

Interior Point Algorithms II: Computation of Analytic Center

Yinyu Ye

Department of Management Science and Engineering

Stanford University

Stanford, CA 94305, U.S.A.

<http://www.stanford.edu/~yyye>

Introduction

As we mentioned in the preceding chapter, a favorable property of the analytic center is that it is relatively easy to compute. In this lecture, we discuss how to compute the analytic center using the Newton method.

Then we proceed to discuss how to solve linear programming problems by the analytic center method.

Analytic Center Condition

$$\mathcal{B}(\mathbf{s}) = \sum_{j=1}^n \ln s_j = \sum_{j=1}^n \ln(\mathbf{c} - A^T \mathbf{y})_j$$

for $\mathcal{F} = \{\mathbf{y} \in \mathbf{R}^m : \mathbf{s} = \mathbf{c} - A^T \mathbf{y} \geq \mathbf{0}\}$. Its derivatives with respect to \mathbf{y} :

$$\nabla \mathcal{B}_y(\mathbf{s}) = -AS^{-1}\mathbf{e}, \quad \text{and} \quad \nabla^2 \mathcal{B}_y(\mathbf{s}) = -AS^{-2}A^T.$$

Analytic center conditions: $AS^{-1}\mathbf{e} = \mathbf{0}$, $\mathbf{s} = \mathbf{c} - A^T \mathbf{y} > \mathbf{0}$ or

$$\begin{aligned} S\mathbf{x} &= \mathbf{e} \\ A\mathbf{x} &= \mathbf{0} \\ -A^T \mathbf{y} - \mathbf{s} &= -\mathbf{c} \\ \mathbf{s} &> \mathbf{0}. \end{aligned} \tag{1}$$

Proximity to Analytic Center

$$\begin{aligned}
 \mathbf{p}(\mathbf{s}) &:= (-\nabla^2 \mathcal{B}(\mathbf{y}))^{-1/2} \nabla \mathcal{B}(\mathbf{s}) \\
 &= S^{-1} A^T (AS^{-2} A^T)^{-1} \nabla \mathcal{B}(\mathbf{s}) \\
 &= -S^{-1} A^T (AS^{-2} A^T)^{-1} AS^{-1} \mathbf{e}.
 \end{aligned} \tag{2}$$

$$\begin{aligned}
 \|\mathbf{p}(\mathbf{s})\|^2 &= \nabla \mathcal{B}(\mathbf{s})^T (-\nabla^2 \mathcal{B}(\mathbf{s}))^{-1} \nabla \mathcal{B}(\mathbf{s}) \\
 &= \mathbf{e}^T S^{-1} A^T (AS^{-2} A^T)^{-1} AS^{-1} \mathbf{e}.
 \end{aligned} \tag{3}$$

$\|\mathbf{p}(\mathbf{s})\| = 0$ is a centrality measure of \mathbf{y} . If $\|\mathbf{p}(\mathbf{s})\| = 0$, then $AS^{-1} \mathbf{e} = \mathbf{0}$ and \mathbf{y} is the analytic center.

Primal-Dual Representation

Consider the problem

$$\begin{aligned} &\text{minimize} && \|S\mathbf{x} - \mathbf{e}\|^2 \\ &\text{s.t.} && A\mathbf{x} = \mathbf{0}. \end{aligned}$$

The minimizer is

$$\mathbf{x}(\mathbf{s}) = S^{-1}(I - S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1})\mathbf{e}. \quad (4)$$

Note that

$$A\mathbf{x}(\mathbf{s}) = \mathbf{0}, \quad \text{and} \quad S\mathbf{x}(\mathbf{s}) = \mathbf{e} - S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}\mathbf{e} = \mathbf{e} + \mathbf{p}(\mathbf{s}),$$

so that

$$\mathbf{p}(\mathbf{s}) = S\mathbf{x}(\mathbf{s}) - \mathbf{e}.$$

The Newton Procedure

Given $(\mathbf{y}, \mathbf{s}) \in \mathcal{F}$, we call it an η -approximate (analytic) center if $\|\mathbf{p}(\mathbf{s})\| \leq \eta < 1$. Starting from such a (\mathbf{y}, \mathbf{s}) , the Newton procedure would be used to generate a better quality center.

Rewrite the system of equations for analytic center as

$$\begin{aligned} \mathbf{x} - S^{-1}\mathbf{e} &= \mathbf{0} \\ A\mathbf{x} &= \mathbf{0} \\ -A^T\mathbf{y} - \mathbf{s} + \mathbf{c} &= \mathbf{0}, \end{aligned}$$

and apply the Newton step for direction $(\mathbf{d}_y, \mathbf{d}_s, \mathbf{d}_x)$

$$\begin{aligned} \mathbf{d}_x + S^{-2}\mathbf{d}_s &= -\mathbf{x} + S^{-1}\mathbf{e} \\ A\mathbf{d}_x &= \mathbf{0} \\ -A^T\mathbf{d}_y - \mathbf{d}_s &= \mathbf{0}. \end{aligned}$$

Multiplying A to the top equation and noting $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{d}_x = \mathbf{0}$, we have

$$AS^{-2}\mathbf{d}_s = AS^{-1}\mathbf{e},$$

which together with the third equation give

$$\mathbf{d}_y = -(AS^{-2}A^T)^{-1}AS^{-1}\mathbf{e} \quad \text{and} \quad \mathbf{d}_s = A^T(AS^{-2}A^T)^{-1}AS^{-1}\mathbf{e}. \quad (5)$$

(Thus, to compute $(\mathbf{d}_y, \mathbf{d}_s)$ we don't need any \mathbf{x} .) Finally, we update (\mathbf{y}, \mathbf{s}) to

$$\mathbf{y}^+ := \mathbf{y} + \mathbf{d}_y \quad \text{and} \quad \mathbf{s}^+ := \mathbf{s} + \mathbf{d}_s.$$

Note that from (2) and (4), we see

$$\mathbf{d}_s = -S\mathbf{p}(\mathbf{s}) = -S(S\mathbf{x}(\mathbf{s}) - \mathbf{e}).$$

Quadratic Convergence

Theorem 1 Let (\mathbf{y}, \mathbf{s}) be an interior point and $\|\mathbf{p}(\mathbf{s})\| < 1$. Then,

$$\mathbf{s}^+ > \mathbf{0} \quad \text{and} \quad \|\mathbf{p}(\mathbf{s}^+)\| \leq \|\mathbf{p}(\mathbf{s})\|^2.$$

$$\|\mathbf{p}(\mathbf{s}^+)\| = \|S^+ \mathbf{x}(\mathbf{s}^+) - \mathbf{e}\| \leq \|S^+ \mathbf{x}(\mathbf{s}) - \mathbf{e}\|,$$

where $\mathbf{x}(\mathbf{s}) > \mathbf{0}$. But

$$\begin{aligned} \|S^+ \mathbf{x}(\mathbf{s}) - \mathbf{e}\|^2 &= \|(2S - S^2 X(\mathbf{s}))\mathbf{x}(\mathbf{s}) - \mathbf{e}\|^2 \\ &= \sum_{j=1}^n (s_j x(\mathbf{s})_j - 1)^4 \\ &\leq \left(\sum_{j=1}^n (s_j x(\mathbf{s})_j - 1)^2\right)^2 \\ &= \|S\mathbf{x}(\mathbf{s}) - \mathbf{e}\|^4 = \|\mathbf{p}(\mathbf{s})\|^4 < 1. \end{aligned}$$

Cutting Plane Again

Let $(\bar{\mathbf{y}}, \bar{\mathbf{s}})$ be an η -approximate (analytic) center of \mathcal{F} , that is, $\|\mathbf{p}(\mathbf{s})\| \leq \eta < 1$.

If one inequality in \mathcal{F} , say the first one, needs to be translated: change $\mathbf{a}_1^T \mathbf{y} \leq c_1$ to $a_1^T y \leq \gamma \mathbf{a}_1^T \bar{\mathbf{y}} + (1 - \gamma)c_1$. Let

$$\mathcal{F}^+ := \{\mathbf{y} : \mathbf{a}_j^T \mathbf{y} \leq c_j^+, j = 1, \dots, n\},$$

where $c_j^+ = c_j$ for $j = 2, \dots, n$ and $c_1^+ = \gamma \mathbf{a}_1^T \bar{\mathbf{y}} + (1 - \gamma)c_1$ for some constant $0 \leq \gamma \leq 1$.

We know there would be a sizable reduction in terms of the analytic volume from \mathcal{F} to \mathcal{F}^+ . But, could we computer an η -approximate (analytic) center of \mathcal{F}^+ efficiently ?

Cutting Plane Again

Since $\|\mathbf{p}(\bar{\mathbf{s}})\| \leq \eta < 1$, there is $\bar{\mathbf{x}}$ such that $A\bar{\mathbf{x}} = \mathbf{0}$ and $\|\bar{S}\bar{\mathbf{x}} - \mathbf{e}\| \leq \eta$, which implies $\bar{s}_1\bar{x}_1 \leq (1 + \eta)$.

Note that $\bar{\mathbf{y}}$ remains an interior point of the new polytope \mathcal{F}^+ with new slacks $\bar{s}_j^+ = \bar{s}_j$ for $j = 2, \dots, n$ and

$$\bar{s}_1^+ = \gamma \mathbf{a}_1^T \bar{\mathbf{y}} + (1 - \gamma)c_1 - \mathbf{a}_1^T \bar{\mathbf{y}} = (1 - \gamma)\bar{s}_1.$$

Thus, the centrality measure of $\bar{\mathbf{y}}$ for the new polytope

$$\|\bar{S}^+ \mathbf{x}(\bar{\mathbf{s}}^+) - \mathbf{e}\| \leq \|\bar{S}^+ \bar{\mathbf{x}} - \mathbf{e}\| \leq \|\bar{S}\bar{\mathbf{x}} - \mathbf{e}\| + \gamma(\bar{s}_1\bar{x}_1) = (1 + \gamma)\eta + \gamma.$$

That is, $\bar{\mathbf{y}}$ is a $((1 + \gamma)\eta + \gamma)$ -approximate (analytic) center of \mathcal{F}^+ .

Let $\eta = 1/4$ and $\gamma = 1/5$. Then, $(1 + \gamma)\eta + \gamma = 1/2$, so that one Newton step will update $\bar{\mathbf{y}}$ to an η -approximate (analytic) center of \mathcal{F}^+ .

Multiple Cutting Planes

If k inequalities in \mathcal{F} , say the first k , needs to be translated: change $\mathbf{a}_j^T \mathbf{y} \leq c_j$ to $\mathbf{a}_j^T \mathbf{y} \leq \gamma \mathbf{a}_j^T \bar{\mathbf{y}} + (1 - \gamma)c_j$. Let

$$\mathcal{F}^+ := \{\mathbf{y} : \mathbf{a}_j^T \mathbf{y} \leq c_j^+, j = 1, \dots, n\},$$

where $c_j^+ = c_j$ for $j = k + 1, \dots, n$ and $c_j^+ = \gamma \mathbf{a}_j^T \bar{\mathbf{y}} + (1 - \gamma)c_j$ for $j = 1, \dots, k$.

Again, $\bar{\mathbf{y}}$ remains an interior point of the new polytope with new slacks $\bar{s}_j^+ = \bar{s}_j$ for $j = k + 1, \dots, n$, and for $j = 1, \dots, k$

$$\bar{s}_j^+ = \gamma \mathbf{a}_j^T \bar{\mathbf{y}} + (1 - \gamma)c_j - \mathbf{a}_j^T \bar{\mathbf{y}} = (1 - \gamma)\bar{s}_j.$$

Thus, the centrality measure of $\bar{\mathbf{y}}$ for the new polytope

$$\|\bar{S}^+ \mathbf{x}(\bar{\mathbf{s}}^+) - \mathbf{e}\| \leq \|\bar{S}^+ \bar{\mathbf{x}} - \mathbf{e}\| \leq \|\bar{S} \bar{\mathbf{x}} - \mathbf{e}\| + \gamma \|\bar{S} \bar{\mathbf{x}} - \mathbf{e}\| + \sqrt{k}\gamma = (1 + \gamma)\eta + \sqrt{k}\gamma.$$

That is, $\bar{\mathbf{y}}$ is a $((1 + \gamma)\eta + \sqrt{k}\gamma)$ -approximate (analytic) center of \mathcal{F}^+ .

Let $\eta = 1/4$ and $\gamma = 1/(5\sqrt{k})$. Then, $(1 + \gamma)\eta + \sqrt{k}\gamma \leq 1/2$, so that one Newton step will update $\bar{\mathbf{y}}$ to an η -approximate (analytic) center of \mathcal{F}^+ .

But the analytic volume would be reduced at a rate $\exp(-\sqrt{k}/5)$. When $k = O(n)$, the reduction rate would be $\exp(-O(\sqrt{n})/5)$.

The Analytic Center Method for LP

Consider the problem

$$\begin{aligned} &\text{maximize} && \mathbf{b}^T \mathbf{y} \\ &\text{s.t.} && A^T \mathbf{y} \leq \mathbf{c}. \end{aligned}$$

Assume that the feasible region is bounded, and the analytic center of the region is \mathbf{y}^0 .

Start with a polytope

$$\mathcal{F}^0 := \{ \mathbf{y} : A^T \mathbf{y} \leq \mathbf{c}, \overbrace{-\mathbf{b}^T \mathbf{y} \leq R, \dots, -\mathbf{b}^T \mathbf{y} \leq R}^{n \text{ times}} \}$$

where R is so large such that \mathbf{y}^0 is an (approximate) analytic center of \mathcal{F}^0 .

Define a new polytope \mathcal{F}^{k+1} by translating the n objective hyperplanes near \mathbf{y}^k , and compute the (approximate) analytic center \mathbf{y}^{k+1} of \mathcal{F}^{k+1} .