Lecture 6

Thursday, March 18, 2021 9:54 AM

Properties of the logarithmic barrier function

$$P = \{ x \in \mathbb{R}^n : \langle a_i, x \rangle \le b_i \ \forall\ i \in [m] \}$$

$$F : \operatorname{int}(P) \to \mathbb{R}, \quad F(x) := -\sum_{i=1}^{m} \log(b_i - \langle a_i, x \rangle)$$

Lemma: (a) F satisfies the NL condition with $\delta_0 = 1/6$.

(b) $\langle \nabla F(x), y - x \rangle \le m \quad \forall\ x \in \operatorname{int}(P)\ \&\ y \in P$

(c) $\|N(x)\|_x := \|H_F(x)^{-1} \nabla F(x)\|_x = \|\nabla F(x)\|_{H_F(x)^{-1}} \le \sqrt{m}$

$\forall\ x \in \operatorname{int}(P)$.

Proof: (a) $s_i(x) := b_i - \langle a_i, x \rangle$

$$\nabla F(x) = \sum_{i=1}^{m} \frac{a_i}{s_i(x)}, \quad H_F(x) = \sum_{i=1}^{m} \frac{a_i a_i^T}{s_i(x)^2}$$

Now suppose $\|y - x\|_x \le \delta \le 1/6$, want to prove that

$$(1-3\delta) H_F(x) \preceq H_F(y) \preceq (1+3\delta) H_F(x)$$

$$\|y - x\|_x^2 = \langle y - x, H_F(x)\, y - x \rangle = \sum_{i=1}^{m} \frac{\langle a_i, y - x \rangle^2}{s_i(x)^2}$$

$$= \sum_i \frac{(s_i(x) - s_i(y))^2}{s_i(x)^2}$$

we have that $\displaystyle\sum_i \frac{(s_i(x) - s_i(y))^2}{s_i(x)^2} \le \delta^2$

$$\Rightarrow \forall\ i \quad \left| \frac{s_i(x) - s_i(y)}{s_i(x)} \right| \le \delta$$

$a_i a_i^T$ is PSD

$$\Rightarrow \frac{s_i(y)}{s_i(x)} \in [1-\delta, 1+\delta]$$

$$H_F(y) = \sum_i \frac{a_i a_i^T}{s_i(y)^2} \succeq \left(\sum_i \frac{a_i a_i^T}{s_i(x)^2} \right) (1+\delta)^{-2}$$

$$\&\ \text{also} \preceq \left(\sum_i \frac{a_i a_i^T}{s_i(x)^2} \right) (1-\delta)^{-2}$$

hence $(1-3\delta) H_F(x) \preceq H_F(y) \preceq (1+3\delta) H_F(x)$

& hence $(1-3\delta) H_F(x) \preceq H_F(y) \preceq (1+3\delta) H_F(x)$

$$\left((1+\delta)^{-2} \geq 1-3\delta \quad \forall\, \delta \leq 1/6 \right.$$
$$\left. \& \quad (1-\delta)^{-2} \leq 1+3\delta \quad \forall\, \delta \leq 1/6 \right)$$

(b) <u>To prove:</u> $\langle \nabla F(x), y-x \rangle \leq m \quad \forall\, x \in \text{int}(P),\ y \in P.$

$$\nabla F(x) = \sum_i \frac{a_i}{s_i(x)}, \quad \langle \nabla F(x), y-x \rangle = \sum_i \frac{\langle a_i, y-x \rangle}{s_i(x)}$$

$$= \sum_i \frac{s_i(x) - s_i(y)}{s_i(x)}$$

$$\leq m \quad \text{because } s_i(y) \geq 0 \ \forall\, i \text{ since } y \in P$$

(c) <u>To prove:</u> $\| N(x) \|_x \leq \sqrt{m} \quad \forall\, x \in \text{int}(P)$

Consider $z := H_F(x)^{-1} \nabla F(x) \qquad (z = -N(x))$

$$\langle z, \nabla F(x) \rangle = \sum_i \frac{\langle a_i, z \rangle}{s_i(x)}$$

Cauchy-Schwarz

$$\leq \sqrt{m} \sqrt{\sum_i \frac{\langle a_i, z \rangle^2}{s_i(x)^2}}$$

$$= \sqrt{m} \sqrt{\langle z, H_F(x) z \rangle}$$

$$= \sqrt{m} \sqrt{\langle z, \nabla F(x) \rangle}$$

Hence $\langle z, \nabla F(x) \rangle \leq m$

Now note that $\langle z, \nabla F(x) \rangle = \| N(x) \|_x^2$

$$\Rightarrow \| N(x) \|_x \leq \sqrt{m} \quad \forall\, x \in \text{int}(P)$$

recall: $N_{\eta_t}(x_t) := -H_F(x_t)^{-1} \nabla f_{\eta_t}(x_t)$

recall: $N_{\eta_t}(x_t) := -H_F(x_t)^{-1} \nabla f_{\eta_t}(x_t)$

Effect of Newton step & updating η_t

At the start of $(t+1)^{th}$ iteration, we have $\|N_{\eta_t}(x_t)\|_{x_t} \le 1/6$

Then we update:
$$x_{t+1} := x_t + N_{\eta_t}(x_t)$$
$$\eta_{t+1} := \eta_t \left(1 + \tfrac{1}{20\sqrt{m}}\right)$$

& we want to maintain that $\|N_{\eta_{t+1}}(x_{t+1})\|_{x_{t+1}} \le 1/6$

$$\|N_{\eta_{t+1}}(x_{t+1})\|_{x_{t+1}} = \left\| H_F(x_{t+1})^{-1} \left(\eta_{t+1} c + \nabla F(x_{t+1})\right)\right\|_{x_{t+1}}$$

$$= \left\| \frac{\eta_{t+1}}{\eta_t} H_F(x_{t+1})^{-1} \left(\eta_t c + \nabla F(x_{t+1})\right) + \left(1 - \frac{\eta_{t+1}}{\eta_t}\right) H_F(x_{t+1})^{-1} \nabla F(x_{t+1}) \right\|_{x_{t+1}}$$

triangle inequality for $\|\cdot\|_{x_{t+1}}$

$$\le \frac{\eta_{t+1}}{\eta_t} \|N_{\eta_t}(x_{t+1})\|_{x_{t+1}} + \left(\frac{\eta_{t+1}}{\eta_t} - 1\right) \|H_F(x_{t+1})^{-1} \nabla F(x_{t+1})\|_{x_{t+1}}$$

$$\le \frac{3\eta_{t+1}}{\eta_t} \|N_{\eta_t}(x_t)\|_{x_t}^2 + \frac{1}{20\sqrt{m}} \cdot \sqrt{m}$$

(by analysis of Newton's iteration & property (c) of the logarithmic barrier function)

$$\le 3\left(1 + \frac{1}{20\sqrt{m}}\right) \cdot \left(\frac{1}{6}\right)^2 + \frac{1}{20}$$

$$\le 1/6$$

Termination:

(1) $\hat{x} :=$ Two Newton steps on x_T w.r.t. f_{η_T}

(2) $\|N_{\eta_T}(x_T)\|_{x_T} \le 1/6$

Termination: ① $\hat{x} :=$ two Newton steps on x_T w.r.t. f_{η_T}

② $\| N_{\eta_T}(x_T) \|_{x_T} \le 1/6$

③ $\eta_T \geq m/\varepsilon$

Want to prove: $\langle c, \hat{x} \rangle - \langle c, x^* \rangle \le 2\varepsilon$

Claim: $\langle c, x^*_{\eta_T} \rangle - \langle c, x^* \rangle \le \varepsilon$

Pf: $\nabla f_{\eta_T}(x^*_{\eta_T}) = 0$ since $x^*_{\eta_T}$ is the optimum for f_{η_T} (and also P is full-dimensional)

$$\eta_T c + \nabla F(x^*_{\eta_T}) \implies c = -\frac{1}{\eta_T} \nabla F(x^*_{\eta_T})$$

$$\implies \langle c, x^*_{\eta_T} \rangle - \langle c, x^* \rangle = \frac{1}{\eta_T} \langle \nabla F(x^*_{\eta_T}), x^* - x^*_{\eta_T} \rangle$$

$$\le \frac{m}{\eta_T} \quad \text{(property (b) of the barrier function)}$$

$$\le \varepsilon$$

Claim: For any $\eta > 0$, $\langle c, x \rangle - \langle c, x^* \rangle \le \dfrac{m}{\eta(1 - \|x - x^*_\eta\|_x)}$ (if $\|x - x^*_\eta\|_x < 1$)

Pf: suffices to prove

$$\langle c, x \rangle - \langle c, x^* \rangle \le \frac{\langle c, x^*_\eta \rangle - \langle c, x^* \rangle}{(1 - \|x - x^*_\eta\|_x)}$$

in fact we will prove something stronger i.e.

$\forall\, x, y \in \text{int}(P)$

$$\langle c, x \rangle - \langle c, x^* \rangle \le \frac{\langle c, y \rangle - \langle c, x^* \rangle}{(1 - \|x - y\|_x)}$$

$$\langle c, x - y \rangle = \langle H_F(x)^{-1/2} c, H_F(x)^{1/2}(x - y) \rangle$$

$$< \|H_F(x)^{-1/2} c\|_2 \, \|H_F(x)^{1/2}(x - y)\|_2$$

$$\leq \| H_F(x)^{-1/2} c \|_2 \; \| H_F(x)^{1/2} (x-y) \|_2$$

exercise

$$= \| H_F(x)^{-1} c \|_x \; \| x - y \|_x$$

Proposition: $E_x := \{ y : \underbrace{\langle y - x, H_F(x) \, y - x \rangle}_{\| y - x \|_x^2} \leq 1 \}$

Dikin's Ellipsoid $\Rightarrow E_x \subseteq P$

continuing the proof $\Rightarrow \; x - \dfrac{H_F(x)^{-1} c}{\| H_F(x)^{-1} c \|_x} \in P$

$$\Rightarrow \left\langle c, x - \frac{H_F(x)^{-1} c}{\| H_F(x)^{-1} c \|_x} \right\rangle \geq \langle c, x^* \rangle$$

Rearranging $\langle c, x \rangle - \langle c, x^* \rangle \geq \dfrac{\langle c, H_F(x)^{-1} c \rangle}{\| H_F(x)^{-1} c \|_x}$

$$= \| H_F(x)^{-1} c \|_x$$

Hence $\langle c, x - y \rangle \leq (\langle c, x \rangle - \langle c, x^* \rangle) \, \| y - x \|_x$

$$= (\langle c, x - y \rangle + \langle c, y \rangle - \langle c, x^* \rangle) \, \| y - x \|_x$$

$$\Rightarrow \langle c, x - y \rangle \leq \frac{(\langle c, y \rangle - \langle c, x^* \rangle) \, \| y - x \|_x}{1 - \| y - x \|_x}$$

$$\leq \frac{\langle c, y \rangle - \langle c, x^* \rangle}{1 - \| y - x \|_x} \quad \text{if } \| y - x \|_x < 1$$

Combining the two claims, we get that

$$\langle c, \hat{x} \rangle - \langle c, x^* \rangle \leq \frac{\varepsilon}{1 - \| \hat{x} - x^*_{\eta_T} \|_{\hat{x}}}$$

Lemma: Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a strictly convex function that satisfies the NL condition with $\delta_0 = 1/6$.

If $\|N_f(x)\|_x \leq 1/24 \Rightarrow \|x - x^*\|_x \leq 4\,\|N(x)\|_x$

Continuing the analysis of the termination.

$\hat{x} :=$ Two Newton steps on x_{η_T}.

$$\|N(x_{\eta_T})\|_{x_{\eta_T}} \leq 1/6 \Rightarrow \|N(\hat{x})\|_{\hat{x}} \leq \frac{1}{48}$$

Using the above lemma $\|\hat{x} - x^*_{\eta_T}\|_{\hat{x}} \leq 1/12$

& then plugging this in $\langle c, \hat{x}\rangle - \langle c, x^*\rangle \leq \dfrac{\varepsilon}{1 - \|\hat{x} - x^*_{\eta_T}\|_{\hat{x}}}$

gives us $\langle c, \hat{x}\rangle - \langle c, x^*\rangle \leq 2\varepsilon$

Proof of above lemma: Let h be s.t. $\|h\|_x \leq 1/6$

By the mean value theorem,

$$f(x+h) = f(x) + \langle \nabla f(x), h\rangle + \frac{1}{2}\langle h, H_f(\theta) h\rangle$$

for some θ lying the line connecting x & $x+h$

$$|\langle \nabla f(x), h\rangle| = \left|\left\langle H_f(x)^{-1/2}\nabla f(x),\ H_f(x)^{1/2} h\right\rangle\right|$$

$$\overset{\text{Cauchy-Schwarz}}{\leq} \|H_f(x)^{-1/2}\nabla f(x)\|_2\ \|H_f(x)^{1/2} h\|_x$$

$$= \|N_f(x)\|_x\ \|h\|_x$$

By the NL property, $H_f(\theta) \succcurlyeq \left(1 - 3\cdot\frac{1}{6}\right) H_f(x)$

$$\Rightarrow \langle h, H_f(\theta) h\rangle \geq \frac{1}{2}\langle h, H_f(x) h\rangle$$

$$\Rightarrow \boxed{f(x+h) \geq f(x) - \|N_f(x)\|_x\,\|h\|_x + \frac{1}{4}\langle h, H_f(x) h\rangle}$$

$$\Rightarrow \boxed{f(x+h) \geq f(x) - \|N_f(x)\|_x \|h\|_x + \frac{1}{4}\langle h, H_f(x) h\rangle}$$

Suppose y is s.t. $\|y\|_x = 4\|N_f(x)\|_x$

$\Rightarrow \|y\|_x \leq 1/6$ & hence

$$f(x+y) \geq f(x) - \|N_f(x)\|_x \|y\|_x + \frac{1}{4}\langle y, H_f(x) y\rangle$$

$$= f(x) - 4\|N_f(x)\|_x^2 + \frac{1}{4}\|y\|_x^2$$

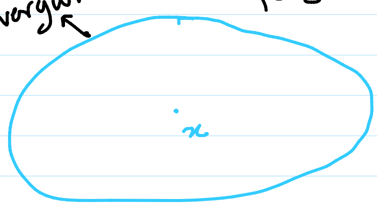
$$= f(x) - 4\|N_f(x)\|_x^2 + 4\|N_f(x)\|_x^2$$

$$= f(x).$$

So we get that $\forall\, y$ s.t. $\|y\|_x = 4\|N_f(x)\|_x$

$$f(x+y) \geq f(x)$$

everywhere on the boundary $f(x+y) \geq f(x)$



Since f is strictly convex, the unique min. of f should be achieved inside this blue region.

$$\Rightarrow \|x - x^*\|_x \leq 4\|N_f(x)\|_x$$

Next lecture: Initialization i.e. how to find some initial $\eta_0 > 0$ & x_0 s.t. $\|N_{\eta_0}(x_0)\|_{x_0} \leq 1/6$.