

Chapter 11 Interior-point methods

Last update on 2024-05-31 09:01

Table of contents

Inequality constrained minimization problems

Logarithmic barrier function and central path

The barrier method

Complexity analysis via self-concordance

Feasibility and phase I methods

Problems with generalized inequalities

Inequality constrained minimization problems

Logarithmic barrier function and central path

The barrier method

Complexity analysis via self-concordance

Feasibility and phase I methods

Problems with generalized inequalities

inequality constrained minimization

$$\begin{array}{ll}\text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b\end{array}$$

general assumptions

- ▶ f_i convex and twice continuously differentiable
- ▶ $A \in \mathbb{R}^{p \times n}$ and **rank** $A = p$
- ▶ p^* is finite and attained
- ▶ problem is strictly feasible: there exists \tilde{x} with

$$\tilde{x} \in \mathbf{dom} f_0, \quad f_i(\tilde{x}) < 0, \quad i = 1, \dots, m, \quad A\tilde{x} = b$$

hence strong duality holds and dual optimum is attained

examples

- ▶ LP, QP, QCQP, GP
- ▶ entropy maximization with linear inequality constraints ($\mathcal{D} = \mathbb{R}_{++}^n$)

$$\begin{array}{ll}\text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Fx \preceq g \\ & Ax = b\end{array}$$

- ▶ differentiability may require reformulating the problem, e.g. piecewise-linear minimization or ℓ_∞ -norm approximation via LP
- ▶ SDPs and SOCPs are better handled as problems with generalized inequalities

Inequality constrained minimization problems

Logarithmic barrier function and central path

The barrier method

Complexity analysis via self-concordance

Feasibility and phase I methods

Problems with generalized inequalities

reformulation via indicator function

$$\begin{array}{ll}\text{minimize} & f_0(x) + \sum_{i=1}^m I_-(f_i(x)) \\ \text{subject to} & Ax = b\end{array}$$

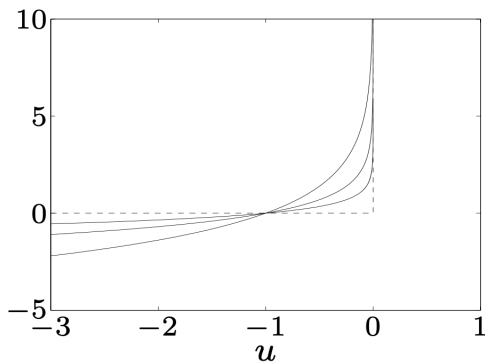
where I_- is the indicator function of \mathbb{R}_-

$$I_-(u) = \begin{cases} 0, & \text{if } u \leq 0 \\ \infty, & \text{if } u > 0 \end{cases}$$

approximation via logarithmic barrier

$$\begin{array}{ll}\text{minimize} & f_0(x) - \frac{1}{t} \sum_{i=1}^m \log(-f_i(x)) \\ \text{subject to} & Ax = b\end{array}$$

- ▶ an equality constrained problem
- ▶ for $t > 0$, the term $-(1/t) \log(-u)$ is a smooth approximation of I_-
- ▶ approximation improves as $t \rightarrow \infty$



- ▶ dashed line: function $I_-(u)$
- ▶ solid curves: function $-(1/t) \log(-u)$ for $t = 0.5, 1, 2$
- ▶ $t = 2$ gives the best approximation

logarithmic barrier function

$$\phi(x) = -\sum_{i=1}^m \log(-f_i(x)), \quad \text{dom } \phi = \{x \mid f_i(x) < 0, \ i = 1, \dots, m\}$$

- ▶ convex function (follows from composition rule)
- ▶ twice continuously differentiable (can be easily computed)

$$\nabla \phi(x) = \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla f_i(x)$$

$$\nabla^2 \phi(x) = \sum_{i=1}^m \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla^2 f_i(x)$$

centering problem

$$\begin{array}{ll}\text{minimize} & tf_0(x) + \phi(x) \\ \text{subject to} & Ax = b\end{array}$$

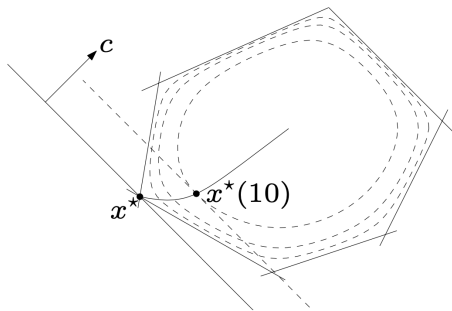
- ▶ assume it has a unique solution $x^*(t)$ for each $t > 0$
- ▶ the curve $\{x^*(t) \mid t > 0\}$ is called the central path
- ▶ there exists some w such that $(x = x^*(t), \nu = w)$ satisfies

$$t\nabla f_0(x) + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla f_i(x) + A^T \nu = 0, \quad Ax = b$$

example central path for an LP

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \dots, 6\end{array}$$

the hyperplane $c^T x = c^T x^*(t)$ is tangent to the level curve of ϕ through $x^*(t)$



dual points from central path

- ▶ by the optimality condition, $x^*(t)$ minimizes the Lagrangian

$$L(x, \lambda^*(t), \nu^*(t)) = f_0(x) + \sum_{i=1}^m \lambda_i^*(t) f_i(x) + \nu^*(t)^T (Ax - b)$$

where we define $\lambda_i^*(t) = 1/(-t f_i(x^*(t)))$ and $\nu^*(t) = w/t$

- ▶ the duality gap for the original problem associated to these values

$$g(\lambda^*(t), \nu^*(t)) = L(x^*(t), \lambda^*(t), \nu^*(t)) = f_0(x^*(t)) - m/t$$

as a consequence

$$f_0(x^*(t)) - p^* \leq m/t$$

which confirms the intuitive idea that $f_0(x^*(t)) \rightarrow p^*$ as $t \rightarrow \infty$

interpretation via KKT conditions

$x = x^*(t)$, $\lambda = \lambda^*(t)$, $\nu = \nu^*(t)$ satisfy

1. primal constraints $f_i(x) \leq 0$, $i = 1, \dots, m$, $Ax = b$
2. dual constraints $\lambda \succeq 0$
3. approximate complementary slackness $-\lambda_i f_i(x) = 1/t$, $i = 1, \dots, m$
4. gradient of Lagrangian with respect to x vanishes

$$\nabla f_0(x) + \sum_{i=1}^m \lambda_i \nabla f_i(x) + A^T \nu = 0$$

the only difference with KKT is that condition 3 replaces $\lambda_i f_i(x) = 0$

centering problem without equality constraints

$$\text{minimize} \quad tf_0(x) - \sum_{i=1}^m \log(-f_i(x))$$

force field interpretation

- ▶ $tf_0(x)$ is potential of force field $F_0(x) = -t\nabla f_0(x)$
- ▶ $-\log(-f_i(x))$ is potential of force field $F_i(x) = (1/f_i(x)) \nabla f_i(x)$

the forces balance at $x^*(t)$

$$F_0(x^*(t)) + \sum_{i=1}^m F_i(x^*(t)) = 0$$

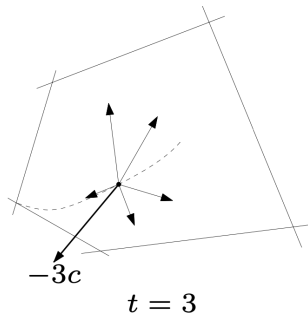
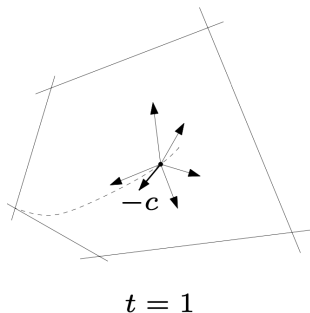
example

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \dots, m\end{array}$$

- ▶ objective force field is constant $F_0(x) = -tc$
- ▶ constraint force decays as inverse distance to constraint hyperplane

$$F_i(x) = \frac{-a_i}{b_i - a_i^T x}, \quad \|F_i(x)\|_2 = \frac{1}{\text{dist}(x, \mathcal{H}_i)}$$

where $\mathcal{H}_i = \{x \mid a_i^T x = b_i\}$



- ▶ a small LP example with $n = 2$ and $m = 5$
- ▶ the equilibrium position of the particle traces out the central path
- ▶ larger value of objective force moves the particle closer to the optimal point

Inequality constrained minimization problems

Logarithmic barrier function and central path

The barrier method

Complexity analysis via self-concordance

Feasibility and phase I methods

Problems with generalized inequalities

given strictly feasible x , $t := t^{(0)} > 0$, $\mu > 1$, tolerance $\epsilon > 0$

repeat

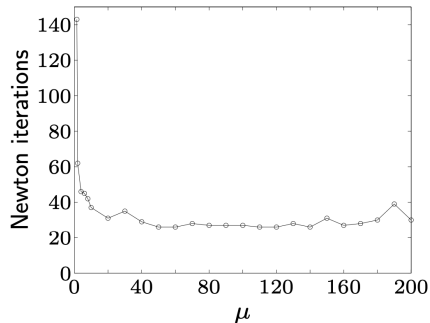
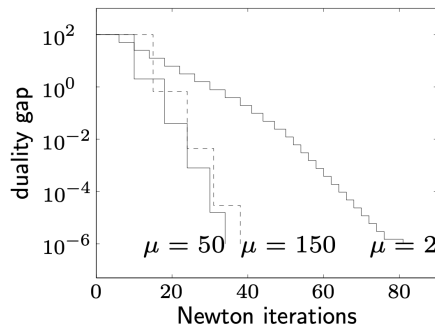
1. *Centering step.* Compute $x^*(t)$ by minimizing $tf_0 + \phi$ subject to $Ax = b$
 2. *Update.* $x := x^*(t)$
 3. *Stopping criterion.* **quit** if $m/t < \epsilon$
 4. *Increase t .* $t := \mu t$
-

remarks

- ▶ terminates with $f_0(x) - p^* \leq \epsilon$
- ▶ centering usually done using Newton's method, starting at current x
- ▶ choice of μ involves a trade-off: larger μ means fewer outer (centering) iterations and more inner (Newton) iterations; typical values $10 \leq \mu \leq 20$
- ▶ several heuristics for choice of $t^{(0)}$

Examples

inequality form LP ($m = 100$ inequalities, $n = 50$ variables)



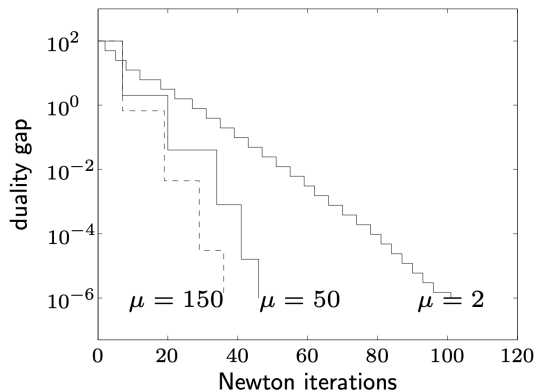
- ▶ starts with x on central path ($t^{(0)} = 1$, duality gap 100)
- ▶ terminates when $t = 10^8$ (gap 10^{-6})
- ▶ centering uses Newton's method with backtracking
- ▶ total number of Newton iterations not very sensitive for $\mu \geq 10$

geometric program

($m = 100$ inequalities and $n = 50$ variables)

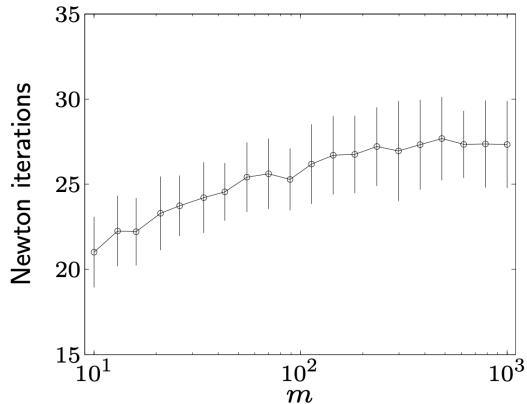
minimize $\log \left(\sum_{k=1}^5 \exp (a_{0k}^T x + b_{0k}) \right)$

subject to $\log \left(\sum_{k=1}^5 \exp (a_{ik}^T x + b_{ik}) \right) \leq 0, \quad i = 1, \dots, m$



family of standard LPs $(A \in \mathbb{R}^{m \times 2m})$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \succeq 0 \end{array}$$



- ▶ solve 100 randomly generated instances for each m between 10 and 1000
- ▶ number of iterations grows very slowly as m ranges over a 100 : 1 ratio

Convergence analysis

outer (centering) iterations number is exactly

$$\left\lceil \frac{\log(m/\epsilon t^{(0)})}{\log \mu} \right\rceil$$

plus the initial centering step for computing $x^*(t^{(0)})$

inner (Newton) iterations

$$\text{minimize} \quad t f_0(x) + \phi(x)$$

see convergence analysis of Newton's method

- ▶ $t f_0 + \phi$ must have closed sublevel sets for $t \geq t^{(0)}$
- ▶ classical analysis requires strong convexity and Lipschitz condition
- ▶ analysis via self-concordance requires self-concordance of $t f_0 + \phi$

Inequality constrained minimization problems

Logarithmic barrier function and central path

The barrier method

Complexity analysis via self-concordance

Feasibility and phase I methods

Problems with generalized inequalities

Self-concordance assumptions

same general assumptions in this chapter plus

- ▶ sublevel sets (of f_0 on the feasible set) are bounded
- ▶ $tf_0 + \phi$ is self-concordant with closed sublevel sets for all $t \geq t^{(0)}$

the second condition above

- ▶ holds for LP, QP, QCQP
- ▶ may require reformulating the problem, e.g.

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Fx \preceq g \end{array} \quad \implies \quad \begin{array}{ll} \text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Fx \preceq g, \quad x \succeq 0 \end{array}$$

- ▶ needed for complexity analysis; barrier method works even when self-concordance assumption does not apply

Newton iterations per centering step

general result for closed strictly convex self-concordant function f

$$\# \text{ Newton iterations} \leq \frac{f(x) - p^*}{\gamma} + c$$

where γ and c are constants depending only on Newton algorithm parameters

barrier method effort of computing $x^+ = x^*(\mu t)$ starting at $x = x^*(t)$

$$\# \text{ Newton iterations} \leq \frac{\mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+)}{\gamma} + c$$

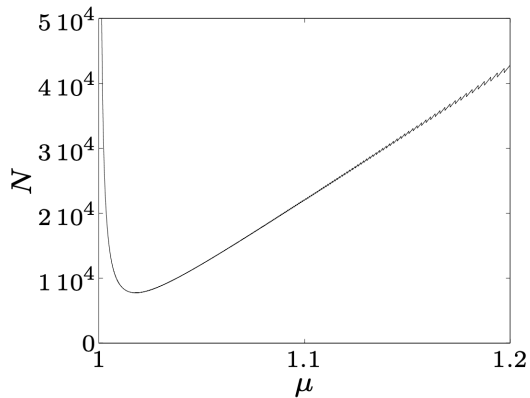
deriving an upper bound with $\lambda = \lambda^*(t)$ and $\nu = \nu^*(t)$

$$\begin{aligned}
 & \mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+) \\
 = & \mu t f_0(x) - \mu t f_0(x^+) + \sum_{i=1}^m \log(-\mu t \lambda_i f_i(x^+)) - m \log \mu \\
 \leq & \mu t f_0(x) - \mu t f_0(x^+) - \mu \sum_{i=1}^m \lambda f_i(x^+) - m - m \log \mu \\
 \leq & \mu t f_0(x) - \mu t g(\lambda, \nu) - m - m \log \mu \\
 = & m(\mu - 1 - \log \mu)
 \end{aligned}$$

Total number of Newton iterations

total number of Newton steps in barrier method excluding initial centering step

$$\# \text{ Newton iterations} \leq N = \left\lceil \frac{\log(m/\epsilon t^{(0)})}{\log \mu} \right\rceil \left(\frac{m(\mu - 1 - \log \mu)}{\gamma} + c \right)$$



- ▶ figure shows N for typical values of γ , c , $m = 100$, $m/\epsilon t^{(0)} = 10^5$
- ▶ confirms trade-off in choice of μ
- ▶ in practice, number of iterations is in the tens; not very sensitive for $\mu \geq 10$

polynomial-time complexity of barrier method

- ▶ we choose $\mu = 1 + 1/\sqrt{m}$, which approximately optimizes worst-case complexity
- ▶ for such μ simple calculation shows $N = O(\sqrt{m} \log(m/\epsilon t^{(0)}))$
- ▶ number of Newton iterations for fixed gap reduction is $O(\sqrt{m})$
- ▶ multiply with cost of one Newton iteration (a polynomial function of problem dimensions) to get bound on number of flops
- ▶ in practice we choose μ fixed (between 10 and 20)

Inequality constrained minimization problems

Logarithmic barrier function and central path

The barrier method

Complexity analysis via self-concordance

Feasibility and phase I methods

Problems with generalized inequalities

feasibility problem find x such that

$$\begin{aligned} f_i(x) &\leq 0, & i = 1, \dots, m \\ Ax &= b \end{aligned}$$

phase I computes strictly feasible starting point for barrier method

Basic phase I method

basic phase I method (with optimal value \bar{p}^*)

$$\begin{array}{ll}\text{minimize} & s \\ \text{subject to} & f_i(x) \leq s, \quad i = 1, \dots, m \\ & Ax = b\end{array}$$

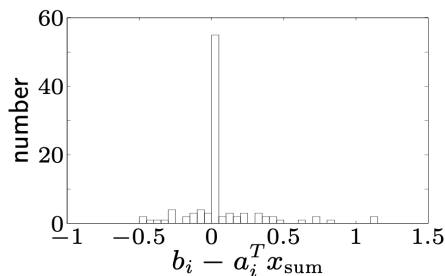
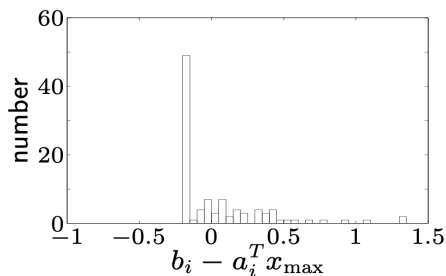
- ▶ if (x, s) feasible with $s < 0$, then x is strictly feasible for feasibility problem
- ▶ if $\bar{p}^* > 0$, then feasibility problem is infeasible
- ▶ if $\bar{p}^* = 0$ and not attained, then feasibility problem is infeasible
- ▶ if $\bar{p}^* = 0$ and attained, then feasibility problem is feasible, but not strictly

sum of infeasibilities phase I method

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T s \\ \text{subject to} & f_i(x) \leq s_i, \quad i = 1, \dots, m \\ & Ax = b \\ & s \succeq 0 \end{array}$$

comparison of methods

infeasible set of 100 linear inequalities in 50 variables



- ▶ left: basic phase I solution; satisfies 39 inequalities
- ▶ right: sum of infeasibilities phase I solution; satisfies 79 inequalities
- ▶ for infeasible problems, second method produces a solution that satisfies many more inequalities than first method

Example

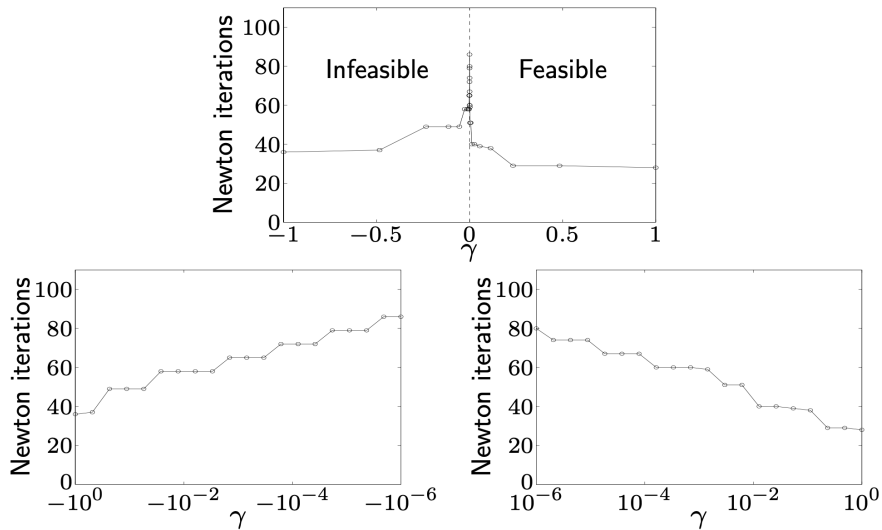
family of linear feasibility problems

$$Ax \preceq b + \gamma \Delta b$$

- ▶ data chosen to be strictly feasible for $\gamma > 0$, infeasible for $\gamma < 0$, feasible but not strictly feasible for $\gamma = 0$
- ▶ use basic phase I method, terminate when $s < 0$ (find a strictly feasible point) or when dual objective > 0 (produce a certificate of infeasibility)

conclusion

- ▶ cost of solving a convex feasibility problem using barrier method is modest when the problem is not close to the boundary between feasibility and infeasibility
- ▶ cost grows when the problem is very close to the boundary
- ▶ cost becomes infinite when the problem is exactly on the boundary



number of iterations roughly proportional to $\log(1/|\gamma|)$

Inequality constrained minimization problems

Logarithmic barrier function and central path

The barrier method

Complexity analysis via self-concordance

Feasibility and phase I methods

Problems with generalized inequalities

minimization with generalized inequalities

$$\begin{array}{ll}\text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & Ax = b\end{array}$$

assumptions

- ▶ f_0 convex function
- ▶ $f_i: \mathbb{R}^n \rightarrow \mathbb{R}^{k_i}$ convex with respect to proper cones $K_i \subset \mathbb{R}^{k_i}$ for $i = 1, \dots, m$
- ▶ all f_i twice continuously differentiable
- ▶ $A \in \mathbb{R}^{p \times n}$ with **rank** $A = p$
- ▶ p^* is finite and attained
- ▶ problem is strictly feasible, hence strong duality holds and dual optimum is attained

examples of greatest interest SOCP, SDP

generalized logarithm for a proper cone

function $\psi: \mathbb{R}^q \rightarrow \mathbb{R}$ is a generalized logarithm for a proper cone $K \subseteq \mathbb{R}^q$ if

1. $\text{dom } \psi = \text{int } K$
2. ψ is concave, closed, twice continuously differentiable
3. $\nabla^2 \psi(y) \prec 0$ for $y \succ_K 0$
4. there exists a constant $\theta > 0$ (degree of ψ) such that for $y \succ_K 0$ and $s > 0$

$$\psi(sy) = \psi(y) + \theta \log s$$

properties $\nabla \psi(y) \succeq_{K^*} 0$ and $y^T \nabla \psi(y) = \theta$ for any $y \succ_K 0$

examples

- ▶ nonnegative orthant $K = \mathbb{R}_+^n$

$$\psi(y) = \sum_{i=1}^n \log y_i, \quad (\theta = n)$$

$$\nabla \psi(y) = (1/y_1, \dots, 1/y_n), \quad y^T \nabla \psi(y) = n$$

- ▶ positive semidefinite cone $K = \mathbb{S}_+^n$

$$\psi(Y) = \log \det Y, \quad (\theta = n)$$

$$\nabla \psi(Y) = Y^{-1}, \quad \text{tr}(Y \nabla \psi(Y)) = n$$

► second-order cone $K = \{y \in \mathbb{R}^{n+1} \mid (y_1^2 + \cdots + y_n^2)^{1/2} \leq y_{n+1}\}$

$$\psi(y) = \log(y_{n+1}^2 - y_1^2 - \cdots - y_n^2), \quad (\theta = 2)$$

$$\nabla\psi(y) = \frac{2}{y_{n+1}^2 - y_1^2 - \cdots - y_n^2} \begin{bmatrix} -y_1 \\ \vdots \\ -y_n \\ y_{n+1} \end{bmatrix}, \quad y^T \nabla\psi(y) = 2$$

logarithmic barrier function for $f_1(x) \preceq_{K_1} 0, \dots, f_m(x) \preceq_{K_m} 0$

$$\phi(x) = - \sum_{i=1}^m \psi_i(-f_i(x)),$$

$$\mathbf{dom} \phi = \{x \mid f_i(x) \prec_{K_i} 0, \ i = 1, \dots, m\}$$

- ▶ ψ_i is generalized logarithm for K_i with degree θ_i
- ▶ ϕ is convex and twice continuously differentiable

central path

- $\{x^*(t) \mid t > 0\}$ where $x^*(t)$ solves

$$\begin{array}{ll} \text{minimize} & t f_0(x) + \phi(x) \\ \text{subject to} & Ax = b \end{array}$$

- $x = x^*(t)$ if there exists $w \in \mathbb{R}^p$ such that

$$t \nabla f_0(x) + \sum_{i=1}^m Df_i(x)^T \nabla \psi_i(-f_i(x)) + A^T w = 0$$

where $Df_i(x) \in \mathbb{R}^{k_i \times n}$ is derivative (Jacobian) matrix of f_i at x

dual points on central path

- ▶ $x^*(t)$ minimizes Lagrangian $L(x, \lambda^*(t), \nu^*(t))$, where

$$\lambda_i^*(t) = \frac{1}{t} \nabla \psi_i(-f_i(x^*(t))), \quad \nu^*(t) = \frac{w}{t}$$

- ▶ $\lambda_i^*(t) \succ_{K_i^*} 0$ from properties of ψ_i , therefore duality gap

$$f_0(x^*(t)) - g(\lambda^*(t), \nu^*(t)) = \frac{1}{t} \sum_{i=1}^m \theta_i$$

example SDP with $F_i, G \in \mathbb{S}^p$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & F(x) = \sum_{i=1}^n x_i F_i + G \preceq 0 \end{array}$$

- ▶ logarithmic barrier: $\phi(x) = \log \det(-F(x)^{-1})$
- ▶ central path: $x^*(t)$ minimizes $tc^T x - \log \det(-F(x))$, hence

$$tc_i - \text{tr}(F_i F(x^*(t))^{-1}) = 0, \quad i = 1, \dots, n$$

- ▶ dual point on central path: $Z^*(t) = -(1/t)F(x^*(t))^{-1}$ is feasible for

$$\begin{array}{ll} \text{maximize} & \text{tr}(GZ) \\ \text{subject to} & \text{tr}(F_i Z) + c_i = 0, \quad i = 1, \dots, n \\ & Z \succeq 0 \end{array}$$

- ▶ duality gap on central path: $c^T x^*(t) - \text{tr}(GZ^*(t)) = p/t$

given strictly feasible x , $t := t^{(0)} > 0$, $\mu > 1$, tolerance $\epsilon > 0$

repeat

1. *Centering step.* Compute $x^*(t)$ by minimizing $tf_0 + \phi$ subject to $Ax = b$
 2. *Update.* $x := x^*(t)$
 3. *Stopping criterion.* **quit** if $(\sum_i \theta_i) / t < \epsilon$
 4. *Increase t .* $t := \mu t$
-

remarks

- ▶ only difference is duality gap m/t on central path is replaced by $\sum_i \theta_i/t$
- ▶ number of outer iterations

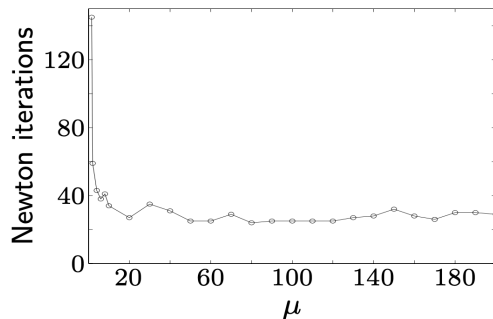
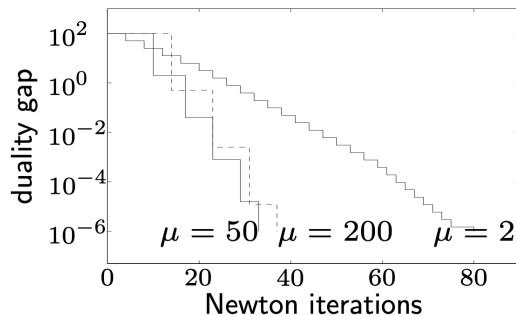
$$\left\lceil \frac{\log((\sum_i \theta_i) / (\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

- ▶ complexity analysis via self-concordance applies to SDP and SOCP

Examples

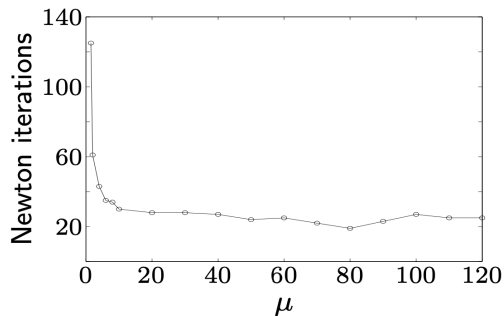
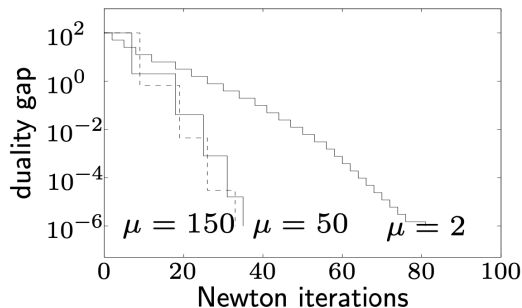
SOCP (50 variables, 50 SOC constraints in \mathbb{R}^6)

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \dots, m \end{array}$$



SDP (100 variables, LMI constraints in \mathbb{S}^{100})

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & \sum_{i=1}^n x_i F_i + G \preceq 0\end{array}$$



family of SDPs

$$(A \in \mathbb{S}^n, x \in \mathbb{R}^n)$$

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T x \\ \text{subject to} & A + \mathbf{diag}(x) \succeq 0 \end{array}$$

solve 100 randomly generated instances for each n between 10 and 1000

