

An Example With Two Lagrange Multipliers

In these notes, we consider an example of a problem of the form “maximize (or minimize) $f(x, y, z)$ subject to the constraints $g(x, y, z) = 0$ and $h(x, y, z) = 0$ ”. We use the technique of Lagrange multipliers. To do so, we define the auxiliary function

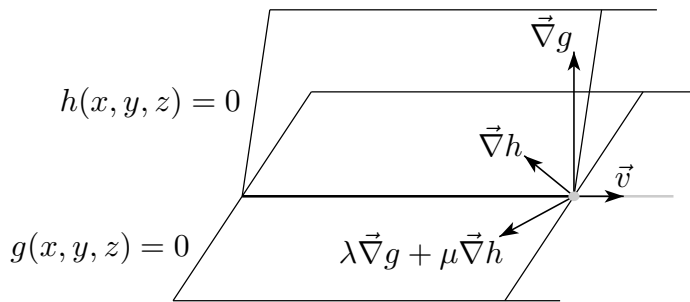
$$L(x, y, z, \lambda, \mu) = f(x, y, z) + \lambda g(x, y, z) + \mu h(x, y, z)$$

It is a function of five variables — the original variables x , y and z , and two auxiliary variables λ and μ . Under suitable assumptions[†] on f , g and h , if the maximum or minimum is achieved at (x_0, y_0, z_0) then (x_0, y_0, z_0) must obey

$$\begin{aligned} 0 &= L_x(x_0, y_0, z_0, \lambda, \mu) = f_x(x_0, y_0, z_0) + \lambda g_x(x_0, y_0, z_0) + \mu h_x(x_0, y_0, z_0) \\ 0 &= L_y(x_0, y_0, z_0, \lambda, \mu) = f_y(x_0, y_0, z_0) + \lambda g_y(x_0, y_0, z_0) + \mu h_y(x_0, y_0, z_0) \\ 0 &= L_z(x_0, y_0, z_0, \lambda, \mu) = f_z(x_0, y_0, z_0) + \lambda g_z(x_0, y_0, z_0) + \mu h_z(x_0, y_0, z_0) \\ 0 &= L_\lambda(x_0, y_0, z_0, \lambda, \mu) = g(x_0, y_0, z_0) \\ 0 &= L_\mu(x_0, y_0, z_0, \lambda, \mu) = h(x_0, y_0, z_0) \end{aligned} \tag{1}$$

for some λ and μ . So solving this system of five equations in five unknowns gives all possible candidates for the locations of local maxima and minima. We’ll go through an example shortly.

Before we get to the example itself, here is why the algorithm above works. Assume that a local minimum occurs at (x_0, y_0, z_0) , which is the grey point in the schematic figure below. Imagine that you start walking away from (x_0, y_0, z_0) along the curve $g = h = 0$. Your path is the grey line in the schematic figure below. Call your velocity vector \vec{v} .



[†] Technically, we assume that all first partial derivatives of f , g and h are continuous and that $\vec{\nabla}g(x_0, y_0, z_0) \times \vec{\nabla}h(x_0, y_0, z_0) \neq 0$. The latter condition says that the normal vectors to $g = 0$ and $h = 0$ at (x_0, y_0, z_0) are not parallel. This ensures that the surfaces $g = 0$ and $h = 0$ are not tangent to each other at (x_0, y_0, z_0) . They intersect in a curve.

It is tangent to the curve $g(x, y, z) = h(x, y, z) = 0$. Because f has a local minimum at (x_0, y_0, z_0) , f must be increasing (or constant) as we leave (x_0, y_0, z_0) . So the directional derivative

$$D_{\vec{v}}f(x_0, y_0, z_0) = \vec{\nabla}f(x_0, y_0, z_0) \cdot \vec{v} \geq 0$$

Now start over. Again walk away from (x_0, y_0, z_0) along the curve $g = h = 0$, but this time moving in the opposite direction, with velocity vector $-\vec{v}$. Again f must be increasing (or constant) as we leave (x_0, y_0, z_0) , so the directional derivative

$$D_{-\vec{v}}f(x_0, y_0, z_0) = \vec{\nabla}f(x_0, y_0, z_0) \cdot (-\vec{v}) \geq 0$$

As both $\vec{\nabla}f(x_0, y_0, z_0) \cdot \vec{v}$ and $-\vec{\nabla}f(x_0, y_0, z_0) \cdot \vec{v}$ are at least zero, we now have that

$$\vec{\nabla}f(x_0, y_0, z_0) \cdot \vec{v} = 0 \tag{2}$$

for all vectors \vec{v} that are tangent to the curve $g = h = 0$ at (x_0, y_0, z_0) . Let's denote by \mathcal{T} the set of all vectors \vec{v} that are tangent to the curve $g = h = 0$ at (x_0, y_0, z_0) and let's denote by \mathcal{T}^\perp the set of all vectors that are perpendicular to all vectors in \mathcal{T} . So (2) says that $\vec{\nabla}f(x_0, y_0, z_0)$ must be in \mathcal{T}^\perp .

We now find all vectors in \mathcal{T}^\perp . We can easily guess two such vectors. Since the curve $g = h = 0$ lies inside the surface $g = 0$ and $\vec{\nabla}g(x_0, y_0, z_0)$ is normal to $g = 0$ at (x_0, y_0, z_0) , we have

$$\vec{\nabla}g(x_0, y_0, z_0) \cdot \vec{v} = 0 \tag{3}$$

Similarly, since the the curve $g = h = 0$ lies inside the surface $h = 0$ and $\vec{\nabla}h(x_0, y_0, z_0)$ is normal to $h = 0$ at (x_0, y_0, z_0) , we have

$$\vec{\nabla}h(x_0, y_0, z_0) \cdot \vec{v} = 0 \tag{4}$$

Picking any two constants λ and μ , multiplying (3) by $-\lambda$, multiplying (4) by $-\mu$ and adding gives that

$$(-\lambda\vec{\nabla}g(x_0, y_0, z_0) - \mu\vec{\nabla}h(x_0, y_0, z_0)) \cdot \vec{v} = 0$$

for all vectors \vec{v} in \mathcal{T} . Thus, for all λ and μ , the vector $-\lambda\vec{\nabla}g(x_0, y_0, z_0) - \mu\vec{\nabla}h(x_0, y_0, z_0)$ is in \mathcal{T}^\perp .

Now the vectors in \mathcal{T} form a line. (They are all tangent to the same curve at the same point.) So, \mathcal{T}^\perp , the set of all vectors perpendicular to \mathcal{T} , forms a plane. As λ and μ run over all real numbers, the vectors $-\lambda\vec{\nabla}g(x_0, y_0, z_0) - \mu\vec{\nabla}h(x_0, y_0, z_0)$ form a plane. Thus we have found all vector in \mathcal{T}^\perp and we conclude that $\vec{\nabla}f(x_0, y_0, z_0)$ must be of the form $-\lambda\vec{\nabla}g(x_0, y_0, z_0) - \mu\vec{\nabla}h(x_0, y_0, z_0)$ for some real numbers λ and μ . The three components of the equation

$$\vec{\nabla}f(x_0, y_0, z_0) = -\lambda\vec{\nabla}g(x_0, y_0, z_0) - \mu\vec{\nabla}h(x_0, y_0, z_0)$$

are exactly the first three equations of (1). This completes the explanation of why the Lagrange multiplier algorithm works.

Example. In this example, we find the distance from the origin to the curve $z^2 = x^2 + y^2$, $x - 2z = 3$. That is, we minimize

$$f(x, y, z) = x^2 + y^2 + z^2$$

subject to the constraints

$$0 = g(x, y, z) = x^2 + y^2 - z^2 \quad 0 = h(x, y, z) = x - 2z - 3$$

To do so we define the auxiliary function

$$\begin{aligned} L(x, y, z, \lambda, \mu) &= f(x, y, z) + \lambda g(x, y, z) + \mu h(x, y, z) \\ &= x^2 + y^2 + z^2 + \lambda[x^2 + y^2 - z^2] + \mu[x - 2z - 3] \end{aligned}$$

and solve the system of equations

$$\begin{aligned} (1) \quad 0 &= L_x(x, y, z, \lambda, \mu) = 2x + 2\lambda x + \mu = 2(1 + \lambda)x + \mu \\ (2) \quad 0 &= L_y(x, y, z, \lambda, \mu) = 2y + 2\lambda y = 2(1 + \lambda)y \\ (3) \quad 0 &= L_z(x, y, z, \lambda, \mu) = 2z - 2\lambda z - 2\mu = 2(1 - \lambda)z - 2\mu \\ (4) \quad 0 &= L_\lambda(x, y, z, \lambda, \mu) = x^2 + y^2 - z^2 \\ (5) \quad 0 &= L_\mu(x, y, z, \lambda, \mu) = x - 2z - 3 \end{aligned}$$

Equation (2) tells us that either $y = 0$ or $\lambda = -1$.

Case $\lambda = -1$: When $\lambda = -1$ the remaining equations reduce to

$$\begin{aligned} (1) \quad 0 &= \mu \\ (3) \quad 0 &= 4z - 2\mu \\ (4) \quad 0 &= x^2 + y^2 - z^2 \\ (5) \quad 0 &= x - 2z - 3 \end{aligned}$$

So

$$(1) \implies \mu = 0 \xrightarrow{(3)} z = 0 \xrightarrow{(5)} x = 3 \xrightarrow{(4)} 0 = 9 + y^2 - 0$$

which is impossible, so we can't have $\lambda = -1$.

Case $y = 0$: When $y = 0$ the remaining equations reduce to

$$\begin{aligned} (1) \quad 0 &= 2(1 + \lambda)x + \mu \\ (3) \quad 0 &= 2(1 - \lambda)z - 2\mu \\ (4) \quad 0 &= x^2 - z^2 \\ (5) \quad 0 &= x - 2z - 3 \end{aligned}$$

Equation (4) tells us that $z = \pm x$.

Subcase $y = 0, z = x$: When $y = 0$ and $z = x$, the remaining equations reduce to

$$(1) \quad 0 = 2(1 + \lambda)x + \mu$$

$$(3) \quad 0 = 2(1 - \lambda)x - 2\mu$$

$$(5) \quad 0 = x - 2x - 3$$

So equation (5) now tells us that $x = -3$ so that $(x, y, z) = (-3, 0, -3)$. (We don't really care what λ and μ are. But as they obey $0 = -6(1 + \lambda) + \mu$, $0 = -6(1 - \lambda) - 2\mu$ we have, adding the two equations together $\mu = -12$, and then, subbing into either equation, $\lambda = -3$.)

Subcase $y = 0, z = -x$: When $y = 0$ and $z = -x$, the remaining equations reduce to

$$(1) \quad 0 = 2(1 + \lambda)x + \mu$$

$$(3) \quad 0 = -2(1 - \lambda)x - 2\mu$$

$$(5) \quad 0 = x + 2x - 3$$

So equation (5) now tells us that $x = 1$ so that $(x, y, z) = (1, 0, -1)$. (We don't really care what λ and μ are. But as they obey $0 = 2(1 + \lambda) + \mu$, $0 = -2(1 - \lambda) - 2\mu$ we have, subtracting the second equation from the first, $\mu = -\frac{4}{3}$, and then, subbing into either equation, $\lambda = -\frac{1}{3}$.)

Conclusion: We have two candidates for the location of the max and min, namely $(-3, 0, -3)$ and $(1, 0, -1)$. The first is a distance $3\sqrt{2}$ from the origin, giving the maximum, and the second is a distance $\sqrt{2}$ from the origin, giving the minimum.