On Approximate Solutions of Infinite Systems of Linear Inequalities

Hui Hu
Department of Mathematical Sciences
Northern Illinois University
DeKalb, Illinois 60115-2888

and

Qing Wang
Department of Mathematics
Yale University
Box 2155, Yale Station
New Haven, Connecticut 06520

Dedicated to Alan J. Hoffman on the occasion of his 65th birthday.

Submitted by Robert E. Bixby

ABSTRACT

This paper extends A. J. Hoffman's results on approximate solutions of finite systems of linear inequalities to infinite systems of linear inequalities. It is shown that for a given infinite system of linear inequalities (satisfying certain conditions), the Euclidean distance from a vector x to the solution set of the system is equivalent to the "biggest violation" by x of the system. Thus, if a vector x "almost" satisfies the system, then x is "close" to a solution of the system.

1. INTRODUCTION

For any real number a, we define

$$a^+ = \begin{cases} a & \text{if} \quad a \geqslant 0, \\ 0 & \text{if} \quad a < 0. \end{cases}$$

For any vector $x = (x_1, \dots, x_n)^T \in R^n$, we define $x^+ = (x_1^+, \dots, x_n^+)^T$, $\|x\| = (x_1^2 + \cdots + x_n^2)^{1/2}$, and $\|x\|_\infty = \max\{|x_i| : i = 1, \dots, n\}$.

655 Avenue of the Americas, New York, NY 10010 0024-3795/89/$3.50

Given a finite consistent system of linear inequalities

$$Ax \leqslant b, \tag{I}$$

where A is an $m \times n$ real matrix, x in R^n, and b in R^m, Hoffman (1952) proved:

THEOREM (Hoffman). *There exists a constant* $\tau > 0$ *such that for any* $x \in R^n$ *there exists a solution* x^* *of* (I) *satisfying*

$$\|x - x^*\| \leqslant \tau \|(Ax - b)^+\|_\infty .$$

This theorem tells us that the distance from x to $\{x: Ax \leqslant b\}$ is dominated by the infinity norm of $(Ax - b)^+$. Hence, if a vector x "almost" satisfies (I), then x is "close" to a solution x^* of (I). In this paper we extend the above theorem to infinite systems of linear inequalities.

2. NOTATION AND PRELIMINARIES

Given a nonempty and bounded set C in R^{n+1}, consider the infinite system of linear inequalities

$$ax \leqslant b \qquad \text{for all} \quad (a, b)^T \in C, \tag{II}$$

where $a^T, x \in R^n$ and $b \in R^1$.

Let $S \equiv \{x \in R^n: ax \leqslant b$ for all $(a, b)^T \in C\}$ be the solution set of (II). We assume that S is not empty.

Let $f: S \to R^n$. A vector $x \in S$ is a stationary point of the pair (S, f) if $x^T f(x) \leqslant y^T f(x)$ for all $y \in S$.

Let $d(x) = \min\{\|x - y\|: y \in S\}$ be the Euclidean distance from x to S.

Let $H(x) = \sup\{(ax - b)^+ : (a, b)^T \in C\}$ be the biggest violation by x of (II). $H(x) < \infty$, since we assume that C is bounded. $H(x) \geqslant 0$, and $H(x) = 0$ if and only if $x \in S$. Moreover, $H(x)$ is a continuous convex function.

We want to show that if the infinite system of linear inequalities (II) satisfies certain conditions, then $H(x)$ is equivalent to $d(x)$, i.e., there exist constants $\tau > 0$ and $\gamma > 0$ such that for all x in R^n we have

$$d(x) \leqslant \tau H(x) \quad \text{and} \quad H(x) \leqslant \gamma d(x).$$

The following lemma will be used in the proofs of the main results. This lemma is a generalization of Farkas's lemma to infinite systems of linear inequalities.

LEMMA. *A given half space $a^*x \leqslant b^*$ (where $a^* \neq 0$) contains S if and only if there exist vectors $(a^{kj}, b^{kj})^T \in C$ and coefficients $\lambda_{kj} \geqslant 0$, $k = 1,2,\ldots,$ $j = 0,\ldots,n$, such that*

$$\lim_{k \to \infty} \left(\sum_{j=1}^{n} \lambda_{kj} a^{kj} \right) = a^*$$

and

$$\lim_{k \to \infty} \left(\lambda_{k0} + \sum_{j=1}^{n} \lambda_{kj} b^{kj} \right) = b^*$$

[*see, e.g., Rockafellar* (1970, *pp.* 159–160)].

3. THE MAIN RESULTS

THEOREM 0. *There exists a constant $\gamma > 0$ such that $H(x) \leqslant \gamma d(x)$ for all $x \in R^n$.*

Proof. Let $x \in R^n$ be any vector exterior to S, and x^* be the vector in S nearest to x. By the definition of $H(x)$, there exists a sequence $(a^i, b^i)^T \in C$, $i = 1,2,\ldots,$ such that $H(x) = \lim_{i \to \infty}(a^i x - b^i)^+$. Then we have

$$\begin{aligned} H(x) &= \lim_{i \to \infty} \left(a^i(x - x^*) + a^i x^* - b^i\right)^+ \\ &\leqslant \limsup_{i \to \infty} \left(a^i(x - x^*)\right)^+ \\ &\leqslant \limsup_{i \to \infty} \|a^i\| \cdot \|x - x^*\|. \end{aligned}$$

It follows that $H(x) \leqslant \gamma d(x)$, where $\gamma = \sup\{\|a\| : (a, b)^T \in C\}$. ■

THEOREM 1. *Suppose that S is bounded by $K > 0$, i.e., $\|x\| \leqslant K$ for all $x \in S$, and that $b \geqslant \delta > 0$ for all $(a,b)^T \in C$. Then there exists a constant $\tau > 0$ such that $d(x) \leqslant \tau H(x)$ for all $x \in R^n$.*

Proof. Let $x \in R^n$ be any vector exterior to S, and x^* be the vector in S nearest to x. Then x^* is the optimal solution of the program

(P) minimize $\|y - x\|^2$
subject to

$$ay \leqslant b \qquad \text{for all} \quad (a,b)^T \in C.$$

It is easy to see that x^* is a stationary point of $(S, \nabla \|y - x\|^2)$. Therefore, the half space

$$(x - x^*)^T y \leqslant (x - x^*)^T x^*$$

contains S. It follows from the lemma that there exist vectors $(a^{kj}, b^{kj})^T \in C$ and coefficients $\lambda_{kj} \geqslant 0$, $k = 1,2,\dots,$ $j = 0,\dots,n$, such that

$$\lim_{k \to \infty} \sum_{j=1}^{n} \lambda_{kj} a^{kj} = x - x^* \tag{1}$$

and

$$\lim_{k \to \infty} \left(\lambda_{k0} + \sum_{j=1}^{n} \lambda_{kj} b^{kj} \right) = (x - x^*)^T x^*. \tag{2}$$

(1) and (2) imply

$$\lim_{k \to \infty} \left(\lambda_{k0} + \sum_{j=1}^{n} \lambda_{kj} (b^{kj} - a^{kj} x^*) \right) = 0. \tag{3}$$

Since $\lambda_{kj} \geqslant 0$ and $b^{kj} - a^{kj}x^* \geqslant 0$ for all $k = 1,2,\dots,$ $j = 0,\dots,n$, we have

$$\lim_{k \to \infty} \lambda_{k0} = 0 \tag{4}$$

and

$$\lim_{k\to\infty} \sum_{j=1}^{n} \lambda_{kj}(b^{kj} - a^{kj}x^*) = 0. \tag{5}$$

From (2), (4), and the assumptions of Theorem 1, we have

$$\limsup_{k\to\infty} \sum_{j=1}^{n} \lambda_{kj} \leqslant \delta^{-1}K\|x - x^*\|. \tag{6}$$

From (1), (5), and (6), we have

$$\begin{aligned}\|x - x^*\|^2 &= \lim_{k\to\infty} \sum_{j=1}^{n} \lambda_{kj} a^{kj}(x - x^*)\\ &= \lim_{k\to\infty}\left(\sum_{j=1}^{n} \lambda_{kj}(a^{kj}x - b^{kj}) + \sum_{j=1}^{n} \lambda_{kj}(b^{kj} - a^{kj}x^*)\right)\\ &\leqslant \left(\limsup_{k\to\infty} \sum_{j=1}^{n} \lambda_{kj}\right) H(x)\\ &\leqslant \delta^{-1}K\|x - x^*\|H(x).\end{aligned}$$

It follows that $d(x) \leqslant \tau H(x)$ for all $x \in R^n$, where $\tau = \delta^{-1}K$. ∎

COROLLARY. *If C is compact, $0 \notin C$, and S is bounded and has an interior point, then there exists a constant $\tau > 0$ such that $d(x) \leqslant \tau H(x)$ for all $x \in R^n$.*

Proof. Let x^0 be an interior point of S. Since $0 \notin C$, we know that $ax^0 < b$ for all $(a, b)^T \in C$. It follows from the compactness of C that $b - ax^0 \geqslant \delta > 0$ for all $(a, b)^T \in C$. Define

$$\hat{S} = \left\{ y \in R^n : ay \leqslant b - ax^0 \text{ for all } (a, b)^T \in C \right\}.$$

Then $\hat{S}$ is bounded, and $x \in S$ if and only if $y = x - x^0 \in \hat{S}$. Define $\hat{d}(y) = \min\{\|y - z\| : z \in \hat{S}\}$ and $\hat{H}(y) = \sup\{(ay - b + ax^0)^+ : (a, b)^T \in C\}$. By Theorem 1, there exists a constant $\tau > 0$ for $\hat{S}$ such that $\hat{d}(y) \leqslant \tau\hat{H}(y)$ for all

$y \in R^n$. The same τ works for S because $\hat{d}(y) = d(x)$ and $\hat{H}(y) = H(x)$, where $y = x - x^0$. ■

THEOREM 2. *Suppose that S is unbounded, and that there is a unit vector $v \in R^n$ such that $av \geqslant \epsilon > 0$ for all $(a, b)^T \in C$. Then there exists a constant $\tau > 0$ such that $d(x) \leqslant \tau H(x)$ for all $x \in R^n$.*

Proof. Let $x \in R^n$ be any vector exterior to S, and x^* and (P) be defined as in the proof of Theorem 1. It is clear that Equations (1) through (5) still hold. From (1) and the assumptions of Theorem 2, we have

$$\limsup_{k \to \infty} \sum_{j=1}^{n} \lambda_{jk} \leqslant \epsilon^{-1} \|x - x^*\|. \tag{7}$$

From (1), (5), and (7) we have

$$\begin{aligned}
\|x - x^*\|^2 &= \lim_{k \to \infty} \sum_{j=1}^{n} \lambda_{kj} a^{kj}(x - x^*) \\
&= \lim_{k \to \infty} \left(\sum_{j=1}^{n} \lambda_{kj}(a^{kj}x - b^{kj}) + \sum_{j=1}^{n} \lambda_{kj}(b^{kj} - a^{kj}x^*) \right) \\
&\leqslant \left(\limsup_{k \to \infty} \sum_{j=1}^{n} \lambda_{kj} \right) H(x) \\
&\leqslant \epsilon^{-1} \|x - x^*\| H(x).
\end{aligned}$$

It follows that $d(x) \leqslant \tau H(x)$ for all $x \in R^n$, where $\tau = \epsilon^{-1}$. ■

REMARKS. S is unbounded if and only if there exists a unit vector v such that $av \geqslant 0$ for all $(a, b)^T \in C$. Indeed, if S is unbounded, then it contains an unbounded ray, i.e., there exist $w \neq 0$ and $\bar{x} \in S$ such that $\bar{x} + tw \in S$ for all $t \geqslant 0$ [see, e.g., Rockafellar (1970, Chapter 8)]. Therefore, $aw \leqslant 0$ holds for all $(a, b)^T \in C$, since $aw \leqslant t^{-1}(b - a\bar{x})$ for all $t > 0$. Conversely, if there exists a unit vector v such that $av \geqslant 0$ for all $(a, b)^T \in C$, then $\bar{x} + t(-v) \in S$ for all $t \geqslant 0$, where $\bar{x} \in S$.

In the rest of this section, we provide two examples. These examples show that the conclusions of Theorems 1 and 2 need not hold when their hypotheses are not satisfied. They are inspired by Duffin and Karlovitz (1965).

EXAMPLE 1. Consider the infinite system of linear inequalities

$$\begin{aligned}
&x_1+0x_2+0x_3\leqslant 1,\\
&0x_1+x_2+0x_3\leqslant 1,\\
&0x_1-x_2+0x_3\leqslant 1,\\
&0x_1+0x_2+x_3\leqslant 0,\\
&0x_1+0x_2-x_3\leqslant 0,\\
&-\frac{x_1}{n}-\frac{x_2}{n^2}-x_3\leqslant 0 \qquad \text{for all} \quad n=1,2,\ldots.
\end{aligned} \tag{8}$$

In this case,

$$C=\left\{\left(\frac{-1}{n},\frac{-1}{n^2},-1,0\right)^T:n=1,2,\ldots\right\}$$
$$\cup\left\{(1,0,0,1)^T,(0,\pm 1,0,1)^T,(0,0,\pm 1,0)^T\right\}$$

is compact and bounded away from the origin, and the condition $b\geqslant\delta>0$ given in Theorem 1 does not hold. It is easy to see that the solution set S of (8) is bounded and

$$\left(0,\tfrac{1}{4},0\right)^T\in S\subset\left\{(x_1,x_2,0)^T\in R^3:0\leqslant x_1\leqslant 1,\ -1\leqslant x_2\leqslant 1\right\}.$$

Let $x(t)=(-t,\frac{1}{4},0)^T$ for all $t>0$; we have

$$d(x(t))=t \quad\text{and}\quad H(x(t))=\sup\left\{\left(\frac{t}{n}-\frac{1}{4n^2}\right)^+:n=1,2,\ldots\right\}\leqslant t^2.$$

Consequently, it is impossible to find a constant $\tau>0$ such that $d(x(t))\leqslant \tau H(x(t))$ for all $t>0$ sufficiently small.

EXAMPLE 2. Consider the infinite system of linear inequalities:

$$
\begin{aligned}
&0x_1+0x_2+x_3\leqslant 0,\\
&0x_1+0x_2-x_3\leqslant 0,\\
&-\frac{x_1}{n}-\frac{x_2}{n^2}-x_3\leqslant 0 \qquad \text{for all} \quad n=1,2,\dots .
\end{aligned}
\tag{9}
$$

In this case, $C=\{(-1/n,\ -1/n^2,\ -1,0)^T: n=1,2,\dots\}\cup\{(0,0,\pm 1,0)^T\}$ is also compact and bounded away from the origin. Note that the condition $av\geqslant \epsilon>0$ given by Theorem 2 does not hold, and

$$\left\{(0,x_2,0)^T\in R^3: x_2\geqslant 0\right\}\subset S\subset\left\{(x_1,x_2,0)^T\in R^3: x_1\geqslant 0\right\}.$$

Now let $x(t)=(1/\sqrt{t})(-t,t^2/4,0)^T$ for all $t>0$; we have

$$d(x(t))=\sqrt{t}\to+\infty \qquad \text{as} \quad t\to+\infty,$$

while

$$
\begin{aligned}
H(x(t)) &= \sup\left\{\frac{1}{\sqrt{t}}\left(\frac{t}{n}-\frac{t^2}{4n^2}\right)^+ : n=1,2,\dots\right\}\\
&\leqslant \frac{1}{\sqrt{t}}\to 0 \qquad \text{as} \quad t\to+\infty .
\end{aligned}
$$

Therefore, the constant $\tau>0$ satisfying $d(x)\leqslant \tau H(x)$ for all $x\in R^3$ does not exist.

For an unbounded S, we see (in Example 2) that a vector may be far away from the solution set even if it almost satisfies the system, while for a bounded S, we can show that $H(x^n)\to 0$ implies $d(x^n)\to 0$. Indeed, fix any $\epsilon>0$ and consider the compact set $D(\epsilon)=\{x\in R^n: d(x)=\epsilon\}$. Since $H(x)$ is continuous and never vanishes on $D(\epsilon)$,

$$\sup\left\{\frac{\epsilon}{H(x)}: x\in D(\epsilon)\right\}\equiv\tau(\epsilon)<+\infty$$

and thus $d(x)\leqslant\tau(\epsilon)H(x)$ for all $x\in D(\epsilon)$. Now pick any x such that $d(x) > \epsilon$, and let x^* be the nearest point from x to S and $y = \lambda x + (1-\lambda)x^* \in D(\epsilon)$, where $0 < \lambda < 1$. Then x^* is also the nearest point from y to S and $d(y) = \lambda d(x)$. From the convexity of $H(x)$, we have $H(y) \leqslant \lambda H(x) + (1-\lambda)H(x^*) = \lambda H(x)$. Hence,

$$d(x) = \lambda^{-1} d(y) \leqslant \lambda^{-1} \tau(\epsilon) H(y) \leqslant \tau(\epsilon) H(x). \tag{10}$$

Let $\{x^n\}$ be a sequence such that $H(x^n) \to 0$ but $d(x^n) \nrightarrow 0$. Then $d(x^{n_j}) > \epsilon_0 > 0$ for some subsequence $\{x^{n_j}\}$. From (10),

$$0 < \epsilon_0 < d(x^{n_j}) \leqslant \tau(\epsilon_0) H(x^{n_j}),$$

which contradicts $H(x^{n_j}) \to 0$.

4. APPLICATIONS TO SEMIINFINITE LINEAR PROGRAMMING

The primal problem of semiinfinite linear programming is defined as

$$\text{(SIL)} \quad \begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & ax - b \leqslant 0 & \text{for all} \quad (a,b)^T \in C, \end{array}$$

where $x, c^T, a^T \in R^n$, $b \in R^1$, and $C \subseteq R^{n+1}$ is a nonempty and bounded set containing infinitely many vectors. There are many practical as well as theoretical problems that can be formulated as semiinfinite linear programs [see, e.g., Gustafson and Kortanek (1973)]. In general, it is very difficult to find an optimal solution of (SIL) in a finite number of steps. The algorithms for solving (SIL) can only provide an approximate solution at each step. One class of algorithms generates an approximate solution x^k at each iteration k with the properties:

(a) $ax^k - b \leqslant \delta_k$ for all $(a,b)^T \in C$, where $\delta_k > 0$, and

(b) $c^T x^k \leqslant v(\text{SIL})$, where $v(\text{SIL})$ is the optimal value of the objective function $c^T x$ of (SIL) [see, e.g., Hettich (1979) and Hu (1988)].

In order to stop the algorithm as soon as a satisfactory approximate solution is found, one needs to estimate the distance from x^k to the feasible set and the difference between $v(\text{SIL})$ and $c^T x^k$. Suppose, for instance, that the condition of Theorem 2 is satisfied, namely, $av \geqslant \epsilon > 0$ for all $(a,b)^T \in C$. Then by

Theorem 2,

$$d(x^k) \leqslant \epsilon^{-1}H(x^k) \leqslant \epsilon^{-1}\delta_k.$$

Let y^k be the feasible solution such that $d(x^k) = \|y^k - x^k\|$. By the feasibility of y^k and the properties of x^k, we have

$$0 < v(\text{SIL}) - c^T x^k \leqslant c^T y^k - c^T x^k \leqslant \|c\| \cdot \|y^k - x^k\| \leqslant \|c\|\epsilon^{-1}\delta_k.$$

Hence, the distance from x^k to the feasible set is dominated by $\epsilon^{-1}\delta_k$, and the difference between $v(\text{SIL})$ and $c^T x^k$ by $\|c\|\epsilon^{-1}\delta_k$.

We wish to thank Dr. A. J. Hoffman for his encouragement and suggestions.

REFERENCES

Agmon, S. 1954. The relaxation method for linear inequalities, *Canad. J. Math.* 6:382–392.

Avriel, M. 1976. *Nonlinear Programming: Analysis and Methods*, Prentice-Hall, Englewood Cliffs, N.J.

Duffin, R. J. and Karlovitz, L. A. 1965. An infinite linear program with a duality gap, *Management Sci.* 12(1):122–134.

Gustafson, S.-A. and Kortanek, K. O. 1973. Numerical treatment of a class of semi-infinite programming problems, *Naval Res. Logist. Quart.* 20:477–504.

Hettich, R. 1979. A comparison of some numerical methods for semi-infinite programming, in *Semi-infinite Programming* (R. Hettich, Ed.), Springer-Verlag, New York, pp. 112–125.

Hoffman, A. J. 1952. On approximate solutions of systems of linear inequalities, *J. Res. Nat. Bur. Standards* 49:263–265.

Hu, H. 1988. A one phase algorithm for semi-infinite linear programming, *Math. Programming*, to appear.

Rockafellar, R. T. 1970. *Convex Analysis*, Princeton U.P., Princeton, N.J.

Received 25 January 1988; final manuscript accepted 30 August 1988