

Yuhao

$$\min_{x \in K} c^T x \rightarrow \min t c^T x + \phi(x)$$

$\phi(x)$: convex

barrier for K

self-concordant $|\nabla^3 \phi(x)_{[h,h,h]}| \leq 2 (\nabla^2 \phi(x)_{[h,h]})^{3/2}$

$$(g(t) = \phi(x+th) \quad |g'''(t)| \leq 2 (g''(t))^{3/2})$$

$$\frac{\|\nabla \phi(x)\|^2}{\nabla^2 \phi(x)} \leq 2 \quad \leftarrow \text{barrier parameter.}$$

$$\phi(x) = -\ln x$$

$$-\sum \ln x_i, \quad -\sum_i \ln (b_i - a_i^T x), \quad -\ln \det X$$

Easy to optimize by Newton Iteration.

Lemma 1. For self-concordant f ,

$$\text{if } \frac{\|\nabla f(x)\|}{\nabla^2 f(x)^{-1}} = r < 1, \text{ then } \hat{x} \leftarrow x - \nabla^2 f(x)^{-1} \nabla f(x)$$

has $\frac{\|\nabla f(\hat{x})\|}{\nabla^2 f(\hat{x})^{-1}} \leq \frac{r^2}{(1-r)^2}$.

Note: if ϕ has barrier parameter γ

then $\frac{\phi}{2} \geq \frac{\gamma}{2}$

Proof uses the following crucial observation that Hessian of a self-concordant function changes slowly.

Lemma 2 If x, y $\|x-y\|_x, \|x-y\|_y < 1$,

$$(1 - \|x-y\|_x^2) \nabla^2 f(x) \preceq \nabla^2 f(y) \preceq \frac{\nabla^2 f(x)}{(1 - \|x-y\|_x^2)}$$

Pf. Let $\alpha(t) = \langle \nabla^2 f(x + t(y-x)) u, u \rangle$.

$$\alpha'(t) = D^3 f(x + t(y-x)) [y-x, u, u]$$

By S.C. $|\alpha'(t)| \leq 2\|y-x\|_{x+t(y-x)} \|u\|_{x+t(y-x)}^2$ (*)

For $u = y-x$

$$|\alpha'(t)| \leq 2 \alpha(t)^{3/2}$$

$$\frac{d}{dt} \left(\frac{1}{\sqrt{\alpha(t)}} \right) \geq -1 \quad \left(\frac{d}{dt} \frac{1}{\sqrt{x}} = \frac{1}{2} \frac{1}{x^{3/2}} \right)$$

$$\implies \frac{1}{\sqrt{\alpha(t)}} \geq -t = \frac{1}{\|u\|} - t$$

$$\frac{1}{\sqrt{\alpha(t)}} \geq \frac{1}{\sqrt{\alpha(0)}} - t = \frac{1}{\|y-x\|_x} - t$$

$$\frac{\|y-x\|}{x+t(y-x)} \leq \frac{\|y-x\|_x}{(1-t)\|y-x\|_x}$$

$$\therefore (*) \Rightarrow |\alpha'(t)| \leq \frac{2\|y-x\|_x}{(1-t)\|y-x\|_x} \alpha(t)$$

$$\left| \frac{d}{dt} \ln \alpha(t) \right| \leq \frac{2\|y-x\|_x}{(1-t)\|y-x\|_x} = \frac{d}{dt} (-2\ln(1-t\|y-x\|_x))$$

Integrating from $t=0$ to 1:

$$(1-\|y-x\|_x)^2 \leq \frac{\alpha(t)}{\alpha(0)} \leq (1-\|y-x\|_x)^2$$

Pf. (L1). By L2, $\|\nabla f(\hat{x})\|_{\nabla^2 f(\hat{x})^{-1}} \leq \frac{\|\nabla f(\hat{x})\| \nabla^2 f(x)^{-1}}{(1-r)}$

$$\begin{aligned} \text{Since } \|\hat{x}-x\|_x^2 &= (\nabla^2 f(x)^{-1} \nabla f(x))^T \nabla^2 f(x) \nabla^2 f(x)^{-1} \nabla f(x) \\ &= \|\nabla f(x)\|_{\nabla^2 f(x)^{-1}}^2 = r^2 \end{aligned}$$

we have $\nabla^2 f(\hat{x})^{-1} \preceq (1-r)^2 \nabla^2 f(x)^{-1}$.

$$\nabla f(\hat{x}) = \nabla f(x) + \int_0^1 \nabla^2 f(x+t(\hat{x}-x))(\hat{x}-x) dt$$

$$\begin{aligned}
&= \nabla f(x) - \int_0^1 \nabla^2 f(x + t(\hat{x} - x)) \nabla^2 f(x)^{-1} \nabla f(x) dt \\
&= \left(\nabla^2 f(x) - \int_0^1 \nabla^2 f(x + t(\hat{x} - x)) dt \right) \nabla^2 f(x)^{-1} \nabla f(x) \\
\nabla^2 f(x) \int_0^1 (1-t)^2 dt &\lesssim \int_0^1 \nabla^2 f(x + t(\hat{x} - x)) dt \lesssim \nabla^2 f(x) \int_0^1 \frac{dt}{(1-t)^2} \\
&\lesssim \nabla^2 f(x) \left(1 - r + \frac{r^2}{3}\right) \lesssim \nabla^2 f(x) \cdot \frac{1}{1-r} \\
\therefore \left\| \nabla^2 f(x)^{-1/2} \left(\nabla^2 f(x) - \int_0^1 \nabla^2 f(x + t(\hat{x} - x)) dt \right) \nabla^2 f(x)^{-1/2} \right\|_{op} &\leq \max \left\{ \frac{r}{1-r}, r - \frac{r^2}{3} \right\} \\
&= \frac{r}{1-r}
\end{aligned}$$

$$\begin{aligned}
\left\| \nabla f(\hat{x}) \right\|_{\nabla^2 f(x)^{-1}} &= \left\| \nabla^2 f(x)^{-1/2} \nabla f(\hat{x}) \right\|_2 \\
&= \left\| \nabla^2 f(x)^{-1/2} \left(\nabla^2 f(x) - \int_0^1 \nabla^2 f(x + t(\hat{x} - x)) dt \right) \nabla^2 f(x)^{-1/2} \nabla^2 f(x)^{-1/2} \nabla f(x) \right\|_2 \\
&\leq \left\| \nabla^2 f(x)^{-1/2} \left(\quad \right) \nabla^2 f(x)^{-1/2} \right\|_{op} \left\| \nabla^2 f(x)^{-1/2} \nabla f(x) \right\|_2 \\
&\leq \frac{r}{1-r} \cdot r = \frac{r^2}{1-r}.
\end{aligned}$$

Our goal is to reach $\nabla f(x) \rightarrow 0$.
 We can measure progress in terms of $\|\nabla f(x)\|$.

We can measure progress in terms "to"

Lemma 3. Suppose $\|\nabla f(x)\|_{\nabla^2 f(x)^{-1}} \leq \frac{1}{6}$. Then

$$\|x^* - x\|_x \leq \frac{4}{3} \|\nabla f(x)\|_{\nabla^2 f(x)^{-1}}$$

$$\|x^* - x\|_{x^*} \leq 2 \|\nabla f(x)\|_{\nabla^2 f(x)^{-1}}$$

$$f(x) - f(x^*) \leq \|\nabla f(x)\|_{\nabla^2 f(x)^{-1}}^2$$

Pf. Let $r = \|x^* - x\|_x$.

$$\nabla f(x) = \nabla f(x^*) + \int_0^1 \nabla^2 f(x^* + t(x-x^*)) (x-x^*) dt \\ (= 0)$$

$$\begin{aligned} \|\nabla f(x)\|_{\nabla^2 f(x)^{-1}} &= \left\| \int_0^1 \nabla^2 f(x^* + t(x-x^*)) (x-x^*) dt \right\|_{\nabla^2 f(x)^{-1}} \\ &\geq \left\| \nabla^2 f(0)(x-x^*) \int_0^1 (1-(1-t)r)^2 dt \right\|_{\nabla^2 f(x)^{-1}} \\ &= \int_0^1 (1-(1-t)r)^2 dt \|x-x^*\|_x = \left(1-r+\frac{r^2}{3}\right)r \end{aligned}$$

$$\Rightarrow r \leq \frac{4}{3} \|\nabla f(x)\|_{\nabla^2 f(x)^{-1}} \leq \frac{4}{3} \cdot \frac{1}{6} < \frac{1}{4}$$

$$\|x-x^*\|_{x^*} \leq \frac{\|x-x^*\|_x}{1-r} \leq \frac{4}{3} \frac{\|\nabla f(x)\|_{\nabla^2 f(x)^{-1}}}{\left(r-\frac{1}{4}\right)} < 2 \|\nabla f(x)\|_{\nabla^2 f(x)^{-1}}$$

Finally,

$$\int_0^1 \nabla^2 f(x^* + t(x-x^*)) (x-x^*) dt$$

Finally

$$\begin{aligned} f(x) &= f(x^*) + \langle \nabla f(x^*), x - x^* \rangle + \int_0^1 (1-t)(x-x^*)^\top \nabla^2 f(x^* + t(x-x^*)) (x-x^*) dt \\ f(x) - f(x^*) &\leq \|x-x^*\|_x^2 \int_0^1 \frac{1-t}{(1-(1-t)\gamma)^2} dt \\ &= \|x-x^*\|_x^2 \cdot \frac{1}{\gamma^2} \left(\frac{\gamma}{1-\gamma} + \ln(1-\gamma) \right) \\ &\leq \left(\frac{1}{2} + \gamma \right) \|x-x^*\|_x^2 \leq \frac{\|\nabla f(x)\|_{\nabla^2 f(x)}^2}{\nabla^2 f(x)^{-1}}. \end{aligned}$$

IPM algo. so barrier $\phi(x)$, γ

start at $x_0 = \arg\min \phi(x)$.

$$\text{Min } f_t(x) = t c^\top x + \phi(x)$$

$$t_0 = \frac{1}{6} \|c\|_{\nabla^2 \phi(x_0)^{-1}}^{-1}$$

$$\text{while } t \leq \frac{\gamma + \sqrt{2}}{\epsilon}$$

$$\begin{cases} x \leftarrow x - \nabla^2 f_t(x)^{-1} \nabla f_t(x) \\ t \leftarrow (1+h)t \quad h = \frac{1}{9\sqrt{2}} \end{cases}$$

Lema. $\phi: K \rightarrow \mathbb{R}$ \mathcal{V} -sc.

$\forall x, y \in K \quad \langle \nabla \phi(x), y - x \rangle \leq \mathcal{V}.$

Pf. Let $\alpha(t) = \langle \nabla \phi(x + t(y-x)), y - x \rangle$

$$\alpha'(t) = \|y-x\|^2_{x+t(y-x)}$$

$$|\alpha(t)| \leq \|\nabla \phi(x + t(y-x))\|_{\nabla \phi(x + t(y-x))} \|y-x\|_{x+t(y-x)}$$

$$\leq \sqrt{\mathcal{V}} \sqrt{\alpha'(t)}$$

$$\alpha'(t) \geq \frac{1}{2} \alpha(t)^2 \text{, i.e. } \frac{d}{dt} \left(\frac{1}{\alpha(t)} \right) \leq -\frac{1}{2}$$

$\alpha(0) \leq 0$ ✓ else α is increasing $\Rightarrow \alpha(1) > 0$

$$0 \leq \frac{1}{\alpha(1)} \leq \frac{1}{\alpha(0)} - \frac{1}{2} \Rightarrow \alpha(0) \leq \mathcal{V}.$$

Lema. (1) $C^T x_t \leq C^T x^* + \frac{\mathcal{V}}{t}$

(2) $\exists x$ with $\|tC + \nabla \phi(x)\|_{\nabla \phi(x)} \leq \frac{1}{t}$

$$C^T x \leq C^T x^* + \frac{\mathcal{V} + \sqrt{\mathcal{V}}}{t}.$$

Pf. By optimality of x_t , $\nabla F_t(x_t) = tC + \nabla \phi(x_t) = 0$

Pf. By optimality of x_t , $\nabla F_t(x_t) = tC + \nabla \phi(x_t) = 0$

$$C^T(x_t - x^*) = -\frac{1}{t} \langle \nabla \phi(x_t), x_t - x^* \rangle \leq \frac{\nu}{t}$$

For f_t , $\|x - x_t\|_X \leq 2 \|\nabla f_t(x)\|_{\nabla^2 f_t(x)^{-1}} \leq \frac{1}{3}$. $\nabla^2 f_t = \nabla^2 \phi$

$$C^T(x - x_t) \leq \|C\| \frac{\|x - x_t\|_X}{\nabla^2 \phi(x)^{-1}} \leq \frac{1}{3} \|x - x_t\|_X$$

Using

$$C = \frac{\nabla f_t(x)}{t} - \frac{\nabla \phi(x)}{t}, \quad C^T(x - x_t) \leq \frac{1}{t} \left(\frac{\|tC + \nabla \phi(x)\|}{\nabla^2 \phi(x)^{-1}} + \frac{\|\nabla \phi\|}{\nabla^2 \phi(x)} \cdot \|x - x_t\|_X \right)$$

$$\leq \frac{1}{3t} \left(\frac{1}{6} + \sqrt{\nu} \right) \leq \frac{\sqrt{\nu}}{t}.$$

Thm, IPM finds $x \in K$ with $Cx \leq \bar{C}x^* + \varepsilon$
in $O(\sqrt{\nu} \log \left(\frac{2}{\varepsilon} \|C\|_{\nabla^2 \phi(x^*)^{-1}} \right))$ iterations.

Pf We maintain $\|\nabla f_t(x)\|_{\nabla^2 f_t(x)^{-1}} \leq \frac{1}{6}$.

Now at $t = \frac{\nu + \sqrt{\nu}}{\varepsilon}$, we have the conclusion.

Given such an x , after one Newton step, we have
 $\|x - x_t\|_X \leq \left(\frac{1}{6}\right)^2 \leq \frac{1}{36} < \frac{1}{3}$.

$$\text{Given some } \alpha \quad \text{and} \quad \|\nabla f_t(x)\|_{\nabla^2 f_t(x)^{-1}} \leq \frac{\left(\frac{1}{6}\right)^2}{1 - \frac{1}{6}} \leq \frac{1}{25}.$$

$$\text{Now } \hat{t} = (1+h)t \quad f_{\hat{t}}(x) = (1+h)t Cx + \phi(x)$$

$$\begin{aligned} \nabla f_{\hat{t}}(x) &= (1+h)t C + \nabla \phi(x) \\ &= (1+h) \nabla f_t(x) + h \nabla \phi(x) \end{aligned}$$

$$\begin{aligned} \text{So, } \|\nabla f_{\hat{t}}(x)\|_{\nabla^2 f_{\hat{t}}(x)^{-1}} &\leq (1+h) \|\nabla f_t(x)\|_{\nabla^2 f_t(x)^{-1}} + h \|\nabla \phi(x)\|_{\nabla^2 \phi(x)^{-1}} \\ &\leq \frac{1+h}{25} + h \sqrt{25} \leq \frac{1+h}{25} + \frac{1}{9} \leq \frac{1}{6}. \end{aligned}$$
