U.S.-Mexico Workshop 2007

Inexact Primal-Dual Methods for Equality Constrained Optimization

Frank Edward Curtis
Northwestern University

with Richard Byrd and Jorge Nocedal

January 9, 2007

Outline

- Description of an Algorithm
  - Step computation
  - Step acceptance
- Global Convergence Analysis
  - Merit function and sufficient decrease
  - Satisfying first-order conditions
- Model Problem (Haber)
  - Problem formulation
  - Numerical Results
- Final Remarks
  - Future work
  - Negative Curvature

Outline

- Description of an Algorithm
  - Step computation
  - Step acceptance
- Global Convergence Analysis
  - Merit function and sufficient decrease
  - Satisfying first-order conditions
- Model Problem (Haber)
  - Problem formulation
  - Numerical Results
- Final Remarks
  - Future work
  - Negative Curvature

Line Search SQP Framework

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

- Define merit function

$$\Phi(x) = f(x) + \pi \|c(x)\|$$

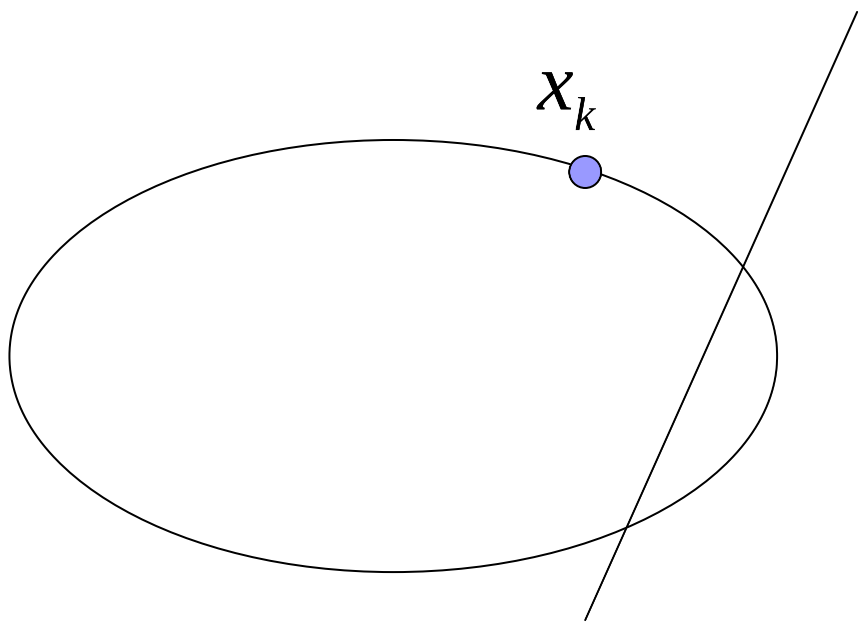
- Implement a line search

$$\Phi(x + \alpha d) \le \Phi(x) + \eta \alpha D_{\Phi}(d)$$

Exact Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \cancelto{0}{\begin{bmatrix} \rho \\ r \end{bmatrix}}$$

$$\begin{array}{ll} \displaystyle\min_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

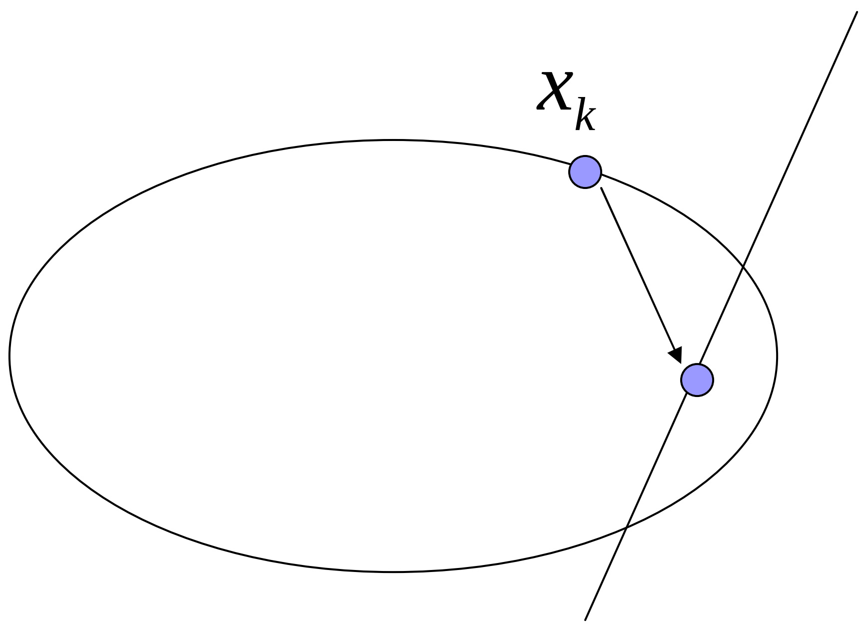


Exact Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \cancelto{0}{\begin{bmatrix} \rho \\ r \end{bmatrix}}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

Exact step minimizes the objective on the linearized constraints



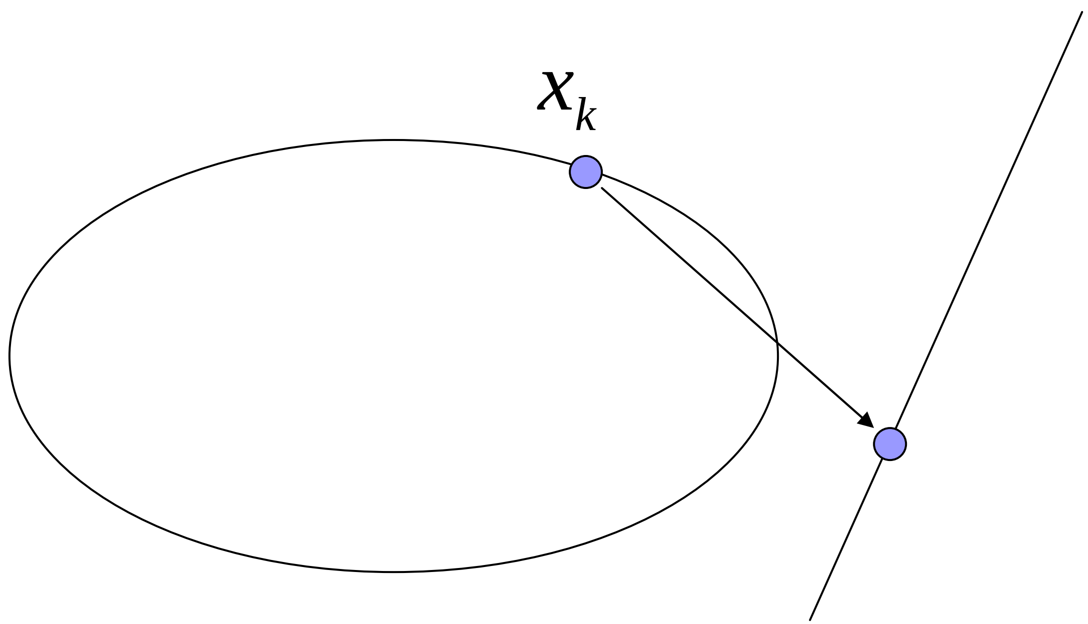
Exact Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \cancelto{0}{\begin{bmatrix} \rho \\ r \end{bmatrix}}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

Exact step minimizes the objective on the linearized constraints

… which may lead to an increase in the model objective



Exact Case

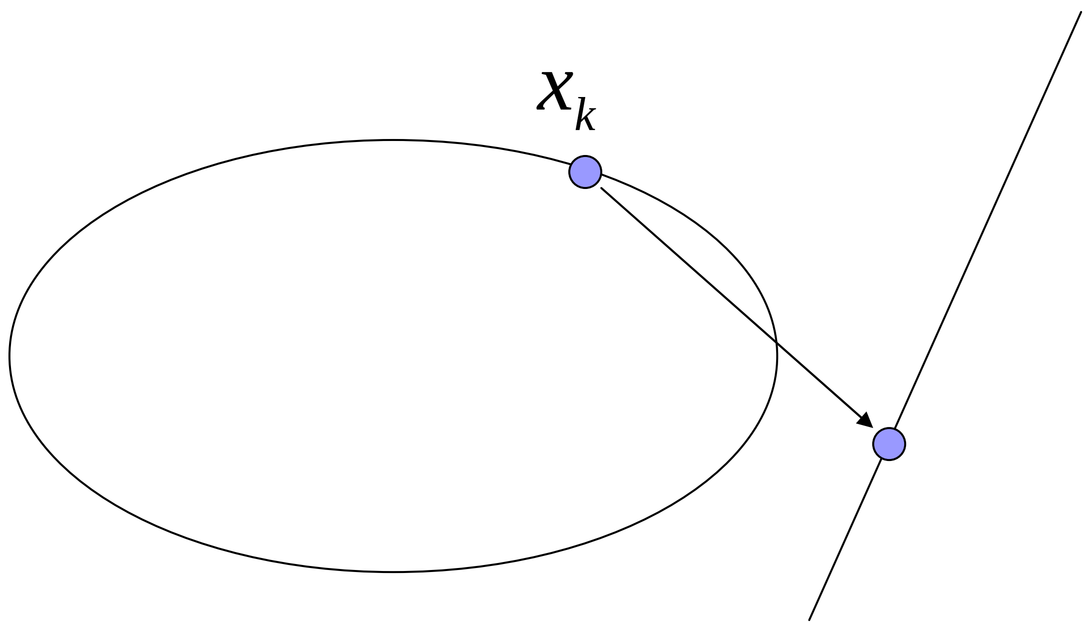
$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \cancelto{0}{\begin{bmatrix} \rho \\ r \end{bmatrix}}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

Exact step minimizes the objective on the linearized constraints

… which may lead to an increase in the model objective

… but this is ok since we can account for this conflict by increasing the penalty parameter

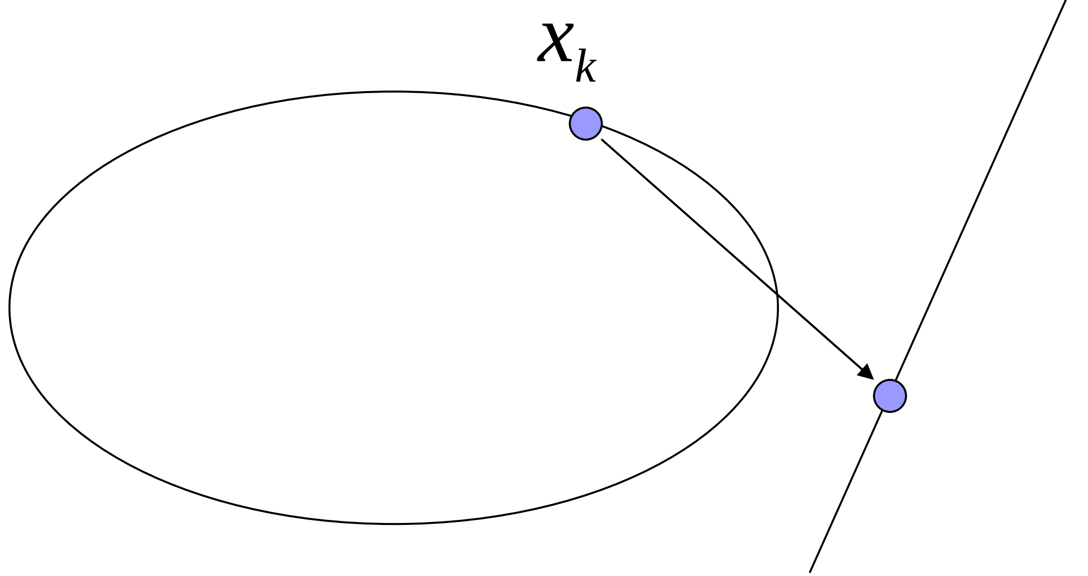


$$\pi \geq \frac{g^T d + \frac{1}{2} d^T W d}{(1-\sigma)\|c\|}, \ 0 < \sigma < 1$$

Exact Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \cancelto{0}{\begin{bmatrix} \rho \\ r \end{bmatrix}}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$



We go directly from solving the quadratic program to obtaining a reduction in the model of the merit function!

That is, either for the most recent penalty parameter or for a higher one, we satisfy the condition:

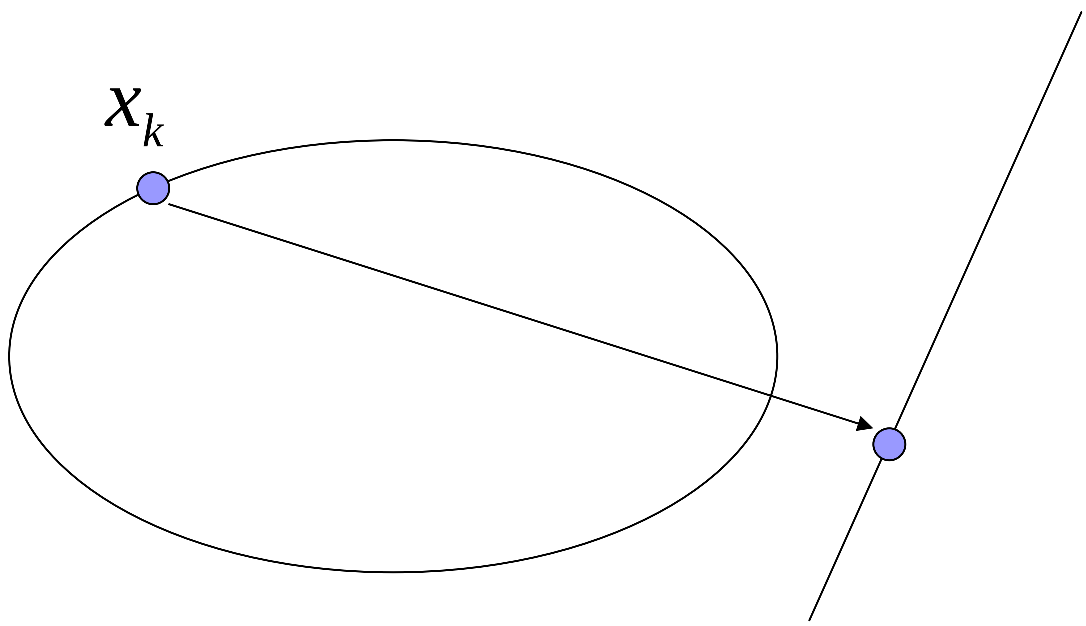
$$-g^T d - \tfrac{1}{2} d^T W d + \pi\|c\| \geq \sigma\pi\|c\|, \; 0 < \sigma < 1$$

Exact Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \cancelto{0}{\begin{bmatrix} \rho \\ r \end{bmatrix}}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

Observe that the quadratic term can significantly influence the penalty parameter



$$\pi_l \equiv \frac{g^T d}{(1-\sigma)\|c\|} < 0$$

$$\pi_q \equiv \frac{g^T d + \frac{1}{2} d^T W d}{(1-\sigma)\|c\|} > 0$$

Inexact Case

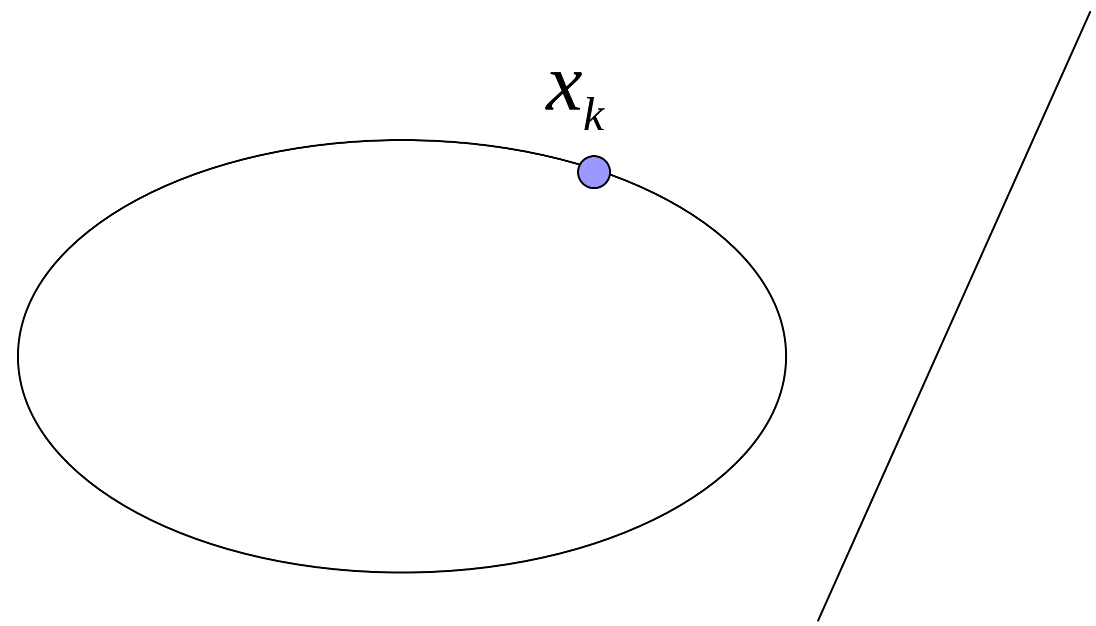
$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

$$\begin{aligned} \min_d \quad & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} \quad & c + Ad = 0 \end{aligned}$$

Inexact Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \frac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$



$$-g^T d - \tfrac{1}{2} d^T W d + \pi\left(\|c\| - \|r\|\right) \geq \sigma\pi\|c\|$$

Inexact Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

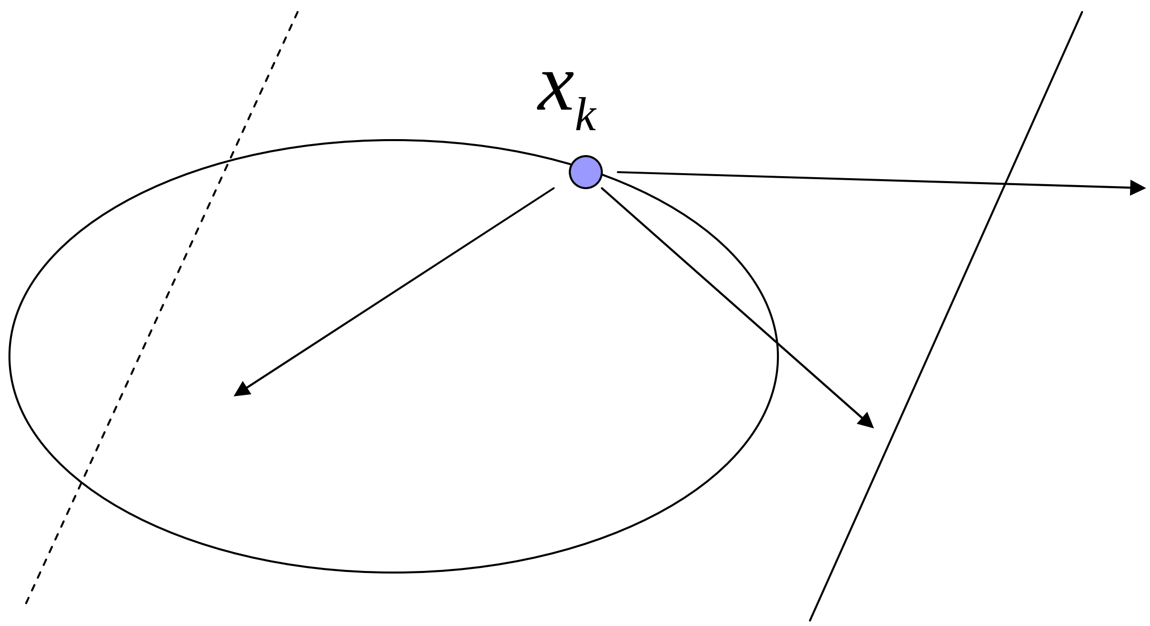
$$\begin{array}{ll} \min_d & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

Step is acceptable if for

$$0 < \varepsilon < 1, \; 0 < \xi :$$

$$\|r\| \leq \xi,$$

$$\|\rho\| \leq \varepsilon \|g + A^T\lambda\|$$



$$-g^T d - \tfrac{1}{2} d^T W d + \pi(\|c\| - \|r\|) \geq \sigma\pi\|c\|$$

Inexact Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T \lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

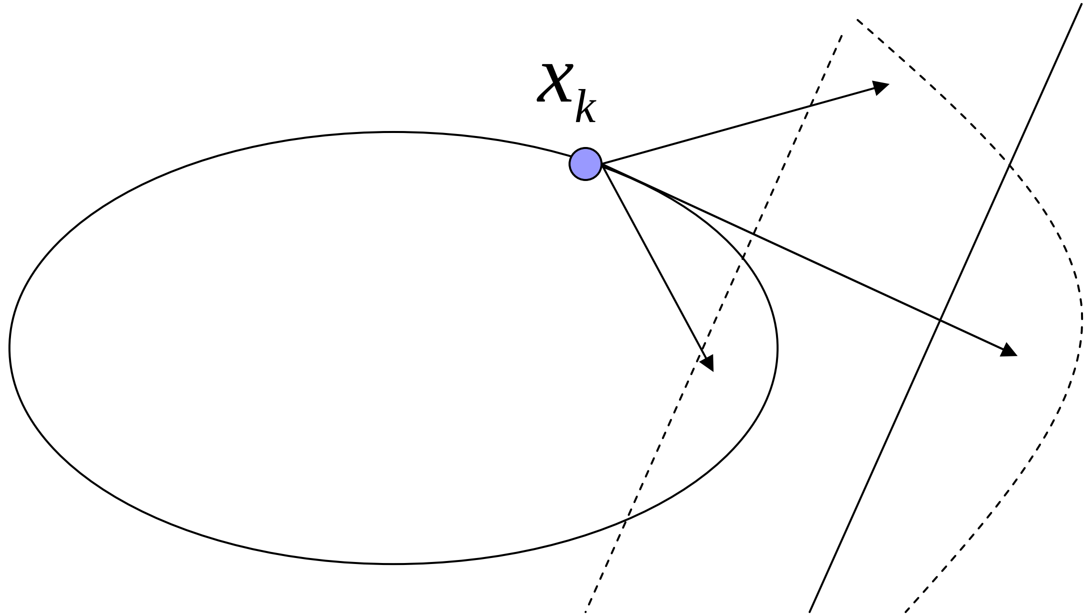
$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

Step is acceptable if for

$$0 < \varepsilon < 1, \; 0 < \beta :$$

$$\|r\| \leq \varepsilon \|c\|,$$

$$\|\rho\| \leq \beta \|c\|$$



$$-g^T d - \tfrac{1}{2} d^T W d + \pi\left(\|c\| - \|r\|\right) \geq \sigma \pi \|c\|$$

Inexact Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

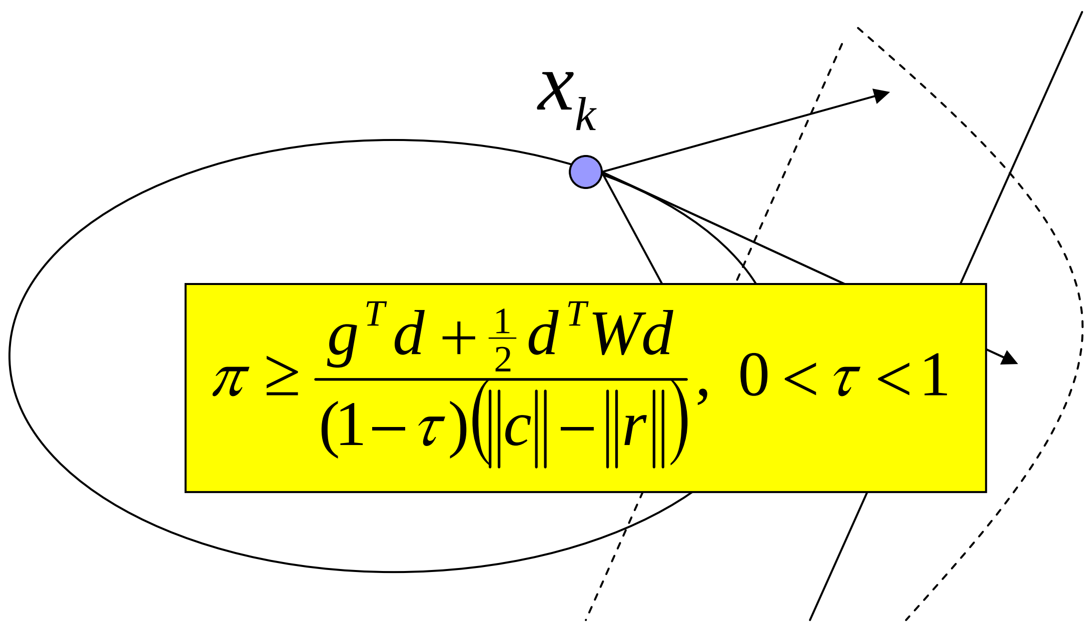
$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

Step is acceptable if for

$$0 < \varepsilon < 1,\ 0 < \beta :$$

$$\|r\| \le \varepsilon \|c\|,$$

$$\|\rho\| \le \beta \|c\|$$



$$\pi \ge \frac{g^T d + \frac{1}{2} d^T W d}{(1-\tau)\left(\|c\| - \|r\|\right)},\ 0 < \tau < 1$$

$$-g^T d - \tfrac{1}{2} d^T W d + \pi\left(\|c\| - \|r\|\right) \ge \sigma\pi\|c\|$$

Algorithm Outline

- for k = 0, 1, 2, …
  - Iteratively solve

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

  - Until

$$\|r\| \le \varepsilon \|c\|,\ 0 < \varepsilon < 1$$
$$\|\rho\| \le \beta \|c\|,\ 0 < \beta$$

or

$$\|r\| \le \xi,\ 0 < \xi$$
$$\|\rho\| \le \varepsilon \|g + A^T\lambda\|,\ 0 < \varepsilon < 1$$
$$mred(d) \ge \sigma\pi\|c\|$$

  - Update penalty parameter
  - Perform backtracking line search
  - Update iterate

Termination Test

- Observe KKT conditions

$$\left\| g + A^T \lambda \right\| \le \max\left\{1, \left\| g \right\|\right\} \varepsilon_{opt} \quad , \; 0 < \varepsilon_{opt} < 1$$

$$\left\| c \right\| \le \max\left\{1, \left\| c(x_0) \right\|\right\} \varepsilon_{feas} \; , \; 0 < \varepsilon_{feas} < 1$$

Outline

- Description of an Algorithm
  - Step computation
  - Step acceptance
- Global Convergence Analysis
  - Merit function and sufficient decrease
  - Satisfying first-order conditions
- Model Problem (Haber)
  - Problem formulation
  - Numerical Results
- Final Remarks
  - Future work
  - Negative Curvature

Assumptions

- The sequence of iterates is contained in a convex set and the following conditions hold:
  - the objective and constraint functions and their first and second derivatives are bounded
  - the multiplier estimates are bounded
  - the constraint Jacobians have full row rank and their smallest singular values are bounded below by a positive constant
  - the Hessian of the Lagrangian is positive definite with smallest eigenvalue bounded below by a positive constant

Sufficient Reduction to Sufficient Decrease

- Taylor expansion of merit function yields

$$\boxed{D_{\Phi}(d) \le g^T d - \pi\left(\|c\| - \|r\|\right)}$$

- Accepted step satisfies

$$\boxed{mred(d) \ge \sigma\pi\|c\|, \; 0 < \sigma < 1}$$

$$\boxed{g^T d - \pi\left(\|c\| - \|r\|\right) \le -\tfrac{1}{2} d^T W d - \sigma\pi\|c\|}$$

$$\boxed{\begin{aligned} D_{\Phi}(d) &\le -\tfrac{1}{2} d^T W d - \sigma\pi\|c\| \\ &\le -\gamma\left(\|d\|^2 + \|c\|\right) \end{aligned}}$$

Intermediate Results

$$\|r\| \le \varepsilon \|c\|,\ 0 < \varepsilon < 1$$

$$\|\rho\| \le \beta \|c\|,\ 0 < \beta$$

$$\pi \ge \frac{g^T d + \frac{1}{2} d^T W d}{(1-\tau)\left(\|c\| - \|r\|\right)},\ 0 < \tau < 1$$

$$\|r\| \le \xi,\ 0 < \xi$$

$$\|\rho\| \le \beta \left\|g + A^T \lambda\right\|,\ 0 < \beta$$

$$mred(d) \ge \sigma\pi\|c\|$$

$\|d\|$ is bounded above

$\{\pi\}$ is bounded above

α is bounded below by a positive constant

Sufficient Decrease in Merit Function

$$D_\Phi(d; x, \pi) \le -\gamma\left(\|d\|^2 + \|c\|\right)$$

$$\Phi(x; \pi) - \Phi(x + \alpha d; \pi) \ge \gamma'\left(\|d\|^2 + \|c\|\right)$$

$$\lim_{k\to\infty}\left(\|d_k\|^2 + \|c_k\|\right) = 0$$

$$\lim_{k\to\infty}\left\|Z_k^T g_k\right\| = 0$$

Step in Dual Space

- We converge to an optimal primal solution, and

$$\boxed{\left\| g^{+} + A^{+^{T}} \lambda^{+} \right\| \leq \varepsilon' \left\| g + A^{T} \lambda \right\|, \ 0 < \varepsilon' < 1}$$

(for sufficiently small $\| c \|$ and $\| d \|$)

Therefore,

$$\boxed{\begin{aligned} &\lim_{k \to \infty} \left\| c_k \right\| = 0 \\ &\lim_{k \to \infty} \left\| g_k + A_k^T \lambda_k \right\| = 0 \end{aligned}}$$

Outline

- Description of an Algorithm
  - Step computation
  - Step acceptance
- Global Convergence Analysis
  - Merit function and sufficient decrease
  - Satisfying first-order conditions
- Model Problem (Haber)
  - Problem formulation
  - Numerical Results
- Final Remarks
  - Future work
  - Negative Curvature

Problem Formulation

- Tikhonov-style regularized inverse problem
  - Want to solve for a reasonably large mesh size $|\Omega|$
  - Want to solve for small regularization parameter γ
- SymQMR for linear system solves
- Input parameters:

$$\left\| \begin{bmatrix} \rho \\ r \end{bmatrix} \right\| \le \kappa \left\| \begin{bmatrix} g + A^T \lambda \\ c \end{bmatrix} \right\|$$

$$\varepsilon \leftarrow 0.1, \; \beta \leftarrow 1, \quad \xi \leftarrow 1, \; \sigma \leftarrow 0.1$$

Recall:

$$\|r\| \le \varepsilon \|c\|, \; 0 < \varepsilon < 1$$
$$\|\rho\| \le \beta \|c\|, \; 0 < \beta$$

or

$$\|r\| \le \xi, \; 0 < \xi$$
$$\|\rho\| \le \varepsilon \|g + A^T \lambda\|, \; 0 < \varepsilon < 1$$
$$mred(d) \ge \sigma\pi\|c\|$$

Numerical Results

n	1024
m	512
γ	1e-6

$$\left\| \begin{bmatrix} \rho \\ r \end{bmatrix} \right\| \le \kappa \left\| \begin{bmatrix} g + A^T \lambda \\ c \end{bmatrix} \right\|$$

κ	Iters.	Time	Total LS Iters.	Avg. LS Iters.	Avg. Rel. Res.
0.5	29	29.5s	1452	50.1	3.12e-1
0.1	12	11.37s	654	54.5	6.90e-2
0.01	9	11.60s	681	75.7	6.27e-3

Numerical Results

n	1024
m	512
γ	1e-6

$$\left\| \begin{bmatrix} \rho \\ r \end{bmatrix} \right\| \le \kappa \left\| \begin{bmatrix} g + A^T \lambda \\ c \end{bmatrix} \right\|$$

κ	Iters.	Time	Total LS Iters.	Avg. LS Iters.	Avg. Rel. Res.
0.5	29	29.5s	1452	50.1	3.12e-1
0.1	12	11.37s	654	54.5	6.90e-2
0.01	9	11.60s	681	75.7	6.27e-3

Numerical Results

n	1024
m	512
κ	1e-1

γ	Iters.	Time	Total LS Iters.	Avg. LS Iters.	Avg. Rel. Res.
1e-6	12	11.40s	654	54.5	6.90e-2
1e-7	11	14.52s	840	76.4	6.99e-2
1e-8	8	10.57s	639	79.9	6.15e-2
1e-9	11	18.52s	1139	104	8.65e-2
1e-10	19	44.41s	2708	143	8.90e-2

Numerical Results

n	8192
m	4096
κ	1e-1

γ	Iters.	Time	Total LS Iters.	Avg. LS Iters.	Avg. Rel. Res.
1e-6	15	264.47s	1992	133	8.13e-2
1e-7	11	236.51s	1776	161	6.89e-2
1e-8	9	204.51s	1567	174	6.77e-2
1e-9	11	347.66s	2681	244	8.29e-2
1e-10	16	805.14s	6249	391	8.93e-2

Numerical Results

n	65536
m	32768
κ	1e-1

γ	Iters.	Time	Total LS Iters.	Avg. LS Iters.	Avg. Rel. Res.
1e-6	15	5055.9s	4365	291	8.46e-2
1e-7	10	4202.6s	3630	363	8.87e-2
1e-8	12	5686.2s	4825	402	7.96e-2
1e-9	12	6678.7s	5633	469	8.77e-2
1e-10	14	14783s	12525	895	8.63e-2

Outline

- Description of an Algorithm
  - Step computation
  - Step acceptance
- Global Convergence Analysis
  - Merit function and sufficient decrease
  - Satisfying first-order conditions
- Model Problem (Haber)
  - Problem formulation
  - Numerical Results
- Final Remarks
  - Future work
  - Negative Curvature

Review and Future Challenges

- Review
  - Defined a globally convergent inexact SQP algorithm
  - Require only inexact solutions of primal-dual system
  - Require only matrix-vector products involving objective and constraint function derivatives
  - Results also apply when only reduced Hessian of Lagrangian is assumed to be positive definite
  - Numerical experience on model problem is promising
- Future challenges
  - (Nearly) Singular constraint Jacobians
  - Inexact derivative information
  - Negative curvature
  - etc., etc., etc….

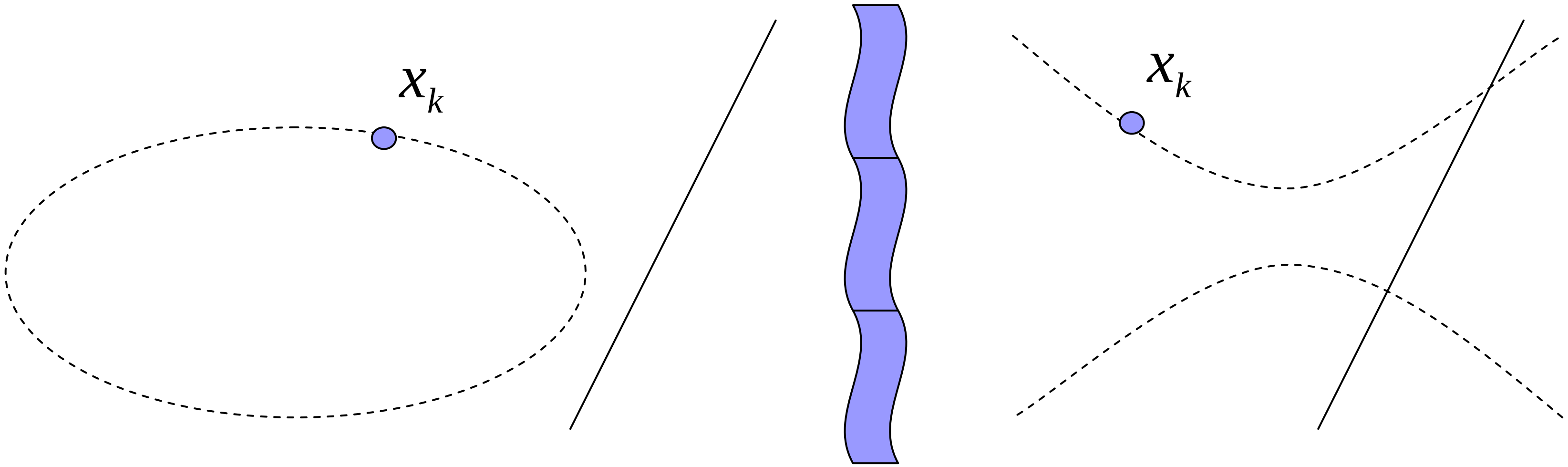
Negative Curvature

- Big question
  - What is the best way to handle negative curvature (i.e., when the reduced Hessian may be indefinite)?
- Small question
  - What is the best way to handle negative curvature in the context of our inexact SQP algorithm?
  - We have no inertia information!
- Smaller question
  - When can we handle negative curvature in the context of our inexact SQP algorithm with NO algorithmic modifications?
  - When do we know that a given step is OK?
  - Our analysis of the inexact case leads to a few observations…

Why Quadratic Models?

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

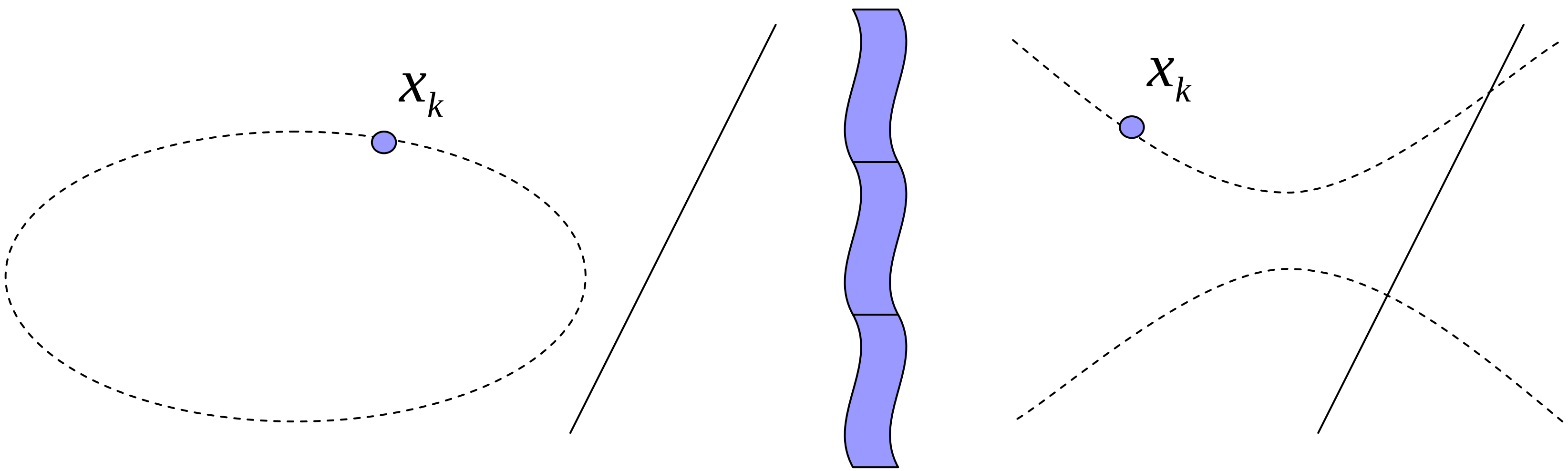
$$\begin{array}{ll} \min\limits_{d} & g^T d + \frac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$



Why Quadratic Models?

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$



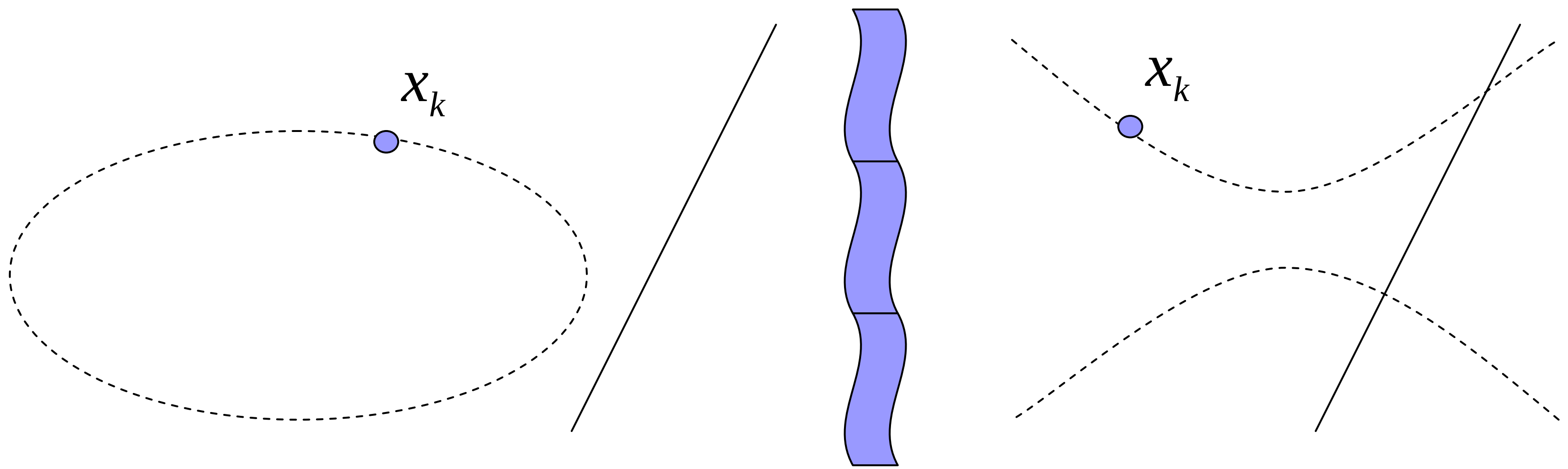
Provides a good…
- direction? Yes
- step length? Yes

Provides a good…
- direction? Maybe
- step length? Maybe

Why Quadratic Models?

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \frac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$



- One can use our stopping criteria as a mechanism for determining which are good directions
- All that needs to be determined is whether the step lengths are acceptable

Unconstrained Optimization

$$\boxed{Hd = -g + \rho} \qquad \boxed{\min_d \; g^T d + \tfrac{1}{2} d^T H d}$$

- Direct method is the angle test

$$\boxed{-g^T d \geq \theta \|g\| \|d\|}$$

- Indirect method is to check the conditions

$$\boxed{\|\rho\| \leq \varepsilon \|g\|, \; d^T H d \geq \mu \|d\|^2}$$

or

$$\boxed{-g^T d \geq \theta \|g\|^2, \; \|d\| \leq \omega \|g\|}$$

Unconstrained Optimization

$$\boxed{Hd = -g + \rho} \qquad \boxed{\min_d \; g^T d + \tfrac{1}{2} d^T H d}$$

- Direct method is the angle test

$$\boxed{-g^T d \geq \theta \|g\| \|d\|}$$

- Indirect method is to check the conditions

$$\text{step quality} \left\{ \begin{array}{c} \boxed{\|\rho\| \leq \varepsilon \|g\|, \; d^T H d \geq \mu \|d\|^2} \\ \text{or} \\ \boxed{-g^T d \geq \theta \|g\|^2, \; \|d\| \leq \omega \|g\|} \end{array} \right\} \text{step length}$$

Constrained Optimization

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

- Step quality determined by

$$\begin{array}{l} \|r\| \le \varepsilon \|c\|, \; 0 < \varepsilon < 1 \\ \|\rho\| \le \beta \|c\|, \; 0 < \beta \end{array}$$

or

$$\begin{array}{l} \|r\| \le \xi, \; 0 < \xi \\ \|\rho\| \le \varepsilon \|g + A^T\lambda\|, \; 0 < \varepsilon < 1 \\ mred(d) \ge \sigma\pi\|c\| \end{array}$$

- Step length determined by

$$d^T W d \ge \mu \|d\|^2$$

or

$$\|d\| \le \omega \max\{\|c\|, \|r\|\}$$

Thanks!

Actual Stopping Criteria

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

- Stopping conditions: $0 < \varepsilon, \sigma < 1,\ 0 < \beta, \xi$

$$\begin{aligned} \|r\| &\le \varepsilon \|c\| \\ \|\rho\| &\le \beta \|c\| \end{aligned}$$

or

$$\begin{aligned} \|r\| &\le \xi \max\{\|c\|, 1\} \\ \|\rho\| &\le \max\left\{\beta\|c\|, \varepsilon\left\|g + A^T\lambda\right\|\right\} \\ mred(d) &\ge \sigma\pi \max\{\|c\|, \|r\| - \|c\|\} \end{aligned}$$

- Model reduction condition

$$-g^T d - \tfrac{\omega}{2} d^T W d + \pi\left(\|c\| - \|r\|\right) \ge \sigma\pi \max\{\|c\|, \|r\| - \|c\|\}$$

Constraint Feasible Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

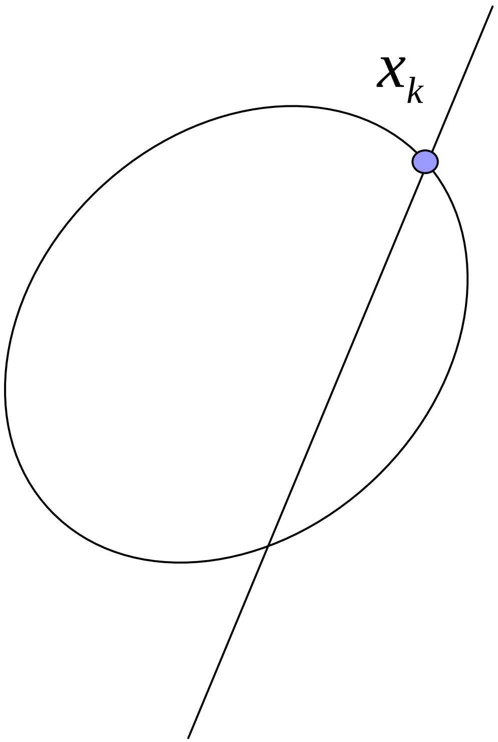
$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

- If feasible, conditions reduce to

$$\|r\| \le \xi$$

$$\|\rho\| \le \varepsilon \left\| g + A^T \lambda \right\|$$

$$(1+\sigma)\pi\|r\| \le -g^T d - \tfrac{\omega}{2} d^T W d$$



Constraint Feasible Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

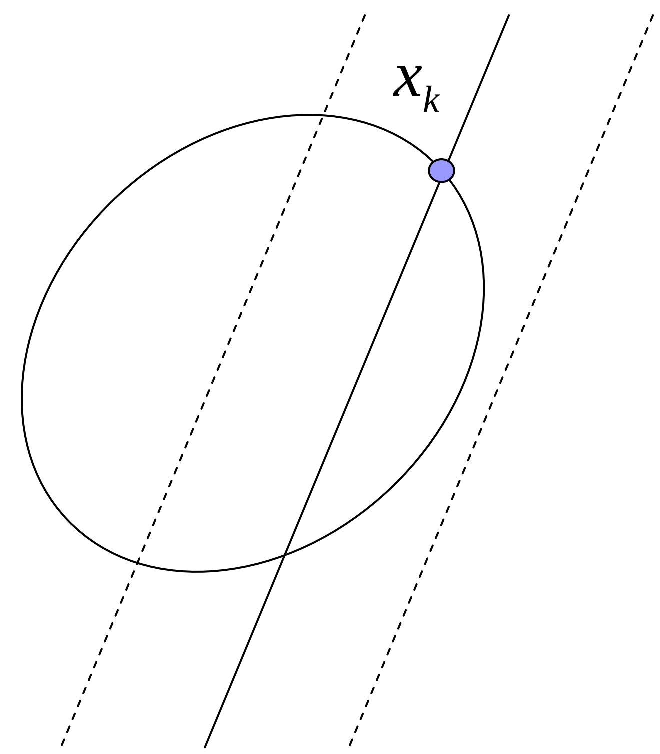
$$\begin{array}{ll} \min\limits_{d} & g^T d + \frac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

- If feasible, conditions reduce to

$$\|r\| \le \xi$$

$$\|\rho\| \le \varepsilon \|g + A^T\lambda\|$$

$$(1+\sigma)\pi\|r\| \le -g^T d - \frac{\omega}{2} d^T W d$$



Constraint Feasible Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

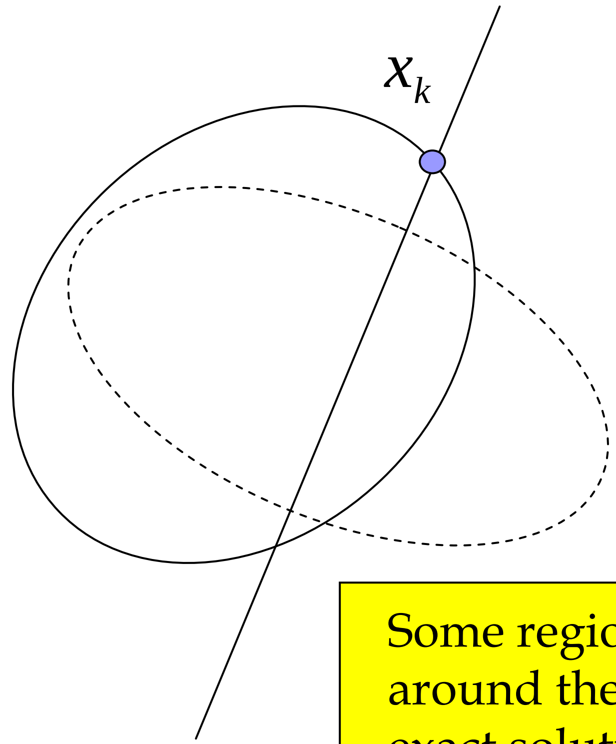
$$\begin{array}{ll} \min\limits_{d} & g^T d + \frac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

- If feasible, conditions reduce to

$$\|r\| \leq \xi$$

$$\|\rho\| \leq \varepsilon \|g + A^T\lambda\|$$

$$(1+\sigma)\pi\|r\| \leq -g^T d - \frac{\omega}{2} d^T W d$$



Some region around the exact solution

Constraint Feasible Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

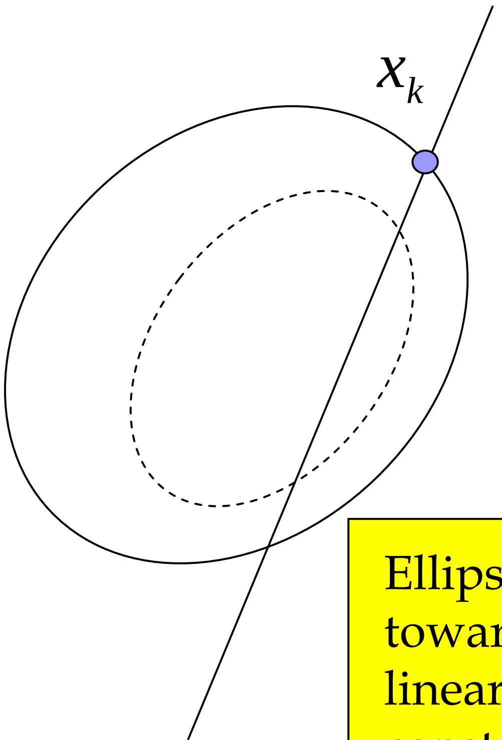
$$\begin{array}{ll} \min\limits_{d} & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} & c + Ad = 0 \end{array}$$

- If feasible, conditions reduce to

$$\|r\| \le \xi$$

$$\|\rho\| \le \varepsilon \|g + A^T\lambda\|$$

$$(1+\sigma)\pi\|r\| \le -g^T d - \tfrac{\omega}{2} d^T W d$$



Ellipse distorted toward the linearized constraints

Constraint Feasible Case

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ \delta \end{bmatrix} = -\begin{bmatrix} g + A^T\lambda \\ c \end{bmatrix} + \begin{bmatrix} \rho \\ r \end{bmatrix}$$

$$\begin{aligned} \min_d \quad & g^T d + \tfrac{1}{2} d^T W d \\ \text{s.t.} \quad & c + Ad = 0 \end{aligned}$$

- If feasible, conditions reduce to

$$\|r\| \le \xi$$

$$\|\rho\| \le \varepsilon \|g + A^T\lambda\|$$

$$(1+\sigma)\pi\|r\| \le -g^T d - \tfrac{\omega}{2} d^T W d$$