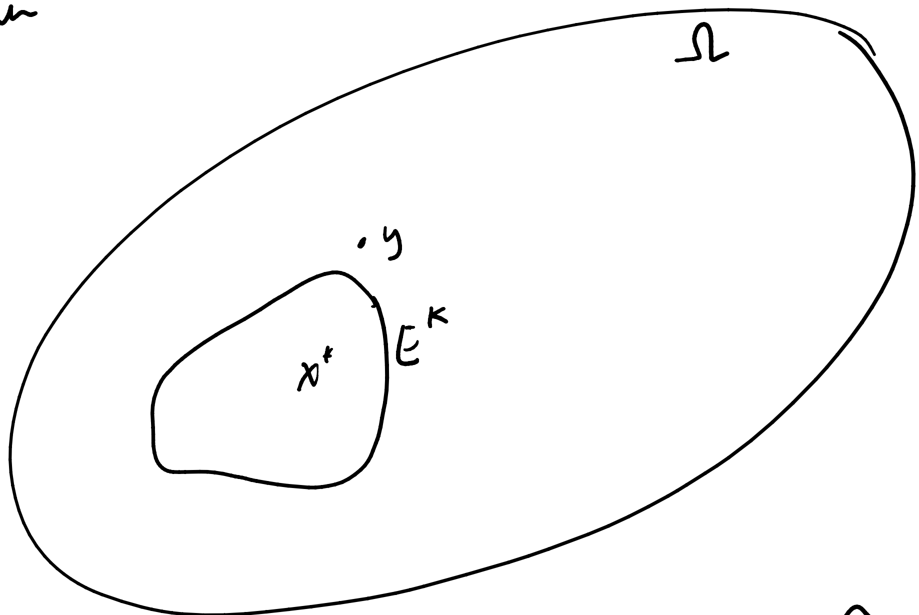
The Cutting Plane Method

Wednesday, January 15, 2020 6:34 PM

Yulin



$\alpha > \dfrac{\nu(E_k)}{\nu(\Omega)}$. $\quad S = (1-\alpha)x^* + \alpha\Omega$

$$y \in S \setminus E_k \;\Leftarrow\; \nu(S) = \alpha\,\nu(\Omega) > \nu(E_k)$$

By separation: $f(y) \geq f(x^i)$ for some $i \leq k$.

$$\exists z \in \Omega : \; y = (1-\alpha)x^* + \alpha z$$

$$f(x^k) \leq f(y) \leq (1-\alpha) f(x^*) + \alpha f(z)$$

$$f(x^k) - f(x^*) \leq \alpha\left(f(z) - f(x^*)\right)$$

$\Rightarrow$ **Thm.**

$$f(x^k) - f(x^*) \leq \frac{\nu(\Omega)}{\nu(E^k)}\left(f(x^0) - f(x^*)\right) \quad (*)$$

$\square$

Let $E_0 = \Omega, E_1, \ldots, E_k$ be the sequence of sets and $x^0, \ldots, x^k$ be the queries.

and $x^0 \dots x^k$ be the queries.

Let $\nu : 2^{\mathbb{R}^n} \longrightarrow \mathbb{R}_+$ be s.t.

(1) $\nu(\alpha E + x) = \alpha \nu(E)$

and (2) $S \subseteq E \Rightarrow \nu(S) \le \nu(E)$.

then $(*)$ holds.

Cutting plane Method.

Start with E_0.

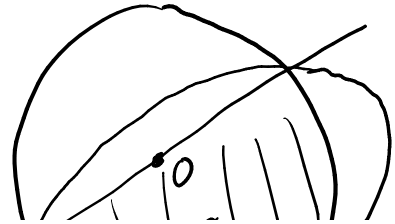
Repeat.
$$\left[\begin{array}{l} \text{Choose } x^k \in E_k \\ \text{Using } \nabla f,\ E_{k+1} \subseteq \{ y : \langle \nabla f(x), y - x \rangle \le 0 \} \end{array}\right.$$

- How to choose x^k, E^k?
- How to measure progress?
- Rate of convergence?
- Time to implement each step?

Ellipsoid algorithm.

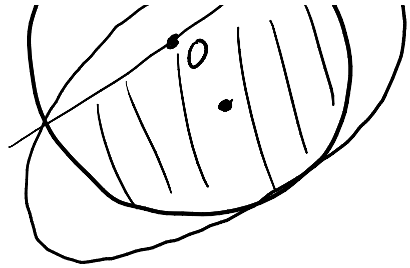
E_k is an ellipsoid.

$E_0 = B(0, R)$.



$E_0 = B(0,R)$.

$x^0 = 0$.



$\nabla f(x_0)(x - x^0) \le 0$

E_1 = min volume E containing $E_0 \cap \underbrace{\{x : \nabla f(x^0)^T (x - x^0) \le 0\}}_{H}$.

x^1 = center of E_1.

Maintain x^k, $E_k = \{x : (x - x^k)^T A_k^{-1} (x - x^k) \le 1\}$

$A_0 = R^2 I$

Repeat:

$$\begin{cases} x^{k+1} = x^k - \dfrac{1}{n+1} \dfrac{A_k \nabla f(x^k)}{\sqrt{\nabla f(x^k) A_k \nabla f(x^k)}} \\[2ex] A_{k+1} = \left(\dfrac{n^2}{n^2-1}\right) \cdot \left(A_k - \dfrac{2}{n+1} \cdot \dfrac{A_k \nabla f(x^k) \nabla f(x^k)^T A_k}{\nabla f(x^k) A_k \nabla f(x^k)}\right) \end{cases}$$

Lemma. (1) $\mathrm{vol}(E_{k+1}) \le e^{-\frac{1}{2n+2}} \cdot \mathrm{vol}(E_k)$

(2) $E_k \cap H_k \subseteq E_{k+1}$.

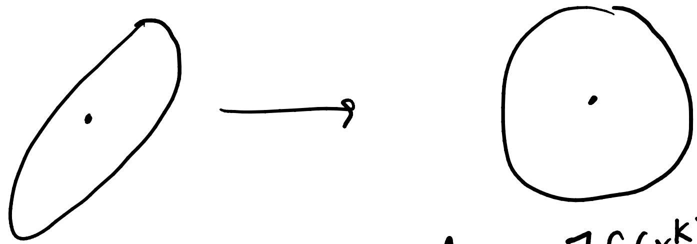
Pf. $\dfrac{\mathrm{vol}(E_{k+1})}{\mathrm{vol}(E_k)}$ is maintained by affine transformation.

$A^{-\frac{1}{2}}(x - x^k)$

$\overline{\mathrm{Vol}(E_k)}$

$E_k = A_k^{\frac{1}{2}} B(0,1) + x^k$, so apply $A_k^{-\frac{1}{2}} \cdot (x - x^*)$

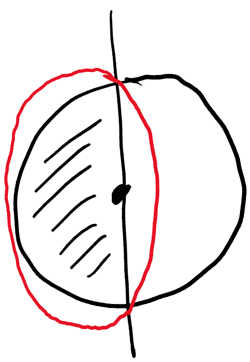
so that $E_k \to B(0,1)$ and $A_k = I$.



WLOG assume that $\nabla f(x^k) = e_1$

s.t. $H_k = \{x : e_1^T (x - 0) \le 0\}$

i.e. $x_1 \le 0$



So $A_{k+1} = \dfrac{n^2}{n^2-1}\left(I - \dfrac{2}{n+1} e_1 e_1^T\right)$

$$= \frac{n^2}{n^2-1}\begin{pmatrix} \frac{n-1}{n+1} & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & \ddots \end{pmatrix}$$

$$\left(\frac{\mathrm{Vol}(E_{k+1})}{\mathrm{Vol}(E_k)}\right)^2 = \left|\frac{\det(A_{k+1})}{\det(A_k)}\right|$$

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

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

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

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

$$\leq e^{\frac{1}{n^2-1}\cdot(n-1) - \frac{2}{n+1}} = e^{\frac{1}{n+1} - \frac{2}{n+1}} = e^{\frac{-1}{n+1}}$$

$$\mathrm{Vol}(E_{k+1}) \leq e^{-\frac{1}{2(n+2)}} \cdot \mathrm{Vol}(E_k).$$

---

(2) $E_k \cap H_k \subseteq E_{k+1}$.

$\forall x \in E_k \cap H_k$, we have $\|x\|_2 \leq 1$, $x_1 \leq 0$.

We need to check: $(x - x_{k+1})^T \cdot A_k^{-1} (x - x_{k+1}) \leq 1$.

$$x_{k+1} = \left(-\frac{1}{n+1}, 0, \cdots 0\right)^T$$

$$A_{k+1}^{-1} = \frac{n^2-1}{n^2} \cdot \begin{pmatrix} \frac{n+1}{n-1} & & 0 \\ & \ddots & \\ 0 & & \ddots \end{pmatrix}$$

So

$$\left(x_1 + \frac{1}{n+1}, x_2 \cdots x_n\right) \cdot \frac{n^2-1}{n^2} \cdot \begin{pmatrix} \frac{n+1}{n-1} & 0 \\ 0 & \ddots \end{pmatrix} \begin{pmatrix} x_1 + \frac{1}{n+1} \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \frac{n^2-1}{n^2} \cdot \left(\frac{n+1}{n-1} \cdot \left(x_1 + \frac{1}{n+1}\right)^2 + x_2^2 + \cdots x_n^2\right)$$

$$= \frac{n^2-1}{n^2} \cdot \left(\sum x_i^2 + \frac{2}{n-1} x_1^2 + \frac{2x_1}{n-1} + \frac{1}{n^2-1}\right)$$

$$\leq \frac{n^2-1}{n^2} \left(1 + \frac{1}{n^2-1}\right) \leq 1.$$

$$\le \frac{n^2-1}{n^2}\left(1+\frac{1}{n^2-1}\right) \le 1.$$

In fact, E_{k+1} is <u>min</u> volume E containing

$$E_k \cap H_k.$$

<u>Note</u>. ∇f is not necessary!

Any g: $g^T(x-x^k) \le 0$ contains x^* suffices.

i.e. any <u>separating</u> hyperplane of $\{x : f(x) \le f(x^*)\}$.

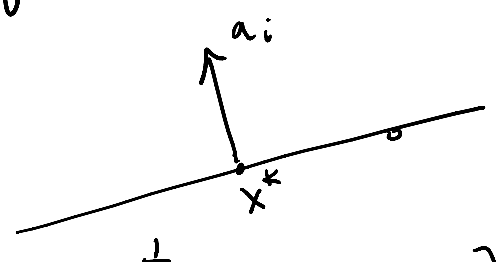
<u>App1</u>. LP. $\min\ c^T x$, $Ax \geq b$.

Set $f(x) = c^T x + \ell_{Ax\geq b}(x)$

$$\ell(x) = \begin{cases} 0 & Ax \geq b \\ \infty & \text{o.w.} \end{cases}$$

$$\nabla f(x) = \begin{cases} c & \text{if } Ax \geq b \\ -a_i & \text{if } a_i x < b \quad (\text{if many, pick any one}) \end{cases}$$

$-a_i^T (x - x^k) \le 0$



$R = \text{Diam}\{x : Ax \geq b\}$

$E_0 = B(0,R)$ $\quad \nu(E) = \text{vol}(E)^{\frac{1}{n}}$, $\nu(\{x : Ax \le b\}) = r$

$E_0 = B(0,R) \qquad \nu(E) = \mathrm{Vol}(E)^n, \quad \nu(\{x : Ax \le b\}) = r$

then

$$f(x) - f^* \le \varepsilon\,(f(x_0) - f^*)$$

in at most $n^2 \log \frac{R}{r\varepsilon}$ steps.

Time $= O\left((n^2 + \mathrm{nnz}(A)) \cdot n^2 \log \frac{R}{r\varepsilon} \right)$ steps.