Sequential Quadratic Optimization with Inexact Subproblem Solves

Frank E. Curtis, Lehigh University

involving joint work with

Travis Johnson, Northwestern University
Daniel P. Robinson, Johns Hopkins University
Andreas Wächter, Northwestern University

ICCOPT 2013 — Lisbon, Portugal

29 July 2013

Outline

Motivation

"Exact" Algorithm

"Inexact" Algorithm

Numerical Experiments

Summary

Problem formulation

Our goal is to solve a generic constrained nonlinear optimization problem:

$$\begin{aligned} \min_x \ & f(x) \\ \text{s.t. } & c(x) = 0,\ \bar{c}(x) \leq 0. \end{aligned} \tag{NLP}$$

If (NLP) is infeasible, then at least we want to minimize constraint violation:

$$\min_x \ v(x), \text{ where } v(x) := \|c(x)\|_1 + \|[\bar{c}(x)]^+\|_1. \tag{FP}$$

(A minimizer of (NLP) is always a minimizer of (FP).)

Sequential quadratic optimization

Advantages:

- "Parameter free" search direction computation (ideally)
- Strong global convergence properties and behavior
- Active-set identification $\Longrightarrow$ Newton-like local convergence

Disadvantage:

- Quadratic subproblems (QPs) are expensive to solve exactly

Sequential quadratic optimization w/ inexactness

Contributions:

- Implementable termination conditions for inexact QP solves
- No specific QP solver required; any will do
- Global convergence guarantees (feasible and infeasible problems)
- Future work: Local convergence (feasible and infeasible problems)[1]

Algorithmic features:

- Allows generic inexactness in QP solutions
- Negative curvature handled with dynamic Hessian modifications
- Separate multipliers for (NLP) and (FP)
- Dynamic updates for penalty parameter and Lagrange multipliers

[1] Avoid using "Cauchy points" that only yield minimal progress for global convergence.

iSQP vs. iSQO

August 22, 2012: 19:45:00

Photo courtesy of Jim Luedtke.

iSQP vs. **iSQO**

August 22, 2012: 19:45:01

Photo courtesy of Jim Luedtke.

Fritz John and penalty functions

(NLP):

$$\min_x \; f(x)$$
$$\text{s.t. } c(x) = 0,\; \bar{c}(x) \leq 0$$

(FP):

$$\min_x \; v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

Define the Fritz John (FJ) function

$$\mathcal{F}(x, y, \bar{y}, \mu) := \mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

and the ℓ_1-norm exact penalty function

$$\phi(x, \mu) := \mu f(x) + v(x).$$

($\mu \geq 0$ acts as objective multiplier/penalty parameter.)

Optimality conditions

(NLP):

$$\min_x \; f(x)$$
$$\text{s.t. } c(x) = 0, \; \bar{c}(x) \le 0$$

(FP):

$$\min_x \; v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \; \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) := \mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT conditions for (FP) and (PP) expressed with residual

$$\rho(x,y,\bar{y},\mu) := \begin{bmatrix} \mu g(x) + J(x)y + \bar{J}(x)\bar{y} \\ \min\{[c(x)]^+, e - y\} \\ \min\{[c(x)]^-, e + y\} \\ \min\{[\bar{c}(x)]^+, e - \bar{y}\} \\ \min\{[\bar{c}(x)]^-, \quad \bar{y}\} \end{bmatrix}$$

- FJ point:

$$\rho(x,y,\bar{y},\mu) = 0, \; v(x) = 0, \; (y,\bar{y},\mu) \ne 0$$

- KKT point:

$$\rho(x,y,\bar{y},\mu) = 0, \; v(x) = 0, \; \mu > 0$$

- Infeasible stationary point:

$$\rho(x,y,\bar{y},0) = 0, \; v(x) > 0$$

Penalty function model and QP subproblems

(NLP):

$$\min_x f(x)$$
$$\text{s.t. } c(x) = 0, \;\; \bar{c}(x) \leq 0$$

(FP):

$$\min_x \; v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \; \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) := \mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residual:

$$\rho(x,y,\bar{y},\mu)$$

Define a local model of $\phi(\cdot,\mu)$ at x_k:

$$l_k(d,\mu) := \mu(f_k + g_k^T d) + \|c_k + J_k^T d\|_1 + \|[\bar{c}_k + \bar{J}_k^T d]^+\|_1$$

Reduction in this model yielded by a given d:

$$\Delta l_k(d,\mu) := \Delta l(0,\mu) - \Delta l(d,\mu)$$

Two subproblems of interest:

$$\min_d \; -\Delta l_k(d,\mu_k) + \tfrac{1}{2} d^T H_k' d \qquad \text{(PQP)}$$

$$\min_d \; -\Delta l_k(d,0) + \tfrac{1}{2} d^T H_k'' d \qquad \text{(FQP)}$$

$\Delta l_k(d,\mu) > 0$ implies d is a direction of strict descent for $\phi(\cdot,\mu)$ from x_k

Optimality conditions (for QPs)

(NLP):

$$\min_x f(x)$$

$$\text{s.t. } c(x) = 0, \ \bar{c}(x) \le 0$$

(FP):

$$\min_x v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) :=$$

$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residual:

$$\rho(x,y,\bar{y},\mu)$$

Local model of ϕ at x_k:

$$l_k(d,\mu)$$

KKT conditions for (PQP) and (FQP) expressed with

$$\rho_k(d,y,\bar{y},\mu,H) := \begin{bmatrix} \mu g_k + Hd + J_k y + \bar{J}_k \bar{y} \\ \min\{[c_k + J_k^T d]^+, e - y\} \\ \min\{[c_k + J_k^T d]^-, e + y\} \\ \min\{[\bar{c}_k + \bar{J}_k^T d]^+, e - \bar{y}\} \\ \min\{[\bar{c}_k + \bar{J}_k^T d]^-, \quad \bar{y}\} \end{bmatrix}$$

- "Exact" solution of (PQP):

$$\rho_k(d,y,\bar{y},\mu,H) = 0$$

- "Exact" solution of (FQP):

$$\rho_k(d,y,\bar{y},0,H) = 0$$

Scenario A

(NLP):

$$\min_x f(x)$$
$$\text{s.t. } c(x) = 0,\ \bar{c}(x) \le 0$$

(FP):

$$\min_x v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x,y,\bar{y},\mu)$$
$$\rho_k(d,y,\bar{y},\mu,H)$$

Local model of ϕ at x_k:

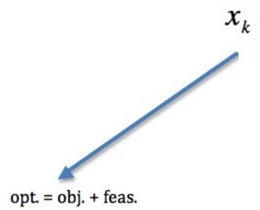
$$l_k(d,\mu)$$

Solution $(d_k', y_{k+1}', \bar{y}_{k+1}')$ of (PQP) satisfies

- $\rho(d_k', y_{k+1}', \bar{y}_{k+1}', \mu_k, H_k') = 0$
- $\Delta l_k(d_k', \mu_k) \ge \epsilon v_k$ for $\epsilon \in (0,1)$

then

- $d_k \leftarrow d_k'$ is the search direction
- $\mu_{k+1} \leftarrow \mu_k$



Scenario B

(NLP):

$$\min_x f(x)$$
$$\text{s.t. } c(x) = 0, \; \bar{c}(x) \leq 0$$

(FP):

$$\min_x v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \phi(x, \mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x, y, \bar{y}, \mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x, y, \bar{y}, \mu)$$
$$\rho_k(d, y, \bar{y}, \mu, H)$$

Local model of ϕ at x_k:

$$l_k(d, \mu)$$

Solution $(d'_k, y'_{k+1}, \bar{y}'_{k+1})$ of (PQP) satisfies

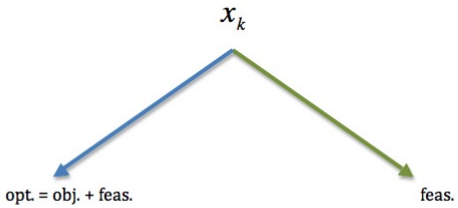
- $\rho(d'_k, y'_{k+1}, \bar{y}'_{k+1}, \mu_k, H'_k) = 0$
- $\Delta l_k(d'_k, \mu_k) \geq \epsilon \Delta l_k(d''_k, 0)$ for $\epsilon \in (0, 1)$

where solution $(d''_k, y''_{k+1}, \bar{y}''_{k+1})$ of (FQP) satisfies

- $\rho(d''_k, y''_{k+1}, \bar{y}''_{k+1}, 0, H''_k) = 0$

then

- $d_k \leftarrow d'_k$ is the search direction
- $\mu_{k+1} \leftarrow \mu_k$



Scenario C

(NLP):

$$\min_x f(x)$$
$$\text{s.t. } c(x) = 0, \ \bar{c}(x) \le 0$$

(FP):

$$\min_x v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x,y,\bar{y},\mu)$$
$$\rho_k(d,y,\bar{y},\mu,H)$$

Local model of ϕ at x_k:

$$l_k(d,\mu)$$

Solution $(d_k', y_{k+1}', \bar{y}_{k+1}')$ of (PQP) satisfies

- $\rho(d_k', y_{k+1}', \bar{y}_{k+1}', \mu_k, H_k') = 0$

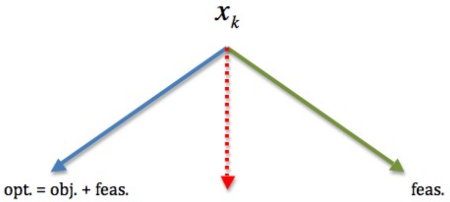
and solution $(d_k'', y_{k+1}'', \bar{y}_{k+1}'')$ of (FQP) satisfies

- $\rho(d_k'', y_{k+1}'', \bar{y}_{k+1}'', 0, H_k'') = 0$

then

- $d_k \leftarrow \tau d_k' + (1-\tau) d_k''$ so $\Delta l_k(d_k, 0) \ge \epsilon \Delta l_k(d_k'', 0)$
- $\mu_{k+1} < \mu_k$ so $\Delta l_k(d_k, \mu_k) \ge \beta \Delta l_k(d_k, 0)$ for $\beta \in (0,1)$
- Overall:

$$\Delta l_k(d_k,\mu_k) \ge \beta \Delta l_k(d_k, 0) \ge \beta\epsilon \Delta l_k(d_k'', 0)$$



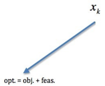
"Exact" SQO

repeat

(1) Check whether KKT point or infeasible stationary point has been obtained.

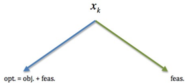
(2) Compute an exact solution of (PQP).

   (a) If Scenario A occurs, then go to step 4.

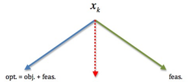


(3) Compute an exact solution of (FQP).

   (a) If Scenario B occurs, then go to step 4.



   (b) Otherwise, Scenario C occurs.



(4) Perform a backtracking line search to reduce $\phi(\cdot, \mu_{k+1})$.

endrepeat

Preview

Our inexact algorithm involves 6 (wow!) scenarios:

- 3 are derived from scenarios in the "exact" algorithm
- 3 focus on penalty parameter and Lagrange multiplier updates

Two critical questions to answer:

- When can we terminate the QP solver?
- When is a given inexact solution good enough?

(Some details will be suppressed; only highlights given for simplicity)

Terminating the QP solver: Test P1

(NLP):

$$\min_x \; f(x)$$
$$\text{s.t. } c(x) = 0, \;\; \bar{c}(x) \le 0$$

(FP):

$$\min_x \; v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \; \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x,y,\bar{y},\mu)$$
$$\rho_k(d,y,\bar{y},\mu,H)$$

Local model of ϕ at x_k:

$$l_k(d,\mu)$$

Solution $(d'_k, y'_{k+1}, \bar{y}'_{k+1})$ of (PQP) satisfies

- $y'_{k+1} \in [-e, e]$, $\bar{y}'_{k+1} \in [0, e]$
- $\Delta l_k(d'_k, \mu_k) \ge \theta \|d'_k\|^2 > 0$ for $\theta \in (0,1)$
- $\|\rho_k(d'_k, y'_{k+1}, \bar{y}'_{k+1}, \mu_k, H'_k)\| \le \kappa \left\| \begin{bmatrix} \rho(x_k, y'_k, \bar{y}'_k, \mu_k) \\ \rho(x_k, y''_k, \bar{y}''_k, 0) \end{bmatrix} \right\|$

Terminating the QP solver: Test P2

(NLP):

$$\min_x f(x)$$
$$\text{s.t. } c(x) = 0, \; \bar{c}(x) \leq 0$$

(FP):

$$\min_x v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x,y,\bar{y},\mu)$$
$$\rho_k(d,y,\bar{y},\mu,H)$$

Local model of ϕ at x_k:

$$l_k(d,\mu)$$

Solution $(d_k', y_{k+1}', \bar{y}_{k+1}')$ of (PQP) satisfies

- $y_{k+1}' \in [-e,e]$, $\bar{y}_{k+1}' \in [0,e]$
- $\Delta l_k(d_k',\mu_k) \geq \theta \|d_k'\|^2 > 0$ for $\theta \in (0,1)$
- $\|\rho_k(d_k', y_{k+1}', \bar{y}_{k+1}', \mu_k, H_k')\| \leq \kappa \|\rho(x_k, y_k'', \bar{y}_k'', 0)\|$

Furthermore,

- If $\Delta l_k(d_k', 0) < \epsilon v_k$, then
$$\|(y_{k+1}', \bar{y}_{k+1}')\|_\infty \gg 0$$
- If $\Delta l_k(d_k', \mu_k) < \beta \Delta l_k(d_k', 0)$, then
$$\|(y_{k+1}', \bar{y}_{k+1}')\|_\infty \gg 0$$

Terminating the QP solver: Test P3

(NLP):

$$\min_x \; f(x)$$
$$\text{s.t. } c(x) = 0,\; \bar{c}(x) \le 0$$

(FP):

$$\min_x \; v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \; \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x,y,\bar{y},\mu)$$
$$\rho_k(d,y,\bar{y},\mu,H)$$

Local model of ϕ at x_k:

$$l_k(d,\mu)$$

Solution $(d_k', y_{k+1}', \bar{y}_{k+1}')$ of (PQP) satisfies

- $y_{k+1}' \in [-e,e]$, $\bar{y}_{k+1}' \in [0,e]$
- $\|\rho_k(0, y_{k+1}', \bar{y}_{k+1}', \mu_k, H_k')\| \le \kappa \|\rho(x_k, y_k', \bar{y}_k', \mu_k)\|$

Furthermore, for $\zeta \in (0,1)$,

- If $\|\rho(x_k, y_k', \bar{y}_k', \mu_k)\| < \zeta \|\rho(x_k, y_k'', \bar{y}_k'', 0)\|$, then

$$\|(y_{k+1}', \bar{y}_{k+1}')\|_\infty \gg 0$$

Terminating the QP solver: Test F

(NLP):

$$\min_x f(x)$$
$$\text{s.t. } c(x) = 0, \ \bar{c}(x) \leq 0$$

(FP):

$$\min_x v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \phi(x, \mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x, y, \bar{y}, \mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x, y, \bar{y}, \mu)$$
$$\rho_k(d, y, \bar{y}, \mu, H)$$

Local model of ϕ at x_k:

$$l_k(d, \mu)$$

Solution $(d_k'', y_{k+1}'', \bar{y}_{k+1}'')$ of (FQP) satisfies

- $y_{k+1}'' \in [-e, e]$, $\bar{y}_{k+1}'' \in [0, e]$
- $\max\{\Delta l_k(d_k', \mu_k), \Delta l_k(d_k'', 0)\} \geq \theta \|d_k''\|^2$ for $\theta \in (0, 1)$
- $\|\rho_k(d_k'', y_{k+1}'', \bar{y}_{k+1}'', 0, H_k'')\| \leq \kappa \|\rho(x_k, y_k'', \bar{y}_k'', 0)\|$

Scenario 1

(NLP):

$$\min_x \; f(x)$$
$$\text{s.t. } c(x) = 0, \;\; \bar{c}(x) \leq 0$$

(FP):

$$\min_x \; v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \; \phi(x, \mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x, y, \bar{y}, \mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x, y, \bar{y}, \mu)$$
$$\rho_k(d, y, \bar{y}, \mu, H)$$

Local model of ϕ at x_k:

$$l_k(d, \mu)$$

The current iterate satisfies

- $\|\rho(x_k, y_k', \bar{y}_k', \mu_k)\| = 0$
- $v(x_k) > 0$

then

- $d_k \leftarrow 0$ is the search direction
- $\mu_{k+1} < \delta \mu_k$ for $\delta \in (0, 1)$

Scenario 2

(NLP):

$$\min_x \; f(x)$$
$$\text{s.t. } c(x) = 0, \;\; \bar{c}(x) \leq 0$$

(FP):

$$\min_x \; v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \; \phi(x, \mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x, y, \bar{y}, \mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x, y, \bar{y}, \mu)$$
$$\rho_k(d, y, \bar{y}, \mu, H)$$

Local model of ϕ at x_k:

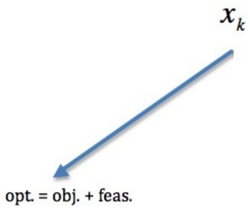
$$l_k(d, \mu)$$

Solution $(d_k', y_{k+1}', \bar{y}_{k+1}')$ of (PQP) satisfies

- Test P1
- $\Delta l_k(d_k', \mu_k) \geq \epsilon v_k$ for $\epsilon \in (0,1)$

then

- $d_k \leftarrow d_k'$ is the search direction
- $\mu_{k+1} \leftarrow \mu_k$



Scenario 3

(NLP):

$$\min_x \; f(x)$$
$$\text{s.t. } c(x) = 0, \;\; \bar{c}(x) \le 0$$

(FP):

$$\min_x \; v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \; \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x,y,\bar{y},\mu)$$
$$\rho_k(d,y,\bar{y},\mu,H)$$

Local model of ϕ at x_k:

$$l_k(d,\mu)$$

Solution $(d'_k, y'_{k+1}, \bar{y}'_{k+1})$ of (PQP) satisfies

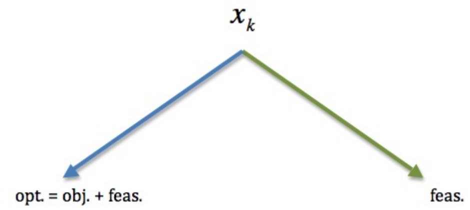
- Test P1
- $\Delta l_k(d'_k, \mu_k) \ge \epsilon \Delta l_k(d''_k, 0)$ for $\epsilon \in (0,1)$

where solution $(d''_k, y''_{k+1}, \bar{y}''_{k+1})$ of (FQP) satisfies

- Test F

then

- $d_k \leftarrow d'_k$ is the search direction
- $\mu_{k+1} \leftarrow \mu_k$



Scenario 4

(NLP):

$$\min_x \; f(x)$$
$$\text{s.t. } c(x) = 0, \;\; \bar{c}(x) \le 0$$

(FP):

$$\min_x \; v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \; \phi(x,\mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x,y,\bar{y},\mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x,y,\bar{y},\mu)$$
$$\rho_k(d,y,\bar{y},\mu,H)$$

Local model of ϕ at x_k:

$$l_k(d,\mu)$$

Solution $(d_k', y_{k+1}', \bar{y}_{k+1}')$ of (PQP) satisfies

- Test P2

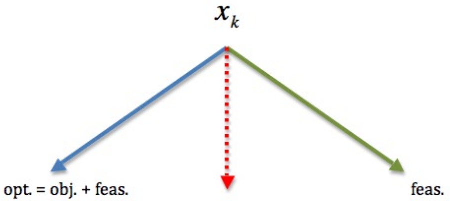
and solution $(d_k'', y_{k+1}'', \bar{y}_{k+1}'')$ of (FQP) satisfies

- Test F

then

- $d_k \leftarrow \tau d_k' + (1-\tau) d_k''$ so $\Delta l_k(d_k, 0) \ge \epsilon \Delta l_k(d_k'', 0)$
- $\mu_{k+1} < \mu_k$ so $\Delta l_k(d_k, \mu_k) \ge \beta \Delta l_k(d_k, 0)$ for $\beta \in (0,1)$
- Overall:

$$\Delta l_k(d_k, \mu_k) \ge \beta \Delta l_k(d_k, 0) \ge \beta\epsilon \Delta l_k(d_k'', 0)$$



Scenarios 5 & 6

(NLP):

$$\min_x f(x)$$
$$\text{s.t. } c(x) = 0, \ \bar{c}(x) \le 0$$

(FP):

$$\min_x v(x) := \left\| \begin{bmatrix} c(x) \\ [\bar{c}(x)]^+ \end{bmatrix} \right\|_1$$

(PP):

$$\min_x \phi(x, \mu) := \mu f(x) + v(x)$$

(FJ):

$$\mathcal{F}(x, y, \bar{y}, \mu) :=$$
$$\mu f(x) + c(x)^T y + \bar{c}(x)^T \bar{y}$$

KKT residuals:

$$\rho(x, y, \bar{y}, \mu)$$
$$\rho_k(d, y, \bar{y}, \mu, H)$$

Local model of ϕ at x_k:

$$l_k(d, \mu)$$

Solution $(d'_k, y'_{k+1}, \bar{y}'_{k+1})$ of (PQP) satisfies

- Test P3

If $\|\rho(x_k, y'_k, \bar{y}'_k, \mu_k)\| < \zeta \|\rho(x_k, y''_k, \bar{y}''_k, 0)\|$, then

- $\mu_{k+1} \leftarrow \mu_k$ (Scenario 5)

else

- $\mu_{k+1} < \mu_k$ (Scenario 6)

"Inexact" SQO (iSQO)

repeat

(1) Check whether KKT point or infeasible stationary point has been obtained.

(2) Check for a trivial iteration.

  (a) If Scenario 1 occurs, then go to step 6.

(3) Compute an inexact solution of (PQP) satisfying Test P1 or P3.

  (a) If Scenario 2 occurs, then go to step 6.

(4) Compute an inexact solution of (FQP) satisfying Test F.

  (a) If Scenario 3 occurs, then go to step 6.
  (b) If Scenario 5 occurs, then go to step 6.

(5) Compute an inexact solution of (PQP) satisfying Test P2 or P3.
If Test 1 holds, then compute an inexact solution of (FQP) satisfying Test F.

  (a) If Scenario 3 occurs, then go to step 6.
  (b) If Scenario 4 occurs, then go to step 6.
  (c) If Scenario 5 occurs, then go to step 6.
  (d) Otherwise, Scenario 6 occurs.

(6) Perform a backtracking line search to reduce $\phi(\cdot, \mu_{k+1})$.

endrepeat

Well-posedness

Assumption

The following hold:

(1) *The functions f, c, and $\bar{c}$ are continuously differentiable in an open convex set Ω containing the sequences $\{x_k\}$ and $\{x_k + d_k\}$.*

(2) *The QP solver can solve* (PQP) *arbitrarily accurately.*

(3) *The QP solver can solve* (FQP) *arbitrarily accurately.*

Lemma

In iteration k, either iSQO *terminates or exactly one of Scenario 1–6 will occur.*

Theorem

One of the following holds:

1. iSQO *terminates with a KKT point or infeasible stationary point.*
2. iSQO *generates an infinite sequence of iterates*

$$\left(x_k, \begin{bmatrix} y_k' \\ \bar{y}_k' \end{bmatrix}, \begin{bmatrix} y_k'' \\ \bar{y}_k'' \end{bmatrix}, \mu_k\right) \text{ where } \begin{bmatrix} y_k' \\ y_k'' \end{bmatrix} \in [-e, e], \ \begin{bmatrix} \bar{y}_k' \\ \bar{y}_k'' \end{bmatrix} \in [0, e], \text{ and } \mu_k > 0.$$

Global convergence

Assumption

(1) *The well-posedness assumptions still apply.*

(2) *The functions f, c, and $\bar{c}$ and their first derivatives are bounded and Lipschitz continuous in Ω (the open convex set containing $\{x_k\}$ and $\{x_k + d_k\}$).*

(3) *The sequences $\{H'_k\}$ and $\{H''_k\}$—including the initial and any modified values of H'_k and H''_k—are bounded.*

Lemma

The KKT residual for (FP) *vanishes:*

$$\lim_{k\to\infty} \|\rho(x_k, y''_k, \bar{y}''_k, 0)\| = 0.$$

Lemma

The KKT residual for (PP) vanishes:

$$\lim_{k\to\infty} \|\rho(x_k, y'_k, \bar{y}'_k, \mu_k)\| = 0.$$

Global convergence: Vanishing penalty parameter

(If $\{\mu_k\} \geq \underline{\mu}$ for some $\underline{\mu} > 0$, then we know the KKT residual for (PP) vanishes.)

Lemma

If $\mu_k \to 0$, then either all limit points of $\{x_k\}$ are feasible or all are infeasible.

Lemma

If $\mu_k \to 0$ and all limit points of $\{x_k\}$ are feasible, then, with

$$K_\mu := \{k : \mu_{k+1} < \mu_k\},$$

all limit points of $\{x_k\}_{k \in K_\mu}$ are FJ points.

Lemma

If $\rho(x_*, y_*, \bar{y}_*, 0) = 0$, x_* is feasible, and $(y_*, \bar{y}_*) \neq 0$, then the MFCQ fails at x_*.

Global convergence

Theorem

One of the following holds:

(a) $\mu_k = \underline{\mu}$ *for some* $\underline{\mu} > 0$ *for all large* k *and either every limit point of* $\{x_k\}$ *corresponds to a KKT point or is an infeasible stationary point;*

(b) $\mu_k \to 0$ *and every limit point of* $\{x_k\}$ *is an infeasible stationary point;*

(c) $\mu_k \to 0$*, all limit points of* $\{x_k\}$ *are feasible, and, with*

$$K_\mu := \{k : \mu_{k+1} < \mu_k\},$$

every limit point of $\{x_k\}_{k \in K_\mu}$ *corresponds to an FJ point where the MFCQ fails.*

Corollary

If $\{x_k\}$ *is bounded and every limit point of this sequence is a feasible point at which the MFCQ holds, then* $\mu_k = \underline{\mu}$ *for some* $\underline{\mu} > 0$ *for all large* k *and every limit point of* $\{x_k\}$ *corresponds to a KKT point.*

Implementation details

- Matlab implementation
- `BQPD` for QP solves with indefinite Hessians; see (Fletcher, 2000)
- *Simulated* inexactness by perturbing QP solutions
- Test set involves 307 `CUTEr` problems with
  - at least one free variable
  - at least one general (non-bound) constraint
  - at most 200 variables and constraints (because it's Matlab!)
- Termination conditions ($\epsilon_{tol} = 10^{-6}$ and $\epsilon_\mu = 10^{-8}$):

$$\begin{aligned}
\left\|\rho(x_k, y_k', \bar{y}_k', \mu_k)\right\| &\leq \epsilon_{tol} \quad \text{and} \quad v_k \leq \epsilon_{tol}; && \text{(Optimal)} \\
\left\|\rho(x_k, y_k'', \bar{y}_k'', 0)\right\| &= 0 \quad \text{and} \quad v_k > 0; && \text{(Infeasible)} \\
\left\|\rho(x_k, y_k'', \bar{y}_k'', 0)\right\| &\leq \epsilon_{tol} \quad \text{and} \quad v_k > \epsilon_{tol} \quad \text{and} \quad \mu_k \leq \epsilon_\mu. && \text{(Infeasible)}
\end{aligned}$$

- Investigate performance of inexact algorithm with $\kappa = 0.01$, 0.1, and 0.5.

Success statistics

Counts of termination messages for exact and three variants of inexact algorithm:

Termination message	Exact	Inexact		
		$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 0.5$
Optimal solution found	271	269	272	275
Infeasible stationary point found	4	3	2	2
Iteration limit reached	12	10	11	9
Subproblem solver failure	18	23	20	19

Termination statistics and reliability do not degrade with inexactness.

Inexactness levels

Observe relative residuals for QP solves:

$$\kappa_I := \frac{\|\rho_k(d, y, \bar{y}, \mu_k, H'_k)\|}{\|\rho(x_k, y'_k, \bar{y}'_k, \mu_k)\|} \quad \text{or} \quad \kappa_I := \frac{\|\rho_k(d, y, \bar{y}, 0, H''_k)\|}{\|\rho(x_k, y''_k, \bar{y}''_k, 0)\|},$$

For problem j, we compute minimum ($\kappa_I(j)$) and mean ($\bar{\kappa}_I(j)$) values over run:

min	κ	$\kappa_{I,\text{mean}}$	$[0,10^{-8})$	$[10^{-8},10^{-6})$	$[10^{-6},10^{-4})$	$[10^{-4},10^{-3})$	$[10^{-3},0.01)$	$[0.01,0.1)$	$[0.1,0.5)$	$[0.5,1)$	$[1,\infty)$
$\kappa_I(j)$	0.01	3.5e-03	0	2	10	7	253	0	0	0	0
	0.1	2.8e-02	0	0	2	10	30	232	0	0	0
	0.5	8.8e-02	0	0	2	4	23	69	179	0	0
mean	κ	$\bar{\kappa}_{I,\text{mean}}$									
$\bar{\kappa}_I(j)$	0.01	7.3e-03	0	0	0	0	254	18	0	0	0
	0.1	6.9e-02	0	0	0	0	0	261	13	0	0
	0.5	3.5e-01	0	0	0	0	0	1	264	12	0

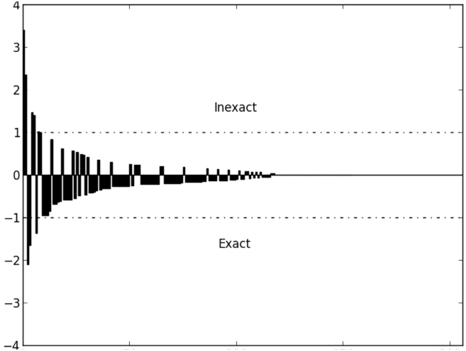
Relative residuals generally need only be moderately smaller than parameter κ.

Iteration comparison

Considering the logarithmic outperforming factor

$$r^j := -\log_2(\text{iter}^j_{\text{inexact}}/\text{iter}^j_{\text{exact}}),$$

we compare iteration counts of our inexact ($\kappa = 0.01$) and exact algorithms:



Iteration counts do not degrade significantly with inexactness.

Summary

- Developed, analyzed, and experimented with an inexact SQO method
- Allows generic inexactness in QP subproblem solves
- No specific QP solver required
- Global convergence guarantees established
- Numerical experiments suggest inexact algorithm is reliable
- Inexact solutions allowed without degradation of performance

Thanks!

"Exact" Algorithms:

- J. V. Burke, F. E. Curtis, and H. Wang, "A Sequential Quadratic Optimization Algorithm with Rapid Infeasibility Detection," in first revision for *SIAM Journal on Optimization*, originally submitted 2012.
- R. H. Byrd, F. E. Curtis, and J. Nocedal, "Infeasibility Detection and SQP Methods for Nonlinear Optimization," *SIAM Journal on Optimization*, Volume 20, Issue 5, pg. 2281-2299, 2010.

"Inexact" Algorithms:

- F. E. Curtis, T. C. Johnson, D. P. Robinson, and A. Wächter, "An Inexact Sequential Quadratic Optimization Algorithm for Large-Scale Nonlinear Optimization," submitted to *SIAM Journal on Optimization*, 2013.
- F. E. Curtis, J. Huber, O. Schenk, and A. Wächter, "A Note on the Implementation of an Interior-Point Algorithm for Nonlinear Optimization with Inexact Step Computations," *Mathematical Programming, Series B*, Volume 136, Issue 1, pg. 209–227, 2012.
- F. E. Curtis, O. Schenk, and A. Wächter, "An Interior-Point Algorithm for Large-Scale Nonlinear Optimization with Inexact Step Computations," *SIAM Journal on Scientific Computing*, Volume 32, Issue 6, pg. 3447-3475, 2010.
- R. H. Byrd, F. E. Curtis, and J. Nocedal, "An Inexact Newton Method for Nonconvex Equality Constrained Optimization," *Mathematical Programming*, Volume 122, Issue 2, pg. 273-299, 2010.
- F. E. Curtis, J. Nocedal, and A. Wächter, "A Matrix-free Algorithm for Equality Constrained Optimization Problems with Rank-Deficient Jacobians," *SIAM Journal on Optimization*, Volume 20, Issue 3, pg. 1224-1249, 2009.
- R. H. Byrd, F. E. Curtis, and J. Nocedal, "An Inexact SQP Method for Equality Constrained Optimization," *SIAM Journal on Optimization*, Volume 19, Issue 1, pg. 351-369, 2008.