Elimination: The Cutting Plane Method

Sunday, January 26, 2025 6:33 PM

Intro

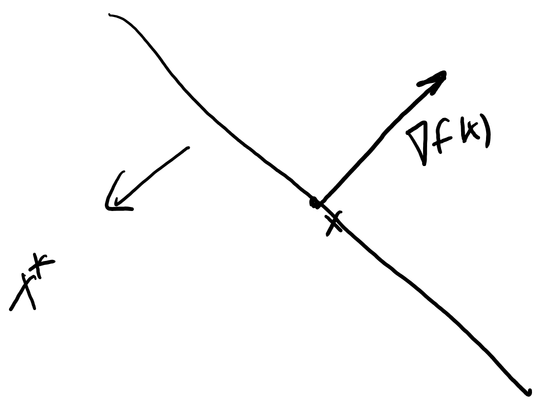
<u>OPT:</u> $\min_{x \in K} f(x)$.

<u>Convex OPT</u> f, K convex.

So $\forall x, y \quad f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$

and for any optimum x^*

$$0 \geq f(x^*) - f(x) \geq \langle \nabla f(x), x^* - x \rangle$$



So we can eliminate a halfspace and focus on its compliment.

"Binary search" in $\mathbb{R}^n$.

<u>Separation Oracle</u>. Convex set $K \subseteq \mathbb{R}^n$ closed.

$\forall x \in \mathbb{R}^n$, oracle returns $g(x) \in \mathbb{R}^n$ st.

$\forall y \in K \quad \langle g(x), y - x \rangle \leq 0$

---

for $K \subseteq \mathbb{R}^n$

Problem 1. Given $\varepsilon > 0$, Separation oracle g for $K \subseteq \mathbb{R}^n$. find $x \in K$ or declare $vol(K) < \varepsilon^n$.

Example. To minimize convex f, we can set $g(x) = \nabla f$ and $K = \{x : x \text{ is } \varepsilon\text{-approx minimizer}\}$

Basic Algorithm

Start with convex set $E^{(0)} \subseteq \mathbb{R}^n$ s.t. $K \subseteq E^{(0)}$.

Repeat:
- Choose $x \in E$
- Query $SEP_K(x)$
- If $x \in K \longrightarrow$ DONE ✓
  Else $H = \{y : \langle g(x), y - x \rangle \leq 0\}$
- Set $E \leftarrow E \cap H$.

Q1. How to choose E ?!

x ?

2. ______ x?

3. ______ measure $E^{(i)}$?

4. Rate of progress?

5. Complexity?

Ellipsoid Algorithm

Maintain Ellipsoid E containing K.

$E^{(0)} = B(0,R)$ (assumption: $K \subseteq$ Ball of radius R)

$$E(A, z) = \{x : (x-z)^T A^{-1} (x-z) \le 1\}$$

$A \succ 0$

x = center of E.

What about next E?

$\hat{E}$ = min volume Ellipsoid containing $E \cap H$.

Lemma! $E(A, z)$ $\quad H = \{x \mid v^T x \le v^T z\}$

$$\hat{z} = z - \frac{1}{n+1} \frac{Av}{\sqrt{v^T A v}}$$

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

$E(\hat{A}, \hat{z})$ is min volume Ellipsoid containing $E \cap H$.

$E(A, z)$ is min volume ellipsoid ...

Lemma 2. (i) $E \cap H \subseteq E(\hat{A}, \hat{z})$

(ii) $\mathrm{vol}(E(\hat{A}, \hat{z})) \le e^{-\frac{1}{2n+2}} \mathrm{vol}(E(A, z))$.

Fact. For $A, B \subseteq \mathbb{R}^n$, affine transformation T,

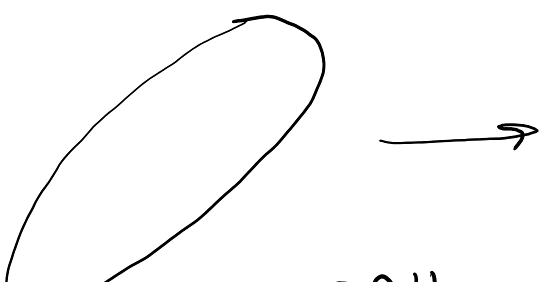
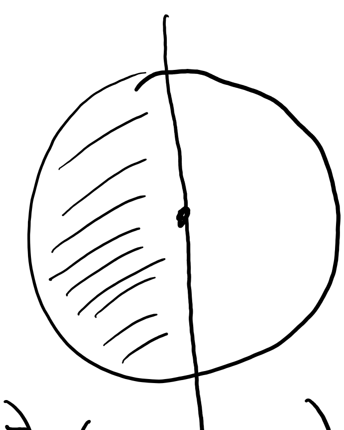
$$A \subseteq B \Rightarrow TA \subseteq TB$$

$$\frac{\mathrm{vol}(TA)}{\mathrm{vol}(TB)} = \frac{\cancel{|\det(T)|}}{\cancel{|\det(T)|}} \cdot \frac{\mathrm{vol}(A)}{\mathrm{vol}(B)}.$$

Pf. WLOG. we can assume $z = 0$, $A = I$.

$E(A, z)$ is $B(0, 1)$. Also $v = e_1$.

$$\hat{z} = -\frac{1}{n+1} e_1 \qquad \hat{A} = \frac{n^2}{n^2-1}\left(I - \frac{2}{n+1} e_1 e_1^T\right)$$

For $x \in E \cap H$

$$\left(x + \frac{1}{n+1} e_1\right)^T \frac{n^2-1}{n^2} \cdot \left(I + \frac{2}{n-1} e_1 e_1^T\right)\left(x + \frac{1}{n+1} e_1\right)$$

$$= \frac{n^2-1}{n^2}\left(\|x\|^2 + x_1^2\left(\frac{2}{n-1}\right) + \frac{4x_1}{n^2-1} + \frac{2x_1}{n+1} + \frac{1}{(n+1)^2} + \frac{2}{(n+1)^2(n-1)}\right)$$

$$= \frac{n^2-1}{n^2}\left(1 + \frac{1}{n^2-1} + \frac{2x_1^2}{n-1} + \frac{x_1}{n^2-1}(4+2n-2)\right)$$

$$= \frac{n^2-1}{n^2}\left(\frac{n^2}{n^2-1} + \frac{2x_1(1+x_1)}{n-1}\right) \le 1 \quad \text{since } x_1(1+x_1) \le 0.$$

$x \le 0 \quad 1+x_1 \ge 0.$

$$\mathrm{vol}(E) = \sqrt{\det(A)} \cdot \mathrm{vol}(B^n)$$

$$\frac{\mathrm{vol}(\hat{E})}{\mathrm{vol}(E)} = \det\left(\frac{n^2}{n^2-1}\left(I - \frac{2}{n+1}e_1e_1^T\right)\right)^{\frac{1}{2}}$$

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

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

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

$1+x \le e^x$

$$\le e^{-\frac{1}{n+1} + \frac{1}{2(n+1)}} = e^{-\frac{1}{2n+2}}.$$

So, after T iterations, $\mathrm{vol}(E^{(T)}) \le e^{-\frac{T}{2n+2}} \mathrm{vol}(E^{(0)})$

If $E^{(0)} = B(0, R)$ then at most

If $E^{(0)} = B(0, R)$ then at most

$$O\left(n \ln\left(\frac{R^n}{\varepsilon^n}\right)\right) = O\left(n^2 \ln \frac{R}{\varepsilon}\right)$$

iterations to terminate.

Time per iteration $= O(n^2)$

What about function value?

$$\min_{x \in K} f(x) \iff \min f(x) + \delta_K(x) \qquad \delta_K(x) = \begin{cases} 0 & x \in K \\ \infty & x \notin K \end{cases}$$

convex!

For convex f, we can define a gradient oracle $\mathrm{GRAD}_f(x) = \{ v : \forall y,\ f(y) \le f(x) \Rightarrow v^T(y-x) \le 0 \}$.

If is differentiable, then $\{\nabla f(x)\} = \mathrm{GRAD}_f(x)$

_____ continuous, ___ subdifferential.

For $f = \delta_K$, $\mathrm{GRAD}_f(x)$ is defined for $x \in \partial K$.

and is any supporting hyperplane

for $x \notin K$, any separating plane.

For $x \notin K$, any separating plane (normal).

Thm. $E^{(0)}, \Omega \subseteq \mathbb{R}^n$. f convex, $\text{SEP}_f(\cdot) = \text{GRAD}_f(\cdot)$

Then after k iterations $x^* = \arg\min_\Omega f$.

$$\min_{i \in [k]} f(x^{(i)}) - f(x^*) \le \left(\frac{\text{Vol}(E^{(k)})}{\text{Vol}(\Omega)}\right)^{\frac{1}{n}} \left(\max_\Omega f(x) - \min_\Omega f(x)\right)$$

"optimality gap drops by a constant factor in $O(n^2)$ iterations."

Pf. Let $\alpha > \left(\frac{\text{Vol}(E^{(k)})}{\text{Vol}(\Omega)}\right)^{\frac{1}{n}}$. $S = (1-\alpha)x^* + \alpha\Omega$

$$\text{Vol}(S) = \alpha^n \text{Vol}(\Omega) > \text{Vol}(E^{(k)})^{\frac{1}{n}}.$$

$\therefore \exists y \in S \setminus E^{(k)}$, i.e. y is separated from $E^{(k)}$ by some query $x^{(i)}$. $\langle \nabla f(x^{(i)}), y - x^* \rangle \ge 0$

$$f((1-\alpha)x^* + \alpha z) \ge f(y) \ge f(x^{(i)}) \qquad z \in \Omega$$

$f((1-\alpha)x+\alpha z) \geq f(\cdot)$ ...

$$(1-\alpha)f(x^*)+\alpha f(z) \geq f(x^{(i)})$$

$$\alpha\left(f(z)-f(x^*)\right) \geq f(x^{(i)})-f(x^*). \quad \blacksquare$$

Thm. We can use any progress function $\nu \geq 0$ with the following properties:

$S \subseteq \mathbb{R}^n \quad x \in \mathbb{R}^n \quad \alpha \geq 0$

Linear (1) $\nu(x+\alpha S) = \alpha \nu(S)$

Monotone (2) $T \subseteq S \Rightarrow \nu(T) \leq \nu(S)$.

Then using Ellipsoid Method for convex f with SEP_f gives

$$\min_{i \in [k]} f(x^{(i)}) - f(x^*) \leq \left(\frac{\nu(E^{(k)})}{\nu(\Omega)}\right)^{\frac{1}{n}} \left(\max_{\Omega} f(x) - \min_{\Omega} f(x)\right).$$

$\nu(\cdot) = \mathrm{vol}(\cdot)^{\frac{1}{n}}$, width, meanwidth, diameter etc...

Cor. [Linear Prog.] To solve $\min_{x \in P = \{x : Ax \geq b\} \subseteq \mathbb{R}^n} c^T x$

with $R = \max_{x \in P} \|x\|_2$, $r = \mathrm{vol}(P)^{\frac{1}{n}}$, $\forall \varepsilon > 0$,

Ellipsoid method finds $x \in P$ s.t.

$c^T x - \min c^T x \leq \varepsilon \left(\max c^T x - \min c^T x\right)$

Ellipsoid method gives x s.t.

$$c^T x - \min_P c^T x \le \varepsilon \left(\max_P c^T x - \min_P c^T x \right)$$

in time $O\left(n^2 \left(n^2 + nnz(A)\right) \log\left(\frac{R}{r\varepsilon}\right)\right)$.

<u>Pf.</u> Set $f(x) = c^T x + \delta_P(x)$. $E^{(0)} = B(0, R)$.

$$SEP_f(x) = \begin{cases} c & \text{if } Ax \ge b \\ -a_i & \text{for any } i \text{ s.t. } a_i^T x < b \end{cases}.$$
