5
D

Understanding Lagrange Multipliers

The method of Lagrange multipliers can be stated (for a function of two variables) as follows:

To find the maximum or minimum of $f(x,y)$ subject to the constraint equation $\phi(x,y) = C$,

1) Let $F(x,y) = f(x,y) - \lambda\,\phi(x,y)$.

(This is sometimes written as $F(x,y) = f(x,y) + \lambda\,\phi(x,y)$. We will see below why it does not matter.)

2) Set $\dfrac{\partial F}{\partial x} = 0$ and $\dfrac{\partial F}{\partial y} = 0$.

3) Solve these two equations together with the constraint equation to find x and y (and λ if desired – but finding λ is usually not necessary or interesting, as explained below).

Why does this method work, and what is the meaning of λ? For two variables, there is a simple way to visualize what is happening.

For conciseness, we let f_x denote the first partial derivative of f with respect to x, etc.

The requirement that $f(x,y)$ is an extremum and the constraint equation lead to two differential relations:

(1) $df = f_x\,dx + f_y\,dy = 0$

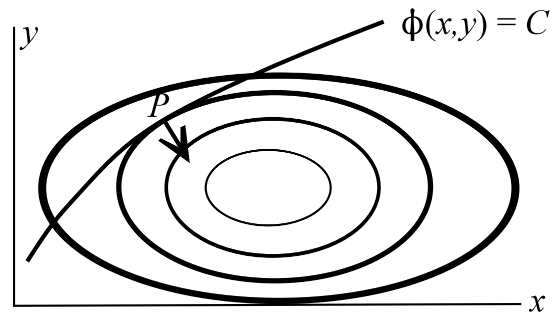
(2) $d\phi = \phi_x\,dx + \phi_y\,dy = 0$

Rearranging and taking ratios, we find that $f_x/\phi_x = f_y/\phi_y$, and we call this ratio λ:

(3) $$\frac{f_x}{\phi_x} = \frac{f_y}{\phi_y} = \lambda$$

Rearranging again, we obtain the fundamental equations of the method,

(4) $f_x - \lambda\,\phi_x = 0$ and $f_y - \lambda\,\phi_y = 0$



To see why this procedure works (at least for a function of two variables), consider the figure. The closed loops are contours of $f(x,y)$, i.e. curves of $f(x,y) = constant$, in the xy plane. The bold curve is the graph of the constraint equation $\phi(x,y) = C$. Assume that f increases toward the center of the figure.

The problem can be restated as follows:

To find an extremum of $f(x,y)$ subject to the constraint $\phi(x,y) = C$, move along the constraint curve until you arrive at the point P at which **the constraint curve is tangent to the local contour** of $f(x,y)$. (Of course one can draw a contour through any point x,y, not just the discrete contours shown in the figure.) In the case illustrated here, if you move forward or backward from that point along the constraint curve, the value of $f(x,y)$ will decrease from its value at P, so point P is the desired *maximum* of f. (Exercise: Convince yourself that if the constraint curve had a greater curvature than the contour curve at P, the extremum of f at P would be a *minimum*.)

To show that this condition is equivalent to the Lagrange multiplier method, we first note that if the two curves are *tangent* at P, the *normals* to the two curves are *parallel* at P. To find the normal to a curve, it is easiest to consider the **gradient** of a function:

$$\vec{\nabla} f \equiv \frac{\partial f}{\partial x}\hat{\boldsymbol{\imath}} + \frac{\partial f}{\partial y}\hat{\boldsymbol{\jmath}} \quad \text{(in 2 dimensions)}.$$

Using the gradient, we can write $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = \vec{\nabla} f \cdot d\vec{s} = \left|\vec{\nabla} f\right| \cos\theta \, ds$, where

$d\vec{s} = \hat{\boldsymbol{\imath}}\,dx + \hat{\boldsymbol{\jmath}}\,dy$ is a step in direction θ relative to the direction of $\vec{\nabla} f$. The direction of $\vec{\nabla} f$ is the direction (corresponding to $\theta = 0$) in which f is increasing most rapidly, i.e. the direction in which the directional derivative df/ds is largest. This is the direction normal to the local contour of f. The arrow in the figure represents this direction.

Similarly, $\vec{\nabla}\phi$ is a vector normal to the curve $\phi(x,y) = C$. Hence, when the curves $\phi(x,y) = C$ and $f(x,y) = \text{constant}$ are tangent to each other, $\vec{\nabla} f$ is parallel to $\pm\vec{\nabla}\phi$. This means that the two gradient vectors are proportional to each other, with some constant of proportionality:

$$\begin{aligned} \vec{\nabla} f \quad &= (\text{constant}) \times \vec{\nabla}\phi \\ \hat{\boldsymbol{\imath}} f_x + \hat{\boldsymbol{\jmath}} f_y &= (\text{constant}) \times \left\{\hat{\boldsymbol{\imath}}\,\phi_x + \hat{\boldsymbol{\jmath}}\,\phi_y\right\} \end{aligned}$$

If we call the constant λ, the second of these relations can be rearranged to give

$$f_x - \lambda\,\phi_x = 0 \text{ and } f_x - \lambda\,\phi_y = 0,$$

i.e. equations (4) above, which are the fundamental equations of the Lagrange multiplier method.

Thus, in two dimensions, the Lagrange multiplier λ is equal to the ratio of $\vec{\nabla} f$ to $\vec{\nabla}\phi$ at the point at which the curve $\phi(x,y) = C$ is tangent to a contour $f(x,y) = \text{constant}$:

$$\lambda = \frac{\vec{\nabla} f}{\vec{\nabla}\phi} \quad \text{at the point } (x\,,\,y) \text{ where } \vec{\nabla} f \text{ is parallel to } \vec{\nabla}\phi\ .$$

Note that the constraint equation $\phi(x,y) = C$ is always somewhat arbitrary, because the same constraint can be written in many equivalent ways (e.g. $x^2 - 2y = 1$ or $4y - 2x^2 = -2$), leading to functions $\phi(x,y)$ which differ from one another by a constant factor. The Lagrange multiplier λ must be freely adjustable (in both magnitude and sign) to follow the scaling of ϕ, and thus contains no fundamental information. For this reason, it is not interesting to solve for λ once x and y have been found. Accordingly, this method is often called the method of "Lagrange undetermined multipliers." (Note that in many references, the sign of λ is opposite to that used here. This has no effect on the method, since λ is arbitrary and never appears in the result.)

The method is easily generalized to functions of more than two variables, as long as the number of constraints is less than the number of variables. For example, to find the extrema of $f(x,y,z)$ subject to two constraints, $g(x,y,z) = c_1$ and $h(x,y,z) = c_2$, the equations become

$$\frac{\partial}{\partial x}(f - \lambda g - \mu h) = 0, \quad \frac{\partial}{\partial y}(f - \lambda g - \mu h) = 0, \quad \frac{\partial}{\partial z}(f - \lambda g - \mu h) = 0,$$

which must be solved together with the two constraint equations to yield the desired values of x,y,z (and, if desired, the usually unwanted multipliers λ and μ.) In this case it is difficult to draw the appropriate picture, but the result can be visualized by stating that the two constraint equations define a curve in 3-dimensional space, and this curve must be tangent to a surface of constant f at the desired extremum. In general, depending on the number of dimensions and the number of constraints, the constraint "curve" may become a surface or hypersurface, and is tangent to a surface or hypersurface corresponding to a fixed value of f.

.