Constrained Optimization in Two Variables

James K. Peterson

Department of Biological Sciences and Department of Mathematical Sciences
Clemson University

November 17, 2016

Outline

Constrained Optimization

What Does the Lagrange Multiplier Mean?
Let's look at the problem of finding the minimum or maximum of a function $f(x, y)$ subject to the constraint $g(x, y) = c$ where $c$ is a constant. Let's suppose we have a point $(x_0, y_0)$ where an extremum value occurs and assume $f$ and $g$ are differentiable at that point.

Then, at another point $(x, y)$ which satisfies the constraint, we have

$$g(x, y) = g(x_0, y_0) + g_x(x_0, y_0)(x - x_0) + g_y(x_0, y_0)(y - y_0) + E_g(x, y, x_0, y_0)$$

But $g(x, y) = g(x_0, y_0) = c$, so we have

$$0 = g_x^0(x - x_0) + g_y^0(y - y_0) + E_g(x, y, x_0, y_0)$$

where we let $g_x^0 = g_x(x_0, y_0)$ and $g_y^0 = g_y(x_0, y_0)$.

Now given an $x$, we assume we can find a $y$ values so that $g(x, y) = c$ in $B_r(x)$ for some $r > 0$. Of course, this value need not be unique. Let $\phi(x)$ be the function defined by

$$\phi(x) = \{ y | g(x, y) = c \}$$

For example, if $g(x, y) = c$ was the function $x^2 + y^2 = 25$, then

$$\phi(x) = \pm \sqrt{25 - x^2} \text{ for } -5 \leq x \leq 5.$$  

Clearly, we can't find a full circle $B_r(x)$ when $x = -5$ or $x = 5$, so let's assume the point $(x_0, y_0)$ where the extremum value occurs does have such a local circle around it where we can find corresponding $y$ values for all $x$ values in $B_r(x_0)$.

So locally, we have a function $\phi(x)$ defined around $x_0$ so that $g(x, \phi(x)) = c$ for $x \in B_r(x_0)$. Then we have

$$0 = g_x^0(x - x_0) + g_y^0(\phi(x) - \phi(x_0)) + E_g(x, \phi(x), x_0, \phi(x_0))$$

Now divide through by $x - x_0$ to get

$$0 = \frac{g_x^0}{x - x_0} + \frac{g_y^0}{x - x_0} \left( \frac{\phi(x) - \phi(x_0)}{x - x_0} \right) + \frac{E_g(x, \phi(x), x_0, \phi(x_0))}{x - x_0}$$

Thus,

$$g_y^0 \left( \frac{\phi(x) - \phi(x_0)}{x - x_0} \right) = -g_x^0 \frac{E_g(x, \phi(x), x_0, \phi(x_0))}{x - x_0}$$

and assuming $g_x^0 \neq 0$ and $g_y^0 \neq 0$, we can solve to find
\[
\frac{\phi(x) - \phi(x_0)}{x - x_0} = -\frac{g_x^0}{g_y^0} - \frac{1}{g_y^0} \frac{E_g(x, \phi(x), x_0, \phi(x_0))}{x - x_0}
\]

Since \( g \) is differentiable at \((x_0, \phi(x_0))\), as \( x \to x_0 \), \( \frac{E_g(x, \phi(x), x_0, \phi(x_0))}{x - x_0} \to 0 \). Thus, we have \( \phi \) is differentiable and therefore also continuous with

\[
\phi'(x_0) = \lim_{x \to x_0} \frac{\phi(x) - \phi(x_0)}{x - x_0} = -\frac{g_x^0}{g_y^0}
\]

Now since both \( g_x^0 \neq 0 \) and \( g_y^0 \neq 0 \), the fraction \( -\frac{g_x^0}{g_y^0} \neq 0 \) too. This means locally \( \phi(x) \) is either strictly increasing or strictly decreasing; i.e. is a strictly monotone function.

Let’s assume \( \phi'(x_0) > 0 \) and so \( \phi \) is increasing on some interval \((x_0 - R, x_0 + R)\). Let’s also assume the extreme value is a minimum, so we know on this interval \( f(x, y) \geq f(x_0, y_0) \) with \( g(x, y) = c \). This means

\[
f(x, \phi(x)) - f(x_0, \phi(x_0)) \geq 0 \quad \text{on} \quad (x_0 - R, x_0 + R). \quad \text{Now do an expansion for } f \text{ to get}
\]

\[
f(x, \phi(x)) = f(x, \phi(x_0)) + f_x^0(x - x_0) + f_y^0(\phi(x) - \phi(x_0)) + E_f(x, \phi(x), x_0, \phi(x_0))
\]

where \( f_x^0 = f_x(x_0, \phi(x_0)) \) and \( f_y^0 = f_y(x_0, \phi(x_0)) \). This implies

\[
f(x, \phi(x)) - f(x_0, \phi(x_0)) = f_x^0(x - x_0) + f_y^0(\phi(x) - \phi(x_0)) + E_f(x, \phi(x), x_0, \phi(x_0))
\]

Now we have assumed \( f(x, \phi(x)) - f(x_0, \phi(x_0)) \geq 0 \), so we have

\[
f_x^0(x - x_0) + f_y^0(\phi(x) - \phi(x_0)) + E_f(x, \phi(x), x_0, \phi(x_0)) \geq 0
\]

If \( x - x_0 > 0 \) we find

\[
f_x^0 + f_y^0 \left( \frac{\phi(x) - \phi(x_0)}{x - x_0} \right) + \frac{E_f(x, \phi(x), x_0, \phi(x_0))}{x - x_0} \geq 0
\]

Now take the limit as \( x \to x_0^+ \) to find

\[
f_x^0 + f_y^0 \phi'(x_0) \geq 0
\]
When \( x - x_0 < 0 \), we argue similarly to find
\[
f_x^0 + f_y^0 \phi'(x_0) \leq 0
\]
Combining, we have
\[
0 \leq f_x^0 + f_y^0 \phi'(x_0) \leq 0
\]
which tells us at this extreme value we have the equation
\[
f_x^0 + f_y^0 \phi'(x_0) = 0
\]
This implies \( f_y^0 = -\frac{f_x^0}{\phi'(x_0)} \). Next note
\[
\begin{bmatrix} f_x^0 \\ f_y^0 \end{bmatrix} = \begin{bmatrix} -\frac{f_x^0}{\phi'(x_0)} \end{bmatrix} = \begin{bmatrix} -f_x^0 \left\{ -\frac{g_y^0}{g_x^0} \right\} \end{bmatrix} = \frac{f_x^0}{g_x^0} \begin{bmatrix} g_x^0 \\ g_y^0 \end{bmatrix}
\]
This says there is a scalar \( \lambda = \frac{f_x^0}{g_x^0} \) so that
\[
\nabla(f)(x_0, y_0) = \lambda \nabla(g)(x_0, y_0)
\]
where \( g(x_0, y_0) = c \). The value \( \frac{f_x^0}{g_x^0} \) is called the Lagrange Multiplier for this extremal Problem. This is the basis for the Lagrange Multiplier Technique for a constrained optimization problem.

The value \( \lambda \) above is called a Lagrange Multiplier of the extremum problem we are trying to solve.

We can do a similar sort of analysis in the case the extremum is a maximum too. Our analysis assumes that the the point \((x_0, y_0)\) where the extremum occurs is like an interior point in the \( \{(x, y) | g(x, y) = c\} \).
That is, we assume for such an \( x_0 \) there is an interval \( B_r(x_0) \) with any \( x \in B_r(x_0) \) having a corresponding value \( y = \phi(x) \) so that \( g(x, \phi(x)) = c \). So the argument does not handle boundary points such as the \( \pm 5 \) is our previous example.

To implement the Lagrange Multiplier technique, we define a new function
\[
H(x, y, \lambda) = f(x, y) + \lambda (g(x, y) - c)
\]
The critical points we see are the ones where the gradient of \( f \) is a multiple of the gradient of \( g \) for the reasons we discussed above. Note the \( \lambda \) we use here is the negative of the Lagrange Multiplier.
To find them, we set the partials of $H$ equal to zero.

$$\frac{\partial H}{\partial x} = 0 \implies \frac{\partial f}{\partial x} = -\lambda \frac{\partial g}{\partial x}$$

$$\frac{\partial H}{\partial y} = 0 \implies \frac{\partial f}{\partial y} = -\lambda \frac{\partial g}{\partial x}$$

$$\frac{\partial H}{\partial \lambda} = 0 \implies g(x, y) = c$$

The first two lines are the statement that the gradient of $f$ is a multiple of the gradient of $g$ and the third line is the statement that the constraint must be satisfied. As usual, solving these three nonlinear equations gives the interior point solutions and we always must also include the boundary points that solve the constraints as well.

Example
Extremize $2x + 3y$ subject to $x^2 + y^2 = 25$.

Solution
Set $H(x, y, \lambda) = (2x + 3y) + \lambda(x^2 + y^2 - 25)$. The critical point equations are then

$$\frac{\partial H}{\partial x} = 2 + \lambda(2x) = 0$$

$$\frac{\partial H}{\partial y} = 3 + \lambda(2y) = 0$$

$$\frac{\partial H}{\partial \lambda} = x^2 + y^2 - 25 = 0$$
Solution
Solving for \( \lambda \), we find \( \lambda = -1/x = -3/(2y) \). Thus, \( x = 2y/3 \).

Using this relationship in the constraint, we find \((2y/3)^2 + y^2 = 25\) or \(13y^2/9 = 25\). Thus, \( y^2 = 225/13 \) or \( y = \pm 15/\sqrt{13} \). This means \( x = \pm 10/\sqrt{13} \).

We find
\[
\begin{align*}
&f(10/\sqrt{13}, 15/\sqrt{13}) = 45/\sqrt{13} = 12.48 \\
&f(-10/\sqrt{13}, -15/\sqrt{13}) = -45/\sqrt{13} = -12.48 \\
&f(-5, 0) = -10 \\
&f(5, 0) = 10 \\
&f(0, -5) = -15 \\
&f(0, 5) = 15.
\end{align*}
\]

So the extreme values here occur at the boundary points \((0, \pm 5)\).

Example
Extremize \( \sin(xy) \) subject to \( x^2 + y^2 = 1 \).

Solution
Set \( H(x, y, \lambda) = (\sin(xy)) + \lambda(x^2 + y^2 - 1) \). The critical point equations are then
\[
\begin{align*}
\frac{\partial H}{\partial x} &= y \cos(xy) + \lambda(2x) = 0 \\
\frac{\partial H}{\partial y} &= x \cos(xy) + \lambda(2y) = 0 \\
\frac{\partial H}{\partial \lambda} &= x^2 + y^2 - 1 = 0
\end{align*}
\]
Solution

Solving for $\lambda$ in the first equation, we find $\lambda = -y \cos(xy)/2x$. Now use this in the second equation to find $x \cos(xy) - 2y^2 \cos(xy)/2x = 0$. Thus, $\cos(xy)(x - y^2/x) = 0$.

The first case is $\cos(xy) = 0$. This implies $xy = \pi/2 + 2n\pi$. These are rectangular hyperbolas and the closest one to the origin is $(\sqrt{\pi/2}, \sqrt{\pi/2}) = (1, 25, 1, 25)$ whose closest point to the origin is $(\sqrt{\pi/2}, \sqrt{\pi/2}) = (1, 25, 1, 25)$. This point is outside of the constraint set, so this case does not matter.

The second case is $x^2 = y^2$ which, using the constraint, implies $x^2 = 1/2$. Thus there are four possibilities: $(1/\sqrt{2}, 1/\sqrt{2})$, $(1/\sqrt{2}, -1/\sqrt{2})$, $(-1/\sqrt{2}, 1/\sqrt{2})$ and $(-1/\sqrt{2}, 1/\sqrt{2})$.

The other possibilities are the circle’s boundary points $(0, \pm 1)$ and $(\pm 1, 0)$.

We find

$\sin(1/\sqrt{2}, 1/\sqrt{2}) = \sin(-1/\sqrt{2}, -1/\sqrt{2}) = \sin(1/2) = .48$
$\sin(1/\sqrt{2}, -1/\sqrt{2}) = \sin(-1/\sqrt{2}, 1/\sqrt{2}) = -\sin(1/2) = -.48$
$\sin(\pm 1, 0) = 0$, $\sin(0, \pm 1) = 0$

So the extreme values occur here at the interior points of the constraint.

Let’s change our extremum problem by assuming the constraint constant is changing as a function of $x$; i.e. the problem is now find the extreme values of $f(x, y)$ subject to $g(x, y) = C(x)$. We can use our tangent plane analysis like before. We assume the extreme value for constraint value $C(x_0)$ occurs at an interior point of the constraint. We have

$$g(x, \phi(x)) - C(x) = 0$$

Then

$$0 = g_x^0 + g_y^0 \left( \frac{\phi(x) - \phi(x_0)}{x - x_0} \right) + \left( \frac{E_g(x, \phi(x), x_0, \phi(x_0))}{x - x_0} \right) - \left( \frac{C(x) - C(x_0)}{x - x_0} \right)$$

We assume the constraint bound function $C$ is differentiable and so letting $x \to x_0$, we find

$$g_x^0 + g_y^0 \phi'(x_0) = C'(x_0)$$

Now assume $(x_1, \phi(x_1))$ extremizes $f$ for the constraint bound value $C(x_1)$. 
We have
\[
f(x_1, \phi(x_1)) - f(x_0, \phi(x_0)) = f_x^0(x_1 - x_0) + f_y^0(\phi(x_1) - \phi(x_0)) + E_f(x_1, \phi(x_1), x_0, \phi(x_0))
\]
or
\[
\left(\frac{f(x_1, \phi(x_1)) - f(x_0, \phi(x_0))}{x_1 - x_0}\right) = f_x^0 + f_y^0\left(\frac{\phi(x_1) - \phi(x_0)}{x_1 - x_0}\right) + \left(\frac{E_f(x_1, \phi(x_1), x_0, \phi(x_0))}{x_1 - x_0}\right)
\]

Now assuming \(C(x) \neq 0\) locally around \(x_0\), we can write
\[
\left(\frac{f(x_1, \phi(x_1)) - f(x_0, \phi(x_0))}{C(x_1) - C(x_0)}\right)\left(\frac{C(x_1) - C(x_0)}{x_1 - x_0}\right) = f_x^0 + f_y^0\left(\frac{\phi(x_1) - \phi(x_0)}{x_1 - x_0}\right) + \left(\frac{E_f(x_1, \phi(x_1), x_0, \phi(x_0))}{x_1 - x_0}\right)
\]

Now let \(\Theta(x)\) be defined by \(\Theta(x) = f(x, \phi(x))\) when \((x, \phi(x))\) is the extreme value for the problem of extremizing \(f(x, y)\) subject to \(g(x, y) = C(x)\). This is called the **Optimal Value Function** for this problem. Rewriting, we have
\[
\left(\frac{\Theta(x_1) - \Theta(x_0)}{C(x_1) - C(x_0)}\right)\left(\frac{C(x_1) - C(x_0)}{x_1 - x_0}\right) = f_x^0 + f_y^0\left(\frac{\phi(x_1) - \phi(x_0)}{x_1 - x_0}\right) + \left(\frac{E_f(x_1, \phi(x_1), x_0, \phi(x_0))}{x_1 - x_0}\right)
\]

Now let \(x_1 \to x_0\). We obtain
\[
\lim_{x_1 \to x_0} \left(\frac{\Theta(x_1) - \Theta(x_0)}{C(x_1) - C(x_0)}\right) C'(x_0) = f_x^0 + f_y^0 \phi'(x_0)
\]

Thus, the rate of change of the optimal value with respect to change in the constraint bound, \(\frac{d\Theta}{dC}\) is well - defined and
\[
\frac{d\Theta}{dC}(C_0) C'(x_0) = f_x^0 + f_y^0 \phi'(x_0)
\]

where \(C_0\) is the value \(C(x_0)\).
Assuming \( C'(x_0) \neq 0 \), we have

\[
\frac{d\Theta}{dC}(C_0) = \frac{f^0_x}{C'(x_0)} + \frac{f^0_y}{C'(x_0)} \phi'(x_0)
\]

But we know at the extreme value for \( x_0 \) that \( f^0_y = \frac{g^0_y}{g^0_x} f^0_x \). Thus,

\[
\frac{d\Theta}{dC}(C_0) = \frac{f^0_x}{C'(x_0)} + \frac{g^0_y}{g^0_x} \frac{f^0_x}{C'(x_0)} \phi'(x_0) = \frac{f^0_x}{C'(x_0)} \left( 1 + \frac{g^0_y}{g^0_x} \phi'(x_0) \right)
\]

But the Lagrange Multiplier \( \lambda_0 \) for extremal value for \( x_0 \) is \( \lambda_0 = \frac{f^0_x}{g^0_x} \). So

\[
\frac{d\Theta}{dC}(C_0) = \lambda_0 \left( g^0_x + g^0_y \phi'(x_0) \right) \frac{1}{C'(x_0)}
\]

But \( C'(x_0) = g^0_x + g^0_y \phi'(x_0) \). Hence, we find \( \frac{d\Theta}{dC}(C_0) = \lambda_0 \).

We see from our argument that the Lagrange Multiplier \( \lambda_0 \) is the rate of change of optimal value with respect to the constraint bound. There are cases:

- \( \text{sign}(\lambda_0) = \text{sign}(\frac{f^0_x}{g^0_x}) = \frac{+}{+} = + \) which implies the optimal value goes up if the constraint bound is perturbed.
- \( \text{sign}(\lambda_0) = \text{sign}(\frac{f^0_x}{g^0_x}) = \frac{+}{-} = + \) which implies the optimal value goes up if the constrained bound is perturbed.
- \( \text{sign}(\lambda_0) = \text{sign}(\frac{f^0_x}{g^0_x}) = \frac{-}{-} = - \) which implies the optimal value goes down if the constrained bound is perturbed.
- \( \text{sign}(\lambda_0) = \text{sign}(\frac{f^0_x}{g^0_x}) = \frac{-}{+} = - \) which implies the optimal value goes down if the constrained bound is perturbed.

Hence, we can interpret the Lagrange Multiplier as a **price**: the change in optimal value due to a change in constraint is essentially the **loss of value** experienced due to the constraint change. For example, if \( |\lambda_0| \) is very small, it says the rate of change of the optimal value with respect to constraint modification is small too. This implies the optimal value is insensitive to constraint modification.
37.1 Find the extreme values of $\sin(xy)$ subject to $x^2 + 2y^2 = 1$ using the Lagrange Multiplier Technique.

37.2 Find the extreme values of $\cos(xy)$ subject to $x^2 + 2y^2 = 1$ using the Lagrange Multiplier Technique.

37.3 If $a_n \to 1$ and $b_n \to 3$, use an $\epsilon - N$ proof to show $5a_n/b_n \to 5/3$.

37.4 If $a_n \to 10$ and $b_n \to 2$, use an $\epsilon - N$ proof to show $5a_n b_n \to 100$. 