Matrix monotonicity and self-concordance: how to handle quantum entropy in optimization problems

Leonid Faybusovich [*] Takashi Tsuchiya[†]

August 2014

Abstract

Let g be a continuously differentiable function whose derivative is matrix monotone on positive semi-axis. Such a function induces a function $\phi(x) = tr(g(x))$ on the cone of squares of an arbitrary Euclidean Jordan algebra. We show that $\phi(x) - \ln\det(x)$ is a self-concordant function on the interior of the cone. We also show that $-\ln(t-\phi(x))-\ln\det(x)$ is $\sqrt{\frac{5}{3}}(r+1)$-self-concordant barrier on the epigraph of ϕ, where r is the rank of the Jordan algebra. The case $\phi(x) = tr(x\ln x)$ is discussed in detail.

Key words: quantum entropy, matrix monotonicity, self-concordance

1 Introduction

A substantial amount of optimization problems arising in quantum information theory,quantum statistical physics ,information geometry ([9]) and other areas (for the impressive list of various applications see [2])requires dealing with the so-called quantum or von Neumann entropy which is the function of the form $Tr(XLnX)$, where X is positive semi-definite complex Hermitian matrix. In present paper we propose a self-concordant barrier for this function which, in principle, allows one to include the above mentioned problems into general interior point polynomial framework. We develop the corresponding theory within the formalism of Euclidean Jordan algebras which indicates possible generalizations of this concept. Moreover, proposed approach allows us to consider relatively broad class of functions, namely, primitives of matrix monotone functions on positive semi-axis. It turns out that the standard *lndet* barrier is compatible (in the sense of the theory of self-concordant functions) with functions from this class. The plan of the paper is as follows. In section 2 we briefly describe some Jordan-algebraic concepts. In section 3 we introduce a class functions on the cone of squares of an Euclidean Jordan algebra associated with matrix monotone functions on positive semi-axis. We then show that the standard *lndet* self-concordant barrier is compatible with every

[*]University of Notre Dame, Department of Mathematics (Email: `leonid.faybusovich.1@nd.edu`) This author is supported in part by Simmons Foundation Grant 275013.

[†]National Graduate Institute for Policy Studies (Email: `tsuchiya@grips.ac.jp`). This author is supported in part with Grant-in-Aid for Scientific Research (B), 24310112 from the Japan Society for the Promotion of Sciences.

function in this class. The self-concordance of associated barriers (along with barrier parameter estimate) on the epigraphs of functions of this class is also shown.

2 Jordan-algebraic Concepts

We adhere to the notation of an excellent book [4].

Let **F** be the field **R** or **C**. A vector space V over **F** is called an algebra over **F** if a bilinear mapping $(x, y) \to xy$ from $V \times V$ into V is defined. For an element x in V let $L(x) : V \to V$ be the linear map such that

$$L(x)y = xy.$$

An algebra V over **F** is a Jordan algebra if

$$xy = yx, x(x^2 y) = x^2(xy), \forall x, y \in V.$$

In other words, Jordan algebra is always commutative but typically not associative. In an algebra V one defines x^n recursively by $x^n = x \cdot x^{n-1}$. An algebra V is said to be power assotiative if $x^p \cdot x^q = x^{p+q}$ for any $x \in V$ and integers p, q.

Proposition 2.1 *A Jordan algebra is power associative. Besides,*

$$[L(x^p), L(x^q)] = 0, \forall x \in V,$$

and any positive integers p and q. (In other words, corresponding linear operators commute).

This is Proposition II.1.2 in [4]. We will always assume that the Jordan algebra has an identity element e (i.e. , $xe = x, \forall x \in V$).

Let V be a finite-dimensional power associative algebra over **F** with an identity element e, and let $F[X]$ denote the algebra over **F** of polynomials in one variable with coefficients in **F**. For $x \in V$ we define

$$\mathbf{F}[x] = \{p(x) : p \in \mathbf{F}[X]\}.$$

A polynomial $p \in \mathbf{F}[X]$ of minimal possible degree such that $p(x) = 0$ is called the minimal polynomial of x. Given $x \in V$, let $m(x)$ be the degree of the minimal polynomial of x. We define the rank of V as

$$r = \max\{m(x) : x \in V\}.$$

An element x is called regular if $m(x) = r$.

Proposition 2.2 *The set of regulr elements is open and dense in V. There exist polynomials $a_1, \ldots a_r$ on V such that the minimal polynomial of every regular element x is given by*

$$f(\lambda; x) = \lambda^r - a_1(x)\lambda^{r-1} + a_2(x)\lambda^{r-2} + \ldots + (-1)^r a_r(x).$$

The polynomials $a_1, \ldots a_r$ are unique and a_j is homogeneous of degree j.

This is Proposition II.2.1 in [4]. The coefficient $a_1(x)$ is called the trace of x and is denoted $tr(x)$ (in particular, trace is linear). The coefficient $a_r(x)$ is called the determinant of x and is denoted $\det(x)$. An element x is said to be invertible if there exists an element $y \in \mathbf{F}[x]$ such that $xy = e$. The set $\lambda \in \mathbf{F}$ such that $x - \lambda e$ is not invertible is called the spectrum of x and is denoted $spec(x)$.

Given $x \in V$, we define

$$P(x) = 2L(x)^2 - L(x^2).$$

The map P is called the quadratic representation of V. We denote $DP(x)y$ by $2P(x,y)$. Here $DP(x)y$ is the Frechet derivative of the map P at point $x \in V$ evaluated on $y \in V$. It is easy to see that

$$P(x,y) = L(x)L(y) + L(y)L(x) - L(xy), x, y \in V.$$

Proposition 2.3 *Let V be a finite-dimensional Jordan algebra over $\mathbf{F}$. An element $x \in V$ is invertible if and only if $P(x)$ is invertible. In this case*

$$P(x)x^{-1} = x, P(x)^{-1} = P(x^{-1}).$$

This is Proposition II.3.1 in [4].

Proposition 2.4 *Let $\mathcal{J}$ be the (open) set of invertible elements in V. The map $x \to x^{-1} : \mathcal{J} \to \mathcal{J}$ is Frechet differentiable and*

i)$D(x^{-1})u = -P(x^{-1})u, x \in \mathcal{J}, u \in V$.

ii) If x and y are invertible, then $P(x)y$ is invertible and $(P(x)y)^{-1} = P(x^{-1})y^{-1}$.

iii)

$$P(P(x)y) = P(x)P(y)P(x), \forall x, y \in V.$$

iv)

$$P(P(x)y, P(x)z) = P(x)P(y,z)P(x), \forall x, y, z \in V.$$

This is Proposition II.3.3 in [4]. A bilinear form β on V is called associative if

$$\beta(xy, z) = \beta(x, yz), \forall x, y, z \in V.$$

Proposition 2.5 *The symmetric bilinear forms $TrL(xy)$ and $tr(xy)$ are associative.*

This is Proposition II.4.3 in [4].

In case , where $\mathbf{F} = \mathbf{R}$ we consider an important class of Euclidean Jordan algebras. A Jordan algebra V over $\mathbf{R}$ is called Euclidean if $tr(x^2) > 0, \forall x \in V \setminus \{0\}$. An element $c \in V$ is called idempotent if $c^2 = c$. Two idempotents are orthogonal if $cd = 0$. A system of idempotents $c_1, \dots c_k$ is a complete system of orthogonal idempotents If $c_i^2 = c_i, c_i c_j = 0, i \neq j$, and $c_1 + \ldots + c_k = e$.

Theorem 2.6 *Let V be an Euclidean Jordan algebra. Given $x \in V$,there exist unique real numbers $\lambda_1, \ldots \lambda_k$, all distinct, and a unique complete system of orthogonal idempotents $c_1, \ldots, c_k$ such that*

$$x = \lambda_1 c_1 + \ldots + \lambda_k c_k.$$

In this case $spec(x) = \{\lambda_1, \ldots, \lambda_k\}, c_1, \ldots, c_k \in \mathbf{R}[x]$.

This is Theorem III.1.1 in [4].

An idempotent is primitive if it is non-zero and cannot be written as a sum of two non-zero idempotents. We say that $c_1, \ldots, c_m$ is a complete system of orthogonal primitive idempotents, or Jordan frame, if each c_j is primitive idempotent and if

$$c_j c_k = 0, j \neq k, c_1 + \ldots + c_m = e.$$

Note that in this case $m = r$ (rank of V).

Theorem 2.7 *Suppose V has rank r. Then for $x \in V$ there exists a Jordan frame $c_1, \ldots c_r$ and real numbers $\lambda_1, \ldots \lambda_r$ such that*

$$x = \sum_{j=1}^{r} \lambda_j c_j.$$

The numbers λ_j (with multiplicities) are uniquely determined by x. Furthermore,

$$det(x) = \prod_{j=1}^{r} \lambda_j, tr(x) = \sum_{j=1}^{r} \lambda_j.$$

This is Theorem III.1.2 in [4].

Given a function f which is defined at least on $spec(x)$, we can define

$$f(x) = \sum_{i=1}^{r} f(\lambda_i) c_i,$$

if $x = \sum_{i=1}^{r} \lambda_i c_i$. In particular,

$$\exp(x) = \sum_{i=1}^{r} \exp(\lambda_i) c_i, \ln x = \sum_{i=1}^{r} \ln \lambda_i c_i, \lambda_i > 0.$$

Convexity and differentiability ofsuch functions on Euclidean Jordan algebras have been studied in [1],[12] (see also [7]). We extensively use these properties in the paper.

Let

$$Q = \{x^2 : x \in V\}.$$

Theorem 2.8 *Let V be an Euclidean Jordan algebra.The interior Ω of Q is a symmetric (i.e. , self-dual, homogeneous) convex cone. Furthermore,Ω is the connected component of e in the set $\mathcal{J}$ of invertible elements, and also Ω is the set of elements x in V for which $L(x)$ is positive definite. In particular, the group of linear automorphisms $GL(\Omega)$ of Ω acts transitively on it. Moreover, $P(x) \in GL(\Omega)$ for any invertible x.*

This is Proposition III.2.2 in [4].

Let $c_1, \dots c_k$ be complete system of orthogonal idempotents. For each idempotent c, denote $V(c,0), V(c,1), V(c,1/2)$ the eigenspaces of $L(c)$ corresponding to eigenvalues $0, 1, 1/2$, respectively. Then $L(c_1), \dots, L(c_k)$ pairwise commute and

$$V = \bigoplus_{1 \leq i \leq j} V_{ij},$$

where $V_{ii} = V(c_i, 1), V_{ij} = V(c_i, 1/2) \cap V(c_j, 1/2)$. Such a decomposition of V corresponding to a complete system of orthogonal idempotents is called the Peirce decomposition. It is studied in detail in Section 1 of Chapter IV in [4]. A typical example of a Jordan algebra over a field $\mathbf{F}$ is the vector space of symmetric matrices over $\mathbf{F}$ with multiplication operation

$$A \cdot B = \frac{AB + BA}{2},$$

where on the right we have a usual matrix multiplication. In case $\mathbf{F} = \mathbf{R}$ we get an exmple of an Euclidean Jordan algebra.

3 Main results

For real symmetric matrices A, B of the same size $n \times n$ we use notation $A \succeq B$ if

$$x^T A x \geq x^T B x, \quad \forall x \in \mathbf{R}^\mathbf{n}.$$

A function $f : [0, \infty) \to \mathbf{R}$ is said to be matrix monotone if $A \succeq B$ implies $f(A) \succeq f(B)$ provided $A \succeq 0, B \succeq 0$. The following theorem is a version of the classical Löwner's result.

Theorem 3.1 *Let f be continuous matrix monotone function on $[0, \infty)$. Then there exists a positive measure μ on $(0, \infty)$ and $\beta \geq 0$ such that*

$$f(\lambda) = f(0) + \beta\lambda + \int_0^{+\infty} \frac{\tau\lambda}{\lambda + \tau} d\mu(\tau) \tag{1}$$

and

$$\int_0^{+\infty} \frac{\tau}{1 + \tau} d\mu(\tau) < +\infty.$$

(see, e.g., Theorem 4.41 in [8]). Let $g : [0, \infty) \to \mathbf{R}$ be a continuously differentiable function such that

$$f(\lambda) = g'(\lambda)$$

is matrix monotone on $[0, \infty)$. Given an Euclidean Jordan algebra V, consider the function

$$\phi : \bar{\Omega} \to \mathbf{R}, \phi(x) = \mathrm{tr}(g(x)).$$

The standard self-concordant barrier function $\mu : \Omega \to \mathbf{R}$ has the following form ([5]):

$$\mu(x) = -\text{lndet}x. \tag{2}$$

Our first result establish compatibility ϕ with μ in the sense of ([10], p.66).

Theorem 3.2 *Given $x \in \Omega, h \in V$:*

$$|D^3(\phi(x)(h,h,h) \leq 2D^2\phi(h,h)[D^2\mu(x)(h,h)]^{1/2} \tag{3}$$

Here $D\phi$, $D^2\phi$, $D^3\phi$ are standard Fŕechet derivtives of function ϕ.

We need two elementary lemmas.

Lemma 3.3 *Given $h \in V$, we have*

$$|\text{tr}(h^3) \leq [\text{tr}(h^2)]^{3/2} \tag{4}$$

Proof. We have:

$$\text{tr}(h^3) = \langle h^2, h\rangle$$

Consequently,

$$|\text{tr}(h^3) \leq \sqrt{\langle h^2, h^2\rangle}\sqrt{\langle h, h\rangle}.$$

Let

$$h = \sum_{i=1}^{r} \lambda_i c_i$$

be the spectral decomposition of h. Then

$$\langle h^2, h^2\rangle = \text{tr}(h^4) = \sum_{i=1}^{r} \lambda_i^4, \quad \langle h, h\rangle = \sum_{i=1}^{r} \lambda_i^2.$$

Since

$$\sum_{i=1}^{r} \lambda_i^4 \leq (\sum_{i=1}^{r} \lambda_i^2)^2,$$

The result follows. ∎

Lemma 3.4 *Consider the function*

$$\psi := \Omega \to V, \quad \psi(x) = P(x + \tau e)^{-1}h.$$

Here $\tau \geq 0, h \in V$ are fixed. Then,

$$D\psi(x)h = -2P(x + \tau e)^{-1/2}[P(x + \tau e)^{-1/2}h^2].$$

Proof. Consider $\psi_1 : V \to V$, $\psi_1(x) = P(x)h;$, $\psi_2 : \Omega \to \Omega$, $\psi_2(x) = (x + \tau e)^{-1}$. Then, $\psi(x) = \psi_1(\psi_2(x))$. Using chain rule, we obtain

$$D\psi(x)h = D\psi_1(\psi_2(x))\dot{(}D\psi_2(x)h) = 2P(\psi_2(x), D\psi_2(x)h)h =$$

$$-2P((x+\tau e)^{-1}, P(x+\tau e)^{-1}h)h = 2P(P(x+\tau e)^{-1/2}e, P(x+\tau e)^{-1}\alpha)h =$$

$$-2P(x+\tau e)^{-1/2}P(e,\alpha)P(x+\tau e)^{-1/2}h,$$

where $\alpha = P(x+\tau e)^{-1/2}h$.

Hence,

$$D\psi(x)h = -2P(x+\tau e)^{-1/2}L(\alpha)\alpha = -2P(x+\tau e)^{-1/2}(\alpha^2).$$

Here we used Proposition 2.4 i). ∎

Proof of Theorem 3.2. We have (see [1]):

$$D\varphi(x)h = \mathrm{tr}(f(x)h),$$

where $f(\lambda) = g'(\lambda)$. Using representation (1):

$$D\varphi(x)h = f(0)\mathrm{tr}(h) + \beta\mathrm{tr}(xh) + \int_0^\infty \mathrm{tr}(\tau h - \tau^2(x+\tau e)^{-1}h)d\mu(\tau).$$

Consequently,

$$D^2\psi(x)(h,h) = \beta\mathrm{tr}(h^2) + \int_0^\infty \mathrm{tr}((P(x+\tau e)^{-1/2}h)^2)\tau^2 d\mu(\tau) = \beta\mathrm{tr}(h^2) + \int_0^\infty \langle h, P(x+\tau e)^{-1}h\rangle \tau^2 d\mu(\tau). \tag{5}$$

By Lemma 3.4

$$D^3\varphi(x)(h,h,h) = -2\int_0^\infty \mathrm{tr}(\alpha(\tau)^3)\tau^2 d\mu(\tau), \tag{6}$$

where $\alpha(\tau) = P(x+\tau e)^{-1/2}h$.

Using Lemma 3.3:

$$|D^3\varphi(x)(h,h,h)| \le 2\int_0^\infty |\mathrm{tr}(\alpha(\tau)^3)|\tau^2 \mathrm{d}\mu(\tau) \le 2\int_0^\infty [\mathrm{tr}(\alpha(\tau)^2)]^{3/2}\tau^2 \mathrm{d}\mu(\tau). \tag{7}$$

Let

$$x = \sum_{i=1}^k \beta_i c_i$$

be the spectral decomposition of x, where $\beta_1 > \beta_2 > \cdots > \beta_k > 0$. Consider the corresponding Peirce decomposition of V:

$$V = \bigoplus_{i \le j} V_{ij}$$

and corresponding decomposition of h:

$$h = \sum_{i \le j} h_{ij}.$$

Then

$$\mathrm{tr}(\alpha(\tau)^2) = \langle h, P(x+\tau e)^{-1}h\rangle = \sum_{i\le j} \frac{1}{(\beta_i+\tau)(\beta_j+\tau)}\|h_{ij}\|^2 \le \sum_{i\le j}\frac{1}{\beta_i\beta_j}\|h_{ij}\|^2, \quad \forall \tau \ge 0.$$

But

$$\sum_{i\le j}\frac{1}{\beta_i\beta_j}\|h_{ij}\|^2 = \langle h, P(x)^{-1}h\rangle = \mathrm{tr}(\alpha(0)^2),$$

where $\alpha(0) = P(x)^{-1/2}h$. Note that

$$\begin{aligned} D\mu(x)h &= -\mathrm{tr}(x^{-1}h) &(8)\\ D^2\mu(x)(h,h) &= \langle P(x^{-1})h, h\rangle = \mathrm{tr}(\alpha(0)^2) &(9)\\ D^3\mu(x)(h,h,h) &= -2\mathrm{tr}(\alpha(0)^3),\ \mu(x) = -\mathrm{lndet}x. &(10)\end{aligned}$$

Consequently, using (7):

$$|\varphi(x)(h,h,h)| \le 2\sqrt{D^2\mu(x)(h,h)}\int_0^\infty \mathrm{tr}(\alpha(\tau)^2)\tau^2 d\mu(\tau) \le 2\sqrt{D^2\mu(x)(h,h)}D^2\varphi(x)(h,h).$$

The last inequality is due to (5).

Our next theorem shows that

$$\Gamma_1(x) = \varphi(x) + \mu(x)$$

is a standard self-concordant function on Ω.

Theorem 3.5 *Given* $x \in \Omega, h \in V$*:*

$$|D^3\Gamma_1(x)(h,h,h)| \le 2[D^2\Gamma_1(x)(h,h)]^{3/2}.$$

Proof. Let $a = [D^2\varphi(x)(h,h)]^{1/2}, b = [D^2\mu(x)(h,h)]^{1/2}$. Then,

$$D^2\Gamma_1(h,h) = a^2 + b^2$$

and by Theorem 3.2

$$|D^3\varphi(x)(h,h,h)| \le 2a^2 b.$$

Furthermore,

$$|D^3\mu(x)(h,h,h)| \le 2b^3.$$

Consequently,

$$|D^3\Gamma_1(x)(h,h,h)| \le 2a^2b + 2b^3.$$

However,

$$(2a^2b+2b^3)^2 \le 4(a^2+b^2)^3, \quad \forall a, b \ge 0.$$

The result follows. ∎

Theorem 3.6 *Let*

$$\begin{aligned}
&epi(\varphi) = \{(t,x) \in \mathbf{R} \times \Omega : t > \varphi(x)\},\\
&\Gamma_2 : ep(\varphi) \to \mathbf{R},\\
&\Gamma_2(t,x) = -\ln(t - \varphi(x)) + \mu(x).
\end{aligned}$$

Then,

$$|D^3\Gamma_2(t,x)(\tilde{h},\tilde{h},\tilde{h})| \leq \frac{10}{3}[D^2\Gamma_2(t,x)(\tilde{h},\tilde{h})]^{3/2},\ \tilde{x} = (t,x) \in epi(\varphi),\ \tilde{h} = (s,h) \in \mathbf{R} \times V.$$

Proof. The idea of this proof is taken from [3], Appendix A. Let $\chi(t,x) = t - \varphi(x)$. Then,

$$D\Gamma_2(\tilde{x})\tilde{h} = -\frac{D\chi(\tilde{x})\tilde{h}}{\chi(\tilde{x})} + D\mu(x)h, \tag{11}$$

$$D^2\Gamma_2(\tilde{x})(\tilde{h},\tilde{h}) = -\frac{D^2\chi(\tilde{x})(\tilde{h},\tilde{h})}{\chi(\tilde{x})} + \frac{[D\chi(\tilde{x})\tilde{h}]^2}{\chi^2(\tilde{x})} + D^2\mu(x)(h,h), \tag{12}$$

$$D^3\Gamma_2(\tilde{x})(\tilde{h},\tilde{h},\tilde{h}) = -\frac{D^3\chi(\tilde{x})(\tilde{h},\tilde{h},\tilde{h})}{\chi(\tilde{x})} + 3\frac{D^2\chi(\tilde{x})(\tilde{h},\tilde{h})(D\chi(\tilde{x})\tilde{h})}{\chi^2(\tilde{x})} - 2\frac{(D\chi(\tilde{x})\tilde{h})^3}{\chi(\tilde{x})^3} + D^3\mu(x)(h,h,h). \tag{13}$$

Since χ is concave on $epi(\varphi)$, all three terms in the right-hand side of (12) are nonnegative.

Let

$$a = \left[-\frac{D^2\chi(\tilde{x})(\tilde{h},\tilde{h})}{\chi(\tilde{x})}\right]^{1/2}, \quad b = \frac{|D\chi(\tilde{x})\tilde{h}|}{\chi(\tilde{x})}, \quad c = [D^2\mu(x)(h,h)]^{1/2}.$$

Then

$$D^2\Gamma_2(\tilde{x})(\tilde{h},\tilde{h}) = a^2 + b^2 + c^2.$$

Note that

$$D^2\chi(\tilde{x})(\tilde{h},\tilde{h}) = -D^2\varphi(x)(h,h), \quad D^3\chi(\tilde{x})(\tilde{h},\tilde{h},\tilde{h}) = -D^3\varphi(x)(h,h,h).$$

Hence, (13) and Theorem 3.2 implies:

$$|D^3\Gamma_2(\tilde{x})(\tilde{h},\tilde{h},\tilde{h})| \leq 2a^2c + 3a^2b + 2b^3 + 2c^3.$$

However,

$$(2a^2c + 3a^2b + 2b^3 + 2c^3)^2 \leq \frac{100}{9}(a^2 + b^2 + c^2)^3.$$

for any nonnegative a, b, c. This is verified by estimating odd degree powers using inequalities of the type $2ab \leq a^2 + b^2$. Hence,

$$|D^3\Gamma_2(\tilde{x})(\tilde{h},\tilde{h},\tilde{h})| \leq \frac{10}{3}[D^2\Gamma_2(\tilde{x})(\tilde{h},\tilde{h})]^{3/2}.$$

Recall ([11] p.193) that the barrier parameter for the self-concordant function Γ_2 is defined as

$$\nu = \sup_{\tilde{x} \in epi(\varphi)} \sup_{\tilde{h} \in \mathbf{R} \times V} [2D\Gamma_2(\tilde{x})\tilde{h} - D^2\Gamma_2(\tilde{x})(\tilde{h},\tilde{h})].$$

Proposition 3.7 *We have $\nu \leq r+1$, where $r = rank(V)$.*

Proof. We use notations of the proof of Theorem 3.6. Since χ is concave on $epi(\varphi)$,

$$2D\Gamma_2(\tilde{x})\tilde{h} - D^2\Gamma_2(\tilde{x})(\tilde{h},\tilde{h}) \leq$$
$$[-2\frac{D\chi(\tilde{x})\tilde{h}}{\chi(\tilde{x})} - (\frac{D\chi(\tilde{x})\tilde{h}}{\chi(\tilde{x})})^2] + [2D\mu(x)h - D^2\mu(x)(h,h))] \leq$$
$$1 - 2\langle x^{-1}, h\rangle - \langle P(x^{-1})h, h\rangle \leq r+1,$$

since

$$\max\{(-2\langle x^{-1}, h\rangle - \langle P(x^{-1})h, h\rangle) : h \in V\} = r.$$

■

Remark 3.8 *The barrier parameter usually defined for normalized barrier functions. Hence,our estimate shows that for normalized* Γ_2

$$\nu \leq \sqrt{\frac{5}{3}}(r+1).$$

■

The most natural examples satisfying our assumptions are

$$g_1(\lambda) = 2\ln\lambda,\ \lambda \geq 0, g_2(\lambda) = \lambda^p,\ 1 \leq p \leq 2.$$

Indeed,

$$g_1'(\lambda) = \ln\lambda + 1,\ g_2'(\lambda) = p\lambda^{p-1}.$$

Both functions $\ln\lambda$ and $\lambda^q, 0 \leq q \leq 1$ are known to be matrix monotone. The corresponding functions

$$\varphi_1(x) = \mathrm{tr}(x\ln x),\ \varphi_2(x) = \mathrm{tr}(x^p),\ 1 \leq p \leq 2$$

arise in various applications typically for the Euclidean Jordan algebra of complex Hermitian matrices; φ_1 is known as quantum or von Neumann Entropy and arise in numerous applications ([2]). Tha function φ_2 appears in compressed sensing (and is known as nuclear norm for $p = 1$).

One way to incorporate a new self-concordant barrier to general interior-point scheme is to modify it for epigraphs of corresponding functions. A general convex programming problem can be easily reformulated as a problem of minimization of a linear function over intersection of epigraphs of constraints and objective function. A clear and straightforward path-following scheme for such a reformulation is presented in [11], section 4.2.6. The resulting algorithm is polynomial provided computable self-concordant barriers for the corresponding epigraphs are known. One advantage of this approach is that epigraph version of a self-concordant barrier frequently diminishes the value of the barrier parameter and consequently improves the theoretical complexity of the algorithm. A substantial factor in the overall complexity of the scheme is the computational cost for the evaluation of Hessians of barrier functions and their inverses. In case of $\phi(x) = x \ln x$ the Hessian and its inverse can be expressed as:

$$H_{\phi}(x) = \int_0^{+\infty} P(x + e\tau)^{-1} d\tau,$$

$$H_{\phi}^{-1} = \int_0^1 P(x^{\tau}, x^{1-\tau}) d\tau,$$

as was shown in [6].

■

References

[1] M. Baes:Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras. Linear Algebra Appl., 422, no 2-3,664-700,2007.

[2] V. Chandrasecaran and P. Shah: Conic geometric programming. Manuscript, arXiv:math/1310.099v2, October,2013.

[3] D. den Hertog:*Interior point approach to linear, quadratic and convex programming: algorithms and complexity*. Kluwer,Boston,1994.

[4] J. Faraut and A. Korányi: *Analysis on Symmetric Cones*, Cambridge University Press, Cambridge, UK, 1994.

[5] L. Faybusovich: Jordan algebras and interior-point algorithms. *Positivity*, Vol. 1 (1997), pp. 331–357.

[6] L. Faybusovich: E.Lieb convexity inequalities and noncommutative Bernstein inequality in Jordan-algebraic setting. Preprint, University of Notre Dame,August 2014.

[7] L.Faybusovich:Several Jordan-algebraic aspects of optimization. Optimization, 57,no 3,379-393,2008.

[8] F. Hiai and D.Petz: *Introduction to matrix analysis and applications*. Springer,New York,2014.

[9] S. Kakihara, A. Ohara and T. Tsuchiya: Information geometry and interior-point algorithms in semidefinite programs and symmetric cone programs. J.Optim. Theory Appl. 157, issue 3, 749-780, 2013.

[10] Yu. Nesterov and A. Nemirovskii: *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, 1994.

[11] Yu. Nesterov : *Introductory lectures on convex optimization* Kluwer,Boston,2004.

[12] D.Sun and J. Sun : Lowner's operator and spectral functions in Euclidean Jordan algebras. Math. Oper. Res. 33, no 2, 421-445,2008.