On the complexity of the primal self-concordant barrier method

Jan Brinkhuis

Econometrisch Instituut

Erasmus Universiteit Rotterdam

Postbus 1738

3000 DR Rotterdam

The Netherlands

E-mail: brinkhuis@few.eur.nl

Econometric Institute Report EI 2000–36/A

1 Introduction

In his Introductory Lectures on Convex Programming Nesterov has given an algorithm to find the analytic centre x_F^* for a given ν-self-concordant barrier F with bounded domain and a given interior point of this domain. The intended use of this algorithm is as an auxiliary phase in a primal short-step path-following method for solving convex programming problems. For the number of iterations in this auxiliary phase an upperbound is given in [N] which for ν much bigger than 1 is essentially

$$7.2\sqrt{\nu}\left(\ln\nu + \tfrac{1}{2}\ln F'(y_0)^T F''(x_F^*)^{-1} F'(y_0)\right)$$

where T denotes transpose.

In this note it is shown that the term $\ln\nu$ can be omitted. Moreover we make the easy observation that the constant 7.2 can be replaced by 3.2. The $\ln\nu$-improvement is achieved in the following way. Using certain inequalities from [N] we obtain a lower bound for the total decrease of the penalty parameter in the last two steps of the algorithm *which does not depend on ν*.

Concerning the constant 7.2 it is clear from [N] how it could be improved: by optimizing the choice of the centering parameter β. A routine optimization shows that $\beta \approx 0.088$ gives the constant 3.2.

2 Statement of the result

In this paper we will use notations, definitions and results from chapter 4 of [N]. We begin by recalling from [N] a scheme to approximate an analytic centre; we use a slightly different stopping criterion. Let F be a ν-self concordant barrier with bounded domain and let a point y_0 in this domain be given. Choose a centering parameter $\beta < \frac{3}{2} - \frac{1}{2}\sqrt{5} \approx 0.4$ and write $\gamma = \frac{\sqrt{\beta}}{1+\sqrt{\beta}} - \beta$. Then $\gamma > 0$. We consider the following scheme.

0. Set $t_0 = 1$

1. k-th iteration ($k \geq 0$). Set
$$t_{k+1} = \max\left(0, t_k - \tfrac{\gamma}{\|F'(y_0)\|^*_{y_k}}\right)$$
$$y_{k+1} = y_k - F''(y_k)^{-1}\left(-t_{k+1}F'(y_0) + F'(y_k)\right)$$

2. Stop the process if $t_k = 0$. Set $\overline{x} = x_k$ and $N = k$.

Theorem 2.1. *The scheme above terminates and*

$$N \leq 2 + \max\left[0, \frac{1}{\gamma}(\beta + \sqrt{\nu})\ln\left(\frac{(1+\sqrt{\beta})\|F'(y_0)\|^*_{x^*_F}}{\gamma(1-\sqrt{\beta})}\right)\right]$$

The vector $\overline{x}$ which is the result of this scheme satisfies

$$\|F'(\overline{x})\|^*_{\overline{x}} \leq \beta.$$

Remark 2.2. If ν, the parameter of the barrier, is much bigger than 1, then it is 'optimal' to choose β such that $\gamma = \gamma(\beta)$ is maximal. A routine calculation shows that this choice is $\beta \approx 0.088$, the unique real root of the equation $4x^3 - 8x^2 + 12x - 1 = 0$. Then $\gamma = 0.317$ and so the upperbound in the theorem is essentially

$$3.2\ \sqrt{\nu}\ln\|F'(y_0)\|^*_{x^*_F}.$$

3 Proof of the result

We write

$$\lambda(t,y) = [(-tF'(y_0) + F'(y))^T F''(y)^{-1}(-tF'(y_0) + F'(y))]^{\frac{1}{2}}$$

for all $t \in \mathbb{R}$ and all $y \in \text{dom } F$. This is well-defined: dom F is bounded, so it contains no straight lines and so, by theorem 4.1.3. of [N] the hessian $F''(y)$ is non- degenerate for all $y \in \text{dom } F$.

Step 1 $\lambda(t_k, y_k) \leq \beta$ for all k and $\lambda(t_{k+1}, y_k) \leq \frac{\sqrt{\beta}}{1+\sqrt{\beta}}$ for all k with $t_k > 0$.
Start induction: $\lambda(t_0, y_0)$ is seen to be 0.
Induction step: assume $\lambda(t_k, y_k) \leq \beta$ for some k with $t_k > 0$. Then $\lambda(t_{k+1}, y_k)$ is by the triangle inequality

$$\leq (t_k - t_{k+1})\|F'(y_0)\|^*_{y_k} + \lambda(t_k, y_k).$$

This is $\leq \frac{\sqrt{\beta}}{1+\sqrt{\beta}}$ as $\lambda(t_k, y_k) \leq \beta$ and $t_k - t_{k+1} \leq \frac{\gamma}{\|F'(y_0)\|^*_{y_k}}$.
Applying theorem 4.1.12 of [N] we get

$$\lambda(t_{k+1}, y_{k+1}) \leq \left(\frac{\lambda(t_{k+1}, y_k)}{1 - \lambda(t_{k+1}, y_k)}\right)^2.$$

This is seen to be $\leq \beta$ as

$$\lambda(t_{k+1}, y_k) \leq \frac{\sqrt{\beta}}{1+\sqrt{\beta}}.$$

Step 2 $\|F'(y_0)\|^*_{y_k} \leq \frac{\beta+\sqrt{\nu}}{t_k}$ for all k with $t_k > 0$.
One has $t_k\|F'(y_0)\|^*_{y_k} = \| - t_k F'(y_0) + F'(y_k) - F'(y_k)\|^*_{y_k}$.
By the triangle inequality this is $\leq \lambda(t_k, y_k) + \|F'(y_k)\|^*_{y_k}$. By $\lambda(t_k, y_k) \leq \beta$ and the definition of self-concordant barriers this is $\leq \beta + \sqrt{\nu}$.

Step 3. $t_k \leq \left(1 - \frac{\gamma}{\beta+\sqrt{\nu}}\right)^k$ for all k with $t_{k+1} > 0$.
Start induction: $t_0 = 1$.
Induction step: for all k with $t_{k+1} > 0$, we have $t_k - \frac{\gamma}{\|F'(y_0)\|^*_{y_k}} > 0$.
It follows that

$$t_{k+1} \leq (1 - \frac{\gamma}{\beta + \sqrt{\nu}})t_k.$$

The rest is clear.

Step 4. The algorithm terminates and the resulting vector $\overline{x}$ satisfies $\|F'(\overline{x})\|_{\overline{x}}^* \leq \beta$.

By Corollary 4.2.1 of [N] one has

$$\|F'(y_0)\|_{y_k}^* \leq (\nu + 2\sqrt{\nu})\|F'(y_0)\|_{x_F^*}^*.$$

Therefore for each k with $t_{k+1} > 0$ one gets

$$t_k > \frac{\gamma}{(\nu + 2\sqrt{\nu})\|F'(y_0)\|_{x_F^*}^*}.$$

Combining this with step 3 it follows that the algorithm terminates, say after N iterations.
We write $\overline{x} = y_N$. Then $t_N = 0$, and $\|F'(\overline{x})\|_{\overline{x}}^* = \lambda(t_N, y_N) \leq \beta$.

Step 5. $\|F'(y_0)\|_{y_{N-2}}^* \leq (1 + \sqrt{\beta})\|F'(y_0)\|_{y_{N-1}}^*$.
By definition

$$y_{N-1} - y_{N-2} = -F''(y_{N-2})^{-1}(-t_{N-1}F'(y_0) + F'(y_{N-2})).$$

Taking the $\| \quad \|_{y_{N-2}}$ norm we get $\|y_{N-1} - y_{N-2}\|_{y_{N-2}} = \| - t_{N-1}F'(y_0) + F'(y_{N-2})\|_{y_{N-2}}^*$.
This is by definition $\lambda(t_{N-1}, y_{N-2})$; this is $\leq \frac{\sqrt{\beta}}{1+\sqrt{\beta}}$ by step 1.
This proves $\|y_{N-1} - y_{N-2}\|_{y_{N-2}} \leq \frac{\sqrt{\beta}}{1+\sqrt{\beta}}$.
Applying theorem 4.1.6. of [N] we get

$$F''(y_{N-1}) \preceq (1 - \|y_{N-1} - y_{N-2}\|_{y_{N-2}})^{-1}F''(y_{N-2}).$$

It follows, on taking inverses, that

$$F''(y_{N-2})^{-1} \preceq (1 + \sqrt{\beta})^2 F''(y_{N-1})^{-1}.$$

Therefore

$$\|F'(y_0)\|_{y_{N-2}}^* \leq (1 + \sqrt{\beta})\|F'(y_0)\|_{y_{N-1}}^*.$$

Step 6. $\|F'(y_0)\|_{y_{N-1}}^* \leq (1 - \sqrt{\beta})^{-1}\|F'(y_0)\|_{x_F^*}^*$.
By step 1 $\lambda(t_N, y_{N-1}) \leq \frac{\sqrt{\beta}}{1+\sqrt{\beta}}$, so as $t_N = 0$, we get by theorem 4.1.11 of [N] that

$$\|y_{N-1} - x_F^*\|_{y_{N-1}} \leq \frac{\lambda(0, y_{N-1})}{1 - \lambda(0, y_{N-1})};$$

this is $\leq \sqrt{\beta}$. Therefore by theorem 4.1.6 of [N]

$$F''(x_F^*) \preceq (1 - \|y_{N-1} - x_F^*\|_{y_{N-1}})^{-2} F''(y_{N-1}).$$

It follows on taking inverses that

$$F''(y_{N-1})^{-1} \preceq (1 - \sqrt{\beta})^{-2} F''(x_F^*).$$

Therefore

$$\|F'(y_0)\|_{y_{N-1}}^* \leq (1 - \sqrt{\beta})^{-1} \|F'(y_0)\|_{x_F^*}^*.$$

Step 7. $N \leq 2 + \max\left[0, \frac{1}{\gamma}(\beta + \sqrt{\nu}) \ln\left[\frac{(1+\sqrt{\beta})\|F'(y_0)\|_{x_F^*}^*}{\gamma(1-\sqrt{\beta})}\right]\right]$.

On the one hand, by step 3

$$t_{N-2} \leq (1 - \frac{\gamma}{\beta + \sqrt{\nu}})^{N-2}.$$

On the other hand, by $t_{N-1} > 0$, we have

$$t_{N-2} > \frac{\gamma}{\|F'(y_0)\|_{y_{N-2}}^*}.$$

Therefore by step 5 and 6 we get

$$t_{N-2} > \frac{(1 - \sqrt{\beta})\gamma}{(1 + \sqrt{\beta})\|F'(y_0)\|_{x_F^*}^*}.$$

Combining this upperbound and lowerbound for t_{N-2} gives an inequality; on taking the logarithm and on using the inequality $\ln(1 + \tau) \leq \tau$ we get the required upperbound for N.

Acknowledgement

I would like to express my thanks to Prof. Yurii Nesterov at CORE, Université Catholique de Louvain for helpful discussions on interior point methods and in particular for suggesting the problem which is solved in this paper.

References

[N] Nesterov, *Introductory Lectures on Convex Programming*, Volume I: Basic course, July 2, 1998 (to be published by Kluwer in 2001).