Optimization Theory and Algorithms

Instructor: Prof. LIAO, Guocheng （廖国成）

Email: liaogch6@mail.sysu.edu.cn

School of Software Engineering
Sun Yat-sen University

Outline

- Inequality constrained minimization
- Logarithmic barrier function and central path
- Barrier method

Equality constrained minimization problem

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t.}\ & f_i(x) \leq 0, i = 1, \dots, m \\ & Ax = b \end{aligned}$$

- f is convex and twice continuously differentiable
- Assume optimal point $x^\star$ exists. Let $p^\star = f(x^\star)$ be the optimal value.
- Assume Slater's condition holds, i.e., strong duality holds.

Optimality condition (KKT conditions): $x^\star$ is optimal iff there exists a $\lambda^\star$ and $\nu^\star$ such that

- $\lambda_i^\star f_i(x^\star) = 0, i = 1, \dots, m$
- $\nabla f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla f_i(x^\star) + A^T \nu^\star = 0$
- $\lambda^\star \geq 0$
- $f_i(x^\star) \leq 0, i = 1, \dots, m, Ax^\star = b$

Logarithmic barrier

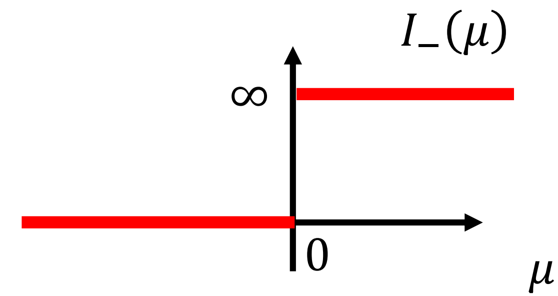
min $f_0(x)$
s.t. $f_i(x) \le 0, i = 1, \dots, m$
$Ax = b$

not differentiable

min $f_0(x) + \sum_{i=1}^{m} I_-(f_i(x))$
s.t. $Ax = b$

I_- is the indicator function for the nonpositive reals

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



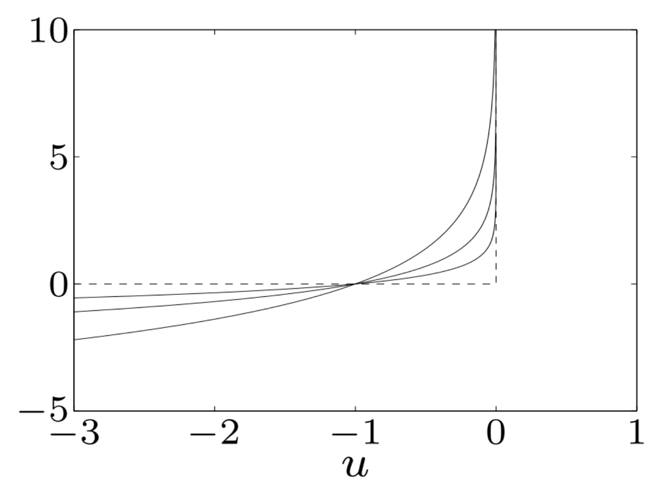
approximation

min $f_0(x) + \sum_{i=1}^{m} I_-(f_i(x))$
s.t. $Ax = b$

min $f_0(x) - (1/t)\sum_{i=1}^{m} \log(-f_i(x))$
s.t. $Ax = b$

$$\widehat{I_-}(\mu) = -(1/t)\sum_{i=1}^{m} \log(-\mu)$$

- Convex
- Differentiable
- As t increases, the approximation is more accurate



Central path

$$\begin{aligned} &\text{min} \;\; f_0(x) - (1/t)\textstyle\sum_{i=1}^{m}\log(-f_i(x)) \\ &\text{s.t.} \;\;\; Ax = b \end{aligned}$$

- $\phi(x) = -\sum_{i=1}^{m}\log(-f_i(x))$: logarithmic barrier function

Multiply the objective with t

$$\begin{aligned} &\text{min} \;\; tf_0(x) + \phi(x) \\ &\text{s.t.} \;\;\; Ax = b \end{aligned}$$

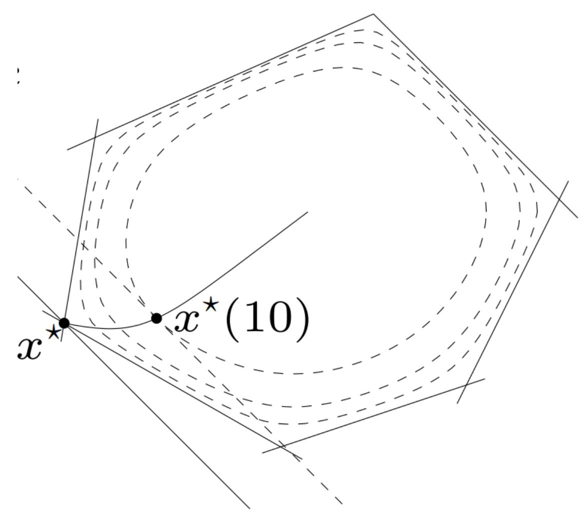
- For $t > 0$, $x^\star(t)$ is the solution of the above problem
- Central path: $x^\star(t)$, $t > 0$:

$$Ax^\star(t) = b$$

$$f_i(x^\star(t)) < 0$$

$$\boxed{t\nabla f_0(x^\star(t)) + \nabla\phi(x^\star(t)) + A^T\nu'} = 0$$

$$t\nabla f_0(x^\star(t)) + \sum_{i=1}^{m}\frac{1}{-f_i\big(x^\star(t)\big)}\nabla f_i(x^\star(t)) + A^T\nu' = 0$$



Approximation gap

$$\min \; tf_0(x) + \phi(x)$$
$$\text{s.t.} \quad Ax = b$$

$$p^\star = \min \; f_0(x)$$
$$\text{s.t.} \quad f_i(x) \le 0, i = 1, \dots, m$$
$$Ax = b$$

Lagrangian

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

- $Ax^\star(t) = b$
- $f_i(x^\star(t)) < 0$
- $t\nabla f_0(x^\star(t)) + \sum_{i=1}^m \frac{1}{-f_i(x^\star(t))} \nabla f_i(x^\star(t)) + A^T\nu' = 0$

Lower bound of the optimal value $p^\star$: $f_0(x^\star(t)) \le p^\star + m/t$

convergence as $t \to \infty$

- $\nabla f_0(x^\star(t)) + \sum_{i=1}^m \frac{1}{-tf_i(x^\star(t))} \nabla f_i(x^\star(t)) + A^T\nu'/t = 0$
- Define $\lambda_i^\star(t) = -1/tf_i\big(x^\star(t)\big)$, and $\nu_i^\star(t) = \nu'/t$
- $x^\star(t)$ minimizes Lagrangian $L(x, \lambda^\star(t), \nu^\star(t)) = f_0(x) + \sum_{i=1}^m \lambda_i^\star(t) f_i(x^\star) + \nu_i^\star(t)^T(Ax - b)$
- $g(\lambda^\star(t), \nu^\star(t)) = f_0(x^\star(t)) + \sum_{i=1}^m \lambda_i^\star(t) f_i(x^\star) + \nu_i^\star(t)^T(Ax - b) = f_0\big(x^\star(t)\big) - m/t$

$$f_0\big(x^\star(t)\big) - m/t = g(\lambda^\star(t), \nu^\star(t)) \le p^\star$$

Interpretation via KKT conditions

$x^\star(t), \lambda^\star(t), \nu^\star(t)$ satisfy

- Approximate complementary slackness: $\lambda_i^\star(t) f_i(x^\star(t)) = 1/t$, $i = 1, \dots, m$
- Lagrangian optimality: $\nabla f_0(x^\star(t)) + \sum_{i=1}^m \lambda_i^\star(t) \nabla f_i(x^\star(t)) + A^T \nu^\star(t) = 0$
- Dual feasibility: $\lambda^\star(t) \geq 0$
- Primal feasibility: $f_i(x^\star(t)) \leq 0, i = 1, \dots, m, Ax^\star(t) = b$

Barrier method

- Given strictly feasible x, $t > 0$,$u > 1$, tolerance $\epsilon > 0$
- **Repeat**

1. *Centering step.*
   Starting at x, compute $x^\star(t)$ by solving the following problem (Newton's method)

$$\begin{array}{ll} \text{min} & tf_0(x) + \phi(x) \\ \text{s.t.} & Ax = b \end{array}$$

2. Update: $x \leftarrow x^\star(t)$

3. Stopping criterion: if $m/t \leq \varepsilon$, break

4. *Increase* t. $t \leftarrow ut$