Second-order cone programming

Imre Pólik, PhD

Lehigh University
Department of Industrial and Systems Engineering

February 10, 2009

Outline

A new product

For $u, v \in \mathbb{R}^n$ define:

$$u \circ v = (u^T v; u_1 v_{2:n} + v_1 u_{2:n}).$$

Theorem (Properties of $\circ$)

1. *Distributive law:* $u \circ (v + w) = u \circ v + u \circ w$.
2. *Commutative law:* $u \circ v = v \circ u$.
3. *The unit element is* $\iota = (1; 0)$*, i.e.,* $u \circ \iota = \iota \circ u = u$.
4. *Using the notation* $u^2 = u \circ u$ *we have* $u \circ (u^2 \circ v) = u^2 \circ (u \circ v)$.
5. *Power associativity:* $u^p = u \circ \cdots \circ u$ *is well-defined.*
6. *Associativity does not hold in general.*

Spectral decomposition

Every vector $u \in \mathbb{R}^n$ can be written as

$$u = \lambda_1 c^{(1)} + \lambda_2 c^{(2)},$$

where $c^{(1)}$ and $c^{(2)}$ are on the boundary of the cone, and

$$\begin{aligned} {c^{(1)}}^T c^{(2)} &= 0 \\ c^{(1)} \circ c^{(2)} &= 0 \\ c^{(1)} \circ c^{(1)} &= c^{(1)} \\ c^{(2)} \circ c^{(2)} &= c^{(2)} \\ c^{(1)} + c^{(2)} &= \iota \end{aligned}$$

$c^{(1)}$, $c^{(2)}$: Jordan frame
λ_1, λ_2: eigenvalues or spectral values:
$\lambda_{1,2}(u) = u_1 \pm \|u_{2:n}\|_2$
Naturally: $u \in \mathbb{L} \Leftrightarrow \lambda_{1,2}(u) \geq 0$

The cone of squares

Theorem

A vector x is in a second order cone (i.e., $x_1 \geq \|x_{2:n}\|_2$) if and only if it can be written as the square of a vector under the multiplication $\circ$, i.e., $x = u \circ u$.

$$\begin{aligned}
\|u\|_F &= \sqrt{\lambda_1^2 + \lambda_2^2} = \sqrt{2}\,\|u\|_2\,, \\
\|u\|_2 &= \max\left\{|\lambda_1|\,, |\lambda_2|\right\} = |u_1| + \|u_{2:n}\|_2\,, \\
u^{-1} &= \lambda_1^{-1} c^{(1)} + \lambda_2^{-1} c^{(2)}, \\
u^{\frac{1}{2}} &= \lambda_1^{\frac{1}{2}} c^{(1)} + \lambda_2^{\frac{1}{2}} c^{(2)},
\end{aligned}$$

where $u \circ u^{-1} = u^{-1} \circ u = \iota$ and $u^{\frac{1}{2}} \circ u^{\frac{1}{2}} = u$.

The arrowhead operator

Since the mapping $v \mapsto u \circ v$ is linear, it can be represented with a matrix.

$$\operatorname{Arr}(u) = \begin{pmatrix} u_1 & u_2 & \dots & u_n \\ u_2 & u_1 & & \\ \vdots & & \ddots & \\ u_n & & & u_1 \end{pmatrix},$$

Now we have $u \circ v = \operatorname{Arr}(u)\, v = \operatorname{Arr}(u) \operatorname{Arr}(v)\, \iota$.
Quadratic representation:

$$Q_u = 2 \operatorname{Arr}(u)^2 - \operatorname{Arr}\left(u^2\right),$$

thus $Q_u(v) = 2u \circ (u \circ v) - u^2 \circ v$ is a quadratic function.

Primal-dual interior-point methods: notation

$$\mathcal{K} = \mathbb{L}_{n_1} \times \cdots \times \mathbb{L}_{n_k},$$

$$A = \left(A^1, \ldots, A^k\right),$$

$$x = \left(x^1; \ldots; x^k\right),$$

$$s = \left(s^1; \ldots; s^k\right),$$

$$c = \left(c^1; \ldots; c^k\right).$$

With this notation we can write

$$Ax = \sum_{i=1}^{k} A^i x^i,$$

$$A^T y = \left({A^1}^T y; , \ldots; {A^k}^T y\right).$$

$\mathrm{Arr}\,(u)$ and Q_u are block diagonal matrices built from the blocks $\mathrm{Arr}\,(u^i)$ and Q_{u^i}, respectively.

Duality and optimality

- Weak duality always holds
- Primal (dual) strict feasibility implies strong duality and dual (primal) solvability

Under strong duality, the optimality conditions for second order conic optimization are

$$\begin{aligned} Ax = b,\ x \in \mathcal{K} \\ A^T y + s = c,\ s \in \mathcal{K} \\ x \circ s = 0. \end{aligned}$$

An equivalent form of the complementarity condition is $c^T x - b^T y = x^T s = 0$.

The central path using barrier functions

If $x \in \operatorname{int} \mathbb{L}$, consider

$$\phi(x) = -\ln\left(x_1^2 - \|x_{2:n}\|_2^2\right) = -\ln \lambda_1(x) - \ln \lambda_2(x),$$

Goes to ∞ if x is getting close to the boundary of the cone. Derivatives:

$$\nabla \phi(x) = -2\frac{(x_1; -x_{2:n})^T}{x_1^2 - \|x_{2:n}\|_2^2} = -2\left(x^{-1}\right)^T,$$

where the inverse is taken in the Jordan algebra.

The central path

Perturbed optimality conditions:

$$
\begin{aligned}
Ax &= b,\ x \in \mathcal{K} \\
A^T y + s &= c,\ s \in \mathcal{K} \\
x^i \circ s^i &= 2\mu\iota^i,\ i = 1, \ldots, k,
\end{aligned}
$$

where $\iota^i = (1; 0; \ldots; 0) \in \mathbb{R}^{n_i}$.
Newton system:

$$
\begin{aligned}
A\Delta x &= 0 \\
A^T \Delta y + \Delta s &= 0, \\
x^i \circ \Delta s^i + \Delta x^i \circ s^i &= 2\mu\iota^i - x^i \circ s^i,\ i = 1, \ldots, k,
\end{aligned}
$$

where $\Delta x = (\Delta x^1; \ldots; \Delta x^k)$ and $\Delta s = (\Delta s^1; \ldots; \Delta s^k)$.

Newton system - rewritten

$$\begin{pmatrix} & A & \\ A^T & & I \\ & \operatorname{Arr}(s) & \operatorname{Arr}(x) \end{pmatrix} \begin{pmatrix} \Delta y \\ \Delta x \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2\mu\iota - x \circ s \end{pmatrix},$$

where $\iota = (\iota^1; \dots; \iota^k)$. Eliminating Δx and Δs:

$$\left(A \operatorname{Arr}(s)^{-1} \operatorname{Arr}(x) A^T\right) \Delta y = -A \operatorname{Arr}(s)^{-1} (2\mu\iota - x \circ s).$$

Problems:

- Not symmetric
- May be singular

Solution: symmetrization!

Scaling for SOCP

$$\begin{aligned} \min\ & \left(Q_{p^{-1}}c\right)^T \left(Q_p x\right) \\ \left(AQ_{p^{-1}}\right)\left(Q_p x\right) &= b \\ Q_p x &\in \mathcal{K} \end{aligned} \qquad \begin{aligned} \max\ & b^T y \\ \left(AQ_{p^{-1}}\right)^T y + Q_{p^{-1}} s &= Q_{p^{-1}} c \\ Q_{p^{-1}} s &\in \mathcal{K} \end{aligned}$$

Lemma

If $p \in \operatorname{int} \mathcal{K}$, *then*

1. $Q_p Q_{p^{-1}} = I$.
2. *The cone* $\mathcal{K}$ *is invariant, i.e.,* $Q_p(\mathcal{K}) = \mathcal{K}$.
3. *The scaled and the original problems are equivalent.*

Scaled optimality conditions

$$
\begin{aligned}
\left(AQ_{p^{-1}}\right)\left(Q_p \Delta x\right) &= 0 \\
\left(AQ_{p^{-1}}\right)^T \Delta y + Q_{p^{-1}} \Delta s &= 0, \\
\left(Q_p x\right) \circ \left(Q_{p^{-1}} \Delta s\right) + \left(Q_p \Delta x\right) \circ \left(Q_{p^{-1}} s\right) &= 2\mu\iota - \left(Q_p x\right) \circ \left(Q_{p^{-1}} s\right).
\end{aligned}
$$

Simplifies to

$$
\begin{aligned}
A \Delta x &= 0 \\
A^T \Delta y + \Delta s &= 0, \\
\left(Q_p x\right) \circ \left(Q_{p^{-1}} \Delta s\right) + \left(Q_p \Delta x\right) \circ \left(Q_{p^{-1}} s\right) &= 2\mu\iota - \left(Q_p x\right) \circ \left(Q_{p^{-1}} s\right).
\end{aligned}
$$

The last equation cannot be simplified!

The choice of p

AHO: $p = \iota$: does not provide a nonsingular Newton system

HKM:

$$p = s^{1/2} \text{ or } p = x^{1/2},$$

in which case

$$Q_{p^{-1}} s = \iota \text{ or } Q_p x = \iota.$$

Implemented in SDPT3.

NT: Most popular one.

$$p = \left(Q_{x^{1/2}} \left(Q_{x^{1/2}} s \right)^{-1/2} \right)^{-1/2} = \left(Q_{s^{-1/2}} \left(Q_{s^{1/2}} x \right)^{1/2} \right)^{-1/2}.$$

Simplifies to

$$Q_p x = Q_{p^{-1}} s.$$

Implemented in SeDuMi, MOSEK, SDPT3.

Centrality measures

$$\mu(x,s) = \sum_{i=1}^{k} \frac{{x^i}^T s^i}{n_i}.$$

$w = (w_1; \dots; w_k)$, where $w_i = Q_{x_i^{1/2}} s_i$.

$$\delta_F(x,s) := \left\| Q_{x^{1/2}} s - \mu\iota \right\|_F := \sqrt{\sum_{i=1}^{k} (\lambda_1(w_i) - \mu)^2 + (\lambda_2(w_i) - \mu)^2}$$

$$\delta_\infty(x,s) := \left\| Q_{x^{1/2}} s - \mu\iota \right\|_2 := \max_{i=1,\dots,k} \left\{ |\lambda_1(w_i) - \mu|, |\lambda_2(w_i) - \mu| \right\}$$

$$\delta_\infty^-(x,s) := \left\| (Q_{x^{1/2}} s - \mu\iota)^- \right\|_\infty := \mu - \min_{i=1,\dots,k} \left\{ \lambda_1(w_i), \lambda_2(w_2) \right\},$$

$$\delta_\infty^-(x,s) \le \delta_\infty(x,s) \le \delta_F(x,s).$$

Neighbourhoods

$$\mathcal{N}(\gamma) := \left\{ (x,y,s) \text{ strictly feasible} : \delta(x,s) \le \gamma\mu(x,s) \right\}.$$

$\delta(x,s) = \delta_F(x,s)$: narrow neighbourhood

$\delta(x,s) = \delta_\infty^-(x,s)$ wide neighbourhood

IPM for SOCP

Theorem (Short-step IPM for SOCO)

Choose $\gamma = 0.088$ *and* $\zeta = 0.06$. *Assume that we have a starting point* $(x^0, y^0, s^0) \in \mathcal{N}_F(\gamma)$. *Compute the Newton step from the scaled Newton system. In every iteration,* μ *is decreased to* $\left(1 - \frac{\zeta}{\sqrt{k}}\right)\mu$, *i.e.,* $\theta = \frac{\zeta}{\sqrt{k}}$, *and the stepsize is* $\alpha = 1$. *This algorithm finds an* ε*-optimal solution for the second order conic optimization problem with* k *second order cones in at most*

$$\mathcal{O}\left(\sqrt{k}\log\frac{1}{\varepsilon}\right)$$

iterations. (Independent of m, n*!) The cost of one iteration is*

$$\mathcal{O}\left(m^3 + m^2 n + \sum_{i=1}^{k} n_i^2\right).$$