Existence and uniqueness of the regularized primal-dual central path

Jordi Cuesta
Dept. of Chemical Engineering
Universitat Rovira i Virgili
jordi.cuesta@urv.cat

Jordi Castro
Dept. of Stat. and Operations Research
Universitat Politècnica de Catalunya
jordi.castro@upc.edu

Research Report UPC-DEIO DR 2009-08
July 2009

Report available from http://www-eio.upc.es/~jcastro

Existence and uniqueness of the regularized primal-dual central path

Jordi Cuesta[a], Jordi Castro[*,b]

[a]*Operations Research unit, Dept. of Chemical Engineering, Universitat Rovira i Virgili, Tarragona, Catalonia, Spain*
[b]*Dept. of Statistics and Operations Research, Universitat Politècnica de Catalunya, Barcelona, Catalonia, Spain*

Abstract

In a recent work [3] the authors improved one of the most efficient interior-point approaches for some classes of block-angular problems. This was achieved by adding a quadratic regularization to the logarithmic barrier. This regularized barrier was shown to be self-concordant, thus fitting the general structural optimization interior-point framework. In practice, however, most codes implement primal-dual path-following algorithms. This short paper shows that the primal-dual regularized central path is well defined, i.e., it exists and it is unique.

Key words:
interior-point methods, primal-dual central path, path-following methods, regularizations

---

1. Introduction

Let us consider the linear programming problem

$$\begin{array}{ll} \min & c^T x \\ \text{s. to} & Ax = b \\ & 0 \le x \le u, \end{array} \tag{1}$$

where $x, c, u \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m\times n}$. Note that any bounded problem can be formulated as (1). The standard logarithmic barrier problem, used in interior-point methods, associated to (1) is

$$\begin{array}{ll} \min & B(x,\mu) \triangleq c^T x + \mu\left(-\sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln(u_i - x_i)\right) \\ \text{s. to} & Ax = b, \end{array} \tag{2}$$

μ being the barrier parameter. Previously used regularized variants replaced $B(x,\mu)$ by

$$\begin{array}{rcl} B_P(x,\mu) & \triangleq & c^T x + \frac{1}{2}(x-\overline{x})^T Q_P (x-\overline{x}) \\ & + & \mu\left(-\sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln(u_i - x_i)\right), \end{array} \tag{3}$$

[*]Corresponding address: Dept. of Statistics and Operations Research, Universitat Politècnica de Catalunya, Jordi Girona 1–3, 08034 Barcelona, Catalonia, Spain.
Email address: `jordi.castro@upc.edu` (Jordi Castro)

Q_P being a positive definite matrix and $\overline{x}$ the current point obtained by the interior-point algorithm. For instance, Q_P was the identity matrix in [5]; and Q_P was a diagonal matrix with small entries—dynamically updated at each interior-point iteration—in [1]. Unfortunately, these proximal point regularizations depend on the current point $\overline{x}$, and then they do not fit the general theory of structural optimization for interior-point methods [4]. In [3] the authors suggested the alternative regularized barrier problem

$$B_Q(x,\mu) \triangleq c^T x + \mu\left(\frac{1}{2}x^T Qx - \sum_{i=1}^{n} \ln x_i - \sum_{i=1}^{n} \ln(u_i - x_i)\right) \tag{4}$$

Q being a diagonal positive semidefinite matrix. This regularized barrier function was shown to be a self-concordant barrier [3] for upper-bounded problems and thus it fits the general interior-point theory of [4]. It was shown in [3] than, due to this regularization term, the spectral properties of a preconditioned system were significantly improved. This allowed the efficient solution of the normal equations of some very large primal block-angular problems by means of a scheme that combines Cholesky factorizations and preconditioned conjugate gradients [2].

The KKT conditions for (2) are [6]:

$$\begin{array}{rcl} Ax & = & b, \\ A^T y + z - w & = & c, \\ XZe & = & \mu e, \\ (U-X)We & = & \mu e, \\ (z,w) & > & 0 \quad u > x > 0; \end{array} \tag{5}$$

$e \in \mathbb{R}^n$ is a vector of 1's; $y \in \mathbb{R}^m$, $z, w \in \mathbb{R}^n$ are the Lagrange multipliers (or dual variables) of $Ax = b$, $x \geq 0$ and $x \leq u$, respectively; and matrices $X, Z, U, W \in \mathbb{R}^{n \times n}$ are diagonal matrices made up of vectors x, z, u, w. The first two sets of equations of (5) impose, respectively, primal and dual feasibility; the last two impose complementarity. The solutions of system (5) for different μ values gives rise to an arc of strictly feasible primal-dual points known as the primal-dual central path. As μ tends to 0, the solutions of (5) converge to those of (1) and its dual. A primal-dual path-following algorithm attempts to follow the primal-dual central path. This is the algorithm implemented in packages like, e.g., CPLEX, XPress, MOSEK, etc.

The KKT conditions for (2) replacing $B(x,\mu)$ by the regularized version (4) are

$$\begin{array}{rcl} Ax & = & b, \\ A^T y - \mu Qx + z - w & = & c, \\ XZe & = & \mu e, \\ (U-X)We & = & \mu e, \\ (z,w) & > & 0 \quad u > x > 0. \end{array} \tag{6}$$

Note (5) and (6) only differ in the dual feasibility. System (6) will be referred as the the regularized KKT conditions, and the arc of primal-dual solutions for different μ values as the regularized primal-dual central path.

The purpose of this short paper is to show that the regularized primal-dual central path is well defined for (primal and dual) feasible problems: it exists and it is unique (i.e., for any μ there is a solution to (6), and this solution is unique). Section 2 shows the existence and uniqueness. We extend previous results [6] for the standard central path defined by (5) to the new regularized version (6).

2. Existence and uniqueness

To simplify the notation, we will consider that (1) has been transformed to an equivalent problem without upper bounds (i.e., adding slacks $s \in \mathbb{R}^n$, and constraints $x + s = u$ to $Ax = b$, and including slacks in the vector of variables). The simplified regularized KKT conditions (6) are

$$\begin{aligned} Ax &= b, \\ A^T y - \mu Q x + z &= c, \\ XZe &= \mu e, \\ (x, z) &> 0. \end{aligned} \tag{7}$$

The primal-dual feasible set $\mathcal{F}$ and the strictly feasible set $\mathcal{F}^0$ are defined by

$$\mathcal{F} = \left\{(x, y, z) | Ax = b,\ A^T y - \mu Q x + z = c,\ (x,\ z) \geq 0\right\}, \tag{8}$$

$$\mathcal{F}^0 = \left\{(x, y, z) | Ax = b,\ A^T y - \mu Q x + z = c,\ (x,\ z) > 0\right\}. \tag{9}$$

We start by proving the following preliminary Lemma, to be used later:

Lemma 1. *If $\mathcal{F}^0 \neq \oslash$ (i.e., the problem is strictly feasible) then for each $K \in \mathbb{R}, K \geq 0$, the set*

$$\left\{(x, z)\,|\,(x, y, z) \in \mathcal{F} \text{ for some } y,\ \text{and } x^T z \leq K\right\} \tag{10}$$

is bounded.

Proof. Let $(\bar{x}, \bar{y}, \bar{z})$ be any point in $\mathcal{F}^0$ and (x, y, z) be any point in $\mathcal{F}$ such that $x^T z \leq K$. Since $A\bar{x} = b$ and $Ax = b$ then $A(\bar{x} - x) = 0$. Similarly, $A^T(\bar{y} - y) + (\bar{z} - z) - \mu Q(\bar{x} - x) = 0$. Therefore,

$$\begin{aligned} (\bar{x} - x)^T (\bar{z} - z) &= (\bar{x} - x)^T \left(-A^T (\bar{y} - y) + \mu Q (\bar{x} - x)\right) \\ &= \mu (\bar{x} - x)^T Q (\bar{x} - x) \\ &\quad - (\bar{x} - x)^T A^T (\bar{y} - y) \\ &= \mu (\bar{x} - x)^T Q (\bar{x} - x) - 0 \cdot (\bar{y} - y) \\ &= \mu (\bar{x} - x)^T Q (\bar{x} - x), \end{aligned}$$

which can be recast as

$$\bar{x}^T z + \bar{z}^T x = \bar{x}^T \bar{z} + x^T z - \mu (\bar{x} - x)^T Q (\bar{x} - x).$$

Since $x^T z \leq K$ and $\mu (\bar{x} - x)^T Q (\bar{x} - x) \geq 0$ because Q is positive semidefinite,

$$\bar{x}^T z + \bar{z}^T x \leq K + \bar{x}^T \bar{z} - \mu (\bar{x} - x)^T Q (\bar{x} - x) \leq K + \bar{x}^T \bar{z}. \tag{11}$$

The value $\xi = \min_{i=1,\ldots,n} \min(\bar{x}_i, \bar{z}_i)$ is positive, since $(\bar{x}, \bar{z}) > 0$. Then from (11) we have

$$\xi e^T (x + z) \leq \bar{x}^T z + \bar{z}^T x \leq K + \bar{x}^T \bar{z},$$

which means

$$0 \leq x_i \leq \frac{1}{\xi} \left(K + \bar{x}^T \bar{z}\right),\ 0 \leq z_i \leq \frac{1}{\xi} \left(K + \bar{x}^T \bar{z}\right), i = 1, \ldots, n,$$

and hence (10) is bounded. □

To show existence and uniqueness we first define the new set

$$\mathcal{H}^0 = \left\{(x,z) \mid (x,y,z) \in \mathcal{F}^0 \text{ for some } y\right\}.$$

We also define the barrier function

$$f_\mu(x,z) = \frac{1}{\mu} x^T z - \sum_{i=1}^{n} \log\left(x_i z_i\right), \tag{12}$$

with the following properties:

Lemma 2. 1. *f_μ tends to $+\infty$ whenever (x,z) approaches the boundary of $\mathcal{H}^0$, i.e., when any x_j or z_j approaches 0.*

2. *f_μ is strictly convex on $\mathcal{H}^0$.*
3. *f_μ is bounded below on $\mathcal{H}^0$.*
4. *Given $\mu > 0$, and any $\kappa \in \mathbb{R}$, points (x,z) of the level set $\mathcal{L}_\kappa = \left\{(x,z) \in \mathcal{H}^0 \mid f_\mu(x,z) \le \kappa\right\}$ satisfy*

$$x_i \in [M_l, M_u], \quad z_i \in [M_l, M_u], \quad i = 1, \ldots, n, \tag{13}$$

for some positive numbers M_l and M_u, and thus they are contained in compact subsets.

Proof. (We remark that the regularization term does not intervene in proofs of properties 1, 3 and 4, and they are the same than for the standard central path; anyway, we recall them here for completeness).

Property 1 is straightforward.

For property 2, note that the second term $-\sum_{i=1}^n \log(x_i z_i)$ is strictly convex (since its Hessian is positive definite. The first term is shown to be convex on $\mathcal{H}^0$. Indeed, if $\bar{x}$ is any point for which $A\bar{x} = b$, we have for any $(x,z) \in \mathcal{H}^0$ that $x^T z = x^T(c - A^T y + \mu Q x) = c^T x - \bar{x}^T A^T y + \mu x^T Q x = c^T x - \bar{x}^T(c - z + \mu Q x) + \mu x^T Q x = c^T x - c^T \bar{x} + \bar{x}^T z - \mu \bar{x}^T Q x + \mu x^T Q x$, which is convex in (x,z) since $Q \ge 0$. Hence, $f_\mu(x,z)$ is the sum of a convex and a strictly convex function, thus it is strictly convex.

To show property 3, we define $g(t) = t - \log t - 1$ and rewrite $f_\mu(x,z)$ as

$$f_\mu(x,z) = \sum_{j=1}^{n} g\left(\frac{x_j z_j}{\mu}\right) + n - n \log \mu. \tag{14}$$

Function $g(t)$ is strictly convex in $(0,\infty)$, $g(t) \ge 0$ for $t \in (0,\infty)$, and tends to ∞ when either $t \to 0$ or $t \to \infty$. Using $g(t) \ge 0$ in (14) we have

$$f_\mu(x,z) \ge n - n \log \mu = n\left(1 - \log \mu\right),$$

i.e., $f_\mu(x,z)$ is bounded below.

Property 4 is shown by noting that by (14) $f_\mu(x,\ z) \le \kappa$ if and only if

$$\sum_{j=1}^{n} g\left(\frac{x_j z_j}{\mu}\right) \le \bar{\kappa},$$

where $\bar{\kappa} = \kappa - n + n \log \mu$. Choosing a particular index $i = j$, and using that $g(t) \ge 0$, we have

$$g\left(\frac{x_i z_i}{\mu}\right) \le \bar{\kappa} - \sum_{j \ne i} g\left(\frac{x_j z_j}{\mu}\right) \le \bar{\kappa}.$$

Therefore, using that $g(t) \to \infty$ when either $t \to 0$ or $t \to \infty$, there exists a value M such that

$$\frac{1}{M} \le x_i z_i \le M, \quad i = 1, \ldots, n. \tag{15}$$

Adding the terms in this expression we get

$$x^T z = \sum_{i=1}^{n} x_i z_i \le nM. \tag{16}$$

By (16) and the boundedness established by Lemma 1 we know there exists a number M_u such that $x_i \in (0, M_u]$ and $z_i \in (0, M_u]$ for all $i = 1, \ldots, n$. Using (15), we have that $x_i \ge 1/(Mz_i) \ge 1/(MM_u)$ for all i; for z_i we obtain the same lower bound. (13) holds by setting $M_l = 1/(MM_u)$. □

Finally, next Theorem 1 shows that for any $\mu > 0$ the barrier function $f_\mu(x,z)$ defined by (12) reaches its minimum in $\mathcal{H}^0$, that the minimizer is unique, and that this means that the regularized KKT conditions (7) have a unique solution.

Theorem 1. *If $F^0 \neq \emptyset$ and $\mu > 0$, then $f_\mu(x,z)$ has a unique minimizer in $\mathcal{H}^0$, and (6) has a unique solution.*

Proof. By property 4 of Lemma 2 we have that level sets $\mathcal{L}_\kappa = \left\{(x,z) \in \mathcal{H}^0 \,\middle|\, f_\mu(x,z) \le \kappa\right\}$ of $f_\mu(x,z)$ are contained in a compact subset of $\mathcal{H}^0$, and thus $f_\mu(x,z)$ has a minimizer in $\mathcal{H}^0$. By property 2 of Lemma 2, $f_\mu(x,z)$ is strictly convex, thus the minimizer will be unique.

We next show this unique minimizer corresponds to the unique solution of (6). This minimizer solves the linearly constrained minimization problem

$$\min f_\mu(x,\ z) \text{ s. to } Ax = b, A^T y + z - \mu Q x = c,\ (x,z) > 0. \tag{17}$$

From the Lagrangian

$$\begin{aligned} \mathcal{L}(x,y,z,v,w) = \quad & f_\mu(x,\ z) + v^T(Ax - b) \\ & + w^T(A^T y + z - \mu Q x - c) \end{aligned}$$

we obtain the KKT conditions of (17)

$$\begin{aligned} \frac{d\mathcal{L}}{dx} &= \frac{df_\mu}{dx} + A^T v - \mu Q w \\ &= \frac{1}{\mu} Ze - X^{-1}e + A^T v - \mu Q w = 0, \\ \frac{d\mathcal{L}}{dy} &= Aw = 0, \\ \frac{d\mathcal{L}}{dz} &= \frac{df_\mu}{dz} + w = \frac{1}{\mu} Xe - Z^{-1}e + w = 0. \end{aligned} \tag{18}$$

By combining the first and third equalities of (18) we obtain

$$A^T v = X^{-1}e - \frac{1}{\mu} Ze + \mu Q(Z^{-1}e - \frac{1}{\mu} Xe). \tag{19}$$

By combining the second and third we find that

$$A(Z^{-1}e - \frac{1}{\mu}Xe) = 0,$$

which means

$$(Z^{-1}e - \frac{1}{\mu}Xe)^T A^T v = 0.$$

Using the above result in (19) we have

$$(Z^{-1}e - \frac{1}{\mu}Xe)^T (X^{-1}e - \frac{1}{\mu}Ze + \mu Q(Z^{-1}e - \frac{1}{\mu}Xe)) = 0,$$

or equivalently,

$$\begin{aligned} & (Z^{-1}e - \frac{1}{\mu}Xe)^T (X^{-1}e - \frac{1}{\mu}Ze) \\ + \quad & (Z^{-1}e - \frac{1}{\mu}Xe)^T \mu Q(Z^{-1}e - \frac{1}{\mu}Xe) = 0 \end{aligned} \tag{20}$$

The first term of (20) can be written as

$$\begin{aligned} (Z^{-1}e - \frac{1}{\mu}Xe)^T (X^{\frac{-1}{2}} Z^{\frac{1}{2}})(X^{\frac{1}{2}} Z^{\frac{-1}{2}})(X^{-1}e - \frac{1}{\mu}Ze) \quad &= \\ (X^{\frac{-1}{2}} Z^{\frac{-1}{2}} e - \frac{1}{\mu}X^{\frac{1}{2}} Z^{\frac{1}{2}} e)^T (X^{\frac{-1}{2}} Z^{\frac{-1}{2}} e - \frac{1}{\mu}X^{\frac{1}{2}} Z^{\frac{1}{2}} e) \quad &= \\ \left\| (XZ)^{-1/2}e - \frac{1}{\mu}(XZ)^{1/2}e \right\|_2^2 \quad &\geq \quad 0. \end{aligned}$$

$\| \ \|_2$ being the Euclidean norm. Using that $Q \geq 0$ and $\mu > 0$, the second term of (20) is equivalent to

$$\left\| Z^{-1}e - \frac{1}{\mu}Xe \right\|_{\mu Q}^2 \geq 0,$$

$\| \ \|_{\mu Q}$ being the norm induced by μQ. Therefore (20) holds if and only if

$$\begin{aligned} \left\| (XZ)^{-1/2}e - \frac{1}{\mu}(XZ)^{1/2}e \right\|_2^2 \quad &= \quad 0, \\ \left\| Z^{-1}e - \frac{1}{\mu}Xe \right\|_{\mu Q}^2 \quad &= \quad 0. \end{aligned}$$

From the first equality we have $(XZ)^{-1/2}e = \frac{1}{\mu}(XZ)^{1/2}e$, and therefore $XZe = \mu e$. The second equality means $Z^{-1}e = \frac{1}{\mu}Xe$ and we obtain the same result. Therefore, the unique minimizer of (17) satisfy not only the feasibility conditions of (7), but also the μ-complementarity condition, and the proof is complete. □

Acknowledgments

This work has been supported by grant MTM2006-05550 of the Spanish Ministry of Science and Education.

References

[1] A. Altman, J. Gondzio, Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization, Optimization Methods and Software 11 (1999) 275–302.

[2] J. Castro, A specialized interior-point algorithm for multicommodity network flows, SIAM Journal on Optimization 10 (2000) 852-877.

[3] J. Castro, J. Cuesta, Quadratic regularizations in an interior-point method for primal block-angular problems, Mathematical Programming, conditionally accepted (2009).

[4] Y. Nesterov, Introductory Lectures on Convex Optimization: A Basic Course, Kluwer, 2004.

[5] R. Setiono, Interior proximal point algorithm for linear programs, Journal of Optimization Theory and Applications 74 (1992) 425–444.

[6] S.J. Wright, Primal-Dual Interior-Point Methods, SIAM, 1996.