Geodesic Newton method

Constantin Udrişte and Gabriel Bercu

Abstract

Many optimization problems require naturally the structure of a Riemannian manifold. This structure is involved by means of Riemannian connection, induced distance, geodesics, sectional curvature.

In this paper we present the complexity analysis of the logarithmic barrier algorithm using tools of Riemannian geometry. For generality, the step is taken as $\tau = \dfrac{1}{\lambda k}$, where $\lambda \geq 1$ is a constant and $k \geq 1$ is the constant of self-concordance.

Mathematics Subject Classification: 90C51, 90C53, 53C21.
Key words: Riemannian convex programming, Newton geodesic algorithm, barrier function, self-concordant function.

1 Preliminary Discussion

The Riemannian Geometry has been used in mathematical optimization as an important instrument to obtain new classess of methods. The interior point methods for convex optimization is such an example, being developed due their computational and theoretical utility; see for example [2], [5], [10], [11].

From the point of view of the interior point theory, the self-concordance of functions have a major role.

The paper is organized as follows. In Section 2 we give a general Riemannian framework for the logarithmic barrier method. In Section 3 we analyse a special case: Riemannian convex programming. The self-concordance property plays a key role in this analysis. We study the properties of the Newton method near the central path of a convex program using simultaneously the initial Riemannian metric and a Hessian Riemannian metric. In Section 4 we present the complexity analysis of the logarithmic barrier algorithm applied to convex program.

Proceedings of The 3-rd International Colloquium "Mathematics in Engineering and Numerical Physics" October 7-9 , 2004, Bucharest, Romania, pp. 285-298.

2 The primal central path in convex programming

Let (M, g) be a complete n-dimensional Riemannian manifold. We consider the primal convex programming problem

$$(P) \qquad \max\{f(x) | a_\alpha(x) \leq 0,\ \alpha = 1, \ldots, m,\ x \in M\}$$

satisfying the following conditions:

1) the interior of the feasible region $\mathcal{F} : a_\alpha(x) \leq 0$ denoted by $\mathcal{F}^\circ$, is nonempty and bounded

2) the functions $-f$, a_α are C^2 convex functions on $(\mathcal{F}^\circ, g)$.

The Riemannian convexity of the functions a_α implies the total Riemannian convexity of the set $\mathcal{F}^\circ$.

The logarithmic barrier function associated to (P) is

$$\phi(x, \mu) = -\frac{f(x)}{\mu} - \sum_{\alpha=1}^{m} \ln(-a_\alpha(x)),$$

where μ is the barrier strictly positive parameter. The existence of the logarithmic imposes $a_\alpha(x) < 0$. That is why the barrier function will prevent the iterate from going outside the feasible region. For this reason the logarithmic barrier function method is called an interior point method.

By supposing the functions f, $a_1, \ldots, a_n$ as being differentiable, we may define:

the gradient

$$\nabla\phi(x, \mu) = \frac{-\nabla f(x)}{\mu} - \sum_{\alpha=1}^{m} \frac{\nabla a_\alpha(x)}{-a_\alpha(x)},$$

where

$$\nabla f = g^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j}, \quad \nabla a_\alpha = g^{ij} \frac{\partial a_\alpha}{\partial x^i} \frac{\partial}{\partial x^j}, \quad \alpha = 1, \ldots, m.$$

the Hessian

$$\nabla^2\phi(x, \mu) = \frac{-\nabla^2 f(x)}{\mu} + \sum_{i=1}^{m} \left[\frac{\nabla^2 a_\alpha(x)}{-a_\alpha(x)} + \frac{\nabla a_\alpha(x) \otimes \nabla a_\alpha(x)}{a_\alpha^2(x)} \right],$$

where

$$\nabla^2 f = \left(\frac{\partial^2 f}{\partial x^i \partial x^j} - \Gamma_{ij}^m \frac{\partial f}{\partial x^m} \right) dx^i \otimes dx^j$$

and

$$\nabla^2 a_\alpha = \left(\frac{\partial^2 a_\alpha}{\partial x^i \partial x^j} - \Gamma_{ij}^m \frac{\partial a_\alpha}{\partial x^m} \right) dx^i \otimes dx^j.$$

Assume that the Hessian $\nabla^2\phi$ is positive definite. Then $\nabla^2\phi$ is used as a new Riemannian metric on $\mathcal{F}^\circ \subset M$. The function ϕ is strictly convex on $\mathcal{F}^\circ$ in (M, g) and takes infinite values on the boundary $\partial\mathcal{F}$.

We start by giving two definition.

Definition 1. The unique critical point of ϕ, i.e., the solution $\mathcal{X}(\mu)$ of the system

$$\frac{-\nabla f(x)}{\mu} - \sum_{\alpha=1}^{m} \frac{\nabla a_\alpha(x)}{-a_\alpha(x)} = 0,$$

is called μ-center.

By the implicit function theorem the preceeding system defines the curve $x(\mu)$ of class C^1.

Definition 2. The set of all μ-centers, when μ runs from ∞ to 0 is called the primal central path.

In the logarithmic barrier method, the original constrained problem (P) is replaced by a sequence of unconstrained minimization problems, i.e. minimizing $\phi(x, \mu)$ succesively for a sequence of positive decreasing values of the barrier parameter μ.

We introduce a geodesic Newton algorithm doing a search along the geodesic tangent to Newton vector $N_\phi = -(\nabla^2\phi)^{-1}d\phi = -\operatorname{grad}\phi$. By hypothesis, the norm of a vector $X_x \in T_xM$ is built using the Riemannian metric $\nabla^2\phi$ i.e.

$$||X_x||_{\nabla^2\phi} = (\nabla^2\phi(x)(X_x, X_x))^{\frac{1}{2}}.$$

We stop the minimizing procedure if $||N_\phi||_{\nabla^2\phi} \le \tau < 1$, where τ is a certain tolerance. Obviously, $||N_\phi||_{\nabla^2\phi} = 0$ iff $x = x(\mu)$.

For finding an ε-optimal solution, we proceed as follows

Logarithmic barrier algorithm

Data: ε is the accuracy parameter, τ is the proximity parameter, $\theta \in (0, 1)$ is the reduction parameter, μ_0 is the initial barrier value, x_0 is a given interior feasible point such that $||N_\phi(x_0, \mu_0)||_{\nabla^2\phi(x_0,\mu_0)} \le \tau$.

Step 1: Set $x = x_0$, $\mu = \mu_0$.

Step 2: If $\mu > \frac{\varepsilon}{4m}$, then set $\bar{\mu} = (1 - \theta)\mu$.

Step 3: While $||N_\phi(x, \mu)||_{\nabla^2\phi(x,\mu)} \ge \tau$, do

$$\bar{t} = \operatorname*{argmin}_{t>0}\{\phi(\gamma(t), \mu) | \gamma(t) \in \mathcal{F}^\circ,\ \gamma(t) = \exp_x(tN_\phi)\}$$

and set $x = \gamma(\bar{t})$.

3 Properties near the central path of a convex program

We introduce the self-concordance Riemannian condition.

Definition 3. The function $\phi : \mathcal{F}^\circ \to \mathbf{R}$ is called k-self concordant on $\mathcal{F}^\circ$, $k \ge 0$ if ϕ is of class C^3 on $\mathcal{F}^\circ$ and for all $x \in \mathcal{F}^\circ$ and $X_x \in T_xM$ the next inequality is true

$$(\nabla^3\phi(x)(X_x, X_x, X_x))^2 \le 2k(\nabla^2\phi(x)(X_x, X_x))^3,$$

where $\nabla^2\phi$ is the second covariant derivative of ϕ and $\nabla^3\phi$ is the third covariant derivative of ϕ.

In this section we assume that ϕ is k-self-concordant with $k \geq 1$. Since $\mathcal{F}^\circ$ is bounded and the logarithmic barrier function ϕ is self-concordant it follows that $\nabla^2\phi$ is positive definite and hence ϕ is strict convex.

The next analysis refers to logarithmic barrier algorithm for the general case $\tau = \frac{1}{\lambda k}$, where λ is a real constant, $\lambda \geq 1$.

For the proofs of the next two teorems see [10].

Theorem 4. *Let $x \in \mathcal{F}^\circ$ and $X \in T_xM$. Let us consider the geodesic $\gamma(t) = \exp_x(tX)$, $t \in [0,1]$ determined by $\gamma(0) = x$, $\dot\gamma(0) = X$.*

1) For an arbitrary vector field $Y \in \mathcal{X}(M)$, the next inequality holds

$$(1 - tk||X||_{\nabla^2\phi(x)}) \cdot ||Y||_{\nabla^2\phi(x)} \leq ||Y||_{\nabla^2\phi(\gamma(t))} \leq$$

$$\leq \frac{1}{1 - tk||X||_{\nabla^2\phi(x)}} \cdot ||Y||_{\nabla^2\phi(x)}.$$

2) If $||X||_{\nabla^2\phi(x)} < \frac{1}{\lambda k}$, where $\lambda \geq 1$ is a real constant, then $\gamma(1) \in \mathcal{F}^\circ$.

Theorem 5. *Let $N_\phi = N_\phi(x, \mu)$ be the Newton vector at x and $\gamma(t) = \exp_x(tN_\phi)$, $t \in [0,1]$ be the corresponding geodesic. If $||N_\phi(x)||_{\nabla^2\phi(x)} < \frac{1}{\lambda k}$, where $\lambda \geq 1$ is a real constant, then*

$$||N_\phi(\gamma(1), \mu)||_{\nabla^2\phi(\gamma(1),\mu)} \leq \frac{k \cdot ||N_\phi(x,\mu)||^2_{\nabla^2\phi(x,\mu)}}{(1 - k||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)})^2}.$$

Let us denote $\lambda_0 = \frac{3 + \sqrt{5}}{2}$.

Corollary 6. *1) If $\lambda \geq \lambda_0$, then the condition $||N_\phi(x)||_{\nabla^2\phi(x)} < \frac{1}{\lambda k}$ implies that the geodesic Newton method is convergent.*

2) If $||N_\phi(x)||_{\nabla^2\phi(x)} \leq \frac{1}{\lambda k}$, with $\lambda \geq \lambda_0$, then

$$||N_\phi(\gamma(1))||_{\nabla^2\phi(\gamma(1))} \leq k \cdot \left(\frac{\lambda}{\lambda - 1}\right)^2 \cdot ||N_\phi(x)||^2_{\nabla^2\phi(x)}.$$

Proof. 1) From $||N_\phi(x)||_{\nabla^2\phi(x)} < \frac{1}{\lambda k}$, we obtain

$$\frac{k}{(1 - k||N_\phi(x)||_{\nabla^2\phi(x)})^2} \leq k\lambda \cdot \frac{\lambda}{(\lambda - 1)^2}.$$

The last theorem implies

$$||N_\phi(\gamma(1))||_{\nabla^2\phi(\gamma(1))} \leq \frac{k \cdot ||N_\phi(x)||^2_{\nabla^2\phi(x)}}{(1 - k||N_\phi(x)||_{\nabla^2\phi(x)})^2} \leq$$

$$\leq k\lambda \cdot \frac{\lambda}{(\lambda - 1)^2} \cdot ||N_\phi(x)||^2_{\nabla^2\phi(x)} \leq \frac{\lambda}{(\lambda - 1)^2} \cdot ||N_\phi(x)||_{\nabla^2\phi(x)}.$$

From the condition $\lambda \geq \lambda_0$, it follows

$$\frac{\lambda}{(\lambda-1)^2} \cdot ||N_\phi(x)||_{\nabla^2\phi(x)} \leq ||N_\phi(x)||_{\nabla^2\phi(x)},$$

hence the geodesic Newton method is convergent.

2) In the first part of corollary we deduced

$$||N_\phi(\gamma(1))||_{\nabla^2\phi(\gamma(1))} \leq k \cdot \left(\frac{\lambda}{\lambda-1}\right)^2 \cdot ||N_\phi(x)||^2_{\nabla^2\phi(x)}, \quad \text{for} \quad \lambda \geq \lambda_0.$$

Theorem 7. *If* $||N_\phi(x)||_{\nabla^2\phi(x)} \leq \dfrac{1}{\lambda k}$, $\lambda \geq \lambda_0$ *and* x *is an approximation of the exact center* $x(\mu)$*, then*

$$\phi(x,\mu) - \phi(x(\mu),\mu) \leq \frac{||N_\phi(x)||^2_{\nabla^2\phi(x)}}{\left(1 - \left(\frac{\lambda}{\lambda-1}\right)^2 \cdot k \cdot ||N_\phi(x)||_{\nabla^2\phi(x)}\right)^2}.$$

Proof. The convexity of the function $x \to \phi(x,\mu)$ along $\gamma(t) = \exp_x(tN_\phi)$, $t \in [0,1]$ implies $\phi(\gamma(1),\mu) \geq \phi(x,\mu) + d\phi(N_\phi)$.

But $N_\phi = -(\nabla^2\phi)^{-1}d\phi$ and we deduce $d\phi(N_\phi) = -||d\phi||^2_{\nabla^2\phi}$, so that

$$\phi(x,\mu) - \phi(\gamma(1),\mu) \leq -d\phi(N_\phi) = ||d\phi||^2_{\nabla^2\phi} = ||N_\phi||^2_{\nabla^2\phi}.$$

Let $\{x_i\}$ be the Newton sequence having the first term x_1. From the corollary 6, if $||N_\phi||_{\nabla^2\phi} \leq \dfrac{1}{\lambda k}$, $\lambda \geq \lambda_0$, then

$$||N_\phi(\gamma(1))||_{\nabla^2\phi(\gamma(1))} \leq k \cdot \frac{\lambda}{\lambda-1} \cdot ||N_\phi(x)||^2_{\nabla^2\phi(x)}.$$

Hence

$$||N_\phi(x_i)||_{\nabla^2\phi(x_i)} \leq \left[\left(\frac{\lambda}{\lambda-1}\right)^2 \cdot k\right]^{2^i-1} \cdot ||N_\phi(x_1)||^{2^i}_{\nabla^2\phi(x_1)} =$$

$$= \left[\left(\frac{\lambda}{\lambda-1}\right)^2 \cdot k\right]^{2^i-1} \cdot ||N_\phi||^{2^i}_{\nabla^2\phi} \leq \left[\left(\frac{\lambda}{\lambda-1}\right)^2 \cdot k\right]^{2^i-1} \cdot \left(\frac{1}{\lambda k}\right)^{2^i} =$$

$$= \frac{1}{\lambda k}\left[\frac{\lambda}{(\lambda-1)^2}\right]^{2^i-1}.$$

For $\lambda \geq \lambda_0$, we find $0 < \dfrac{\lambda}{(\lambda-1)^2} < 1$ and $\lim_{i\to\infty} ||N_\phi(x_i)||_{\nabla^2\phi(x_i)} = 0$. The terms of the sequence $\{x_i\}$ lie in a bounded region and $||N_\phi||_{\nabla^2\phi} = 0$ iff $\nabla\phi = 0$ iff $x = x(\mu)$. Hence $\{x(\mu)\}$ is the set of all limit points of the sequence $\{x_i\}$, i.e. the Newton sequence converges to $x(\mu)$.

Moreover, using the inequality $\phi(x,\mu) - \phi(\gamma(1),\mu) \leq ||d\phi||^2_{\nabla^2\phi}$ we can write

$$\phi(x,\mu)-\phi(x(\mu),\mu)=\sum_{i=1}^{\infty}[\phi(x_i,\mu)-\phi(x_{i+1},\mu)]\leq$$

$$\leq\sum_{i=1}^{\infty}||N_\phi(x_i)||^2_{\nabla^2\phi(x_i)}\leq\sum_{i=1}^{\infty}\left[\left(\frac{\lambda}{\lambda-1}\right)^2\cdot k\right]^{2^i-1}\cdot||N_\phi(x)||^{2^i}_{\nabla^2\phi(x)}\leq$$

$$\leq\frac{||N_\phi(x)||^2_{\nabla^2\phi(x)}}{1-\left[\left(\frac{\lambda}{\lambda-1}\right)^2\cdot k\cdot||N_\phi(x)||_{\nabla^2\phi(x)}\right]^2}.$$

Lemma 8. *If the real numbers $a_1,\dots,a_m$ satisfies the inequality $a_1^2+\dots+a_m^2\leq r^2$, then $-r\sqrt{m}\leq a_1^2+\dots+a_m^2\leq r\sqrt{m}$.*

The equality is attained only for

$$a_1=\dots=a_m=\frac{-r}{\sqrt{m}}\quad\text{or}\quad a_1=\dots=a_m=\frac{r}{\sqrt{m}}.$$

Proof. We apply the Cauchy-Schwartz inequality:

$$(a_1+\dots+a_m)^2=(a_1\cdot1+\dots+a_m\cdot1)^2\leq(a_1^2+\dots+a_m^2)\cdot m\leq mr^2.$$

Therefore, the extrema values are $\pm r\sqrt{m}$ and they are obtained for $a_1=\dots=a_m=\pm\frac{r}{\sqrt{m}}$.

Lemma 9. *If x is an approximation of the exact center $x(\mu)$, then*

$$df(x)(N_\phi)\leq\mu\left(||N_\phi||^2_{\nabla^2\phi}+\sqrt{m}\cdot||N_\phi||_{\nabla^2\phi}\right).$$

Proof. We use again that $d\phi(N_\phi)=-||d\phi||^2_{\nabla^2\phi}$. From the formulas for $\nabla\phi$ and $\nabla^2\phi$, we get

$$\nabla\phi(x,\mu)=-\frac{df(x)(N_\phi)}{\mu}+\sum_{i=1}^{m}\frac{da_i(x)}{-a_i(x)}(N_\phi),$$

or equivalent

$$\frac{df(x)(N_\phi)}{\mu}=||d\phi||^2_{\nabla^2\phi}+\sum_{i=1}^{m}\frac{da_i(x)}{-a_i(x)}(N_\phi).$$

Also, we find

$$||N_\phi||^2_{\nabla^2\phi}=\nabla^2\phi(x,\mu)(N_\phi,N_\phi)=$$

$$=-\frac{\nabla^2f(x)(N_\phi,N_\phi)}{\mu}+\sum_{i=1}^{m}\left[\frac{\nabla^2a_i(x)(N_\phi,N_\phi)}{-a_i(x)}+\left(\frac{da_i(x)(N_\phi)}{a_i(x)}\right)^2\right].$$

The functions $-f$ and a_i, $i=\overline{1,m}$ are Riemannian convex. Hence

$$\nabla^2f(x)(N_\phi,N_\phi)\leq0,\quad\nabla^2a_i(x)(N_\phi,N_\phi)\geq0.$$

Having in mind that $a_i(x)\leq0$, we deduce

$$||N_\phi||^2_{\nabla^2\phi} \geq \sum_{i=1}^{m} \left(\frac{da_i(x)(N_\phi)}{a_i(x)} \right)^2$$

and consequently from Lemma 8,

$$df(x)(N_\phi) \leq \mu \left(||N_\phi||^2_{\nabla^2\phi} + \sqrt{m} \cdot ||N_\phi||_{\nabla^2\phi} \right).$$

Lemma 10. *If x is an approximation of the exact center $x(\mu)$ and $\gamma : [0,1] \to M$, $\gamma(t) = \exp_x(tN_\phi)$, then*

$$|df(\gamma(1)(N_\phi)| \leq \frac{\mu k \cdot ||N_\phi||_{\nabla^2\phi}(||N_\phi||_{\nabla^2\phi} + \sqrt{m})}{1 - k \cdot ||N_\phi||^2_{\nabla^2\phi}}.$$

Proof. Like in Lemma 9, we find

$$\frac{df(\gamma(1))(N_\phi)}{\mu} = -d\phi(\gamma(1))(N_\phi) + \sum_{i=1}^{m} \frac{da_i(\gamma(1))}{-a_i(\gamma(1))}(N_\phi).$$

From the proof of Lemma 9, it follows

$$\left| \sum_{i=1}^{m} \frac{da_i(\gamma(1))}{-a_i(\gamma(1))}(N_\phi) \right| \leq \sqrt{m} \cdot ||N_\phi||_{\nabla^2(\gamma(1))}.$$

In the first part of the Theorem 4 we take $X = Y = N_\phi$ and find

$$||N_\phi||_{\nabla^2\phi(t)} \leq \frac{1}{1 - tk||N_\phi||_{\nabla^2\phi(x)}} \cdot ||N_\phi||_{\nabla^2\phi(x)} \leq \frac{k \cdot ||N_\phi||_{\nabla^2\phi(x)}}{1 - tk||N_\phi||_{\nabla^2\phi(x)}},$$

since $k \geq 1$. For $t = 1$, it follows

$$||N_\phi||_{\nabla^2\phi(1)} \leq \frac{k||N_\phi||_{\nabla^2\phi(x)}}{1 - k||N_\phi||_{\nabla^2\phi(x)}}.$$

Therefore

$$\left| \sum_{i=1}^{m} \frac{\nabla a_i(\gamma(1))}{-a_i(\gamma(1))}(N_\phi) \right| \leq \sqrt{m} \cdot ||N_\phi||_{\nabla^2(\gamma(1))} \leq \sqrt{m} \cdot \frac{k||N_\phi||_{\nabla^2\phi(x)}}{1 - k||N_\phi||_{\nabla^2\phi(x)}}.$$

From the proof of the Theorem 5 we have

$$|df(\gamma(1))(N_\phi)| \leq \frac{k||N_\phi||^2_{\nabla^2\phi(x)}}{1 - k||N_\phi||_{\nabla^2\phi(x)}}.$$

Finally we conclude

$$|df(\gamma(1))(N_\phi)| \leq \mu \left[|df(\gamma(1))(N_\phi)| + \left| \sum_{i=1}^{m} \frac{da_i(\gamma(1))}{-a_i(\gamma(1))}(N_\phi) \right| \right] \leq$$

$$\leq \mu \cdot \left[\frac{k||N_\phi||^2_{\nabla^2\phi(x)}}{1 - k||N_\phi||_{\nabla^2\phi(x)}} + \frac{\sqrt{m} \cdot k \cdot ||N_\phi||_{\nabla^2\phi(x)}}{1 - k||N_\phi||_{\nabla^2\phi(x)}} \right] =$$

$$= \frac{\mu k||N_\phi||_{\nabla^2\phi(x)}(k||N_\phi||_{\nabla^2\phi(x)} + \sqrt{m})}{1 - k||N_\phi||_{\nabla^2\phi(x)}}.$$

The following theorem gives an upper bound for the difference of the objective value in the exact center and an approximately centered iterate.

Theorem 11. *If x is an approximation of the exact center $x(\mu)$ and $||N_\phi||_{\nabla^2\phi(x)} \leq \frac{1}{\lambda k}$, $\lambda \geq \lambda_0$, then*

$$|f(x) - f(x,\mu)| \leq \mu\sqrt{m} \cdot \frac{1 + k||N_\phi(x,\mu)||^2_{\nabla^2\phi(x,\mu)}}{1 - k||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)}} \cdot \frac{||N_\phi(x,\mu)||^2_{\nabla^2\phi(x,\mu)}}{1 - \left(\frac{\lambda}{\lambda - 1}\right)^2 \cdot k \cdot ||N_\phi||_{\nabla^2\phi(x,\mu)}}.$$

Proof. Let us consider the geodesic $\gamma : [0,1] \to M$, $\gamma(t) = \exp_x(tN_\phi)$. The convexity of $-f(x)$ implies

$$|df(\gamma(1))(N_\phi)| \leq f(\gamma(1)) - f(x) \leq df(x)(N_\phi).$$

Using the results from the Lemmas 9 and 10 we obtain

$$|f(\gamma(1)) - f(x)| \leq \mu\sqrt{m} \cdot ||N_\phi||_{\nabla^2\phi(x)} \cdot \frac{1 + k||N_\phi||^2_{\nabla^2\phi}}{1 - k||N_\phi||_{\nabla^2\phi}}.$$

Let $\{x_i\}$ be the Newton sequence starting at x.

$$|f(\gamma(1)) - f(x)| = \left| \sum_{i=1}^{\infty} f(x_{i+1} - f(x_i) \right| \leq \sum_{i=1}^{\infty} |f(x_{i+1} - f(x_i)| \leq$$

$$\leq \sum_{i=1}^{\infty} \mu\sqrt{m} \cdot ||N_\phi(x_i,\mu)||_{\nabla^2\phi(x_i,\mu)} \cdot \frac{1 + k||N_\phi(x_i,\mu)||^2_{\nabla^2\phi(x_i,\mu)}}{1 - k||N_\phi(x_i,\mu)||_{\nabla^2\phi(x_i,\mu)}} \leq$$

$$\leq \mu\sqrt{m} \cdot \frac{1 + k||N_\phi(x,\mu)||^2_{\nabla^2\phi(x,\mu)}}{1 - k||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)}} \cdot \sum_{i=1}^{\infty} ||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)} \leq$$

$$\leq \mu\sqrt{m} \cdot \frac{1 + k||N_\phi(x,\mu)||^2_{\nabla^2\phi(x,\mu)}}{1 - k||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)}} \cdot \sum_{i=1}^{\infty} \left[\left(\frac{\lambda}{\lambda - 1}\right)^2 \cdot k \right]^{2^i - 1} \cdot ||N_\phi(x,\mu)||^{2^i}_{\nabla^2\phi(x,\mu)} \leq$$

$$\leq \mu\sqrt{m} \cdot \frac{1 + k||N_\phi(x,\mu)||^2_{\nabla^2\phi(x,\mu)}}{1 - k||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)}} \cdot \frac{||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)}}{1 - \left(\frac{\lambda}{\lambda - 1}\right)^2 \cdot k \cdot ||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)}}.$$

4 Complexity analysis of the logarithmic barrier algorithm

Let us refer to upper bounds for the total number of outer and inner iterations for the logarithmic barrier algortihm (LBA) applied to the program (P). For the analysis we consider the step $\tau = \frac{1}{\lambda k}$, $\lambda \geq \lambda_0$.

Theorem 12. *After at most* $\frac{1}{\theta + \frac{\theta^2}{2} + \ldots + \frac{\theta^s}{2}} \cdot \ln \frac{4m\mu_0}{\varepsilon}$ *($s \in \mathbf{N}$ fixed) outer iterations and $\lambda \geq 2,9$ the LBA ends up with an approximate solution x of the program (P) such that $z^* - f(x) \leq \varepsilon$, where $x^*, z^* = f(x^*)$ is the solution of the problem (P).*

Proof. The algorithm stops when $\mu_p = (1-\theta)^p \cdot \mu_0 \leq \frac{\varepsilon}{4m}$ or $-p\ln(1-\theta) \geq \ln \frac{4m\mu_0}{\varepsilon}$. Since $\theta + \frac{\theta^2}{2} + \ldots + \frac{\theta^s}{s} \leq -\ln(1-\theta)$, we find

$$p \geq \frac{1}{\theta + \frac{\theta^2}{2} + \ldots + \frac{\theta^s}{s}} \cdot \ln \frac{4m\mu_0}{\varepsilon}.$$

Also we have (see [2]) $z^* - f(x(\mu_p)) \leq m \cdot \mu_p$. From theorem 11, for $||N_\phi||_{\nabla^2\phi(x)} \leq \frac{1}{\lambda k}$, $\lambda \geq \lambda_0$ we obtain

$$f(x(\mu_p)) - f(x) \leq \mu_p\sqrt{m} \cdot \frac{1 + k \cdot \left(\frac{1}{\lambda k}\right)^2}{1 - k \cdot \frac{1}{\lambda k}} = \frac{\frac{1}{\lambda k}}{1 - \left(\frac{1}{\lambda - 1}\right)^2}$$

or equivalent

$$f(x(\mu_p)) - f(x) \leq \mu_p\sqrt{m} \cdot \left(\frac{1}{k} + \frac{1}{\lambda^2 k^2}\right) \cdot \frac{\lambda - 1}{\lambda^2 - 3\lambda + 1}.$$

Since $k \geq 1$, it follows $\frac{1}{k} + \frac{1}{\lambda^2 k^2} \leq 1 + \frac{1}{\lambda^2}$ and hence

$$f(x(\mu_p)) - f(x) \leq \mu_p\sqrt{m} \cdot \frac{(\lambda^2+1)(\lambda-1)}{\lambda^2(\lambda^2 - 3\lambda + 1)}.$$

The real function $a : (\lambda_0, \infty) \to \mathbf{R}$, $a(t) = \frac{(t^2+1)(t-1)}{t^2(t^2-3t+1)}$ has the derivative $a'(t) = \frac{(t-2)(-t^4-5t^2-10)-18}{(t^4-3t^3+t^2)^2}$, which is strictly negative for $t > \lambda_0 = = \frac{3+\sqrt{5}}{2}$.

Therefore the function a decreases from $+\infty$ to 0 on (λ_0, ∞). But $a(2,9) = 2,9942556 < 3$, hence for all $t \geq 2,9$, $a(t) \leq a(2,9) < 3$ and $f(x(\mu_p)) - f(x) < 3\mu_p\sqrt{m}$.

Finally we deduce that

$$z^* - f(x) = z^* - f(x(\mu_p)) + f(x(\mu_p)) - f(x) \leq \mu_p m + 3\mu_p\sqrt{m} \leq$$

$$\leq \mu_p \cdot 4m \leq \frac{\varepsilon}{4m} \cdot 4m = \varepsilon.$$

The next theorem enssures a sufficient decrease of the value of logarithmic barrier function.

Theorem 13. *If* $\gamma(t) = \exp_x(tN_\phi)$, $t \in [0,1]$ *and* $\bar{t} = \dfrac{1}{1 + t||N_\phi||_{\nabla^2\phi}}$, *then*

$$\phi(x,\mu) - \phi(\gamma(\bar{t}),\mu) \geq \frac{1}{k^2}\left(k||N_\phi||_{\nabla^2\phi} - \ln(1 + k||N_\phi||_{\nabla^2\phi})\right).$$

The proof of this theorem is presented in [10].

Theorem 14. *Each outer iteration requires at most*

$$\frac{k^2}{\frac{1}{2\lambda^2} - \frac{1}{3\lambda^3} - \frac{1}{4\lambda^4} - \frac{1}{5\lambda^5}} \cdot \left\{ \frac{(\lambda-1)^4}{\lambda^2 k^2 (\lambda^2 - 3\lambda + 1)(\lambda^2 - \lambda + 1)} + \right.$$
$$\left. + \frac{\theta}{(1-\theta)^2} \cdot \left[\frac{\sqrt{m}(\lambda^2 k + 1)(\lambda - 1)}{\lambda^2 k^2 (\lambda^2 - 3\lambda + 1)} + \theta m \right] \right\}$$

inner iterations, $\lambda > \lambda_0$.

Proof. We use the notations: $\bar{\bar{\mu}}$ = the barrier parameter value in an arbitrary outer iteration, $\bar{\mu}$ = the barrier parameter value in the previous outer iteration, $x =$ the iterate at the beginning of the outer iteration. Therefore x is centered with respect to $x(\bar{\mu})$ and $\bar{\bar{\mu}} = (1-\theta)\bar{\mu}$.

By Theorem 13, at each inner iteration, the decrease of the barrier fuction is at least

$$\frac{1}{k^2}[k||N_\phi||_{\nabla^2\phi} - \ln(1 + k||N_\phi||_{\nabla^2\phi})].$$

The last function is increasing in $||N_\phi||_{\nabla^2\phi}$ and $||N_\phi||_{\nabla^2\phi} \geq \dfrac{1}{\lambda k}$, during each iteration, hence

$$\frac{1}{k^2}[k||N_\phi||_{\nabla^2\phi} - \ln(1 + k||N_\phi||_{\nabla^2\phi})] \geq \frac{1}{k^2}\left[\frac{1}{\lambda} - \ln\left(1 + \frac{1}{\lambda}\right)\right].$$

We use the inequality $\ln(1+t) < t - \dfrac{t^2}{2} + \dfrac{t^3}{3} - \dfrac{t^4}{4} + \dfrac{t^5}{5}$, which is true for all $t > 0$ and deduce that

$$\frac{1}{\lambda} - \ln\left(1 + \frac{1}{\lambda}\right) > \frac{1}{2\lambda^2} - \frac{1}{3\lambda^3} + \frac{1}{4\lambda^4} - \frac{1}{5\lambda^5}.$$

Let's denote $E = \dfrac{1}{2\lambda^2} - \dfrac{1}{3\lambda^3} + \dfrac{1}{4\lambda^4} - \dfrac{1}{5\lambda^5}$. If N is the number of inner iterations during one outer iteration, then $\dfrac{NE}{k^2} \leq \phi(x,\bar{\bar{\mu}}) - \phi(x(\bar{\bar{\mu}}),\bar{\bar{\mu}})$.

We also consider the function $\psi(x,\bar{\bar{\mu}}) = \phi(x,\bar{\bar{\mu}}) - \phi(x(\bar{\bar{\mu}}),\bar{\bar{\mu}})$ and apply the Mean-value theorem:

$$\psi(x,\bar{\bar{\mu}}) = \psi(x,\bar{\mu}) + \frac{d\psi}{d\mu}(x,\mu)|_{\mu=\hat{\mu}}(\bar{\bar{\mu}} - \bar{\mu}),$$

where $\hat{\mu} \in (\bar{\bar{\mu}},\bar{\mu})$. From $\phi(x,\mu) = -\dfrac{f(x)}{\mu} - \sum_{i=1}^{m} \ln(-a_i(x))$ it follows $\dfrac{d\phi}{d\mu}(x,\mu) = \dfrac{f(x)}{\mu^2}$ and

$$\frac{d\phi}{d\mu}(x(\mu),\mu) = \frac{f(x(\mu))}{\mu^2} - \frac{\nabla(x(\mu)) \cdot x'(\mu)}{\mu} + \sum_{i=1}^{m} \frac{\nabla a_i(x(\mu)) \cdot x'(\mu)}{-a_i(x(\mu))}.$$

But $x(\mu)$ satisfies the Kuhn-Tucker conditions:

$$\begin{cases} a_i(x) \le 0, \quad \forall i = \overline{1,m} \\ \nabla f(x) - \sum_{i=1}^{m} u_i \nabla a_i(x) = 0, \quad u_i \ge 0 \\ -a_i(x) u_i = \mu. \end{cases}$$

It follows that

$$\frac{d\phi}{d\mu}(x(\mu),\mu) = \frac{f(x(\mu))}{\mu^2} - \frac{\nabla f(x(\mu))}{\mu} \cdot x'(\mu) + \sum_{i=1}^{m} \frac{\nabla a_i(x(\mu))}{\frac{\mu}{u_i}} \cdot x'(\mu) =$$

$$= \frac{f(x(\mu))}{\mu^2} - \frac{x'(\mu)}{\mu} \cdot \left[\nabla f(x(\mu)) - \sum_{i=1}^{m} u_i \nabla a_i(x(\mu))\right] = \frac{f(x(\mu))}{\mu^2}.$$

Hence

$$-\left.\frac{d\psi}{d\mu}(x,\mu)\right|_{\bar{\bar{\mu}}=\hat{\mu}} = \left.\frac{f(x(\mu)) - f(x)}{\mu^2}\right|_{\bar{\bar{\mu}}=\hat{\mu}} \le \frac{|f(x(\bar{\bar{\mu}})) - f(x)|}{\bar{\bar{\mu}}^2},$$

the last inequality beig a consequence of the monotonicity of the objective function along the central path. We may write

$$\psi(x,\bar{\bar{\mu}}) \le \psi(x,\bar{\mu}) + \frac{|f(x(\bar{\bar{\mu}})) - f(x)|}{\bar{\bar{\mu}}^2} \cdot (\bar{\mu} - \bar{\bar{\mu}}) \le$$

$$\le \psi(x,\bar{\mu}) + \frac{|f(x(\bar{\mu})) - f(x)|}{\bar{\bar{\mu}}} + \frac{|f(x(\bar{\bar{\mu}})) - f(x(\bar{\mu}))|}{\bar{\bar{\mu}}} \cdot \frac{\bar{\mu} - \bar{\bar{\mu}}}{\bar{\bar{\mu}}}.$$

The function $h_1 : \left(0, \frac{1}{\lambda_0 k}\right] \to (0,\infty)$, $h_1(t) = \dfrac{t^2}{1 - \left[\left(\frac{\lambda}{\lambda-1}\right)^2 kt\right]^2}$ has the derivative $h_1'(t) = \dfrac{2t}{\left\{1 - \left[\left(\frac{\lambda}{\lambda-1}\right)^2 kt\right]^2\right\}^2}$, which is strictly positive for all $t > 0$. Therefore h_1 is increasing for $t \in \left(0, \frac{1}{\lambda_0 k}\right]$. From the Theorem 7 it follows that for $||N_\phi(x,\bar{\mu})||_{\nabla^2\phi(x,\mu)} \le \frac{1}{\lambda k}$, $\lambda \ge \lambda_0$, we have

$$\psi(x,\bar{\mu}) = \phi(x,\bar{\mu}) - \phi(x(\bar{\mu}),\bar{\mu}) \le \frac{||N_\phi(x)||^2_{\nabla^2\phi(x)}}{1 - \left[\left(\frac{\lambda}{\lambda-1}\right)^2 k||N_\phi(x)||_{\nabla^2\phi(x)}\right]^2} \le$$

$$\leq \frac{\frac{1}{\lambda^2 k^2}}{1-\left[\left(\frac{\lambda}{\lambda-1}\right)^2 \cdot k \cdot \frac{1}{\lambda k}\right]^2} = \frac{(\lambda-1)^4}{\lambda^2 k^2(\lambda^2-3\lambda+1)(\lambda^2-\lambda+1)},$$

the last expression make sense if $\lambda \neq \lambda_0$. So we have to impose the condition $\lambda > \lambda_0$.

The functions $h_2, h_3 : \left(0, \frac{1}{\lambda_0 k}\right] \to (0,\infty)$, $h_2(t) = \frac{1+kt^2}{1-kt}$ and $h_3(t) = \frac{t}{1-\left(\frac{\lambda}{\lambda-1}\right)^2 kt}$ have the derivatives $h_2'(t) = \frac{k(-kt^2+2t+1)}{(1-kt)^2}$, $h_3'(t) = \frac{1}{\left[1-\left(\frac{\lambda}{\lambda-1}\right)^2 kt\right]^2}$, which are positive for $t \in \left(0, \frac{1}{\lambda_0 k}\right]$. Therefore the functions h_2 and h_3 are increasing and positive, hence $h_2 \cdot h_3$ is increasing. From the Theorem 11 it follows that for $||N_\phi(x,\bar{\mu})||_{\nabla^2\phi(x,\mu)} \leq \frac{1}{\lambda k}$, $\lambda > \lambda_0$, we have

$$|f(x) - f(x,\bar{\mu})| \leq \bar{\mu}\sqrt{m} \cdot \frac{1+k||N_\phi(x,\mu)||^2_{\nabla^2\phi(x,\mu)}}{1-k||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)}} \cdot \frac{||N_\phi(x,\mu)||_{\nabla^2\phi(x,\mu)}}{1-\left(\frac{\lambda}{\lambda-1}\right)^2 k} \leq$$

$$\leq \bar{\mu}\sqrt{m} \cdot \frac{(\lambda^2 k+1)(\lambda-1)}{\lambda^2 k^2(\lambda^2-3\lambda+1)}.$$

Moreover,

$$f(x(\bar{\bar{\mu}})) - f(x(\bar{\mu})) \leq \sum_{i=1}^m u_i(\bar{\bar{\mu}}) \cdot a_i(x(\bar{\bar{\mu}})) - \sum_{i=1}^m u_i(\bar{\mu}) a_i(x(\bar{\mu})) =$$
$$= m(\bar{\mu} - \bar{\bar{\mu}}) = \theta m \bar{\mu},$$

since $\bar{\bar{\mu}} = (1-\theta)\bar{\mu}$. Finally we obtain

$$\psi(x,\bar{\bar{\mu}}) \leq \frac{(\lambda-1)^4}{\lambda^2 k^2(\lambda^2-3\lambda+1)(\lambda^2-\lambda+1)} +$$
$$+\left[\frac{\bar{\mu}\sqrt{m}}{\bar{\bar{\mu}}} \cdot \frac{(\lambda^2 k+1)(\lambda-1)}{\lambda^2 k^2(\lambda^2-3\lambda+1)} + \frac{\theta m \bar{\mu}}{\bar{\bar{\mu}}}\right] \cdot \frac{\bar{\mu}-\bar{\bar{\mu}}}{\bar{\bar{\mu}}} =$$
$$= \frac{(\lambda-1)^4}{\lambda^2 k^2(\lambda^2-3\lambda+1)(\lambda^2-\lambda+1)} + \frac{\theta}{(1-\theta)^2} \cdot \left[\frac{\sqrt{m}(\lambda^2 k+1)(\lambda-1)}{\lambda^2 k^2(\lambda^2-3\lambda+1)} + \theta m\right].$$

We denote the last expression by F. We may write that $\frac{NE}{k^2} \leq \psi(x,\bar{\bar{\mu}}) \leq F$, or equivalent, $N \leq \frac{k^2 F}{E}$, since $E > 0$ for $\lambda > \lambda_0$. Hence

$$N \leq k^2 \cdot \frac{1}{\frac{1}{2\lambda^2} - \frac{1}{3\lambda^3} - \frac{1}{4\lambda^4} - \frac{1}{5\lambda^5}} \cdot \left\{\frac{(\lambda-1)^4}{\lambda^2 k^2(\lambda^2-3\lambda+1)(\lambda^2-\lambda+1)} + \right.$$
$$\left. + \frac{\theta}{(1-\theta)^2} \cdot \left[\frac{\sqrt{m}(\lambda^2 k+1)(\lambda-1)}{\lambda^2 k^2(\lambda^2-3\lambda+1)} + \theta m\right]\right\}$$

Corolary 16. *An upper bound for the total number of Newton iterations is given by*

$$k^2 \cdot \frac{1}{\frac{1}{2\lambda^2} - \frac{1}{3\lambda^3} - \frac{1}{4\lambda^4} - \frac{1}{5\lambda^5}} \cdot \left\{ \frac{(\lambda-1)^4}{\lambda^2 k^2 (\lambda^2 - 3\lambda + 1)(\lambda^2 - \lambda + 1)} + \right.$$
$$\left. + \frac{\theta}{(1-\theta)^2} \cdot \left[\frac{\sqrt{m}(\lambda^2 k + 1)(\lambda - 1)}{\lambda^2 k^2 (\lambda^2 - 3\lambda + 1)} + \theta m \right] \right\} \cdot$$
$$\cdot \frac{1}{\theta + \frac{\theta^2}{2} + \ldots + \frac{\theta^s}{s}} \cdot \ln \frac{4m\mu_0}{\varepsilon},$$

where $s \in \mathbf{N}$ *is fixed and* $\lambda \in \mathbf{R}$, $\lambda \geq 2,9$.

Acknowledgement. The present work is developed within the Grant CNCSIS MEN 1107/2005.

References

[1] J.X. da Cruz Neto, O.P. Ferreira, L.R. Lucambio Perez, S.Z. Németh, *Convex and Monotone Transformable Mathematical Programming Problems and a Proximal Like Point Method*, JOGO, 2003, 11-32.

[2] D. den Hertog, *Interior point aproach to linear quadratic and convex programming*, Mathematics and Its Applications, 277, Kluwer Academic Publishers, 1994.

[3] R.K. Frisch, *The logarithmic potential method for solving linear programming problems*, Memorandum, 1955, Institute of Economics, Oslo, Norway.

[4] S. Mehrota, J. Sun, *An interior point based on Newton method*, Mathematical Developments Arising from Linear Programming, J.C. Logarias, M. J. Todd, eds., Contemporary Mathematics 114 (1990), 184-265.

[5] Y. Nesterov, A. Nemirovsky, *Self-concordant functions and polynomial time methods in convex programming*, Report 1989, Central Economical and Mathematical Institute, Academy of Science, Moscow, USSR.

[6] Y. Nesterov, A. Nemirovsky, *Interior point polynomial algorithms in convex programming*, Siam, 1994.

[7] Y. Nesterov, M.J. Todd, *On the Riemanian Geometry Defined by Self-Concordant barriers and Interior Point Methods*, Found. of Comput. Math., 2 (2002), 333-361.

[8] C. Udrişte, *Minimization algortihms on Riemannian manifolds*, Proceedings of the 23-rd Conference on Geometry and Topology, Cluj-Napoca, Romania, 1993, 185-193.

[9] C. Udrişte, *Convex functions and optimization methods on Riemannian manifolds*, Mathematics and Its Applications, 297, Kluwer Academic Publishers, 1994.

[10] C. Udrişte, *Optimization methods on Riemannian manifolds*, Proceedings of IRB International Workshop, Monteroduni, Italy, August 8-12, 1995; Algebras, Groups and Geometries 14 (1997), 339-359.

[11] C. Udrişte, *Riemannian Convexity in Programming (I)*, Analele Ştiinţifice ale Univ. AL.I. Cuza, Iaşi, 42, 1-a, Matematică (1996), 123-136.

[12] C. Udrişte, *Riemannian Convexity in Programming (II)*, The Second International Workshop on Differential Geometry and Its Applications, Ovidius University of Constantza, Romania, September 25-28 (1995); Balkan Journal of Geometry and Its Applications, 1,1 (1996), 99-109.

Constantin Udrişte
Department of Mathematics I, Splaiul Independenţei 313,
RO-060042 Bucharest, Romania.
email: udriste@mathem.pub.ro

Gabriel Bercu
Universitatea "Dunărea de Jos", Galaţi, Romania.
email: gbercu@ugal.ro