Lagrange Multiplier in optimization problems with multiple constraints

Jingxiang Zou

October 2023

1 Scenario One: Single Constraint

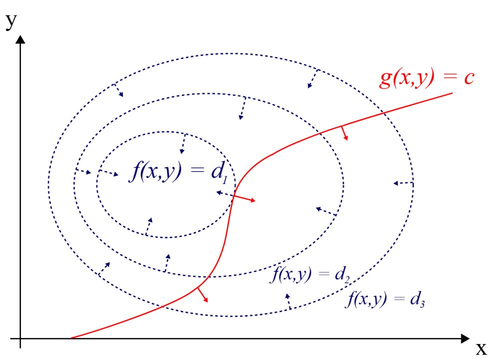

Figure 1: This picture is from wikipedia.

consider a optimization problem in two-dimensional space:
max $f(x,y)$
given one constraint: $g(x,y) = 0$
This is very simple. We go along the level = 0 contour of g(x, y). It is not feasible to seek extremum directly because that involves comparison among uncountable infinite values for f(x, y). We instead consider the necessary condition: any local maximum must be a stationary point, which means when we move for a small distance along g(x, y) = 0, f(x, y) hardly changes, or equivalently, moving in a direction that is perpendicular to the corresponding gradient of f(x, y).

Suppose such a point exists, denoted as point P := (X0, y0), then the gradient of f(x, y) and g(x, y) at this point would have to be parallel, thus:

$$\nabla f(P) = \lambda \nabla g(P)$$

where λ is characterized by P.

Assume λ is know, then this equation is equivalent to the first order condition of function

$$F(x,y) := f(x,y,z) + \lambda g(x,y)$$

that $\frac{\partial F}{\partial x} = 0$ and $\frac{\partial F}{\partial y} = 0$
However λ is not known.
Consider function

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

apparently, $\frac{\partial F}{\partial \lambda} = g(x,y) = 0$ regardless of what λ is.
Thus it suffice to obtain a solution for $(\bar{x}, \bar{y}, \lambda)$ with the FOC of function L that $\frac{\partial F}{\partial x,y,\lambda} = 0$

2 Scenario Two: two constraint

consider a optimization problem in a 3D space:
max $f(x, y, z)$
given two constraints:
1. $g_1(x, y, z) = 0$
2. $g_2(x, y, z) = 0$
where we aim to get a local maximum of function f under the constraints

consider the space comprise of the set of points that satisfy both constraints, denoted as

$$Space(x, y, z)$$

suppose one extremum point we require does exist, denoted as $\bar{P} := (\bar{x}, \bar{y}, \bar{z})$

consider a vector inside $Space(x, y, z)$ denoted as v^* that points from $\bar{P}$ to $P^* = (\bar{x} + \delta(x), \bar{y} + \delta(y), \bar{z} + \delta(z))$ where $\delta(x), \delta(y), \delta(z)$ are arbitrary small value such that $f(\bar{P}) = f(P^*)$.

here are what we know about the vector v^*:

1. Along this vector, the function f changes essentially close to zero, meaning that:

$$v^* \perp \triangledown f(\bar{P})$$

2. This vector is inside $Space(x, y, z)$, which means that values for functions g_1 and g_2 should remain unchanged along its direction, thus we have

$$v^* \perp \triangledown g_1(\bar{P})$$

and

$$v^* \perp \triangledown g_2(\bar{P})$$

Notice that our discussion happens in a 3D space, and all three vectors $\triangledown f(\bar{P})$, $\triangledown g_1(\bar{P})$ and $\triangledown g_2(\bar{P})$ are perpendicular to the same vector v^*. We immediately know that there exist a unique representation of $v^* \perp \triangledown f(\bar{P})$ using $\triangledown g_1(\bar{P})$ and $\triangledown g_1(\bar{P})$ as basis, we denote such representation as follows:

$$\triangledown f(\bar{P}) = \lambda_1 \triangledown g_1(\bar{P}) + \lambda_2 \triangledown g_2(\bar{P})$$

where λ_1 and λ_2 are characterized by $\bar{P}$.

Assume λ_1 and λ_2 are both know, then this equation is equivalent to the first order condition of function

$$F(x, y, z) := f(x, y, z) + \lambda_1 g_1(x, y, z) + \lambda_2 g_2(x, y, z)$$

that $\frac{\partial F}{\partial x} = 0$ and $\frac{\partial F}{\partial y} = 0$ and $\frac{\partial F}{\partial z} = 0$

However λ_1 and λ_2 are not known.

Consider function

$$L(x, y, z, \lambda_1, \lambda_2) := f(x, y, z) + \lambda_1 g_1(x, y, z) + \lambda_2 g_2(x, y, z)$$

apparently, $\frac{\partial F}{\partial \lambda_1} = g_1(x, y, z) = 0$ and $\frac{\partial F}{\partial \lambda_2} = g_2(x, y, z) = 0$ regardless of what λ_1 and λ_2 are.

Thus it suffice to obtain a solution for $(\bar{x}, \bar{y}, \bar{z}, \lambda_1, \lambda_2)$ with the FOC of function L that $\frac{\partial F}{\partial x, y, z, \lambda_1, \lambda_2} = 0$

3 References

Lagrange Multiplier: Wikipedia