Local Convergence of Newton's Method

$f: Q \to \mathbb{R}$ convex, differentiable $\quad Q \subseteq \mathbb{R}^n$ - open convex set

Main Lemma Let f be self-concordant with $M \geq 0$. $x \in Q$. Let $y \in Q$ belong to Dikin's ellipsoid:

$$y \in \mathcal{E}_x = \{ y : \|y - x\|_x < \tfrac{2}{M} \}$$

Then, the Hessian is **stable**:

$$r := \|y - x\|_x < \frac{2}{M}$$

$$\left(1 - \frac{M}{2} r\right)^2 f''(x) \preceq f''(y) \preceq \left(1 - \frac{M}{2} r\right)^{-2} f''(x).$$

Newton's step

$$x^+ = x - f''(x)^{-1} f'(x) \iff f'(x) + f''(x)(x^+ - x) = 0$$

Why $x^+ \in Q$?

Newton's Decrement:

$$\lambda(x) := \underline{\|f'(x)\|_x^*} \equiv \underbrace{\langle f'(x), f''(x)^{-1} f'(x) \rangle^{1/2}}_{\text{Dual Norm}} = \|f''(x)^{-1/2} f'(x)\|_2 =$$

$$= \underline{\|x^+ - x\|_x} \equiv \langle f''(x)(x^+ - x), x^+ - x \rangle^{1/2}$$

$$= \langle \underline{x - x^+, f'(x)} \rangle^{1/2}$$

Remark: $\lambda(x^*) = 0$.

Theorem Let $\lambda(x) \leq \frac{1}{M}$. Then $\quad Mr \leq 1$

$$\lambda(x^+) \leq M \lambda(x)^2.$$

Region of quadratic convergence.

Proof Denote $r := \lambda(x) \equiv \|x^+ - x\|_x$

$$\lambda(x^+) = \|f'(x^+)\|_{x^+}^* \leq \frac{1}{1 - \frac{M}{2} r} \|f'(x^+)\|_x^*$$

$$f'(x^+) = f'(x^+) - f'(x) - f''(x)(x^+ - x) \underset{\substack{\uparrow \\ \text{Fundamental} \\ \text{Theorem} \\ \text{of calculus}}}{=} (G-H)(x^+ - x)$$

where $H = f''(x)$

$$G = \int_0^1 f''(x + t(x^+ - x))\,dt$$

$$\|f'(x^+)\|_x = \|(G-H)(x^+-x)\|_x = \|H^{-1/2}(G-H)(x^+-x)\|_2 =$$

$$= \|H^{-1/2}(G-H)H^{-1/2}H^{1/2}(x^+-x)\|_2 \leq$$

$$\leq \underbrace{\|H^{-1/2}(G-H)H^{-1/2}\|_2}_{\leq ?} \cdot \underbrace{\|H^{1/2}(x^+-x)\|_2}_{= r}$$

$$G = \int_0^1 f''(x+\tau(x^+-x))\,d\tau \succeq H \cdot \int_0^1 \left(1-\tau\frac{M}{2}r\right)^2 d\tau = H\cdot\left(1-\frac{Mr}{2}+\frac{1}{12}M^2r^2\right)$$

$$H^{-1/2}(G-H)H^{-1/2} \succeq I \cdot \frac{Mr}{2}\Big(\underbrace{\frac{Mr}{6}-1}_{<0}\Big)$$

$$G = \int_0^1 f''(x+\tau(x^+-x))\,d\tau \preceq H \cdot \int_0^1 \frac{d\tau}{\left(1-\tau\frac{Mr}{2}\right)^2} = H\cdot\left(\frac{1}{1-\frac{Mr}{2}}\right)$$

$$H^{-1/2}(G-H)H^{-1/2} \preceq I\cdot\left[\frac{1}{1-\frac{Mr}{2}}-1\right] = I\cdot\frac{Mr/2}{1-Mr/2}.$$

$$\|H^{-1/2}(G-H)H^{-1/2}\|_2 \leq \frac{Mr}{2}\max\left\{\frac{1}{1-\frac{Mr}{2}}, 1-\frac{Mr}{6}\right\} \leq \frac{Mr/2}{1-Mr/2}.$$

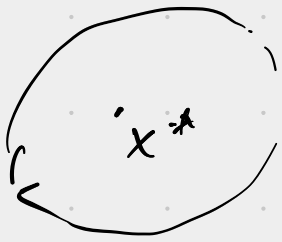
$$\lambda(x^+) \leq \frac{M}{2-Mr}r^2 = \frac{M}{2-M\lambda(x)}\lambda(x)^2 \leq M\lambda(x)^2 \qquad \square$$

Discussion

1. The region of local quadratic convergence for Newton's Method:

$$\lambda(x)^2 = \langle f'(x), f''(x)^{-1} f'(x)\rangle < \frac{1}{M^2}$$

2. Function is Quadratic: $M=0$. $x^+ = x^*$.



3. Standard Self-Concordant Function: $M=2$.

4. Fix Euclidean norm:

$$\left.\begin{array}{l} \cdot \text{ Hessian is Lipschitz: } \|f'''(x)\| \le L \\ \cdot \text{ Strong convex: } \underline{f''(x) \succeq \mu I} \end{array}\right\} \Rightarrow M = \frac{L}{\mu^{3/2}}$$

Strict local minimum: $f''(x^*) \succ 0$

$$\langle f'(x), f''(x)^{-1} f'(x)\rangle = \lambda(x)^2 \le \frac{1}{\mu}\|f'(x)\|_2^2$$

<u>Sufficient condition:</u>

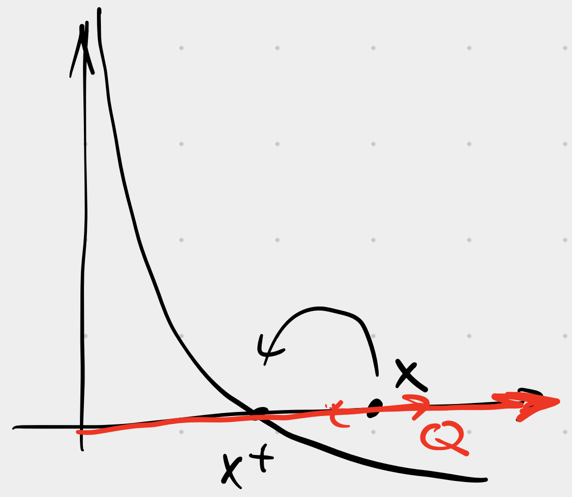
<u>Corollary</u> Let $\underline{\|f'(x)\|_2 < \frac{\mu^2}{L}}$ Then

$$\lambda(x) \le \frac{1}{\sqrt{\mu}}\|f'(x)\|_2 < \frac{\mu^{3/2}}{L} = \frac{1}{M}$$

$\Rightarrow$ x is in a region of quadratic convergence of Newton's method.

Quadratic function:

$$f(y) = \frac{1}{2}\langle Ay, y\rangle - \langle b, y\rangle \equiv f(x) + \langle f'(x), y-x\rangle + \frac{1}{2}\langle f''(x)(y-x), y-x\rangle$$



We use for $Q \subseteq \mathbb{R}^n$

– "the largest possible set"

$f: Q \to \mathbb{R}$ "a natural domain" of f.

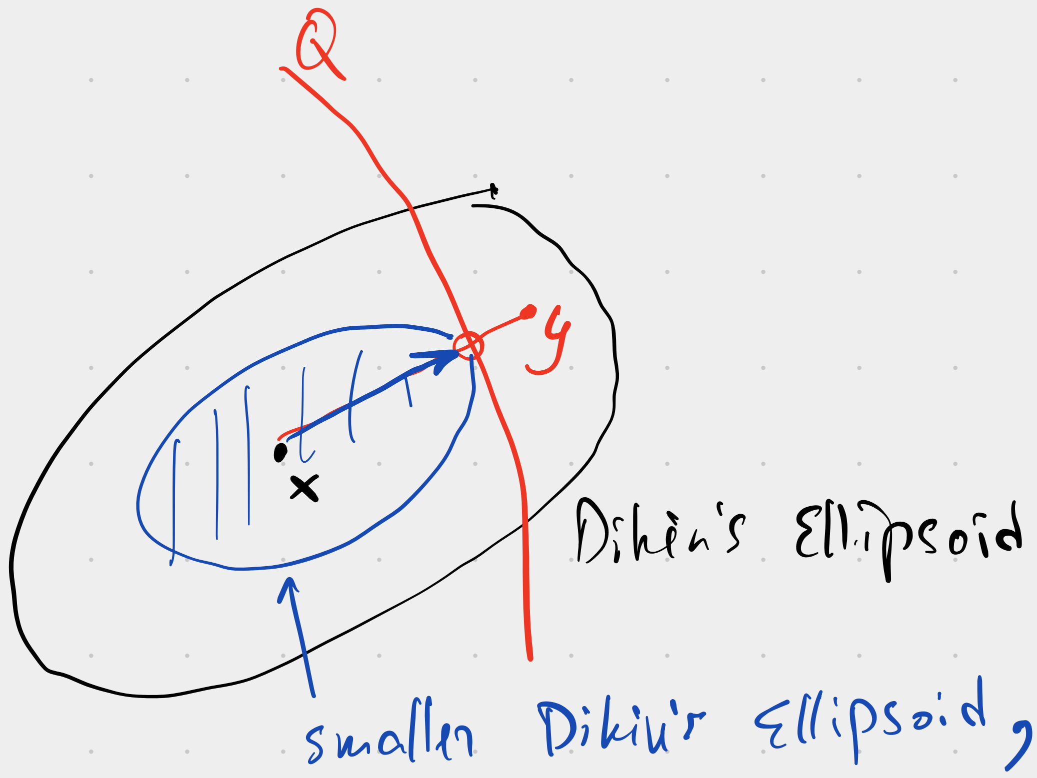
- $Q = \mathrm{dom} f$
- f is closed convex function.
- $\boxed{\text{For any } x_k \to \partial Q \quad f(x_k) \to +\infty}$

<u>Proposition</u> Let $f: Q \to \mathbb{R}$ be self-concordant, Q - natural domain.

Then, $x \in Q$:

$$\mathcal{E}_x = \{ y : \|y - x\|_x < \tfrac{2}{M} \} \subseteq Q.$$

<u>Proof sketch</u>



Main lemma:

- Hessians are bounded
- f is bounded. (?!) □

We proved: locally, $\lambda(x_k) \to 0$, $k \to +\infty$

$f(x_k) - f^*$

$\|x_k - x^*\|_{x_k}$ locally, they are related.

$\|x_k - x^*\|_{x^*}$

Proposition Assume $\lambda(x) < \frac{2}{M}$ Then

$$\|x - x^*\|_x \leq \frac{\lambda(x)}{1 - \frac{M}{2}\lambda(x)} .$$

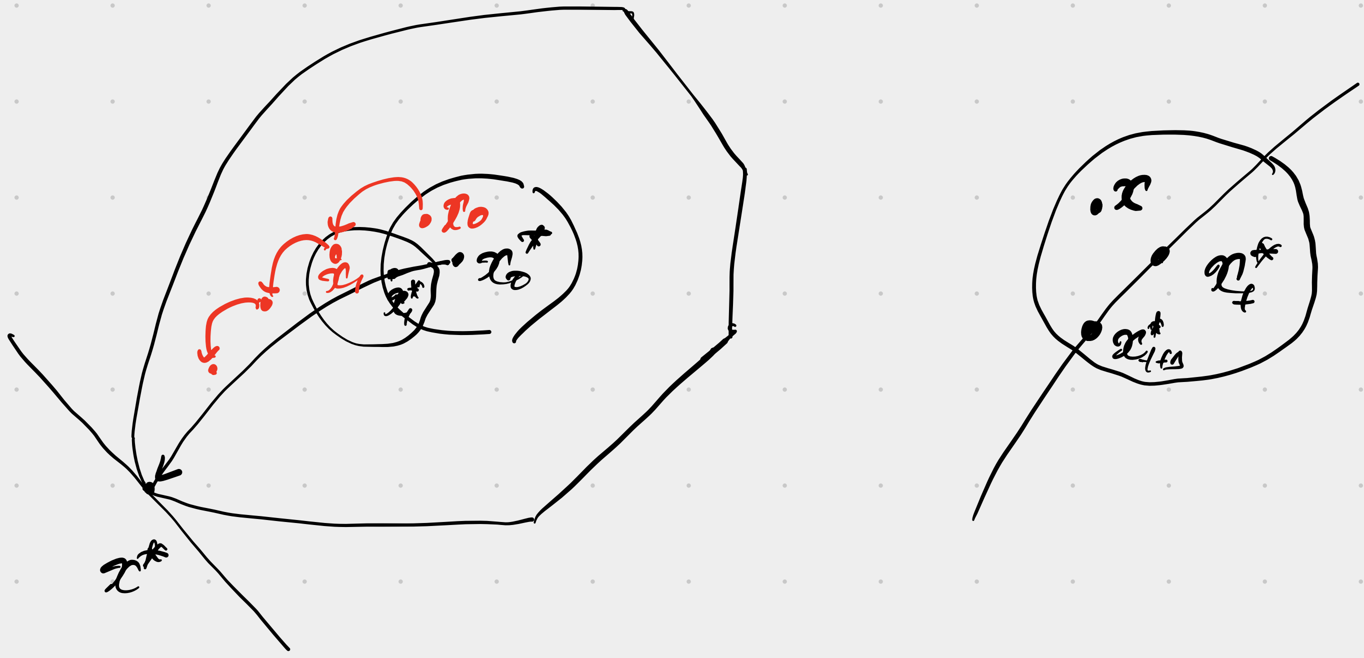
$\min\limits_{x \in \Omega} f(x)$ $\quad\quad \Omega \subseteq \operatorname{dom} f = Q$

A similar result can be proved for other accuracy measures as well.

Path-Following Scheme

$$\min_{x \in \bar{Q}} \langle c, x \rangle$$

$Q \subset \mathbb{R}^n$ - a bounded open convex set.



Idea: Introduce a "barrier function": $F : Q \to \mathbb{R}$
- helps to remain in the set.

We consider a family problems, $t \geq 0$:

$$f_t(x) = t\langle c, x \rangle + F(x)$$

Denote

$$x_t^* = \operatorname*{argmin}_{x \in Q} f_t(x) \equiv \operatorname*{argmin}_x f_t(x)$$

$$\boxed{tc + F'(x_t^*) = 0}$$

x_t^* - "central path"

$t = 0$: $x_0^* = \operatorname*{argmin}_x F(x)$ - an analytic center of Q.

$t \to +\infty$: $x_t^* \to x^*$