§13.3 Lagrange Multipliers

We will consider constrained extreme value problems, i.e. the maximizing/minimizing problems where the variables are related by an equation called the constraint. Functions and constraints in concern may involve two or three variables. We may have different types of constrained extreme value problems such as

maximize/minimize $f(x, y)$ subject to $g(x, y) = 0$

maximize/minimize $f(x, y, z)$ subject to $g(x, y, z) \leq 0$ (or $g(x, y, z) \geq 0$)

In the second example above the constraint is given as $g \leq 0$. In this case we can seperate the region as $g < 0$ and $g = 0$, and then find the extreme values seperately for each of these regions. For the case $g < 0$, it is enough to find the ciritical and singular points of $f(x, y, z)$ satisfying the inequality $g(x, y, z) < 0$ (Recall Theorem 1 of §13.2). So our main focus is the set of points satisfying $g = 0$ which is clarified by the following theorem.

Theorem 1. Let $f(x, y)$ and $g(x, y)$ be functions with continuous first order partial derivatives around a point $P_0 = (x_0, y_0)$ satisfying $g(x_0, y_0) = 0$. Suppose that $f(x_0, y_0)$ is a local extremum of f on the set of points satisfying $g(x, y) = 0$. Also suppose that

i) P_0 is not an endpoint of the curve given by $g(x, y) = 0$, and

ii) $\nabla g(P_0) \neq 0$.

Then there exists λ_0 such that (x_0, y_0, λ_0) is a critical point of the function

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y).$$

Note that if the functions f and g involve three variables, then the theorem still holds.

Now let's explore the theorem. The critical points of $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ (Lagrange function) are the points (x, y, λ) satisfying

$$\begin{aligned}
L_1(x, y, \lambda) &= f_1(x, y) + \lambda g_1(x, y) = 0\\
L_2(x, y, \lambda) &= f_2(x, y) + \lambda g_2(x, y) = 0\\
L_3(x, y, \lambda) &= g(x, y) = 0.
\end{aligned}$$

The third equation guarantees that our point obeys the constraint. The first and second conditions can be combined and written in terms of gradients of f and g as

$$\nabla f(x, y) + \lambda \nabla g(x, y) = 0$$

So we look for the points on $g(x, y) = 0$ at which the gradients of f and g differ by a constant multiple ($-\lambda$ in the above notation).

Indeed this is the idea behind the proof of the above theorem; if we set $\mathbf{u}$ to be the projection of $\nabla f(a, b)$ on $\nabla g(a, b)$

$$\mathbf{u} = \text{ Proj }_{\nabla g} \nabla f$$

for some point (a, b) then f has a positive (respectively negative) directional derivative along $(\nabla f(a, b) - \mathbf{u})$ (respectively $-(\nabla f - \mathbf{u})$). So f can not have a local extreme value at (a, b).

Example 1. Find the absolute maximum of $f(x, y) = xy$ on the ellipse $x^2 + 2y^2 = 1$.

Solution: We need to

$$\text{maximize} \;\; f(x, y) = xy \;\; \text{subject to} \;\; g(x, y) = x^2 + 2y^2 - 1 = 0$$

Note that we can parametrize the ellipse $x^2 + 2y^2 = 1$, and then reduce the problem to maximizing a function in single variable (See §13.2). But let's solve the question by the method of Lgrange multipliers. Let

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y) = xy + \lambda(x^2 + 2y^2 - 1).$$

Then we have

$$\begin{aligned} L_1(x, y, \lambda) &= y + 2\lambda x = 0, \\ L_2(x, y, \lambda) &= x + 4\lambda y = 0, \\ L_3(x, y, \lambda) &= x^2 + 2y^2 - 1 = 0. \end{aligned}$$

By the first and the second equalites we obtain

$$y(1 - 8\lambda^2) = 0 \implies y = 0 \text{ or } \lambda = \pm 1/2\sqrt{2}$$

But $y = 0$ also impleis that $x = 0$ which violates the constraint $x^2 + 2y^2 = 1$. So we must have $\lambda = \pm 1/2\sqrt{2}$. So in any case we find $x = \pm\sqrt{2}y$. Plugging in the constraint we find the four points

$$(\pm 1/\sqrt{2}, \pm 1/2).$$

Our theorem gurantees that the maximum will be attained at one of these points, and in this case the maximum is

$$f(1/\sqrt{2}, 1/2) = f(-1/\sqrt{2}, -1/2) = 1/2\sqrt{2}$$

Example 2. Find the minimum of $f(x, y, z) = xy + z$ on the unit sphere.

Solution: Since the unit sphere $x^2+y^2+z^2 = 1$ is closed and bounded f has an absolute minimum on it. Our problem is to

$$\text{minimize} \quad f(x, y, z) = xy+z \quad \text{subject to } g(x, y, z) = x^2+y^2+z^2-1 = 0$$

The Lagrange function is $L(x, y, z, \lambda) = xy + z + \lambda(x^2 + y^2 + z^2 - 1)$. Then we have

$$\begin{aligned}
L_1(x, y, z, \lambda) &= y + 2\lambda x = 0, \\
L_2(x, y, z, \lambda) &= x + 2\lambda y = 0, \\
L_3(x, y, z, \lambda) &= 1 + 2\lambda z = 0, \\
L_4(x, y, z, \lambda) &= x^2 + y^2 + z^2 - 1 = 0.
\end{aligned}$$

The first and the second equations imply that $x = y = 0$ or $\lambda = \pm 1/2$. In the second case we see that $z = -2\lambda = \pm 1$ and again have that $x = y = 0$. We have two points $(0, 0, \pm 1)$, so the minimum is $f(0, 0, -1) = -1$

We may also consider problems with more than one constraint. For example if we assume that the problem

$$\begin{aligned}&\text{maximize (or minimize)}\ \ f(x,y,z)\\ &\text{subject to}\ \ g(x,y,z)=0 \text{ and } \ h(x,y,z)=0\end{aligned}$$

has a solution at $P_0 = (x_0, y_0, z_0)$ then the intersection of the constraints will be a curve with the tangent vector $\mathbf{T} = \nabla g(P_0) \times \nabla h(P_0)$ at P_0. Then $\nabla f(P_0)$ must be perpendicular to $\mathbf{T}$ (Otherwise f has nonzero directional derivatives along $\pm\mathbf{T}$, and so can not have a local extreme value). But since $\mathbf{T}$ is already perpendicular to $\nabla g(P_0)$ and $\nabla h(P_0)$, then $\nabla f(P_0)$ must be in the plane determined by $\nabla g(P_0)$ and $\nabla h(P_0)$. In short there exist λ_0 and μ_0 such that $(x_0, y_0, z_0, \lambda_0, \mu_0)$ is a critical point of the Lagrange function

$$L(x,y,z,\lambda,\mu) = f(x,y,z) + \lambda g(x,y,z) + \mu h(x,y,z).$$

Example 3. Maximize $f(x, y, z) = xyz$ subject to $x + y + z = 0$ and $x^2 + y^2 + z^2 = 6$.

Solution: The Lagrange function is $L(x, y, z, \lambda, \mu) = xyz + \lambda(x + y + z) + \mu(x^2 + y^2 + z^2 - 6)$. So we have

$$
\begin{aligned}
L_1 &= yz + \lambda + 2\mu x = 0, \\
L_2 &= xz + \lambda + 2\mu y = 0, \\
L_3 &= xy + \lambda + 2\mu z = 0, \\
L_4 &= x + y + z = 0, \\
L_5 &= x^2 + y^2 + z^2 - 6 = 0
\end{aligned}
$$

Subtracting L_2 from L_1 we find that $(y - x)(z - 2\mu) = 0$. Similarly we have

$$(z - x)(y - 2\mu) = 0, \;\; (z - y)(x - 2\mu) = 0$$

Considering all possible cases we find the following points

$$(1, 1, -2), (1, -2, 1), (-2, 1, 1), (-1, -1, 2), (-1, 2, -1), (2, -1, -1).$$

(Exercise: Verify these points). So the minimum (the maximum) is $f(1, 1, -2) = -2$ ($f(-1, -1, 2) = 2$).