18. Primal-dual interior-point methods

- primal-dual central path equations
- infeasible primal-dual method
- primal-dual method for self-dual embedding

Symmetric cone program

primal and dual problem

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax + s = b \\
& s \succeq 0
\end{array}
\qquad\qquad
\begin{array}{ll}
\text{maximize} & -b^T z \\
\text{subject to} & A^T z + c = 0 \\
& z \succeq 0
\end{array}
$$

inequalities are with respect to a symmetric cone

optimality conditions

$$
\begin{bmatrix} 0 \\ s \end{bmatrix} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix}
$$

$$
(s, z) \succeq 0, \qquad s^T z = 0
$$

Central path equations

barrier function: we use the log-det barrier of lecture 17

$$\phi(x) = -\log\det x$$

- a θ-normal barrier for K
- gradient is $\nabla\phi(x) = -x^{-1}$ (see page 17-14)

primal-dual central path equations

$$\begin{bmatrix} 0 \\ s \end{bmatrix} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix}$$

$$(s, z) \succ 0, \qquad z = -\mu\nabla\phi(s) = \mu s^{-1}$$

last condition can be written symmetrically as $s \circ z = \mu\, \mathbf{e}$

Scaling

scaling matrix: we call a nonsingular W a scaling matrix if

- multiplications with W and W^T preserve the cone

$$W\,\mathbf{int}\,K = \mathbf{int}\,K, \qquad W^T\,\mathbf{int}\,K = \mathbf{int}\,K$$

- inverses are transformed as $Wx^{-1} = (W^{-T}x)^{-1}$

scaled central path equations: for any scaling, central path is solution of

$$\begin{bmatrix} 0 \\ s \end{bmatrix} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix}$$

$$(s, z) \succ 0, \qquad (W^{-T}s) \circ (Wz) = \mu\,\mathbf{e}$$

Nesterov-Todd scaling

for a given pair $(\hat{s}, \hat{z}) \succ 0$, define

$$W = W^T = P(w^{1/2})$$

where w satisfies $\hat{s} = P(w)\hat{z}$

- from page 17-21,

$$w = P(\hat{z}^{-1/2}) \left(P(\hat{z}^{1/2})\hat{s} \right)^{1/2}$$

- multiplications by W and W^{-1} map $\hat{s}$ and $\hat{z}$ to the same point:

$$W^{-1}\hat{s} = W\hat{z} = \lambda$$

  this implies that $\|\lambda\|_2^2 = \hat{s}^T\hat{z}$

NT scaling for nonnegative orthant

W is a positive diagonal scaling

$$W = P(w^{1/2}) = \begin{bmatrix} \sqrt{\hat{s}_1/\hat{z}_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\hat{s}_2/\hat{z}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\hat{s}_p/\hat{z}_p} \end{bmatrix}$$

- scaling point is

$$w = \left(\sqrt{\hat{s}_1/\hat{z}_1}, \sqrt{\hat{s}_2/\hat{z}_2}, \ldots, \sqrt{\hat{s}_p/\hat{z}_p} \right)$$

- scaled $\hat{s}$, $\hat{z}$ are

$$\lambda = W^{-T}\hat{s} = W\hat{z} = \left(\sqrt{\hat{s}_1\hat{z}_1},\ \sqrt{\hat{s}_2\hat{z}_2},\ \ldots,\ \sqrt{\hat{s}_p\hat{z}_p} \right)$$

NT scaling for second-order cone

W is a hyperbolic Householder transformation

$$W = P(w^{1/2}) = \beta(2vv^T - J), \qquad J = \begin{bmatrix} 1 & 0 \\ 0 & -I \end{bmatrix}$$

and

$$\beta = \frac{\bar{w}^T J \bar{w}}{2}, \qquad v = \frac{1}{\sqrt{\bar{w}^T J \bar{w}}} \bar{w}, \qquad \bar{w} = w^{1/2}$$

scaling point w can be computed from

$$w = P(\hat{z}^{-1/2}) \left(P(\hat{z}^{1/2}) \hat{s} \right)^{1/2}$$

using the expressions for P and squareroot on pages 17-15, 17-17

NT scaling for positive semidefinite cone

W is a symmetric congruence transformation

$$Wy = \mathbf{vec}\left(T^{1/2}\,\mathbf{mat}(y)T^{1/2}\right)$$

where

$$T = \hat{Z}^{-1/2}\left(\hat{Z}^{1/2}\hat{S}\hat{Z}^{1/2}\right)^{1/2}\hat{Z}^{-1/2}$$

- $T = RR^T$ with R computed as on page 17-24
- a simpler, nonsymmetric scaling is

$$Wy = \mathbf{vec}\left(R^T\,\mathbf{mat}(y)R\right), \qquad W^Ty = \mathbf{vec}\left(R\,\mathbf{mat}(y)R^T\right)$$

Outline

- primal-dual central path equations
- **infeasible primal-dual method**
- primal-dual method for self-dual embedding

Basic primal-dual update

suppose the current iterates are $\hat{s}$, $\hat{x}$, $\hat{z}$ with $\hat{s} \succ 0$, $\hat{z} \succ 0$

- define $\mu = \hat{s}^T\hat{z}/\theta$ and compute the NT scaling matrix W for $\hat{s}$, $\hat{z}$
- compute Δs, Δx, Δz by linearizing the central path equation

$$\begin{bmatrix} 0 \\ s \end{bmatrix} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix}$$

$$(W^{-T}s) \circ (Wz) = \sigma\mu\,\mathbf{e}$$

  around $\hat{s}$, $\hat{x}$, $\hat{z}$, for some $\sigma < 1$

- make an update

$$(\hat{s}, \hat{x}) := (\hat{s}, \hat{x}) + \alpha_{\rm p}(\Delta x, \Delta s), \qquad \hat{z} := \hat{z} + \alpha_{\rm d}\Delta z$$

  that preserves positivity of $\hat{s}$, $\hat{z}$

Linearized central path equation

define $\lambda = W^{-T}\hat{s} = W\hat{z}$ and

$$r = \begin{bmatrix} 0 \\ \hat{s} \end{bmatrix} - \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{z} \end{bmatrix} - \begin{bmatrix} c \\ b \end{bmatrix}$$

linearized central path equation

$$\begin{bmatrix} 0 \\ \Delta s \end{bmatrix} - \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix} = -r$$

$$\lambda \circ \left(W\Delta z + W^{-T}\Delta s\right) = \sigma\mu\,\mathbf{e} - \lambda \circ \lambda$$

second equation is linearization of

$$(W^{-T}(\hat{s} + \Delta s)) \circ (W(\hat{z} + \Delta z)) = \sigma\mu\,\mathbf{e}$$

Path-following algorithm

choose starting points $\hat{s}$, $\hat{x}$, $\hat{z}$ with $\hat{s} \succ 0$, $\hat{z} \succ 0$

1. **compute residuals and evaluate stopping criteria**

$$r = \begin{bmatrix} 0 \\ \hat{s} \end{bmatrix} - \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{z} \end{bmatrix} - \begin{bmatrix} c \\ b \end{bmatrix}$$

   terminate if r and $\hat{s}^T \hat{z}$ are sufficiently small

2. **compute scaling matrix** W associated with $(\hat{s}, \hat{z})$ and set

$$\lambda := W^{-T} \hat{s} = W \hat{z}, \qquad \mu := \frac{\lambda^T \lambda}{\theta} = \frac{\hat{s}^T \hat{z}}{\theta}$$

3. **compute affine scaling direction** by solving the linear equation

$$\begin{bmatrix} 0 \\ \Delta s_{\rm a} \end{bmatrix} - \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\rm a} \\ \Delta z_{\rm a} \end{bmatrix} = -r$$

$$\lambda \circ \left(W \Delta z_{\rm a} + W^{-T} \Delta s_{\rm a} \right) = -\lambda \circ \lambda$$

4. **select barrier parameter**

$$\sigma = \left(\frac{(\hat{s} + \alpha_{\rm p} \Delta s_{\rm a})^T (\hat{z} + \alpha_{\rm d} \Delta z_{\rm a})}{\hat{s}^T \hat{z}} \right)^{\delta}$$

where δ is an algorithm parameter (a typical value is $\delta = 3$) and

$$\begin{aligned} \alpha_{\rm p} &= \sup\{\alpha \in [0,1] \mid \hat{s} + \alpha \Delta s_{\rm a} \succeq 0\} \\ \alpha_{\rm d} &= \sup\{\alpha \in [0,1] \mid \hat{z} + \alpha \Delta z_{\rm a} \succeq 0\} \end{aligned}$$

5. **compute search direction** by solving the linear equation

$$\begin{bmatrix} 0 \\ \Delta s \end{bmatrix} - \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix} = -r$$

$$\lambda \circ (W \Delta z + W^{-T} \Delta s) = \sigma \mu \, \mathbf{e} - \lambda \circ \lambda$$

6. **update iterates**

$$\begin{aligned} (\hat{x}, \hat{s}) &:= (\hat{x}, \hat{s}) + \min\{1, 0.99\alpha_{\mathrm{p}}\}(\Delta x, \Delta s) \\ \hat{z} &:= \hat{z} + \min\{1, 0.99\alpha_{\mathrm{d}}\}\Delta z \end{aligned}$$

where

$$\alpha_{\mathrm{p}} = \sup\{\alpha \geq 0 \mid \hat{s} + \alpha \Delta s \succeq 0\}, \qquad \alpha_{\mathrm{d}} = \sup\{\alpha \geq 0 \mid \hat{z} + \alpha \Delta z \succeq 0\}$$

return to step 1

interpretation and discussion

- step 3: affine scaling direction solves linearized central path equation with $\sigma = 0$, $i.e.$, the linearized optimality conditions

- step 4 is a heuristic for choosing σ based on an estimate of the quality of the affine scaling direction

  σ is small if a step in the affine scaling direction gives a large reduction in $\hat{s}^T\hat{z}$

- step 5: linear equation has same coefficient matrix as equation in step 3

  if a direct method is used, we can reuse the factorization used in step 3, and solve the two equations at the cost of one

Mehrotra correction

in step 5, solve

$$\begin{bmatrix} 0 \\ \Delta s \end{bmatrix} - \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix} = -r$$

$$\lambda \circ (W\Delta z + W^{-T}\Delta s) = \sigma\mu\,\mathbf{e} - \lambda \circ \lambda - (W^{-T}\Delta s_{\mathrm{a}}) \circ (W\Delta z_{\mathrm{a}})$$

- extra term on the r.h.s. is approximation of the second-order term in

$$(W^{-T}(\hat{s} + \Delta s)) \circ (W(\hat{z} + \Delta z)) = \sigma\mu\,\mathbf{e}$$

- adding the correction typically saves a few iterations

Newton equations

steps 3 and 5 reduce to equations

$$\begin{bmatrix} 0 & A^T \\ A & -W^TW \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix} = \begin{bmatrix} d_x \\ d_z \end{bmatrix}$$

usually solved by eliminating Δz:

$$A^TW^{-1}W^{-T}A\,\Delta x = d_x + A^TW^{-1}W^{-T}d_z$$

- a KKT system (see §10.4.2 in BV for a discussion of solution methods)
- since $W^TW = P(w) = \nabla^2\phi(w)^{-1}$,

$$A^TW^{-1}W^{-T}A = A^T\nabla^2\phi(w)A,$$

  the Hessian of the barrier function $\phi(b - Ax)$ at the scaling point w

Quadratic cone program

$$\begin{array}{ll} \text{minimize} & (1/2)x^T Q x + q^T x \\ \text{subject to} & Ax + s = b \\ & s \succeq 0 \end{array}$$

optimality conditions

$$\left[\begin{array}{c} 0 \\ s \end{array}\right] = \left[\begin{array}{cc} Q & A^T \\ -A & 0 \end{array}\right] \left[\begin{array}{c} x \\ z \end{array}\right] + \left[\begin{array}{c} q \\ b \end{array}\right], \quad (s, z) \succeq 0, \quad s^T z = 0$$

central path

$$\left[\begin{array}{c} 0 \\ s \end{array}\right] = \left[\begin{array}{cc} Q & A^T \\ -A & 0 \end{array}\right] \left[\begin{array}{c} x \\ z \end{array}\right] + \left[\begin{array}{c} q \\ b \end{array}\right], \quad (s, z) \succ 0, \quad s \circ z = \mu \mathbf{e}$$

Path-following algorithm

algorithm is almost identical to algorithm on page 18-11

- compute search directions from linearized central path equation;

  for example, step 5 becomes

$$\begin{bmatrix} 0 \\ \Delta s \end{bmatrix} - \begin{bmatrix} Q & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix} = -r$$

$$\lambda \circ (W\Delta z + W^{-T}\Delta s) = \sigma\mu\,\mathbf{e} - \lambda \circ \lambda$$

- use equal primal and dual step sizes

  for example, in step 6,

$$\alpha_{\rm p} = \alpha_{\rm d} = \sup\{\alpha \geq 0 \mid \hat{s} + \alpha\Delta s \succeq 0, \hat{z} + \alpha\Delta z \succeq 0\}$$

Outline

- primal-dual central path equations
- infeasible primal-dual method
- **primal-dual method for self-dual embedding**

Extended self-dual embedding

$$\begin{array}{ll} \text{minimize} & (\theta+1)\gamma \\ \text{subject to} & \begin{bmatrix} 0 \\ s \\ \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & A^T & c & q_x \\ -A & 0 & b & q_z \\ -c^T & -b^T & 0 & q_\tau \\ -q_x^T & -q_z^T & -q_\tau & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ \tau \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \theta+1 \end{bmatrix} \\ & (s, \kappa, z, \tau) \succeq 0 \end{array}$$

- θ is the logarithmic degree or rank of the cone
- q_x, q_z, q_τ defined as

$$\begin{bmatrix} q_x \\ q_z \\ q_\tau \end{bmatrix} = \frac{\theta+1}{s_0^T z_0 + 1} \left(\begin{bmatrix} 0 \\ s_0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & A^T & c \\ -A & 0 & b \\ -c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ z_0 \\ 1 \end{bmatrix} \right)$$

s_0, x_0, z_0 are arbitrary with $s_0 \succ 0$, $z_0 \succ 0$

Optimality condition

$$\begin{bmatrix} 0 \\ s \\ \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & A^T & c & q_x \\ -A & 0 & b & q_z \\ -c^T & -b^T & 0 & q_\tau \\ -q_x^T & -q_z^T & -q_\tau & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ \tau \\ \gamma \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \theta + 1 \end{bmatrix}$$

$$(s, \kappa, z, \tau) \succeq 0, \qquad s^T z + \kappa\tau = 0$$

- follows from self-dual property
- shows that $\gamma = 0$ at optimum
- optimal solution gives nonzero solution of embedding of p.14-30

Properties of extended self-dual embedding

- problem is strictly feasible; a strictly feasible point is given by

$$(s,\, \kappa,\, x,\, z,\, \tau,\, \gamma) = (s_0,\, 1,\, x_0,\, z_0,\, 1,\, \frac{s_0^T z_0 + 1}{\theta + 1}) \qquad (1)$$

- if $s,\, \kappa,\, x,\, z,\, \tau,\, \gamma$ satisfy equality constraint, then

$$\gamma = \frac{s^T z + \kappa\tau}{\theta + 1}$$

(take inner product with (x, z, τ, γ) on two sides of the equality)

- this is the extended embedding of page 14-34, but using variable γ instead of θ, and with a coefficient $\theta + 1$ in objective and r.h.s.

Central path for extended embedding

$$\begin{bmatrix} 0 \\ s \\ \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & A^T & c & q_x \\ -A & 0 & b & q_z \\ -c^T & -b^T & 0 & q_\tau \\ -q_x^T & -q_z^T & -q_\tau & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ \tau \\ \gamma \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \theta + 1 \end{bmatrix}$$

$$(s, \kappa, z, \tau) \succeq 0, \qquad s \circ z = \mu \mathbf{e}, \qquad \kappa\tau = \mu$$

- inner product with (x, z, τ, γ) shows that on the central path

$$\gamma = \frac{z^T s + \kappa\tau}{\theta + 1} = \mu$$

- initial point (1) is on the central path with $\mu = (s_0^T z_0 + 1)/(\theta + 1)$

Simplified central path equations

$$\begin{bmatrix} 0 \\ s \\ \kappa \end{bmatrix} = \begin{bmatrix} 0 & A^T & c \\ -A & 0 & b \\ -c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ \tau \end{bmatrix} + \mu \begin{bmatrix} q_x \\ q_z \\ q_\tau \end{bmatrix}$$

$$(s, \kappa, z, \tau) \succeq 0, \qquad s \circ z = \mu\, \mathbf{e}, \qquad \kappa\tau = \mu$$

- we eliminated variable γ because $\gamma = \mu$ on the central path
- we removed the 4th equality, because it is implied by the first three (this follows by taking inner product with (x, z, τ))
- can be seen as a 'shifted central path' for the embedding on p.14-30

Path-following algorithm

choose starting points $\hat{s}$, $\hat{x}$, $\hat{z}$, with $\hat{s} \succ 0$, $\hat{z} \succ 0$; set $\hat{\kappa} := 1$, $\hat{\tau} := 1$

1. **compute residuals and evaluate stopping criteria**

$$r = \begin{bmatrix} 0 \\ \hat{s} \\ \hat{\kappa} \end{bmatrix} - \begin{bmatrix} 0 & A^T & c \\ -A & 0 & b \\ -c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{z} \\ \hat{\tau} \end{bmatrix}$$

   terminate if r and $\hat{s}^T \hat{z}/\tau^2$ are sufficiently small, or an approximate certificate of primal or dual infeasibility has been found

2. **compute scaling matrix** W associated with $(\hat{s}, \hat{z})$ and set

$$\lambda := W^{-T}\hat{s} = W\hat{z}, \qquad \mu := \frac{\hat{s}^T\hat{z} + \hat{\kappa}\hat{\tau}}{\theta + 1}$$

3. **compute affine scaling direction** by solving the linear equation

$$\begin{bmatrix} 0 \\ \Delta s_{\mathrm{a}} \\ \Delta \kappa_{\mathrm{a}} \end{bmatrix} - \begin{bmatrix} 0 & A^T & c \\ -A & 0 & b \\ -c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\mathrm{a}} \\ \Delta z_{\mathrm{a}} \\ \Delta \tau_{\mathrm{a}} \end{bmatrix} = -r$$

$$\lambda \circ \left(W \Delta z_{\mathrm{a}} + W^{-T} \Delta s_{\mathrm{a}} \right) = -\lambda \circ \lambda, \qquad \hat{\kappa} \Delta \tau_{\mathrm{a}} + \hat{\tau} \Delta \kappa_{\mathrm{a}} = -\hat{\kappa} \hat{\tau}$$

4. **select barrier parameter**

$$\sigma := (1 - \alpha)^{\delta}$$

where δ is an algorithm parameter (typical value is $\delta = 3$) and

$$\alpha = \sup \{ \alpha \in [0, 1] \mid (\hat{s}, \hat{\kappa}, \hat{z}, \hat{\tau}) + \alpha(\Delta s_{\mathrm{a}}, \Delta \kappa_{\mathrm{a}}, \Delta z_{\mathrm{a}}, \Delta \tau_{\mathrm{a}}) \succeq 0 \}$$

5. **compute search direction** by solving the linear equation

$$\begin{bmatrix} 0 \\ \Delta s \\ \Delta\kappa \end{bmatrix} - \begin{bmatrix} 0 & A^T & c \\ -A & 0 & b \\ -c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \\ \Delta\tau \end{bmatrix} = -(1-\sigma)r$$

$$\lambda \circ (W\Delta z + W^{-T}\Delta s) = \sigma\mu\,\mathbf{e} - \lambda \circ \lambda - (W^{-T}\Delta s_{\mathrm{a}}) \circ (W\Delta z_{\mathrm{a}})$$

$$\hat{\kappa}\Delta\tau + \hat{\tau}\Delta\kappa = \sigma\mu - \hat{\kappa}\hat{\tau} - \Delta\kappa_{\mathrm{a}}\Delta\tau_{\mathrm{a}}$$

6. **update iterates**

$$(\hat{s}, \hat{\kappa}, \hat{x}, \hat{z}, \hat{\tau}) := (\hat{s}, \hat{\kappa}, \hat{x}, \hat{z}, \hat{\tau}) + \min\{1, 0.99\alpha\}\,(\Delta s, \Delta\kappa, \Delta x, \Delta z, \Delta\tau)$$

where $\alpha = \sup\{\alpha \in [0,1] \mid (\hat{s}, \hat{\kappa}, \hat{z}, \hat{\tau}) + \alpha(\Delta s, \Delta\kappa, \Delta z, \Delta\tau) \succeq 0\}$

return to step 1

properties (without proof)

- step 3: affine scaling direction satisfies

$$\hat{s}^T\Delta z_{\mathrm{a}} + \hat{z}^T\Delta s_{\mathrm{a}} = -\hat{s}^T\hat{z}, \qquad \hat{\kappa}\Delta\tau_{\mathrm{a}} + \hat{\tau}\Delta\kappa_{\mathrm{a}} = -\hat{\kappa}\hat{\tau}$$

$$\Delta s_{\mathrm{a}}^T\Delta z_{\mathrm{a}} + \Delta\tau_{\mathrm{a}}\Delta\kappa_{\mathrm{a}} = 0$$

- step 5: search direction satisfies

$$\hat{s}^T\Delta z + \hat{\kappa}\Delta\tau + \hat{z}^T\Delta s + \hat{\tau}\Delta\kappa = -(1-\sigma)(\hat{s}^T\hat{z} + \hat{\kappa}\hat{\tau})$$

$$\Delta s^T\Delta z + \Delta\tau\Delta\kappa = 0$$

discussion

- step 4: expression for σ is based on simplifiying

$$\sigma = \left(\frac{(\hat{s} + \alpha \Delta s_{\mathrm{a}})^T (\hat{z} + \alpha \Delta z_{\mathrm{a}}) + (\hat{\kappa} + \alpha \Delta \kappa_{\mathrm{a}})(\hat{\tau} + \alpha \Delta \tau_{\mathrm{a}})}{\hat{s}^T \hat{z} + \hat{\kappa} \hat{\tau}} \right)^{\delta}$$

- steps 5 and 6: gap and residual decrease linearly with α:

$$\mu^+ = (1 - \alpha(1 - \sigma))\mu, \qquad r^+ = (1 - \alpha(1 - \sigma))r,$$

  if μ^+ and r^+ are the values of μ and r at the next iteration

- $r = \mu q$, with q defined on p.14-34 (a multiple of the initial residual)
- in step 5, $-(1 - \sigma)r = -r + \sigma \mu q$: the equation is the linearization of the central path equation of p.18-23 for barrier parameter $\sigma \mu$

Linear algebra complexity

- essentially the same as for the method on page 18-11
- eliminating $\Delta\tau$, $\Delta\kappa$ in steps 3 and 5 requires solution of an extra system

$$\begin{bmatrix} 0 & A^T \\ A & -W^TW \end{bmatrix} \begin{bmatrix} \Delta\tilde{x} \\ \Delta\tilde{z} \end{bmatrix} = \begin{bmatrix} c \\ b \end{bmatrix}$$

- this increases the number of KKT systems solved per iteration to 3 (as opposed to 2 in the method on page 18-11)

References

implementations of primal-dual algorithms based on Nesterov-Todd scaling

- J.F. Sturm, *Implementation of interior-point methods for mixed semidefinite and second order cone optimization problems*, Optimization methods and Software (2002)

  an overview of Sedumi

- R.H. Tütüncü, K.C. Toh, M.J. Todd, *Solving semidefinite-quadratic-linear programs using SDPT3*, Mathematical Programming (2003)

  an overview of SDPT3

- CVXOPT (`cvxopt.org`)

  the `conelp` and `coneqp` solvers implement the algorithms in on page 18-24 and page 18-18

- M.S. Andersen, J. Dahl, L. Vandenberghe, *Interior-point methods for large-scale cone programming*, in: S. Sra, S. Nowozin, S.J. Wright (eds.), *Optimization for Machine Learning*, 2011

  numerical results with the CVXOPT solvers on problems with structure