Detecting Infeasibility in Conic Optimization: Theory and Practice

Imre Pólik

McMaster University
Advanced Optimization Lab

CORS 2005, Halifax
Outline

1 Introduction
   - Importance of infeasibility

2 Theory
   - Classical results
   - Approximate results
   - Interior Point Methods (IPMs)
   - Self-dual embedding

3 Practice
   - Stopping criteria
   - Experiments
   - Results
   - Future plans
Motivation

- Importance of infeasibility
  - No solution.
    - Why? Certificate!
  - What does it mean?
    - Good news?
    - Wrong model? Wrong data?
    - Numerical problems?
    - Bug in the code?

- Practical problems
  - Not known a priori
  - Feasible but impractical solution
    - Missing constraints
  - Weakly infeasible problems
Classical results

- The cone: $\mathcal{K} \subset \mathbb{R}^n$ closed, convex, pointed, solid
- Dual cone: $\mathcal{K}^* = \{ s \in \mathbb{R}^n : x^T s \geq 0, \forall x \in \mathcal{K} \}$
- Ordering: $x \succeq_K 0 \iff x \in \mathcal{K}$
- Conic optimization problem (CO)
  \[
  \begin{align*}
    \min & \quad c^T x \\
    \text{subject to} & \quad Ax = b \\
    & \quad A^T y + s = c \\
    & \quad x \succeq_K 0 \\
    & \quad s \succeq_{\mathcal{K}^*} 0
  \end{align*}
  \]

- Weak duality: $c^T x - b^T y = x^T s \geq 0$
- Primal-dual strict feasibility implies strong duality.
- Dual improving direction
  \[
  A^T y \succeq_{\mathcal{K}^*} 0 \\
  b^T y = 1.
  \]
  - Primal is feasible $\Rightarrow$ No dual improving direction
  - Primal is infeasible $\Rightarrow$ Almost improving direction
Self-dual cones

- Linear
  \[ \mathbb{R}^n \]

- Second order
  \[ \mathcal{K}^q = \left\{ x \in \mathbb{R}^n : x_1^2 \geq \|x_{2:n}\|^2, x_1 \geq 0 \right\} \]

- Rotated second order
  \[ \mathcal{K}^r = \left\{ x \in \mathbb{R}^n : 2x_1x_2 \geq \|x_{3:n}\|^2, x_1, x_2 \geq 0 \right\} \]

- Positive semidefinite
  \[ \mathcal{S}^n = \left\{ X \in \mathbb{R}^{n \times n} : X \succeq 0 \right\} \]

- And any product of these!
Approximate Farkas Lemma for CO

Near-solvability
\[ \alpha_x = \min \|x\|_\infty \quad \beta_u = \min \|u\|_1 \]
\[ Ax = b \]
\[ x \succeq_K 0 \]

Theorem
\[ \alpha_x \beta_u = 1. \]

Perturbed system:
\[ \alpha_x^\varepsilon := \min \|x\|_\infty \]
\[ Ax = b^\varepsilon \]
\[ x \succeq_K v^\varepsilon \]
\[ \|b - b^\varepsilon\|_\infty \leq \varepsilon \]
\[ \|v^\varepsilon\|_\infty \leq \varepsilon. \]

- \[ \alpha_x^\varepsilon \rightarrow \alpha_x \ (\varepsilon \rightarrow 0) \]
- If the original is feasible then \[ \alpha_x^\varepsilon \]
  and \[ \alpha_x \] are realized.
IPMs for Conic Optimization

- Optimality conditions
  
  \[
  \begin{align*}
  Ax &= b \\
  A^T y + s &= c \\
  xs &= 0 \\
  x, s &\succeq 0,
  \end{align*}
  \]

- The Newton-system
  
  \[
  \begin{align*}
  A\Delta x &= 0 \\
  A^T \Delta y + \Delta s &= 0 \\
  \Delta xs + x\Delta s &= \mu I - xs
  \end{align*}
  \]

- Starting point?
IPMs for Conic Optimization

- Perturbed optimality conditions
  \[ Ax = b \]
  \[ A^T y + s = c \]
  \[ xs = \mu I \]
  \[ x, s \succ 0, \]

- The Newton-system
  \[ A\Delta x = 0 \]
  \[ A^T \Delta y + \Delta s = 0 \]
  \[ \Delta xs + x\Delta s = \mu I - xs \]

- Starting point?
IPMs for Conic Optimization

- Perturbed optimality conditions
  \[ Ax = b \]
  \[ A^T y + s = c \]
  \[ x s = \mu I \]
  \[ x, s \succ 0, \]

- The Newton-system
  \[ A \Delta x = 0 \]
  \[ A^T \Delta y + \Delta s = 0 \]
  \[ \Delta x s + x \Delta s = \mu I - x s \]

- Starting point?
IPMs for Conic Optimization

- Perturbed optimality conditions

\[
\begin{align*}
Ax & = b \\
A^T y + s & = c \\
x s & = \mu I \\
x, s & \succ 0,
\end{align*}
\]

- The Newton-system

\[
\begin{align*}
A \Delta x & = 0 \\
A^T \Delta y + \Delta s & = 0 \\
\Delta x s + x \Delta s & = \mu I - x s
\end{align*}
\]

- Starting point?
Self-dual embedding model for CO

- Strictly interior starting point

\[
\min (x_0^T s_0 + 1) \theta \\
Ax - b \tau + \bar{b} \theta = 0 \\
-A^T y + c \tau - \bar{c} \theta = s \\
b^T y - c^T x + \bar{z} \theta = \kappa \\
-\bar{b}^T y - \bar{c}^T x - \bar{z}^T \tau = -x_0^T s_0 - 1
\]

\[x \succeq_{\kappa} 0, \ \tau \geq 0, \ s \succeq_{\kappa^*} 0, \ \kappa \geq 0,\]

where
- \(x_0, s_0 \in \mathbb{R}^n, \ y_0 \in \mathbb{R}^m, \ \tau, \theta, \kappa \in \mathbb{R}\)
- \(\bar{b} = b - Ax_0, \ \bar{c} = c - A^T y_0 - s_0, \ \bar{z} = c^T x_0 - b^T y_0 + 1\)

- \(\tau > 0, \ \theta = 0: \ x/\tau, y/\tau, s/\tau \) are optimal
- \(\tau = 0, \ \theta > 0: \) infeasibility
- \(\tau = 0, \ \theta = 0: \) ?
Stopping criteria for the SD model I.

\[
\min (x_0^T s_0 + 1) \theta \\
Ax - b\tau + \bar{b}\theta = 0 \\
-A^T y + c\tau - \bar{c}\theta = s \\
b^T y - c^T x + \bar{z}\theta = \kappa \\
-\bar{b}^T y - \bar{c}^T x - \bar{z}^T \tau = -x_0^T s_0 - 1 \\
x \succeq_K 0, \ \tau \geq 0, \ s \succeq_K^* 0, \ \kappa \geq 0,
\]

Stop if

- optimality:
  \[
  \max \left( \|Ax - b\tau\|, \left\| A^T y + s - c\tau \right\|_*, \ c^T x - b^T y \right) < \overline{\varepsilon} \tau
  \]

- large primal feasible solutions:
  \[
  b^T y > (\tau \|c\|_* + \theta \|\bar{c}\|_*) \overline{\rho}
  \]

- large dual feasible solutions:
  \[
  c^T x > -(\tau \|b\|_* + \theta \|\bar{b}\|_*) \overline{\rho}
  \]

Complexity: \( O \left( \sqrt{n} \ln \frac{\overline{\rho}}{\overline{\varepsilon}} \right) \)
Stopping criteria for the SD model II.

- Assume $\tau \kappa \geq (1 - \beta)\theta$
- $\min (x_0^T s_0 + 1)\theta$
- $Ax - b\tau + \bar{b}\theta = 0$
- $-A^Ty + c\tau - \bar{c}\theta = s$
- $b^Ty - c^Tx + \bar{z}\theta = \kappa$
- $-\bar{b}^Ty - \bar{c}^Tx - \bar{z}^T\tau = -x_0^T s_0 - 1$
- $x \succeq \kappa 0$, $\tau \geq 0$, $s \succeq \kappa^* 0$, $\kappa \geq 0$,

- Stop if
  - optimality:
    \[
    \max (\|Ax - b\tau\|, \|A^Ty + s - c\tau\|_*, c^Tx - b^Ty) < \bar{\varepsilon}\tau
    \]
  - large optimal solutions ($x^*^T s_0 + s^*^T x_0 \geq \rho$):
    \[
    \tau \leq \frac{1 - \beta}{1 + \rho}
    \]

- Complexity: $\mathcal{O} \left( \sqrt{n} \ln \frac{\bar{\rho}}{\bar{\varepsilon}} \right)$
Experiments using SeDuMi

What is SeDuMi?
- IPM solver for symmetric cone optimization
- Self-dual embedding
- Open source (GPL), Matlab and C
- Originally by Jos F. Sturm (until 2003)
- Now by AdvOL (McMaster University)
- See http://sedumi.mcmaster.ca

Starting point: scaled identity
- Only the first criterion
- Difficulty: few good(!) infeasible problems
### A typical infeasible run

SeDuMi 1.1 by AdvOL, 2005

- **eqs m = 10, order n = 31, dim = 901, blocks = 2**

<table>
<thead>
<tr>
<th>it</th>
<th>b*y</th>
<th>gap</th>
<th>feas</th>
<th>prec</th>
<th>pnorm</th>
<th>tau</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>3.56E+1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7.44E+0</td>
<td>2.40E+1</td>
<td>1.71</td>
<td>3.7E+2</td>
<td>0.0441</td>
<td>1.1137</td>
</tr>
<tr>
<td>2</td>
<td>1.38E+1</td>
<td>1.16E+1</td>
<td>1.69</td>
<td>1.5E+2</td>
<td>0.1150</td>
<td>1.3545</td>
</tr>
<tr>
<td>3</td>
<td>1.81E+1</td>
<td>4.51E+0</td>
<td>1.42</td>
<td>5.2E+1</td>
<td>0.1810</td>
<td>1.4779</td>
</tr>
<tr>
<td>4</td>
<td>4.69E+1</td>
<td>1.47E+0</td>
<td>0.34</td>
<td>3.6E+1</td>
<td>0.5307</td>
<td>0.6770</td>
</tr>
<tr>
<td>5</td>
<td>3.79E+2</td>
<td>4.73E-1</td>
<td>-1.87</td>
<td>1.0E+2</td>
<td>4.2900</td>
<td>0.0761</td>
</tr>
<tr>
<td>6</td>
<td>7.04E+2</td>
<td>1.00E-1</td>
<td>-2.15</td>
<td>4.1E+1</td>
<td>7.9618</td>
<td>0.0402</td>
</tr>
<tr>
<td>7</td>
<td>2.78E+3</td>
<td>3.23E-2</td>
<td>-0.90</td>
<td>5.3E+1</td>
<td>31.4465</td>
<td>0.0313</td>
</tr>
<tr>
<td>8</td>
<td>1.30E+5</td>
<td>8.78E-4</td>
<td>-1.02</td>
<td>6.7E+1</td>
<td>1.4671E+3</td>
<td>0.0006</td>
</tr>
<tr>
<td>9</td>
<td>2.31E+5</td>
<td>3.72E-4</td>
<td>-1.12</td>
<td>5.1E+1</td>
<td>2.6118E+3</td>
<td>0.0004</td>
</tr>
<tr>
<td>10</td>
<td>1.14E+6</td>
<td>7.89E-5</td>
<td>-0.98</td>
<td>5.3E+1</td>
<td>1.2870E+4</td>
<td>0.0003</td>
</tr>
<tr>
<td>11</td>
<td>1.14E+8</td>
<td>8.14E-7</td>
<td>-1.00</td>
<td>5.4E+1</td>
<td>1.2888E+6</td>
<td>0</td>
</tr>
</tbody>
</table>

Primal infeasible, dual improving direction found.
Observations, conclusions

- It works!
- Useful information
- Special cases, strong bounds
  - SDP, bounded diagonal
  - Natural bounds
- Theoretical properties
  - Polynomial-time stopping criteria
  - Robust, reliable
  - No prior knowledge
  - Numerical accuracy matters
Future work

- Utilize the $x^* s_0 + s^* x_0 \geq \rho$ bound
  - Depends on the starting point
  - Primal or dual?
- Weak infeasibility
  - Not implemented anywhere
  - Decision problem
- Derive strong bounds on solutions (preprocessing)
- Statistical methods
  - Similar runs
  - Time series analysis, prediction
  - Algorithm modification
  - No certificate! (uncertainty)
Detecting Infeasibility in Conic Optimization: Theory and Practice

Imre Pólik

McMaster University
Advanced Optimization Lab

CORS 2005, Halifax