SM286A – Mathematics for Economics
Asst. Prof. Nelson Uhan

Fall 2015

Lesson 21. Optimization of Functions of n Variables

0 Warm up

Example 1. Find the local optima of $f(x,y) = 12x + 18y - 2x^2 - xy - 2y^2$.

1 An economic application: profit maximization for a multiproduct firm

- There are <u>many</u> applications of optimization to economics
- A classic example: profit maximization
- Consider a firm that produces and sells two products
- Prices of these products are exogenously determined
- Variables:

R = revenue

C = cost

Q_1 = quantity of product 1 produced

Q_2 = quantity of product 2 produced

- Model:

$$\begin{aligned} \text{maximize} \quad & R - C \\ \text{subject to} \quad & R = 12Q_1 + 18Q_2 \\ & C = 2Q_1^2 + Q_1Q_2 + 2Q_2^2 \end{aligned}$$

- The unit price of product 1 is and the unit price of product 2 is
- The marginal cost of product 1 is
- The marginal cost of product 2 is
- The production costs of the two products are related to each other!
- We can write profit as a function of Q_1 and Q_2:
- We want to maximize profit π – we already did this in Example 1!

$$f \leftrightarrow \pi \qquad\qquad x \leftrightarrow Q_1 \qquad\qquad y \leftrightarrow Q_2$$

- Locally optimal production plan and profit:
- Looking ahead:
    - How do we know if the local maximum we found is an **absolute maximum**?
    - What if we have 3 products? 100 products? n products?

2 The gradient and the first-order necessary condition

- Let f be a function of n variables
- Let's call these variables $x_1, x_2, \dots, x_n$
- $f(a_1, a_2, \dots, a_n)$ is a **local minimum** of f if

$$f(a_1, a_2, \dots, a_n) \leq f(x_1, x_2, \dots, x_n) \quad \text{for all } (x_1, x_2, \dots, x_n) \text{ “near” } (a_1, a_2, \dots, a_n)$$

- $f(a_1, a_2, \dots, a_n)$ is a **local maximum** of f if

$$f(a_1, a_2, \dots, a_n) \geq f(x_1, x_2, \dots, x_n) \quad \text{for all } (x_1, x_2, \dots, x_n) \text{ “near” } (a_1, a_2, \dots, a_n)$$

- Let's assume that all the first and second partial derivatives always exist
- The **gradient** of f is the vector

- In words, $\dfrac{\partial f}{\partial x_i}(a_1, a_2, \ldots, a_n)$ is

- Intuitively, the rate of change at a local minimum or local maximum should be zero in all directions
- **First-order necessary condition.** If $(a_1, a_2, \ldots, a_n)$ is a local minimum or local maximum of f, then $\nabla f(a_1, a_2, \ldots, a_n) = 0$, or equivalently

$$\frac{\partial f}{\partial x_1}(a_1, \ldots, a_n) = 0 \qquad \frac{\partial f}{\partial x_2}(a_1, \ldots, a_n) = 0 \qquad \cdots \qquad \frac{\partial f}{\partial x_n}(a_1, \ldots, a_n) = 0$$

- The points that satisfy the first-order necessary condition are called **critical points**
- Note that this is just a more general version of what we had for functions with 1 or 2 variables

Example 2. Find the critical points of $f(x_1, x_2, x_3) = e^{2x_1} + e^{-x_2} + e^{x_3^2} - 2x_1 - 2e^{x_3} + x_2$.

3 The Hessian and the second-order sufficient condition

- How do we know if a critical point is a local minimum or a local maximum?
- We need a "second derivative test" for n variables
- The **Hessian matrix** of f is

$$H = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Example 3. Find the Hessian matrix of $f(x_1, x_2, x_3) = e^{2x_1} + e^{-x_2} + e^{x_3^2} - 2x_1 - 2e^{x_3} + x_2$. Recall from Example 2 that

$$\frac{\partial f}{\partial x_1} = 2e^{2x_1} - 2 \qquad \frac{\partial f}{\partial x_2} = -e^{-x_2} + 1 \qquad \frac{\partial f}{\partial x_3} = 2x_3 e^{x_3^2} - 2e^{x_3}$$

- The i**th leading principal minor** of H – denoted by $|H_i|$ – is the determinant of the square submatrix formed by the first i rows and columns of H

Example 4. Find all of the leading principal minors of H from Example 3 at $(x_1, x_2, x_3) = (0, 0, 1)$, the critical point of f found in Example 2.

- **Second-order sufficient condition.** If $(a_1, a_2, \dots, a_n)$ satisfies the first-order necessary condition, then
  (i) $f(a_1, a_2, \dots, a_n)$ is a local minimum if all the leading principal minors of H are positive, i.e.

$$|H_1| > 0 \quad |H_2| > 0 \quad \cdots \quad |H_n| > 0$$

  (ii) $f(a_1, a_2, \dots, a_n)$ is a local maximum if the first leading principal minor is negative, and the remaining leading principal minors alternate in sign, i.e.

$$|H_1| < 0 \quad |H_2| > 0 \quad |H_3| < 0 \quad \text{etc.}$$

Example 5. Is $(x_1, x_2, x_3) = (0, 0, 1)$, the critical point of f found in Example 2, a local minimum or a local maximum?

- Note that the second-order sufficient condition is just a more general version of the second derivative test we had for functions with 1 or 2 variables
- For a function $f(x)$ with 1 variable, the Hessian is and so
  - (i) "all the leading principal minors of H are positive" means
  - (ii) "the first leading principal minor is negative, and the remaining leading principal minors alternate in sign" means
- For a function $f(x, y)$ with 2 variables, the Hessian is

  and so
  - (i) "all the leading principal minors of H are positive" means
  - (ii) "the first leading principal minor is negative, and the remaining leading principal minors alternate in sign" means