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

Oliver Hinder, Yinyu Ye

ohinder@stanford.edu

July 13, 2016

Outline

1. Introduction
   - Current non-convex IPMs
   - Local infeasibility certificates
   - Algorithm outline
2. Theory
   - Global convergence
   - Local convergence
   - How to return L1 minimizer of constraint violation
3. Practice
   - Implementation details
   - Preliminary empirical results with CUTEst

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 $\theta \in \mathbb{R}^m$

One phase IPMs for convex optimization

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

1. Infeasible start IPM (Mehortra, 1992)
2. Homogenous algorithm (Anderson and Ye, 1998, 1999)

State of the art for convex optimization

Why one phase non-convex IPM do not exist (yet)

1. Some infeasible start algorithms have been tested, but no theoretical guarantees (i.e. LOQO)
2. Ideally our algorithm would either find a local optimum or a local minimizer the violation of the constraints
3. Infeasible start IPMs may fail this criterion (Wachter and Biegler in 2000).

Solutions in literature to this problem

1. Two phases (IPOPT)
    1. Each phase has a different variables
    2. Awkward convergence assumptions
    3. Poor performance on infeasible problems
2. Compute two directions (KNITRO)
3. Penalty (or big-M) method (Curtis, 2012)

Solutions in literature to this problem

1. Two phases (IPOPT)
    1. Each phase has a different variables
    2. Awkward convergence assumptions
    3. Poor performance on infeasible problems
2. Compute two directions (KNITRO)
3. Penalty (or big-M) method (Curtis, 2012)

It doesn't need to be this complicated!

Local infeasibility certificates

Farkas certificate, there exists $y \geq 0$ such that x^* is a stationary point of:

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

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

L1 constraint violation, x^* is a stationary point of:

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

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

One phase barrier problem

Original problem:

$$\begin{aligned} \min f(x) \\ a(x) \leq 0 \end{aligned}$$

Barrier problem:

$$\begin{aligned} &\min f(x) - (1-\eta^k)\mu^k \sum_i \log s_i \\ &a(x) + s = (1-\eta^k)(a(x^k) + s^k) \\ &s \geq 0 \end{aligned}$$

Direction computation

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

$$
\begin{aligned}
H^k &= \nabla_x^2 L(x^k, y^k) + \delta^k I \\
A^k &= \nabla a(x^k)
\end{aligned}
$$

One phase non-convex algorithm

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

Factor newton system using IPOPT strategy to compute δ^k

if $\|\nabla f(x^k) + (y^k)^T \nabla a(x^k)\| < \mu^k$ **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

Iterate update

Backtrack on α and update iterates:

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

Only accept iterates that approximately satisfy complementary:

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

Merit function line search

Backtrack on α until armijo rule is satisfied for the merit function:

$$\psi^k(x) = f(x) - \mu^k \sum_i \log\left(a_i(x^k) + s_i^k - a_i(x)\right)$$

Termination criterion

Declare problem first-order ϵ-locally optimal if:

$$\left\|\nabla f(x) + y^T \nabla a(x)\right\| + \|a(x) + s\| + \mu < \epsilon_{opt}$$

Declare problem first-order ϵ-locally (farkas) infeasible if:

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

Global convergence

If δ^k and η^k are judiciously chosen then it possible to show the following result:

Theorem

For any $\epsilon > 0$, after a finite number of steps the algorithm has found either a first-order ϵ-local optimum or a first-order ϵ-local (farkas) infeasibility certificate.

Local convergence

Assume:

1. Converging towards a KKT solution that satisfies strict complementary.
2. The sufficient conditions for local optimality are satisfied at the KKT solution.

If η^k is judiciously chosen then it is possible to show:

Theorem

If (1-2) hold, the algorithm converges Q-quadratically to a local optimality certificate.

How to return L1 minimizer of constraint violation

We wish to solve:

$$\begin{aligned} \min f(x) \\ a(x) \leq 0 \end{aligned}$$

This problem is trivially equivalent to:

$$\begin{aligned} \min f(x) \\ a(x) \leq z \\ e^T z \leq 0 \\ z \geq 0 \end{aligned}$$

The local Farkas infeasibility certificate to this problem is an certificate of L1 minimization of the constraint violation.

Implementation details

1. Use fraction to boundary rule i.e. $y^{k+1} \geq 0.01 y^k$ and $s^{k+1} \geq 0.01 s^k$
2. Choose η^k on aggressive steps using predictor-corrector technique
3. Search on negative eigenvector to guarantee second-order necessary conditions are satisfied

Scaled KKT merit function

Scaled KKT residuals is an alternate merit function:

$$\frac{\|\nabla f(x) + y^T \nabla a(x)\|}{\|y\|}$$

Accept step that makes progress on either:

1. Classic merit function
2. Scaled KKT residuals

Using a filter to prevent cycling

Empirical results

Table: Results on 14 selected CUTEst problems

	One phase	IPOPT (2004)	KNITRO (2004)
Mean iterations	71	152	740
Standard error	5	25	85

Summary

1. Introduction
   - Current non-convex IPMs
   - Local infeasibility certificates
   - Algorithm outline
2. Theory
   - Global convergence
   - Local convergence
   - How to return L1 minimizer of constraint violation
3. Practice
   - Implementation details
   - Preliminary empirical results with CUTEst

The End

References

Itai Ashlagi, Mark Braverman, and Avinatan Hassidim (2014)
Stability in Large Matching Markets with Complementarities
Operations Research 62:4, 713-732.

Conjectures

Conjecture

Assume (x^*, s^*) is a limit point of the algorithm and there exists finite dual multipliers at (x^*, s^*) then for some subsequence of the dual multipliers $y^{\pi(k)} \to y^*$.

Conjecture

After a polynomial number of steps a simple variant of the algorithm has found either a first-order ϵ-local optimum or a first-order ϵ-local infeasibility certificate.