boor: "Elementary numerical analysis", McGraw-Hill, New York, 1980.
2. D. Gale: "The theory of linear economic models", McGraw-Hill, New York 1960.
3. A. J. Hoffman: "On approximate solutions of systems of linear inequalities", Journal of Research of the National Bureau of Standards 49, 1952, 263-265.
4. A. S. Householder: "The theory of matrices in numerical analysis", Blaisdell Publishing, New York, 1964.
5. O. L. Mangasarian: "Nonlinear programming", McGraw-Hill, New York, 1969.
6. O. L. Mangasarian: "A condition number for linear inequalities and linear programs", in "Methods of Operations Research", Proceedings of 6. Symposium uber Operations Research, Augsburg 7-9 September 1981, G. Bamberg & O. Opitz (eds.), Verlagsgruppe Athenaum/Hain/ Scriptor/Hanstein, Konigstein 1981, 3-15.
7. J. S. Ortega: "Numerical analysis", Academic Press, New York, 1972.
8. S. M. Robinson: "An application of error bounds for convex programming in a linear space", SIAM Journal on Control and Optimization 13, 1975, 271-273.