A symmetric extension of the NT direction

Florian Jarre
Mathematisches Institut
Universität Düsseldorf
–
based on joint work with C. Hergenroeder

November 19, 2012

Outline

- Why do iterative approaches for general purpose interior point methods pretty much fail so far?

Outline

- Why do iterative approaches for general purpose interior point methods pretty much fail so far?
- Can we overcome this?

Outline

- Why do iterative approaches for general purpose interior point methods pretty much fail so far?
- Can we overcome this?
- **I)** LP

Outline

- Why do iterative approaches for general purpose interior point methods pretty much fail so far?
- Can we overcome this?
- **I)** LP
- **II)** SDP and the NT direction

Iterative approaches

What do we need?

1. Avoid systematic cancellation errors in the transformations of the right hand side.
2. Generate linear systems with *bounded condition numbers*.
3. Preserve the *sparsity structure* of the constraints.
4. Generate *symmetric* systems.
5. Generate *positive definite* systems.

Apology

Analysis of the linear systems will be a bit technical

To get this implemented you just need to be stubborn and push it through

(I try to concentrate on the main points)

I) Linear Programs

$$\text{minimize}_{x \in IR^n} \; c^T x \qquad \text{subject to } \; Ax = b \; \text{ and } \; x \geq 0,$$

Interior point system:

$$\begin{aligned} Ax &= b, \\ A^T y + s &= c, \\ Xs &= \mu e, \end{aligned}$$

$x > 0$, $s > 0$ and $X := \text{Diag}(x_1, x_2, \ldots, x_n)$

Search direction

Let

$$p := b - Ax, \quad q := c - A^T y - s, \quad \text{and} \quad r := \mu^+ e - Xs.$$

"Standard System", **(Std.S)**:

$$\begin{bmatrix} 0 & A & 0 \\ A^T & 0 & I \\ 0 & S & X \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta s \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix}.$$

Reduction to 2 by 2

Eliminating with any of the red diagonal matrices

$$\begin{bmatrix} 0 & A & 0 \\ A^T & 0 & I \\ 0 & S & X \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta s \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix}.$$

and diagonal scaling leads to

$$\begin{bmatrix} 0 & A \\ A^T & -SX^{-1} \end{bmatrix} \begin{bmatrix} \Delta y \\ u \end{bmatrix} = \begin{bmatrix} rhs1 \\ rhs2 \end{bmatrix},$$

with $u = \Delta x$ or $u = -XS^{-1}\Delta s$ and

$$\begin{bmatrix} rhs1 \\ rhs2 \end{bmatrix} = \begin{bmatrix} p \\ q - X^{-1} r \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} rhs1 \\ rhs2 \end{bmatrix} = \begin{bmatrix} p - AS^{-1} r \\ q \end{bmatrix}$$

Systematic cancellation

The small entries of X or the small entries of S in

$$\begin{bmatrix} rhs1 \\ rhs2 \end{bmatrix} = \begin{bmatrix} p \\ q - X^{-1} r \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} rhs1 \\ rhs2 \end{bmatrix} = \begin{bmatrix} p - AS^{-1} r \\ q \end{bmatrix}$$

lead to systematic cancellation error of entries in q or in p.

Pivoting:
Partition x and s into two parts, x_1, x_2 and s_1, s_2 such that

$$x_1 \geq s_1 \quad \text{and} \quad x_2 < s_2.$$

(x_1 and x_2 each have many components.)

Stable reduction (Freund, J', 1993)

Defining

$$\tilde{q} := \begin{bmatrix} q_1 - X_1^{-1} r_1 \\ q_2 \end{bmatrix} \text{ and } \tilde{p} := p - A_2 S_2^{-1} r_2,$$

the "Stable Reduction" **(Stb.R)** is given by

$$\begin{bmatrix} 0 & A \\ A^T & -SX^{-1} \end{bmatrix} \begin{bmatrix} \Delta y \\ u \end{bmatrix} = \begin{bmatrix} \tilde{p} \\ \tilde{q} \end{bmatrix}$$

where

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad u_1 := \Delta x_1, \text{ and } u_2 := -X_2 S_2^{-1} \Delta s_2.$$

Normal equations

Eliminating again with the diagonal matrix SX^{-1}, any of the 2-by-2 systems leads to the same set of normal equations, **(Nrm.E)**

$$AXS^{-1}A^T\Delta y = p + AS^{-1}(Xq - r).$$

Small entries of X cancel components in q and small components of S cancel p.

Most commonly used with direct solvers.

Comparison

Condition numbers for small nondegenerate LPs (100 x 250)
100 random problems, A unitary, x, s nearly central:

μ	cond. **(Std.S)**	cond. **(Stb.R)**	cond. **(Nrm.E)**
1	210 (320/120)	1300 (2200/610)	10 (15/8)
10^{-4}	1.5e4 (3.5e5/2400)	3.8e8 (1.8e10/5.1e7)	6.1e5 (9.2e7/1.5e6)
10^{-8}	1.4e4 (1.1e6/3400)	2.7e12 (2.1e14/5.8e11)	5.0e5 (3.0e9/4.7e4)

(This does not reflect systematic cancellation – or symmetry)

Consequence: None of the systems seems suitable
(and **(Stb.R)** seems worst)

Generally, no/few general purpose interior point methods using iterative solvers.

Extending instead of reducing?

Introduce a "dummy variable" $\Delta z := -X^{-1/2}\Delta x$.

Then,

$$\begin{bmatrix} 0 & A & 0 & 0 \\ A^T & 0 & I & 0 \\ 0 & I & 0 & X^{1/2} \\ 0 & 0 & X^{1/2} & -S \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta s \\ \Delta z \end{bmatrix} = \begin{bmatrix} p \\ q \\ 0 \\ X^{-1/2} r \end{bmatrix}.$$

(Still some cancellation with small entries of X)

Diagonal pivoting:

Pivoting as for **(Stb.R)** leads to a first symmetric extension

$$\begin{bmatrix} 0 & A_1 & A_2 & 0 & 0 & 0 & 0 \\ A_1^T & 0 & 0 & I & 0 & 0 & 0 \\ A_2^T & 0 & 0 & 0 & I & 0 & 0 \\ 0 & I & 0 & 0 & 0 & X_1^{1/2} & 0 \\ 0 & 0 & I & 0 & 0 & 0 & X_2^{1/2} \\ 0 & 0 & 0 & X_1^{1/2} & 0 & -S_1 & 0 \\ 0 & 0 & 0 & 0 & X_2^{1/2} & 0 & -S_2 \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x_1 \\ \Delta x_2 \\ \Delta s_1 \\ \Delta s_2 \\ \Delta z_1 \\ \Delta z_2 \end{bmatrix} = \begin{bmatrix} p \\ q_1 \\ q_2 \\ 0 \\ S_2^{-1} r_2 \\ X_1^{-1/2} r_1 \\ 0 \end{bmatrix}.$$

In case anybody likes to see:

Δz has changed its meaning but $\Delta x, \Delta y, \Delta s$ remain the same: Solving the last two block rows for Δz we obtain

$$\Delta z = S^{-1} X^{1/2} \Delta s - \begin{bmatrix} S_1^{-1} X_1^{-1/2} r_1 \\ 0 \end{bmatrix}$$

and inserting this in the third and fourth block row yields

$$\Delta x + X^{1/2} \left(S^{-1} X^{1/2} \Delta s - \begin{bmatrix} S_1^{-1} X_1^{-1/2} r_1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ S_2^{-1} r_2 \end{bmatrix}.$$

Multiplying this from left with S yields $S \Delta x + X \Delta s = r$.

Symmetric diagonal scaling (last block row/column)

with $\mathrm{Diag}(d)$ where $d_1 = x_1^{-1/2}$ and $d_2 = s_2^{-1/2}$ – leads to

$$\begin{bmatrix} 0 & A_1 & A_2 & 0 & 0 & 0 & 0 \\ A_1^T & 0 & 0 & I & 0 & 0 & 0 \\ A_2^T & 0 & 0 & 0 & I & 0 & 0 \\ 0 & I & 0 & 0 & 0 & I & 0 \\ 0 & 0 & I & 0 & 0 & 0 & S_2^{-1/2}X_2^{1/2} \\ 0 & 0 & 0 & I & 0 & -X_1^{-1}S_1 & 0 \\ 0 & 0 & 0 & 0 & S_2^{-1/2}X_2^{1/2} & 0 & -I \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x_1 \\ \Delta x_2 \\ \Delta s_1 \\ \Delta s_2 \\ \Delta z_1 \\ \Delta z_2 \end{bmatrix} = \begin{bmatrix} p \\ q_1 \\ q_2 \\ 0 \\ S_2^{-1} r_2 \\ X_1^{-1} r_1 \\ 0 \end{bmatrix}$$

the symmetric extension **(Sm.Ex)**
(The meaning of Δz has changed again).

Reducing this again

Let

$$D_x := \begin{bmatrix} I & 0 \\ 0 & S_2^{-1/2} X_2^{1/2} \end{bmatrix}, \qquad D_s := \begin{bmatrix} X_1^{-1} S_1 & 0 \\ 0 & I \end{bmatrix}.$$

The entries of D_x, D_s are nonnegative and at most 1.

Eliminating Δs with the second and third block row and Δx with the fourth and fifth block row leads to

(Rd.Ex)

$$\begin{bmatrix} 0 & AD_x \\ D_x A^T & D_s \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} A \begin{bmatrix} 0 \\ S_2^{-1} r_2 \end{bmatrix} - p \\ D_x q - \begin{bmatrix} X_1^{-1} r_1 \\ 0 \end{bmatrix} \end{bmatrix}.$$

Given the (approximate) solution, the components Δx and Δs can be recovered in a stable fashion via

$$\Delta x = \begin{bmatrix} 0 \\ S_2^{-1} r_2 \end{bmatrix} - D_x \Delta z, \qquad \Delta s = q - A^T \Delta y.$$

This system will be denoted by "reduced extension" **(Rd.Ex)**.

Only canclellation error in q_2; recovered in back substitution!

Comparison

μ	cond. **(Sm.Ex)**	cond **(Rd.Ex)**	cond. opt. basis
1	43 (61/32)	18 (26/14)	320 (1.1e4/58)
1.0e-4	1100 (4.6e4/170)	350 (1.4e4/54)	280 (5700/52)
1.0e-8	820 (6.4e4/250)	250 (2.0e4/76)	250 (1.9e4/75)

Even far away from the central path, the condition of **(Rd.Ex)** and of the optimal basis are about the same when μ is small!

(Only about 30% of sQMR iterations compared to **(Sm.Ex)**)

Theorem

Assume that the LP has a unique (primal-dual) optimal solution and let $\bar{\kappa}$ be the 2-norm-condition number of the optimal basis. Assume further that the 2-norm of the optimal basis matrix and of its inverse are both at least 1.
For $\mu > 0$ let x_μ, y_μ, s_μ be any (not necessarily strictly) primal dual feasible solution with $x_\mu^T s_\mu \leq n\mu$. Let $\kappa(\mu)$ be the condition number of system **(Rd.Ex)** evaluated at x_μ, s_μ. Then, $\lim_{\mu \to 0} \kappa(\mu) = \bar{\kappa}$.

Moreover, (with $y_\mu := (AA^T)^{-1}A(c - s_\mu)$), the reduction

$$(\mathbf{Std.S}) \longrightarrow (\mathbf{Rd.Ex})$$

is asymptotically stable for $\mu \to 0$.

Theorem

Assume that the LP has a unique (primal-dual) optimal solution and let $\bar{\kappa}$ be the 2-norm-condition number of the optimal basis. Assume further that the 2-norm of the optimal basis matrix and of its inverse are both at least 1.
For $\mu > 0$ let x_μ, y_μ, s_μ be any (not necessarily strictly) primal dual feasible solution with $x_\mu^T s_\mu \le n\mu$. Let $\kappa(\mu)$ be the condition number of system **(Rd.Ex)** evaluated at x_μ, s_μ. Then, $\lim_{\mu\to 0} \kappa(\mu) = \bar{\kappa}$.

Moreover, (with $y_\mu := (AA^T)^{-1}A(c - s_\mu)$), the reduction

$$(\mathbf{Std.S}) \longrightarrow (\mathbf{Rd.Ex})$$

is asymptotically stable for $\mu \to 0$.

Proof:

Theorem

Assume that the LP has a unique (primal-dual) optimal solution and let $\bar{\kappa}$ be the 2-norm-condition number of the optimal basis. Assume further that the 2-norm of the optimal basis matrix and of its inverse are both at least 1.
For $\mu > 0$ let x_μ, y_μ, s_μ be any (not necessarily strictly) primal dual feasible solution with $x_\mu^T s_\mu \leq n\mu$. Let $\kappa(\mu)$ be the condition number of system **(Rd.Ex)** evaluated at x_μ, s_μ. Then, $\lim_{\mu \to 0} \kappa(\mu) = \bar{\kappa}$.

Moreover, (with $y_\mu := (AA^T)^{-1}A(c - s_\mu)$), the reduction

$$(\textbf{Std.S}) \longrightarrow (\textbf{Rd.Ex})$$

is asymptotically stable for $\mu \to 0$.

Proof: Straightforward. –

Question:

- Could we have gotten there directly
  (much simpler – without using the extension first)?

Question:

- Could we have gotten there directly
  (much simpler – without using the extension first)?
- Scaling the second block row and block column of **(Stb.R)** with D_x leads to the same system matrix.
  D_x is singular in the limit
  $\Longrightarrow$ cancellation in the right hand side.

Question:

- Could we have gotten there directly
  (much simpler – without using the extension first)?
- Scaling the second block row and block column of **(Stb.R)**
  with D_x leads to the same system matrix.
  D_x is singular in the limit
  $\Longrightarrow$ cancellation in the right hand side.
- The original variables Δx and Δs disappeared,
  only the "extension variable" Δz remains in a rescaled form.

Question:

- Could we have gotten there directly
  (much simpler – without using the extension first)?
- Scaling the second block row and block column of **(Stb.R)** with D_x leads to the same system matrix.
  D_x is singular in the limit
  $\Longrightarrow$ cancellation in the right hand side.
- The original variables Δx and Δs disappeared,
  only the "extension variable" Δz remains in a rescaled form.
- The definition of Δz and the right hand side is a bit "involved"
  – how to come up with it without extending first is unclear.

Question:

- Could we have gotten there directly (much simpler – without using the extension first)?
- Scaling the second block row and block column of **(Stb.R)** with D_x leads to the same system matrix.
  D_x is singular in the limit
  $\Longrightarrow$ cancellation in the right hand side.
- The original variables Δx and Δs disappeared, only the "extension variable" Δz remains in a rescaled form.
- The definition of Δz and the right hand side is a bit "involved"
  ~~– how to come up with it without extending first is unclear.~~

II) SDP

Notation:

$$\text{minimize}_{X \in \mathcal{S}^n} \; C \bullet X \qquad \text{subject to } \; \mathcal{A}(X) = b \; \text{ and } \; X \succeq 0,$$

where $\mathcal{S}^n$ is the space of symmetric $n \times n$-matrices,
$b \in IR^m$, $C \in \mathcal{S}^n$, $C \bullet X := \text{trace}(CX)$, and

$$\mathcal{A}(X) = \begin{bmatrix} A^{(1)} \bullet X \\ \vdots \\ A^{(m)} \bullet X \end{bmatrix}, \qquad \mathcal{A}^*(y) = \sum_{i=1}^{m} y_i A^{(i)}.$$

Central Path

Let $\mathbf{S}_P$ denote the symmetrization operator introduced by Monteiro and Zhang,

$$\mathbf{S}_P(U) := \frac{1}{2}(PUP^{-1} + (PUP^{-1})^T)$$

with some nonsingular matrix P.
The central path is defined by $X, S \succ 0$, $y \in IR^m$ with

$$\begin{aligned} \mathcal{A}(X) &= b, \\ \mathcal{A}^*(y) + S &= C, \\ \mathbf{S}_P(XS) &= \mu I. \end{aligned}$$

(Std.S)

Let $X, S \succ 0$, $y \in I\!R^m$ be given, and set

$$p := b - \mathcal{A}(X), \;\; Q := C - \mathcal{A}^*(y) - S, \;\; \text{and} \;\; R := \mu^+ I - \mathbf{S}_P(XS)$$

with some "target value" $\mu^+ (\leq X \bullet S/n)$. Newton's method yields

$$\begin{bmatrix} 0 & \mathcal{A} & 0 \\ \mathcal{A}^* & 0 & I \\ 0 & \mathbf{S}_P(\;.\;S) & \mathbf{S}_P(X\;.\;) \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta X \\ \Delta S \end{bmatrix} = \begin{bmatrix} p \\ Q \\ R \end{bmatrix}.$$

NT-direction

Let

$$W := S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2}$$

be the metric geometric mean of X and S^{-1}. Then,

$$WSW = X, \qquad S = W^{-1}XW^{-1},$$

$$W^{1/2}SW^{1/2} = W^{-1/2}XW^{-1/2} =: E \quad (= E(X,S)),$$

and when $XS \approx \mu I$, we have

$$E \approx \sqrt{\mu} I.$$

Set $P := W^{-1/2}$.

“Scaling” operator $\mathcal{D}_P : \mathcal{S}^n \to \mathcal{S}^n$ and
“Lyapunov operator” $\mathcal{L}_E : \mathcal{S}^n \to \mathcal{S}^n$ defined via

$$\mathcal{D}_P(\Delta X) := P\Delta X P \qquad \text{and} \qquad \mathcal{L}_E(\Delta X) := \frac{1}{2}(E\Delta X + \Delta X E).$$

Then,

$$\mathbf{S}_P(\Delta X S) = \frac{1}{2}(P\Delta X S P^{-1} + (P\Delta X S P^{-1})^T)$$

$$= \frac{1}{2}(P\Delta X P \underbrace{P^{-1} S P^{-1}}_{E} + (P\Delta X P \underbrace{P^{-1} S P^{-1}}_{E})^T) = \mathcal{L}_E \mathcal{D}_P(\Delta X)$$

and, likewise,

$$\mathbf{S}_P(X\Delta S) = \mathcal{L}_E \mathcal{D}_{P^{-1}}(\Delta S).$$

Same system:

$$\begin{bmatrix} 0 & \mathcal{A} & 0 \\ \mathcal{A}^* & 0 & I \\ 0 & \mathcal{L}_E\mathcal{D}_P & \mathcal{L}_E\mathcal{D}_{P^{-1}} \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta X \\ \Delta S \end{bmatrix} = \begin{bmatrix} p \\ Q \\ R \end{bmatrix}.$$

or

$$\begin{bmatrix} 0 & \mathcal{A} & 0 \\ \mathcal{A}^* & 0 & I \\ 0 & \sqrt{\mu}\mathcal{D}_P & \sqrt{\mu}\mathcal{D}_{P^{-1}} \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta X \\ \Delta S \end{bmatrix} = \begin{bmatrix} p \\ Q \\ \sqrt{\mu}(\mathcal{L}_E)^{-1}R \end{bmatrix}.$$

P and P^{-1} have same eigenbasis!
(Unique to NT)

Unitary rescaling

Let $P = U\Lambda U^T$ be the eigenvalue decomposition of P, and $\mathcal{U} : \mathcal{S}^n \to \mathcal{S}^n$ be defined by $\mathcal{U}(Z) = UZU^T$ with adjoint $\mathcal{U}^*(Y) = \mathcal{U}^{-1}(Y) = U^T YU$.

$$\begin{bmatrix} 0 & \mathcal{A}\mathcal{U} & 0 \\ \mathcal{U}^*\mathcal{A}^* & 0 & I \\ 0 & \sqrt{\mu}\mathcal{D}_\Lambda & \sqrt{\mu}\mathcal{D}_{\Lambda^{-1}} \end{bmatrix} \begin{bmatrix} \Delta y \\ \mathcal{U}^*\Delta X \\ \mathcal{U}^*\Delta S \end{bmatrix} = \begin{bmatrix} p \\ \mathcal{U}^* Q \\ \mathcal{U}^*\sqrt{\mu}(\mathcal{L}_E)^{-1} R \end{bmatrix},$$

Here, $\mathcal{D}_\Lambda$, $\mathcal{D}_{\Lambda^{-1}}$ are "true" diagonal scalings.
$\mathcal{D}_\Lambda(M) = \Lambda M \Lambda = M \circ (\text{diag}(\Lambda)e^T + e(\text{diag}(\Lambda))^T)$

(Hadamard product, $e = (1, 1, \ldots, 1)^T$.)

Note

Iterative schemes only require multiplication with this matrix, not its representation.
In particular, $\mathcal{A}\mathcal{U}$ and $\mathcal{U}^*\mathcal{A}*$ are not formed explicitly.

Let $\widehat{\Delta X} := \mathcal{U}^*\Delta X$, $\widehat{\Delta S} := \mathcal{U}^*\Delta S$, and $\widehat{\Delta Z}$ be a dummy variable,
set $\tilde{R} := \sqrt{\mu}(\mathcal{L}_E)^{-1}R$
then, (up to diagonal scaling) the symmetric extension is given by

Symmetric Extension

$$\begin{bmatrix} 0 & \mathcal{A}\mathcal{U} & 0 & 0 \\ \mathcal{U}^*\mathcal{A}^* & 0 & I & 0 \\ 0 & I & 0 & \sqrt[4]{\mu}\mathcal{D}_{\Lambda^{-1/2}} \\ 0 & 0 & \sqrt[4]{\mu}\mathcal{D}_{\Lambda^{-1/2}} & -\sqrt{\mu}\mathcal{D}_{\Lambda} \end{bmatrix} \begin{bmatrix} \Delta y \\ \widehat{\Delta X} \\ \widehat{\Delta S} \\ \widehat{\Delta Z} \end{bmatrix} = \begin{bmatrix} p \\ \mathcal{U}^* Q \\ \mathcal{U}^* R_2 \\ \mathcal{U}^* R_1 \end{bmatrix},$$

where

$$\sqrt{\mu}\mathcal{D}_{\Lambda}\mathcal{U}^* R_2 + \sqrt[4]{\mu}\mathcal{D}_{\Lambda^{-1/2}}\mathcal{U}^* R_1 = \mathcal{U}^* \tilde{R}$$

Changing R_1 and R_2 s.t. to above constraint

changes $\widehat{\Delta Z}$ but not $\Delta y, \widehat{\Delta X}, \widehat{\Delta S}$.

Chioce of ΔZ

- Minimizing $\|R_1\|_F^2 + \|R_2\|_F^2$ subject to above constraint can be solved explicitly (and cheaply),

$$(\mathcal{U}^* R_2)_{i,j} = (\mathcal{U}^* \tilde{R})_{i,j} \frac{\sqrt{\mu}\lambda_i \lambda_j}{\mu \lambda_i^2 \lambda_j^2 + \sqrt{\mu}\lambda_i^{-1}\lambda_j^{-1}},$$

$$(\mathcal{U}^* R_1)_{i,j} = (\mathcal{U}^* \tilde{R})_{i,j} \frac{\sqrt[4]{\mu}\lambda_i^{-1/2}\lambda_j^{-1/2}}{\mu \lambda_i^2 \lambda_j^2 + \sqrt{\mu}\lambda_i^{-1}\lambda_j^{-1}}.$$

– replacing the partition into x_1 and x_2 for LPs,

Chioce of ΔZ

- Minimizing $\|R_1\|_F^2 + \|R_2\|_F^2$ subject to above constraint can be solved explicitly (and cheaply),

$$(\mathcal{U}^* R_2)_{i,j} = (\mathcal{U}^* \tilde{R})_{i,j} \frac{\sqrt{\mu}\lambda_i \lambda_j}{\mu \lambda_i^2 \lambda_j^2 + \sqrt{\mu}\lambda_i^{-1}\lambda_j^{-1}},$$

$$(\mathcal{U}^* R_1)_{i,j} = (\mathcal{U}^* \tilde{R})_{i,j} \frac{\sqrt[4]{\mu}\lambda_i^{-1/2}\lambda_j^{-1/2}}{\mu \lambda_i^2 \lambda_j^2 + \sqrt{\mu}\lambda_i^{-1}\lambda_j^{-1}}.$$

  – replacing the partition into x_1 and x_2 for LPs,
- So, all the same as for LPs?

The Theorem ?

The partition $x_1 > s_1$ and $x_2 < s_2$
was crucial for the proof of the theorem for LPs.

Here, $\sqrt{\mu}\lambda_i^2 \to 0$ for some i,
$\sqrt{\mu}\lambda_j^2 \to \infty$ for some j,
and $\sqrt{\mu}\lambda_i\lambda_j \to \mathit{const}$ for some i, j.

$\Longrightarrow$ Somewhat worse result,
the right hand side may increase by a factor of $1/\sqrt[3]{\mu}$.

(Rd.Ex)

What we really aim for is the reduced extension which takes the form

$$\begin{bmatrix} 0 & L^{-1}\mathcal{A}\mathcal{U}\mathcal{D}_x \\ \mathcal{D}_x\mathcal{U}^*\mathcal{A}^*L^{-T} & \mathcal{D}_s \end{bmatrix} \begin{bmatrix} \widehat{\Delta y} \\ \widehat{\Delta Z} \end{bmatrix} = \begin{bmatrix} L^{-1}\mathcal{A}R_2 - L^{-1}p \\ \mathcal{D}_x\mathcal{U}^*Q - \mathcal{D}_M\mathcal{U}^*R_1 \end{bmatrix},$$

where we assume that a sparse Cholesky factor L of $\mathcal{A}\mathcal{A}^*$ is available. (Not really crucial, but if available we like to use it.)
Given $\widehat{\Delta y}$ get

$$\widehat{\Delta X} = \mathcal{U}^*R_2 - \mathcal{D}_x\widehat{\Delta Z}, \quad \widehat{\Delta S} = \mathcal{U}^*Q - \mathcal{U}^*\mathcal{A}^*L^{-T}\widehat{\Delta y}.$$

Details

Choose R_1 abd R_2 as to minimize the right hand side of this system subject to the constraint

$$\sqrt{\mu}\mathcal{D}_\Lambda\mathcal{U}^* R_2 + \sqrt[4]{\mu}\mathcal{D}_{\Lambda^{-1/2}}\mathcal{U}^* R_1 = \mathcal{U}^* \tilde{R}.$$

Use an adapted predictor corrector interior-point method.
(Project the search direction onto the equations $\mathcal{A}\Delta X = p$ and $\mathcal{A}^*\Delta y + \Delta S = Q$.)

10% accuracy in the computation of the search direction (!)

Numerical Example:

- Not yet competitive with SDPNAL
  (but much faster than SEDUMI or SDPT3 for large problems)

Numerical Example:

- Not yet competitive with SDPNAL
  (but much faster than SEDUMI or SDPT3 for large problems)
- Needs some form of preconditioning

Numerical Example:

- Not yet competitive with SDPNAL
  (but much faster than SEDUMI or SDPT3 for large problems)
- Needs some form of preconditioning
- possibly recycling the previous predictor step in the initial approximation for QMR in the next predictor step.

Numerical Example:

- Not yet competitive with SDPNAL
  (but much faster than SEDUMI or SDPT3 for large problems)
- Needs some form of preconditioning
- possibly recycling the previous predictor step in the initial approximation for QMR in the next predictor step.
- A factor 10 faster than the “unbalanced symmetric extension”.

Small dense SDP 50x50, 230 linear constraints

iteration	corr. qmr-steps	centrality	step length	$-\log_{10}\mu$
1	3	.88	.70	.58
2	10	2.1	.78	1.3
5	103	2.6	.74	3.1
6	121	1.5	.68	3.7
10	495	3.0	.66	6.1
11	485	3.0	.63	6.6
15	1000	3.2	.41	8.5
16	2000	2.6	.59	9.0

The overall number of QMR Iterations was 15477.

Solution generated by the algorithm: $X^{alg}, y^{alg}, S^{alg}$
Exact optimal solution: $X^{opt}, y^{opt}, S^{opt}$.

$$\frac{\|\mathcal{A}(X^{alg}) - b\|_2}{\|b\|_2} = 1,4 \cdot 10^{-9}, \qquad \frac{X^{alg} \bullet S^{alg}}{\|X^{alg}\|_F \|S^{alg}\|_F} = 2.9 \cdot 10^{-8},$$

$$\frac{\|\mathcal{A}^* y^{alg} + S^{alg} - C\|_F}{\|C\|_F} = 5.0 \cdot 10^{-13},$$

$$\frac{\|X^{alg} - X^{opt}\|_F}{\|X^{alg}\|_F} = 0.008, \qquad \frac{\|S^{alg} - S^{opt}\|_F}{\|S^{alg}\|_F} = 1.1 \cdot 10^{-5}.$$

Relative errors significantly larger (by a factor of more than 10^5) than the relative residuals.
$\Longrightarrow$ optimal solution is not well conditioned.

Small dense LP 2500 variables, 230 linear constraints

iteration	corr. qmr-steps	centrality	step length	$-\log_{10}\mu$
1	3	.48	.54	.43
5	9	4,4	.27	2.4
10	110	3.3	.60	4.2
15	125	3.9	.42	5.3
20	287	2.0	.39	6.5
25	699	2.1	.34	7.9
28	1205	1.4	.69	9.0

The overall number of QMR Iterations was 18700.

Larger SDP

X of dimension 2000×2000,
100000 sparse linear constraints
38 Interior-Point-Iterations
4469 QMR iterations
5 hours (older desktop):

relative primal infeasibility: 5.2e-11,
relative dual infeasibility: 1.6e-11,
relative complementarity: 1.7e-08.

(Optimal solution and condition number unknown.)

Outlook

Further details on Optimization Online (tonight?)

Seems to work also for AHO (joint work with T. Davi)

Outlook

Further details on Optimization Online ~~(tonight?)~~

Seems to work also for AHO (joint work with T. Davi)