Definition

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

f - differentiable, strictly convex

Def. f is <u>self-concordant</u> with $M \geq 0$:

$$D^3 f(x)[h, h, u] \leq M \|h\|_x^2 \cdot \|u\|_x \qquad \forall x \in Q, \ \forall h, u \in \mathbb{R}^n$$

$$\|h\|_x = \sqrt{\underbrace{D^2 f(x)[h, h]}_{>0}}$$

$$\Updownarrow$$

$$D^3 f(x)[h, h, h] \leq M \cdot \|h\|_x^3 \qquad \forall x \in Q, h \in \mathbb{R}^n$$

$$D^3 f(x)[h, h, h] = D^3 f(x)[h]^3 = \sum_{1 \leq i, j, k \leq n} \frac{\partial^3 f(x)}{\partial x_i \partial x_j \partial x_k} h^{(i)} h^{(j)} h^{(k)}$$

<u>Example</u> $f(x) = -\ln x \Rightarrow$ self-concordant with $M = 2$.

<u>Example</u> (Log. Barrier for semidefinite cone)

$$S^n = \{ X \in \mathbb{R}^{n \times n} \mid X = X^T \}$$

$$Q = \{ X \in S^n \mid X \succ 0 \}$$

$$f(X) = -\ln \det X = -\sum_{i=1}^n \ln \lambda_i(X)$$

$$Df(X)[H] = -\mathrm{tr}(X^{-1} H) \Rightarrow f'(X) = -X^{-1}$$

$H \in S^n$

$$D^2 f(X)[H,H] = tr(X^{-1} H X^{-1} H) = tr(\underbrace{X^{-1/2} H X^{-1/2}}_{S} X^{-1/2} H X^{-1/2})$$
$$= tr(S^2)$$

$$D^3 f(X)[H,H,H] = -2tr(X^{-1} H X^{-1} H X^{-1} H) = -2tr(S^3)$$

To check:

$$|2tr(S^3)| = 2\left|\sum_{i=1}^n \lambda_i(S)^3\right| \leq \underline{2\sum_{i=1}^n |\lambda_i(S)|^3}$$
$$\overset{?}{\leq} M \cdot (tr(S^2))^{3/2} = \underline{M \cdot \left(\sum_{i=1}^n |\lambda_i(S)|^2\right)^{3/2}}$$

$$\boxed{M = 2}$$

$$\left(\sum |x_i|^3\right)^{1/3} \leq \left(\sum |x_i|^2\right)^{1/2}$$

$$\|\cdot\|_3 \leq \|\cdot\|_2 \qquad \square$$

Basic Properties:

- **Sum of self-concordant functions**

Proposition $f(x) = \sum_{i=1}^m f_i(x)$, $f_i(x)$ – SC, $M_i \geq 0$

Then, f – SC, $M = \max_{1 \leq i \leq m} M_i$.

Proof.

$$D^3 f(x)[h]^3 = \sum_{i=1}^m D^3 f_i(x)[h]^3 \leq \sum_{i=1}^m M_i \cdot (D^2 f_i(x)[h]^2)^{3/2} \leq$$
$$\leq M \cdot \sum_{i=1}^m (D^2_i f(x)[h]^2)^{3/2} \overset{?}{\leq} M \cdot \left(\sum_{i=1}^m D^2_i f(x)[h]^2\right)^{3/2}$$

$$\left(\sum_{i=1}^{m} a_i^{3/2}\right) \le \left(\sum_{i=1}^{m} a_i\right)^{3/2}, \qquad a_i \geq 0.$$

$$\|\cdot\|_{3/2} \le \|\cdot\|_1 \qquad \square$$

Example $Q = \mathbb{R}^n_{>0} = \{ x \in \mathbb{R}^n \mid x_1 > 0, \dots, x_n > 0 \}$

$$f(x) = -\sum_{i=1}^{n} \ln x_i \quad \Rightarrow \quad \boxed{M = 2}$$

• Affine Invariance

Define $g(y) = f(Ay + b)$, $\qquad M_g = M_f$

Example

$$Q = \{ \langle a_1, x \rangle < b_1, \dots, \langle a_m, x \rangle < b_m \}$$

$$f(x) = -\sum_{i=1}^{m} \ln(b_i - \langle a_i, x \rangle) \qquad \boxed{M = 2.}$$

Main lemma

$$\|h\|_x \approx \|h\|_y \qquad x \approx y$$

Lemma Let $x, y \in Q$, denote $r = \|x-y\|_x$. Assume

$$r < \frac{2}{M}.$$

Then:

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

Proof

First, $h := y-x$.

$$\left(1-\frac{M}{2}r\right)\|y-x\|_x \leq \|y-x\|_y \leq \frac{1}{\left(1-\frac{M}{2}r\right)} \|y-x\|_x$$

$$\Leftrightarrow \quad \left(1-\frac{M}{2}r\right) r \leq \|y-x\|_y \leq \frac{r}{1-\frac{M}{2}r} \qquad (I)$$

$$\varphi(t) = \langle f''(x+t(y-x))(y-x), y-x \rangle = \|y-x\|^2_{x+t(y-x)} \qquad 0 \leq t \leq 1$$

$$|\varphi'(t)| = \left|D^3 f(x+t(y-x))[y-x]^3\right| \leq M\,\varphi(t)^{3/2}$$

$$\left|\frac{d}{dt}\left(\frac{-2}{\sqrt{\varphi(t)}}\right)\right| = \frac{|\varphi'(t)|}{\varphi(t)^{3/2}} \leq M$$

For $0 \le t \le 1$:

$$\underline{\left| \frac{2}{\sqrt{\varphi(t)}} - \frac{2}{\sqrt{\varphi(0)}} \right|} = \left| \int_0^t \frac{d}{d\tau}\left(\frac{-2}{\sqrt{\varphi(\tau)}} \right) \right| \le \int_0^t \left| \frac{d}{d\tau}\left(\frac{-2}{\sqrt{\varphi(\tau)}} \right) \right| \le$$

$$\le \underline{Mt}$$

Take $t=1$:

$$M \ge \left| \frac{2}{\sqrt{\varphi(1)}} - \frac{2}{\sqrt{\varphi(0)}} \right| = \left| \frac{2}{\|y-x\|_y} - \frac{2}{\|y-x\|_x} \right|$$

① $$\frac{1}{\|y-x\|_y} \le \frac{M}{2} + \frac{1}{\|y-x\|_x} = \frac{1}{r} + \frac{M}{2} = \frac{1+\frac{M}{2}r}{r} \le$$

$$(1-\alpha)(1+\alpha) = 1-\alpha^2 \le 1 \quad , \quad 0 \le \alpha \le 1 \quad , \quad \alpha = \frac{M}{2}r < 1$$

$$1+\frac{M}{2}r \le \frac{1}{1-\frac{M}{2}r}$$

$$\le \frac{1}{r(1-\frac{M}{2}r)}$$

$$\|y-x\|_y \ge r\left(1-\frac{M}{2}r\right)$$

② $$\frac{1}{\|x-y\|_y} \ge \frac{1}{\|x-y\|_x} - \frac{M}{2} = \frac{1}{r} - \frac{M}{2} = \frac{1-\frac{M}{2}r}{r}$$

$$\Leftrightarrow \quad \|x-y\|_y \le \frac{r}{(1-\frac{M}{2}r)} .$$

We also get:

$$\frac{1}{\sqrt{\varphi(t)}} \geq \frac{1}{\sqrt{\varphi(0)}} - \frac{tM}{2} = \frac{1}{r} - \frac{tM}{2} = \frac{1-\frac{tM}{2}r}{r}$$

$$\Leftrightarrow \boxed{\sqrt{\varphi(t)} \leq \frac{r}{1-\frac{tM}{2}r}}$$

II. $\psi(t) = D^2 f(x+t(y-x))[h]^2 \qquad h \in \mathbb{R}^n$

$$|\psi'(t)| = |D^3 f(x+t(y-x))[h,h,y-x]|$$

$$\leq M \cdot \|h\|^2_{x+t(y-x)} \cdot \|y-x\|_{x+t(y-x)}$$

$$= M\,\psi(t)\sqrt{\varphi(t)} \leq M\,\psi(t) \cdot \frac{r}{1-\frac{tM}{2}r}$$

$$\left|\frac{d}{dt}\ln\psi(t)\right| = \left|\frac{\psi'(t)}{\psi(t)}\right| \leq M \cdot \frac{r}{1-\frac{tM}{2}r} = -2\frac{d}{dt}\ln\left(1-t\frac{Mr}{2}\right)$$

$$\left|\ln\frac{\|h\|_y}{\|h\|_x}\right| = |\ln\psi(1) - \ln\psi(0)| \leq -2\ln\left(1-\frac{Mr}{2}\right)$$

$$\ln\left(\left[1-\frac{Mr}{2}\right]^2\right) \leq \ln\frac{\|h\|_y}{\|h\|_x} \leq \ln\left(\left[1-\frac{Mr}{2}\right]^{-2}\right) \qquad \square.$$

__Remark__ $y \in Q$. If y belongs to Dikin's Ellipsoid

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

Then $\nabla^2 f(y)$ and $\nabla^2 f(x)$ are __comparable__ or __stable__.

Local Convergence of Newton's Method

$\|\cdot\|_x \qquad h \in \mathbb{R}^n$

$g \in \mathbb{R}^n, \quad \langle g, \cdot \rangle$

$$\|\langle g, \cdot \rangle\|_x = (\|g\|_x)_* := \max_{\|h\|_x \le 1} \langle g, h \rangle =$$

$$= \langle g, f''(x)^{-1} g \rangle^{1/2} = \| f''(x)^{-1/2} g \|_2 .$$

$$\boxed{g := f'(x)}$$

The Newton Decrement:

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

Newton's step:

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

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

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

Region of quadratic convergence.

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

$\Rightarrow$ locally, we have quadratic convergence!

At iteration $k \geq 0$: $\delta_k := M\lambda(x_k)$

$$\delta_{k+1} \leq \delta_k^2 \leq (\delta_{k-1}^2)^2 \leq \left[(\delta_{k-2}^2)^2\right]^2 \leq \cdots$$

$$\delta_k \leq (\delta_0)^{2^k}$$

Assume: $\delta_0 \leq \frac{1}{2} \iff \lambda(x_0) \leq \frac{1}{2M}$

Then, $\delta_k \leq \left(\frac{1}{2}\right)^{2^k} \leftarrow$ Quadratic Convergence.

$\delta_k \leq \varepsilon \Rightarrow K = \log_2 \log_2 \frac{1}{\varepsilon}$ steps.