18. Primal-dual interior-point methods II

- self-dual embedding
- path-following algorithm

Initialization and infeasibility detection

barrier method (EE236B)

- assumes problem is primal and dual feasible
- requires phase I to find initial primal feasible point

primal-dual path-following method (last lecture)

- assumes problem is primal and dual feasible
- allows infeasible starting points

methods based on self-dual embedding

- can detect primal and dual infeasibility
- embed cone program in slightly larger problem that is always feasible
- from solution of embedded problem, extract solution of original problem, or certificates of primal or dual infeasibility

Infeasibility

primal infeasibility: a solution y, z of

$$A^T y + G^T z = 0, \qquad h^T z + b^T y = -1, \qquad z \succeq 0$$

is a certificate of infeasibility of $Gx \preceq h$, $Ax = b$

dual infeasibility: a solution x of

$$Gx \preceq 0, \qquad Ax = 0, \qquad c^T x = -1$$

is a certificate of infeasibility of $A^T y + G^T z + c = 0$, $z \succeq 0$

these are strong alternatives if a constraint qualification holds

Self-dual embedding

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & \begin{bmatrix} 0 \\ 0 \\ s \\ \kappa \end{bmatrix} = \begin{bmatrix} 0 & A^T & G^T & c \\ -A & 0 & 0 & b \\ -G & 0 & 0 & h \\ -c^T & -b^T & -h^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ \tau \end{bmatrix} \\ & (s, \kappa, z, \tau) \succeq 0 \end{array}$$

- problem has a trivial solution (all variables zero)
- equality constraint implies $s^T z + \kappa\tau = 0$ at feasible points
- problem is not strictly feasible (hence, central path does not exist)

Optimality condition for embedded problem

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

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

- follows from self-dual property
- a (mixed) linear complementarity problem

Classification of nonzero solution

let s, κ, x, y, z, τ be a nonzero solution of the self-dual embedding

case 1: $\tau > 0$, $\kappa = 0$

$$\hat{s} = s/\tau, \qquad \hat{x} = x/\tau, \qquad \hat{y} = y/\tau, \qquad \hat{z} = z/\tau$$

are primal and dual solutions of the cone program, $i.e.$, satisfy

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

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

case 2: $\tau = 0$, $\kappa > 0$; this implies $c^Tx + b^Ty + h^Tz < 0$

- if $b^Ty + h^Tz < 0$, then

  $$\hat{z} = \frac{1}{-(h^Tz + b^Ty)}z, \qquad \hat{y} = \frac{1}{-(h^Tz + b^Ty)}y$$

  is a proof of primal infeasibility, $i.e.$, satisfies

  $$A^T\hat{y} + G^T\hat{z} = 0, \qquad h^T\hat{z} + b^T\hat{y} = -1, \qquad \hat{z} \succeq 0$$

- if $c^Tx < 0$, then

  $$\hat{x} = \frac{1}{-c^Tx}x, \qquad \hat{x} = \frac{1}{-c^Tx}s$$

  is a proof of dual infeasibility, $i.e.$, satisfies

  $$G\hat{x} \preceq 0, \qquad A\hat{x} = 0, \qquad c^T\hat{x} = -1$$

case 3: $\tau = \kappa = 0$; no conclusion can be made about the original problem

Extended self-dual embedding

$$\begin{array}{ll}
\text{min.} & (m+1)\theta \\
\text{s.t.} & \begin{bmatrix} 0 \\ 0 \\ s \\ \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & A^T & G^T & c & q_x \\ -A & 0 & 0 & b & q_y \\ -G & 0 & 0 & h & q_z \\ -c^T & -b^T & -h^T & 0 & q_\tau \\ -q_x^T & -q_y^T & -q_z^T & -q_\tau & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ \tau \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ m+1 \end{bmatrix} \\
& (s, \kappa, z, \tau) \succeq 0
\end{array}$$

here, m is the logarithmic degree of the cone and

$$\begin{bmatrix} q_x \\ q_y \\ q_z \\ q_\tau \end{bmatrix} = \frac{m+1}{s_0^Tz_0 + 1}\left(\begin{bmatrix} 0 \\ 0 \\ s_0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & A^T & G^T & c \\ -A & 0 & 0 & b \\ -G & 0 & 0 & h \\ -c^T & -b^T & -h^T & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix} \right)$$

s_0, x_0, y_0, z_0 are arbitrary with $s_0 \succ 0$, $z_0 \succ 0$

Optimality condition

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

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

- follows from self-dual property
- a (mixed) linear complementarity problem

Properties of extended self-dual embedding

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

$$(s,\, \kappa,\, x,\, y,\, z,\, \tau,\, \theta) = (s_0,\, 1,\, x_0,\, y_0,\, z_0,\, 1,\, \frac{s_0^T z_0 + 1}{m+1}) \qquad (1)$$

- if s, κ, x, y, z, τ, θ satisfy equality constraint, then

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

  (take inner product with (x, y, z, τ, θ) of each side of the equality)

- at optimum, $\theta = 0$ and problem reduces to the embedding on p.18–4
- classification of p.18–6 also applies to solutions of extended embedding

Central path for extended embedding

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

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

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

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

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

Simplified central path equations

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

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

- we eliminated variable θ because $\theta = \mu$ on the central path
- we removed the 5th equality, because it is implied by the first four (this follows by taking inner product with (x, y, z, τ))
- can be seen as a 'shifted central path' for the embedding on p.18–4

Path-following algorithm

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

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

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

   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

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

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

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

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

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

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

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

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

6. **update iterates**

$$\begin{aligned} &(\hat s, \hat\kappa, \hat x, \hat y, \hat z, \hat\tau) := \\ &\qquad (\hat s, \hat\kappa, \hat x, \hat y, \hat z, \hat\tau) + \min\{1, 0.99\alpha\}\,(\Delta s, \Delta\kappa, \Delta x, \Delta y, \Delta z, \Delta\tau) \end{aligned}$$

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

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

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

- 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.18–8 (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.18–12 for barrier parameter $\sigma \mu$

Linear algebra complexity

- essentially the same as for the method of lecture 17
- eliminating $\Delta\tau$, $\Delta\kappa$ in steps 3 and 5 requires solution of an extra system

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

  so number of KKT systems solved per iteration is 3 (as opposed to 2 in the method of lecture 17)

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 (`abel.ee.ucla.edu/cvxopt`)

  the `conelp` and `coneqp` solvers implement the algorithms in lectures 17 and 18