

$$\cancel{\text{why}} \quad \text{OPT} \rightarrow \text{SEP} \quad \checkmark$$

$$\text{SEP} \rightarrow \text{MEM}$$

$$\cancel{\text{x}} \quad \text{GRAD} \rightarrow \text{EVAL}$$

$$\nabla f(x)_i = \lim_{h \rightarrow 0} \frac{f(x + h e_i) - f(x)}{h} \quad O(n) \text{ calls.}$$

But what if f is not differentiable, e.g. $f = \delta_K$?

Thm. For an L -Lipschitz convex $f: B_n(0,1) \rightarrow \mathbb{R}$
 $\nabla^2 f(x)$ exists almost everywhere and
 $E_{B^n}(\|\nabla^2 f\|_F) \leq nL$.

Pf. Existence a.e. is a classical theorem of Lebesgue.

$$\int_{B^n} \|\nabla^2 f(x)\|_F dx \leq \int_{B^n} \operatorname{Tr}(\nabla^2 f(x)) dx = \int_{B^n} \sum_i \frac{\partial^2 f}{\partial x_i^2} dx$$

$$\left(\sqrt{\sum x_i^2} \leq \varepsilon x_i, x_i \geq 0 \right)$$

$$= \int_{\partial B^n} \langle \nabla f(x), \vec{n}(x) \rangle dx \leq \int_{\partial B^n} \|\nabla f(x)\| \|\vec{n}(x)\| dx$$

$$\leq L \cdot \operatorname{Vol}(\partial B^n)$$

$$\therefore \|\nabla^2 f(x)\|_F \leq L \cdot \underline{\operatorname{Vol}(\partial B^n)} = nL.$$

$$\therefore \mathbb{E}_{B^n} (\|\nabla^2 f(x)\|_F) \leq L \cdot \frac{\text{Vol}(\partial B^n)}{\text{Vol}(B^n)} = nL.$$

$$\left(\text{Vol}(B^n) = \int_0^1 r^{n-1} \text{Vol}(\partial B^n) dr = \frac{\text{Vol}(\partial B^n)}{n} \right). \quad \text{Diagram of concentric circles}$$

We use this idea to estimate the gradient of any Lipschitz convex function.

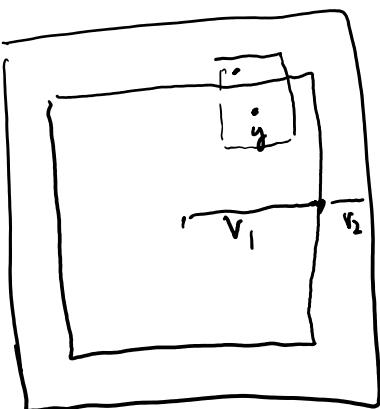
$$B_\infty(x, r) = \{y : \|x - y\|_\infty \leq r\}.$$

Lemma. $r_1 \geq r_2 > 0$. f is convex, L -Lipschitz, $f: B_\infty(x, r_1 + r_2) \rightarrow \mathbb{R}$

$$g(y) = \mathbb{E}_{\substack{x \sim B_\infty(y, r_2)}} (\nabla f(x)). \quad \|\nabla f(x)\|_\infty \leq L.$$

$$\left| \mathbb{E}_{\substack{y \sim B_\infty(x, r_1)}} \right| \mathbb{E}_{\substack{z \sim B_\infty(y, r_2)}} \|\nabla f(z) - g(y)\|_1 \leq n^{\frac{r_2}{r_1}} L$$

Pf.



$$\omega_i(z) = \langle \nabla f(z) - g(y), e_i \rangle$$

$$\int_{B_\infty(y, r_2)} \|\nabla f(z) - g(y)\|_1 dz = \sum_i \int_{B_\infty(y, r_2)} |\omega_i(z)| dz$$

$$\int_{B_\infty(y, r_2)} \omega_i(z) dz = 0.$$

Thm (L₁-Poincaré) $f: \Omega \rightarrow \mathbb{R}$

$$\left\| f - \frac{1}{\Omega} \int_\Omega f \right\|_{L^1} \leq \sup_{S \subseteq \Omega} \frac{2|S| |\Omega \setminus S|}{|\partial S| |\Omega|} \cdot \|\nabla f\|_L$$

$$\|g\|_{L_1} = \int_{\Omega} \|g\| .$$

$$\text{Applying this to } w_i, \quad \Omega = B_\infty(y, r_2)$$

$$\int |w_i(z)| \leq r_2 \int \|\nabla^2 f(z) e_i\|_2$$

$$\sum_i \int |w_i(z)| \leq r_2 \sqrt{n} \int \|\nabla^2 f\|_F$$

$$\leq r_2 \sqrt{n} \int \text{Tr}(\nabla^2 f)$$

$$|E_{B_\infty(y, r_2)} \|\nabla f(z) - g(y)\|_1| \leq r_2 \sqrt{n} |E_{B_\infty(y, r_2)} \sum_i \frac{\partial^2 f(x)}{\partial x_i^2}| = r_2 \sqrt{n} |E_{B_\infty(y, r_2)} \Delta f|$$

Let $h = \frac{1}{(2r_2)^n} f * \mathbf{1}_{B_\infty(0, r_2)}$. Then $|E_{B_\infty(y, r_2)} \Delta f| = \Delta h$.

$$\begin{aligned} \text{So, } & |E_{y \sim B_\infty(x, r_1)} |E_{B_\infty(y, r_2)} \|\nabla f(z) - g(y)\|_1| \leq r_2 \sqrt{n} \cdot |E_{B_\infty(x, r_1)} \Delta h| \\ &= r_2 \sqrt{n} \cdot \frac{1}{(2r_1)^n} \cdot \int_{\partial B_\infty(x, r_1)} \langle \nabla h(y), \vec{n}(y) \rangle dy \\ &\leq r_2 \sqrt{n} \cdot \frac{1}{(2r_1)^n} \cdot \int_{\partial B_\infty(x, r_1)} \|\nabla h(y)\|_\infty \|\vec{n}(y)\|_1 dy \\ &\leq r_2 \sqrt{n} \cdot \frac{1}{(2r_1)^n} L \cdot 1 \cdot (2n) (2r_1)^{n-1} \\ &= n^{3/2} \cdot \frac{r_2}{r_1} \cdot L \end{aligned}$$

Lemma. For $\Omega = \mathbb{D}^2$

$$\sup_{S \subseteq \Omega} \frac{2 |S| |\Omega \setminus S|}{|\partial S| |\Omega|} = r_2 \dots$$

Algo: SubGrad (f, x, r_1, ε)

INPUT $r_1 > 0$, $\|\nabla f(z)\|_\infty \leq L \quad \forall z \in B_\infty(x, 2r_1)$.

Set $r_2 = \sqrt{\frac{\varepsilon r_1}{\sqrt{n} L}}$

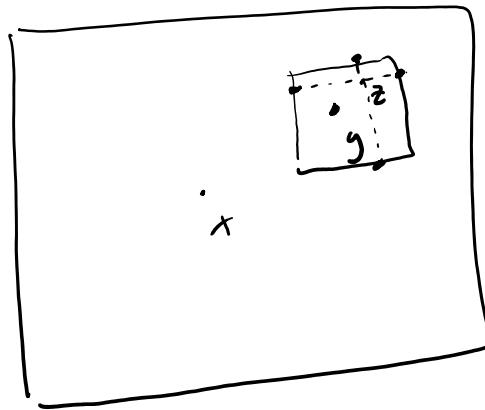
Pick random $y \in B_\infty(x, r_1)$ and $z \in B_\infty(y, r_2)$.

For $i: 1 \dots n$

Let α_i, β_i be endpoints of $B_\infty(y, r_2) \cap \{z + s e_i : s \in \mathbb{R}\}$

Set $\tilde{g}_i = \frac{f(\beta_i) - f(\alpha_i)}{2r_2}$

Output \tilde{g} .



Lemma. f convex, $\|\nabla f(z)\|_\infty \leq L \quad \forall z \in B_\infty(x, 2r_1)$.

$r_1 \wedge r_2 \geq \text{dist}(x, \partial \Omega) \geq r_1 / \sqrt{n} L$. $\tilde{g} = \text{SubGrad}(f, x, r_1, \varepsilon)$.

Lemma- EVAL(f) with error $\epsilon \leq r_1 \sqrt{n} L$. $\tilde{g} = \text{SubGrad}(f, x, r_1, \epsilon)$.
 \exists random variable $\xi \geq 0$ with $E\xi \leq 2 \sqrt{\frac{L\epsilon}{r_1}} n^{3/4}$ s.t.

$$f(q) \geq f(x) + \langle \tilde{g}, q - x \rangle - \xi \|q - x\|_\infty - 4nr_1L$$

Approximate Gradient Oracle!

Pf. Assume f twice-differentiable by viewing as limit
of twice-differentiable functions.

$$g(y) = E_{B_{20}(y, r_2)} (\nabla f)$$

$$\begin{aligned} |E_z | \tilde{g}_i - g(y)_i | &= E_z \left| \frac{f(\beta_i) - f(d_i)}{2r_2} - g(u)_i \right| \\ &\leq E_z \frac{1}{2r_2} \int \left| \frac{df}{dx_i} (z + \delta e_i) - g(y)_i \right| \\ &\quad z + \delta e_i \in B_{20}(y, r_2) \\ &= E_z \left| \frac{df}{dx_i} (z) - g(y)_i \right| \end{aligned}$$

$z, z + \delta e_i$ are both uniform in $B_{20}(y, r_2)$.

$$\begin{aligned} E_z \|\nabla f(z) - \tilde{g}\|_1 &\leq E \|\nabla f(z) - g(y)\|_1 + E \|g - g(y)\|_1 \\ &\leq 2 E \|\nabla f(z) - g(y)\|_1 \left(\leq 2 n^{3/2} \cdot \frac{r_2}{r_1} L \right. \\ &\quad \left. \text{taking } E_y \text{ as well} \right) \end{aligned}$$

Hence, by convexity of f

$\forall z$,

$$f(q) \geq f(z) + \langle \nabla f(z), q - z \rangle$$

$$\begin{aligned} &\geq f(x) - \|\nabla f(z)\|_\infty \|x - z\|_1 + \langle \tilde{g}, q - x \rangle \\ &\quad + \langle \nabla f(z) - \tilde{g}, q - x \rangle + \langle \nabla f(z), x - z \rangle \end{aligned}$$

$$f(q) \geq f(x) + \langle \tilde{g}, q - x \rangle - \xi \|q - x\|_\infty - L r_1 n$$

$$(\|x - z\|_1 \leq (r_1 + r_2) n \leq 2r_1 n)$$

$$\text{where } |\xi| \leq 2n^{3/2} \frac{r_2}{r_1} L.$$

If f is evaluated only to $+\epsilon$ error, then error in \tilde{g}_i can be $+\frac{\epsilon}{2r_2}$, so $|\xi| \leq 2n^{3/2} \frac{r_2}{r_1} L + \frac{\epsilon n}{2r_2}$

$$r_2 = \frac{1}{2} \sqrt{\frac{\epsilon r_1}{n L}} < r_1$$

Using this we will design an (approximate) SEP oracle for K , given an (approximate) MEM oracle for K .

$$\text{MEM}_K^S : \begin{cases} \text{YES} & x \in B(K, \delta) \quad \exists y \in K : \|x - y\| \leq \delta \\ \text{NO} & x \notin B(K, \delta) \quad \forall y \notin K : \|x - y\| \leq \delta \end{cases}$$

$$\text{SEP}_K^\delta : \begin{cases} \text{YES} & " \\ \text{NO} & \bar{A}^T u < \bar{A}^T \bar{u} + \delta \quad \text{if } x \in B(K, \delta). \end{cases}$$

$$\text{SEP}_K : \quad \begin{cases} \text{No, } \theta : \quad \theta^T x \leq \theta^T y + \delta \quad \forall x \in B(K, -\delta). \end{cases}$$

Thm $B(0, r) \subseteq K \subseteq B(0, R)$. For any $\eta \in [0, \frac{1}{2})$

$$\text{SEP}_K^\eta = O\left(n \log\left(\frac{nR}{\eta r}\right)\right) \text{MEM}_K^{(nR/\eta r)^c}.$$

Cor. $B(0, r) \subseteq K \subseteq B(0, R)$. Convex f , $\epsilon > 0$.
 MEM EVAL .

Then we can find $z \in B(K, \epsilon)$ s.t.

$$f(z) \leq \min_K f(x) + \epsilon \left(\max_K f(x) - \min_K f(x) \right)$$

with prob 0.99, using $O(n^2 \log^2 \frac{nR}{\epsilon r})$ calls to MEM/EVAL

and $O(n^3 \log^{O(1)} \frac{nR}{\epsilon r})$ arithmetic operations in total.

We need a Lipschitz convex f whose GRAD is
 SEP for K . δ_K doesn't work, not Lipschitz!

Given $x \notin K$, define

$$h_x : K \rightarrow \mathbb{R} \quad h_x(y) = -\max \{\alpha \mid y + \alpha x \in K\}$$



Lema 1. h_x is convex.

Pf.

$$y_1, y_2 \quad y_1 + \alpha_1 x \in K$$

$$t \in [0,1]$$

$$y_2 + \alpha_2 x \in K$$

$$h_x(ty_1 + (1-t)y_2) - \max_{t \in [0,1]} \alpha : \\ ty_1 + (1-t)y_2 + \alpha x \in K$$

$$\alpha \geq t\alpha_1 + (1-t)\alpha_2$$

$$\text{So } h_x(ty_1 + (1-t)y_2) \leq t h_x(y_1) + (1-t) h_x(y_2)$$

Lema 2. If $\delta < r$ h_x is $\left(\frac{R+\delta}{r-\delta}\right)$ -Lipschitz over $B(0, \delta)$.

Pf. Need to show that for $y_1, y_2 \in B(0, \delta)$

$$|h_x(y_1) - h_x(y_2)| \leq L \|y_1 - y_2\|.$$

assume $\alpha(y_1) > \alpha(y_2)$.

$$\text{So } |h_x(y_1) - h_x(y_2)| = (\alpha(y_1) - \alpha(y_2)) \|x\|_2$$

(1) if $\|y_1 - y_2\| \geq r - \delta$, and $0 \geq h_x \geq -R - \delta$

$$|h_x(y_1) - h_x(y_2)| \leq R + \delta \leq \left(\frac{R+\delta}{r-\delta}\right) \cdot \|y_1 - y_2\|.$$

(2) Else $\|y_1 - y_2\| < r - \delta$

$$\text{let } y_3 = y_1 + \frac{y_2 - y_1}{\lambda} \quad \lambda = \frac{\|y_2 - y_1\|}{r - \delta}$$

$$\|y_3\| \leq \|y_1\| + r - \delta \leq r \Rightarrow y_3 \in K.$$

$$\lambda y_3 + (1-\lambda)(y_1 + \alpha(y_1)x) \in K$$

$$y_2 + (1-\lambda)\alpha(y_1)x \in K \Rightarrow \alpha(y_2) \geq (1-\lambda)\alpha(y_1) \\ \geq \left(1 - \frac{\|y_2 - y_1\|}{r - \delta}\right) \alpha(y_1)$$

$$|h_x(y_1) - h_x(y_2)| = (\alpha(y_1) - \alpha(y_2)) \|x\|_2$$

$$\leq \alpha(y_1) \frac{\|y_1 - y_2\|}{r - \delta} \|x\|_2$$

$$\leq \frac{R + \delta}{r - \delta} \cdot \|y_1 - y_2\|.$$

Algo. $x \in B(K, \varepsilon) \rightarrow \text{YES} \checkmark$

Algo. $x \in B(K, \varepsilon) \rightarrow \text{YES}^*$
 $x \notin B(0, R) - \text{NO}, \quad \langle y-x, x \rangle \leq 0.$

else SubGrad($h_x, 0, r_1, 4\varepsilon$) $\rightarrow \tilde{g}$
report $_{H_x}: \langle \tilde{g}, y-x \rangle \leq C \cdot n^{\frac{3}{6}} \varepsilon^{\frac{1}{3}} \frac{r}{R^{\frac{1}{3}}}.$

Lemma: $K \subseteq H_x.$

Pf. set $\delta = r_2$ h_x is $\frac{R+r_2}{r_2} \leq \frac{3R}{r}$ - Lipschitz.
 $3K \quad \left(\frac{R}{r} = K\right)$
over $B(0, \frac{r}{2})$.

We will set r_1 small enough st. $B_\infty(0, 2r_1) \subseteq B(0, \frac{r}{2}).$

Then $h_x(y) \geq h_x(0) + \langle \tilde{g}, y \rangle - \xi \|y\|_\infty - 12nr_1K \quad (*)$

$\forall y \in K.$

since $\|x\| \leq R, \quad -\frac{x}{K} \in B(0, r) \subseteq K$

and $h_x\left(-\frac{x}{K}\right) = h_x(0) - \frac{1}{K} \|x\|_2$

$\geq h_x(0) + \langle \tilde{g}, -\frac{x}{K} \rangle - \frac{1}{K} \xi \|x\|_\infty - 12nr_1K.$

$\therefore \langle \tilde{g}, x \rangle \geq \|x\|_2 - \xi \|x\|_\infty - 24nr_1K^2.$

$$\cdots \langle \tilde{g}, x \rangle \geq \|x\|_2 - \xi \|x\|_\infty - 24nr_1K.$$

$$\overline{x \notin B(k, -\varepsilon)}, \quad B(0, r) \subseteq K \Rightarrow \left(1 - \frac{\varepsilon}{r}\right)K \subset B(k, -\varepsilon)$$

$$h_x(0) \geq -\left(1 - \frac{\varepsilon}{r}\right)\|x\|_2 \geq -\|x\|_2$$

$$\text{so } h_x(0) + \langle \tilde{g}, x \rangle \geq -\xi \|x\|_\infty - 12nr_1k^2$$

$$(*) \Rightarrow 0 \geq h_x(y) \geq \langle \tilde{g}, y-x \rangle - 2\xi R - 24nr_1k^2$$

$$\therefore \langle \tilde{g}, y-x \rangle \leq 2\xi R + 24nr_1k^2$$

$$\mathbb{E}(\xi) \leq 2 \sqrt{\frac{3K\varepsilon}{r_1}} n^{5/4}. \quad r_1 = n^{1/6} \varepsilon^{1/3} R^{2/3}, \quad \varepsilon \leq r.$$

$$\mathbb{E}(\text{RHS}) \leq 31 n^{7/6} R^{2/3} \varepsilon^{1/3} k.$$

(1)

(OP). Deterministic Algorithm?