Adaptive Barrier Procedures for Nonlinear Interior Methods

Richard A. Waltz

Northwestern University

Collaborators —

Jorge Nocedal, Northwestern University
Andreas Waechter, IBM
Interior-Point Methods: Background

NLP

\[
\begin{align*}
\min & \quad f(x) \\
\text{s.t.} & \quad c(x) = 0 \\
& \quad x \geq 0
\end{align*}
\]

\[
\begin{align*}
\min & \quad f(x) - \mu \sum \ln(x_i) \\
\text{s.t.} & \quad c(x) = 0
\end{align*}
\]

\[
\nabla f(x) - \nabla c(x)y - z = 0 \\
c(x) = 0 \\
Xz = \mu e \\
x, z \geq 0
\]

Global convergence issues:
Ignoring \(f\) may lead to
Stationary points or maxima
Interior Point Methods: Background

How to drive $\mu$ to 0? \[ \min f(x) - \mu \sum \ln(x_i) \]
Theory vs. practice? \[ \text{s.t. } c(x) = 0 \]

1. Monotone decrease
   - Solve barrier problem for fixed $\mu$, then decrease
   - Global convergence guarantees, robust
   - May be too cautious or slow in practice

2. Adaptive $\mu = \sigma \frac{x^T z}{n}$
   - Choose $\mu$ every iteration; may increase or decrease
   - More flexible; often more efficient
   - Typically no global convergence guarantees
Adaptive barrier update rules

General Form:

$$\mu = \sigma \frac{x^T z}{n}$$

- $\sigma$ balances decrease in complementarity and centrality

Loqo: deviation of smallest complementarity pair from ave.

$$\sigma = 0.1 \min \left( 0.05 \frac{1 - \xi}{\xi}, 2 \right)^3 \quad \xi = \frac{\min_i x_i z_i}{x^T z / n}$$

Mehrotra probing

$$\sigma = \left( \frac{(x + \alpha_x \Delta x^{aff})^T (z + \alpha_z \Delta z^{aff})}{x^T z} \right)^3$$
Quality function

- Create a function which characterizes the “quality” of a particular choice of \( \mu \)
- Minimize the function to get optimal \( \mu \)

\[
\mu = \sigma \frac{x^T z}{n}
\]

Define

\[
x(\sigma) = x + \alpha_x(\sigma) \Delta x(\sigma)
\]
\[
y(\sigma) = y + \alpha_y(\sigma) \Delta y(\sigma)
\]
\[
z(\sigma) = z + \alpha_z(\sigma) \Delta z(\sigma)
\]

\[
q_N(\sigma) = \left\| \nabla f(x(\sigma)) - \nabla c(x(\sigma)) y(\sigma) - z(\sigma) \right\| + \left\| c(x(\sigma)) \right\| + \left\| X(\sigma) z(\sigma) \right\|
\]

\[
\nabla f(x) - \nabla c(x) y - z = 0
\]
\[
c(x) = 0
\]
\[
Xz = \mu e
\]
\[
x, z \geq 0
\]

Too expensive!
Quality function

• Use linear models for primal and dual feasibility
• Square terms (emphasis on largest term)

\[ q_L(\sigma) = (1 - \alpha_z(\sigma))^2 \left\| \nabla f(x) - \nabla c(x)y - z \right\|^2 + \\
(1 - \alpha_x(\sigma))^2 \left\| c(x) \right\|^2 + \left\| X(\sigma)z(\sigma) \right\|^2 \]

Choose \( \mu = \sigma \frac{x^Tz}{n} \) to (approximately) minimize \( q_L(\sigma) \)

Nonconvex!
Globalization Framework

Repeat until termination test satisfied

*Adaptive Mode*
- Choose $\mu$ to minimize $q_L(\sigma)$
- Compute primal dual step $(\Delta_x, \Delta_y, \Delta_x)$ for given $\mu$
- Compute stepsizes, $\alpha_x, \alpha_z$ (descent in merit function)
- Update iterate
- Compute KKT error

If KKT error has decreased in the last $k$ iterations
  - Accept step and loop

Else
  - Run *Monotone Mode* until KKT error has decreased
  - Goto the beginning of the loop

Endif

End Repeat
KNITRO Results: NETLIB LPs

![Graph showing the performance of different algorithms for NETLIB LPs.](image-url)
KNITRO Results: CUTEr NLPs

CUTEr - function evaluations

% of problems

not more than x-times worse than the best option

Monotone
Log-opt
Mehrotra-probing
Quality-globalized
Linear Programming and Corrector steps

Predictor (probing) step to determine $\mu$:
1. Affine-scaling ($\mu=0$) step, $\Delta_{aff}$
2. Find maximum steplengths to boundary

$$\sigma = \left( \frac{(x + \alpha_x \Delta_{aff})^T (z + \alpha_z \Delta_{aff})}{x^T z} \right)^3$$

$$\mu = \sigma \frac{x^T z}{n}$$

3. Add Corrector step to satisfy complementarity better

$$\nabla f(x) - \nabla c(x)y - z = 0$$
$$c(x) = 0$$
$$Xz = \mu e$$
$$x, z \geq 0$$
Mehrotra Predictor-Corrector

What you hope for!
Linear Programming and Corrector steps

- PCx, Czyzk, Wagner, Wright, Mehrotra
- Standard predictor/corrector
- 93 problems from Netlib
- Default solves 89 (+4)

Initial point \( x_0 = e \quad z_0 = e \)

Solves 28 (+3)
Mehrotra’s predictor-corrector algorithm: what can go wrong
Remedies: control increase in complementarity

![Iteration Count Graph](image)

- default MPC corrector
- no corrector
- conditional MPC corrector

% of problems vs. not more than $2^x$-times worse than best solver
Mehrotra Predictor-Corrector Extension to NLP:

- Corrector step for everything or just complementarity?
- Barrier value obtained by affine-scaling step not reliable if nonconvex
- Ensure descent direction for a merit function
- Dynamic barrier formula tuned for LP problems (maybe something better for NLP)
- Safeguards needed
KNITRO Results: CUTEr NLPs

CUTEr - function evaluations

% of problems

not more than x-times worse than the best option

May 10, 2005
Montreal, 2005
Future Work

1. Refine quality function
   - Centrality term
   - Use quadratic models
   - Balance feasibility progress versus complementarity progress

2. Improve search procedure for optimal $\sigma$

3. Extend quality function ideas to the computation of the corrector step

Thank You!