Interior Point Algorithms for Constrained Convex Optimization

Chee Wei Tan

Convex Optimization and its Applications to Computer Science

Outline

- Inequality constrained minimization problems
- Barrier function and central path
- Barrier method

Inequality Constrained Minimization

Let $f_0, \ldots, f_m : \mathbf{R}^n \to \mathbf{R}$ be convex and twice continuously differentiable and $A \in \mathbf{R}^{p \times n}$ with rank $p < n$:

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

Assume the problem is strictly feasible and an optimal x^* exists

Idea: reduce it to a sequence of linear equality constrained problems and apply Newton's method

First, need to approximately formulate inequality constrained problem as an equality constrained problem

Barrier Function

Make inequality constraints implicit in the objective:

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

where I_- is indicator function:

$$I_-(u) = \begin{cases} 0 & u \le 0 \\ \infty & u > 0 \end{cases}$$

No inequality constraints, but objective function not differentiable

Approximate indicator function by a differentiable, closed, and convex function:

$$\hat{I}_-(u) = -(1/t)\log(-u), \quad \mathbf{dom}\,\hat{I}_- = -\mathbf{R}_{++}$$

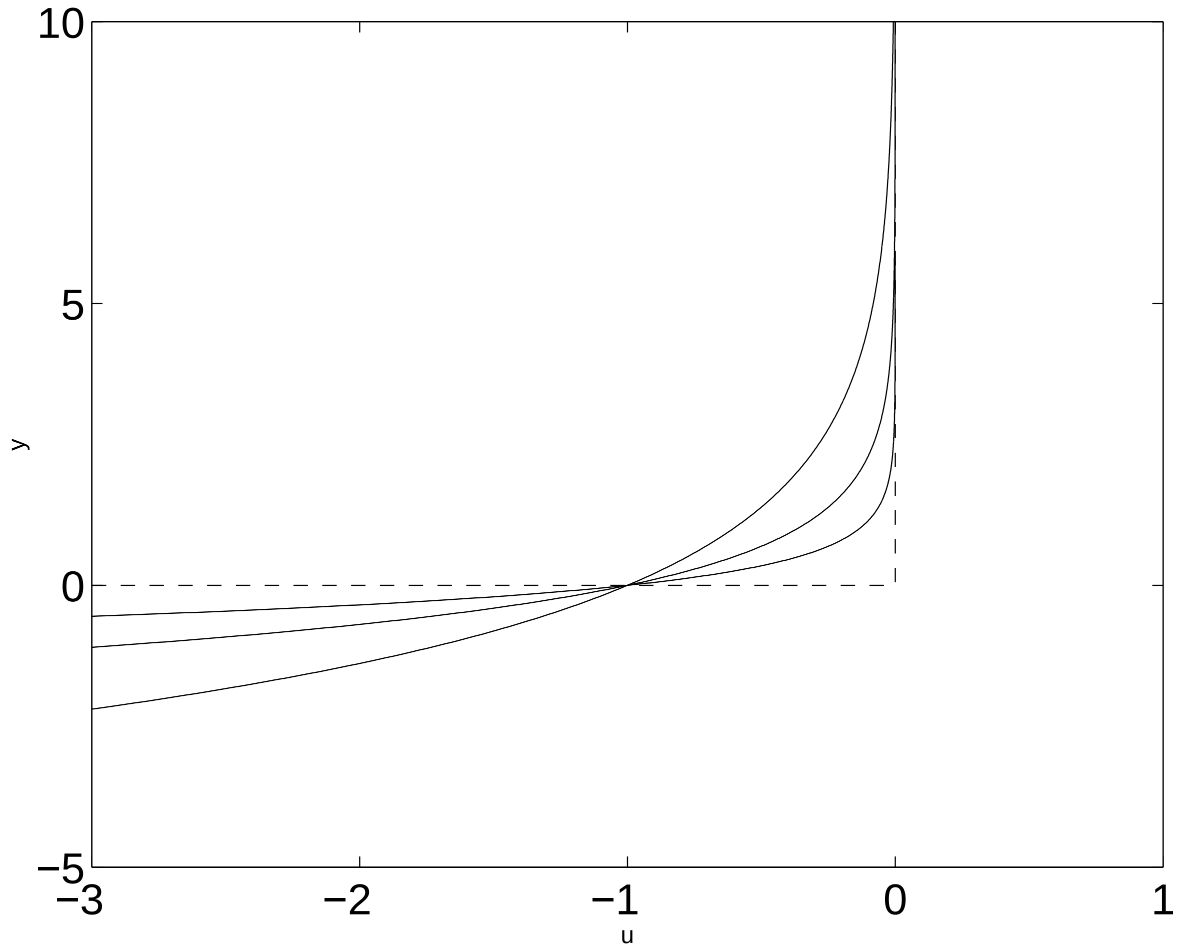
where a larger parameter t gives more accurate approximation

$\hat{I}_-$ increases to ∞ as u increases to 0

Log Barrier

Use Newton method to solve approximation:

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



Log Barrier

Log barrier function:

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

Approximation better if t is large, but then Hessian of $f_0 + (1/t)\phi$ varies rapidly near boundary of feasible set. Accuracy Stability tradeoff

Solve a sequence of approximation with larger t, using Newton method for each step of the sequence

Gradient and Hessian of log barrier function:

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

Central Path

Consider the family of optimization problems parameterized by $t > 0$:

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

Central path: solutions to above problem $x^*(t)$, characterized by:

1. Strict feasibility:

$$Ax^*(t) = b, \quad f_i(x^*(t)) < 0, \;\; i = 1, \ldots, m$$

2. Centrality condition: there exists $\hat{\nu} \in \mathbf{R}^p$ such that

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

Every central point gives a dual feasible point. Let

$$\lambda_i^*(t) = -\frac{1}{tf_i(x^*(t))}, \quad i = 1, \ldots, m \quad \nu^*(t) = \frac{\hat{\nu}}{t}$$

Central Path

Dual function

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

which implies duality gap is m/t. Therefore, suboptimality gap

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

Interpretation as modified KKT condition:

x is a central point $x^*(t)$ iff there exits λ, ν such that

$$
\begin{aligned}
Ax = b, \quad f_i(x) \le 0, & \qquad i = 1, \dots, m \\
\lambda \succeq 0 & \\
\nabla f_0(x) + \sum_{i=1}^m \lambda_i \nabla f_i(x) + A^T \nu = 0 & \\
-\lambda_i f_i(x) = 1/t, & \qquad i = 1, \dots, m
\end{aligned}
$$

Complementary slackness is relaxed from 0 to $1/t$

Example

Inequality form LP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

Log barrier function:

$$\phi(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$$

with gradient and Hessian:

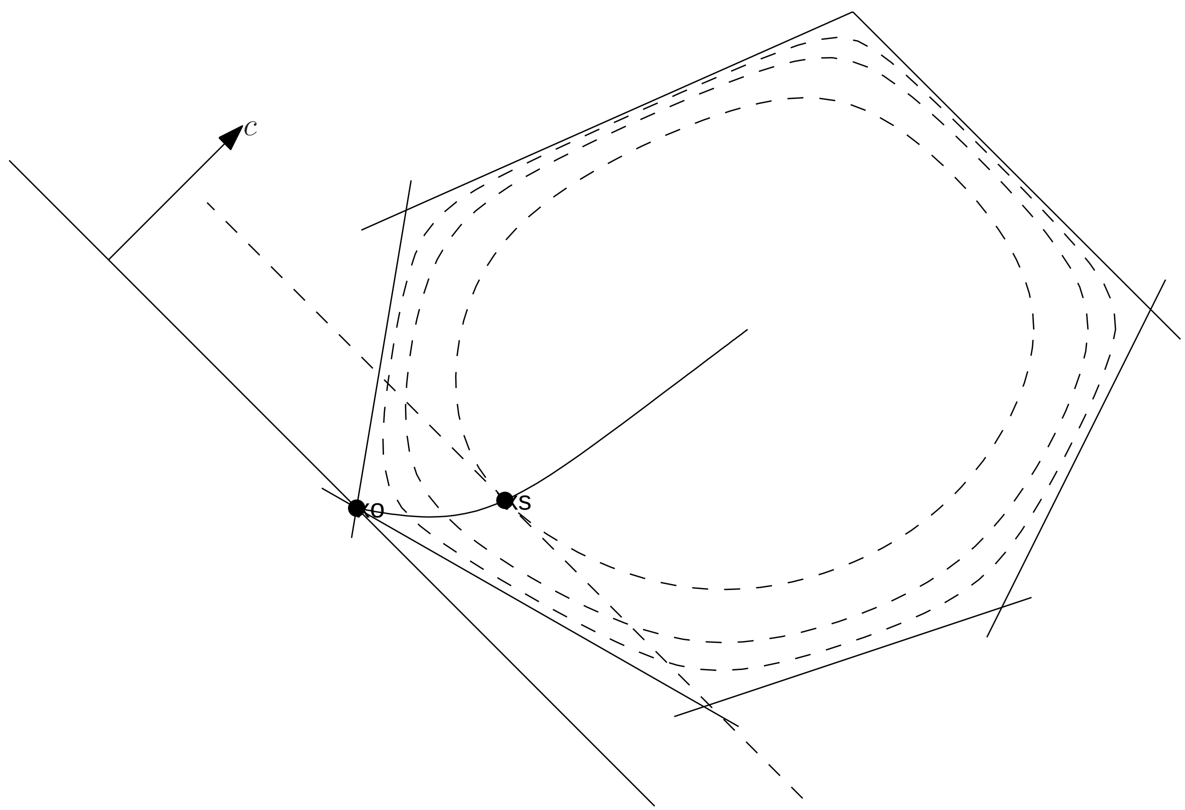
$$\begin{aligned} \nabla \phi(x) &= \sum_{i=1}^{m} \frac{a_i}{b_i - a_i^T x} = A^T d \\ \nabla^2 \phi(x) &= \sum_{i=1}^{m} \frac{a_i a_i^T}{(b_i - a_i^T x)^2} = A^T \operatorname{diag}(d^2) A \end{aligned}$$

where $d_i = 1/(b_i - a_i^T x)$

Centrality condition becomes: $tc + A^T d = 0$

c is parallel to $\nabla\phi(x)$.

Therefore, hyperplane $c^T x^*(t)$ is tangent to level set of ϕ



Barrier Method

GIVEN a strictly feasible point $x \in \mathbf{dom}\, f, t := t^{(0)} > 0, \mu > 1$ and tolerance $\epsilon > 0$

REPEAT

1. Centering step: compute $x^*(t)$ by minimizing $tf_0 + \phi$ subject to $Ax = b$, starting at x

2. Update: $x := x^*(t)$

3. Stopping criterion: QUIT if $\frac{m}{t} \leq \epsilon$

4. Increase t: $t := \mu t$

Other names: Sequential Unconstrained Minimization Technique (SUMT) or path-following method

Usually use Newton method for Centering Step

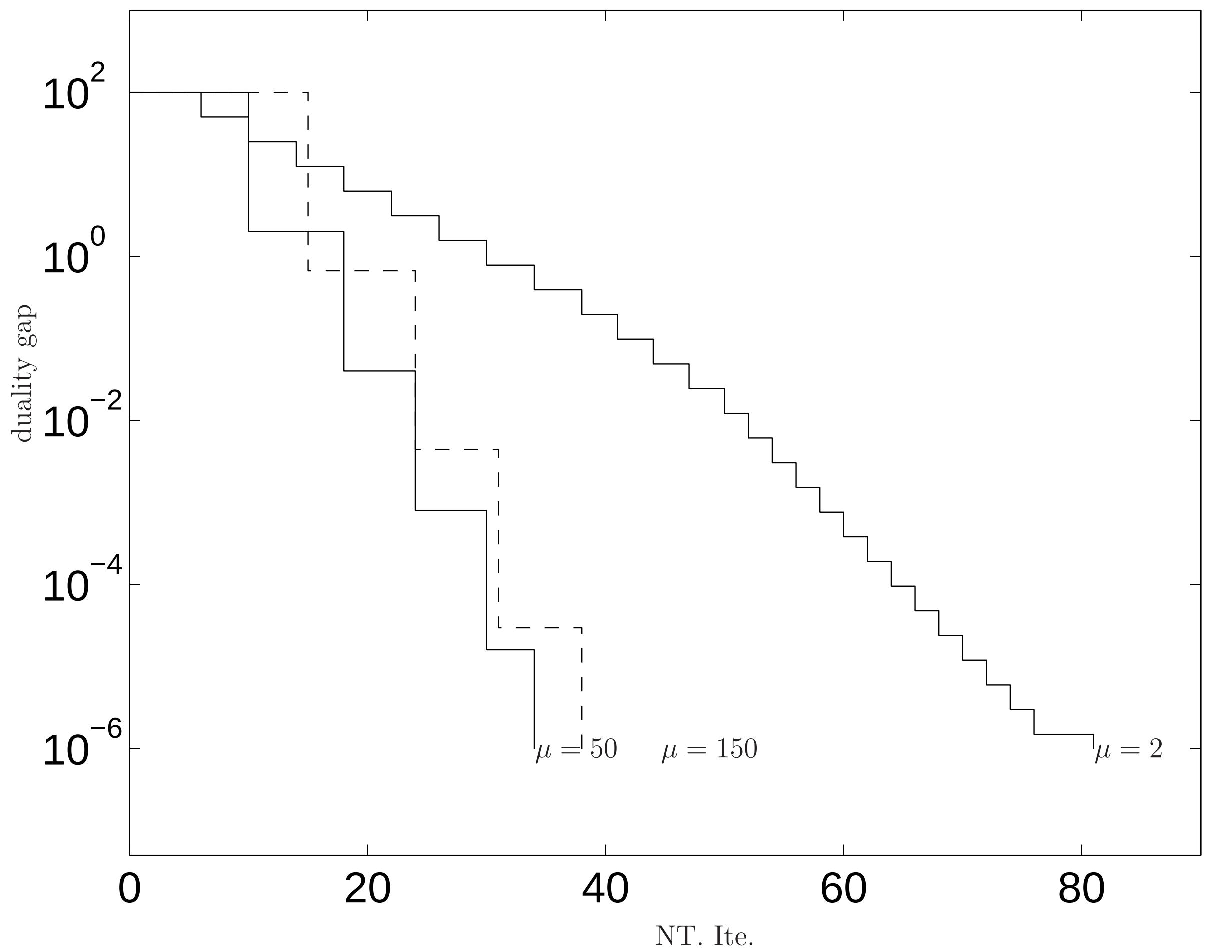
Remarks

- Each centering step does not need to be exact
- Choice of μ: tradeoff number of inner iterations with number of outer iterations
- Choice of $t^{(0)}$: tradeoff number of inner iterations within the first outer iteration with number of outer iterations
- Number of centering steps required:

$$\frac{\log(m/(\epsilon t^{(0)}))}{\log \mu}$$

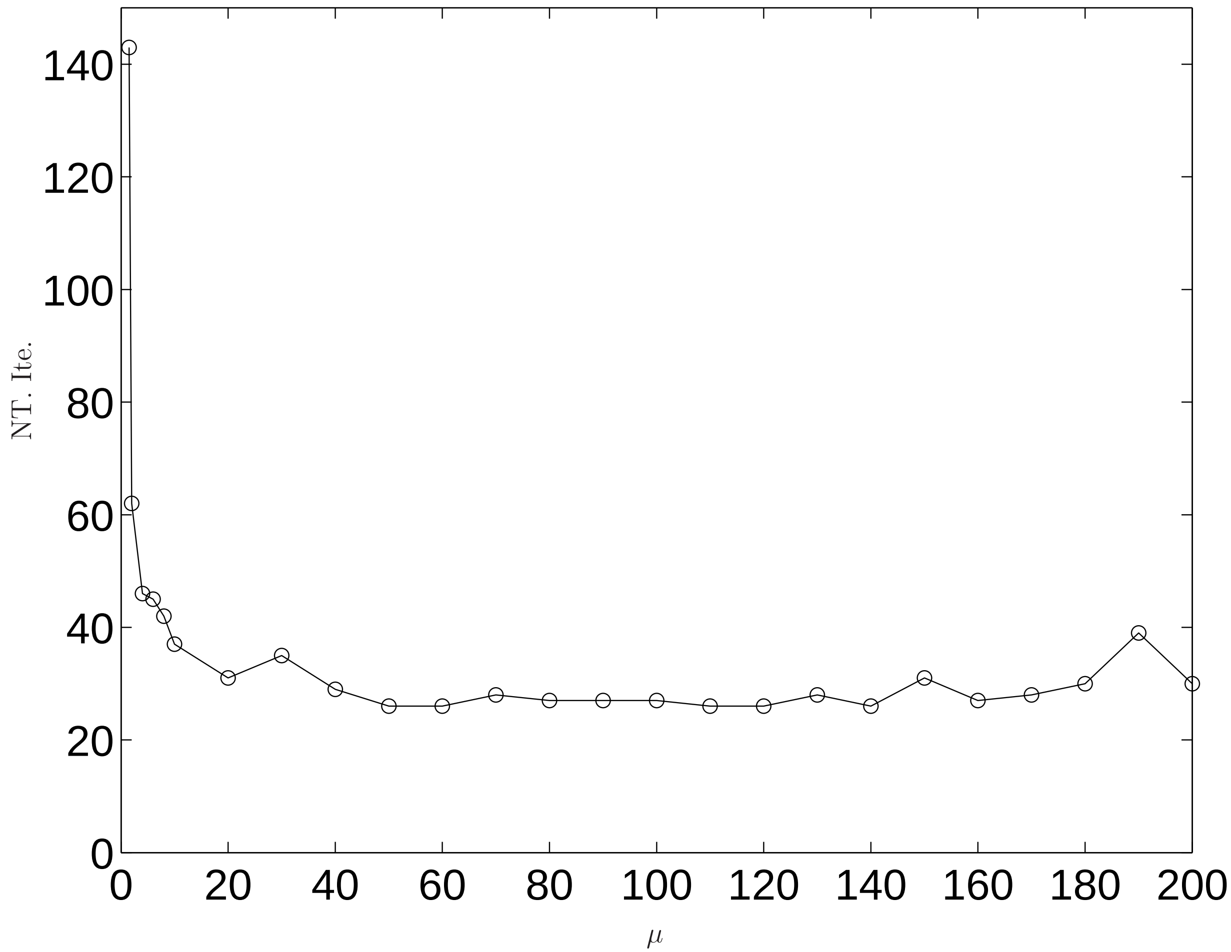
where m is the number of inequality constraints and ϵ is desired accuracy

Progress of Barrier Method for an LP Example



Three curves for $\mu = 2, 50, 150$

Tradeoff of μ Parameter for a Small LP



Phase I Method

How to compute a strictly feasible point to start barrier method?

Consider a phase I optimization problem in variables $x \in \mathbf{R}^n, s \in \mathbf{R}$:

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

Strictly feasible point: for any $x^{(0)}$, let $s = \max f_i(x^{(0)})$

1. Apply barrier method to solve phase I problem (stop when $s < 0$)

2. Use the resulted strictly feasible point for the original problem to start barrier method for the original problem

Not Covered

- Self-concordance analysis and complexity analysis (polynomial time)
- Numerical linear algebra (large scale implementation)
- Generalized inequalities (SDP)

Other (sometimes more efficient) algorithms:

- Primal-dual interior point method
- Ellipsoid methods
- Analytic center cutting plane methods

Summary

- Solve a general convex optimization with interior point methods
- Turn inequality-constrained problem into a sequence of equality constrained problems that are increasingly accurate approximation of the original problem
- Polynomial time (in theory) and much faster (in practice): about 25-50 least-squares effort for a wide range of problem sizes

Reading assignment: Sections 11.1-11.4 of textbook.