KNITRO: An Integrated Approach for Nonlinear Optimization

◆ Richard Waltz, Northwestern University

Outline

- PART I:

  Adaptive Barrier Updates for NLP

- PART II:

  Overview of KNITRO Software Package

Adaptive barrier updates

NLP

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s.t.} \quad & h_i(x) = 0, \; i \in E \\ & g_i(x) \geq 0, \; i \in I \\ & x \in \Re^n \end{aligned}$$

- Functions twice continuously differentiable

Adaptive barrier updates

Solve a sequence of barrier subproblems

$$\min_{x,s} \quad f(x) - \mu \sum_{i \in I} \ln s_i$$

$$\text{s.t.} \quad h_i(x) = 0, \qquad i \in E$$

$$g_i(x) - s_i = 0, i \in I$$

- Approach solution to NLP as $\mu \to 0$

Adaptive barrier updates

Optimality conditions for barrier:

$$F(x,s,y,z) \equiv \begin{pmatrix} \nabla f(x) - A_h(x)y - A_g(x)z \\ SZe \\ h(x) \\ g(x) - s \end{pmatrix} = \begin{pmatrix} 0 \\ \mu e \\ 0 \\ 0 \end{pmatrix}$$

$$s, z \geq 0$$

$$\bar{x} = (x, s, y, z)^{\mathrm{T}}$$

Adaptive barrier updates (LP)

Predictor (probing) step to determine μ:

1. Affine-scaling step $F(\bar{x})' d_{\mathrm{AS}} = -F(\bar{x})$

2. Find maximum steplengths to boundary

$$\mu_{\mathrm{AS}} = (s + \alpha_{\mathrm{AS}}^{s} d_{\mathrm{AS}}^{s})^{\mathrm{T}} (z + \alpha_{\mathrm{AS}}^{z} d_{\mathrm{AS}}^{z}) / n_s$$

$$\mu^{+} = \left(\frac{\mu_{\mathrm{AS}}}{\mu} \right)^{3} \mu$$

Adaptive barrier updates (LP)

Predictor-corrector:

- μ determined adaptively
- Efficient to compute corrector step
  - One additional backsolve
- Very effective in practice
- not globally convergent

Adaptive barrier updates (NLP)

Pred-Corr Extension to NLP:

- Corrector step requires one additional evaluation of functions and gradients
- Barrier value obtained by affine-scaling step not reliable if nonconvex
- Ensure descent direction for a merit function
- Dynamic barrier formula for LP not as effective for NLP

Adaptive barrier updates (NLP)

Overview of Barrier Strategies:

1. Fixed decrease with barrier stop test (e.g. KNITRO)
2. Centrality-based strategies
3. Probing strategies (e.g. Mehrotra PC)

Adaptive barrier updates (NLP)

KNITRO

- Conservative rule
  - Initially $\mu=0.1$
  - Decrease μ linearly
  - Fastlinear decrease near solution
- Globally convergent
- Robust but trade-off some efficiency
- Initial point option

Adaptive barrier updates (NLP)

- Develop a more flexible adaptive rule
  - Allow increases in barrier parameter!

$$\mu_{k+1} = \theta \left[\frac{s^T z}{n_s} \right]$$

θ : function of:

Spread of complementarity pairs

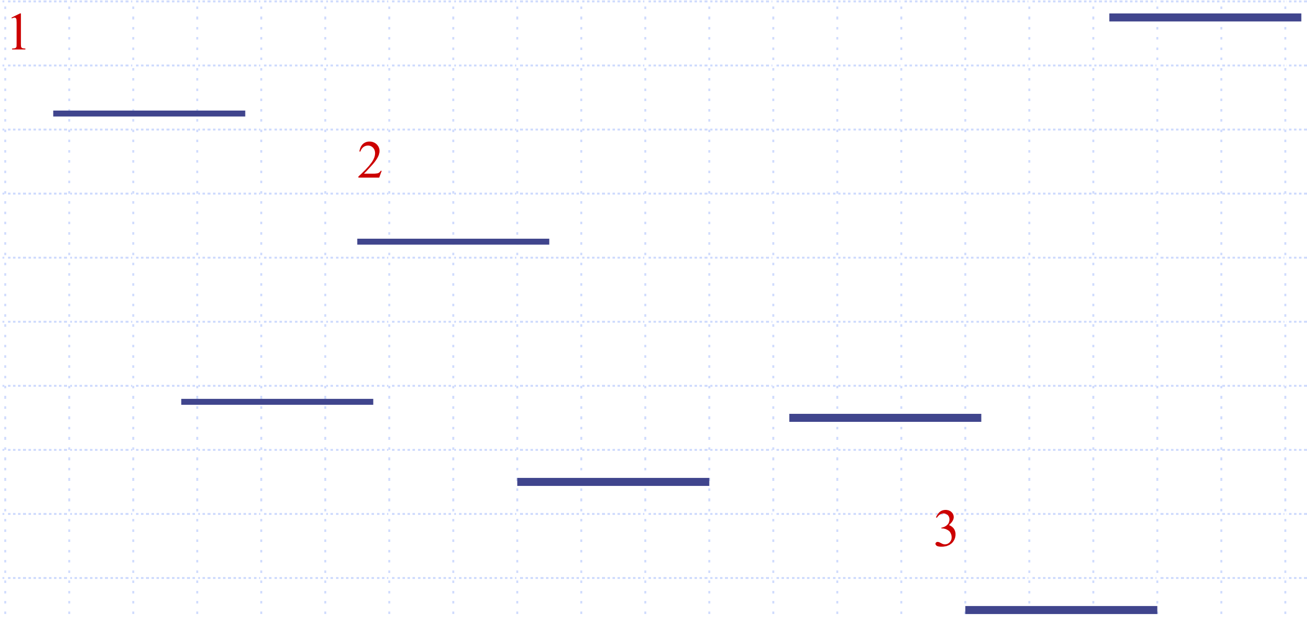
Recent steplengths

Ease of meeting a barrier stop test

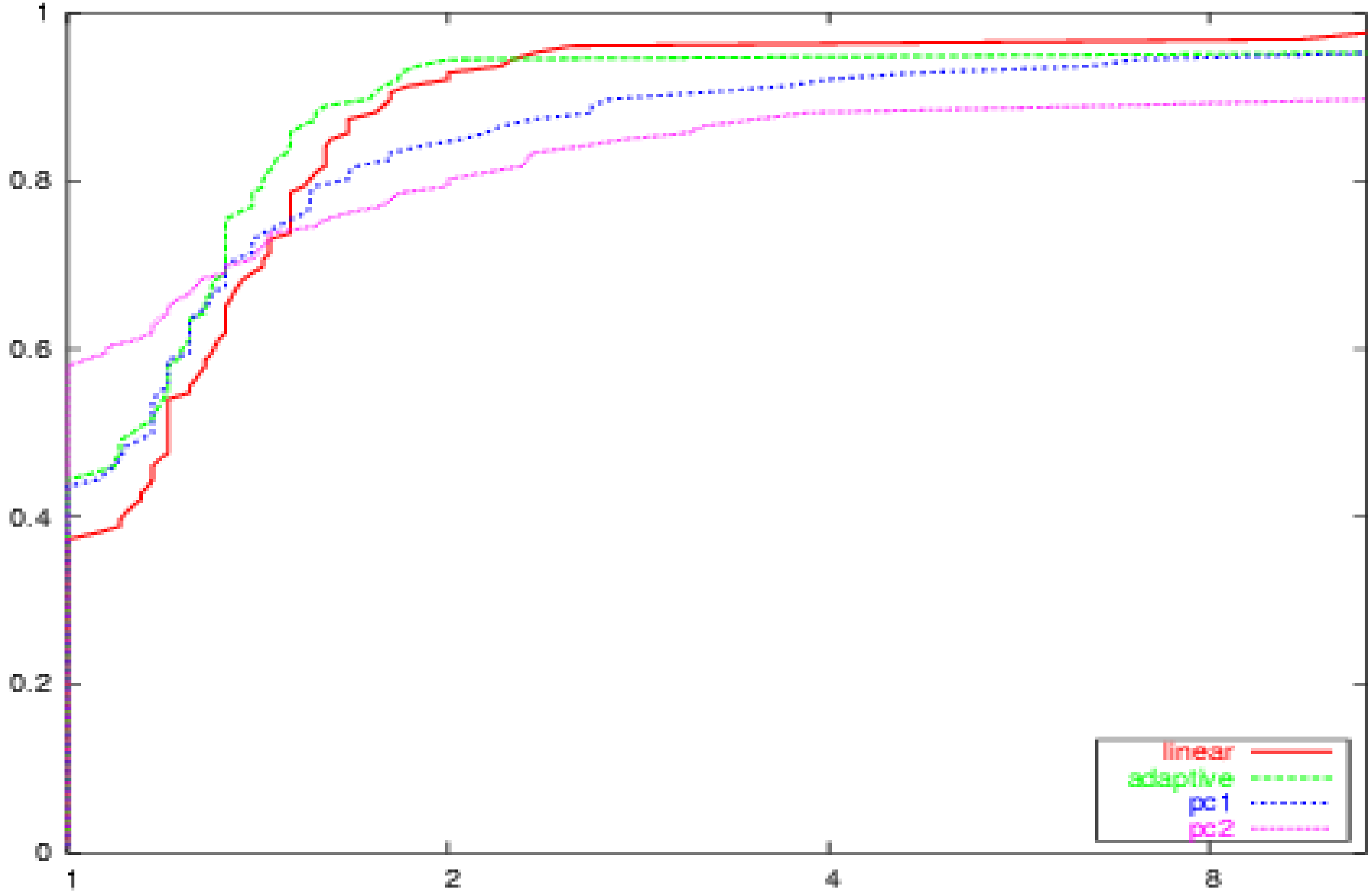
Probing step (e.g. predictor step)

Globally Convergent Framework

1. Official μ for global conv (satisfies barrier stop test)
2. Trial μ for flexibility



Adaptive barrier updates (NLP)



PART II: KNITRO Software

PART II: KNITRO Software

MOTIVATION

- Diversity of nonlinear optimization probs. requires diversity of algorithms/features

GOAL

- Comprehensive nonlinear optimization package offering a variety of algorithms/options
  - Interior-point and Active-set
  - Iterative and Direct approaches
  - 1st and 2nd derivative options
  - Adaptive techniques

Three Algorithms

1. KNITRO-CG:
   - Barrier iterative approach
2. KNITRO-Direct
   - Barrier direct approach
3. KNITRO-Active (not yet available)
   - SLPEQP active-set approach
   - Coming Summer 2004

KNITRO-CG Overview

- Iterative trust-region interior method
- Step: $d = v + t$
  - Feasibility/normal component, v
    - Min. norm step to linearized constraints
    - Computed uses Powell dogleg method
  - Minimization/tangential component, t
    - Minimize quad. model of barrier function
    - Use iterative CG approach (Steihaug)

KNITRO-CG Overview

- Strengths:
  - Use exact 2nd derivatives always
    - Negative curvature handled by trust-region
  - Do not need to form/factor Hessian!
  - Only need Hessian-vector products
- Weaknesses
  - CG inefficient if Hessian ill-conditioned

KNITRO-Direct Overview

- Direct line-search interior method (safeguarded by trust-region steps)
- Compute step by direct solve of KKT system

$$\begin{bmatrix} \mathbf{H}(x,z) & \mathbf{A}^{\mathrm{T}}(x) \\ \mathbf{A}(x) & 0 \end{bmatrix} \begin{bmatrix} d_x \\ d_z \end{bmatrix} = -\begin{bmatrix} \nabla \mathrm{L}(x,z) \\ c(x) \end{bmatrix}$$

$$x^+ = x + \alpha_x d_x, \; z^+ = z + \alpha_z d_z$$

KNITRO-Direct Overview

- Always try direct step first
- Fall back on KNITRO-CG if:
  1. Negative curvature detected (~14%)
  2. Backtrack LS results in small α (~7%)
     - robustness when singularities in Hessian or Jacobian

- Globally convergent

KNITRO-Direct Overview

- Strengths:
  - Better able to handle ill-conditioning
  - Often more efficient on easier problems
- Weaknesses
  - Needs to form/factor Hessian
  - Inefficiencies if falling back on CG often

CG-Direct Comparison

- CVXQP2
  - n=10,000,
  - m=2,500,
  - nnzH = 40,000
  - 99.6% time spent on factor in KNITRO-Dir

Code	iters	time	time/it
CG	11	401	36.5
Direct	14	2638	188.4

- BQPGAUSS
  - n = 2003
  - bound-constrained
  - H ill-conditioned but not expensive

Code	Iters	time	time/it
CG	27	1310	48.5
Direct	19	3	0.16

KNITRO-Active Overview

Trust-region active-set SLPEQP algorithm

0. **Given**: x
1. **Solve LP** to get working set $\mathcal{W}$.
2. **Compute a step, *d***, by solving an equality constrained QP using constraints in $\mathcal{W}$.
3. **Set**: $x_T = x+d$.

KNITRO-Active Overview

- Strengths:
  - Warm starts
  - Better active-set/sensitivity info
  - Only factors systems with active constraints
  - Crossover techniques
- Weaknesses:
  - Not yet as efficient as barrier approach on really large problems
  - Requires good simplex solver

Other problem classes

Prob class	CG	Direct	Active
Uncon	Newton-CG	Newton	Newton-CG
Non. Eq.	Powell-dogleg	Newton	?
Least Sq.	Levenb-Marq.	Gauss-Newt	?
Bound Con	Barrier-CG	Barrier-Direct	Gradient projection
Equal Con	SQP-iterative	SQP-direct	SQP-iterative
LP	Barrier-CG?	Barrier-Direct	simplex
QP	Barrier-CG	Barrier-Direct	SLPEQP (SQP like)

Second Derivative options

- Exact 2nd derivatives
- Quasi-Newton
  - SR1 (dense)
  - BFGS (dense)
  - Limited memory BFGS (large-scale)
- Hessian-vector products (KNITRO-CG/Active)
  - exact Hessian-vector products
  - via finite-differencing of gradients

Feasible Option

- By default constraints may be violated during optimization process
- Feasible option: enforces feasibility with respect to inequalities given initial point satisfying inequalities
- Honor bounds: special case of feasible option
- Constraints may be undefined outside feasible region
- Allows for early termination with feasible solution

Future Work

- Active-set code
- Mixed Integer NLP
- Crossover, adaptive rules
- Preprocessing NLP's

References

- www.ece.northwestern.edu/~rwaltz
- www.ziena.com/knitro.html (student edition)