Lecture 9: Interior-Point Methods

Xiugang Wu

University of Delaware

Fall 2019

Outline

- Inequality Constrained Minimization
- Logarithmic Barrier Function and Central Path
- Barrier Method
- Generalized Inequalities

Outline

- Inequality Constrained Minimization
- Logarithmic Barrier Function and Central Path
- Barrier Method
- Generalized Inequalities

Inequality Constrained Minimization

$$\begin{aligned} \text{minimize } & f_0(x) \\ \text{subject to } & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & Ax = b \end{aligned}$$

- f_i convex, twice continuously differentiable

- p^* is finite and attained

- $A \in \mathbf{R}^{p \times n}$ with rank p

- problem is strictly feasible: there exists $\tilde{x}$ with

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

hence strong duality holds and dual optimum is attained

Examples

- LP, QP, QCQP, GP

- entropy maximization with linear inequality constraints

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

- differentiability may require reformulating the problem, e.g., piecewise-linear minimization via LP

- SDP and SOCP are better handled as problems with generalized inequalities

Outline

- Inequality Constrained Minimization
- Logarithmic Barrier Function and Central Path
- Barrier Method
- Generalized Inequalities

Logarithmic Barrier

reformulation via indicator function:

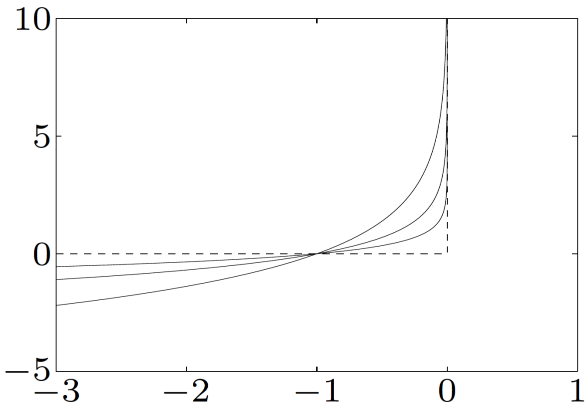
$$\begin{array}{ll} \text{minimize} & f_0(x) + \sum_{i=1}^{m} I_-(f_i(x)) \\ \text{subject to} & Ax = b \end{array}$$

where $I_-(u)$ equals 0 if $u \leq 0$, and ∞ otherwise

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 $t > 0$, $-(1/t)\log(-u)$ is a smooth approximation of I_-
- approximation improves as $t \to \infty$



Logarithmic Barrier

logarithmic barrier function:

$$\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x)), \quad \mathbf{dom}\ \phi = \{x | f_i(x) < 0, i = 1, \ldots, m\}$$

- convex (follows from composition rules)

- twice continuously differentiable, with derivative

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

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

Central Path

- for $t > 0$ define $x^*(t)$ as the solution of

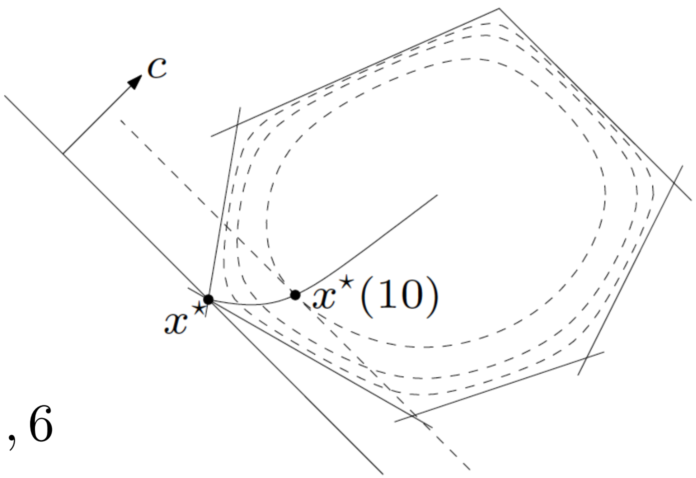
$$\begin{array}{ll} \text{minimize} & t f_0(x) + \phi(x) \\ \text{subject to} & Ax = b \end{array}$$

(for now assume $x^*(t)$ exists and is unique for each $t > 0$)

- central path is $\{x^*(t) | t > 0\}$

example: central path for an LP

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



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

Dual Points on Central Path

$x = x^*(t)$ if there exists a w such that

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

- therefore $x^*(t)$ minimizes the Lagrangian

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

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

- this confirms that $f_0(x^*(t)) \to p^*$ as $t \to \infty$

$$\begin{aligned} p^* &\geq g(\lambda^*(t), \nu^*(t)) \\ &= L(x^*(t), \lambda^*(t), \nu^*(t)) \\ &= f_0(x^*(t)) - m/t \end{aligned}$$

Interpretation Via KKT Conditions

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

- primal constraints: $f_i(x) \leq 0, i = 1, \ldots, m,\ Ax = b$

- dual constraints: $\lambda \succeq 0$

- approximate complementary slackness: $-\lambda_i f_i(x) = 1/t, i = 1, \ldots, m$

- gradient of Lagrangian w.r.t. x vanishes:

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

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

Outline

- Inequality Constrained Minimization
- Logarithmic Barrier Function and Central Path
- Barrier Method
- Generalized Inequalities

Barrier Method

---

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

repeat

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

- terminates with $f_0(x) - p^* \leq \epsilon$; stopping criterion follows from $f_0(x^*(t)) - p^* \leq m/t$

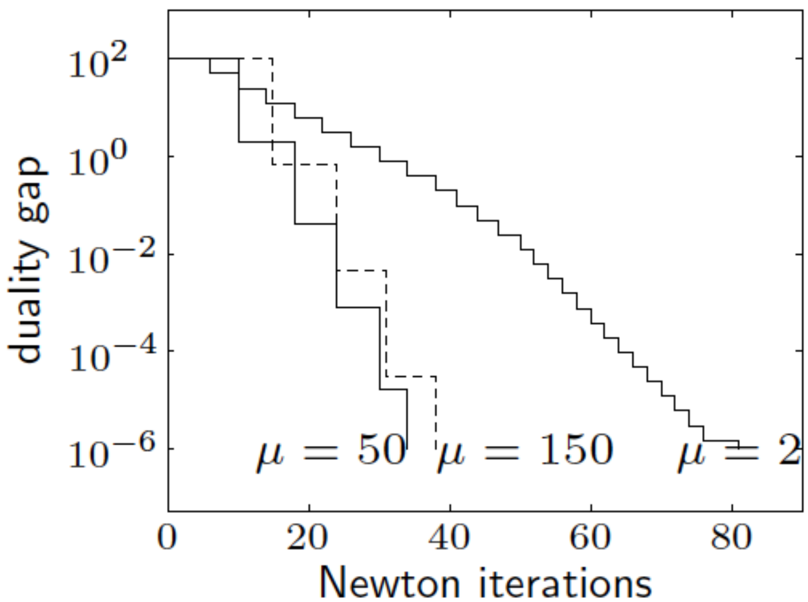
- centering usually done using Newton's method, starting at current x

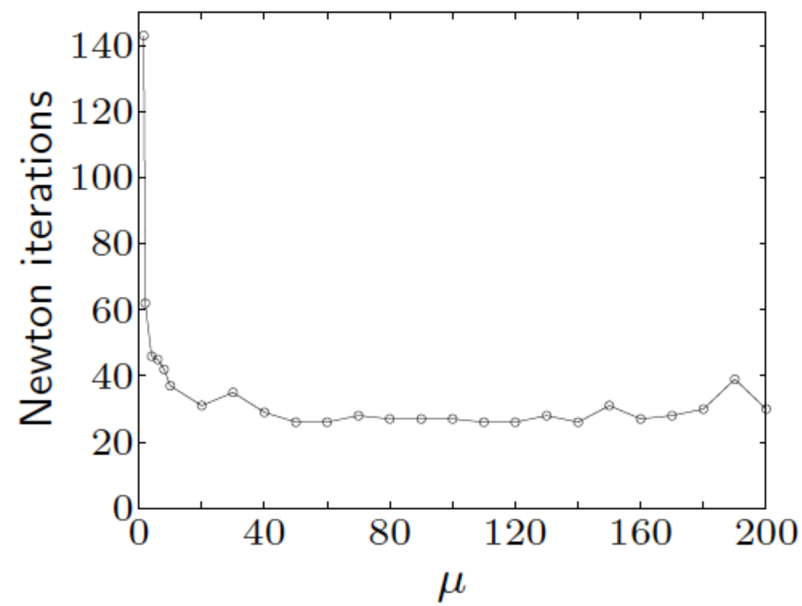
- choice of μ involves a trade-off: large μ means fewer outer iterations, more inner (Newton) iterations; typical value: $\mu = 15 - 20$

- several heuristics for choice of $t^{(0)}$

- number of outer (centering) iterations: exactly $\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$ plus the initial centering step (to compute $x^*(t^{(0)})$)

Examples

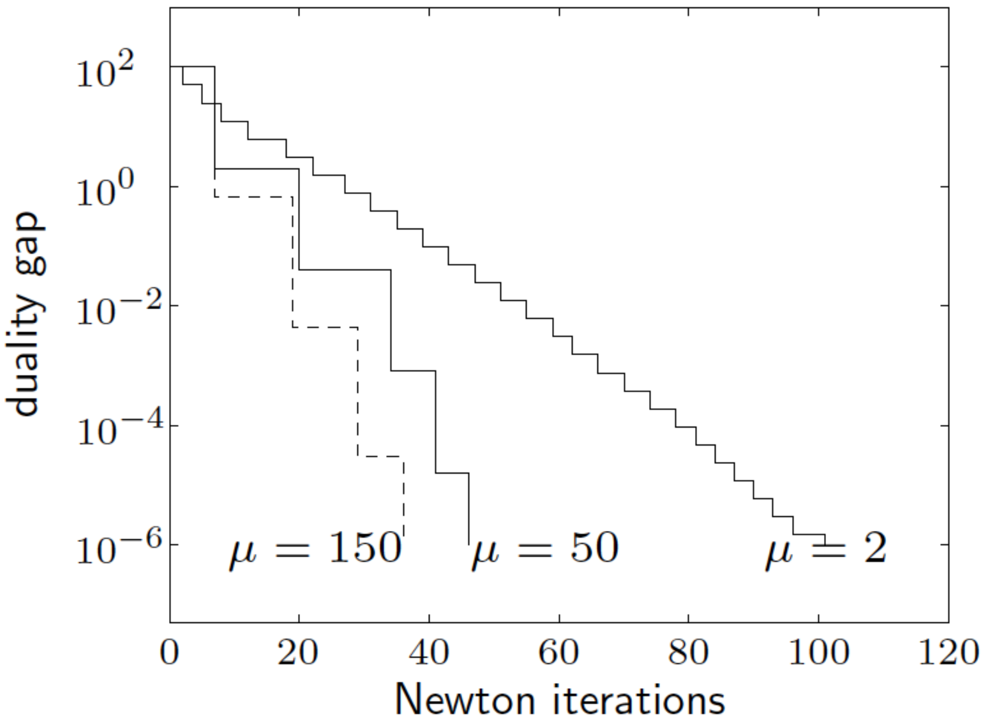




inequality form LP ($m = 100$ inequalities, $n = 50$ variables)
- starts with x on central path ($t^{(0)} = 1$, duality gap 100)
- terminate when $t = 10^8$ (gap 10^{-6})
- centering uses Newton's method with backtracking
- total number of Newton iterations not very sensitive for $\mu \geq 10$

Examples

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



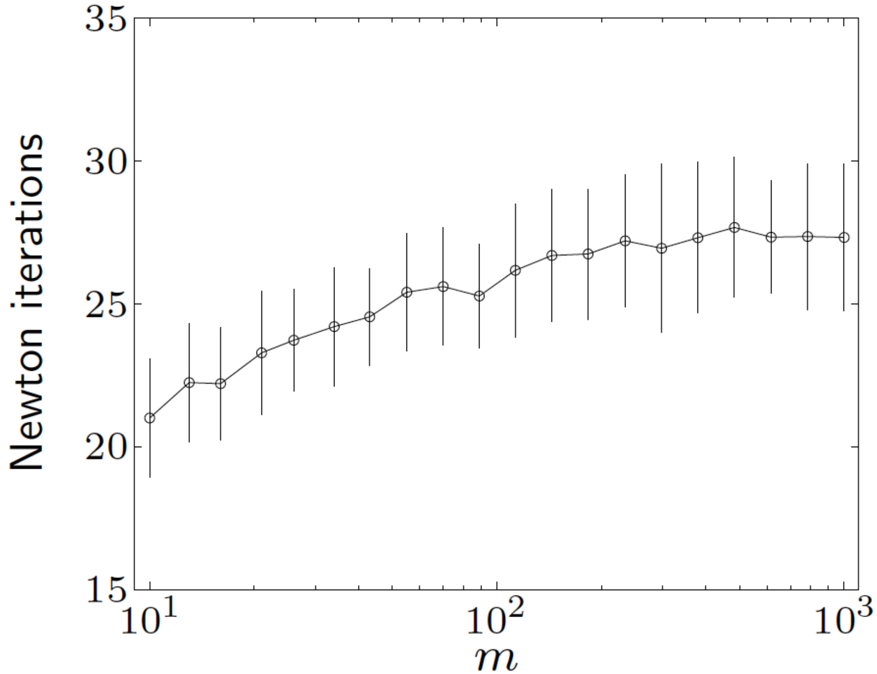
Examples

family of standard LP's ($A \in \mathbf{R}^{m \times 2m}$)

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

$m = 10, \ldots, 1000$; for each m solve 100 randomly generated instances

number of iterations grows very slowly as m ranges over a $100 : 1$ ratio



Outline

- Inequality Constrained Minimization
- Logarithmic Barrier Function and Central Path
- Barrier Method
- Generalized Inequalities

Generalized Inequalities

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

- f_0 convex, $f_i : \mathbf{R}^n \to \mathbf{R}^{k_i}, i = 1, \ldots, m$, convex with respect to proper cones $K_i \in \mathbf{R}^{k_i}$

- f_i twice continuously differentiable

- p^* is finite and attained

- $A \in \mathbf{R}^{p \times n}$ with rank p

- problem is strictly feasible; hence strong duality holds and dual optimum is attained

- examples of greatest interest: SOCP, SDP

Generalized Logarithm for Proper Cone

$\psi : \mathbf{R}^q \to \mathbf{R}$ is generalized logarithm for proper cone $K \subseteq \mathbf{R}^q$ if:

- $\mathbf{dom}\psi = \mathbf{int}K$ and $\nabla^2\psi(y) \prec 0$ for $y \succ_K 0$

- $\psi(sy) = \psi(y) + \theta \log s$ for $y \succ_K 0$, $s > 0$; θ is the degree of ψ

Examples:

- nonnegative orthant $K = \mathbf{R}^n_+$: $\psi(y) = \sum_{i=1}^n \log y_i$, with degree $\theta = n$

- positive semidefinite cone $K = \mathbf{S}^n_+$:

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

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

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

Generalized Logarithm for Proper Cone

properties (without proof): for $y \succ_K 0$,

$$\nabla\psi(y) \succeq_{K^*} 0, \quad y^T\nabla\psi(y) = \theta$$

- nonnegative orthant $\mathbf{R}^n_+$: $\psi(y) = \sum_{i=1}^n \log y_i$

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

- positive semidefinite cone $\mathbf{S}^n_+$: $\psi(Y) = \log\det Y$

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

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

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

and

$$\nabla\psi(y) = \frac{2}{y_{n+1}^2 - y_1^2 - \cdots - y_n^2}(-y_1, -y_2, \ldots, -y_n, y_{n+1}), \quad y^T\nabla\psi(y) = 2$$

Logarithmic Barrier and Central Path

logarithmic barrier for $f_i(x) \preceq_{K_i} 0$, $i = 1, \ldots, m$:

$$\phi(x) = -\sum_{i=1}^{m} \psi_i(-f_i(x)), \quad \mathrm{dom}\phi = \{x | f_i(x) \prec_{K_i} 0, i = 1, \ldots, m\}$$

- ψ_i is generalized logarithm for K_i, with degree θ_i

- $\phi(x)$ is convex, twice continuously differentiable

central path: $\{x^*(t) | t > 0\}$, where $x^*(t)$ solves

$$\text{minimize } t f_0(x) + \phi(x) \;\; \text{subject to } Ax = b$$

Dual Points on Central Path

$x = x^*(t)$ if there exists $w \in \mathbf{R}^p$ 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 $Df_i(x) \in \mathbf{R}^{k_i \times n}$ is derivative matrix of f_i

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

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

- from properties of ψ_i: $\lambda_i^*(t) \succ_{K^*} 0$ and $-f_i(x)^T\lambda_i^*(t) = \theta_i$, we have duality gap

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

and therefore

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

Barrier Method

---

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

repeat

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

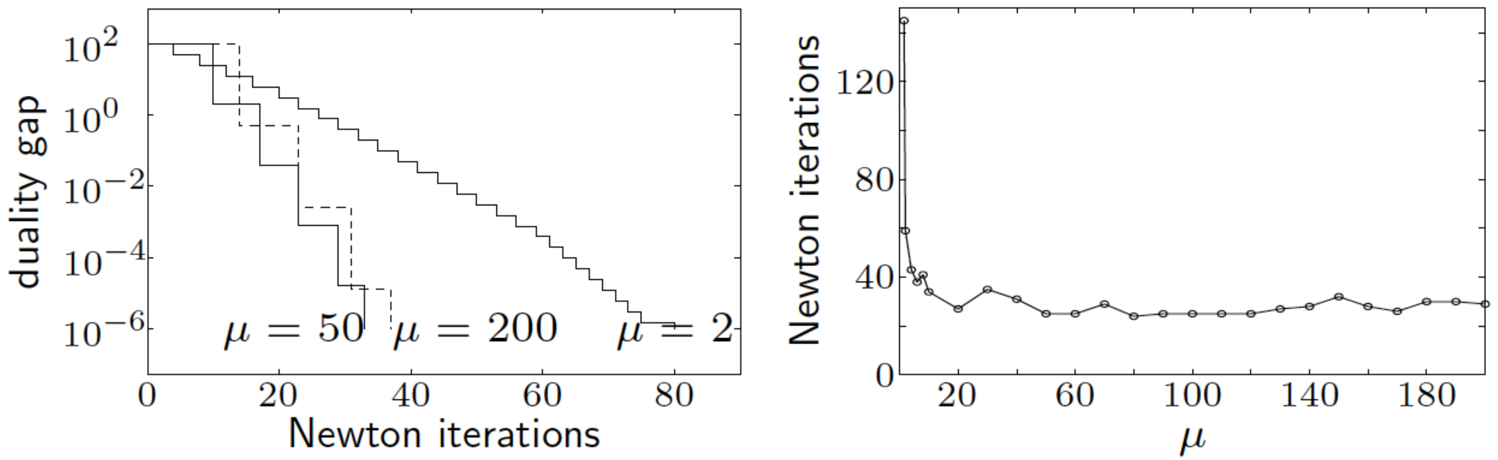
- only difference is that duality gap m/t on central path is replaced by $\sum_{i=1}^m \theta_i/t$

- number of outer iterations:

$$\left\lceil \frac{\log(\sum_{i=1}^m \theta_i/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

Examples

second-order cone program (50 variables, 50 SOC constraints in $\mathbf{R}^6$)



semidefinite program (100 variables, LMI constraint in $\mathbf{S}^{100}$)