Lecture Four: Constrained Optimization II - Kuhn Tucker Theory

1 Some Examples

A general problem that arises countless times in economics takes the form:

(Verbally): Choose $x$ to give the highest payoff where $x$ is restricted to some set of alternatives.

(Formally):

$$\max_x f(x)$$

such that

$$x \in C.$$ 

Typically, the constraint set $C$ is restricted so that it can be described mathematically or in a functional form of the type:

$$C = \{ x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, 2, \ldots m, g_i : \mathbb{R}^n \rightarrow \mathbb{R} \}$$

or even more concisely

$$C = \{ x \in \mathbb{R}^n \mid G(x) \geq 0, G : \mathbb{R}^n \rightarrow \mathbb{R}^m \}.$$ 

Example: With linear constraints, the set of points which satisfy the inequalities simultaneously

$$3x_1 + 2x_2 - 5x_3 \geq 0$$

$$2x_1 + 5x_2 + 6 \geq 0$$

$$x_1 + x_3 \geq 0$$

describes a set. Note that from our earlier discussions, each of these inequalities defines a halfspace. Therefore each set of points satisfying each inequality is a convex set, and what the three together yield is the intersection of three half-spaces.

Question: How can we incorporate equality constraints in this format? Consider the two inequalities:
The only points which satisfy both must satisfy

\[ 3x_1 + 2x_2 - 5x_3 = 0 \]

**Technical Note:** This is a little too easy in some sense. A formulation like this will typically violate a Constraint Qualification.

**Question:** What if our constraints are of the form, \( G(x) \leq 0 \)? Then define a new function, \( H(x) = -G(x) \) and write the constraint as \( H(x) \geq 0 \). Notice that we can also incorporate non-negativity constraints simply by writing some of the \( g_i \)'s as \( g_i(x) = x_i \geq 0 \).

### 1.1 A general typology of Constrained maximization problems

**Note:** I will call these constrained maximization problems. However, this is without loss of generality since our Problem Set 2 showed that any MINIMIZATION problem

\[
\min_x f(x), \text{ such that } x \in C.
\]

can always be formulated as a MAXIMIZATION problem,

\[
\max_x -f(x), \text{ such that } x \in C
\]

Certain types of optimization problems arise frequently. Here is a partial list:

#### 1.1.1 Unconstrained Maximization

These are the simplest of course. In this case, the constraint set \( C \) is just the whole vector space that \( x \) is assumed to lie in. (For us, this will generally be \( \mathbb{R}^n \))

#### 1.1.2 Lagrange Maximization Problems

In these cases, the constraint set \( C \) is defined solely by equality constraints:

\[ C = \{ x \in \mathbb{R}^n \mid G(x) = 0, G: \mathbb{R}^n \to \mathbb{R}^m \} . \]
1.1.3 Linear Programming Problems

In these cases, \( g_i(x) = a'_i x - b_i \) for all \( i = 1, 2, \ldots, m \) and the constraint is an inequality constraint. Also, \( f(x) = c'x \).

1.1.4 Kuhn-Tucker Problems

These problems take a class of constraints (non-negativity constraints) away from the constraint set \( C \) and describes them directly

\[
\max_{x \in \mathbb{R}^n} f(x)
\]

such that

\[
G(x) \geq 0, \quad G : \mathbb{R}^n \to \mathbb{R}^m
\]

Sometimes equality constraints are also allowed but I am not going to focus on that. This general framework with general \( f \) and \( G \) is often referred to as NONLINEAR PROGRAMMING. We have already looked at unconstrained problems. We next examine the Lagrange case.

2 Lagrange Problems

This section is really mostly useful as a context or background for the K-T problems we look at later. The canonical Lagrange problem is of the form

\[
\max_{x \in \mathbb{R}^n} f(x), \quad G(x) = 0, \quad G : \mathbb{R}^n \to \mathbb{R}^m
\]  

That is, the constraints are all equality constraints. In many economic problems, even though this is not the formally correct way to model the problem, we may know enough about the problem to assume that all the (inequality) constraints will bind anyway and so any solution will be as if it was a solution to this type. (For example, consumer optimization with utility functions which are always increasing in every good.)

**Theorem 1. Lagrange** In the constrained maximization problem (1) , suppose that \( f \) and \( G \) are \( C^1 \) and suppose that the \( n \) by \( m \) matrix \( \nabla G(x^*) \) has rank \( m \). Then if \( x^* \) solves (1), there exists a vector, \( \lambda^* \in \mathbb{R}^m \), such that

\[
\nabla f(x^*) + \nabla G(x^*) \lambda^* = 0
\]

Alternatively, we can write this as

\[
\nabla f(x^*) + \sum_{j=0}^{m} \nabla g_j(x^*) \lambda_j^* = 0
\]
where $g_j$ is the $j^{th}$ component of the $m$-vector function, $G$.

Since we can also write the first order necessary conditions from the theorem as $\nabla f(x^*) = -\sum_{j=1}^{m} \nabla g_j(x^*) \lambda_j^*$, this says that we can express the gradient of the objective function as a linear combination of the gradients of the constraint functions. The weights of the linear combination are determined by $\lambda^*$. Note, also that no claims are made about the SIGN of $\lambda^*$. Also, the rank $m$ condition in the Theorem is a version of what is called a CONSTRAINT QUALIFICATION. I will come back to this in the K-T part. Notice that as the theorem is stated, we can only count on the existence of Lagrange Multipliers if the CQ is satisfied. Otherwise, all bets are off.
3  Kuhn-Tucker and Differentiability

The most common form of a constrained optimization problem that we encounter in economics takes the form

$$\max_x f(x)$$

such that

$$G(x) \geq 0, x \geq 0.$$ 

That is, the equality constraints that characterize the classical Lagrange problem, are inequality constraints in this problem, and there are often non-negativity constraints as well. In what follows, I will refer to problems of this sort as Kuhn-Tucker problems or K-T problems after Harold Kuhn and Robert Tucker who developed the theory.

3.1 Example 1: Utility Maximization

A consumer with preferences over the consumption of $n$ goods, represented by a utility function, $U : \mathbb{R}^n_+ \to \mathbb{R}$, with income $I$ and facing (fixed) prices, $p$, solves the problem

$$\max_{x \in \mathbb{R}^n} U(x)$$

subject to

$$I - p'x \geq 0,$$

$$x_i \geq 0, \forall i = 1, 2, ..., n$$

Here, $G(x) \equiv I - p'x$.

3.2 Example 2: Cost Minimization

A manager with a technology represented by a production function, $f : \mathbb{R}^n \to \mathbb{R}^m$, which describes how a vector of inputs $x \in \mathbb{R}^n$ can be used to produce a vector of outputs $y \in \mathbb{R}^m$ and facing (fixed) input prices, $w$, is instructed to produce a given vector of output, $Y$, at minimum cost, solves the problem:

$$\min_x w'x, \text{ such that } f(x) \geq Y, x \geq 0$$

Here, $G(x) \equiv f(x) - Y \in \mathbb{R}^m$. 
3.3 Example 3: Nonlinear Least-Squares Estimation

An econometrician observes output decisions by \( l \) firms and also observes input decisions by the same firms. (Each input decision is an \( m \) dimensional vector.) She knows that the underlying technology in the industry is (stochastically) Cobb-Douglas, that is, it is given by \( y_i = f(x_i; \beta) + \epsilon_1 \), for firm \( i \), where

\[
f(x_i, \beta) = x_{i1}^{\beta_1} \times x_{i1}^{\beta_1} \times x_{i1}^{\beta_1} \times \ldots \times x_{i1}^{\beta_m}
\]

and

\[
\text{and } \sum_{j=1}^{m} \beta_j \leq 1
\]

That is, the function exhibits decreasing returns to scale. Her problem is to find the best estimate of the parameter of the production function, \( \beta \). One way to determine best, is the value of \( \beta \) that minimizes the sum of squared differences:

\[
\max_{\beta} - \sum_{i=1}^{l} [y_i - f(x_i; \beta)]^2
\]

subject to

\[
\sum_{j=1}^{m} \beta_j \leq 1
\]

Note that here the constraint \( G(\beta) \) can be written as

\[
G(\beta) \equiv 1 - \sum_{j=1}^{m} \beta_j \geq 0
\]

3.4 The Lagrangian

The tool we use to solve all of these problems, and many others, is called a LAGRANGIAN. When \( f : \mathbb{R}_+^n \to \mathbb{R} \) and \( G : \mathbb{R}_+^n \to \mathbb{R}^m \), the Lagrangian of the Kuhn-Tucker problem is a new function

\( L : \mathbb{R}_+^{n+m} \to \mathbb{R} \). It is constructed as:

\[
L(x, \lambda) = f(x) + \lambda G(x)
\]

Notice that a new, \( m \) vector, \( \lambda \), is introduced in this definition. It is referred to as a Lagrange Multiplier. Observe the important limit on the domain space of \( L \). The Lagrange multiplier here is non-negative. This is an important difference between the Kuhn-Tucker problem and the Lagrange Problem we looked at earlier.
Of course, the Lagrangian can also be written as

\[ L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) , \]

where \( G(x) = (g_1(x), g_2(x), ..., g_m(x))' \).

In unconstrained problems, we already know that there is a relationship between stationary points, (points where the Jacobian is zero) of a function and its optima. How does this extend to constrained problems? We can get an idea from the following graph which concerns the problem

\[
\max_{x} f(x) \\
\text{s.t.}  \\
10 - x \geq 0 \text{ and } x \geq 0.
\]

By eye we can see that A, C, E, G are local and possibly global maxima and B, D, F are local minima. B, C, D, E, F are ALL stationary points and not all global maxima of course. We already have seen the problems involved in distinguishing maximums from minimums and local from global. What is new is that A and G are not stationary points in the usual sense, but they can still be global maxima (within the constraint set.)

The Kuhn Tucker (Karush) theory has been developed to deal with this. Consider our standard Kuhn-Tucker problem with the assumption that that \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) and \( G : \mathbb{R}^n \rightarrow \mathbb{R}^m \) are \( C^1 \).
For any point, $x^*$, define the indices of $G$ where the constraint exactly binds, and the indices of $x$ where the non-negativity constraint DOES NOT bind. That is,

$$K = \{ i \mid g_i(x^*) = 0 \} \text{ and } M = \{ i \mid x_i^* > 0 \}$$

Now construct a sort of Jacobian of the constraint set by:

1. Looking only at the binding constraints and
2. Only differentiating w.r.t to the $x_j$'s which are strictly positive.

That is,

$$H(x^*) = \nabla_M G_K(x)$$

where $\nabla_M$ represents partial differentiation with respect only to the components, $j \in M$ and $G_K$ is $G$ with only the $K$ components.

**Theorem 2. (Karush, Kuhn-Tucker):** Suppose that $x^*$ solves K-T as a local maximum and suppose that $H(x^*)$ has maximal rank (the Constraint Qualification). Then there exists $\lambda^* \geq 0$ such that

$$\frac{\partial L(x^*, \lambda^*)}{\partial x_i} \leq 0$$

and

$$\frac{\partial L(x^*, \lambda^*)}{\partial x_i} x_i^* = 0$$

for all $i = 1, 2, \ldots, n$ and

$$\frac{\partial L(x^*, \lambda^*)}{\partial \lambda_j^*} = g_j(x^*) \geq 0$$

and

$$\frac{\partial L(x^*, \lambda^*)}{\partial \lambda_j^*} \lambda_j^* = g_j(x^*) \lambda_j^* = 0$$

for all $j = 1, 2, \ldots, m$.

The second and fourth equations are referred to as Complementary Slackness conditions. In words, it says that either the inequality holds as an equality or else the corresponding variable, ($x_i^*$ or $\lambda_j^*$) is zero.

Note that the weird condition on the $H(x^*)$ is yet another example of a Constraint Qualification. It is not very pretty and typically ignored in applications, but to be correct we need to include it. Kuhn-Tucker have a more peculiar but weaker Constraint Qualification. The way the theorem is
stated shows that if the Constraint Qualification is NOT satisfied, then all bets are off. We cannot
conclude anything about First Order Necessary Conditions (FONCs). Also, notice that the theorem
says that satisfying the Karush-Kuhn-Tucker first order conditions is a NECESSARY CONDITION
for an optimum. Therefore, we would be taking a BIG risk assuming that just because the KKT
conditions are satisfied we have solved the problem.

### 3.5 Example

Consider the diagrammatic example above. In that case, the optimization problem can be stated as

\[
\max_x f(x) \text{ s.t } 10 - x \geq 0, \ x \geq 0
\]

The lagrangian is

\[
L(x, \lambda) = f(x) + \lambda(10 - x)
\]

and the first order conditions are

\[
\begin{align*}
\frac{\partial L(x^*, \lambda^*)}{\partial x} &= f'(x^*) - \lambda^* \leq 0, \ \left[f'(x^*) - \lambda^* \right] x^* = 0 \\
\frac{\partial L(x^*, \lambda^*)}{\partial \lambda} &= 10 - x^* \geq 0, \ [10 - x^*]\lambda^* = 0
\end{align*}
\]

We break these conditions down into three cases:

1. **Strict interior.** \(x^* > 0, 10 - x^* > 0\). At a point of this type, since \(10 - x^* > 0\), we must have \(\lambda^* = 0\), otherwise, \((10 - x^*)\lambda^* = 0\) could not hold. Since \(x^* > 0\), the first FONC must hold with equality, otherwise, \((f'(x^*) - \lambda^*)x^* = 0\) could not hold. Combining gives us that in cases like this, we are like an unconstrained optimum. That is, \(f'(x^*) = 0\). Note that points B,C,D,E,F all satisfy these conditions. Thus we are not ruling out minima with this test. To rule those out, we need to do the second order test. Also, we are not ruling out local maxima.

2. **Left boundary:** \(x^* = 0\). At this point, \(10 - x^* > 0\), so again, we must have \(\lambda^* = 0\) and the first condition becomes, \(f'(x^*) \leq 0, x^* = 0\). Thus all we can do is check at \(x^* = 0\) to see if \(f'(0) \leq 0\).

3. **Right boundary:** \(10 - x^* = 0\). Obviously, \(x^* > 0\). We cannot say anything about \(\lambda^*\), but we do know that we must have \(f'(x^*) = \lambda^* \geq 0\) so at \(x^* = 10\) we need to have the slope of the objective function be non-negative. So we have to check that \(f'(10) \geq 0\). which it is.
Notice that as we can see in the graph, we still need to worry about

1. Multiple solutions (A,C,G)

2. Local vs. Global maxima (E vs A,C,G)

3. Min versus max (B,D,F vs A,C,G)

As in the case of unconstrained optimization, if we put more structure on the problem, we can resolve some of these issues:

**Theorem 3.** Suppose \( G(x) \) is concave, and \( f(x) \) is strictly quasi-concave, then if \( x^* \) solves KT, \( x^* \) is unique. Furthermore, if \( \{ x : G(x) \geq 0, x \geq 0 \} \) is compact and non-empty and \( f(x) \) is continuous, then there exists \( x^* \) which solves KT.

**Proof.** The second part of the theorem can be proved using Weierstrass Theorem which states that a continuous function achieves its maximum on any non-empty, compact set. We prove the first part by contradiction. Suppose, there exists two maximizers \( x \neq x' \) s.t \( f(x) = f(x') \), \( G(x) \geq 0, G(x') \geq 0 \). Since \( G \) is concave, for \( t \in [0,1] \):

\[
G(tx + (1-t)x') \geq gG(x) + (1-t)G(x') \geq 0
\]

Thus, \( tx + (1-t)x' \) is feasible. Since \( f \) is strictly quasi-concave

\[
f(tx + (1-t)x') > \min\{f(x), f(x')\} = f(x)
\]

which is a contradiction. \( \square \)

Notice that if the global concavity and quasiconcavity conditions do not hold, we still need to check local second order conditions to rule out local minima. There exist ways of doing this by looking at the leading principal minors of the matrix

\[
H = \begin{pmatrix}
\frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \cdots & \nabla G \\
\frac{\partial^2 L}{\partial x_1 \partial x_2} & \ddots & \ddots & \\
\vdots & \ddots & \ddots & \\
\nabla G & 0 & & \\
\end{pmatrix}
\]

which is an \((n + m)\) squared matrix. Intuitively, we require the Hessian to be NSD for any vectors along the constraint. See SB pp 457-468.
3.6 The Constraint Qualification

Consider the problem,

\[
\max_x x_1 \text{ s.t. } (1 - x_1)^3 - x_2 \geq 0, x \geq 0.
\]

The picture illustrates the constraint set, so it is apparent that \((1, 0)\) is the solution to this problem. However, at this point \(x_2 \geq 0\), is a binding constraint. Notice that \(\nabla g(1, 0) = (2(x_1 - 1), -1) = (0, -1)\) at the solution. But this gives us a matrix of binding constraints of

\[
\begin{bmatrix}
0 & 1 \\
0 & -1
\end{bmatrix}
\]

which has rank 1 a violation of the rank condition. The gradient of \(f(1, 0)\) is \((1,0)\) and there is no way to express \((1,0)\) as a linear combination of \((0,1)\) or \((0,-1)\). Thus, here is a simple example of a constrained optimization problem where there do not exist Lagrange multipliers that satisfy the KT FONC at the optimum.
3.7 Complementary Slackness

Recall that at an optimum of a constrained optimization problem, we have a collection of \( n \) inequalities of the form

\[
\frac{\partial f(x)}{\partial x_1} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_1} \leq 0
\]

\[
\frac{\partial f(x)}{\partial x_2} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_2} \leq 0
\]

\[\vdots\]

\[
\frac{\partial f(x)}{\partial x_n} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_n} \leq 0
\]

plus one equality of the form

\[
\sum_{i=1}^{n} \left( \frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_i} \right) x_i = 0
\]

and then a collection of \( m \) inequalities of the form,

\[
g_1(x) \geq 0
\]

\[
g_2(x) \geq 0
\]

\[\vdots\]

\[
g_m(x) \geq 0
\]

plus one equality of the form

\[
\sum_{j=1}^{m} \lambda_j g_j(x) = 0
\]

This is a system of \( n + m + 2 \) conditions which we use to find values for \( n \) values of the \( x \) variables and \( m \) values of the Lagrange multiplier variables, \( \lambda \). We use these together to draw inferences about the nature of the solution.

Focus on the first conditions and consider (say) the \( i^{th} \) partial in conjunction with the final equality. Suppose that it holds with strict inequality

\[
\frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_i} < 0
\]

Also,

\[
\left( \frac{\partial f(x)}{\partial x_k} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_k} \right) x_k = 0
\]
for all \( k \neq i \). If in addition \( x_i > 0 \),

\[
\left( \frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_i} \right) x_i < 0
\]

which implies that the equality

\[
\sum_{i=1}^{n} \left( \frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_i} \right) x_i = 0
\]

cannot hold. This argument implies that we cannot have both \( x_i > 0 \) and

\[
\frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_i} < 0
\]

at the same time. Another way to say this is if

\[
\frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_i} < 0
\]

then the non-negativity constraint on \( x_i \) must BIND. If the non-negativity constraint on \( x_i \) is SLACK, then we must have

\[
\frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_i} = 0
\]

A similar argument holds for the final \( m \) inequalities and equality. In this case, we have, if \( g_j(x) > 0 \), (that is, if the \( j^{th} \) constraint is SLACK) then the corresponding Lagrange multiplier, \( \lambda_j = 0 \). If the lagrange multiplier on the \( j^{th} \) constraint \( \lambda_j > 0 \), then the \( j^{th} \) constraint BINDS, \( g_j(x) = 0 \).

### 3.8 Saddle-Value Theorem and Linear Programming Theorem. (SB 21.5)

Note that the Lagrangian is defined on a “bigger space” then \( f \). \( f : \mathbb{R}^n_+ \to \mathbb{R} \) while \( f : \mathbb{R}^n_+ \times \mathbb{R}^m_+ \to \mathbb{R} \)

**Definition 1.** A **saddle point** of the Lagrangian \( L \) is an element in \( \mathbb{R}^n_+ \times \mathbb{R}^m_+ \), say \((x^*, \lambda^*)\), such that

\[
L(x, \lambda^*) \leq L(x^*, \lambda^*) \leq L(x^*, \lambda)
\]

for all \( x \in \mathbb{R}^n_+ \) and all \( \lambda \in \mathbb{R}^m_+ \).

Notice that the definition incorporates two optimization problems simultaneously. Holding \( x^* \) fixed, we minimize with respect to \( \lambda \) and holding \( \lambda^* \) fixed, we maximize with respect to \( x \).

We now have a very powerful sufficiency result:
Theorem 4. Suppose that \((x^*, \lambda^*)\) is a Saddle Point of \(L(x, \lambda)\) on \(\mathbb{R}_+^n \times \mathbb{R}_+^m\). Then \(x^*\) solves the K-T problem.

Proof. Using the definition of a saddle point:

\[
f(x^*) + \lambda'G(x^*) \geq f(x^*) + \lambda^*G(x^*) \geq f(x) + \lambda^*G(x)
\]

for all non-negative \(x\) and \(\lambda\). We can rewrite the first inequality as

\[
(\lambda - \lambda^*)'G(x^*) \geq 0
\]

This must be true for all nonnegative \(\lambda\). Suppose that there is some strictly negative component of \(G(x^*)\), call it \(g_j(x^*) < 0\) If we select \(\lambda = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_j^* + 1, \ldots, \lambda_m^*)'\) and plug that in for \(\lambda\), we will get for the left side a value of \(g_j(x^*) < 0\) which violates the inequality, so we conclude that \(g_j(x^*) \geq 0\) for all \(j\). This implies that \(x^*\) is feasible for the K-T problem. Now consider what happens when we plug \(\lambda = (0, 0, \ldots, 0)\) in the equation above. This gives

\[
-\lambda^*G(x^*) \geq 0
\]

Since \(g_j(x^*) \geq 0\) and \(\lambda_j^* \geq 0\) for all \(j\), \(\lambda_j^* g_j(x^*) \geq 0\). This implies that

\[
\lambda^*G(x^*) = \sum_{j=1}^{m} \lambda_j^* g_j(x^*) \geq 0
\]

and these two inequalities imply that

\[
\lambda^*G(x^*) = \sum_{j=1}^{m} \lambda_j^* g_j(x^*) = 0
\]

i.e., either \(g_j(x^*) = 0\) or \(\lambda_j^* = 0\) or both. Using the fact that \(\lambda^*G(x^*) = 0\) in the second inequality in the definition of a saddle point yields:

\[
f(x^*) \geq f(x) + \lambda^*f(x), \ \forall x \in \mathbb{R}_+^n
\]

Restrict attention only to the \(x\)'s that lie in the feasible set for the K-T problem. That is, only \(x\)'s such that \(G(x) \geq 0\). this means that the summation term is nonnegative, which implies that

\[
f(x^*) \geq f(x), \ \forall x \text{ s.t. } G(x) \geq 0
\]

i.e., \(x^*\) solves the Kuhn-Tucker problem. 

\qed
The theorem says that a sufficient condition for \( x^* \) to be a solution to the K-T problem is that it form part of a solution to the Saddle Value problem of the Lagrangian. What about the other direction? That is, is it both necessary and sufficient?

Recall that in the Lagrange theorem we needed a condition called the Constraint Qualification. There are a variety of such CQs. The next one is a very simple one:

**Definition 2. Slaters Condition or Strong Consistency (SC):** The constraint set of the K-T problem satisfies SC if there exists an \( x > 0 \) such that \( G(x) > 0 \).

Note that this just requires the interior of the constraint set be non-empty.

**Theorem 5.** Suppose that \( f \) is quasi-concave and \( G \) is made up of \( m \) concave functions. If \( G \) also satisfies SC the \( x^* \) solves K-T implies there exists a \( \lambda^* \geq 0 \) such that \((x^*, \lambda^*)\) forms a Saddle Point of the Lagrangian.

Observe that extra conditions are needed to get the reverse direction.

The saddle value problems in the above section have had very powerful applications in economics. In particular, it can be shown that a general equilibrium of an economy can be expressed as a solution to a saddle value problem (it turns out the the dual to maximizing the “income” of an economy has as its dual the problem of minimizing expenditures on inputs). An attraction of the approach is that we do not need to assume anything about differentiability. However, finding saddle points can be very difficult and often impossible. If we make use of differentiability, then we can often make the problem of explicitly solving constrained optimization problems easier.