The Q Method for Symmetric Cone Programmin

Farid Alizadeh and Yu Xia
alizadeh@rutcor.rutgers.edu, xiay@optlab.mcma

Large Scale Nonlinear and Semidefinite Progra
University of Waterloo, Canada
May 2004

⇒ • Basics of Jordan Algebras . . . . . .
• The Symmetric Cone Programs . . .
• The Q Method for SDP . . . .
• Polar Decomposition for V . . . .
• Optimal Conditions for Symmetric Cone Programs . . . .
• An Interior Point Algorithm . . . . .
• A Newton-type Method . . . .

Basics of Jordan Algebras

- **Euclidean Jordan Algebra.**
  1) A finite dimensional vector space V over $\mathbb{R}$ with
  mapping (product) from $V \times V$ into V:

$$\mathbf{x}\mathbf{y} = \mathbf{y}\mathbf{x}$$

$$\mathbf{x}(\mathbf{x}^2\mathbf{y}) = \mathbf{x}^2(\mathbf{x}\mathbf{y}).$$

  2) There exists an associative, positive definite symm
  bilinear form on V.

- **Jordan Frame.** $\mathbf{f}_i$ is non-zero and cannot be writte
  sum of two (necessarily orthogonal) non-zero idempo

$$\mathbf{f}_i^2 = \mathbf{f}_i$$

$$\mathbf{f}_i\mathbf{f}_j = \mathbf{0} \quad (i \neq j)$$

$$\mathbf{f}_1 + \cdots + \mathbf{f}_k = \mathbf{e}.$$

- **Spectral theorem.** For $\mathbf{x} \in V$, there exist a Jorda
$\mathbf{f}_1, \ldots, \mathbf{f}_r$ and real numbers $\lambda_1, \ldots, \lambda_r$ such that

$$\mathbf{x} = \sum_{j=1}^{r} \lambda_j \mathbf{f}_j. \text{ (For symmetric matrix } X = \sum_{j=1}^{r} \lambda$$

- **Symmetric cone.** Open, convex, homogeneous, se
cone. There are one-to-one correspondences betwee
cones and Euclidean Jordan algebras.

$$Sq \stackrel{\text{def}}{=} \{\mathbf{x}^2 : \mathbf{x} \in V\}, \quad \operatorname{int} Sq \text{ is a symmetric c}$$

$$\mathbf{x} \in Sq \Leftrightarrow \lambda_i \geq 0.$$

The Symmetric Cone Programs

Primal		**Dual**	
$\min_{\mathbf{x}}$	$\langle \mathbf{c}, \mathbf{x} \rangle$	$\max_{\mathbf{z},\,\mathbf{y}}$	$\langle \mathbf{b}, \mathbf{y} \rangle$
s.t.	$A\mathbf{x} = \mathbf{b}$	s.t.	$A^*\mathbf{y} + \mathbf{z} = \mathbf{c}$
	$\mathbf{x} \geq_{Sq} \mathbf{0}$		$\mathbf{z} \geq_{Sq} \mathbf{0}$

Notations

- $\mathbf{x} \in V$, $\mathbf{y} \in Y$.
- A: linear mapping from V to Y^*.
- $\langle \cdot, \cdot \rangle \stackrel{\text{def}}{=} \operatorname{tr}(xy) = \sum_{i=1}^{r} \lambda_i(\mathbf{xy})$.

The Q Method for SDP

We extend the Q method for SDP {Alizadeh, Haeberl
Overton: *SIAM J. Optim. 8 (1998)* } to symmetric cone

- **Semidefinite Programming (SDP)**

Primal		**Dual**	
$\min_X$	$C \bullet X$	$\max_{\mathbf{y},Z}$	$\mathbf{b}^T\mathbf{y}$
s.t.	$A_i \bullet X = b_i \; (i = 1, \ldots, m)$	s.t.	$\sum_{i=}^{m}$
	$X \succeq 0$		$Z \succeq$

- **Interior Point Method**

$$A_i \bullet X = b_i \; (i = 1, \ldots, m)$$

$$\sum_{i=1}^{m} y_i A_i + Z = C$$

$$XZ = \mu I.$$

- **Facts**
  - $XZ = \mu I$ iff: $\exists\, Q$ real orthogonal, Λ and Ω diagonal:
    $X = Q\Lambda Q^T,\ Z = Q\Omega Q^T,\ \Lambda\Omega = \mu I.$
  - **Surjections from skew-symmetric to real orth matrices:**
    1. Exponential Function: $Q = \exp(S) = I + S + \frac{1}{2!}S^2$
    2. Cayley Transform: $Q = (I + \frac{1}{2}S)(I - \frac{1}{2}S)^{-1}$

    Linear approximation of both: $I + S$.
- **Basic Ideas of the Q Method**
  - Replace (X, Z) by $(Q, \boldsymbol{\lambda}, \boldsymbol{\omega})$.
  - Replace Q by either of the above mappings and appl

- **The Newton System**

$$\Delta\Omega + S\Omega - \Omega S + \sum_{i=1}^{m} \Delta y_i B_i = \mathbf{r}_d$$
$$B_i \bullet (\Delta\Lambda + S\Lambda - \Lambda S) = (r_p)_i \quad (i = 1, \ldots$$
$$\Lambda\Delta\Omega + \Omega\Delta\Lambda = \mu I - \Lambda\Omega.$$

Notation: $B_i = A_i Q$.

- **Properties**

1. No eigenvalue factorization;
2. the Schur complement is symmetric positive definite calculated by Cholesky factorization;
3. the Jacobian is nonsingular ⇒ stable and highly acc
4. Storage space is small.

Polar Decomposition for V

- **Notation**

$$G(Sq) \stackrel{\text{def}}{=} \{g \in GL(V) \colon g(Sq) = Sq\}.$$

G: connected component of the identity in $G(Sq)$.

$$K \stackrel{\text{def}}{=} G \cap O(V), \quad K = Aut_0(V).$$

- **Polar Decomposition**

For any Jordan frames $\mathbf{f}_1, \ldots, \mathbf{f}_r$ and $\mathbf{d}_1, \ldots, \mathbf{d}_r$, $\exists\, A \in K$

$$A\mathbf{f}_i = \mathbf{d}_i \quad (1 \leq i \leq r).$$

- $A \in Aut(V) \Rightarrow A\mathbf{f}_i$ is a Jordan frame.

- Fix a Jordan frame $\mathbf{f}_1, \ldots, \mathbf{f}_r$, $\forall\, \mathbf{x} \in V$ can be decompose

$$\mathbf{x} = Q \sum_{i=1}^{r} \lambda_i \mathbf{f}_i \ (Q \in K).$$

Optimal Conditions for Symmetric C Programs

- Assume Slater conditions for Primal and Dual ⇒ (1) sol $\langle \mathbf{x}, \mathbf{z} \rangle = 0$.
- Assume $\mathbf{x}$, $\mathbf{z}$ in Sq. Then $\langle \mathbf{x}, \mathbf{z} \rangle = 0 \Leftrightarrow \mathbf{xz} = \mathbf{0} \Leftrightarrow$, $\mathbf{x}$ and $\mathbf{z}$ share a same Jordan fr $\mathbf{x} = \sum_{i=1}^{r} \lambda_i \mathbf{d}_i,\ \mathbf{z} = \sum_{i=1}^{r} \omega_i \mathbf{d}_i,\ \lambda_i \omega_i = 0.$
- Fix a Jordan frame $\mathbf{f}_1, \ldots, \mathbf{f}_r$. At optimum, replace $\mathbf{x}$ an decomposition (2):

$$
\begin{aligned}
AQ\left(\sum_{i=1}^{r} \lambda_i \mathbf{f}_i\right) &= \mathbf{b} \\
A^* \mathbf{y} + Q\left(\sum_{i=1}^{r} \omega_i \mathbf{f}_i\right) &= \mathbf{c} \\
\lambda_i \omega_i &= 0 \\
\lambda_i \geq 0 \quad \omega_i &\geq 0.
\end{aligned}
$$

An Interior Point Algorithm

- Perturbed system on the central path:

$$
\begin{aligned}
A\mathbf{x} &= \mathbf{b} \\
A^*\mathbf{y} + \mathbf{z} &= \mathbf{c} \\
\langle \mathbf{x}, \mathbf{z} \rangle &= \mu.
\end{aligned}
$$

- $\mathbf{x} \in \operatorname{int} Sq,\ \mu \in \mathbb{R}\, \mathbf{xz} = \mu\mathbf{e} \Rightarrow$
  $\mathbf{x}$ and $\mathbf{z}$ share a same Jordan frame.
  Hence $\exists\, Q \in K:\ \mathbf{x} = Q\sum_{i=1}^{r} \mathbf{f}_i,\ \mathbf{z} = Q\sum_{i=1}^{r} \mathbf{f}_i,\ \lambda_i\omega_i = \mu$
- Replace $\mathbf{x}$ and $\mathbf{z}$ by their decomposition (2).

- **Stepsize Update**

$$\mathbf{x} = Q\sum_{i=1}^{r}\lambda_i\mathbf{f}_i \in Sq,\ \lambda_i + \Delta\,\lambda_i \geq 0,\ \Delta\,Q \in K$$

$$\Rightarrow Q\,\Delta\,Q\sum_{i=1}^{r}(\lambda_i + \Delta$$

- **Linearly Update Q**
  - **Notations**

    Lie algebra of K: $\mathfrak{k} = \mathfrak{m} \oplus \mathfrak{l}$.

    $W \stackrel{\text{def}}{=} \left\{\mathbf{a} = \sum_{i=1}^{r}\lambda_i\mathbf{f}_i : \lambda_i \in \mathbb{R}\right\}$, $\mathfrak{m} \stackrel{\text{def}}{=} \{S \in \mathfrak{k} \colon \forall\, a \in$
  - **Exponential Map**

    $\exp : \mathfrak{k} \mapsto K$.

    V is power associative $\Rightarrow \exp\xi = \sum_{i=0}^{\infty}\frac{\xi^i}{i!}$.
  - $\Delta\,Q = \sum_{i=0}^{\infty}\frac{S^i}{i!}$ (for $S \in \mathfrak{k}$) $\Rightarrow$ linearization of $\Delta\,Q$ is

- **The Newton System**

$$B^k S \sum_{i=1}^{r} \lambda_i^k \mathbf{f}_i^k + B^k \sum_{i=1}^{r} \Delta \lambda_i \mathbf{f}_i = \mathbf{r}_p^k$$

$$(B^k)^* \Delta \mathbf{y} + S \sum_{i=1}^{r} \omega_i^k \mathbf{f}_i + \sum_{i=1}^{r} \Delta \omega_i \mathbf{f}_i = \mathbf{r}_d^k$$

$$\Lambda^k \Delta \boldsymbol{\omega} + \Omega^k \Delta \boldsymbol{\lambda} = \mathbf{r}_c^k.$$

Notations

- $B^k = AQ^k$,
- $\mathbf{r}_p^k = \mathbf{b} - A\mathbf{x}^k$,
- $\mathbf{r}_d^k = (Q^k)^* \left(\mathbf{c} - \mathbf{z}^k - A^* \mathbf{y}^k\right)$,
- $(r_c)_i^k = \mu^k - \lambda_i^k \omega_i^k$.

Properties

- $\lambda_i \neq \lambda_j \Rightarrow$ unique solution for each iteration.
- Further assume strict complementarity and primal-d nondegeneracy $\Rightarrow$ unique solution at optimum.

The Algorithm

- Based on an infeasible interior point method {Kojima, Mizuno: *Math. Programming 61 (1993)* }
- Stop when $\left[\|\gamma_p\| < \epsilon_p,\ \|\gamma_d\| < \epsilon_d,\ \boldsymbol{\lambda}^T\boldsymbol{\omega} < \epsilon_c\right]$ or $\left[\|(\boldsymbol{\lambda},\boldsymbol{\omega}\right.$
- Restrict each iterate in the set

$$\mathcal{N}(\gamma_c,\gamma_p,\gamma_d) \stackrel{\text{def}}{=} \{(\boldsymbol{\lambda},\boldsymbol{\omega},\mathbf{y},Q)\colon \boldsymbol{\lambda} > \mathbf{0},\ \boldsymbol{\omega} > \mathbf{0},\ \lambda_i\omega_i \geq$$

$$\boldsymbol{\lambda}^T\boldsymbol{\omega} \geq \gamma_p\|A\mathbf{x}-\mathbf{b}\| \text{ or } \|A\mathbf{x}-\mathbf{b}\| \leq \epsilon_p,$$

$$\boldsymbol{\lambda}^T\boldsymbol{\omega} \geq \gamma_d\|A^*\mathbf{y}+\mathbf{z}-\mathbf{c}\| \text{ or } \|A^*\mathbf{y}+\mathbf{z}-$$

- Update each iterate:

$$\boldsymbol{\lambda}+\alpha\,\Delta\,\boldsymbol{\lambda} \longrightarrow \boldsymbol{\lambda} \qquad \boldsymbol{\omega}+\beta\,\Delta\,\boldsymbol{\omega} \longrightarrow \boldsymbol{\omega}$$

$$\mathbf{y}+\theta\,\Delta\,\mathbf{y} \longrightarrow \mathbf{y} \qquad Q\exp(\gamma S) \longrightarrow Q$$

- Reduce μ at each iteration.

- **Convergence Results.**
  If $\exists\, g > 0$ such that $\forall\, N > 0,\ \exists\, n \geq N$ so that $\forall\, \|\mathbf{w}\| = 1,\ \|F^n \mathbf{w}\| \geq g$, where F^n is the linear mapping the Newton system at the *nth* iteration. Then the algor after finite steps.

- **Boundedness of Iterates.**
  - Based on {FREUND, JARRE, MIZUNO: *Math. Oper.* *(1999)*}.
  - Assume an interior feasible solution exists such that $\omega_i \geq \delta_d$.
  - Restrict each iterate in the set

    $$\tilde{\mathcal{N}} = \mathcal{N} \cap \{\|A\mathbf{x} - \mathbf{b}\| \leq \zeta_p,\ \|A^*\mathbf{y} + \mathbf{z} - \mathbf{c}\| \leq$$

    where $\zeta_p = \frac{1}{2}\frac{\delta_p}{\|A^+\|}, \quad \zeta_d = \frac{1}{2}\delta_d$.
  - Then the iterates are bounded.

A Newton-type Method

- Replace the complementarity conditions by some comple
  function φ. Choice of φ:

$$\min(\lambda_i, \omega_i), \quad \sqrt{\lambda_i^2 + \omega_i^2} - \lambda_i - \omega_i.$$

- **Optimum Conditions**

$$AQ \sum_{i=1}^{r} \lambda_i \mathbf{f}_i = \mathbf{b}$$

$$A^* \mathbf{y} + Q \sum_{i=1}^{r} \omega_i \mathbf{f}_i = \mathbf{c}$$

$$\varphi(\lambda_i, \omega_i) = 0 \ (i = 1, \ldots, r).$$

- **Newton Direction**

$$BS\sum_{i=1}^{n}\lambda_i\mathbf{f}_i + B\sum_{i=1}^{n}\Delta\,\lambda_i\mathbf{f}_i = \mathbf{r}_p$$

$$B^*\,\Delta\,\mathbf{y} + S\sum_{i=1}^{n}\omega_i\mathbf{f}_i + \sum_{i=1}^{n}\Delta\,\omega_i\mathbf{f}_i = \mathbf{r}_d$$

$$p_i\,\Delta\,\lambda_i + q_i\,\Delta\,\omega_i = \mathbf{r}_c.$$

- Split the index set at optimum into

$$L_\lambda = \{i\colon \lambda_i = 0,\, \omega_i \neq 0\},\quad L_\omega = \{i\colon \lambda_i \neq 0,\, \omega_i$$

Assume $\Delta\,\varphi$ satisfy

$$\begin{cases} p_i \neq 0 \\ q_i = 0 \end{cases} \quad (i \in L_\lambda), \qquad \begin{cases} p_i = 0 \\ q_i \neq 0 \end{cases} \quad (i \in L_\omega)$$

Then the mapping is one-to-one $\Rightarrow$ Pure Newton's meth
Q-quadratic convergence rate.

Numerical Examples (SOCP)

The Steiner Minimal Tree Problem

{Xue & Ye: *SIAM J. Optim., 7 (1997)* }

- **Xue & Ye:** 23 iterations, feasible initial point, gap redu
- **Interior Point Method:**
  Starting point: $\mathbf{x}_i = (2; 1; 0)$; $\mathbf{z}_i = (2; -1; 0)$, $\mathbf{y} = \mathbf{0}$.
  34 iterations, total network cost at $28th$ is better. Accu
  $\|\mathbf{r}_p\|_2$, $\|\mathbf{r}_d\|_2$, $\|\mathbf{r}_c\|_1 \leq 5.0e - 12$.
  Modified Q Method: 29 iterations, total network cost at
  better.
- **Pure Newton's Method:**
  Perturb each coordinate by a scalar in $(-0.1, 0.1)$: 4 ite
  Set point 9 from $(2.328223, 9.139549)$ to $(2.5, 9.0)$: 2 ite