# Short-Step Interior Point Algorithm

## A Primal Barrier Newton-Iteration Technique

Avinash N. Madavan

Univeristy of Illinois, Urbana-Champaign

2017

# Table of Contents

# Table of Contents

# Overview

The proof outlined in the next slides will depend on several important steps:

- Self-concordant local analogues for strong convexity and Lipschitz hessian properties.
- Bound self-concordant function deviation from its quadratic approximation.
- Bound the magnitude of a Newton step as a function of the magnitude of the previous step.
- Bound for the distance to the minimizer as a function of the magnitude of the Newton step.
- Define a set of barrier parameters that will convergence to the optimal.

# Table of Contents

# Definitions

$$H(x) := \text{hessian evaluated at } x$$

**Definition 1 (Intrinsic Inner Product)**

The intrinsic inner product over a function $f$, denoted $\langle \cdot, \cdot \rangle_x$, is the weighted inner product, such that

$$\langle u, v \rangle_x = \langle u, H(x)v \rangle$$

**Definition 2 (Induced Norm)**

The induced norm, denoted $\|\cdot\|_x$, represents the norm induced by the intrinsic inner product,

$$\|v\|_x = \|v\|_{H(x)} = \langle v, H(x)v \rangle$$

# Table of Contents

# Self-Concordant Functions

$D_f$: open convex set (i.e. $\{x > 0\}$)
$B_x(y, r)$: open ball of radius $r$, centered at $y$, measured with $\|\cdot\|_x$

## Definition 3 (Self-Concordance)

A function, $f$, is said to be self-concordant if for all $x \in D_f$ we have that $B_x(x, 1) \subseteq D_f$ and if $\forall y \in B_x(x, 1)$ we have,

$$1 - \|y - x\|_x \leq \frac{\|v\|_y}{\|v\|_x} \leq \frac{1}{1 - \|y - x\|_x}$$

# Hessian Bound I

$$H_x(y) = H(x)^{-1}H(y)$$

**Theorem 4**

*Assume that the function $f$ has the property that $B_x(x,1) \subseteq D_f$ for all $x \in D_f$. Then $f$ is self-concordant iff for all $x \in D_f$ and $y \in B_x(x,1)$,*

$$\|H_x(y)\|_x, \|H_x(y)^{-1}\|_x \leq \frac{1}{(1-\|y-x\|)^2} \quad (1)$$

*and likewise iff,*

$$\|I - H_x(y)\|_x, \|I - H_x(y)^{-1}\|_x \leq \frac{1}{(1-\|y-x\|_x)^2} - 1 \quad (2)$$

# Hessian Bound II

*proof.* Let $\lambda_1 \leq \ldots \leq \lambda_n$ denote the eigenvalues of $H_x(y)$. Since,

$$\max_v \frac{\|v\|_y^2}{\|v\|_x^2} = \max_v \frac{\langle v, H_x(y)v\rangle_x}{\|v\|_x^2} = \lambda_n = \|H_x(y)\|_x$$

and similarly,

$$\min_v \frac{\|v\|_y^2}{\|v\|_x^2} = \min_v \frac{\langle v, H_x(y)v\rangle_x}{\|v\|_x^2} = \lambda_1 = \frac{1}{\|H_x(y)^{-1}\|_x}$$

from the definition of self-concordance,

$$\|H_x(y)\|_x = \max_v \frac{\|v\|_y^2}{\|v\|_x^2} \leq \frac{1}{(1 - \|y - x\|)^2}$$

# Hessian Bound III

and similarly,

$$\frac{1}{\|H_x(y)^{-1}\|_x} = \min_v \frac{\|v\|_y^2}{\|v\|_x^2} \geq (1 - \|y - x\|)^2$$
$$\implies \|H_x(y)^{-1}\|_x \leq \frac{1}{(1 - \|y - x\|)^2}$$

Recall the eigenvalues of $I - H_x(y)$ are $1 - \lambda_i$, so

$$\begin{aligned} \|I - H_x(y)\|_x &= \max\{\lambda_n - 1, 1 - \lambda_1\} \\ &\leq \max\left\{\lambda_n - 1, \frac{1}{\lambda_1} - 1\right\} \\ &= \max\{\|H_x(y)\|_x - 1, \|H_x(y)^{-1}\|_x - 1\} \end{aligned}$$

□

Recall $H_x(x) = I$, so bounding $\|I - H_x(y)\|_x$ has essentially given us local analogues for Lipschitz continuity and strong convexity.

# Table of Contents

# Log-Barrier

$g(x)$: gradient of function self-concordant function $f$.
$g_x(y) = H_x(^{-1} g(y)$

Recall, $g(x)$ for the log-barrier is the vector with the $j$-th entry $-1/x_j$ and $H(x)$ is the diagonal matrix with the $j$-th diagonal entry $1/x_j^2$. So,

$$\|g_x(x)\|_x^2 = \langle g(x), H(x)^{-1} g(x) \rangle = n$$

# Table of Contents

# Problem Definition

The problem which we aim to solve is the following linear program,

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

In the augmented form,

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \eta c^T x - \sum_{i=1}^{n} \ln(x_i) \\ \text{subject to} \quad & Ax = b \end{aligned}$$

# Quadratic Approximation I

$$q_x(y) := f(x) + \langle g(x), y - x\rangle + \frac{1}{2}\langle y - x, H(x)(y - x)\rangle$$

**Theorem 5**

*If $f$ is a self-concordant function, $x \in D_f$ and $y \in B_x(x, 1)$ then*

$$\|f(y) - q_x(y)\| \leq \frac{\|y - x\|_x^3}{3(1 - \|y - x\|_x)}$$

*proof.* From the fundamental theorem of calculus,

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2}\phi''(0) + \int_0^1 \int_0^t \phi''(s) - \phi''(0) ds dt$$

# Quadratic Approximation II

So,

$$f(y) = f(x) + \langle g(x), y - x\rangle + \frac{1}{2}\langle y - x, H(x)(y - x)\rangle$$
$$+ \int_0^1 \int_0^t \langle y - x, [H(x + s(y - x)) - H(x)](y - x)\rangle dsdt$$

$$\begin{aligned}
\|f(y) - q_x(y)\| &\leq \|y - x\|_x^2 \int_0^1 \int_0^t \|I - H_x(x + s(y - x))\|_x dsdt \\
&\leq \|y - x\|_x^2 \int_0^1 \int_0^t \frac{1}{(1 - s\|y - x\|_x)^2} - 1 dsdt \\
&= \|y - x\|_x^3 \int_0^1 \frac{t^2}{1 - t\|y - x\|_x} dt \\
&\leq \frac{\|y - x\|_x^3}{1 - \|y - x\|_x} \int_0^1 t^2 dt \\
&= \frac{\|y - x\|_x^3}{3(1 - \|y - x\|_x)}
\end{aligned}$$

# Newton Step Bound I

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

## Theorem 6

*Assume $f$ is self-concordant. If $\|n(x)\|_x < 1$, then*

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

# Newton Step Bound II

*proof.*

$$\begin{aligned}\|n(x_+)\|_{x_+}^2 &= \|H_x(x_+)^{-1}g_x(x_+)\|_{x_+}^2 \\ &= \langle H_x(x_+)^{-1}g_x(x_+), H_x(x_+)^{-1}g_x(x_+)\rangle_{x_+} \\ &= \langle H_x(x_+)^{-1}g_x(x_+), g_x(x_+)\rangle_x \\ &\leq \|H_x(x_+)^{-1}\|_x \|g_x(x_+)\|_x^2\end{aligned}$$

From 1,

$$\|H_x(x_+)^{-1}\|_x \leq \frac{1}{(1 - \|n(x)\|_x)^2}$$

Therefore,

$$\|n(x_+)\|_{x_+} \leq \frac{\|g_x(x_+)\|_x}{1 - \|n(x)\|_x}$$

# Newton Step Bound III

Recall $g_x(x) = -n(x)$,

$$\begin{aligned}
\|g_x(x_+)\|_x &= \|g_x(x_+) - g(x) - n(x)\|_x \\
&= \left\| \int_0^1 [H_x(x + tn(x)) - I]n(x)dt \right\|_x \\
&\leq \|n(x)\|_x \int_0^1 \|I - H_x(x + tn(x))\|_x dt \\
&\leq \|n(x)\|_x \int_0^1 \frac{1}{(1 - t\|n(x)\|_x)^2} - 1dt \\
&= \frac{\|n(x)\|_x^2}{1 - \|n(x)\|_x}
\end{aligned}$$

□

# Convergence to Minimizer I

## Theorem 7

*Assume $f$ is self-concordant. If $\|n(x)\|_x \leq \frac{1}{4}$ for some $x \in D_f$, then $f$ has a minimizer $z$ and*

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

*proof.* We first prove the weaker result, if $\|n(x)\|_x \leq \frac{1}{9}$, then $f$ has a minimizer $z$ and $\|x - z\|_x \leq 3\|n(x)\|_x$. From Theorem 5, for all $y \in \bar{B}_x(x, \frac{1}{3})$,

$$\|f(y) - q_x(y)\| \leq \frac{1}{6}\|y - x\|_x^2$$

$$\begin{aligned} f(y) &\geq f(x) - \|n(x)\|_x \|y - x\|_x + \frac{1}{2}\|y - x\|_x^2 - \frac{1}{6}\|y - x\|^2 \\ &= f(x) - \|n(x)\|_x \|y - x\|_x + \frac{1}{3}\|y - x\|_x^2 \end{aligned}$$

# Convergence to Minimizer II

So if $\|n(x)\|_x \leq \frac{1}{9}$ and $\|y - x\|_x = 3\|n(x)\|_x \leq \frac{1}{3}$, then $f(y) \geq f(x)$. Assuming a minimizer, $z$, in $S$ exists, then $\|x - z\|_x \leq 3\|n(x)\|_x$. Now, assume $\|n(x)\|_x \leq \frac{1}{4}$. Then from 6

$$\|n(x+)\|_{x+} \leq \left(\frac{\|n(x)\|_x}{1 - \|n(x)\|_x}\right)^2 \leq \frac{1}{9}$$

Using the previous conclusion $\|z - x_+\|_{x_+} \leq 3\|n(x_+)\|_{x_+}$. Thus

$$\begin{aligned}
\|z - x_+\|_x &\leq \frac{\|z - x_+\|_{x_+}}{1 - \|n(x)\|_x} \\
&\leq \frac{3\|n(x_+)\|_{x_+}}{1 - \|n(x)\|_x} \\
&\leq \frac{3\|n(x)\|_x^2}{(1 - \|n(x)\|_x)^3}
\end{aligned}$$

# The Algorithm I

$$n_\eta(x) := -H(x)^{-1}(\eta c + g(x)) = -(\eta c_x + g_x(x))$$
$$x_2 := x_1 + n_{\eta_2}(x_1)$$

**Lemma 8**

*For the step defined above,* $\|n_{\eta_2}\|_x \leq \frac{\eta_2}{\eta_1}\|n_{\eta_1}(x)\| + |\frac{\eta_2}{\eta_1} - 1|\sqrt{n}$

# The Algorithm II

*proof.*

$$\frac{1}{\eta_2}(n_{\eta_2}(x) + g_x(x)) = \frac{1}{\eta_1}(n_{\eta_1}(x) + g_x(x))$$

$$n_{\eta_2}(x) = \frac{\eta_2}{\eta_1} n_{\eta_1}(x) + \left(\frac{\eta_2}{\eta_1} - 1\right) g_x(x)$$

$$\implies \quad \|n_{\eta_2}(x)\|_x \leq \frac{\eta_2}{\eta_1}\|n_{\eta_1}(x)\|_x + \left|\frac{\eta_2}{\eta_1} - 1\right| \|g_x(x)\|_x$$

$$\leq \frac{\eta_2}{\eta_1}\|n_{\eta_1}(x)\|_x + \left|\frac{\eta_2}{\eta_1} - 1\right| \sqrt{n}$$

□

# The Algorithm III

## Lemma 9

*For* $\beta := \frac{\eta_2}{\eta_1} = 1 + \frac{1}{8\sqrt{n}}$,

$$\|n_{\eta_1}(x_1)\|_{x_1} \leq \frac{1}{9} \implies \|n_{\eta_2}(x_2)\|_{x_2} \leq \frac{1}{9}$$

*proof.*

$$\begin{aligned}
\|n_{\eta_2}(x_1)\|_{x_1} &\leq \beta\|n_{\eta_1}(x_1)\|_{x_1} + |\beta - 1|\sqrt{n} \\
&\leq \frac{1}{9}\left(1 + \frac{1}{8\sqrt{n}}\right) + \left|\left(1 + \frac{1}{8\sqrt{n}}\right) - 1\right|\sqrt{n} \\
&= \frac{1}{9} + \frac{1}{8} + \frac{1}{72\sqrt{n}} \\
&\leq \frac{1}{9} + \frac{1}{8} + \frac{1}{72} \\
&= \frac{1}{4}
\end{aligned}$$

# The Algorithm IV

By applying Theorem 6,

$$\begin{aligned}\|n_{\eta_2}(x_2)\|_{x_2} &\leq \left(\frac{\|n_{\eta_2}(x_1)\|_{x_1}}{1 - \|n_{\eta_2}(x_1)\|_{x_1}}\right)^2 \\ &\leq \frac{1}{9}\end{aligned}$$

By applying Theorem 7,

$$\begin{aligned}\|x - z(\eta)\|_x &\leq \frac{3\|n(x)\|_x^2}{(1 - \|n(x)\|_x)^3} \\ &\leq \frac{3\left(\frac{1}{9}\right)^2}{\left(1 - \frac{1}{9}\right)^3} \\ &\leq \frac{1}{6}\end{aligned}$$

# The Algorithm V

By applying the definition of self-concordance,

$$\begin{aligned}\|x - z(\eta)\|_{z(\eta)} &\leq \frac{\|x - z(\eta)\|_x}{1 - \|x - z(\eta)\|_x} \\ &\leq \frac{1}{5}\end{aligned}$$

Thus, all points generated are within distance $\frac{1}{5}$ of the central path. □

# Additional Points

- For convergence to a desired $\eta_f$, from an initial $\eta_0$, we require

$$k := \frac{\log(\frac{\eta_f}{\eta_0})}{\log(\beta)}$$

iterations.

- In practice the step is assigned as

$$n_\mu(x) := -H(x)^{-1}(c + \mu g(x)) = -(c_x + \mu g_x(x))$$
$$x_2 := x_1 + n_{\mu_2}(x_1)$$

where $\mu := \eta^{-1}$. Note this means that $\mu_2 = \beta^{-1}\mu_1$.

# Example

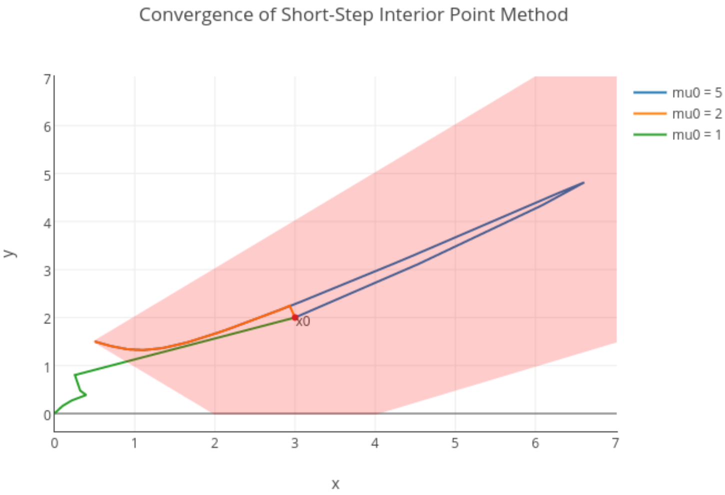
$$\begin{aligned} \underset{x}{\text{minimize}} \quad & 2x_1 + x_2 \\ \text{subject to} \quad & -x_1 + x_2 \leq 1 \\ & x_1 + x_2 \geq 2 \\ & x_1 - 2x_2 \leq 4 \\ & x_2 \geq 0 \end{aligned}$$

We attempt to solve this problem using the short-step technique outlined above. 3 slack variables are introduced to replace the first three inequality constraints. The solver parameters are,

$$x_0 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \qquad \mu_0 \in \{1, 2, 5\}$$

with the slack variables assigned to ensure feasibility.

# Visualization



# References I

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