Lagrange Multiplier (Part II)

Liming Pang

1 Why the Method of Lagrange Multiplier Works

Now let’s explain why the method of Lagrange multiplier works. We will study the situation for optimizing a to variable function $f(x, y)$ under the constraint $g(x, y) = c$. The cases with more variables or more constraints can be shown in the similar fashion.

If $(x_0, y_0)$ is a maximum/minimum point of $f(x, y)$ under the constraint $g(x, y) = c$, we need to show that $\nabla f(x_0, y_0)$ is parallel to $\nabla g(x_0, y_0)$. Suppose they are not parallel. Since gradient is normal to level sets, we see that the assumption implies the level set $f(x, y) = f(x_0, y_0)$ and the level set $g(x, y) = c$ are not tangent at the intersection $(x_0, y_0)$. Then it means on each side of $f(x, y) = f(x_0, y_0)$, there are points on $g(x, y) = c$, and for points on one of the two sides, the values under $f$ is larger than $f(x_0, y_0)$. This can be shown as in the following figure. The following figure shows what’s going on.
We can also draw a 3-dimensional figure to illustrate:

\[ z = f(x, y) \]

Algebraically, we can apply the Implicit Function Theorem. By assumption, \( \nabla g(x_0, y_0) = (\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0)) \neq \vec{0} \), we may assume \( \frac{\partial f}{\partial y}(x_0, y_0) \neq 0 \). This condition implies \( g(x, y) = c \) implicitly defines a function \( y = h(x) \) near \( (x_0, y_0) \) such that \( g(x, h(x)) = 0 \) and \( y_0 = h(x_0) \).

By Implicit Function Theorem, we know that \( h'(x_0) = -\frac{\frac{\partial g}{\partial x}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)} \). Define \( \phi(x) = f(x, h(x)) \), then the assumption that \( (x_0, y_0) \) is an extreme point for the constraint optimization indicates that \( x_0 \) is an extreme point for \( \phi(x) \), so \( \phi'(x) = 0 \).

By the Chain Rule, we see

\[
0 = \phi'(x) = \frac{\partial f}{\partial x}(x_0, y_0) + \frac{\partial f}{\partial y}(x_0, y_0)h'(x_0) = \frac{\partial f}{\partial x}(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)\frac{\partial g}{\partial x}(x_0, y_0)
\]

i.e.

\[
\frac{\partial f}{\partial x}(x_0, y_0) \frac{\partial g}{\partial y}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) \frac{\partial g}{\partial x}(x_0, y_0)
\]

This implies \( \nabla f(x_0, y_0) \) is parallel to \( \nabla g(x_0, y_0) \), since if \( \frac{\partial g}{\partial x}(x_0, y_0) \neq 0 \), we see \( \frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) \frac{\partial g}{\partial x}(x_0, y_0) \). Let \( \lambda \neq 0 \) be this ratio; if \( \frac{\partial g}{\partial x}(x_0, y_0) = 0 \), we get \( \frac{\partial f}{\partial x}(x_0, y_0) = 0 \) as well, this is the case for \( \lambda = 0 \).
2 Interpretation of the Lagrange Multiplier

In this section we will explore what information the Lagrange multiplier $\lambda$ can bring to us.

Assume $(x_0, y_0)$ is a maximal point for $f(x, y)$ subject to constraint $g(x, y) = c$. We thus know:

$$\begin{align*}
\nabla f(x_0, y_0) &= \lambda \nabla g(x_0, y_0) \\
g(x_0, y_0) &= c
\end{align*}$$

Now we would like to know how the constraint optimal value changes if we change the constraint constant from $c$ by a small change to some $c'$. Recall that given a function $f(x, y)$, there is a differential

$$df = \frac{\partial f}{\partial x}(x_0, y_0)dx + \frac{\partial f}{\partial y}(x_0, y_0)dy$$

which gives a good estimation of how the value of the function changes if $x, y$ are changed by small amount $dx, dy$ respectively from $x_0, y_0$.

Now assume the new extreme point of $f$ subject to the constraint $c'$ is obtained from $(x_0, y_0)$ by a change of $(dx, dy)$, we get the corresponding change in the value of $f$ to be

$$df = \frac{\partial f}{\partial x}(x_0, y_0)dx + \frac{\partial f}{\partial y}(x_0, y_0)dy$$

$$= \left( \frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right) \cdot (dx, dy)$$

$$= \lambda \left( \frac{\partial g}{\partial x}(x_0, y_0), \frac{\partial g}{\partial y}(x_0, y_0) \right) \cdot (dx, dy)$$

$$= \lambda \left( \frac{\partial g}{\partial x}(x_0, y_0)dx + \frac{\partial g}{\partial y}(x_0, y_0)dy \right)$$

$$= \lambda dg$$

$$= \lambda (c' - c)$$

The above computation leads to the following theorem:

**Theorem 1.** If the constraint is changed by a small amount from $g(x, y) = c$ to $g(x, y) = c'$, then the optimal value will change by about $\lambda(c' - c)$, where $\lambda$ is the Lagrange multiplier.
Another way of explanation is that if we denote $P$ to be the optimal value of $f(x, y)$ subject to the constraint $g(x, y) = c$, then $P$ can be regarded as a function of $c$, i.e. $P = P(c)$, then the above argument indicates $\lambda = \frac{dP}{dc}$.

**Example 2.** There is an economic understanding of the above discussion:

A company is producing two brands, $A$ and $B$. When producing $x$ units of $A$ and $y$ units of $B$, the cost is $g(x, y)$ dollars and the profit is $f(x, y)$ dollars. Currently the company inputs $c$ dollars everyday, and the daily profit is $P$ (i.e. the maximum value for $f(x, y)$ subject to the constraint $g(x, y) = c$ is $P$). If we know at this point the Lagrange multiplier $\lambda$, then if the daily input increases by 1 dollar, the daily profit will increase by about $\lambda$ dollars.

In economics, the Lagrange multiplier is called the **shadow price** of the resource. It is the marginal profit with respect to the budget.

**Example 3.** Now let’s compute a concrete example to see how well the Lagrange multiplier estimates the actual change of optimal value. We are re-visiting Example 4 in Part I:

A person has utility function $u(x, y) = 10xy + 5x + 2y$. Suppose the price for one unit of $x$ is 2 dollars and the price for one unit of $y$ is 5 dollars. If the person has 100 dollars that can be spent on $x$ and $y$, find $x$ and $y$ that maximize the utility.

The question is to maximize $u(x, y) = 10xy + 5x + 2y$ under the constraint $2x + 5y = 100$.

Let $g(x, y) = 2x + 5y$, by the Lagrange method, we have

$$
\begin{align*}
\nabla u(x, y) &= \lambda \nabla g(x, y) \\
\n\nabla g(x, y) &= 100
\end{align*}
$$

i.e.

$$
\begin{align*}
10y + 5 &= \lambda \times 2 \\
10x + 2 &= \lambda \times 5 \\
2x + 5y &= 100
\end{align*}
$$

We get $\lambda = 51.45$, $x = 25.525$, $y = 9.79$, the maximal utility is $u(25.525, 9.79) = 2646.1025$.

Now assume the person has 1 more dollar in budget, we have

$$
\begin{align*}
\nabla u(x, y) &= \lambda \nabla g(x, y) \\
\n\nabla g(x, y) &= 101
\end{align*}
$$
i.e.

\[
\begin{align*}
10y + 5 &= \lambda \times 2 \\
10x + 2 &= \lambda \times 5 \\
2x + 5y &= 101
\end{align*}
\]

We get $\lambda = 51.95$, $x = 25.7755$, $y = 9.89$, the maximal utility is $u(25.7755, 9.89) = 2697.85445$

So the actual increase of maximal utility is $2697.85445 - 2646.1025 = 51.75195$, and the initial Lagrange multiplier for budget being 100 dollars is 51.45, which is close to the actual increase.