Journal of Computational and Applied Mathematics 236 (2012) 3704–3709

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Regularization techniques in interior point methods[☆]

Csaba Mészáros

Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest, P.O. BOX 63. Hungary

ARTICLE INFO

Article history:
Received 24 September 2010
Received in revised form 1 July 2011

Keywords:
Interior point methods
Symmetric factorization
Regularization

ABSTRACT

Regularization techniques, i.e., modifications on the diagonal elements of the scaling matrix, are considered to be important methods in interior point implementations. So far, regularization in interior point methods has been described for linear programming problems, in which case the scaling matrix is diagonal. It was shown that by regularization, free variables can be handled in a numerically stable way by avoiding column splitting that makes the set of optimal solutions unbounded. Regularization also proved to be efficient for increasing the numerical stability of the computations during the solutions of ill-posed linear programming problems. In this paper, we study the factorization of the augmented system arising in interior point methods. In our investigation, we generalize the methods developed and used in linear programming to the case when the scaling matrix is positive semidefinite, but not diagonal. We show that regularization techniques may be applied beyond the linear programming case.

1. Introduction

Karmarkar's publication on the projective scaling algorithm in 1984 [1] led to an explosion of interest among researchers and practitioners of linear programming. Nowadays, after 25 years of evolution of the algorithms and their implementation techniques, interior point methods (IPMs) have become reliable and efficient practical tools for solving large-scale optimization problems. The most efficient implementations follow the primal–dual method with predictor–corrector techniques [2], and are based on the direct Cholesky or quasidefinite factorization approaches [3]. Very significant results were achieved in the fields that are not directly connected to the interior point algorithms but contribute to the efficiency when solving problems in practice, such as in presolve [4,5], sparse matrix orderings [6], vector and parallel processing in the numerical kernels [7–9]. Numerical stability proved to be an important issue of the implementations of IPMs that also gained great attention in the literature [10–13]. In this paper, we generalize the regularization techniques which were developed for linear programming for handling free variables [14] and improving numerical stability of the computations [13]. We show that the basic properties are also valid where the optimization problem is convex, but nonlinear and when the objective function is nonseparable.

The paper is organized as follows. Section 2 introduces the novel interior point algorithm and the necessary operations to compute its iterations. Section 3 summarizes the regularization techniques that were developed for linear programming. In Section 4, we generalize these methods to the case when the objective function is convex but nonseparable. Section 5 presents numerical results and final remarks are given in Section 6.

☆ The work was supported in part by the Hungarian Research Fund OTKA K-77420.
E-mail address: meszaros@sztaki.hu.
URL: http://www.oplab.sztaki.hu.

0377-0427/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2011.07.012

2. Interior point methods and symmetric factorizations

For simplicity, we consider here the linearly constrained convex optimization problem and the infeasible primal–dual log barrier interior point method for further investigations. Let the primal problem be

$$\begin{aligned} &\min \quad f(x), \\ &Ax = b, \\ &x \geq 0, \end{aligned} \tag{1}$$

where $x \in R^n$, $b \in R^m$, $A \in R^{m\times n}$, $f : R^n \Longrightarrow R$ twice continuously differentiable, and assume that A is of full row rank and the feasible region $\{x : Ax = b, x \geq 0\}$ is nonempty.

The primal–dual logarithmic barrier method, introduced in [15,16] and its infeasible version [17] are widely considered as the most efficient IPM in practice. The algorithm can be described by the following simple scheme:

- replace nonnegativity with logarithmic barrier terms with a μ barrier parameter,
- derive the first order necessary conditions,
- solve the resulting nonlinear system of equations by the Newton method.

The barrier problem, corresponding to (1) can be written as follows:

$$\begin{aligned} &\min \quad f(x) - \mu \sum_{i=1}^{n} \ln x_i, \\ &Ax = b, \end{aligned} \tag{2}$$

where $\mu > 0$ is called the barrier parameter. The Lagrangian function of the barrier problem can be written as

$$L(x, y) = f(x) - \mu \sum_{i=1}^{n} \ln x_i - y^T(Ax - b),$$

from which the following first order optimality conditions can be derived:

$$\begin{aligned} &\nabla_x L(x, y) = \nabla_x f(x) - \mu X^{-1}e - A^T y = 0, \\ &\nabla_y L(x, y) = Ax - b = 0, \end{aligned}$$

where $X = \operatorname{diag}(x_1, \ldots, x_n)$, $e = (1, \ldots, 1)$. After introducing

$$z = \mu X^{-1} e,$$

the first order optimality conditions will be

$$\begin{aligned} &\nabla f(x) - z - A^T y = 0, \\ &Ax - b = 0, \\ &Xz - \mu e = 0. \end{aligned} \tag{3}$$

In this way the optimization problem was transformed to a system of nonlinear equations. The main idea behind the primal–dual logarithmic barrier methods is that (3) is solved iteratively by the Newton method, while μ is decreased simultaneously to zero. It is easy to derive that one iteration of the Newton method from (x, y, z) to solve (3) requires solving the system of linear equations as follows:

$$\begin{bmatrix} A & 0 & 0 \\ -H_f(x) & A^T & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} b - Ax \\ A^T y + z - \nabla f(x) \\ \mu e - XZe \end{bmatrix},$$

where $Z = \operatorname{diag}(z_1, \ldots, z_n)$, H_f denotes the Hessian and ∇f the gradient of f. The above set of equations can be easily reduced to

$$\begin{bmatrix} -ZX^{-1} - H_f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} A^T y + z - \nabla f(x) - \mu X^{-1} e + Ze \\ b - Ax \end{bmatrix} \tag{4}$$

by pivoting on the diagonal matrix X. In the paper we will use the notations

$$\begin{aligned} &D := ZX^{-1} + H_f(x) \quad \text{and} \\ &M := \begin{bmatrix} -D & A^T \\ A & 0 \end{bmatrix}. \end{aligned} \tag{5}$$

In matrix M the term ZX^{-1}, which changes during the iteration of the interior point method, is referred as the scaling matrix.