Sven O. Krumke

# On the Convergence of a Modified Barrier Function Method for Convex Quadratic and Linear Programming

*We analyze convergence properties of the Modified Barrier Function Method (MBFM), which operates like the usual Augmented Lagrangian method, except that it uses a shifted logarithmic barrier function in place of the usual quadratic penalty function. The MBFM avoids some of the drawbacks of classical barrier methods. The convergence of the MBFM had previously been established under strict regularity assumptions for general convex programs. We focus on Convex Quadratic and Linear Programs and strengthen substantially the existing convergence results. For Linear Programs, we also analyze the speed of convergence of the dual sequence.*

## 1. Foundations

In the following we are going to consider the solution of the convex quadratic program

$$
\begin{array}{lll}
\text{(QP)} & \min & q(x) \equiv \frac{1}{2}x^TQx + c^Tx \\
 & x \in S := & \{x \in \mathbb{R}^n : Ax \geq b\} = \{x \in \mathbb{R}^n : a_j^Tx - b_j \geq 0, j = 1, \dots, m\},
\end{array}
$$

using the *Modified Barrier Function Method (MBFM)* introduced by R. Polyak in [3]. Here $Q$ is a symmetric positive semidefinite $n \times n$ matrix, $A$ is a $m \times n$ matrix with the rows $a_1^T, \dots, a_m^T$. Thus (QP) is a convex program. We denote by $q^* := \inf\{q(x) : x \in S\}$ the optimal function value and by $S^* := \{x \in S : q(x) = q^*\}$ the set of optimal solutions for (QP). Throughout the paper we will make the following assumption:

**Assumption A:** The optimal solution set $S^*$ for (QP) is nonempty and bounded.

The Lagrangian $L$ associated with (QP) is given by $L(x, \lambda) := \frac{1}{2}x^TQx + c^Tx + \lambda^T(b - Ax)$ if $\lambda \geq 0$ and $L(x, \lambda) := -\infty$ otherwise. By the Kuhn–Tucker Theorem from a feasible vector $x^* \in S$ is an optimal solution of (QP) if and only if there is a vector $\lambda^* \geq 0$ such that $\nabla_x L(x^*, \lambda^*) = Qx^* + c - A^T\lambda^* = 0$ and the complementary slackness condition $(\lambda^*)^T(Ax^* - b)^T = 0$ is satisfied. The *(Lagrange) dual program* (D) associated with (QP) is given by

$$
\text{(D)} \qquad \max_{\lambda \geq 0} \quad d(\lambda), \tag{1}
$$

where $d(\lambda) := \inf_{x \in \mathbb{R}^n} L(x, \lambda)$ is the *dual functional* for (QP). It is easy to see that under our assumptions $d$ is concave, proper and closed. The weak duality theorem asserts that the value $d(\lambda)$ of any vector $\lambda \in \mathbb{R}^m$ is less than or equal to the cost $q(x)$ of any primal feasible vector $x^*$. Moreover, there is no duality gap, i.e. the optimal value of (D) is equal to $q^*$ and there exists a dual optimal solution (cf. [1]).

The *Modified Barrier Function* $\mathcal{F}$ for (QP) is defined by $\mathcal{F}(x, \lambda, k) = \frac{1}{2}x^TQx + c^Tx - \frac{1}{k}\sum_{j=1}^m \lambda_j \log(k \cdot (a_j^Tx - b_j) + 1)$. Given a sequence $(k^\nu)_\nu$ of positive penalty parameters which are bounded away from zero by a constant $\varepsilon > 0$, the MBFM performs a sequence of unconstrained minimizations steps, and iterates on the multipliers at the end of each minimization step. At the $\nu$–th iteration, we are given a vector $\lambda^\nu > 0$ (with the initial $\lambda^0 > 0$ chosen arbitrarily); the method then computes $x^\nu \in \arg\min\{\mathcal{F}(x, \lambda^\nu, k^\nu) : x \in \mathbb{R}^n\}$ and updates the multiplier $\lambda^\nu$ according to $\lambda_j^{\nu+1} := \lambda_j^\nu / [k^\nu \cdot (a_j^Tx^\nu - b_j) + 1]$ for $j = 1, \dots, m$. The choice of the update formula is motivated by comparing the Kuhn–Tucker condition $\nabla_x L(x^*, \lambda^*) = 0$ with the fact that the minimizer $x^\nu$ satisfies $0 = \nabla_x \mathcal{F}(x^\nu, \lambda^\nu, k^\nu) = Qx^\nu + c - \sum_{j=1}^m \lambda_j^\nu / [k \cdot (a_j^Tx^\nu - b_j) + 1] \cdot a_j$, since then we obtain that

$$
0 = Qx^\nu + c - A^T\lambda^{\nu+1} = \nabla_x L(x^\nu, \lambda^{\nu+1}). \tag{2}
$$

Using elementary properties of the two sequences that are generated one can establish

Proposition 1. *If $\lim_{\nu \to \infty} x^\nu = x^*$ and $\lim_{\nu \to \infty} \lambda^\nu = \lambda^*$, then $x^* \in S^*$ and $\lambda^*$ is a corresponding optimal multiplier.*

In [3] it has been shown for general convex programs that under strict regularity assumptions, the primal sequence $(x^\nu)_\nu$ and the dual sequence $(\lambda^\nu)_\nu$ generated by the MBFM converges to an optimal solution and a corresponding multiplier respectively, provided that $k^\nu \geq \omega$ for some large $\omega > 0$. It is also proved that the condition number of the Hessian $\nabla^2_{xx}\mathcal{F}$ remains stable in a neighborhood of the optimal solution, thus avoiding one of the major drawbacks of classical barrier functions.

## 2. Convergence of the Dual Sequence

From (2) and the convexity of $L(., \lambda^{\nu+1})$ we obtain that $x^\nu$ attains the minimum in the dual function with $\lambda$ set to $\lambda^{\nu+1}$. Thus it follows that $d(\lambda^{\nu+1}) = q(x^\nu) + (\lambda^{\nu+1})^T(b - Ax^\nu)$. By observing that for all $\lambda \in \mathbb{R}^m$ the estimation $d(\lambda) \leq q(x^\nu) + \lambda^T(b - Ax^n u) = d(\lambda^{\nu+1}) + (b - Ax^\nu)^T(\lambda - \lambda^{\nu+1})$ does hold, we see that $b - Ax^\nu \in \partial d(\lambda^{\nu+1})$. Using this property, we can establish the following:

Lemma 1. $\lambda^{\nu+1} \in \arg\max_{\lambda \geq 0}[d(\lambda) - \frac{1}{k^\nu}\sum_{j=1}^{m} \lambda_j^\nu \beta\left(\frac{\lambda_j}{\lambda_j^\nu}\right)]$, *where* $\beta(t) = -\log t + t - 1$.

In the next step we prove that the sequence $(d(\lambda^\nu))_\nu$ is monotonically nondecreasing. This is done by using $b - Ax^\nu \in \partial d(\lambda^{\nu+1})$ and $(\lambda_j^{\nu+1} - \lambda_j^\nu)(1 - \lambda_j^\nu/\lambda_j^{\nu+1}) \geq 0$ for $j = 1, \ldots, m$. Since by the duality theorem $(d(\lambda^\nu))_\nu$ is bounded from above, we conclude that it must converge to some $d^\infty \leq q^*$. The set $M^\infty := \{\lambda \geq 0 : d(\lambda) \geq d^\infty\}$ is nonempty.

Lemma 2. *For each $\hat{\lambda} \in M^\infty$, the sequence $(\mathcal{D}(\hat{\lambda}, \lambda^\nu))_\nu$ is monotonically nonincreasing. Here $\mathcal{D}$ denotes the* Bregman–distance $\mathcal{D}(\lambda, \mu) := \sum_{j=1}^m \lambda_j \log(\lambda_j/\mu_j) - \lambda_j + \mu_j$.

Using the fact that $\mathcal{D}$ is nonnegative and $(\mathcal{D}(\hat{\lambda}, \lambda^\nu))_\nu$ converges to zero if and only if $(\lambda^\nu)_\nu$ converges to $\hat{\lambda}$ we show that the dual sequence is convergent. The next theorem states one of our main results

Theorem 1. *The dual sequence $(\lambda^\nu)_\nu$ converges to an optimal solution $\lambda^*$ of* (D).

## 3. Convergence Properties of the Primal Sequence

If $Q$ is strictly positive definite, the convergence of the dual sequence implies that the $x^\nu$ must also converge. Thus the result of proposition 1 tells us that the limit must be an optimal solution of (QP). On the other hand, we have:

Theorem 2. *Suppose* (QP) *has a unique solution $x^*$. Suppose further that either $Q = 0$ or that there exists a strict complementary multiplier $\lambda^*$ correspoding to $x^*$. Then $\lim_{\nu \to \infty} x^\nu = x^*$.*

## 4. Speed of Convergence for Linear Programs

If $Q = 0$ then the dual functional becomes $d(\lambda) = b^T\lambda$ if $A^T\lambda = c$ and $\lambda = 0$, and $d(\lambda) = -\infty$ otherwise. Thus (QP) and its dual (D) reduce to the well known pair of dual linear programs. To analyze the speed of convergence, we use a lemma which has been proved in [4]. Thus we establish the following:

Theorem 3. *If $k^\nu = k$ for $\nu = 0, 1, \ldots$ and $k > 0$ is an arbitrary constant, the sequence $(d(\lambda^\nu))_\nu = (b^T\lambda^\nu)_\nu$ converges at least linearly to the optimal function value $d^* = q^*$. If $\lim_{\nu \to \infty} k^\nu = +\infty$, then the sequence $(b^T\lambda^\nu)_\nu$ converges superlinearly to $q^*$.*


## Acknowledgements

*The author would like to thank Prof. Dr. Josef Stoer for his valuable ideas and his support during the work.*


## 5. References


1 R. Horst: Nichtlineare Optimierung, Carl Hanser Verlag, München Wien, 1979
2 S.O. Krumke: Eine modifizierte Barrieremethode für konvexe quadratische Optimierungsprobleme, Diplomathesis, Universität Würzburg, 1994
3 R. Polyak: Modified Barrier Functions (Theory and Methods), Mathematical Programming 54 (1992), 177–222
4 P. Tseng, D. P. Bertsekas: On the convergence of the Exponential Multiplier Method for Convex Programming, Mathematical Programming 60 (1993), 1–19



*Addresses:* Sven O. Krumke, Lehrstuhl für Informatik I, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany