

**UNIVERSIDAD DEL CEMA  
Buenos Aires  
Argentina**

Serie  
**DOCUMENTOS DE TRABAJO**

**Área: Matemática y Economía**

**Intuitive Mathematical Economics Series.  
Constrained Maximization and the Method of Lagrange Multipliers**

**Sergio A. Pernice**

**Diciembre 2018  
Nro. 680**

**[www.cema.edu.ar/publicaciones/doc\\_trabajo.html](http://www.cema.edu.ar/publicaciones/doc_trabajo.html)  
UCEMA: Av. Córdoba 374, C1054AAP Buenos Aires, Argentina  
ISSN 1668-4575 (impreso), ISSN 1668-4583 (en línea)  
Editor: Jorge M. Streb; asistente editorial: Valeria Dowding [jae@cema.edu.ar](mailto:jae@cema.edu.ar)**



# Intuitive Mathematical Economics Series

## Constrained Maximization and the Method of Lagrange Multipliers

SERGIO A. PERNICE<sup>1</sup>

*Universidad del CEMA*  
*Av. Córdoba 374, Buenos Aires, 1054, Argentina*

December 26, 2018

### Abstract

A fundamental assumption in most of economic modeling is that people maximize their utility subject to a budget constraint. This, as well as many other economic problems, mathematically translate into problems of maximization with constraints. A powerful and widely used method to tackle some of these problems is the method of Lagrange multipliers. Yet, the exposition of such method in standard textbooks is rather formal and utilitarian. In this paper we try to present it emphasising the fundamental intuitions behind the method.

*Keywords:* Maximization with constraints, Lagrange multipliers.

## 1 Introduction

This paper belongs to the “Intuitive Mathematical Economics Series”, whose objective is to present mathematical economics in a more intuitive way than the one usually encountered in the standard textbooks on the subject. In this case we cover constrained maximization in two variables.

Perhaps the most basic problem in economics concerns the one-period, rational choice between two goods, given a finite budget. In this paper we will use this problem to motivate the presentation of a geometrical approach to constrained maximization problems that on the one hand easily generalizes to many variables and many periods (and the continuum time limit), and on the other it is highly intuitive and its intuitive nature remains so in these higher dimensional problems.

The paper is organized as follows...

---

<sup>1</sup>sp@ucema.edu.ar

The points of view of the author do not necessarily represent the position of Universidad del CEMA.

## 2 Standard ways of solving maximization with constraints problems

The basic problem of optimizing utility with a budget constraint is this: there are two goods, say food and clothes, a unit food costs  $P_f$  and a unit of clothes costs  $P_c$ , your budget is  $i$ , how much do you buy of each good so as to maximize your utility?

Suppose the utility function is:

$$U(f, c) = f^\alpha c^\beta \quad (1)$$

where  $f$  represents the units of food and  $c$  the units of clothes that the person consumes. We assume that  $f$  and  $c$  can be any real, positive number. The exponents  $\alpha$  and  $\beta$  can also be any positive number in principle. However if we assume that they are numbers between zero and one, we have the more or less standard situation for normal goods in which, on the one hand, the more the better, and on the other, the more the person consumes of one item, the less utility she gets from an additional unit of that same item.

Assume we have  $i$  dollars to spend on food and clothing, and define the budget function  $I$  of  $f$  units of food and  $c$  units of clothes as:

$$I(f, c) = fP_f + cP_c \quad (2)$$

We are modeling a one-period problem and we assume that we do not get any utility from keeping unspent money, therefore we will choose  $f$  and  $c$  such that  $I(f, c) = i$ .

The problem then is to maximize the utility (1) subject to the budget constraint:

$$\max[U(f, c)] \quad (3)$$

$$I(f, c) = i \quad (4)$$

Mathematically, problem (3-4) is a simple two dimensional problem of maximization with constraints, and it can be solved in many different ways. In the rest of this section we present three standard ways of solving it.

### 2.1 Direct solution

The direct way, given the particular form of the constraint, is to use (4) to find an explicit expression of  $c$  in terms of  $f$ :

$$c = -\frac{P_f}{P_c}f + \frac{i}{P_c} \quad (5)$$

Replacing (5) in (1) and rearranging:

$$U(f, c(f)) = \left(\frac{P_f}{P_c}\right)^\beta f^\alpha \left(-f + \frac{i}{P_f}\right)^\beta \quad (6)$$

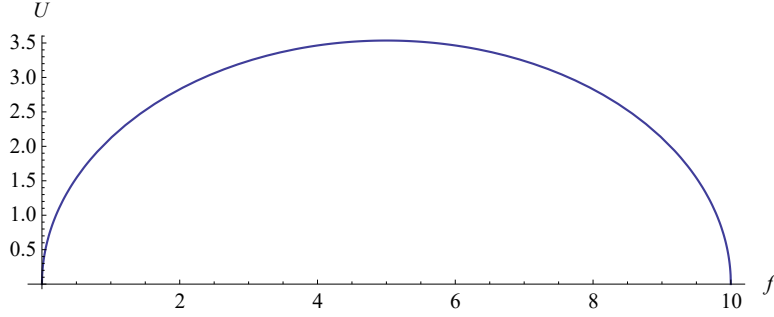


Figure 1:  $U(f, c(f))$  in (6) for  $\alpha = \beta = 1/2$ ,  $P_f = 1$ ,  $P_c = 2$ ,  $I = 10$ .

we end up with a trivial one-dimensional maximization problem. In Fig. 1,  $U(f, c(f))$  is plotted for  $\alpha = \beta = 1/2$ ,  $P_f = 1$ ,  $P_c = 2$ ,  $I = 10$ . The symmetry of the graph makes clear that the maximization happens at  $f = 5$ . (5) then implies that  $c = 2.5$ , and of course  $I = 5 \times 1 + 2.5 \times 2 = 10$ , satisfying the constraint (see Problems 2.1: 1 and 2).

In general, the maximization of  $U(f, c(f))$ , where  $c(f)$  is an *explicitly known* function of  $f$ , corresponds to the value of  $f$  such that  $dU/df = 0$  (see [Pernice (2018 a)]) (see Problem 2.1: 3):

$$\frac{dU(f, c(f))}{df} = \frac{\partial U(f, c)}{\partial f} + \frac{\partial U(f, c)}{\partial c} \frac{dc}{df} = 0 \quad (7)$$

From (5),  $dc/df = -P_f/P_c$ , so, at the maximum

$$\frac{\partial U(f, c)}{\partial f} - \frac{\partial U(f, c)}{\partial c} \frac{P_f}{P_c} = 0 \quad (8)$$

which can be rearranged in the form

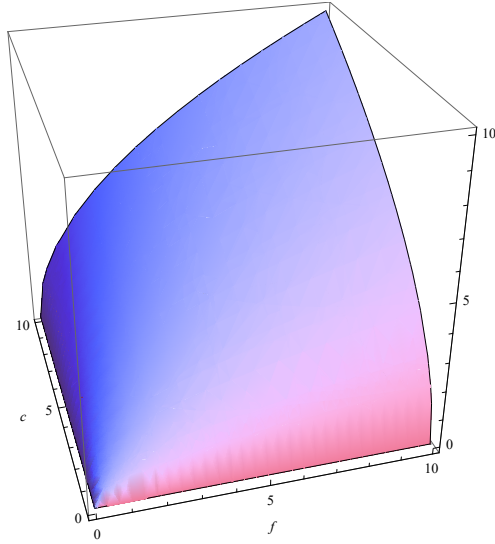
$$\frac{\frac{\partial U(f, c)}{\partial f}}{P_f} = \frac{\frac{\partial U(f, c)}{\partial c}}{P_c} \quad (9)$$

(9) is interpreted as the fact that, at the maximum, the marginal utility per dollar of expenditure is the same for each good.

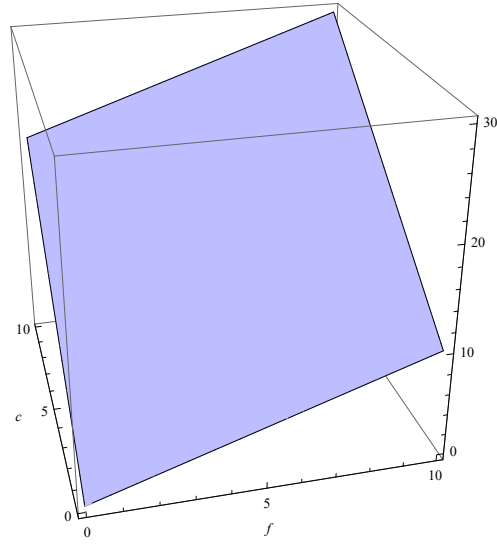
Another way of expressing the same that we will use later is

$$\frac{\frac{\partial U(f, c)}{\partial f}}{\frac{\partial U(f, c)}{\partial c}} = \frac{P_f}{P_c} \quad (10)$$

This direct method of solving maximization problems with constraints has the advantage that requires a minimum knowledge of calculus. Unfortunately, however, it generally becomes useless for higher dimensional problems, or when the constraint is not so easily resolved as in (5).



(a) Utility function (1).



(b) Budget function (2).

Figure 2: 3-D representation of the utility function and the budget function.  $\alpha = \beta = 1/2$ ,  $P_f = 1$ ,  $P_c = 2$ ,  $I = 10$ ,

## 2.2 Graphical method for maximization with constraints

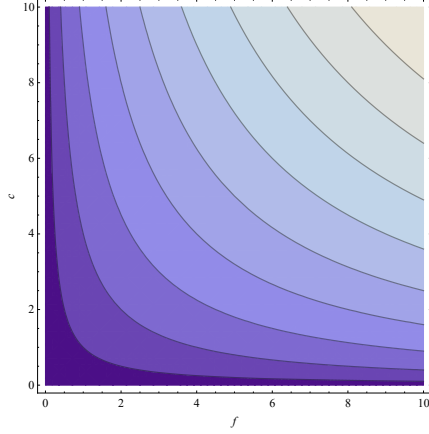
For simple 2 dimensional problems like (3), a very intuitive graphical method is usually enough to understand the essence of the problem and derive from this intuition the corresponding equations. We present in this section this graphical method, which is also used in elementary economic textbooks to explain maximization with constraints.

The utility function (1) and the budget function (2) can be represented in 3 dimensions, as in Fig. 2, or in 2 dimensions, by their level curves, as in Fig. 3. The level curves representation is particularly useful here.

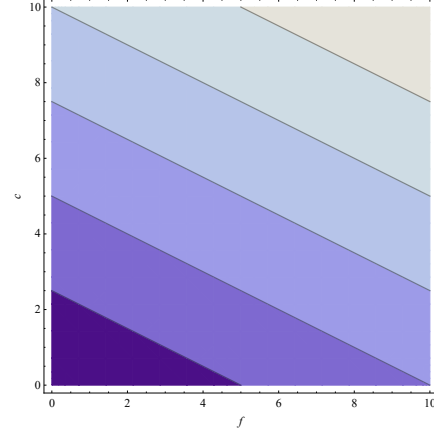
Comparing Fig. 2a with Fig. 3a, and Fig. 2b with Fig. 3b, it is obvious that in both cases, as we move in the up-right direction, the level curves correspond to increasing utility (Fig. 3a) and budget (Fig. 3b), respectively.

Consider, as we did in the previous subsection, the level curve for the budget function (2) corresponding to  $I(f, c) = 10$ . We found then that the utility is maximized for  $f = 5$  and  $c = 5/2$ , corresponding to a utility  $U = \sqrt{5 \times 5/2} = 5/\sqrt{2}$ . Therefore, the corresponding budget and

<sup>2</sup>The level curves of a function  $G(f, c)$  correspond to the curves in the  $(f, c)$  plane defined implicitly by the equation  $G(f, c) = g$ , where here  $g$  is just a number.



(a) Utility function (1) level curves.



(b) Budget function (2) level curves.

Figure 3: 2-D representation of the utility function and the budget function by their level curves.  $\alpha = \beta = 1/2$ ,  $P_f = 1$ ,  $P_c = 2$ ,  $I = 10$ ,

utility curves in the  $(f, c)$  plane are, respectively:

$$2 \times c + 1 \times f = 10 \Rightarrow c = -\frac{1}{2}f + 5 \quad (11)$$

$$\sqrt{fc} = \frac{5}{\sqrt{2}} \Rightarrow c = \frac{25}{2} \frac{1}{f} \quad (12)$$

In Fig. 4 we represent in the  $(f, c)$  plane three utility level curves ( $U = 5/\sqrt{2} - 0.5$ ,  $U = 5/\sqrt{2}$ , and  $U = 5/\sqrt{2} + 0.5$ ), and the budget line  $I = 10$ .

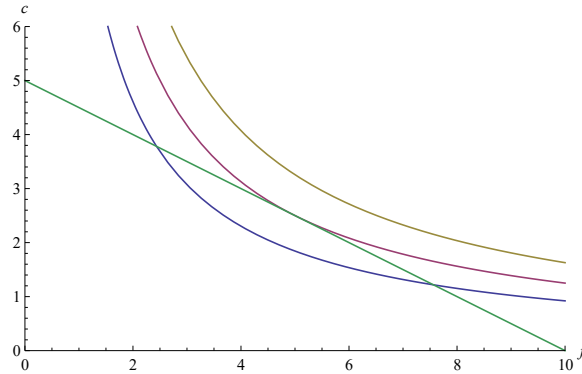


Figure 4:  $U = 5/\sqrt{2} - 0.5$  (blue),  $U = 5/\sqrt{2}$  (purple),  $U = 5/\sqrt{2} + 0.5$  (yellow) and  $I = 10$  (green).

Just staring at Fig. 4, given the convex form of the level curves of the utility function, it is obvious that the maximization of utility compatible with the budget constraint will happen at the point in the  $(f, c)$  plane in which the budget line is tangent to the level curve of the utility function, intercepting it only once.

If the level curve of the utility function intercepts the budget line at two points, as in the blue curve, it is clear that we can increase utility by choosing higher level curves. If it does not intercept the budget line, as in the yellow curve, then it is not compatible with our budget constraint. Therefore the optimum is the level curve of the utility function that intercepts the budget line just once, as in the purple curve. And since the level curves are smooth, at that point the budget line must coincide with the tangent line of the level curve of the utility function.

Let us translate all this graphical intuition into equations. The level curves of the utility function in the  $(f, c)$  plane are curves  $c(f)$  implicitly given by the equation  $U(f, c(f)) = u$ , where  $u$  is just a number. By the rules of differentiation of implicit functions [Pernice (2018 a)], the slope  $dc/df$  of such curve, is given by  $\partial U/\partial f + (\partial U/\partial c)(dc/df) = 0$ , or

$$\frac{dc}{df} = -\frac{\frac{\partial U}{\partial f}}{\frac{\partial U}{\partial c}} \quad (13)$$

The intuitive graphical argument above indicates that at the optimum, this slope coincides with the slope of the budget line (5)

$$\frac{dc}{df} = -\frac{P_f}{P_c} \quad (14)$$

From (13) and (14) we arrive at the equation (see problem 2.2: 1)

$$\frac{\frac{\partial U}{\partial f}}{\frac{\partial U}{\partial c}} = \frac{P_f}{P_c} \quad (15)$$

In section 2.1 we reduced the problem to a one dimensional problem for which equation (10) determines the solution. But our graphical approach in this section “lives” in the two dimensional  $(f, c)$  plane, so, in addition to (15), we need a second equation to determine the answer. To find the second equation it helps to look again at figures 3 and 4. All the level lines of the budget function (given by (5) for different values of  $i$ ) have the same slope  $-P_f/P_c$  (see Fig. 3b), so the right hand side of (15) does not single out our specific budget constraint. It seems clear then that the second equation is our specific budget constraint.

But before that, let us see the same from a different angle. Note that every level curve of the utility function (see Fig. 3a) will have a point where the tangent line also has slope  $-P_f/P_c$ . Therefore equation (15) determines a curve  $c^*(f)$  whose points correspond to the optimum utility for every possible budget value. To determine the optimum for our budget we need to find the interception of that curve with our budget line.

So our constrained maximization problem is solved by finding the solution to the system of two equations (see problems 2.2: 2 and 3)

$$\frac{\frac{\partial U}{\partial f}}{\frac{\partial U}{\partial c}} = \frac{P_f}{P_c} \quad (16)$$

$$fP_f + cP_c = i \quad (17)$$



Equation (17), and the right hand side of (16), depend explicitly on the specific form of the budget function (2). In the rest of this section, to emphasise their generality, we extend the graphical arguments presented here to more general functions and constraints. We will still use the notation  $U(f, c)$  for the function to be maximized and  $I(f, c) = i$  for the constraint, but they no longer necessarily represent utility and budget functions.

The key idea behind equation (16) is that, at the constrained maximum of  $U(f, c)$ , the constraining curve  $c_I(f)$  should be *tangent* to the level curve of  $U(f, c)$  corresponding to that maximum. This fact, it turns out, is more general than our particular example suggest. Let us try to see intuitively why this is so.

Presenting our graphical arguments in a more geometric context will help our intuition. Suppose that a generic function  $U(f, c)$  represents the vertical height of a mountain above sea level. We walk through a trekking trail in that mountain. Let us map our walk in a Cartesian 3-D system with the positive  $f$  axis pointing east, the positive  $c$  axis pointing north (we are assuming a small enough region so that the spherical nature of earth's surface can be well approximated by an Euclidean plane), and the positive  $U$  axis pointing vertically up.  $U = 0$  is at the sea level.

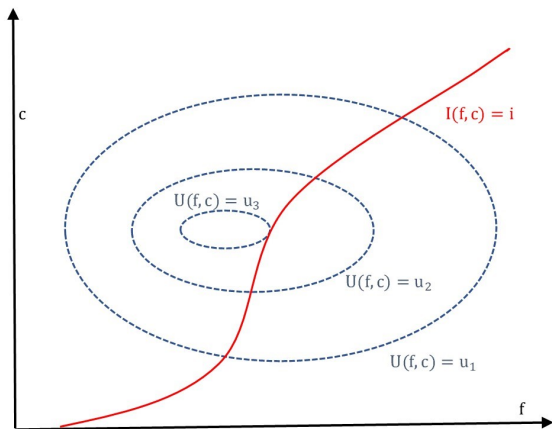


Figure 5: The blue, dashed, curves are the level curves of  $U$ .  $u_1 < u_2 < u_3$ . The red curve is the projection on the  $U = 0$  plane of the trekking trail in the mountain. The walk is in the north-east direction.

The red curve in Fig. 5 represents the projection on the  $U = 0$  plane of the trekking trail, it is a curve  $c_I(f)$  on that plane. It can always be given implicitly in the form  $I(f, c) = i$  (knowing explicitly  $c_I(f)$ , this is trivially the case with  $I(f, c) = c - c_I(f)$  and  $i = 0$ ).

We are moving roughly in the north-east direction. At the beginning we go up. Eventually we reach our highest point in the trail, at height  $u_3$ , and then we go down. Suppose each step is very small. Each step represents a vector in 3-D. The projection of this vector on the  $U = 0$  plane is tangent to the red line. The  $U$  component of those vectors at the beginning is positive (we are going up). When we reach the highest point, our small step right at this point represents a vector with zero component in the vertical ( $U$ ) direction. The projection on the  $U = 0$  plane, as always, is tangent to the red line. The fact that at that point we make a small displacement without vertical component means that, at that point, we are momentarily walking on a level curve of the

mountain itself. So, at that point, the level curve of the mountain and the projection on the  $(f, c)$  plane of our trekking trail are tangent to each other. Therefore in the  $(f, c)$  plane, the slope of the curve  $c_U(f)$  given implicitly by the level curve  $U(f, c) = u_3$  and the curve  $c_I(f)$  given implicitly by the level curve  $I(f, c) = i$  should be identical at the maximum. This is the intuition behind the generality of the tangency of the level curve  $c_U(f)$  of  $U(f, c)$  and the constraining curve  $c_I(f)$  at the constrained maximum.

So we need to generalize the right hand side of (16) to include the slope of a generic constraint given implicitly by the level curve  $I(f, c) = i$  of a function  $I(f, c)$ . But we know (see (13)) that the slope of any level curve  $c_I(f)$  of a function  $I(f, c)$  is given by

$$\frac{dc_I}{df} = -\frac{\frac{\partial I}{\partial f}}{\frac{\partial I}{\partial c}} \quad (18)$$

So the generalization of (16) is

$$\frac{\frac{\partial U}{\partial f}}{\frac{\partial U}{\partial c}} = \frac{\frac{\partial I}{\partial f}}{\frac{\partial I}{\partial c}} \quad (19)$$

In addition to (16), the generalization of equation (17) is as necessary in the generic case as it was in the specific one. Let us see why in this more general context. First, as it was the case before, (19) is only one equation in the two variables  $f$  and  $c$ , so it won't determine the solution. Second, in the same way in which the right hand side of (16) was the same for every  $i$ , and so (17) was necessary to fix the real constraint, it is also the case that the right hand side of (18) is the same for every value  $i$  of the constrain  $I(f, c) = i$  that defines implicitly  $c_I(f)$  (see problem 2.2: 4).

Therefore we need to impose explicitly the equation  $I(f, c) = i$ , for the  $i$  corresponding to our real constraint. So, finally, the generalization of the system (16-17) is

$$\frac{\frac{\partial U}{\partial f}}{\frac{\partial U}{\partial c}} = \frac{\frac{\partial I}{\partial f}}{\frac{\partial I}{\partial c}} \quad (20)$$

$$I(f, c) = i \quad (21)$$

The graphical method described in this section has the advantage that is very intuitive, but, at least in its present form, it is very difficult to generalize to higher, arbitrary dimensions, where we are no longer able to "see" graphically the problem. However, as we will see in section 4, with the help of some new mathematical concept, we will be able to expand our intuition to many dimensions.

### 2.3 The Lagrange multipliers method: standard treatment

The two methods presented so far are not very good at generalizations to many dimensions and many constraints. And it just happens that the more important economic applications involving

maximization with constraints do involve many dimensions and constraints. So we present here a method invented by the Italian mathematician Joseph-Louis Lagrange, born on January 25, 1736, that does generalize easily to many dimensions and many constraints.

Consider the function

$$L(f, c, \lambda) = U(f, c) - \lambda(I(f, c) - i) \quad (22)$$

where  $\lambda$  is a new variable called Lagrange multiplier. Maximize  $L$  it with respect to the *three* variables  $f$ ,  $c$ , and  $\lambda$ :

$$\frac{\partial L}{\partial f} = \frac{\partial U}{\partial f} - \lambda \frac{\partial I}{\partial f} = 0 \quad (23)$$

$$\frac{\partial L}{\partial c} = \frac{\partial U}{\partial c} - \lambda \frac{\partial I}{\partial c} = 0 \quad (24)$$

$$\frac{\partial L}{\partial \lambda} = I(f, c) - i = 0 \quad (25)$$

These are three equations with three unknowns. Except in special cases, it will have a unique solution. The  $f$  and  $c$  solutions will correspond to the optimum of  $U(f, c)$  under the constraint  $I(f, c) = i$ . To see this, simply pass the  $\lambda$  term in (23) and (24) to the right, and divide (23) by (24), obtaining (20). And (25) is identical to (21).

This is the method, it reproduces the right equations, and it just happens that this method almost trivially generalizes to many dimensions and many constraints, as we will see in another work. The fact that the method reproduces the right equations is the formal, utilitarian way of presenting and justifying the method in the typical textbook. This leaving students in the dark with regard to the deep reasons of why it works, the nature of the additional  $\lambda$  variable, etc.

In the rest of the paper we provide the elements to endow the Lagrange multipliers method with the same intuitive appeal as the graphical method described in section 2.2. Moreover, as we will see in later work, with small doses of abstraction, we will learn to extend the geometric intuition to many dimensions, where our ability to visualize, in the standard sense of the word, is just impossible.

### 3 Gradient and level curves

The underlying mathematical concepts of vector spaces, vector fields, gradients, etc., will be covered in a different work. Here we provide a summary of the concepts needed to make the paper self contained, since the geometrical interpretation of Lagrange multipliers uses gradients and their geometric properties.

Given any sufficiently smooth function  $U(f, c)$ , its “gradient”, for which we use the symbol  $\vec{\nabla}U$ , is the vector field

$$\vec{\nabla}U(f, c) = \frac{\partial U(f, c)}{\partial f} \hat{f} + \frac{\partial U(f, c)}{\partial c} \hat{c} \quad (26)$$

This equation should be interpreted as follows:  $\hat{f}$  and  $\hat{c}$  are unit vectors in the horizontal and vertical directions respectively. At each point in the  $(f, c)$  plane we assign a vector, whose horizontal component is  $\partial U/\partial f$  evaluated at  $(f, c)$ , and its vertical component is  $\partial U/\partial c$  evaluated at the same point. This is then a *vector field*, i.e., a vector for each point in the  $(f, c)$  plane.

For example, consider the function  $U(f, c) = 50 - (f^2 + c^2)$ , its level curves, and the corresponding implicit curves in the  $(f, c)$  plane in Fig. 6. Its gradient is

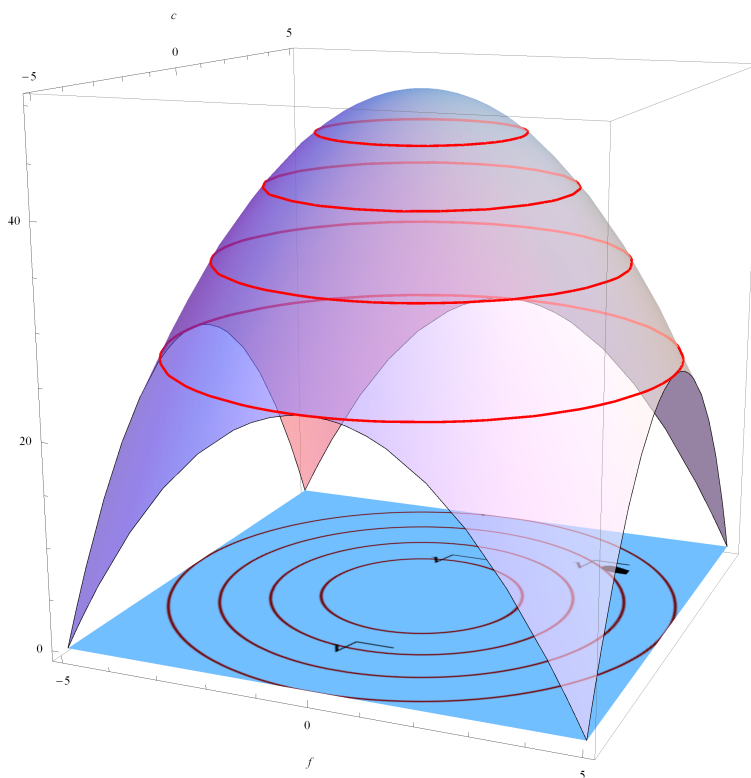


Figure 6:  $U(f, c) = 50 - (f^2 + c^2)$  and its level curves  $U(f, c) = u$ , with  $u = 46, 41, 34, 25$ . In the plane  $(f, c)$ , the curves implicitly defined by these level curves are shown.

$$\vec{\nabla}U = -2f \hat{f} - 2c \hat{c} \tag{27}$$

So, for example, at the point  $(2, 3)$  in the  $(f, c)$  plane, the gradient is the vector  $-4 \hat{f} - 6 \hat{c}$ , and at  $(1, -1)$ , it is  $-2 \hat{f} + 2 \hat{c}$ .

In Fig. 7 we plot the gradient (27), rescaled for presentation purposes, and the curves in the  $(f, c)$  plane implicitly defined by  $U(f, c) = u$  for various values of  $u$ .

Looking at Fig. 7, it becomes obvious that the gradient vector field of this function has the following properties:

1. It is perpendicular (orthogonal) to the level curves at every point.

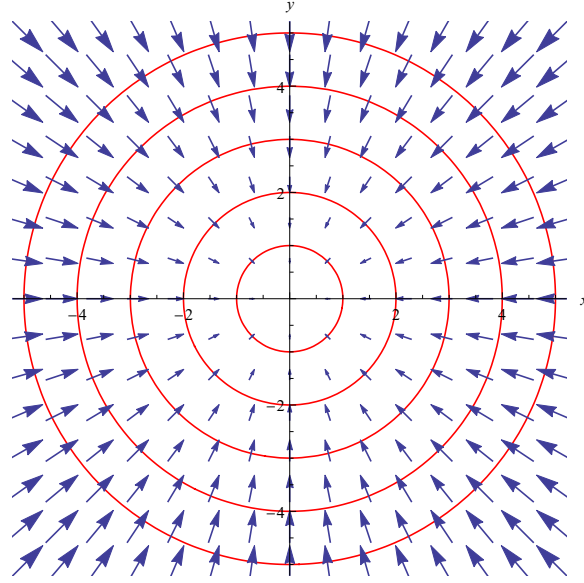


Figure 7:  $\vec{\nabla}U(f, c) = -2f \hat{f} - 2c \hat{c}$  (rescaled), and curves in the  $(f, c)$  plane defined by  $U(f, c) = u$  for different values of  $u$ .

2. The direction of the gradient is the direction where  $U$  increases most. It is the direction of steepest ascent.
3. The magnitude of the gradient is proportional to the “steepness” of  $U$  at each point.
4. In particular, at a local maximum (or minimum), where the steepness is zero, the gradient is zero.

These properties, it turns out, are valid for every sufficiently smooth function  $U(f, c)$ , and in fact extend to any number of variables. Let us prove the first property.

The equation  $U(f, c) = u$  defines locally an implicit function  $c(f)$ . In paper [\[Pernice \(2018 a\)\]](#) we studied how to compute the derivative of this implicit function  $c(f)$  with respect to  $f$  using the chain rule for the total derivative of  $U$  with respect to  $f$ . If  $c(f)$  is a level curve, and we change  $f$  and  $c$  so as to remain in the level curve, then  $dU/df = 0$ . On the other hand, from the chain rule

$$\frac{dU}{df} = \frac{\partial U}{\partial f} + \frac{\partial U}{\partial c} \frac{dc}{df} = 0 \quad (28)$$

This equation can be interpreted vectorially as

$$\left( \frac{\partial U}{\partial f} \hat{f} + \frac{\partial U}{\partial c} \hat{c} \right) \cdot \left( \hat{f} + \frac{dc}{df} \hat{c} \right) = \vec{\nabla}U \cdot \left( \hat{f} + \frac{dc}{df} \hat{c} \right) = 0 \quad (29)$$

where “ $\cdot$ ” is the “dot”, or “scalar” product between the two vectors. As can be seen for example in mentioned reference, the second vector,  $\hat{f} + \frac{dc}{df} \hat{c}$ , is tangent to the curve  $c(f)$ .

Remember that, given two vectors  $\vec{r}$  and  $\vec{s}$ ,  $\vec{r} \cdot \vec{s} = \|\vec{r}\| \|\vec{s}\| \cos(\theta)$ , where  $\theta$  is the angle between the two vectors, and  $\|\vec{r}\|$  ( $\|\vec{s}\|$ ) is the magnitude of the vector  $\vec{r}$  ( $\vec{s}$ )<sup>3</sup>, so,  $\hat{f} \cdot \hat{f} = 1$ ,  $\hat{c} \cdot \hat{c} = 1$  and  $\hat{f} \cdot \hat{c} = \hat{c} \cdot \hat{f} = 0$  (since the angle between  $\hat{f}$  and  $\hat{c}$  is  $90^\circ$ , and  $\cos 90^\circ = 0$ ). From these relations we see that (28) can be written vectorially as (29).

The fact that the scalar product of  $\vec{\nabla}U(f, c)$  and a vector tangent to the curve  $c(f)$  (defined implicitly by  $U(f, c) = u$ ) is zero, means that they are perpendicular to each other. This finishes the proof of the first property. As will be seen in another work, this proof trivially generalizes to many variables.

## 4 Geometric interpretation of Lagrange multiplier method

Remember that our problem consisted in maximizing a function  $U(f, c)$  (that we interpreted as a utility function, but can be any general function), not by exploring all possible points in the  $(f, c)$  plane, but by exploring only those points for which the function  $I(f, c) = i$ . We interpreted this constraint as the budget constraint, but again, it can be any constraint.

Think now about  $I(f, c)$  not as a constraint but as a function in its own right. Moving along the constraint  $I(f, c) = i$  means moving along a level line of  $I$ . So the gradient of  $I$ ,  $\vec{\nabla}I$ , is perpendicular to the tangent of these level lines.

Equations (23-24) can then be interpreted as

$$\vec{\nabla}U = \lambda \vec{\nabla}I \quad (30)$$

which means, geometrically, that at the constrained optimum, the gradients of  $U$  and  $I$  are proportional to each other. Parallel or antiparallel depending on whether  $\lambda$  is, respectively, positive or negative. The Lagrange multiplier is just the constant of proportionality.

Remember that in section 2.2 we arrived intuitively at the conclusion that, at the constrained maximum of  $U(f, c)$ , the constraining curve  $c_I(f)$  should be tangent to the level curve of  $U(f, c)$  corresponding to that maximum. Or, what amounts to the same thing, vectors tangent to the level curves of  $U(f, c)$  and  $I(f, c)$  should be proportional (parallel or antiparallel) to each other. But if the tangent vectors are proportional to each other, the vectors in the two dimensional  $(f, c)$  plane that are perpendicular to them, should also be proportional to each other at the maximum, see Fig. 8. This is what equation 30 is saying. Since equations (23-24) are equivalent to 30, this gives the geometrical meaning of the first two equations of the Lagrange multiplier method.

The need for the third equation (25) was already explained in section 2.2. In terms of gradients, note that, as we saw in Fig. 7,  $\vec{\nabla}I$  gives a vector perpendicular to a level line of  $I$  at every point in the  $(f, c)$  plane, i.e., for every possible value of  $i$ . So equation (25) selects the level line  $I(f, c) = i$  with the particular  $i$  corresponding to our problem.

What about the nature of the new variable  $\lambda$ ? The geometric perspective presented here also

---

<sup>3</sup>This is elementary vector algebra. It will be covered in a different work in vector spaces.

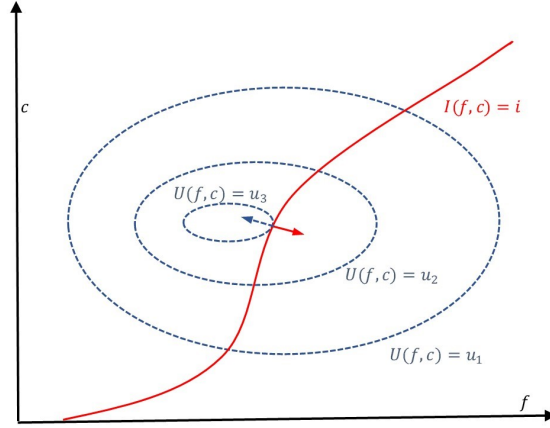


Figure 8: The tangent of the level curves of  $U$  and  $I$  should be proportional to each other. So, in the two dimensional  $(f, c)$  plane, the vectors perpendicular to them should also be proportional to each other.

explains the necessity of this new variable. As we saw, the magnitude of the gradient is proportional to the steepness of the underlying function at each point, and the geometric intuitive argument says that at the maximum, the gradients of  $U$  and  $I$  should be parallel or antiparallel to each other. But if what we demand from these vectors is just to be parallel or antiparallel, the relative magnitude of these vectors is completely undefined. This ambiguity is what demands the introduction of the new variable  $\lambda$ . In fact,  $|\lambda|$  is the ratio of the magnitudes of the two gradients (see problem 4: 1).

This relative magnitude is of course determined once the gradients of the specific functions  $U$  and  $I$ , at a specific point  $(f, c)$ , are fixed. This is what is done by solving simultaneously the 3 equations with 3 unknowns (23-25).

Very much related to  $\lambda$  being the constant of proportionality between the two gradients, there is another relationship with a nice economic interpretation. So let us return now to the interpretation of  $U$  as utility and of  $I$  as budget. Remember that for general changes  $df$  and  $dc$  in the variables  $f$  and  $c$ ,

$$dI = \frac{\partial I}{\partial f}df + \frac{\partial I}{\partial c}dc = \vec{\nabla}I \cdot (df\hat{f} + dc\hat{c}) \quad (31)$$

$$dU = \frac{\partial U}{\partial f}df + \frac{\partial U}{\partial c}dc = \vec{\nabla}U \cdot (df\hat{f} + dc\hat{c}) \quad (32)$$

Suppose we want to see how the optimum utility changes when the budget changes in  $di$ . If we fix  $dI = di$ , the vector  $df\hat{f} + dc\hat{c}$  in (31) is not arbitrary anymore. There has to be a relationship between  $df$  and  $dc$  so that  $\vec{\nabla}I \cdot (df\hat{f} + dc\hat{c}) = di$ . On the other hand, at the constrained optimum, equation (31) becomes  $dU = \lambda\vec{\nabla}I \cdot (df\hat{f} + dc\hat{c}) = \lambda di$ .

This result is sufficiently important to number the corresponding equation: when the budget increases by  $di$ , the optimal choice with the new budget produces a change in utility give by

$$dU = \lambda di \quad (33)$$

i.e., the Lagrange multiplier gives the increase in utility per unit change in the budget constraint.

How about the generalization to many dimensions and many constraints? As mentioned before, this will be covered in another work, but just to present a glimpse of why this method generalizes trivially to many dimensions while the other methods presented here do not, let us consider one peculiar aspect of the arguments presented so far. While the graphical argument presented in section 2.2 naturally lead us to guess the proportionality of the *tangent* vectors of the level curves of  $U$  and  $I$ , the Lagrange multipliers method points to the proportionality of the gradients of  $U$  and  $I$ , which are *perpendicular* to the level curves of  $U$  and  $I$ . Now, in two dimensions, as we saw, one tangent direction determines (up to a sign) a unique perpendicular direction, so the two arguments are equivalent. But this is not so in three or more dimensions!

Consider for example the maximization of a three variables function  $U(f, c, w)$  subject to the constraint  $I(f, c, w) = i$ . The equivalent to the level curves of  $U$  and  $I$  of Fig. 3 will now be level surfaces in the three dimensional space  $(f, c, w)$ . But in surfaces there are infinitely many directions tangent to the surface at any given point, while there is still only one (up to a sign) perpendicular direction. So, while an equation between vectors perpendicular to these surfaces (as gradients are) is unambiguous, it would not be so for tangent vectors.

## 5 Conclusions

We have provided an intuitive interpretations of the Lagrange multipliers method, and of the multiplier itself. For that, we have used concepts like vectors, vector fields, gradients, etc. Although we tried to include sufficient content to make the paper self contained, we will present these concepts in detail in a different work.



## Problems:

### Section 2.1:

- For  $\alpha = \beta = 1/2$  and  $0 \leq f \leq I/P_f$ , find the maximum of (6) by taking the derivative and equaling to zero.
  - By taking the second derivative confirm that it is a maximum
- Do the same for arbitrary  $\alpha$  and  $\beta$ .
- Do the derivation of (7) step by step, explicitly stating the properties of calculus used in each step.

### Section 2.2:

- Equation (15) looks identical to (10). However, while in (10)  $c(f)$  was assumed explicitly known, in (15) we do not make such assumption. Convince yourself, and write down your arguments explicitly, of the different meaning of these two equations.
- For  $\alpha = \beta = 1/2$  and  $0 \leq f \leq i/P_f$ , find the explicit function  $c^*(f)$  determined by (16).
  - Solve the system (16-17). c) Make graphs of  $c^*(f)$  and the budget level line for  $P_f = 1, P_c = 2, I = 10$ .
- Do (a) and (b) of the previous problem for arbitrary  $\alpha$  and  $\beta$ .
- Explain why the right hand side of (18) is the same for every value  $i$  of the constrain  $I(f, c) = i$  that defines implicitly  $c_I(f)$ .

### Section 4:

- Show that the absolute value of the Lagrange multiplier  $\lambda$ , is the ratio of the magnitudes of the two gradients, i.e.,  $|\lambda| = \frac{\|\vec{\nabla} U\|}{\|\vec{\nabla} I\|}$ .

## References

[Pernice (2018 a)] Pernice S., (2018). "Intuitive Mathematical Economics Series: Chain Rule and Derivatives of Functions Defined Implicitly", *Ucema Working Papers* 679 (2018).