Intuitions About Lagrangian Optimization

Department of Economics
University of Guelph
Discussion Paper Number 2010-3

January 2010

Dan Kalman
Department of Mathematics and Statistics
American University
4400 Massachusetts Avenue, NW
Washington, DC 20016-8050
USA
Email: kalman@american.edu

Michael Hoy
Department of Economics
University of Guelph
Guelph, ON N1G 2W1
Canada
Email: mhoy@uoguelph.ca

Keywords: Constrained optimization, Lagrange method, transformation fallacy

JEL Codes: A2, C02

Intuitions About Lagrangian Optimization

Abstract: We expose a weakness in an intuitive description popularly associated with the method of Lagrange multipliers and propose an alternative intuition. According to the deficient intuition, the Lagrange technique transforms a constrained optimization problem into an unconstrained optimization problem. This is both mathematically incorrect, and in some contexts contrary to a basic understanding of economic principles. In fact, as is probably understood by most instructors, solutions to the Lagrange conditions for a constrained optimization problem are generally saddle points, an observation typically included in advanced treatments of mathematical economics. At the introductory level, however, instructors often 'cut corners' and emphasize that the first-order conditions associated with the method of Lagrange multipliers are essentially the same as for an unconstrained optimization problem, hence leading to an incorrect intuition. We propose an alternative intuition that introduces the Lagrangian function as a perturbation of the original objective function. We characterize a constrained optimum as a point at which all the derivatives of a suitable perturbation must vanish. The paper is both useful for instructors of introductory courses in mathematical methods for economics and also can be used to provide enrichment to students for this very important mathematical technique.

Intuitions About Lagrangian Optimization

The method of Lagrange multipliers is a common topic in elementary courses in mathematical economics and continues as one of the most important tools used in more advanced undergraduate and graduate courses. In a typical application, we wish to maximize (or minimize) a function $f(\mathbf{x})$ subject to a constraint $g(\mathbf{x}) = 0$. Here, $\mathbf{x}$ can be thought of as a point $(x_1, x_2, \ldots, x_n)$ with any number of entries, although for concreteness it is often convenient to assume $n = 2$. Under certain technical assumptions about the existence of derivatives of f and g, the Lagrange technique can be stated as follows. Define the Lagrangian function $F(\mathbf{x},\lambda) = f(\mathbf{x}) + \lambda\, g(\mathbf{x})$, where λ is a real variable. If, among the points where $g(\mathbf{x}) = 0$, $f(\mathbf{x})$ assumes a maximum at $\mathbf{x}^*$, then there exists a λ^* such that all of the partial derivatives of F vanish at the point $(\mathbf{x}^*, \lambda^*)$. The same applies for a constrained minimum. In symbols, $\nabla F(\mathbf{x}^*, \lambda^*) = 0$, defining what are sometimes called the *first order conditions* for a constrained optimum.

The first order conditions provide a means for identifying potential solutions of the constrained optimization problem. Among all the points $(\mathbf{x}, \lambda)$ where the partial derivatives of F equal zero, the desired solution $(\mathbf{x}^*, \lambda^*)$ must occur. Frequently the context of the constrained problem tells us that there will be exactly one local maximum[1] and the Lagrange conditions provide a single candidate $\mathbf{x}$, which must therefore be the desired solution. This is the situation most often seen in introductory courses. And in this situation, solving a constrained optimization problem amounts to finding the solution to the Lagrange conditions. Understandably, to many students (and some instructors),

[1] For ease of discussion, we will not always mention that the same statements apply for a minimum of f($\mathbf{x}$). Furthermore, we will mostly ignore the possibility of nonuniqueness.

mastering the Lagrange method is synonymous with remembering how to formulate the first order conditions.

Here there is the tempting (although unfortunately erroneous) intuitive conception: *The constrained maximum of* $f(\mathbf{x})$ *occurs at the same* $\mathbf{x}$ *as the unconstrained maximum of* $F(\mathbf{x}, \lambda)$. Although this idea is algorithmically attractive, it is conceptually misleading, as we shall see. In the sequel, this incorrect conception will be referred to as the *transformation fallacy*, because it suggests that the Lagrangian technique transforms a constrained optimization problem into an unconstrained optimization problem. The transformation fallacy has been analyzed mathematically by Kalman (2009).

From an algorithmic standpoint, the transformation fallacy does little if any harm. It leads to the correct first order conditions, because an unconstrained maximum of F must occur at a point where $\nabla F(\mathbf{x}^*, \lambda^*) = 0$. So if correctly solving textbook problems is all that matters, the fallacy should raise no objection. However, mathematical methods in economics are only as valid as the conceptual frameworks in which they arise. Accordingly, students should be encouraged to develop a deeper understanding of mathematical techniques than simple algorithmic mastery. This is particularly true at the introductory level, where the audience includes many students who will not go on to more advanced study of mathematical economics.

So what exactly is wrong with the transformation fallacy? As a conceptual image of the Lagrangian technique, it is simply untrue. The fact is, the constrained maximum of f does *not* occur at the same $\mathbf{x}$ as the unconstrained maximum of F, but rather at an $\mathbf{x}$ where F has a saddle point. To be more precise, if $\mathbf{x}^*$ is the solution to the constrained

optimization problem, and if ($\mathbf{x}^*$, λ^*) is a solution to the first order conditions $\nabla F(\mathbf{x}^*, \lambda^*) = 0$, then as a function of both $\mathbf{x}$ and λ, F has a saddle point at ($\mathbf{x}^*$, λ^*), not a maximum (Kalman, 2009). Indeed, as we will argue shortly, F need not even have a maximum. Even if you hold λ fixed at λ^*, and consider F as a function solely of $\mathbf{x}$, it need not have a maximum at $\mathbf{x}^*$ (Hestenes, 1975). Thus, imagining a correspondence between a constrained maximum of f and unconstrained maximum of F is a false intuition. Perhaps it is the result of some instructors (overzealous) efforts to make it easy for students to remember the algorithm required to find the solution to a constrained optimization that has led to such an incorrect intuition. We feel adopting a correct intuition reinforces the nature of the constrained optimization problem and graphical depictions that relate to the technique for finding the solution.

The above objections are mathematical in nature. Some may wish to dismiss them as technicalities. But there is also a simple economic context in which the transformation fallacy is clearly revealed. In many situations, it is the very presence of a constraint that gives rise to a (finite) maximum value. For example, the variable $\mathbf{x}$ can represent choices for consumption of goods, with f measuring the utility of the goods (i.e., as an ordinal representation of preferences) and g the budget constraint of the consumer. Hence the consumer wishes to optimize his or her choices, receiving the greatest utility $f(\mathbf{x})$, (i.e., most preferred consumption bundle possible), subject to the budget constraint ($g(\mathbf{x}) = 0$). Suppose we write the Lagrangian function $F(\mathbf{x},\lambda) = f(\mathbf{x}) + \lambda\, g(\mathbf{x})$ and imagine that what we want to do to solve the consumer's optimization problem is to choose $\mathbf{x}^*$, λ^* in order to maximize $F(\mathbf{x},\lambda)$. To do this one can just choose $\lambda^* = 0$ and effectively ignore the constraint. Now observe that in the absence of the

constraint, there will be no maximum. The amount consumed can increase without bound, and the utility will go to infinity as seen by the choice of **x*** unbounded (positive infinity) and $\lambda^* = 0$. Thus, the very idea of a maximum in this context is dependent on the existence of the constraint which would be effectively neutralized in the unconstrained optimization of $F(\mathbf{x},\lambda)$.

More generally, it is clear that the Lagrangian function cannot have an unconstrained maximum. For consider the possibility that $\lambda = 0$. Then, since $F(\mathbf{x},0) = f(\mathbf{x})$, it is clear that F will go to infinity as the components of **x** increase without bound whenever $f(\mathbf{x})$ is increasing in at least one x_i. If F has no maximum for $\lambda = 0$, it certainly cannot have a maximum when **x** and λ are free to vary without constraint. Nevertheless, we expect there to be a maximum utility when **x** is constrained. This shows that the transformation fallacy contradicts common sense. The constrained maximum of f cannot be an unconstrained maximum of F, because there will be no such point.

Another common constrained optimization problem in economics is determining the combination of inputs for production of a given output level (the constraint) that will minimize cost of production (minimize $C = r_1x_1 + r_2x_2$ subject to $g(x_1, x_2) - \bar{y} = 0$ where x_i are inputs used, r_i respective unit prices of inputs, $g(x_1, x_2)$ is the production function, and $\bar{y} > 0$ is the required output to be produced. With the usual requirement that inputs must be non-negative one candidate for the solution to the unconstrained problem of minimizing the Lagrangian $F(x_1, x_2, \lambda) = r_1x_1 + r_2x_2 + \lambda[g(x_1, x_2) - \bar{y}]$ is to simply choose $x_1^* = x_2^* = \lambda^* = 0$ which delivers a cost of zero. The problem, of course, is that the constraint is in fact not satisfied at this point. Moreover, since $[g(0,0) - \bar{y}] < 0$, one

can even seem to do better by choosing λ^* to a positive number (without bound), delivering a result of negative infinity.

As these examples show, the transformation fallacy is incompatible with common sense reasoning in some simple and familiar economic models. When and how did the fallacy arise? One speculates that the idea is natural enough, given a certain amount of mathematical license. If one equates the Lagrange method with the solution of the first order equations, and likewise identifies the unconstrained optimization of F with solution of the very same first order equations, then that seems to imply that constrained optimization of f and unconstrained optimization of F are the same problem. Of course, the problem here is that unconstrained optimization is not equivalent to solving first order equations. For example, a maximum must be a solution to the first order equations, but a solution to first order equations need not be a maximum. Nevertheless, subtleties of this sort are often neglected in favor of a simpler and more memorable (if technically less accurate) formulation. If instructors are not vigilant in dispelling this fallacy, students may well end up believing in some dubious rationale for the transformation fallacy.

We suspect that there may be another *deeper* explanation for the persistence of this fallacy, going back to Lagrange's own description of the multiplier method. In translation, he stated the following general principle:

> When a function of several variables must be a maximum or a minimum, and there are one or more equations relating these variables, it will suffice to add to the proposed function the functions that must be equal to zero, multiplied each one by an indeterminate quantity, and next to find the maximum or minimum as if

> these variables were independent; the equations that one will find combined with the given equations will serve to determine all the unknowns. (Lagrange, 1797, pp 198-199).

This is almost a statement of the transformation fallacy. However, in the context of Lagrange's earlier discussion of multiplier techniques in the area of statics, it is clear that he was well aware that the points to be found, where all the partial derivatives equal zero, need not correspond to maxima or minima. From that context, it seems evident that what Lagrange meant was to proceed *as if* seeking a maximum or minimum, and that this would lead to the correct equations. This idea is developed more fully in (Kalman, 2009).

If this interpretation is correct, Lagrange stated something very close to the transformation fallacy as a kind of rhetorical expedient. It was more efficient to say "seek an unconstrained maximum" than to say "solve the equations that would arise if you were to seek an unconstrained maximum." He can be forgiven for believing his readers would all know what he meant. But perhaps over time, the more correct meaning faded, leaving some of the users of Lagrange's method to take his words literally.

Whether or not that explanation is true, the transformation fallacy has become part of the folklore of mathematics and economics, and is often repeated in informal conversation. It appears explicitly or implicitly in elementary books on mathematical economics (Archibald and Lipsey, 1976; Chiang, 1984) and calculus (Hughes-Hallett et al, 1998). In some cases, it is asserted quite directly, as in the following statement:

> We now set out the *Lagrange multiplier technique* ... We proceed by introducing a new variable, λ the **Lagrange multiplier**, and by forming the **Lagrange function** or **Lagrangean**
>
> $$\mathcal{L}(x_1,x_2, \lambda) = f(x_1,x_2) + \lambda\, g(x_1,x_2)$$
>
> We then carry out the unconstrained maximization of $\mathcal{L}$ with respect to x_1, x_2 and λ . (Hoy *et al*., 2001, p. 588)

In contrast, the identification of ($\mathbf{x}^*$, λ^*) as a saddle point of F is common in more advanced treatments on mathematical economics (Jehle, 1991; Takayama, 1974; later in the same text Hoy et al., 2001, pp. 678-679 when presenting more formally the Kuhn-Tucker conditions). Even at this level, though, some imprecision can be observed, as in the following excerpt:

> The optimal solution (x_1^*,x_2^*, λ^*), however, is likely to be a saddle point rather than a maximizer of $\mathcal{L}$ (the Lagrange function): Whereas $\mathbf{x}^*$ does maximize $\mathcal{L}(\mathbf{x}^*,\lambda^*)$, it will not be true in general that λ^* maximizes $\mathcal{L}(\mathbf{x}^*,\lambda^*)$. (de la Fuente, 2000, p. 283)

As mentioned earlier, the point ($\mathbf{x}^*$,λ^*) is essentially *always* a saddle point, although technically it can still be the case that F has a maximum with respect to both $\mathbf{x}$ and λ at that point. In any case, given the fact that advanced mathematical works in

economics generally avoid the transformation fallacy, introductory courses and texts should be careful to steer students away from it.

At the same time, we recognize how appealing the fallacy is to the untutored intuition. At first encounter it is so plausible that it appears to provide a rationale for why the method works. And it is certainly desirable for students to have some such rationale. If the transformation fallacy is to be suppressed, an alternative rationale or intuition is needed.

Happily, there is an intuitive explanation of the Lagrangian technique that fills this need. The transformation fallacy says that the Lagrangian function is a device by which a constrained problem becomes an unconstrained problem. An equally memorable idea, and one that is correct, is that the Lagrangian is an approach that *levels the playing field.*

To make this idea clear, let us consider once again the fundamental problem, to maximize $f(x,y)$ subject to $g(x,y) = 0$. We can interpret this situation graphically. The graph of $z = f(x,y)$ is a surface in three dimensional space, showing how f varies with its two variables. An unconstrained maximum, if one exists, occurs at the top of a hill, where the surface flattens out and, under extreme magnification, takes on the appearance of a horizontal plane. At such a point all of the partial derivatives are zero, because a flat horizontal plane has zero slope in all directions. That is why we can find unconstrained maxima by setting all of the partial derivatives to zero.

But we have to take into consideration the constraint equation $g(x,y) = 0$. Because of this restriction, we cannot wander freely over the surface for f, but rather we are confined to a curve on that surface. Envision a foot path threading through hilly

terrain. The constrained maximum will be at the high point on the path, where it crosses a ridge perhaps, but not at the top of a hill. In particular, the slope along the path will be zero at such a point, but the slope in the cross-path direction will not. For this reason, the constrained maximum will not be revealed by the familiar unconstrained procedure of finding points where the partial derivatives are all zero.

Can we rectify this problem? Is it possible to flatten out the surface for f at the constrained maximum? In any such attempt, we have to be careful not to alter the curve or *foot path* to which we are confined. That has to be left unchanged. But it should be perfectly alright to alter the values of f at points that are not on the constraint curve. In fact, in the direction of this curve, the surface is already flat at the constrained maximum. We can imagine pivoting the surface as if the curve were a hinge. That would leave the constraint curve fixed, and if we pivot by just the right amount, the surface should flatten in the cross-curve direction.

How do we put this idea into effect mathematically? The simplest way to alter the values of the function f is to add some perturbation. But we want the amount we add to be 0 along the constraint curve, so as not to alter f for those points. On the other hand, we know that g is 0 along the constraint curve. So, adding g to f is exactly the sort of perturbation we need. It leaves the values of f unchanged along the constraint curve, but modifies the values away from that curve. More generally, we can add any multiple of g to f and achieve the same effect. Moreover, by varying the multiplier, we can increase or decrease the amount of the perturbation. This introduces the pivoting effect described above, and is the motivation behind defining for each λ a perturbed function

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

In this conception, we do not think of F as a function of three variables. Rather, we have in mind an entire family of two variable functions, one for each value of the parameter λ. A representation of this situation is shown in Figure 1. The surface marked $\lambda = 0$ is the unperturbed graph for f. The other surfaces are graphs of different members of the family of F functions. All of the F functions have to agree along the constraint curve, which is therefore visible in the figure as the curve on which all the surfaces intersect. The high point on that curve is the constrained maximum of f.

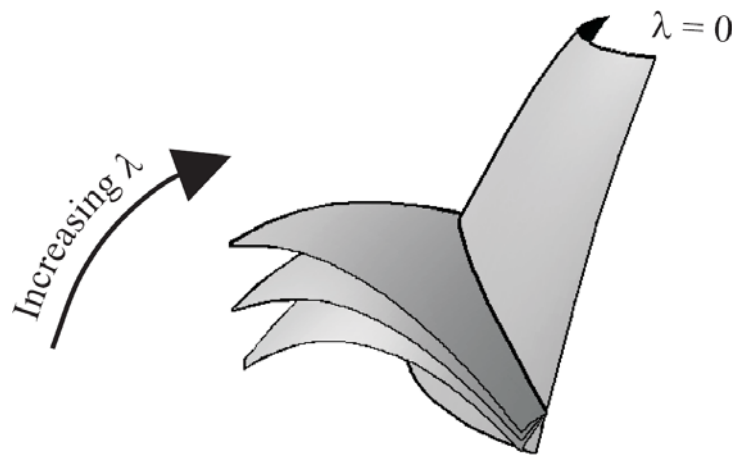

Figure 1: Graphs of several members of the family of F functions.

The figure shows how the variation of λ imposes the pivoting action described earlier. Intuition suggests (and a little further analysis proves) that there must actually be a choice of λ that makes the surface for F horizontal. That is, if (x^*,y^*) is the location of the maximum value of f subject to the constraint $g = 0$, then for some λ^* it will be the case that $\nabla F(x^*,y^*) = 0$. That means the following equations must hold:

$$\frac{\partial f}{\partial x}(x^*, y^*) + (\lambda^*)\frac{\partial g}{\partial x}(x^*, y^*) = 0$$
$$\frac{\partial f}{\partial y}(x^*, y^*) + (\lambda^*)\frac{\partial g}{\partial y}(x^*, y^*) = 0$$
$$g(x^*, y^*) = 0$$

These are the familiar first order equations from the Lagrangian technique, equivalent to the equation $\nabla F(\mathbf{x}^*, \lambda^*) = 0$ cited at the start of the paper. Thus, the perturbation approach explains why, for a suitable λ, the constrained maximum must be a solution to the first order conditions.

This concept of *leveling* also leads immediately to the interpretation of λ at the solution point. Geometrically, we can envision the slope of F in any direction. If the original surface for f has slope m_f in a specified direction, and if the surface for g has slope m_g in that direction, then the surface for F will have slope $m_f + \lambda\, m_g$. At the constrained maximum, we want this to be zero in the direction perpendicular to the constraint curve. Therefore we need to choose $\lambda^* = - m_f / m_g$. This ratio is constant in all directions if the first order conditions are satisfied. Neglecting the sign, this ratio represents the change in f relative to a change in g, and so gives the marginal rate of increase of f as we relax the constraint g. That is, for each unit increase in the constraint g, we expect an increase of m_f / m_g in the constrained maximum of f.

Let us return to the consumer choice problem discussed earlier, but with a change in notation to reflect common usage. An individual wishes to optimize over his or her choices, i.e., to obtain the greatest utility $u(\mathbf{x})$ (i.e., most preferred consumption bundle possible) subject to the budget constraint that expenditure not exceed income available ($g(\mathbf{x}) = 0$ or $p_1x_1 + p_2x_2 = M$). The variables x_i represent the amount of each of two

goods, i = 1, 2 consumed, p_i represents the price per unit of each good, and *M* the consumer's income. Notice that the constraint can be written either as $p_1x_1 + p_2x_2 - M = 0$ or, upon multiplying by *-1*, as $M - p_1x_1 - p_2x_2 = 0$. This makes it clear why the sign of the multiplier can be either negative or positive in the solution. Suppose the utility function is $u(x_1, x_2) = x_1 \cdot x_2^2$, the specific prices are $p_1 = 5$ and $p_2 = 4$, and income is $M = 60$. The Lagrangian function then becomes $F(x_1, x_2, \lambda) = x_1 \cdot x_2^2 + \lambda[60 - 5x_1 - 4x_2]$ and the first-order conditions from the Lagrangian technique means the following equations must hold.

$$\frac{\partial F}{\partial x_1} = \frac{\partial u}{\partial x_1}(x_1^*, x_2^*) + (\lambda^*)\frac{\partial g}{\partial x_1}(x_1^*, x_2^*) = (x_2^*)^2 - 5\lambda^* = 0$$
$$\frac{\partial F}{\partial x_2} = \frac{\partial u}{\partial x_2}(x_1^*, x_2^*) + (\lambda)\frac{\partial g}{\partial x_2}(x_1^*, x_2^*) = 2x_1^*x_2^* - 4\lambda^* = 0$$
$$\frac{\partial F}{\partial \lambda} = g(x_1^*, x_2^*) = 60 - 5x_1^* - 4x_2^* = 0$$

The solution is easy to find and is $x_1^* = 4, x_2^* = 10, \lambda^* = 20$. It follows that treating the function $F(x_1, x_2, \lambda)$ as a family of perturbed functions of $u(x_1, x_2)$ we can see that as we increase λ from $\lambda = 0$ to $\lambda^* = 20$ we eventually generate a function *F* for which the surface is horizontal at the optimal choice $x_1^* = 4, x_2^* = 10$.

As an intuitive device to understand the meaning of the solution λ^* it is sometimes suggested in economics terminology that the result of choosing a point (x_1, x_2) off the constraint that requires more than available income *M* to purchase leads to a sort of *penalty*. This *penalty* is seen in our diagram to reduce value of the function *F* at any point off the constraint such that $p_1x_1 + p_2x_2 > M$. The size of λ determines the size of the *penalty* or, in our terminology, the extent of the leveling of the playing field. The

appropriate penalty corresponds to the value $\lambda^* = 20$. One can think of "the other side of the coin" in interpreting this penalty. If the penalty from breaking the constraint by one (marginal) unit is 20 lost 'utils' of satisfaction (the implied reduction in *u*) then having an extra unit of income, which expands the constraint by one unit, allows for a gain of 20 'utils' once the consumer optimizes again with the new higher income level. So $\lambda^* = 20$ is effectively the marginal value (marginal utility) from having an extra dollar in the budget – or the so-called marginal utility of income. More generally, this explains why the solution of the Lagrange multiplier is often referred to as the shadow value of the constraint.[2] As a check, notice that although increasing *M* from 60 to 61 represents a discrete amount rather than a marginal amount, working through the example with general level of income *M* delivers approximately the result that utility goes up by approximately 20 utils as a result of income rising by a dollar – i.e., the value of an extra dollar is 20 utils. To see this first note that $x_1^* = \frac{M}{15}, x_2^* = \frac{2.5M}{15}$. Substituting these optimal values into the objective function $u(x_1, x_2)$ gives $u^* = 400$ for $M = 60$ and $u^* \approx 420.3$ for $M = 61$. This implies that an increase in income of \$1 from $M = 60$ to 61 generates an increase in the optimal value of the objective function of approximately 20 utils, which is the value of λ^* when $M = 60$.

Along the lines similar to that of Figure 1, the four panels (a through d) of Figure 2 demonstrate how the shape of the Lagrangian function, $F(x_1, x_2, \lambda) = x_1 \cdot x_2^2 + \lambda[60 - 5x_1 - 4x_2]$, changes as the value of λ changes. In each of the graphs the constraint is represented by the (black) vertical plane (i.e., the plane emanating

[2] If the constraint had been written as $p_1x_1 + p_2x_2 - M = 0$ rather than as $M - p_1x_1 - p_2x_2 = 0$ then the solution for the multiplier would have been λ* = -20. The sign just depends on the way the constraint is written and so it is just the magnitude of the solution that is relevant.

vertically from the two-dimensional budget line $5x_1 + 4x_2 = 60$) while the curved "blue-purple" shape represents the Lagrangian function for the various values of the multiplier λ. For λ = 0 (Figure 2a), the Lagrangian is simply the objective function, $u(x_1, x_2) = x_1 \cdot x_2^2$. The intersection of the constraint with the objective function traces out the 'footpath threading through hilly terrain' as described earlier. The constrained maximum is the highest point on this path. Notice that for all values of λ in Figures 2a through 2d the shape of the Lagrangian doesn't change (as in Figure 1) and so the constrained maximum is identifiable as the same point in each of these.

For λ = 10, the Lagrangian function is 'flattened out' relative to the objective function in that it appears to be increasing 'less steeply' along the constraint.[3] For λ = 20, the flattening out is 'just right' in that the penalty from breaking the constraint is such that no increase in the value of the Lagrangian function $F(x_1, x_2, \lambda)$ results from a marginal increase (or decrease) in the choice variables x_1 and/or x_2 relative to the optimal solution values. Finally, if one chooses a value of λ that is 'too large' (e.g., λ = 30 in Figure 2d) then the penalty is excessive and increasing the value of (x_1, x_2) actually leads to a reduction in the value of the Lagrangian while decreasing the value of (x_1, x_2) would lead to an increase in the value of the Lagrangian.

[3] Of course, as λ changes, the shape of the function changes everywhere else in its domain. In particular, as x_1 and/or x_2 approaches zero, the fact that $[60 - 5x_1 - 4x_2 > 0]$ is increasing means that for greater values of (positive) λ the part of the Lagrangian that depends on the constraint , $+ \lambda[60 - 5x_1 - 4x_2]$, leads to increases in the overall value of the Lagranginan. This explains why the Lagrangian function twists upward more in this region as one considers higher value of the multiplier λ.

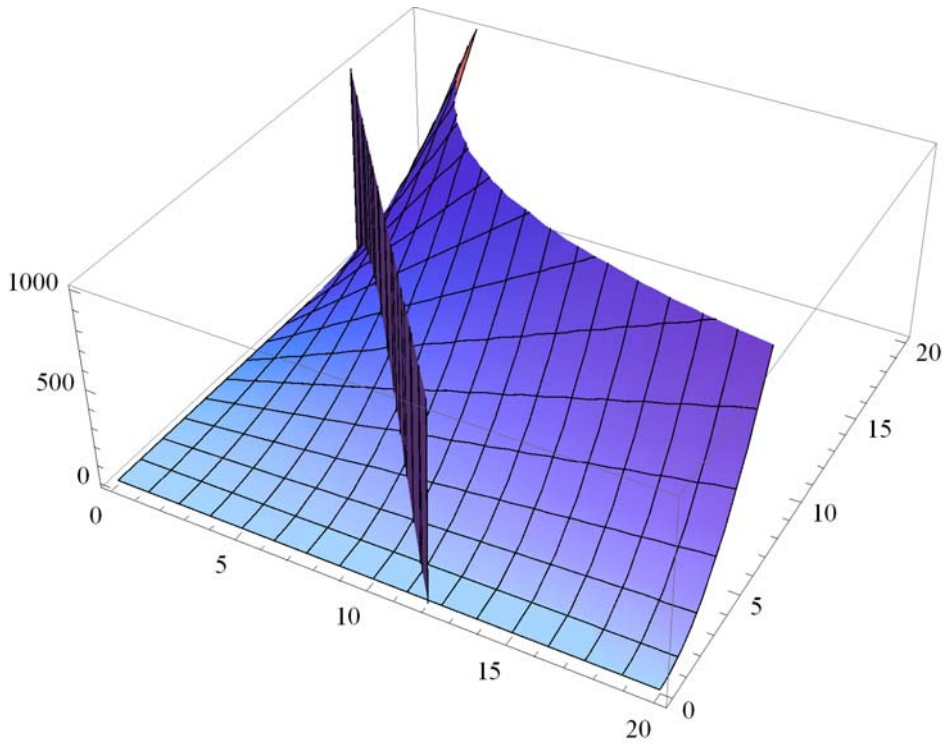

Figure 2a: Lagrange function with $\lambda = 0$.

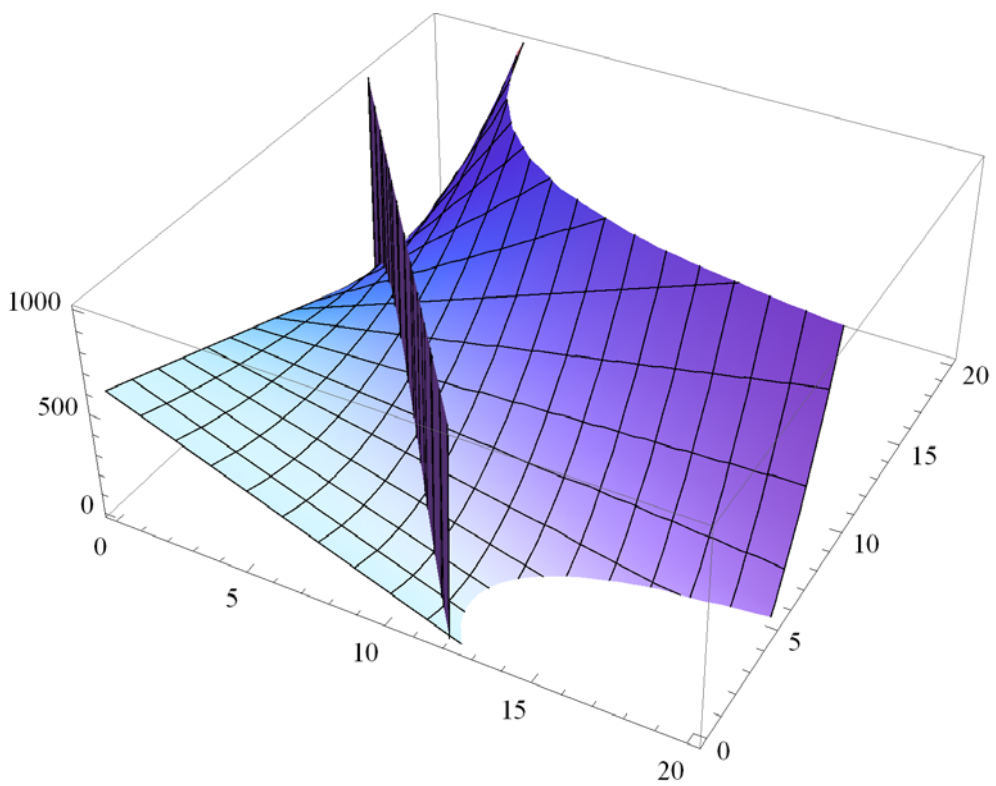

Figure 2b: Lagrange function with $\lambda = 10$.

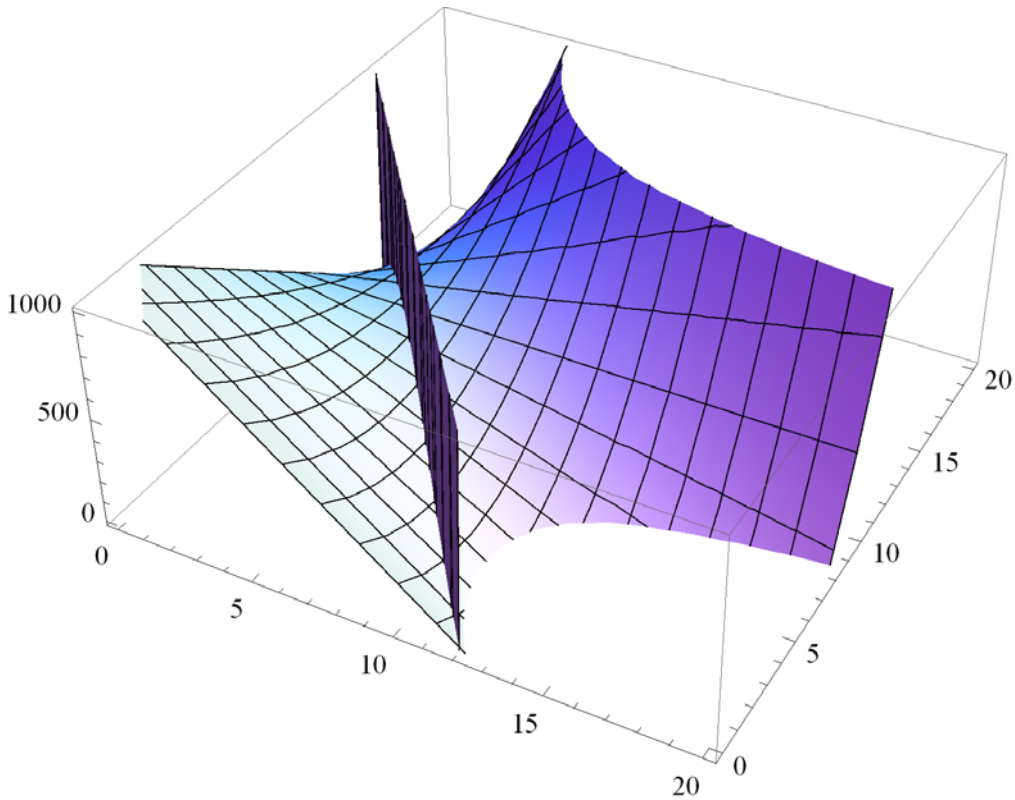

Figure 2c: Lagrange function with λ = 20.

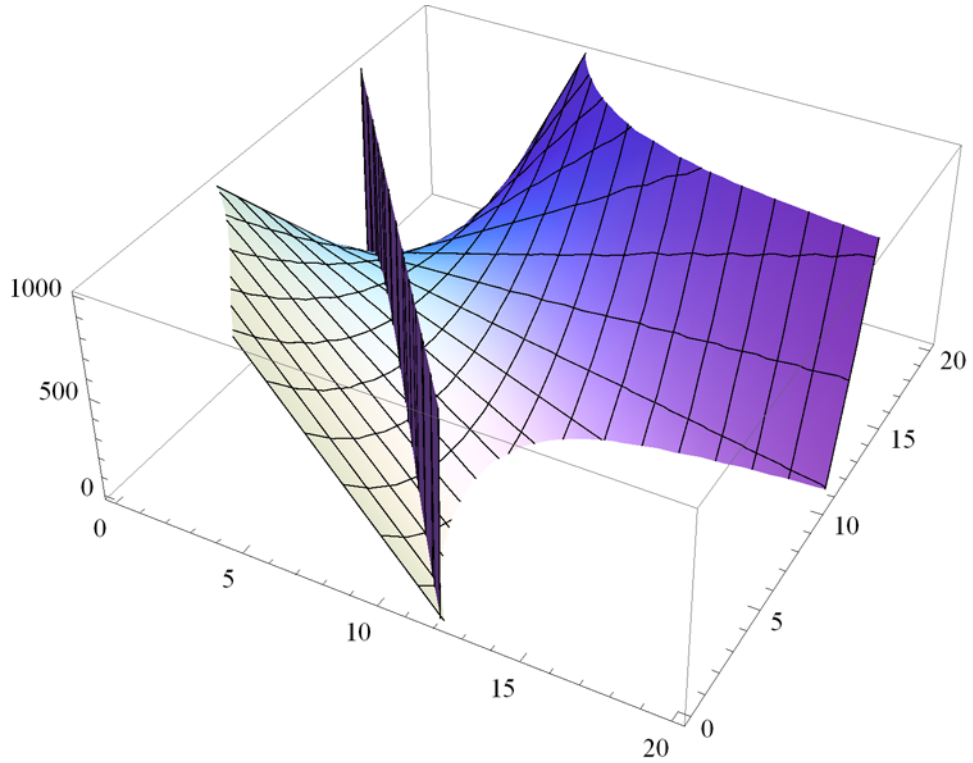

Figure 2d: Lagrange function with λ = 30.

In order not to promote a different sort of fallacy, which we will call the marginal ordinal utility of income fallacy, we want to emphasize an important point about the consumer problem. In modern consumer theory, the utility function used in the standard consumer choice problem is an ordinal measure representing the individual's preference over alternative commodity bundles. All we can or need say is that if one consumption bundle *A* is preferred to another bundle *B*, then the value of the utility function is greater for *A* than for *B*. The amount by which it is bigger is not relevant. Therefore, for example, one could double the utility function used in our example (i.e., use $u(x_1, x_2) = 2x_1 \cdot x_2^2$) and perform the same constrained optimization exercise.[4] The optimal solution to this revised problem is $x_1^* = 4, x_2^* = 10, \lambda^* = 40.$ The optimal consumption decision is the same but the multiplier becomes double what it was for the original problem. This follows simply because the utility function is now double what it was before and so the 'marginal utility of income' is also double. From a mathematical perspective this makes perfect sense but from an economic perspective it represents an important lesson: namely, one must take care when interpreting a solution to the lagrange multiplier as a shadow value of a constraint if the value function being optimized has only ordinal and not cardinal significance. For objectives like costs (to be minimized) or profits (to be maximized), such a concern does not arise as a doubling of costs, for example, is actually a meaningful statement since the concept of costs has cardinal significance.

In general, the image of leveling the graph of *f* using the Lagrangian family *F* offers an attractive way to think about Lagrange multipliers. In this formulation, we find a clear and convincing intuitive justification of the Lagrangian approach. It also makes

[4] Once can perform any positive monotonic transformation (that preserves continuity and differentiability) of the utility function, i.e., $G(u(x_1, x_2))$, with $G' > 0$, and not affect the solution (x_1^*, x_2^*).

the existence of the necessary λ completely transparent, while leading naturally to the interpretation of λ as the marginal rate of change of the objective function relative to the constraint. This has the potential of enriching the insight of students, providing an intuitive understanding of the Lagrangian first order conditions. And it is certainly superior to perpetuating the transformation fallacy.

References

Archibald, G. C. and Richard G. Lipsey. 1976. *An Introduction to Mathematical Economics: Methods and Applications,* Harper and Row, New York.

Chiang, Alpha C. 1984. *Fundamental Methods of Mathematical Economics,* third ed., McGraw-Hill, New York.

de la Fuente, Angel. 2000. *Mathematical Methods and Models for Economists*, Cambridge University Press, Cambridge, UK, page 283.

Hestenes, Magnus R. 1975. *Optimization Theory the Finite Dimensional Case,* Wiley, New York.

Hoy, Michael, John Livernois, Chris McKenna, Ray Rees, and Thanasis Stengos. 2001. *Mathematics for Economics,* second ed., MIT Press, Cambridge, MA.

Hughes-Hallett, Deborah, Andrew M. Gleason, William G. McCallum, et al. 1998. *Calculus Single and Multivariable,* second ed., Wiley, New York.

Jehle, Geoffrey A. 1991. *Advanced Microeconomic Theory*, Prentice-Hall, Reading, MA, p. 106.

Kalman, Dan. 2009. Leveling with Lagrange: an alternate view of constrained optimization. *Mathematics Magazine* , Volume 82, Number 3, pp 186 - 196.

Lagrange, J. L. 1797. *Théorie des fonctions analytiques contenant les principes du calcul différentiel, dégagés de toute considération d'infiniment petits ou d'évanouissans, de limites ou de fluxions, et réduits à l'analyse algébrique des quantités finies*, De l'imprimerie de la République, Paris.

Takayama, Akira. 1974. *Mathematical Economics,* Dryden Press, Hinsdale, IL, p. 96.