

## 9. Analytic center cutting-plane method

- analytic center cutting-plane method
- computing the analytic center
- pruning constraints
- lower bound and stopping criterion

9-1

### Analytic center and ACCPM

**analytic center** of a set of inequalities  $Ax \preceq b$

$$x_{\text{ac}} = \underset{z}{\operatorname{argmin}} - \sum_{i=1}^m \log(b_i - a_i^T z)$$

### analytic center cutting-plane method (ACCPM)

localization method that

- represents  $\mathcal{P}_k$  by set of inequalities  $A^{(k)}x \preceq b^{(k)}$
- selects analytic center of  $A^{(k)}x \preceq b^{(k)}$  as query point  $x^{(k+1)}$

## ACCPM algorithm outline

**given** an initial polyhedron  $\mathcal{P}_0 = \{x \mid A^{(0)}x \preceq b^{(0)}\}$  known to contain  $X$

**repeat** for  $k = 1, 2, \dots$

1. compute  $x^{(k)}$ , the analytic center of  $A^{(k-1)}x \preceq b^{(k-1)}$
2. query cutting-plane oracle at  $x^{(k)}$
3. if  $x^{(k)} \in X$ , quit; otherwise, add returned cutting plane  $a^T z \leq b$ :

$$A^{(k)} = \begin{bmatrix} A^{(k-1)} \\ a^T \end{bmatrix}, \quad b^{(k)} = \begin{bmatrix} b^{(k-1)} \\ b \end{bmatrix}$$

if  $\mathcal{P}_k = \{x \mid A^{(k)}x \preceq b^{(k)}\} = \emptyset$ , quit

## Constructing cutting-planes

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

$f_0, \dots, f_m : \mathbf{R}^n \rightarrow \mathbf{R}$  convex

**cutting-plane** for optimal set  $X$

- if  $x^{(k)}$  is not feasible, say  $f_j(x^{(k)}) > 0$ , we have (deep) *feasibility cut*

$$f_j(x^{(k)}) + g_j^T(z - x^{(k)}) \leq 0, \quad g_j \in \partial f_j(x^{(k)})$$

- if  $x^{(k)}$  is feasible, we have (deep) *objective cut*

$$g_0^T(z - x^{(k)}) + f_0(x^{(k)}) - f_{\text{best}}^{(k)} \leq 0, \quad g_0 \in \partial f_0(x^{(k)})$$

where  $f_{\text{best}}^{(k)} = \min\{f_0(x^{(i)}) \mid i \leq k, x^{(i)} \text{ feasible}\}$

## Computing the analytic center

$$\text{minimize } \phi(x) = - \sum_{i=1}^m \log(b_i - a_i^T x)$$

$$\text{dom } \phi = \{x \mid a_i^T x < b_i, \ i = 1, \dots, m\}$$

**challenge:** we are not given a point in  $\text{dom } \phi$

### some options

- use phase I to find  $x \in \text{dom } \phi$ , followed by standard Newton method
- standard Newton method applied to dual problem
- infeasible start Newton method (EE236B lecture 11, BV §10.3)

## Dual Newton method

### dual analytic centering problem

$$\begin{aligned} \text{maximize } & g(z) = \sum_{i=1}^m \log z_i - b^T z + m \\ \text{subject to } & A^T z = 0 \end{aligned}$$

### optimality conditions

$x, z$  are primal and dual optimal if

$$b_i - a_i^T x = 1/z_i, \quad A^T z = 0, \quad z \succ 0, \quad Ax \prec b$$

## initialization of Newton's method

dual method is interesting when a strictly feasible  $z$  is easy to find, *e.g.*,

$$A^T = [ B^T \quad I \quad -I ]$$

- can take  $z = (u, v, w) \succ 0$ , with  $u$  arbitrary;  $v, w$  s.t.  $w - v = B^T u$
- (primal) analytic centering problem includes variable bounds  $l \preceq x \preceq u$

## Newton equation

$$\begin{bmatrix} -\mathbf{diag}(z)^{-2} & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} \Delta z \\ w \end{bmatrix} = \begin{bmatrix} b - \mathbf{diag}(z)^{-1}\mathbf{1} \\ 0 \end{bmatrix}$$

can be solved by elimination of  $\Delta z$ : solve

$$(A^T \mathbf{diag}(z)^2 A) w = A^T (\mathbf{diag}(z)^2 b - z)$$

and take  $\Delta z = z - \mathbf{diag}(z)^2 (b - Aw)$

## stopping criterion for Newton's method

Newton decrement at  $z$  is

$$\lambda(z) = (\Delta z^T \nabla g(z))^{1/2} = \|\mathbf{diag}(z)^{-1} \Delta z\|_2$$

- $\lambda(z) = 0$  implies  $w$  is the analytic center:

$$b - Aw = \mathbf{diag}(z)^{-1} \mathbf{1}$$

- $\lambda(z) < 1$  implies  $x = w$  is primal feasible

$$b - Aw = \mathbf{diag}(z)^{-1} (\mathbf{1} - \mathbf{diag}(z)^{-1} \Delta z) \succ 0$$

terminating with small  $\lambda(z)$  gives strictly feasible, approximate center

## Infeasible start Newton method

$$\begin{aligned} & \text{minimize} && -\sum_{i=1}^m \log y_i \\ & \text{subject to} && y = b - Ax \end{aligned}$$

variables  $x$  and  $y$

### optimality conditions

$$y \succ 0, \quad z \succ 0, \quad r(x, y, z) = \begin{bmatrix} r_p(x, y, z) \\ r_d(x, y, z) \end{bmatrix} = 0$$

where

$$r_p(x, y, z) = y + Ax - b, \quad r_d(x, y, z) = \begin{bmatrix} A^T z \\ z - \mathbf{diag}(y)^{-1} \mathbf{1} \end{bmatrix}$$

**initialization:** can start from *any*  $x$  and *any*  $y \succ 0$

example: take previous analytic center as  $x$ , and choose  $y$  as

$$y_i = \begin{cases} b_i - a_i^T x & b_i - a_i^T x > 0 \\ 1 & \text{otherwise} \end{cases}$$

**Newton equation** at  $(x, y, z)$ :

$$Dr(x, y, z) \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = -r(x, y, z)$$

$$\begin{bmatrix} A & I & 0 \\ 0 & 0 & A^T \\ 0 & \mathbf{diag}(y)^{-2} & I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = - \begin{bmatrix} y + Ax - b \\ A^T z \\ z - \mathbf{diag}(y)^{-1} \mathbf{1} \end{bmatrix}$$

## solving the Newton equations

can be solved by block elimination: solve

$$(A^T \mathbf{diag}(y)^{-2} A) \Delta x = A^T \mathbf{diag}(y)^{-2} (b - Ax - 2y)$$

and take

$$\Delta y = b - y - Ax - A\Delta x, \quad \Delta z = \mathbf{diag}(y)^{-2} (y - \Delta y) - z$$

- form  $A^T \mathbf{diag}(y)^{-2} A$ , then use dense or sparse Cholesky factorization
- solve (diagonally scaled) least-squares problem

$$\Delta x = \underset{z}{\operatorname{argmin}} \left\| \mathbf{diag}(y)^{-1} (Az + \mathbf{diag}(y)^{-1} (b - Ax) - \mathbf{21}) \right\|_2$$

- solve (approximately) using iterative method (*e.g.*, conjugate gradients)

## Pruning constraints

### enclosing ellipsoid at analytic center

if  $x_{ac}$  is the analytic center of  $a_i^T x \leq b_i$ ,  $i = 1, \dots, m$ , then the ellipsoid

$$\mathcal{E} = \{z \mid (z - x_{ac})^T \nabla^2 \phi(x_{ac}) (z - x_{ac}) \leq m^2\}$$

contains  $\mathcal{P} = \{z \mid a_i^T x \leq b_i, i = 1, \dots, m\}$

- proof in BV page 420
- from expression for Hessian,

$$\mathcal{E} = \left\{ z \mid \sum_{i=1}^m \left( \frac{a_i^T (z - x_{ac})}{b_i - a_i^T x_{ac}} \right)^2 \leq m^2 \right\}$$

**(ir)relevance measure** for constraint  $a_i^T x \leq b_i$

$$\eta_i = \frac{b_i - a_i^T x_{ac}}{(a_i^T \nabla^2 \phi(x_{ac})^{-1} a_i)^{1/2}}$$

if  $\eta_i \geq m$ , then constraint  $a_i^T x \leq b_i$  is redundant

proof: the optimal value of

$$\begin{aligned} & \text{maximize} && a_i^T z \\ & \text{subject to} && (z - x_{ac})^T H (z - x_{ac}) \leq m^2 \end{aligned}$$

(with  $H = \nabla^2 \phi(x_{ac})$ ) is

$$m \sqrt{a_i^T H^{-1} a_i} + a_i^T x_{ac}$$

the constraint is redundant if this is less than  $b_i$

### common ACCPM constraint dropping schemes

- drop all constraints with  $\eta_i \geq m$  (guaranteed to not change  $\mathcal{P}$ )
- drop constraints in order of irrelevance, keeping constant number, usually  $3n - 5n$

## Lower bound in ACCPM

suppose we apply ACCPM to a convex problem

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && Gx \preceq h \end{aligned} \tag{1}$$

the inequalities  $A^{(k-1)}x \preceq b^{(k-1)}$  at iteration  $k$  can be divided in two sets

- $A_f x \preceq b_f$  includes the constraints  $Gx \preceq h$  plus the feasibility cuts
- $A_o x \preceq b_o + c_o$  includes the objective cuts

$$f_0(x^{(i)}) + g_0^{(i)T}(x - x^{(i)}) \leq f_{\text{best}}^{(i)},$$

with  $g_0^{(i)T}x^{(i)} - f_0(x^{(i)})$  stored in the vector  $b_o$  and  $f_{\text{best}}^{(i)}$  in  $c_o$

**piecewise-linear relaxation:** the problem

$$\begin{aligned} & \text{minimize} && \max(A_o x - b_o) \\ & \text{subject to} && A_f x \preceq b_f \end{aligned}$$

is a relaxation of the problem (1) ( $\max(y)$  for vector  $y$  means  $\max_i y_i$ )

- $f_0(x) \geq \max(A_o x - b_o)$  for all  $x$  (by convexity)
- optimal set is contained in the polyhedron  $A_f x \preceq b_f$  (by construction)

**dual of PWL relaxation**

$$\begin{aligned} & \text{maximize} && -b_o^T u - b_f^T v \\ & \text{subject to} && A_o^T u + A_f^T v = 0 \\ & && \mathbf{1}^T u = 1 \\ & && u \succeq 0, \quad v \succeq 0 \end{aligned}$$

dual feasible points give lower bounds on optimal value of (1)

## dual feasible point from analytic centering

$x^{(k)}$  is the analytic center of  $A_o x \preceq b_o + c_o$ ,  $A_f x \preceq b_f$ ; hence

$$A_o^T z_o + A_f^T z_f = 0,$$

where

$$z_o = \mathbf{diag}(b_o + c_o - A_o x^{(k)})^{-1} \mathbf{1}, \quad z_f = \mathbf{diag}(b_f - A_f x^{(k)})^{-1} \mathbf{1}$$

- normalizing gives a dual feasible point for the PWL relaxation:

$$u = \frac{1}{\mathbf{1}^T z_o} z_o, \quad v = \frac{1}{\mathbf{1}^T z_o} z_f$$

- $l^{(k)} = -b_o^T u - b_f^T v$  is a lower bound on optimal value of (1)

from  $x^{(k)}$  we get a readily computed lower bound

## Stopping criterion

keep track of best point found, and best lower bound

- best function value so far

$$f_{\text{best}}^{(k)} = \min_{\substack{i=1, \dots, k \\ x^{(i)} \text{ feasible}}} f_0(x^{(i)})$$

- best lower bound so far

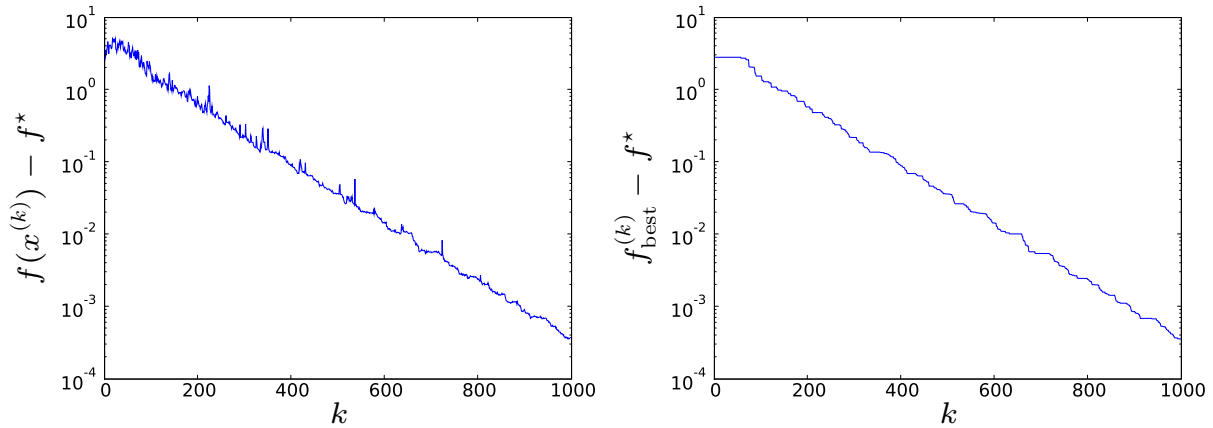
$$l_{\text{best}}^{(k)} = \max_{i=1, \dots, k} l^{(i)}$$

can stop when  $f_{\text{best}}^{(k)} - l_{\text{best}}^{(k)} \leq \epsilon$  to guarantee  $\epsilon$ -suboptimality

## Example: piecewise linear minimization

$$\begin{aligned} & \text{minimize} && \max_{i=1,\dots,m} (a_i^T x + b_i) \\ & \text{subject to} && -\mathbf{1} \preceq x \preceq \mathbf{1} \end{aligned}$$

$n = 100$  variables,  $m = 200$  terms,  $f^* \approx 0.36$

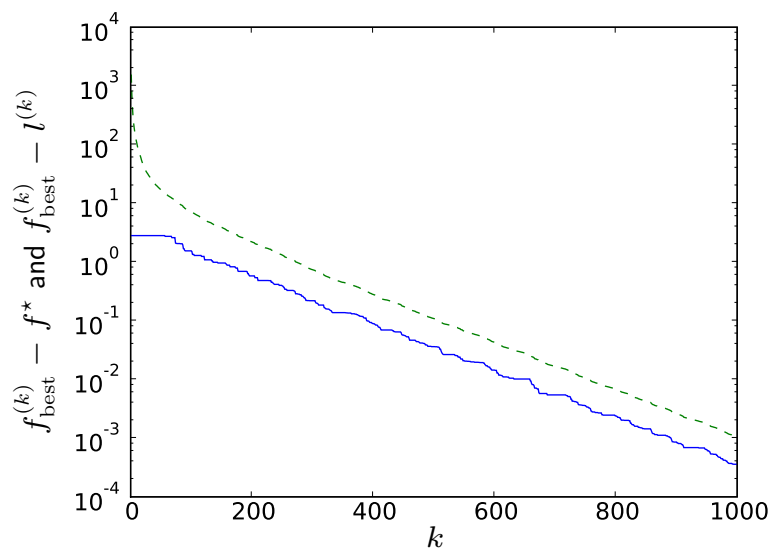


Analytic center cutting-plane method

9-19

### computed lower bound on optimal value

$f_{\text{best}}^{(k)} - l^{(k)}$  (dashed line) and  $f_{\text{best}}^{(k)} - f^*$  (solid line)

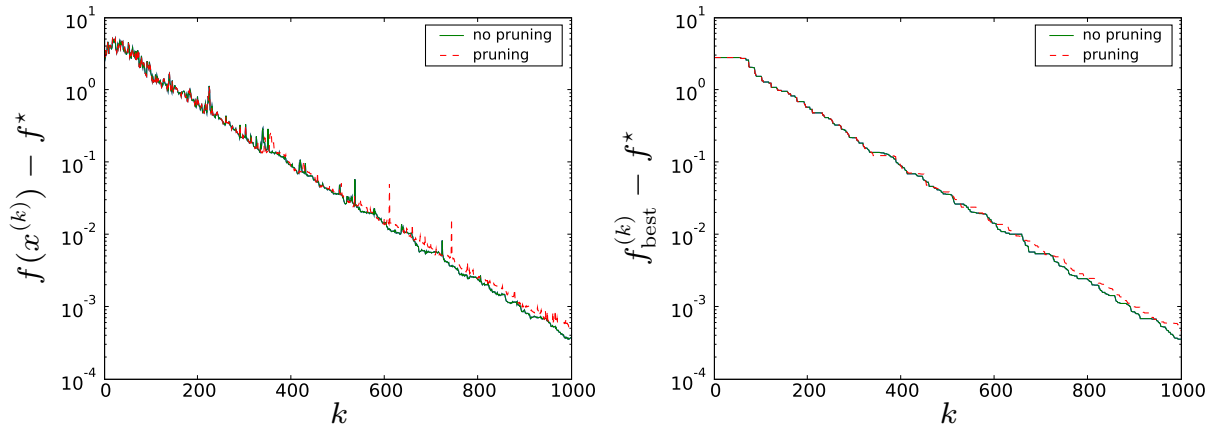


Analytic center cutting-plane method

9-20

# ACCPM with constraint dropping

same problem; convergence with and without pruning (to  $3n$  constraints)



## References

- Y. Ye, *Interior-Point Algorithms. Theory and Analysis* (1997)  
§8.2.3 gives a convergence proof of ACCPM (the bound on the number of iterations is  $n^2$  times a function of  $R/r$ )
- S. Boyd, course notes for EE364b, Convex Optimization II