Analytic center cutting-plane method

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

Analytic center and ACCPM

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

$$x_{\rm ac} = \mathop{\rm argmin}_z - \sum_{i=1}^m \log(b_i - a_i^T z)$$

analytic center cutting-plane method (ACCPM)

localization method that

- represents $\mathcal{P}_k$ by set of inequalities $A^{(k)}$, $b^{(k)}$
- 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 C

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

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 $x^{(k)} \in C$, quit; otherwise, add returned cutting plane $a^T z \leq b$:

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

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

Constructing cutting planes

cutting planes for optimal set C of convex problem

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

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

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

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

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

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

Computing the analytic center

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

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

challenge: we are not given a point in $\mathbf{dom}\,\phi$

some options

- use phase I to find $x \in \mathbf{dom}\,\phi$, 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

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

optimality conditions

x, z are primal and dual optimal if

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

Initialization of dual Newton method

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

$$A = \begin{bmatrix} I \\ -I \\ B \end{bmatrix}$$

- dual feasibility requires

$$A^T z = z_1 - z_2 + B^T z_3 = 0, \qquad z = (z_1, z_2, z_3) \succeq 0$$

(for example, can pick any $z_3 \succ 0$ and find corresponding z_1, z_2)

- this corresponds to variable bounds in (primal) centering problem, *e.g.*,

$$\mathcal{P}_0 = \{x \mid l \preceq x \preceq u\}$$

Dual Newton equation

analytic centering problem

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^{m} \log z_i + b^T z \\ \text{subject to} & A^T z = 0 \end{array}$$

Newton equation

$$\left[\begin{array}{cc} -\mathbf{diag}(z)^{-2} & A \\ A^T & 0 \end{array}\right] \left[\begin{array}{c} \Delta z \\ w \end{array}\right] = \left[\begin{array}{c} b - \mathbf{diag}(z)^{-1}\mathbf{1} \\ 0 \end{array}\right]$$

can be solved by elimination of Δz: solve

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

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

Stopping criterion for dual Newton method

Newton decrement at z is

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

- $\lambda(z) = 0$ 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

reformulated analytic centering problem (variables x and y)

$$\text{minimize} \quad -\sum_{i=1}^{m} \log y_i, \qquad \text{subject to} \quad y = b - Ax$$

optimality conditions

$$y \succ 0, \qquad z \succ 0, \qquad r(x, y, z) = \begin{bmatrix} y + Ax - b \\ A^T z \\ z - \mathbf{diag}(y)^{-1}\mathbf{1} \end{bmatrix} = 0$$

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

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

$$y_i = b_i - a_i^T x \quad \text{if } b_i - a_i^T x > 0, \qquad y_i = 1 \quad \text{otherwise}$$

Newton equation for infeasible Newton method

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

can be solved by block elimination of Δy, Δz: solve

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

and take

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

Pruning constraints

enclosing ellipsoid at analytic center

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

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

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

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

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

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

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

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

proof: the optimal value of

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

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

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

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

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

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

- $A_{\rm f}x \preceq b_{\rm f}$ includes the constraints $Gx \preceq h$ plus the feasibility cuts
- $A_{\rm o}x \preceq b_{\rm o} + c_{\rm o}$ includes the objective cuts

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

with $g_0^{(i)T}x^{(i)} - f_0(x^{(i)})$ stored in the vector $b_{\rm o}$ and $f_{\rm best}^{(i)}$ in $c_{\rm o}$

piecewise-linear relaxation: the problem

$$\begin{array}{ll} \text{minimize} & \max(A_\text{o}x - b_\text{o}) \\ \text{subject to} & A_\text{f}x \preceq b_\text{f} \end{array}$$

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

- $f_0(x) \geq \max(A_\text{o}x - b_\text{o})$ for all x (by convexity)
- optimal set is contained in the polyhedron $A_\text{f}x \preceq b_\text{f}$ (by construction)

dual of PWL relaxation

$$\begin{array}{ll} \text{maximize} & -b_\text{o}^T u - b_\text{f}^T v \\ \text{subject to} & A_\text{o}^T u + A_\text{f}^T v = 0 \\ & \mathbf{1}^T u = 1 \\ & u \succeq 0, \quad v \succeq 0 \end{array}$$

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

dual feasible point from analytic centering

$x^{(k)}$ is the analytic center of $A_\mathrm{o}x \preceq b_\mathrm{o} + c_\mathrm{o}$, $A_\mathrm{f}x \preceq b_\mathrm{f}$; hence

$$A_\mathrm{o}^T z_\mathrm{o} + A_\mathrm{f}^T z_\mathrm{f} = 0,$$

where

$$z_\mathrm{o} = \mathbf{diag}(b_\mathrm{o} + c_\mathrm{o} - A_\mathrm{o}x^{(k)})^{-1}\mathbf{1}, \qquad z_\mathrm{f} = \mathbf{diag}(b_\mathrm{f} - A_\mathrm{f}x^{(k)})^{-1}\mathbf{1}$$

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

$$u = \frac{1}{\mathbf{1}^T z_\mathrm{o}} z_\mathrm{o}, \qquad v = \frac{1}{\mathbf{1}^T z_\mathrm{o}} z_\mathrm{f}$$

- $l^{(k)} = -b_\mathrm{o}^T u - b_\mathrm{f}^T v$ 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

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

- best lower bound so far

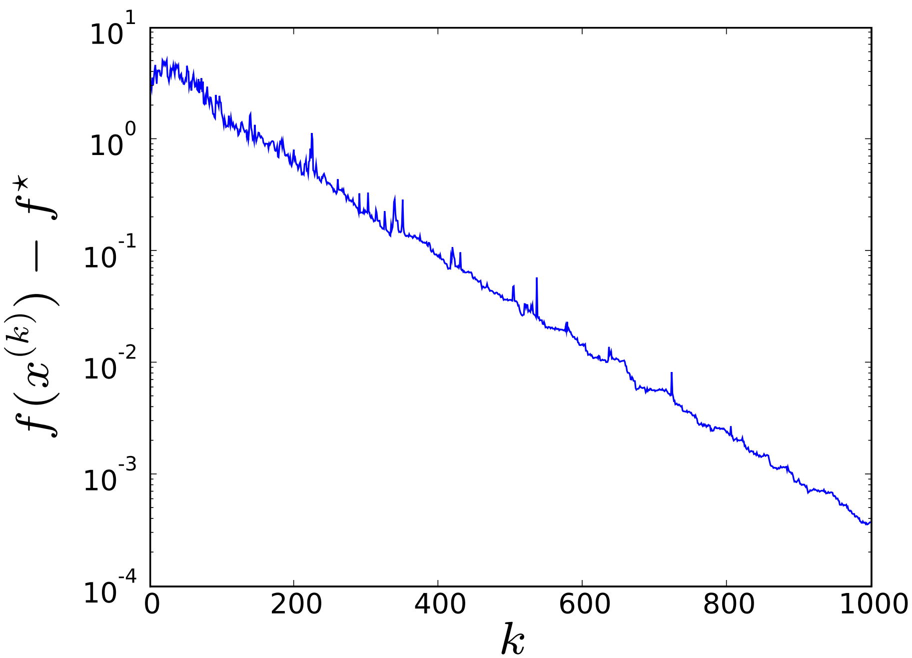
$$l_{\rm best}^{(k)} = \max_{i=1,\ldots,k} l^{(i)}$$

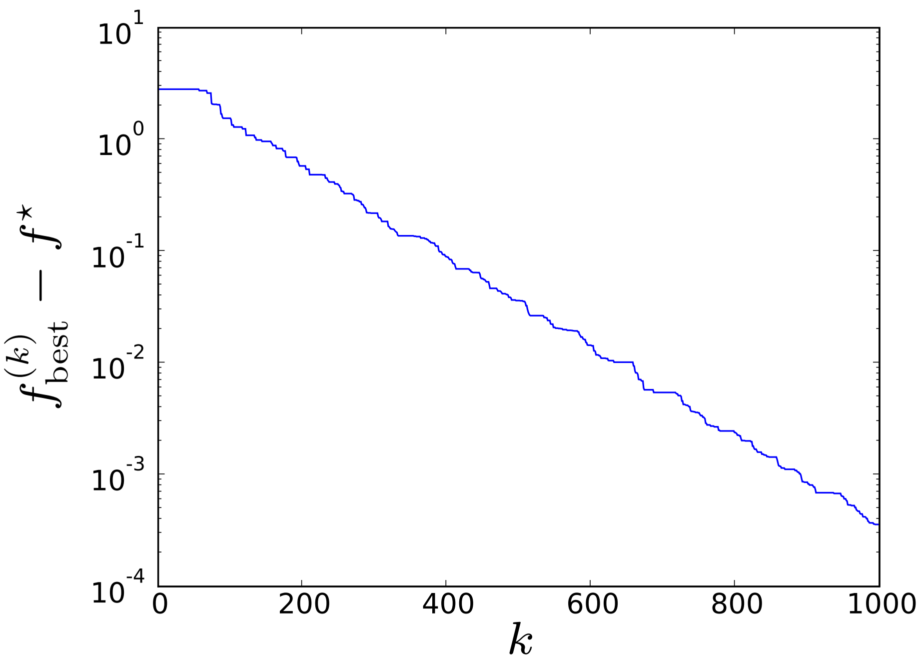
can stop when $f_{\rm best}^{(k)} - l_{\rm best}^{(k)} \leq \epsilon$ to guarantee ϵ-suboptimality

Example: piecewise linear minimization

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

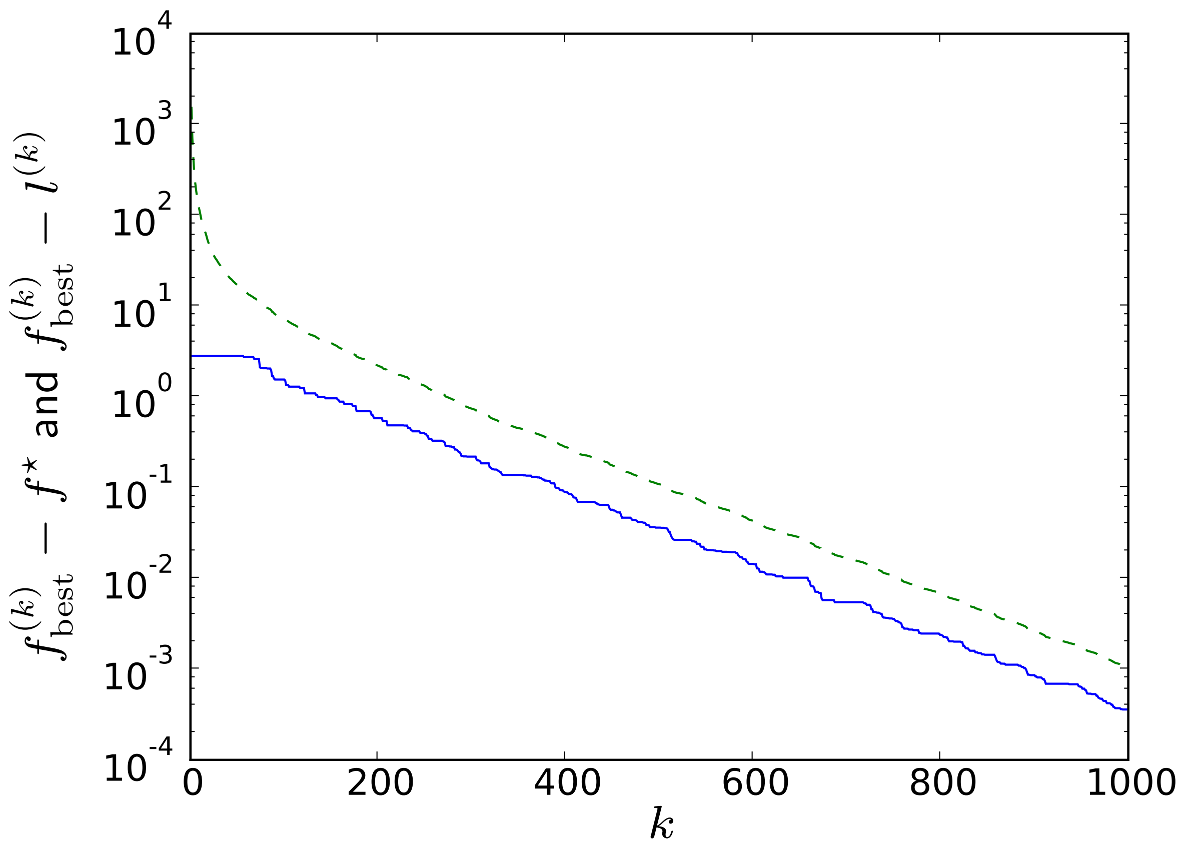
$n = 100$ variables, $m = 200$ terms, $f^\star \approx 0.36$





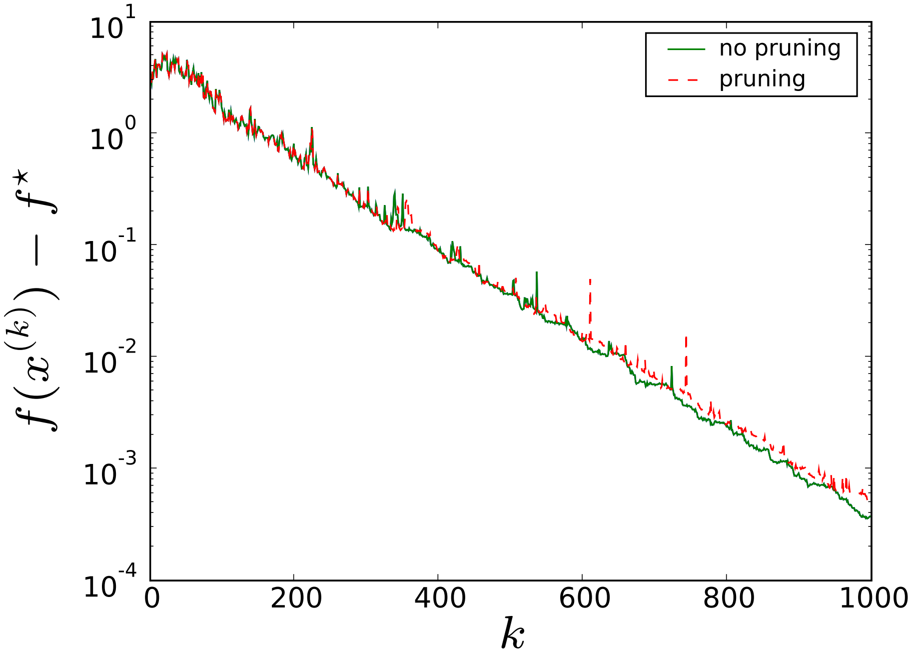
computed lower bound on optimal value

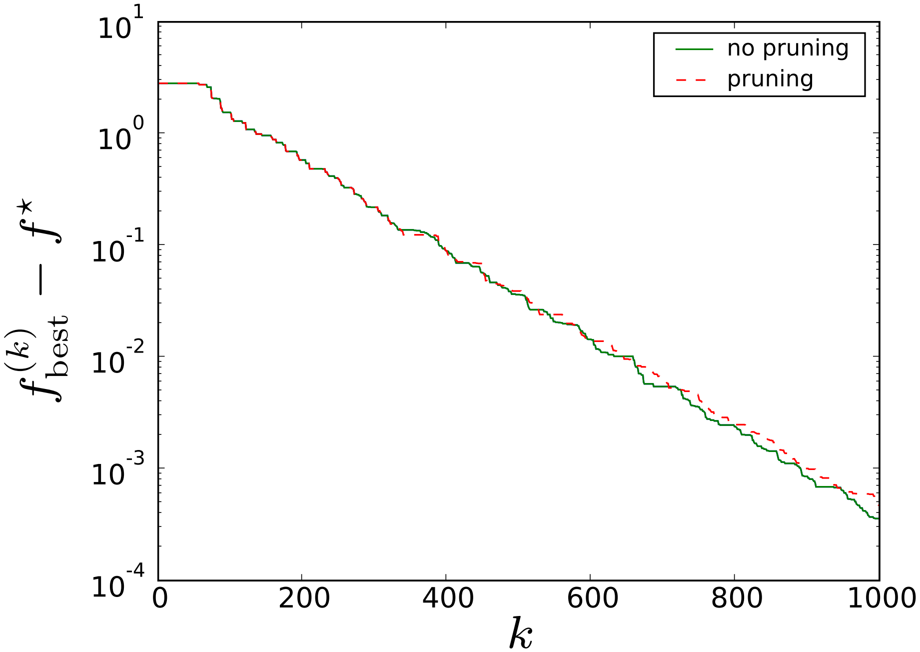
$f_{\rm best}^{(k)} - l^{(k)}$ (dashed line) and $f_{\rm best}^{(k)} - f^\star$ (solid line)



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