Lecture 7.

Self-Concordant Barriers

– Motivation.

– Definition of Self-Concordant Barriers.

– Main Properties.

– Standard Minimization Problem.

– Central Path.

– Path-Following Method.

– How to initialize the process?

– Problems with functional constraints.

Motivation

What we have now:

- An open convex set $\mathrm{dom}\, f$, which contains no straight line.
- A self-concordant function f. This means that $f \in C^3(\mathrm{dom}\, f)$, it is closed and convex and

$$D^3 f(x)[h, h, h] \leq 2 \parallel h \parallel_x^3,$$

  for all $x \in \mathrm{dom}\, f$ and $h \in R^n$.

  (For this function $M_f = 2$; we call such functions the *standard* self-concordant functions).
- f is a barrier for $\mathrm{dom}\, f$.
- The Newton Method is very efficient in minimization f over $\mathrm{dom}\, f$.

We want to use these objects in the scheme of *Sequential Unconstrained Minimization.*

Standard Problem

Denote

$$\operatorname{Dom} f = \operatorname{cl}(\operatorname{dom} f).$$

We call a constrained minimization problem *standard* if:

- it has the following form:

$$\min\{\langle c, x\rangle \mid x \in Q\}, \tag{7.1}$$

  where Q is a closed convex set.
- We have a self-concordant function f such that

$$\operatorname{Dom} f = Q.$$

Central path

Define the parametric penalty function

$$f(t; x) = t\langle c, x\rangle + f(x)$$

with $t \geq 0$. Note that $f(t; x)$ is self-concordant in x (see Corollary 6.1).

Denote

$$x^*(t) = \arg\min_{x \in \operatorname{dom} f} f(t; x).$$

This trajectory is called *the central path* of the problem (7.1).

Our goals:

- We can expect that $x^*(t) \to x^*$ as $t \to \infty$.
- We want to follow the central path.

Note:

- The Newton Method for function $f(t;x)$ has a local quadratic convergence.
- We have an exact description of the region of quadratic convergence ($\lambda_{f(t;\cdot)}(x) \leq \beta < \bar{\lambda}$).

Assume that we have computed $x = x^*(t)$.

We can increase t:

$$t_+ = t + \Delta, \quad \Delta > 0,$$

but we want to keep x in the region of quadratic convergence of the Newton Method for the function $f(t+\Delta;\cdot)$:

$$\lambda_{f(t+\Delta;\cdot)}(x) \leq \beta < \bar{\lambda}.$$

Note that the update $t \to t_+$ does not change the Hessian:

$$f''(t+\Delta;x) = f''(t;x).$$

Question: How large can be Δ?

Central path equation:

$$tc + f'(x^*(t)) = 0. \tag{7.2}$$

Since $tc + f'(x) = 0$, we have:

$$\begin{aligned} \lambda_{f(t+\Delta;\cdot)}(x) &= \parallel t_+ c + f'(x) \parallel_x^* \\ &= \Delta \parallel c \parallel_x = \frac{\Delta}{t} \parallel f'(x) \parallel_x^* \leq \beta. \end{aligned}$$

Conclusion:

If we want to increase t in a *linear rate*, we need to assume that the value

$$\parallel f'(x) \parallel_x^{*2} \equiv \langle [f''(x)]^{-1} f'(x), f'(x) \rangle$$

is *uniformly bounded* on dom f.

Thus, we come to the definition of *self-concordant barrier*.

Definition of Self-Concordant Barriers

Definition 7.1 *Let $F(x)$ be a standard self-concordant function. We call it a ν-*self-concordant barrier *for the set* $\mathrm{Dom}\, F$, *if*

$$\max_{u \in R^n} [2\langle F'(x), u\rangle - \langle F''(x)u, u\rangle] \leq \nu \qquad (7.3)$$

for all $x \in \mathrm{dom}\, F$.

The value ν is called the *parameter* of the barrier.

Note:

1. We don't assume that $F''(x)$ is non-degenerate.

2. If $F''(x)$ is non-degenerate, then (7.3) is equivalent to the following:

$$\langle [F''(x)]^{-1} F'(x), F'(x)\rangle \leq \nu. \qquad (7.4)$$

3. We have the following consequence of the inequality (7.3): $\forall u \in R^n$

$$\langle F'(x), u\rangle^2 \leq \nu \langle F''(x)u, u\rangle, \qquad (7.5)$$

(To see that for u with $\langle F''(x)u, u\rangle > 0$, replace u in (7.3) by λu and find the maximum of LHS in λ.)

At this moment:

- We are almost sure that we can follow the central path defined by a self-concordant barrier.
- We don't have any examples of these barriers.
- We don't know the properties of these barriers.
- We don't know for which convex sets such barriers exist.

That is our program.

Let us check our examples of self-concordant functions from Lecture 6.

1. Linear function

$$f(x) = \alpha + \langle a, x \rangle, \quad \operatorname{dom} f = R^n.$$

Clearly, it is not a self-concordant barrier since $f''(x) = 0$.

2. Convex quadratic function:

$$\begin{aligned} f(x) &= \alpha + \langle a, x \rangle + \tfrac{1}{2}\langle Ax, x \rangle, \\ & A = A^T > 0, \quad \operatorname{dom} f = R^n, \\ f'(x) &= a + Ax, \quad f''(x) = A. \end{aligned}$$

Then

$$\begin{aligned} & \langle [f(x)]^{-1} f'(x), f'(x) \rangle = \langle A^{-1}(Ax - a), Ax - a \rangle \\ & = \langle Ax, x \rangle - 2\langle a, x \rangle + \langle A^{-1}a, a \rangle. \end{aligned}$$

Clearly, this value is unbounded from above on R^n.

3. Logarithmic barrier for a ray:

$$\begin{aligned} F(x) &= -\ln x, \quad \operatorname{dom} F = \{x \in R^1 \mid \ x > 0\}, \\ F'(x) &= -\tfrac{1}{x}, \quad F''(x) = \tfrac{1}{x^2} > 0. \end{aligned}$$

Therefore

$$\frac{(F'(x))^2}{F''(x)} = \frac{1}{x^2} \cdot x^2 = 1.$$

Thus, $F(x)$ is a ν-self-concordant barrier with $\nu = 1$.

4. Logarithmic barrier for a quadratic region.

Let $A = A^T \geq 0$. Consider the *concave* function

$$\phi(x) = \alpha + \langle a, x \rangle - \tfrac{1}{2}\langle Ax, x \rangle.$$

Define

$$F(x) = -\ln \phi(x), \quad \mathrm{dom}\, f = \{x \in R^n \mid \ \phi(x) > 0\}.$$

Then

$$\begin{aligned} \langle F'(x), u \rangle = & \ -\tfrac{1}{\phi(x)}[\langle a, u \rangle - \langle Ax, u \rangle], \\ \langle F''(x)u, u \rangle = & \ \tfrac{1}{\phi^2(x)}[\langle a, u \rangle - \langle Ax, u \rangle]^2 \\ & + \tfrac{1}{\phi(x)}\langle Au, u \rangle. \end{aligned}$$

Denote

$$\omega_1 = \langle F'(x), u \rangle, \quad \omega_2 = \frac{1}{\phi(x)}\langle Au, u \rangle.$$

Then

$$\langle F''(x)u, u \rangle = \omega_1^2 + \omega_2 \geq \omega_1^2.$$

Therefore

$$2\langle F'(x), u \rangle - \langle F''(x)u, u \rangle \leq 2\omega_1 - \omega_1^2 \leq 1.$$

Thus, $F(x)$ is a ν-self-concordant barrier with $\nu = 1$.

Simple Properties

Theorem 7.1 *Let $F(x)$ be a self-concordant barrier. Then the function*

$$\langle c, x\rangle + F(x)$$

is self-concordant on $\operatorname{dom} F$.

Proof:

Since $F(x)$ is a self-concordant function, we just refer to Corollary 6.1. $\square$

Note:

This property is very important for path-following schemes.

Theorem 7.2 *Let F_i are ν_i-self-concordant barriers, $i = 1, 2$. Then*

$$F(x) = F_1(x) + F_2(x)$$

is a self-concordant barrier for the set

$$\mathrm{Dom}\, F = \mathrm{Dom}\, F_1 \bigcap \mathrm{Dom}\, F_2$$

with the parameter

$$\nu = \nu_1 + \nu_2.$$

Proof:
Note that F is a standard self-concordant function in view of Theorem 6.1.

Let us fix $x \in \mathrm{dom}\, F$. Then

$$\begin{aligned}
& \max_{u \in R^n} [\ 2\langle F'(x)u, u\rangle - \langle F''(x)u, u\rangle] \\
= & \max_{u \in R^n} [\ 2\langle F_1'(x)u, u\rangle - \langle F_1''(x)u, u\rangle \\
& \qquad +2\langle F_1'(x)u, u\rangle - \langle F_1''(x)u, u\rangle] \\
\leq & \max_{u \in R^n} [\ 2\langle F_1'(x)u, u\rangle - \langle F_1''(x)u, u\rangle] \\
+ & \max_{u \in R^n} [\ 2\langle F_2'(x)u, u\rangle - \langle F_2''(x)u, u\rangle] \leq \nu_1 + \nu_2.
\end{aligned}$$

□

Theorem 7.3 *Let $\mathcal{A}(x) = Ax+b$ be a linear operator:*

$$\mathcal{A}(x): \quad R^n \to R^m.$$

Assume that the function $F(y)$ is a ν-self-concordant barrier.
Then $\Phi(x) = F(\mathcal{A}(x))$ is a ν-self-concordant barrier for the set

$$\mathrm{Dom}\,\Phi = \{x \in R^n \mid \ \mathcal{A}(x) \in \mathrm{Dom}\,F\}.$$

Proof:
The function $\Phi(x)$ is a standard self-concordant function in view of Theorem 10.2.

Let us fix $x \in \mathrm{dom}\,\Phi$. Then $y = \mathcal{A}(x) \in \mathrm{dom}\,F$.

Note that for any $u \in R^n$ we have:

$$\begin{aligned} \langle \Phi'(x), u\rangle \ &= \langle F'(y), Au\rangle, \\ \langle \Phi''(x)u, u\rangle \ &= \langle F''(y)Au, Au\rangle. \end{aligned}$$

Therefore

$$\begin{aligned} &\max_{u\in R^n} [2\langle \Phi'(x), u\rangle - \langle \Phi''(x)u, u\rangle] \\ &= \max_{u\in R^n} [2\langle F'(y), Au\rangle - \langle F''(y)Au, Au\rangle] \\ &\leq \max_{v\in R^m} [2\langle F'(y), v\rangle - \langle F''(y)v, v\rangle] \leq \nu. \end{aligned}$$

□

Main Inequalities

Theorem 7.4 *Let $F(x)$ be a ν-self-concordant barrier.*

Then for any $x \in \operatorname{dom} F$ and $y \in \operatorname{Dom} F$ we have:

$$\langle F'(x), y - x \rangle \leq \nu. \tag{7.6}$$

Proof:
Let $x \in \operatorname{dom} F$ and $y \in \operatorname{Dom} F$. Consider the function

$$\phi(t) = \langle F'(x + t(y - x)), y - x \rangle, \quad t \in [0, 1).$$

If $\phi(0) \leq 0$, then (7.6) is trivial.

Suppose that $\phi(0) > 0$. Note that in view of (7.5) we have:

$$\begin{aligned} \phi'(t) &= \langle F''(x + t(y - x))(y - x), y - x \rangle \\ &\geq \tfrac{1}{\nu} \langle F'(x + t(y - x)), y - x \rangle^2 = \tfrac{1}{\nu} \phi^2(t). \end{aligned}$$

Therefore $\phi(t)$ increases and it is positive for $t \in [0, 1)$. Moreover,

$$-\tfrac{1}{\phi(t)} + \tfrac{1}{\phi(0)} \geq \tfrac{1}{\nu} t, \quad t \in [0, 1).$$

This means that

$$\langle F'(x), y - x \rangle = \phi(0) \leq \frac{\nu}{t}$$

for all $t \in [0, 1)$. $\square$

Theorem 7.5 *Let $F(x)$ be a ν-self-concordant barrier.*

Then for any $x \in \operatorname{dom} F$ and $y \in \operatorname{Dom} F$ such that

$$\langle F'(x), y - x\rangle \geq 0 \tag{7.7}$$

we have:

$$\| y - x \|_x \leq \nu + 2\sqrt{\nu}. \tag{7.8}$$

Proof:
Denote $r = \| y - x \|_x$. Let $r > \sqrt{\nu}$. Consider

$$y_\alpha = x + \alpha(y - x), \quad \alpha = \frac{\sqrt{\nu}}{r} < 1.$$

In view of our assumption (7.7) and Theorem 10.7 we have:

$$\begin{aligned}\omega \equiv\ & \langle F'(y_\alpha), y - x\rangle \geq \langle F'(y_\alpha) - F'(x), y - x\rangle \\ =\ & \tfrac{1}{\alpha}\langle F'(y_\alpha) - F'(x), y_\alpha - x\rangle \\ \geq\ & \tfrac{1}{\alpha} \cdot \tfrac{\|y_\alpha - x\|_x^2}{1 + \|y_\alpha - x\|_x^2} = \tfrac{\alpha \|y - x\|_x^2}{1 + \alpha\|y - x\|_x} = \tfrac{r\sqrt{\nu}}{1 + \sqrt{\nu}}.\end{aligned}$$

On the other hand, in view of (7.6), we obtain:

$$(1 - \alpha)\omega = \langle F'(y_\alpha), y - y_\alpha\rangle \leq \nu.$$

Thus,

$$\left(1 - \frac{\sqrt{\nu}}{r}\right) \frac{r\sqrt{\nu}}{1 + \sqrt{\nu}} \leq \nu,$$

and that is exactly (7.8). □

Analytic Center

Let $F(x)$ be a ν-self-concordant barrier for the set $\operatorname{Dom} F$. The point

$$x_F^* = \arg \min_{x \in \operatorname{dom} F} F(x),$$

(if exists), is called the *analytic center* of the convex set $\operatorname{Dom} F$, generated by the barrier $F(x)$.

Theorem 7.6 *Assume that the analytic center of a ν-self-concordant barrier $F(x)$ exists.*

Then for any $x \in \operatorname{Dom} F$ we have:

$$\| x - x_F^* \|_{x_F^*} \leq \nu + 2\sqrt{\nu}.$$

Moreover, for any $x \in R^n$ such that

$$\| x - x_F^* \|_{x_F^*} \leq 1$$

we have $x \in \operatorname{Dom} F$.

Proof:
The first statement follows from Theorem 7.5 since $F'(x_F^*) = 0$.
The second statement follows from Theorem 6.5. □

Thus, the *asphericity* of the set $\operatorname{Dom} F$ in the metric $\| \cdot \|_{x_F^*}$ with respect to x_F^* is

$$\nu + 2\sqrt{\nu}.$$

Remarks:

1. If Dom F contains no straight line, the existence of x_F^* implies the boundedness of Dom F.

(Since then $F''(x_F^*)$ is nondegenerate, see Theorem 6.3.)

2. For a general convex set in R^n the asphericity does not exceed n (John Theorem).

3. We obtained a bound for the asphericity in terms of *parameter* of the barrier.

This value does not depend on the dimension.

Corollary 7.1 *Let* $\mathrm{Dom}\, F$ *be bounded. Then for any* $x \in \mathrm{dom}\, F$ *and* $v \in R^n$ *we have:*

$$\| v \|_x^* \leq (\nu + 2\sqrt{\nu}) \| v \|_{x_F^*}^* .$$

Proof:
Let us show that

$$\begin{aligned}\| v \|_x^* &\equiv \langle [F''(x)]^{-1} v, v\rangle^{1/2} \\ &= \max\{\langle v, u\rangle \mid \langle F''(x)u, u\rangle \leq 1\}.\end{aligned}$$

Indeed, in view of Theorem 4.17, the solution of this problem u^* satisfies the condition:

$$v = \lambda^* F''(x) u^*, \quad \langle F''(x)u^*, u^*\rangle = 1.$$

Therefore $\langle v, u^*\rangle = \langle [F''(x)]^{-1} v, v\rangle^{1/2}$.

Further, in view of Theorem 6.5 and Theorem 7.6,

$$\begin{aligned} B &\equiv \{y \in R^n \mid \| y - x \|_x \leq 1\} \subseteq \mathrm{Dom}\, F \\ &\subseteq \{y \in R^n \mid \| y - x_F^* \|_x \leq \nu + 2\sqrt{\nu}\} \equiv B_*.\end{aligned}$$

Therefore, using again Theorem 7.6, we have:

$$\begin{aligned}\| v \|_x^* &= \max\{\langle v, y - x\rangle \mid y \in B\} \\ &\leq \max\{\langle v, y - x\rangle \mid y \in B_*\} \\ &= \langle v, x_F^* - x\rangle + (\nu + 2\sqrt{\nu}) \| v \|_{x_F^*}^* .\end{aligned}$$

Note that $\| v \|_x^* = \| -v \|_x^*$. Therefore we can assume that $\langle v, x_F^* - x\rangle \leq 0$. □

Path-following Scheme

Standard Problem: (Mediator)

$$\min\{\langle c, x\rangle \mid \ x \in Q\}, \tag{7.9}$$

where

- $Q \equiv \operatorname{Dom} F$ is a *bounded* closed convex set with nonempty interior;
- F is a ν-self-concordant barrier.

We are going to solve (7.9) by following the *central path*:

$$x^*(t) = \arg \min_{x \in \operatorname{dom} F} \ f(t; x), \tag{7.10}$$

where $f(t; x) = t\langle c, x\rangle + F(x)$ and $t \geq 0$.

Note:

1. Any point of the central path satisfies the equation

$$tc + F'(x^*(t)) = 0. \tag{7.11}$$

2. Since Q is bounded, x^*_F exists and

$$x^*(0) = x^*_F. \tag{7.12}$$

We are going to update the points, satisfying the *approximate centering condition*:

$$\lambda_{f(t;\cdot)}(x) \equiv \| f'(t;x) \|_x^* = \| tc + F'(x) \|_x^* \leq \beta, \quad (7.13)$$

where β is small enough.

Let us show that this is a reasonable goal.

Theorem 7.7 *For any $t > 0$ we have*

$$\langle c, x^*(t) \rangle - c^* \leq \frac{\nu}{t}, \quad (7.14)$$

where c^* is the optimal value of (7.9).

If x satisfies (7.13), then

$$\langle c, x \rangle - c^* \leq \tfrac{1}{t}\left(\nu + \tfrac{(\beta+\sqrt{\nu})\beta}{1-\beta}\right). \quad (7.15)$$

Proof:
Let x^* be a solution to (7.9). In view of (7.11) and (7.6) we have:

$$\langle c, x^*(t) - x^* \rangle = \tfrac{1}{t}\langle F'(x^*(t)), x^* - x^*(t) \rangle \leq \tfrac{\nu}{t}.$$

Further, let x satisfy (7.13). Denote $\lambda = \lambda_{f(t;\cdot)}(x)$. Then

$$t\langle c, x - x^*(t) \rangle = \langle f'(t;x) - F'(x), x - x^*(t) \rangle$$

$$\leq (\lambda + \sqrt{\nu}) \| x - x^*(t) \|_x \leq (\lambda + \sqrt{\nu})\tfrac{\lambda}{1-\lambda} \leq \tfrac{(\beta+\sqrt{\nu})\beta}{1-\beta}$$

in view of (7.4), Theorem 6.11 and (7.13). □

One step of Path-following Scheme

Assume that $x \in \operatorname{dom} F$.
Let us analyze the following iterate:

$$\left.\begin{aligned} t_+ &= t + \tfrac{\gamma}{\|c\|_x^*}, \\ x_+ &= x - [F''(x)]^{-1}(t_+ c + F'(x)). \end{aligned}\right\} \qquad (7.16)$$

Theorem 7.8 *Let x satisfy (7.13):*

$$\| \ tc + F'(x) \ \|_x^* \leq \beta$$

with $\beta < \bar{\lambda} = \frac{3-\sqrt{5}}{2}$. Then for γ, such that

$$| \ \gamma \ | \leq \tfrac{\sqrt{\beta}}{1+\sqrt{\beta}} - \beta, \qquad (7.17)$$

we have again $\| \ t_+ c + F'(x_+) \ \|_x^* \leq \beta$.

Proof:
Denote $\lambda_0 = \| \ tc + F'(x) \ \|_x^* \leq \beta$,

$$\lambda_1 = \| \ t_+ c + F'(x) \ \|_x^*, \quad \lambda_+ = \| \ t_+ c + F'(x_+) \ \|_{x_+}^* .$$

Then $\lambda_1 \leq \lambda_0 + | \ \gamma \ | \leq \beta + | \ \gamma \ |$ and

$$\lambda_+ \leq \left(\tfrac{\lambda_1}{1-\lambda_1} \right)^2 \equiv [\omega_*'(\lambda_1)]^2$$

in view of Theorem 6.12. It remains to note that (7.17) is equivalent to

$$\omega_*'(\beta + | \ \gamma \ |) \leq \sqrt{\beta},$$

(recall that $\omega'(\omega_*'(\tau)) = \tau$). □

Lemma 7.1 *Let x satisfy (7.13) then*

$$\| c \|_x^* \leq \frac{1}{t}(\beta + \sqrt{\nu}). \tag{7.18}$$

Proof:
Indeed, in view of (7.13) and (7.4), we have

$$\begin{aligned} t \| c \|_x^* &= \| f'(t;x) - F'(x) \|_x^* \\ &\leq \| f'(t;x) \|_x^* + \| F'(x) \|_x^* \leq \beta + \sqrt{\nu}. \end{aligned}$$

□

Let us fix the parameters:

$$\beta = \tfrac{1}{9}, \quad \gamma = \tfrac{\sqrt{\beta}}{1+\sqrt{\beta}} - \beta = \tfrac{5}{36}. \tag{7.19}$$

Conclusion:

1. We can follow the central path, using the rule (7.16).

2. The possible rate for *increasing* t is

$$t_+ \geq \left(1 + \frac{5}{4 + 36\sqrt{\nu}}\right) \cdot t.$$

3. The possible rate for *decreasing* t is

$$t_+ \geq \left(1 - \frac{5}{4 + 36\sqrt{\nu}}\right) \cdot t.$$

Main Process (7.20)

0. Set $t_0 = 0$. Choose an accuracy $\epsilon > 0$ and $x_0 \in \mathrm{dom}\, F$ such that
$$\| F'(x_0) \|^*_{x_0} \leq \beta.$$

1. kth iteration ($k \geq 0$).
$$\text{Set} \quad t_{k+1} = t_k + \frac{\gamma}{\|c\|^*_{x_k}},$$
$$x_{k+1} = x_k - [F''(x_k)]^{-1}(t_{k+1}c + F'(x_k)).$$

2. Stop the process if $\nu + \frac{(\beta+\sqrt{\nu})\beta}{1-\beta} \leq \epsilon t_k$. □

Theorem 7.9 *The scheme (7.20) terminates no more than after N steps,*
$$N \leq O\left(\sqrt{\nu} \ln \frac{\nu \| c \|^*_{x^*_F}}{\epsilon}\right).$$
Moreover, $\langle c, x_N \rangle - c^* \leq \epsilon$.

Proof:
Note that $r_0 \equiv \| x_0 - x^*_F \|_{x_0} \leq \frac{\beta}{1-\beta}$ (T. 10.11). Therefore, in view of Theorem 10.6, we have:
$$\frac{\gamma}{t_1} = \| c \|^*_{x_0} \leq \frac{1}{1-r_0} \| c \|^*_{x^*_F} \leq \frac{1-\beta}{1-2\beta} \| c \|^*_{x^*_F}.$$
Thus, $t_k \geq \frac{\gamma(1-2\beta)}{(1-\beta)\|c\|^*_{x^*_F}} \left(1 + \frac{\gamma}{\beta+\sqrt{\nu}}\right)^{k-1}$, $k \geq 1$. □

Remarks:

1. The main term in the complexity estimate is

$$7.2\sqrt{\nu}\ln\frac{\nu \parallel c \parallel^*_{x^*_F}}{\epsilon}.$$

The value $\nu \parallel c \parallel^*_{x^*_F}$ estimates the variation of $\langle c, x\rangle$ on $\mathrm{Dom}\, F$ (see Theorem 7.6). Thus,

$$\frac{\epsilon}{\nu \parallel c \parallel^*_{x^*_F}}$$

is a *relative accuracy* of the solution.

2. It is not too often, when the starting condition

$$\parallel F'(x_0) \parallel^*_{x_0} \leq \beta$$

can be easily satisfied.

Therefore we need an additional process for *finding* such starting point.

Namely, we need a process for finding an approximation to the *analytic center* of the set $\mathrm{Dom}\, F$.

Finding the analytic center

Problem:

$$\min\{F(x) \mid \; x \in \operatorname{dom} F\}, \tag{7.21}$$

where F is a ν-self-concordant barrier.

Approximate solution:

Find $\bar{x} \in \operatorname{dom} F$ such that

$$\| F'(\bar{x}) \|^*_{\bar{x}} \leq \beta,$$

where $\beta \in (0,1)$.

There are two possibilities:

1. Damped Newton Method.
2. Path-following scheme.

Damped Newton Method

0. Choose $y_0 \in \mathrm{dom}\, F$.
1. kth iteration ($k \geq 0$).

   Set
$$y_{k+1} = y_k - \frac{[F''(y_k)]^{-1} F'(y_k)}{1+ \parallel F'(y_k) \parallel_{y_k}^*}. \tag{7.22}$$
2. Stop the process if $\parallel F'(y_k) \parallel_{y_k}^* \leq \beta$. □

Theorem 7.10 *The process (7.22) terminates no more than after*
$$\frac{1}{\omega(\beta)}(F(y_0) - F(x_F^*))$$
iterations.

Proof:
Indeed, in view of Theorem 6.10, we have:
$$F(y_{k+1}) \leq F(y_k) - \omega(\lambda_F(y_k)) \leq F(y_k) - \omega(\beta).$$
Therefore
$$F(y_0) - k\,\omega(\beta) \geq F(y_k) \geq F(x_F^*).$$
□

Auxiliary path-following scheme

Let us choose $y_0 \in \operatorname{dom} F$.

Define the *auxiliary* central path:

$$y^*(t) = \arg \min_{y \in \operatorname{dom} F} [-t\langle F'(y_0), y\rangle + F(y)],$$

where $t \geq 0$.

This trajectory is a solution of the equation:

$$F'(y^*(t)) = tF'(y_0). \tag{7.23}$$

Therefore

$$\begin{aligned} y^*(1) &= y_0, \\ y^*(0) &= x_F^*. \end{aligned}$$

Note:

We can follow this trajectory by the process (7.16) with *decreasing* t.

Let us describe the convergence of $y^*(t)$ to the analytic center.

Lemma 7.2 *For any $t \geq 0$ we have:*

$$\| F'(y^*(t)) \|_{y^*(t)}^* \leq (\nu + 2\sqrt{\nu}) \| F'(x_0) \|_{x_F^*}^* \cdot t.$$

Proof:

This estimate follows from the equation (7.23) and Corollary 7.1. □

Auxiliary Process (7.24)

0. Choose $y_0 \in \operatorname{Dom} F$. Set $t_0 = 1$.
1. kth iteration ($k \geq 0$).

   Set

$$\begin{aligned} t_{k+1} &= t_k - \tfrac{\gamma}{\|F'(y_0)\|^*_{y_k}}, \\ y_{k+1} &= y_k - [F''(y_k)]^{-1}(t_{k+1}F'(y_0) + F'(y_k)). \end{aligned}$$

2. Stop the process if $\| F'(y_k) \|_{y_k} \leq \frac{\sqrt{\beta}}{1+\sqrt{\beta}}$.

   Set $\bar{x} = y_k - [F''(y_k)]^{-1}F'(y_k)$. □

Comments:

1. This scheme follows the trajectory $y^*(t)$:

$$\| t_k F'(y_0) + F'(y_k) \|_{y_k} \leq \beta,$$

as $t_k \to 0$.

2. The termination criterion,

$$\lambda_k = \| F'(y_k) \|_{y_k} \leq \frac{\sqrt{\beta}}{1+\sqrt{\beta}},$$

guarantees that

$$\| F'(\bar{x}) \|_{\bar{x}} \leq \left(\frac{\lambda_k}{1-\lambda_k}\right)^2 \leq \beta$$

(see Theorem 6.12).

Theorem 7.11 *The process (7.24) terminates at most after*

$$\frac{1}{\gamma}(\beta+\sqrt{\nu})\ln\left[\frac{1}{\gamma}(\nu+2\sqrt{\nu}) \parallel F'(x_0) \parallel^*_{x^*_F}\right]$$

iterations.

Proof:
Recall that we have fixed the parameters:

$$\beta = \frac{1}{9}, \quad \gamma = \frac{\sqrt{\beta}}{1+\sqrt{\beta}} - \beta = \frac{5}{36}.$$

Note that $t_0 = 1$. Therefore, in view of Theorem 7.8 and Lemma 7.1, we have:

$$t_{k+1} \leq \left(1 - \frac{\gamma}{\beta+\sqrt{\nu}}\right) t_k \leq \exp\left(-\frac{\gamma(k+1)}{\beta+\sqrt{\nu}}\right).$$

Further, in view of Lemma 7.2, we obtain:

$$\begin{aligned} \parallel F'(y_k) \parallel^*_{y_k} &= \parallel (t_k F'(x_0) + F'(y_k)) - t_k F'(x_0) \parallel^*_{y_k} \\ &\leq \beta + t_k \parallel F'(x_0) \parallel^*_{y_k} \\ &\leq \beta + t_k(\nu + 2\sqrt{\nu}) \parallel F'(x_0) \parallel^*_{x^*_F}. \end{aligned}$$

Thus, the process is terminated at most when the following inequality holds:

$$t_k(\nu+2\sqrt{\nu}) \parallel F'(x_0) \parallel^*_{x^*_F} \leq \frac{\sqrt{\beta}}{1+\sqrt{\beta}} - \beta = \gamma.$$

□

Remarks:

1. The principal term in the complexity is

$$7.2\sqrt{\nu}[\ln \nu + \ln \parallel F'(x_0) \parallel^*_{x^*_F}]$$

for the auxiliary path-following scheme and

$$O(F(y_0) - F(x^*_F))$$

for the auxiliary Damped Newton Method.

Now we cannot compare these estimates.

However, a more sophisticated analysis demonstrates the advantages of path-following scheme.

2. The principal term in the total complexity of the path-following scheme (7.20) with the auxiliary process (7.24) is

$$7.2\sqrt{\nu}\left[2\ln \nu + \ln \parallel F'(x_0) \parallel^*_{x^*_F} + \ln \parallel c \parallel^*_{x^*_F} + \ln \tfrac{1}{\epsilon}\right].$$

3. For some problems it is difficult to find a starting point $y_0 \in \operatorname{dom} F$.

To do that, we can introduce one more auxiliary process.

Problems with functional constraints

Problem:

$$\begin{aligned} &\min \; f_0(x), \\ \text{s.t} \; &f_j(x) \leq 0, \; j = 1 \ldots m, \\ &x \in Q, \end{aligned} \tag{7.25}$$

where

- Q is a bounded closed convex set, $\operatorname{int} Q \neq \emptyset$.
- f_j, $j = 0 \ldots m$, are convex.

Assumptions:

- $\exists \bar{x} \in \operatorname{int} Q$: $f_j(\bar{x}) < 0$, $j = 1 \ldots m$.
- We know $\bar{\tau}$: $f_0(x) < \bar{\tau}$ for all $x \in Q$.

Equivalent standard problem:

$$\begin{aligned} &\min \; \tau, \\ \text{s.t} \; &f_0(x) \leq \tau, \\ &f_j(x) \leq \kappa, \; j = 1 \ldots m, \\ &x \in Q, \; \tau \leq \bar{\tau}, \; \kappa \leq 0. \end{aligned} \tag{7.26}$$

We need the following self-concordant barriers:

- $F_Q(x)$ for the set Q.
- $F_0(x, \tau)$ for the epigraph of $f_0(x)$.
- $F_j(x, \kappa)$ for the epigraphs of $f_j(x)$, $j = 1 \dots m$.

Assume that we can do that. Then we can use

$$\hat{F}(x, \tau, \kappa) = F_Q(x) + F_0(x, \tau) + \sum_{j=1}^{m} F_j(x, \kappa) - \ln(\bar{\tau} - \tau) - \ln(-\kappa).$$

$$\hat{\nu} = \nu_Q + \nu_0 + \sum_{j=1}^{m} \nu_j + 2. \tag{7.27}$$

Note:

- It could be difficult to find a point from $\text{dom}\, \hat{F}$.
- If $x_0 \in \text{int}\, Q$, then for large τ_0 and κ_0 we have:

  $$f_0(x_0) < \tau_0 < \bar{\tau}, \quad f_j(x_0) < \kappa_0, \; j = 1 \dots m.$$

  But $\kappa \leq 0$ could be violated.

New notation:

$$\begin{aligned} & \min \langle c, z\rangle, \\ \text{s.t. } & z \in S, \\ & \langle d, z\rangle \le 0, \end{aligned} \tag{7.28}$$

where

- $z = (x, \tau, \kappa)$,
- $\langle c, z\rangle \equiv \tau$,
- $\langle d, z\rangle \equiv \kappa$,
- S is the feasible set (7.26) without the constraint $\kappa \le 0$.

We know:

- A self-concordant barrier $F(z)$ for S.
- A point $z_0 \in \operatorname{int} S$.
- The set $S(\alpha) = \{z \in S \mid \langle d, z\rangle \le \alpha\}$ is bounded.
- For α large enough $\operatorname{int} S(\alpha) \ne \emptyset$.

We solve (7.28) in *three* stages.

First Stage:

1. Choose a starting point $z_0 \in \mathrm{int}\, S$.

2. Choose an initial gap $\Delta > 0$.
Set $\alpha = \langle d, z_0 \rangle + \Delta$.
If $\alpha \leq 0$, apply the two-stage scheme.

3. Otherwise, find an approximate analytic center of $S(\alpha)$, using

$$\tilde{F}(z) = F(z) - \ln(\alpha - \langle d, z \rangle).$$

We find a point $\tilde{z}$ satisfying the condition

$$\langle \tilde{F}''(\tilde{z})^{-1} \left(F'(\tilde{z}) + \tfrac{d}{\alpha - \langle d, \tilde{z} \rangle} \right), F'(\tilde{z}) + \tfrac{d}{\alpha - \langle d, \tilde{z} \rangle} \rangle^{1/2}$$

$$\equiv \lambda_{\tilde{F}}(\tilde{z}) \leq \beta.$$

To do that, we can apply the auxiliary path-following schemes.

Second Stage:

Follow the central path $z(t)$:

$$td + \tilde{F}'(z(t)) = 0, \quad t \geq 0.$$

Note:

- The first stage computes an approximation to $z(0)$, the analytic center $S(\alpha)$.
- We can follow this path, using (7.16).
- This trajectory leads us to the solution of
$$\min\{\langle d, z\rangle \mid \ z \in S(\alpha)\} \quad (< 0)$$
(Slater condition for problem (7.28)).

Goal:

Find an approximation to the analytic center of

$$\bar{S} = \{z \in S(\alpha) \mid \ \langle d, z\rangle \leq 0\},$$

generated by the barrier $\bar{F}(z) = \tilde{F}(z) - \ln(-\langle d, z\rangle)$.

For this point, z_*, we have:

$$\tilde{F}'(z_*) - \tfrac{d}{\langle d, z_*\rangle} = 0.$$

Therefore $\exists t_* : \ z_* = z(t_*)$. Namely, $t_* = -\frac{1}{\langle d, z_*\rangle} > 0$.

This stage computes $\bar{z}$:

$$\langle \tilde{F}''(\tilde{z})^{-1}\left(\tilde{F}'(\bar{z}) - \tfrac{d}{\langle d, \bar{z}\rangle}\right), \tilde{F}'(\bar{z}) - \tfrac{d}{\langle d, \bar{z}\rangle}\rangle^{1/2}$$

$$\equiv \lambda_{\tilde{F}}(\bar{z}) \leq \beta.$$

Third Stage:

Since $\bar{F}''(z) > \tilde{F}''(z)$, we have:

$$\langle \bar{F}''(\tilde{z})^{-1}\left(\tilde{F}'(\bar{z}) - \tfrac{d}{\langle d,\bar{z}\rangle}\right), \tilde{F}'(\bar{z}) - \tfrac{d}{\langle d,\bar{z}\rangle}\rangle^{1/2}$$
$$\equiv \lambda_{\bar{F}}(\bar{z}) \leq \beta.$$

Therefore:

- $\bar{z}$ is a good approximation to the analytic center of the set $\bar{S}$.
- We can apply the main path-following scheme (7.20) to solve the problem

  $$\min\{\langle c, z\rangle \mid z \in \bar{S}\}.$$

  This problem is equivalent to (7.28).

Remarks:

1. We omit the detailed complexity analysis.

2. The main term in the complexity of this scheme

$$\sqrt{\hat{\nu}} \cdot \left(\ln \tfrac{1}{\epsilon} + \ \ldots\ \right),$$

(see (7.27)), where "…" are the logarithms of some structural characteristics of the problem (size of the region, deepness of the Slater condition, etc.).

Note: We need to know how to construct the self-concordant barriers, hopefully with *small* value of the parameter.