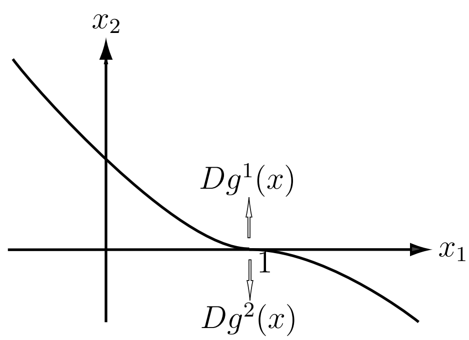


Constraint qualification and sufficient conditions

1 Constraint qualification

We derived the Kuhn-Tucker theorem by assuming regularity of the binding constraints. With the possibility of corner solutions, regularity is too stringent. The regularity condition will not necessarily apply throughout the constraint set. Verifying the regularity condition requires knowing the optimal solution, but the first-order conditions will not identify the optimal solution unless the regularity condition is satisfied – "Catch 22".
To illustrate, consider the problem

$$\max_{x_1, x_2} x_1$$
$$\text{subject to } x_2 - (1 - x_1)^3 \leq 0$$
$$-x_2 \leq 0$$



The solution is obviously $(1, 0)$, but this does not satisfy the Kuhn-Tucker conditions. The gradients of the two constraints at $(1, 0)$ are $(0, 1)$ and $(0, -1)$ which are linearly dependent. The constraints are not regular at the optimum $(1, 0)$. The problem is that the set

$$L(\mathbf{x}^*) = \{\mathbf{dx} \in X : \nabla g_j(\mathbf{x}^*)^T \mathbf{dx} \leq \mathbf{0}, \quad j = 1, 2\} = \{(x_1, 0)\}$$

does not represent the set of feasible perturbations $T(\mathbf{x}^*) = \{(x_1, 0) : x_1 \leq 1\}$. If we add an additional constraint $g_3(x_1, x_2) = x_1 + x_2 \leq 1$, the three gradients are still linearly dependent, but now

$$L(\mathbf{x}^*) = \{\mathbf{dx} \in X : \nabla g_j(\mathbf{x}^*)^T \mathbf{dx} \leq \mathbf{0}, \quad j = 1, 2, 3\} = T(\mathbf{x}^*)$$

and the optimum $(1, 0)$ satisfies the Kuhn-Tucker conditions. We seek conditions on the constraints $\mathbf{g}$ such that the set $L(\mathbf{x}^*)$ correctly represents the set of feasible perturbations $T(\mathbf{x}^*)$, which is known as the **cone of tangents**.

1.1 Constraint qualification conditions (Theorem 5.4)

Suppose that $\mathbf{x}^*$ is a local solution of

$$\max_{\mathbf{x} \in X} f(\mathbf{x}) \text{ subject to } g(\mathbf{x}) \leq \mathbf{0}$$

at which the binding constraints $B(\mathbf{x}^*)$ satisfy one of the following constraint qualification conditions.

Concave CQ g_j is concave for every $j \in B(\mathbf{x}^*)$

Pseudoconvex CQ g_j is pseudoconvex and there exists $\hat{\mathbf{x}} \in X$ such that $g_j(\hat{\mathbf{x}}) < 0$ for every $j \in B(\mathbf{x}^*)$

Quasiconvex CQ g_j is quasiconvex, $\nabla g_j(\mathbf{x}^*) \neq \mathbf{0}$ and there exists $\hat{\mathbf{x}} \in X$ such that $g_j(\hat{\mathbf{x}}) < 0$ for every $j \in B(\mathbf{x}^*)$

Regularity The set $\{\ \nabla g_j(\mathbf{x}^*) : j \in B(\mathbf{x}^*)\ \}$ is linearly independent

Then the Kuhn-Tucker conditions are necessary for an optimal solution.

The most common constraint qualification conditions encountered are

- g_j linear $\implies$ g_j concave (e.g. linear programming)
- **Slater condition:** g_j convex and there exists $\hat{\mathbf{x}} \in X$ such that $g_j(\hat{\mathbf{x}}) < 0$ for every j

1.2 Constraint qualification with nonnegative variables (Corollary 5.4.1)

Provided the binding constraints $j \in B(\mathbf{x}^*)$ in the problem

$$\max_{\mathbf{x} \geq \mathbf{0}} f(\mathbf{x}) \text{ subject to } \mathbf{g}(\mathbf{x}) \leq \mathbf{0}$$

satisfy any one of the following constraint qualification conditions:

Concave CQ

Pseudoconvex CQ

Quasiconvex CQ

the Kuhn-Tucker conditions are necessary for an optimal solution.
Proof. The problem can be specified as

$$\begin{aligned} &\max_{\mathbf{x}} f(\mathbf{x}) \\ \text{subject to } &\mathbf{g}(\mathbf{x}) \leq \mathbf{0} \\ &\mathbf{h}(\mathbf{x}) = -\mathbf{x} \leq \mathbf{0} \end{aligned}$$

We note that $\mathbf{h}$ is linear, and is therefore both concave and convex. Further $D\mathbf{h}[\mathbf{x}] \neq \mathbf{0}$ for every $\mathbf{x}$. Therefore, if $\mathbf{g}$ satisfies one of the three constraint qualification conditions, so does the combined constraint $(\mathbf{g}, \mathbf{h})$. By Theorem 5.4, the Kuhn-Tucker conditions are necessary for a local optimum. □

Example 5.41: The consumer's problem The consumer's problem

$$\max_{\mathbf{x} \geq \mathbf{0}} u(\mathbf{x})$$
$$\text{subject to } \mathbf{p}^T\mathbf{x} \leq m$$

has one functional constraint $g(\mathbf{x}) = \mathbf{p}^T\mathbf{x} \leq m$ and n inequality constraints $h_i(\mathbf{x}) = -x_i \leq 0$, the gradients of which are

$$\nabla g = \mathbf{p} \qquad \nabla h_i = \mathbf{e}_i, \qquad i = 1, 2, \ldots, n$$

where $\mathbf{e}_i$ is the i unit vector (Example 1.79). Provided all prices are positive $\mathbf{p} > \mathbf{0}$, it is clear that these are linearly independent and the regularity condition of Corollary 5.3.2 is always satisfied. However, it is easier to appeal directly to Corollary 5.4.1, and observe that the budget constraint $g(\mathbf{x}) = \mathbf{p}^T\mathbf{x} \leq m$ is linear and therefore concave.

2 Sufficient conditions

We know that the KT conditions are sufficient when f is concave and g is convex. The following is a signficant generalization (e.g. consumer theory).

2.1 Sufficient conditions for a global optimum (Theorem 5.5)

Suppose that $\mathbf{x}^*$ satisfies the KT conditions and

- f is pseudoconcave
- g is quasiconvex

Then $\mathbf{x}^*$ is a global maximum.
Proof. For every j

either $\lambda_j = 0$ which implies $\lambda_j \nabla g_j(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) = 0$ for every $\mathbf{x} \in X$

or $g_j(\mathbf{x}^*) = 0$ and therefore $g_j(\mathbf{x}) \leq 0 = g_j(\mathbf{x}^*)$ for every $\mathbf{x} \in G = \{\mathbf{x} \in X : g_j(\mathbf{x}) \leq 0, \quad j = 1, 2, \ldots, m\}$. Quasiconvexity implies

$$\nabla g_j(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \leq 0 \text{ for every } \mathbf{x} \in G$$

and since $\lambda_j \geq 0$

$$\lambda_j \nabla g_j(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \leq 0 \text{ for every } \mathbf{x} \in G$$

The KT condition is

$$\nabla f(\mathbf{x}^*) = \sum_{j=1}^{m} \lambda_j \nabla g_j(\mathbf{x}^*)$$

and therefore

$$\nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) = \sum_{j=1}^{m} \lambda_j \nabla g_j(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \leq 0$$

Since f is pseudoconcave, this implies that $f(\mathbf{x}^*) \geq f(\mathbf{x})$ □

2.2 Arrow-Enthoven theorem (Corollary 5.5.1)

Suppose that $\mathbf{x}^*$ satisfies the KT conditions and

- f is quasiconcave and $\nabla f(\mathbf{x}^*) \neq \mathbf{0}$
- g is quasiconvex

Then $\mathbf{x}^*$ is a global maximum.

2.3 Necessary and sufficient conditions (Corollary 5.5.3)

Suppose that

- f is quasiconcave and g quasiconvex
- $\nabla f(\mathbf{x}^*) \neq \mathbf{0}$ and $\nabla g_j(\mathbf{x}^*) \neq \mathbf{0}$ for every $j \in B(\mathbf{x}^*)$
- there exists $\hat{\mathbf{x}}$ such that $g_j(\hat{\mathbf{x}}) < 0$ for every $j \in B(\mathbf{x}^*)$

Then the KT conditions are necessary and sufficient for a global optimum.

Example 5.43: The consumer's problem Convex preferences and nonsatiation imply u quasiconcave and $\nabla u(\mathbf{x}) \neq \mathbf{0}$ for every $\mathbf{x} \in X$. Therefore the KT conditions

$$\nabla u(\mathbf{x}^*) = \lambda \mathbf{p}$$

are necessary and sufficient for utility maximization.

Example: Linear programming (Section 5.4.4) A linear programming problem is a special case of the general constrained optimization problem

$$\max_{\mathbf{x} \geq \mathbf{0}} f(\mathbf{x}) \tag{1}$$
$$\text{subject to } g(\mathbf{x}) \leq \mathbf{0}$$

in which both the objective function f and the constraint function $\mathbf{g}$ are linear. Consequently, the Kuhn-Tucker conditions are both necessary and sufficient for a global optimum. The simplex algorithm is an efficient algorithm for solving the Kuhn-Tucker conditions.

3 Homework

1. Suppose that $\mathbf{x}^*$ is a local solution of

$$\max_{\mathbf{x} \in X} f(\mathbf{x}) \text{ subject to } g(\mathbf{x}) \leq \mathbf{0}$$

at which the binding constraints $B(\mathbf{x}^*)$ satisfy one of the above constraint qualification conditions, so that $\mathbf{x}^*$ satisfies the Kuhn-Tucker conditions

$$\nabla f(\mathbf{x}^*) = \sum_{j=1}^{m} \lambda_j \nabla g_j(\mathbf{x}^*) \text{ and } \lambda_j g_j(\mathbf{x}^*) = 0 \quad j = 1, 2 \ldots, m$$

Show that the Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$ are unique if and only if $B(\mathbf{x}^*)$ satisfies the Regularity condition, that is $\{ \ \nabla g_j(\mathbf{x}^*) : j \in B(\mathbf{x}^*) \ \}$ is linearly independent.

2. Suppose that a firm has contracted with its union to hire at least l units of labour at rate w_1 per unit. It can also hire non-union labour at $w_2 < w_1$ per hour. Assume that labour is the only input. Union and non-union labour are equally productive, with diminishing marginal product. Output is sold at a fixed price p.

   (a) Derive and interpret the first-order conditions for maximizing profit. (Note: Although the optimal solution is obvious, please show how it can be derived from the first-order conditions.)

   (b) Are these conditions necessary for a solution?

   (c) Are they sufficient to identify a global optimum?

Solutions 5

1 Since $\lambda_j = 0$ for every $j \notin B(\mathbf{x}^*)$, the Kuhn-Tucker conditions imply

$$\nabla f(\mathbf{x}^*) = \sum_{j \in B(\mathbf{x}^*)} \lambda_j \nabla g_j(\mathbf{x}^*)$$

If $\nabla g_j(\mathbf{x}^*), j \in B(\mathbf{x}^*)$ are independent, then the λ_j are unique (Exercise 1.137). Conversely, if there exist $\mu_1, \mu_2, \dots, \mu_m$ such that with $\mu_j \neq \lambda_j$ for some j and

$$\nabla f(\mathbf{x}^*) = \sum_{j \in B(\mathbf{x}^*)} \mu_j \nabla g_j(\mathbf{x}^*)$$

then

$$\nabla f(\mathbf{x}^*) - \nabla f(\mathbf{x}^*) = \sum_{j \in B(\mathbf{x}^*)} (\lambda_j - \mu_j) \nabla g_j(\mathbf{x}^*) = \mathbf{0}$$

which implies that $\nabla g_j(\mathbf{x}^*), j \in B(\mathbf{x}^*)$ are dependent (Exercise 1.133).

2 (a) Let x_1 denote the quantity of union labour and x_2 the quantity of non-union labour hired by the firm. The firm's optimisation problem is

$$\max_{x_1 \geq l, x_2 \geq 0} pf(x_1 + x_2) - w_1 x_1 - w_2 x_2$$

which can be written as

$$\max_{x_1 \text{ and } x_2 \geq 0} pf(x_1 + x_2) - w_1 x_1 - w_2 x_2$$
$$\text{subject to } g(x_1) = l - x_1 \leq 0$$

Forming the Lagrangean

$$L(x_1, x_2, \lambda) = pf(x_1 + x_2) - w_1 x_1 - w_2 x_2 - \lambda(l - x_1)$$

the Kuhn-Tucker conditions for an optimum are

$$D_{x_1} L = pf'(x_1 + x_2) - w_1 + \lambda = 0 \tag{1}$$
$$D_{x_2} L = pf'(x_1 + x_2) - w_2 \leq 0 \quad x_2 \geq 0 \quad \Big(pf'(x) - w_2\Big) x_2 = 0 \tag{2}$$
$$x_1 \geq l \quad \lambda \geq 0 \qquad \lambda(l - x_1) = 0$$

There are two cases:

$\boxed{x_2 > 0}$ (1) and (2) imply

$$w_1 + \lambda = w_2 \implies \lambda > 0 \implies x_1 = l$$

and

$$pf'(x_2 + l) = w_2$$

$\boxed{x_2 = 0}$ (2) implies

$$pf'(x_1) \leq w_2 < w_1 \implies \lambda > 0 \implies x_1 = l$$

In both cases, the firm hires exactly l units of union labour (x_1). If $pf'(l) > w_2$, then the firm hires additional units of non-union labour (x_2).

(b) Since the constraint function $g(x_1) = l - x_1$ is linear, it is satisfies the CQ condition. The Kuhn-Tucker conditions are necessary.

(c) Diminishing marginal product means that the production function f is concave. Therefore the objective function $pf(x_1 + x_2) - w_1x_1 - w_2x_2$ is concave. Since the constraint function g is convex (linear), the Kuhn-Tucker conditions are also sufficient.