Interior-Point Theory for Convex Optimization

Robert M. Freund

April, 2007

©2007 Massachusetts Institute of Technology. All rights reserved.
1 Background

The material presented herein is based on the following two research texts:

• *Interior-Point Polynomial Algorithms in Convex Programming* by Yurii Nesterov and Arkadii Nemirovskii, SIAM 1994, and


2 Barrier Scheme for Solving Convex Optimization

Our problem of interest is

$$P : \text{minimize}_x \ c^T x$$

s.t. \( x \in S \),

where \( S \) is some closed convex set, and denote the optimal objective value by \( V^* \). Let \( f(\cdot) \) be a barrier function for \( S \), namely \( f(\cdot) \) satisfies:

• \( f(\cdot) \) is strictly convex on its domain \( D_f := \text{int}S \)

• \( f(x) \to \infty \) as \( x \to \partial S \).

The idea of the barrier method is to dissuade the algorithm from computing points too close to \( \partial S \), effectively eliminating the complicating factors of dealing with \( \partial S \). For every value of \( \mu > 0 \) we create the barrier problem:

$$P_\mu : \text{minimize}_x \ \mu c^T x + f(x)$$

s.t. \( x \in D_f \).

Note that \( P_\mu \) is effectively unconstrained, since the boundary of the feasible region will never be encountered. The solution of \( P_\mu \) is denoted \( z(\mu) \):
\[ z(\mu) := \arg \min_x \{ \mu c^T x + f(x) : x \in D_f \} . \]

Intuitively, as \( \mu \to \infty \), the impact of the barrier function on the solution of \( P_\mu \) should become less and less, so we should have \( c^T z(\mu) \to V^* \) as \( \mu \to \infty \). Presuming this is the case, the barrier scheme tries to use Newton’s method to solve for approximate solutions \( x^i \) of \( P_{\mu_i} \) for an increasing sequence of values of \( \mu_i \to \infty \).

In order to be more specific about how the barrier scheme might work, let us assume that at each iteration we have some value \( x \in D_f \) that is an approximate solution of \( P_\mu \) for a given value \( \mu > 0 \). We will, of course, need a way to define “is an approximate solution of \( P_\mu \)” that will be developed later. We then will increase the barrier parameter \( \mu \) by a multiplicative factor \( \alpha > 1 \):

\[ \hat{\mu} \leftarrow \alpha \mu . \]

Then we will take a Newton step at \( x \) for the problem \( P_{\hat{\mu}} \) to obtain a new point \( \hat{x} \) that we would like to then be an approximate solution of \( P_{\hat{\mu}} \). If so, we can continue the scheme recursively.

We typically use \( g(\cdot) \) and \( H(\cdot) \) to denote the gradient and Hessian of \( f(\cdot) \). Note that the Newton iterate for \( P_{\hat{\mu}} \) has the formula:

\[ \hat{x} \leftarrow x - H(x)^{-1}(\hat{\mu}c + g(x)) . \]

The general algorithmic scheme is presented in Table 1.

### 3 Some Plain Facts

Let \( f(\cdot) : \mathbb{R}^n \to \mathbb{R} \) be a twice-differentiable function. We typically use \( g(\cdot) \) and \( H(\cdot) \) denote the gradient and Hessian of \( f(\cdot) \).

**Fact 3.1** \( g(y) = g(x) + \int_0^1 H(x + t(y-x))(y-x)dt \)

**Fact 3.2** Let \( h(t) := f(x + tv) \). Then
**General Barrier Scheme**

**Step 1 Initialize.** Define $\alpha > 1$.
Initialize with $\mu_0 > 0$, $x^0 \in D_f$ that is “an approximate solution of $P_{\mu_0}$.” Set $x \leftarrow x^0$, $\mu \leftarrow \mu_0$, and $i \leftarrow 0$.

**Step 2 Increase $\mu$ and take Newton step.**
$\hat{\mu} \leftarrow \alpha \mu$

$\hat{x} \leftarrow x - H(x)^{-1}(\hat{\mu} c + g(x))$

**Step 3 Update counter and repeat.** $x \leftarrow \hat{x}$, $\mu \leftarrow \hat{\mu}$,
$i \leftarrow i + 1$ and Goto Step 2.

| Table 1: The General Barrier Scheme. |

- $h'(t) = g(x + tv)^T v$
- $h''(t) = v^T H(x + tv)v$

**Fact 3.3**

$$f(y) = f(x) + g(x)^T (y - x) + \frac{1}{2} (y - x)^T H(x)(y - x)$$

$$+ \int_0^1 \int_0^t (y - x)[H(x + s(y - x)) - H(x)](y - x) \, ds \, dt$$

**Fact 3.4**

$$\int_0^r \frac{1}{(1-at)^2} - 1 \, dt = \frac{ar^2}{1-ar}$$

This follows by observing that $\int \frac{1}{(1-at)^2} \, dt = \frac{1}{a(1-at)}$. 

4
Fact 3.5 Suppose $f(\cdot)$ is a convex function on $\mathbb{R}^n$, and $S \subset \mathbb{R}^n$ is a compact convex set, and suppose $x \in \text{int}S$ satisfies $f(x) \leq f(y)$ for all $y \in \partial S$. Then $f(\cdot)$ attains its global minimizer on $S$.

Fact 3.6 Let $\|v\| := \sqrt{v^Tv}$ be the Euclidean norm. Let $\lambda_1 \leq \ldots \leq \lambda_n$ be the ordered eigenvalues of the symmetric matrix $M$, and define $\|M\| := \max\{|\lambda_i| : \|v\| \leq 1\}$. Then $\|M\| = \max_i\{\lambda_i\} = \max\{|\lambda_n|, |\lambda_1|\}$.

Fact 3.7 Suppose $A, B$ are symmetric and $A + B = \theta I$ for some $\theta \in \mathbb{R}$. Then $AB = BA$. Furthermore, if $A \succeq 0, B \succeq 0$, then $A^{\frac{1}{2}}B^{\frac{1}{2}} = B^{\frac{1}{2}}A^{\frac{1}{2}}$.

To see why this is true, decompose $A = PDP^T$ where $P$ is orthonormal ($P^T = P^{-1}$) and $D$ is diagonal. Then $B = P(\theta I - D)^{-1}P^T$, whereby $AB = PD^{\frac{1}{2}}P^T = P(\theta I - D)P^T = P(\theta I - D)D^{\frac{1}{2}}P^T = BA$. If $A \succeq 0, B \succeq 0$, then $A^{\frac{1}{2}} = PD^{\frac{1}{2}}P^T, B^{\frac{1}{2}} = (\theta I - D)^{\frac{1}{2}}P^T$, and similarly $A^{\frac{1}{2}}B^{\frac{1}{2}} = PD^{\frac{1}{2}}P^T(\theta I - D)^{\frac{1}{2}}P^T = PD^{\frac{1}{2}}(\theta I - D)^{\frac{1}{2}}P^T = P(\theta I - D)D^{\frac{1}{2}}P^T = P(\theta I - D)^{\frac{1}{2}}P^TPD^{\frac{1}{2}}P^T = B^{\frac{1}{2}}A^{\frac{1}{2}}$.

Fact 3.8 Suppose $\lambda_n \geq \ldots \geq \lambda_1 > 0$. Then

$$\max_i\{\lambda_i - 1\} \leq \max\{\lambda_n - 1, 1/\lambda_1 - 1\}.$$

Fact 3.9 Suppose $a, b, c, d > 0$. Then

$$\min\left\{\frac{a}{b}, \frac{c}{d}\right\} \leq \frac{a + c}{b + d} \leq \max\left\{\frac{a}{b}, \frac{c}{d}\right\}.$$

4 Self-Concordant Functions and Properties

Let $f(\cdot)$ be a strictly convex twice-differentiable function defined on the open set $D_f := \text{domain}f(\cdot)$ and let $\bar{D}_f := \text{cl} D_f$. Consider $x \in D_f$. We will often abbreviate $H_x := H(x)$ for the Hessian at $x$. Notice that $H_x \succ 0$ and so can be used to define the norm

$$\|v\|_x := \sqrt{v^T H_x v}.$$
which is the “local norm” at $x$. Let
\[ B_x(x,1) := \{ y : \|y - x\|_x < 1 \} . \]
This is called the open Dikin ball at $x$ after the Russian mathematician I.I.Dikin.

**Definition 4.1** $f(\cdot)$ is said to be (strongly nondegenerate) self-concordant if for all $x \in D_f$ we have $B_x(x,1) \subset D_f$, and for all $y \in B_x(x,1)$ we have:
\[
1 - \|y - x\|_x \leq \frac{\|v\|_y}{\|v\|_x} \leq \frac{1}{1 - \|y - x\|_x}
\]
for all $v \neq 0$.

Let $SC$ denote the class of all such functions.

**Remark 1** The following are the most-used self-concordant functions:

- $f(x) = -\ln(x)$ for $x \in D_f = \{ x \in \mathbb{R} : x > 0 \}$
- $f(X) = -\ln \det(X)$ for $X \in D_f = \{ X \in S^{k \times k} : X \succ 0 \}$
- $f(x) = -\ln(x_1^2 - \sum_{i=2}^{\infty} x_j^2)$ for $x \in D_f := \{ x : \|(x_2, \ldots, x_n)\| \leq x_1 \}$

Before showing that these functions are self-concordant, let us see how we can combine self-concordant functions to obtain other self-concordant functions.

**Proposition 4.1** (self-concordance under addition/intersection) Suppose that $f_i(\cdot) \in SC$ with domain $D_i := D_{f_i}$ for $i = 1, 2$, and suppose that $D := D_1 \cap D_2 \neq \emptyset$. Define $f(\cdot) = f_1(\cdot) + f_2(\cdot)$. Then $D_f = D$ and $f(\cdot) \in SC$.

**Proof:** Consider $x \in D = D_1 \cap D_2$. Let $B^i_x(c,r)$ denote the Dikin ball centered at $c$ with radius $r$ defined by $f_i(\cdot)$ and let $\|\cdot\|_{x,i}$ denote the norm induced at $x$ using the Hessian $H_i(x)$ of $f_i(x)$ for $i = 1, 2$. Then since $x \in D_i$ we have $B^i_x(x,1) \subset D_i$ and $\|v\|^2_2 = \|v\|^2_{x,1} + \|v\|^2_{x,2}$. Because $H(x) =$
\[ H_1(x) + H_2(x). \text{ Therefore if } \|y - x\|_x < 1 \text{ it follows that } \|y - x\|_{x,1} < 1 \text{ and } \|y - x\|_{x,2} < 1, \text{ whereby } y \in B^i_x(x, 1) \subset D_i \text{ and hence } y \in D_1 \cap D_2 = D. \]

Also, for any \( v \neq 0 \), using Fact 3.9 we have

\[
\frac{\|v\|_y^2}{\|v\|_x^2} = \frac{\|v\|_{y,1}^2 + \|v\|_{y,2}^2}{\|v\|_{x,1}^2 + \|v\|_{x,2}^2} \leq \max \left\{ \frac{\|v\|_{y,1}^2}{\|v\|_{x,1}^2}, \frac{\|v\|_{y,2}^2}{\|v\|_{x,2}^2} \right\}
\]

\[
\leq \max \left\{ \left( \frac{1}{1 - \|y - x\|_{x,1}} \right)^2, \left( \frac{1}{1 - \|y - x\|_{x,2}} \right)^2 \right\}
\]

\[
\leq \left( \frac{1}{1 - \|y - x\|_x} \right)^2.
\]

The virtually identical argument can also be applied to prove the “\( \geq \)” inequality of the definition of self-concordance by replacing “max” by “min” above and applying the other inequality of Fact 3.9. \( \blacksquare \)

**Proposition 4.2 (self-concordance under affine transformation)** Let \( A \in \mathbb{R}^{m \times n} \) satisfy \( \text{rank } A = n \leq m \). Suppose that \( f(\cdot) \in SC \) with domain \( D_f \subset \mathbb{R}^m \) and define \( \hat{f}(\cdot) : \hat{f}(x) = f(Ax - b) \). Then \( \hat{f}(\cdot) \in SC \) with domain \( \hat{D} := \{ x : Ax - b \in D_f \} \).

**Proof:** Consider \( x \in \hat{D} \) and \( s = Ax - b \). Letting \( g(s) \) and \( H(s) \) denote the gradient and Hessian of \( f(s) \) and \( \hat{g}(x) \) and \( H(x) \) the gradient and Hessian of \( \hat{f}(x) \), we have \( \hat{g}(x) = A^T g(s) \) and \( H(x) = A^T H(s) A \). Suppose that \( \|y - x\|_x < 1 \). Then defining \( t := Ay - b \) we have \( 1 > \|y - x\|_x = \sqrt{(y^T A^T - x^T A^T) H(s)(Ay - Ax)} = \|t - s\|_s \), whereby \( t \in D_f \) and so \( y \in \hat{D} \). Therefore \( B_x(x, 1) \subset \hat{D} \). Also, for any \( v \neq 0 \), we have

\[
\frac{\|v\|_y}{\|v\|_x} = \sqrt{\frac{v^T A^T H(t) A v}{v^T A^T H(s) A v}} = \frac{\|Av\|_s}{\|Av\|_s} \leq \frac{1}{1 - \|s - t\|_s} = \frac{1}{1 - \|y - x\|_x}.
\]

The exact same argument can also be applied to prove the “\( \geq \)” inequality of the definition of self-concordance. \( \blacksquare \)

**Proposition 4.3** The three functions defined in Remark 1 are self-concordant.
Proof: We will prove that \( f(X) := -\ln \det(X) \) is self-concordant on its domain \( \{ X \in S^{k \times k} : X \succ 0 \} \). When \( k = 1 \), this is the logarithmic barrier function. Although it is true, we will not prove that \( f(x) = -\ln(x_1^2 - \sum_{j=2}^n x_j^2) \) is a self-concordant barrier for the interior of the second-order cone \( Q^n := \{ x : \| (x_2, \ldots, x_n) \| \leq x_1 \} \), as this proof is arithmetically uninspiring.

To prove \( f(X) := -\ln \det(X) \) is self-concordant, let \( X \succ 0 \) be given, and let \( Y \in B_X(X, 1) \) and \( V \in S^{k \times k} \) be given. We need to verify three statements:

1. \( Y \succ 0 \),
2. \( \frac{\| V \|_Y}{\| V \|_X} \leq \frac{1}{1 - \| Y - X \|_X} \), and
3. \( \frac{\| V \|_Y}{\| V \|_X} \geq 1 - \| Y - X \|_X \)

To get started, direct expansion yields the following second-order expansion of \( f(X) \):

\[
 f(X + \Delta X) \approx f(X) - X^{-1} \cdot \Delta X + \frac{1}{2} \Delta X \cdot X^{-1} \Delta XX^{-1}
\]

and indeed it is easy to derive:

- \( g(X) = -X^{-1} \) and
- \( H(X) \Delta X = X^{-1} \Delta XX^{-1} \)

It therefore follows that

\[
\| \Delta X \|_X = \sqrt{\text{Tr}(\Delta XX^{-1} \Delta XX^{-1})} = \sqrt{\text{Tr}(X^{-1} \Delta XX^{-1} \Delta XX^{-1} \Delta XX^{-1})} = \sqrt{\text{Tr}([X^{-1} \Delta XX^{-1}]^2)}.
\]

Now define two auxiliary matrices:

\[
 F := X^{-\frac{1}{2}} Y X^{-\frac{1}{2}} \quad \text{and} \quad S := X^{-\frac{1}{2}} V X^{-\frac{1}{2}}.
\]
Note that
\[ \|S\| = \sqrt{\text{Tr}(X^{-\frac{1}{2}}VX^{-\frac{1}{2}}X^{-\frac{1}{2}}VX^{-\frac{1}{2}})} = \|V\|_X. \] (1)
Furthermore let us write \( F = QDQ^T \) where the diagonal matrix \( D \) is comprised of the eigenvalues of \( F \) and let \( \lambda \) denote the vector of eigenvalues, with minima and maxima \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \). To prove item (1.) above, we observe:

\[
1 > \|Y - X\|_X^2 = \text{Tr}(X^{-\frac{1}{2}}(Y - X)X^{-\frac{1}{2}}X^{-\frac{1}{2}}(Y - X)X^{-\frac{1}{2}}) \]
\[
= \text{Tr}(F - I)(F - I)) \]
\[
= \text{Tr}(Q(D - I)Q^TQ(D - I)Q^T) \]
\[
= \text{Tr}((D - I)(D - I)) \]
\[
= \sum_{j=1}^{k} (\lambda_j - 1)^2 \]
\[
= \|\lambda - e\|_2^2 \]

where \( e = (1, \ldots, 1) \). Since the last quantity above is less than 1, it follows that \( \lambda > 0 \) and hence \( F > 0 \) and therefore \( Y > 0 \), establishing (1.). In order to establish (2.) and (3.) we will need the following

\[\|F^{-\frac{1}{2}}SF^{-\frac{1}{2}}\| \leq \frac{1}{\lambda_{\text{min}}} \|S\| \quad \text{and} \quad \|F^{-\frac{1}{2}}SF^{-\frac{1}{2}}\| \geq \frac{1}{\lambda_{\text{max}}} \|S\|. \] (3)
To prove (3), we proceed as follows:

\[ \|F^{-\frac{1}{2}}SF^{-\frac{1}{2}}\| = \sqrt{\text{Tr}(QD^{-\frac{1}{2}}Q^TSD^{-\frac{1}{2}}QD^{-\frac{1}{2}}Q^TSD^{-\frac{1}{2}}Q^T)} \]

\[ = \sqrt{\text{Tr}(D^{-1}Q^TSD^{-1}Q^T)} \]

\[ \leq \frac{1}{\sqrt{\lambda_{\min}}} \sqrt{\text{Tr}(Q^TSD^{-1}Q^T)} \]

\[ = \frac{1}{\sqrt{\lambda_{\min}}} \sqrt{\text{Tr}(D^{-1}SS)} \]

\[ \leq \frac{1}{\lambda_{\min}} \sqrt{\text{Tr}(SS)} = \frac{1}{\lambda_{\min}} \|S\| \]

The other inequality of (3) follows by substituting \(\lambda_{\max}\) for \(\lambda_{\min}\) and switching \(\geq\) for \(\leq\) in the above chain of equalities and inequalities. We now have:

\[ \|V\|_Y^2 = \text{Tr}(VY^{-1}V^Y^{-1}) \]

\[ = \text{Tr}(X^{-\frac{1}{2}}VX^{-\frac{1}{2}}X^{\frac{1}{2}}Y^{-1}X^{\frac{1}{2}}X^{-\frac{1}{2}}VX^{-\frac{1}{2}}X^{\frac{1}{2}}Y^{-1}X^{\frac{1}{2}}) \]

\[ = \text{Tr}(SF^{-1}SF) \]

\[ = \text{Tr}(F^{-\frac{1}{2}}SF^{-\frac{1}{2}}F^{-\frac{1}{2}}SF^{-\frac{1}{2}}) \]

\[ = \|F^{-\frac{1}{2}}SF^{-\frac{1}{2}}\|^2 \leq \frac{1}{\lambda_{\min}} \|S\|^2 = \frac{1}{\lambda_{\min}} \|V\|_X^2 \]

where the last inequality follows from (3) and the last equality from (1). Therefore

\[ \frac{\|V\|_Y}{\|V\|_X} \leq \frac{1}{\lambda_{\min}} \leq \frac{1}{1 - |1 - \lambda_{\min}|} \leq \frac{1}{1 - \|e - \lambda\|_2} = \frac{1}{1 - \|Y - X\|_X} \]

where the last equality is from (2). This proves (2.). To prove (3.), use the same equalities as above and the second inequality of (3) to obtain:

\[ \|V\|_Y^2 = \|F^{-\frac{1}{2}}SF^{-\frac{1}{2}}\|^2 \geq \frac{1}{\lambda_{\max}} \|V\|_X^2 \]
and therefore $\|V\|_V \geq \frac{1}{\lambda_{\max}}$. If $\lambda_{\max} \leq 1$ it follows directly that $\frac{1}{\lambda_{\max}} \geq 1 \geq 1 - \|Y - X\|_X$, while if $\lambda_{\max} > 1$ we have:

$$\|Y - X\|_X = \|\lambda - e\|_2 \geq \lambda_{\max} - 1,$$

from which it follows that

$$\lambda_{\max}\|Y - X\|_X \geq \|Y - X\|_X \geq \lambda_{\max} - 1$$

and so

$$\lambda_{\max} (1 - \|Y - X\|_X) \leq 1.$$  

From this it then follows that $\frac{\lambda_{\max}}{\lambda_{\max}} \geq 1 - \|Y - X\|_X$, thus completing the proof of (3.).

Our next result is rather technical, as it shows further properties of changes in Hessian matrices under self-concordance:

**Lemma 4.1** Suppose that $f(\cdot) \in \mathcal{SC}$ and $x \in D_f$. If $\|y - x\|_x < 1$, then

- $\|H^{-\frac{1}{2}}_x H_y H^{-\frac{1}{2}}_x\| \leq \left(\frac{1}{1 - \|y - x\|_x}\right)^2$
- $\|H^{-\frac{1}{2}}_y H^{-\frac{1}{2}}_y\| \leq \left(\frac{1}{1 - \|y - x\|_x}\right)^2$
- $\|I - H^{-\frac{1}{2}}_x H_y H^{-\frac{1}{2}}_x\| \leq \left(\frac{1}{1 - \|y - x\|_x}\right)^2 - 1$
- $\|I - H^{-\frac{1}{2}}_x H_y H^{-\frac{1}{2}}_x\| \leq \left(\frac{1}{1 - \|y - x\|_x}\right)^2 - 1$

**Proof:** Let $Q := H^{-\frac{1}{2}}_x H_y H^{-\frac{1}{2}}_x$, and observe that $Q \succ 0$ with eigenvalues $\lambda_n \geq \ldots \geq \lambda_1 > 0$. From Fact 3.6 we have

$$\sqrt{\|Q\|} = \sqrt{\lambda_n} = \max_w \sqrt{w^T Q w} = \max_v \sqrt{v^T H_y v} = \max_v \|v\|_y \leq \frac{1}{1 - \|y - x\|_x}$$

(where the third equality uses the substitution $v = H^{-\frac{1}{2}}_x w$) and squaring yields the first assertion. Similarly, we have

$$\frac{1}{\sqrt{\|Q^{-1}\|}} = \sqrt{\lambda_1} = \min_w \sqrt{w^T Q w} = \min_v \sqrt{v^T H_y v} = \min_v \|v\|_y \geq 1 - \|y - x\|_x$$

11
(where the third equality again uses the substitution $v = H_{\frac{1}{2}}w$) and squaring and rearranging yields the second assertion. Next observe

$$\|I - Q\| = \max_i{|\lambda_i - 1|} \leq \max\{\lambda_n - 1, 1/\lambda_1 - 1\} \leq \left(\frac{1}{1 - \|y - x\|_x}\right)^2 - 1$$

where the first inequality is from Fact 3.8 and the second inequality follows from the two equation streams above, thus showing the third assertion of the lemma. Finally, we have

$$\|I - Q^{-1}\| = \max_i{|1/\lambda_i - 1|} \leq \max\{1/\lambda_1 - 1, \lambda_n - 1\} \leq \left(\frac{1}{1 - \|y - x\|_x}\right)^2 - 1$$

where the first inequality is from Fact 3.8 and the second inequality follows from the two equation streams above, thus showing the fourth assertion of the lemma. □

Recall Newton’s method to minimize $f(\cdot)$. At $x \in D_f$ we compute the Newton step:

$$n(x) := -H(x)^{-1}g(x)$$

and compute the Newton iterate:

$$x_+ := x + n(x) = x - H(x)^{-1}g(x).$$

When $f(\cdot) \in SC$, Newton’s method has some very wonderful properties as we now show.

**Theorem 4.1** Suppose that $f(\cdot) \in SC$ and $x \in D_f$. If $\|n(x)\|_x < 1$, then

$$\|n(x_+)\|_{x_+} \leq \left(\frac{\|n(x)\|_x}{1 - \|n(x)\|_x}\right)^2.$$

**Proof:** We will prove this by proving the following two results which together establish the result:

(I) $\|n(x_+)\|_{x_+} \leq \frac{\|H_{\frac{1}{2}}g(x_+)\|}{1 - \|n(x)\|_x}$, and
\[ (II) \|H_x^{-\frac{1}{2}}g(x_+)\| \leq \frac{\|n(x)\|_2^2}{1 - \|n(x)\|_x} \]

First we prove (I):
\[
\|n(x_+)\|_2^2 = g(x_+)^T H^{-1}(x_+)H(x_+)H^{-1}(x_+)g(x_+)
\]
\[
= g(x_+)^T H_x^{-\frac{1}{2}} H_f^{-\frac{1}{2}} H^{-1}(x_+)H_f^{-\frac{1}{2}} H_x^{-\frac{1}{2}} g(x_+)
\]
\[
\leq \|H_x^{-\frac{1}{2}} H^{-1}(x_+)H_f^{-\frac{1}{2}} \| \|H_x^{-\frac{1}{2}} g(x_+)\|^2
\]
\[
\leq \left( \frac{1}{1 - \|n(x)\|_x} \right)^2 \|H_x^{-\frac{1}{2}} g(x_+)\|^2 \quad \text{(from Lemma 4.1)}
\]

which proves (I). To prove (II), observe first that
\[
g(x_+) = g(x_+) - g(x) + g(x)
\]
\[
= g(x_+) - g(x) - H_x n(x)
\]
\[
= \int_0^1 H(x + t(x_+ - x))(x_+ - x) dt - H_x n(x) \quad \text{(from Fact 3.1)}
\]
\[
= \int_0^1 H(x + tn(x)) - H_x n(x) dt
\]
\[
= \int_0^1 [H(x + tn(x)) - H_x] H_x^{-\frac{1}{2}} H_f^{-\frac{1}{2}} n(x) dt.
\]

Therefore
\[
H_x^{-\frac{1}{2}} g(x_+) = \int_0^1 \left[ H_x^{-\frac{1}{2}} H(x + tn(x)) H_x^{-\frac{1}{2}} - I \right] H_f^{-\frac{1}{2}} n(x) dt
\]

which then implies
\[
\|H_x^{-\frac{1}{2}} g(x_+)\| \leq \int_0^1 \|H_x^{-\frac{1}{2}} H(x + tn(x)) H_x^{-\frac{1}{2}} - I\| \|H_f^{-\frac{1}{2}} n(x)\| dt
\]
\[
\leq \|H_x^{-\frac{1}{2}} n(x)\| \int_0^1 \left( \frac{1}{1 - \|n(x)\|_x} \right)^2 - 1 dt \quad \text{(from Lemma 4.1)}
\]
\[
= \|n(x)\|_x \frac{\|n(x)\|_x^2}{(1 - \|n(x)\|_x)^2} \quad \text{(from Fact 3.4)}
\]

which proves (II). \[\blacksquare\]

**Proposition 4.4** Suppose that \(f(\cdot) \in SC\) and \(x \in D_f\). If \(\|y - x\|_x < 1\), then
\[
\left| f(y) - \left[ f(x) + g(x)^T (y - x) + \frac{1}{2} (y - x)^T H_x (y - x) \right] \right| \leq \frac{\|y - x\|_x^3}{3(1 - \|y - x\|_x)}.
\]
**Proof:** Let $L$ denote the left-hand side of the inequality to be proved. From Fact 3.3 we have

$$L = \left| \int_0^1 \int_0^t (y - x)^T [H(x + s(y - x)) - H(x)](y - x) \, ds \, dt \right|$$

$$= \left| \int_0^1 \int_0^t (y - x)^T H \left( \frac{1}{2} H(x + s(y - x)) - \frac{1}{2} I \right) (y - x) \, ds \, dt \right|$$

$$\leq \|y - x\|_x^2 \left| \int_0^1 \int_0^t \left( \frac{1}{1 - s\|y - x\|_x} \right)^2 - 1 \, ds \, dt \right|$$

(from Lemma 4.1)

$$= \|y - x\|_x^2 \int_0^1 \int_0^t \left( \frac{1}{1 - s\|y - x\|_x} \right)^2 - 1 \, ds \, dt$$

(from Fact 3.4)

$$\leq \frac{\|y - x\|_x^3}{1 - \|y - x\|_x} \int_0^1 t^2 \, dt = \frac{\|y - x\|_x^3}{3(1 - \|y - x\|_x)}.$$

\[
\text{Theorem 4.2} \quad \text{Suppose that } f(\cdot) \in SC \text{ and } x \in D_f. \text{ If } \|n(x)\|_x < \frac{1}{4}, \text{ then } f(\cdot) \text{ has a minimizer } z, \text{ and}
\]

$$\|z - x\|_x \leq \frac{3\|n(x)\|_x^2}{(1 - \|n(x)\|_x)^3}.$$

**Proof:** First suppose that $\|n(x)\|_x \leq 1/9$, and define $Q_x(y) := f(x) + g(x)^T(y - x) + \frac{1}{2}(y - x)^T H_x(y - x)$. Let $y$ satisfy $\|y - x\|_x \leq 1/3$. Then from Proposition 4.4 we have

$$|f(y) - Q_x(y)| \leq \frac{\|y - x\|_x^3}{3(1 - 1/3)} \leq \frac{\|y - x\|_x^2}{9(2/3)} = \frac{\|y - x\|_x^2}{6}.$$
and therefore
\[ f(y) \geq f(x) + g(x)^T H_x^{-1} H_x^{\frac{1}{2}} H_x^{\frac{1}{2}} (y - x) + \frac{1}{2} \|y - x\|_x^2 - \frac{1}{6} \|y - x\|_x^2 \]
\[ \geq f(x) - \|n(x)\|_x \|y - x\|_x + \frac{1}{3} \|y - x\|_x^2 \]
\[ = f(x) + \frac{1}{3} \|y - x\|_x (-3\|n(x)\|_x + \|y - x\|_x) . \]

Now if \( \tilde{y} \in \partial S := \partial \{ y : \|y - x\|_x \leq 3\|n(x)\|_x \} \), it follows that \( f(\tilde{y}) \geq f(x) \).
So, by Fact 3.5, \( f(\cdot) \) has a global minimizer \( z \in S \), and so \( \|z - x\|_x \leq 3\|n(x)\|_x \).

Now suppose that \( \|n(x)\|_x \leq 1/4 \). From Theorem 4.1 we have
\[ \|n(x_+)\|_{x_+} \leq \left( \frac{1/4}{1 - 1/4} \right)^2 = 1/9 , \]
so \( f(\cdot) \) has a global minimizer \( z \) and \( \|z - x_+\|_{x_+} \leq 3\|n(x_+)\|_{x_+} \). Therefore
\[ \|z - x_+\|_x \leq \frac{\|z - x_+\|_{x_+}}{1 - \|x - x_+\|_x} \quad \text{(from Definition 4.1)} \]
\[ = \frac{\|z - x_+\|_{x_+}}{1 - \|n(x)\|_x} \]
\[ \leq \frac{3\|n(x_+)\|_{x_+}}{1 - \|n(x)\|_x} \]
\[ \leq \frac{3\|n(x)\|_x^2}{1 - \|n(x)\|_x^3} . \quad \text{(from Theorem 4.1)} \]

5 Self-Concordant Barriers

We begin with another definition.
Definition 5.1 \( f(\cdot) \) is a \( \vartheta \)-(strongly nondegenerate self-concordant)-barrier if \( f(\cdot) \in \mathcal{SC} \) and
\[
\vartheta = \vartheta_f := \max_{x \in D_f} \|n(x)\|_2^2 < \infty .
\]

Note that \( \|n(x)\|_2^2 = (-g(x)^TH(x)^{-1}H(x)H(x)^{-1}(-g(x)) = g(x)^TH(x)^{-1}g(x) \),
so we can equivalently define
\[
\vartheta_f := \max_{x \in D_f} g(x)^TH(x)^{-1}g(x) .
\]
The quantity \( \vartheta_f \) is called the complexity value of the barrier \( f(\cdot) \).

Let \( \mathcal{SCB} \) denote the class of all such functions. The following property is very important.

Theorem 5.1 Suppose that \( f(\cdot) \in \mathcal{SCB} \) and \( x, y \in D_f \). Then
\[
g(x)^T(y - x) < \vartheta_f .
\]

Proof: Define \( \phi(t) := f(x+t(y-x)) \), whereby \( \phi'(t) = g(x+t(y-x))^T(y-x) \) and \( \phi''(t) = (y-x)^TH(x+t(y-x))(y-x) \). We want to prove that \( \phi'(0) < \vartheta_f \).
If \( \phi'(0) \leq 0 \) there is nothing further to prove, so we can assume that \( \phi'(0) > 0 \) whereby from convexity it also follows that \( \phi'(t) > 0 \) for all \( t > 0 \) in the domain of \( \phi(\cdot) \). Let \( t > 0 \) be in the domain of \( \phi(\cdot) \) and let \( v = x + t(y - x) \).
Then
\[
\phi'(t) = g(v)^T(y - x)
\]
\[
= g(v)^THv^{-1}H_{\frac{1}{2}}H_{\frac{1}{2}}(y - x)
\]
\[
= -n(v)H_{\frac{1}{2}}H_{\frac{1}{2}}(y - x)
\]
\[
\leq \|H_{\frac{1}{2}}n(v)\| \|H_{\frac{1}{2}}(y - x)\|
\]
\[
= \|n(v)\|_v \|y - x\|_v \leq \sqrt{\vartheta_f} \|y - x\|_v .
\]
Also \( \phi''(t) = (y-x)H_v(y-x) = \|y-x\|_v^2 \), whereby
\[
\frac{\phi''(t)}{\phi'(t)^2} \geq \frac{\|y-x\|_v^2}{\vartheta_f \|y-x\|_v^2} = \frac{1}{\vartheta_f}
\]
for all \( t > 0 \) in the domain of \( \phi(\cdot) \). In fact, for all \( s > 0 \) that are in the domain of \( \phi(\cdot) \) it follows that

\[
\frac{s}{\partial_f} \leq \int_0^s \frac{\phi''(t)}{\phi'(t)^2} \, dt = \frac{1}{\phi'(0)} - \frac{1}{\phi'(s)},
\]

and since \( s = 1 \) is in the domain of \( \phi(\cdot) \) we have

\[
\frac{1}{\phi'(0)} \geq \frac{1}{\phi'(1)} + \frac{1}{\partial_f},
\]

which proves the result. ■

The next two results show how the complexity value \( \vartheta \) behaves under addition/intersection and affine transformation.

**Theorem 5.2 (self-concordant barriers under addition/intersection)**

Suppose that \( f_i(\cdot) \in SCB \) with domain \( D_i := D_{f_i} \) and complexity values \( \vartheta_i := \vartheta_{f_i} \) for \( i = 1,2 \), and suppose that \( D := D_1 \cap D_2 \neq \emptyset \). Define \( f(\cdot) = f_1(\cdot) + f_2(\cdot) \). Then \( f(\cdot) \in SCB \) with domain \( D \), and \( \vartheta_f \leq \vartheta_1 + \vartheta_2 \).

**Proof:** Fix \( x \in D \) and let \( g, g_1, g_2 \) and \( H, H_1, H_2 \) denote the gradients and Hessians of \( f(\cdot), f_1(\cdot), f_2(\cdot) \) at \( x \), whereby \( g = g_1 + g_2 \) and \( H = H_1 + H_2 \). Define \( A_i = H^{-\frac{1}{2}}H_iH^{-\frac{1}{2}} \) for \( i = 1,2 \). Then \( A_i \succ 0 \), \( A_1 + A_2 = I \), so \( A_1, A_2 \) commute and \( A_1^\frac{1}{2}, A_2^\frac{1}{2} \) commute, from Fact 3.7. Also define \( u_i = A_i^{-\frac{1}{2}}H^{-\frac{1}{2}}g_i \)
for $i = 1, 2$. We have
\[
g^T H^{-1} g = g_1^T H^{-1} g_1 + g_2^T H^{-1} g_2 + 2g_1^T H^{-1} g_2 \\
= u_1^T A_1 u_1 + u_2^T A_2 u_2 + 2u_1^T A_1^\frac{1}{2} A_2^\frac{1}{2} u_2 \\
= u_1^T [I - A_2] u_1 + u_2^T [I - A_1] u_2 + 2u_1^T A_2^\frac{1}{2} A_1^\frac{1}{2} u_2 \\
= u_1^T u_1 + u_2^T u_2 - \left[ u_1^T A_2 u_1 + u_2^T A_1 u_2 - 2u_1^T A_2^\frac{1}{2} A_1^\frac{1}{2} u_2 \right] \\
= g_1^T H^{-\frac{1}{2}} A_1^{-1} H^{-\frac{1}{2}} g_1 + g_2^T H^{-\frac{1}{2}} A_2^{-1} H^{-\frac{1}{2}} g_2 - \| A_2^\frac{1}{2} u_1 - A_1^\frac{1}{2} u_2 \|^2 \\
\leq g_1^T H^{-\frac{1}{2}} H_1^\frac{1}{2} H_1^{-1} H_2^\frac{1}{2} H_2^\frac{1}{2} H^{-\frac{1}{2}} g_1 + g_2^T H^{-\frac{1}{2}} H_2^\frac{1}{2} H_2^{-1} H_1^\frac{1}{2} H_1^\frac{1}{2} H^{-\frac{1}{2}} g_2 \\
\leq \vartheta_1 + \vartheta_2
\] thereby showing that $\vartheta_f \leq \vartheta_1 + \vartheta_2$.

**Theorem 5.3 (self-concordant barriers under affine transformation)**

Let $A \in \mathbb{R}^{m \times n}$ satisfy $\text{rank} A = n \leq m$. Suppose that $f(\cdot) \in \text{SCB}$ with complexity value $\vartheta_f$, with domain $D_f \subset \mathbb{R}^m$ and define $\hat{f}(\cdot)$ by $\hat{f}(x) = f(Ax - b)$. Then $\hat{f}(\cdot) \in \text{SCB}$ and $\vartheta_f \leq \vartheta_f$.

**Proof:** Fix $x \in \hat{D}$ and define $s = Ax - b$. Letting $g$ and $H$ denote the gradient and Hessian of $f(s)$ at $s$ and $\hat{g}$ and $\hat{H}$ the gradient and Hessian of $\hat{f}(x)$ at $x$, we have $\hat{g} = A^T g$ and $\hat{H} = A^T H A$. Then
\[
\hat{g}^T \hat{H}^{-1} \hat{g} = g^T A (A^T H A)^{-1} A^T g = g^T H^{-\frac{1}{2}} H_1^\frac{1}{2} A (A^T H_1^\frac{1}{2} H_2^\frac{1}{2} H_1^{-1} H_2^\frac{1}{2} A)^{-1} A^T H_2^\frac{1}{2} H^{-\frac{1}{2}} g \\
\leq g^T H^{-\frac{1}{2}} H^{-\frac{1}{2}} g = g^T H^{-1} g \leq \vartheta_f
\] since the matrix $H_1^\frac{1}{2} A (A^T H_1^\frac{1}{2} H_2^\frac{1}{2} A)^{-1} A^T H_2^\frac{1}{2}$ is a projection matrix.

**Remark 2** The complexity values of the three most-used barriers are as follows:
1. \( \vartheta_f = 1 \) for the barrier \( f(x) = -\ln(x) \) defined on \( D_f = \{ x : x > 0 \} \)

2. \( \vartheta_f = k \) for the barrier \( f(X) = -\ln \det(X) \) defined on \( D_f = \{ X \in S^{k \times k} : X > 0 \} \}

3. \( \vartheta_f = 2 \) for the barrier \( f(x) = -\ln(x_1^2 - \sum_{j=2}^{n} x_j^2) \) defined on \( D_f = \{ x : \| (x_2, \ldots, x_n) \| \leq x_1 \} \)

**Proof:** Item (1.) follows from item (2.) so we first prove (2.). Recall that for \( f(X) = -\ln \det(X) \) we have \( g(X) = -X^{-1} \) and \( H(X) \Delta X = X^{-1} \Delta X X^{-1} \). Therefore the Newton step at \( X \), denoted by \( n(X) \), is the solution of the following equation:

\[
X^{-1} [n(X)] X^{-1} = X^{-1}
\]

and it follows that \( n(X) = X \). Therefore

\[
\| n(X) \|_X^2 = \text{Tr}([X^{-\frac{1}{2}} [n(X)] X^{-\frac{1}{2}}]^2) = \text{Tr}(I) = k
\]

and therefore \( \vartheta_f = \max_{X > 0} \| n(X) \|_X^2 = k \), which proves (2.) and hence (1.).

In order to prove (3.) we amend our notation a bit, letting \( Q^n = \{ (t, x) \in \mathbb{R}^1 \times \mathbb{R}^{n-1} : \| x \| \leq t \} \). For \((t, x) \in \text{int} Q^n \) we have \( \| x \| < t \) and mechanically we can derive:

\[
g(t, x) = \begin{pmatrix}
-2t \\
t^2 - x^T x
\end{pmatrix}
\]

\[H(t, x) = \begin{pmatrix}
\frac{-2t^2 + 2x^T x}{(t^2 - x^T x)^2} & \frac{-4tx^T}{(t^2 - x^T x)^2} \\
\frac{-2tx^T}{(t^2 - x^T x)^2} & \frac{2(t^2 - x^T x)}{(t^2 - x^T x)^2}
\end{pmatrix}
\]

and the Hessian inverse is given by

\[
H(t, x)^{-1} = \begin{pmatrix}
\frac{t^2 + x^T x}{2} & \frac{tx^T}{2} \\
\frac{t^2 - x^T x}{2} & \frac{t^2 - x^T x}{2} I + xx^T
\end{pmatrix}
\]

Directly plugging in yields

\[
g(t, x)^T H^{-1}(t, x) g(t, x) = 2
\]

from which it follows that \( \vartheta_f = \max_{(t, x) \in Q^n} g(x)^T H(x)^{-1} g(x) = 2 \).
6 The Barrier Method and its Analysis

Our original problem of interest is

\[ P : \text{ minimize } c^T x \]
\[ \text{ s.t. } x \in S, \]

whose optimal objective value we denote by \( V^* \). Let \( f(\cdot) \) be a self-concordant barrier on \( D_f = \text{int} S \). For every \( \mu > 0 \) we create the barrier problem:

\[ P_\mu : \text{ minimize } \mu c^T x + f(x) \]
\[ \text{ s.t. } x \in D_f. \]

The solution of this problem for each \( \mu \) is denoted \( z(\mu) \):

\[ z(\mu) := \text{arg min}_{x} \{ \mu c^T x + f(x) : x \in D_f \}. \]

Intuitively, as \( \mu \to \infty \), the impact of the barrier function on the solution of \( P_\mu \) should become less and less, so we should have \( c^T z(\mu) \to V^* \) as \( \mu \to \infty \). Presuming this is the case, the barrier scheme will use Newton’s method to solve for approximate solutions \( x^i \) of \( P_{\mu_i} \) for an increasing sequence of values of \( \mu_i \to \infty \).

In order to be more specific about how the barrier scheme might work, let us assume that at each iteration we have some value \( x \in D_f \) that is an approximate solution of \( P_\mu \) for a given value \( \mu > 0 \). (We will define “an approximate solution of \( P_\mu \)” shortly.) We then will increase the barrier parameter \( \mu \) by a multiplicative factor \( \alpha > 1 \):

\[ \hat{\mu} \leftarrow \alpha \mu. \]

Then we will take a Newton step at \( x \) for the problem \( P_{\hat{\mu}} \) to obtain a new point \( \hat{x} \) that we would like to then be an approximate solution of \( P_{\hat{\mu}} \). If so, we can continue the scheme recursively.
Let $x \in D_f$ be given, and let us compute the Newton step for $P_\mu$ at $x$. The objective function of $P_\mu$ is

$$h_\mu(x) := \mu c^T x + f(x),$$

whereby we have:

- $\nabla h_\mu(x) = \mu c + g(x)$
- $\nabla^2 h_\mu(x) = H(x) = H_x$

Therefore the Newton step for $h_\mu(\cdot)$ at $x$ is:

$$n_\mu(x) := -H_x^{-1}(\mu c + g(x)) = n(x) - \mu H_x^{-1}c$$

and the new iterate is:

$$\hat{x} := x + n_\mu(x) = x - H_x^{-1}(\mu c + g(x)).$$

**Remark 3** Notice that $h_\mu(\cdot) \in \mathcal{SC}$, since membership in $\mathcal{SC}$ has only to do with Hessians, and $h_\mu(\cdot)$ and $f(\cdot)$ have the same Hessian. However, membership in $\mathcal{SCB}$ depends also on gradients, and $h_\mu(x)$ and $f(x)$ have different gradients, and $h_\mu(\cdot) \notin \mathcal{SCB}$ (unless $c = 0$).

We now define what we mean for $y$ to be an “approximate solution” of $P_\mu$.

**Definition 6.1** Let $\gamma \in [0,1)$ be given. We say that $y \in D_f$ is a $\gamma$-approximate solution of $P_\mu$ if

$$\|n_\mu(y)\|_y \leq \gamma.$$

Essentially, the above definition states that $y$ is a $\gamma$-approximate solution of $P_\mu$ if the Newton step for $P_\mu$ at $y$ is small (measured using the local norm at $y$). The following theorem gives an explicit optimality gap bound for $y$ if $y$ is a $\gamma = 1/4$-approximate solution of $P_\mu$.  

21
Theorem 6.1 Suppose $\gamma \leq \frac{1}{4}$, and $y \in D_f$ is a $\gamma$-approximate solution of $P_\mu$. Then

$$c^T y \leq V^* + \frac{\vartheta_f}{\mu} \left( \frac{1}{1 - \delta} \right)$$

where $\delta = \gamma + \frac{3\gamma^2}{(1-\gamma)^3}$.

Proof: From Theorem 4.1 we know that $z(\mu)$ exists and furthermore

$$\|y - z(\mu)\|_y = \|y + n_\mu(y) - z(\mu) - n_\mu(y)\|_y$$

$$\leq \|y_+ - z(\mu)\|_y + \|n_\mu(y)\|_y$$

$$\leq \frac{3\gamma^2}{(1-\gamma)^3} + \gamma = \delta .$$

From basic first-order optimality conditions we know that $z(\mu)$ satisfies

$$\mu c + g(z(\mu)) = 0$$

and from Theorem 5.1 we have

$$-\mu c^T (w - z(\mu)) = g(z(\mu))^T (w - z(\mu)) < \vartheta_f \quad \text{for all } w \in D_f.$$ 

Rearranging we have

$$c^T w + \frac{\vartheta_f}{\mu} > c^T z(\mu) \quad \text{for all } w \in D_f,$$
whereby $V^* + \frac{\vartheta_f}{\mu} \geq c^T z(\mu)$. Now for convenience let $z = z(\mu)$ and observe

$$c^T y = c^T z + c^T (y - z)$$

$$\leq V^* + \frac{\vartheta_f}{\mu} + c^T H^{-\frac{1}{2}} H^{-\frac{1}{2}} (y - z)$$

$$\leq V^* + \frac{\vartheta_f}{\mu} + \|H^{-\frac{1}{2}} c\| \|(y - z)\|_z$$

$$\leq V^* + \frac{\vartheta_f}{\mu} + \left(\sqrt{c^T H^{-1} c}\right) \frac{\|(y - z)\|_y}{1 - \|(y - z)\|_y}$$

$$\leq V^* + \frac{\vartheta_f}{\mu} + \left(\sqrt{(g(z)/\mu)^T H^{-1} (g(z)/\mu)}\right) \frac{\delta}{1 - \delta}$$

$$= V^* + \frac{\vartheta_f}{\mu} + \frac{\sqrt{(g(z))^T H^{-1} (g(z))}}{\mu} \frac{\delta}{1 - \delta}$$

$$\leq V^* + \frac{\vartheta_f}{\mu} + \frac{\delta}{1 - \delta}$$

$$\leq V^* + \frac{\vartheta_f}{\mu} \left(1 + \frac{\delta}{1 - \delta}\right)$$

$$= V^* + \frac{\vartheta_f}{\mu(1 - \delta)}.$$

The last inequality above follows from the fact (which we will not prove) that $\vartheta_f \geq 1$ for any $f(\cdot) \in SCB$. ■

Note that with $\gamma = 1/9$ we have $1/(1 - \delta) \leq 6/5$ and $c^T y \leq V^* + 1.2 \vartheta_f / \mu$.

**Theorem 6.2** Let $\beta := \frac{1}{4}$, $\gamma := \frac{1}{3}$, and $\alpha := \frac{\sqrt{\beta + \gamma}}{\sqrt{\beta + \gamma}}$. Suppose $x$ is a $\gamma$-approximate solution of $P_\mu$. Define $\hat{\mu} := \alpha \mu$, and let $\hat{x}$ be the Newton iterate for $P_{\hat{\mu}}$ at $x$, namely

$$\hat{x} := x - H(x)^{-1} (\hat{\mu} c + g(x)).$$

Then

1. $x$ is a $\beta$-approximate solution of $P_\mu$, and
**Barrier Method**

### Step 1 Initialize.
Define $\gamma := 1/9$, $\beta := 1/4$, $\alpha := \sqrt{\frac{\beta}{\beta+\gamma}}$. Initialize with $\mu_0 > 0$, $x^0 \in D_f$ that is a $\gamma$-approximate solution of $P_{\mu_0}$. Set $x \leftarrow x^0$, $\mu \leftarrow \mu_0$, and $i \leftarrow 0$.

### Step 2 Increase $\mu$ and take Newton step.
$\bar{\mu} \leftarrow \alpha \mu$

$\hat{x} \leftarrow x - H(x)^{-1}(\bar{\mu}c + g(x))$

### Step 3 Update counter and repeat.
Set $x \leftarrow \hat{x}$, $\mu \leftarrow \bar{\mu}$, $i \leftarrow i + 1$ and Goto Step 2.

Table 2: The Barrier Method.

2. $\hat{x}$ is a $\gamma$-approximate solution of $P_{\mu}$.

**Proof:** To prove (1.) we have:

$$\|n_{\bar{\mu}}(x)\|_x = \|n_{\alpha \mu}(x)\|_x$$

$$= \|H(x)^{-1}(\alpha \mu c + g(x))\|_x$$

$$= \|\alpha [H(x)^{-1}(\mu c + g(x))] + (1 - \alpha)H(x)^{-1}g(x)\|_x$$

$$\leq \alpha \|H(x)^{-1}(\mu c + g(x))\|_x + (\alpha - 1)\|H(x)^{-1}g(x)\|_x$$

$$\leq \alpha \gamma + (\alpha - 1)\|n(x)\|_x$$

$$\leq \alpha \gamma + (\alpha - 1)\sqrt{\bar{f}} = \beta.$$
To prove (2.) we invoke Theorem 4.1:

\[
\|n_\mu(\hat{x})\|_2 \leq \frac{\|n_\mu(x)\|_2^2}{(1 - \|n_\mu(x)\|_2)^2} \leq \frac{\beta^2}{(1 - \beta)^2} = \frac{1}{9} = \gamma.
\]

Applying Theorem 6.2 recursively we obtain the basic barrier method for self-concordant barriers as presented in Table 2. The complexity of this scheme is presented below.

**Theorem 6.3** Let \( \varepsilon > 0 \) be the desired optimality tolerance, and define

\[
J := \left\lceil 9\sqrt{\delta} \ln \left( \frac{6\delta}{5\mu_0\varepsilon} \right) \right\rceil.
\]

Then by iteration \( J \) of the barrier method the current iterate \( x \) satisfies \( c^T x \leq V^* + \varepsilon \).

**Proof:** With the given values of \( \gamma, \beta, \alpha \) we have \( \delta := \gamma + \frac{2\gamma^2}{(1-\gamma)^2} \) satisfies \( \frac{1}{1-\delta} \leq 6/5 \) and

\[
1 - \frac{1}{\alpha} = \frac{1}{7.2\sqrt{\delta} + 1.8} \geq \frac{1}{9\sqrt{\delta}}.
\]

After \( J \) iterations the current iterate \( x \) is a \( \gamma \)-approximate solution of \( P_\mu \) where \( \mu = \alpha J \mu_0 \). Therefore

\[
\ln \mu_0 - \ln \mu = J \ln \left( \frac{1}{\alpha} \right) \leq J \left( \frac{1}{\alpha} - 1 \right).
\]

Therefore

\[
\ln \mu \geq \ln \mu_0 + J \left( 1 - \frac{1}{\alpha} \right) \geq \ln \mu_0 + J \frac{J}{9\sqrt{\delta}} \geq \ln \mu_0 + \ln \left( \frac{6\delta}{5\mu_0\varepsilon} \right) \geq \ln \left( \frac{\delta}{(1-\delta)\varepsilon} \right).
\]

Therefore \( \mu \geq \frac{\delta}{(1-\delta)\varepsilon} \), and from Theorem 6.1 we have

\[
c^T x \leq V^* + \frac{\delta}{\mu(1-\delta)} \leq V^* + \varepsilon.
\]
7 Remarks and other Matters

- Nesterov-Nemirovskii definition of self-concordance
- getting started
- other formats for convex optimization
- \( \vartheta \)-logarithmic homogeneous barriers for cones
- universal barrier
- primal-dual methods
- computational practice
- other self-concordant functions and self-concordant calculus