INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 7, Issue 3 - V7I3-1347)
Available online at: https://www.ijariit.com

A comparative analysis of interior-point methods in convex optimization

Mohd Aquib
aquib01011992@gmail.com
Indian Institute of Technology, Kanpur, Uttar Pradesh

ABSTRACT

Interior point methods or sometimes called Barrier methods are class of optimization algorithms that are useful in solving inequality constrained linear or nonlinear programming problems. The optimization problem with inequality constraint is solved by using Newtons method by converting it into series of equality constraint or KKT conditions. This paper presents the comparative study of the classical logarithmic barrier method and the primal-dual interior point method along with their convergence analysis. A kernel function based interior point method for second order cone problems is also examined. Computational complexity of the algorithms is also discussed along with the implementation issues.

Keywords*: Interior point methods, KKT condition, Kernel function.*

1. INTRODUCTION

Interior point methods have been in practice for the past few decades for solving various class of optimization problems including both convex and non-convex categories. Its history dates back to the early 1960s when Fiacco and Mccormick proposed a method which used logarithmic barrier function for solving nonlinear programming problems. But it was the effort of Karmarkar in 1984, who developed a projective method for solving linear programs [1], which proved a boon to the researchers working in the field of interior point methods. Karmarkar's algorithm provided a polynomial time complexity for linear programming problems. As a result, over 1300 research papers were published since 1984 in the area of interior point methods [1], which proved its betterment over the classical simplex algorithm.

Further probe into the area of interior point methods, led to the development of series of variants of the barrier method. The primal log barrier method for linear programming was developed by Gill et al, which provided a faster convergence rate. Some years later its dual problem was formulated and solved with O ($\sqrt{nL}$) complexity bound. Subsequently the primal-dual point method was devised using the previous variants of the method. This method provided a faster convergence rate.

At present the interior point method has progressed a lot in the area of linear as well as convex programming with high quality codes and software to solve the problem. Current implementations are mostly based on Mehrotra's Predictor Corrector method (1992) involving Cholesky factorization along with Newton's method to estimate penalty parameters. The primal dual interior point method has now been extended for solving second order cone problems based on kernel functions. IPMs have proved unparallel in solving semidefinite programming problems. Software packages like SeDuMi [Stu99], SDPT3 [TTT02], SDPA [FKN98], CSDP [Bor02] [5] etc are very useful in solving such problems. This paper is structured to give a comparative analysis of the two widely used interior point methods namely Barrier method and its primal-dual version. Section 2 of the paper presents an analytical frame of the two methods along with their convergence study. Section 3 deals with kernel function based IPM for SOCP along with its complexity. Section 4 into account the issues which are confronted while implementing interior point methods. Finally, section 5 provides the conclusion and future scope of research in the area of IPMs.

2. INTERIOR POINT METHODS

2.1. The Barrier Method

The barrier method also called as path following method was developed by Fiacco and Mccormick [5] in the 1960s. It was mainly focussed on solving nonlinear optimization problems. Barrier method mainly uses an indicator function, on the inequality constraint, which gets absorbed in the objective function, yielding an equality constraint optimization problem. The most common and useful approximation of the indicator function is logarithmic function called as Logarithmic Barrier Function.

1) Problem formulation and working procedure*:* Consider a simple optimization problem as under:

$$\begin{aligned} &\min_{x} f_o(x) \\ \textbf{subject to} \quad & f_i(x) \le 0, \qquad i = 1, \dots . m \\ & Ax = b \end{aligned} \tag{1}$$

Where $f_o(x)$ and $f_i(x)$ are convex and twice differentiable objective and inequality constraint functions. Applying the concept of barrier function and central path to (1) gives an equality constraint problem of the form,

$$\begin{aligned} &\underset{x,t}{\textbf{minimize}} \;\; t f_o(x) + \phi(x) \\ &\textbf{subject to} \quad Ax = b \end{aligned} \tag{2}$$

For $t > 0, x^*(t)$ is the solution of (2) and the collection of points $x^*(t), t > 0$ are called the central points. The logarithmic barrier function $\phi(x)$ is given by

$$\phi(x) = -\sum_{i=1}^{n} \log(-f_i(x)) \tag{3}$$

Now equation (2) can be solved using Newton's method as follows;

Step 1 – Initializing: Choose $t := t^{(0)} > 0, \mu > 1$, and tolerance $\epsilon > 0$ for strictly feasible x.

Step 2 – Centring: Compute $x^* = argmin\big(t f_o(x) + \phi(x)\big)\ subject\ to\ Ax = b$.

Step 3 – Stopping: If $(m/t) < \epsilon$ then stop.

Step 4 – Incrementing: Else set $t = \mu t$ and go to step 2.

2) Convergence Analysis: The major criteria over which an algorithm's success is judged is its convergence. A fast-converging algorithm and small complexity bound is desired in any iterative process. It has been shown that the barrier method converges under some assumptions. If we assume that equation (2) can be solved using Newton's algorithm then for $t = t^0, \mu t^0, \mu^2 t^0$ ... after *k* centring steps the duality gap ϵ becomes

$$\epsilon = \frac{m}{\mu^k t^0} \tag{4}$$

$$\begin{aligned} \mu^k &= \frac{m}{\epsilon t^0} \\ \log \mu^k &= \log\left(m/\epsilon t^0\right) \\ k &= \frac{\log\left(m/\epsilon t^0\right)}{\log \mu} \end{aligned} \tag{5}$$

which is the iteration complexity for barrier method. As the parameter t increases the centring problem becomes more complex because the number of iterations increases. In order to sort out this problem self-concordant assumptions are made which makes the bound more rigorous on total number of iterations required to solve the problem [5]. The assumptions are as follows:

(1) Its initial sublevel set should be closed.

(2) Its inverse KKT matrix should be bounded, i.e.

$$\left\| \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix}^{-1} \right\|_2 \le K \tag{6}$$

(3) It satisfies Lipschitz condition i.e.; it must be L smooth.

$$\|\nabla^2 f(x) - \nabla^2 f(\tilde{x})\|_2 \le L \|x - \tilde{x}\|_2 \tag{7}$$

(4) It must be strongly convex or μ convex.

$$\nabla^2 f(x) \ge \mu I, \;\; \mu > 0 \tag{8}$$

With these self-concordance assumptions, we can ensure an upper bound on the total number of steps required to solve the problem, which also guarantees a linear convergence of the barrier method.

2.2. Primal-Dual Interior Point Method

Primal-dual interior point methods are very closely related to the barrier method, with the difference that there is only one iteration in this method without any contrast between outer and inner iteration iterations. Each time an iteration is executed, the primal and dual variables are updated. The Newton's method when applied to modified KKT conditions, gives the search direction for the primal-dual interior point method. This method can also be used when the problem is not strictly feasible (but feasible). A better convergence rate of primal-dual interior point method over the barrier method makes it more efficient in solving several classes of problems like LPs, QPs, QCQPs, SOCPs and SDPs.

1) Problem formulation and working procedure: Consider an optimization problem given by:

$$\begin{gathered} \min_{x} \boldsymbol{f(x)} \\ \boldsymbol{subject\ to\ \ g(x) \leq 0} \\ \boldsymbol{Ax = b} \end{gathered} \tag{9}$$

Where *f* and *g* convex differentiable functions. Assuming strong duality the central path equations can be written as:

$$\begin{gathered} \boldsymbol{\nabla f(x) + \nabla g(x)\lambda + A^T \nu = 0} \\ \boldsymbol{Ug(x) + \tau 1 = 0} \\ \boldsymbol{Ax - b = 0} \\ \boldsymbol{\lambda, -g(x) > 0} \end{gathered} \tag{10}$$

Where λ and ν are the dual variables. Now writing the residuals for $w = (x, \lambda, \nu)$

$$\boldsymbol{r(x, \lambda, \nu) = r(w) := \begin{bmatrix} \nabla f(x) + \nabla g(x)\lambda + A^T \nu \\ Ug(x) + \tau 1 \\ Ax - b \end{bmatrix}} \tag{11}$$

Since solving the centrality equation by Newton's method forms the basis of primal-dual method, we have to compute Newton step $(\Delta x, \Delta\lambda, \Delta\nu)$ by solving the matrix equation given by:

$$\begin{bmatrix} \nabla^2 f(x) + \sum_i \lambda_i \nabla^2 g_i(x) & \nabla g(x) A^T \\ U\nabla g(x)^T & G(x) \\ A & \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta\nu \end{bmatrix} = -r(x, \lambda, \nu) \tag{12}$$

Finally for a step length α the update is given by:

$$(x^*, \lambda^*, \nu^*) = (x, \lambda, \nu) + \alpha(\Delta x, \Delta\lambda, \Delta\nu) \tag{13}$$

Here the iterates $x^{(k)}, \lambda^{(k)}, \nu^{(k)}$ may or may not be feasible, unless the algorithm converges to a optimal value. Hence, we define the surrogate duality gap [5] for primal-dual IPM as under:

$$\hat{\eta}(x, \lambda) = -f(x)^T \lambda \tag{14}$$

The process terminates when the values (x, λ, ν) becomes primal and dual feasible and the surrogate duality gap becomes less than the tolerance ϵ. Generally, the tolerance value is taken to be smaller than the barrier method.

2) Convergence Analysis

The convergence of primal-dual interior point method follows from the line search called as backtracking based on l_2 norm regularization of the residuals. This often results in super linear asymptotic convergence. If we denote the step size by *s*, then the largest value of *s* [5] such that $\lambda^* \succcurlyeq 0$ is:

$$s^{max} = sup\{s \in [0, 1] | \lambda + s\Delta\lambda \succcurlyeq 0\} \tag{15}$$

$$s^{max} = \min\{1, min\{-\lambda_i / \Delta\lambda_i \,|\, \Delta\lambda_i < 0\}\} \tag{16}$$

Now to ensure convergence we minimize the l_2 norm of the residuals by multiplying the step size *s* by regularizing parameter β until we achieve the bound [5] given by:

$$\|r_t(x^*, \lambda^*, \nu^*)\|_2 \leq (1 - \alpha s)\|r_t(x, \lambda, \nu)\|_2 \tag{17}$$

3) Convergence Comparison

The following figures taken from [5] shows the variations of the duality gap with the total number of Newton iterations required.

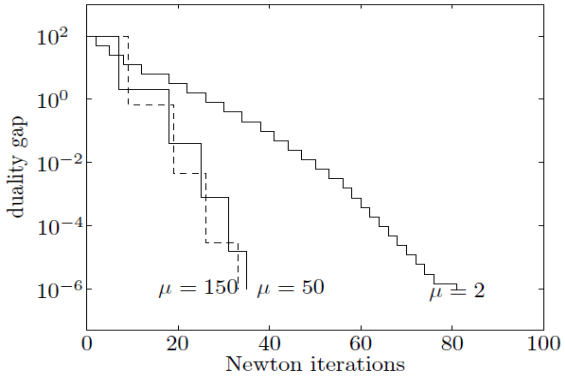

Fig 1. Barrier duality gap

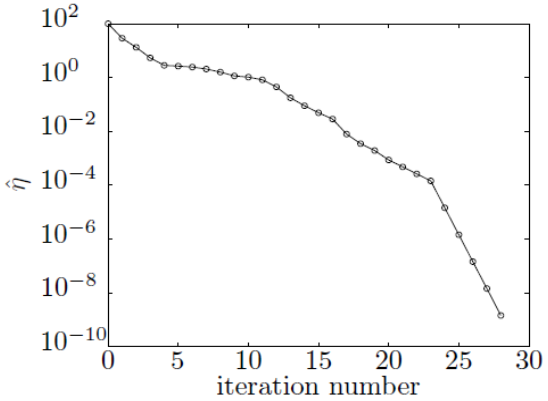

Fig 2. Primal-dual surrogate duality gap

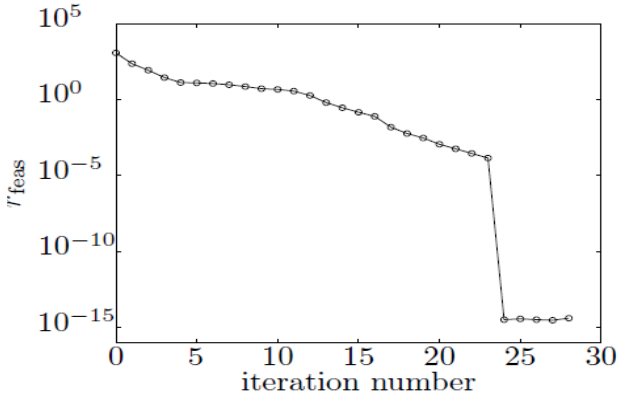

Fig 3. Primal-dual feasibility gap

3. KERNEL FUNCTION BASED IPM FOR SOCPs

Kernel functions are a class of functions that are used to solve nonlinear programs using linear classifier. Kernel function maps original nonlinear observations to a space in higher dimension where they become separable. Authors of [3] have analysed the performance of a kernel function as barrier function in primal dual interior point algorithm for solving second order cone problems. The kernel function used by the author is given by:

$$\psi(t) = t - 1 + \frac{t^{1-q} - 1}{q - 1}, t > 0, q \geq 2 \tag{18}$$

This kernel function was proposed by Bai and Roos [6] in the year 2002 which provided a linear growth rate.

3.1 Primal-dual interior point algorithm for SOCPs

After analysing the search direction for SOCP, the authors formulated the barrier function based on the given kernel function for primal-dual algorithm.

$$\Psi(x, s, \mu) := \Psi(v) = \sum_{i=1}^{n} \mathrm{Tr}(\psi(v^i)) \tag{19}$$

where x and s are optimization variables and μ is the barrier parameter. The algorithm [3] can now be written as below:

Primal-Dual interior point algorithm for SOCP

Input:

A threshold parameter $\tau > 1$;
An accuracy parameter $\epsilon > 0$;
A fixed barrier update parameter $\theta \in (0,1)$;
A strictly feasible pair (x^o, s^o) *and* $\mu^o > 0$ *such that* $\Psi(x^o, s^o, \mu^o) \leq \tau$.

Begin
$x := x^o, s := s^o, \mu := \mu^o$
While $n\mu \geq \epsilon$ **do**
Begin
$\mu := (1 - \theta)\mu$;
While $\Psi(x, s, \mu) > \tau$ **do**
Begin
$x := x + \alpha\delta; s := s + \alpha\delta s; y := y + \alpha\delta\, y$;

End
End
End

3.2 Convergence Analysis

The inner iteration bound of the given algorithm can be calculated by counting the total number of inner iterations such that $\Psi(v) \leq \tau$, after μ update. If Ψ_k denote the subsequent values in the same outer iteration for k = 1, 2,...T where T is the total number of inner iterations in the outer iterations then we can show that

$$\boldsymbol{T \leq 78q^2\left(\frac{\theta\sqrt{n}+\sqrt{\frac{\tau^2}{4n}+\tau}}{\sqrt{1-\theta}}\right)} \quad \textbf{(20)}$$

Since the number of outer iterations is bounded by $\frac{1}{\theta}\log\frac{n}{\epsilon}$ the final expression of upper bound for total number of iterations is given by:

$$\boldsymbol{78q^2\left(\frac{\theta\sqrt{n}+\sqrt{\frac{\tau^2}{4n}+\tau}}{\sqrt{1-\theta}}\right)\frac{1}{\theta}\log\frac{n}{\epsilon}} \quad \textbf{(21)}$$

Since q is a constant independent of n, the iteration bound for large update methods becomes $O\left(n\log\frac{n}{\epsilon}\right)$ For small update taking $\theta = \Theta\left(1/\sqrt{n}\right), q = 2$ and $\tau = O(1)$ the complexity becomes $O\left(\sqrt{n}\log\frac{n}{\epsilon}\right)$ which is same as the classical best-known complexity bound for interior point algorithms.

4. ISSUES IN IMPLEMENTIING INTERIOR POINT METHODS

This section describes the challenges that are encountered in implementing the interior point methods. They can be summarized as below:

4.1 Cholesky Decomposition

The major computational part in most of the interior point algorithms is to solve system of linear equations with sparse positive definite matrix, which is performed by Cholesky decomposition. When the solution approaches to the optimal, ill conditioning of these system of equations occurs which results in the breakdown of Cholesky factorization. Cholesky factorization is generally done by computing a lower triangular matrix L such that, where is diagonal matrix whose entries are given by. The computing cost of Cholesky factorization of a dense matrix is (1/3)n3 flops, half the cost of LU factorization. The breakdown generally arises when the pivot elements become very small or zero during factorization. This problem was identified by several researchers who developed modified Cholesky factorization in which they skipped factors related to small pivots by setting large pivot values called as Pivot Boosting or Regularization technique.

4.2 Problem Size

For interior point methods to solve large linear programs, requires a large run time. For faster implementation of computational linear algebra, the problem must be pre-processed in order to reduce its size. The pre-processor used in many software like OB1, SeDuMi etc aim to reduce problem size by:
1) Removing empty rows that are redundant in the large sized matrix, which make the problem infeasible.
2) Problem infeasibility is determined by infeasible bounds.
3) Dual infeasibility is determined by monitoring dual redundancies.
4) Identifying redundant rows by substituting lower and upper bounds for row variables.
5) Columns that contain only one non zero coefficient are used to change row type or eliminate the variable from the model.

4.3 Starting Point

Initial estimate of the starting point x(0) is still a challenging issue in all interior point algorithms. An arbitrary point cannot form a successful start of interior point algorithm. An ideal starting point should possess the following;
1) It should be close to primal and dual feasible.
2) It should be centred properly.
3) It should be close to the optimal solution.

Most of the interior point codes selects the starting point by taking into account the equality constraint and leaving behind the inequality constraint. As a result, the optimality of the solution is altered. Several solutions have been proposed by researchers in the context of starting point of interior point methods, that satisfy all the three conditions. Crash start technique proposed by Jack Gondzio in 2015, provided a better starting point for medium sized problems and saves 20-40% of iterations required by primal dual algorithm.

5. CONCLUSION

In this paper the most widely used interior point methods have been discussed, presenting a detailed overview of the barrier method and its primal dual version. The analysis shows that the primal-dual interior point methods are quite efficient in solving large nonlinear convex optimization problems, owing to its better convergence rate.

As compared to the barrier method the primal dual version possess several advantages one of them being the ability to work in a situation when the problem is not is not strictly feasible. This allows it to extend its range in solving second order cone problems and semidefinite programming problems. The kernel function-based approach of the IPM for SOCPs, shows the same complexity bound as the classical barrier method. Challenges in implementing the IPMs discussed at the end reveals that, though progress have been made to reduce the problem size and lessen the computational time, the choice of starting point infeasibility issues still remain a topic of active research. Commendable progress has been made in the recent years in the area of convergence of primal-dual variants of the interior point methods. Improvements in primal and dual regularization in the context of line search can be done by careful and precise update of the dual variable. Such improvement can make the solver more robust.

6. REFERENCES

[1] Lustig, Irvin and Marsten, Roy and F. Shanno, David. (1994). "Feature Article-Interior Point Methods for Linear Programming: Computational State of the Art." INFORMS Journal on Computing 6. 1-14. 10.1287/ijoc.6.1.1.
[2] Y.Q. Bai, G.Q. Wang, C. Roos, "Primal-dual interior-point algorithms for second-order cone optimization based on kernel functions", Nonlinear Analysis, vol. 70, pp. 3584-3602, 2009.
[3] Dong, Li and Tang, Jingyong, "Interior-point method for second-order cone programming based on a simple kernel function", in2nd IEEE International Conference on Computational Intelligence and Natural Computing CINC. 2010. 5643888.
[4] Nemirovski, Arkadi and Todd, Michael. (2008). "Interior point methods for optimization", Acta Numerica. 17.191-234. 10.1017/S0962492906370018.
[5] S.Boyd and L.Vandenberghe, "Convex Optimization", Cambridge University Press,2009 [Online]. http://www.stanford.edu/boyd/cvxbook/
[6] Y.Q.Bai and C.Roos, "A primal-dual interior-point method based on a new kernel function with linear growth rate", in Proceedings of Industrial Optimization Symposium and Optimization Day, Australia, November 2002.
[7] Behrouz Kheirfam. "Primal-dual interior-point algorithm for semidefinite optimization based on a new kernel function with trigonometric barrier term", in Numerical Algorithms, March, 2012.
[8] Lecture notes on "Primal-dual interior point methods", [online]. http://www.users.ece.cmu.edu/

BIOGRAPHY

Mohd Aquib
Research Scholar,
Indian Institute of Technology, Kanpur, Uttar Pradesh, India

