Polynomial Convergence of an Infeasible-Interior-Point Method for Self-Scaled Conic Programming

Bharath Kumar Rangarajan and Michael J. Todd

Cornell University, Ithaca, NY

http://www.orie.cornell.edu/~{bharath|miketodd}/

Outline

- Problem Definition
- Algorithm
- Outline of Analysis
- Indicators of Infeasibility
- Summary

Problem Definition

$$(P) \qquad \min\{\langle c, x\rangle : \quad Ax = b,\ x \in K\}.$$

$$(D) \qquad \max\{\langle b, y\rangle : \quad A^*y + s = c,\ s \in K^*\}.$$

E and Y are finite-dimensional real vector spaces, $b \in Y^*$ and $c \in E^*$.

- K is a *regular* closed convex cone in E.
- K^* is the dual cone in E^*.

Examples

Linear Programming

$$
\begin{array}{ll}
(P) & \min\{c^T x : \quad Ax = b,\ x \geq 0\}, \\
(D) & \max\{b^T y : \quad A^T y + s = c,\ s \geq 0\},
\end{array}
$$

where $A \in \Re^{m \times n}$, $b \in \Re^m$, and $c \in \Re^n$.

Semidefinite Programming

$$
\begin{array}{ll}
(P) & \min \left\{ \text{trace}(C^T X) : \mathcal{A}(X) = b, X \succeq 0 \right\}, \\
(D) & \max\{b^T y : \ \sum_i A_i y_i + S = C,\ S \succeq 0\}
\end{array}
$$

$\mathcal{A}(X) = (\text{trace}(A_i^T X))_{i=1}^m$, each A_i and C are $n \times n$ symmetric matrices, and $b \in \Re^m$.

Weak Duality: For x feasible in (P) and (y, s) feasible in (D)

$$\langle c, x\rangle - \langle b, y\rangle = \text{complementarity} := \langle s, x\rangle \geq 0.$$

Assumptions : The problems (P) and (D) have strictly feasible solutions and A is surjective.

F and F_* are $\underline{\nu\text{-self-scaled}}$ barriers for K and K^* respectively.

Lemma 1 *[**NT**] For any $(x, s) \in int\ K \times int\ K^*$, there exists a unique scaling point $w := w(x, s) \in int\ K$ such that $F''(w)x = s$.*

Outline

- Problem Definition
- Algorithm
- Outline of Analysis
- Indicators of Infeasibility
- Summary

Central Path Equations

$$\begin{array}{rcccl} & A^* y & + & s & = & c, \\ Ax & & & & = & b, \\ \mu F'(x) & & + & s & = & 0, \\ x \in \text{int } K, & & & s \in \text{int } K^*. & & \end{array}$$

$(x(\mu), y(\mu), s(\mu))$ - unique minimizer and maximizer of the barrier problems.

- central path := $\{(x(\mu), y(\mu), s(\mu)) : \mu > 0\}$.
- $\langle s(\mu), x(\mu) \rangle = \mu\nu$.

Relevant Literature

- Linear Programming (infeasible-interior-point)
  - Kojima, Megiddo and Mizuno - Global convergence
  - Zhang; Mizuno; Potra - Polynomial convergence
- Semidefinite Programming (infeasible-interior-point)
  - Zhang; Potra and Sheng
- Self-Scaled Conic Programs (feasible-interior-point)
  - Nesterov and Todd; Schmieta and Alizadeh
- Our Work: Self-Scaled Conic Programs (infeasible-interior-point)

Outline of Interior-Point Methods

- Start with given initial point (x_0, y_0, s_0).
- From (x_k, y_k, s_k) to $(x_{k+1}, y_{k+1}, s_{k+1})$:
  - Set $\mu_k = \frac{\langle s_k, x_k \rangle}{\nu}$.
  - Aim towards $(x(\beta_1 \mu_k), y(\beta_1 \mu_k), s(\beta_1 \mu_k))$ for $\beta_1 \in (0,1)$ by moving in a Newton direction.
  - Take a positive step α_k along that direction.
  - Set this point to be $(x_{k+1}, y_{k+1}, s_{k+1})$.
- Repeat until some termination criterion is met.

Newton Equations **[NT]**

$$
\begin{array}{rcrcl}
 A^*\triangle y & + & \triangle s & = & c - A^*y - s, \\
A\triangle x & & & = & b - Ax, \\
F''(w)\triangle x & + & \triangle s & = & h \;:=\; -\beta_1 \mu F'(x) - s,
\end{array}
$$

where $\mu = \frac{\langle s,x \rangle}{\nu}$ and $\beta_1 \in (0,1)$ is a given parameter.

Note : If $x(\alpha) = x + \alpha \triangle x$, $(y(\alpha), s(\alpha)) = (y,s) + \alpha(\triangle y, \triangle s)$.

$$
\begin{aligned}
&Ax(\alpha) - b = (1-\alpha)(Ax - b); \\
&A^*y(\alpha) + s(\alpha) - c = (1-\alpha)(A^*y + s - c).
\end{aligned}
$$

Step Length and Termination

1. Compute step length $0 < \alpha_k < 1$ such that for all $\alpha \in [0, \alpha_k]$
   - stay in the neighborhood
   - linear decrease in complementarity
   - "relative complementarity" $\geq$ "relative infeasibility" (ϕ_k).
2. If $\langle s_k, x_k \rangle < \epsilon_* \langle s_0, x_0 \rangle$, then STOP.

- Linear decrease in complementarity:

for $1 > \beta_2 > \beta_1$,

$$\langle s(\alpha), x(\alpha) \rangle \leq \langle s_k, x_k \rangle (1 - (1 - \beta_2)\alpha).$$

- "Relative complementarity" $\geq$ "relative infeasibility" (ϕ_k).

$$\frac{\langle s(\alpha), x(\alpha) \rangle}{\langle s_0, x_0 \rangle} \geq \phi_k(1 - \alpha) \text{ and}$$
$$\phi_k(1 - \alpha) = \frac{\|Ax(\alpha) - b\|}{\|Ax_0 - b\|} = \frac{\|A^* y(\alpha) + s(\alpha) - c\|^*}{\|A^* y_0 + s_0 - c\|^*}.$$

Outline

- Problem Definition
- Algorithm
- Outline of Analysis
- Indicators of Infeasibility
- Summary

Main Theorem *Given* (A, b, c, K, K^*) *and* $\beta_1, \theta_G, \epsilon_* > 0$, *we can obtain a solution* (x_*, y_*, s_*) *such that* $\langle s_*, x_* \rangle \leq \epsilon_* \langle s_0, x_0 \rangle$ *and* $\phi_* \leq \epsilon_*$ *in* $O(\nu^{2.5} \ln \left(\frac{1}{\epsilon_*}\right))$ *iterations.*

Proof Outline: If $\alpha_k \geq \alpha_* = \Omega(\nu^{-2.5})$ for every k, then for $k = \left\lceil \frac{1}{(1-\beta_2)\alpha_*} \ln \left(\frac{1}{\epsilon_*}\right) \right\rceil = O(\nu^{2.5} \ln \left(\frac{1}{\epsilon_*}\right))$,

$$\begin{aligned} \langle s_k, x_k \rangle &\leq \langle s_0, x_0 \rangle \left(1 - \alpha_*(1 - \beta_2)\right)^k \leq \epsilon_* \langle s_0, x_0 \rangle \\ \phi_k &\leq \frac{\langle s_k, x_k \rangle}{\langle s_0, x_0 \rangle} \leq \epsilon_*. \end{aligned}$$

$$\|Ax_k - b\| \leq \epsilon_* \|Ax_0 - b\|, \text{ and}$$
$$\|A^* y_k + s_k - c\|^* \leq \epsilon_* \|A^* y_0 + s_0 - c\|^*.$$

$\square$

Bounding the Search Direction

Proposition 2 *There exists ω independent of k such that,*

$$\begin{aligned} \|\triangle x_k\|_w^2 + \|\triangle s_k\|_w^{*2} &\leq \omega \langle s_k, x_k \rangle \ \textit{and} \\ |\langle \triangle s_k, \triangle x_k \rangle| &\leq \frac{\omega}{2} \langle s_k, x_k \rangle . \end{aligned}$$

Feasible-interior-point methods: $\langle \triangle s_k, \triangle x_k \rangle = 0$.

Lower bound on α_*

- Stay in the neighborhood $\mathcal{N}_G(\theta_G)$.

$$\mathcal{N}_G(\theta_G) \quad := \quad \{(x,y,s) \in \text{int } K \times Y \times \text{int } K^* : \gamma_G(x,s) \le \theta_G\},$$

where $\gamma_G(x,s) := \mu \langle F'(x), F'_*(s)\rangle - \nu$.

$$\gamma_G(x(\alpha), s(\alpha)) \quad \le \quad \gamma_G - \alpha\beta_1 \frac{\gamma_G(\gamma_G + \nu)}{\nu} + \alpha^2 \tau$$

for all $\alpha \in [0, \bar{\alpha}_1]$, where $\bar{\alpha}_1 := \left(2\sqrt{(\theta_G + 2)\nu\omega}\right)^{-1}$.

$$\tau \quad = \quad O(\nu^{3/2}\omega).$$

For $\bar{\alpha}_2 := \frac{\beta_1 \theta_G}{\tau}$,

$(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_G$ for all $\alpha \in [0, \bar{\alpha}_2]$.

- "Relative complementarity" $\geq$ "relative infeasibility"

$$\text{For } \bar{\alpha}_3 := \frac{2\beta_1}{\omega},$$
$$\langle s(\alpha), x(\alpha) \rangle \geq \phi(1-\alpha) \langle s_0, x_0 \rangle \text{ for all } \alpha \in [0, \bar{\alpha}_3].$$

- Linear decrease in complementarity

$$\text{For } \bar{\alpha}_4 := \frac{2(\beta_2 - \beta_1)}{\omega},$$
$$\langle s(\alpha), x(\alpha) \rangle \leq \langle s, x \rangle (1 - \alpha(1-\beta_2)) \text{ for all } \alpha \in [0, \bar{\alpha}_4].$$

Polynomial bound on ω

Let (u_0, r_0, v_0) be the least-squares solution to $Au = b$, $A^* r + v = c$ using $\|u\| + \|v\|^*$.
Let $x_0 := \rho_0 e \in \text{int } K,\ s_0 := -\rho_0 F'(e) \in \text{int } K^*$ for $\rho_0 > \|u_0\| + \|v_0\|^*$.

$$\rho_* := \min\{\max(|x_*|_e, |s_*|_e^*) : (x_*, y_*, s_*) \text{ solves (P) and (D)}\}.$$

Assumption There exists a constant $\Psi > 0$ such that

$$\boxed{\rho_0 \geq \frac{\rho_*}{\Psi}.}$$

- $\omega = O(\nu)$.
- $\tau = O(\omega \nu^{1.5}) = O(\nu^{2.5})$.
- $\boxed{\alpha_* = \min(1, \bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3, \bar{\alpha}_4) = \Omega(\nu^{-2.5}).}$

Recall Main Theorem *Given* (A, b, c, K, K^*) *and* $\beta_1, \theta_G, \epsilon_* > 0$, *we can obtain a solution* (x_*, y_*, s_*) *such that* $\langle s_*, x_* \rangle \leq \epsilon_* \langle s_0, x_0 \rangle$ *and* $\phi_* \leq \epsilon_*$ *in* $O(\nu^{2.5} \ln \left(\frac{1}{\epsilon_*}\right))$ *iterations.*

Proof Outline: $\alpha_* = \Omega(\nu^{-2.5})$. After
$k = \left\lceil \frac{1}{(1-\beta_2)\alpha_*} \ln \left(\frac{1}{\epsilon_*}\right) \right\rceil = O(\nu^{2.5} \ln \left(\frac{1}{\epsilon_*}\right))$ iterations,

$$\begin{aligned} \langle s_k, x_k \rangle &\leq \epsilon_* \langle s_0, x_0 \rangle, \text{ and} \\ \phi_k &\leq \frac{\langle s_k, x_k \rangle}{\langle s_0, x_0 \rangle} \leq \epsilon_*. \end{aligned}$$

$$\begin{aligned} \|Ax_k - b\| &\leq \epsilon_* \|Ax_0 - b\|, \text{ and} \\ \|A^* y_k + s_k - c\|^* &\leq \epsilon_* \|A^* y_0 + s_0 - c\|^*. \end{aligned}$$

Outline

- Problem Definition
- Algorithm
- Outline of Analysis
- Indicators of Infeasibility
- Summary

Indicators of Infeasibility

- Large optimal solutions

$$\rho := \max(|x_0 - u_0|_e, |s_0 - v_0|_e^*), \quad \underline{\phi} = \min(\phi_p, \phi_d).$$

* Stopping Rule 1. For some $\tilde{\rho}$, stop if

$$\frac{\phi_p \langle s, x_0 - u_0 \rangle + \phi_d \langle s_0 - v_0, x \rangle}{\langle s, x \rangle} \geq \left(1 + \frac{\rho(2\tilde{\rho} + \underline{\phi}\rho)}{\rho_0^2}\right).$$

Theorem 3 *If stopping rule* 1 *applies, then there is no optimal solution pair* x_* *and* (y_*, s_*) *for* (P) *and* (D) *with* $|x_*|_e \leq \tilde{\rho}$ *and* $|s_*|_e^* \leq \tilde{\rho}$.

- Large feasible solutions

* Stopping Rule 2_p. Let $r = y - \phi_d(y_0 - r_0)$. Then, for some $\bar{\rho}_p > 0$, stop if

$$\langle b, r \rangle \geq \|c + \phi_d(s_0 - v_0)\|^* \ \bar{\rho}_p.$$

* Stopping Rule 2_d. Let $u = x - \phi_p(x_0 - u_0)$. Then, for some $\bar{\rho}_d > 0$, stop if

$$\langle c, u \rangle \leq -\max\left(\|b\|^*, \ \phi_p\|x_0 - u_0\|\right) \ \bar{\rho}_d.$$

Theorem 4 *If stopping rule 2_p applies, then any feasible solution to (P) has norm at least $\bar{\rho}_p$; if stopping rule 2_d applies, then any feasible solution to (D) has $\|y\| + \|s\|^*$ at least $\bar{\rho}_d$.*

- Large optimal solutions imply large feasible solutions

Theorem 5 *If*

$$\tilde{\rho} \geq \frac{1}{2\rho\bar{\phi}\nu} \left[\|c + \phi_d(s_0 - v_0)\|^* \ \bar{\rho}_p + \max\left(\|b\|^*, \ \phi_p\|x_0 - u_0\|\right) \ \bar{\rho}_d \right],$$

where $\bar{\phi} = \max(\phi_p, \phi_d)$*, then if stopping rule* 1 *applies, so does either* 2_p *or* 2_d*.*

Outline

- Problem Definition
- Algorithm
- Outline of Analysis
- Indicators of Infeasibility
- Summary

Summary

- $O(\nu^{2.5})$ convergence using the $\mathcal{N}_G$ neighborhood.
- In practice, binary (or line) searches can be done to improve step-sizes.
- Can allow different (primal and dual) step sizes.
- Lower bounds on size of optimal as well as feasible solutions.
- Can obtain $O(\nu^4)$ convergence for a given $\mathcal{N}_\infty$ neighborhood.