A UNIFIED INTERIOR POINT FRAMEWORK FOR OPTIMIZATION ALGORITHMS

M. A. Ibiejugba
Department of Mathematical Sciences, Kogi State University, Anyigba, Nigeria

J. O. Fakiya
Department of Mathematics/Statistics/ College of Natural Sciences, Computer & Information Systems, Achievers University, Owo, Nigeria

Abstract
Interior Point algorithms are optimization methods developed over the last three decades following the 1984 fundamental paper of Karmarkar. Over this period, IPM algorithms have had a profound impact on optimization theory as well as practice and have been successfully applied to many problems of business, engineering and science. Because of their operational simplicity and wide applicability, IPM algorithms are now playing an increasingly important role in computational optimization and operations research. This article provides unified interior point algorithms to optimization problems as well as comparing performances with classical algorithms.
Keywords Interior Point methods, Optimization algorithms, Lagrangian Multipliers, Barrier methods, Newton's method, Matrix-free method.

1. Introduction

In 1984, Karmarkar [5] inspired the new era of mathematical programming with the publication of his landmark paper that sparked the research on polynomial interior point methods. In brief, while the simplex method goes along the edges of the polyhedron corresponding to the feasible region, interior point methods pass through the interior of this polyhedron [2]. A major attraction of these interior point methods is that its application is not limited to the linear optimization problem alone but has found applications in other areas of large scale optimization problems in systems and control theory [12]. They are much easier to derive, motivate, and understand than they at first appear, Marsten et al (1990), Lesaja (2009), Gondzio and Grothey (2004).

Lagrange told us how to convert a minimization with equality constraints into an unconstrained minimization; Fiacco and McCormick (1968) told us how to convert a minimization with inequality constraints into a sequence of unconstrained minimizations, and Newton told us how to solve unconstrained minimizations [7]. Consequently, subject to minor modifications, the same linear algebra kernel may be used to implement interior point methods for all three classes of optimization problems [4, 7]. Hence, interior point methods provide a unified framework for optimization algorithms for Linear, Quadratic and Nonlinear programming and are the right way to solve large linear programs.

The theoretical foundation for interior point methods consists of three crucial building blocks Marsten et al (1990). First is the Newton's [8] method for solving nonlinear equations which is applicable to the solution of unconstrained optimization. Second, Lagrange's method for transforming optimization with equality constraints into an unconstrained problem leading to a system of (n+m) equations in (n+m) variables. Thirdly, Fiacco and McCormick's [3] barrier method for optimization with inequality constraints which are converted to equations by adding nonnegative slack (or surplus) variables; so that the only essential inequalities are the non-negativity conditions; ≥ 0 .

Adapting the use of the above three building blocks, it is shown in [4, 7, 9] how to construct the primal-dual interior point method, which is considered the most elegant theoretically of the many variants of the interior point techniques and also the most successful computationally. The main idea of interior point methods is essentially to change the constrained problem into a succession of unconstrained problems using a logarithmic barrier penalty function, which was initially introduced by Fiacco and McCormick [11].

This paper is divided into six sections. In sections 2, 3 and 4 we discuss the unified interior point approach to the solution of the linear, quadratic and general nonlinear programming problems, while in section 5, we compare the KKT requirements for optimality with the unified form of interior point methods and section 6 gives a conclusion of the study.

2. Linear Programming

Armed with Newton's method for unconstrained minimization and the Lagrangian and barrier methods for converting constrained problems into unconstrained ones, we take a fresh look at linear programming and its relatives.

We consider the primal-dual pair of linear programming problem

$$\text{(P)} \quad \text{minimize} \quad c^T x \quad \text{subject to} \quad Ax = b, \quad x \geq 0 \tag{1}$$

where $A \in \mathcal{R}^{m\times n}$ is the full rank matrix of linear constraints and vectors x, c and b have appropriate dimensions.

$$\text{(D)} \quad \text{maximize} \quad b^T y \quad \text{subject to} \quad A^T y + z = c, \quad z \geq 0 \tag{2}$$

The equations can be handled by Lagrange's method and the nonnegativity conditions by Fiacco and McCormick's barrier method and the resulting unconstrained functions can be optimized by Newton's method [7].

The usual transformation in interior point methods consists in replacing inequality constraints with the logarithmic barrier and then form the Lagrangian [3, 4, 7] to get

$$\min \quad c^T x - \mu \sum_{j=1}^{n} \ln x_j \quad \text{s.t.} \quad Ax = b, \tag{3}$$

where $\mu \geq 0$ is a barrier parameter. The Lagrangian associated with the primal problem has the form

$$L(x, y, \mu) = c^T x - y^T (Ax - b) - \mu \sum_{j=1}^{n} \ln x_j \tag{4}$$

and the conditions for a stationary point given by

$$\begin{aligned} \nabla_x L(x, y, \mu) &= c - A^T y - \mu X^{-1} e = 0 \\ \nabla_y L(x, y, \mu) &= \quad Ax - b = 0, \end{aligned} \tag{5}$$

where $X^{-1} = \operatorname{diag}\{x_1^{-1}, x_2^{-1}, \dots, x_n^{-1}\}$.

Having denoted

$$s = \mu X^{-1} e, \qquad \text{i.e.} \qquad XSe = \mu e,$$

where $S = \operatorname{diag}\{s_1, s_2 \dots, s_n\}$ and $e = (1,1 \dots, 1)^T$, the first order optimality conditions (for the barrier problem) are:

$$\begin{aligned} Ax &= b, \\ A^T y + s &= c, \\ XSe &= \mu e \\ (x, s) &\geq 0. \end{aligned} \tag{6}$$

Interior point algorithm for linear programming [9] applies Newton method to solve this system of nonlinear equations and gradually reduces the barrier parameter μ to guarantee the convergence to the optimal solution of the original problem. The Newton direction is obtained by solving the system of linear equations:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} \xi_p \\ \xi_d \\ \xi_\mu \end{bmatrix}, \tag{7}$$

where

$$\xi_p = b - Ax, \qquad \xi_d = A^T y - s, \qquad \xi_\mu = \mu e - XSe.$$

By elimination of

$$\Delta s = A^{-1}(\xi_\mu - S\Delta x) = -X^{-1} S \Delta x + X^{-1} \xi_\mu,$$

from equation (7) we get the symmetric indefinite augmented system of linear equations

$$\begin{bmatrix} -\ominus_p^{-1} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \xi_d - X^{-1}\xi_\mu \\ \xi_p \end{bmatrix}. \tag{8}$$

where $\ominus_p = XS^{-1}$ is a diagonal scaling matrix. By eliminating Δx from equation (6) we can reduce (8) further to the form of normal equations

$$\left(A \ominus_p A^T\right) \vartriangle y = \boldsymbol{b}_{LP}.$$

3. Quadratic Programming

Following the presentation in [4] we consider the convex quadratic programming problem

$$\min \quad c^T x + \frac{1}{2} x^T Q x \quad \text{s.t.} \quad Ax = b, \quad x \geq 0, \tag{9}$$

where $A \in \mathcal{R}^{n\times n}$ is positive semidefinite matrix, $A \in \mathcal{R}^{m\times n}$ is the full matrix of linear constraints and vectors x, and b have appropriate dimensions. The inequality constraints are again replaced with the logarithmic barriers

$$\min \quad c^T x + \frac{1}{2} x^T Q x - \sum_{j=1}^{n} \ln x_j$$
$$\text{s.t.} \qquad Ax = b, \qquad (10)$$

where $\mu \geq 0$ is a barrier parameter and the associated Lagrangian has the form:

$$L(x, y, \mu) = c^T x + \tfrac{1}{2} x^T Q x - y^T(Ax - b) - \mu \textstyle\sum_{j=1}^{n} \ln x_j, \qquad (11)$$

The conditions for a stationary point [3], are as usual

$$\begin{aligned} \nabla_x L(x, y, \mu) &= c - A^{\mathrm{T}} y - \mu X^{-1} e + Qx = 0 \\ \nabla_y L(x, y, \mu) &= \qquad Ax - b = 0. \end{aligned} \qquad (12)$$

With the usual notation for diagonal matrices: $X^{-1} = diag\{x_1^{-1}, x_2^{-1}, \dots, x_n^{-1}\}$ and $S = diag\{s_1, s_2 \dots, s_n\}$, the first order optimality conditions (for the barrier problem) are:

$$\begin{aligned} Ax \qquad &= b, \\ A^T y + s - Qx &= c, \\ XSe &= \mu e \\ (x, s) &\geq 0. \end{aligned} \qquad (13)$$

Interior point algorithm for quadratic programming [3, 9] applies Newton method to solve this system of nonlinear equations and gradually reduces the barrier parameter μ to guarantee the convergence to the optimal solution of the original problem. The Newton direction is obtained by solving the system of linear equations:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} \xi_p \\ \xi_d \\ \xi_\mu \end{bmatrix}, \qquad (14)$$

where

$$\xi_p = b - Ax, \qquad \xi_d = A^T y - s + Qx, \qquad \xi_\mu = \mu e - XSe.$$

By elimination of

$$\Delta s = A^{-1}(\xi_\mu - S\Delta x) = -X^{-1} S \Delta x + X^{-1} \xi_\mu,$$

from equation (14) we get the symmetric indefinite augmented system of linear equations

$$\begin{bmatrix} -Q - \Theta_p^{-1} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \xi_d - X^{-1}\xi_\mu \\ \xi_p \end{bmatrix}. \qquad (15)$$

where $\Theta_p = XS^{-1}$ is a diagonal scaling matrix. By eliminating Δx from the first equation we could reduce (15) further to the form of normal equations

$$(A(Q + \Theta_P^{-1})^{-1} A^T) \triangle y = \boldsymbol{b}_{QP}. \qquad (16)$$

Nonlinear Programming

We begin by considering the convex optimization problem

$$\begin{aligned} \min \qquad & f(x) \\ \text{s.t.} \qquad & g(x) \leq 0, \end{aligned} \qquad (17)$$

where $x \in \mathcal{R}^n$, $f\colon \mathcal{R}^n \mapsto \mathcal{R}$ and $\mathrm{g}\colon \mathcal{R}^n \mapsto \mathcal{R}^m$ are convex, twice differentiable. Replacing the inequality constraints with an equality $\mathrm{g}(x) + z = 0$, where $z \in \mathcal{R}^m$ is a nonnegative slack variable, we can formulate the associated barrier problem

$$\min \quad f(x) - \mu \sum_{j=1}^{n} \ln z_j$$
$$\text{s.t.} \qquad g(x) + z = 0. \qquad (18)$$

and obtain the Lagrangian as

$$L(x, y, z, \mu) = f(x) + y^T(g(x) + z - \mu \textstyle\sum_{i=1}^{m} \ln x_{j.} \qquad (19)$$

The condition for a stationary point given by

$$\begin{aligned} \mathrm{L}(x, y, \mu) &= c - A^T y - \mu X^{-1} e = 0 \\ \nabla_y L(x, y, z, \mu) &= \qquad g(x) + z = 0 \\ \nabla_z L(x, y, z, \mu) &= \qquad y - \mu Z^{-1} e = 0, \end{aligned} \qquad (20)$$

$where$ $Z^{-1} = diag\{z_1^{-1}, z_2^{-1}, \dots, z_m^{-1}\}$. The first order optimality conditions (for the barrier problem) have the form

$$\begin{aligned} &\nabla_x f(x) + \nabla g(x)^T y = 0, \\ &g(x) + z \qquad\quad = 0, \\ &YZe \ = \qquad\quad = \mu e \\ &(y,z) \qquad\qquad \geq 0 \end{aligned} \tag{21}$$

where $Y = diag\{y_1, y_2 \ldots, y_m\}$. Interior point algorithm for linear programming, Nocedal and Wright (2006) applies Newton method to solve this system of equations and equally reduces the barrier parameter μ to guarantee the convergence to the optimal solution of the original problem. The Newton direction is obtained by solving the system of linear equations:

$$\begin{bmatrix} Q(x,y) & A(x)^T & 0 \\ A(x) & 0 & I \\ 0 & Z & Y \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} -\nabla f(x) - A(x)^T y \\ -g(x) - z \\ \mu e - YZe, \end{bmatrix}, \tag{22}$$

where

$$A(x) = \nabla g(x) \ \in \mathcal{R}^{m \times n}$$

$$Q(x,y) = \nabla^2 f(x) + \sum_{i=1}^{m} y_i \nabla^2 g_i(x) \qquad \in \quad \mathcal{R}^{n \times n} \ .$$

Using the third equation in (22) we eliminate

$$\Delta z = \mu Y^{-1} e - Ze - ZY^{-1}\Delta y,$$

from the second equation and get

$$\begin{bmatrix} -Q(x,y) & A(x)^T \\ A(x) & \ominus_D \end{bmatrix} \begin{bmatrix} \Delta x \\ -\Delta y \end{bmatrix} = \begin{bmatrix} \nabla f(x) + A(x)^T y \\ -g(x) - \mu Y^{-1} e \end{bmatrix}, \tag{23}$$

where $\ominus_D \ ZY^{-1}$ is a diagonal scaling matrix. The matrix involved in this set of linear equation is symmetric and indefinite. For convex optimization problem (that is, when f and g are convex), the matrix Q is positive semindefinite and if f is strictly convex, Q is positive definite. Similarly to the case of quadratic programming by eliminating Δx from the first equation in (14) we could reduce this system further to the form of normal equations

$$(A(x)Q(x,y)^{-1})A(x)^T + ZY^{-1})\Delta y = \boldsymbol{b}_{NLP}. \tag{24}$$

Hence, the application of the lagrangian method coupled with the logarithm barrier approach lead to the system of equations for which the Newton method is applicable for their solutions

4. The K-K-T Conditions.

Following the work of Q_i and Jiang (1997), the nonlinearly constrained programming problem (NLP)

$$\text{Min } \{f(x): g(x) \leq 0, h(x) = 0\} \tag{25}$$

where f, g *and* h are continuously differentiable functions from $\mathcal{R}^n$ to $\mathcal{R}$, $\mathcal{R}^p$ and $\mathcal{R}^q$ respectively.

Let $N = n + p + q$.

The Karush-Kuhn-Tucker (KKT) system for this problem is:

$$\nabla f(x) + \textstyle\sum_{j=1}^{p} u_j \nabla g_j + \sum_{j=1}^{q} \vartheta_j \nabla h_j(x) \tag{26}$$

$$u \geq 0, \quad g(x) \leq 0, u^T g(x) = 0, h(x) = 0.$$

The KKT system plays a central role in the theory and algorithm for the NLP in (25).

In [10] two forms of the NLP is considered. One form is that f, g and h are twice continuously differentiable and $\nabla^2 f$, $\nabla^2 g$, $\nabla^2 h$ are locally Lipschitzian. Such an NLP are referred to as LC^2 NLP or simply and LC^2 problem [10]. Similarly, smooth functions are referred to as an LC^1 function if its derivative is locally Lipschitzian.

The case when f, g and h are not necessary twice continuously differentiable but their derivatives are semismooth are called an NLP and SC^1 NLP or simply an SC^1 problem, Qi and Jiang (1997). The applications of the SC^1 problem include the LC^2 problem; the stochastic quadratic program and the minimax problem (see [10] and quoted references therein).

Different methods have been developed to formulate the KKT system as a system of nonsmooth equations (NE) (see [10], [3], [4], [7]). A recent approach to construct generalizations of classical Newton and quasi-Newton methods for solving these nonsmooth KKT equations are due to several authors (see [10] and references quoted). Such methods are called NE methods for solving the NLP defined in (25). In 1994, Pang [32] constructed a sequential quadratic such problems and merit functions with these equivalent KKP equations. These methods are called NE/SQP methods. It is also possible to form smooth KKT equations; however a drawback is that these smooth KKT equations may be singular or ill-conditioned at a solution if strict

Complementarity is not satisfied. On the other hand, singularity occurs less for nonsmooth versions of KKT equations.

Variants of generalized Newton methods have been developed for nonsmooth KKT equations. Some of them are Q-quadratically convergent under suitable conditions for the LC^2 problem [1, 4]. However, a general theory is still needed for Q-quadratic convergence of generalized Newton methods for nonsmooth KKT equations Q_i and Jiang (1997). Similarly, Negin and Nezan (2014) indicate that effective methods for computing the positive definite matrix satisfying KKT conditions have yet to be developed. However, the new matrix-free method in which the normal equations are solved by the conjugate gradient method may have significant computational advantage over the simplex for linear programming instance, see [13].

5. Conclusion

We observe that each iteration of interior point method for linear, quadratic or nonlinear programming requires the solution of a possibly large and most always sparse linear system of the form

$$\begin{bmatrix} -Q - \ominus_p^{-1} & A^T \\ A & \ominus_D \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \xi_d - X^{-1}\xi_\mu \\ \xi_p \end{bmatrix}. \quad (27)$$

Therefore the three symmetric indefinite augmented system of linear equations formed for linear, quadratic and nonlinear programming have essentially the same structure, and only the diagonal scaling matrix $\ominus_p = XS^{-1}$ changes from iteration to iteration.

In expression (27), $\ominus_p \in \mathcal{R}^{n\times n}$ and $\ominus_D \in \mathcal{R}^{m\times m}$ are diagonal scaling matrices with strictly positive elements. Depending on the problem type, one or both matrices $\ominus_p$ and $\ominus_D$ may be present in this system.

For linear and quadratic programs with equality constraints, $\ominus_D = 0$. For nonlinear programs with inequality constraints (and variables without sign restriction) $\ominus_p^{-1} = 0$. With assumption of convexity, the Hessian $Q \in \mathcal{R}^{n\times n}$ is a symmetric positive definite matrix, where $A \in \mathcal{R}^{m\times n}$ is the matrix of linear constraints (or the linearization of nonlinear constraints), with full rank.

In the area of implementation, the most surprising and important characteristic of the interior point methods is that the number of iterations required is very insensitive to problem size [7]. Given this behavior, the advantage of an interior method as compared to the simplex method depends on how efficiently the individual steps of the interior method can be executed.

The main weakness of the interior point methods involves the issue of dense columns. For example, a single column in A with a nonzero in every will cause AA^T to be completely dense. However, a small number of dense columns (say 10) can be handled successfully, but that a large number presents a very real stumbling block to the use of interior point methods [10]. However, further research on the new matrix-free method in which the normal equations are solved by the conjugate gradient method may offer some advantage as noted in [13].

Acknowledgement
We gratefully acknowledge the comments and suggestions of anonymous Reviewers for the improvement of this paper.

REFERENCES

[1] Bertsekas, D. P. (2004): Lagrange Multipliers with Optimal Sensitivity Properties in Constrained Optimization, Report LIDS **2632,** Dept. of Electrical Engineering and Computer Science, M.I.T., Cambridge, Mass.

[2 Deza, A., Nematollahi, E. and Terlaky, T.(2004): How good are interior point methods? Klee-Minty cubes tighten iteration-complexity bounds, AdVOL – Report # **2004/20**. Advanced Optimization Laboratory, Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada

[3 Fiacco, A. V. and McCormick, G. P. (1968): Non-linear Programming: Sequential Unconstrained Minimization Techniques, John Wiley and Sons, New York.

[4 Gondzio J. and Grothey A. (2004): Exploiting Structure in Parallel Implementation of Interior Point Methods for Optimization, Technical Report **MS04 – 00**4, School of Mathematics, The University of Edinburgh, Edinburg E H9 3JZ, UK

[5 Karmarkar, N. (1984): A new polynomial-time algorithm for linear programming, Combinatorica **4**, 373-395

[6] Lesaja, G. (2009): Introducing Interior-Point Methods for Introductory Operations Research Courses and/or Linear Programming Courses, The Open Operational Research Journal, **3**, 1-12

ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.5, No.12, 2015

[7] Marsten, R., Subramanian, R,, Saltzman, M., Lustig, I. and Shanno, D. (1990): Interior Point Methods for Linear Programming: Just Call Newton, Lagrange, and Fiacco and McCormick!, Interfaces,**20**: pp.105-116.
[8] Newton, I. (1687): Principia, London, England
[9] Nocedal, J. and Wright, S. J.(2006): Numerical Optimization, Springer Series in Operations Research: Springer, New York
[10] Qi, L. and Jiang, H., Semi-smooth Karush-Kuhn-Tucker equations and convergence analysis of Newton and quasi- Newton methods for solving these equations, Mathematics of Operations Research, 22(2) (1997) 301-325.
[11 Tanoh, G., Renard, Y. and Noll, D. (…): Computational Experience with an Interior point Algorithm for Large Scale Contact Problems, Research Memoranda, Mathematiques pour l'Industrie et la Physique, UMR CNRS 5640, Universite Paul Sabatier, Toulouse cedex 4, France, e-mail; tanoh@cict.fr
[12] Vandenberghe, L and Boyd, S (1994): A primal-dual potential reduction method for problems involving matrix inequalities, Internet: boyd@isl.stanford.edu
[13] Zukowski, M., Laskowski, W., Hall, J. A. J., Gruca, J. A., and Gondzio J. (2014): Solving Large-Scale Optimization Problems Related to Bell's Theorem, Journal of Computational and Applied Mathematics.