# Conic quadratic optimization - part 5

**Erling D. Andersen**
MOSEK ApS,
Fruebjergvej 3, Box 16,
2100 Copenhagen,
Denmark.
Email: e.d.andersen@mosek.com
Personal WWW: http://erling.andersen.name
Company WWW: http://www.mosek.com

April 10, 2008

# Introduction

Introduction
Topics
A symmetric primal-dual algorithm
Computational results
Summary
References

# Topics

Topics:

- Learn about the symmetric primal-dual algorithm of Nesterov-Todd [5], [7].
- What goes into a state of the art implementation.
- See some computational results.

Major reference

- Andersen et. al . [2]

# A symmetric primal-dual algorithm

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

# Recap.

The primal problem

$$\begin{array}{lccc} \min & \sum_k (c^k)^T x^k & & \\ \text{st} & \sum_k A^k x^k & = & b, \\ & x^k \in \mathcal{K}^k & & \end{array} \tag{1}$$

and the dual problem

$$\begin{array}{lccc} \max & b^T y & & \\ \text{st} & (A^k)^T y + s^k & = & c^k, \\ & s^k \in (\mathcal{K}^k)^*. & & \end{array} \tag{2}$$

# Simplifying notation

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Let us simplify the notation i.e.

$$
\begin{aligned}
c &:= \begin{bmatrix} c^1 \\ c^2 \\ \vdots \\ c^r \end{bmatrix}, \\
A &:= \begin{bmatrix} A^1 & A^2 & \ldots & A^k \end{bmatrix}, \\
\mathcal{K} &:= \mathcal{K}^1 \times \mathcal{K}^2 \times \cdots \times \mathcal{K}^r, \\
\mathcal{K}^* &:= (\mathcal{K}^1)^* \times (\mathcal{K}^2)^* \times \cdots \times (\mathcal{K}^r)^*.
\end{aligned}
$$

and

$$A \in \mathcal{R}^{m \times n}.$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Moreover let

$$x := \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^r \end{bmatrix} \text{ and } s := \begin{bmatrix} s^1 \\ s^2 \\ \vdots \\ s^r \end{bmatrix}.$$

Observe that

$$\mathcal{K}^* = \{s \in \mathcal{R}^n : \; x^T s \geq 0, \; \forall x \in \mathcal{K}\}.$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

The primal problem

$$\begin{array}{lccc} \min & c^T x & & \\ \text{st} & Ax & = & b, \\ & x \in \mathcal{K}, & & \end{array} \tag{3}$$

and the dual problem

$$\begin{array}{lccc} \max & b^T y & & \\ \text{st} & A^T y + s & = & c, \\ & s \in \mathcal{K}^*. & & \end{array} \tag{4}$$

# Central theorem

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

**Weak duality:** Let $x$ be a feasible solution to $(P)$ and $(y, s)$ be a feasible solution to $(D)$, then

$$c^T x - b^T y = x^T s \geq 0.$$

**Strong duality:** If $(P)$ is strictly feasible and its optimal objective value is bounded or $(D)$ is strictly feasible and its optimal objective value is bounded, then $(x, y, s)$ is an optimal solution if and only if

$$c^T x - b^T y = x^T s = 0$$

and $x$ is primal feasible and $(y, s)$ is dual feasible.

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

**Primal infeasibility:** If

$$\exists (y, s): \; s \in K_*, \; A^T y + s = 0, \; b^T y > 0, \tag{5}$$

then $(P)$ is infeasible.

**Dual infeasibility:** If

$$\exists x: \; x \in K, \; Ax = 0, \; c^T x < 0, \tag{6}$$

then $(D)$ is infeasible.

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

# Goals

Algorithm should:

- Feasible case: Find an optimal solution.
- Infeasible case: Find an infeasibility certificate.

# A homogeneous model

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Generalized Goldman-Tucker homogeneous model:

$$
(H) \qquad \begin{array}{rcl}
Ax - b\tau & = & 0, \\
A^T y + s - c\tau & = & 0, \\
-c^T x + b^T y - \kappa & = & 0, \\
(x;\tau) \in \bar{\mathcal{K}}, (s;\kappa) \in \bar{\mathcal{K}}_* & &
\end{array}
$$

where

$$
\bar{\mathcal{K}} := \mathcal{K} \times \mathcal{R}_+ \quad \text{and} \quad \bar{\mathcal{K}}_* := \mathcal{K}_* \times \mathcal{R}_+.
$$

- The homogeneous model always have a solution
- Partial list of references:
  - Linear case: [4], [3], [8].
  - Nonlinear case: [6].

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

**Lemma 1**

*Let $(x^*, \tau^*, y^*, s^*, \kappa^*)$ be any feasible solution to (H), then*

**i)**

$$(x^*)^T s^* + \tau^* \kappa^* = 0.$$

**ii)** *If $\tau^* > 0$, then*

$$(x^*, y^*, s^*)/\tau^*$$

*is an optimal solution.*

**iii)** *If $\kappa^* > 0$, then at least one of the strict inequalities*

$$b^T y^* > 0 \tag{7}$$

*and*

$$c^T x^* < 0 \tag{8}$$

*holds. If the first inequality holds, then $(P)$ is infeasible. If the second inequality holds, then $(D)$ is infeasible.*

Introduction

A symmetric primal-dual algorithm

Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Summary:

- Compute a nontrivial solution to $(H)$.
- Provides required information in most cases.
- Bad case:

$$\tau^* = \kappa^* = 0.$$

- Bad case cannot occur for linear problems.

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

# Quadratic cones

## Definition 1

$\mathcal{R}_+$:

$$\mathcal{R}_+ := \{x \in \mathcal{R} : \ x \geq 0\}.$$

## Quadratic cone:

$$\mathcal{K}_q := \{x \in \mathcal{R}^n : \ x_1^2 \geq \|x_{2:n}\|^2, \ x_1 \geq 0\}.$$

## Rotated quadratic cone:

$$\mathcal{K}_r := \{x \in \mathcal{R}^n : \ 2x_1x_2 \geq \|x_{3:n}\|^2, \ x_1, x_2 \geq 0\}.$$

Notes:

- Allowed cone types.
- Are homogeneous and self-dual.

# Some definitions

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

**i.)** If $\mathcal{K}^k$ is $\mathcal{R}_+$, then

$$T^k := 1 \text{ and } Q^k = 1.$$

**ii.)** If $\mathcal{K}^k$ is the quadratic cone, then

$$T^k := I_{n^k} \text{ and } Q^k := diag(1, -1, \ldots, -1).$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

**iii.)** If $\mathcal{K}^k$ is the rotated quadratic cone, then

$$T^k := \begin{bmatrix} \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} & 0 & \cdots & 0 \\ \sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{2}} & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

and

$$Q^k := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{bmatrix}.$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

# Facts

- Orthogonality:

$$Q^k Q^k = I \quad \text{and} \quad T^k T^k = I.$$

- Linear and quadratic cone:

$$K^k = \{x^k \in \mathcal{R}^{n^k} : \ (x^k)^T Q^k x^k \geq 0, \ x_1^k \geq 0\}$$

- Rotated quadratic cone:

$$\mathcal{K}^k = \{x^k \in \mathcal{R}^{n^k} : \ (x^k)^T Q^k x^k \geq 0, \ x_1^k, x_2^k \geq 0\}.$$

- Equivalence:

$$x^k \in \mathcal{K}_q \ \Leftrightarrow \ T^k x^k \in \mathcal{K}_r.$$

# Complementarity

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

**Definition:**

$$V := \mathrm{mat}\,(v) = \begin{bmatrix} v_1 & v_{2:n}^T \\ v_{2:n} & v_1 I_{n-1} \end{bmatrix}.$$

Given $x, s \in K$ then

$$x^T s = 0 \Leftrightarrow X^k S^k e^k = S^k X^k e^k = 0, \quad i = 1, \ldots, k,$$

where

$$X^k := \mathrm{mat}\left(T^k x^k\right) \text{ and } S^k := \mathrm{mat}\left(T^k s^k\right).$$

**Definition:**

$$\begin{array}{rcl} X & := & diag(X^1, \ldots, X^k), \\ S & := & diag(S^1, \ldots, S^k). \end{array}$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

# The central path

Let

$$(x^{(0)}, \tau^{(0)}, y^{(0)}, s^{(0)}, \kappa^{(0)})$$

be given such that

$$(x^{(0)}; \tau^{(0)}), (s^{(0)}; \kappa^{(0)}) \in \mathrm{int}(\bar{\mathcal{K}}).$$

Central path definition:

$$\begin{array}{rcl} Ax - b\tau & = & \gamma(Ax^{(0)} - b\tau^{(0)}), \\ A^T y + s - c\tau & = & \gamma(A^T y^{(0)} + s^{(0)} - c\tau^{(0)}), \\ -c^T x + b^T y - \kappa & = & \gamma(-c^T x^{(0)} + b^T y^{(0)} - \kappa^{(0)}), \\ XSe & = & \gamma\mu^{(0)} e, \\ \tau\kappa & = & \gamma\mu^{(0)}, \end{array} \tag{9}$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

where $\gamma \in [0,1]$ and

$$\mu^{(0)} := \frac{(x^{(0)})^T s^{(0)} + \tau^{(0)}\kappa^{(0)}}{r+1} \quad \text{and} \quad e := \begin{bmatrix} e^1 \\ \vdots \\ e^k \end{bmatrix}.$$

Observe:

- For instance choose

$$(x^{(0)}, \tau^{(0)}, y^{(0)}, s^{(0)}, \kappa^{(0)}) = (e, 1, 0, e, 1).$$

- That point is on the central path for $\gamma = 1$.

# Neighborhood definition

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Nesterov and Todd proves for $(x;\tau),(s,\kappa) \in \mathrm{int}(\mathcal{K})$:

$$(x^T s + \tau\kappa)\left(\left(\sum_{k=1}^{r} \frac{(x^k)^T s^k}{(x^k)^T Q^k x^k (s^k)^T Q^k s^k}\right) + \frac{1}{\tau\kappa}\right) \geq (r+1)^2$$

- If the inequality holds as equality if the point is on the central path.

If $\beta \in (0,1]$ and

$$\begin{array}{rcl} \frac{(x^k)^T Q^k x^k (s^k)^T Q^k s^k}{(x^k)^T s^k} & \geq & \beta \frac{x^T s + \tau\kappa}{r+1}, \quad \forall k \\ \tau\kappa & \geq & \beta \frac{x^T s + \tau\kappa}{r+1} \end{array}$$

then

$$(x^T s + \tau\kappa)\left(\left(\sum_{k=1}^{r} \frac{(x^k)^T s^k}{(x^k)^T Q^k x^k (s^k)^T Q^k s^k}\right) + \frac{1}{\tau\kappa}\right) \leq \frac{1}{\beta}(r+1)^2.$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Central path neighborhood ($\mathcal{N}(\beta)$):

$$\min \begin{pmatrix} \frac{(x^1)^T Q^1 x^1 (s^1)^T Q^1 s^1}{(x^1)^T s^1} \\ \vdots \\ \frac{(x^k)^T Q^k x^k (s^k)^T Q^k s^k}{(x^k)^T s^k} \\ \tau\kappa \end{pmatrix} \geq \beta\mu$$

and

$$\mu := \frac{x^T s + \tau\kappa}{r+1}$$

where $\beta \in [0, 1]$.

# Algorithm outline

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

- Follow the central path to the optimum.
  - I.e. stay in the neighborhood of the central path.
- Use Newton's method to compute points in the neighborhood.

# The (unscaled) Newton direction

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

$$
\begin{array}{rcl}
Ad_x - bd_\tau & = & \eta(Ax^{(0)} - b\tau^{(0)}), \\
A^T d_y + d_s - cd_\tau & = & \eta(A^T y^{(0)} + s^{(0)} - c\tau^{(0)}), \\
-c^T d_x + b^T d_y - d_\kappa & = & \eta(-c^T x^{(0)} + b^T y^{(0)} - \kappa), \\
X^{(0)} T d_s + S^{(0)} T d_x & = & -X^{(0)} S^{(0)} e + \gamma\mu^{(0)} e, \\
\tau^{(0)} d_\kappa + \kappa^{(0)} d_\tau & = & -\tau^{(0)}\kappa^{(0)} + \gamma\mu^{(0)}.
\end{array}
$$

where $\eta = \gamma - 1$.

Problems:

- The search direction is not well-defined everywhere.
- Hard to prove polynomial convergence.
- (The search direction is hard to compute because of lack of symmetry. Makes the linear algebra expensive.)
- Solution: Perform Newtons method in a scaled space.

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

# What is scaling?

**Definition 2** $W^k \in \mathcal{R}^{n^k \times n^k}$ *is a scaling matrix if it satisfies the conditions*

$$\begin{array}{rcl} W^k & \succ & 0, \\ W^k Q^k W^k & = & Q^k. \end{array}$$

*A scaled point* $\bar{x}, \bar{s}$ *is obtained by the transformation*

$$\bar{x} := \Theta W x \quad \textit{and} \quad \bar{s} := (\Theta W)^{-1} s,$$

*where*

$$\begin{array}{rcl} W & := & diag(W^1, \ldots, W^k), \\ \Theta & := & diag(\theta^1 1_{n^1}; \ldots; \theta^k 1_{n^k}). \end{array}$$

*and* $\theta^k > 0$.

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

## Lemma 2

*i)* $(x^k)^T s^k = (\bar{x}^k)^T \bar{s}^k$.

*ii)* $\theta_k^2 (x^k)^T Q^k x^k = (\bar{x}^k)^T Q^k \bar{x}^k$.

*iii)* $\theta_k^{-2} (s^k)^T Q^k s^k = (\bar{s}^k)^T Q^k \bar{s}^k$.

*iv)* $x \in \mathcal{K} \Leftrightarrow \bar{x} \in \mathcal{K}$ $(x \in \mathit{int}(\mathcal{K}) \Leftrightarrow \bar{x} \in \mathit{int}(K))$.

*v)* *Given a* $\beta \in (0,1)$ *then*

$$(x, \tau, s, \kappa) \in \mathcal{N}(\beta) \Rightarrow (\bar{x}, \tau, \bar{s}, \kappa) \in \mathcal{N}(\beta).$$

# Nesterov-Todd scaling

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Comments:

- Many choices for a scaling has been suggested.
- Many of them leads polynomial complexity.
- The most satisfactory one is the Nesterov-Todd scaling which chooses the scaling such that

$$\Theta W x = \bar{x} = \bar{s} = (\Theta W)^{-1} s$$

or equivalently

$$s = W \Theta^2 W x.$$

# How to compute the scaling

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Assume that $x^k, s^k \in \text{int}(\mathcal{K}^k)$ then

$$\theta_k^2 = \sqrt{\frac{(s^k)^T Q^k s^k}{(x^k)^T Q^k x^k}}. \tag{10}$$

Moreover, if $\mathcal{K}^k$ is

**i)** the positive half-line $\mathbf{R}_+$, then:

$$W^k = \frac{1}{\theta_k}((X^k)^{-1} S^k)^{\frac{1}{2}}.$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

**ii)** a quadratic cone, then:

$$\begin{aligned} W^k &= \begin{bmatrix} w_1^k & \left(w_{2:n^k}^k\right)^T \\ w_{2:n^k}^k & I + \frac{w_{2:n^k}^k\left(w_{2:n^k}^k\right)^T}{1+w_1^k} \end{bmatrix} \\ &= -Q^k + \frac{(e_1^k+w^k)(e_1^k+w^k)^T}{1+(e_1^k)^T w^k} \end{aligned} \tag{11}$$

where

$$w^k = \frac{\theta_k^{-1} s^k + \theta_k Q^k x^k}{\sqrt{2}\sqrt{(x^k)^T s^k + \sqrt{(x^k)^T Q^k x^k (s^k)^T Q^k s^k}}}. \tag{12}$$

Furthermore,

$$(W^k)^2 = -Q^k + 2w^k(w^k)^T. \tag{13}$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

**iii)** a rotated quadratic cone, then:

$$W^k = -Q^k + \frac{(T^k e_1^k + w^k)(T^k e_1^k + w^k)^T}{1 + (e_1^k)^T T^k w^k} \tag{14}$$

where $w^k$ is given by (12). Furthermore,

$$(W^k)^2 = -Q^k + 2w^k(w^k)^T. \tag{15}$$

# Important facts for computations

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

**Lemma 3**

$$(\theta_k W^k)^{-2} = \theta_k^{-2} Q^k (W^k)^2 Q^k.$$

Notes:

- $W^k$ can be stored using a $n^k$ dimensional vector.
- Multiplications with $W^k$ and $(W^k)^{-1}$ can be carried out in $O(n^k)$ complexity.
- $(W^k)$ has the simple structure

$$-Q^k + 2w^k(w^k)^T.$$

# The Nesterov-Todd search direction

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

$$
\begin{aligned}
A d_x - b d_\tau &= \eta(A x^{(0)} - b\tau^{(0)}), \\
A^T d_y + d_s - c d_\tau &= \eta(A^T y^{(0)} + s^{(0)} - c\tau^{(0)}), \\
-c^T d_x + b^T d_y - d_\kappa &= \eta(-c^T x^{(0)} + b^T y^{(0)} - \kappa), \\
\bar{X}^{(0)} T (\Theta W)^{-1} d_s + \bar{S}^{(0)} T \Theta W d_x &= -\bar{X}^{(0)} \bar{S}^{(0)} e + \gamma\mu^{(0)} e, \\
\tau^{(0)} d_\kappa + \kappa^{(0)} d_\tau &= -\tau^{(0)}\kappa^{(0)} + \gamma\mu^{(0)}.
\end{aligned}
$$

where $\eta := \gamma - 1$.

New iterate:

$$
\begin{bmatrix} x^{(1)} \\ \tau^{(1)} \\ y^{(1)} \\ s^{(1)} \\ \kappa^{(1)} \end{bmatrix} = \begin{bmatrix} x^{(0)} \\ \tau^{(0)} \\ y^{(0)} \\ s^{(0)} \\ \kappa^{(0)} \end{bmatrix} + \alpha \begin{bmatrix} d_x \\ d_\tau \\ d_y \\ d_s \\ d_\kappa \end{bmatrix}.
$$

# Properties of the search direction

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

## Lemma 4

$$
\begin{aligned}
Ax^{(1)} - b\tau^{(1)} &= (1+\alpha\eta)(Ax^{(0)} - b\tau^{(0)}), \\
A^T y^{(1)} + s^{(1)} - c\tau^{(1)} &= (1+\alpha\eta)(A^T y^{(0)} + s^{(0)} - c\tau^{(0)}), \\
-c^T x^{(1)} + b^T y^{(1)} - \kappa^{(1)} &= (1+\alpha\eta)(-c^T x^{(0)} + b^T y^{(0)} - \kappa^{(0)}), \\
d_x^T d_s^T + d_\tau d_\kappa &= 0, \\
(x^{(1)})^T s^{(1)} + \tau^{(1)}\kappa^{(1)} &= (1+\alpha\eta)((x^{(0)})^T s^{(0)} + \tau^{(0)}\kappa^{(0)}).
\end{aligned}
$$

Observations:

- The complementarity gap is reduced by a factor of $(1+\alpha\eta) \in [0,1)$.
- The infeasibility is reduced by the same factor.
- High advantageous property.
- Implies convergence.

# Practical issues

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

- Step-size computation
  - Back-tracking line search type.
  - Computational cheap.
- Mehrotra predictor-corrector extension.
  - Estimate $\gamma$.
  - High-order correction.

# Practical stopping criteria

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

- A solution

$$(x, y, s) = (x^{(k)}, y^{(k)}, s^{(k)})/\tau^{(k)}$$

is said to be primal dual optimal solution if

$$\begin{array}{rcl} \left\| Ax^{(k)} - b\tau^{(k)} \right\|_\infty & \leq & \varepsilon_p(1 + \|b\|_\infty)\tau^{(k)}, \\ \left\| A^T y^{(k)} + s^{(k)} - c\tau^{(k)} \right\|_\infty & \leq & \varepsilon_d(1 + \|c\|_\infty)\tau^{(k)}, \\ \frac{|c^T x^{(k)} - b^T y^{(k)}|}{\tau^{(k)} + \max(|c^T x^{(k)}|, |b^T y^{(k)}|)} & \leq & \varepsilon_g \end{array}$$

where $\varepsilon_p, \varepsilon_d$ and $\varepsilon_g$ all are small user specified constants.

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

- If

$$b^T y^{(k)} > 0 \text{ and } b^T y^{(k)} \varepsilon_p \geq \frac{\|b\|_\infty \left\| A^T y^{(k)} + s^{(k)} \right\|_\infty}{\max(1, \|c\|_\infty, |a_{ij}|)}$$

the problem is denoted to be primal infeasible and the certificate is $(y^{(k)}, s^{(k)})$ is reported.

- If

$$-c^T x^{(k)} > 0 \text{ and } -c^T x^{(k)} \varepsilon_d \geq \frac{\|c\|_\infty \left\| A x^{(k)} \right\|_\infty}{\max(1, \|b\|_\infty, |a_{ij}|)}$$

is said denoted to be dual infeasible and the certificate is $x^{(k)}$ is reported.

# Computation of search direction

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

The computational most expensive operation in the algorithm is the search direction computation:

$$
\begin{array}{rcl}
Ad_x - bd_\tau & = & f^1, \\
A^T d_y + d_s - cd_\tau & = & f^2, \\
-c^T d_x + b^T d_y - d_\kappa & = & f^3, \\
\bar{X}^{(0)} T (\Theta W)^{-1} d_s + \bar{S}^{(0)} T \Theta W d_x & = & f^4, \\
\tau^{(0)} d_\kappa + \kappa^{(0)} d_\tau & = & f^5
\end{array}
$$

where $f^i$ represents an arbitrary right-hand side.
This implies

$$
\begin{array}{rcl}
d_s & = & (\bar{X}^{(0)} T (\Theta W)^{-1})^{-1} (f^4 - \bar{S}^{(0)} T \Theta W d_x) \\
 & = & (\bar{X}^{(0)} T (\Theta W)^{-1})^{-1} f^4 - W \Theta^2 W d_x, \\
d_\kappa & = & (\tau^{(0)})^{-1} (f^5 - \kappa^{(0)} d_\tau).
\end{array}
$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Hence,

$$\begin{aligned} Ad_x - bd_\tau &= f^1, \\ A^T d_y - W\Theta^2 W d_x - cd_\tau &= \hat{f}^2, \\ -c^T d_x + b^T d_y + (\tau^{(0)})^{-1}\kappa^{(0)} d_\tau &= \hat{f}^3, \end{aligned}$$

and

$$d_x = -(W\Theta^2 W)^{-1}(\hat{f}^2 - A^T d_y + cd_\tau).$$

Thus

$$\begin{aligned} A(W\Theta^2 W)^{-1} A^T d_y - (b + A(W\Theta^2 W)^{-1} c) d_\tau &= \\ (b - A(W\Theta^2 W)^{-1} c)^T d_y + (c^T (W\Theta^2 W)^{-1} c + (\tau^{(0)})^{-1}\kappa^{(0)}) d_\tau &= \end{aligned}$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

Given

$$M = A(W\Theta^2 W)A^T = \sum_{k=1}^{r} \theta_k^{-2} A^k (W^k)^{-2} (A^k)^T,$$

and

$$\begin{aligned} Mv^1 &= (b + A(W\Theta^2 W)^{-1}c), \\ Mv^2 &= \hat{f}^1 \end{aligned}$$

we reach the easy solvable linear system

$$\begin{aligned} d_y - v^1 d_\tau &= \\ (b - A(W\Theta^2 W)^{-1}c)^T d_y + (c^T (W\Theta^2 W)^{-1} c + (\tau^{(0)})^{-1} \kappa^{(0)}) d_\tau &= \end{aligned}$$

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

- The hard part is the linear equation systems involving $M$.
- Observe that:

$$M = A(W\Theta^2 W)^{-1}A^T = \sum_{k=1}^{r} \theta_k^{-2} A^k (W^k)^{-2} (A^k)^T,$$

where

$$\begin{array}{rcl} A^k (W^k)^{-2} (A^k)^T & = & A^k Q^k (-Q^k + 2 w^k (w^k)^T) Q^k (A^k)^T \\ & = & -A^k Q^k (A^k)^T \\ & & +2(A^k Q^k w^k)(A^k Q^k w^k)^T, \end{array}$$

- $M = M^T$.
- $M$ is positive definite.
- Use Cholesky factorization $M = LL^T$.

# Can we use sparse computations?

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

- Is $M$ sparse? Yes, if

  $$-A^k Q^k (A^k)^T$$

  and

  $$(A^k Q^k w^k)(A^k Q^k w^k)^T$$

  is sparse. Likely to be the case if
  - $A^k$ is sparse.
  - $A^k$ contains no dense columns.
  - $w^k$ is not high dimensional.
- $M$ is usually very sparse in the linear case.

Problematic:

- Big cones and/or dense columns in $A$ are trouble some.
- It is possible to deal with dense columns and large cones see [1] for details.

# Practical implementation

- Employs presolve to reduce problem size.
- Exploit problem structure:
  - Upper bounds on linear variables: $x_j \leq u_j$.
  - Fixed variables: $x_j = u_j$.
- Sparse Cholesky (minimum degree or GP ordering).

Introduction
A symmetric primal-dual algorithm
Recap.
Simplifying notation
Central theorem
Goals
A homogeneous model
Quadratic cones
Some definitions
Complementarity
The central path
Neighborhood definition
Algorithm outline
What is scaling?
Nesterov-Todd scaling
The Nesterov-Todd search direction
Properties of the search direction
Practical stopping criteria
Computation of search direction
Can we use sparse computations?
Practical implementation

# Computational results

Introduction
A symmetric primal-dual algorithm
Computational results
Hardware and software
Problems
Optimized problems
Accuracy and efficiency
Summary
References

# Hardware and software

- MOSEK.
- Windows X64.
- AMD 64bit 2.21 Ghz, 4GB RAM.

http://www.mosek.com

Introduction
A symmetric primal-dual algorithm
Computational results
Hardware and software
Problems
Optimized problems
Accuracy and efficiency
Summary
References

# Problems

| | | | Presolved | |
|---|---|---|---|---|
| Name | Constraints | Variables | Constraints | Variables |
| arkadi-bug | 5 | 13 | 5 | 13 |
| arkadi1 | 1 | 4 | 1 | 4 |
| arkadi6 | 758 | 1452 | 758 | 1452 |
| bart1 | 14 | 33 | 14 | 33 |
| c-glineur3 | 698 | 1049 | 698 | 1049 |
| c-nql180 | 42300 | 106622 | 42300 | 106622 |
| c-qssp180 | 64799 | 261364 | 64799 | 261364 |
| c-qssp90 | 16199 | 65884 | 16199 | 65884 |
| c-traffic-36 | 4035 | 6730 | 4035 | 6730 |
| clinton3 | 17208 | 30741 | 17208 | 30741 |
| dsNRL | 405 | 15897 | 405 | 15897 |
| dttd13-13-2 | 14744 | 84708 | 14744 | 84708 |
| ex4-mark | 23 | 32 | 23 | 32 |
| firL1 | 301 | 17766 | 301 | 17766 |
| firL1Linfalph | 302 | 35532 | 302 | 35532 |
| firL1Linfeps | 5867 | 11532 | 5867 | 11532 |
| firL2L1alph | 5868 | 9612 | 5868 | 9612 |
| firL2Linfeps | 6086 | 14711 | 6086 | 14711 |
| firLinf | 402 | 11886 | 402 | 11886 |
| ivor1 | 900 | 2709 | 900 | 2709 |
| nb | 121 | 2379 | 121 | 2379 |
| nb_L1 | 913 | 3172 | 913 | 3172 |
| nb_L2 | 120 | 4191 | 120 | 4191 |
| nb_L2_bessel | 122 | 2639 | 122 | 2639 |
| rjabr1 | 210 | 344 | 210 | 344 |
| sched_100_100_scaled | 8337 | 18238 | 8337 | 18238 |
| sched_100_50_scaled | 4843 | 9744 | 4843 | 9744 |
| sched_200_100_scaled | 18086 | 37887 | 18086 | 37887 |
| sched_50_50_scaled | 2526 | 4977 | 2526 | 4977 |
| than-x1-1 | 1270 | 1422 | 1270 | 1422 |
| wbNRL | 459 | 18295 | 459 | 18295 |

# Optimized problems

Introduction
A symmetric primal-dual algorithm
Computational results
Hardware and software
Problems
Optimized problems
Accuracy and efficiency
Summary
References

| Name | Con- straints | Quad. cones | Varia- bles | Cone var. | Bnd. cone var. |
|---|---|---|---|---|---|
| arkadi-bug | 5 | 13 | 3 | 11 | 0 |
| arkadi1 | 1 | 4 | 1 | 3 | 0 |
| arkadi6 | 758 | 1452 | 270 | 1164 | 0 |
| bart1 | 14 | 33 | 5 | 25 | 1 |
| c-glineur3 | 698 | 1049 | 1 | 350 | 0 |
| c-nql180 | 42300 | 106622 | 32400 | 97199 | 0 |
| c-qssp180 | 64799 | 261364 | 65341 | 261364 | 0 |
| c-qssp90 | 16199 | 65884 | 16471 | 65884 | 0 |
| c-traffic-36 | 4035 | 6730 | 1365 | 4025 | 1330 |
| clinton3 | 17208 | 30741 | 3614 | 13609 | 3621 |
| dsNRL | 405 | 15897 | 5254 | 15897 | 0 |
| dttd13-13-2 | 14744 | 84708 | 28236 | 84708 | 0 |
| ex4-mark | 23 | 32 | 1 | 24 | 0 |
| firL1 | 301 | 17766 | 5922 | 17766 | 0 |
| firL1Linfalph | 302 | 35532 | 11844 | 35532 | 0 |
| firL1Linfeps | 5867 | 11532 | 3844 | 11532 | 0 |
| firL2L1alph | 5868 | 9612 | 1923 | 9611 | 1 |
| firL2Linfeps | 6086 | 14711 | 2943 | 14711 | 0 |
| firLinf | 402 | 11886 | 3962 | 11886 | 0 |
| ivor1 | 900 | 2709 | 4 | 1808 | 0 |
| nb | 121 | 2379 | 793 | 2379 | 0 |
| nb_L1 | 913 | 3172 | 793 | 2379 | 793 |
| nb_L2 | 120 | 4191 | 839 | 4191 | 0 |
| nb_L2_bessel | 122 | 2639 | 839 | 2637 | 0 |
| rjabr1 | 210 | 344 | 57 | 238 | 41 |
| sched_100_100_scaled | 8337 | 18238 | 1 | 8236 | 2592 |
| sched_100_50_scaled | 4843 | 9744 | 1 | 4742 | 1553 |
| sched_200_100_scaled | 18086 | 37887 | 1 | 17885 | 5655 |
| sched_50_50_scaled | 2526 | 4977 | 1 | 2475 | 808 |
| than-x1-1 | 1270 | 1422 | 487 | 1422 | 0 |
| wbNRL | 459 | 18295 | 9 | 1118 | 0 |

# Accuracy and efficiency

Introduction
A symmetric primal-dual algorithm
Computational results
Hardware and software
Problems
Optimized problems
Accuracy and efficiency
Summary
References

| Name | Primal obj. | Sig. fig. | Iter. | Time |
|---|---|---|---|---|
| arkadi-bug | 4.097343e-001 | 10 | 7 | 0.00 |
| arkadi1 | -1.414214e-001 | 10 | 8 | 0.01 |
| arkadi6 | 9.627900e+006 | 9 | 25 | 0.12 |
| bart1 | 1.348270e+002 | 9 | 14 | 0.01 |
| c-glineur3 | 1.842864e-004 | 9 | 11 | 0.03 |
| c-nql180 | -9.276925e-001 | 9 | 21 | 64.46 |
| c-qssp180 | -6.639490e+000 | 10 | 21 | 61.76 |
| c-qssp90 | -6.594398e+000 | 11 | 22 | 9.73 |
| c-traffic-36 | -5.390246e+003 | 10 | 26 | 0.71 |
| clinton3 | -3.295814e-005 | 10 | 28 | 5.35 |
| dsNRL | -5.574582e-005 | 9 | 34 | 166.45 |
| dttd13-13-2 | 1.917137e+004 | 9 | 40 | 10.59 |
| ex4-mark | -1.098362e+000 | 9 | 9 | 0.01 |
| firL1 | -3.522637e+000 | 13 | 17 | 58.93 |
| firL1Linfalph | -3.181250e+000 | 12 | 26 | 180.37 |
| firL1Linfeps | -1.556303e-002 | 10 | 63 | 13.96 |
| firL2L1alph | -2.328451e-001 | 9 | 11 | 3.31 |
| firL2Linfeps | -1.033366e-002 | 6 | 22 | 25.25 |
| firLinf | -1.006898e-002 | 7 | 28 | 111.31 |
| ivor1 | -1.801961e-001 | 8 | 18 | 23.39 |
| nb | -5.070309e-002 | 9 | 20 | 1.96 |
| nb_L1 | -1.301227e+001 | 9 | 17 | 1.79 |
| nb_L2 | -1.628972e+000 | 8 | 13 | 3.31 |
| nb_L2_bessel | -1.025695e-001 | 9 | 11 | 1.18 |
| rjabr1 | 3.397210e+000 | 9 | 9 | 0.01 |
| sched_100_100_scaled | 2.733459e+001 | 8 | 41 | 2.59 |
| sched_100_50_scaled | 6.716503e+001 | 9 | 31 | 1.04 |
| sched_200_100_scaled | 5.181196e+001 | 8 | 48 | 8.23 |
| sched_50_50_scaled | 7.852038e+000 | 10 | 21 | 0.34 |
| than-x1-1 | -2.275539e+001 | 9 | 18 | 0.12 |
| wbNRL | -4.149784e-005 | 9 | 22 | 152.43 |

# Summary

# Conclusion

Introduction
A symmetric primal-dual algorithm
Computational results
Summary
Conclusion
References

- The symmetric primal-dual algorithm of Nesterov-Todd is theoretical attractive algorithm for conic quadratic optimization.
- The algorithm is works very well in practice.
- Hence, large-scale sparse conic quadratic problems can be solved efficiently.

# References

Introduction
A symmetric primal-dual algorithm
Computational results
Summary
References

[1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Math. Programming*, 95(1):3–51, 2003.

[2] E. D. Andersen, C. Roos, and T. Terlaky. On implementing a primal-dual interior-point method for conic quadratic optimization. *Math. Programming*, 95(2), February 2003.

[3] A. J. Goldman and A. W. Tucker. Polyhedral convex cones. In H. W. Kuhn and A. W. Tucker, editors, *Linear Inequalities and related Systems*, pages 19–40, Princeton, New Jersey, 1956. Princeton University Press.

[4] A. J. Goldman and A. W. Tucker. Theory of linear programming. In H. W. Kuhn and A. W. Tucker, editors, *Linear Inequalities and related Systems*, pages 53–97, Princeton, New Jersey, 1956. Princeton University Press.

[5] Y. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.*, 22(1):1–42, February 1997.

[6] Yu. Nesterov, M. J. Todd, and Y. Ye. Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming