Local Quadratic Convergence of Polynomial-time Interior-Point Methods[a]

Levent Tunçel

Dept. of Combinatorics and Optimization,

Faculty of Mathematics,

University of Waterloo,

Canada.

ltuncel@math.uwaterloo.ca

March 2, 2010

[a] Joint work with **Yu. Nesterov**.

Interior-Point Methods

Modern theory of interior-point methods pushes all *difficult* constraints into the *convex set constraints* and/or *convex cone constraints*. Then, each of these *convex inclusion constraints* is treated via a strictly convex *barrier function* with very special properties.

Theory of interior-point methods for beautiful and powerful special cases:

- Symmetric (self-scaled) cones (Nesterov and Todd [1996-...] and others)
- Homogeneous cones (Güler and T. [1998], Chua [2008], Chua and T. [2008])
- Hyperbolic cones (Güler [1997], Renegar [2006-...], today's talk, ...)

We denote by $\mathbb{E}$ a finite-dimensional linear space (other variants: $\mathbb{H}$, $\mathbb{V}$), and by $\mathbb{E}^*$ its *dual* space, composed by linear functions on $\mathbb{E}$. The value of function $s \in \mathbb{E}^*$ at point $x \in \mathbb{E}$ is denoted by $\langle s, x \rangle$.

For a linear transformation $A : \mathbb{E} \to \mathbb{H}^*$ we denote by A^* the corresponding *adjoint* transformation:

$$\langle Ax, y \rangle \;\; = \;\; \langle A^* y, x \rangle, \quad x \in \mathbb{E}, \; y \in \mathbb{H}.$$

Thus, $A^* : \mathbb{H} \to \mathbb{E}^*$. A self-adjoint positive-definite linear transformation $B : \mathbb{E} \to \mathbb{E}^*$ (notation $B \succ 0$) defines the Euclidean norms for the primal and dual spaces:

$$\| x \|_B \;\; = \;\; \langle Bx, x \rangle^{1/2}, \; x \in \mathbb{E}, \quad \| s \|_B \; = \; \langle s, B^{-1} s \rangle^{1/2}, \; s \in \mathbb{E}^*.$$

The sense of this notation is determined by the space of arguments.

Let Φ be a self-concordant function defined on the interior of a convex set $Q \subset \mathbb{E}$:

$$\nabla^3 \Phi(x)[h,h,h] \;\; \leq \;\; 2 \langle \nabla^2 \Phi(x) h, h \rangle^{3/2}, \quad x \in \operatorname{int} Q, \;\; h \in \mathbb{E}.$$

Note that $\nabla^3 \Phi(x)[h_1, h_2, h_3]$ is a trilinear symmetric form. Thus,

$$\nabla^3 \Phi(x)[h_1, h_2] \;\; = \;\; \nabla^3 \Phi(x)[h_2, h_1] \; \in \; \mathbb{V}^*,$$

and $\nabla^3 \Phi(x)[h]$ is a self-adjoint linear transformation from $\mathbb{V}$ to $\mathbb{V}^*$.

We call cone $K \subset \mathbb{E}$ *regular* if it is a closed, convex, and pointed cone with nonempty interior. Sometimes it is convenient to write inclusion $x \in K$ in the form $x \succeq_K 0$.

If K is regular, then the *dual cone*

$$K^* \quad := \quad \{ s \in \mathbb{E}^* : \; \langle s, x \rangle \geq 0, \; \forall x \in K \; \},$$

is also regular.

Every such convex cone K, admits a ν-normal barrier $F(x)$ (that is, F is self-concordant and ν-logarithmically homogeneous:

$$F(\tau x) \quad = \quad F(x) - \nu \ln \tau, \quad x \in \operatorname{int} K, \ \tau > 0.)$$

Note that $-\nabla F(x) \in \operatorname{int} K^*$ for every $x \in \operatorname{int} K$.

Note that the *dual barrier*

$$F_*(s) \quad := \quad \max_{x \in \operatorname{int} K} \{ -\langle s, x \rangle - F(x) \}$$

is a ν-normal barrier for cone K^*.

1 Conic Programming, Central Path

Consider the standard conic optimization problem:

$$\min_{x \in K} \{ \langle c, x \rangle : \; Ax = b \} ,$$

where $c \in \mathbb{E}^*$, $b \in \mathbb{H}^*$, A is a linear transformation from $\mathbb{E}$ to $\mathbb{H}^*$, and $K \subset \mathbb{E}$ is a regular cone. The dual problem is then

$$\max_{s \in K^*, \; y \in \mathbb{H}} \{ \langle b, y \rangle : \; s + A^* y = c \} .$$

$\mathcal{F}_p := \{x \in K : \; Ax = b\}$, and $\mathcal{F}_d := \{s \in K^* : \; s + A^* y = c\}$.

Under the *strict feasibility* assumption,

$$\exists\, x_0 \in \operatorname{int} K,\; s_0 \in \operatorname{int} K^*,\; y_0 \in \mathbb{H} : \quad Ax_0 = b, \quad s_0 + A^* y_0 = c,$$

the optimal sets of the primal and dual problems are nonempty and bounded, and there is no duality gap.

Primal-dual central path $z_\mu := (x_\mu, s_\mu, y_\mu)$:

$$\left.\begin{array}{rcl} Ax_\mu & = & b, \\ c + \mu \nabla F(x_\mu) & = & A^* y_\mu, \\ s_\mu & = & -\mu \nabla F(x_\mu) \end{array}\right\}, \quad \mu > 0,$$

is well defined. Note that

$$\langle c, x_\mu \rangle - \langle b, y_\mu \rangle = \langle s_\mu, x_\mu \rangle = \nu \cdot \mu.$$

2 Prediction from a neighborhood of the central path

For a fast local convergence of a path-following scheme, we need to show that the predicted point

$$\hat{z}_\mu \;=\; z_\mu - z'_\mu \cdot \mu$$

enters a small neighborhood of the solution point

$$z_* \;=\; \lim_{\mu \to 0} z_\mu \;=\; (x_*, s_*, y_*).$$

It is more convenient to analyze this situation by looking at y-component of the central path.

Note that s-component of the dual problem can be easily eliminated:

$$s \;=\; s(y) \;:=\; c - A^* y.$$

More concise full-dimensional form:

$$f^* \;:=\; \max_{y \in \mathbb{H}} \{ \langle b, y \rangle : \; y \in Q \},$$

$$Q \;:=\; \{ y \in \mathbb{H} : \; c - A^* y \in K^* \}.$$

In view of the strict feasibility assumption, the interior of set Q is nonempty. Moreover, for this set we have a ν-self-concordant barrier

$$f(y) \quad := \quad F_*(c - A^* y), \quad y \in \operatorname{int} Q.$$

We will use the predictor step

$$y_{k+1} \quad := \quad y_k + \alpha[\nabla^2 f(y)]^{-1} \nabla f(y),$$

for $\alpha > 0$ and prove that under mild assumptions, in addition to guaranteeing global linear rate convergence (poly. iteration complexity), we can take $\alpha \to 1$ in the limit in a way that we can attain local quadratic convergence.

Local quadratic convergence is a natural and very desired property of many methods in Nonlinear Optimization.

Local quadratic convergence is a natural and very desired property of many methods in Nonlinear Optimization.

However, for interior-point methods the corresponding analysis does not seem to be trivial.

The reason is that the barrier function is not defined in a neighborhood of the solution. Therefore, in order to study the behavior of the central path, we need to employ somehow the separable structure of the functional inequality constraints. From the very beginning (going back to Fiacco and McCormick [1968]), this analysis was based on the Implicit Function Theorem as applied to Karush-Kuhn-Tucker conditions.

What about an appropriate framework for analyzing the local behavior of general polynomial-time interior-point methods?.

What about an appropriate framework for analyzing the local behavior of general polynomial-time interior-point methods?.

Indeed, in the theory of self-concordant functions it seems difficult to analyze the local structure of the solution since we have no access to the components (individual coordinates) of the barrier function. Moreover, in general, it seems difficult to relate the self-concordant barrier with functional inequality constraints of the initial optimization problem.

Previously, the local superlinear convergence for polynomial-time path-following methods was proved only for Linear Programming Ye, Güler, Tapia and Y. Zhang [1993], Mehrotra [1993] and for Semidefinite Programming problems Kojima, Shindoh and Shida [1998] ,Potra and Sheng [1998], Luo, Sturm and S. Zhang [1998], Ji, Potra and Sheng [1999].

In the current work,

- we investigate the possibility of a local quadratic convergence theory for modern interior-point methods,
- we establish the local quadratic convergence of interior-point path-following methods by utilizing some geometric properties of the general conic optimization problem
- the main structural property used in our analysis is the logarithmic homogeneity of self-concordant barrier functions

- we propose new path-following predictor-corrector schemes which work only in the dual space
- the algorithms are based on an easily computable gradient proximity measure, which ensures an automatic transformation of the global linear rate of convergence to the local quadratic rate (under a mild assumption)
- our step-size procedure for the predictor step is related to the maximum step size to stay feasible.

Now we can introduce our main assumptions.

Assumption 1 *There exists a constant* $\gamma_d > 0$ *such that*

$$f^* - \langle b, y \rangle \; = \; \langle s, x_* \rangle \;\; \geq \;\; \gamma_d \|s - s_*\|_B \; \equiv \; \gamma_d \|y - y_*\|_G,$$

for every $y \in Q$ *(that is* $s = s(y) \in \mathcal{F}_d$*).*

Thus, we assume that the dual problem admits a *sharp* optimal solution.

We need one more assumption.

Assumption 2 *There exists a constant* σ_d *such that for any* $\mu \leq 1$ *we have*

$$\|\nabla^2 F_*(s_\mu) s_*\|_B \; \leq \; \sigma_d.$$

Example 1 *For the cone of positive-semidefinite matrices $K = K^* = \mathbb{S}^n_+$, we choose*

$$F(X) = -\ln\det X, \quad F_*(S) = n - \ln\det S.$$

Then,

$$\langle I, \nabla^2 F_*(S_\mu) S_* \rangle \quad = \quad \langle I, S_\mu^{-1} S_* S_\mu^{-1} \rangle.$$

We have

$$\langle I, S_\mu^{-1} S_* S_\mu^{-1} \rangle \quad = \quad \mu^{-2} \langle X_\mu^2, S_* \rangle \ = \ \mu^{-2} \langle (X_\mu - X_*)^2, S_* \rangle.$$

Thus, we get an upper bound for $\|\nabla^2 F_*(S_\mu) S_*\|$ assuming $\|X_\mu - X_*\| \leq O(\mu)$. This condition is weaker than assuming that the primal problem admits a sharp solution. It is also weaker than assuming differentiability of the primal central path at $\mu = 0$. ■

3 Neighborhood, proximity measure

For $\mu \in (0,1],\ \beta \in (0,\frac{1}{2})$,

$$\gamma(y,\mu) := \left\| \nabla f(y) - \frac{1}{\mu} b \right\|_y ,$$

and

$$\mathcal{N}(\mu,\beta) \quad := \quad \{ y \in \mathbb{H} : \ \gamma(y,\mu) \leq \beta \} .$$

This proximity measure has a very familiar interpretation in the special case of Linear Programming. Denoting by S the diagonal matrix made up from the slack variable $s = c - A^T y$, notice that Dikin's affine scaling direction in this case is given by $\left(AS^{-2}A^T\right)^{-1} b$. Our predictor step corresponds to the search direction $\left(AS^{-2}A^T\right)^{-1} AS^{-1}e$. Our proximity measure becomes

$$\left\| AS^{-1}e - \frac{b}{\mu} \right\|_{AS^{-2}A^T}.$$

$$p(y) \quad := \quad y + v(y), \quad y \in \operatorname{int} Q,$$

$$v(y) \quad := \quad [\nabla^2 f(y)]^{-1} \nabla f(y).$$

Theorem 1 *Let dual problem satisfy the assumptions 1 and 2. If for some* $\mu \in (0, 1]$ *and* $\beta \in (0, \frac{1}{2})$ *we have* $y \in \mathcal{N}(\mu, \beta)$, *then*

$$\|p(y) - y_*\|_G \quad \leq \quad \frac{4\sigma_d(1-\beta)^2}{\gamma_d^2(1-2\beta)^2} \langle b, y - y_* \rangle^2 \; \leq \; \frac{4\nu(1-\beta)^2\sigma_d}{\gamma_d^2(1-2\beta)^2} \cdot \|y - y_*\|_G^2.$$

4 Efficiency of the Predictor Step

$$y(\alpha) \quad := \quad y + \alpha v(y), \quad \alpha \in [0, 1].$$

Denote by $\bar{\alpha}(y)$, the maximal feasible step size along direction $v(y)$:

$$\bar{\alpha}(y) \quad := \quad \max_{\alpha \geq 0} \{\alpha : \ y + \alpha v(y) \in Q\}.$$

Theorem 2 *Let* $y \in \mathcal{N}(\mu, \beta)$ *with* $\mu \in (0, 1]$ *and* $\beta \in (0, \frac{1}{2})$*. Then*

$$1 - \bar{\alpha}(y) \leq \frac{\kappa\mu}{1+\kappa\mu},$$

$$\|y(\bar{\alpha}) - y^*\|_y \leq (1 + \sqrt{\nu})\kappa\mu.$$

5 Polynomial-time path-following method

Note that at the predictor stage, we need to choose the rate of decrease of the penalty parameter (central path parameter) μ as a function of the predictor step size α.

We denote

$$\xi_{\bar{\alpha}}(\alpha) \;=\; 1 + \tfrac{\alpha\bar{\alpha}}{\bar{\alpha}-\alpha}, \quad \alpha \in [0, \bar{\alpha}),$$

and choose

$$\mu(\alpha) \;\approx\; \tfrac{\mu}{\xi_{\bar{\alpha}}(\alpha)}.$$

$$\eta_{\bar{\alpha}}(\alpha) \quad := \quad \begin{cases} 2\alpha, & \alpha \in [0, \frac{1}{3}\bar{\alpha}], \\ \frac{\alpha+\bar{\alpha}}{2}, & \alpha \in [\frac{1}{3}\bar{\alpha}, \bar{\alpha}]. \end{cases}$$

This function will be used for updating the length of our predictor step.

Lemma 1 *If $\alpha \geq 0$ and $\alpha_+ = \eta_{\bar{\alpha}}(\alpha)$, then $\xi_{\bar{\alpha}}(\alpha_+) \geq 2\xi_{\bar{\alpha}}(\alpha) - 1$. Hence, for the recurrence*

$$\alpha_{i+1} \quad = \quad \eta_{\bar{\alpha}}(\alpha_i), \quad i \geq 0,$$

we have $\xi_{\bar{\alpha}}(\alpha_i) \geq 1 + \alpha_0 \cdot 2^i$.

Path-following method for general barriers

1. Set $\mu_0 = 1$ and find point $y_0 \in \mathcal{N}(\mu_0, \frac{1}{18})$.

2. For $k \geq 0$ iterate:

a) Compute $\bar{\alpha}_k = \bar{\alpha}(y_k)$.

b) Using recurrence

$$\alpha_{k,0} = \frac{1}{6 \cdot \max\{1, \|v(y_k)\|_{y_k}\}}, \quad \alpha_{k,i+1} = \eta_{\bar{\alpha}_k}(\alpha_{k,i}),$$

find the maximal $i \equiv i_k$ such that $\gamma\left(y_k(\alpha_{k,i}), \frac{\mu_k}{\xi_{\bar{\alpha}_k}(\alpha_{k,i})}\right) \leq \frac{1}{6}$.

c) Set $\alpha_k = \alpha_{k,i_k}$, $p_k = y_k + \alpha_k v(y_k)$, $\mu_{k+1} = \frac{\mu_k}{\xi_{\bar{\alpha}_k}(\alpha_k)}$.

d) Starting from p_k, apply the Newton method for

finding $y_{k+1} \in \mathcal{N}(\mu_{k+1}, \frac{1}{18}\mu_{k+1})$.

Theorem 3 *Let K be a regular cone and F_* be a normal barrier for K^*. Also let $y_0 \in \mathcal{N}(\mu_0, \frac{1}{18})$ for some $\mu_0 > 0$. Then, above method generates a sequence of feasible points such that*

$$f^* - \langle b, y_k \rangle \quad \leq \mu_0 \kappa_1 \exp\left\{-\frac{k}{1+6\nu^{1/2}}\right\}.$$

This is an

$$O\left(\sqrt{\nu}\ln\left(\frac{1}{\epsilon}\right)\right)$$

iteration complexity bound. Moreover, if Assumptions 1 and 2 hold, the method attains local quadratic convergence.

In this scheme, for computing the value of gradient proximity measure at new points, we need to compute and ”invert” the Hessian of barrier function. However, the step size in this procedure is rapidly increased. Therefore, it is easy to prove that the total number of auxiliary steps i_k, which is necessary for computing an ϵ-solution to our problem is bounded by $O(\nu^{1/2} \ln \frac{\nu}{\epsilon})$. The number of steps at the correction stage (Step 2d) cannot be large since p_k belongs to the region of quadratic convergence of the Newton method. In any case, if Assumptions 1 and 2 are satisfied, then the above method is locally quadratically convergent.

6 Recession coefficient of barrier function

Definition 1 *We call* γ_F recession coefficient *of the normal barrier* F *if it is the smallest positive constant such that for every* $x \in \operatorname{int} K$ *and* $u \in K$ *we have*

$$\nabla^2 F(x+u) \quad \preceq \quad \gamma_F \cdot \nabla^2 F(x).$$

Proposition 1 *For every normal barrier* F*,*

$$1 \leq \gamma_F \leq 4\nu^2.$$

However, very often this upper bound is very pessimistic. Note that the following main operations with convex cones *do not increase* this coefficient.

Theorem 4 *1. Let F be a normal barrier for the cone K, and*

$$K_A = \{x \in K : \ Ax = 0\}.$$

Denote by f the restriction of F onto the relative interior of K_A. Then $\gamma_f \leq \gamma_F$.

2. Let F_i, $i = 1, 2$, be normal barriers for cones $K_i \subset \mathbb{E}$. Denote $F = F_1 + F_2$. If $\mathrm{int}\,(K_1 \bigcap K_2) \neq \emptyset$, then $\gamma_F \leq \max\{\gamma_{F_1}, \gamma_{F_2}\}$.

3. Let F_i, $i = 1, 2$, be normal barriers for cones $K_i \subset \mathbb{E}_i$. Denote $F(x, y) = F_1(x) + F_2(y)$. Then $\gamma_F \leq \max\{\gamma_{F_1}, \gamma_{F_2}\}$.

Thus, all barriers constructed as sums or direct products of small-dimensional cones have small recession coefficients. On the other hand, restriction of such barriers onto linear subspaces does not increase the recession coefficient. It remains to note that there exists an important family of normal barriers with *minimal* value of the recession coefficient.

Definition 2 *Let F be a normal barrier for the regular cone K. We say that F has* negative curvature *if for every $x \in \operatorname{int} K$ and $h \in K$ we have*

$$\nabla^3 F(x)[h] \quad \preceq \quad 0.$$

Thus, for such a barrier $\gamma_F = 1$. It is clear that self-scaled barriers have negative curvature (Nesterov and Todd [1997]). Some other important barriers, like the logarithms of *hyperbolic polynomials* (Güler [1997]) also share this property.

Theorem 5 *Let K be a regular cone and F be a normal barrier for K. Then, TFAE:*

1. *F has negative curvature;*
2. *for every $x \in \operatorname{int} K$ and $h \in \mathbb{E}$ we have*
$$-\nabla^3 F(x)[h,h] \quad \in \quad K^*;$$
3. *for every $x \in \operatorname{int} K$ and for every $h \in \mathbb{E}$ such that $x + h \in \operatorname{int} K$, we have*
$$\nabla F(x+h) - \nabla F(x) \;\preceq_{K^*}\; \nabla^2 F(x) h.$$

Theorem 6 *Let the curvature of F be negative. Then for every $x \in K$, we have*

$$\nabla^2 F(x)h \quad \succeq_{K^*} \quad 0, \quad \forall h \in K,$$

and, consequently,

$$\nabla F(x+h) - \nabla F(x) \quad \succeq_{K^*} \quad 0.$$

Lemma 2 *Let both F and F^* have negative curvature. Then K is a symmetric cone.*

Proof: Indeed, for every $x \in \operatorname{int} K$ we have $\nabla^2 F(x) K \subseteq K^*$. Denote $s = -\nabla F(x)$. Since F^* has negative curvature, then $\nabla^2 F_*(s) K^* \subseteq K$. However, since $\nabla^2 F_*(s) = [\nabla^2 F(x)]^{-1}$, this means $K^* \subseteq \nabla^2 F(x) K$. Thus $K^* = \nabla^2 F(x) K$. Now, using standard arguments, it is easy to prove that for every pair $x \in \operatorname{int} K$ and $s \in \operatorname{int} K^*$ there exists a scaling point $w \in \operatorname{int} K$ such that $s = \nabla^2 F(w) x$ (this w can be taken as the minimizer of the convex function $-\langle \nabla F(w), x \rangle + \langle s, w \rangle$). Thus, we have proved that K is homogeneous and self-dual. Hence, it is symmetric. ■

Remark 1 *Lemma 2 shows that the value*

$$\min\{\gamma_F + \gamma_{F^*} - 2 : F \text{ is a normal barrier for } K\}$$

can be seen as a measure of distance *between the cone K and the family of symmetric cones.*

$$\sigma_x(h) \quad := \quad \min_{\rho \geq 0} \{\rho : \; \rho \cdot x - h \in K\}.$$

Theorem 7 *Let K be a regular cone and F be a normal barrier for K. Further let $x, x + h \in \operatorname{int} K$. Then for every $\alpha \in [0, 1)$ we have*

$$\tfrac{1}{\gamma_F(1+\alpha\sigma_x(h))^2} \nabla^2 F(x) \; \preceq \; \nabla^2 F(x + \alpha h) \quad \preceq \quad \tfrac{\gamma_F}{(1-\alpha)^2} \nabla^2 F(x).$$

7 Bounding the growth of our proximity measure

Let us analyze now our predictor step

$$y(\alpha) = y + \alpha v(y), \quad \alpha \in [0, \bar{\alpha}],$$

where $\bar{\alpha} = \bar{\alpha}(y)$. Denote $\bar{s} = s(y(\bar{\alpha})) \in K^*$.

Lemma 3 *For every $\alpha \in [0, \bar{\alpha})$, we have*

$$\delta_y(\alpha) \quad := \quad \|\nabla f(y(\alpha)) - \tfrac{\bar{\alpha}}{\bar{\alpha}-\alpha} \nabla f(y)\|_G \;\leq\; \tfrac{\alpha \gamma_{F_*}}{\bar{\alpha}-\alpha} \|\nabla^2 F_*(s(y))\bar{s}\|_B .$$

$$\Gamma_\mu(y,\alpha) := \gamma_{F_*}^{1/2}\left(1+\alpha\cdot\sigma_{s(y)}\left(-A^* v(y)\right)\right)\left\|\nabla f(y(\alpha)) - \frac{\xi_{\bar\alpha}(\alpha)}{\mu}\cdot b\right\|_y .$$

Theorem 8 *Let $y \in \mathcal{N}(\mu,\beta)$ with $\mu \in (0,1]$ and $\beta \in (0,\frac{1}{2})$. Then for $y(\alpha) = y + \alpha v(y)$ with $\alpha \in (0,\bar\alpha)$ we have*

$$\begin{aligned}\gamma\left(y(\alpha), \tfrac{\mu}{\xi_{\bar\alpha}(\alpha)}\right) &\leq \Gamma_\mu(y,\alpha)\\ &\leq \gamma_{F_*}^{1/2}(1+\alpha\cdot\sigma_{s(y)}(-A^* v(y)))\left[\gamma_1(\alpha)+\beta\cdot\left(1+\tfrac{\alpha\bar\alpha}{\bar\alpha-\alpha}\right)\right].\end{aligned}$$

Lemma 4 *Let* $y \in \mathcal{N}(\mu, \beta)$ *with* $\beta \leq \frac{1}{18\gamma_{F_*}^{1/2}}$. *Then for all*

$$\alpha \in \left[0, \frac{1}{6\,\gamma_{F_*}^{1/2} \max\{1, \|v(y)\|_y\}}\right]$$

we have $\Gamma_\mu(y, \alpha) \leq \frac{1}{6}$.

Path-following method based on recession coefficient

1. Set $\mu_0 = 1$ and find point $y_0 \in \mathcal{N}\left(\mu_0, \frac{1}{18\sqrt{\gamma_{F_*}}}\right)$.

2. For $k \geq 0$ iterate:

a) Compute $\bar{\alpha}_k = \bar{\alpha}(y_k)$.

b) Using recurrence

$$\alpha_{k,0} = \frac{1}{6 \cdot \sqrt{\gamma_{F_*}} \max\{1, \|v(y_k)\|_{y_k}\}}, \quad \alpha_{k,i+1} = \eta_{\bar{\alpha}_k}(\alpha_{k,i}),$$

find the maximal $i \equiv i_k$ such that $\Gamma_{\mu_k}(y_k, \alpha_{k,i}) \leq \frac{1}{6}$.

c) Set $\alpha_k = \alpha_{k,i_k}$, $p_k = y_k + \alpha_k v(y_k)$, $\mu_{k+1} = \frac{\mu_k}{\xi_{\bar{\alpha}_k}(\alpha_k)}$.

d) Starting from p_k, apply the Newton method for finding $y_{k+1} \in \mathcal{N}\left(\mu_{k+1}, \frac{\mu_{k+1}}{18\sqrt{\gamma_{F_*}}}\right)$.

For the above method, we can prove a polynomial complexity bound:

$$f^* - \langle b, y_k \rangle \quad \leq \mu_0 \kappa_1 \exp \left\{ - \frac{k}{1+6\sqrt{\gamma_{F_*} \nu}} \right\}.$$

On the other hand, in a small neighborhood of the solution, the method accelerates to the quadratic convergence rate.

Theorem 9 *Let K be a regular cone and F_* be a normal barrier for K^* with negative curvature. Also let $y_0 \in \mathcal{N}(\mu_0, \frac{1}{18})$ for some $\mu_0 > 0$. Then, above method generates a sequence of feasible points such that*

$$f^* - \langle b, y_k \rangle \quad \leq \mu_0 \kappa_1 \exp\left\{-\frac{k}{1+6\nu^{1/2}}\right\}.$$

This is an

$$O\left(\sqrt{\nu}\ln\left(\frac{1}{\epsilon}\right)\right)$$

iteration complexity bound. Moreover, if Assumptions 1 and 2 hold, the method attains local quadratic convergence.

7.1 Examples of cones with negative curvature

Negative curvature of barrier functions is preserved by the following operations.

- If barriers F_i for cones $K_i \subset \mathbb{E}_i$, $i = 1, 2$, have negative curvature, then the curvature of the barrier $F_1 + F_2$ for the cone $K_1 \oplus K_2$ is negative.

- If barriers F_i for cones $K_i \subset \mathbb{E}$, $i = 1, 2$, have negative curvature, then the curvature of the barrier $F_1 + F_2$ for the cone $K_1 \bigcap K_2$ is negative.

- If barrier F for cone K has negative curvature, then the curvature of the barrier $f(y) = F(A^* y)$ for the cone $K_y = \{y \in \mathbb{H} : \; A^* y \in K\}$ is negative.

- If barrier $F(x)$ for cone K has negative curvature, then the curvature of its restriction onto the linear subspace $\{x \in \mathbb{E} : \ Ax = 0\}$ is negative.

At the same time, we know two important families of cones with negative curvature.

- Self-scaled barriers have negative curvature (Nesterov and Todd [1996]).
- Let $p(x)$ be hyperbolic polynomial. Then the barrier $F(x) = -\ln p(x)$ has negative curvature (Güler [1997]).

Thus, using above mentioned operations, we can construct barriers with negative curvature for many interesting cones.

8 Exploiting the Primal-Dual Asymmetry of a Problem

In some situations we can argue that currently, some nonsymmetric treatments of the primal-dual problem pair have better complexity bounds than the primal-dual symmetric treatments.

Example 2 *Consider the cone of nonnegative univariate polynomials:*

$$K = \left\{ p \in \mathbb{R}^{2n+1} : \sum_{i=0}^{2n} p_i t^i \geq 0, \ \forall t \in \mathbb{R} \right\}.$$

The dual to this cone is the cone of positive semidefinite Hankel matrices.

For $k = 0 \dots, 2n$, *denote*

$$H_k \in \mathbb{R}^{(n+1)\times(n+1)} : \; H_k^{(i,j)} \;=\; \begin{cases} 1, & \text{if } i + j = k + 2 \\ 0, & \text{otherwise} \end{cases}$$

For $s \in \mathbb{R}^{2n+1}$ *we can define now the following linear transformation:*

$$H(s) \;=\; \sum_{i=0}^{2n} s_i \cdot H_i.$$

Then the cone dual to K *can be represented as follows:*

$$K^* \;=\; \{s \in \mathbb{R}^{2n+1} : \; H(s) \succeq 0\}.$$

The natural barrier for the dual cone is $f(s) = -\ln \det H(s)$. *Clearly, it has negative curvature.*

Note that we can lift the primal cone to a higher dimensional space (see Nesterov [1999]):

$$K = \{p \in \mathbb{R}^{2n+1} : \; p_i = \langle H_i, Y \rangle, \; Y \succeq 0, \; i = 0, \ldots, 2n\},$$

and use $F(Y) = -\ln \det Y$ *as a barrier function for the extended feasible set. However, in this case we significantly increase the number of variables. Moreover, we need* $O(n^3)$ *operations for computing the value of the barrier* $F(Y)$ *and its gradient. On the other hand, in the dual space the cost of all necessary computations is very low (*$O(n \ln^2 n)$ *for the function value and* $O(n^2 \ln^2 n)$ *for solution of the Newton system, see Genin et al. [2003]). On top of these advantages, for non-degenerate dual problems, now we have a locally quadratically convergent path-following scheme.*

9 Conclusion

- Negative curvature seems to be a natural property of self-concordant barriers.

- Negative curvature seems to be a natural property of self-concordant barriers.
- Is it possible to construct a self-concordant barrier with negative curvature for any convex cone?

- Negative curvature seems to be a natural property of self-concordant barriers.
- Is it possible to construct a self-concordant barrier with negative curvature for any convex cone? However, we have already seen that for nonsymmetric cones this property is *not* dual-invariant.

- Negative curvature seems to be a natural property of self-concordant barriers.
- Is it possible to construct a self-concordant barrier with negative curvature for any regular cone? However, we have already seen that for nonsymmetric cones this property is *not* dual-invariant.
- Do there exist self-concordant barrier functions which have small recession coefficients for arbitrary regular cone?

Strictly speaking we have,

$$LP \subset SOCP \subset SDP \subset SymCP \subset HomCP \subset HypCP \subset CP.$$

Strictly speaking we have,

$$LP \subset SOCP \subset SDP \subset SymCP \subset HomCP \subset HypCP \subset CP.$$

However, in some sense,

$$LP \subset SOCP \subseteq SDP = SymCP = HomCP \subseteq HypCP \subset CP.$$

Strictly speaking we have,

$$LP \subset SOCP \subset SDP \subset SymCP \subset HomCP \subset HypCP \subset CP.$$

However, in some sense,

$$LP \subset SOCP \subseteq SDP = SymCP = HomCP \subseteq HypCP \subset CP.$$

Yet in an another sense,

$$LP = SOCP \subseteq SDP = SymCP = HomCP \subseteq HypCP \subseteq CP.$$

See, Ben-Tal and Nemirovski, Chua, Faybusovich, Nesterov and Nemirovski, Vinnikov, Helton and Vinnikov, Lewis, Parillo and Ramana, Gurvits.