Active Constraint Indicators in Interior Point Algorithms for Conic Optimization

Alexander Guldemond

July 2025

Mosek Aps

Outline

Table of Contents

Introduction

A typical interior-point algorithm, on success, returns a solution that satisfies some approximate termination criteria. Both the iterative nature of the algorithm itself and the nature of finite-precision arithmetic make it extremely difficult to return an exact solution. In a well-posed problem (one that is either linear or has strictly feasible primal and dual formulations), the returned primal-dual solution will lie within the strict interior of the constraint cone. It will be close to complementary, but not exactly. This makes determining which constraints are active very difficult when given just the computed solution.

Consider the standard form primal-dual conic optimization problems:

Primal	**Dual**
$$\begin{aligned} \min_{x} \quad & \langle c, x \rangle \\ \text{s.t.} \quad & Ax = b \\ & x \in \mathcal{K} \end{aligned}$$	$$\begin{aligned} \max_{y,s} \quad & \langle b, y \rangle \\ \text{s.t.} \quad & A^{\mathrm{T}} y + s = c \\ & s \in \mathcal{K}^*, \end{aligned}$$

where $\mathcal{K}$ is a proper cone and $\mathcal{K}^*$ is its dual.

Mosek's conic optimizer solves the conic optimization problem using the **Simplified Homogeneous Model**:

$$
\begin{aligned}
Ax - \tau b &= 0 \\
A^{\mathrm{T}} y + s - \tau c &= 0 \\
-c^{\mathrm{T}} x + b^{T} y - \kappa &= 0 \\
x \in \mathcal{K},\ s &\in \mathcal{K}^{\star} \\
\tau \geq 0,\ \kappa &\geq 0.
\end{aligned}
$$

It seeks to find a solution to the above that satisfies an additional complementarity condition

$$
x^{\mathrm{T}} s + \tau\kappa = 0.
$$

Properties of the Simplified Homogeneous Model

The homogeneous model has a number of convenient properties that make it a good choice for a computational solver. Let $(x^\star, y^\star, \tau^\star, s^\star\kappa^\star)$ be a complementary solution to the model. Then:

1. If $\tau^\star > 0$, $x^\star/\tau^\star$ is an optimal primal solution and $(y^\star, s^\star)/\tau^\star$ is an optimal dual solution to the original problem.
2. If $\kappa^\star > 0$, $\langle b, y^\star \rangle > 0$ and the primal problem is infeasible, or $\langle c, x \rangle < 0$ and the dual problem is infeasible, or both.

Thus, a complementary solution to the simplified homogeneous model with $\tau^\star + \kappa^\star > 0$ provides either a scaled solution to the original problem or a certificate of infeasibility. If no such solution exists, we say that the original problem is **ill-posed**.

However, users may not appreciate that a model may take a long time to run, only to return with a certificate of infeasibility. It is helpful to provide them with a tool to evaluate whether the model should continue running depending on which way the solution is trending.

The termination criteria is based on the following quantities:

$$\rho_p^k = \min\{\ \rho\ :\ \left|\left|A\frac{x}{\tau} - b\right|\right|_\infty \leq \rho\varepsilon_p(1 + ||b||_\infty)\}$$

$$\rho_d^k = \min\{\ \rho\ :\ \left|\left|A^T\frac{y}{\tau} + \frac{s}{\tau} - c\right|\right|_\infty \leq \rho\varepsilon_p(1 + ||c||_\infty)\}$$

$$\rho_g^k = \min\left\{\rho\ :\ \min\left(\frac{\langle s, x\rangle}{\tau^2}, \left|\frac{\langle c, x, \rangle - \langle b, y\rangle}{\tau}\right|\right) \leq \rho\varepsilon_g \max\left(1, \frac{\min\{|\langle c, x\rangle|, |\langle b, y\rangle|\}}{\tau}\right)\right\}$$

Feasibility Indicators

A crucial step in any interior-point solver is the computation of the **affine-search direction**. In the homogeneous model, the affine-search direction (using Nesterov-Todd scaling) is given by the solution to the following equations:

$$
\begin{aligned}
Ad_x - d_\tau b &= -(Ax - \tau b) \\
A^T d_y + d_s - d\tau c &= -(A^T y + s - \tau c) \\
-c^T d_x + b^T d_y - d\kappa &= -(-c^T x + b^T y - \kappa) \\
Wdx + W^{-T} ds &= -v \\
\kappa d\tau + \tau d\kappa &= -\tau\kappa.
\end{aligned}
$$

where W and v are determined by scaling.

Consider the final equation $\kappa d\tau + \tau d\kappa = -\tau\kappa$. Then we have

$$\frac{d\tau}{\tau} + \frac{d\kappa}{\kappa} = -1$$

for every affine search direction $(d\tau, d\kappa)$ with corresponding iterate (τ, κ). Let $(x^{(k)}, y^{(k)}, \tau^{(k)}, s^{(k)}, \kappa^{(k)})$ denote the k^{th} iterate, and $(d_x^{(k)}, d_y^{(k)}, d\tau^{(k)}, ds^{(k)}, d\kappa^{(k)})$ the k^{th} affine search direction. Then, as $(\tau^{(k)}, \kappa^{(k)}) \to (\tau^\star, \kappa^\star)$, $(d\tau^{(k)}, d\kappa^{(k)}) \to (0, 0)$. If $\tau^\star > 0$, then $\kappa^\star = 0$, and

$$\frac{d\kappa^{(k)}}{\kappa^{(k)}} \to -1.$$

Alternatively, if $\kappa^\star > 0$, then $\tau^\star = 0$, and

$$\frac{d\tau^{(k)}}{\tau^{(k)}} \to -1.$$

Feasibility Indicator

At each iteration, we compute a **feasibility indicator** given by

$$\frac{d\tau^{(k)}}{\tau^{(k)}} - \frac{d\kappa^{(k)}}{\kappa^{(k)}},$$

and provide this value to the user via an interface to a callback. If the iterates are converging to a point with $\tau^\star > 0$, then this value will converge to -1. If either the primal or dual problem is infeasible, it will converge to 1. If the value doesn't seem to be converging, the user may conclude the model is ill-posed.

Table of Contents

In many instances, users also wish to know, or at least have a heuristic of, which constraints are active at a solution. Particularly, in the case of linear programs, this can be an import aspect of using the solution from an interior-point solver to find a true basic solution. In particular, Mosek implements an optimal basis identification procedure after the interior-point algorithm has terminated. To this end, we extend the feasibility indicator framework to all conic constraints.

In the simplest case, consider a single linear constraint $x \geq 0$, with corresponding dual constraint $s \geq 0$. As before, it holds that

$$\frac{dx^{(k)}}{x^{(k)}} - \frac{ds^{(k)}}{s^{(k)}} \to -1$$

if $x \geq 0$ is active, and

$$\frac{dx^{(k)}}{x^{(k)}} - \frac{ds^{(k)}}{s^{(k)}} \to 1$$

if $s \geq 0$ is active. Internally, this is tracked at each iteration by simply checking if

$$\frac{dx^{(k)}}{x^{(k)}} < \frac{ds^{(k)}}{s^{(k)}}.$$

If this holds, an internal flag is set to indicate that the primal constraint is the active constraint, and if not, the dual constraint is active.

Second Order Cone Constraints

In the case of non-trivial cones, the situation becomes slightly more complicated. Consider second-order cone problem

$$
\begin{aligned}
\min_{x \in \mathbb{R}^5} \quad & 2x_2 + 3x_3 + 4x_4 + 5x_5 \\
\text{s.t.} \quad & x_2 + 2x_3 - x_4 + x_5 = 1 \\
& x_2 - 2x_4 - x_5 = 1 \\
& x_1 = 2 \\
& x_2 \geq -0.5, \quad x_3 \geq 0.1, \quad x_4 \geq 0.2 \\
\text{and} \quad & x_1 \geq \sqrt{x_2^2 + x_3^2 + x_4^2 + x_5^2}
\end{aligned}
$$

After forming the homogeneous model and solving, the Mosek solver returns a primal-dual solution

Mosek Output

```
x = [ 7.943537542447e-01, -2.211664031425e-02, 5.384464234853e-01,
7.943532564971e-02, -5.781640863102e-01 ]
...
s = [ 1.1082006528498896e+00, 3.0920695408106994e-02,
-7.5120087698657878e-01, -1.1076909538909678e-01,
8.0658542468112682e-01 ]
```

Evaluating

$$f(x) = \frac{x_1^2 - (x_2^2 + x_3^2 + x_4^2 + x_5^2)}{2}$$

yields $f(x) = 2.54 \times 10^{-7}$ and $f(s) = 9.282 \times 10^{-11}$. Analytically, it can in fact be shown that both primal and dual constraints are active.

Second-Order Cone Indicators

In second-order cone programming using Nesterov-Todd scaling, the affine search direction satisfies

$$Wdx + W^{-1}ds = -v,$$

Where W is the symmetric positive define scaling matrix satisfying $Wx = W^{-1}s = v$. For any vector x, let x^{-1} denote

$$x^{-1} = \frac{2}{x_1^2 - ||x_{(2:n)}||_2^2} \begin{pmatrix} x_1 \\ -x_{(2:n)} \end{pmatrix}.$$

Then for all vectors x, it holds that $\langle x^{-1}, x \rangle = 2$, (2 being the complexity of the quadratic cone).

The scaling matrix additionally satisfies

$$Ws^{-1} = W^{-1}x^{-1} = v^{-1}.$$

Thus,

$$\begin{aligned}\langle v^{-1}, Wdx + W^{-1}ds\rangle &= \langle W^{-1}x^{-1}, Wdx\rangle + \langle Ws^{-1}, W^{-1}ds\rangle \\ &= \langle x^{-1}, dx\rangle + \langle s^{-1}, ds\rangle \\ &= -2\end{aligned}$$

Our active-constraint indicator is therefore

$$\langle x^{-1}, dx\rangle - \langle s^{-1}, ds\rangle.$$

As before, the primal constraint is internally flagged as active if

$$\langle x^{-1}, dx\rangle < \langle s^{-1}, ds\rangle,$$

otherwise the dual constraint is flagged as active. The principle difference is that this is insufficient for concluding that the unflagged constraint is inactive.

Consider the previous example. This problem is solved in 10 iterations.

Iteration	**Second-Order Cone Indicator**
0	0.806654
1	-1.67049
2	-23.2541
3	0.0683571
4	-0.128352
5	-2.81298
6	-4.36849
7	-14.1536
8	-38.8899
9	-0.0142842

The behavior in this case is fairly erratic. However, the solver does give us enough information to make sense of this data. First, the solver converges to a complementary primal-dual feasible solution, and it concludes the problem is well-posed. Since both constraints cannot be inactive (as this would violate the complementarity of the solution), the lack of convergence to 2 or -2 indicates that both constraints are active. Let's examine another example.

Two Quadratic Cones

$$
\begin{aligned}
\min_{x \in \mathbb{R}^8} \quad & 2x_2 + 3x_3 + 4x_4 + 6x_6 + 7x_7 \\
\text{s.t.} \quad & x_2 + 2x_3 - x_4 - x_6 + x_7 = 1 \\
& x_2 - 2x_4 + 2x_7 = 1 \\
& x_1 = 1, \quad x_5 = 0.5, \quad x_8 = -0.5 \\
& 1 \leq x_6 \leq 2 \\
& x_1 \geq \sqrt{x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2} \\
& x_6 \geq \sqrt{x_7^2 + x_8^2}
\end{aligned}
$$

This problem has two second-order cone constraints. In the following table, we show the indicator values for the last 10 iterations:

Iteration	Second-Order Cone Indicator 1	Second-Order Cone Indicator 2
k-9	0.0106311	1.99998
k-8	-0.581197	2.0
k-7	0.0149345	2.00002
k-6	-0.597839	2.0
k-5	0.0263113	1.99998
k-4	0.0262816	1.99998
k-3	-0.567517	2.0
k-2	0.0318507	2.00001
k-1	-0.540986	2.0
k	0.0370409	1.99999

This data, combined with the fact that the solver converges to a complementary primal-dual feasible solution leads us to conclude that both primal and dual constraints are active in the first case, and only the dual constraint is active in the second case. Analytically solving the problem validates this conclusion.

Table of Contents

Non-symmetric Cones

The case of non-symmetric cones, such as the exponential cone and power cones, is remarkably similar to the second-order cone case. The affine search direction satisfies the equation

$$Wdx + W^{-T}ds = -v,$$

where the scaling matrix W is chosen from many possible matrices to satisfy

$$Wx = W^{-T}s = v, \quad \text{and} \quad W\bar{x} = W^{-T}\bar{s} = \bar{v},$$

where $\bar{x}$ and $\bar{s}$ are the so-called *shadow iterates* given by

$$\bar{s} = -\nabla F(x), \quad \bar{s} = -\nabla F_{\star}(s),$$

where F is the ν-logarithmically homogeneous self-concordant barrier function used by the non-symmetric cone.

A function is logarithmically homogeneous if it satisfies

$$F(ax) = F(x) - \nu \log(x) \quad \text{for all } x \in \mathcal{K}.$$

The conjugate barrier function $F_\star(s)$ is also logarithmically homogeneous. It additionally holds that

$$\langle -\nabla F(x), x \rangle = \nu, \quad \langle s, -\nabla F_\star(s) \rangle = \nu, \quad -\nabla F(x) \in \mathcal{K}^\star, \quad \text{and} \quad -\nabla F_\star(s) \in \mathcal{K}.$$

The indicator used is then

$$\langle \bar{s}, dx \rangle - \langle dx, \bar{x} \rangle,$$

which, in the case of one inactive constraint, will converge to either ν or $-\nu$.

Table of Contents

Future Work

We have found the most useful aspect of this active-set indicator framework to be its use in basis identification for linear problems. For now, we have only deeply experimented with indicating active constraints for symmetric cones. For non-symmetric cones, we have only begun some initial experimentation with the exponential cone, however initial results show promise.

The main difficulty with this approach comes from the ill-conditioning of the KKT system as the interior-point iterates converge to a solution. In particular, as the solver nears the solution, the accuracy of the affine search direction typically decreases. In particular, the accuracy of the affine search direction is considered a secondary task. As long as the final search direction generates a sufficient decrease in the termination criteria, the algorithm will not spend too much time refining the affine search direction.

Furthermore, in the case where both x and s converge to a boundary point, the computation of $\bar{s}$ and $\bar{x}$ becomes increasingly ill-conditioned, leading to a potential breakdown in

$$\bar{s}^T dx + \bar{x}^T ds = -\nu.$$

One potential way to mitigate this problem is to take a scaled-space approach. In a scaled-space approach, the updates to the primal and dual variables are handles in the scaled space, and are based on the updates

$$v_x^+ = v + W dx, \quad \text{and} \quad v_s^+ = v + W^{-T} ds.$$

The procedure for recovering x^+ and s^+ depends on the cone $\mathcal{K}$. Instead of solving for dx and ds directly, Wdx and $W^{-T}ds$ are explicitly solved by modifying the KKT system.

The active-set indicator then becomes

$$\bar{s}^T dx - \bar{x}^T ds = \bar{s}^T W^{-1} W dx - \bar{x}^T W^T W {-} T ds = \bar{v}^T W dx - \bar{v}^T W^{-T} ds$$

Since the iterates $\{v_i\}$ are theoretically much better conditioned than $\{x_i\}$ and $\{s_i\}$, the computation of

$$\bar{v}^T W dx - \bar{v}^T W^{-t} ds$$

should be far less prone to breakdown.