

# Lagrange Multipliers

Gabriela Georgieva  
Angelina Kuznetsova  
Fedor Shafranov

December 14, 2022

## Abstract

Oftentimes one wants to find a minima or maxima of a (differentiable) function subject to one or more constraints. An elegant way to find an extremum is by using the so-called Lagrange Multipliers. Lagrange Multipliers are handy when solving optimization problems in Economics, Business, Computer Science, etc.

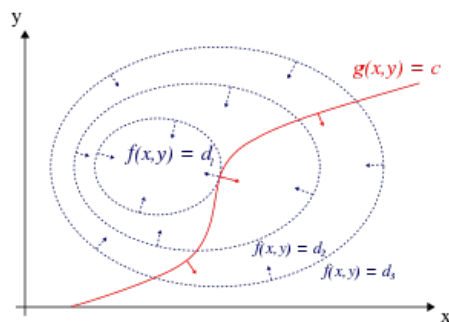


Figure 1: Lagrange multiplier. Wikipedia

## Contents

<b>1</b>	<b>Lagrange Multipliers</b>	<b>3</b>
1.1	Introduction . . . . .	3
1.2	Single Constraint . . . . .	4
1.3	Multiple Constraints . . . . .	6
1.4	Second Derivative Test . . . . .	7
1.5	Lagrangian . . . . .	7
<b>2</b>	<b>Applications</b>	<b>7</b>
2.1	Utility Maximization . . . . .	7
2.2	Cost Minimization . . . . .	8

# 1 Lagrange Multipliers

## 1.1 Introduction

**Theorem 1.** (*Implicit Function Theorem*)

Suppose that  $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$  is  $C^1$ . We will denote points in  $\mathbb{R}^{n+1}$  by  $(\mathbf{x}, z)$ , where  $\mathbf{x} \in \mathbb{R}^n$  and  $z \in \mathbb{R}$ . Assume

$$F(\mathbf{x}_0, z_0) = c \quad \text{and} \quad \nabla F(\mathbf{x}_0, z_0) \neq \mathbf{0}$$

Then there is a ball  $U$  that contains  $\mathbf{x}_0$  and a neighborhood  $V$  of  $z_0$  in  $\mathbb{R}$  such that there is a function  $z = g(\mathbf{x})$  defined for  $\mathbf{x}$  in  $U$  and  $z$  in  $V$  that satisfies

$$F(\mathbf{x}, g(\mathbf{x})) = c$$

**Theorem 2.** (*Method of Lagrange Multipliers*)

Suppose that  $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$  and  $g : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$  are  $C^1$ . Let  $\mathbf{x}_0 \in U$  and  $g(\mathbf{x}_0) = c$ , and let  $S$  be the level set for  $g$  with value  $c$  (i.e these are the set of points  $\mathbf{x} \in \mathbb{R}^n$  that satisfy  $g(\mathbf{x}) = c$ ). Assume  $\nabla g(\mathbf{x}_0) \neq \mathbf{0}$ .

If  $f$  achieves a local extremum at  $\mathbf{x}_0$ , then there exists  $\lambda$  such that

$$\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0)$$

*Proof.* For the sake of simplicity, take  $n=3$ . Then we are dealing with a level surface of the function  $g(x, y, z) = c$  through the point  $(x_0, y_0, z_0)$ . By the Implicit Function Theorem we know that there is a function  $z = \phi(x, y)$  satisfying  $g(x, y, \phi(x, y)) = c$  for  $(x, y)$  near  $(x_0, y_0)$  and  $z$  near  $z_0$ . It follows that locally (near  $z_0$ ), the surface  $S$  is the graph of the function  $\phi$ . For  $\phi$  differentiable and continuous, the tangent plane at  $(x_0, y_0, z_0)$  to  $S$  is given by:

$$z = z_0 + \left[ \frac{\partial \phi}{\partial x}(x_0, y_0) \right] (x - x_0) + \left[ \frac{\partial \phi}{\partial y}(x_0, y_0) \right] (y - y_0) \quad (1)$$

We can substitute

$$\frac{\partial \phi}{\partial x} = -\frac{g_x}{g_z} \quad \frac{\partial \phi}{\partial y} = -\frac{g_y}{g_z}$$

in (1) and obtain:

$$(x - x_0)g_x + (y - y_0)g_y + (z - z_0)g_z = 0$$

$$(x - x_0, y - y_0, z - z_0) \cdot \nabla g(x_0, y_0, z_0) = 0 \quad (2)$$

At  $(x_0, y_0, z_0)$  the tangent plane to the level surface  $g$  is orthogonal to  $\nabla g(x_0, y_0, z_0)$ . Now we need to show that every vector tangent to  $S$  at  $(x_0, y_0, z_0)$  is tangent to every curve in  $S$ . If  $\mathbf{v} = (x - x_0, y - y_0, z - z_0)$  is tangent to  $S$ , then  $(\mathbf{v})$  is tangent to every path in  $S$  given by

$$c(t) = (x_0 + t(x - x_0), y_0 + t(y - y_0), \phi(x_0 + t(x - x_0), y_0 + t(y - y_0)))$$

at  $t = 0$ . Now, if  $(x_0, y_0, z_0)$  is an extremum, then  $f(c(t))$  is an extremum when  $t = 0$  and  $c'(0)$  is a tangent vector to  $S$  at  $(x_0, y_0, z_0)$

$$\left. \frac{d}{dt} f(c(t)) \right|_{t=0} = \nabla f(x_0) \cdot c'(0) = 0$$

Thus,  $\nabla f(x_0)$  is orthogonal to every tangent vector to  $S$  at  $(x_0, y_0, z_0)$ . Since the space orthogonal to this tangent space is a line,  $\nabla f(x_0, y_0, z_0)$  and  $\nabla g(x_0, y_0, z_0)$  are parallel and  $\nabla f(x_0, y_0, z_0)$  is a multiple of  $\nabla g(x_0, y_0, z_0)$

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$$

The multiple  $\lambda$  is called a Lagrange Multiplier. □

## 1.2 Single Constraint

**Example 1.** Find the points closest to the origin on  $xy + 3x + z^2 = 9$

The distance from any point  $(x, y, z)$  to the origin can be expressed as  $\sqrt{x^2 + y^2 + z^2}$ . We want to minimize this distance subject to the constraint  $xy + 3x + z^2 = 9$ . We will denote the function we want to minimize as  $f$ . We will use the square of the distance formula as it does not change the result but makes the calculations simpler. We have

$$f(x, y, z) = x^2 + y^2 + z^2 \qquad g(x, y, z) = xy + 3x + z^2 - 9$$

The Method of Lagrange Multipliers tells us that at the critical points of a function

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) \tag{1}$$

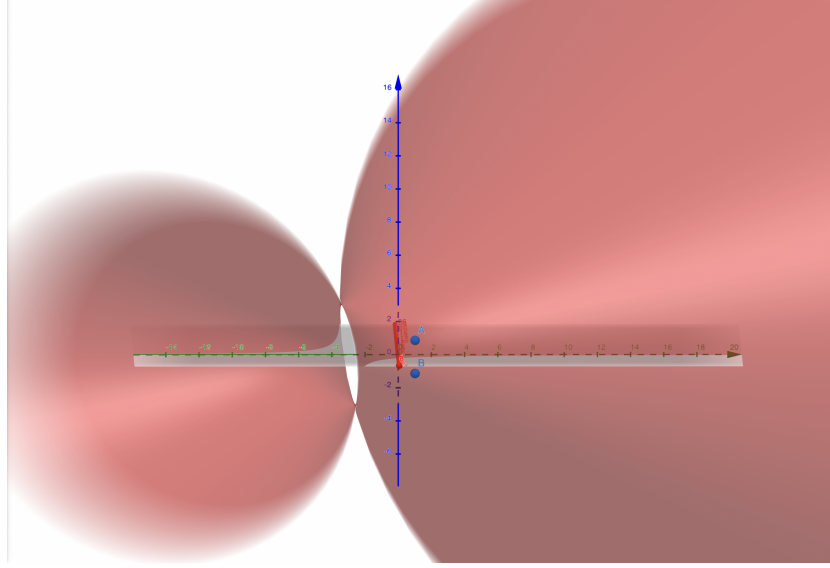
We find the gradient of  $f$  and  $g$

$$\nabla f(x, y, z) = \langle 2x, 2y, 2z \rangle \quad \nabla g(x, y, z) = \langle y + 3, x, 2z \rangle$$

After substituting in (1), we get a system of three equations and three variables. For the fourth equation we use the constraint.

$$\begin{cases} 2x = \lambda(y + 3) \\ 2y = \lambda x \\ 2z = \lambda 2z \\ xy + 3x + z^2 = 9 \end{cases}$$

From the third equation we see that  $\lambda = 1$  and substituting in the first and second we get  $x = 2$  and  $y = 1$ . Finally, substituting in the last equation gives us  $z = \pm 1$ . So, the two critical points are  $A = (2, 1, 1)$  and  $B = (2, 1, -1)$  and the distance to the origin is  $\sqrt{6}$ . We know that at this points on  $f$  we will be closest to the origin because any other point we choose on  $f$  will give us a bigger distance.



**Example 2.** Find the ellipse  $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$  that passes through  $(3, 1)$  and has the smallest area.

The function we are striving to minimize is the area of an ellipse function  $A = \pi ab$  that is subject to the constraint  $\frac{9}{a^2} + \frac{1}{b^2} = 1$ . We need  $a, b > 0$ . We will denote the function we want to minimize  $f(a, b) = \pi ab$  and the constraint  $g(a, b) = \frac{9}{a^2} + \frac{1}{b^2} - 1$ . For  $f$  and  $g$   $C^1$ , if  $(a_0, b_0)$  is an extremum, then there exists  $\lambda$  such that

$$\begin{aligned} \nabla f(a_0, b_0) &= \lambda \nabla g(a_0, b_0) \\ \langle \pi a_0, \pi b_0 \rangle &= \lambda \langle -18a_0^{-3}, -2b_0^{-3} \rangle \\ \begin{cases} \pi b_0 = \lambda(-18)a_0^{-3} \\ \pi a_0 = \lambda(-2)b_0^{-3} \\ \frac{9}{a_0^2} + \frac{1}{b_0^2} = 1 \end{cases} \end{aligned}$$

Solving for  $\lambda$  in the first equation and substituting in the second we get

$$\begin{cases} \lambda = \frac{\pi b_0 a_0^3}{-18} \\ a_0 = \frac{a_0^3}{9b_0^2} \\ \frac{9}{a_0^2} + \frac{1}{b_0^2} = 1 \end{cases}$$

Solving the second equation gives us  $a_0 = 0$  and  $9b_0^2 - a_0^2 = 0$ .  $a_0 = 0$  is not in our domain. Solving for  $b_0$  gives us  $b_0 = \pm \frac{a_0}{3}$ . Now, we solve the third

equation for  $a_0$  and obtain  $a_0 = \pm 3\sqrt{2}$ . For  $a_0 = \pm 3\sqrt{2}$  and  $b_0 = \pm \frac{a_0}{3}$  we get the following choices of  $(a_0, b_0)$ :  $(3\sqrt{2}, \sqrt{2})$ ,  $(3\sqrt{2}, -\sqrt{2})$ ,  $(-3\sqrt{2}, \sqrt{2})$ ,  $(-3\sqrt{2}, -\sqrt{2})$ . Since we are interested in  $a, b > 0$  the only choice is  $(3\sqrt{2}, \sqrt{2})$ . The method of Lagrange Multipliers guarantees that the result we obtain is an extremum but does not tell us whether it is a minima or maxima. To find out, we pick another  $(a_1, b_1)$  that satisfies the constraint, for example  $(\sqrt{\frac{27}{2}}, \sqrt{3})$ . Since for  $(a_1, b_1)$  we get a bigger area, then we conclude that, indeed, for  $(3\sqrt{2}, \sqrt{2})$  we achieve the smallest area considering the constraint.

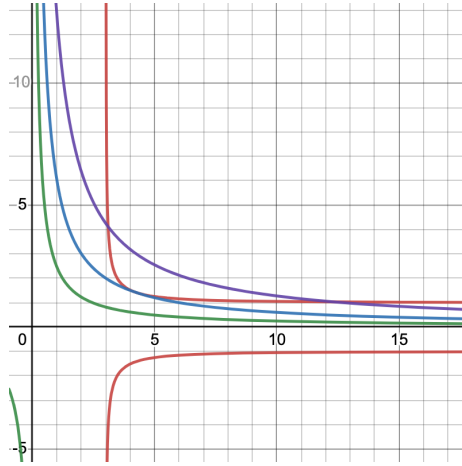


Figure 2: Constraint curve  $g(a, b)$  with contour plot of the surface  $\pi ab$

### 1.3 Multiple Constraints

If we have  $n$  number of constraints, then the method of Lagrangian Multipliers tells us that if there is an extremum at point  $P$  on  $S$ , then there exist  $\lambda_1, \lambda_2, \dots, \lambda_n$  such that

$$\nabla f(x, y, z) = \lambda_1 \nabla g_1(x, y, z) + \lambda_2 \nabla g_2(x, y, z) + \dots + \lambda_n \nabla g_n(x, y, z)$$

**Example 3.** Find the minimum distance from the origin to the line of intersection of the two planes  $x + y + z = 8$  and  $2x - y + 3z = 28$

We will again use the squared distance formula, so we set  $f(x, y, z) = x^2 + y^2 + z^2$ . We write the two constraints in the form  $g(x, y, z) = x + y + z - 8$  and  $h(x, y, z) = 2x - y + 3z - 28$ . Then we have

$$\nabla f(x, y, z) = \lambda_1 \nabla g(x, y, z) + \lambda_2 \nabla h(x, y, z)$$

$$\langle 2x, 2y, 2z \rangle = \lambda_1 \langle 1, 1, 1 \rangle + \lambda_2 \langle 2, -1, 3 \rangle$$

To find the point on the line closest to the origin we need to solve the system

$$\begin{cases} 2x = \lambda_1 + 2\lambda_2 \\ 2y = \lambda_1 - \lambda_2 \\ 2z = \lambda_1 + 3\lambda_2 \\ x + y + z = 8 \\ 2x - y + 3z = 28 \end{cases}$$

From the first three equations we get  $x = \frac{\lambda_1 + 2\lambda_2}{2}$ ,  $y = \frac{\lambda_1 - \lambda_2}{2}$  and  $z = \frac{\lambda_1 + 3\lambda_2}{2}$ . After substituting in the 4th and 5th equation we get  $\lambda_1 = 0$  and  $\lambda_2 = 4$ . After substituting back for  $x, y, z$  we get the point  $(4, -2, 6)$ . Thus, the distance from  $(4, -2, 6)$  to the origin is  $\sqrt{56}$ . We know that the point we got is closest to the origin because  $\max(\sqrt{x^2 + y^2 + z^2}) \rightarrow \infty$ .

## 1.4 Second Derivative Test

## 1.5 Lagrangian

The idea behind the Lagrangian is to reduce constrained optimization to unconstrained optimization, and to take the (functional) constraints into account by augmenting the objective function with a weighted sum of them. If we have a function  $f(x)$  and a constraint  $g(x) = c$  the Lagrangian is defined as

$$L(x, \lambda) = f(x) - \lambda(g(x) - c)$$

# 2 Applications

## 2.1 Utility Maximization

Lagrange Multipliers are widely used in microeconomic theory to find the optimal consumption of production bundles under a certain constraint. Even though the examples are trivial, we see how basic understanding of Lagrange Multipliers can help solving Economics problems for optimization.

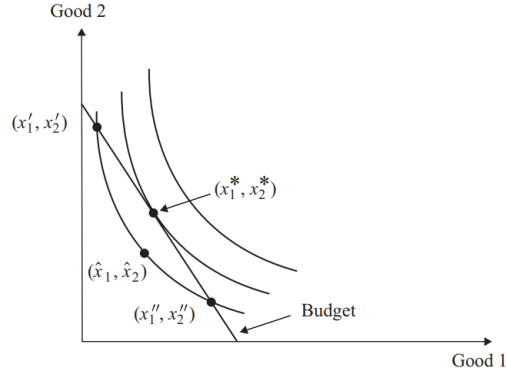
Let us take a utility function that will represent a utility for consuming a bundle of goods where  $x_1$  and  $x_2$  are some goods.

$$u = u(x_1, x_2)$$

The budget of the agent is limited, so we need to introduce a budget constraint. Where  $x_1, x_2$  are goods and  $p_1, p_2$  is their price

$$p_1 x_1 + p_2 x_2 = B$$

Now, we need to find an optimum point of consumption that lies on the edge of the intersection between budget and utility.



Here we introduce a Lagrangian multiplier to find a point where gradients are parallel.

Maximize

$$L = U(x_1, x_2) + \lambda[B - p_1x_1 - p_2x_2]$$

FOC:

$$\begin{aligned} \frac{\partial L}{\partial x_1} = 0 &\implies \frac{\partial U}{\partial x_1} - \lambda p_1 = 0 \implies \frac{\partial U}{\partial x_1} = \lambda p_1 \\ \frac{\partial L}{\partial x_2} = 0 &\implies \frac{\partial U}{\partial x_2} - \lambda p_2 = 0 \implies \frac{\partial U}{\partial x_2} = \lambda p_2 \\ \frac{\partial L}{\partial \lambda} = 0 &\implies B = p_1x_1 + p_2x_2 \end{aligned}$$

In Economics, a corresponding change in utility  $U$  over an increase in  $x_i$  is called marginal utility  $MU_i$ . Now, finishing calculations, we see

$$\frac{\frac{\partial U}{\partial x_1}}{\frac{\partial U}{\partial x_2}} = \frac{p_1}{p_2} \implies \frac{MU_1}{p_1} = \frac{MU_2}{p_2} \implies \frac{MU_1}{MU_2} = \frac{p_1}{p_2}$$

And we can see that it is exactly the point where the constraint touches the curve.

## 2.2 Cost Minimization

The next example is similar, but now we view the production decision of a firm that minimizes its costs. A firm fixes its output level, making it a constraint, and chooses a budget at which to produce.

$$L = w_1x_1 + w_2x_2 + \lambda[y - f(x, y)]$$

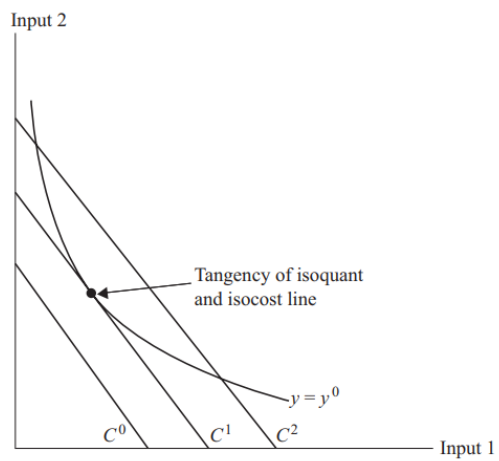
FOC:

$$w_1 - \lambda \frac{\partial f}{\partial x_1} = 0$$



$$w_2 - \lambda \frac{\partial f}{\partial x_2} = 0$$

$$\frac{w_1}{w_2} = \frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}}$$



## References

- [1] J.E. MARSDEN AND A.J. TROMBA. Vector Calculus: 5th Edition. W.H.Freeman and Company, 2003.
- [2] J. STEWART. Calculus: 7th edition. Brooks/Cole, 2012.