Lecture 21

21.1 Barrier Approach for Constrained Minimization

We consider the convex optimization problem in the following form,

$$\min\Big\{ \langle c, x \rangle \; : \; x \in \bar{Q} \Big\}, \tag{21.1}$$

where $Q \subset \mathbb{R}^n$ is an open convex set, which, we assume to be bounded, for simplicity.

The idea of interior-point methods is to substitute constrained problem (21.1) by a *sequence of unconstrained minimization* subproblems. For that, we introduce a "barrier" function F defined on $\mathrm{dom}\, F = Q$, that prevents us from going outside the set. Then, we consider a family of objectives:

$$f_t(x) \quad = \quad t\langle c, x \rangle + F(x), \qquad t \geq 0. \tag{21.2}$$

We denote the minimum of (21.2) by

$$x_t^\star \quad := \quad \operatorname*{arg\,min}_{x \in \mathrm{dom}\, F} \Big[\, t\langle c, x \rangle + F(x) \, \Big], \tag{21.3}$$

which is called the *central path*. The initial point of the central path,

$$x_0^\star \quad = \quad \operatorname*{arg\,min}_{x \in \mathrm{dom}\, F} F(x),$$

is called an *analytic center* of the set, and $x_t^\star \to x^\star$ with $t \to +\infty$, where $x^\star \in \partial Q$ is a solution to (21.1). Note that a solution to (21.1) is always at the boundary, while the central path belong to the interior of our feasible set:

$$x_t^\star \quad \in \quad Q, \qquad t \geq 0.$$

The idea of interior-point methods is to trace the central path, by approximately solving (21.3) for an increasing sequence of parameters t, starting from $t = 0$.

Since

$$\nabla^2 f_t(x) \quad \equiv \quad \nabla^2 F(x),$$

we see that the second derivative of f_t does not depend on t. It appears that the geometry induced by $\nabla^2 F(x)$ is crucial for being able to solve (21.2) efficiently.

Note that even though the subproblem in (21.3) is formally still a *constrained* minimization, with the right choice of the barrier F, we can use the plain Newton's method to solve it. In fact, after each change of $t \mapsto t^+ = t + \Delta$ in (21.2), we will use only *one Newton's step*, ensuring that every new point belongs to the local region of quadratic convergence around the central path. The "right" choice of F is described by the following formal definition of *self-concordant barriers*.

21.2 Self-Concordant Barriers

21.2.1 Definition

We say that a strictly convex differentiable function $F : Q \to \mathbb{R}$ is a *self-concordant barrier* for a convex open set Q with parameter $\theta > 0$ if the following three conditions are satisfied:

1. *Q is the domain of F*: for any sequence $\{x_k\}_{k \geq 0}$ such that $x_k \to \partial Q$ it holds $F(x_k) \to +\infty$.

2. F is a *standard self-concordant function* (the constant of self-concordance is $M = 2$; otherwise we can rescale the barrier):

$$D^3 F(x)[h,h,h] \;\; \leq \;\; 2\|h\|_x^3 \;\; \equiv \;\; 2\langle \nabla^2 F(x) h, h\rangle^{3/2}, \qquad \forall x \in Q, \; h \in \mathbb{R}^n. \tag{21.4}$$

3. *F is Lipschitz with respect to the local norm*[1]*:*

$$\|DF(x)\|_x^2 \;\; \equiv \;\; \langle \nabla F(x), \nabla^2 F(x)^{-1} \nabla F(x)\rangle \;\; \leq \;\; \theta, \qquad \forall x \in Q. \tag{21.5}$$

The first two conditions are already familiar to us. In particular, they imply that with every point $x \in Q$, the entire *Dikin's ellipsoid* belong to the domain (see Proposition 19.2.1 in Lecture 19 for the proof):

$$\mathcal{E}_x \;\; = \;\; \Big\{ \, y \;\in\; \mathbb{R}^n \;:\; \|y - x\|_x < 1 \, \Big\} \;\; \subseteq \;\; Q,$$

and one Newton's step $x \mapsto x^+ = x - \nabla^2 F(x)^{-1} g$, for some $g \in \mathbb{R}^n$, remains in the ellipsoid, as soon as

$$\|x^+ - x\|_x \;\; = \;\; \langle g, \nabla^2 F(x)^{-1} g\rangle^{1/2} \;\; < \;\; 1. \tag{21.6}$$

This explains why we can treat optimization subproblem (21.3) as unconstrained one: assuming that g is sufficiently small (21.6) and using the Hessian of the barrier as a preconditioner ensures the feasibility of all points *without the need to perform any projections*. If first-order methods are used to solve (21.3), this property is lost. Note that computing the projection onto Q is harder than solving the original linear minimization problem (21.1).

The third condition (21.5) of the definition is new. It enable the efficient tracing of the central path (21.3). The main result of the interior-point method theory we aim to establish is the following:

Theorem 21.2.1. *We can solve* (21.1) *to within accuracy $\varepsilon > 0$, i.e., find a point $\bar{x} \in Q$ such that*

$$\langle c, \bar{x} - x^\star \rangle \;\; \leq \;\; \varepsilon$$

using a total of

$$K \;\; = \;\; O\Big(\sqrt{\theta} \ln \tfrac{1}{\varepsilon}\Big). \tag{21.7}$$

Newton steps.

Therefore, the barrier parameter $\theta > 0$ is crucial for describing the complexity of the problem (21.1).

[1] Recall that we use the dual norm to measure the size of the linear form $DF(x)[\cdot] = \langle \nabla F(x), \cdot \rangle$. That is why the Hessian in (21.5) is inverted, while in (21.4) we use the primal local norm.

21.2.2 Equivalent Conditions

The barrier condition (21.5) is equivalent to the following global inequality:

$$2\langle \nabla F(x), u\rangle - \langle \nabla^2 F(x) u, u\rangle \quad \leq \quad \theta, \qquad \forall u \in \mathbb{R}^n, \; x \in Q. \tag{21.8}$$

Indeed, (21.5) is the maximum of the left-hand-side of (21.8) over $u \in \mathbb{R}^n$. Furthermore, the new condition (21.8) can be used as a definition in cases when $\nabla^2 F(x)$ has *degenerate directions*, implying that F is not strictly convex along them[2].

Finally, substituting $u := \tau h$ for $\tau > 0$ and $h \in \mathbb{R}^n$ in (21.8), and maximizing the left-hand-side over $\tau > 0$, which corresponds to *homogenization* of the inequality, we obtain the following equivalent global bound:

$$\max_{\tau > 0} \Big[\, 2\tau \langle \nabla F(x), h\rangle - \tau^2 \langle \nabla^2 F(x) h, h\rangle \, \Big] \quad = \quad \tfrac{\langle \nabla F(x), h\rangle^2}{\langle \nabla^2 F(x) h, h\rangle} \quad \overset{(21.8)}{\leq} \quad \theta, \tag{21.9}$$

where the maximum is achieved for $\tau^\star = \frac{\langle \nabla F(x), h\rangle}{\langle \nabla^2 F(x) h, h\rangle}$. The last condition is the most convenient for checking the third barrier property. It can be rewritten as follows:

$$\langle \nabla F(x), h\rangle^2 \quad \leq \quad \theta \langle \nabla^2 F(x) h, h\rangle, \qquad \forall h \in \mathbb{R}^n, \; x \in Q, \tag{21.10}$$

or

$$\nabla^2 F(x) \quad \succeq \quad \tfrac{1}{\theta} \nabla F(x) \nabla F(x)^\top. \tag{21.11}$$

Therefore, we can interpret the last barrier property as a certain form of "self-concordant strong convexity" with parameter $\mu_{\mathfrak{sc}} := \frac{1}{\theta}$, while the self-concordant parameter $L_{\mathfrak{sc}} := 2$ is a measure of "smoothness". Under this speculative interpretation, the complexity (21.7) resembles that of the fast gradient method for smooth and strongly convex functions:

$$O\Big(\sqrt{\tfrac{L_{\mathfrak{sc}}}{\mu_{\mathfrak{sc}}}} \ln \tfrac{1}{\varepsilon} \Big),$$

even though the analysis of path-following schemes looks very different.

21.2.3 Examples

We have the following important examples of self-concordant barriers. We already know that these functions are standard self-concordant. Therefore, we check only the new barrier condition (21.5).

Example 21.2.2. $F(x) = -\log x$ is self-concordant barrier for $\mathbb{R}_{>0}$ with $\boxed{\theta = 1}$. Indeed,

$$F'(x) \quad = \quad -\tfrac{1}{x}, \qquad F''(x) \quad = \quad \tfrac{1}{x^2}.$$

Hence, (21.5) is satisfied.

Example 21.2.3. $F(X) = -\log\det(X)$ is self-concordant barrier for $\mathbb{S}^n_{\succ 0}$ with $\boxed{\theta = n}$. We have,

$$DF(X)[H] \quad = \quad \mathrm{tr}(X^{-1}H) \quad = \quad \mathrm{tr}(S), \quad \text{for} \quad S := X^{-1/2} H X^{-1/2}, \quad \text{and}$$

$$D^2 F(x)[H, H] \quad = \quad \mathrm{tr}(X^{-1} H X^{-1} H) \quad = \quad \mathrm{tr}(S^2).$$

Therefore, condition (21.10) implies that we need to check

$$\mathrm{tr}(S)^2 \quad = \quad \Big[\sum_{i=1}^n \lambda_i(S) \Big]^2 \quad \leq \quad \theta \Big[\sum_{i=1}^n \lambda_i(S)^2 \Big] \quad = \quad \mathrm{tr}(S^2),$$

which is true for $\theta = n$ due to the standard inequality between norms: $\| \cdot \|_1 \leq \sqrt{n} \| \cdot \|_2$.

[2] This might happen if Q contains a line: $\{x + \tau h \;:\; \tau \in \mathbb{R}\} \subseteq Q$ for some $x \in Q$ and direction $h \in \mathbb{R}^n$. In this case, self-concordance of F implies $\nabla^2 F(y) h \equiv 0$ *for all* $y \in Q$ along this direction h. Assuming that Q does not contain lines (e.g., it is bounded), we prevent this situation, ensuring $\nabla^2 F(y) \succ 0$ for all $y \in Q$.

21.2.4 Basic Operations

Let us look at some simple properties of the self-concordant barriers.

Proposition 21.2.4. *Let $F_1, \ldots, F_m$ be self-concordant barriers for $Q_1, \ldots, Q_m$. Then,*

$$F(x) \quad = \quad \sum_{i=1}^{m} F_i(x)$$

is self-concordant barrier for

$$Q \quad = \quad \bigcap_{1 \leq i \leq m} Q_i.$$

with

$$\boxed{\theta \quad = \quad \sum_{i=1}^{m} \theta_i.}$$

Proof. It follows immediately from the bound (21.8) and the fact that the maximum of a sum does not exceed the sum of the maxima. □

Example 21.2.5. The logarithmic barrier for the positive orthant $Q = \mathbb{R}^n_{>0}$ given as the sum of barriers for rays (21.2.2),

$$F(x) \quad = \quad -\sum_{i=1}^{n} \log x^{(i)},$$

is a self-concordant barrier with parameter $\boxed{\theta = n}$.

Note that, similarly to the definition of self-concordant functions, the definition of self-concordant barriers is *affine-invariant* (check it!), and thus the parameter of self-concordance θ is preserved under affine transformation:

Proposition 21.2.6. *Let $\mathcal{A}(x) = Ax + b$ be an affine transformation, and let $F : Q \to \mathbb{R}$ be a self-concordant barrier for Q with parameter θ_F. Then,*

$$\Phi(x) \quad = \quad F(\mathcal{A}(x)), \qquad x \in \Omega,$$

is a self-concordant barrier for the affine preimage of Q:

$$\Omega \quad = \quad \Bigl\{ x \quad : \quad \mathcal{A}(x) \in Q \Bigr\},$$

with parameter $\boxed{\theta_\Phi = \theta_F}$.

Example 21.2.7. The logarithmic barrier for the polyhedron $Q = \Bigl\{ x \in \mathbb{R}^n \quad : \quad \langle a_1, x \rangle \leq b_1, \ldots, \langle a_m, x \rangle \leq b_m \Bigr\}$:

$$F(x) \quad = \quad -\sum_{i=1}^{m} \ln(b_i - \langle a_i, x \rangle)$$

is a self-concordant barrier with parameter $\boxed{\theta = m}$.

21.2.5 Key Property

It appears that the following property is central for the theory of interior-point methods. Functions that satisfy (21.12) are called *set-limited* [Nes24].

Lemma 21.2.8. *For any $x, y \in Q$, we have*

$$\langle \nabla F(x), y - x \rangle \quad \leq \quad \theta. \tag{21.12}$$

Proof. Consider $\varphi(t) := \langle \nabla F(x + t(y - x)), y - x \rangle$, for $t \in [0, 1]$. Then, our goal is to show that $\varphi(0) \leq \theta$. We have

$$\begin{aligned} \varphi'(t) \quad &= \quad \langle \nabla^2 F(x + t(y-x))(y-x), y-x \rangle \\ &\overset{(21.10)}{\geq} \quad \tfrac{1}{\theta} \langle \nabla F(x + t(y-x)), y - x \rangle^2 \quad = \quad \tfrac{1}{\theta} \varphi(t)^2. \end{aligned} \tag{21.13}$$

Thus, $\varphi(t)$ is increasing with t. We can assume that $\varphi(0) > 0$ (otherwise, (21.12) trivially holds). Therefore, we conclude

$$\tfrac{1}{\varphi(0)} \quad > \quad \tfrac{1}{\varphi(0)} - \tfrac{1}{\varphi(1)} \quad = \quad \int\limits_0^1 \tfrac{d}{dt} \Big[-\tfrac{1}{\varphi(t)} \Big] \quad = \quad \int\limits_0^1 \tfrac{\varphi'(t) dt}{\varphi(t)^2} \quad \overset{(21.13)}{\geq} \quad \tfrac{1}{\theta},$$

which completes the proof. □

Geometric interpretation of inequality (21.12) is that normalized gradient direction $\frac{1}{\theta} \nabla F(x)$ belongs to the *polar* to Q at point $x \in Q$:

$$\tfrac{1}{\theta} \nabla F(x) \quad \overset{(21.12)}{\in} \quad P_Q(x) \quad := \quad \Big\{ \, g \in \mathbb{R}^n \quad : \quad \langle g, y - x \rangle \leq 1 \;\; \forall y \in Q \Big\}.$$

Literature

[Nes24] Yurii Nesterov. Set-limited functions and polynomial-time interior-point methods. *Journal of Optimization Theory and Applications*, 202(1):11–26, 2024.

[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.