MA3484—Methods of Mathematical Economics
School of Mathematics, Trinity College
Hilary Term 2017
Lecture 23 (March 23, 2017)

David R. Wilkins

5.9. Kuhn-Tucker Theory

We consider the *General Maximum Problem* of nonlinear programming. This problem may be stated as follows:

(The General Maximum Problem of Nonlinear Programming) *Let* g, $f_1, f_2, \ldots, f_m$ *be differentiable real-valued functions on the set*

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\},$$

and let

$$X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0} \text{ and } f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\}.$$

Determine $\mathbf{x}^* \in X$ *such that* $g(\mathbf{x}^*) \geq g(\mathbf{x})$ *for all* $\mathbf{x} \in X$.

Let g, $f_1, f_2, \ldots, f_m$ be differentiable real-valued functions on the set

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\},$$

and let

$$X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0} \text{ and } f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\}.$$

Let $\mathbf{x}^*$ be an element of X with the property that $g(\mathbf{x}^*) \geq g(\mathbf{x})$ for all $\mathbf{x} \in X$.
Let $\gamma\colon (-\delta, \delta) \to \mathbb{R}^m$ be a differentiable path in $\mathbb{R}^m$, defined over an open interval $(-\delta, \delta)$ centred on 0, where $\delta > 0$, with the properties that $\gamma(t) \in X$ for all real numbers t satisfying $0 \leq t < \delta$ and $\gamma(0) = \mathbf{x}^*$, and let

$$\mathbf{v} = \gamma'(0) = \left. \frac{\partial \gamma(t)}{\partial t} \right|_{t=0}.$$

Let

$$I^0 = \{i \in \mathbb{N} : 1 \leq i \leq m \text{ and } f_i(\mathbf{x}^*) = 0\}$$

and

$$J^0 = \{j \in \mathbb{N} : 1 \leq j \leq n \text{ and } (\mathbf{x}^*)_j = 0\}.$$

If $i \in I^*$ then $f_i(\gamma(0)) = 0$ and $f_i(\gamma(t)) \geq 0$ for all $t \in [0, \delta)$. It follows from the Chain Rule of multivariable differential calculus that

$$(Df_i)_{\mathbf{x}^*}(\mathbf{v}) = (\mathrm{grad}\, f_i)^T_{\mathbf{x}^*}\mathbf{v} = \sum_{j=1}^{n} (\mathbf{v})_j \left.\frac{\partial f_i}{\partial x_j}\right|_{\mathbf{x}^*} = \left.\frac{df_i(\gamma(t))}{dt}\right|_{t=0} \geq 0.$$

Thus $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for all $i \in I^0$.
Also if $j \in J^0$ then $(\gamma(0))_j = 0$ and $(\gamma(t))_j \geq 0$ for all $t \in [0, \delta)$ and therefore $(\mathbf{v})_j \geq 0$.

Definition

Let X be a subset of $\mathbb{R}^n$, let g be a differentiable real-valued function defined throughout some open neighbourhood of X, and let $\mathbf{x}^*$ be a point of X. We say that the function g achieves a local maximum on X at the point $\mathbf{x}^*$, if the inequality $g(\mathbf{x}) \leq g(\mathbf{x}^*)$ for all points $\mathbf{x}$ of X that lie sufficiently close to the point $\mathbf{x}^*$.

Let g be a differentiable real-valued function defined throughout some open neighbourhood of the set X, and let $\mathbf{x}^*$ be a point of X. Suppose that the function g achieves a local maximum on X at the point $\mathbf{x}^*$. Let $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$ be a differentiable curve, where $\delta > 0$, $\gamma(0) = \mathbf{x}^*$, and $\gamma(t) \in X$ for all real numbers t satisfying $0 \leq t < \delta$. Then $g(\gamma(t)) \leq g(\gamma(0))$ for all real numbers t satisfying $0 \leq t < \delta$, and therefore

$$(Dg)_{\mathbf{x}^*}(\mathbf{v}) = \left. \frac{d(g(\gamma(t)))}{dt} \right|_{t=0} \leq 0,$$

where

$$\mathbf{v} = \gamma'(0) = \left. \frac{d(\gamma(t))}{dt} \right|_{t=0}.$$

We have shown that if a vector $\mathbf{v}$ is tangent to a differentiable curve $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$ for which $\gamma(0) = \mathbf{x}^*$ and $\gamma(t) \in X$ when $0 \leq t < \delta$ then $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for all $i \in I^0$ and $(\mathbf{v})_j \geq 0$ for all $j \in J^0$. Those points $\mathbf{x}^*$ where these properties characterize tangent vectors to diffentiable curves entering the region X at $\mathbf{x}^*$ are said to satisfy the *constraint qualification* (CQ). This constraint qualification is thus formally defined as follows.

Definition

Let $f_1, f_2, \ldots, f_m$ be differentiable real-valued functions $\mathbb{R}^n$, let

$$X = \{\mathbf{x} \in \mathbb{R}^n : f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\},$$

and let $\mathbf{x}^* \in X$. The *constraint qualification* (CQ) is said to be satisfied at $\mathbf{x}^*$ if, given any vector $\mathbf{v} \in \mathbb{R}^n$ with the properties that $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for all $i \in I^0$, where the set I^0 consists of those indices i between 1 and m for which $f_i(\mathbf{x}^*) = 0$, there exists a differentiable curve $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$, where $\delta > 0$, with the property that

$$\mathbf{v} = \left.\frac{d\gamma(t)}{dt}\right|_{t=0}.$$

Theorem 5.21 (Karush-Kuhn-Tucker)

Let $f_1, f_2, \ldots, f_m$ be differentiable real-valued functions on the set $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$, let

$$X = \{\mathbf{x} \in \mathbb{R}^n : f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\}$$

and let $g\colon X \to \mathbb{R}$ be a real-valued function on X. Suppose that the function g achieves a local maximum at some point $\mathbf{x}^*$ of X and is differentiable there. Suppose also that $f_i(\mathbf{x}^*) = 0$ for $i = 1, 2, \ldots, m$ and that the constraint qualification (CQ) is satisfied at the point $\mathbf{x}^*$. Then there exist non-negative real numbers $\lambda_1, \lambda_2, \ldots \lambda_m$ such that

$$\left.\frac{\partial g}{\partial x_j}\right|_{\mathbf{x}^*} + \sum_{i=1}^{m} \lambda_i \left.\frac{\partial f_i}{\partial x_j}\right|_{\mathbf{x}^*} = 0$$

for $j = 1, 2, \ldots, n$.

Proof

Let C be the subset of $\mathbb{R}^n$ consisting of those vectors $\mathbf{v} \in \mathbb{R}^n$ with the properties that $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for $i = 1, 2, \ldots, m$. Then C is a closed convex cone in $\mathbb{R}^n$. Let $\mathbf{v} \in C$. The constraint qualification (CQ) ensures that there exists a differentiable curve $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$, where $\delta > 0$, such that $\gamma(0) = \mathbf{x}^*$, $\gamma(t) \in X$ when $0 \leq t < \delta$ and

$$\left.\frac{d\gamma(t)}{dt}\right|_{t=0} = \mathbf{v}.$$

But then

$$(Dg)_{\mathbf{x}^*}(\mathbf{v}) = \left.\frac{dg(\gamma(t))}{dt}\right|_{t=0} \leq 0.$$

Let A be the $m \times n$ matrix whose coefficient in the ith row and jth column is $\dfrac{\partial f_i}{\partial x_j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, and let $\mathbf{c}$ be the n-dimensional vector whose jth component is $\dfrac{\partial g}{\partial x_j}$ for $j = 1, 2, \ldots, n$. Then $\mathbf{c}^T\mathbf{v} \leq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ satisfying $A\mathbf{v} \geq \mathbf{0}$. It then follows from Corollary 5.16 that there exists $\mathbf{y} \in \mathbb{R}^m$ for which $\mathbf{y}^T A = -\mathbf{c}$. Let

$$\mathbf{y}^T = (\lambda_1, \lambda_2, \ldots, \lambda_m).$$

Then

$$\sum_{i=1}^{m} \lambda_j \frac{\partial f_i}{\partial x_j} = -\frac{\partial g}{\partial x_j}.$$

The result follows. ∎

Let $f_1, f_2, \ldots, f_m$ be differentiable real-valued functions on the set $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$, let

$$X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0} \text{ and } f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\},$$

and let $\mathbf{x}^* \in X$. The *constraint qualification* (CQ) is said to be satisfied at $\mathbf{x}^*$ if, given any vector $\mathbf{v} \in \mathbb{R}^n$ with the properties that $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for all $i \in I^0$ and $(\mathbf{v})_j \geq 0$ for all $j \in J^0$, where the set I^0 consists of those indices i between 1 and m for which $f_i(\mathbf{x}^*) = 0$ and the set J^0 consists of those indices j between 1 and n for which $(\mathbf{x}^*)_j = 0$, there exists a differentiable curve $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$ (where $\delta > 0$) with the property that

$$\mathbf{v} = \left.\frac{d\gamma(t)}{dt}\right|_{t=0}.$$

Corollary 5.22

Let $f_1, f_2, \ldots, f_m$ be differentiable real-valued functions on the set $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$, let

$$X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0} \text{ and } f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\}$$

and let $g \colon X \to \mathbb{R}$ be a real-valued function on X. Suppose that the function g achieves a local maximum at some point $\mathbf{x}^*$ of X and is differentiable there. Let I^0 be the set consisting of those indices i between 1 and m for which $f_i(\mathbf{x}^*) = 0$ and let J^0 be the set consisting of those indices j between 1 and n for which $(\mathbf{x}^*)_j = 0$. Suppose that the constraint qualification (CQ) is satisfied at the point $\mathbf{x}^*$. Then there exist real numbers $\lambda_1, \lambda_2, \ldots \lambda_m$, and $\mu_1, \mu_2, \ldots, \mu_n$ for which the following properties are satisfied:—

(i) $\left.\dfrac{\partial g}{\partial x_j}\right|_{\mathbf{x}^*} + \displaystyle\sum_{i=1}^{m} \lambda_i \left.\dfrac{\partial f_i}{\partial x_j}\right|_{\mathbf{x}^*} + \mu_j = 0$ for $j = 1, 2, \ldots, n$;

(ii) $\lambda_i \geq 0$ for $i = 1, 2, \ldots, m$ and $\mu_j \geq 0$ for $j = 1, 2, \ldots, n$;

(iii) $\lambda_i = 0$ unless $i \in I^0$, and $\mu_j = 0$ unless $j \in J^0$.

Proof

We may assume, without loss of generality, that $I^0 = \{1, 2, \ldots, m\}$ and that if j is an index between 1 and n for which $(\mathbf{x}^*)_j = 0$ then the coordinate function $\mathbf{x} \mapsto (\mathbf{x})_j$ is included amongst the functions $f_1, f_2, \ldots, f_m$. This follows from the observation that we can, without loss of generality, ignore those functions f_i for which $f_i(\mathbf{x}^*) > 0$. Also we can augment the functions f_i for $i \in I^0$ with the functions $\mathbf{x} \mapsto (\mathbf{x})_j$ for all $j \in J^0$ in order to reduce the general problem to one in which the function g is defined over a subset X of $\mathbb{R}^n$ of the form

$$X = \{\mathbf{x} \in \mathbb{R}^n : f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\},$$

where $X \subset \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$ and $f_i(\mathbf{x}^*) = 0$ for $i = 1, 2, \ldots, m$. The result then follows on application of Theorem 5.21. ∎

Example

This example was presented by Kuhn and Tucker in 1950. Let

$$X = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, \;\; x_2 \geq 0 \text{ and } f(x_1, x_2) \geq 0\},$$

where

$$f(x_1, x_2) = (1 - x_1)^3 - x_2.$$

and let $g \colon \mathbb{R}^2 \to \mathbb{R}$ be defined so that $g(x_1, x_2) = x_1$. Then the maximum value of the function g on X is achieved at $(1, 0)$. At this point the gradient of g is $(1, 0)$ and the gradient of f is $(0, -1)$. These gradients are not collinear. This is not a counter example to the Kuhn-Tucker conditions stated in Theorem 5.21 because the constraint qualification (CQ) is not satisfied at $(1, 0)$.