Feasible Interior-Point Methods Using Slacks for Nonlinear Optimization

ISMP 2000, Atlanta, 7-11 Aug 2000

Richard Waltz
Dept. of Electrical and Computer Engineering
Northwestern University
Evanston, IL USA

in collaboration with

Jorge Nocedal
Dept. of Electrical and Computer Engineering
Northwestern University
Evanston, IL USA

Richard Byrd
Department of Computer Science
University of Colorado
Boulder, CO USA

Background

Interior Methods for solving:

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \geq 0 \end{aligned}$$

Definition:

Feasible meth. satisfies inequalities, $g(x)$, but not necessarily equalities, at each iteration.

Feasible Methods

- Gay, Overton, Wright (1997)
- Lawrence, Tits (1998)
- Armand, Gilbert, Jan-Jégou (1998)
- Forsgren, Gill (1998)
- ...

Feasible $\Rightarrow$ Infeasible

$$g(x) \geq 0 \quad \Rightarrow \quad g(x) - s = 0, \; s \geq 0$$

Infeasible Methods

Infeasible $\Rightarrow$ Feasible

- General framework (LS or TR)
- Incorporates slacks (uniform structure)

Infeasible framework

$$\begin{aligned} \min_{x,s} \quad & f(x) - \mu \sum_{i \in \mathcal{I}} \ln(s_i) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) - s = 0 \\ & s > 0 \end{aligned}$$

Infeasible Methods

- Vanderbei, Shanno [LOQO] (2000)
- Toint (1998)
- Yamashita, Yabe, Tanabe (1997)
- Byrd, Hribar, Nocedal [NITRO] (1999)
- ...

Feasible Method Using Slacks

Assume merit function, ϕ, has the form

$$\phi(x,s) \quad = \quad f(x) - \mu \sum_{i \in \mathcal{I}} \ln(s_i) + \chi(h(x), g(x,s))$$
$$\phi(x,s) \quad \to \quad \infty \quad \text{as} \quad s \to 0$$

Feasible Algorithm Step Computation

1. Assume current point (x_k, s_k) is feasible.
2. Set $s_k = g(x_k)$
3. After taking a trial step, $x_\mathsf{T} = x_k + d_x$ redefine the slacks as

$$\boxed{s_\mathsf{T} \leftarrow g(x_\mathsf{T})}$$

4. While ϕ does not decrease sufficiently

   Compute a shorter step d_x

   Set $x_\mathsf{T} = x_k + d_x$; $\boxed{s_\mathsf{T} = g(x_\mathsf{T})}$

   End While

Feasible Method Using Slacks

Line Search methods:

- everything fine

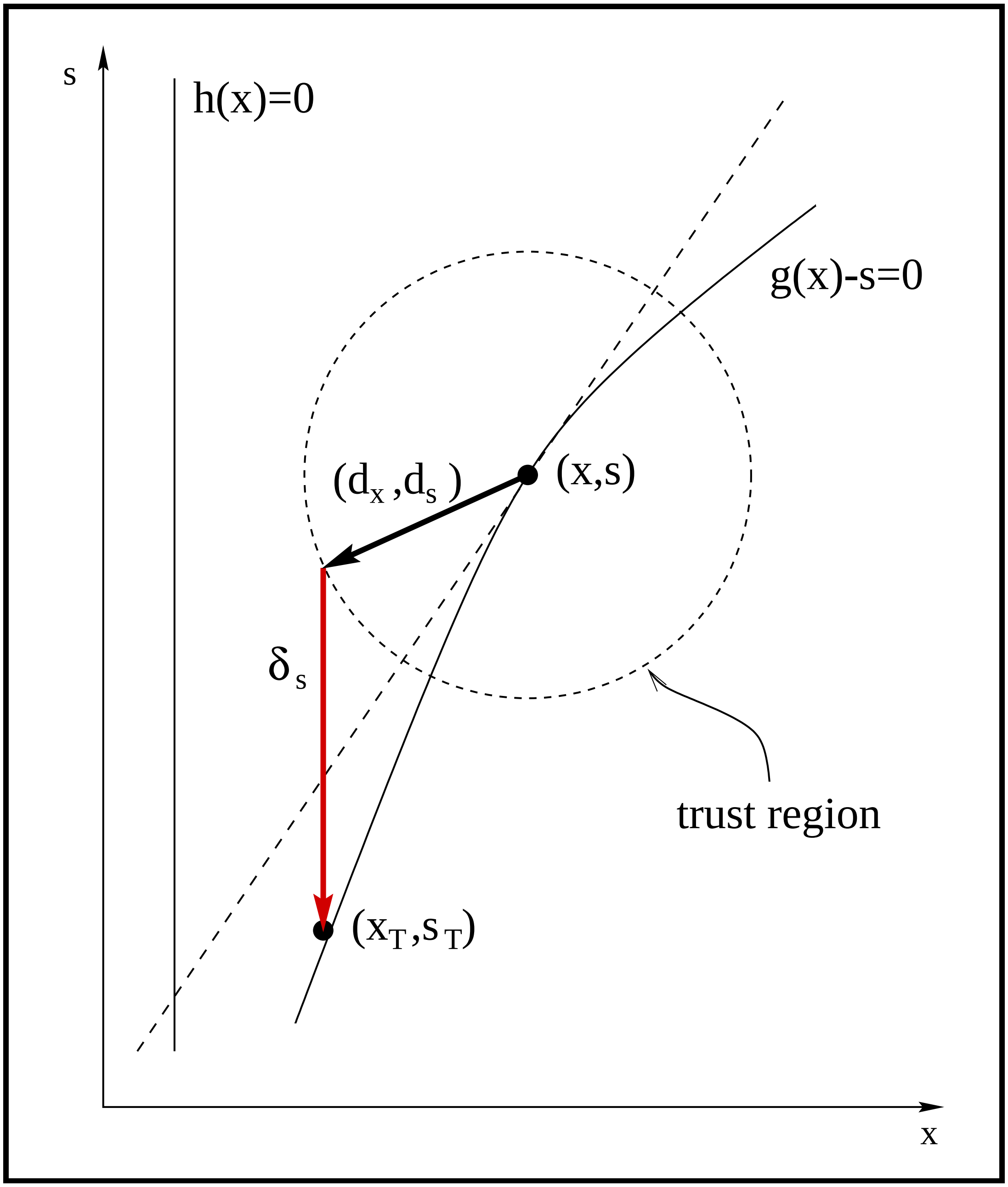
Trust Region methods:

- everything usually fine
- potential problem if Cauchy component active

Trust Region Reset Difficulty

Problem: Slack reset may increase merit function!

$$s_\top = g(x_\top) = s + d_s + \delta_s$$



No Equalities

- Assume there are **no** equalities
- From the slack reset, $g(x) - s = 0$
- The step (d_x, d_s) will maintain feasibility with respect to the linearization of $g(x) - s = 0$

$$\|g(x + d_x) - (s + d_s)\| = O(\|d_x\|^2).$$

- We know from the slack reset that

$$s_T = g(x + d_x) = s + d_s + \delta_s$$

Therefore

$$\boxed{\|\delta_s\| = O(\|d_x\|^2)}$$

- Merit function, ϕ, will decrease for small d_x

With Equalities

1. Line Search

   - Step direction satisfies linearized constraints

   - No problem: $\|\delta_s\| = O(\|d_x\|^2)$

2. Trust Region

   - Trust region constraint and satisfaction of linearized constraints may be incompatible

$$\|g(x + d_x) - (s + d_s)\| = O(\|d_x\|)$$

$$\boxed{\|\delta_s\| = O(\|d_x\|)}$$

   - **No guarantee merit function, $\phi(x, s)$, will decrease for small d_x!**

Solution

Ensure step (d_x, d_s) satisfies 2 conditions:

1. Satisfy linearization of $g(x) - s$ always

$$A_g(x)^T d_x - d_s = 0$$

$$A_g(x) = [\nabla g_1(x), \nabla g_2(x), ..., \nabla g_m(x)]$$

2. Make sufficient progress towards feasibility of violated equality constraints.

- Solution is algorithm dependent

NITRO solution

$$d = n + t$$

- $n = (n_x, n_s) =$ "Normal (feasibility) step"
- Normal step subproblem:

$$\begin{array}{ll} \min & \|A_{\mathsf{h}}^T n_x + h\|_2^2 + \|A_{\mathsf{g}}^T n_x + g - n_s - s\|_2^2 \\ \text{s.t.} & \|(n_x, n_s)\|_2 \le \Delta \\ & s + n_s > 0 \end{array}$$

- Newton step satisfies the linear. constraints
- Cauchy direction may violate linearized constraints

$$\begin{aligned} n_x^C &= -A_{\mathsf{h}} h(x) + A_{\mathsf{g}} \overbrace{(g(x) - s)}^{0} \\ &= -A_{\mathsf{h}} h(x) \end{aligned}$$

- Need a new Cauchy-type step (n_x^C, n_s^C)
- Solve via Dogleg method.

Modified Cauchy Step

Conditions:

1. $(n_x^C, n_s^C) \in \text{range}(A^T)$

2. (n_x^C, n_s^C) satisfy linearization of $g(x) - s = 0$

$$A_{\mathsf{g}}^T n_x^C - n_s^C = 0$$

3. n_x^C provide a reduction in the linearized equality constraints.

   Normal Cauchy step:

$$n_x^C = -A_{\mathsf{h}} h(x)$$

   Modified Cauchy step

$$\boxed{A_{\mathsf{h}}^T n_x^C = -A_{\mathsf{h}}^T A_{\mathsf{h}} h}$$

$$\begin{aligned} \|h + \alpha A_{\mathsf{h}}^T n_x^C\|^2 &= \|h + \alpha A_{\mathsf{h}}^T A_{\mathsf{h}} h\|^2 \\ &= \|h\|^2 - 2\alpha \|A_{\mathsf{h}} h\|^2 + O(\alpha^2) \end{aligned}$$

Modified Cauchy Step

- *Modified Cauchy step* is the solution of

$$\begin{bmatrix} I & 0 & A_{\mathsf{h}} & A_{\mathsf{g}} \\ 0 & I & 0 & -I \\ A_{\mathsf{h}}^T & 0 & 0 & 0 \\ A_{\mathsf{g}}^T & -I & 0 & 0 \end{bmatrix} \begin{bmatrix} n_x^C \\ n_s^C \\ p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -A_{\mathsf{h}}^T A_{\mathsf{h}} h \\ 0 \end{bmatrix}$$

- Cauchy step - no matrix inversion

- Coefficient matrix already exists (tangential step, Lagrange multipliers)

- Extra cost = one backsolve

Equivalence of Slack-Based and Classical Feasible Methods

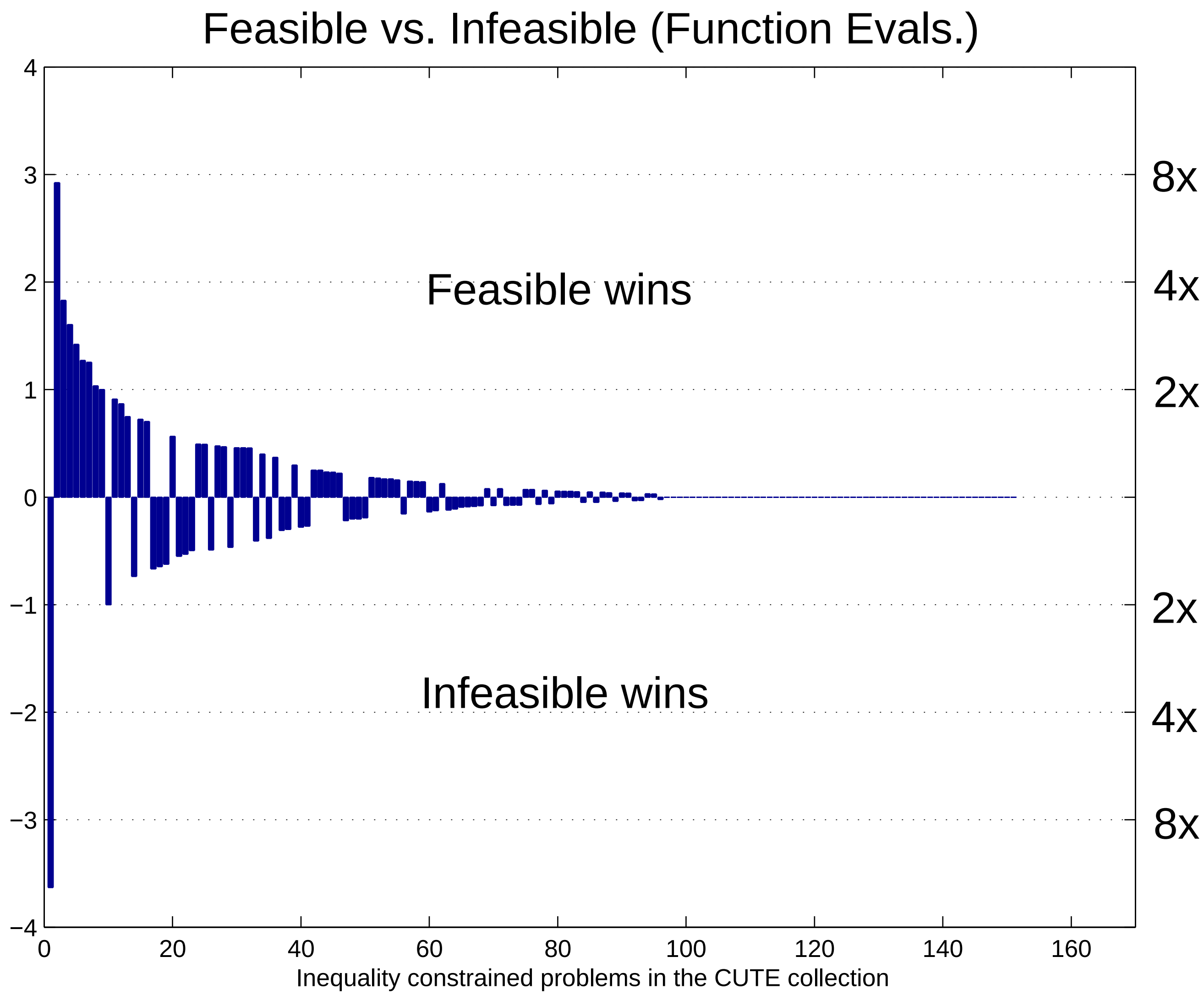
Newton on primal-dual equations of

$$\begin{aligned} \min_{x} \quad & \psi_{CB}(x;\mu) \equiv f(x) - \mu \sum_{i \in \mathcal{I}} \ln(g_i(x)) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) > 0 \end{aligned}$$

equivalent to Newton on primal-dual equations of

$$\begin{aligned} \min_{x,s} \quad & \psi_{SB}(x,s;\mu) \equiv f(x) - \mu \sum_{i \in \mathcal{I}} \ln(s_i) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) - s = 0 \\ & s > 0 \end{aligned}$$

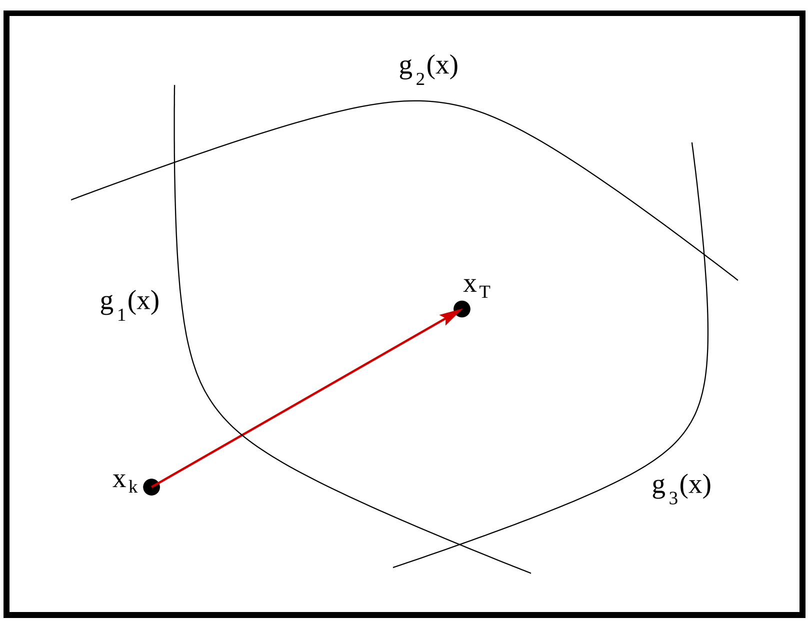
Numerical Results



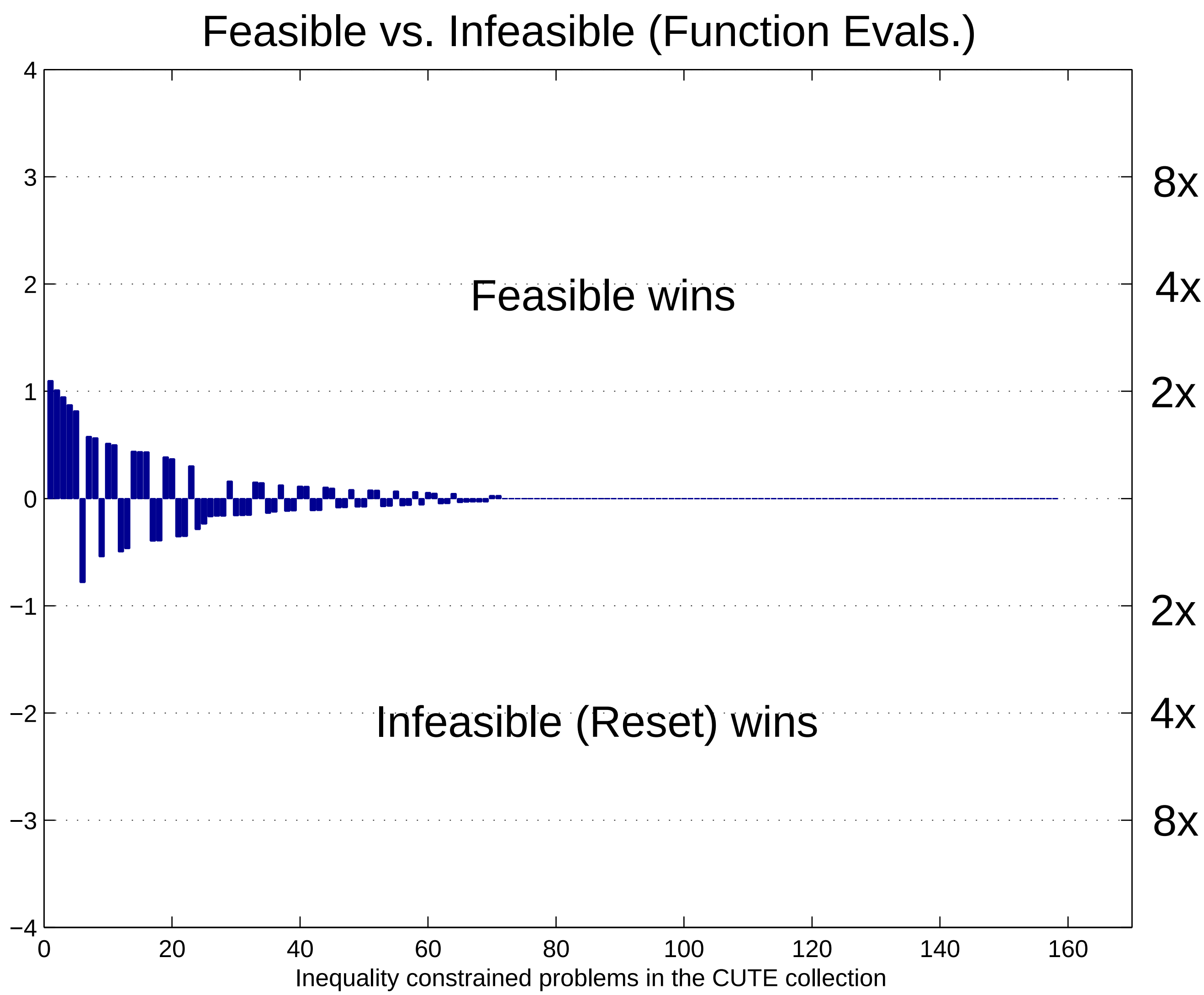
Infeasible Slack Reset Scheme

$$s_\top \leftarrow \max(g(x_\top), s + d_s)$$

- if $g(x_\top) > s + d_s$: feasible reset
  - go ahead and satisfy $g(x_\top) - s_\top = 0$
- if $g(x_\top) \leq s + d_s$: no reset
  (may take infeasible steps)
- only reset slacks if it will further decrease the merit function (can do no harm)



Numerical Results



Numerical Results

Robustness Comparison

Method	# of Problems Solved
Infeasible (NoReset)	170
Infeasible (Reset)	172
Feasible	174

- Test Set: 216 CUTE problems
- Not much difference

Summary/Conclusions

1. Slack-based feasible methods are derivable from infeasible methods.

   - LS: trivial modification

   - TR: requires *modified Cauchy step*

2. Slack-based feasible alg $\equiv$ classical method

3. You can improve the efficiency of an infeasible method by manipulating the slack variables.

4. Feasible vs. Infeasible?
   - Not a big difference overall
   - Feasible better on group of problems