Semidefinite programming Interior-point methods: the Newton system, symmetrization

Imre Pólik, PhD

Lehigh University
Department of Industrial and Systems Engineering

February 3, 2009

Outline

The central path

- Optimality conditions

$$
\begin{array}{rcrcl}
 & \mathcal{A}X & & = & b \\
\mathcal{A}^*y & & +S & = & C \\
 & XS & & = & 0 \\
 & X,\, S & & \succeq & 0,
\end{array}
$$

- Perturbation:

$$
\begin{array}{rcrcl}
 & \mathcal{A}X & & = & b \\
\mathcal{A}^*y & & +S & = & C \\
 & XS & & = & \mu I \\
 & X,\, S & & \succ & 0,
\end{array}
$$

Exists if we have primal-dual strict feasibility, or equivalently, an interior point.

The Newton-step

- The Newton-system

$$\begin{array}{rrrcl} & \mathcal{A}\Delta X & & = & 0 \\ \mathcal{A}^*\Delta y & & +\Delta S & = & 0 \\ & \Delta X S & +X\Delta S & = & \mu I - XS \end{array}$$

- The Newton-step

$$\Delta S = -\mathcal{A}^*\Delta y = -\sum_{i=1}^{m} A_i \Delta y_i$$

$$\Delta X = \mu S^{-1} - X - X\Delta S S^{-1} \quad \text{not symmetric!}$$

- Force symmetry: no solution, overdetermined system
- What went wrong? (XS is not necessarily symmetric)

The AHO method

- Replace $XS = \mu I$ with

$$\frac{1}{2}(XS + SX) = \mu I,$$

- The step is defined by:

$$\begin{aligned} \mathcal{A}\Delta X &= r_p \\ \mathcal{A}^*\Delta y + \Delta S &= R_d \\ \frac{\Delta XS + S\Delta X}{2} + \frac{X\Delta S + \Delta SX}{2} &= R_M \end{aligned}$$

- Not well-defined in general

The Monteiro–Zhang family

- Replace $XS = \mu I$ with

$$\frac{1}{2}\left(PXSP^{-1} + P^{-T}SXP^{T}\right) = \mu I,$$

- Alternatively, with $M = P^T P$:

$$\frac{1}{2}\left(MXS + SXM\right) = \mu M,$$

- P may change in each iteration.
- The step is defined by:

$$\begin{aligned} \mathcal{A}\Delta X &= r_p \\ \mathcal{A}^*\Delta y + \Delta S &= R_d \\ \frac{M\Delta XS + S\Delta XM}{2} + \frac{MX\Delta S + \Delta SXM}{2} &= R_M \end{aligned}$$

Some important members

- AHO Alizadeh–Haeberly–Overton
  - $P = M = I, \qquad \frac{1}{2}(XS + SX) = \mu I$
  - Each MZ direction is a scaled AHO direction
  - $X' = PXP^T$
- NT Nesterov–Todd
  - $P = W^{-1/2}$, $M = W^{-1}$
  - $WSW = X$, scaling matrix
- HKM
  - Helmberg–Rendl–Haeberly–Vanderbei–Wolkowicz
  - Kojima–Shindoh–Hara
  - Monteiro
  - $P = S^{1/2}$, $M = S$
- dual HKM
  - $P = X^{-1/2}$, $M = X^{-1}$

Properties

- Extends LP Diagonal X, S: $S\Delta X + X\Delta S = \mu I - XS$
- Predicts duality gap $S \bullet \Delta X + X \bullet \Delta S = \mu n - X \bullet S$
- Scale-invariance Take $X' = PXP^T$
- Well defined
- Primal-dual symmetry

Property	**AHO**	**HKM**	**dual-HKM**	**NT**
Extends LP	✓	✓	✓	✓
Predicts gap	✓	✓	✓	✓
Well-defined		✓	✓	✓
Scale invariant	Q	✓	✓	✓
PD-symmetric	✓			✓

Q: invariant under orthogonal transformation

See Mike Todd's paper on all the symmetrization schemes.

Computing the directions I.

- General Newton-system

$$
\begin{array}{rrrcl}
 & \mathcal{A}\Delta X & & = & r_p \\
\mathcal{A}^*\Delta y & & +\Delta S & = & R_d \\
 & \mathcal{E}\Delta X & +\mathcal{F}\Delta S & = & R_M
\end{array}
$$

- Solution

$$
\begin{aligned}
\left(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*\right)\Delta y &= \mathcal{A}\mathcal{E}^{-1}\left(R_{EF} - \mathcal{F}R_d\right) \\
\Delta S &= R_d - \mathcal{A}^*\Delta y \\
\Delta X &= \mathcal{E}^{-1}\left(R_{EF} - \mathcal{F}\Delta S\right)
\end{aligned}
$$

- How to find $\mathcal{E}^{-1}$?
- How much does it cost?

Computing the directions II.

- What is $\mathcal{E}^{-1}$ for MZ?

$$\begin{aligned} \mathcal{E} : \mathbb{R}^{n\times n} &\rightarrow \mathbb{R}^{n\times n} \\ U &\mapsto \frac{1}{2}(SUM + MUS) \end{aligned}$$

- The resulting equation:

$$MUS + SUM = 2R$$

1. Compute $S^{\pm 1/2} \Rightarrow \qquad M'U' + U'M' = 2R'$
2. Write M' as $Q'D'Q'^T \Rightarrow \qquad \bar{D}\bar{U} + \bar{U}\bar{D} = 2\bar{R}$
3. $\bar{U}_{ij} = \frac{2\bar{R}_{ij}}{\bar{d}_i + \bar{d}_j}$

- $\Rightarrow$ cheap! $\mathcal{O}(n^3)$
- How to form $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$? How much does it cost?

Computing the directions III.

- Forming $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$

$$\mathbb{R}^m \xrightarrow{\mathcal{A}^*} \mathbb{R}^{n^2} \xrightarrow{\mathcal{F}} \mathbb{R}^{n^2} \xrightarrow{\mathcal{E}^{-1}} \mathbb{R}^{n^2} \xrightarrow{\mathcal{A}} \mathbb{R}^m$$

$$\left(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*\right)_{i,j} = A_i \bullet \mathcal{E}^{-1}\mathcal{F}A_j,\ i,j = 1,\ldots,m$$

  - $\mathcal{E}^{-1}\mathcal{F}A_j$: (Cholesky, product) $\mathcal{O}(n^3)$, m-times
  - $A_i \bullet \mathcal{E}^{-1}\mathcal{F}A_j$ (dot product) $\mathcal{O}(n^2)$, m^2-times
- Overall cost is $\mathcal{O}((m+n)mn^2) \Rightarrow$ expensive
- Total operation count:

$$\mathcal{O}\left((m+n)mn^{5/2}\log\left(\frac{1}{\varepsilon}\right)\right)$$

- Memory usage: $\mathcal{O}\left(mn^2 + m^2\right)$
- $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is dense