Lecture 16

Cutting plane methods

Outline

- cutting plane methods intuition
- the center of gravity method
- ellipsoid method
- Vaidya's cutting plane algorithm

- cutting plane methods

Problem: Check if a convex set X is empty or not. A membership oracle is given.

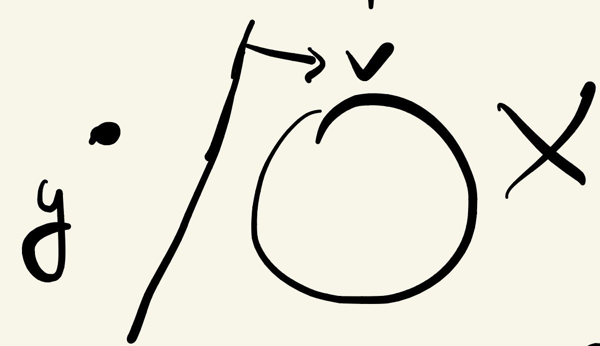
example: X is a polytope, which is given as the intersection of linear inequalities.

Oracle: checks if a point satisfies all the inequalities

This abstraction is important. Because there are polytopes which are described by exponentially many inequalities, but the membership oracle can be computed in polynomial time. (e.g. perfect matching in general graphs)

Cutting plane method approach

We assume a separation oracle.

That is  if $y \notin X$ then

the separation oracle returns an hyperplane v s.t $\langle v, x\rangle \geq \langle v, y\rangle$ $\forall x \in X$

Procedure

(1) Take $S \supseteq X$

(2) choose $y \in C$ and check if $y \in X$

(3) If yes then we output IT IS NOT EMPTY

(4) If no use the separating hyperplane v to cutoff $S \to S'$ s.t

$S \supset S' \supseteq X$

(5) repeat (2)

Observations

(1) we should assume that if X is non-empty then it is contained in some small ball of volume $= \alpha$ so that we know when to stop cutting off S (when $vol(S) < \alpha$)

(2) ideally we would like to cutoff S as quickly as possible and thus to choose $y \simeq$ center of S

Different cutting plane methods consist in how to choose y and how S is select.

Center of gravity method

convex function

$$\min \; f(x)$$

convex set

$$\text{s.t} \quad x \in X \subseteq \mathbb{R}^n$$

we assume that we have a:

(1) value oracle $f(y) \; \forall y \in X$

(2) subgradient oracle $\Rightarrow$

$$v_y \in \mathbb{R}^n \text{ st } f(x) \geq f(y) + \langle v_y, x-y \rangle \quad \forall x \in X$$

Algorithm

$S_1 = X$

for $t = 1$ to T do

- $c_t = \frac{1}{vol(S_t)} \int_{x \in S_t} x \, dx$
- $v_t \leftarrow$ subgradient at c_t $f(x) \geq f(c_t) + \langle v_t, x - c_t \rangle$
- $S_{t+1} \leftarrow S_t \cap \{ x \in \mathbb{R}^n \mid \langle v_t, x - c_t \rangle \leq 0 \}$

return $x_{alg} \in \arg\min_{c \in \{c_1, \dots, c_T\}} f(c)$

Analysis

Theorem (Grünbaum)

Let K be a centered convex set

(i.e. $\int_{x \in K} x \, dx = 0$). Then $\forall v \in \mathbb{R}^n$

$v \neq 0$, $\Rightarrow vol(K \cap \{x \in \mathbb{R}^n \mid \langle v, x \rangle \geq 0\}) \geq \frac{1}{e} vol(K)$

So in our case $vol(S_{t+1}) \leq (1-\frac{1}{e}) vol(S_t)$

The idea is that we cut off all the points for which we are sure they are not optimal

$$f(x) \geq f(c_t) + \langle v_t, x - c_t \rangle$$

If $\langle v_t, x - c_t \rangle \Rightarrow f(x) > f(c_t)$

and at every iteration we cut off a constant fraction of the volume!!!

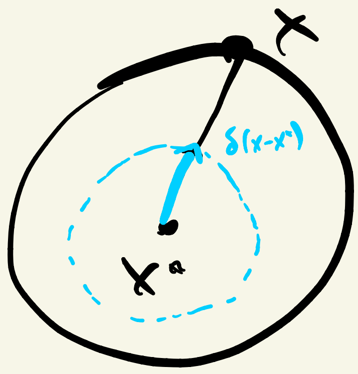
When do we stop?

x^* optimal point

and assume that $|f(x)| < B \ \forall x \in X$

let $X_\delta = \{(1-\delta)x^* + \delta x \mid \forall x \in X\} =$

$$= \{x^* + \delta(x - x^*) \mid \forall x \in X\}$$



X

$$\text{vol}(X_\delta) = \delta^n \text{vol}(X)$$

$$\forall y \in X_\delta \quad f(y) \leq (1-\delta) f(x^*) + \delta f(x) \leq$$

$$\leq f(x^*) + \delta B$$

$\Rightarrow \delta = \frac{\varepsilon}{B}$ we get an ε-additive error

we want

$$\Rightarrow \text{vol}(S_T) \leq \text{vol}(X_{\varepsilon/B}) = \left(\frac{\varepsilon}{B}\right)^n \text{vol}(x)$$

$\Downarrow$

It suffices

$$\left(1 - \frac{1}{e}\right)^T \text{vol}(X) \leq \left(\frac{\varepsilon}{B}\right)^n \text{vol}(x) \Rightarrow T = O\left(n \log \frac{B}{\varepsilon}\right)$$

iterations.

Discussion

The "expensive" computationally part of the algorithm is the calculation of $c_t = \frac{1}{vol(S_t)} \int_{x \in S_t} x \, dx$. It can be approximate by sampling random points (and also argue, that on expectation a constant fraction of S_t is cut off).

We still need $\Omega(n)$ random points.

Ellipsoid method

Same idea. But the containing sets are ellipsoids, for which is easier to compute the center. However the the volume of the containing set does not decrease so fast $\Rightarrow$

$\Rightarrow$ more iterations

Ellipsoid method

$H_1 = R^2 \cdot I$, $c_1 = 0$

R big enough such that $X \subseteq \{x \in \mathbb{R}^n \mid (x-c_1)^T H_1^{-1} (x-c_1) \leq 1\}$

for $t = 1$ to T do

- c_t center of $\mathcal{E}_t = \{x \in \mathbb{R}^n \mid (x-c_t)^T H_t^{-1} (x-c_t) \leq 1\}$
- If $c_t \in X \Rightarrow v_t$ subgradient of f at c_t

  $c_t \notin X \Rightarrow v_t$ separating hyperplane

- $\mathcal{E}_{t+1}$ = minimum ellipsoid that contains

  $$\mathcal{E}_t \cap \{x \in \mathbb{R}^n \mid \langle v_t, x - c_t \rangle \leq 0\}$$

return $x_{alg} \in \operatorname{argmin}_{c \in \{c_1, \dots, c_T\}} f(c)$

Analysis

we will prove that $\text{vol}(\mathcal{E}_{t+1}) \leq e^{-\frac{1}{2n}} \text{vol}(\mathcal{E}_t)$

$\Rightarrow n$ iterations to decrease by a constant fraction $\rightarrow O\left(n \cdot n \log \frac{B \cdot R}{\varepsilon}\right) = O\left(n^2 \log \frac{B \cdot R}{\varepsilon}\right)$

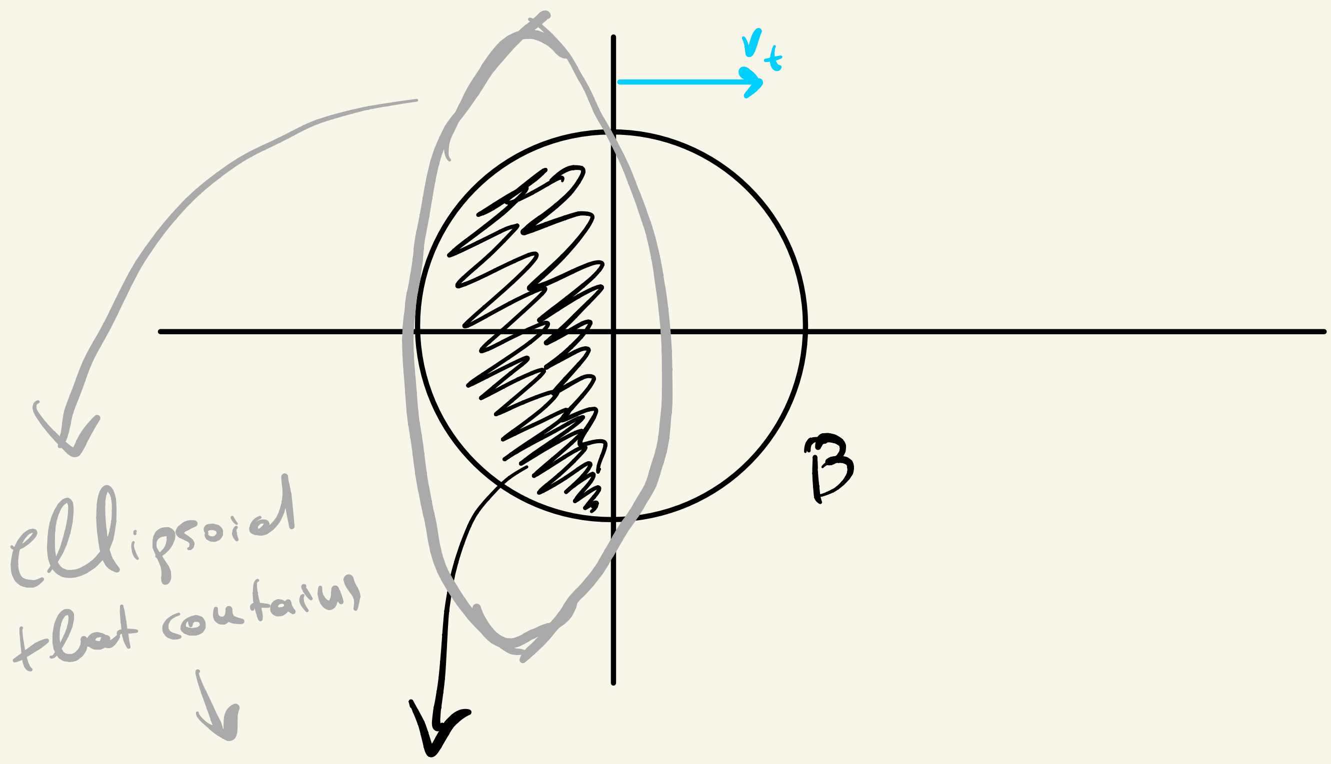
Finding $\mathcal{E}_{t+1}$ (and prove that $vol(\mathcal{E}_{t+1}) \le e^{-1/2n} vol(\mathcal{E}_t)$)

special case

assume that $\mathcal{E}_t = B$ ← a unit ball ($c_t = 0$)

and v_t defines the cutting plane

w.l.o.g. $\|v_t\| = 1$



$$B \cap \{x \in \mathbb{R}^n \mid \langle v_t, x \rangle \le 0\}$$

observations

① by symmetry the center of the new ellipsoid should be $-av$, $a \in [0,1]$

② v is on semiaxis and the rest of the semiaxes are orthogonal to v

and they have the same length

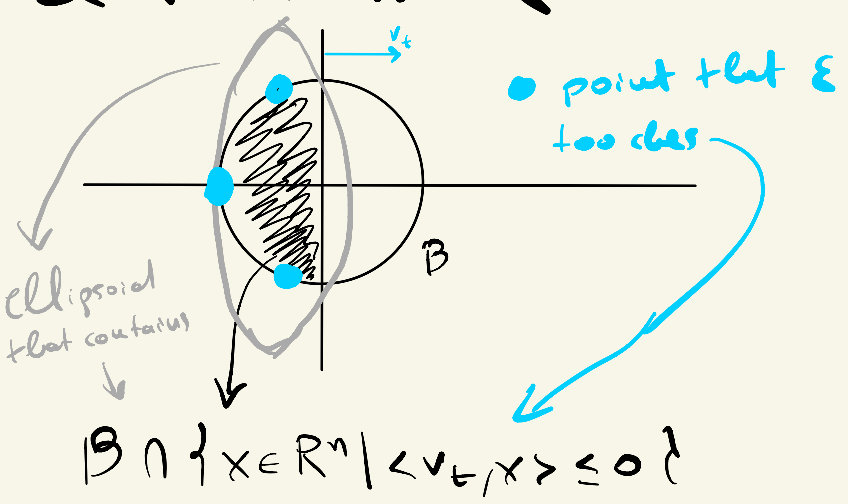
$$\Rightarrow H^{-1} = l_1 v_t v_t^T + l_2 (I - v_t v_t^T)$$

$$c = -av$$

now we want to minimize the volume of such ellipsoid such that it contains $B \cap \{x \in \mathbb{R}^n \mid \langle v_t, x \rangle \leq 0\}$

$$\mathcal{E} = \{x \in \mathbb{R}^n : \|x - c\|^2_{H^{-1}} \leq 1\} \subseteq$$

a minimal such ellipsoid touches



$$-v \Rightarrow (-v + av)^T (l_1 v v^T + l_2 (I - v v^T))(-v + av) = 1$$

$$\Rightarrow (a-1)^2 l_1 = 1$$

$$\partial B \cap v^\perp \Rightarrow (y + av)^T (l_1 v v^T + l_2 (I - v v^T))(y + av) = 1$$

$$\forall y \perp v, \|y\|_2 = 1 \Rightarrow l_2 + l_1 a^2 = 1$$

$$\mathrm{vol}(\mathcal{E}) = \frac{1}{\sqrt{\ell_1}} \cdot \left(\frac{1}{\sqrt{\ell_2}}\right)^{n-1}$$

$$(a-1)^2 \ell_1 = 1$$

$$\ell_2 + \ell_1 a^2 = 1$$

$$\min \mathrm{vol}(\mathcal{E})$$

$$a = \frac{1}{n+1}$$

$$\ell_1 = \left(1+\frac{1}{n}\right)^2$$

$$\ell_2 = 1 - \frac{1}{n^2}$$

$$\mathcal{E} = \left\{ x \in \mathbb{R}^n \;\middle|\; \left(x + \frac{v/\|v\|_2}{n+1}\right)^T \left(\left(1-\frac{1}{n^2}\right)I + \frac{2}{n}\left(1+\frac{1}{n}\right)\frac{vv^T}{\|v\|_2^2}\right) \cdot \left(x + \frac{v/\|v\|_2}{n+1}\right) \le 1 \right\}$$

general case

given $\mathcal{E} = \{x \mid (x-c)^T H^{-1}(x-c) \le 1\}$ and v

we would like to find an ellipsoid $\mathcal{E}'$

s.t.

$\mathcal{E}' \supseteq \mathcal{E} \cap \{x \mid \langle v, x-c\rangle \le 0\}$ and $\mathcal{E}'$

has minimal volume

Do an affine transformation to get back to the ball case

$$x = H^{1/2}y + c \Rightarrow \mathcal{E} = \{y \mid y^T y \le 1\}$$

$$\mathcal{E} \cap \{y \mid \langle vH^{1/2}, y\rangle \le 0\} \Rightarrow$$

$$\mathcal{E}' = \left\{ x \in \mathbb{R}^n \;\middle|\; \left(x + \frac{v/\|v\|_2}{n+1}\right)^T \left(\left(1 - \frac{1}{n^2}\right) I + \frac{2}{n}\left(1 + \frac{1}{n}\right) \frac{vv^T}{\|v\|_2^2} \right) \cdot \left(x + \frac{v/\|v\|_2}{n+1}\right) \leq 1 \right\}$$

where $v = H^{1/2} v$

$$\mathcal{E}' = \left\{ y \in \mathbb{R}^n \;\middle|\; \left(y + \frac{H^{1/2}v}{(n+1)\|v\|_H}\right)^T \left(\left(1 - \frac{1}{n^2}\right) I + \frac{2}{n}\left(1 + \frac{1}{n}\right) \frac{H^{1/2}vv^TH^{1/2}}{\|v\|_H^2} \right) \left(y + \frac{H^{1/2}v}{(n+1)\|v\|_H}\right) \leq 1 \right\}$$

$$y = H^{-1/2}(x - c)$$

$$\mathcal{E}' = \{ x \in \mathbb{R}^n \mid \dots \}$$

The decrease in the volume is the same as the ratio remains unchanged under affine transformations!!!

LPs
\# iterations $\Omega(n^2 \log 1/\varepsilon)$

per iteration $\Rightarrow$ only the separation oracle
If we check every constraint
$\Rightarrow \Omega(m \cdot n)$

$\Rightarrow \Omega(mn^3 \log n)$ super slow but still theoretically important. It permits to solve LPs with exponentially many constraints if an oracle can be computed efficiently