Lecture 19

19.1 Self-Concordant Analysis

19.1.1 Definition

We say that a convex differentiable function $f : Q \to \mathbb{R}$ defined on an open convex set $Q \subseteq \mathbb{R}^n$ is *self-concordant with constant* $M \geq 0$ if

$$D^3 f(x)[h, h, u] \quad \leq \quad M \|h\|_x^2 \|u\|_x, \qquad \forall h, u \in \mathbb{R}^n, x \in Q. \tag{19.1}$$

In the last lecture, we discussed how this definition is equivalent to the one bounding the third derivative along one direction $h = u \in \mathbb{R}^n$.

19.1.2 Main Lemma

Our goal is to compare two Hessians, $\nabla^2 f(x)$ and $\nabla^2 f(y)$, when the points are close to each other, $x \approx y$. The following lemma is the main consequence of self-concordance, which, in fact, can be substituted for its definition (see [Ren01]). It underpins most of the other important results about self-concordant functions.

Lemma 19.1.1. *Let $x, y \in Q$ and denote $r := \|y - x\|_x$. Assume that x and y are close:*

$$r \quad < \quad \tfrac{2}{M}. \tag{19.2}$$

Then,

$$(1 - \tfrac{M}{2} r)^2 \nabla^2 f(x) \quad \preceq \quad \nabla^2 f(y) \quad \preceq \quad \tfrac{1}{(1 - \frac{M}{2} r)^2} \nabla^2 f(x). \tag{19.3}$$

Proof. First, we prove (19.3) along direction $h := y - x$, that is

$$(1 - \tfrac{M}{2} r) r \quad \leq \quad \|y - x\|_y \quad \leq \quad \tfrac{r}{(1 - \frac{M}{2} r)}. \tag{19.4}$$

For that, we consider a function

$$\varphi(t) \quad := \quad D^2 f(x + t(y - x))[y - x]^2 \quad = \quad \|y - x\|^2_{x + t(y - x)}.$$

We have

$$|\varphi'(t)| \quad = \quad D^3 f(x + t(y - x))[y - x]^3 \quad \leq \quad M \varphi(t)^{3/2}. \tag{19.5}$$

Hence,

$$\Big| \tfrac{d}{dt} \Big[\tfrac{-2}{\sqrt{\varphi(t)}} \Big] \Big| \quad = \quad \tfrac{|\varphi'(t)|}{\varphi(t)^{3/2}} \quad \overset{(19.5)}{\leq} \quad M. \tag{19.6}$$

Therefore, by the fundamental theorem of calculus, for $0 \le t \le 1$:

$$\left| \frac{2}{\sqrt{\varphi(t)}} - \frac{2}{\sqrt{\varphi(0)}} \right| \;=\; \left| \int_0^t \frac{d}{d\tau} \frac{-2}{\sqrt{\varphi(\tau)}} \right| \;\le\; \int_0^t \left| \frac{d}{d\tau} \frac{-2}{\sqrt{\varphi(\tau)}} \right| \;\overset{(19.6)}{\le}\; tM. \tag{19.7}$$

Substituting $t = 1$ and using the definition of φ, we immediately obtain:

$$\left| \frac{1}{\|y-x\|_y} - \frac{1}{\|y-x\|_x} \right| \;\le\; \frac{M}{2},$$

or, the corresponding pair of inequalities, the first one:

$$\frac{1}{\|y-x\|_y} \;\le\; \frac{1}{\|y-x\|_x} + \frac{M}{2} \;=\; \frac{1}{r} + \frac{M}{2} \;=\; \frac{1+\frac{M}{2}r}{r} \;\le\; \frac{1}{r(1-\frac{M}{2}r)}, \tag{19.8}$$

where we used that $(1-t)(1+t) = 1 - t^2 \le 1$ for $t := \frac{M}{2}r \overset{(19.2)}{<} 1$, and the second one:

$$\frac{1}{\|y-x\|_y} \;\ge\; \frac{1}{\|y-x\|_x} - \frac{M}{2} \;=\; \frac{1}{r} - \frac{M}{2} \;=\; \frac{1-\frac{M}{2}r}{r}. \tag{19.9}$$

Inequalities (19.8) and (19.9) together give (19.4).

Note that from (19.7), we also have, for $0 \le t \le 1$:

$$\frac{1}{\sqrt{\varphi(t)}} \;\ge\; \frac{1}{\sqrt{\varphi(0)}} - \frac{tM}{2} \;=\; \frac{1-tMr/2}{r} \qquad \Leftrightarrow \qquad \sqrt{\varphi(t)} \;\le\; \frac{r}{1-tMr/2}. \tag{19.10}$$

Finally, to prove (19.3) along arbitrary direction h, we denote

$$\psi(t) \;:=\; D^2 f(x + t(y-x))[h]^2 \;=\; \|h\|^2_{x+t(y-x)}.$$

Differentiating it gives

$$\begin{aligned} |\psi'(t)| \;&=\; |D^3 f(x+t(y-x))[h,h,y-x]| \;\overset{(19.1)}{\le}\; M\|h\|_{x+t(y-x)}\|y-x\|_{x+t(y-x)} \\ &=\; M\psi(t)\sqrt{\varphi(t)} \;\overset{(19.10)}{\le}\; \frac{Mr}{1-tMr/2}\psi(t). \end{aligned} \tag{19.11}$$

Thus,

$$\left| \frac{d}{dt} \ln \psi(t) \right| \;=\; \left| \frac{\psi'(t)}{\psi(t)} \right| \;\overset{(19.11)}{\le}\; \frac{Mr}{1-tMr/2} \;=\; 2\frac{d}{dt} \ln\Big(1 - t\frac{Mr}{2}\Big). \tag{19.12}$$

Integrating this inequality, we conclude:

$$\left| \ln \frac{\|h\|_y}{\|h\|_x} \right| \;=\; \Big| \ln \psi(1) - \ln \psi(0) \Big| \;\le\; -2\ln(1 - \tfrac{Mr}{2}),$$

or, equivalently,

$$\ln\Big(\big[1 - \tfrac{Mr}{2}\big]^2 \Big) \;\le\; \ln \frac{\|h\|_y}{\|h\|_x} \;\le\; \ln\Big(\big[1 - \tfrac{Mr}{2}\big]^{-2} \Big)$$

which completes the proof. □

19.2 Dikin's Ellipsoid

The previous lemma states that as long as $y \in Q$ belongs to *Dikin's ellipsoid*:

$$y \quad \in \quad \mathcal{E}_x \;:=\; \Big\{ y \in \mathbb{R}^n \;:\; \|y - x\|_x < \tfrac{2}{M} \Big\}, \tag{19.13}$$

the Hessians $\nabla^2 f(y)$ and $\nabla^2 f(x)$ are *comparable* or *stable*: that is (19.3) holds.

Note that, in general, it might happen that $\mathcal{E}_x \not\subseteq Q$, as Q can be *any open convex set* on which f is defined. However, instead of taking an arbitrary set, it is natural to assume that Q is "as large as possible". Specifically, we define Q as the "domain" of f, meaning that the following property holds:

For any $y^\star \in \partial Q$ and any sequence $\{y_k\}_{k \geq 0}$ of points in Q such that $y_k \to y^\star$, we have

$$f(y_k) \quad \to \quad +\infty. \tag{19.14}$$

Thus, under condition (19.14), we assume either that $Q = \mathbb{R}^n$ is the whole space, or that f blows up at the boundary.

With this condition, we can show that self-concordance implies that with every point $x \in Q$, *the entire Dikin's ellipsoid* (19.13) belongs to it.

Proposition 19.2.1. *Let $Q \subseteq \mathbb{R}^n$ be an open set such that condition* (19.14) *holds for $f : Q \to \mathbb{R}$. Assume that f is self-concordant with constant $M > 0$. Then, for every $x \in Q$:*

$$\mathcal{E}_x \quad \subseteq \quad Q. \tag{19.15}$$

Proof. Assume (19.15) does not hold. Thus, there exists $y \in \mathcal{E}_x$ such that $y \notin Q$. Denote

$$t^\star \quad = \quad \inf \Big\{ \; t \in (0, 1] \;:\; x + t(y - x) \notin Q \; \Big\}.$$

By definition of $t^\star$, for any $\epsilon > 0$ there exists $y(\epsilon) = x + (t^\star - \epsilon)(y - x) \in Q$. Therefore, selecting a sequence of small positive $\epsilon_k \to 0$, we generate the sequence of points $y_k = y(\epsilon_k) \in Q$ such that $y_k \to y^\star := x + t^\star(y - x) \notin Q$. Hence, $y^\star \in \partial Q$, and by assumption (19.14) we have $f(y_k) \to +\infty$.

At the same time, $\|y_k - x\|_x = (t^\star - \epsilon_k)\|y - x\|_x < \|y - x\|_x < \frac{2}{M}$ and we conclude that the Hessian is uniformly bounded over segments $z = x + \alpha(y_k - x) \in Q$, $0 \leq \alpha \leq 1$:

$$\|\nabla^2 f(z)\| \quad \overset{(19.3)}{\leq} \quad \tfrac{1}{(1 - \frac{M}{2}\|z - x\|_x)^2} \|\nabla^2 f(x)\| \quad < \quad L \quad := \quad \tfrac{1}{(1 - \frac{M}{2}\|y - x\|_x)^2} \|\nabla^2 f(x)\|.$$

Therefore, all function values $f(y_k)$ are uniformly bounded, which contradicts $f(y_k) \to +\infty$. □

Literature

[NN94] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming.* SIAM, 1994.

[Ren01] James Renegar. *A mathematical view of interior-point methods in convex optimization.* SIAM, 2001.