# LINMA2471 : Optimization models and models : course 11

Maxime Duyck
Rémi Chauvenne

December 2015

## 1 Self-concordant function

**Definition 1.** A function $f$ is self-concordant if and only if

- $f \in C^3(x)$ (on an open domain X)

- $f$ is convex

- $\nabla^3 f(x)[h, h, h] \leq 2[h^T \nabla^2 f(x)h]^{\frac{3}{2}}$

Let's define $F_{x,h}(t) : t \to f(x + th)$. We have

$$\nabla^3 f(x)[h, h, h] = F'''_{x,h}(0) \overset{1D}{=} f'''(x)h^3$$

$$\nabla^2 f(x)[h, h] = F''_{x,h}(0) \overset{1D}{=} f''(x)h^2$$

$$\nabla f(x)[h] = F'_{x,h}(0) \overset{1D}{=} f'(x)h$$

where 1D means that we are studying the one-dimensional case. Let's look at the third condition of **definition 1** in 1D :

$$f'''(x)h^3 \leq 2(f''(x)h^2)^{\frac{3}{2}}$$

which is equivalent to

$$\begin{cases} f'''(x) \leq 2f''(x)^{\frac{3}{2}} & \text{(h positive)} \\ -f'''(x) \leq 2f''(x)^{\frac{3}{2}} & \text{(h negative)} \end{cases}$$

Then one can say that in the one-dimensional case :

$$|f'''(x)| \leq 2f''(x)^{\frac{3}{2}}$$

for a self-concordant function.

**Example 1.** $f(x) = -\log(x)$.

$$f'(x) = -\frac{1}{x} \quad f''(x) = \frac{1}{x^2} \quad f'''(x) = -\frac{2}{x^3}$$

$$\Rightarrow \left| -\frac{2}{x^3} \right| = 2 \left( \frac{1}{x^2} \right)^{\frac{3}{2}}$$

**Property 1.** Let $f$ and $g$ be self-concordant (s.c.) functions. Then $f + g$ is a s.c. function.

**Example 2.**
$$f(x) : \mathbb{R}^n_{++} \to \mathbb{R} : -\sum \log (x_i)$$
is s.c. as it is a sum of s.c. functions.

**Property 2.** Let $x \to f(x)$ be a s.c. function. Then $y \to f(c - A^T y)$ is s.c.

**Example 3.**
$$f(y) = -\sum_{i=1}^{m} \log (c_i - a_i^T y)$$
on domain $\{y | c - A^T y > 0\}$ is s.c.

**Example 4.**
$$f(x_0, x_1, ..., x_n) = -\log (x_0^2 - x_1^2 - ... - x_n^2)$$
is s.c. on $\text{int}(\mathbb{L}^n)$ with $\mathbb{L}^n = \left\{ (x_0, x_1, ..., x_n) | x_0 > \sqrt{x_1^2 + ... + x_n^2} \right\}$

**Example 5.** Let $X \in \mathbb{S}^n$.
$$f(X) = -\log \det X$$
is s.c. on $\text{int}(\mathbb{S}^n_+)$

**Property 3.** Let X be a set that contains no line. Then one can say that

- any s.c. function in X satisfies $\nabla^2 f(x) > 0$

- $f(x) \to +\infty$ as $x \to \delta(X)$

where $\delta(X)$ is the boundary of set X.

## 2 Minimisation of s.c. functions with Newton's method

S.c. functions are easy to minimize with Newton's method. But the optimality measure given by $||\nabla f(x)||$ is bad as it is not affine-invariant. Let's try with another norm.

**Definition 2.** Local norm (given a s.c. function f) : $(x \in X)$
$$||z||_x = (z^T \nabla^2 f(x) z)^{\frac{1}{2}}$$

The dual of the local norm is given by
$$||z||_x^* = (z^T \nabla^2 f(x)^{-1} z)^{\frac{1}{2}}$$

Using the dual of the local norm, we have the optimality measure given by
$$\begin{aligned} \delta(x) &= ||\nabla f(x)||_x^* \\ &= (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{\frac{1}{2}} \\ &= ||n(x)||_x \end{aligned}$$

2

**Property 4.** Given a s.c. function (on domain X), $x \in X$. If $\delta(x) < 1$ then

- min $x^*$ of f exists

- $f(x) \leq f(x^*) - \delta - \log(1 - \delta) \approx f(x^*) - \frac{\delta^2}{2}$

- $\|x - x^*\|_x \leq \frac{\delta}{1 - \delta}$

- $x^+ = x + n(x)$ is feasible $(x^+ \in X)$

- $\delta(x^+) \leq \left( \frac{\delta(x)}{1 - \delta(x)} \right)^2$

And for any $\delta(x)$, we have

- $x^+ = x + \left( \frac{\delta(x)}{1 + \delta(x)} \right) n(x)$

- $f(x) - f(x^+) \geq \delta(x) - \log(1 + \delta(x)) \geq 0$

with $\left( \frac{\delta(x)}{1 + \delta(x)} \right) \leq 1$ which is the damped Newton step.

**Application.** Suppose $\delta(x_0) > \frac{1}{\sqrt{2}}$, we have

$$f(x_0) - f(x_1) \geq 0.16.$$

If again $\delta(x_1) > \frac{1}{\sqrt{2}}$, then

$$f(x_1) - f(x_2) \geq 0.16.$$

Hence, after at most $\frac{f(x_0) - f(x^*)}{0.16}$ iterations, we have $\leq \frac{1}{\sqrt{2}}$.

After applying $O(\log \log \frac{1}{\epsilon})$ pure Newton steps, we obtain an $\epsilon-$solution.