Natalia Lazzati

Mathematics for Economics

Note 8: Nonlinear Programming - The Kuhn-Tucker Problem

Note 8 is based on de la Fuente (2000, Ch. 7) and Simon and Blume (1994, Ch. 18 and 19).

## The Kuhn-Tucker Problem

Let $f : \mathbb{R}^n \to \mathbb{R}$, and consider the problem

$$\max_{\mathbf{x}} \{f(\mathbf{x}) : \mathbf{x} \in U\} \tag{P.K-T}$$

where

$$U = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \leq b_1, ..., g_k(\mathbf{x}) \leq b_k\}.$$

We assume $k \leq n$, that is, the number of constraints is at most equal to the number of decision variables. As you can notice, the only difference from the Lagrange problem is that the constraints are now written as weak inequalities, rather than equalities.[1]

Let us start again with a simple case of two variables of choice and one constraint

$$\max_{x_1,x_2} \{f(x_1, x_2) : g(x_1, x_2) \leq b\}. \tag{1}$$

In Figure 1 the thicker curves are the points that satisfy $g(x_1, x_2) = b$, and the region to the left and below these curves is the constraint set $g(x_1, x_2) \leq b$. The thinner lines are the level sets of the objective function $f$.

In the left panel, the highest level curve that meets the constraint set meets it at the point $(x_1^*, x_2^*)$. Since this point satisfies $g(x_1^*, x_2^*) = b$, we say that the constraint is binding. At $(x_1^*, x_2^*)$ the level sets of $f$ and $g$ are tangent. Therefore, as in the Lagrange problem, their slopes satisfy, for some $\lambda$,

$$Df(x_1^*, x_2^*) - \lambda Dg(x_1^*, x_2^*) = (0, 0). \tag{2}$$

However, this time the sign of the multiplier is important! Recall that $Df(x_1^*, x_2^*)$ points in the direction in which $f$ increases most rapidly at $(x_1^*, x_2^*)$. In particular, $Dg(x_1^*, x_2^*)$ points to the set

[1] All the results in this note remain valid if $f : X \to \mathbb{R}$ where $X$ is an open set in $\mathbb{R}^n$.

$g(x_1, x_2) \geq b$ not to the set $g(x_1, x_2) \leq b$. Since $(x_1^*, x_2^*)$ maximizes $f$ on the set $g(x_1, x_2) \leq b$, the gradient of $f$ cannot point to the constraint set. If it did, we could increase $f$ and still satisfy $g(x_1, x_2) \leq b$. This means that at a maximum like $(x_1^*, x_2^*)$, $Df(x_1^*, x_2^*)$ and $Dg(x_1^*, x_2^*)$ must point in the same direction, i.e. $\lambda$ must be positive.

It follows that for this first case we still form the Lagrangian function

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda [g(x_1, x_2) - b] \tag{3}$$

and set

$$\frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} = 0 \text{ and } \frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} - \lambda \frac{\partial g}{\partial x_2} = 0. \tag{4}$$

Here we also require the constraint qualification that the maximizer not be a critical point of the constraint function $g$. Before considering $\partial L / \partial \lambda$ we need to think about the situation in the right panel of Figure 1.

Suppose that the maximum of $f$ on the constraint set $g(x_1, x_2) \leq b$ occurs at a point $(x_1^{**}, x_2^{**})$ where $g(x_1^{**}, x_2^{**}) < b$. In this case we say the constraint is not binding. Since the constraint is not effective, then $(x_1^{**}, x_2^{**})$ must be an optimal point of the unconstrained problem. Hence, from Note 6, we know that

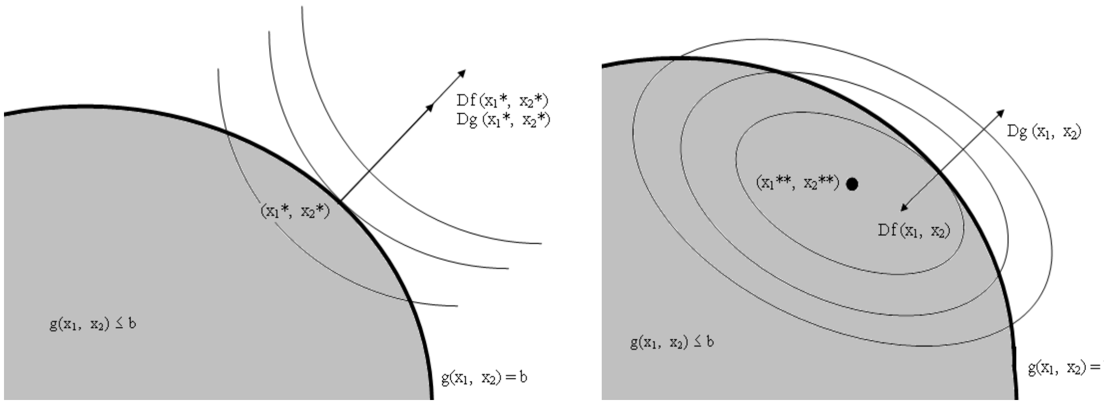
$$\frac{\partial f}{\partial x_1} = 0 \text{ and } \frac{\partial f}{\partial x_2} = 0 \tag{5}$$

at $(x_1^{**}, x_2^{**})$. Notice that we can still use the Lagrangian function (3) and set $\partial L / x_1$ and $\partial L / x_2$ equal zero provided that $\lambda$ equals zero too.

In summary, either the constraint is binding, that is, $g(x_1, x_2) - b = 0$, as in the left panel of Figure 1, in which case the multiplier $\lambda$ must be $\geq 0$, or the constraint is not binding as in the right panel of Figure 1, in which case the multiplier $\lambda$ must be 0. Such a condition in which one of two inequalities must be binding is called a complementary slackness condition. The criterion that either $g(x_1, x_2) - b = 0$ or $\lambda = 0$ holds can be stated as

$$\lambda [g(x_1, x_2) - b] = 0. \tag{6}$$

Figure 1. Graphical Representation of the Kuhn-Tucker Problem




All these observations are captured by the next theorem.

**Theorem 1. (Kuhn-Tucker conditions for $n = 2$ and $k = 1$)** Let $f$ and $g$ be $C^1$ functions. Suppose that $(x_1^*, x_2^*)$ is a (local) maximum of $f$ on the constraint set $\left\{(x_1, x_2) \in \mathbb{R}^2 : g(x_1, x_2) \leq b\right\}$. If $g(x_1^*, x_2^*) = b$, suppose that

$$\frac{\partial g}{\partial x_1}(x_1^*, x_2^*) \neq 0 \text{ or } \frac{\partial g}{\partial x_2}(x_1^*, x_2^*) \neq 0.$$

Then, there exists a multiplier $\lambda^*$ such that the conditions below hold.

| Marginal Conditions | Complementary Slackness Conditions |
|---|---|
| $Df(x_1^*, x_2^*) - \lambda^* Dg(x_1^*, x_2^*) = (0,0)$ | |
| $\lambda^* \geq 0$ | $\lambda^* [g(x_1^*, x_2^*) - b] = 0$ |

## Necessary and Sufficient Conditions for Local Maxima

The generalization of Theorem 1 to more variables of choice and several inequality constraints is quite simple. Let us first motivate the first-order conditions.

An inequality constraint, $g_i(\mathbf{x}) \leq b_i$, is binding or active at a feasible point $\mathbf{x}'$ if it holds with equality $[g_i(\mathbf{x}') = b_i]$, and not binding or inactive if it holds with strict inequality $[g_i(\mathbf{x}') < b_i]$. Intuitively, it is clear that only binding inequalities matter and that the others have no effect on the local properties of the maximizer. Hence, if we knew from the beginning which inequalities were binding at an optimum, then the Kuhn-Tucker problem would reduce to a Lagrange problem in which we take the active constraints as equalities and ignore the rest.

A good recipe to remembering the first-order conditions consists in introducing a vector of multipliers $\lambda_1, ..., \lambda_k$, one for each constraint, and writing the Lagrangian

$$L(\mathbf{x}, \boldsymbol{\lambda}) \equiv f(\mathbf{x}) - \lambda_1 [g_1(\mathbf{x}) - b_1] - ... - \lambda_k [g_k(\mathbf{x}) - b_k]. \tag{7}$$

Next we proceed as if we wanted to maximize $L(\mathbf{x}, \boldsymbol{\lambda})$ with respect to $\mathbf{x}$ and $\boldsymbol{\lambda}$ without constraints. This yields the following first-order conditions

$$D_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = Df(\mathbf{x}^*) - \sum_{i=1}^{k} \lambda_i^* Dg_i(\mathbf{x}^*) = (0, ..., 0) \tag{8}$$

$$D_{\lambda_i} L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = g_i(\mathbf{x}^*) - b_i \leq 0 \text{ and } g_i(\mathbf{x}^*) - b_i = 0 \text{ if } \lambda_i^* > 0 \text{ for } i = 1, ..., k \tag{9}$$

$$\lambda_i^* \geq 0 \text{ and } \lambda_i^* = 0 \text{ if } g_i(\mathbf{x}^*) - b_i < 0 \text{ for } i = 1, ..., k \tag{10}$$

where the last condition is the complementary slackness condition. There is a very natural economic interpretation for the latter. Remember the motivation for the Lagrangian function: instead of directly forcing the agent to respect the constraint, we allowed him to choose the value of the choice variables freely, but made him pay a fine $\mu$ "per unit violation" of the restriction. If at his optimal choice of $\mathbf{x}$, $\mathbf{x}^*$, he does not violate the restrictions, then we do not need to make him pay the fee!

The only extra thing we need to consider is the NDCQ. As the next theorem states, the qualification constraint applies only to the binding constraints and it is identical to the one introduced for the Lagrange problem.

**Theorem 2. (Kuhn-Tucker conditions)** Let $f, g_1, ..., g_k$ be $C^1$ functions. Suppose that $\mathbf{x}^*$ is a local maximum of $f$ on $U$. For ease of notation, assume that the first $e$ constraints are binding at $\mathbf{x}^*$ and that the last $k - e$ constraints are not binding. Suppose the following NDCQ is satisfied at $\mathbf{x}^*$ : the rank of the Jacobian matrix of the binding constraints

$$D\mathbf{g}_e(\mathbf{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_e}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial g_e}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

evaluated at $\mathbf{x}^*$ is $e$, where $\mathbf{g}_e = (g_1, ..., g_e)^T$. In other words, $D\mathbf{g}_e(\mathbf{x}^*)$ has full rank $e$.

Then, there exist multipliers $\lambda_1^*, ..., \lambda_k^*$ such that the conditions below hold.

| | Marginal Conditions | Complementary Slackness Conditions |
|---|---|---|
| (K-T 1) | $Df(\mathbf{x}^*) - \sum_{i=1}^{k} \lambda_i^* Dg_i(\mathbf{x}^*) = (0, ..., 0)$ | |
| (K-T 2) | $\lambda_1^* \geq 0, ..., \lambda_k^* \geq 0$ | $\lambda_1^* [g_1(\mathbf{x}^*) - b_1] = 0, ..., \lambda_k^* [g_k(\mathbf{x}^*) - b_k] = 0$ |

**Proof.** The theorem assumes that $\mathbf{x}^*$ maximizes $f$ on the constraint set

$$U = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \leq b_1, ..., g_k(\mathbf{x}) \leq b_k\};$$

that only $\mathbf{g}_e = (g_1, ..., g_e)$ are binding at $\mathbf{x}^*$

$$g_1(\mathbf{x}^*) = b_1, ..., g_e(\mathbf{x}^*) = b_e, g_{e+1}(\mathbf{x}^*) < b_{e+1}, ..., g_k(\mathbf{x}^*) < b_k;$$

and that the $e \times n$ Jacobian matrix

$$D\mathbf{g}_e(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_e}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial g_e}{\partial x_n}(\mathbf{x}^*) \end{pmatrix}. \tag{11}$$

has maximal rank $e$.

Since the $g_i$'s are continuous functions, there is an open ball $B = B_r(\mathbf{x}^*)$ of radius $r > 0$ about $\mathbf{x}^*$ such that $g_i(\mathbf{x}) < b_i$ for all $\mathbf{x} \in B$ and for $i = e+1, ..., n$. We will work in the open set $B$ for the rest of this proof.

Note that $\mathbf{x}^*$ maximizes $f$ in $B$ on the constraint set

$$g_1(\mathbf{x}) = b_1, ..., g_e(\mathbf{x}) = b_e \tag{12}$$

as if there were another point $\mathbf{x}^{**}$ in $B$ that satisfied (12) and gave a higher value of $f$, then this point would yield a higher value of $f$ on the original constraint set $U$ and this would contradict the definition of $\mathbf{x}^*$. Moreover, by (11), $\mathbf{x}^*$ satisfies the NDCQ for the problem of maximizing $f$ on the constraint set (12). Therefore, by Theorem 1 in Note 7, there exist $\mu_1^*, ..., \mu_e^*$ such that

$$\begin{aligned} &\frac{\partial \widehat{L}}{\partial x_1}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0, ..., \frac{\partial \widehat{L}}{\partial x_n}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0 \\ &g_1(\mathbf{x}) = b_1, ..., g_e(\mathbf{x}) = b_e \end{aligned} \tag{13}$$

where $\widehat{L} \equiv f(\mathbf{x}) - \mu_1[g_1(\mathbf{x}) - b_1] - ... - \mu_e[g_e(\mathbf{x}) - b_e]$.

Now consider the original Lagrangian function

$$L(\mathbf{x}, \boldsymbol{\lambda}) \equiv f(\mathbf{x}) - \lambda_1 [g_1(\mathbf{x}) - b_1] - \ldots - \lambda_k [g_k(\mathbf{x}) - b_k].$$

Let $\lambda_1^* = \mu_1^*$ for $i = 1, ..., e$ and set $\lambda_i^* = 0$ for $i = e+1, ..., k$. Using these values for $\boldsymbol{\lambda}^*$ and noting equation (13), we see that $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ is a solution of the $n+k$ equations in $n+k$ unknowns

$$\begin{aligned} &\frac{\partial L}{\partial x_1}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0, ..., \frac{\partial L}{\partial x_n}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0 \\ &\lambda_1^* [g_1(\mathbf{x}^*) - b_1] = 0, ..., \lambda_e^* [g_e(\mathbf{x}^*) - b_e] = 0 \\ &\lambda_{e+1}^* [g_{e+1}(\mathbf{x}^*) - b_{e+1}] = 0, ..., \lambda_k^* [g_k(\mathbf{x}^*) - b_k] = 0. \end{aligned} \tag{14}$$

Except for the condition that $\lambda_1^* \geq 0, ..., \lambda_k^* \geq 0$ must hold, the proof is complete. To show the latter consider the system of $e$ equations in $n+e$ variables

$$\begin{aligned} g_1(\mathbf{x}) &= b_1 \\ \vdots &= \vdots \\ g_e(\mathbf{x}) &= b_e. \end{aligned} \tag{15}$$

By the rank condition of (11) and the IFT, there exist $e$ coordinates $x_{i_1}, ..., x_{i_e}$ such that we can consider the system (15) as implicitly defining $x_{i_1}, ..., x_{i_e}$ in terms of the rest $x_i$'s and all the $b$'s. In this latter set of exogenous variables, hold $b_2, ..., b_e$ constant, hold the exogenous $x_i$'s constant, and let $b_1$ decrease linearly: $t \to b_1 - t$ for $t \geq 0$. By the IFT, as the exogenous variable $b_1$ varies, we can still solve the system (15) for $x_{i_1}, ..., x_{i_e}$. This means, in particular, that there is a $C^1$ curve $\mathbf{x}(t)$ defined for $t \in [0, e)$ such that $\mathbf{x}(0) = \mathbf{x}^*$ and, for all $t \in [0, e)$,

$$g_1[\mathbf{x}(t)] = b_1 - t \text{ and } g_i[\mathbf{x}(t)] = b_i \text{ for } i = 2, ..., e. \tag{16}$$

Let $\mathbf{v} = \mathbf{x}'(0)$. Applying the Chain Rule to (16), we conclude that

$$Dg_1(\mathbf{x}^*)\mathbf{v} = -1 \text{ and } Dg_i(\mathbf{x}^*)\mathbf{v} = 0 \text{ for } i = 2, ..., e. \tag{17}$$

Since $\mathbf{x}(t)$ lies in the constraint set for all $t$ and $\mathbf{x}^*$ maximizes $f$ in the constraint set, $f$ must be nonincreasing along $\mathbf{x}(t)$. Therefore,

$$\frac{d}{dt} f[\mathbf{x}(t)] \mid_{t=0} = Df(\mathbf{x}^*)\mathbf{v} \leq 0.$$

Let $D_{\mathbf{x}}L(\mathbf{x}^*)$ denote the derivative of the Lagrangian with respect to $\mathbf{x}$. By our first order conditions (14) and (17)

$$\begin{aligned} 0 &= D_{\mathbf{x}}L(\mathbf{x}^*)\,\mathbf{v} \\ &= Df(\mathbf{x}^*)\,\mathbf{v} - \sum\nolimits_i \lambda_i Dg_i(\mathbf{x}^*)\,\mathbf{v} \\ &= Df(\mathbf{x}^*)\,\mathbf{v} - \lambda_1 Dg_1(\mathbf{x}^*)\,\mathbf{v} \\ &= Df(\mathbf{x}^*)\,\mathbf{v} + \lambda_1. \end{aligned}$$

Since $Df(\mathbf{x}^*)\,\mathbf{v} \leq 0$ we conclude that $\lambda_1 \geq 0$. A similar argument shows that $\lambda_i \geq 0$ for $i = 2, ..., e$. This completes the proof. ■

**Example 1.** Let us consider the problem

$$\max_{x_1,x_2} \{f(x_1, x_2) = x_1 x_2 : x_1 + x_2 \leq 100, x_1 \leq 40, x_1 \geq 0, x_2 \geq 0\}.$$

Notice that the last two constraints can be rewritten as $-x_1 \leq 0$ and $-x_2 \leq 0$ respectively. The Kuhn-Tucker conditions are displayed in the table below.

$$\begin{array}{lll} & \text{Marginal Conditions} & \text{Complementary Slackness Conditions} \\ (\text{K-T } 1') & \left\{\begin{array}{l} x_2 - \lambda_1 - \lambda_2 + \lambda_3 = 0 \\ x_1 - \lambda_1 + \lambda_4 = 0 \end{array}\right. & \\ (\text{K-T } 2') & \left\{\begin{array}{ll} \lambda_1 \geq 0 & x_1 + x_2 \leq 100 \\ \lambda_2 \geq 0 & x_1 \leq 40 \\ \lambda_3 \geq 0 & -x_1 \leq 0 \\ \lambda_4 \geq 0 & -x_2 \leq 0 \end{array}\right. & \begin{array}{l} \lambda_1 (x_1 + x_2 - 100) = 0 \\ \lambda_2 (x_1 - 40) = 0 \\ \lambda_3 (-x_1) = 0 \\ \lambda_4 (-x_2) = 0 \end{array} \end{array}$$

There are 16 possibilities to consider! (**Why?**) By simple inspection we can discard some of them. Here it does not make any sense to select $x_1$ or $x_2$ equal to zero, as then $f(x_1, x_2) = 0$. Therefore, by (K-T 2′), we know that $\lambda_3$ and $\lambda_4$ should be both zero at a global maximum. In addition, since $f$ is strictly increasing on $\mathbb{R}^2_+$ the first constraint will be binding, i.e. $x_1 + x_2 = 100$ (**Why?**). By these two arguments the 16 possibilities reduce to only 2! (**Why?**)

$$
\begin{array}{lll}
 & \text{Marginal Conditions} & \text{Complementary Slackness Conditions} \\
(\text{K-T } 1') & \left\{ \begin{array}{l} x_2 - \lambda_1 - \lambda_2 = 0 \\ x_1 - \lambda_1 = 0 \end{array} \right. & \\
(\text{K-T } 2') & \left\{ \begin{array}{ll} \lambda_1 \geq 0 & x_1 + x_2 = 100 \\ \lambda_2 \geq 0 & x_1 \leq 40 \\ \lambda_3 = 0 & -x_1 < 0 \\ \lambda_4 = 0 & -x_2 < 0 \end{array} \right. & \begin{array}{l} \checkmark \\ \lambda_2 (x_1 - 40) = 0 \\ \checkmark \\ \checkmark \end{array}
\end{array}
$$

Let us study the two final possibilities. Let $\lambda_2 = 0$. From (K-T $1'$) we get $x_1 = x_2 = \lambda_1$. Since $-x_1 < 0$ and $-x_2 < 0$, then the candidate is $x_1 = x_2 = \lambda_1 = 50$. This contradicts the fact that $x_1 \leq 40$.

The other possibility is $x_1 = 40$. Since $x_1 + x_2 = 100$, then $x_2 = 60$. From (K-T $1'$) $\lambda_1 = 40$ and $\lambda_2 = 20$. This second case satisfies all the Kuhn-Tucker conditions.

Let us check now the NDCQ. There are only three cases to test, associated to the two possible binding restrictions:

$D\mathbf{g}(\mathbf{x}) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $Dg_1(\mathbf{x}) = \begin{pmatrix} 1 & 1 \end{pmatrix}$ and $Dg_2(\mathbf{x}) = \begin{pmatrix} 1 & 0 \end{pmatrix}$.

Since these three matrices have full rank, then the NDCQ holds for all vector in the feasible set.

We conclude that $x_1 = 40$ and $x_2 = 60$ is our candidate for a global maximizer.

Many problems in economics involve just one inequality constraint. But even in this case, the Kuhn-Tucker conditions can be quite involved. If the objective function is nonstationary, then the required conditions adopt a simpler form. (We will use this result in ECON 501A.)

**Corollary 3.** Let $f$ and $g$ be $C^1$ functions and let $f$ be nonstationary, i.e. $Df(\mathbf{x}) \neq (0, ..., 0)$ for all $\mathbf{x}$. Suppose that $\mathbf{x}^*$ is a local maximum of $f$ on $C$ and that the NDCQ holds at $\mathbf{x}^*$.

Then, there exists a multiplier $\lambda^*$ such that

$$
\begin{array}{lll}
 & \text{Marginal Conditions} & \text{Complementary Slackness Conditions} \\
(\text{K-T } 1') & Df(\mathbf{x}^*) - \lambda^* Dg(\mathbf{x}^*) = (0, ..., 0) & \\
(\text{K-T } 2') & \lambda^* > 0 & \lambda^* [g(\mathbf{x}^*) - b] = 0
\end{array}
$$

The next example applies Corollary 3 to a standard problem in Microeconomics.

**Example 2.** Consider the standard utility maximization problem. Our goal is to maximize a $C^1$ utility function $U(x_1, x_2)$ subject to the budget constraint $p_1x_1 + p_2x_2 \leq m$, where $p_1$ and $p_2$ represent positive unit prices. Assume that for every commodity bundle $(x_1, x_2)$

$$\frac{\partial U}{\partial x_1}(x_1, x_2) > 0 \text{ or } \frac{\partial U}{\partial x_2}(x_1, x_2) > 0.$$

This is a version of the nonstationary assumption. The Lagrangian function is

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

The NDCQ holds at any bundle $(x_1, x_2)$ in the constraint set. (**Why?**) By Corollary 3, the necessary first order conditions are, for some $\lambda^* > 0$,

$$\frac{\partial U}{\partial x_1}(x_1^*, x_2^*) = \lambda^* p_1, \ \frac{\partial U}{\partial x_2}(x_1^*, x_2^*) = \lambda^* p_2 \text{ and } p_1x_1^* + p_2x_2^* - m = 0.$$

In particular, the first order conditions imply that (at the optimal bundle) the marginal rate of substitution $(\partial U/\partial x_1)/(\partial U/\partial x_2)$ must equal the ratio of prices $p_1/p_2$. (**Provide the K-T conditions for local maxima without assuming the nonstationary condition.**) ▲

Let us discuss now the second order sufficient conditions for global maxima. To include the inequality constraints in the sufficient conditions, we proceed as before. Given a solution $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ of the first order conditions, we divide the inequality constraints into binding and not binding at $\mathbf{x}^*$. On the one hand, we treat the binding constraints like equality constraints. On the other hand, the multipliers for the nonbinding constraints must be zero and these constraints drop out from the Lagrangian. The following theorem formalizes these comments.

**Theorem 4. (Sufficient conditions for a strict local maximum)** Let $f, g_1, ..., g_k$ be $C^2$ functions, and assume $\mathbf{x}^*$ is a feasible point satisfying the Kuhn-Tucker conditions for some $\boldsymbol{\lambda}^*$.

For the sake of notation, suppose that $g_1, ..., g_e$ are binding at $\mathbf{x}^*$ and that the other constraints $g_{e+1}, ..., g_k$ are not binding. Suppose that the Hessian of $L$ with respect to $\mathbf{x}$ at $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ is negative definite on the linear constraint set $\{\mathbf{v} : D\mathbf{g}_e(\mathbf{x}^*)\mathbf{v} = 0\}$, that is,

$$\mathbf{v} \neq \mathbf{0} \text{ and } D\mathbf{g}_e(\mathbf{x}^*)\mathbf{v} = 0 \Rightarrow \mathbf{v}^T D_{\mathbf{x}}^2 L(\mathbf{x}^*, \boldsymbol{\lambda}^*)\mathbf{v} < 0.$$

Then, $\mathbf{x}^*$ is a strict local maximum of $f$ on $U$.

## Concavity and Optimal Solutions to Problem (P.K-T)

The next theorem provides sufficient conditions for global maxima. As in Note 7 we assume the objective function is concave and the constraint functions are convex.

**Theorem 5. (Sufficient conditions for a global maximum)** Let $f, -g_1, ..., -g_k$ be $C^1$ concave functions. Let $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ be a pair of vectors that satisfy the necessary Kuhn-Tucker conditions. Then $\mathbf{x}^*$ is a solution to problem (P.K-T).

**Proof.** From the first order conditions

$$Df(\mathbf{x}^*) - \sum_{i=1}^{k} \lambda_i Dg_i(\mathbf{x}^*) = (0, ..., 0). \tag{18}$$

Let $\mathbf{x}$ be an arbitrary point in the constraint set $U$. For each binding constraint $g_i$, $g_i(\mathbf{x}) \leq g_i(\mathbf{x}^*)$. Since $-g_i$ is $C^1$ and concave

$$-Dg_i(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \geq 0. \text{ } (\mathbf{Why?})$$

Since $\lambda_i^* = 0$ for the nonbinding constraints, then

$$-\sum_{i=1}^{k} \lambda Dg_i(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \geq 0.$$

By (18) it follows that

$$Df(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \leq 0. \text{ } (\mathbf{Why?})$$

Since $f$ is $C^1$ and concave

$$f(\mathbf{x}^*) \geq f(\mathbf{x}). \text{ } (\mathbf{Why?})$$

The result follows as $\mathbf{x}$ is an arbitrary point in the constraint set $U$. ∎

The next theorem assumes the objective function is strictly concave. The added benefit of this condition is that now, if there exists a solution to problem (P.K-T) then it is unique.

**Theorem 6. (Uniqueness)** Let $\mathbf{x}^*$ be an optimal solution to problem (P.K-T). If $f$ is strictly concave and $-g_1, ..., -g_k$ are all concave, then $\mathbf{x}^*$ is the only optimal solution to problem (P.K-T).