A Matrix-free Algorithm for Optimization

Frank E. Curtis
New York University

involving joint work with
Richard H. Byrd, Jorge Nocedal, and Andreas Wächter

SIAM Optimization, 2008

Outline

Outline

Equality constrained optimization

- We consider *very large* problems of the form

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } c(x) = 0$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $c : \mathbb{R}^n \to \mathbb{R}^t$ are smooth functions

Equality constrained optimization

- We consider *very large* problems of the form

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } c(x) = 0$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $c : \mathbb{R}^n \to \mathbb{R}^t$ are smooth functions

- Defining the Lagrangian

$$\mathcal{L}(x, \lambda) \triangleq f(x) + \lambda^T c(x)$$

we are interested in finding a first-order optimal point; i.e., one satisfying

$$\nabla \mathcal{L} = \begin{bmatrix} g(x) + A(x)^T \lambda \\ c(x) \end{bmatrix} = 0$$

where $g(x)$ is the gradient of $f(x)$ and $A(x)$ is the Jacobian of $c(x)$

Method of choice: Newton/SQP

A Newton iteration from the point (x_k, λ_k) has the form

$$\begin{bmatrix} W(x_k,\lambda_k) & A(x_k)^T \\ A(x_k) & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = - \begin{bmatrix} g(x_k) + A(x_k)^T \lambda_k \\ c(x_k) \end{bmatrix}$$

where $W(x_k, \lambda_k) \approx \nabla^2_{xx}\mathcal{L}(x_k, \lambda_k)$, which is equivalent to solving the sequential quadratic programming (SQP) subproblem

$$\begin{aligned} \min_{d \in \mathbb{R}^n} \ & f(x_k) + g(x_k)^T d + \tfrac{1}{2} d^T W(x_k, \lambda_k) d \\ \text{s.t. } & c(x_k) + A(x_k) d = 0 \end{aligned}$$

- Multipliers λ_k provide signature of optimality
- Step computation is a single linear system
- Primal-dual equations are scale invariant

Globalization mechanism: Exact penalty function

- We judge progress toward a solution with the exact penalty function

$$\phi(x;\pi) \triangleq f(x) + \pi\|c(x)\|$$

where $\pi > 0$ is a penalty parameter increased dynamically

- The directional derivative of ϕ at x_k along d_k is

$$D\phi(d_k;\pi) = g_k^T d_k - \pi\|c_k\|$$

- Since d_k from SQP is known to be a good step, we can always set

$$\pi_k \geq \frac{g_k^T d_k + \frac{1}{2} d_k^T W_k d_k}{(1-\tau)\|c_k\|}$$

for some $\tau \in (0,1)$ to ensure $D\phi(d_k;\pi_k) \ll 0$

Line search SQP

for $k = 0, 1, 2, \ldots$

- Evaluate f_k, g_k, c_k, A_k, and W_k
- Solve the *primal-dual* equations

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = - \begin{bmatrix} g_k + A_k^T \lambda_k \\ c_k \end{bmatrix}$$

$$\min_{d \in \mathbb{R}^n} f_k + g_k^T d + \tfrac{1}{2} d^T W_k d \quad \text{s.t. } c_k + A_k d = 0$$

- Set the penalty parameter

$$\pi_k \geq \frac{g_k^T d_k + \frac{1}{2} d_k^T W_k d_k}{(1-\tau)\|c_k\|}$$

- Perform a line search to find $\alpha_k \in (0, 1]$ satisfying the Armijo condition

$$\phi(x_k + \alpha_k d_k; \pi_k) \leq \phi(x_k; \pi_k) + \eta \alpha_k D\phi(d_k; \pi_k)$$

- Update iterate $(x_k, \lambda_k) \leftarrow (x_k, \lambda_k) + \alpha_k (d_k, \delta_k)$

Outline

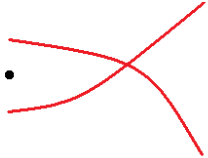
Handling ill-conditioned and rank-deficient Jacobians

- If at any point the Jacobian A of c is ill-conditioned or rank deficient, the step computation

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = - \begin{bmatrix} g_k + A_k^T \lambda_k \\ c_k \end{bmatrix}$$

$$\begin{aligned} \min_{d \in \mathbb{R}^n} \ & f_k + g_k^T d + \tfrac{1}{2} d^T W_k d \\ \text{s.t. } & c_k + A_k d = 0 \end{aligned}$$

may not be well-defined or may lead to $\|d_k\| \gg 0$ and $\alpha_k \approx 0$



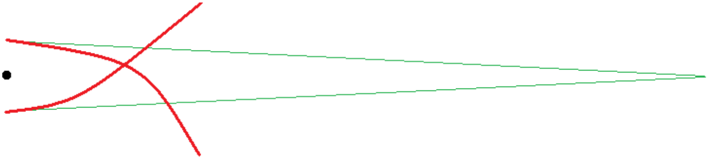
Handling ill-conditioned and rank-deficient Jacobians

- If at any point the Jacobian A of c is ill-conditioned or rank deficient, the step computation

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = -\begin{bmatrix} g_k + A_k^T \lambda_k \\ c_k \end{bmatrix}$$

$$\min_{d \in \mathbb{R}^n} \ f_k + g_k^T d + \tfrac{1}{2} d^T W_k d \quad \text{s.t. } c_k + A_k d = 0$$

may not be well-defined or may lead to $\|d_k\| \gg 0$ and $\alpha_k \approx 0$



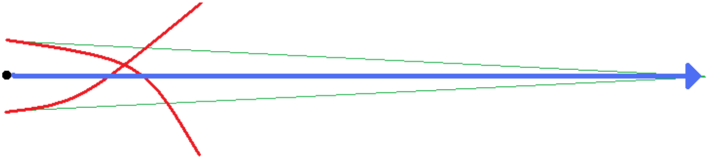
Handling ill-conditioned and rank-deficient Jacobians

- If at any point the Jacobian A of c is ill-conditioned or rank deficient, the step computation

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = - \begin{bmatrix} g_k + A_k^T \lambda_k \\ c_k \end{bmatrix}$$

$$\min_{d \in \mathbb{R}^n} \; f_k + g_k^T d + \tfrac{1}{2} d^T W_k d \quad \text{s.t.} \; c_k + A_k d = 0$$

may not be well-defined or may lead to $\|d_k\| \gg 0$ and $\alpha_k \approx 0$



Handling ill-conditioned and rank-deficient Jacobians

- If at any point the Jacobian A of c is ill-conditioned or rank deficient, the step computation

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = - \begin{bmatrix} g_k + A_k^T \lambda_k \\ c_k \end{bmatrix}$$

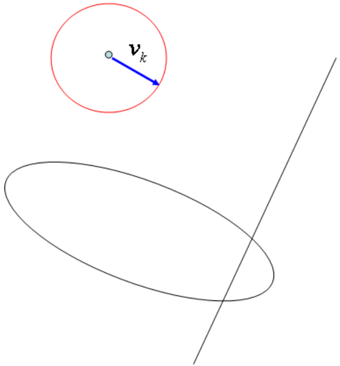
$$\min_{d \in \mathbb{R}^n} \ f_k + g_k^T d + \tfrac{1}{2} d^T W_k d \quad \text{s.t.} \ c_k + A_k d = 0$$

may not be well-defined or may lead to $\|d_k\| \gg 0$ and $\alpha_k \approx 0$

- We decompose the step by first considering the trust region subproblem

$$\min_{v \in \mathbb{R}^n} \ \tfrac{1}{2} \|c_k + A_k v\|^2 \quad \text{s.t.} \ \|v\| \leq \Omega_k$$

Trust region for the normal step v_k



The trust region keeps us in a local region of the search space

$$\min_{v \in \mathbb{R}^n} \tfrac{1}{2}\|c_k + A_k v\|^2$$
$$\text{s.t. } \|v\| \leq \Omega_k$$

Trust region for the tangential step $u_k = d_k - v_k$

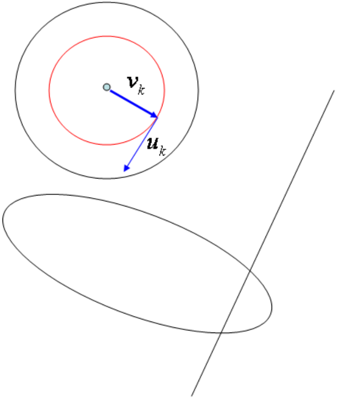
Once v_k is computed, we could compute a step toward optimality within a larger trust region

$$\min_{d \in \mathbb{R}^n} \; g_k^T d + \tfrac{1}{2} d^T W_k d$$
$$\text{s.t. } A_k d = A_k v_k$$
$$\|d\| \leq \Omega_k'$$

but then we may need

$$Z_k \quad \text{s.t.} \quad A_k Z_k = 0$$

or to project vectors onto $\text{Null}(A_k)$



Trust region only for the normal step v_k

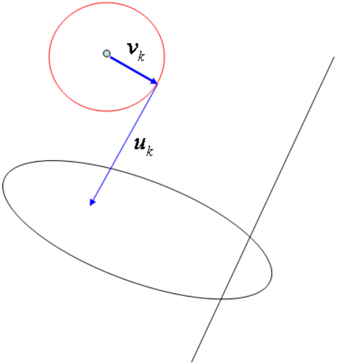
Instead, we set no trust region for u

$$\min_{d\in\mathbb{R}^n} \; g_k^T d + \tfrac{1}{2} d^T W_k d$$
$$\text{s.t. } A_k d = A_k v_k$$

which has the same solutions as

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{bmatrix} -(g_k + A_k^T \lambda_k) \\ A_k v_k \end{bmatrix}$$

Notice that this system is consistent (though perhaps (near) singular)



Setting the trust region radius

We propose a very specific form for the trust region radius; i.e., we solve

$$\min_{v\in\mathbb{R}^n} \ \tfrac{1}{2}\|c_k + A_k v\|^2$$
$$\text{s.t. } \|v\| \le \omega\|A_k^T c_k\|$$

for a given *constant* $\omega > 0$

- Problem information incorporated into the right-hand-side (note that a stationary point for the feasibility measure $\|c(x)\|$ has $\|A(x)^T c(x)\| = 0$)
- Radius is set dynamically without a heuristic update
- ω should correspond to the reciprocal of the smallest allowable singular value of A_k

Handling nonconvexity

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{bmatrix} -(g_k + A_k^T \lambda_k) \\ A_k v_k \end{bmatrix}$$

- Recall

$$\phi(x;\pi) = f(x) + \pi\|c(x)\|$$

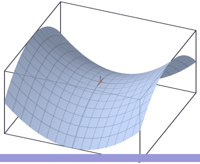
If W_k is not positive definite in the null space of A_k, then we may find

$$D\phi(d_k;\pi) > 0 \quad \forall \pi \geq 0$$

- As in unconstrained optimization

$$\nabla^2 f_k d_k = -\nabla f_k \quad \Rightarrow \quad \nabla f_k^T d_k = -d_k^T \nabla^2 f_k d_k$$

may not lead to a descent direction if $\nabla^2 f_k$ is not positive definite



Handling nonconvexity

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{bmatrix} -(g_k + A_k^T \lambda_k) \\ A_k v_k \end{bmatrix}$$

- Recall
$$\phi(x;\pi) = f(x) + \pi\|c(x)\|$$
If W_k is not positive definite in the null space of A_k, then we may find
$$D\phi(d_k;\pi) > 0 \quad \forall \pi \geq 0$$
- As in unconstrained optimization
$$\nabla^2 f_k d_k = -\nabla f_k \quad \Rightarrow \quad \nabla f_k^T d_k = -d_k^T \nabla^2 f_k d_k$$
may not lead to a descent direction if $\nabla^2 f_k$ is not positive definite
- However, it may be a descent direction, and for any direction descent is sufficient if
$$\nabla f_k^T d_k \leq -\max\{\tfrac{1}{2} d_k^T \nabla^2 f_k d_k, \theta\|d_k\|^2\}$$
for some $\theta > 0$

Handling nonconvexity

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{bmatrix} -(g_k + A_k^T \lambda_k) \\ A_k v_k \end{bmatrix}$$

Similarly, we can iteratively modify W_k to attempt to create a convex model, but we may stop whenever

$$\tfrac{1}{2} u_k^T W_k u_k \geq \theta \|u_k\|^2 \quad \text{or} \quad \|u_k\| \leq \psi \|v_k\|, \qquad (\psi \in (0,1))$$

or when the resulting step (d_k, δ_k) yields a sufficiently large reduction in a local linear model of $\phi(x;\pi)$:

$$m_k(d;\pi_k) \triangleq f_k + g_k^T d + \pi \|c_k + A_k d_k\|$$

i.e., whenever (d_k, δ_k) yields

$$\begin{aligned} \Delta m_k(d_k;\pi_k) = {} & m_k(0;\pi_k) - m_k(d_k;\pi_k) \\ = {} & -g_k^T d_k + \pi_k(\|c_k\| - \|c_k + A_k d_k\|) \gg 0 \end{aligned}$$

as in this case we have

$$D\phi(d_k;\pi_k) \leq -\Delta m_k(d_k;\pi_k) \ll 0$$

Regularized line search SQP

for $k = 0, 1, 2 \ldots$

- Solve
$$\min_{v \in \mathbb{R}^n} \tfrac{1}{2}\|c_k + A_k v\|^2, \quad \text{s.t. } \|v\| \le \omega \|A_k^T c_k\|$$
- Solve the *primal-dual* equations
$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{bmatrix} -(g_k + A_k^T \lambda_k) \\ A_k v_k \end{bmatrix},$$
iteratively modifying the Hessian until $d_k = u_k + v_k$ yields
$$\Delta m_k(d_k; \pi_k) \ge \max\{\tfrac{1}{2} u_k^T W_k u_k, \theta \|u_k\|^2\} + \sigma \pi_k (\|c_k\| - \|c_k + A_k v_k\|)$$
for an appropriate
$$\pi_k \ge \frac{g_k^T d_k + \max\{\frac{1}{2} u_k^T W_k u_k, \theta\|u_k\|^2\}}{(1-\tau)(\|c_k\| - \|c_k + A_k d_k\|)}$$
- Perform a line search to find $\alpha_k \in (0, 1]$ satisfying
$$\phi(x_k + \alpha_k d_k; \pi_k) \le \phi(x_k; \pi_k) - \eta \alpha_k \Delta m_k(d_k; \pi_k)$$

Outline

Working with matrices may be impractical

We compute search directions via

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = - \begin{bmatrix} g_k + A_k^T \lambda_k \\ c_k \end{bmatrix}$$

or

$$\min_{v \in \mathbb{R}^n} \ \tfrac{1}{2}\|c_k + A_k v\|^2 \quad \text{s.t.} \ \|v\| \le \omega \|A_k^T c_k\|$$

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{bmatrix} -(g_k + A_k^T \lambda_k) \\ A_k v_k \end{bmatrix}$$

What if...

- A_k, A_k^T, and W_k cannot be computed explicitly?
- A_k, A_k^T, and W_k cannot be stored?
- the *primal-dual matrix* cannot be factored?
- iterative methods may be more efficient?

Inexact normal step computation

Methods for the approximate solution of the trust region subproblem

$$\min_{v \in \mathbb{R}^n} \tfrac{1}{2}\|c_k + A_k v\|^2$$
$$\text{s.t. } \|v\| \leq \omega \|A_k^T c_k\|$$

are available

- For example, CG/LSQR with Steihaug/Toint stop tests
- However, we simply require $v_k \in \text{range}(A_k^T)$ and Cauchy decrease:

$$\|c_k\| - \|c_k + A_k v_k\| \geq \gamma(\|c_k\| - \|c_k + \alpha_k^c A_k v_k^c\|)$$

for some $\gamma > 0$ (note $v_k^c = -A_k^T c_k$)

Inexact primal-dual step computation

Apply an iterative linear solver to the symmetric indefinite system

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{bmatrix} -(g_k + A_k^T \lambda_k) \\ A_k v_k \end{bmatrix} + \begin{bmatrix} \rho_k \\ r_k \end{bmatrix}$$

- How can we be sure that a given inexact step is *acceptable*?
- How small do the residuals need to be?
- When should we modify W_k without being too conservative?

Central idea: Sufficient Model Reductions

Modern optimization algorithms work with models.

Again, take the penalty function

$$\phi(x;\pi) \triangleq f(x) + \pi\|c(x)\|$$

and consider the model

$$m_k(d;\pi) \triangleq f_k + g_k^T d + \pi\|c_k + A_k d\|$$

The reduction in m_k attained by d_k is computed easily as

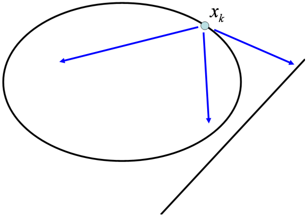
$$\begin{aligned}\Delta m_k(d_k;\pi) &\triangleq m_k(0;\pi) - m_k(d_k;\pi)\\ &= -g_k^T d_k + \pi(\|c_k\| - \|c_k + A_k d_k\|)\end{aligned}$$

and yields

$$D\phi(d_k;\pi) \leq -\Delta m_k(d_k;\pi)$$

Main tool: "SMART" Tests

Sufficient Merit function Approximation Reduction Termination Tests.

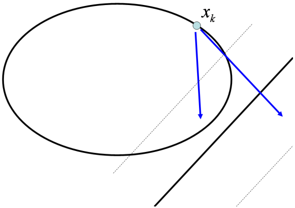


Termination Test I: A sufficient model reduction is attained for π_{k-1} (i.e., the most recent penalty parameter value):

$$\Delta m_k(d_k; \pi_{k-1}) = -g_k^T d_k + \pi_{k-1}(\|c_k\| - \|c_k + A_k d_k\|) \gg 0$$

Main tool: “SMART” Tests

Sufficient Merit function Approximation Reduction Termination Tests.



Termination Test II: A sufficient reduction in the constraint model is attained

$$\|c_k + A_k d_k\| \ll \|c_k\|$$

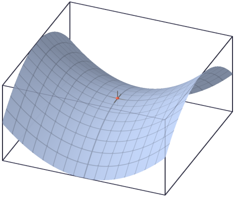
Hessian modifications

- We may need to modify the Hessian in order to guarantee that a descent direction will eventually be computed

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{bmatrix} -(g_k + A_k^T \lambda_k) \\ A_k v_k \end{bmatrix} + \begin{bmatrix} \rho_k \\ r_k \end{bmatrix}$$

- However, we care about the directions we are computing, not necessarily about the entire null space of the constraint Jacobian
- Current Hessian need not be modified if

$$\|u_k\| \leq \psi \|v_k\| \quad \text{or} \quad \tfrac{1}{2} u_k^T W_k u_k \geq \theta \|u_k\|^2$$



Step acceptance criteria:

Tangential Component Condition. The component u_k must satisfy

$$\|u_k\| \leq \psi\|v_k\| \text{ or } \tfrac{1}{2}u_k^T W_k u_k \geq \theta\|u_k\|^2 \text{ and } (g_k + W_k v_k)^T u_k + \tfrac{1}{2}u_k^T W_k u_k \leq 0$$

Model Reduction Condition. A step (d_k, δ_k) is acceptable if and only if

$$\Delta m_k(d_k;\pi_k) \geq \max\{\tfrac{1}{2}u_k^T W_k u_k, \theta\|u_k\|^2\} + \sigma\pi_k(\|c_k\| - \|c_k + A_k v_k\|)$$

for some $\sigma \in (0,1)$ and an appropriate $\pi_k > 0$.

Termination Test I. For some $\sigma \in (0,1)$ and $\pi_k = \pi_{k-1}$ the Tangential Component Condition holds, the Model Reduction Condition is satisfied, and for some $\kappa \in (0,1)$ we have

$$\left\|\begin{bmatrix}\rho_k \\ r_k\end{bmatrix}\right\| \leq \kappa \min\left\{\left\|\begin{bmatrix}g_k + A_k^T\lambda_k \\ A_k v_k\end{bmatrix}\right\|, \left\|\begin{bmatrix}g_{k-1} + A_{k-1}^T\lambda_k \\ A_{k-1}v_{k-1}\end{bmatrix}\right\|\right\}$$

Termination Test II. For some $\epsilon \in (0,1)$ and $\beta > 0$, the Tangential Component Condition holds and we have

$$\begin{aligned} \|c_k\| - \|c_k + A_k d_k\| &\geq \epsilon(\|c_k\| - \|c_k + A_k v_k\|) \\ \text{and} \quad \|\rho_k\| &\leq \beta(\|c_k\| - \|c_k + A_k v_k\|), \\ \text{and we set} \quad \pi_k &\geq (g_k^T d_k + \tfrac{1}{2}u_k^T W_k u_k)/((1-\tau)(\|c_k\| - \|c_k + A_k d_k\|)) \end{aligned}$$

Regularized line search SQP with SMART Tests

for $k = 0, 1, 2 \ldots$

- Solve
$$\min_{v \in \mathbb{R}^n} \tfrac{1}{2}\|c_k + A_k v\|^2, \quad \text{s.t. } \|v\| \leq \omega \|A_k^T c_k\|$$
- Solve the *primal-dual* equations
$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{bmatrix} -(g_k + A_k^T \lambda_k) \\ A_k v_k \end{bmatrix},$$
until a termination test is satisfied, modifying the Hessian when needed; i.e., until
$$\Delta m_k(d_k; \pi_k) \geq \max\{\tfrac{1}{2} u_k^T W_k u_k, \theta \|u_k\|^2\} + \sigma \pi_k (\|c_k\| - \|c_k + A_k v_k\|)$$
for
$$\pi_k \geq \frac{g_k^T d_k + \max\{\tfrac{1}{2} u_k^T W_k u_k, \theta \|u_k\|^2\}}{(1-\tau)(\|c_k\| - \|c_k + A_k d_k\|)}$$
- Perform a line search to find $\alpha_k \in (0, 1]$ satisfying
$$\phi(x_k + \alpha_k d_k; \pi_k) \leq \phi(x_k; \pi_k) - \eta \alpha_k \Delta m_k(d_k; \pi_k)$$

Outline

Main result

Assumptions: The generated sequence $\{x_k, \lambda_k\}$ is contained in a convex set over which f and c and their first derivatives are bounded, and the iterative linear system solver can solve the primal-dual equations to an arbitrary accuracy

Theorem: If all limit points satisfy the linear independence constraint qualification (LICQ), then $\{\pi_k\}$ is bounded and

$$\lim_{k \to \infty} \left\| \begin{bmatrix} g_k + A_k^T \lambda_{k+1} \\ c_k \end{bmatrix} \right\| = 0$$

Otherwise,

$$\lim_{k \to \infty} \left\| A_k^T c_k \right\| = 0$$

and if $\{\pi_k\}$ is bounded then

$$\lim_{k \to \infty} \left\| g_k + A_k^T \lambda_{k+1} \right\| = 0$$

Brief overview of analysis

- The step length (d_k, v_k, u_k) is explicitly or implicitly controlled...
- The reduction in the model of the penalty function satisfies

$$\Delta m_k(d_k; \pi_k) \geq \gamma(\|u_k\|^2 + \pi_k \|A_k^T c_k\|^2)$$

- In particular

$$\Delta m_k(d_k; \pi_k) \geq \gamma' \|A_k^T c_k\|^2 \Rightarrow \lim_{k\to\infty} \|A_k^T c_k\| = 0$$

- If $\{\pi_k\}$ remains bounded (guaranteed if LICQ holds), then

$$\lim_{k\to\infty} \left\| g_k + A_k^T \lambda_{k+1} \right\| = 0,$$

and otherwise $\pi \to \infty$

Implementation details

We use MINRES to solve the primal-dual equations

$$\begin{bmatrix} W_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} d_k \\ \delta_k \end{bmatrix} = \begin{cases} -\begin{bmatrix} g_k + A_k^T \lambda_k \\ c_k \end{bmatrix} \\ -\begin{bmatrix} g_k + A_k^T \lambda_k \\ -A_k v_k \end{bmatrix} \end{cases}$$

and LSQR (algebraically equivalent to CG, but with better numerical properties) with Steihaug-Toint stop tests to solve the trust region subproblem

$$\min_{v \in \mathbb{R}^n} \tfrac{1}{2}\|c_k + A_k v\|^2$$
$$\text{s.t. } \|v\| \leq \omega \|A_k^T c_k\|$$

All experiments performed in Matlab

Problems with rank-deficiency

Total of 73 problems from the CUTEr collection

- Original and perturbed models have

$$c_1(x) = 0 \quad \text{and} \quad \begin{cases} c_1(x) = 0 \\ c_1(x) - c_1^2(x) = 0 \end{cases}$$

  respectively

- Success rates:

Original	100%
Perturbed	93%

- A few of the failures were due to the Maratos effect, so second-order correction steps may be beneficial

Outline

Conclusion

We have:

- focused on a particular class of problems to which contemporary optimization techniques cannot be applied
- considered the fundamental question of how to ensure global convergence via a type of inexact SQP/Newton approach
- developed a methodology where inexact solutions are appraised based on the reductions obtained in linear models of an exact penalty function
- extended the algorithm and analysis for cases involving rank deficiency and nonconvexity

Papers:

- R. H. Byrd, F. E. Curtis, and J. Nocedal, "An Inexact SQP Method for Equality Constrained Optimization," to appear in SIAM Journal on Optimization.
- R. H. Byrd, F. E. Curtis, and J. Nocedal, "An Inexact Newton Method for Nonconvex Equality Constrained Optimization," to appear in Mathematical Programming.
- F. E. Curtis, J. Nocedal, and A. Wächter, "A Matrix-free Algorithm for Equality Constrained Optimization Problems with Rank-deficient Jacobians," in preparation.