An easily computable upper bound on the Hoffman constant for homogeneous inequality systems

Javier F. Peña*

February 7, 2023

Abstract

Let $A \in \mathbb{R}^{m\times n} \setminus \{0\}$ and $P := \{x : Ax \leq 0\}$. This paper provides a procedure to compute an upper bound on the following *homogeneous Hoffman constant*

$$H_0(A) := \sup_{u\in\mathbb{R}^n\setminus P} \frac{\operatorname{dist}(u,P)}{\operatorname{dist}(Au,\mathbb{R}^m_-)}.$$

In sharp contrast to the intractability of computing more general Hoffman constants, the procedure described in this paper is entirely tractable and easily implementable.

1 Introduction

Hoffman constants for systems of linear inequalities, and more general error bounds for feasibility problems, play a central role in mathematical programming. In particular, Hoffman constants provide a key building block for the convergence of a variety of algorithms [1, 3, 9, 10, 12, 19]. Since Hoffman's seminal work [6], Hoffman constants and more general error bounds has been widely studied [2,4,5,11,13,16,20,21]. However, there has been very limited work on algorithmic procedures that compute or bound Hoffman constants. The only two references that appear to tackle this computational challenge are the 1995 article by Klatte and Thiere [8] and the more recent 2021 article by Peña, Vera, and Zuluaga [15]. However, as it is discussed in both [8] and [15], there are limitations on the algorithmic schemes proposed in both these articles.

The central goal of this paper is to devise a procedure that computes an upper bound on the following *homogeneous Hoffman constant* $H_0(A)$. Suppose $A \in \mathbb{R}^{m\times n}$. Let $P := \{x : Ax \leq 0\}$ and define $H_0(A)$ as

$$H_0(A) := \sup_{u\in\mathbb{R}^n\setminus P} \frac{\operatorname{dist}(u,P)}{\operatorname{dist}(Au,\mathbb{R}^m_-)}.$$

For notational convenience, by convention let $H_0(A) := 0$ when $P = \mathbb{R}^n$. This occurs precisely when $A = 0$.

*Tepper School of Business, Carnegie Mellon University, USA, `jfp@andrew.cmu.edu`

To position this work in the context of Hoffman constants, we next recall the *local* and *global* Hoffman constants $H(A,b)$ and $H(A)$ associated to linear systems of inequalities defined by A. The homogeneous Hoffman constant $H_0(A)$ is a special case of the following *local Hoffman constant* $H(A,b)$. Suppose $A \in \mathbb{R}^{m\times n}$ and $b \in A\mathbb{R}^n + \mathbb{R}^m_+$. Let $P_A(b) := \{x \in \mathbb{R}^n : Ax \leq b\}$ and define $H(A,b)$ as

$$H(A,b) := \sup_{u\in\mathbb{R}^n\setminus P_A(b)} \frac{\text{dist}(u, P_A(b))}{\text{dist}(Au, \mathbb{R}^m_-)}.$$

It is evident that $H_0(A) = H(A,0)$ and thus $H_0(A)$ is bounded above by the following *global Hoffman constant* $H(A)$. Suppose $A \in \mathbb{R}^{m\times n}$. Define

$$H(A) := \sup_{b\in A\mathbb{R}^n+\mathbb{R}^m_+} H(A,b).$$

In his seminal paper [6], Hoffman showed that $H(A)$ is finite and consequently so are $H_0(A)$ and $H(A,b)$ for all $b \in A\mathbb{R}^n + \mathbb{R}^m_+$.

The articles [8,15] propose algorithms to compute or estimate the global Hoffman constant $H(A)$. These algorithms readily yield a computational procedure to bound $H_0(A)$. However, as it is detailed in [8,15], except for very special cases the computation or even approximation of $H(A)$ is an extremely challenging problem. Indeed, the recent results in [14] show that the Stewart-Todd condition measure $\chi(A)$ [17,18] is the same as $H(\mathbf{A})$ where $\mathbf{A} = \begin{bmatrix} A \\ -A \end{bmatrix}$. Since the quantity $\chi(A)$ is known to be NP-hard to approximate [7], so is $H(A)$. In sharp contrast, the procedure proposed in this paper for upper bounding the more specialized Hoffman constant $H_0(A)$ is entirely tractable and easily implementable for any $A \in \mathbb{R}^{m\times n}$. The bound is a formalization of the following three-step approach detailed in Section 2.

First, upper bound $H_0(A)$ in the following two special cases:

(i) When $A\hat{x} < 0$ for some $\hat{x} \in \mathbb{R}^n$ or equivalently when $A^\mathsf{T} y = 0, y \geq 0 \Rightarrow y = 0$. (See Proposition 1.)

(ii) When $A^\mathsf{T}\hat{y} = 0$ for some $\hat{y} > 0$ or equivalently when $Ax \leq 0 \Rightarrow Ax = 0$. (See Proposition 2.)

Second, use a canonical partition $A = \begin{bmatrix} A_B \\ A_N \end{bmatrix}$ of the rows of A such that A_N is as in case (i) and A_B is as in case (ii) above. (See Proposition 3.)

Third, upper bound $H_0(A)$ by stitching together the Hoffman constants $H_0(A_B)$, $H_0(A_N)$, and a third Hoffman constant $\mathcal{H}(L,K)$ associated to the intersection of the subspace $L := \{x : A_B x = 0\}$ and the cone $K := \{x : A_N x \leq 0\}$. (See Theorem 1.)

The above steps suggest the following computational procedure to upper bound $H_0(A)$: First, compute the partition B, N. Second, compute upper bounds on $H_0(A_B)$ and on $H_0(A_N)$. Third, upper bound $\mathcal{H}(L,K)$. Section 3 details this procedure. As explained in Section 3, the total computational work in the entire procedure consists of two linear programs, two quadratic programs, a convex program, and a singular value calculation, all of which are computationally tractable. This is noteworthy in light of the challenges associated to estimating the Hoffman constants $H(A)$ and $H(A,b)$. A Python implementation and some illustrative examples of this procedure are publicly available at

`https://github.com/javi-pena`

For ease of notation and computability, we assume throughout the paper that the norm in $\mathbb{R}^m$ satisfies the following *componentwise compatibility condition:* if $y,z \in \mathbb{R}^m$ and $|y| \le |z|$ componentwise then $\|y\| \le \|z\|$. The componentwise compatibility condition in particular implies that for all $u \in \mathbb{R}^n$

$$\mathrm{dist}(Au, \mathbb{R}^n_-) = \|(Au)^+\|$$

where $(Au)^+ = \max\{Au, 0\}$ componentwise. Consequently,

$$H_0(A) = \sup_{u \in \mathbb{R}^n \setminus P} \frac{\mathrm{dist}(u,P)}{\|(Au)^+\|}.$$

Observe that most of the usual norms in $\mathbb{R}^m$, including the ℓ_p norms for $1 \le p \le \infty$ satisfy the componentwise compatibility condition.

2 Upper bounds on $H_0(A)$

2.1 Upper bounds on $H_0(A)$ in two special cases

We next consider two special cases that can be seen as dual counterparts of each other.

Proposition 1. *Suppose $A \in \mathbb{R}^{m\times n}$ and $A\hat{x} < 0$ for some $\hat{x} \in \mathbb{R}^n$ or equivalently $A^\mathsf{T} y = 0, y \ge 0 \Rightarrow y = 0$. Then*

$$H_0(A) \le \max_{\substack{y \in \mathbb{R}^m \\ \|y\| \le 1}} \min_{\substack{x \in \mathbb{R}^n \\ Ax \le y}} \|x\|. \tag{1}$$

Proof. For ease of notation, let H denote the right-hand side expression in (1). Observe that $H < +\infty$ because the assumption on A implies that $A\mathbb{R}^n + \mathbb{R}^m_+ = \mathbb{R}^m$.

We need to show that $H_0(A) \le H$. To that end, let $P := \{x \in \mathbb{R}^n : Ax \le 0\}$ and suppose that $u \in \mathbb{R}^n \setminus P$. Let $y := (Au)^+ \in \mathbb{R}^m$. The construction of H implies that there exists $x \in \mathbb{R}^n$ such that $Ax \le -y$ and $\|x\| \le H \cdot \|y\| = H \cdot \|(Au)^+\|$. Thus $x + u \in P$ because

$$A(x+u) = Ax + Au \le -y + Au = -(Au)^+ + Au \le 0.$$

Furthermore $\|(x+u) - u\| = \|x\| \le H \cdot \|(Au)^+\|$. Since this holds for all $u \in \mathbb{R}^n \setminus P$, it follows that $H_0(A) \le H$. □

In addition to the simple direct proof above, an alternative proof of Proposition 1 can also be obtained from [15]. Indeed, [15, Proposition 2] implies that when $A \in \mathbb{R}^{m\times n}$ satisfies the assumption in Proposition 1, the right-hand side in (1) is precisely the global Hoffman constant $H(A)$ which is at least as large as $H_0(A)$ as previously noted.

For computational purposes, it is useful to note that when $\mathbb{R}^m$ is endowed with the ℓ_∞ norm, the upper bound in Proposition 1 can be computed via the following convex optimization problem:

$$\min\{\|x\| : Ax \geq \mathbf{1}\}.$$

In particular, any $\bar{x} \in \mathbb{R}^n$ such that $A\bar{x} \geq \mathbf{1}$ yields the upper bound

$$H_0(A) \leq \|\bar{x}\|.$$

The following proposition, which can be seen as a dual counterpart of Proposition 1, relies on the dual norms in $\mathbb{R}^m$ and $\mathbb{R}^n$. More precisely, suppose both $\mathbb{R}^m$ and $\mathbb{R}^n$ are endowed with their canonical inner products. In each case let $\|\cdot\|^*$ denote the norm defined as

$$\|u\|^* = \max_{\|x\|\leq 1} \langle u, x \rangle .$$

Proposition 2. *Suppose $A \in \mathbb{R}^{m\times n}$ is such that $A^\mathsf{T}\hat{y} = 0$ for some $\hat{y} > 0$ or equivalently $Ax \leq 0 \Rightarrow Ax = 0$. Then*

$$H_0(A) \leq \max_{\substack{v\in A^\mathsf{T}(\mathbb{R}^m)\\ \|v\|^*\leq 1}} \min_{y\in\mathbb{R}^m_+, A^\mathsf{T}y=v} \|y\|^*. \tag{2}$$

Proof. Again for ease of notation, let H denote the right-hand side expression in (2). Observe that $H < +\infty$ because the assumption on A implies that $A^\mathsf{T}\mathbb{R}^m_+ = A^\mathsf{T}\mathbb{R}^m$.

We need to show that $H_0(A) \leq H$. To that end, let $P := \{x \in \mathbb{R}^n : Ax \leq 0\} = \{x \in \mathbb{R}^n : Ax = 0\}$ and suppose that $u \in \mathbb{R}^n \setminus P$. Let

$$\bar{x} := \arg\min_{x\in P} \|u - x\| = \arg\min_{x:Ax=0} \|u - x\|.$$

The optimality conditions of the latter problem imply that there exists $v \in A^\mathsf{T}\mathbb{R}^m$ with $\|v\|^* = 1$ such that

$$\|u - \bar{x}\| = \langle v, u - \bar{x} \rangle .$$

The construction of H implies that there exists $y \in \mathbb{R}^m_+$ such that $A^\mathsf{T}y = v$ and $\|y\|^* \leq H$. Since $v = A^\mathsf{T}y$ we have

$$\|u - \bar{x}\| = \langle v, u - \bar{x} \rangle = \left\langle A^\mathsf{T}y, u - \bar{x} \right\rangle = \langle y, A(u - \bar{x}) \rangle = \langle y, Au \rangle .$$

In addition, since $y \in \mathbb{R}^m_+$ and $\|y\|^* \leq H$, we also have

$$\|u - \bar{x}\| = \langle y, Au \rangle \leq \left\langle y, (Au)^+ \right\rangle \leq \|y\|^* \cdot \|(Au)^+\| \leq H \cdot \|(Au)^+\|.$$

Since this holds for all $u \in \mathbb{R}^n \setminus P$, it follows that $H_0(A) \leq H$. $\square$

For computational purposes, it is useful to note that when $\mathbb{R}^m$ is endowed with the ℓ_∞ norm, the upper bound in Proposition 2 can be computed as follows

$$\max_{\substack{v\in A^\mathsf{T}(\mathbb{R}^m)\\ \|v\|^*\leq 1}} \min_{\substack{y\in\mathbb{R}^m_+\\ A^\mathsf{T}y=v}} \mathbf{1}^\mathsf{T}y.$$

The reciprocal of the latter quantity in turn is the radius of the largest ball in $A^\mathsf{T}(\mathbb{R}^m)$ centered at 0 and contained in the set

$$\{A^\mathsf{T} y : y \in \mathbb{R}^m_+, \mathbf{1}^\mathsf{T} y = 1\} = \{A^\mathsf{T} y : y \in \mathbb{R}^m_+, \mathbf{1}^\mathsf{T} y \leq 1\}.$$

Therefore, if in addition $\mathbb{R}^n$ is endowed with the ℓ_2 norm then any $\bar{y} \in \mathbb{R}^m_{++}$ with $\mathbf{1}^\mathsf{T}\bar{y} = 1$ and $A^\mathsf{T}\bar{y} = 0$ yields the upper bound

$$H_0(A) \leq \frac{2}{\sigma^+_{\min}(A^\mathsf{T}\bar{Y})}, \tag{3}$$

where $\bar{Y} = \mathrm{Diag}(\bar{y})$ and $\sigma^+_{\min}(A^\mathsf{T}\bar{Y})$ denotes the smallest positive singular value of $A^\mathsf{T}\bar{Y}$. To see why (3) holds, observe that if $v \in A^\mathsf{T}\mathbb{R}^m$ and $\|v\|_2 \leq \frac{\sigma^+_{\min}(A^\mathsf{T}\bar{Y})}{2}$ then $2v = A^\mathsf{T}\bar{Y}z$ for some $\|z\|_2 \leq 1$. The latter implies that $|\bar{Y}z| \leq \bar{y}$ componentwise and thus $2v = A^\mathsf{T}(\bar{y} + \bar{Y}z)$ with

$$\bar{y} + \bar{Y}z \in \mathbb{R}^m_+ \text{ and } \mathbf{1}^\mathsf{T}(\bar{y} + \bar{Y}z) \leq 2 \cdot \mathbf{1}^\mathsf{T}\bar{y} = 2.$$

In particular, $v \in \{A^\mathsf{T} y : y \in \mathbb{R}^m_+, \mathbf{1}^\mathsf{T} y \leq 1\}$. Since this holds for any $v \in A^\mathsf{T}\mathbb{R}^m$ with $\|v\|_2 \leq \frac{\sigma^+_{\min}(A^\mathsf{T}\bar{Y})}{2}$, it follows that the radius of the largest ball in $A^\mathsf{T}(\mathbb{R}^m)$ centered at 0 and contained in the set

$$\{A^\mathsf{T} y : y \in \mathbb{R}^m_+, \mathbf{1}^\mathsf{T} y = 1\} = \{A^\mathsf{T} y : y \in \mathbb{R}^m_+, \mathbf{1}^\mathsf{T} y \leq 1\}.$$

is at least $\frac{\sigma^+_{\min}(A^\mathsf{T}\bar{Y})}{2}$.

2.2 Upper bound on $H_0(A)$ for general A

An upper bound on $H(A)$ for general $A \in \mathbb{R}^{m\times n}$ follows by stitching together the cases in the above two propositions via the the canonical partition result in Proposition 3 and the additional Hoffman constant $\mathcal{H}(L, K)$ defined in (4) below.

The following result is a consequence of the classical Goldman-Tucker partition theorem. To make our exposition self-contained, we include a proof.

Proposition 3. *Let $A \in \mathbb{R}^{m\times n}$. There exists a unique partition $B \cup N = \{1, \ldots, m\}$ such that*

$$A_B\hat{x} = 0,\ A_N\hat{x} < 0 \text{ for some } \hat{x} \in \mathbb{R}^n$$

and

$$A_B^\mathsf{T}\hat{y}_B = 0 \text{ for some } \hat{y}_B > 0.$$

Proof. Let $N \subseteq \{1, \ldots, m\}$ be the largest subset of $\{1, \ldots, m\}$ such that

$$Ax \leq 0 \text{ and } A_N x < 0$$

has a solution and let $B := \{1, \ldots, m\} \setminus N$. Observe that N is well-defined and unique and thus so is B. To finish the proof it suffices to show that

$$A_B^\mathsf{T} y_B,\ y_B > 0$$

has a solution. To that end, for $i \in \{1,\dots,m\}$ let $e_i \in \mathbb{R}^n$ is the vector with i-th component equal to one and all other equal to zero. Observe that $i \in B$ if and only if the following system of equations and inequalities does not have a solution:

$$\begin{bmatrix} A & e_i \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \le 0, \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} > 0.$$

Farkas Lemma thus implies that $i \in B$ if and only if the following system of equations and inequalities has a solution:

$$\begin{bmatrix} A^\mathsf{T} \\ e_i^\mathsf{T} \end{bmatrix} y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, y \ge 0.$$

Since this holds for each $i \in B$, it follows that $A_B^\mathsf{T} y_B = 0, y_B > 0$ has a solution. □

Suppose $L \subseteq \mathbb{R}^n$ is a linear subspace and $K \subseteq \mathbb{R}^n$ is a regular cone such that $L \cap \mathrm{int}(K) \neq \emptyset$. Let

$$\mathcal{H}(L,K) = \sup_{u \in \mathbb{R}^n \setminus L \cap K} \frac{\mathrm{dist}(u, L \cap K)}{\max\{\mathrm{dist}(u,L), \mathrm{dist}(u,K)\}}. \tag{4}$$

In the remainder of this paper, we will use the following notation for $A \in \mathbb{R}^{m \times n}$: Let B, N denote the canonical partition defined by A as in Proposition 3 and let $L \subseteq \mathbb{R}^n, K \subseteq \mathbb{R}^n$ be defined as

$$L := \{x : A_B x = 0\}, K := \{x : A_N x \le 0\}.$$

Observe that L is a linear subspace, K is a regular cone, and $\{x : Ax \le 0\} = L \cap K$. We now have all the necessary ingredients to upper bound $H_0(A)$.

Theorem 1. *Suppose $A \in \mathbb{R}^{m \times n}$ and the componentwise compatibility condition holds. Let B, N and L, K be as above. Then*

$$H_0(A) \le \mathcal{H}(L,K) \cdot \max\{H_0(A_N), H_0(A_B)\}. \tag{5}$$

Proof. Suppose $u \in \mathbb{R}^n \setminus P$. The construction of $\mathcal{H}(\cdot,\cdot)$ and $H_0(\cdot)$, and the componentwise compatibility condition imply that there exists $x \in P = L \cap K$ such that

$$\begin{aligned} \|x - u\| &\le \mathcal{H}(L,K) \cdot \max\{\mathrm{dist}(u,L), \mathrm{dist}(u,K)\} \\ &\le \mathcal{H}(L,K) \cdot \max\{H_0(A_B) \cdot \|(A_B u)^+\|, H_0(A_N) \cdot \|(A_N u)^+\|\} \\ &\le \mathcal{H}(L,K) \cdot \max\{H_0(A_B), H_0(A_N)\} \cdot \|(Au)^+\|. \end{aligned}$$

Since this holds for all $u \in \mathbb{R}^n \setminus P$, the inequality in (5) follows. □

Observe that unlike $H_0(A)$ that depends on the *data representation* $A \in \mathbb{R}^{m \times n}$ of the cone $P = \{x : Ax \le 0\}$, the constant $\mathcal{H}(L,K)$ only depends on the sets $L \subseteq \mathbb{R}^n$ and $K \subseteq \mathbb{R}^n$. In particular, $\mathcal{H}(L,K)$ does not depend on the norm in $\mathbb{R}^m$ while $H_0(A)$ evidently does.

The next proposition provides an upper bound on $\mathcal{H}(L,K)$ analogous to the upper bounds on $H_0(A)$ in Proposition 1 and Proposition 2. It will be useful for the computational procedure in Section 3.

Proposition 4. *Suppose $L \subseteq \mathbb{R}^n$ is a linear subspace and $K \subseteq \mathbb{R}^n$ is a regular cone such that $L \cap \mathrm{int}(K) \neq \emptyset$. Then*

$$\mathcal{H}(L, K) \leq 1 + 2 \cdot \max_{\substack{u \in \mathbb{R}^n \\ \|u\| \leq 1}} \min_{\substack{x \in L, y \in K \\ x - y = u}} \|x\|.$$

Proof. To ease notation, let

$$H := \max_{\substack{u \in \mathbb{R}^n \\ \|u\| \leq 1}} \min_{\substack{x \in L, y \in K \\ x - y = u}} \|x\|.$$

We need to show that $\mathcal{H}(L, K) \leq 1 + 2H$. To that end, suppose $u \in \mathbb{R}^n \setminus L \cap K$. Let $u_L := \arg\min_v \{\|u - v\| : v \in L\}$ and $u_K := \arg\min_v \{\|u - v\| : v \in K\}$. The construction of H implies that there exist $x \in L, y \in K$ such that $\|x\| \leq H \cdot \|u_K - u_L\|$ and $x - y = u_K - u_L$. Hence $u_L + x = u_K + y \in L \cap K$ and

$$\begin{aligned}
\mathrm{dist}(u, L \cap K) &\leq \|u - u_L - x\| \\
&\leq \|u - u_L\| + \|x\| \\
&\leq \|u - u_L\| + H \cdot \|u_K - u_L\| \\
&\leq \max\{\mathrm{dist}(u, L), \mathrm{dist}(u, K)\} + H \cdot (\mathrm{dist}(u, K) + \mathrm{dist}(u, L)) \\
&\leq (1 + 2H) \cdot \max\{\mathrm{dist}(u, L), \mathrm{dist}(u, K)\}.
\end{aligned}$$

Since this holds for any $u \in \mathbb{R}^n \setminus L \cap K$, it follows that

$$\mathcal{H}(L, K) \leq 1 + 2H.$$

□

For computational purposes, it is useful to note that if $\bar{x} \in L \cap \mathrm{int}(K)$ is such that $\bar{x} + u \in K$ for all $\|u\| \leq 1$ then Proposition 4 implies that

$$\mathcal{H}(L, K) \leq 1 + 2\|\bar{x}\|.$$

3 A computable procedure to bound $H_0(A)$

We next describe a procedure to compute an upper bound on $H_0(A)$. The procedure consists of four main steps. First, compute the partition B, N. Second, compute an upper bound on $H_0(A_B)$. Third, compute an upper bound on $H_0(A_N)$. Fourth, compute an upper bound on $\mathcal{H}(L, K)$. An upper bound on $H_0(A)$ thereby follows from Theorem 1. For computational convenience, throughout this section we assume that $\mathbb{R}^m$ is endowed with the ℓ_∞ norm and $\mathbb{R}^n$ is endowed with the ℓ_2 norm. A Python implementation and some illustrative examples of this procedure are publicly available at

`https://github.com/javi-pena`

Step 1: partition B, N

The partition B, N can be obtained from any point (x, y, s, t) that satisfies the following systems of equations and inequalities for some $t > 0$:

$$\begin{aligned} &A^\mathsf{T} y = 0 \\ &Ax + s = 0 \\ &y + s - t\mathbf{1} \geq 0 \\ &\mathbf{1}^\mathsf{T} y + \mathbf{1}^\mathsf{T} s = 1 \\ &y \geq 0, s \geq 0. \end{aligned} \tag{6}$$

More precisely, if (x, y, s, t) satisfies (6) with $t > 0$ then B, N can be obtained as follows:

$$B := \{i : y_i > 0\}, \quad N := \{i : s_i > 0\}.$$

Proposition 3 guarantees that a solution (x, y, s, t) to (6) with $t > 0$ always exists and that the associated partition B, N is unique. Such a point (x, y, s, t) can be computed via the following linear program:

$$\begin{aligned} \max_{x,y,s,t} \quad & t \\ & A^\mathsf{T} y = 0 \\ & Ax + s = 0 \\ & y + s - t\mathbf{1} \geq 0 \\ & \mathbf{1}^\mathsf{T} y + \mathbf{1}^\mathsf{T} s = 1 \\ & y \geq 0, s \geq 0. \end{aligned} \tag{7}$$

Step 2: upper bound on $H_0(A_N)$

The remarks following Proposition 1 show that

$$H_0(A_N) \leq \|\bar{x}\|_2$$

for any $\bar{x} \in \mathbb{R}^n$ such that $A_N \bar{x} \geq \mathbf{1}$. The best such upper bound can be computed via the following quadratic program

$$\bar{x} := \arg\min\{\|x\|_2^2 : A_N x \geq \mathbf{1}\}. \tag{8}$$

Step 3: upper bound on $H_0(A_B)$

The remarks following Proposition 2 show that

$$H_0(A_B) \leq \frac{2}{\sigma^+_{\min}(A_B^\mathsf{T} \bar{Y})}$$

for any $\bar{y} \in \mathbb{R}^B_{++}$ such that $\mathbf{1}_B^\mathsf{T} \bar{y} = 1$ and $A_B^\mathsf{T} \bar{y} = 0$. Although the best such upper bound is challenging to compute, an upper bound of this kind that is within a factor of $\sqrt{|B|}$ of the best possible one can be computed via the following convex program

$$\bar{y} := \arg\min_{y \in \mathbb{R}^B_{++}} \left\{ -\sum_{i \in B} \log(y_i) : \mathbf{1}_B^\mathsf{T} y = 1, A_B^\mathsf{T} y = 0 \right\}. \tag{9}$$

Step 4: upper bound on $\mathcal{H}(L, K)$

Let Q be an orthonormal basis for $L := \{x : A_B x = 0\}$ and $M = DA_N Q$ where D is the diagonal matrix with positive diagonal entries such that all rows of DA_N have Euclidean norm equal to one. Then the remarks following Proposition 4 imply that

$$\mathcal{H}(L, K) \leq 1 + 2\|\bar{z}\|_2$$

for any $\bar{z} \geq 0$ such that $M\bar{z} \geq \mathbf{1}$. The best such upper bound can be computed via the following quadratic program

$$\bar{z} := \arg\min\{\|z\|_2^2 : Mz \geq \mathbf{1}\}. \tag{10}$$

Putting it all together: a procedure to bound $H_0(A)$

Theorem 1 allows us to stitch together the partition B, N and the upper bounds on $H_0(A_B)$, $H_0(A_N)$, and $\mathcal{H}(L, K)$ to obtain an upper bound on $H_0(A)$ as detailed in Algorithm 1 below.

4 Funding, conflict of interest, and data information

This research has been supported by the Bajaj Family Chair at the Tepper School of Business. The author has no competing interests to declare that are relevant to the content of this article. The developments on this paper followed a theoretical mathematical approach and did not analyze or generate any datasets.

References

[1] D. Applegate, O. Hinder, H. Lu, and M. Lubin. Faster first-order primal-dual methods for linear programming using restarts and sharpness. *Mathematical Programming*, pages 1–52, 2022.

[2] D. Azé and J. Corvellec. On the sensitivity analysis of Hoffman constants for systems of linear inequalities. *SIAM Journal on Optimization*, 12(4):913–927, 2002.

[3] A. Beck and S. Shtern. Linearly convergent away-step conditional gradient for non-strongly convex functions. *Mathematical Programming*, 164(1):1–27, 2017.

[4] J. Burke and P. Tseng. A unified analysis of Hoffman's bound via Fenchel duality. *SIAM Journal on Optimization*, 6(2):265–282, 1996.

[5] O. Güler, A. Hoffman, and U. Rothblum. Approximations to solutions to systems of linear inequalities. *SIAM Journal on Matrix Analysis and Applications*, 16(2):688–696, 1995.

[6] A. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards*, 49(4):263–265, 1952.

Algorithm 1 Upper bound on $H_0(A)$

input: $A \in \mathbb{R}^{m \times n} \setminus \{0\}$
solve (7) to enough accuracy to get a solution (x, y, s, t) to (6) with $t > 0$
Let $B := \{i : y_i > 0\}, N := \{i : s_i > 0\}$
if $N \neq \emptyset$ **then**
    solve (8) to enough accuracy to get $\bar{x} \in \mathbb{R}^n$ such that $A_N \bar{x} \geq \mathbf{1}$
end if
if $B \neq \emptyset$ **then**
    solve (9) to enough accuracy to get $\bar{y} \in \mathbb{R}^B_{++}$ such that $\mathbf{1}_B^\mathsf{T} \bar{y} = 1$ and $A_B^\mathsf{T} \bar{y} = 0$
end if
if $N = \emptyset$ **then** return the upper bound

$$H_0(A) \leq \frac{2}{\sigma^+_{\min}(A_B^\mathsf{T} \bar{Y})}$$

end if
if $B = \emptyset$ **then** return the upper bound

$$H_0(A) \leq \|\bar{x}\|_2$$

end if
let Q be an orthonormal basis for $L := \{x : A_B x = 0\}$ and $M = DA_N Q$ where D is the diagonal matrix with positive diagonal entries such that all rows of DA_N have Euclidean norm equal to one
solve (10) to enough accuracy to get $\bar{z}$ such that $M\bar{z} \geq \mathbf{1}$
return the upper bound

$$H_0(A) \leq (1 + 2\|\bar{z}\|_2) \cdot \max\left\{\|\bar{x}\|_2, \frac{2}{\sigma^+_{\min}(A_B^\mathsf{T} \bar{Y})}\right\}$$

[7] L. Khachiyan. On the complexity of approximating extremal determinants in matrices. *Journal of Complexity*, 11(1):138–153, 1995.

[8] D. Klatte and G. Thiere. Error bounds for solutions of linear equations and inequalities. *Zeitschrift für Operations Research*, 41(2):191–214, 1995.

[9] S. Lacoste-Julien and M. Jaggi. On the global linear convergence of Frank-Wolfe optimization variants. In *Advances in Neural Information Processing Systems (NIPS)*, 2015.

[10] D. Leventhal and A. Lewis. Randomized methods for linear constraints: Convergence rates and conditioning. *Math. Oper. Res.*, 35:641–654, 2010.

[11] W. Li. The sharp Lipschitz constants for feasible and optimal solutions of a perturbed linear program. *Linear algebra and its applications*, 187:15–40, 1993.

[12] Z. Luo and P. Tseng. Error bounds and convergence analysis of feasible descent methods: a general approach. *Annals of Operations Research*, 46(1):157–178, 1993.

[13] O. Mangasarian and T-H Shiau. Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems. *SIAM Journal on Control and Optimization*, 25(3):583–595, 1987.

[14] J. Peña, J. Vera, and L. Zuluaga. Equivalence and invariance of the chi and Hoffman constants of a matrix. *arXiv preprint arXiv:1905.06366*, 2019.

[15] J. Peña, J. Vera, and L. Zuluaga. New characterizations of Hoffman constants for systems of linear constraints. *Math. Program.*, 187(1):79–109, 2021.

[16] S. Robinson. Bounds for error in the solution set of a perturbed linear program. *Linear Algebra and its applications*, 6:69–81, 1973.

[17] G. Stewart. On scaled projections and pseudoinverses. *Linear Algebra and its Applications*, 112:189–193, 1989.

[18] M. Todd. A Dantzig-Wolfe-like variant of Karmarkar's interior-point linear programming algorithm. *Operations Research*, 38(6):1006–1018, 1990.

[19] P. Wang and C. Lin. Iteration complexity of feasible descent methods for convex optimization. *Journal of Machine Learning Research*, 15(1):1523–1548, 2014.

[20] C. Zalinescu. Sharp estimates for Hoffman's constant for systems of linear inequalities and equalities. *SIAM Journal on Optimization*, 14(2):517–533, 2003.

[21] X. Zheng and K. Ng. Hoffman's least error bounds for systems of linear inequalities. *Journal of Global Optimization*, 30(4):391–403, 2004.