Cutting Plane Methods for Conic Programs

John E. Mitchell
Math Sciences
Rensselaer Polytechnic Institute
mitchj@rpi.edu

May 16, 2005

Coauthors:

Luc Basescu, Kartik Krishnan, Mohammad Oskoorouchi

Outline

1. Warm starting in conic programs
2. Approximating SDP problems by smaller problems.
3. A long step version of Renegar's algorithm.

1: Warm starting in conic programs

Convex optimization

$$\begin{array}{lll} \max & < b, y > & \\ \text{subject to} & y \in Y & (OPT) \end{array}$$

Assumptions:

- Y is a convex, bounded set contained in $\mathbb{R}^m$, with a nonempty interior.
- If $\bar{y} \notin Y$, a separation oracle returns a conic inequality $G^* y + s = h$, $s \in K_0$ satisfied by all $y \in Y$ and violated by $\bar{y}$. K_0 is a full-dimensional self-scaled cone in $\mathbb{R}^p$.
- Assume the cone K_0 has a self-concordant barrier function $f_0(K_0)$.

Conic relaxation

The problem (OPT) is approximated by a **conic program**:

$$\begin{array}{llcccl} \max & < b, y > & & & & \\ \text{subject to} & A^*y & + & s & = & c \\ & & & s & \in & K \end{array} \qquad (CD)$$

where K is a full-dimensional self-scaled cone.

Note that K may be a product of smaller cones. Eg, $K = \mathbb{R}^n_+$, $K =$ SDP cone, $K =$ product of SOCP cones, . . .

Duality

Call (CD) the dual problem. The corresponding primal problem is:

$$\begin{array}{lrcl} \min & & < c, x > & \\ \text{subject to} & Ax & = & b \\ & x & \in & K \end{array} \qquad (CP)$$

Note that we assume that K is self-dual.

Cutting planes

- We find an **approximate analytic center** $(\bar{x}, \bar{y}, \bar{s})$ for (CP) and (CD).
- If $\bar{y} \in Y$, can continue as with a standard interior point method.
- Otherwise, the oracle returns a cut violated by $\bar{y}$ and satisfied by all $y \in Y$:

$$\begin{array}{rcrcl} G^* y & + & s & = & h \\ & & s & \in & K_0 \end{array}$$

- We assume the cut is shifted to be **central**, that is, $h = G^* \bar{y}$.
- Note that the modified primal problem can still be restarted effectively, even if the cut is deep.

Modifying (CP) and (CD)

$$
\begin{array}{llllll}
\max & < b, y > & & & & \\
\text{subject to} & A^* y & + & s & = & c \qquad (\overline{CD}) \\
 & G^* y & + & s_0 & = & c_0 \\
 & s \in K, & & s_0 & \in & K_0
\end{array}
$$

$$
\begin{array}{llllll}
\min & < c, x > & + & < h, x_0 > & & \\
\text{subject to} & Ax & + & Gx_0 & = & b \qquad (\overline{CP}) \\
 & x \in K, & & x_0 & \in & K_0
\end{array}
$$

With $x_0 = 0$ and $s_0 = 0$, the current point $(\bar{x}, \bar{y}, \bar{s})$ is on the boundary of the feasible regions of $(\overline{CD})$ and $(\overline{CP})$.

Dikin Ellipsoids

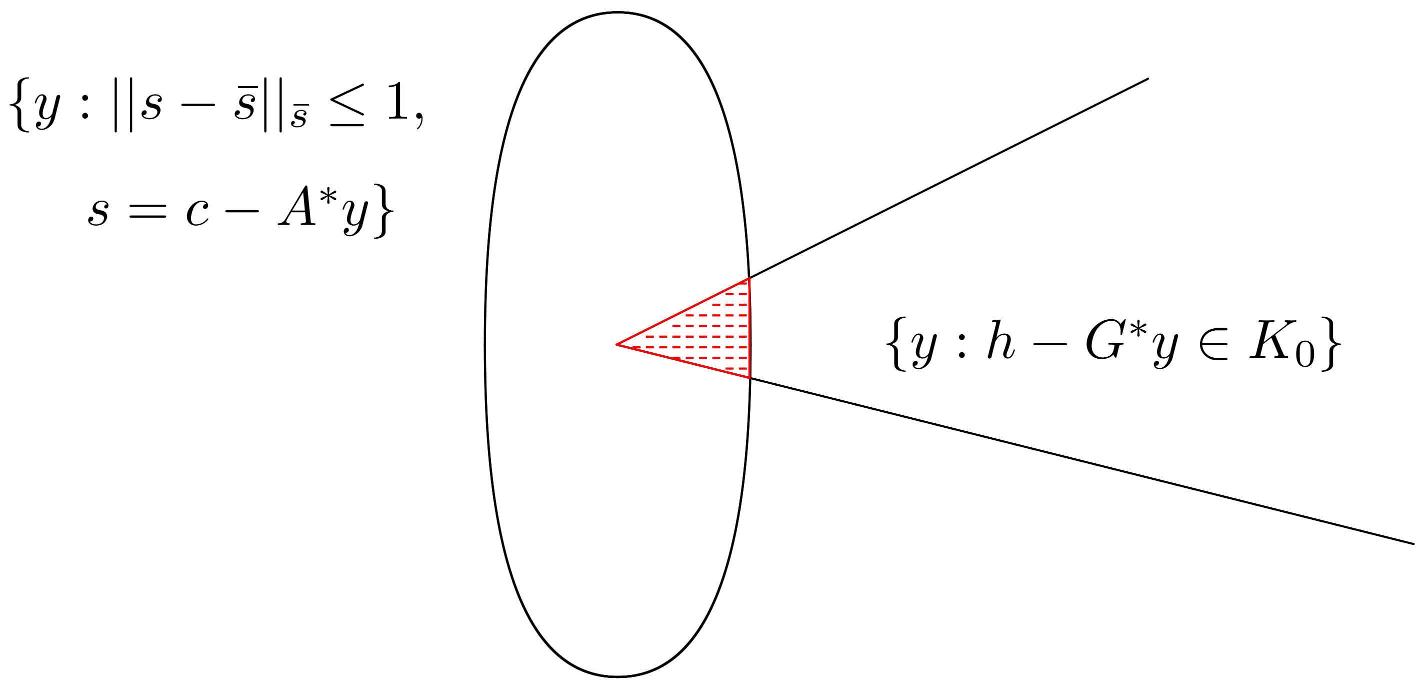
- Want to find s_0 and x_0 in the **interior** of K_0. Assume the cone K_0 has a self-concordant barrier function $f_0(K_0)$, with conjugate function $f_0^*(K_0)$.

- Try to **minimize this barrier function over s_0 (or x_0) subject to keeping y, s (or x) in an appropriately defined Dikin ellipsoid**:

$$
\begin{array}{lrcrcl}
\min & & & f_0^*(s_0) & & \\
\text{subject to} & A^*y & + & s & = & c \\
 & G^*y & + & s_0 & = & c_0 \\
 & & & ||s - \bar{s}||_{\bar{s}} & \leq & 1 \\
 & & & s_0 & \in & K_0
\end{array}
$$

- Note that $||v||_u := ||H(u)^{1/2}v||$, where $H(u)$ is the Hessian of f.

Dikin Ellipsoid Illustration



Due to Goffin and Vial for LP, generalized to SDP and SOCP by Oskoorouchi and Goffin, and to more general problems by Basescu and Mitchell. Generalizes the single cut approach of Mitchell and Todd.

Finding a new approximate analytic center

- Look at feasibility problem (so $b = 0$).

- Let ϑ_f denote the **complexity value of $\mathbf{f}$**, in the terminology of Renegar's text. For an LP with n variables, the complexity value of $-\sum_{i=1}^{n} \log(x_i)$ is $\vartheta_f = n$. For SDP with $n \times n$ matrix X, the complexity value of $-\det(X)$ is also n. For a single SOCP constraint, the complexity value of the barrier function is $\vartheta_f = 2$.

- The **primal-dual potential function** is

$$\Phi_{PD} \;=\; < x, s > + f(x) + f^*(s).$$

It is minimized with value 0 at the analytic center.

Updating the potential function

- The contribution to the potential function of the new point from the additional variables x_0 and s_0 is equal to

$$-\vartheta_{f_2} - 2\vartheta_{f_2} \ln \alpha + \vartheta_{f_2} \ln \vartheta_{f_2}$$

  since gradient $g_2(s_0) = -\vartheta_{f_2} x_0$. (Here, α is the steplength.)

- The change in the potential function due to the old variables x and s is $O(1)$ for appropriate α since the new point sits within Dikin ellipsoids.

- Get convergence to new approximate analytic center in $O(\vartheta_{f_2} \ln \vartheta_{f_2})$ steps.

More on the change in the potential function

- Assume G is injective and G^* is surjective.

- Find x_0 (and hence $s_0 = Vx_0$) by solving

$$\begin{array}{ll} \min & \frac{\vartheta_{f_2}}{2} < x_0, Vx_0 > + f_2(x_0) \\ \text{subject to} & x_0 \in K_0 \end{array}$$

where $V = G^*(AH(\bar{s})A^*)^{-1}G$. Here $H(s)$ is the Hessian of $f^*(s)$.

- Easy to initialize. Can then solve using Newton or quasi-Newton method.

- This is an extension of Goffin/Vial/Oskoorouchi approach.

More on the change in the potential function (continued)

- The primal Dikin ellipsoid problem is

$$\begin{array}{llcrcl} \min & & & f_0(x_0) & & \\ \text{subject to} & A(x-\bar{x}) & + & Gx_0 & = & 0 \\ & ||x-\bar{x}||_{H(\bar{s})^{-1}} & & & \leq & 1 \end{array}$$

- The solutions to the primal and dual Dikin ellipsoid problems are closely related. Get $g_2(s_0) = -\vartheta_{f_2} x_0$. For SDP, this gives $X_0 S_0 = \frac{1}{p} I$ if X_0 and S_0 are $p \times p$ matrices.

2: Approximating SDP problems by smaller problems

Cutting plane model

$$
\begin{array}{llcl}
(SDP) & & & \\
\min & C \bullet X & & \\
\text{subject to} & \mathcal{A}(X) & = & b \\
 & X & \succeq & 0
\end{array}
\qquad
\begin{array}{llcccc}
(SDD) & & & & & \\
\max & b^T y & & & & \\
\text{subject to} & \mathcal{A}^T(y) & + & S & = & C \\
 & & & S & \succeq & 0
\end{array}
$$

- Approximate $S \succeq 0$ by $P_i^T S P_i \succeq 0$ for some matrices P_i.
- The primal matrix X must then satisfy $X = \sum_i P_i V_i P_i^T$ for some positive semidefinite matrices V_i.

Choices for $\mathbf{P_i}$

- Only choose one matrix P_i, update this matrix, and limit the number of columns in this matrix: basis for spectral bundle method of Helmberg and Rendl, and simplex method of Krishnan, Pataki, and Terlaky.

- Each matrix P_i has only one column: basis of LP relaxation method of Krishnan and Mitchell.

- Each matrix P_i has a small number of columns: basis of method of Oskoorouchi and Goffin.

- Each matrix P_i has one or two columns: ongoing research (Oskoorouchi, Krishnan, Mitchell).

SDP constraints with one or two columns

- Columns of P_i typically chosen to be eigenvectors of $C - A^T(\bar{y})$ with negative eigenvalue.

- The single column constraints give linear cuts.

- The SDP constraints where P_i has two columns can be expressed equivalently as rotated SOCP constraints:

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} \succeq 0 \Leftrightarrow ac - b^2 \geq 0,\ a \geq 0,\ c \geq 0 \Leftrightarrow \begin{pmatrix} a + c \\ 2b \\ a - c \end{pmatrix} \succeq_K 0$$

- Computational results: see talks by Oskoorouchi and by Krishnan.

Restarting after adding SOCP constraints in practice

- The added cuts are **deep**, that is, strictly violated by the current point.
- Can work just in (CP), since $\bar{x}$ is still feasible (Oskoorouchi and Mitchell).
- With SDP relaxations of combinatorial optimization problems, it is usually straightforward to find a strictly feasible dual iterate in the modified relaxation, so primal-dual methods can be used.

3: A long step version of Renegar's algorithm

- The results on warm starts can be used to show that a long step version of Renegar's algorithm converges in $O(nL)$ iterations.

- Assume solving standard dual form linear program, and the problem has an optimal value and a strictly feasible solution:

$$\begin{array}{lrcl} \max & b^T y & & \\ \text{subject to} & A^T y & \leq & c \end{array} \qquad (LD)$$

A is $m \times n$.

Renegar's algorithm

1. Get initial lower bound l on optimal value of (LD).

2. Find approximate analytic center $\bar{y}$ of the polyhedron

$$\begin{array}{l} A^T y \leq c \\ \left.\begin{array}{l} b^T y \geq l \\ \ldots \\ b^T y \geq l \end{array}\right\} \quad m \text{ copies} \end{array}$$

3. If $\bar{y}$ not good enough, update l to $(1-\delta)l + \delta b^T \bar{y}$, with $\delta = O(\frac{1}{\sqrt{n}})$.

4. Return to Step 2 and get a new approximate analytic center in one step.

A long step version of Renegar's algorithm

1. Get initial lower bound l on optimal value of (LD).

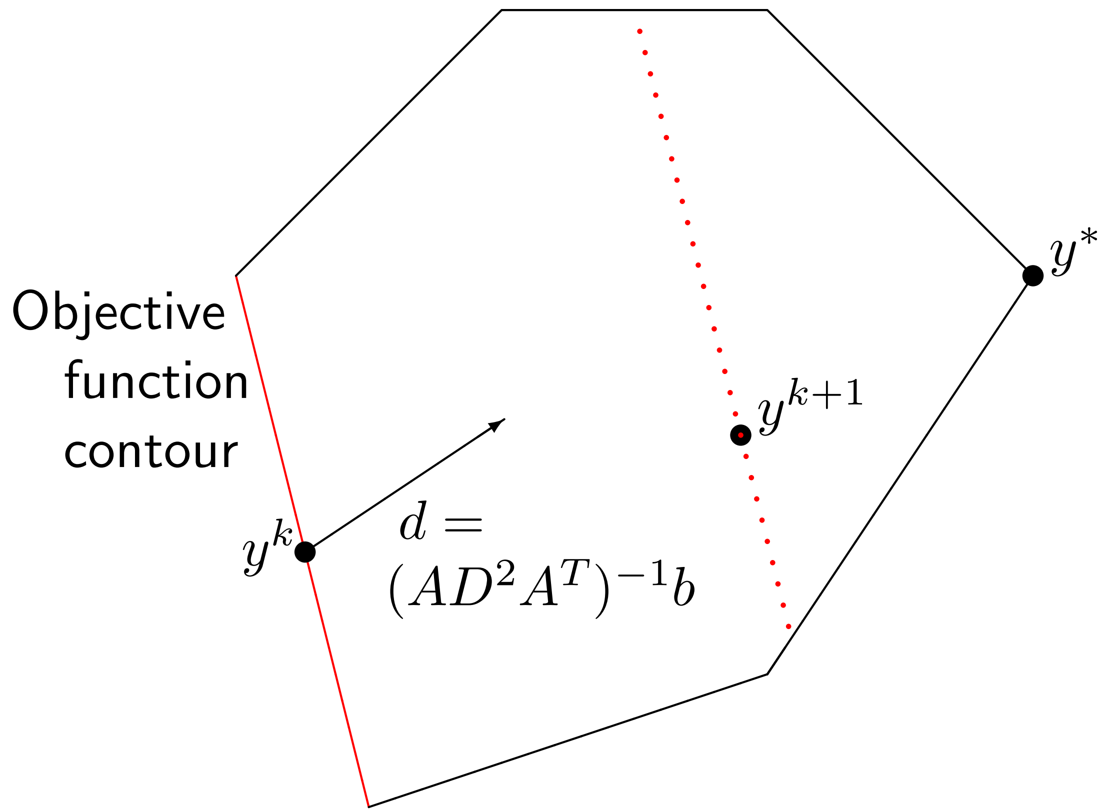
2. Find approximate analytic center $\bar{y}$ of the polyhedron

$$
\begin{array}{l}
A^T y \leq c \\
\left.\begin{array}{l} b^T y \geq l \\ \ldots \\ b^T y \geq l \end{array}\right\} \quad m \text{ copies}
\end{array}
$$

3. If $\bar{y}$ not good enough, **update $\mathbf{l}$ to $\mathbf{b^T \bar{y}}$**.

4. Return to Step 2 and get a new approximate analytic center.

Picture of the long step method



Properties of the long step method

- Let z^{opt} be the optimal value of (LP). Results of Renegar can be used to show that the value of $z^{opt} - b^T\bar{y}$ is at least **halved** at each iteration.

- Results of Goffin and Vial can be used to show that the new approximate analytic center can be found in $O(m\log(m))$ Newton steps.

- The restart direction is exactly the direction for adding the single cut $b^T y \geq b^T\bar{y}$, namely:
$$\Delta y = (AD^2A^T)^{-1}b.$$

- This restart direction is the "predictor" direction. Can then recover an approximate analytic center by taking steps that could be considered "corrector" steps.

Conclusions

- Can efficiently restart conic programs.
- Can solve SDP problems as SOCP problems.
- Can devise a long step version of Renegar's algorithm using the theory of warm starts.