Active Set Identification in SLQP Algorithms for Nonlinear Optimization

Richard A. Waltz
Northwestern University

Collaborator —
Richard Byrd, Colorado University
Part I
Active-Set ID in SLQP Algorithms

SLQP:
• estimate active-set by solving an LP
• then solve EQP to generate step
  • Fletcher, Sainz de la Maza
  • Byrd, Gould, Nocedal, W.
• can we do better?
SLQP Overview

**LP**

\[
\begin{align*}
\text{min} & \quad \nabla f^T d \\
\text{s.t.} & \quad h + \nabla h^T d = 0 \\
& \quad g + \nabla g^T d \geq 0 \\
& \quad \|d\|_\infty \leq \Delta^{LP}
\end{align*}
\]

\[\text{Working set } \mathcal{W}\]

- Equality constrained quadratic program

**EQP**

\[
\begin{align*}
\text{min} & \quad q(d) \\
\text{s.t.} & \quad \nabla h^T d = 0 \\
& \quad \nabla g_i^T d = 0 \quad i \in \mathcal{W} \\
& \quad \|d\| \leq \Delta
\end{align*}
\]

Fletcher et al; KNITRO-Active

May 16, 2005 SIOPT, 2005 3
SLQP Cauchy point and step computation

\[ \text{Cauchy} \quad \min_{\alpha} \quad m(\alpha d^{LP}), \quad x_C = x + \alpha^{LP} d^{LP} \]

Active-set based on \( x_{LP} \), not \( x_C \)!
Short $\alpha^{\text{EQP}}$ caused by poor active-set estimate
Example:
- Poor active-set ID at LP point
- Model increasing from Cauchy to EQP point

\[
\alpha_{\text{EQP}} = 0
\]
SLQP vs Gradient Projection

SLQP approach
• Quadratic info not considered in active-set ID
• Active-set not necessarily at Cauchy point (fewer activities)

Gradient Projection
• Active-set defined at Cauchy point
  (where quadratic model is minimized)

Can we do something more like gradient projection?
Parametric LP approach

\[
\begin{align*}
\text{min} & \quad \nabla f^T d \\
\text{s.t} & \quad g + \nabla g^T d \geq 0 \\
& \quad \|d\|_\infty \leq \Delta^{LP}
\end{align*}
\]

Vary \( \Delta^{LP} \)
Parameterized LP approach

Algorithm

Given: initial $\Delta^{LP}$
Repeat until sufficient model reduction
  • Decrease $\Delta^{LP}$ (allow increases as well)
  • Solve LP with trust region $\Delta^{LP}$
End (Repeat)

Take final LP solve as Cauchy point with corresponding working-set, $W$

Solve an EQP using constraints in $W$

Drawbacks: cost of parameterized (or multiple) LP
Results: Bound Constrained

\[ \log_2 \frac{\text{fevals (parametric LP)}}{\text{fevals (SLQP)}} \]

Standard SLQP Approach better vs. Parametric LP Approach better

Area: 12.4

Standard SLQP Approach better

Area: 28.7
Results: Linearly Constrained

Standard SLQP Approach better
Area: 43.2

Parametric LP Approach better
Area: 85.5
Results: Nonlinear Constraints

Standard SLQP Approach better
Area: 43.9

Parametric LP Approach better
Area: 71.6

...but less robust (-10)
Part II
Crossover for NLP

How to “clean-up” interior solutions for NLP
• clear active-set identification
• accurate multipliers/sensitivities
• accurate primal solutions (i.e. on constraints)

**Definition:** Procedure to determine a solution with exact complementarity and a linearly independent active-set sufficient to “cover” $\nabla f^*$.
How to effectively integrate these different algorithms?
Crossover for NLP

\[ \text{min} \quad \nabla f^T d + \nu \sum_{i \in I} t_i \]
\[ \text{s.t.} \quad g + \nabla g^T d \geq -t, \quad t > 0 \]
\[ \| d \|_{\infty} \leq \Delta_{\text{LP}} \]

Use l1 trust-region LP subproblem to identify activities

- how to choose \( \nu \)?
- how to choose \( \Delta_{\text{LP}} \)?
- how to warm start the LP solution?
Choosing $\Delta^{LP}$

\[
\begin{align*}
\min_{\Delta \in L^P} & \quad \nabla f^T d + \nu \sum_{i \in I} t_i \\
\text{s.t.} & \quad g + \nabla g^T d \geq -t, \quad t > 0 \\
& \quad \| d \|_\infty \leq \Delta^{LP}
\end{align*}
\]

Thereom: C. Oberlin and S. Wright 2005

If MFCQ and close enough to the solution there exists

\[
\Delta^{LP} \in [\| x - x^* \|^{\sigma}, \overline{\Delta}]
\]

Giving $A \subset A^*$ where $A$ “covers” $\nabla f^*$
Choosing $\Delta^{LP}$ For Crossover

$\Delta^{LP} \in [\|x - x^*\|, \Delta]$

Want $\|d\|_\infty < \Delta^{LP} \Rightarrow g_i + \nabla g_i^T d > 0, \ i \notin A$

True if $\Delta^{LP} = \min_{i \notin A} \frac{g_i}{\| \nabla g_i \|_1}$

Concerns:
- May choose $\Delta^{LP}$ too small (also look at steplength)
- Requires estimate of activity
Example output: HS35

Interior-point solution (8 interior-point iterations):

Constraint Vector
\[ c[0] = 9.02258487923e-08, \]
Lagrange Multipliers
\[ \lambda[0] = -2.22222267954e-01 \]

Solution Vector
\[ x[0] = 1.33333335913e+00, \]
\[ \lambda[1] = -1.50037299637e-08 \]
\[ x[1] = 7.77777759871e-01, \]
\[ \lambda[2] = -2.57073125817e-08 \]
\[ x[2] = 4.44444395384e-01, \]
\[ \lambda[3] = -4.49469872002e-08 \]

Crossover solution (1 crossover iteration):

Constraint Vector
\[ c[0] = 1.11022302463e-16, \]
Lagrange Multipliers
\[ \lambda[0] = -2.22222246187e-01 \]

Solution Vector
\[ x[0] = 1.33333333333e+00, \]
\[ \lambda[1] = 0.00000000000e+00 \]
\[ x[1] = 7.77777777778e-01, \]
\[ \lambda[2] = 0.00000000000e+00 \]
\[ x[2] = 4.44444444444e-01, \]
\[ \lambda[3] = 0.00000000000e+00 \]
Crossover Efficiency (616 CUTEr probs)

**Inequality Constrained Problems in the CUTEr collection (616)**

- **% of Problems**
- **Number of Crossover Iterations**

May 16, 2005

SIOPT, 2005
Quality of Solution Value (616 CUTEr probs)

Interior-point solutions:
Ave. primal feasibility error: 0.0000e-00
Ave. dual feasibility error: 1.7250e-09
Ave. complementarity error: 4.2750e-07

After crossover:
Ave. primal feasibility error: 1.6100e-15
Ave. dual feasibility error: 6.1100e-10
Ave. complementarity error: 2.5050e-17
... plus active-set ID
Thank You