Barrier method

- centering problem
- Newton decrement
- local convergence of Newton method
- short-step barrier method
- global convergence of Newton method
- predictor-corrector method

Centering problem

centering problem (with notation and assumptions of page 8–14)

$$\text{minimize} \quad f_t(x) = tc^T x + \phi(x)$$

- $\phi(x) = -\sum_{i=1}^m \log(b_i - a_i^T x)$ is logarithmic barrier of $Ax \leq b$
- minimizer x is point $x^\star(t)$ on central path
- minimizer is m/t-suboptimal solution of LP:

$$c^T x^\star(t) - p^\star \leq \frac{m}{t}$$

barrier method(s):

use Newton's method to (approximately) minimize f_t, for a sequence of t

Properties of centering cost function

gradient and Hessian

$$\nabla f_t(x) = tc + A^T d_x, \qquad \nabla^2 f_t(x) = A^T \mathbf{diag}(d_x)^2 A$$

$d_x = \big(1/(b_1 - a_1^T x), \ldots, 1/(b_m - a_m^T x)\big)$ is a positive m-vector

(strict) convexity and its consequences (see pages 8–7 and 8–10)

- $\nabla^2 f_t(x)$ is positive definite for all $x \in P^\circ$
- first order condition:

$$f_t(y) > f_t(x) + \nabla f_t(x)^T (y - x) \qquad \text{for all } x, y \in P^\circ \text{ with } x \neq y$$

- x minimizes f_t if and only if $\nabla f_t(x) = 0$; if minimizer exists, it is unique

Lower bound for centering problem

centering problem is bounded below if dual LP is strictly feasible:

$$f_t(y) \geq -tb^T z + \sum_{i=1}^{m} \log z_i + m \log t + m \quad \text{for all } y \in P^\circ$$

here z can be *any* strictly dual feasible point ($A^T z + c = 0$, $z > 0$)

proof: difference of left- and right-hand sides is

$$\begin{aligned}
t(c^T y + b^T z) &- \sum_{i=1}^{m} \log(t(b_i - a_i^T y) z_i) - m \\
&= \quad t(b - Ay)^T z - \sum_{i=1}^{m} \log(t(b_i - a_i^T y) z_i) - m \\
&\geq \quad 0 \qquad\qquad (\text{since } u - \log u - 1 \geq 0)
\end{aligned}$$

Newton method for centering problem

Newton step for f_t at $x \in P^\circ = \{y \mid Ay < b\}$

$$\begin{aligned}
\Delta x_{\mathrm{nt}} &= -\nabla^2 f_t(x)^{-1} \nabla f_t(x) \\
&= -\nabla^2 \phi(x)^{-1} (tc + \nabla \phi(x)) \\
&= -\left(A^T \mathbf{diag}(d_x)^2 A\right)^{-1} \left(tc + A^T d_x\right)
\end{aligned}$$

Newton iteration: choose suitable stepsize α and make update

$$x := x + \alpha \Delta x_{\mathrm{nt}}$$

we will show that Newton method converges if f_t is bounded below

Outline

- centering problem
- **Newton decrement**
- local convergence of Newton method
- short-step barrier method
- global convergence of Newton method
- predictor-corrector method

Newton decrement

the **Newton decrement** at $x \in P^\circ$ is

$$\begin{aligned}
\lambda_t(x) &= \left(\Delta x_{\mathrm{nt}}^T \nabla^2 \phi(x) \Delta x_{\mathrm{nt}}\right)^{1/2} \\
&= \|\,\mathbf{diag}(d_x) A \Delta x_{\mathrm{nt}}\| \\
&= \|\Delta x_{\mathrm{nt}}\|_x
\end{aligned}$$

- $-\lambda_t(x)^2$ is the directional derivative of f_t at x in the direction Δx_{nt}:

$$-\lambda_t(x)^2 = \nabla f_t(x)^T \Delta x_{\mathrm{nt}} = \frac{d}{d\alpha} f_t(x + \alpha \Delta x_{\mathrm{nt}}) \Big|_{\alpha=0}$$

- $\lambda_t(x) = 0$ if and only if $x = x^\star(t)$
- we will use $\lambda_t(x)$ to measure proximity of x to $x^\star(t)$

Dual feasible points near central path

on central path (p. 8–17): strictly dual feasible point from $x = x^\star(t)$

$$z^\star(t) = \frac{1}{t} d_x, \qquad z_i^\star(t) = \frac{1}{t(b_i - a_i^T x)}, \quad i = 1, \ldots, m$$

near central path: for $x \in P^\circ$ with $\lambda_t(x) < 1$, define

$$z = \frac{1}{t} \left(d_x + \mathbf{diag}(d_x)^2 A \Delta x_{\mathrm{nt}} \right)$$

- Δx_{nt} is the Newton step for f_t at x
- z is strictly dual feasible (see next page); duality gap with x is

$$(b - Ax)^T z = \frac{m + d_x^T A \Delta x_{\mathrm{nt}}}{t} \leq \left(1 + \frac{\lambda_t(x)}{\sqrt{m}} \right) \frac{m}{t}$$

proof:

- by definition, Newton step $\Delta x_{\rm nt}$ at x satisfies

$$A^T \mathbf{diag}(d_x)^2 A \Delta x_{\rm nt} = -tc - A^T d_x$$

  therefore z satisfies the equality constraints $A^T z + c = 0$

- $z > 0$ if and only if

$$\mathbf{1} + \mathbf{diag}(d_x) A \Delta x_{\rm nt} > 0$$

  a sufficient condition is $\lambda_t(x) = \|\mathbf{diag}(d_x) A \Delta x_{\rm nt}\| < 1$

- bound on duality gap follows from Cauchy-Schwarz inequality

Lower bound for centering problem near central path

- substituting the dual z of p.9–8 in the lower bound of p.9–4 gives

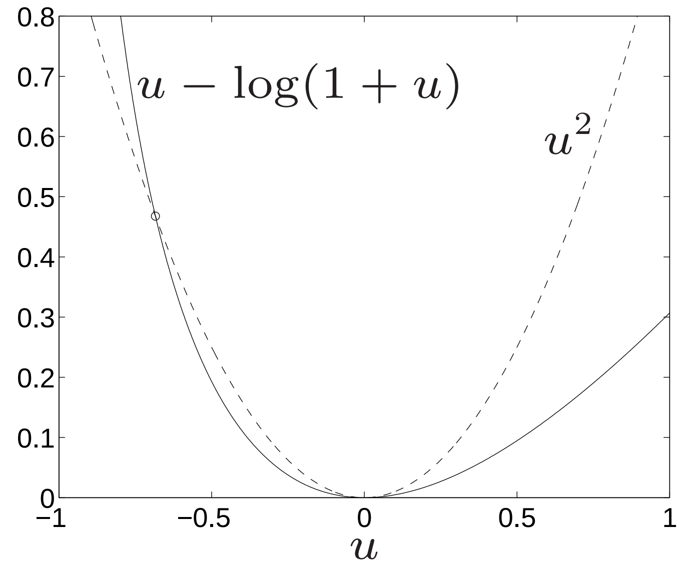
$$f_t(y) \geq f_t(x) - \sum_{i=1}^{m} (w_i - \log(1 + w_i)) \qquad \forall y \in P^\circ$$

where $w_i = (a_i^T \Delta x_{\mathrm{nt}})/(b_i - a_i^T x)$

- this bound holds if $\lambda_t(x) < 1$; a simpler bound holds if $\lambda_t(x) \leq 0.68$:

$$\begin{aligned} f_t(y) &\geq f_t(x) - \sum_{i=1}^{m} w_i^2 \\ &= f_t(x) - \lambda_t(x)^2 \end{aligned}$$

(since $u - \log(1+u) \leq u^2$ for $|u| \leq 0.68$)



Outline

- centering problem
- Newton decrement
- **local convergence of Newton method**
- short-step barrier method
- global convergence of Newton method
- predictor-corrector method

Quadratic convergence of Newton's method

theorem: if $\lambda_t(x) < 1$ and Δx_{nt} is Newton step of f_t at x, then

$$x^+ = x + \Delta x_{\mathrm{nt}} \in P^\circ, \qquad \lambda_t(x^+) \le \lambda_t(x)^2$$

Newton method (with unit stepsize)

$$x^{(k+1)} = x^{(k)} - \nabla^2 f_t(x^{(k)})^{-1} \nabla f_t(x^{(k)})$$

- if $\lambda_t(x^{(0)}) \le 1/2$, Newton decrement after k iterations is

$$\lambda_t(x^{(k)}) \le (1/2)^{2^k}$$

  decreases very quickly: $(1/2)^{2^5} = 2.3 \cdot 10^{-10}$, $(1/2)^{2^6} = 5.4 \cdot 10^{-20}$, $\ldots$

- $\lambda_t(x^{(k)})$ very small after a few iterations *proof of quadratic convergence result*

feasibility of x^+: follows from $\lambda_t(x) = \|\Delta x_{\rm nt}\|_x$ and result on p.8–8

quadratic convergence: define $D = \mathbf{diag}(d_x)$, $D_+ = \mathbf{diag}(d_{x^+})$

$$
\begin{aligned}
\lambda_t(x^+)^2 &= \|D_+A\,\Delta x_{\rm nt}^+\|^2 \\
&\le \|D_+A\,\Delta x_{\rm nt}^+\|^2 + \|(I - D_+^{-1}D)DA\,\Delta x_{\rm nt} + D_+A\,\Delta x_{\rm nt}^+\|^2 \\
&= \|(I - D_+^{-1}D)DA\,\Delta x_{\rm nt}\|^2 \\
&= \|(I - D_+^{-1}D)^2\mathbf{1}\|^2 \\
&\le \|(I - D_+^{-1}D)\mathbf{1}\|^4 \\
&= \|DA\Delta x_{\rm nt}\|^4 \\
&= \lambda_t(x)^4
\end{aligned}
$$

- on lines 4 and 6 we used

$$\begin{array}{rcl} DA\,\Delta x_{\mathrm{nt}} & = & D(b - Ax - b + Ax^{+}) \\ & = & (I - D_{+}^{-1}D)\mathbf{1} \end{array}$$

- line 3 follows from

$$\begin{array}{rl} & A^T D_{+}\left(D_{+}A\,\Delta x_{\mathrm{nt}}^{+} + (I - D_{+}^{-1}D)DA\,\Delta x_{\mathrm{nt}}\right) \\ = & A^T D_{+}^2 A\,\Delta x_{\mathrm{nt}}^{+} - A^T D^2 A\,\Delta x_{\mathrm{nt}} + A^T D_{+} DA\,\Delta x_{\mathrm{nt}} \\ = & -tc - A^T D_{+}\mathbf{1} + tc + A^T D\mathbf{1} + A^T D_{+}(I - D_{+}^{-1}D)\mathbf{1} \\ = & 0 \end{array}$$

- line 5 follows from $(\sum_i y_i^4) \leq (\sum_i y_i^2)^2$

Outline

- centering problem
- Newton decrement
- local convergence of Newton method
- **short-step barrier method**
- global convergence of Newton method
- predictor-corrector method

Short-step barrier method

simplifying assumptions

- a central point $x^\star(t_0)$ is given
- $x^\star(t)$ is computed exactly

algorithm: define tolerance $\epsilon \in (0, 1)$ and parameter

$$\mu = 1 + \frac{1}{2\sqrt{m}}$$

starting at $t = t_0$, repeat until $m/t \leq \epsilon$:

- compute $x^\star(\mu t)$ by Newton's method with unit step started at $x^\star(t)$
- set $t := \mu t$

Newton decrement after update of t

- gradient of f_{t^+} at $x = x^\star(t)$ for new value $t^+ = \mu t$:

$$\nabla f_{t^+}(x) = \mu t c + A^T d_x = -(\mu - 1) A^T d_x$$

- Newton decrement for new value t^+ is

$$\begin{array}{rcl}
\lambda_{t^+}(x) & = & \left(\nabla f_{t^+}(x)^T \nabla^2 \phi(x)^{-1} \nabla f_{t^+}(x)\right)^{1/2} \\
& = & (\mu - 1) \left(\mathbf{1}^T B (B^T B)^{-1} B^T \mathbf{1}\right)^{1/2} \quad (\text{with } B = \mathbf{diag}(d_x) A) \\
& \leq & (\mu - 1)\sqrt{m} \\
& = & 1/2
\end{array}$$

line 3 follows because maximum eigenvalue of $B(B^T B)^{-1} B^T$ is one

$x^\star(t)$ is in region of quadratic convergence of Newton's method for $f_{\mu t}$

Iteration complexity

- Newton iterations per outer iteration: a small constant
- number of outer iterations: we reach $t^{(k)} = \mu^k t_0 \geq m/\epsilon$ when

$$k \geq \frac{\log(m/(\epsilon t_0))}{\log \mu}$$

cumulative number of Newton iterations

$$O\left(\sqrt{m}\log\left(\frac{m}{\epsilon t_0}\right)\right)$$

(we used $\log \mu = \log(1 + 1/(2\sqrt{m})) \geq (\log 2)/(2\sqrt{m})$)

- multiply by flops per Newton iteration to get polynomial complexity
- $\sqrt{m}$ dependence is lowest known complexity for interior-point methods

Outline

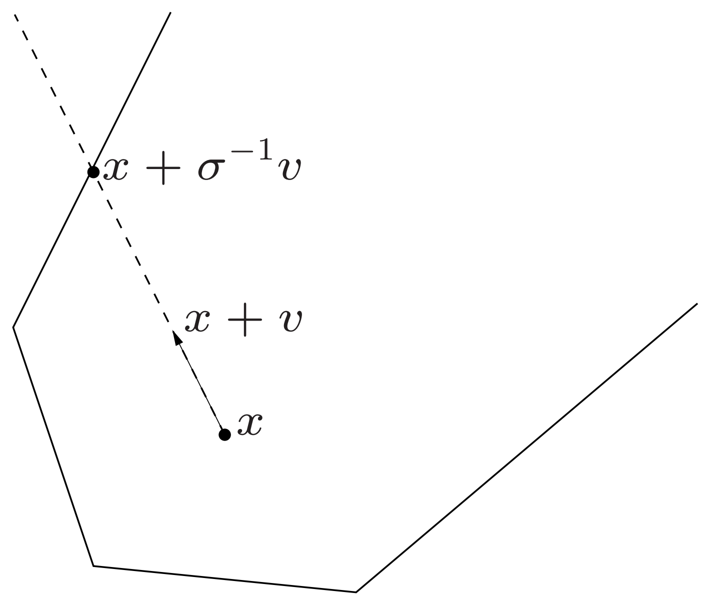
- centering problem
- Newton decrement
- local convergence of Newton method
- short-step barrier method
- **global convergence of Newton method**
- predictor-corrector method

Maximum stepsize to boundary

for $x \in P^\circ$ and arbitrary $v \neq 0$, define

$$\sigma_x(v) = \begin{cases} 0 & \text{if } Av \leq 0 \\ \displaystyle\max_{i=1,\ldots,m} \frac{a_i^T v}{b_i - a_i^T x} & \text{otherwise} \end{cases}$$

point $x + \alpha v$ ($\alpha \geq 0$) is in P if and only if $\alpha \sigma_x(v) \leq 1$



$x + \alpha v \in P$ for

$$\begin{array}{ll} \alpha \in [0, 1/\sigma] & \text{if } \sigma > 0 \\ \alpha \in [0, \infty) & \text{if } \sigma = 0 \end{array}$$

Upper bound on centering cost function

arbitrary direction: for $x \in P^\circ$ and arbitrary $v \neq 0$

- if $\sigma = \sigma_x(v) > 0$ and $\alpha \in [0, 1/\sigma)$:

$$f_t(x + \alpha v) \leq f_t(x) + \alpha \nabla f_t(x)^T v - \frac{\|v\|_x^2}{\sigma^2}(\alpha\sigma + \log(1 - \alpha\sigma))$$

- if $\sigma = \sigma_x(v) = 0$ and $\alpha \in [0, \infty)$:

$$f_t(x + \alpha v) \leq f_t(x) + \alpha \nabla f_t(x)^T v + \frac{\alpha^2}{2}\|v\|_x^2$$

on the right-hand sides, $\|v\|_x = (v^T \nabla^2 f_t(x) v)^{1/2}$

Newton direction: for $v = \Delta x_{\mathrm{nt}}$, substitute $-\nabla f_t(x)^T v = \|v\|_x^2 = \lambda_t(x)^2$

proof: define $w_i = (a_i^T v)/(b_i - a_i^T x)$ and note that $\|w\| = \|v\|_x$

- if $\sigma = \max_i w_i > 0$:

$$
\begin{aligned}
& f_t(x+\alpha v) - f_t(x) - \alpha \nabla f_t(x)^T v \\
&\quad = -\sum_{i=1}^{m} (\alpha w_i + \log(1-\alpha w_i)) \\
&\quad \leq -\sum_{w_i>0} (\alpha w_i + \log(1-\alpha w_i)) + \sum_{w_i \leq 0} \frac{\alpha^2 w_i^2}{2} \qquad (\star) \\
&\quad \leq -\sum_{w_i>0} \frac{w_i^2}{\sigma^2} (\alpha\sigma + \log(1-\alpha\sigma)) + \frac{(\alpha\sigma)^2}{2} \sum_{w_i \leq 0} \frac{w_i^2}{\sigma^2} \\
&\quad \leq -\sum_{i=1}^{m} \frac{w_i^2}{\sigma^2} (\alpha\sigma + \log(1-\alpha\sigma))
\end{aligned}
$$

- if $\sigma = 0$, upper bound follows from $(\star)$

Damped Newton iteration

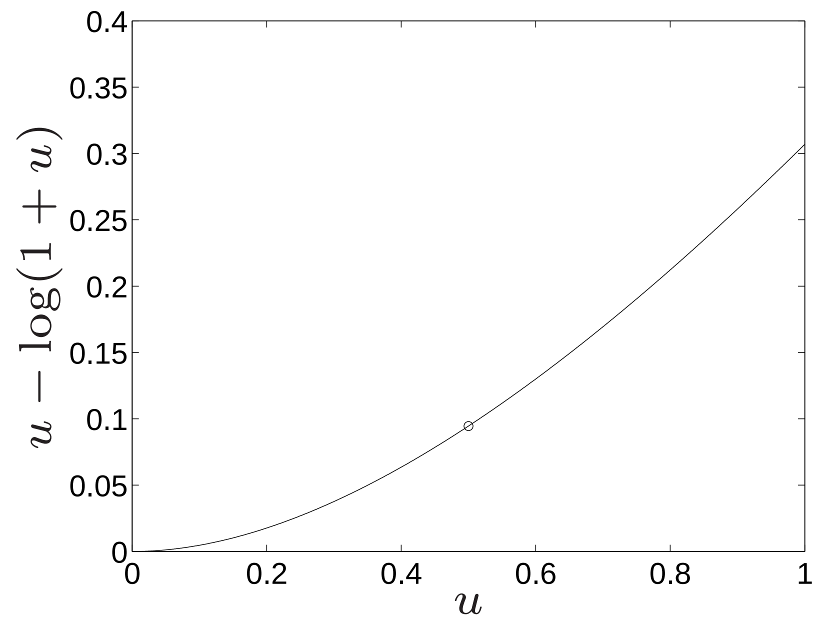
$$x^+ = x + \frac{1}{1 + \sigma_x(\Delta x_{\mathrm{nt}})} \Delta x_{\mathrm{nt}}$$

theorem: damped Newton iteration at any $x \in P^\circ$ decreases cost

$$f_t(x^+) \leq f_t(x) - \lambda_t(x) + \log(1 + \lambda_t(x))$$

- graph shows $u - \log(1 + u)$
- if $\lambda_t(x) \geq 0.5$

$$f_t(x^+) \leq f_t(x) - 0.09$$



proof: apply upper bounds on page 9–21 with $v = \Delta x_{\rm nt}$

- if $\sigma > 0$, the value of the upper bound

$$f_t(x) - \alpha\lambda_t(x)^2 - \frac{\lambda_t(x)^2}{\sigma^2}(\alpha\sigma + \log(1-\alpha\sigma))$$

  at $\alpha = 1/(1+\sigma)$ is

$$f_t(x) - \frac{\lambda_t(x)^2}{\sigma^2}(\sigma - \log(1+\sigma)) \leq f_t(x) - \lambda_t(x) + \log(1+\lambda_t(x))$$

- if $\sigma = 0$, value of upper bound

$$f_t(x) - \alpha\lambda_t(x)^2 + \frac{\alpha^2}{2}\lambda_t(x)^2$$

  at $\alpha = 1$ is

$$f_t(x) - \frac{\lambda_t(x)^2}{2} \leq f_t(x) - \lambda_t(x) + \log(1+\lambda_t(x))$$

Summary: Newton algorithm for centering

centering problem

$$\text{minimize} \quad f_t(x) = tc^T x + \phi(x)$$

algorithm

given: tolerance $\delta \in (0, 1)$, starting point $x := x^{(0)} \in P^\circ$

repeat:

1. compute Newton step Δx_{nt} at x and Newton decrement $\lambda_t(x)$
2. if $\lambda_t(x) \leq \delta$, return x
3. otherwise, set $x := x + \alpha \Delta x_{\text{nt}}$ with

$$\alpha = \begin{cases} \dfrac{1}{1 + \sigma_x(\Delta x_{\text{nt}})} & \text{if } \lambda_t(x) > 1/2 \\ \alpha = 1 & \text{if } \lambda_t(x) \leq 1/2 \end{cases}$$

Convergence

theorem: if $\delta < 1/2$ and $f_t(x)$ is bounded below, algorithm takes at most

$$\frac{f_t(x^{(0)}) - \min_y f_t(y)}{0.09} + \log_2 \log_2(1/\delta) \quad \text{iterations} \tag{1}$$

proof: combine theorems on pages 9–12 and 9–23

- if $\lambda_t(x^{(k)}) > 1/2$, iteration k decreases the function value by at least

$$\lambda_t(x^{(k)}) - \log(1 - \lambda_t(x^{(k)})) \geq 0.09$$

- the first term in (1) bounds the number of iterations with $\lambda_t(x) > 1/2$
- if $\lambda_t(x^{(l)}) \leq 1/2$, quadratic convergence yields $\lambda_t(x^{(k)}) \leq \delta$ after

$$k = l + \log_2 \log_2(1/\delta) \quad \text{iterations}$$

Computable bound on #iterations

replace unknown $\min_y f_t(y)$ in (1) by the lower bound from page 9–4:

$$f_t(x) - \min_y f_t(y) \le V_t(x, z)$$

where z is a strictly dual feasible point and

$$\begin{aligned} V_t(x, z) &= f_t(x) + tb^T z - \sum_{i=1}^m \log z_i - m \log t - m \\ &= t(b - Ax)^T z - \sum_{i=1}^m \log(t(b_i - a_i^T x) z_i) - m \end{aligned}$$

number of Newton iterations to minimize f_t starting at x is bounded by

$$10.6\, V_t(x, z) + \log_2 \log_2(1/\delta)$$

Outline

- centering problem
- Newton decrement
- local convergence of Newton method
- short-step barrier method
- global convergence of Newton method
- **predictor-corrector method**

Predictor-corrector methods

short-step methods

- stay in narrow neighborhood of central path, defined by limit on λ_t
- make small, fixed increases $t^+ = \mu t$
- quite slow in practice

predictor-corrector methods

- select new t using a linear approximation to central path ('predictor')
- recenter with new t ('corrector')
- can make faster and adaptive increases in t

Tangent to central path

central path equation

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

$$s_i^\star(t) z_i^\star(t) = \frac{1}{t}, \quad i = 1, \ldots, m$$

derivatives: $\dot{x} = dx^\star(t)/dt$, $\dot{s} = ds^\star/dt$, $\dot{z} = dz^\star(t)/dt$ satisfy

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

$$s_i^\star(t) \dot{z}_i + z_i^\star(t) \dot{s}_i = -\frac{1}{t^2}, \quad i = 1, \ldots, m$$

tangent direction: defined as $\Delta x_{\mathrm{tg}} = t\dot{x}$, $\Delta s_{\mathrm{tg}} = t\dot{s}$, $\Delta z_{\mathrm{tg}} = t\dot{z}$

Predictor equations

with $x = x^\star(t)$, $s = s^\star(t)$, $z = z^\star(t)$

$$\begin{bmatrix} (1/t)\,\mathbf{diag}(s)^{-2} & 0 & I \\ 0 & 0 & A^T \\ -I & -A & 0 \end{bmatrix} \begin{bmatrix} \Delta s_{\mathrm{tg}} \\ \Delta x_{\mathrm{tg}} \\ \Delta z_{\mathrm{tg}} \end{bmatrix} = \begin{bmatrix} -z \\ 0 \\ 0 \end{bmatrix} \qquad (1)$$

equivalent equation (using $s_i z_i = 1/t$)

$$\begin{bmatrix} I & 0 & (1/t)\,\mathbf{diag}(z)^{-2} \\ 0 & 0 & A^T \\ -I & -A & 0 \end{bmatrix} \begin{bmatrix} \Delta s_{\mathrm{tg}} \\ \Delta x_{\mathrm{tg}} \\ \Delta z_{\mathrm{tg}} \end{bmatrix} = \begin{bmatrix} -s \\ 0 \\ 0 \end{bmatrix} \qquad (2)$$

Properties of tangent direction

- from 2nd and 3rd block in (1): $\Delta s_{\rm tg}^T \Delta z_{\rm tg} = 0$
- take inner product with s on both sides of first block in (1):

$$s^T \Delta z_{\rm tg} + z^T \Delta s_{\rm tg} = -s^T z$$

- hence, gap in tangent direction is

$$(s + \alpha \Delta s_{\rm tg})^T (z + \alpha \Delta z_{\rm tg}) = (1-\alpha) s^T z$$

- take inner product with $\Delta s_{\rm tg}$ on both sides of first block in (1):

$$\Delta s_{\rm tg}^T \mathbf{diag}(s)^{-2} \Delta s_{\rm tg} = -t z^T \Delta s_{\rm tg}$$

- similarly, from first block in (2): $\Delta z_{\rm tg}^T \mathbf{diag}(z)^{-2} \Delta z_{\rm tg} = -t s^T \Delta z_{\rm tg}$

Potential function

definition: for strictly primal, dual feasible x, z, define

$$\begin{aligned}\Psi(x,z) &= m\log\frac{z^Ts}{m} - \sum_{i=1}^m \log(s_iz_i) \quad \text{with } s = b - Ax \\ &= m\log\frac{(\sum_i z_is_i)/m}{(\prod_i z_is_i)^{1/m}} \\ &= m\log\frac{\text{arithmetic mean of } z_1s_1,\ \ldots,\ z_ms_m}{\text{geometric mean of } z_1s_1,\ \ldots,\ s_mz_m}\end{aligned}$$

properties

- $\Psi(x,z)$ is nonnegative for all strictly feasible x, z
- $\Psi(x,z) = 0$ only if x and z are centered, $i.e.$, for some $t > 0$,

$$z_is_i = 1/t, \quad i = 1,\ldots,m$$

Potential function and proximity to central path

for any strictly feasible x, z, with $s = b - Ax$:

$$\begin{aligned}
\Psi(x, z) &= \min_{t>0} V_t(x, z) \qquad \text{(see page 9–27)} \\
&= \min_{t>0} \left(ts^T z - \sum_{i=1}^{m} \log(ts_i z_i) - m \right)
\end{aligned}$$

minimizing t is $t = m/s^T z$

Ψ as global measure of proximity to central path

- $V_t(x,z)$ bounds the effort to compute $x^\star(t)$, starting at x (page 9–27)
- $\Psi(x, z)$ bounds centering effort, without imposing a specific t

Predictor-corrector method with exact centering

simplifying assumptions: exact centering, a central point $x^\star(t_0)$ is given

algorithm: define tolerance $\epsilon \in (0,1)$, parameter $\beta > 0$, and initial values

$$t := t_0, \qquad x := x^\star(t_0), \qquad z := z^\star(t_0), \qquad s := b - Ax^\star(t_0)$$

repeat until $m/t \leq \epsilon$:

- compute tangent direction Δx_{tg}, Δs_{tg}, Δz_{tg} at x, s, z
- determine α by solving $\Psi(x + \alpha\Delta x_{\mathrm{tg}}, z + \alpha\Delta z_{\mathrm{tg}}) = \beta$ and take

$$x := x + \alpha\Delta x_{\mathrm{tg}}, \qquad z := z + \alpha\Delta z_{\mathrm{tg}}, \qquad s := b - Ax$$

- set $t := m/(s^T z)$ and use Newton's method to compute

$$x := x^\star(t), \qquad z := z^\star(t), \qquad s := b - Ax$$

Iteration complexity

- bound on potential function in tangent direction (proof on next page):

$$\Psi(x + \alpha \Delta x_{\mathrm{tg}}, z + \alpha \Delta z_{\mathrm{tg}}) \le -\alpha\sqrt{m} - \log(1 - \alpha\sqrt{m})$$

- lower bound on predictor step length α:

$$\alpha\sqrt{m} \ge \gamma \quad \text{with } \gamma \text{ the solution of } -\gamma - \log(1-\gamma) = \beta$$

- reduction in duality gap after one predictor-corrector cycle:

$$\frac{t}{t^+} = 1 - \alpha \le 1 - \frac{\gamma}{\sqrt{m}} \le \exp(\frac{-\gamma}{\sqrt{m}})$$

- bound on total #Newton iterations to reach $t^{(k)} \ge m/\epsilon$:

$$O\left(\sqrt{m}\log\left(\frac{m}{\epsilon t_0}\right)\right)$$

proof of bound on Ψ: let $s^+ = s + \alpha \Delta s_{\rm tg}$, $z^+ = z + \alpha \Delta z_{\rm tg}$

- from definition of Ψ and $(s^+)^T z^+ = (1-\alpha) s^T z$:

$$\begin{aligned}
\Psi(x^+, z^+) - \Psi(x, z) &= m \log \frac{(s^+)^T z^+}{z^T s} - \sum_{i=1}^m (\log \frac{s_i^+}{s_i} + \log \frac{z_i^+}{z_i}) \\
&= m \log(1-\alpha) - \sum_{i=1}^m (\log \frac{s_i^+}{s_i} + \log \frac{z_i^+}{z_i})
\end{aligned}$$

- define a $(2m)$-vector $w = (\mathbf{diag}(s)^{-1} \Delta s_{\rm tg}, \mathbf{diag}(z)^{-1} \Delta z_{\rm tg})$

$$\begin{aligned}
-\sum_{i=1}^m (\log(s_i^+/s_i) + \log(z_i^+/z_i)) &= -\sum_{i=1}^{2m} \log(1 + \alpha w_i) \\
&\leq -\alpha \mathbf{1}^T w - \alpha \|w\| - \log(1 - \alpha \|w\|)
\end{aligned}$$

last inequality can be proved as on page 9–22

- from the properties on page 9–32 and $sz = 1/t$:

$$\begin{aligned}
\mathbf{1}^T w &= t(s^T \Delta z_{\mathrm{tg}} + z^T \Delta s_{\mathrm{tg}}) \\
&= -ts^T z \\
&= -m \\
\|w\|^2 &= \Delta s_{\mathrm{tg}}^T \,\mathbf{diag}(s)^{-2} \Delta s_{\mathrm{tg}} + \Delta z_{\mathrm{tg}}^T \,\mathbf{diag}(z)^{-2} \Delta z_{\mathrm{tg}} \\
&= -t(z^T \Delta s_{\mathrm{tg}} + s^T \Delta z_{\mathrm{tg}}) \\
&= m
\end{aligned}$$

- substituting this in the upper bound on Ψ gives

$$\begin{aligned}
\Psi(x^+, z^+) - \Psi(x, z) &\leq m \log(1-\alpha) + \alpha m - \alpha\sqrt{m} - \log(1 - \alpha\sqrt{m}) \\
&\leq -\alpha\sqrt{m} - \log(1 - \alpha\sqrt{m})
\end{aligned}$$

Conclusion: barrier methods

started at $x^\star(t_0)$, find ϵ-suboptimal point after

$$O\left(\sqrt{m}\log\left(\frac{m}{\epsilon t_0}\right)\right) \quad \text{Newton iterations}$$

- analysis can be modified to account for inexact centering
- end-to-end complexity analysis must include the cost of phase I
- parameters were chosen to simplify analysis, not for efficiency in practice