A one phase interior point method (IPM) for non-convex optimization

Oliver Hinder, Yinyu Ye

Stanford University

ohinder@stanford.edu

June 9, 2017

Outline

1. Introduction
   - Local infeasibility certificate(s)
   - Current non-convex IPMs
   - Algorithm outline
2. Empirical results
   - Starting point
   - Netlib results
   - CUTEst results

Problem we wish to solve:

$$\min f(x)$$
$$a(x) \leq 0$$

Assume:

1. The constraints and objective are C^2
2. The set $a(x) \leq \theta$ is bounded for any $\theta \in \mathbb{R}^m$

Local infeasibility certificate(s)

1. *First-order local L1-infeasibility certificate*, if x^* is a first-order local optimum of:

$$\min e^T (a(x))^+$$

With $e^T (a(x^*))^+ > 0$.

2. *First-order local farkas infeasibility certificate*, if there exists $y \geq 0$ such that x^* is a first-order local optimum of:

$$\min y^T a(x)^+$$

With $y^T a(x^*)^+ > 0$.

One phase IPMs for conic optimization

Either declare problem infeasible (primal or dual) or optimal in one phase:

1. Homogenous algorithm
   [Ye et al., 1994, Andersen and Ye, 1999]
2. Infeasible start IPM
   [Lustig, 1990, Mehrotra, 1992, Todd, 2002]

Failure of infeasible start IPM for non-convex optimization

[Wächter and Biegler, 2000] showed that if we apply an infeasible start IPM to:

$$
\begin{aligned}
\min x \\
x^2 - s_1 - 1 = 0 \\
x - s_2 - 1/2 = 0 \\
s_1, s_2 \geq 0
\end{aligned}
$$

Fails to converge to either a local optimum or infeasibility certificate

Modifications to the infeasible start IPM

1. Two phases (IPOPT)
2. Compute two directions (KNITRO)
3. Penalty (or big-M) method e.g. [Chen, 2006, Curtris, 2012]

Modifications to the infeasible start IPM

1. Two phases (IPOPT)
2. Compute two directions (KNITRO)
3. Penalty (or big-M) method e.g. [Chen, 2006, Curtris, 2012]

Modifications to the infeasible start IPM

1. Two phases (IPOPT)
2. Compute two directions (KNITRO)
3. Penalty (or big-M) method e.g. [Chen, 2006, Curtris, 2012]

Modifications to the infeasible start IPM

1. Two phases (IPOPT)
2. Compute two directions (KNITRO)
3. Penalty (or big-M) method e.g. [Chen, 2006, Curtris, 2012]

It doesn't need to be this complicated!

One phase barrier problem

$$\min f(x)$$
$$a(x) \leq 0$$

1. Add a slack variable
2. Add a log barrier term
3. Keep the constraint violation the same
4. Reduce the constraint violation and μ^k by $\eta^k \in [0, 1]$
5. Add proximal term

One phase barrier problem

$$\min f(x)$$
$$a(x) + s = 0$$
$$s \geq 0$$

1. Add a slack variable
2. Add a log barrier term
3. Keep the constraint violation the same
4. Reduce the constraint violation and μ^k by $\eta^k \in [0, 1]$
5. Add proximal term

One phase barrier problem

$$\min f(x) - \mu^k \sum_i \log s_i$$
$$a(x) + s = 0$$
$$s \geq 0$$

1. Add a slack variable
2. Add a log barrier term
3. Keep the constraint violation the same
4. Reduce the constraint violation and μ^k by $\eta^k \in [0,1]$
5. Add proximal term

One phase barrier problem

$$\min f(x) - \mu^k \sum_i \log s_i$$
$$a(x) + s = (a(x^k) + s^k)$$
$$s \geq 0$$

1. Add a slack variable
2. Add a log barrier term
3. Keep the constraint violation the same
4. Reduce the constraint violation and μ^k by $\eta^k \in [0, 1]$
5. Add proximal term

One phase barrier problem

$$\min f(x) - (1-\eta^k)\mu^k \sum_i \log s_i$$
$$a(x) + s = (1-\eta^k)(a(x^k) + s^k)$$
$$s \geq 0$$

1. Add a slack variable
2. Add a log barrier term
3. Keep the constraint violation the same
4. Reduce the constraint violation and μ^k by $\eta^k \in [0, 1]$
5. Add proximal term

One phase barrier problem

$$\min f(x) - (1-\eta^k)\mu^k \sum_i \log s_i + \frac{\delta^k}{2}\|x - x^k\|^2$$

$$a(x) + s = (1-\eta^k)(a(x^k) + s^k)$$

$$s \geq 0$$

1. Add a slack variable
2. Add a log barrier term
3. Keep the constraint violation the same
4. Reduce the constraint violation and μ^k by $\eta^k \in [0,1]$
5. Add proximal term

Primal-dual direction computation

$$\nabla f(x) + \delta^k (x - x^k) + y^T \nabla a(x) = 0$$
$$a(x) + s = (1 - \eta^k)(a(x^k) + s^k)$$
$$s_i y_i = (1 - \eta^k)\mu^k$$
$$s, y \geq 0$$

1. Form KKT system for barrier problem
2. Linearize the KKT system
3. Factorize matrix and compute direction d^k

Primal-dual direction computation

$$\begin{bmatrix} H^k & \nabla a(x^k)^T & 0 \\ \nabla a(x^k) & 0 & I \\ 0 & S^k & Y^k \end{bmatrix} \begin{bmatrix} d_x^k \\ d_y^k \\ d_s^k \end{bmatrix} = \begin{bmatrix} -(\nabla f(x^k) + \nabla a(x^k)^T y^k) \\ -\eta^k (a(x^k) + s^k) \\ (1 - \eta^k)\mu^k e - Y^k s^k \end{bmatrix}$$

$$H^k = \nabla_x^2 L(x^k, y^k) + \delta^k I$$

1. Form KKT system for barrier problem
2. Linearize the KKT system
3. Factorize matrix and compute direction d^k

Primal-dual direction computation

$$\begin{bmatrix} H^k & \nabla a(x^k)^T & 0 \\ \nabla a(x^k) & 0 & I \\ 0 & S^k & Y^k \end{bmatrix} \begin{bmatrix} d_x^k \\ d_y^k \\ d_s^k \end{bmatrix} = \begin{bmatrix} -(\nabla f(x^k) + \nabla a(x^k)^T y^k) \\ -\eta^k (a(x^k) + s^k) \\ (1 - \eta^k)\mu^k e - Y^k s^k \end{bmatrix}$$

$$H^k = \nabla_x^2 L(x^k, y^k) + \delta^k I$$

1. Form KKT system for barrier problem
2. Linearize the KKT system
3. Factorize matrix and compute direction d^k

Iterate update

For step size $\alpha^k \in [0, 1]$ update iterates as follows:

$$
\begin{aligned}
x^{k+1} &= x^k + \alpha^k d_x^k \\
y^{k+1} &= y^k + \alpha^k d_y^k \\
s^{k+1} &= (1 - \alpha^k \eta^k)(a(x^k) + s^k) - a(x^{k+1}) \\
\mu^{k+1} &= (1 - \alpha^k \eta^k)\mu^k
\end{aligned}
$$

Only accept iterates that approximately satisfy complementary:

$$
\beta \mu^{k+1} \leq s_i^{k+1} y_i^{k+1} - \mu^{k+1} \leq \frac{1}{\beta} \mu^{k+1}
$$

For some constant $\beta \in (0, 1)$.

One phase non-convex algorithm

for $k \leftarrow 1, \ldots, \infty$ **do**

if $\frac{\|\nabla f(x^k) + (y^k)^T \nabla a(x^k)\|}{\|y^k\| + 1} < \mu^k$ **then**

Aggressive step:

Compute direction with $\eta^k \in (0, 1]$

Take largest step while maintaining complementary

else

Stabilization step:

Compute direction with $\eta^k = 0$

Backtracking line search on direction using merit function

end if

end for

Termination criterion

First-order ϵ-locally optimal if:

$$\max \left\{ \frac{\|\nabla f(x) + y^T \nabla a(x)\|_\infty}{\|y\|_\infty + 1}, \|a(x) + s\|_\infty, \mu \right\} < \epsilon = 1e^{-6}$$

First-order ϵ-locally farkas infeasible if:

$$\frac{\mu + \|y^T \nabla a(x)\|}{y^T (a(x) + s)} < \epsilon$$

One possible starting point selection

1. Set x_1 to be the projection of x_0 onto the interior of the lower and upper bounds, set slack variables so bounds are satisfied.
2. Then compute:

$$\tilde{y}_1 \leftarrow \arg\min_y \|\nabla a(x_1)^T y - \nabla g(x_1)\|_2^2$$

3. Perform guarding strategy to ensure s and y are sufficiently large (i.e. $a(x) - s = \kappa e$).
4. Set $\mu_1 \leftarrow \frac{s^T y}{n}$.

Ratio of complementary to primal feasiblity

Recall we start with

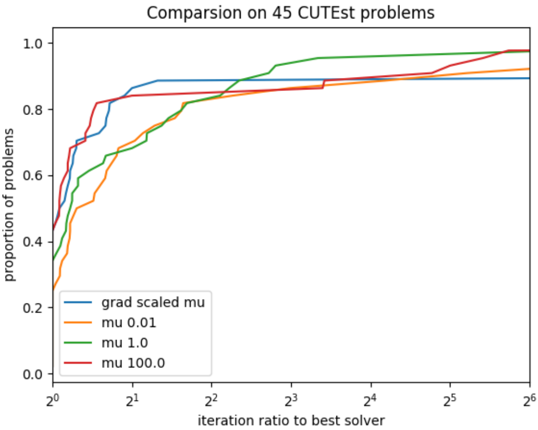
$$a(x) + s = \kappa_1 e$$

And some barrier parameter μ_1. We keep:

$$\frac{\mu_k}{\kappa_k} = \frac{\mu_1}{\kappa_1} \approx 1$$

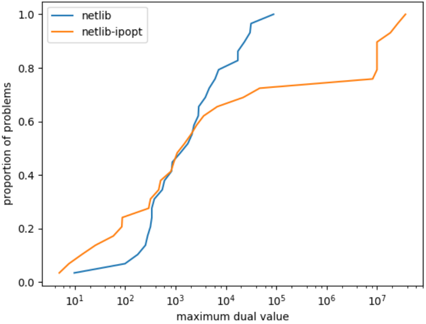
The choice of this ratio has a large effect on performance.

Starting point



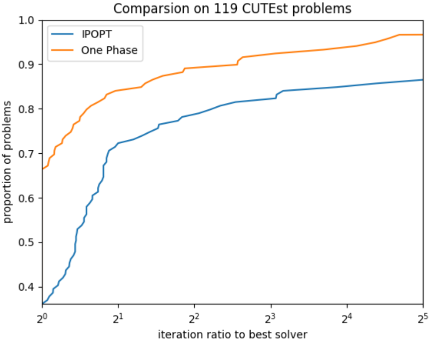
Duals sequence on subset of netlib

1. IPOPT fails on 11/29 problems (most of these are a result of IPOPT stalling far from the optimum)
2. We fail on 4/29 (all due to numerical difficulties near the optimum).



Empirical results with CUTEst

1. Selected problems with 100-1000 variables and 100-1000 constraints.
2. Removed problems with NaNs on domain.



Some infeasible problems

Problem name	IPOPT	One phase
MODEL	primal infeasible (29)	primal infeasible (37)
CRESC50	primal infeasible (491)	factorization failure (1245)
HIMMELBD	primal infeasible (20)	primal infeasible (35)
FLOSP2HL	primal infeasible (11)	primal infeasible (22)
FLOSP2HM	primal infeasible (16)	primal infeasible (36)
WOODSNE	primal infeasible (9)	primal infeasible (13)
ARTIF	primal infeasible (98)	optimal (15)
EQC	Restoration Failed! (35)	optimal (314)
HIMMELBJ	Restoration Failed! (28)	optimal (28)

Conclusions and future work

1. Promising initial results
2. Look at dual values for CUTEst
3. Adaptive μ choice?
4. Better δ choice

The End

References I

Andreas Wächter, Lorenz T. Biegler (2000)
Failure of global convergence for a class of interior point methods for nonlinear programming. *Mathematical Programming*.

Yinyu Ye, Michael J. Todd, and Shinji Mizuno (1994)
An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm *Mathematics of Operations Research*.

Erling D. Andersen and Yinyu Ye (1999)
On a homogeneous algorithm for the monotone complementarity problem. *Mathematical Programming*.

Irvin J Lustig (1990)
Feasibility issues in a primal-dual interior-point method for linear programming. *Mathematical Programming*.

References II

Mehrotra Sanjay (1992)
On the implementation of a primal-dual interior point method. *SIAM Journal on optimization*.

Michael J. Todd (2002)
Detecting infeasibility in infeasible-interior-point methods for optimization. *Foundations of Computational Mathematics*.

Frank E. Curtis (2012)
A penalty-interior-point algorithm for nonlinear constrained optimization. *Mathematical Programming Computation*.

Lifeng Chen and Donald Goldfarb (2006)
A penalty-interior-point algorithm for nonlinear constrained optimization. *Mathematical Programming*.