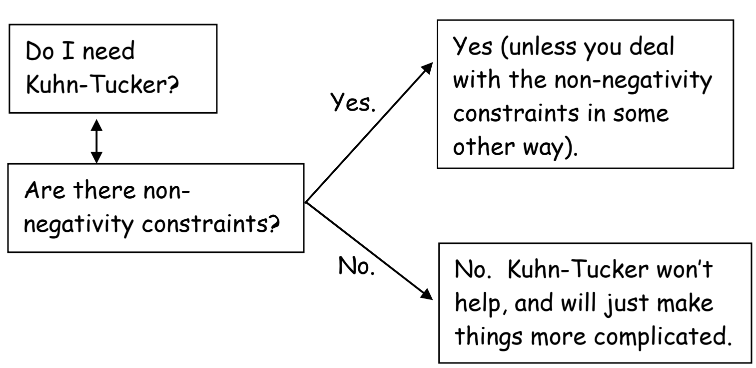
Nolan Miller Sept., 2007

You, the Kuhn-Tucker Conditions, and You

Why should I bother with the Kuhn-Tucker conditions? The Kuhn-Tucker conditions should be neither difficult nor scary. They may, however, be confusing, at least when seeing them for the first time.

What do the Kuhn-Tucker conditions do? Simply put, they are **a short cut for writing down the first-order conditions for a constrained optimization problem when there are also non-negativity constraints on the variables over which you are maximizing.** You could include each of the non-negativity constraints explicitly, adding each as a constraint in the Lagrangian with an associated Lagrange multiplier. However, in a problem with many variables, this can become unwieldy. The Kuhn-Tucker formulation incorporates these non-negativity constraints by imposing additional restrictions on the first-order conditions for the problem, rather than including additional terms in the Lagrangian. But, the important thing to remember is that **the Kuhn-Tucker conditions are nothing more than a short cut for writing down the first-order conditions for a constrained optimization problem when there are non-negativity constraints on the variables over which you are maximizing.**

So, when you ask yourself "Do I need to use Kuhn-Tucker here?" the next question is "Are there non-negativity constraints on the variables that I'm optimizing over?" If the answer is "yes," then you need to take these constraints into account when trying to solve the problem. The Kuhn-Tucker approach is one way to do this, and it is often the most parsimonious way to do it. On the other hand, ***if there are no non-negativity constraints on the variables over which you are optimizing, then the Kuhn-Tucker conditions are not necessary and will not help you.***



Intuition for the Kuhn-Tucker Conditions

I'm going to focus on intuition here. What I say will be less mathematically rigorous than a textbook presentation, but will help develop the intuition more. I suggest using both sources to help get a full understanding of the issues.

Let's think about a very simple problem: max $f(x,y)$ s.t. $g(x,y) \leq k$ and $x \geq 0$ and $y \geq 0$. Suppose the feasible region is as in Figure 1.

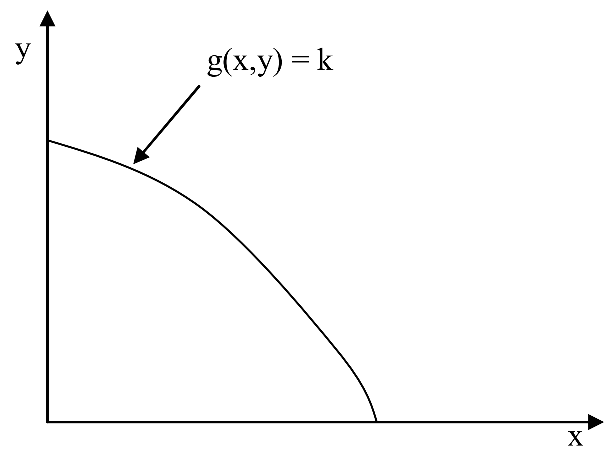


Figure 1

The feasible region here is below the constraint and above the axes. What kinds of solutions can there be? Well, the constraint can either bind or not, and $x^* = 0$ or not, and $y^* = 0$ or not. This gives us a number of cases to consider. Actually, there are seven in total (See Figure 2).

1. $g(x^*,y^*) < k$, $x^* > 0$, $y^* > 0$.
2. $g(x^*,y^*) = k$, $x^* > 0$, $y^* = 0$.
3. $g(x^*,y^*) < k$, $x^* > 0$, $y^* = 0$.
4. $g(x^*,y^*) < k$, $x^* = 0$, $y^* = 0$.
5. $g(x^*,y^*) < k$, $x^* = 0$, $y^* > 0$.
6. $g(x^*,y^*) = k$, $x^* = 0$, $y^* > 0$.
7. $g(x^*,y^*) = k$, $x^* > 0$, $y^* > 0$.

Note: there is in principle an eighth possibility where $g(x,y) = k$, $x = 0$, $y = 0$. But, this is only possible if the constraint goes through the origin. In the diagram as drawn, that is not possible.

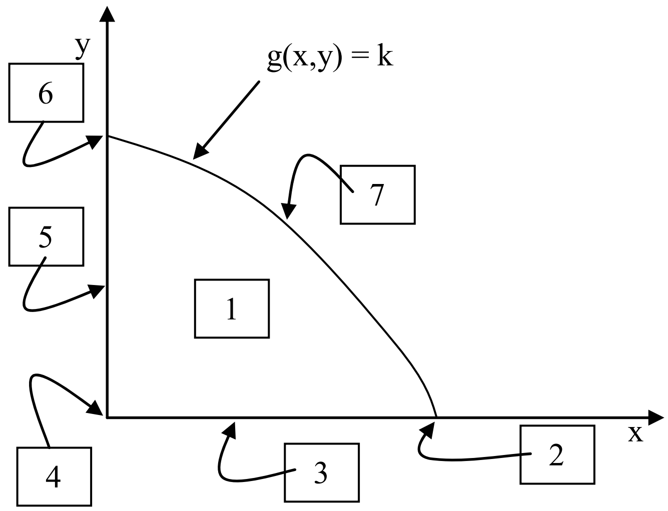


Figure 2

Now, let's take a step back and think a bit about non-negativity constraints in general. Consider maximizing a function subject to a non-negativity constraint. Call the function h(z). The non-negativity constraint is $z \geq 0$. There are two basic types of solutions (See Figure 3). First, consider $h_1(z)$, with maximum at $z^* > 0$. In this case, the maximum is at a flat point of the function. Hence $h_1'(z^*) = 0$. Second, consider $h_2(z)$, with maximum at $z^{**} = 0$. In this case, the function could be flat at $z^{**} = 0$. But, it could also be downward sloping, as in the diagram. So, $h_2'(z^{**}) \leq 0$.

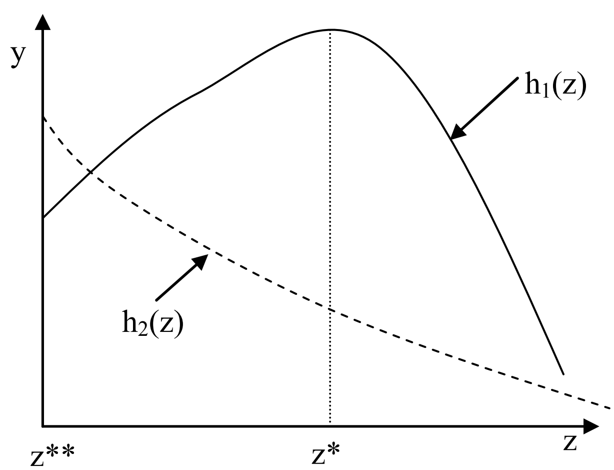


Figure 3

Since we don't know whether the function h(z) looks like $h_1(z)$ or $h_2(z)$, we need a set of first-order conditions that allows for either case. These can be written as: either $z^* = 0$ and $h'(z^*) \leq 0$, or $z^* > 0$ and $h'(z^*) = 0$. An even shorter way of writing this is $h'(z^*) \leq 0$ and $z^* (h'(z^*)) = 0$. This is essentially what the Kunh-Tucker first-order conditions are. **They are a shorthand way of writing down these two cases for each of the variables over which you are maximizing.**

Seen in this light, the Kuhn-Tucker conditions are not all that different from the complementary slackness conditions you already use to deal with the constraint in the usual Lagrangian formulation. Recall, that if you have max f(x,y) s.t. $g(x,y) \leq k$ (without imposing non-negativity on x and y yet), the Lagrangian is written:

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

and the first-order conditions are (using subscripts to denote partial derivatives):

$f_x(x^*,y^*) - \lambda^*\, g_x(x^*,y^*) = 0$,
$f_y(x^*,y^*) - \lambda^*\, g_y(x^*,y^*) = 0$, and
$g(x^*,y^*) - k \leq 0$ and $\lambda^*\,(g(x^*,y^*) - k) = 0$.

What's going on with the last condition? This takes care of the idea that either the constraint binds, or else the Lagrange multiplier is zero. Thus it is a concise way of writing that the solution could either be in region (7) or region (1).

The Kuhn-Tucker conditions are relevant in problems where we add the non-negativity constraints $x \geq 0$ and $y \geq 0$ to the constrained maximization problem. They incorporate the restrictions implied by these constraints directly into the first-order conditions. In other words, they capture the way the first-order conditions change when the solution is in regions (2) – (6).

Intuitively, what do the Kuhn-Tucker conditions do, and where do they come from? Let's think about the intuition for what the additional requirements should be. First, think about x. If $x^* > 0$, then the derivative of the Lagrangian must be zero (like case 1 in figure 3). (Intuitively, if the optimum is not at zero, then the non-negativity constraint on x doesn't matter, so the condition for the optimum on x must be just as in the unconstrained problem.) On the other hand, if $x^* = 0$, then the derivative of the Lagrangian needs only to be less than or equal to zero (like case 2 in figure 3). We can write these two cases with respect to x in short hand as:

$f_x(x^*,y^*) - \lambda^* g_x(x^*,y^*) \leq 0$ and $x^* (f_x(x^*,y^*) - \lambda^* g_x(x^*,y^*)) = 0$.

In our discussion of figure 3, we wrote $h(z^*) \leq 0$ and $z^* (h(z^*) = 0$. So, these conditions really just combine the intuition for the two cases discussed in Figure 3 with the Lagrangian approach we use for any constrained problem.

Similarly, the condition for y should allow for either $y^* = 0$ or $y^* > 0$. This gives us:

$f_y(x^*,y^*) - \lambda^* g_y(x^*,y^*) \leq 0$ and $y^* (f_y(x^*,y^*) - \lambda^* g_y(x^*,y^*)) = 0$.

Since the first-order condition for the Lagrangian with respect to λ already allows for the possibilities of the constraint binding or not, we don't need to adjust it. So, we are left with the Kuhn-Tucker conditions:

$f_x(x^*,y^*) - \lambda^* g_x(x^*,y^*) \leq 0$ and $x^* (f_x(x^*,y^*) - \lambda^* g_x(x^*,y^*)) = 0$,
$f_y(x^*,y^*) - \lambda^* g_y(x^*,y^*) \leq 0$ and $y^* (f_y(x^*,y^*) - \lambda^* g_y(x^*,y^*)) = 0$, and
$g(x^*,y^*) - k \leq 0$ and $\lambda^* (g(x^*,y^*) - k) = 0$.

Again, when you have a constrained optimization problem with non-negativity constraints on the endogenous variables, you can use the Kuhn-Tucker conditions as a shorthand notation that captures all of the possible cases for an optimum.

Let's see how this works. What we want to do is think about the solution to the constrained maximization problem in a particular case, and show that the Kuhn-Tucker conditions incorporate the intuitive conditions that must hold in that case. In doing so, we will make use of the intuition you already know for a constrained optimization problem without non-negativity: namely that if the constraint doesn't bind, then the optimized value of the Lagrange multiplier (λ^*) must be zero.

Consider a particular case, such as (3) above. Here we have that $y = 0$ and $x > 0$ and $g(x,y) < k$. Suppose there is an optimal solution in this region. Intuitively (i.e., just by thinking about the picture), what conditions must the optimal solution satisfy? Since $g(x^*,y^*) < k$ in this region, the Lagrange multiplier must be zero: $\lambda^* = 0$. That is, relaxing the constraint further does not affect the value of the objective function. Since $x^* > 0$, it must be that the objective function is flat with respect to x: $f_x(x^*,y^*) = 0$. Since $y^* = 0$, it cannot be that the objective function would increase if we increased y. Thus, it must **not** be that $f_y(x^*,y^*) > 0$. But, this is the same as saying that $f_y(x^*,y^*) \leq 0$. So, we have the following:

$x^* > 0$ and $f_x(x^*,y^*) = 0$.
$y^* = 0$ and $f_y(x^*,y^*) \leq 0$.
$g(x^*,y^*) - k < 0$ and $\lambda^* = 0$.

Do these intuitive properties of the solution imply the Kuhn-Tucker conditions?

Since $\lambda^* = 0$, $\lambda^* (g_x(x^*,y^*) - k) = 0$. This tells us that $f_x(x^*,y^*) - \lambda^* (g_x(x^*,y^*) - k) = 0$. And, something that is equal to zero is also less than or equal to zero. So, the intuition tells us that:

$f_x(x^*,y^*) - \lambda^* g_x(x^*,y^*) \leq 0$ and $x^* (f_x(x^*,y^*) - \lambda^* g_x(x^*,y^*)) = 0$,

For y, we have that $\lambda^* (g_y(x^*,y^*) - k) = 0$. This tells us that $f_y(x^*,y^*) - \lambda^* (g_y(x^*,y^*) - k) \leq 0$, and since $y^* = 0$, we have that:

$f_y(x^*,y^*) - \lambda^* g_y(x^*,y^*) \leq 0$ and $y^* (f_y(x^*,y^*) - \lambda^* g_y(x^*,y^*)) = 0$.

Finally, since λ^*, we have that:

$g(x^*,y^*) - k \leq 0$ and $\lambda^* (g(x^*,y^*) - k) = 0$.

Thus, the intuitive properties of the solution in this case imply the Kuhn-Tucker conditions. So, we have argued that if there is a solution in region (3) it satisfies the Kuhn-Tucker conditions. You should go through each of the cases to verify that the intuitive properties of the solution imply the Kuhn-Tucker conditions.

As a final note, recall that the Kuhn-Tucker conditions are necessary but not sufficient. That means that any solution must satisfy them, but satisfying them does not mean that the point actually solves the constrained maximization problem. But, what it does mean is that **if a solution exists** and **if there is a unique point satisfying the Kuhn-Tucker conditions, then this is the unique solution to the maximization problem.**

Special Bonus Example:

Consider a situation where the solution to the constrained maximization problem with non-negativity constraints is at (2) in Figure 2. What does this mean? Since this is a corner solution, it must be that the gradient of the objective function at (2) points up and to the right. That is, the objective function increases only by moving out of the feasible region. Such a case is depicted in Figure 4.

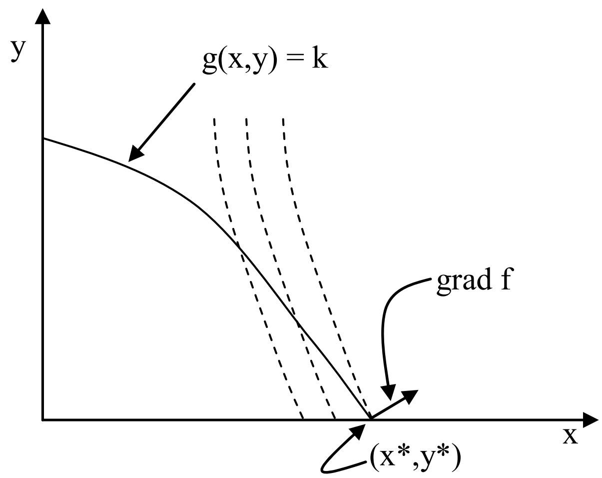


Figure 4

Consider Figure 4. We have assumed that f(x,y) increases as you move up and to the right, and we've drawn the indifference curves for the objective function as dashed lines. Notice that they are steeper than the constraint at (x^*,y^*). This must be the case if (x^*,y^*) is a maximum.

How do we formalize the idea that the indifference curve must be steeper than the constraint at (x^*,y^*)? Consider the constraint: $g(x,y) = k$. The slope of the constraint at a particular point is found by implicitly differentiating. Let $y(x)$ implicitly denote the constraint. Then we have:

$$g(y(x),y) \equiv k.$$

This is an identity, so we can implicitly differentiate both sides with respect to y.

This gives us:

$$g_x(x,y) + g_y(x,y) * dy/dx = 0, \text{ or}$$

$$dy/dx = - g_x(x,y)/g_y(x,y).$$

So, the slope of the constraint at point (x^*,y^*) is $-g_x(x^*,y^*)/g_y(x^*,y^*)$. Similarly, the slope of the indifference curve for $f(x,y)$ that goes through (x^*,y^*) is

$-f_x(x^*,y^*) / f_y(x^*,y^*)$.

Now, the indifference curve for f is steeper than the constraint. This means that the absolute value of the slope of f is greater than the absolute value of the slope of the constraint. This gives:

$$f_x(x^*,y^*) / f_y(x^*,y^*) > g_x(x^*,y^*) / g_y(x^*,y^*)$$

Each of these derivatives is positive, so we can rearrange:

$$f_x(x^*,y^*) / g_x(x^*,y^*) > f_y(x^*,y^*) / g_y(x^*,y^*). \qquad (\#)$$

Now, equation (#) captures our intuition about the solution to the maximization problem. We know (#), $x^* > 0$, $y^* = 0$, and $g(x^*,y^*) = k$. Do these facts imply the Kuhn-Tucker conditions are satisfied?

First, since $x^* > 0$, the usual first-order condition for a constrained maximization problem should hold. So, we have:

$f_x(x^*,y^*) - \lambda^* (g_x(x^*,y^*)) = 0$, or

$\lambda^* = f_x(x^*,y^*) / g_x(x^*,y^*) > 0$.

Combine this last fact with (#) above. This gives:

$f_y(x^*,y^*) / g_y(x^*,y^*) < \lambda^* = f_x(x^*,y^*) / g_x(x^*,y^*)$, or
$f_y(x^*,y^*) - \lambda^* g_y(x^*,y^*) < 0$.

So, to summarize, we have

(a) $x^* > 0$, $f_x(x^*,y^*) - \lambda^* g_x(x^*,y^*) = 0$
(b) $y^* = 0$, $f_y(x^*,y^*) - \lambda^* g_y(x^*,y^*) < 0$
(c) $g(x^*,y^*) - k^* = 0$, $\lambda^* > 0$.

It is easy to confirm that the Kuhn-Tucker first-order conditions (see the double-lined box above) are satisfied here. So, once again we have started with geometric intuition for what the solution must be like in a particular case (in this case the corner solution (2)), and we've shown that this intuition is captured in the Kuhn-Tucker conditions.