Constrained Optimization:
Kuhn-Tucker conditions

Brian Wallace,
Economics dept
b.wallace@rhul.ac.uk

September 23, 2004

Abstract

In this document, we set out the constrained optimisation with inequality constraints and state the Kuhn-Tucker necessary conditions for a solution; after an example, we state the Kuhn-Tucker sufficient conditions for a maximum.

1 The Problem

Suppose we have a function f, which we wish to maximise, together with some constraints, $g_i \leq c_i$, which must be satisfied. This leads us to the problem:

$$\max f(\mathbf{x}) \qquad \text{subject to} \qquad \begin{aligned} &g_1(\mathbf{x}) \leq c_1 \\ &g_2(\mathbf{x}) \leq c_2 \\ &\vdots \\ &g_m(\mathbf{x}) \leq c_m \\ &x_i \geq 0 \end{aligned}$$

We have seen how to solve this problem with equality constraints; we introduce a number of multipliers $\lambda_1, \ldots, \lambda_m$ and form the Lagrangian which we partially differentiate with respect to each x_i and each λ_j and set equal to 0, to get a system of $m + n$ equations with nm variables.
For the inequality case, we do a similar thing. First we form the Lagrangian, $\mathcal{L}$:

$$\mathcal{L} = f(\mathbf{x}) + \lambda_1(c_1 - g_1(\mathbf{x})) + \cdots + \lambda_m(c_m - g_m(\mathbf{x}))$$

Result 1 The Kuhn-Tucker conditions, which are **necessary** (but not sufficient) for **a point** to be a maximum are:

$$\begin{array}{rcrl} \frac{\partial \mathcal{L}}{\partial x_i} \leq 0 & x_i \geq 0 & x_i \frac{\partial \mathcal{L}}{\partial x_i} = 0 & \text{for all } i = 1 \ldots n \\ g_j(\mathbf{x}) \leq c_j & \lambda_j \geq 0 & \lambda_j(c - g_j(\mathbf{x})) = 0 & \text{for all } j = 1 \ldots m \end{array}$$

So, for each variable x_i we have 3 conditions which must be met; similarly for each constraint (and hence λ), we have 3 conditions which must be met. Each point that is a solution to this equation system is a possible candidate for the maximum. Once we have established all the points, we need to check them individually to see which is the real maximum.

The conditions are called the **complementary slackness** conditions. This is because for each set of three conditions, either the first or the second condition can be slack (i.e. not equal to zero), but the third condition ensures that they cannot both be non-zero.

Notes: This is a maximum only problem. To do a minimisation, you need to maximise the function $-f(\mathbf{x})$. Secondly, notation in books varies, so some state the constrant conditions as $g_j(\mathbf{x}) \geq c_j$, in which case the signs of some of the terms in the Lagrangian are altered.

Example 1 $\max f(x_1, x_2) = 4x_1 + 3x_2$ subject to $g(x_1, x_2) = 2x_1 + x_2 \leq 10$ and $x_1,\ x_2 \geq 0$. We first form the Lagrangian:

$$\mathcal{L} = 4x_1 + 3x_2 + \lambda(10 - 2x_1 - x_2)$$

Hence, the necessary conditions for a point to be a maximum are:

$$\begin{array}{rcr} \mathcal{L}_{x_1} = 4 - 2\lambda \leq 0 & x_1 \geq 0 & x_1(4 - 2\lambda) = 0 \\ \mathcal{L}_{x_2} = 3 - \lambda \leq 0 & x_2 \geq 0 & x_2(3 - \lambda) = 0 \\ 2x_1 + x_2 - 10 \leq 0 & \lambda \geq 0 & \lambda(10 - 2x_1 - x_2) = 0 \end{array}$$

We solve this set of inequalities and equations to find points which may be maxima. Let 1 (ii) denote the second condition on the first line etc.

$$1 \text{ (iii)}: \quad x_1(4 - 2\lambda) = 0 \Rightarrow x_1 = 0 \text{ or } \lambda = 2$$

Suppose $\lambda = 2$. Then:

$$2 \text{ (i)}: \quad 3 - \lambda \leq 0 \Rightarrow 1 \leq 0$$

which is clearly false. Hence, we must have that $x_1 = 0$. Then:

$$2 \text{ (iii)}: \quad x_2(3 - \lambda) = 0 \Rightarrow x_2 = 0 \text{ or } \lambda = 3$$

We can see that, if $x_2 = 0$ (together with $x_1 = 0$):

$$3 \text{ (iii)}: \quad 10\lambda = 0 \Rightarrow \lambda = 0$$

But this contradicts 1 (i). So, we have $x_1 = 0$ and $\lambda = 3$. Hence:

$$3 \text{ (iii)}: \quad 3(10 - x_2) = 0 \Rightarrow x_2 = 10$$

So, we have solved the system and found the only solution, which is $x_1 = 0$, $x_2 = 10$, $\lambda = 3$ and hence $f(x_1, x_2) = 30$. As before with equality constraints, λ measures the increase in $f(x_1, x_2)$ for a one unit increase in the constraint (i.e. increasing $c_1 = 10$ by 1).

In general, solving such systems can be very tedious (although not difficult) if there are more than 4 variables and constraints in total. What you need to do is just plug away at every possible combination and eliminate those that do not fit the conditions. When you have eliminated all impossible points, you will be left with a few candidate points which you must then check by substitution into

$f(\mathbf{x})$ to find whether they are the maximum.

For certain types of systems, we can assume the point(s) we find are maxima. This occurs if the function f and the functions g_i satisfy some more conditions, known as the Kuhn-Tucker sufficiency conditions.

Result 2 If the following conditions are satisfied:

1. $f(\mathbf{x})$ is differentiable and concave in the nonnegative orthant
2. each constraint function $g_i(\mathbf{x})$ is differentiable and convex in the nonnegative orthant
3. a point $\mathbf{x}_0$ satisfies the Kuhn-Tucker conditions

then the point $\mathbf{x}_0$ is a maximum. [The nonnegative orthant is the region where each $x_i \geq 0$]

Example 2 Suppose $f(x, y) = 2x + 3y$ and $g(x, y) = x^2 + y^2 \leq 2$. Show that f and g satisfy the Kuhn-Tucker sufficiency conditions and hence find the maxima of $f(x, y)$.

Well, all linear functions are both convex and concave, so f is certainly concave, and is clearly differentiable. As for $g(x, y)$: we can show g is convex by showing the differential d^2g is positive definite. The Hessian matrix associated with g is:

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

So clearly $|H_1|$, $|H_2| > 0$, so g is convex. Hence, any candidate points we find will automatically be maxima. As before, we form the Lagrangian:

$$\mathcal{L} = 2x + 3y + \lambda(2 - x^2 - y^2)$$

The Kuhn-Tucker conditions imply that:

$$\begin{array}{lrcr} (1) & \mathcal{L}_x = 2 - 2\lambda x \leq 0 & x \geq 0 & x(2 - 2\lambda x) = 0 \\ (2) & \mathcal{L}_y = 3 - 2\lambda y \leq 0 & y \geq 0 & y(3 - 2\lambda y) = 0 \\ (3) & x^2 + y^2 \leq 2 & \lambda \geq 0 & \lambda(2 - x^2 - y^2) = 0 \end{array}$$

1 (iii) implies that either $x = 0$ or $\lambda x = 1$. If $x = 0$, then 1 (i) cannot be satisfied, hence $\lambda x = 1$. Similarly, $\lambda y = \frac{2}{3}$ from 2 (iii) (and clearly x, y, $\lambda > 0$). Substituting these values for x and y in 3 (iii), and noting that $\lambda > 0$, we get:

$$2 - x^2 - y^2 = 0 \Rightarrow 2 - \left(\frac{1}{\lambda}\right)^2 - \left(\frac{2}{3\lambda}\right)^2 = 0 \Rightarrow \lambda = \pm\sqrt{\frac{13}{18}}$$

As $\lambda > 0$, we must have $\lambda = \sqrt{\frac{13}{18}}$, hence $y = \sqrt{\frac{8}{13}}$ and $x = \sqrt{\frac{18}{13}}$, and we have a maximum at this point.