Karush-Kuhn-Tucker (KKT) Conditions

In this section, we will extend the Lagrangian technique to deal with a more general class of problems: NLPs with \( \leq \) constraints. Problems of the following form:

\[
\begin{align*}
\text{max (or min)} & \quad z = f(x_1, \cdots, x_n) \\
\text{subject to} & \quad g_i(x_1, \cdots, x_n) \leq b_i, \quad 1 \leq i \leq m,
\end{align*}
\]

To apply the K-T conditions, all constraints must be of \( \leq \) type. Of course, it is always possible to express all constraints as \( \leq \) constraints. If the constraints are of \( \geq \) type, multiply them with \(-1\). A constraint of the \( = \) type (i.e. \( g(X) = b \)) must be replaced with two constraints of the \( \leq \) type (i.e. \(-g(X) \leq -b \) and \( g(X) \leq b \)).

**Theorem 9**

Suppose NLP (1) is a maximization problem. If \( x^* = (x_1^*, x_2^*, \ldots, x_n^*) \) is a solution, then it must satisfy the constraints in the problem and there must exist multipliers \( \lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*) \) satisfying the following:

\[
\begin{align*}
\frac{\partial f(\bar{x})}{\partial x_j} & - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\bar{x})}{\partial x_j} = 0, \quad 1 \leq j \leq n; \\
\lambda_i [b_i - g_i(\bar{x})] & = 0, \quad 1 \leq i \leq m; \\
\lambda_i & \geq 0, \quad 1 \leq i \leq m.
\end{align*}
\]

**Theorem 11**

For a maximization problem, if \( f(x_1, x_2, \ldots, x_n) \) is a concave function, and \( g_1(x), \ldots, g_m(x) \) are convex functions, any point \( x^* \) satisfying the above Kuhn-Tucker conditions is an optimal solution to the problem.

**Theorem 9’**

Similarly if NLP (1) is a minimization problem and \( x^* = (x_1^*, x_2^*, \ldots, x_n^*) \) is a solution, then it must satisfy the constraints of the problem and there must exist multipliers \( \lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*) \) satisfying the following:

\[
\begin{align*}
\frac{\partial f(\bar{x})}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\bar{x})}{\partial x_j} & = 0, \quad 1 \leq j \leq n; \\
\lambda_i [b_i - g_i(\bar{x})] & = 0, \quad 1 \leq i \leq m; \\
\lambda_i & \geq 0, \quad 1 \leq i \leq m.
\end{align*}
\]
Theorem 11’

For a minimization problem, if \( f(x_1, x_2, \ldots, x_n) \) is a convex function, and \( g_1(x), \ldots, g_m(x) \) are convex functions, any point \( x^* \) satisfying the above Kuhn-Tucker conditions is an optimal solution to the problem.

The easiest way to think about the K-K-T conditions is to think of the following situation. Assume that we are interested in minimizing \( f(x) = -x^2 \), subject to some constraints on \( x \).

Now, let’s say that we add in an arbitrary constraint on \( x \). For example, let’s add in a constraint stating that \( x \leq -9 \).

We can see from the graph that the solution to the problem is obviously going to be \( x = -9 \) and that the constraint will be active at the solution, as is the case in an LP.
Let’s say we change our constraint to be \( x \leq 10 \).

![Graph of \( f(x) = -x^2 \)]

In this case, we can see that the solution is obviously going to be 0. In this case, the constraint will not be active at the solution. (This differs from LP.)

So, let’s think about K-K-T conditions as being related to duality in an NLP. Essentially, what we’re saying is that at the solution either the constraint is active, in which case it has a shadow price \( \lambda > 0 \) or it is not active, in which case it has a shadow price of 0.

**Example 1a**

Maximise \(-x^2\)

subject to \( x \leq -9 \)

The K-K-T conditions indicate that the solution to the problem solves the original problem, plus:

\[
\begin{align*}
-2x - \lambda &= 0 \quad (1) \\
\lambda(-9 - x) &= 0 \quad (2) \\
\lambda &\geq 0 \quad (3)
\end{align*}
\]

Let’s try \( \lambda = 0 \), indicating that the constraint is not active the solution.

From (1) we get that: \( x = 0 \)

Unfortunately, this contradicts the original constraint, so our assumption must be incorrect.

Let’s try \( \lambda > 0 \), indicating that the constraint is active at the solution.

From (2) we get that: \( x = -9 \)

From (1) we get that: \( \lambda = 18 \)

Note that this solution is reasonable and it satisfies the original constraints.
Example 1b

Maximise \(-x^2\)

subject to \(x \leq 10\)

The KKT conditions indicate that the solution to the problem solves the original problem, plus:

\[-2x - \lambda = 0\]  \hspace{1cm} (1)
\[\lambda(10 - x) = 0\]  \hspace{1cm} (2)
\[\lambda \geq 0\]  \hspace{1cm} (3)

Let’s try \(\lambda = 0\), indicating that the constraint is not active the solution.

From (1) we get that:
\[x = 0\]

Note that this solution is reasonable and it satisfies the original constraints.

For argument’s sake let’s try \(\lambda > 0\).

From (2) we get \(x = 10\).
From (1) we get \(\lambda = -20\).

Note that this violates constraint (3), so we see that \(\lambda\) must = 0.
Example 2

Maximize \(5x_1 + 4x_2\)

subject to \(x_1 + x_2 \leq 10\), \(x_1 \leq 8\)

The K-K-T conditions for optimality are the original constraints plus:

For \(x_1\):
\[
\frac{df}{dx_1} - \lambda_1 \frac{dg_1}{dx_1} - \lambda_2 \frac{dg_2}{dx_1} = 0 \quad 5 - \lambda_1 - \lambda_2 = 0 \quad (1)
\]

For \(x_2\):
\[
\frac{df}{dx_2} - \lambda_1 \frac{dg_1}{dx_2} - \lambda_2 \frac{dg_2}{dx_2} = 0 \quad 4 - \lambda_1 = 0 \quad (2)
\]

Const 1: \(\lambda_1 [b_1 - g_1(x)] = 0\) \(\lambda_1 (10 - x_1 - x_2) = 0 \quad (3)\)

Const 2: \(\lambda_2 [b_2 - g_2(x)] = 0\) \(\lambda_2 [8 - x_1] = 0 \quad (4)\)

\(\lambda_1 \geq 0 \quad (5)\)

\(\lambda_2 \geq 0 \quad (6)\)

Note

We need to test for conditions where some or all of the \(\lambda\)'s = 0.

If \(\lambda_1 = 0\) condition (2) indicates that 4 = 0. Therefore \(\lambda_1 > 0\).

If \(\lambda_2 = 0\) condition (2) indicates that \(\lambda_1 = 4\), but condition (1) indicates \(\lambda_1 = 5\). Thus, \(\lambda_2 > 0\).

So, solving the equations we find \(\lambda_1 = 4, \lambda_2 = 1\). Thus, \(x_1 = 8\) and \(x_2 = 2\).

Since the objective function is concave and the constraints are convex, we know that we have found a global maximum \((8, 2)\).

Note

Of course, we could also have solved this example as a linear programming problem.
Non-Negativity Constraints

OK, so what do we do with NLPs of the form:

$$\max \ (\text{or min}) \quad z = f(x_1, \ldots, x_n)$$

subject to \quad $$g_i(x_1, \ldots, x_n) \leq b_i \quad 1 \leq i \leq m$$

$$x_j \geq 0 \quad 1 \leq j \leq n$$

(12.8.2)

The rules for applying the K-K-T conditions differ slightly (sort of):

**Theorem 10**

Suppose 12.8.2 is a maximization problem. If $$x^* = (x_1^*, x_2^*, \ldots, x_n^*)$$ is an optimal solution to the problem, then $$x^* = (x_1^*, x_2^*, \ldots, x_n^*)$$ must satisfy the constraints in 12.8.2 and there must exist multipliers $$\lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*)$$ satisfying the following:

$$\lambda_i [b_i - g_i(\bar{x})] = 0 \quad 1 \leq i \leq m$$

$$\left[ \frac{\partial f(\bar{x})}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\bar{x})}{\partial x_j} \right] \bar{x}_j = 0 \quad 1 \leq j \leq n$$

$$\lambda_i \geq 0 \quad 1 \leq i \leq m$$

**Theorem 10'**

10’ extends 10 for minimization cases. Suppose 12.8.2 is a minimization problem. If $$x^* = (x_1^*, x_2^*, \ldots, x_n^*)$$ is an optimal solution to the problem, then $$x^* = (x_1^*, x_2^*, \ldots, x_n^*)$$ must satisfy the constraints in 12.8.2 and there must exist multipliers $$\lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*)$$ satisfying the following:

$$\lambda_i [b_i - g_i(\bar{x})] = 0 \quad 1 \leq i \leq m$$

$$\left[ \frac{\partial f(\bar{x})}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\bar{x})}{\partial x_j} \right] \bar{x}_j = 0 \quad 1 \leq j \leq n$$

$$\lambda_i \geq 0 \quad 1 \leq i \leq m$$
Again, let’s start with a simple case to illustrate this idea. Consider a slightly revised version of our earlier problem.

Maximise \(- (x + 2)^2\)

subject to \[x \leq 10\]
\[x \geq 0\]

Now, essentially, we have the same sort of situation going on here as before. The constraint may be active (in which case \(\lambda\) will be > 0) or it may be inactive (in which case \(\lambda = 0\)). However, we also need to consider the case of \(x\). \(x\) of course could be 0 (implying that the non-negativity constraint is active and has a shadow price \(\mu > 0\)) or it could be > 0 in which case, the non-negativity constraint is not active and has a shadow price \(\mu = 0\).

Let’s try this out in an example:

Maximise \(- (x + 2)^2\)

subject to \[x \leq 10\]
\[x \geq 0\]

The \(K-K-T\) conditions for this problem are the original constraints plus:

\[-2(x + 2) - \lambda \leq 0\]  \hspace{1cm} (1)
\[\lambda (10 - x) = 0\]  \hspace{1cm} (2)
\[-2(x+2) - \lambda x = 0\]  \hspace{1cm} (3)
\[\lambda \geq 0\]  \hspace{1cm} (4)

Let’s start out by trying the cases where \(\lambda > 0\).

\textit{Case 1a}

Assume \(x = 0\)

From (1) \[-2(0 + 2) - \lambda \leq 0\]
\[-4 - \lambda \leq 0\]
\[-4 \leq \lambda\]

From (2) \[x = 10\]

Clearly this violates our constraints on both \(\lambda\) and \(x\), so we can reject this case.
Case 1b
Assume $x > 0$

From (2)  
\[ x = 10 \]

From (3)  
\[ -2(10+2) - \lambda = 0 \]
\[ \lambda = -24 \]

This violates our constraint on $\lambda$, so we can reject this case as well.

Moving on to $\lambda = 0$.

Case 2a  
$x = 0$

\[ -2(0 + 2) - 0 \leq 0 \quad \text{Holds} \]
\[ 0(10 - 0) = 0 \quad \text{Holds} \]
\[ (-2(0 + 2) - 0)0 = 0 \quad \text{Holds} \]
\[ 0 = 0 \quad \text{Holds} \]

Ah, we have found a solution to the problem at $x = 0, \lambda = 0$

Just for fun, let’s consider the final case (though we really don’t have to…)

Case 2b  
$x > 0$

From (3)  
\[ -2(x + 2) = 0 \]
\[ -2x - 4 = 0 \]
\[ x = -2 \]

This violates our assumption of $x > 0$, therefore, we can reject case 2b.
Example

Use the $K$-$K$-$T$ conditions to find the optimal solution to the following NLP:

Minimise $z = (x_1 - 3)^2 + (x_2 - 5)^2$
subject to
$x_1 + x_2 \leq 7$
$x_1, x_2 \geq 0$

Ok, first converting this problem to standard format:

Minimise $z = (x_1 - 3)^2 + (x_2 - 5)^2$
subject to
$x_1 + x_2 \leq 7$

The $K$-$K$-$T$ conditions for optimality are the original constraints plus:

\[
\begin{align*}
2(x_1 - 3) + \lambda_1 &\geq 0 \\
2(x_2 - 5) + \lambda_1 &\geq 0 \\
\lambda_1[7 - x_1 - x_2] &= 0 \\
(2(x_1 - 3) + \lambda_1)x_1 &= 0 \\
(2(x_2 - 5) + \lambda_1)x_2 &= 0
\end{align*}
\]

Try $\lambda_1 = 0$, Assume $x_1, x_2 > 0$

From (1) $2(x_1 - 3) = 0$; $x_1 = 3$
From (2) $2(x_2 - 5) = 0$; $x_2 = 5$
From (3) $0(7 - 3 - 5) = 0$ holds
From (4) $(2(3 - 3) + 0)(3) = 0$ holds

Note, however, that the constraint in the original problem $x_1 + x_2 \leq 7$ does not hold.

Therefore $\lambda_1$ cannot = 0.

Try $\lambda_1 > 0$ (i.e. constraint is active.). Assume $x_1, x_2 > 0$

From (1) $x_1 = 3 - \lambda_1/2$
From (2) $x_2 = 5 - \lambda_1/2$

From (3) $\lambda_1(7 - (3 - \lambda_1/2) - (5 - \lambda_1/2)) = 0$
Thus, $(7 - (3 - \lambda_1/2) - (5 - \lambda_1/2)) = 0$
Solving for $\lambda_1$ we get $\lambda_1 = 1, x_1 = 2.5, x_2 = 4.5$

From (4) $(2(2.5 - 3) + 1)(2.5) = 0$ holds
From (5) $(2(4.5 - 5) + 1)(4.5)$ holds

Thus $\lambda_1 = 1, x_1 = 2.5, x_2 = 4.5$ satisfies the $K$-$K$-$T$ conditions. Furthermore, since the objective function in convex and the constraints are linear (and hence convex) we have found the global minimum of $z = 0.5$. 

9