Linear Programming

(5531)

Lecture 09 Interior Point Methods

Prof. Sheng-I Chen

http://www.iem-omglab.nctu.edu.tw/OMG_LAB/

November 19, 2020

Agenda

- The central path (Chapter 17)
  - Barrier problem / central path
  - Lagrange multipliers
  - The first order optimality conditions / the second order information
  - Existence
- The central-following method (Chapter 18)

The Barrier Problem

- Given an LP:

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \le b \\ & x \ge 0 \end{array}$$

- The associated dual is:

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \ge c \\ & y \ge 0 \end{array}$$

The Barrier Problem

- After adding slack variables, the primal becomes:

$$\begin{aligned} &\text{maximize} \quad c^T x \\ &\text{subject to} \quad Ax + \omega = b \\ &\qquad\qquad\qquad x, \omega \geq 0 \end{aligned}$$

- , and dual is:

$$\begin{aligned} &\text{minimize} \quad b^T y \\ &\text{subject to} \quad A^T y - z = c \\ &\qquad\qquad\qquad y, z \geq 0 \end{aligned}$$

Convert Constrained LP to Unconstrained LP

- Consider the primal:

$$\begin{aligned} \text{maximize} \quad & c^T x \\ \text{subject to} \quad & Ax + \omega = b \\ & x, \omega \geq 0 \end{aligned}$$

- Replace *nonnegative constraints* ($x \geq 0$ and $\omega \geq 0$) with extra terms in the objective function
- However, the new objective function becomes *discontinue or undifferentiable* (*i.e.,* x and $\omega \geq 0$)
- We may replace this *discontinuous function* with another function as follows:
  - **finite** when x or ω is **positive**
  - *, and approaching* **negative infinity** *when* x or ω *is* **closed to zero**
- This conversion smoothes out the discontinuity, thus, we may use *calculus* to study it
- The simplest function to do so is the **logarithm**

The Barrier Problem

- Given the LP:

$$\begin{aligned} \text{maximize} \quad & c^T x \\ \text{subject to} \quad & Ax + \omega = b \\ & x, \omega \geq 0 \end{aligned}$$

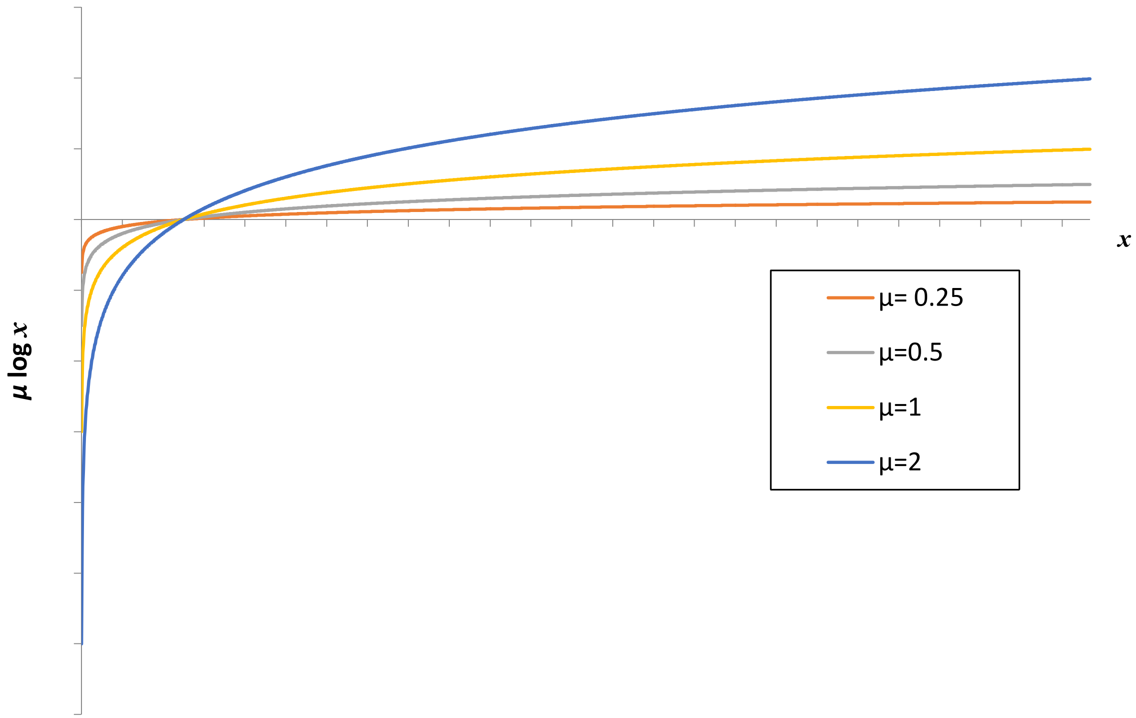
- It's associated *barrier problem*:

$$\begin{aligned} \text{maximize} \quad & c^T x + \mu \sum_j \log x_j + \mu \sum_i \log \omega_i \\ \text{subject to} \quad & Ax + \omega = b \end{aligned}$$

  - , where the parameter ***μ*** is a positive constant, when it gets small, the objective function becomes similar to primal objective function

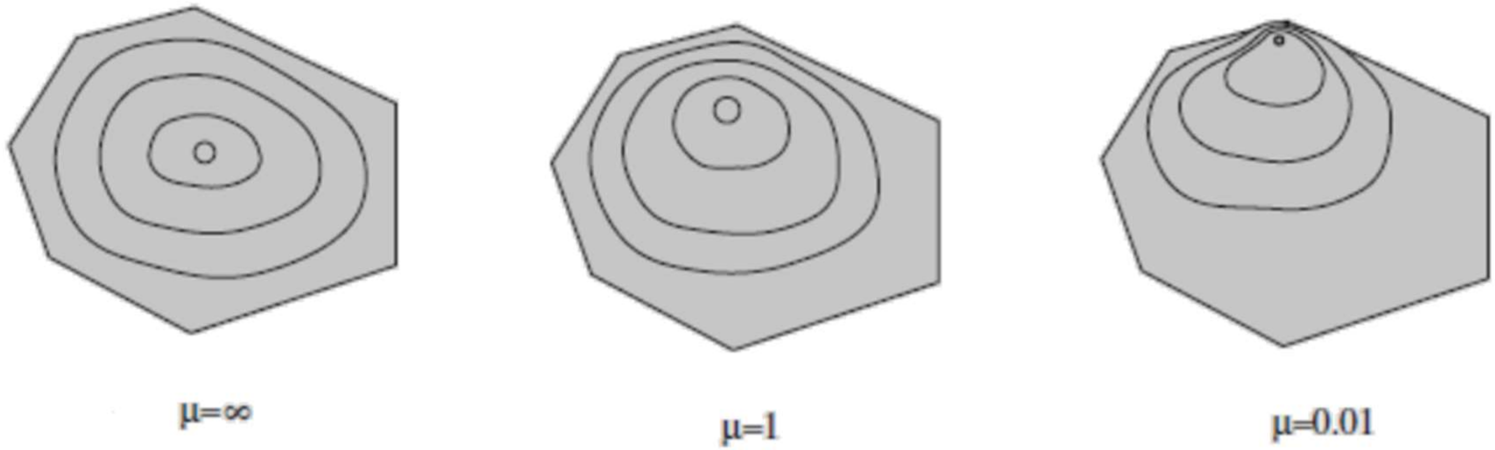
- This problem is a *nonlinear programming* problem as the objective function is nonlinear. The objective function is called a **barrier function** or, more specifically, a **logarithmic barrier function**

Logarithmic Barrier Functions



The Barrier Function

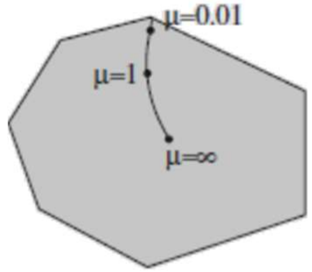
- The barrier functions with different setting of μ:



- The barrier function is *always* at an interior point (why?)
- Thus, the interior of feasible sets must be **nonempty**, and the barrier function method can *only* use for the problems with **inequality constraints**
- When μ gets closer to zero this interior point moves closer to the optimal solution of the original linear programming problem

The Barrier Function

- Recall that the set of feasible solutions is a polyhedron with each face characterized setting one of the variables as zero
- The barrier function is *minus infinity* on each face of the polyhedron (why?)
- It is finite in the interior of the polyhedron, and it approaches minus infinity as the boundary is approached
- The set of optimal solutions to the barrier problems forms a path through the interior of the polyhedron of feasible solutions. This path is called the **central path**



Lagrange Multipliers

- Given a general problem of maximizing a function subject to one or more equality constraints:

$$\text{maximize} \quad f(x)$$
$$\text{subject to} \quad g(x) = 0$$

- Note that the function $f(x)$ can be nonlinear, but are assumed to be smooth and twice differentiable

Lagrange Multipliers

- There is a simple algebraic formalism yielded the same equations, so called **Lagrangian function**:

$$L(x, y) = f(x) - \sum_i y_i g_i(x)$$

- Now, it becomes an **unconstrained** optimization problem, and the **critical point** is determined by simply setting of all the **first derivatives to zero**:

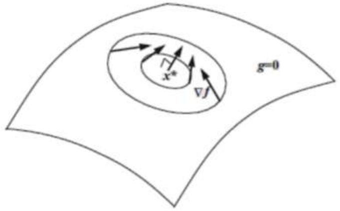
$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} - \sum_i y_i \frac{\partial g_i}{\partial x_j} = 0, j = 1,2,\ldots,n$$

$$\frac{\partial L}{\partial y_i} = -g_i = 0, i = 1,2,\ldots,m$$

- These equations are usually referred as the **first-order optimality conditions**

Lagrange Multipliers

- The **gradient** of f, ∇f, is a *vector* that points in the direction of the *most rapid increase* of f



$$\boxed{\begin{array}{ll} \text{maximize} & f(x) \\ \text{subject to} & g(x)=0 \end{array}}$$

- We simply **set this vector equal to zero** to determine the **critical point** of f, and the maximum, if it exists, would have to be included in this set
- Since the system has constraints, $g(x)$ *= 0, it is no longer* correct just looking at points with $\nabla f = 0$
- At each point x in the feasible set, $\nabla g(x)$, is a vector that is **orthogonal** to the feasible set at this point x
- $\nabla f(x^*)$ lies in the span of the gradients of $\nabla g(x^*)$
- Therefore, our new requirements for a point x^* to be a critical point are: (1) x^* is feasible, (2) $\nabla f(x^*)$ be proportional to $\nabla g(x^*)$. That is:

$$g(x^*)=0$$

$$\nabla f(x^*)=\sum_{i=1}^{m} y_i \nabla g(x^*)$$

Lagrange Multipliers

- Now, we want to check if there exists an optimal solution
- If all constraints are *linear*, then we may look at the matrix of **second derivatives**. This matrix is called the **Hessian** of f at x:

$$Hf(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]$$

- **Theorem 17.1:** *If the constraints are linear, a critical point x^* is a **local maximum** if*

  $\xi^T Hf(x^*)\, \xi < 0$, *for each* $\xi \neq 0$ *satisfying* $\xi^T \nabla g_i(x^*) = 0,\ i = 1, 2, \ldots, m$

- Note that if $x^T Hf(x^*) x < 0$ is satisfied not just at x^* but **at all x**, then x^* is a **unique global maximum**

Proof of Theorem 17.1

Consider the two-term Taylor series expansion of f about x^*:

$$f(x^* + \xi) = f(x^*) + \nabla f(x^*)^{\mathrm{T}} \xi + ½\, \xi^{\mathrm{T}} Hf(x^*)\, \xi + o(\|\xi\|^2)$$

Let ξ be a direction vector satisfying $\xi^T \nabla g_i(x^*) = 0$

Thus, $\nabla f(x^*) = \sum_i \xi_i \nabla gi(x^*) = 0$

If ½ $\xi^{\mathrm{T}} Hf(x^*)\, \xi < 0$, then $f(x^* + \xi) < f(x^*)$

Implies the critical point x^* is local minimum

The Central Path

Lagrange Multipliers Applied to the Barrier Problem

- We will show for each value of the barrier parameter μ, there is a unique solution to the barrier problem
- Also, we will show when μ closes to zero, the solution to the barrier problem tends to the solution of the original LP
- Recall the barrier problem:

$$\begin{aligned} &\text{maximize} \quad c^T x + \mu \sum_j \log x_j + \mu \sum_i \log \omega_i \\ &\text{subject to} \quad Ax + \omega = b \end{aligned}$$

- The Lagrangian for this problem is:

Why plus?

$$L(x,\omega,y) = c^T x + \mu \sum_j \log x_j + \mu \sum_i \log \omega_i + y^T (b - Ax - \omega)$$

- Taking derivatives with respect to each variable and setting them to zero, we obtain the **first-order optimality conditions**:

$$\frac{\partial L}{\partial x_j} = c_j + \mu \frac{1}{x_j} - \sum_i y_i a_{ij} = 0,\ j = 1,2,\ldots,n$$

$$\frac{\partial L}{\partial \omega_i} = \mu \frac{1}{\omega_i} - y_i = 0, i = 1,2,\ldots,m$$

$$\frac{\partial L}{\partial y_i} = b_i - \sum_j a_{ij} x_j - \omega_i = 0, i = 1,2,\ldots,m$$

- These equations are presented in matrix form:

$$A^T y - \mu X^{-1} e = c$$

$$y = \mu W^{-1} e$$

$$Ax + \omega = b$$

Note: X and W are diagonal matrices, e.g. $X = \begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{bmatrix}$

First-order Optimality Conditions

- Introduce a vector $z = \mu X^{-1} e$, then rewrite the *first-order optimality conditions:*

$$Ax + \omega = b$$

$$A^T y - z = c$$

$$z = \mu X^{-1} e$$

$$y = \mu W^{-1} e$$

- Multiply the third equation by X and the fourth equation by W yielded:

$$\begin{aligned} Ax + \omega &= b \\ A^T y - z &= c \\ XZe &= \mu e \\ YWe &= \mu e \end{aligned}$$

$$\boxed{\begin{aligned} &Ax + \omega = b \\ &A^T y - z = c \\ &XZe = \mu e \\ &YWe = \mu e \end{aligned}}$$

- Let's compare the first-order optimality conditions with original primal and dual problems:
  - The first equation is the primal constraint
  - The second equation is the dual constraint
  - Rewrite the third and fourth equations as component-wise:

$$x_j z_j = \mu \qquad j = 1,2,\ldots,n$$
$$y_i \omega_i = \mu \qquad j = 1,2,\ldots,m$$

  - If set μ to zero, then they are exactly the usual complementarity conditions that must be satisfied at optimality:

$$x_j z_j = 0 \qquad j = 1,2,\ldots,n$$
$$y_i \omega_i = 0 \qquad j = 1,2,\ldots,m$$

The Second-Order Information

- To check whether a solution to the first-order optimality conditions is **unique** (if it is exists), we might use *second-order information on the barrier function*: $f(x,\omega) = c^T x + \mu \sum_j \log x_j + \mu \sum_i \log \omega_i$
- The first derivatives are:

$$\frac{\partial f}{\partial x_j} = c_j + \mu \frac{1}{x_j} \qquad , j = 1,2,\ldots,n$$

$$\frac{\partial f}{\partial \omega_i} = \frac{\mu}{\omega_i} \qquad , i = 1,2,\ldots,m$$

- The *pure second derivatives* are:

$$\frac{\partial^2 f}{\partial x_j^2} = -\mu \frac{1}{x_j^2} \qquad , j = 1,2,\ldots,n$$

$$\frac{\partial^2 f}{\partial \omega_i^2} = -\frac{\mu}{\omega_i^2} \qquad , i = 1,2,\ldots,m$$

Why consider pure only for second derivatives?

- All *mixed second derivatives* are vanished. Therefore, the *Hessia*n is a **diagonal matrix with strictly negative entries** Why strictly negative?
- By **Theorem 17.1**, there can be *at most one critical point* and, if it exists, it is a *global maximum*

Existence

- Does a solution to the barrier problem always exist?

- Consider the following LP:

$$\begin{aligned} &\text{maximize} \quad 0 \\ &\text{subject to} \quad x \geq 0 \end{aligned}$$

- It's barrier function is :

$$f(x) = \mu \log x$$

- This function doesn't have a maximum (as the maximum is infinity when x = $+\infty$)

- Let's modify the objective function as:

$$\begin{aligned} &\text{maximize} \quad -x \\ &\text{subject to} \quad x \geq 0 \end{aligned}$$

- It's barrier function is :

$$f(x) = -x + \mu \log x$$

- This function has a maximum with x = μ

Existence

- **Theorem 17.2**: There exists a solution to the barrier problem *if and only if* both the primal and the dual feasible regions have *nonempty interior*
- **Corollary 17.3:** If a primal feasible set (or its dual) has a *nonempty interior* and *bounded*, then for each $\mu > 0$ there *exists a unique solution* $(x_\mu, \omega_\mu, y_\mu, z_\mu)$
- The path $\{(x_\mu, \omega_\mu, y_\mu, z_\mu) : \mu > 0\}$ is called the **primal–dual central path**

Homework

- **Hand-writing questions:**

  Exercise 17.1

  Exercise 17.3