Lesson 24. Optimization with Equality Constraints

1 The effect of a constraint

- Let's model a consumer whose utility depends on his or her consumption of two products
- Define the following variables:

$$x_1 = \text{units of product 1 consumed} \qquad x_2 = \text{units of product 2 consumed}$$

- The consumer's utility function is

$$U(x_1, x_2) = 4x_1^{1/2} + 8x_2^{1/2}$$

- Without any additional information, the consumer can maximize his or her utility by

- To make this model more realistic, we should take into account the consumer's budget
- Suppose the unit prices of products 1 and 2 are \$2 and \$4 respectively
- In addition, suppose the consumer intends to spend \$6 on the two products
- The consumer's budget constraint can be expressed as

- Putting this all together, we obtain the following optimization model:

$$\begin{aligned} \text{maximize} \quad & 4x_1^{1/2} + 8x_2^{1/2} \\ \text{subject to} \quad & 2x_1 + 4x_2 = 6 \end{aligned}$$

- We have seen models like this before, with an **objective function** to be maximized/minimized, and **equality constraints** defining relationships between the variables — e.g. profit maximization
- Sometimes we can solve these models by first substituting the equality constraint into the objective function, and then finding the minimum/maximum of the resulting objective function
- This isn't always possible, especially when the equality constraint is complex
- Instead, we can use **the method of Lagrange multipliers**

2 The Lagrange multiplier method – 1 equality constraint

$$\begin{aligned} \text{minimize/maximize} \quad & f(x_1, \ldots, x_n) \\ \text{subject to} \quad & g(x_1, \ldots, x_n) = c \end{aligned}$$

- **Step 1.** Introduce the **Lagrange multiplier** λ and form the **Lagrangian function** Z:

$$Z(x_1, \ldots, x_n, \lambda) = f(x_1, \ldots, x_n) + \lambda\big[c - g(x_1, \ldots, x_n)\big]$$

- **Step 2.** Find the **critical points** by solving the system of equations implied by the first-order necessary condition:

$$\nabla Z(x_1, \ldots, x_n, \lambda) = 0 \qquad \text{or equivalently} \qquad \begin{aligned} \frac{\partial Z}{\partial x_1}(x_1, \ldots, x_n, \lambda) &= 0 \\ &\vdots \\ \frac{\partial Z}{\partial x_n}(x_1, \ldots, x_n, \lambda) &= 0 \\ \frac{\partial Z}{\partial \lambda}(x_1, \ldots, x_n, \lambda) &= 0 \end{aligned}$$

- **Step 3.** Classify each critical point as a local minimum or local maximum by applying the second-order sufficient condition:
  - The **bordered Hessian matrix** $\overline{H}$ is

$$\overline{H} = \begin{bmatrix} 0 & \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & \cdots & \frac{\partial g}{\partial x_n} \\ \frac{\partial g}{\partial x_1} & \frac{\partial^2 Z}{\partial x_1^2} & \frac{\partial^2 Z}{\partial x_1 x_2} & \cdots & \frac{\partial^2 Z}{\partial x_1 \partial x_n} \\ \frac{\partial g}{\partial x_2} & \frac{\partial^2 Z}{\partial x_2 \partial x_1} & \frac{\partial^2 Z}{\partial x_2^2} & \cdots & \frac{\partial^2 Z}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g}{\partial x_n} & \frac{\partial^2 Z}{\partial x_n \partial x_1} & \frac{\partial^2 Z}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 Z}{\partial x_n^2} \end{bmatrix}$$

  - The i**th bordered leading principal minor** of $\overline{H}$ — denoted by $|\overline{H}_i|$ — is the determinant of the square submatrix formed by the first $i+1$ rows and columns of $\overline{H}$
  - Let $(a_1, \ldots, a_n)$ be a critical point found in Step 2. Then

    (i) $f(a_1, a_2, \ldots, a_n)$ is a local minimum if

$$|\overline{H}_2| < 0 \quad |\overline{H}_3| < 0 \quad \cdots \quad |\overline{H}_n| < 0$$

    (ii) $f(a_1, a_2, \ldots, a_n)$ is a local maximum if

$$|\overline{H}_2| > 0 \quad |\overline{H}_3| < 0 \quad |\overline{H}_4| > 0 \quad \text{etc.}$$

Example 1. Use the Lagrange multiplier method to find the local optima of

$$\begin{aligned} \text{minimize/maximize} \quad & 4x_1^{1/2} + 8x_2^{1/2} \\ \text{subject to} \quad & 2x_1 + 4x_2 = 6 \end{aligned}$$

Step 1. Introduce the Lagrange multiplier λ and form the Lagrangian function Z.

- The Lagrangian function Z is

Step 2. Find the critical points.

- The gradient of Z is

- The first-order necessary condition tells us that the critical points must satisfy

- Therefore, we have one critical point:

Step 3. Classify the critical points as a local minimum or local maximum.

- The bordered Hessian is

- The bordered Hessian at the critical point $(x_1, x_2, \lambda) = (1, 1, 1)$ is

- The bordered leading principal minors $|\overline{H}_2|, |\overline{H}_3|, \ldots$ are

- Therefore,

Example 2. Use the Lagrange multiplier method to find the local optima of

$$\begin{aligned} \text{minimize/maximize} \quad & x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} \quad & 2x_1 + x_2 + 4x_3 = 168 \end{aligned}$$

(next page for more space)

3 What's up with λ?

- Consider the optimization problem in Example 2:

$$\begin{aligned} \text{minimize/maximize} \quad & x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} \quad & 2x_1 + x_2 + 4x_3 = 168 \end{aligned}$$

- Using the first-order necessary condition, we found one critical point at $(x_1, x_2, x_3, \lambda) = (16, 8, 32, 16)$
- Using the second-order sufficient condition, we found that $f(16, 8, 32) = 1344$ is a local minimum
- If we change the problem by increasing the constraint RHS, does the local minimum increase or decrease?
  - This is known as **sensitivity analysis** – how does your optimal solution change when the parameters of your optimization problem change?
  - e.g. What happens if we have a larger budget? Larger production quota?
- It turns out that λ is the rate of change in the optimal value with respect to the constraint RHS
- So for the problem in Example 2, increasing the constraint RHS
- More generally:
  - Consider the generic optimization problem with 1 equality constraint:

$$\begin{aligned} \text{minimize/maximize} \quad & f(x_1, \dots, x_n) \\ \text{subject to} \quad & g(x_1, \dots, x_n) = c \end{aligned}$$

  - Let $(x_1^*, \dots, x_n^*, \lambda^*)$ be a critical point
  - Suppose $f(x_1^*, \dots, x_n^*)$ is a local optimum
  - Then λ^* is the rate of change of this local optimum with respect to c
  - λ^* is known as the **marginal cost** or **shadow price** of the constraint