Convex Optimization

Lecture 8: Lagrange Multiplier

Lecturer: *Dr*. Wan-Lei Zhao

Autumn Semester 2025

Contact: wlzhao@xmu.edu.cn

Outline

Opening Discussion

1. So far we have seen various optimization problems
2. We know how to judge whether a problem is convex
3. Optimization problems can be divided into different categories
4. The problem now is how to solve these problems
5. Namely, how to find the extreme values for an objective function under several constraints

The Popularity of Langrangian Multiplier

- **Differentiable Optimiz.**
  - **Convex**
    - **Linear Programming**
      (Simplex, interior-point)
    - **Quadratic Programming**
      (Active set)
    - **Others**
  - **non-Convex**
- **Descrete Optimiz.**
  - **Integer Programming**
    (branch-bounding, cutting-plane)
  - **Dynamic Programming**
  - **Others**

Lagrangian Multiplier

Opening Example (1)

$$\begin{aligned} \underset{x_1,x_2}{\text{Min.}} \quad & x_1 + x_2 \\ s.t.\ & x_1^2 + x_2^2 - 1 \leq 0 \end{aligned} \tag{1}$$

- Given there is no constraint, the problem is easy to address
- So we should consider how to remove the constraint
- While still being in line with the constraint
- We introduce the constraint as a penalty term into the objective

$$\underset{x_1,x_2}{\text{Min.}} \quad x_1 + x_2 + P(x_1^2 + x_2^2 - 1) \tag{2}$$

- If x_1, x_2 is out of the constraint, big value will be assigned to P

Opening Example (2)

$$\begin{aligned} &\underset{x_1,x_2}{\text{Min.}} \quad x_1 + x_2 \\ &s.t.\ x_1^2 + x_2^2 - 1 \leq 0 \end{aligned} \tag{3}$$

$$\Downarrow$$

$$\underset{x_1,x_2}{\text{Min.}} \quad x_1 + x_2 + P(x_1^2 + x_2^2 - 1) \tag{4}$$

- If x_1, x_2 is out of the constraint, big value will be assigned to P
- Namely, P is given as

$$P(y) = \begin{cases} +\infty & y > 0 \\ 0 & y \leq 0 \end{cases}$$

- However, the objective function is no longer differentiable
- We need a smooth penalty function

Opening Example (3)

$$\underset{x_1,x_2}{\text{Min.}} \quad x_1 + x_2 + P(x_1^2 + x_2^2 - 1) \tag{5}$$

- However, the objective function is no longer differentiable
- We need a smooth penalty function

$$P(y) = \lambda \cdot y$$

- The penalty takes higher value if y violates more
- If we take the maximum of λ, we still approximate the previous one

$$P(y) = \begin{cases} +\infty & y > 0 \\ 0 & y \leq 0 \end{cases}$$

Opening Example (4)

$$\underset{x_1,x_2}{\text{Min.}} \quad x_1 + x_2 + P(x_1^2 + x_2^2 - 1) \tag{6}$$

$$P(y) = \lambda \cdot y \Downarrow$$

$$\underset{x_1,x_2}{\text{Min.}} \quad x_1 + x_2 + \underset{\lambda \geq 0}{\text{Max.}}\, \lambda(x_1^2 + x_2^2 - 1)$$

$$\Downarrow$$

$$\underset{x_1,x_2}{\text{Min.}}\, \underset{\lambda \geq 0}{\text{Max.}} \quad x_1 + x2 + \lambda(x_1^2 + x_2^2 - 1)$$

- Maximize on λ first or later do not impact the overall, so we have
$$\underset{\lambda \geq 0}{\text{Max.}}\, \underset{x_1,x_2}{\text{Min.}} \quad x_1 + x_2 + \lambda(x_1^2 + x_2^2 - 1)$$

Opening Example (5)

$$\underset{\lambda\geq 0}{\text{Max.}}\ \underset{x_1,x_2}{\text{Min.}} \quad x_1 + x_2 + \lambda(x_1^2 + x_2^2 - 1)$$

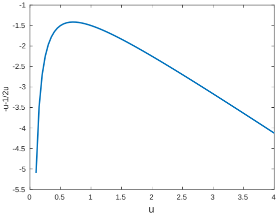
- Now let's solve the minimization first, treating λ as a constant
- We have

$$1 + 2\lambda x_1 = 0 \Rightarrow x_1 = \frac{-1}{2\lambda}$$

$$1 + 2\lambda x_2 = 0 \Rightarrow x_2 = \frac{-1}{2\lambda}$$

$$x_1 = \tfrac{-1}{2\lambda} \Downarrow x_2 = \tfrac{-1}{2\lambda}$$

$$\underset{\lambda\geq 0}{\text{Max.}} -\lambda - \frac{1}{2\lambda}$$



- $x_1 = x_2 = -\frac{\sqrt{2}}{2}, \lambda = \frac{\sqrt{2}}{2}$, the objective value is $-\sqrt{2}$

Lagrange Multiplier (1)

- Given (x_0, y_0) is the extreme point for function $z = f(x, y)$ under the constraint of $\phi(x, y) = 0$
- We can solve y out from $\phi(x, y) = 0$, that is $y = y(x)$
- Then we plug $y = y(x)$ into $f(x, y)$, we have

$$z = f(x, y(x))$$

$$x = x_0 \text{ is the extrema}$$

- We take derivative on $f(x, y(x))$ w.r.t x

$$\frac{dz}{dx}|_{x=x_0} = f_x(x_0, y_0) + f_y(x_0, y_0) \cdot \frac{dy}{dx}|_{x=x_0} = 0$$

$$\tfrac{dy}{dx} = -\tfrac{\phi_x(x_0,y_0)}{\phi_y(x_0,y_0)} \Downarrow$$

$$\frac{dz}{dx}|_{x=x_0} = f_x(x_0, y_0) + f_y(x_0, y_0) \cdot (-\frac{\phi_x(x_0, y_0)}{\phi_y(x_0, y_0)})|_{x=x_0} = 0$$

Lagrange Multiplier (2)

- We take derivative on $f(x, y(x))$ w.r.t x

$$\frac{dz}{dx}|_{x=x_0} = f_x(x_0, y_0) + f_y(x_0, y_0)\cdot\frac{dy}{dx}|_{x=x_0} = 0$$

$$\frac{dy}{dx} = -\frac{\phi_x(x_0,y_0)}{\phi_y(x_0,y_0)} \Downarrow$$

$$\frac{dz}{dx}|_{x=x_0} = f_x(x_0, y_0) + f_y(x_0, y_0)\cdot(-\frac{\phi_x(x_0, y_0)}{\phi_y(x_0, y_0)})|_{x=x_0} = 0$$

$$\frac{f_y(x_0,y_0)}{\phi_y(x_0,y_0)} = -\mu \Downarrow$$

$$\begin{cases} f_x + \mu\phi_x = 0 \\ f_y + \mu\phi_y = 0 \\ \phi(x, y) = 0 \end{cases}$$

Noted that the derivative for implicit function $\phi(x, y)$ is $\frac{dy}{dx} = -\frac{\phi_x(x,y)}{\phi_y(x,y)}$

Lagrange Multiplier (3)

- Given an optimization problem

$$\begin{aligned} Min.\ \ & z = f(x, y) \\ \text{s.t.}\ \ & \phi(x, y) = 0 \end{aligned} \tag{7}$$

$$\Downarrow$$

$$\begin{cases} f_x + \mu\phi_x = 0 \\ f_y + \mu\phi_y = 0 \\ \phi(x, y) = 0 \end{cases} \tag{8}$$

Summarize these necessary $\Downarrow$conditions into one equation

$$L(x, y) = f(x, y) + \mu\phi(x, y) \tag{9}$$

- This is called **Lagrange function**

Lagrange Multiplier (4)

- Given an optimization problem

$$\begin{aligned} Min.\ \ & z = f(x, y) \\ \text{s.t.}\ \ & \phi(x, y) = 0 \end{aligned} \tag{10}$$

$$\Downarrow$$

$$L(x, y) = f(x, y) + \mu\phi(x, y) \tag{11}$$

- Take partial derivative on x, y and μ

$$\begin{cases} L_x = 0 \Rightarrow & f_x + \mu\phi_x = 0 \\ L_y = 0 \Rightarrow & f_y + \mu\phi_y = 0 \\ L_\mu = 0 \Rightarrow & \phi(x, y) = 0 \end{cases} \tag{12}$$

Lagrange Multiplier (5)

- Given an optimization problem under multiple contraints

$$\begin{aligned} Min.\quad & z = f(x,y) \\ \text{s.t.}\quad & h(x,y) = 0 \\ & g(x,y) \leq 0 \end{aligned} \tag{13}$$

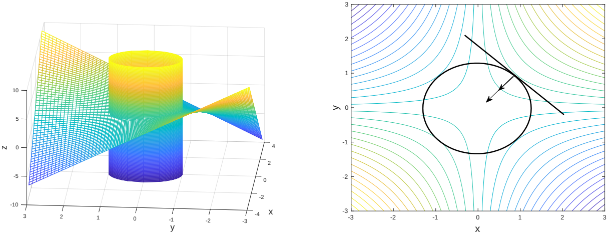
$$\Downarrow$$

$$L(x,y) = f(x,y) + \lambda g(x,y) + \mu h(x,y) \tag{14}$$

- Take partial derivative on x, y, λ and μ

$$\begin{cases} L_x = 0 \Rightarrow & f_x + \lambda g_x + \mu h_x = 0 \\ L_y = 0 \Rightarrow & f_y + \lambda g_y + \mu h_y = 0 \\ L_\lambda = 0 \Rightarrow & g(x,y) = 0 \\ L_\mu = 0 \Rightarrow & h(x,y) = 0 \end{cases} \tag{15}$$

Lagrange Multiplier (6)

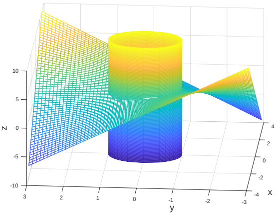


(a) (b)

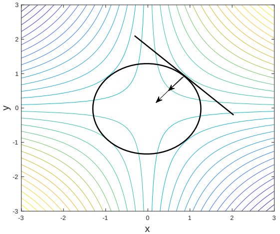
- Given an optimization problem under constraint
- On the potential extremal point, tengants from two functions parallel

$$\begin{aligned} Max. \quad & z = f(x, y) \\ \text{s.t.} \quad & h(x, y) = 0 \end{aligned}$$

Lagrange Multiplier (7)



(a)

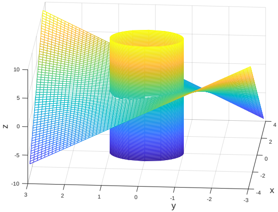


(b)

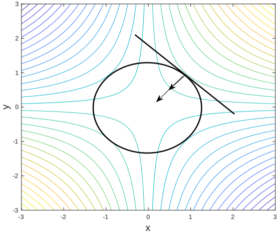
- Given an optimization problem under constraint
- The optimization searches for $(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y}) = \mu(\frac{\partial h(x,y)}{\partial x}, \frac{\partial h(x,y)}{\partial y})$

$$\begin{aligned} Max. \quad & z = f(x,y) \\ \text{s.t.} \quad & h(x,y) = 0 \end{aligned}$$

Lagrange Multiplier (8)



(a)



(b)

- The optimization searches for $(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y}) = \mu(\frac{\partial h(x,y)}{\partial x}, \frac{\partial h(x,y)}{\partial y})$
- Let us solve a concrete problem

$$Max. \quad z = xy + 1$$
$$\text{s.t.} \quad x^2 + y^2 - 1 = 0$$

Lagrange Multiplier (9)

- Let us solve a concrete problem

$$\begin{aligned} Max. \quad & z = xy + 1 \\ \text{s.t.} \quad & x^2 + y^2 - 1 = 0 \end{aligned}$$

$$\Downarrow$$

$$\begin{cases} z = xy + 1 \Rightarrow & \left(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y}\right) = (y, x) \\ x^2 + y^2 - 1 = 0 \Rightarrow & \left(\frac{\partial h(x,y)}{\partial x}, \frac{\partial h(x,y)}{\partial y}\right) = (2x, 2y) \end{cases} \tag{16}$$

$$\Downarrow$$

$$\begin{cases} y = \mu 2x \\ x = \mu 2y \\ x^2 + y^2 - 1 = 0 \end{cases}$$

Lagrange Multiplier (10)

- Let us solve a concrete problem

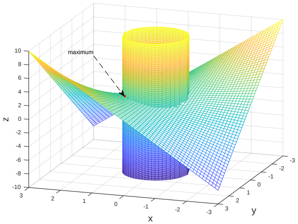
$$\begin{aligned} Max. \quad & z = xy + 1 \\ \text{s.t.} \quad & x^2 + y^2 - 1 = 0 \end{aligned}$$

$$\Downarrow$$

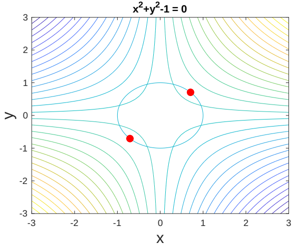
$$\begin{cases} y = \mu 2x \\ x = \mu 2y \\ x^2 + y^2 - 1 = 0 \end{cases}$$

$$y = \mu 2x \Downarrow y = \frac{x}{2\mu}$$

$$\begin{cases} 4\mu^2 x^2 + x^2 - 4\mu^2 = 0 \\ x^2 + 4\mu^2 x^2 - 1 = 0 \end{cases}$$



(c)



(d)

- We have $x = \pm\frac{1}{\sqrt{2}}$, $\mu = \pm\frac{1}{2}$, $y = \pm\frac{1}{\sqrt{2}}$