ECON 772001
MATH FOR ECONOMISTS

Peter Ireland

Boston College

September 3, 2020

The Kuhn-Tucker Theorem

Associated with the problem:

$$\max_x F(x) \text{ subject to } c \geq G(x)$$

Define the Lagrangian $L : \mathbb{R}^2 \to \mathbb{R}$:

$$L(x, \lambda) = F(x) + \lambda[c - G(x)]$$

The Kuhn-Tucker Theorem

Theorem (Kuhn-Tucker) Let x^* maximize $F(x)$ subject to $c \geq G(x)$, where F and G are both continuously differentiable. Assume, as well, that $G'(x^*) \neq 0$. Then there exists a value λ^* or λ that, together with x^*, satisfies

$$L_1(x^*, \lambda^*) = F'(x^*) - \lambda^* G'(x^*) = 0 \tag{1}$$

$$L_2(x^*, \lambda^*) = c - G(x^*) \geq 0 \tag{2}$$

$$\lambda^* \geq 0 \tag{3}$$

$$\lambda^*[c - G(x^*)] = 0 \tag{4}$$

The Kuhn-Tucker Theorem

Proof Case 1: nonbinding constraint.

If x^* is such that $c > G(x^*)$

$$L_1(x^*, \lambda^*) = F'(x^*) - \lambda^* G'(x^*) = 0 \quad (1)$$

$$L_2(x^*, \lambda^*) = c - G(x^*) \geq 0 \quad (2)$$

$$\lambda^* \geq 0 \quad (3)$$

$$\lambda^*[c - G(x^*)] = 0 \quad (4)$$

set $\lambda^* = 0$.

The Kuhn-Tucker Theorem

With $\lambda^* = 0$

$$L_1(x^*, \lambda^*) = F'(x^*) - \lambda^* G'(x^*) = 0 \tag{1}$$

$$L_2(x^*, \lambda^*) = c - G(x^*) \geq 0 \tag{2}$$

$$\lambda^* \geq 0 \tag{3}$$

$$\lambda^*[c - G(x^*)] = 0 \tag{4}$$

(2)-(4) clearly hold.

The Kuhn-Tucker Theorem

Hence, it only remains to show that

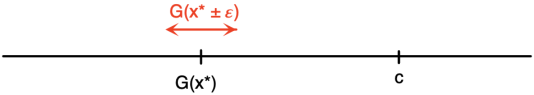
$$L_1(x^*, \lambda^*) = F'(x^*) - \lambda^* G'(x^*) = 0 \tag{1}$$

is satisfied.

But, with $\lambda^* = 0$, (1) holds if and only if

$$F'(x^*) = 0 \tag{5}$$

The Kuhn-Tucker Theorem



Since, by assumption $c > G(x^*)$, and since G is continuously differentiable, small changes in x^* can be made for the sake of argument, without violating the constraint.

The Kuhn-Tucker Theorem

To show that (5) holds, assume instead that $F'(x^*) \neq 0$. One way this can happen is if $F'(x^*) < 0$.

But then, the continuous differentiability of F and G implies that there exists an $\varepsilon > 0$ such that

$$F(x^* - \varepsilon) > F(x^*) \text{ and } c > G(x^* - \varepsilon)$$

contradicting the assumption that x^* solves the problem.

The Kuhn-Tucker Theorem

Similarly, if $F'(x^*) > 0$, then the continuous differentiability of F and G implies that there exists an $\varepsilon > 0$ such that

$$F(x^* + \varepsilon) > F(x^*) \text{ and } c > G(x^* + \varepsilon)$$

again contradicting the assumption that x^* solves the problem.

Hence (5) and therefore (1) hold, completing the proof for this first case.

The Kuhn-Tucker Theorem

Summary: If $c > G(x^*)$, so that the constraint does not bind, "the constraint is not really a constraint."

In this case, set $\lambda^* = 0$ to "zero it out" in the Lagrangian

$$L(x, \lambda) = F(x) + \lambda[c - G(x)]$$

The constraint (2), the nonnegativity condition (3), and the complementary slackness condition "automatically" hold. And the continuous differentiability of G allows us to make small adjustments to x^* in confirming that the FOC holds.

The Kuhn-Tucker Theorem

Proof Case 2: binding constraint.

If x^* is such that $c = G(x^*)$

$$L_1(x^*, \lambda^*) = F'(x^*) - \lambda^* G'(x^*) = 0 \tag{1}$$

$$L_2(x^*, \lambda^*) = c - G(x^*) \geq 0 \tag{2}$$

$$\lambda^* \geq 0 \tag{3}$$

$$\lambda^*[c - G(x^*)] = 0 \tag{4}$$

set $\lambda^* = F'(x^*)/G'(x^*)$.

The Kuhn-Tucker Theorem

If x^* is such that $c = G(x^*)$, set $\lambda^* = F'(x^*)/G'(x^*)$.

This is possible, since the constraint qualification $G'(x^*) \neq 0$ holds.

The Kuhn-Tucker Theorem

With $c = G(x^*)$ and $\lambda^* = F'(x^*)/G'(x^*)$

$$L_1(x^*, \lambda^*) = F'(x^*) - \lambda^* G'(x^*) = 0 \tag{1}$$

$$L_2(x^*, \lambda^*) = c - G(x^*) \geq 0 \tag{2}$$

$$\lambda^* \geq 0 \tag{3}$$

$$\lambda^*[c - G(x^*)] = 0 \tag{4}$$

(1), (2), and (4) clearly hold.

The Kuhn-Tucker Theorem

Hence, it only remains to show that

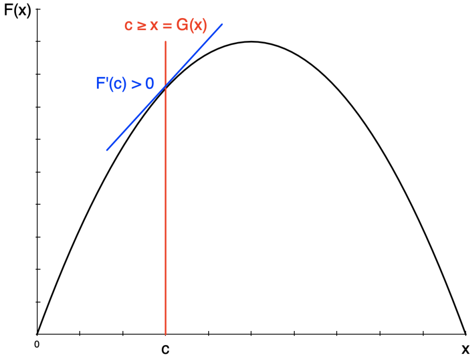
$$\lambda^* \geq 0 \tag{3}$$

is satisfied or, equivalently, that

$$F'(x^*)/G'(x^*) \geq 0 \tag{6}$$

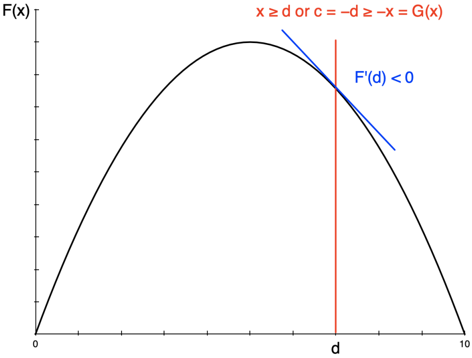
Equation (6) requires that $F'(x^*)$ and $G'(x^*)$ have the same sign. Why must this be true?

The Kuhn-Tucker Theorem



Here: $G(x^*) = x^*$. $F'(x^*) > 0$ means "we'd like to increase x^*." But $G'(x^*) = 1 > 0$ means "the constraint won't let us."

The Kuhn-Tucker Theorem



Here: $c = -d$ and $G(x^*) = -x^*$. $F'(x^*) < 0$ means "we'd like to decrease x^*." But $G'(x^*) = -1 < 0$ means "the constraint won't let us."

The Kuhn-Tucker Theorem

Therefore, $F'(x^*)$ and $G'(x^*)$ having the same sign means that any desirable change in x^* is blocked by the constraint.

To prove that (6) must hold, suppose instead that

$$F'(x^*)/G'(x^*) < 0$$

One way this can happen is if $F'(x^*) < 0$ and $G'(x^*) > 0$. But then, the continuous differentiability of F and G implies that there exists an $\varepsilon > 0$ such that

$$F(x^* - \varepsilon) > F(x^*) \text{ and } c = G(x^*) > G(x^* - \varepsilon)$$

contradicting the assumption that x^* solves the problem.

The Kuhn-Tucker Theorem

Similarly, if $F'(x^*) > 0$ and $G'(x^*) < 0$. But then, the continuous differentiability of F and G implies that there exists an $\varepsilon > 0$ such that

$$F(x^* + \varepsilon) > F(x^*) \text{ and } c = G(x^*) > G(x^* + \varepsilon)$$

again contradicting the assumption that x^* solves the problem.

Hence (6) and therefore (3) must hold, completing the proof.

The Kuhn-Tucker Theorem

Summary: If $c = G(x^*)$, so that the constraint binds, the constraint (2) and complementary slackness condition (4) are "automatically" satisfied.

In this case, the constraint qualification allows us to "solve" the FOC (1) for λ^*.

Then, we can show that unless $\lambda^* \geq 0$, it will be possible to adjust x^* to get a higher value of the objective without violating the constraint.

The Kuhn-Tucker Theorem

In the more general case, with $n \geq 1$ choice variables and m ($n \geq m \geq 1$) constraints, of which $\bar{m} \leq m$ bind at the optimum, there are n first-order conditions and $\bar{m}$ non-zero multipliers.

The n FOCs can be written as

$$F'(x^*) - \lambda^* G'(x^*) = 0_{1\times n}$$

where $F'(x^*)$ is an $1 \times n$ vector of partial derivatives, λ^* is a $1 \times \bar{m}$ vector of non-zero multipliers, $G'(x^*)$ is a $\bar{m} \times n$ matrix of partial derivatives, and $0_{1\times n}$ is an $1 \times n$ vector of zeros.

The Kuhn-Tucker Theorem

The n FOCs can be written as

$$F'(x^*) - \lambda^* G'(x^*) = 0_{1 \times n}$$

where $F'(x^*)$ is an $1 \times n$ vector of partial derivatives, λ^* is a $1 \times \bar{m}$ vector of non-zero multipliers, $G'(x^*)$ is a $\bar{m} \times n$ matrix of partial derivatives, and $0_{1 \times n}$ is an $1 \times n$ vector of zeros.

When $n = 1$ and $m = 1$, the constraint qualification $G'(x^*) \neq 0$ allowed us to find the value $\lambda^* = F'(x^*)/G'(x^*)$ that satisfies the FOC. In the more general case, the constraint qualification will require the $\bar{m} \times n$ matrix $G'(x^*)$ to have maximal rank $\bar{m}$.

Comments on The Kuhn-Tucker Theorem

We can, and will, extend the results to apply to problems with more than one choice variable and more than one constraint.

The theorem, even in its more general form, provides necessary, but not sufficient, conditions for x^*.

Comments on The Kuhn-Tucker Theorem

In practice, how can one use the Kuhn-Tucker conditions to find x^*?

Check the SOCs. At least then you'll know you've found a local maximum. SOCs will be covered on a problem set.

Find all the (x^*, λ^*) pairs that satisfy the KT conditions. Typically, these will be of limited number. Then. pick the pair associated with the largest value of F. See the notes for an example.

Impose the further assumptions that F is concave and G is convex. Then the KT conditions are sufficient. See the notes for a proof.

Comments on The Kuhn-Tucker Theorem

When F is concave and G is convex:

The problem is called a "concave program." The KT conditions are both necessary and sufficient.

(x^*, λ^*) is a saddle point of L:

x^* maximizes $L(x^*, \lambda^*)$: (1) is the FOC for this problem.

λ^* minimizes $L(x^*, \lambda^*)$ subject to $\lambda \geq 0$: (2) is the FOC for this problem.

Comments on The Kuhn-Tucker Theorem

The saddle point result provides intuition for the Kuhn-Tucker theorem. The idea is to convert the constrained problem into an unconstrained problem by forming the Lagrangian and maximizing L rather than F subject to $c \geq G(x)$.

This intuition is exactly right when applied to a concave program.

Comments on The Kuhn-Tucker Theorem

Without the extra concavity and convexity assumptions, x^* will not necessarily maximize L. Still, x^* will be a critical point of L. See the notes for examples.

Extending the results to apply to quasi-concave functions requires further refinements. See Kenneth Arrow and Alain Enthoven. "Quasi-Concave Programming." *Econometrica* (1961).

Comments on The Kuhn-Tucker Theorem

In the Lagrangian, λ^* plays the role of a "shadow price" or "exchange rate."

Consider a utility maximization example:

$$L(c_1, c_2, \lambda) = U(c_1, c_2) + \lambda(Y - p_1 c_1 - p_2 c_2)$$

$U(c_1, c_2)$ is "utility," Y is "dollars."

Thus, λ^* must be "utility per dollar" or the "marginal utility of income." The envelope theorem formalizes this idea.

Comments on The Kuhn-Tucker Theorem

The constraint qualification is needed only for the case of a binding constraint.

It almost always holds in practice.

And when it does not, it is often because the problem has been written in an unnecessarily complicated way. This will be illustrated in a problem set question.

Comments on The Kuhn-Tucker Theorem

Finally, consider the problem, augmented by a nonnegativity constraint on x:

$$\max_x F(x) \text{ subject to } c \geq G(x) \text{ and } x \geq 0.$$

One approach treats the nonnegativity constraint symmetrically, by defining

$$L(x, \lambda, \mu) = F(x) + \lambda[c - G(x)] + \mu x$$

and writing the FOC as

$$F'(x^*) - \lambda^* G'(x^*) + \mu^* = 0 \tag{1'}$$

where $x^* \geq 0$, $\mu^* \geq 0$, and $\mu^* x^* = 0$.

Comments on The Kuhn-Tucker Theorem

$$\max_x F(x) \text{ subject to } c \geq G(x) \text{ and } x \geq 0.$$

Alternatively, following Kuhn and Tucker, define

$$L(x, \lambda) = F(x) + \lambda[c - G(x)]$$

and write the FOC as

$$F'(x^*) - \lambda^* G'(x^*) \leq 0 \text{ with equality if } x^* > 0 \qquad (1'')$$

Comments on The Kuhn-Tucker Theorem

$$F'(x^*) - \lambda^* G'(x^*) + \mu^* = 0 \qquad (1')$$

where $x^* \geq 0$, $\mu^* \geq 0$, and $\mu^* x^* = 0$.

$$F'(x^*) - \lambda^* G'(x^*) \leq 0 \text{ with equality if } x^* > 0 \qquad (1'')$$

If $x^* > 0$, $(1'')$ holds with equality and, since $\mu^* = 0$, $(1')$ says the same thing.

If $x^* = 0$, $(1'')$ holds as an inequality and, since $\mu^* > 0$, $(1')$ still says the same thing.

Comments on The Kuhn-Tucker Theorem

$$\max_x F(x) \text{ subject to } c \geq G(x) \text{ and } x \geq 0.$$

With Kuhn and Tucker's formulation

$$L(x, \lambda) = F(x) + \lambda[c - G(x)]$$

there is additional symmetry across (1) and (2):

$$L_1(x^*, \lambda^*) = F'(x^*) - \lambda^* G'(x^*) \leq 0 \text{ with equality if } x^* > 0$$

$$L_2(x^*, \lambda^*) = c - G(x^*) \geq 0 \text{ with equality if } \lambda^* > 0$$

which also emphasizes the saddle point property of (x^*, λ^*) when F is concave and G is convex.