Power classes of search directions and dualistic structure on symmetric cones

Keiko Uohashi

Abstract. Our original results refer to dualistic structure on primal-dual interior-point methods for symmetric cone programs with linear constraints. It is shown that scalings by the Nesterov-Todd direction are generated by middle points of geodesics joining with primal interior points and dual interior points. Finally we relate power classes of search directions with geodesics and weighted geometric means.

M.S.C. 2010: 17C36, 53A15, 90C51.
Key words: Symmetric cone; Jordan algebra; interior-point method; power class; information geometry; geodesic; mean.

1 Introduction

Commutative classes of search directions for linear programming over symmetric cones were proposed. For instance MZ* family, which is a subfamily of MZ (Monteiro-Zhang) family, for semi-definite programming was extended for linear programming over symmetric cones. In addition the HRVW/KSH/M (Helmberg-Rendl-Vanderbei-Wolkowicz/Kojima-Shindoh-Hara/Monteiro) direction and the NT (Nesterov-Todd) direction for linear programming were extended over symmetric cones by Muramatsu, who defined power classes including these directions [5].

Monotone operators were made use for convex programming by Nesterov and Nemirovskii [6]. Sturm discussed relations between the NT direction and geometric means on symmetric cones [13].

Tools from Riemannian geometry (suitable Riemannian metrics, Riemann-Newton method, Riemann interior point method, exponential map, search along geodesics, covariant differentiation, sectional curvature etc) are now intensively used in Mathematical Programming to obtain deeply theoretical results and practical algorithms [15]-[21].

On these backgrounds, we discuss information geometric (i.e. dualistic structural) properties of power classes of search directions. First we describe linear programming over symmetric cones. In section 3 we talk about dualistic structure of primal-dual

Balkan Journal of Geometry and Its Applications, Vol. 18, No. 2, 2013, pp. 83-89.

interior-point methods for linear programming over symmetric cones. In section 4 we mention that the middle point of the geodesic with respect to the Levi-Civita connection defines the NT direction of primal-dual interior-point methods for symmetric cones. In section 5 we show that the $1:q$ internally dividing point on a geodesics (i.e. weighted geometric means) defines a q power class of search directions of primal-dual interior-point methods for symmetric cones.

2 Linear programming over symmetric cones

In this section we note primal-dual interior-point methods for linear programming over symmetric cones in terms of Jordan algebras [3] [5]. See [2] on Jordan algebras and symmetric cones.

Let V be an Euclidean Jordan algebra and Ω a symmetric cone associated with V. Let X be a vector subspace of V and $X^\perp$ its orthogonal complement relative to an inner product $\langle x, y\rangle = \mathrm{tr}(x \circ y)$. Consider the following optimization problems with linear constraints on symmetric cones; Given $a \in X, b \in X^\perp$,

$$\text{primal problem (P)} : \langle a, x\rangle \to \min, \quad \text{s.t. } x \in (b + X) \cap \bar{\Omega} = P,$$

$$\text{dual problem (D)} : \langle b, y\rangle \to \min, \quad \text{s.t. } y \in (a + X^\perp) \cap \bar{\Omega} = D.$$

Here we assume that relative interiors of feasible regions P, D are not empty sets, i.e., $ri(P) = (b + X) \cap \Omega \neq \varnothing$, $ri(D) = (a + X^\perp) \cap \Omega \neq \varnothing$.

Lemma 2.1. *(Faybusovich, [3]) For a positive number β, give the pair of optimization problems*

$$(P_\beta): \quad f_\beta(x) = \beta\langle a, x\rangle - \log\det(x) \to min, \quad s.t.\ x \in ri(P),$$

$$(D_\beta): \quad g_\beta(y) = \beta\langle b, y\rangle - \log\det(y) \to min, \quad s.t.\ y \in ri(D),$$

where det *is defined in the sense of Jordan algebras. Then the following are necessary and sufficient conditions for that $x(\beta)$, $y(\beta)$ are optimal solutions of minimization problems (P_β), (D_β), respectively:*

$$x(\beta) \in ri(P),\ y(\beta) \in ri(D),\ x(\beta) \circ y(\beta) = \frac{e}{\beta}. \tag{2.1}$$

In addition $x(\beta)$, $y(\beta)$ converge to optimization solutions (P), (D), respectively as $\beta \to +\infty$.

The trajectory of $x(\beta)$ as $\beta \to +\infty$ is called a central path of a primal problem (P). Similarly the trajectory of $y(\beta)$ as $\beta \to +\infty$ is called a central path of the dual problem (D). We often consider the trajectory of a pair $(x(\beta), y(\beta))$ a central path of an optimal problem on $ri(P) \times ri(D)$. (Solving an optimization problem practically, we also define the slack variable for $(x(\beta), y(\beta))$ and trace its trajectory.)

The following proposition is used to prove Lemma 2.1.

Lemma 2.2. *(Faybusovich, [3]) For optimal solutions of $x(\beta)$ and $y(\beta)$,*

$$x(\beta) \in ri(P),\ y(\beta) \in ri(D),$$

$$\beta a - x(\beta)^{-1} \in X^{\perp},\ \beta b - y(\beta)^{-1} \in X,\ y(\beta) = \frac{1}{\beta} x(\beta)^{-1},$$

where x^{-1} is the inverse of $x \in \Omega$ relative to Jordan product. Conversely, optimal solutions of (P_β), (D_β), which hold these conditions, are uniquely determined, respectively.

3 Dualistic structure of primal-dual interior-point methods

Let V be an Euclidean Jordan algebra and Ω be the symmetric cone associated with V. Let $\{e_1, \ldots, e_n\}$ be a basis of V. The formula $x = x^i(x)e_i$, for $x \in V$, defines the component functions $x^1, \ldots, x^n$ on V. We shall consider $\{x^1, \ldots, x^n\}$ as an affine coordinate system on V. Set $\psi = -\log\det = (r/n)\log\varphi$, where r is the rank of V and φ is the characteristic function of Ω. Then the dual affine coordinate system $\{x^{1'}, \ldots, x^{n'}\}$, a Riemannian metric g, the canonical flat affine connection ∇, and the dual flat affine connection ∇' are defined, respectively, by:

$$x^{i'} = x^i \circ \iota = -\frac{\partial \psi}{\partial x^i}, \qquad g = \sum_{i,j} \frac{\partial^2 \psi}{\partial x^i \partial x^j} dx^i dx^j,$$

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = 0, \tag{3.1}$$

$$\nabla'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \iota_*^{-1}(\nabla_{\iota_*(\frac{\partial}{\partial x^i})} \iota_*(\frac{\partial}{\partial x^j})),$$

where ι is the restriction of a map $x \in \{x \mid x : \text{invertible}\} \mapsto x^{-1} \in V$ on Ω. We often call ∇' the dual connection simply. Note that

$$\nabla'_{\frac{\partial}{\partial x^{i'}}} \frac{\partial}{\partial x^{j'}} = 0. \tag{3.2}$$

By (3.1), (3.2), the triple (g, ∇, ∇') is called dually flat structure. The triple (Ω, ∇, g) is a flat statistical manifold, and (Ω, ∇', g) the dually flat statistical manifold. Including the case for non-flat connections ∇ or ∇', we call the similar structure as dualistic structure or information geometric structure [1] [10] [12] [22].

Let $(ri(P), \nabla, g)$ be a statistical submanifold which is the restriction of (Ω, ∇, g) on $ri(P)$, and $(ri(P), \nabla', g)$ a statistical submanifold which is the restriction of (Ω, ∇', g) on $ri(P)$. Ohara and Tsuchiya showed that, as $\beta \to +\infty$, an optimal solution $x(\beta)$ of a minimization problem (P_β) converges to an optimal solution of primal problem (P) along a geodesic on $(ri(P), \nabla', g)$ [11]. This fact includes results similar to LP ([14]) or to SDP ([8]). In [8] [11], it is shown that we have as many iterations on Newton's methods as large embedded curvature of a submanifold $(ri(P), \nabla', g)$ in (Ω, ∇', g).

4 The NT direction and dualistic structure

We treat a set of conditions (2.1) given by Lemma 2.1,

$$x(\beta) \in ri(P),\quad y(\beta) \in ri(D),\quad x(\beta) \circ y(\beta) = \frac{e}{\beta},$$

a pair of the optimization problems (P_β) and (D_β). Researchers on optimization problems are interested in how to seek limits of $(x(\beta), y(\beta))$ for $\beta \to +\infty$ by Newton's method.

The Newton's direction is given as a pair $(\Delta x, \Delta y)$ satisfying an equation

$$L(y)\Delta x + L(x)\Delta y = \frac{e}{\beta} - x \circ y \ \text{ for } \ x \in ri(P),\ y \in ri(D),\ \Delta x \in X,\ \Delta y \in X^\perp.$$

For x, y in V, let $L(x)$ and $P(x)$ be endomorphisms of V defined by

$$L(x)y = x \circ y, \quad P(x) = 2L(x)^2 - L(x^2).$$

Newton's directions scaled by elements in $G(\Omega)$ are useful. For example next directions are suggested in [5] [7] [23].

(1) HRVW/KSH/M direction : primal scaling

$$(\Delta x, \Delta y) := (P(y)^{-\frac{1}{2}}\Delta\tilde{x}, P(y)^{\frac{1}{2}}\Delta\tilde{y}) \in V \times V,$$

where $(\tilde{x}, \tilde{y}) := (P(y)^{\frac{1}{2}}x, P(y)^{-\frac{1}{2}}y) = (P(y)^{\frac{1}{2}}x, e)$.
(2) NT direction : primal-dual scaling

$$(\Delta x, \Delta y) := (P(z)^{-\frac{1}{2}}\Delta\tilde{x}, P(z)^{\frac{1}{2}}\Delta\tilde{y}) \in V \times V,$$

where $(\tilde{x}, \tilde{y}) := (P(z)^{\frac{1}{2}}x, P(z)^{-\frac{1}{2}}y)$ for ${}^{\exists 1}z \in \Omega$ such that $P(z)^{\frac{1}{2}}x = P(z)^{-\frac{1}{2}}y$.

Note that $\Delta\tilde{x} \in X, \Delta\tilde{y} \in X^\perp$ and that for invertible $w, v \in \Omega$

$$w^{\frac{1}{2}} \circ w^{\frac{1}{2}} = w, \quad w^{\frac{1}{2}} \in \Omega,$$

$$P(w)^{\frac{1}{2}} = P(w^{\frac{1}{2}}) \in G, \quad P(w)^{-1} = P(w^{-1}),$$

$$P(w)^{-\frac{1}{2}}w = e, \quad (P(w)v)^{-1} = P(w^{-1})v^{-1},$$

where G is the connected component of the identity of $G(\Omega)$ the automorphism group of Ω [2]. On the NT direction we have

Theorem 4.1. *For an automorphism $P(z)^{\frac{1}{2}}$ appearing in the NT direction, the point*

$$z = P(x^{-\frac{1}{2}})(P(x^{\frac{1}{2}})y)^{\frac{1}{2}} \tag{4.1}$$

is the middle point of the geodesic joining x^{-1} to y, determined by the Levi-Civita connection $(\nabla + \nabla')/2$.

Proof. Sturm described that z by (4.1) is a spectral geometric mean in [13]. Coincidence with a spectral geometric mean and a middle point of a geodesic by the Levi-Civita connection is mentioned in Ohara [9], where "middle" means a parameter $s = 1/2$ on the geodesic $P(x^{-\frac{1}{2}})(P(x^{\frac{1}{2}})y)^s$ $(s \in [0, 1])$ joining x^{-1} $(s = 0)$ to y $(s = 1)$. Thus Theorem 4.1 holds. □

Geometric means are defined on symmetric cones as above, extended from geometric means on matrixes or on linear operators [4]. For matrices x and y, we have

$$z = x^{-\frac{1}{2}}(x^{\frac{1}{2}}yx^{\frac{1}{2}})^{\frac{1}{2}}x^{-\frac{1}{2}}$$

following by $P(w)v = wvw$ for matrices w, v.

Suppose $x(\beta) \in ri(P)$, $y(\beta) \in ri(D)$ again. Setting a parameter $s \in [0, 1]$ along a geodesic $P(y^{\frac{1}{2}})(P(y^{-\frac{1}{2}})x^{-1})^s$, we obtain

Corollary 4.2. *For an automorphism $P(z)^{\frac{1}{2}}$ appearing in the NT direction, the point*

$$z = P(y^{\frac{1}{2}})(P(y^{-\frac{1}{2}})x^{-1})^{\frac{1}{2}}$$

is the middle point of the geodesic joining y to x^{-1}, determined by the Levi-Civita connection $(\nabla + \nabla')\ /\ 2$.

5 Power classes and dualistic structure

In this section we talk about the family of search direction called a power class of search directions and dualistic structure on linear programming over symmetric cones.

Muramatsu proposed a primal-dual interior-point algorithm using

$$G(x, y) = \{g \in G \mid gx \text{ and } g^{-*}y \text{ share a Jordan frame}\} \text{ for } (x, y) \in \Omega \times \Omega,$$

where G is the connected component of the identity of $G(\Omega)$ the automorphism group of Ω, and $^{-}$, * are the inverse, the adjoint, respectively. For example, $g \in G(x, y)$ given by

$$g^{-*}y = (gx)^q \tag{5.1}$$

defines the power class of search directions, which has an index q $(q = 0, 1, 2, \cdots)$ [5]. It is known that, to decide g for the q power class (5.1), we choose

$$g = P(z)^{\frac{1}{2}}, \quad z = P(x^{-\frac{1}{2}})(P(x^{\frac{1}{2}})y)^{\frac{1}{q+1}}.$$

Then the next follows by a description of a geodesic with respect to Levi-Civita connection in proof of Theorem 4.1.

Theorem 5.1. *For an automorphism $P(z)^{\frac{1}{2}}$ appearing in the q power class of search direction, the point*

$$z = P(x^{-\frac{1}{2}})(P(x^{\frac{1}{2}})y)^{\frac{1}{q+1}}$$

is a weighted geometric mean (or a $1/(q + 1)$ power mean) of x^{-1} and y, i.e., the $1 : q$ internally dividing point on a geodesic joining x^{-1} to y, and determined by the Levi-Civita connection $(\nabla + \nabla')/2$.

"$1 : q$" means a parameter $s = 1/(q + 1)$ on the geodesic $P(x^{-\frac{1}{2}})(P(x^{\frac{1}{2}})y)^s$ $(s \in [0, 1])$ joining x^{-1} to y with respect to the Levi-Civita connection. For matrices x and y, we have

$$z = x^{-\frac{1}{2}}(x^{\frac{1}{2}}yx^{\frac{1}{2}})^{\frac{1}{q+1}}x^{-\frac{1}{2}}.$$

In the case of $q = 0$, $q = 1$, Theorem 5.1 is applied to the HRVW/KSH/M direction, the NT direction, respectively [5].

Along a geodesic $P(y^{\frac{1}{2}})(P(y^{-\frac{1}{2}})x^{-1})^s$, we obtain

Corollary 5.2. *For an automorphism $P(z)^{\frac{1}{2}}$ appearing in the q power class of search direction, the point*

$$z = P(y^{\frac{1}{2}})(P(y^{-\frac{1}{2}})x^{-1})^{\frac{q}{q+1}}$$

is a weighted geometric mean of y and x^{-1}, i.e., the $q:1$ internally dividing point on a geodesic joining y to x^{-1}, and determined by the Levi-Civita connection $(\nabla + \nabla')/2$.

We treat an element $z = P(x^{-\frac{1}{2}})(P(x^{\frac{1}{2}})y)^{\frac{1}{q+1}}$ again. If $y = x^{-1}/\beta$ (cf. Lemma 2.2), it holds that

$$g^{-*}y = P(z)^{-\frac{1}{2}}y = \frac{e}{\beta^{\frac{q}{q+1}}} = (P(z)^{\frac{1}{2}}x)^q = (gx)^q.$$

6 Conclusion

Our original results are related to q power classes of search direction of primal-dual interior-point methods for symmetric cones with respect to information geometry. The paper underlines that some surveying (geodesics) plays a central role in the cone optimization problems with linear constraints. For real values q, not only for non negative integers q, we have to investigate q power classes.

Acknowledgements. The author was supported in part by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (B), 16740070, in Japan. Thanks to referees of BJGA for pertinent observations.

References

[1] S. Amari, H. Nagaoka, *Method of Information Geometry*, Amer. Math. Soc., Providence, Oxford University Press, Oxford 2000.

[2] J. Faraut, A. Korányi, *Analysis on Symmetric Cones*, Clarendon Press, Oxford 1994, Oxford University Press, New York 2002.

[3] L. Faybusovich, *Linear systems in Jordan algebras and primal-dual interior-point algorithms*, J. Computational and Applied Math. 86 (1997), 149-175.

[4] F. Kubo, T. Ando, *Means of positive linear operators*, Math. Ann. 246 (1980), 205-224.

[5] M. Muramatsu, *On commutative class of search directions for linear programming over symmetric cones*, J. Optim. Theory Appl. 112, 3 (2002), 595-625.

[6] Y. Nesterov, A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia 1994.

[7] Y. Nesterov, M. J. Todd, *Primal-dual interior-point methods for self-scaled cones*, SIAM J. Optim. 8 (1998), 324-364.

[8] A. Ohara, *Information geometric analysis of an interior point method for semidefinite programming*, O.Barndorf-Nielsen and V.Jensen eds., Proc. of Geometry in Present Day Science, World Sci., Singapore 1999, 49-74.

[9] A. Ohara, *Geodesics for dual connections and means on symmetric cones*, Integr. Eq. Oper. Theory 50 (2004), 537-548.

[10] A. Ohara, N. Suda, S. Amari, *Dualistic differential geometry of positive definite matrices and its applications to related problems*, Linear Algebra Appl. 247 (1996), 31-53.

[11] A. Ohara, T. Tsuchiya, *An information geometric approach to polynomial-time interior-point algorithms: Complexity bound via curvature integral*, Research Memorandum, The Institute of Statistical Mathematics 1055 (2007).

[12] H. Shima, *The Geometry of Hessian Structures*, World Sci., Singapore 2007.

[13] J. F. Sturm, *Similarity and other spectral relations for symmetric cones*, Linear Algebra Appl. 312 (2000), 135-154.

[14] K. Tanabe, *Center flattening transformation and a centered Newton method for linear programming*, Manuscript presented at MP seminar, The Operations Research Society of Japan, July (1987).

[15] C. Udrişte, *Convex Functions and Optimization Methods on Riemannian Manifolds*, Mathematics and Its Applications, 297, Kluwer Academic Publishers, Dordrecht, Boston, London, 1994.

[16] C. Udrişte, *Riemannian convexity in programming (I)*, An. St. Univ. Al. I. Cuza University of Iassy, S. I-a, Mat., 42 (1996), 123-136.

[17] C. Udrişte, *Riemannian convexity in programming (II)*, Balkan J. Geom. Appl., 1, 1 (1996), 99-109.

[18] C. Udrişte, *Sufficient decrease principle on Riemannian manifolds*, Balkan J. Geom. Appl., 1, 2 (1996), 111-123.

[19] C. Udrişte, *Optimization methods on Riemannian manifolds*, Algebras, Groups and Geometries, 14 (1997), 339-359.

[20] C. Udrişte, G. Bercu, *Riemannian Hessian metrics*, Analele Universităţii Bucureşti, 55, 1 (2005), 189-204.

[21] C. Udrişte, G. Bercu, *Complexity analysis for logarithmic barrier algorithm on Riemannian manifolds*, Tensor, NS, 67, 3 (2006), 226-234.

[22] K. Uohashi, A. Ohara, *Jordan algebras and dual affine connections on symmetric cones*, Positivity 8 (2004), 369-378.

[23] S. J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia 1997.

Author's address:

Keiko Uohashi
Department of Mechanical Engineering & Intelligent Systems,
Faculty of Engineering, Tohoku Gakuin University,
Tagajo, Miyagi 985-8537, Japan.
E-mail: uohashi@tjcc.tohoku-gakuin.ac.jp