Journal of Global Optimization **24:** 133–147, 2002.

Optimal Hoffman-Type Estimates in Eigenvalue and Semidefinite Inequality Constraints

DOMINIQUE AZÉ and JEAN-BAPTISTE HIRIART-URRUTY
UMR CNRS MIP, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse Cedex, France (e-mail: aze@mip.ups-tlse.fr; jbhu@cict.fr)

Abstract. Starting from a general Hoffman-type estimate for inequalities defined via convex functions, we derive estimates of the same type for inequality constraints expressed in terms of eigenvalue functions (as in eigenvalue optimization) or positive semidefiniteness (as in semidefinite programming).

Key words: Convex inequalities; Hoffman-type estimates; Eigenvalue inequality constraints; Semidefinite programming

1. Introduction and preliminary results

In his celebrated result dating back to 1952 [16], A.J. Hoffman showed how the distance to a convex polyhedron of $\mathbb{R}^n$ could be estimated (from above) in terms of the affine functions defining it. Numerous works have been devoted since to extensions of this result, we mention here some recent ones [2, 4, 8, 9, 18, 20, 21]. In the present paper, we start by recalling a general Hoffman-type result obtained in the context of convex inequalities. We then make it operate for inequalities defined by some eigenvalue functions, and for inequalities expressed in terms of positive semidefiniteness of some matrices (such as occurring in semidefinite programming).

Let us fix the notations used in the statement of our preliminary result. For a given Banach space $(X, \|\cdot\|)$, we denote by d (resp. d_*) the distance associated with $\|\cdot\|$ (resp. the one associated with the dual norm $\|\cdot\|_*$ of $\|\cdot\|$ in the dual space X^* of X); $B(\bar{x}, r)$ (resp. $\bar{B}(\bar{x}, r)$ stands for the ball (resp. the closed ball) centered at $\bar{x}$ and of radius r. For a function $f: X \to \mathbb{R} \cup \{+\infty\}$, we use the following notations for the sublevel-sets:

$$[f \leqslant a] = \{x \in X : f(x) \leqslant a\}, \qquad [f > a] = \{x \in X : f(x) > a\}.$$

The other definitions and notations, from Convex analysis and optimization (such as the support function σ_C of a convex set C and the distance function $d(\cdot, C)$, the subdifferential ∂f of a convex function f), are the ones commonly used in such an area.

We shall mainly rely on the following theorem taken from [4]; for the convenience of the reader, we provide a self-contained proof of it.

THEOREM 1.1. *Let X be a Banach space, let $f: X \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex, and lower semicontinuous function. Then,*

$$\inf_{x \in [f>0]} d_*(0, \partial f(x)) = \inf_{x \in [f>0]} \frac{f(x)}{d(x, [f \leqslant 0])} \tag{1}$$

(*with the convention that the member on the right is zero if* $[f \leqslant 0] = \emptyset$.)

Proof. Let $\sigma := \inf_{x \in [f>0]} d_*(0, \partial f(x))$. Assume that $\sigma > 0$ and let $0 < \tau < \sigma$. We claim that

$$\tau d(x, [f \leqslant 0]) \leqslant f(x)^+ \quad \text{for all } x \in [f > 0]\,.$$

Otherwise, there exists $\bar{x} \in [f > 0]$ such that

$$\tau d(\bar{x}, [f \leqslant 0]) > f(\bar{x})\,.$$

Introducing $r := d(\bar{x}, [f \leqslant 0]) > 0$ and $g := \sup(f, 0)$, we get

$$g(\bar{x}) < \inf_{\bar{B}(\bar{x},r)} g + \tau r\,.$$

From the Ekeland variational principle [10], there exists $z \in B(\bar{x}, r)$ such that

$$g(z) \leqslant g(y) + \tau\|z - y\| \quad \text{for all } y \in \bar{B}(\bar{x}, r)\,,$$

so that z is a local (and then global) minimum of the convex function $g(\cdot) + \tau\|z - \cdot\|$. It follows that there exists $\zeta \in \partial g(z)$ such that $\|\zeta\|_* \leqslant \tau$. Observe now that $f(z) > 0$ since $\|\bar{x} - z\| < d(\bar{x}, [f \leqslant 0])$, thus $\partial f(z) = \partial g(z)$ yielding the contradiction: $f(z) > 0$ and $d_*(0, \partial f(z)) \leqslant \tau < \sigma$. Letting τ increase to σ, we get

$$\sigma d(x, [f \leqslant 0]) \leqslant f(x)^+ \quad \text{for all } x \in [f > 0]\,,$$

hence

$$\inf_{x \in [f>0]} \frac{f(x)}{d(x, [f \leqslant 0])} \geqslant \sigma = \inf_{x \in [f>0]} d_*(0, \partial f(x))\,,$$

and this inequality is also satisfied if $x = 0$.

Conversely, let $\tau = \inf_{x \in [f>0]} \frac{f(x)}{d(x, [f \leqslant 0])}$. Assuming that $\tau > 0$, we get

$$\tau d(x, [f \leqslant 0]) \leqslant f(x)^+ \quad \text{for all } x \in X\,.$$

Let $x \in [f > 0]$ and let $\xi \in \partial f(x)$. Given $0 < \varepsilon < \tau$, there exists $z \in [f \leqslant 0]$ such that $\|x - z\|(\tau - \varepsilon) \leqslant f(x)$, yielding

$$\|x - z\|(\tau - \varepsilon) \leqslant f(x) - f(z) \leqslant \|\xi\|_*\|x - z\|\,,$$

thus $\|\xi\|_* \geqslant \tau - \varepsilon$, hence $d_*(0, \partial f(x)) \geqslant \tau$ by letting ε go to 0, and then

$$\inf_{x \in [f>0]} d_*(0, \partial f(x)) \geqslant \tau = \inf_{x \in [f>0]} \frac{f(x)}{d(x, [f \leqslant 0])}\,,$$

and the previous inequality also holds true if $\tau = 0$. □

The theorem above tells us that a necessary and sufficient condition for the existence of a positive real number $\tau > 0$ such that

$$d(x, [f \leqslant 0]) \leqslant \frac{1}{\tau} f(x)^+ \quad \text{for all } x \in X , \tag{2}$$

is that

$$\inf_{x \in [f > 0]} d_*(0, \partial f(x)) > 0 . \tag{3}$$

It also proves that $\inf_{x \in [f>0]} d_*(0, \partial f(x))$ is then the optimal (i.e., the largest) positive real number τ such that (2) holds true.

A particular case of Theorem 1.1 will be of a constant use in the sequel.

COROLLARY 1.1. *Let $C \subset X^*$ be a nonempty closed convex set such that $0 \notin C$; let σ_C denote the support function of C. Then, for all $\lambda \in \mathbb{R}$, we have*

$$d(x, [\sigma_C \leqslant \lambda]) \leqslant \frac{1}{\tau_*} (\sigma_C(x) - \lambda)^+ \quad \textit{for all } x \in X , \tag{4}$$

where $\tau_* := d_*(0, C)$.

Proof. Let $f = \sigma_C - \lambda$. Using techniques and results from Convex analysis, the following is easy to derive:

$$\begin{cases} [f > 0] = \emptyset \Rightarrow C = \{0\} ; \\ [f \leqslant 0] = \emptyset \Rightarrow 0 \in C . \end{cases}$$

Hence, our assumption on C makes that both $[f > 0]$ and $[f \leqslant 0]$ are nonempty.

According to (2),

$$d(x, [\sigma_C \leqslant \lambda]) \leqslant \frac{1}{\tau} (\sigma_C(x) - \lambda)^+ \quad \text{for all } x \in X ,$$

where $\tau = \inf_{x \in [f>0]} d_*(0, \partial f(x)))$. Now, due to the specific structure of the function f involved,

$$\partial f(x) = \partial \sigma_C(x) \subset C \quad \text{for all } x \in X .$$

Therefore $\tau \geqslant d_*(0, C) = \tau_*$, and the inequality (4) is proved. □

Theorem 1.1 can be illustrated geometrically. Assuming the three dimensional space $\mathbb{R}^3$ is embedded with the standard euclidean norm,

- the left-hand side of (1) represents the smallest slope of a tangent line to the graph of f, when x runs through $[f > 0]$;
- the right-hand side of (1) measures the smallest tangent value of angles α designed from the triangle x, $f(x)$ and the projection of x on $[f \leqslant 0]$, when x runs through $[f > 0]$.

REMARK 1.1. Corollary does not claim that τ_* is the optimal real number such

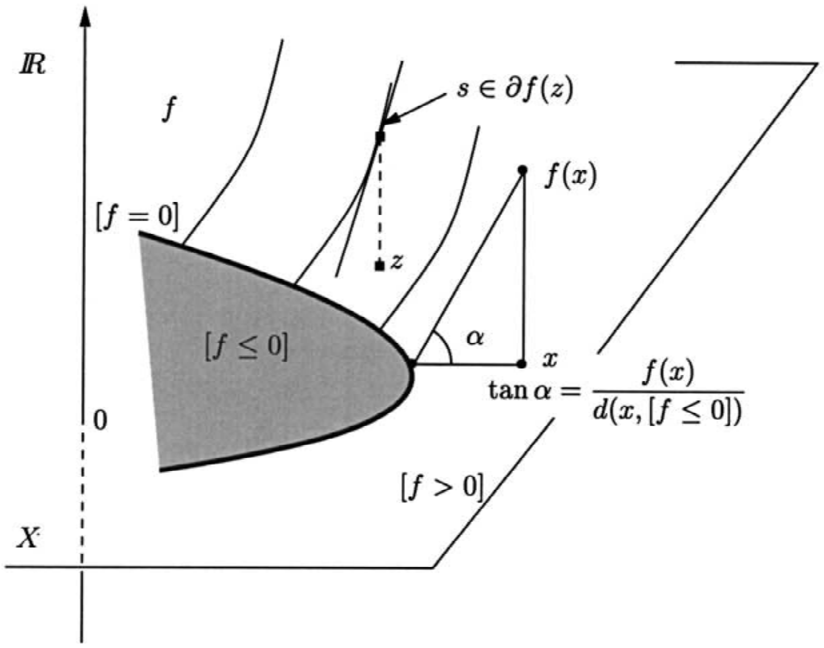

Figure 1. Geometrical interpretation of (1).

that (4) holds true; it just provides a general inequality. In examples below, we however shall see how such τ_* is shown to be optimal.

2. Optimal Hoffman-type estimates in some eigenvalue inequality constraints

Let $\mathscr{S}^n(\mathbb{R})$ denote the space of real (n, n) symmetric matrices, endowed with the standard inner product

$$\langle\langle A, B) \rangle\rangle := \mathrm{Trace}(A^T B) ,$$

(also denoted by $A \cdot B$ in the literature). here, the norm $\|\cdot\|$ associated with $\langle\langle \cdot , \cdot \rangle\rangle$ is self-dual, so that $d = d_*$ for the corresponding distances. Note also that $\|\cdot\|$ is 'orthogonal invariant', i.e., $\|UAU^T\| = \|A\|$ whenever U is orthogonal. Let furthermore posit

$$\mathscr{S}_n^+(\mathbb{R}) = \{A \in \mathscr{S}^n(\mathbb{R}) : A \geq 0\} ,$$

and

$$\mathscr{S}_n^-(\mathbb{R}) = \{A \in \mathscr{S}^n(\mathbb{R}) : A \leq 0\} ,$$

where, as usual, $A \geq 0$ (resp. $A \leq 0$) means A is positive semidefinite (resp. negative semidefinite).

2.1. CONSTRAINT INEQUALITIES INVOLVING THE SUM OF THE m LARGEST EIGENVALUES

For an integer m between 1 and n, let f_m: $\mathcal{S}^n(\mathbb{R}) \to \mathbb{R}$ denote the (convex positively homogeneous) function which assigns to $A \in \mathcal{S}^n(\mathbb{R})$,

$$f_m(A) := \text{sum of the } m \text{ largest eigenvalues of } A\,. \tag{5}$$

We also use the notation $\lambda_{\max}$ for f_1 (the largest eigenvalue function); note that f_n is nothing else than the trace function. The functions f_m are well understood from the subdifferential calculus viewpoint (see [22] or [13] for example); the points we raise here concern the inequality constraint cones

$$\mathcal{K}_m := \{A \in \mathcal{S}^n(\mathbb{R}) : f_m(A) \leqslant 0\}\,; \tag{6}$$

for example, $\mathcal{K}_1$ is $\{A \in \mathcal{S}_n(\mathbb{R}) : \lambda_{\max}(A) \leqslant 0\}$, that is $\mathcal{S}_n^-(\mathbb{R})$.

Given $X \in \mathcal{S}_n(\mathbb{R})$, there is one and only one matrix in $\mathcal{K}_m$ closest to X, however this 'matrix nearness problem' (in the sense of [11]) does not have an explicit solution (apart from the extreme cases $m = 1$ and $m = n$). We therefore provide Hoffman-type estimates of the distance of X to $\mathcal{K}_m$.

2.1.1. *Optimal* Hoffman-*type estimates of* $d(X, \mathcal{K}_m)$

The series of inclusions

$$\mathcal{S}_n^-(\mathbb{R}) = \mathcal{K}_1 \subset \cdots \subset \mathcal{K}_m \subset \cdots \subset \mathcal{K}_n$$

induces a series of inclusions in the other way for polar cones

$$\mathcal{K}_n^\circ \subset \cdots \subset \mathcal{K}_m^\circ \subset \cdots \subset \mathcal{K}_1^\circ = \mathcal{S}_n^+(\mathbb{R})\,, \tag{7}$$

where $\mathcal{K}^\circ$ denotes the negative polar cone of $\mathcal{K}$, that is

$$\mathcal{K}^\circ = \{Y \in \mathcal{S}_n(\mathbb{R}) : \langle\langle Y, X\rangle\rangle \leqslant 0 \quad \text{for all } X \in \mathcal{K}\}\,.$$

Note here that $\mathcal{K}_n^\circ$ is nothing else than the half-line $\mathbb{R}_+ I_n$ (i.e., directed by the identity matrix I_n). As a consequence of that and of (7), the projection of αI_n, $\alpha \geqslant 0$ on $\mathcal{K}_m$ is 0 for all m. See Fig. 2 to support the intuition.

The function f_m is known to be the support function of

$$\Omega_m = \{A \succeq 0 : \text{Trace } A = m \text{ and } \lambda_{\max}(A) \leq 1\}$$

(see [22, 13]). In the particular case when $m = 1$,

$$\Omega_1 = \{A \succeq 0 : \text{Trace } A = 1\}$$

is the so-called *unit spectraplex* of $\mathcal{S}_n(\mathbb{R})$.

We now are ready to derive from Corollary 1.1 the following Hoffman-type estimate for $d(X, \mathcal{K}_m)$.

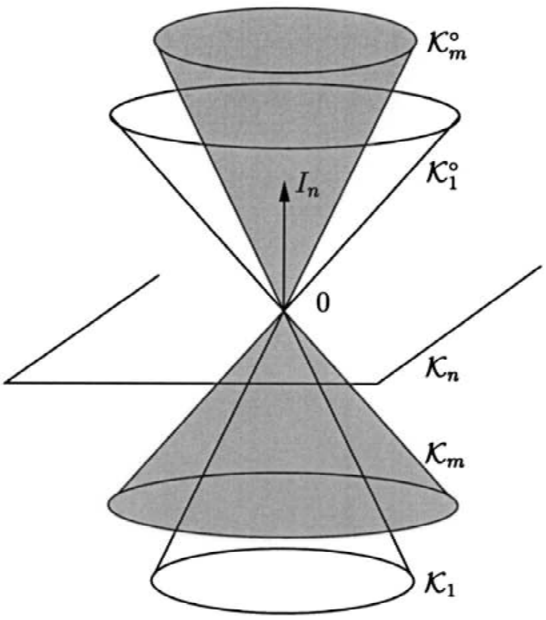

Figure 2.

THEOREM 2.1. *We have*

$$d(X, \mathscr{K}_m) \leqslant \frac{\sqrt{n}}{m}(f_m(X))^+ \quad \textit{for all } X \in \mathscr{S}_n(\mathbb{R}), \tag{8}$$

and one cannot do better than $\sqrt{n}/m$ *in (8).*

Proof. The convex compact set Ω_m does not contain 0; thus according to Corollary 1.1:

$$d(X, \mathscr{K}_m) \leqslant \frac{1}{\tau_*}(f_m(X))^+ \quad \text{for all } X \in \mathscr{S}_n(\mathbb{R}), \tag{9}$$

where $\tau_* = d(0, \Omega_m)$. We easily check that the projection of 0 on Ω_m is $(m/n)I_n$. Indeed, for all $X \in \Omega_m$,

$$\left\langle\!\!\left\langle -\frac{m}{n} I_n, X - \frac{m}{n} I_n \right\rangle\!\!\right\rangle = -\frac{m}{n} \operatorname{Trace} X + \left(\frac{m}{n}\right)^2 \operatorname{Trace} I_n = 0 .$$

Therefore, $\tau_* = m/n\|I_n\| = m/\sqrt{n}$, and (8) follows from (9).

Let $\alpha \geqslant 0$ be such that

$$d(X, \mathscr{K}_m) \leqslant \alpha(f_m(X))^+ \quad \text{for all } X \in \mathscr{S}_n(\mathbb{R}) .$$

Consider the particular $X := I_n/m$. As explained earlier, the projection of X on $\mathscr{K}_m$ is 0, whence $d(X, \mathscr{K}_m) = \|X\| = \sqrt{n}/m$. Since $f_m(X) = 1$, we get that $\sqrt{n}/m \leqslant \alpha$. □

2.1.2. *The extreme cases $m = n$ and $m = 1$*

The case $m = n$ does not offer much interest: the projection $\bar{X}$ of $X \in \mathcal{S}_n(\mathbb{R})$ on the half-space $\mathcal{K}_n$ is known explicitly,

$$\bar{X} = X - \frac{(\text{Trace } X)^+}{n} I_n ,$$

whence $d(X, \mathcal{K}_n) = \|X - \bar{X}\| = (\text{Trace } X)^+ / \sqrt{n}$. The inequality (8) is strengthened to equality in that case.

The case $m = 1$ deserves to be explored more, although here again the projection of any $X \in \mathcal{S}_n(\mathbb{R})$ on $\mathcal{K}_1 = \mathcal{S}^-_n(\mathbb{R})$ can be made explicit from X.

THEOREM 2.2. *For $X \in \mathcal{S}_n(\mathbb{R})$, let $(\lambda_i(X))_{i=1,\dots,n}$ denote the eigenvalues of X. Then*

(*a*) $d(X, \mathcal{K}_1) = \sqrt{\sum_{i=1}^{n} (\lambda_i^+(X))^2}$,

(*b*) $d(X, bd\ \mathcal{K}_1) = -\lambda_{max}(X)$ *whenever X is negative definite.*

Proof. Part (a) of the conclusion of the theorem is well-known in the area of matrix approximation problems (see [11, Section 3] for example). We see how to derive it immediately from the projection operations on the closed convex cones $\mathcal{K}_1 = \mathcal{S}^-_n(\mathbb{R})$ and $\mathcal{K}_1^\circ = \mathcal{S}^+_n(\mathbb{R})$.

Let U be an orthogonal matrix such that

$$X = U^T \text{Diag}(\lambda_1(X), \dots, \lambda_n(X))U . \tag{10}$$

Then, setting

$$X^+ = U^T \text{Diag}(\lambda_1^+(X), \dots, \lambda_n^+(X))U ,$$

and

$$X^- = U^T \text{Diag}(\lambda_1^-(X), \dots, \lambda_n^-(X))U ,$$

we realize that $X = X^+ - X^-$, $X^+ \in \mathcal{K}_1^\circ$, $X^- \in \mathcal{K}_1$ and $\langle\langle X^+, X^- \rangle\rangle = 0$. Thus, we have got at Moreau's decomposition of X (see, e.g., [12, p. 121]), whence X^+ and X^- are the projections of X on $\mathcal{K}_1^\circ$ and $\mathcal{K}_1$ respectively. It then remains to calculate

$$\begin{aligned} d(X, \mathcal{K}_1)^2 &= \|X - X^-\|^2 \\ &= \|U^T \text{Diag}(\lambda_1^+(X), \dots, \lambda_n^+(X))U\|^2 \\ &= \|\text{Diag}(\lambda_1^+(X), \dots, \lambda_n^+(X))\|^2 , \end{aligned}$$

whence (a) follows.

Part (b) for X negative definite, i.e., satisfying $\lambda_{\max}(X) < 0$, the distance of X to the boundary bd $\mathcal{K}_1$ of $\mathcal{K}_1$ and to the complement set of $\mathcal{K}_1$ are the same. Indeed

$$\text{bd } \mathcal{K}_1 = \{X \in \mathcal{S}_n(\mathbb{R}) : \lambda_{\max}(X) = 0\} ;$$

$$\text{int } \mathcal{K}_1 = \{X \in \mathcal{S}_n(\mathbb{R}) : \lambda_{\max}(X) < 0\} ;$$

$$\text{ext } \mathcal{K}_1 = \{X \in \mathcal{S}_n(\mathbb{R}) : \lambda_{\max}(X) > 0\} ;$$

where int $\mathscr{K}_1$ and ext $\mathscr{K}_1$ denote respectively the interior and the interior of the complement of $\mathscr{K}_1$. We have

$$\begin{aligned}(d(X, \operatorname{bd} \mathscr{K}_1))^2 &= \inf\{\|X - A\|^2 : \lambda_{\max}(A) = 0\} \\ &= \inf\{\|\operatorname{Diag}(\lambda_1(X), \ldots, \lambda_n(X)) - B\|^2 : \lambda_{\max}(B) = 0\} \\ &\quad \text{(by making use of (10) for example)} \\ &= \inf\left\{2 \sum_{i<j} b_{ij}^2 + \sum_{i=1}^{n} (\lambda_i(X) - b_{ii})^2 : \lambda_{\max}([b_{ij}]) = 0\right\}.\end{aligned}$$

Restricting to diagonal matrices $B = [b_{ij}]$ does not affect this lower bound. Therefore

$$(d(X, \operatorname{bd} \mathscr{K}_1))^2 = \inf\left\{\sum_{i=1}^{n} (\lambda_i(X) - b_i)^2 : b \in \mathbb{R}^n, \max_{1 \leqslant i \leqslant n} b_i = 0\right\}.$$

The lower bound above is achieved for the following choice of b_i's:

$$\begin{cases} b_{i_0} = 0 & \text{for } i_0 \text{ such that } \lambda_{i_0}(X) = \lambda_{\max}(X) \\ b_i = \lambda_i(X) & \text{for } i \neq i_0 . \end{cases}$$

Consequently

$$d(X, \operatorname{bd} \mathscr{K}_1) = \sqrt{(\lambda_{\max}(X))^2} = -\lambda_{\max}(X) . \qquad \square$$

We infer from (a)

$$d(X, \mathscr{K}_1) \leqslant \sqrt{n}\lambda_{\max}^{+}(X) \quad \text{for all } X \in \mathscr{S}_n(\mathbb{R})$$

which is the optimal Hoffman estimate predicted by Theorem 1.1 for the representation

$$\mathscr{K}_1 = \{X \in \mathscr{S}_n(\mathbb{R}) : \lambda_{\max}(X) \leqslant 0\},$$

by observing that $\partial f_1(\mathscr{S}_n(\mathbb{R})) = \Omega_1$ and that $d(0, \Omega_1) = \sqrt{n}$.

There is another interesting representation of $\mathscr{K}_1$ as an inequality constraint, via the so-called signed distance function. Indeed, let $\Delta_{\mathscr{K}_1}: \mathscr{S}_n(\mathbb{R}) \to \mathbb{R}$ be defined by:

$$\Delta_{\mathscr{K}_1}(X) := d(X, \mathscr{K}_1) - d(X, \mathscr{K}_1^c) \quad \text{for all } X \in \mathscr{S}_n(\mathbb{R}),$$

where $\mathscr{K}_1^c$ denotes the complement set of $\mathscr{K}_1$. In a more explicit form,

$$\Delta_{\mathscr{K}_1}(X) = \begin{cases} \lambda_{\max}(X) & \text{if } \lambda_{\max}(X) \leqslant 0 \\ \sqrt{\sum_{i=1}^{n} (\lambda_i^{+}(X))^2} & \text{if } \lambda_{\max}(X) \geqslant 0 . \end{cases} \tag{11}$$

The signed distance to a convex set has been studied in details in [14, Section I.2] and [15, Section II.B]. By applying results exposed there, we get the following description of $\Delta_{\mathscr{K}_1}$ as a support function.

THEOREM 2.3. *The function* $\Delta_{\mathscr{K}_1}$ *is convex, positively homogeneous, and Lipschitz of rank* 1 *on* $\mathscr{S}_n(\mathbb{R})$. *It is the support function of the set*

$$\tilde{\Omega}_1 := \{A \succeq 0 : \text{Trace } A \geqslant 1, \|A\| \leqslant 1\}\,. \tag{12}$$

When $n = 2$, we can somehow visualize in Figure 3 the three-dimensional euclidean space $(\mathscr{S}_2(\mathbb{R}), \langle\!\langle \cdot, \cdot \rangle\!\rangle)$ and the various convex sets therein:

- $\mathscr{K}_1$ and $\mathscr{K}_1^\circ$ are 'smooth' convex cones (they do not contain faces of dimension 2);
- Ω_1 is a 'pancake' located in the affine plane of equation Trace$(\cdot) = 1$;
- $\tilde{\Omega}_1$ is a sort of 'lunula' containing Ω_1.

Moreover, due to the inequality

$$\lambda_{\max} \leqslant \Delta_{\mathscr{K}_1} \leqslant \sqrt{n}\lambda^{+}_{\max}\,, \tag{13}$$

we have

$$\Omega_1 \subset \tilde{\Omega}_1 \subset \sqrt{2}\,\text{conv}(\{0\} \cup \Omega_1)\,. \tag{14}$$

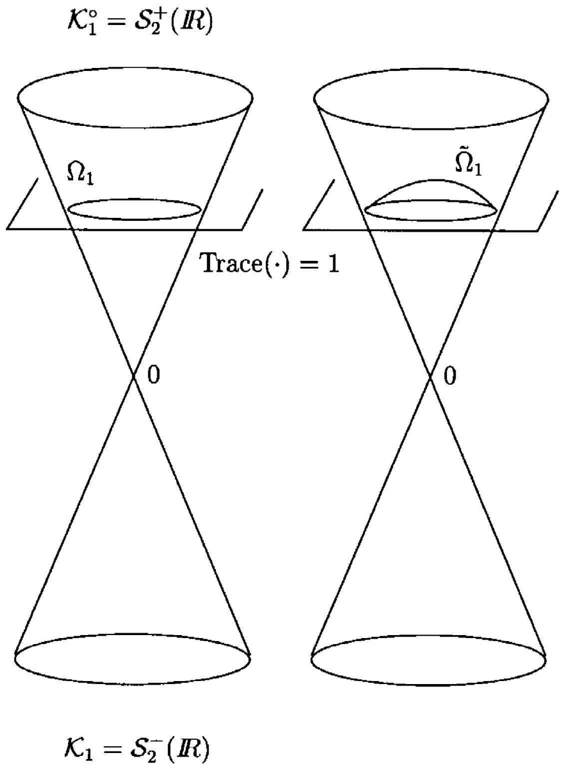

Figure 3.

2.2. CONSTRAINT INEQUALITIES INVOLVING THE BARRIER FUNCTION log(det)

Let us consider $f\colon \mathcal{S}_n^+(\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$ defined as follows

$$f(X) = \begin{cases} \log(\det(X^{-1})) & \text{if } X \succ 0\,, \\ +\infty & \text{otherwise}\,. \end{cases}$$

The function f is known to be proper, convex, and lower semicontinuous (see [12, A.4.3]) with

$$\partial f(X) = \{-X^{-1}\} \quad \text{whenever } X \succ 0\,.$$

THEOREM 2.4. *We have*

$$d(X, \{M \succeq 0 : \det(M) \geqslant 1\}) \leqslant \frac{1}{\sqrt{n}}(\log(\det(X^{-1})))^+ \quad \textit{for all } X \succ 0\,, \tag{15}$$

and one cannot do better than $1/\sqrt{n}$ *in (15).*

Proof. Observe that

$$[f \leqslant 0] = \{M \succeq 0 : \det(M) \geqslant 1\}\,,$$

$$[f > 0] \cap \operatorname{dom} f = \{M \succ 0 : \det(M) \leqslant 1\}\,.$$

According to Theorem 1.1, given $X \succ 0$,

$$d(X, \{M \succeq 0 : \det(M) \geqslant 1\}) \leqslant \frac{1}{\tau}(\log(\det(X^{-1})))^+\,,$$

where

$$\tau = \inf_{\{M \succ 0 : \det(M) \leqslant 1\}} \|M^{-1}\|\,.$$

Let us prove that $\tau = \sqrt{n}$. This amounts to

$$\left\{ \min_{\substack{\lambda_i > 0 \\ \Pi_{i=1}^n \lambda_i \leqslant 1}} \sum_{i=1}^n \frac{1}{\lambda_i^2} = n\,, \right.$$

or, equivalently by taking logarithms,

$$\min_{\Sigma_{i=1}^n x_i \leqslant 0} \sum_{i=1}^n \exp(-2x_i) = n\,.$$

This is easy to derive by using Karush–Kuhn–Tucker optimality conditions. □

3. Hoffman-type estimates in semidefinite inequality constraint

Let $B, A_1, \dots, A_m \in \mathscr{S}_n(\mathbb{R})$ and let $\mathscr{A}: \mathbb{R}^m \to \mathscr{S}_n(\mathbb{R})$ be defined by

$$\mathscr{A}(x) = \sum_{j=1}^{m} x_j A_j \quad \text{for all } x \in \mathbb{R}^m .$$

Let us set

$$\mathscr{F}_{A,B} = \{x \in \mathbb{R}^m : \mathscr{A}(x) - B \preceq 0\} .$$

In other words

$$\mathscr{F}_{A,B} = [f \leqslant 0] ,$$

where $f: \mathbb{R}^m \to \mathbb{R}$ is defined by

$$f(x) = \lambda_{\max}(\mathscr{A}(x) - B) = f_1(\mathscr{A}(x) - B) . \tag{16}$$

Observe that (see, e.g., [12]), for all $y \in \mathbb{R}^n$,

$$f^*(y) = \min_{\{Y \in \mathscr{S}_n(\mathbb{R}) : \mathscr{A}^T(Y) = y\}} f_1^*(Y) .$$

As $f_1 = \sigma_{\Omega_1}$ we get $f_1^* = i_{\Omega_1}$, where i_S stands for the indicator function of the set S. Thus we derive that

$$f^*(y) = \min_{\{Y \in \mathscr{S}_n(\mathbb{R}) : \mathscr{A}^T(Y) = y\}} i_{\Omega_1}(Y) ,$$

so that

$$\operatorname{dom} f^* = \mathscr{A}^T(\Omega_1) . \tag{17}$$

It is proved in [9] that under the assumption

$$\text{there exists } x_0 \in \mathbb{R}^m \text{ such that } \mathscr{A}(x_0) \prec 0 , \tag{18}$$

then $\mathscr{F}_{A,B} \neq \emptyset$ for all $B \in \mathscr{S}_n(\mathbb{R})$, and there exists a positive real number $c > 0$ such that, for all $B \in \mathscr{S}_n(\mathbb{R})$,

$$d(x, \mathscr{F}_{A,B}) \leqslant c(\lambda_{\max}(\mathscr{A}(x) - B))^+ \quad \text{for all } x \in \mathbb{R}^m . \tag{19}$$

It is natural to ask whether or not the estimate (19) holds true when assumption (18) is replaced by a weaker one:

$$\text{there exists } x_0 \in \mathbb{R}^m \text{ such that } \mathscr{A}(x_0) \prec B . \tag{20}$$

Relying on [23], this is the case if moreover $\mathscr{F}_{A,B}$ is bounded, since (20) is equivalent to the Slater qualification condition for f. This is also the case in a case where $\mathscr{F}_{A,B}$ is not necessarily bounded and where (18) is not in force, as shown by the following theorem. Before stating it, we need to recall the notion of 'good asymptotic behaviour' of a convex function, introduced by Auslender and Crouzeix in [1]. Following these authors, we say that the closed convex function $f: X \to \mathbb{R} \cup$

$\{+\infty\}$ has a good asymptotic behaviour whenever $(f(x_i))_{i\in\mathbb{N}}$ converges to $\inf_X f$ for any sequence $(x_i)_{i\in\mathbb{N}} \subset X$ such that $d_*(0, \partial f(x_i))$ goes to 0.

THEOREM 3.1. *Assume that (20) holds true along with*

$$\lambda_{\max}(\mathscr{A}(x)) \leqslant 0 \text{ implies } \mathscr{A}(x) = 0\,. \tag{21}$$

Then there exists a positive real number $c > 0$ such that

$$d(x, \mathscr{F}_{A,B}) \leqslant c(\lambda_{\max}(\mathscr{A}(x) - B))^+ \quad \text{for all } x \in \mathbb{R}^m\,.$$

Proof. We claim that $0 \in \operatorname{int}(\Omega_1 + \ker(\mathscr{A}^T))$. If not, there exists $Y \in \mathscr{S}_n(\mathbb{R})\setminus\{0\}$ such that

$$\langle\langle Y, Z\rangle\rangle + \langle\langle Y, \mathscr{A}^T(X)\rangle\rangle \leqslant 0\,,$$

for all $Z \in \Omega_1$ and $X \in \mathscr{S}_n(\mathbb{R})$, leading to $Y \in \mathscr{S}_n^-(\mathbb{R}) \cap R(\mathscr{A})$ and $Y \neq 0$, which contradicts assumption (21). Thus there exists $\eta > 0$ such that

$$\eta B \subset \Omega_1 + \ker(\mathscr{A}^T)\,. \tag{22}$$

Now, setting $f(x) = \lambda_{\max}(\mathscr{A}(x) - B)$, we have $\operatorname{dom} f^* = \mathscr{A}^T(\Omega_1)$, thus $0 \in \operatorname{dom} f^*$ since

$$\sigma_{\mathscr{A}^T(\Omega_1)}(x) = \sigma_{\Omega_1}(\mathscr{A}(x)) = \lambda_{\max}(\mathscr{A}(x)) \geqslant 0$$

for all $x \in \mathbb{R}^m$. It follows that

$$\operatorname{Aff}(\operatorname{dom} f^*) = \mathbb{R}_+(\operatorname{dom} f^*) - \mathbb{R}_+(\operatorname{dom} f^*) = R(\mathscr{A}^T)\,,$$

since $\mathbb{R}_+(\operatorname{dom} f^*) = \mathscr{A}^T(\mathbb{R}_+\Omega_1) = \mathscr{A}^T(\mathscr{S}_n^+(\mathbb{R}))$. Now there exists $\delta > 0$ such that $R(\mathscr{A}^T) \cap \delta B \subset \mathscr{A}^T(\eta B)$. Relying on (22), we derive that

$$R(\mathscr{A}^T) \cap \delta B \subset \mathscr{A}^T(\Omega_1) = \operatorname{dom} f^*\,,$$

thus $0 \in \operatorname{ri}(\operatorname{dom} f^*)$ and then f has a 'good assymptotic behaviour' (see [3]). Now we claim that $\inf_{x\in[f>0]} d(0, \partial f(x)) > 0$. If not, there would exist sequences $(x_i)_{i\in\mathbb{N}} \subset \mathbb{R}^m$ and $(y_i)_{i\in\mathbb{N}} \subset \mathbb{R}^m$ such that $y_i \in \partial f(x_i)$, $f(x_i) > 0$ and $\lim_{i\to\infty} y_i = 0$. As f has a good asymptotic behaviour, this would lead to $\inf_{x\in\mathbb{R}^m} f(x) \geqslant 0$, contradicting (20). Thus, we can apply Theorem 1.1 yielding the conclusion of the theorem □

We mention that Burke and Tseng derived in [7] Hoffman's type estimates in a general abstract setting. Nevertheless, the general results obtained by these authors, such as their Theorem 6 do not apply to our setting since $\mathscr{S}_n^+(\mathbb{R})$ is not polyhedral and $\mathscr{A}(\mathbb{R}^m) + \mathscr{S}_n^+(\mathbb{R})$ is not equal, in general, to $\mathscr{S}_n(\mathbb{R})$.

REMARK 3.1. Assumption (21) in Theorem 3.1 is equivalent to the existence of $Y \prec 0$ such that $\mathscr{A}^T Y = 0$. Indeed, assuming that this property fails, we have

$\ker \mathscr{A}^T \cap \operatorname{int}(\mathscr{S}_n^+(\mathbb{R})) = \emptyset$. Thus, from the Hahn–Banach Theorem, there exists $Y \in \mathscr{S}_n(\mathbb{R}) \backslash \{0\}$ such that

$$\sup_{Z \in \mathscr{S}_n^+(\mathbb{R})} \langle\langle Y, Z \rangle\rangle \leqslant 0 \leqslant \inf_{X \in \ker \mathscr{A}^T} \langle\langle Y, X \rangle\rangle .$$

It follows that $Y \neq 0$ satisfies $Y \in \mathscr{A}(\mathbb{R}^m)$ and $Y \in \mathscr{S}_n^-(\mathbb{R})$, thus (21) fails too. Conversely, assuming that there exists $Y \prec 0$ such that $\mathscr{A}^T Y = 0$, we derive that $\mathscr{A}^T(\mathscr{S}_n^+(\mathbb{R})) = \mathscr{A}^T(\mathscr{S}_n(\mathbb{R}))$, so that

$$(\mathscr{A}^T(\mathscr{S}_n^+(\mathbb{R})))^- = (\mathscr{A}^T(\mathscr{S}_n(\mathbb{R})))^- ,$$

yielding $\mathscr{F}_{\mathscr{A},0} = \ker \mathscr{A}$, that is (21).

EXAMPLE 3.1. Let $H \in S_n(\mathbb{R})$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}$, and let q be the convex quadratic function defined by $q(x) = x^T Hx/2 - c^T x - d$. Let A be a $n \times n$ matrix such that $H = 2A^T A$. It is easily seen that $[q \leqslant 0]$ is the set of those $x \in \mathbb{R}^n$ such that the matrix

$$\begin{pmatrix} I_n & Ax \\ x^T A^T & c^T x + d \end{pmatrix}$$

is positive semidefinite, or equivalently $\mathscr{A}(x) + B \succeq 0$ where

$$\mathscr{A}(x) = \begin{pmatrix} 0_n & Ax \\ x^T A^T & c^T x \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} I_n & 0 \\ 0 & d \end{pmatrix} .$$

Observing that

$$(z^T, t)\mathscr{A}(x)\begin{pmatrix} z \\ t \end{pmatrix} = 2tz^T Ax + t^2 c^T x ,$$

we derive that $\mathscr{A}(x) \succeq 0$ if and only if $Ax = 0$ and $c^T x \geqslant 0$. Assuming that $c \in R(H)$, we have $c^T x = 0$ whenever $Ax = 0$, thus $\mathscr{A}(x) \succeq 0$ if and only if $\mathscr{A}(x) = 0$, so that assumption (21) is satisfied. On the other hand, we have

$$(z^T, t)(\mathscr{A}(x) + B)\begin{pmatrix} z \\ t \end{pmatrix} = \|z\|^2 + 2tz^T Ax + t^2 c^T x + t^2 d ,$$

and

$$(z^T, t)(\mathscr{A}(x) + B)\begin{pmatrix} z \\ t \end{pmatrix} + q(x)(\|z\|^2 + t^2) = \|z + tAx\|^2 + \|z\|^2 q(x) ,$$

thus we have

$$\lambda_{\min}(\mathscr{A}(x) + B) \geqslant -q(x) \text{ whenever } q(x) \geqslant 0 .$$

Assuming that there exists $x_0 \in \mathbb{R}^n$ such that $q(x_0) < 0$, then Theorem 3.1 applies.

References

[1] Auslender, A. and Crouzeix, J.-P. (1989), Well Behaved Asymptotical Convex Functions, *Ann. Inst. Henri Poincaré, Anal. Non Lineáire* 6, 101–121.
[2] Auslender, A. and Crouzeix, J.-P. (1988), Global Regularity Theorems, *Math. Oper. Res.* 13, 242 − 253.
[3] Auslender, A., Cominetti, R. and Crouzeix, J.-P. (1993), Convex Functions with Unbounded Level Sets, *SIAM J. Optimization* 3, 669–687.
[4] Azé, D. and Corvellec, J.-N. (2001), On the Sensitivity Analysis of Hoffman Constants for Systems of Linear Inequalities, *SIAM J. Optimization* (in press).
[5] Azé, D., Corvellec, J.-N. and Lucchetti, R.E. (2001), Variational Pairs and Applications to Stability in Nonsmooth Analysis, *Nonlinear Anal., Theory Methods Appl.* (in press).
[6] Azé, D. (2001), Characterizations of the Existence of a Global Error Bound for Convex Inequalities in General Banach Spaces, submitted.
[7] Burke, J.W., Tseng, P. (1996), A Unified Analysis of Hoffman's Bound via Fenchel Duality, *SIAM J. Optimization* 6, 265–282.
[8] Deng, S. (1997), Computable Error Bound for Convex Inequality Systems in Reflexive Banach Spaces, *SIAM J. Optimization* 7, 274–279.
[9] Deng, S. and Hu, H. (1999), Computable Error Bounds for Semidefinite Programming, *Journal of Global Optimization* 14, 105–115.
[10] Ekeland, I. (1976), On the Variational Principle, *Journal of Mathematical Analysis and Applications* 47, 324–353.
[11] Higham, I. (1989), Matrix Nearness Problems and Applications, in M.J.C. Glover and S. Barnett (eds.), Application of Matrix Theory, Oxford University Press, pp. 1–27.
[12] Hiriart-Urruty, J.-B. and Lemaréchal, C. (1993), Convex Analysis and Minimization Algorithms I and II, Grundlehren der Mathematischen Wissenschaften 305 & 306, Springer, Berlin.
[13] Hiriart-Urruty, J.-B. and Ye, D. (1995), Sensitivity Analysis of all Eigenvalues of a Symmetric Matrix, *Numerische Mathematik* 70, 45–72.
[14] Hiriart-Urruty, J.-B. (1979), Tangent Cones, Generalized Gradients and Mathematical Programming in Banach Spaces, *Math. Oper. Res.* 4, 79–97.
[15] Hiriart-Urruty, J.-B. (1979), New Concepts in Nondifferentiable Programming, Mémoire 60 du Bulletin de la Société Mathématique de France, 57–85.
[16] Hoffman, A.J. (1952), On Approximate Solutions of Systems of Linear Inequalities, *J. Res. Nat. Bur. Stand.* 49, 263–265.
[17] Jarre, F. (2000), Convex Analysis on Symmetric Matrices, in H. Wolkowicz, R. Saigal and L. Vandenberghe (eds.), *Handbook of Semidefinite programming. Theory, algorithms, and applications*, International Series in Operations Research & Management Science 27, Kluwer Academic Publishers, Dordrecht, pp. 13–27.
[18] Klatte, D. and Li, W. (1999), Asymptotic Constraint Qualifications and Global Error Bound for Convex Inequalities, *Math. Program.* 84, 137–160.
[19] Lewis, A.S. and Pang, J.-S. (1998), Error Bounds for Convex Inequality Systems, *Nonconvex Optim. Appl.* 14, 75–110.
[20] Li, W. and Singer, I. (1998), Global Error Bounds for Convex Multifunctions and Applications, *Math. Oper. Res.* 23, 443–462.
[21] Mangasarian, O.L. (1998), Error Bounds for Nondifferentiable Conex Inequalities Under a Strong Slater Constraint Qualification, *Math. Programming* 83, 187–194
[22] Overton, M.L. and Womersley, R.S. (1993), Optimality Conditions and Duality Theory for Minimizing Sums of the Largest Eigenvalue of Symmetric Matrices, *Math. Programming* 62, 321–357.

[23] Robinson, S.M. (1975), An Application of Error Bounds for Convex Progamming in a Linear Space, *SIAM J. Control Optimization* 13, 271–273.
[24] Simmons, S. (1979), Subdifferential of Convex Functions, *Contemporary Mathematics* 204, 217–246.
[25] Sturm, J.F. (2000), Error Bounds for Linear Matrix Inequalities, *SIAM H. Optimization* 10, 1228–1248.
[26] 2000, *Handbook of Semidefinite Programming, Theory, Algorithms, and Applications*, in H. Wolkowicz, R. Saigal and L. Vandenbergh (eds.), *International Series in Operations Research & Management Science*, 27, Kluwer Academic Publishers, Dordrecht.