COMP 9602: Convex Optimization

Localization Methods

Dr. C Wu

Department of Computer Science
The University of Hong Kong

Roadmap

Theory	convex set convex function standard forms of optimization problems, quasi-convex optimization linear program, integer linear program quadratic program geometric program semidefinite program vector optimization duality
Algorithm	unconstrained optimization equality constrained optimization interior-point method subgradient methods localization methods decomposition methods and more

Localization methods

- also called cutting-plane methods
  - based on the idea of “localizing” desired point in some set, which becomes smaller at each step
- classes of algorithms
  - Ellipsoid method
  - Center of gravity method
  - Analytic centering method
  - Chebyshev center method
  - ...
- less efficient for problems to which interior point methods apply
- but can solve general convex and quasi-convex problems
  - may not require differentiability of the objective and constraint functions
- as compared to subgradient methods
  - can be much more efficient
  - require more memory and computation per step

Cutting plane oracle

- Oracle: a device that can answer question for us; we make no assumption on how an answer is found by the device

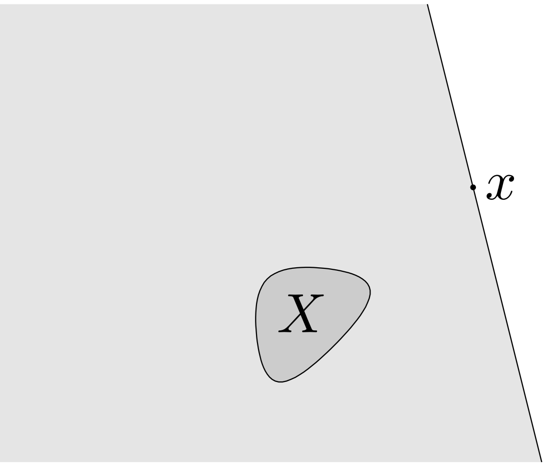
- Cutting plane oracle (also called separation oracle), once queried at x, either
  - asserts $x \in X$ ($X \subseteq \mathbf{R}^n$)
  - or returns a separating hyperplane between x and X : $a \neq 0$,

$$a^T z \leq b \text{ for } z \in X, \qquad a^T x \geq b$$

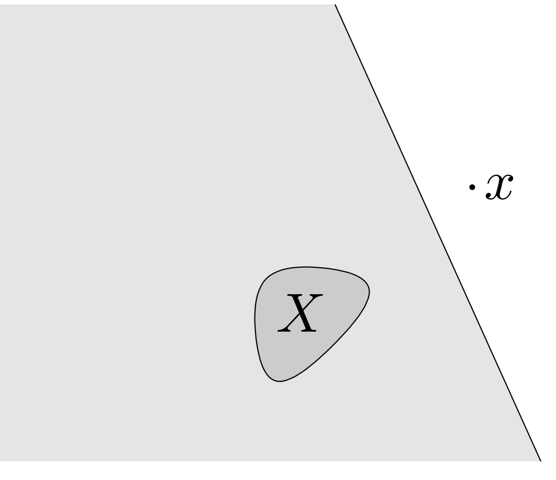
(a, b) called a *cutting-plane*, or *cut*, since it eliminates the halfspace $\{z \mid a^T z > b\}$ from our search for a point in X

Neutral and deep cuts

- If $a^T x = b$ (x is on boundary of the halfspace that is cut), the cutting plane is called a neutral cut



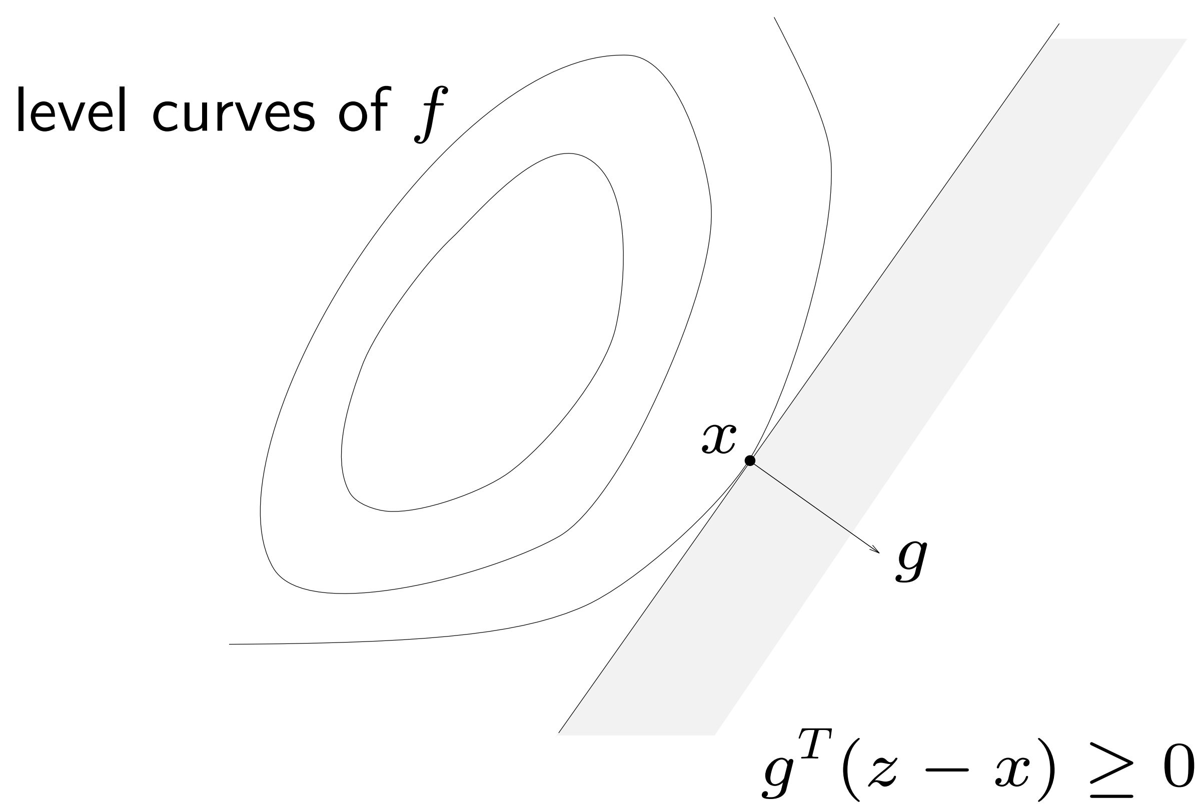
- If $a^T x > b$ (x lies in the interior of the halfspace that is cut), the cutting plane is called a deep cut



Unconstrained optimization

$$\text{minimize} \quad f_0(x)$$

- $f_0 : \mathbf{R}^n \to \mathbf{R}$, convex; x^* is the optimal solution; X is the set of optimal points
- Given x, find $g \in \partial f_0(x)$
- x^* belongs to the halfspace: $\{z | g^T(z - x) \leq 0\}, \forall x$
  - $g^T(z - x) \leq 0$ defines a cutting plane at x ($a = g$, $b = g^T x$)



Unconstrained optimization (cont'd)

$$\text{minimize} \quad f_0(x)$$

- The idea of the localization algorithm is to use hyperplanes to separate the current point from the optimal point

For k=1,2,...

- at $x^{(k)}$, evaluate $g \in \partial f_0(x^{(k)})$

  then $x^* \in \{z | g^T(z - x^{(k)}) \leq 0\}$

  use this cutting plane to cut down the volume of the set of feasible points

- choose $x^{(k+1)}$ at the center of the new polyhedron and hope to further cut down its volume

Inequality constrained optimization

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

$f_0, \ldots, f_m : \mathbf{R}^n \to \mathbf{R}$ convex

- If x is feasible, we have a (neural) objective cut

$$g_0^T(z - x) \leq 0, \qquad g_0 \in \partial f_0(x)$$

- If x is not feasible, e.g., $f_j(x) > 0$, we have a (deep) feasibility cut

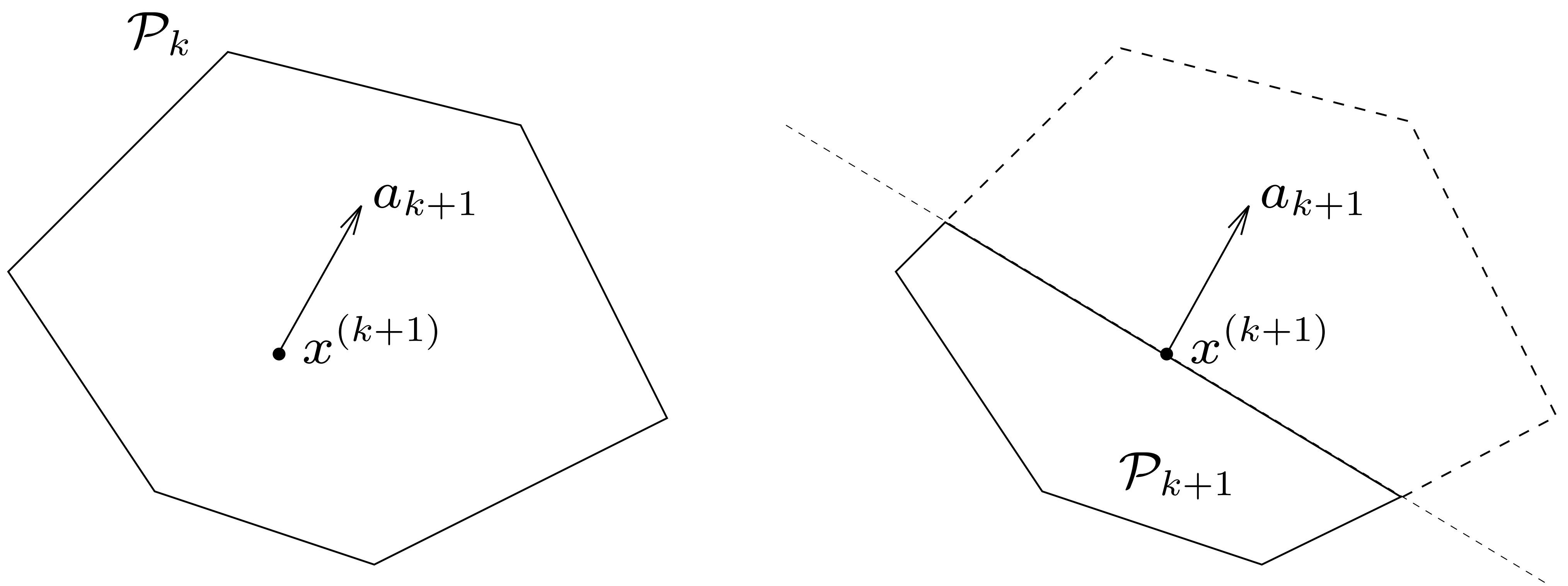
$$f_j(x) + g_j^T(z - x) \leq 0, \qquad g_j \in \partial f_j(x)$$

Localization algorithm

- Basic (conceptual) cutting-plane/localization algorithm

given an initial polyhedron $\mathcal{P}_0 = \{z \mid Cz \preceq d\}$ known to contain X.
$k := 0$.
repeat
  Choose a point $x^{(k+1)}$ in $\mathcal{P}_k$.
  Query the cutting-plane oracle at $x^{(k+1)}$.
  If the oracle determines that $x^{(k+1)} \in X$, quit.
  Else, add new cutting-plane $a_{k+1}^T z \leq b_{k+1}$:
    $\mathcal{P}_{k+1} := \mathcal{P}_k \cap \{z \mid a_{k+1}^T z \leq b_{k+1}\}$
  If $\mathcal{P}_{k+1} = \emptyset$, quit
  $k := k + 1$

Localization algorithm (cont'd)



- want to pick $x^{(k+1)}$ so that $\mathcal{P}_{k+1}$ is as small as possible, no matter what cut is made

- want $x^{(k+1)}$ near center of $\mathcal{P}^{(k)}$

  center of gravity
  analytic center
  Chebyshev center
  ...

Choosing the center point (specific localization methods)

P. 416, textbook

- Center of gravity
  - Center of gravity method (CG algorithm)
  - $$x^{(k)} = \mathbf{cg}(C^{(k-1)}) = \frac{\int_{C^{(k-1)}} z dz}{\int_{C^{(k-1)}} dz}$$
- Analytic centering
  - Analytic centering cutting-plane method (ACCPM)
  - $x^{(k)}$ is the analytic center of $C^{(k-1)}$, i.e.,

$$x^{(k)} = argmin_z - \sum_{i=1}^{m} \log(b_i - a_i^T z)$$

- Chebyshev center
  - Chebyshev center cutting-plane method
  - Chebyshev center is the center of the largest ball inside $C^{(k-1)}$
- Center of the maximum volume ellipsoid (MVE)

Bisection method on R

- minimize convex $f : \mathbf{R} \to \mathbf{R}$
- $\mathcal{P}_k$ is an interval
- obvious choice for query point: $x^{(k+1)} := \text{midpoint}(\mathcal{P}_k)$

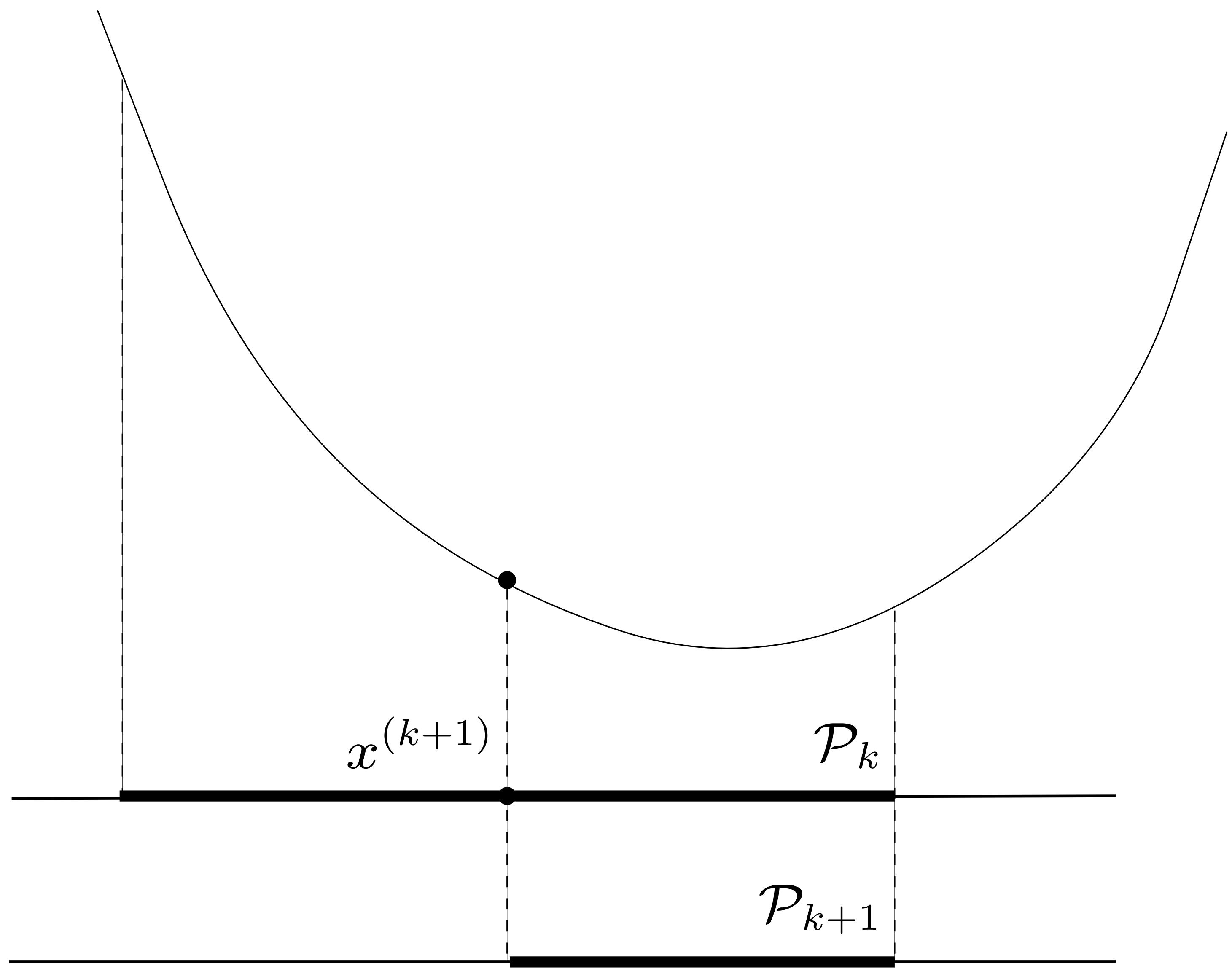
Bisection algorithm for one-dimensional search.

given an initial interval $[l, u]$ known to contain x^*; a required tolerance $r > 0$
repeat
   $x := (l + u)/2$.
   Query the oracle at x.
   If the oracle determines that $x^* \leq x$, $u := x$.
   If the oracle determines that $x^* \geq x$, $l := x$.
until $u - l \leq 2r$

Bisection method on R (cont'd)

for differentiable f: evaluate $f'(x)$

if $f'(x) < 0$, $l := x$; else $u := x$



Ellipsoid method (for unconstrained convex problems)

- Idea
  - Use ellipsoids to approximate polyhedron, find x* in an ellipsoid instead of a polyhedron
- Algorithm sketch

ellipsoid at iteration k

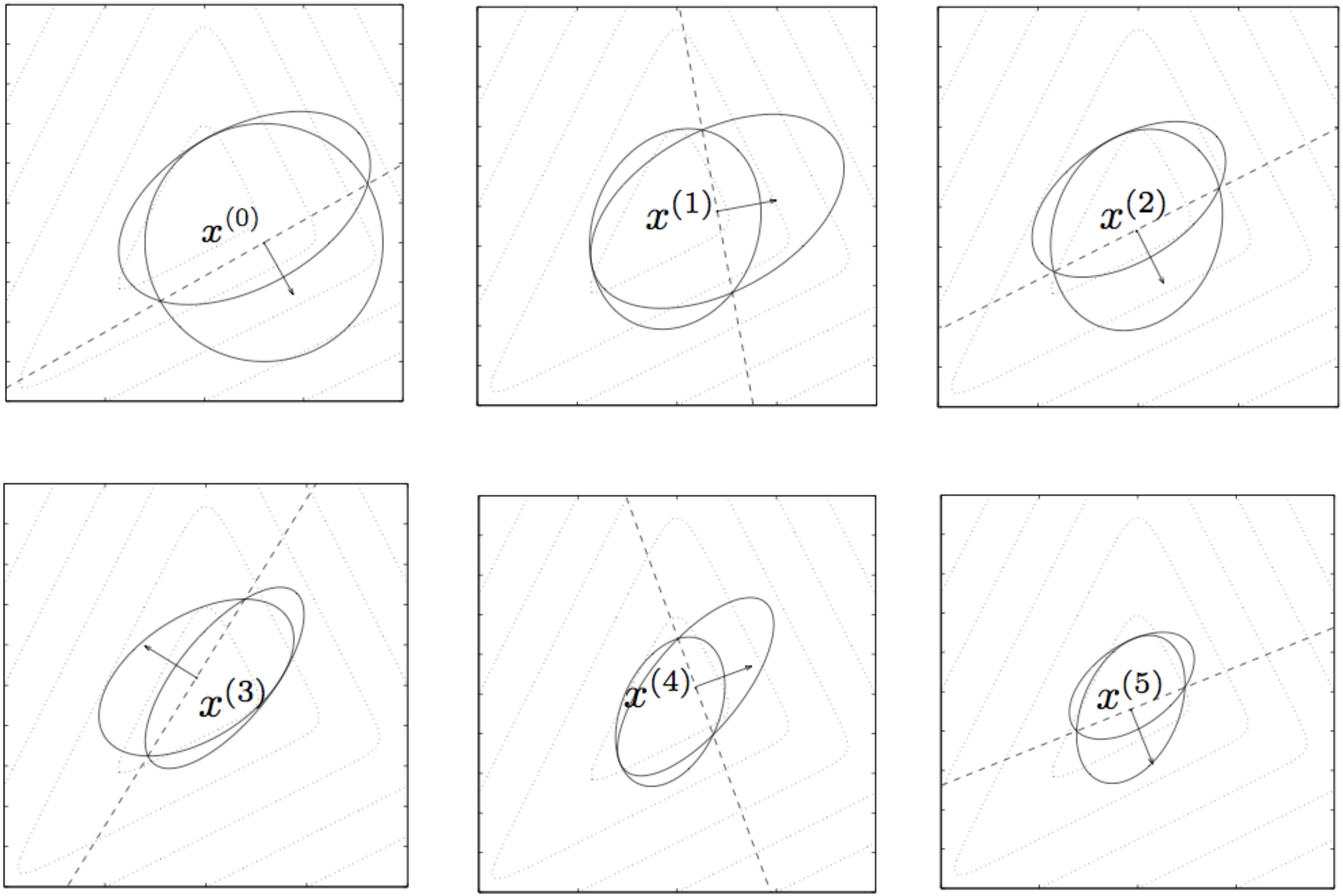
1. at iteration k we know $x^\star \in \mathcal{E}^{(k)}$
2. set $x^{(k+1)} := \text{center}(\mathcal{E}^{(k)})$; evaluate $g^{(k+1)} \in \partial f_0(x^{(k+1)})$
   ($g^{(k)} = \nabla f_0(x^{(k)})$ if f_0 is differentiable)
3. hence we know

$$x^\star \in \mathcal{E}^{(k)} \cap \{z \mid g^{(k+1)T}(z - x^{(k+1)}) \leq 0\}$$

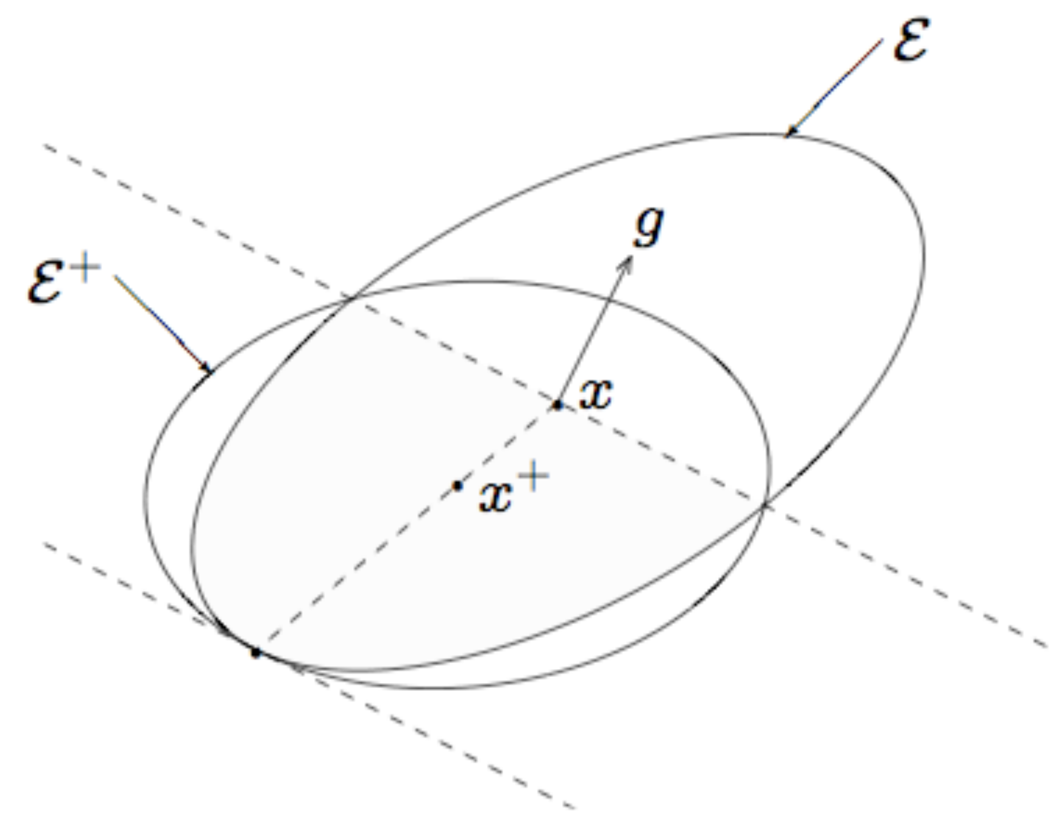
   (a half-ellipsoid)

4. set $\mathcal{E}^{(k+1)} :=$ minimum volume ellipsoid covering
   $\mathcal{E}^{(k)} \cap \{z \mid g^{(k+1)T}(z - x^{(k+1)}) \leq 0\}$

Example



Ellipsoid method (cont'd)



- How to find the minimum volume ellipsoid?

- let $\mathcal{E}(x, P) = \{z \mid (z-x)^T P^{-1}(z-x) \leq 1\}$
- the minimum volume ellipsoid to cover $\mathcal{E} \cap \{z \mid g^T(z-x) \leq 0\}$ is described by: (n is the dimension of x)

$$
\begin{aligned}
x^+ &= x - \frac{1}{n+1} P\tilde{g} \\
P^+ &= \frac{n^2}{n^2-1}\left(P - \frac{2}{n+1} P\tilde{g}\tilde{g}^T P\right)
\end{aligned}
$$

where $\tilde{g} = (1/\sqrt{g^T P g})g$

Ellipsoid method (cont'd)

- Basic ellipsoid algorithm (for unconstrained convex problems)

given ellipsoid $\mathcal{E}(x, P)$ containing $x^\star$, accuracy $\epsilon > 0$

repeat

1. evaluate $g \in \partial f_0(x)$
2. if $\sqrt{g^T P g} \leq \epsilon$, return$(x)$
3. update ellipsoid

   3a. $\tilde{g} := \frac{1}{\sqrt{g^T P g}} g$

   3b. $x := x - \frac{1}{n+1} P\tilde{g}$

   3c. $P := \frac{n^2}{n^2-1}\left(P - \frac{2}{n+1} P\tilde{g}\tilde{g}^T P\right)$

Ellipsoid method (cont'd)

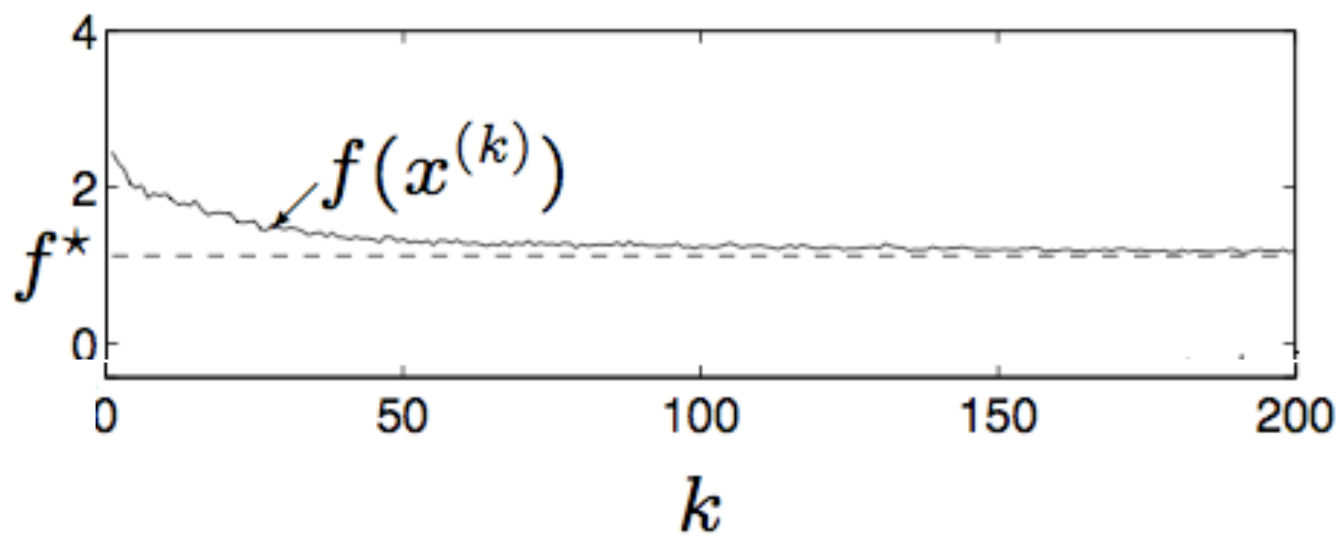
- Stopping criteria

$$\sqrt{g^{(k)T}P^{(k)}g^{(k)}} \leq \epsilon \quad \Longrightarrow \quad f(x^{(k)}) - f(x^\star) \leq \epsilon$$

- Convergence
  - $\mathbf{vol}(\mathcal{E}^{(k+1)}) < e^{-\frac{1}{2n}}\,\mathbf{vol}(\mathcal{E}^{(k)})$
    (volume reduction factor degrades rapidly with n, compared to CG or MVE cutting-plane methods)
  - modest computation per step ($O(n^2)$), via analytical formula
  - efficient in theory; slow but steady in practice

Example

$f(x) = \max_{i=1}^{m}(a_i^T x + b_i)$, with $n = 20$, $m = 100$



Ellipsoid method (for inequality constrained problems)

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

- if $x^{(k)}$ feasible, update ellipsoid with objective cut

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

a deep cut

$f_{\text{best}}^{(k)}$ is best objective value of feasible iterates so far

- if $x^{(k)}$ infeasible, update ellipsoid with feasibility cut

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

a deep cut

assuming $f_j(x^{(k)}) > 0$

Ellipsoid method (for inequality constrained problems)

minimum volume ellipsoid containing ellipsoid intersected with halfspace

$$\mathcal{E} \cap \{z \mid g^T(z-x) + h \le 0\}$$

with $h \ge 0$, is given by

$$\begin{aligned} x^+ &= x - \frac{1+\alpha n}{n+1} P\tilde{g} \\ P^+ &= \frac{n^2(1-\alpha^2)}{n^2-1}\left(P - \frac{2(1+\alpha n)}{(n+1)(1+\alpha)} P\tilde{g}\tilde{g}^T P\right) \end{aligned}$$

where

$$\tilde{g} = \frac{g}{\sqrt{g^T P g}}, \qquad \alpha = \frac{h}{\sqrt{g^T P g}}$$

(if $\alpha > 1$, intersection is empty)

Ellipsoid method (for inequality constrained problems)

- Stopping criteria

  - if $x^{(k)}$ is feasible and $\sqrt{g_0^{(k)T} P^{(k)} g_0^{(k)}} \leq \epsilon$ ($x^{(k)}$ is ϵ-suboptimal)

  - if $f_j(x^{(k)}) - \sqrt{g_j^{(k)T} P^{(k)} g_j^{(k)}} > 0$ (problem is infeasible)

Reference

- Localization methods: localization_methods_notes.pdf (reference 7 on Moodle)
- Ellipsoid method:
  - ellipsoid_method_notes (reference 7 on Moodle)
  - pp. 170-185, C. H. Papadimitrou, K. Steiglitz. Combinatorial Optimization: Algorithms and Complexity, 1998.

Acknowledgement

- Some materials are extracted from the slides created by Prof. Stephen Boyd for EE364b at Standard University