Journal of Global Optimization **30:** 391–403, 2004.

Hoffman's Least Error Bounds for Systems of Linear Inequalities*

XI YIN ZHENG[1] and KUNG FU NG[2]
[1]*Department of Mathematics, Yunnan University, Kunming 650091, P.R. China (e-mail: xyzheng@ynu.edu.cn)*
[2]*Department of Mathematics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong (e-mail: kfng@math.cuhk.edu.hk)*

(Received: 23 April 2003; accepted: 26 November 2003)

Abstract. Let E be a normed space, $a_1^*,\ldots,a_m^*\in E^*$, $c_1,\ldots,c_m\in R$ and $S=\{x\in E|\langle a_i^*,x\rangle-c_i\leqslant 0, 1\leqslant i\leqslant m\}\neq\emptyset$. Let $\tau_*=\inf\{\tau\geqslant 0:\mathrm{dist}(x,S)\leqslant\tau\max\{[\langle a_i^*,x\rangle-c_i]_+: i=1,\ldots,m\}\ \forall x\in E\}$. We give some exact formulas for τ_*.

Mathematics Subject Classifications (2000). 90C05, 90C48.

Key words. normed space, system of linear inequalities, the least error bound.

1. Introduction

Let E be a normed space and E^* the dual space of E. Let $a_1^*,\ldots,a_m^*\in E^*$, $c_1,\ldots,c_m\in R$ and S denote the solution set of the following inequality system

$$\langle a_i^*,x\rangle\leqslant c_i,\quad i=1,\ldots,m. \tag{1}$$

Assuming that S is nonempty, the fundamental result of Hoffman [5] asserts that if $E=R^n$ then there exists a constant $\tau>0$ such that

$$\mathrm{dist}(x,S)\leqslant\tau[\phi(x)]_+\quad\text{for all } x\in E, \tag{2}$$

where $\phi(x):=\max\{\langle a_i^*,x\rangle-c_i: i=1,\ldots,m\}$. A coefficient τ satisfying (2) is called a Lipschitz error bound of the system (1). Let τ_* denote the infimum of all Lipschitz error bounds; namely

$$\tau_*=\sup\left\{\frac{\mathrm{dist}(x,S)}{[\phi(x)]_+}: x\in E\backslash S\right\}.$$

*This research was supported by an Earmarked grant from the Research Grant Council of Hong Kong. Research of the first author was also supported by the National Natural Science Foundation of P.R. China (Grant No. 10361008) and Natural Science Foundation of Yunnan Province, P.R. China (Grant No. 2003A0002M).

Since (1) has a Lipschitz error bound, τ_* does exist of course (in view of Theorem 2.2 in Burke and Ferris [2]), but the interest remains in expressing τ_* in a simple way. We show in Theorem 2.3 that

$$\tau_* = \frac{1}{\min\{\gamma_D\colon D \in \mathcal{W}(I)\}}$$

where $I = \{1, \ldots, m\}, \mathcal{W}(I)$ is a certain (finite) family of subsets of I and γ_D is defined by

$$\gamma_D = \min\left\{\left\|\sum_{i\in D}\alpha_i a_i^*\right\| : \alpha_i \geqslant 0, \sum_{i\in D}\alpha_i = 1\right\}.$$

Note that the subprogramme of computing γ_D is a convex minimization problem over a simplex. In particular if E is finite dimensional with the Euclidean norm then that is a quadratic minimization problem while if E is finite dimensional with the l_1- or the l_∞-norm then that is simply a linear programming problem. Consequently τ_* is given as a finitely computable object.

Lipschitz error bounds are related to the convergence rate of algorithms appearing in many applications. Several authors considered the bounds, seeing Mangasarian and Shiau [9], Bergthaller and Singer [1], Li [8], Guler, Hoffman and Rothblum [4], Burke and Tseng [3] and references therein. In particular, Guler, Hoffman and Rothblum [4] (Theorem 3.2 with the maximum norm in R^m) proved that

$$\tau_1 := \sup\left\{\sum_{i=1}^m \lambda_i : (\lambda_1, \ldots, \lambda_m) \text{ is an extreme point of } \sigma(a_1^*, \ldots, a_m^*)\right\} \tag{3}$$

is a Lipschitz error bound, where

$$\sigma(a_1^*, \ldots, a_m^*) := \left\{(\lambda_1, \ldots, \lambda_m) \in R_+^m : \left\|\sum_{i=1}^m \lambda_i a_i^*\right\| \leqslant 1\right\}.$$

Let

$$\mathcal{J} := \{D \subset I : \{a_i^* : i \in D\} \text{ is linearly independent}\}.$$

Burke and Tseng [3] (Theorem 8), in the case when their cone $K = R_-^m$ and $X = R^n$, proved that

$$\tau_2 := \max\left\{\sum_{i\in D}\lambda_i : \lambda_i \geqslant 0,\ \left\|\sum_{i\in D}\lambda_i a_i^*\right\| = 1,\ D \in \mathcal{J}\right\}. \tag{4}$$

is a Lipschitz error bound. As noted in Burke and Tseng [3], $\tau_1 = \tau_2$. Let

$$\widehat{\mathcal{J}} = \{D \in \mathcal{J}\colon \text{ there exists } x \in S \text{ such that } \langle a_i^*, x\rangle = c_i\ \forall i \in D\}.$$

Bergthaller and Singer [1] (Theorem 1.3) proved that

$$\tau_3 := \max\left\{\sum_{i\in D}\lambda_i\colon \lambda_i \geqslant 0,\ \left\|\sum_{i\in D}\lambda_i a_i^*\right\| = 1,\ D \in \widehat{\mathcal{J}}\right\} \tag{5}$$

is also a Lipschitz error bound. Since $\widehat{\mathcal{J}} \subset \mathcal{J}$, $\tau_2 \geqslant \tau_3$.

Take $E = R$, $a_1^* = 1$, $a_2^* = \frac{1}{2}$, $a_3^* = \frac{1}{3}$, $a_4^* = \frac{1}{4}$, $c_1 = 3$, $c_2 = 1$, $c_3 = \frac{2}{3}$ and $c_4 = 1$. Then

$$\phi(x) = \max\{a_i^* x - c_i : 1 \leqslant i \leqslant 4\} = \begin{cases} \frac{1}{4}x - 1, & x \leqslant -4, \\ \frac{1}{3}x - \frac{2}{3}, & -4 \leqslant x \leqslant 2, \\ \frac{1}{2}x - 1, & 2 \leqslant x \leqslant 4, \\ x - 3, & 4 \leqslant x. \end{cases}$$

Then $S = \{x \in R : \phi(x) \leqslant 0\} = (-\infty, 2]$. In this case, $\mathcal{J} = \{\{i\} : 1 \leqslant i \leqslant 4\}$, $\widehat{\mathcal{J}} = \{\{2\}, \{3\}\}$. It is easy to verify that $\tau_1 = \tau_2 = 4 > \tau_3 = 3$ while 2 obtained by Theorems 2.2 and 2.3 is the least Lipschitz error bound. This shows that in general τ_3 (a fortiori, τ_1 and τ_2) is not the least Lipschitz error bound of the inequality system (1).

2. Main Results

Throughout this paper, let E denote a normed space, $a_1^*, \ldots, a_m^* \in E^*$ and $c_1, \ldots, c_m \in R$; let

$$S := \{x \in E : \langle a_i^*, x \rangle \leqslant c_i,\ i = 1, \ldots, m\}.$$

We always assume that $\emptyset \neq S \neq E$. Let $I = \{1, \ldots, m\}$ and

$$\phi(x) = \max\{\langle a_i^*, x \rangle - c_i : i \in I\} \quad \forall x \in E.$$

Then $S = \{x \in E : \phi(x) \leqslant 0\}$. We use $\mathcal{M}(I)$ to denote the family of all subsets D of I such that $\{a_i^* : i \in D\}$ is a maximal linearly independent subset of $\{a_i^* : i \in I\}$.

DEFINITION 2.1. We say that a subset D of I has property (W) if the following conditions hold.

(i) $D \in \mathcal{M}(I)$.

(ii) Given a solution s_D of the linear equation system

$$\langle a_i^*, x \rangle - c_i = 0 \quad \forall i \in D \tag{6}$$

and a solution e_D of the linear equation system

$$\langle a_i^*, x \rangle = 1 \quad \forall i \in D, \tag{7}$$

one has $\langle a_j^*, s_D \rangle - c_j \leqslant 0$ for each $j \in I \backslash D$ and the strict inequality $\langle a_j^*, s_D \rangle - c_j < 0$ holds whenever $\langle a_j^*, e_D \rangle > 1$ and $j \in I \backslash D$.

Let $\mathcal{W}(I)$ denote the family of all subsets of I with property (W). For each $D \in \mathcal{W}(I)$, let

$$\tau_D := \mathrm{dist}(e_D, C_D) \tag{8}$$

where $C_D := \{x \in E : \langle a_i^*, x\rangle \leqslant 0, i \in D\}$. Note from (i) of Definition 2.1 that each of (6) and (7) has solutions and that $\bigcap_{i\in D}\ker(a_i^*) = \bigcap_{i\in I}\ker(a_i^*)$. Therefore, for another solution s'_D of the linear equation system (6), $\langle a_i^*, s'_D\rangle = \langle a_i^*, s_D\rangle$ for each $i \in I$. In fact, the solution set of the linear equation system (6) is a linear variety paralleled to $\bigcap_{i\in I}\ker(a_i^*)$. Similar remarks hold for e_D and (7). Hence, to verify (ii) one needs only to do so for any one particular pair of solutions s_D, e_D of (6), (7) and hence $\mathcal{W}(I)$ is well defined; moreover τ_D does not depend on the particular choice of e_D.

Throughout, let τ_* denote the least Lipschitz error bound for the inequality system (1), that is,

$$\tau_* = \inf\{\tau > 0\colon \operatorname{dist}(x, S) \leqslant \tau[\phi(x)]_+,\ \forall x \in E\}.$$

THEOREM 2.1. $\tau_* = \max\{\tau_D : D \in \mathcal{W}(I)\}$.

The proof of this result is based on geometrically intuitive arguments but is rather intricate to verify in details. We postpone the proof to the next section.

THEOREM 2.2. *For each* $D \in \mathcal{W}(I)$, *let*

$$\bar{\tau}_D = \max\left\{\sum_{i\in D}\lambda_i : \lambda_i \geqslant 0,\ \left\|\sum_{i\in D}\lambda_i a_i^*\right\| = 1\right\}.$$

Then $\bar{\tau}_D = \tau_D$ *for each* $D \in \mathcal{W}(I)$ *and hence*

$$\tau_* = \max\{\bar{\tau}_D : D \in \mathcal{W}(I)\}.$$

Proof. By [1, Theorem 1.1], for each $D \in \mathcal{W}(I)$

$$\tau_D = \operatorname{dist}(e_D, C_D) = \max\left\{\sum_{i\in D}\lambda_i\langle a_i^*, e_D\rangle : \lambda_i \geqslant 0,\ \left\|\sum_{i\in D}\lambda_i a_i^*\right\| = 1\right\}.$$

Since $\langle a_i^*, e_D\rangle = 1$ for each $i \in D$, it follows that

$$\tau_D = \max\left\{\sum_{i\in D}\lambda_i : \lambda_i \geqslant 0,\ \left\|\sum_{i\in D}\lambda_i a_i^*\right\| = 1\right\} = \bar{\tau}_D.$$

By Theorem 2.1, one has that $\tau_* = \max\{\bar{\tau}_D : D \in \mathcal{W}(I)\}$. □

THEOREM 2.3. *For each* $D \in \mathcal{W}(I)$, *let*

$$\gamma_D := \min\left\{\left\|\sum_{i\in D}\alpha_i a_i^*\right\| : \alpha_i \geqslant 0,\ \sum_{i\in D}\alpha_i = 1\right\}.$$

Then $\bar{\tau}_D = \frac{1}{\gamma_D}$ *for each* $D \in \mathcal{W}(I)$ *and hence*

$$\tau_* = 1/\min\{\gamma_D : D \in \mathcal{W}(I)\}.$$

Proof. Pick $(\alpha_i')_{i\in D}\in R_+^{|D|}$ with $\sum_{i\in D}\alpha_i'=1$ such that $\gamma_D=\|\sum_{i\in D}\alpha_i' a_i^*\|$, where $|D|$ denotes the number of elements in D. Then $\|\sum_{i\in D}\frac{\alpha_i'}{\gamma_D}a_i^*\|=1$ and hence

$$\bar{\tau}_D\geqslant\sum_{i\in D}\frac{\alpha_i'}{\gamma_D}=\frac{1}{\gamma_D}. \tag{9}$$

For each $(\lambda_i)_{i\in D}\in R_+^{|D|}$ with $\|\sum_{i\in D}\lambda_i a_i^*\|=1$, let $\eta_k=\frac{\lambda_k}{\sum_{i\in D}\lambda_i}$ for each $k\in D$. Then $\sum_{k\in d}\eta_k=1$. Therefore,

$$\gamma_D\leqslant\left\|\sum_{k\in D}\eta_k a_k^*\right\|=\frac{1}{\sum_{i\in D}\lambda_i}\left\|\sum_{k\in D}\lambda_k a_k^*\right\|=\frac{1}{\sum_{i\in D}\lambda_i},$$

that is, $\frac{1}{\gamma_D}\geqslant\sum_{i\in D}\lambda_i$. It follows from the definition of $\bar{\tau}_D$ that $\frac{1}{\gamma_D}\geqslant\bar{\tau}_D$. This and (9) imply that $\bar{\tau}_D=\frac{1}{\gamma_D}$. The proof is completed.

Let

$$\tau_4:=\max\left\{\operatorname{dist}(e_D,C_D):D\in\mathcal{M}(I)\right\},$$

$$\tau_5:=\max\left\{\sum_{i\in D}\lambda_i:\lambda_i\geqslant 0,\ \left\|\sum_{i\in D}\lambda_i a_i^*\right\|=1,\ D\in\mathcal{M}(I)\right\}$$

and

$$\tau_6:=1/\min\left\{\left\|\sum_{i\in D}\alpha_i a_i^*\right\|:\alpha_i\geqslant 0,\ \sum_{i\in D}\alpha_i=1,\ D\in\mathcal{M}(I)\right\}.$$

Since $\mathcal{W}(I)\subset\mathcal{M}(J)$, Theorems 2.1, 2.2 and 2.3 imply the following result.

COROLLARY 2.1. *Each of τ_4, τ_5 and τ_6 is a Lipschitz error bound of* (1).

In general the constants in corollary 2.1 are not necessarily the sharpest but, on the other hand, they do have an advantage that they only depend on a_i^*'s (not on c_i's).

The following theorem is taken from Theorem 10 of [10].

THEOREM B. *Let $g_i:R^n\to R$ be a continuous convex function for $i=1,\dots,m$ and $S=\{x\in R^n:g_i(x)\leqslant 0,i=1,\dots,m\}$. Let*

$$\tau_*:=\inf\{\tau\geqslant 0:\operatorname{dist}(x,S)\leqslant\tau\max\{[g_i(x)]_+:i=1,\dots,m\}\ \forall x\in R^n\}.$$

Suppose that Abadie's CQ holds for S at each feasible point x (i.e., $N(x,S)=\{\sum_{i\in I(x)}\lambda_i v^i : \lambda_i \geqslant 0, v^i\in\partial g_i(x), i\in I(x)\}$, *where* $I(x)=\{1\leqslant i\leqslant m: g_i(x)\geqslant g_j(x), 1\leqslant j\leqslant m\}$). *Then*

$$\tau_* = 1/\inf\left\{\sup\{v^T\sum_{j\in I(x)}\lambda_j v^j : v\in\partial g_i(x),\ i\in I(x)\}: \lambda_j\geqslant 0,\right.$$
$$\left.\left\|\sum_{j\in I(x)}\lambda_j v^j\right\|=1,\ v^j\in\partial g_j(x),\ j\in I(x),\ x\in\partial S\right\}, \tag{10}$$

where ∂S *denotes the boundary of* S.

In the case when each $g_i(x)=a_i^{*T}x-c_i$ *with* $a_i^*\in R^n$ *and* $c_i\in R$ *for* $i=1,\dots,m$, *where* a_i^{*T} *denotes the transpose of* a_i^*, $\partial g_i(x)=\{a_i^*\}$ *and* (10) *reads*

$$\tau_* = 1/\inf\left\{\max\left\{a_i^{*T}\sum_{j\in I(x)}\lambda_j a_j^* : i\in I(x)\right\}: \lambda_j\geqslant 0,\right.$$
$$\left.\left\|\sum_{j\in I(x)}\lambda_j a_j^*\right\|=1,\ j\in I(x),\ x\in\partial S\right\}, \tag{11}$$

On the other hand, Theorem 2.2 *and Theorem* 2.3 *imply that*

$$\tau_* = \sup\left\{\sum_{i\in D}\lambda_i : \lambda_i\geqslant 0,\ \left\|\sum_{i\in D}\lambda_i a_i^*\right\|=1,\ D\in\mathcal{W}(I)\right\} \tag{12}$$

$$= 1/\min\left\{\left\|\sum_{i\in D}\alpha_i a_i^*\right\| : \alpha_i\geqslant 0,\ \sum_{i\in D}\alpha_i=1,\ D\in\mathcal{W}(I)\right\}. \tag{13}$$

Though (11), (12) *and* (13) *give the same constant* τ_*, (12) *and* (13) *appear to be simpler: first,* $\mathcal{W}(I)$ *is a finite family, secondly, by the definition of* $\mathcal{W}(I)$, *one has that for each* $D\in\mathcal{W}(I)$ *there exists* $s_D\in\partial S$ *such that* $D\subset I(s_D)$. *Moreover, Theorem* 2.2 *and* 2.3 *are valid in the general setting of normed spaces* (*finite or infinite dimensions*) *while the formulation of* (10) *and* (11) *requires the inner product structure. On the other hand, to compute* τ_* *in terms of the formulas* (11), *one needs to solve a minimization problem over an infinite set. Noting the finiteness of* $\{a_i : 1\leqslant i\leqslant m\}$, *it is possible that this problem can be reduced to minimizing the same objective function over a finite set. But this would not be as explicit as* (12) *and* (13) *as it is not easy to provide concrete steps to identify this finite set.*

We conclude this section with a summary giving steps to determine the least error bound τ_*.

Step 1. Determine $\mathcal{M}(I)$ (e.g., one can determine $\mathcal{M}(I)$ by using Gram–Schmidt process when E is of an inner product structure).

Step 2. Find all elements in $\mathcal{M}(I)$ with the property (W). Let $D\in\mathcal{M}(I)$. To determine whether or not $D\in\mathcal{W}(I)$, one only needs to complete the following procedures (1) and (2)).

(1) Solve two linear equation systems:

$$\langle a_i^*, x\rangle - c_i = 0, \quad i\in D \tag{I}$$

$$\langle a_i^*, x\rangle = 1, \quad i\in D \tag{II}$$

(both (I) and (II) can be solved as $\{a_i^* : i\in D\}$ is linearly independent).

(2) Pick an arbitrary pair (s_D, e_D) such that s_D and e_D are solutions of (I) and (II) respectively. Calculate $\alpha_j = \langle a_j^*, s_D\rangle - c_j$ and $\beta_j = \langle a_j^*, e_D\rangle$ for each $j\in I\backslash D$. If $\alpha_j \leqslant 0$ for each $j\in I\backslash D$ and $\alpha_j < 0$ when $j\in I\backslash D$ with $\beta_j > 1$ then $D\in\mathcal{W}(I)$; otherwise $D\notin\mathcal{W}(I)$. [This criterion does not depend on the particular choice of the pair.]

Step 3. Calculate $\bar{\tau}_D$ or γ_D for each $D\in\mathcal{W}(I)$.
Step 4. Calculate τ_* by virtue of Theorems 2.2 or 2.3.

Remark. Computational works involving in Steps 1 and 2 can be very large (cf. [6, p. 207]).

3. Proof of Theorem 2.1

The proof is divided into several steps which will be presented as lemmas. For convenience we first set out notations (that will be used throughout this section). For a subset Y of E, let $\partial(Y)$ and ext(Y) respectively denote the topological boundary and the extreme point set (consisting of all extreme points) of Y. For each $x\in E$, let

$$I(x) := \{i\in I:\ \langle a_i^*, x\rangle - c_i = \phi(x)\},$$

where ϕ and I are as in the beginning of Section 2. For each $t\in R$, let $S_t := \{x\in E : \phi(x)\leqslant t\}$ (and hence $S = S_0$). Let

$$Z_+ := \{(x,\tau)\in E\times R:\ \langle a_i^*, x\rangle - \tau \leqslant c_i,\ i\in I\}$$

and let β be defined by

$$\beta := \inf\{\tau > 0:\ \text{there exists } x\in E\times R \text{ such that } (x,\tau)\in\ \mathrm{ext}(Z_+)\}.$$

Then Z_+ is a polyhedron in $E\times R$ and so the extreme point set $\mathrm{ext}(Z_+)$ is a finite set; consequently $\beta > 0$.

LEMMA 3.1. *Suppose that* $E^* = \mathrm{span}\{a_i^* : i\in I\}$. *Then there exists finitely many subsets* $I_1,\ldots,I_p$ *of* I *and* $\{s_1,\ldots,s_p,e_1,\ldots,e_p\}\subset E$ *such that the following properties are satisfied* (P *will denote the set* $\{1,\ldots,p\}$).

(i) $E^* = \text{span}\{a_i^* : i \in I_j\}$, $\forall j \in P$.
(ii) $\langle a_i^*, s_j \rangle - c_i = 0$ *and* $\langle a_i^*, e_j \rangle = 1$ $\forall i \in I_j$ *and* $\forall j \in P$.
(iii) $I(s_j + te_j) = I_j$ $\forall t \in (0, \beta)$ *and* $\forall j \in P$.
(iv) $\text{ext}(S_t) = \{s_j + te_j : j \in P\}$ $\forall t \in (0, \beta)$.

Proof. Let $t \in (0, \beta)$ and $u \in \text{ext}(S_t)$. Clearly it suffices to show that

(a) $E^* = \text{span}\{a_i^* : i \in I(u)\}$ and that there exist $s_u, e_u \in E$ with the following properties.
(b) $\langle a_i^*, s_u \rangle - c_i = 0$ and $\langle a_i^*, e_u \rangle = 1$ $\forall i \in I(u)$.
(c) $\langle a_i^*, s_u + t' e_u \rangle - c_i < t'$ $\forall t' \in (0, \beta)$ and $\forall i \in I \setminus I(u)$.
(d) $s_u + t' e_u \in \text{ext}(S_{t'})$ $\forall t' \in (0, \beta)$.

To show (a), suppose to the contrary that $E^* \neq \text{span}\{a_i^* : i \in I(u)\}$. Then there exists $x_0 \in E \setminus \{0\}$ such that $\langle a_i^*, x_0 \rangle = 0$ for all $i \in I(u)$. Noting that $\langle a_i^*, u \rangle - c_i < \phi(u) = t$ (as $u \in \text{ext}(S_t) \subset \partial(S_t)$) for each $i \in I \setminus I(u)$, it follows that there exists $\varepsilon > 0$ small enough such that $\langle a_i^*, u \pm \varepsilon x_0 \rangle - c_i \leqslant t$ for all $i \in I$, that is, $u \pm \varepsilon x_0 \in S_t$, contradicting $u \in \text{ext}(S_t)$. Therefore, (a) holds. Thus there exists a subset D of $I(u)$ such that $\{a_i^* : i \in D\}$ is a basis of E^*. This implies that there exists a unique pair $(s_u, e_u) \in E \times E$ such that

$$\langle a_i^*, s_u \rangle - c_i = 0 \quad \text{and} \quad \langle a_i^*, e_u \rangle = 1 \quad \forall i \in D. \tag{14}$$

Hence

$$\langle a_i^*, s_u + te_u \rangle - c_i = t = \phi(u) = \langle a_i^*, u \rangle - c_i \quad \forall i \in D.$$

Since $\{a_i^* : i \in D\}$ is a basis of E^*, it follows that $u = s_u + te_u$. For each $i \in I$, let π_i denote the hyperplane $\{(x, \tau) \in E \times R : \langle a_i^*, x \rangle - \tau = c_i\}$. Then $\cap_{i \in D} \pi_i$ is a line in $E \times R$. It follows from (14) that

$$(s_u, 0) + R(e_u, 1) = \bigcap_{i \in D} \pi_i, \tag{15}$$

where $R(e_u, 1)$ consists of all $t(e_u, 1)$ with $t \in R$. We claim that

$$(s_u, 0) + R(e_u, 1) = \bigcap_{i \in I(u)} \pi_i. \tag{16}$$

Indeed if (16) is not true, then (15) implies that there exists $i' \in I(u) \setminus D$ such that $(s_u, 0) + R(e_u, 1)$ is not a subset of $\pi_{i'}$. It follows from $u = s_u + te_u$ that $((s_u, 0) + R(e_u, 1)) \bigcap \pi_{i'} = \{(u, t)\}$, that is $(\bigcap_{i \in D} \pi_i) \cap \pi_{i'} = \{(u, t)\}$. Noting that $(\bigcap_{i \in D} \pi_i) \cap \pi_{i'} \cap Z_+$ is an extreme subset of Z_+, it follows that $(u, t) \in \text{ext}(Z_+)$, contradicting $t \in (0, \beta)$ and the definition of β. This shows that (16) holds and hence (b) holds. To prove (c), suppose to the contrary that there exists $i_0 \in I \setminus I(u)$ and $t_0 \in (0, \beta)$ such that

$$\langle a_{i_0}^*, s_u + t_0 e_u \rangle - c_{i_0} \geqslant t_0. \tag{17}$$

Note that $\langle a_i^*, s_u + te_u \rangle - c_i < \phi(u) = t\,(i.e., \langle a_i^*, s_u + te_u \rangle - t < c_i)$ for each $i \in I \backslash I(u)$. Consequently, it holds that for each $i \in I \backslash I(u), \tau_i > 0$ where

$$\tau_i := \sup\{\tau > 0 : \langle a_i^*, s_u + [(1-\lambda)t + \lambda t_0]e_u \rangle - (1-\lambda)t - \lambda t_0 < c_i,\ \forall \lambda \in [0, \tau]\}.$$

Note that $\tau_{i_0} \leqslant 1$ thanks to (17). Take $j \in I \backslash I(u)$ such that $\tau_j = \min\{\tau_i : i \in I \backslash I(u)\}$. Then $\tau_j \in (0, 1]$ and

$$\langle a_i^*, s_u + [(1-\tau_j)t + \tau_j t_0]e_u \rangle - (1-\tau_j)t - \tau_j t_0 \leqslant c_i, \quad \forall i \in I \backslash I(u);$$

moreover, the equality holds if $i = j$. This and (b) imply that

$$(s_u + [(1-\tau_j)t + \tau_j t_0]e_u, (1-\tau_j)t + \tau_j t_0) \in \left(\bigcap_{i \in I(u)} \pi_i \right) \cap \pi_j \cap Z_+.$$

Note that

$$(s_u + [(1-\tau_j)t + \tau_j t_0]e_u, (1-\tau_j)t + \tau_j t_0) \notin \text{ext}(Z_+)$$

(by the definition of β and $0 < (1-\tau_j)t + \tau_j t_0 < \beta$) and that $(\bigcap_{i \in I(u)} \pi_i) \cap \pi_j \cap Z_+$ is an extreme subset of Z_+. Therefore $(\bigcap_{i \in I(u)} \pi_i) \cap \pi_j$ must be a line containing the point $(s_u + [(1-\tau_j)t + \tau_j t_0]e_u, (1-\tau_j)t + \tau_j t_0)$. This and (16) imply that

$$(s_u, 0) + R(e_u, 1) = \left(\bigcap_{i \in I(u)} \pi_i \right) \cap \pi_j.$$

Thus $(u, t) = (s_u + te_u, t) \in \pi_j$, that is, $\langle a_j^*, u \rangle - c_j = t = \phi(u)$, contradicting $j \in I \backslash I(u)$. It remains to show (d). By virtue of (a) and (b), one has that

$$\bigcap_{i \in I(u)} \{x \in E : \langle a_i^*, x \rangle - c_i = t'\} = \{s_u + t'e_u\}, \quad \forall t' \in (0, \beta);$$

it follows from (c) that $s_u + t'e_u \in S_{t'}$ and consequently that $s_u + t'e_u \in \text{ext}(S_{t'})$ for each $t' \in (0, \beta)$. This shows that (d) holds. The proof is completed. □

LEMMA 3.2. *Suppose that* $E^* = \text{span}\{a_i^* : i \in I\}$. *Let subsets* $I_1, \ldots, I_p$ *of* I *and* $\{s_1, \ldots, s_p, e_1, \ldots, e_p\} \subset E$ *be such that* (i)–(iv) *of Lemma* 3.1 *hold. Let* $P := \{1, \ldots, p\}$ *and for each* $j \in P$,

$$C_j := \{x \in E \colon \langle a_i^*, x \rangle \leqslant 0,\ \forall i \in I_j\}.$$

Then $\tau_* = \max\{\text{dist}(e_j, C_j) : j \in P\}$.

Proof. For each $t \in R$ and $j \in P$, define $S_t(j)$ by

$$S_t(j) \colon = \{x \in E : \langle a_i^*, x \rangle - c_i \leqslant t,\ \forall i \in I_j\}.$$

By (i) and (ii) of Lemma 3.1 it is easy to verify that $S_t(j)=s_j+te_j+C_j$ and $S\subset S_0(j)=s_j+C_j$ for each $t\in R$ and each $j\in P$. Take $\delta\in(0,\beta)$. Then

$$\mathrm{dist}(s_j+\delta e_j, s_j+C_j)\leqslant \mathrm{dist}(s_j+\delta e_j, S)\leqslant \tau_*\phi(s_j+\delta e_j)=\tau_*\delta,\quad \forall j\in P \tag{18}$$

where the last equality is due to (ii) and (iii) of Lemma 3.1. Making use of the fact that $C_j=\delta C_j$, (18) implies that

$$\delta\,\mathrm{dist}(e_j, C_j)=\mathrm{dist}(s_j+\delta e_j, s_j+C_j)\leqslant\tau_*\delta.$$

Therefore,

$$\tau_*\geqslant\max\{\mathrm{dist}(e_j, C_j): j\in P\}. \tag{19}$$

Let $z\in E$ with $\phi(z)>0$. Take a sequence $\{\varepsilon_k\}$ in $(0, \min\{\phi(z),\beta\})$ convergent to 0. By (iii) of Lemma 3.1 there exists $\tau_k>0$ such that for each $j\in P$, the inequality

$$\langle a_i^*,\cdot\rangle-c_i<\phi(\cdot),\quad\forall i\in I\backslash I_j \tag{20}$$

holds on the ball $B(s_j+\varepsilon_k e_j,\tau_k)$ with center $s_j+\varepsilon_k e_j$ and radius τ_k. Pick a sequence $\{t_k\}$ convergent to 0 with each $t_k\in(\varepsilon_k,\min\{\phi(z),\beta\})$ such that

$$(t_k-\varepsilon_k)\|e_j\|<\frac{\tau_k}{2},\quad\forall j\in P.$$

We claim that

$$\mathrm{dist}(s_j+t_k e_j, S_{\varepsilon_k})=\mathrm{dist}(s_j+t_k e_j, S_{\varepsilon_k}(j)),\quad\forall j\in P. \tag{21}$$

In fact, pick $s_{kj}\in S_{\varepsilon_k}(j)$ such that

$$\mathrm{dist}(s_j+t_k e_j, S_{\varepsilon_k}(j))=\|s_j+t_k e_j-s_{kj}\|.$$

To verify (21), we need show that $s_{kj}\in S_{\varepsilon_k}$ (noting $S_{\varepsilon_k}\subset S_{\varepsilon_k}(j)$), and equivalently that $\langle a_j^*, s_{kj}\rangle-c_i\leqslant\varepsilon_k$ for each $i\in I\backslash I_j$, which clearly follows from (20) provided we can show that s_{kj} belongs to $B(s_j+\varepsilon_k e_j,\tau_k)$. This later condition is indeed satisifed as

$$\|s_j+t_k e_j-s_{kj}\|\leqslant\mathrm{dist}(s_j+t_k e_j, S_{\varepsilon_k})\leqslant\|s_j+t_k e_j-(s_j+\varepsilon_k e_j)\|$$

(the last inequality holds as $s_j+\varepsilon_k e_j\in S_{\varepsilon_k}$), and so

$$\begin{aligned}\|s_j+\varepsilon_k e_j-s_{kj}\| &\leqslant \|s_j+\varepsilon_k e_j-(s_j+t_k e_j)\|+\|s_j+t_k e_j-s_{kj}\|\\ &\leqslant 2\|s_j+t_k e_j-(s_j+\varepsilon_k e_j)\|=2(t_k-\varepsilon_k)\|e_j\|<\tau_k.\end{aligned}$$

Therefore (21) holds. Take $x_k \in \partial(S_{\varepsilon_k})$ such that $\mathrm{dist}(z, S_{\varepsilon_k}) = \|z - x_k\|$; then $\phi(x_k) = \varepsilon_k$. Define z_k by

$$z_k = \frac{t_k - \varepsilon_k}{\phi(z) - \varepsilon_k} z + \left(1 - \frac{t_k - \varepsilon_k}{\phi(z) - \varepsilon_k}\right) x_k.$$

Then

$$\phi(z_k) \leqslant \frac{t_k - \varepsilon_k}{\phi(z) - \varepsilon_k} \phi(z) + \left(1 - \frac{t_k - \varepsilon_k}{\phi(z) - \varepsilon_k}\right) \phi(x_k) = t_k,$$

showing that $z_k \in S_{t_k}$; moreover

$$\mathrm{dist}(z_k, S_{\varepsilon_k}) = \|z_k - x_k\| = \frac{t_k - \varepsilon_k}{\phi(z) - \varepsilon_k} \mathrm{dist}(z, S_{\varepsilon_k}). \tag{22}$$

We claim that

$$\mathrm{dist}(z_k, S_{\varepsilon_k}) \leqslant \max\{\mathrm{dist}(s_j + t_k e_j, S_{\varepsilon_k}) \colon j \in P\}. \tag{23}$$

Note that S_{t_k} is a polyhedron in E containing no lines; hence by [11, Theorem 18.5] one has that $S_{t_k} = \mathrm{co}(\mathrm{ext}(S_{t_k})) + \mathrm{Rec}(S_{t_k})$, where $\mathrm{Rec}(S_{t_k})$ denotes the recession cone of S_{t_k}. It is easy to verify that

$$\mathrm{Rec}(S_{t_k}) \subset \bigcap_{i \in I} \{x \in E : \langle a_i^*, x \rangle \leqslant 0\},$$

and hence that

$$S_{\varepsilon_k} + \mathrm{Rec}(S_{t_k}) \subset S_{\varepsilon_k}. \tag{24}$$

Pick $x' \in \mathrm{co}(\mathrm{ext}(S_{t_k}))$ and $y' \in \mathrm{rec}(S_{t_k})$ such that $z_k = x' + y'$. By the convexity of S_{ε_k},

$$\mathrm{dist}(x', S_{\varepsilon_k}) \leqslant \max\{\mathrm{dist}(e, S_{\varepsilon_k}) \colon e \in \mathrm{ext}(S_{t_k})\}.$$

Noting that $\mathrm{dist}(z_k, S_{\varepsilon_k}) = \mathrm{dist}(x', S_{\varepsilon_k} - y')$, (24) and $y' \in \mathrm{Rec}(S_{t_k})$ imply that $\mathrm{dist}(z_k, S_{\varepsilon_k}) \leqslant \mathrm{dist}(x', S_{\varepsilon_k})$. Therefore,

$$\mathrm{dist}(z_k, S_{\varepsilon_k}) \leqslant \max\{\mathrm{dist}(e, S_{\varepsilon_k}) \colon e \in \mathrm{ext}(S_{t_k})\}.$$

and (23) is seen to hold by virtue of (iv) of Lemma 3.1. Recalling that $S_{\varepsilon_k}(j) = s_j + \varepsilon_k e_j + C_j$ and making use of (21), one has that

$$\begin{aligned}
\mathrm{dist}(s_j + t_k e_j, S_{\varepsilon_k}) &= \mathrm{dist}(s_j + t_k e_j, S_{\varepsilon_k}(j)) \\
&= \mathrm{dist}(s_j + t_k e_j, s_j + \varepsilon_k e_j + C_j) \\
&= \mathrm{dist}((t_k - \varepsilon_k) e_j, C_j) \\
&= (t_k - \varepsilon_k)\ \mathrm{dist}(e_j, C_j)
\end{aligned}$$

(the last equality holds as C_j is a cone and $t_k-\varepsilon_k>0$). It follows from (22) and (23) that

$$\mathrm{dist}(z,S_{\varepsilon_k})\leqslant(\phi(z)-\varepsilon_k)\max\{\mathrm{dist}(e_j,C_j)\colon j\in P\}. \tag{25}$$

By Hoffman's error bound result, there exists a constant $\tau\in(0,+\infty)$ such that

$$\mathrm{dist}(x_k,S)\leqslant\tau\phi(x_k)=\tau\varepsilon_k,\quad \forall k.$$

Noting that

$$\begin{aligned}\mathrm{dist}(z,S) &\leqslant \|z-x_k\|+\mathrm{dist}(x_k,S)\\ &= \mathrm{dist}(z,S_{\varepsilon_k})+\mathrm{dist}(x_k,S),\end{aligned}$$

it follows from $S\subset S_{\varepsilon_k}$ that $\mathrm{dist}(z,S)=\lim_{k\to\infty}\mathrm{dist}(z,S_{\varepsilon_k})$. Thus passing to the limits in (25) gives

$$\mathrm{dist}(z,S)\leqslant\max\{\mathrm{dist}(e_j,C_j)\colon j\in P\}\phi(z).$$

This shows that $\tau_*\leqslant\max\{\mathrm{dist}(e_j,C_j)\colon j\in P\}$. Combining this with (19), the proof is completed. □

LEMMA 3.3. *Suppose that* $E^*=\mathrm{span}\{a_i^*: i\in I\}$. *Then Theorem* 2.1 *holds.*

Proof. Let P,I_j,s_j and $e_j (j\in P)$ be as in Lemma 3.1. For each $j\in P$, by [1, Corollary 1.1] there exists a subset D' of I_j such that $\{a_i^*: i\in D'\}$ is linearly independent and

$$\mathrm{dist}(e_j,C_j)=\mathrm{dist}(e_j,C_{D'}), \tag{26}$$

where $C_{D'}:=\{x\in E:\langle a_i^*,x\rangle\leqslant 0,\forall i\in D'\}$. Pick $D\subset I_j$ such that $D'\subset D$ and $\{a_i^*: i\in D\}$ is a basis of E^*. Since $C_j\subset C_D\subset C_{D'}$,

$$\mathrm{dist}(e_j,C_j)\geqslant\mathrm{dist}(e_j,C_D)\geqslant\mathrm{dist}(e_j,C_{D'}).$$

It follows from (26) that $\mathrm{dist}(e_j,C_j)=\mathrm{dist}(e_j,C_D)$. By (ii) and (iii) of Lemma 3.1, it is easy to verify that $D\in\mathcal{W}(I)$ and $\mathrm{dist}(e_j,C_D)=\mathrm{dist}(e_D,C_D)$ because e_j is clearly a solution of the following linear equation system

$$\langle a_j^*,x\rangle=1,\quad \forall i\in D.$$

Therefore Lemma 3.2 implies that $\tau_*\leqslant\max\{\mathrm{dist}(e_D,C_D): D\in\mathcal{W}(I)\}$. To prove the converse inequality, let $D\in\mathcal{W}(I)$ and $t_D:=\min\{\frac{\langle a_i^*,s_D\rangle-c_i}{1-\langle a_i^*,e_D\rangle}: i\in I_D\}$, where $I_D:=\{i\in I\backslash D:\langle a_i^*,e_D\rangle>1\}$ and the minimum is understood as $+\infty$ if $I_D=\emptyset$. From Definition 2.1 it is easy to verify that $t_D>0$ and $\phi(s_D+te_D)=t$ for each $t\in[0,t_D)$. Given $t_0\in(0,t_D)$, it follows from the definition of τ_* that

$$\mathrm{dist}(s_D+t_0e_D,S)\leqslant\tau_*t_0. \tag{27}$$

Let $S_D := \{x \in E : \langle a_j, x\rangle - c_i \leqslant 0,\ \forall \in D\}$. Then $S \subset S_D$. Noting that $S_D = s_D + C_D$, it follows that

$$\mathrm{dist}(s_D + t_0 e_D, s_D + C_D) = \mathrm{dist}(s_D + t_0 e_D, S_D) \leqslant \mathrm{dist}(s_D + t_0 e_D, S).$$

Therefore, $t_0 \mathrm{dist}(e_D, C_D) \leqslant \mathrm{dist}(s_D + t_0 e_D, S)$. It follows from (27) that $\mathrm{dist}(e_D, C_D) \leqslant \tau_*$. This completes the proof. □

The Proof of Theorem 2.1. Let $E_0 = \bigcap_{i\in I}\{x \in E: \langle a_i^*, x\rangle = 0\}$. Then E/E_0 is finite dimensional. For each $x \in E$, let $[x]$ denote the equivalence class containing x in E/E_0, that is, $[x] = x + E_0$. Define $\hat{a}_i^* \in (E/E_0)^*$ such that $\langle \hat{a}_i^*, [x]\rangle = \langle a_i^*, x\rangle$ for each $x \in E$ and $i \in I$. It is clear that

$$(E/E_0)^* = \mathrm{span}\{\hat{a}_i^* : i \in I\}.$$

Let $\hat{\phi}([x]) = \max\{\langle \hat{a}_i^*, [x]\rangle - c_i\colon i \in I\}$ for each $[x] \in E/E_0$ and $\widehat{S} = \{[x] \in E/E_0 : \hat{\phi}([x]) \leqslant 0\}$. It is easy to verify that a subset D of I has the property (W) with respect to $\{a_i^*\colon i \in I\}$ if and only if it has the property (W) with respect to $\{\hat{a}_i^* : i \in I\}$. We equip E/E_0 with the norm $|||\cdot||| : |||[x]||| = \inf\{\|y\| : y \in [x]\} \forall x \in E$. Noting that $\hat{S} = \{[x]\colon x \in S\}$ and $E_0 + S = S$, one has that $\mathrm{dist}(x, S) = \mathrm{dist}([x], \widehat{S})$ for each $x \in E$. These and Lemma 3.3 imply that Theorem 2.1 holds.

References

1. Bergthaller, C. and Singer, I. (1992), The distance to a polyhedron, *Linear Algebra and its Applications*, 169, 111–129.
2. Burke, J.V. and Ferris, M.C. (1993), Weak sharp minima in mathematical programming, *SIAM . Control and Optim.*, 31, 1340–1359.
3. Burke, J.V. and Tseng, P. (1996), A unified analysis of Hoffman's bound via Fenchel duality, *SIAM Journal on Optimization*, 6, 265–282.
4. Guler, O., Hoffman, A.J. and Rothblum, U.G. (1995), Approximations to solutions to systems of linear inequalities, *SIAM Journal on Matrix Analysis and Applications*, 16, 688–696.
5. Hoffman, A.J. (1952), On approximate solutions of systems of linear inequalities, *Journal of Research of the National Bureau of Standards*, 49, 263–265.
6. Klatte, D. and Thiere, G. (1995), Error bounds for solutions of linear equations and inequalities, *Mathematical Methods of Operations Research*, 41, 191–214.
7. Lewis, A.S. and Pang, J.S. (1996), Error bounds for convex inequality systems, In: J.P. Couzeix (ed.), *Proceedings of the Fifth Symposium on Generalized Convexity*, Luminy–Marseille
8. Li, W. (1993), The sharp lipschitz constants for feasible and optimal solutions of a perturbed linear program, *Linear Algebra and its Applications*, 187, 15–40.
9. Mangasarian, O.L. and Shiau, T.H. (1986), A variable-complexity norm maximization problem, *SIAM J. Alg. Discrete Methods*, 7, 455–461.
10. Pang, J.S. (1997), Error bounds in mathematical programming, *Mathematical Programming* 79, 299–332.
11. Rockafellar, R.T. (1970), *Convex Analysis*, Princeton University Press, Princeton, NJ.
12. Walkup, D.W. and Wets, R.J.B. (1969), Wets, A Lipschitzian characterization of convex polyhedra, *Proceedings of the American Mathematical Society*, 20, 167–173.