REVISTA DE INVESTIGACIÓN EN MODELOS MATEMATICOS APLICADOS A LA GESTION Y LA ECONOMIA AÑO 6 VOLUMEN II (2019-II)

CMA CENTRO DE INVESTIGACIÓN EN METODOS CUANTITATIVOS APLICADOS A LA ECONOMÍA Y LA GESTIÓN IADCOM - UBA

Universidad de Buenos Aires Facultad de Ciencias Económicas

LATINDEX
http://www.latindex.org/latindex/ficha?folio=2751

http://www.economicas.uba.ar/institutos_y_centros/revista-modelos-matematicos/

ANÁLISIS DE LA CUALIFICACIÓN DE LAS RESTRICCIONES EN EL TEOREMA DE KUHN-TUCKER

Pablo Matías Herrera y Gonzalo Hernán Ballestero

Universidad de Buenos Aires, Facultad de Ciencias Económicas, Centro de Investigación en Metodologías Básicas y Aplicadas a la Gestión (CIMBAGE), Av. Córdoba 2122 – 1120AAQ. Ciudad Autónoma de Buenos Aires, República Argentina

pabloherrera@economicas.uba.ar; gonzaloballestero@economicas.uba.ar

Resumen

Recibido: 04-03-2019

Aceptado: 30-08-2019

Palabras clave

Cualificación de restricciones – Teorema de Kuhn-Tucker – Condiciones necesarias de optimalidad local – Programación no lineal.

Dado un problema de programación no lineal con restricciones de desigualdad, en este artículo se explicitan las propiedades que deben satisfacer las restricciones para asegurar que las condiciones de Kuhn-Tucker sean condiciones necesarias de optimalidad local. Estas propiedades se denominan cualificación de restricciones y constituyen una hipótesis fundamental para establecer la validez del teorema de Kuhn-Tucker. En este trabajo se presentan diversas cualificaciones de restricciones, así como también las relaciones de implicancia que se establecen entre estas. Finalmente, se demuestra que en el problema del consumidor las restricciones cualifican, de modo tal que la canasta óptima se deduce directamente de las condiciones de primer orden.

ISSN (En línea) 2362 3225

CONSTRAINT QUALIFICATION ANALYSIS IN KUHN-TUCKER THEOREM

Abstract

KEYWORDS

Constraint qualifications – Kuhn-Tucker theorem – Necessary conditions for local optimality – Nonlinear programming.

In this paper is set out to determinate, given a nonlinear programming problem, which are the properties the constraints must have to ensure Kuhn-Tucker necessary conditions of local optimum hold. These constraint qualifications proprieties constitute a fundamental hypothesis to establish the validity of Kuhn-Tucker theorem. Here, diverse constraint qualifications, as well as the logical implications that can be established between them will be presented. Finally, it will be proven that in the consumer maximization problem these constraint qualifications are met such that the optimal bundle is deduced directly from the first order conditions.

ISSN (En línea) 2362 3225

INTRODUCCIÓN

Considere el siguiente problema general de programación no lineal

$$\max\ f(\mathbf{x}) \qquad \text{s. a} \quad \begin{matrix} g_j(\mathbf{x}) \leq 0 \quad j = 1, \dots, m \\ \mathbf{x} \in X \end{matrix} \tag{P}$$

donde $f: \mathbb{R}^n \to \mathbb{R}$ y $g_j: \mathbb{R}^n \to \mathbb{R}$ son funciones continuamente diferenciables y sea $X \subseteq \mathbb{R}^n$ un conjunto abierto y no vacío. A su vez, $\mathcal{D} = \{\mathbf{x} \in X : g_j(\mathbf{x}) \leq 0,\ j = 1, \dots, m\}$ representa el conjunto de soluciones factibles. En caso que alguna restricción se cumpla con igualdad en el punto $\mathbf{x} \in \mathbb{R}^n$, se dice que la restricción está *activa* o *saturada* en ese punto. Caso contrario, se dice que está *inactiva*. Sea $A(\mathbf{x}) = \{j :\ g_j(\mathbf{x}) = 0\}$ el conjunto de índices de restricciones activas en el punto $\mathbf{x} \in \mathbb{R}^n$.

Suponiendo que $\mathbf{x}^* \in \mathcal{D}$ es un máximo local, las condiciones de Kuhn-Tucker (KT) establecen que existe un vector no negativo $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ tal que

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

$$\lambda_j^* \left[b_j - g_j(\mathbf{x}^*)\right] = 0 \quad \text{y} \quad \lambda_j^* \geq 0 \qquad j = 1, \dots, m$$

De este modo, se considera que las condiciones de KT son condiciones necesarias de optimalidad local que caracterizan las soluciones del problema (P). No obstante, existen casos, como el ejemplo planteado por Kuhn & Tucker (1951), donde las condiciones de KT conforman un sistema de ecuaciones inconsistente, de manera tal que no identifican soluciones óptimas.

Ejemplo 1. $\max\ f(\mathbf{x}) = x_1 \quad$ s. a $\quad \begin{cases} g_1(\mathbf{x}) = x_2 - (1 - x_1)^3 \leq 0 \\ g_2(\mathbf{x}) = -x_2 \leq 0 \end{cases}$

Para resolver el problema, se plantea el Lagrangeano

$$L(x_1, x_2, \lambda_1, \lambda_2) = x_1 + \lambda_1(-x_2 + (1 - x_1)^3) + \lambda_2 x_2$$

Las condiciones de Kuhn-Tucker son

$$\frac{\partial L}{\partial x_1} = 1 - 3\lambda_1(1 - x_1)^2 = 0 \tag{1.1}$$

$$\frac{\partial L}{\partial x_2} = -\lambda_1 + \lambda_2 = 0 \tag{1.2}$$

$$\frac{\partial L}{\partial \lambda_1} = -x_2 + (1 - x_1)^3 \geq 0,\ \ \lambda_1 \geq 0,\ \ \lambda_1[-x_2 + (1 - x_1)^3] = 0 \tag{1.3}$$

$$\frac{\partial L}{\partial \lambda_2} = x_2 \geq 0,\ \ \lambda_2 \geq 0,\ \ \lambda_2 x_2 = 0 \tag{1.4}$$

Para encontrar los puntos críticos o candidatos a óptimo, se utiliza el método de ensayo y error.
Caso 1. Activa-Activa.

Dado que ambas restricciones se cumplen con igualdad, de la condición (1.4) se tiene que $x_2 = 0$ y reemplazando en (1.3) se deduce $x_1 = 1$. Sin embargo, al reemplazar el vector $\mathbf{x} = (1,0)$ en la condición (1.1) se obtiene una contradicción. Por lo tanto, no se identifican candidatos a óptimo.

Caso 2. Activa-Inactiva.

De las condiciones de holgura complementaria (1.3) y (1.4) se sigue que $\lambda_1 > 0$ y $\lambda_2 = 0$. Reemplazando en (1.2) se sigue una contradicción. Por lo tanto, no se identifican candidatos a óptimo.

Caso 3. Inactiva-Activa.

De la condición de holgura complementaria (1.3) se tiene que $\lambda_1 = 0$. Reemplazando este valor en la condición (1.1) se sigue una contradicción. De este modo, no se identifican candidatos a óptimo.

Caso 4. Inactiva-Inactiva.

Del mismo modo que el caso anterior, de la condición (1.3) y (1.1) se sigue una contradicción, de manera tal que no se identifican candidatos a óptimo.

Si bien a partir de las condiciones de Kuhn-Tucker no se deducen candidatos a óptimo, el mismo puede encontrarse mediante el siguiente razonamiento. De la primera restricción, se tiene $(1-x_1)^3 \geq x_2$. Debido a la condición de no negatividad sobre x_2 se sigue $(1-x_1)^3 \geq 0$, de manera tal que $x_1 \leq 1$. Por lo tanto, se deduce que $x_1 = 1$. Reemplazando el valor obtenido en la primera restricción, se obtiene $x_2 = 0$. Por lo tanto, el vector $\mathbf{x}^* = (1,0)$ es la solución del problema.

Este ejemplo demuestra que para asegurar la validez del teorema de Kuhn-Tucker, las restricciones deben verificar cierta hipótesis adicional, denominada *cualificación de restricciones*. En este artículo se explicitan diversas cualificaciones de este tipo que aseguran que las condiciones de Kuhn-Tucker sean condiciones necesarias de optimalidad local.

Dentro de la literatura de optimización no lineal, se han desarrollado diferentes propiedades que deben satisfacer las restricciones para considerar que éstas cualifican (Abadie, 1966; Arrow, Hurwicz, & Uzawa, 1961; Gould & Tolle, 1971; Mangasarian & Fromovitz, 1967; Kuhn & Tucker, 1951; Slater, 1959). La búsqueda ha sido motivada por hallar condiciones cada vez más generales y especificar cuáles son las interrelaciones que se establecen entre sí (véase Bazaraa et al., 1972 y Peterson, 1973).

El artículo está organizado de acuerdo al siguiente esquema. En la segunda sección se definen los conceptos necesarios para describir las condiciones necesarias de primer orden en problemas de programación no lineal. Luego, en la tercera sección se demuestra el teorema de Kuhn-Tucker utilizando la cualificación de restricciones de Abadie, de manera tal que se muestra explícitamente la relación de implicancia que se establece entre la cualificación de restricciones y la validez de las condiciones de Kuhn-Tucker. En la cuarta sección se enuncian cualificaciones de restricciones alternativas y se presentan esquemáticamente las relaciones que existen entre éstas. Finalmente, mediante dos ejemplos se ilustra el procedimiento para decidir si las restricciones cualifican o no.

1. CONDICIONES NECESARIAS DE OPTIMALIDAD

En esta sección, se introducen las nociones fundamentales para enunciar las propiedades que caracterizan las soluciones óptimas del problema (P). La intuición geométrica detrás del concepto de máximo local en problemas de programación no lineal con restricciones de desigualdad es la siguiente. Supóngase una persona que se encuentra escalando una montaña y, por diversos motivos, existen zonas de acceso restringido. De este modo, la montaña se divide en áreas factibles o posibles de ser visitadas y áreas no factibles. Además, supóngase que las condiciones climáticas son adversas y existe una espesa niebla que recubre la totalidad del paisaje, reduciendo completamente la visión del montañista. En este contexto, ¿cómo sabe que ha alcanzado una cima factible o máximo local factible? Debido a la niebla, sólo puede usar información local, de modo que puede inferir que ha alcanzado un máximo local factible si, al dar un paso en cualquier dirección factible, éste lo lleva a una altura menor con respecto a la posición inicial. Esta es la idea que subyace tras el teorema 1. El teorema 2 es simplemente una extensión generalizada del teorema 1. A continuación, se definen los conceptos que permiten formalizar la descripción mencionada anteriormente.

Definición 1. Dada una función $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$, un vector $\mathbf{d} \in \mathbb{R}^n$ es una *dirección ascendente* de f en un punto $\mathbf{x} \in \mathcal{D}$ si existe un escalar $\delta > 0$ tal que $f(\mathbf{x} + \lambda\mathbf{d}) > f(\mathbf{x})$ para todo $\lambda \in (0, \delta)$.

Una dirección ascendente se denomina también *dirección de mejora*, debido a que un pequeño desplazamiento desde un punto a otro a lo largo de una dirección ascendente lleva a puntos donde aumenta o mejora el valor de la función objetivo.

Definición 2. Dada una función $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$, el *cono de direcciones ascendentes* de f en un punto $\mathbf{x} \in \mathcal{D}$ es el conjunto

$$F(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \; f(\mathbf{x} + \lambda\mathbf{d}) > f(\mathbf{x}), \qquad \forall\, \lambda \in (0, \delta) \qquad \text{para algún } \delta > 0 \}$$

Proposición 1. (*Condición suficiente de dirección ascendente*). Sea f una función diferenciable en $\mathbf{x}$, sea un vector $\mathbf{d} \in \mathbb{R}^n$. Si se verifica que $\nabla f(\mathbf{x})^{\mathrm{T}}\, \mathbf{d} > 0$, entonces $\mathbf{d}$ es una dirección ascendente con respecto a f en $\mathbf{x}$.

Demostración. Debido a que f es diferenciable en $\mathbf{x}$, es posible construir la expansión de Taylor de primer orden alrededor del punto $\mathbf{x}$

$$f(\mathbf{x} + \lambda\mathbf{d}) = f(\mathbf{x}) + \lambda\nabla f(\mathbf{x})^{\mathrm{T}}\,\mathbf{d} + R(\lambda) \qquad \text{con} \qquad \lim_{\lambda \to 0} \frac{R(\lambda)}{\lambda} = 0$$

Considere un valor de λ lo suficientemente pequeño, de manera tal que el error de aproximación tienda a cero. De este modo, se obtiene

$$f(\mathbf{x} + \lambda\mathbf{d}) = f(\mathbf{x}) + \lambda\nabla f(\mathbf{x})^{\mathrm{T}}\,\mathbf{d}$$

Por definición, se sabe que $\lambda > 0$ y, por hipótesis, $\nabla f(\mathbf{x})^{\mathrm{T}}\,\mathbf{d} > 0$, de modo tal que se verifica $f(\mathbf{x} + \lambda\mathbf{d}) - f(\mathbf{x}) > 0$. De esta manera, se deduce

$$f(\mathbf{x} + \lambda\mathbf{d}) > f(\mathbf{x})$$

Por lo tanto, queda demostrado que $\mathbf{d} \in \mathbb{R}^n$ es una dirección ascendente. ■

Definición 3. Dada una función $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$, se dice que un vector no nulo $\mathbf{d} \in \mathbb{R}^n$ es una *dirección factible* de $\mathcal{D}$ en un punto $\mathbf{x}$ si existe un escalar $\delta > 0$ tal que $\mathbf{x} + \lambda\mathbf{d} \in \mathcal{D}$ para todo $\lambda \in (0, \delta)$.

De este modo, si $\mathbf{d}$ es una dirección factible, entonces un pequeño desplazamiento λ desde $\mathbf{x}$ en dirección a $\mathbf{d}$ lleva a puntos que pertenecen al conjunto $\mathcal{D}$.

Definición 4. Dada una función $f: \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$, el *cono de direcciones factibles* de f en un punto $\mathbf{x} \in \mathcal{D}$ es el conjunto

$$D(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \ \mathbf{x} + \lambda\mathbf{d} \in \mathcal{D}, \qquad \forall\, \lambda \in (0, \delta) \qquad \text{para algún } \delta > 0 \qquad \text{con} \qquad \mathbf{d} \neq \mathbf{0}\,\}$$

Proposición 2. (*Condición necesaria de dirección factible*). Sea $g_j: \mathbb{R}^n \to \mathbb{R}^m$ una función diferenciable en $\mathbf{x}$, y sea $\mathbf{d} \in \mathbb{R}^n$ una dirección factible. Entonces se verifica que $\nabla g_j(\mathbf{x})^{\mathrm{T}}\, \mathbf{d} \leq 0$ para todo $j \in A(\mathbf{x})$.

Demostración. (*Por reducción al absurdo*). Supóngase que $\nabla g_j(\mathbf{x})^{\mathrm{T}}\, \mathbf{d} > 0$ para algún $j \in A(\mathbf{x})$. Debido a que g_j es diferenciable en $\mathbf{x}$, es posible construir la expansión de Taylor de primer orden alrededor del punto $\mathbf{x}$

$$g_j(\mathbf{x} + \lambda\mathbf{d}) = g_j(\mathbf{x}) + \lambda\nabla g_j(\mathbf{x})^{\mathrm{T}}\mathbf{d} + R(\lambda) \qquad \text{con} \qquad \lim_{\lambda\to 0}\frac{R(\lambda)}{\lambda} = 0$$

$$g_j(\mathbf{x}+\lambda\mathbf{d}) - g_j(\mathbf{x}) = \lambda\nabla g_j(\mathbf{x})^{\mathrm{T}}\,\mathbf{d} + R(\lambda) \qquad \text{con} \qquad \lim_{\lambda\to 0}\frac{R(\lambda)}{\lambda} = 0$$

Supóngase un valor de λ lo suficientemente pequeño, de manera tal que el error de aproximación tiende a cero. Debido a que λ es un escalar positivo y, por hipótesis, $\nabla g_j(\mathbf{x})^{\mathrm{T}}\, \mathbf{d} > 0$, entonces se tiene que

$$g_j(\mathbf{x} + \lambda\mathbf{d}) > g_j(\mathbf{x})$$

Por lo tanto, se deduce que no existe ningún escalar $\delta > 0$ tal que

$$g_j(\mathbf{x} + \lambda\mathbf{d}) < g_j(\mathbf{x}) \qquad \forall\, \lambda \in (0, \delta)$$

Lo que contradice que $\mathbf{d} \in \mathbb{R}^n$ sea una dirección factible. ■

El siguiente teorema refiere a la condición necesaria de optimalidad local. Básicamente, se establece que si $\mathbf{x}^*$ es un máximo local del problema (P), entonces no existe dirección que sea factible y ascendente simultáneamente.

Teorema 1. Si $\mathbf{x}^* \in \mathcal{D}$ es un máximo local del problema (P), entonces $F_0(\mathbf{x}^*) \cap D(\mathbf{x}^*) = \emptyset$, donde $F_0(\mathbf{x}^*) = \{\mathbf{d} \in \mathbb{R}^n : \nabla f(\mathbf{x}^*)^{\mathrm{T}}\, \mathbf{d} > 0\}$ y $D(\mathbf{x}^*)$ es el cono de direcciones factibles.

Demostración. (*Por reducción al absurdo*). Supóngase que existe una dirección factible de mejora $\mathbf{d} \in F_0(\mathbf{x}^*) \cap D(\mathbf{x}^*)$. Consecuentemente, por la Definición 3 y la Proposición 1, existe un $\delta > 0$ tal que

$$f(\mathbf{x}^* + \lambda\mathbf{d}) > f(\mathbf{x}^*) \quad \text{y} \quad \mathbf{x}^* + \lambda\mathbf{d} \in \mathcal{D}, \qquad \forall\, \lambda \in (0, \delta)$$

A su vez, dado que $\mathbf{x}^*$ es un máximo local, entonces debe existir un $r > 0$ tal que

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \qquad \forall\, \mathbf{x} \in B(\mathbf{x}^*, r) \cap \mathcal{D}$$

donde $B(\mathbf{x}^*, r)$ es una bola abierta con centro en $\mathbf{x}^*$ y radio r. Considérese un valor de lambda tal que $\lambda < \frac{r}{\mathbf{d}}$. De este modo, se sigue

$$f(\mathbf{x}^* + \lambda\mathbf{d}) > f(\mathbf{x}^*) \quad \text{donde} \quad \mathbf{x}^* + \lambda\mathbf{d} \in B(\mathbf{x}^*, r) \cap \mathcal{D}$$

Lo que contradice la hipótesis que $\mathbf{x}^*$ es un máximo local. ■

Sin embargo, si en lugar de considerar el cono de direcciones factibles se tiene en cuenta un conjunto más amplio, denominado *cono de direcciones tangentes* $T(\mathbf{x})$, que contiene al conjunto $D(\mathbf{x})$, entonces se obtiene una versión más general del Teorema 1. De este modo, como se verá a continuación, el Teorema 2 debe ser entendido como una caracterización generalizada de las condiciones necesarias de optimalidad local.

Definición 4. Sea $\mathcal{D} \subseteq \mathbb{R}^n$ un conjunto cerrado y sea $\mathbf{x} \in \mathcal{D}$. Entonces $\mathbf{d} \in \mathbb{R}^n$ es una *dirección tangente* a $\mathcal{D}$ si existe una sucesión de escalares positivos $\{\lambda_k\}_{k\in\mathbb{N}}$ y una sucesión factible $\{\mathbf{x}_k\}_{k\in\mathbb{N}} \subset \mathcal{D}$ tal que

$$\mathbf{x} = \lim_{k\to\infty} \mathbf{x}_k \quad \wedge \quad \mathbf{d} = \lim_{k\to\infty} \lambda_k(\mathbf{x}_k - \mathbf{x})$$

Definición 5. El conjunto de direcciones tangentes a $\mathcal{D}$ en un punto $\mathbf{x} \in \mathbb{R}^n$

$$T(\mathbf{x}) = \left\{\mathbf{d} \in \mathbb{R}^n \mid \exists\{\mathbf{x}_k\} \subset \mathcal{D}, \{\lambda_k\} \subset \mathbb{R}_+ : \mathbf{d} = \lim_{k\to\infty} \lambda_k(\mathbf{x}_k - \mathbf{x}) \wedge \mathbf{x} = \lim_{k\to\infty} \mathbf{x}_k\right\}$$

se denomina *cono direcciones tangentes* a $\mathcal{D}$ en $\mathbf{x}$.

Puede demostrarse que $D(\mathbf{x}) \subseteq T(\mathbf{x})$ (Carter, 2001). En virtud de esta propiedad, es posible generalizar las condiciones de primer orden mediante el teorema que sigue a continuación.

Teorema 2. Si $\mathbf{x}^* \in D$ es un máximo local del problema (P), entonces $F_0(\mathbf{x}^*) \cap T(\mathbf{x}^*) = \emptyset$.

Demostración. Supóngase que $\mathbf{d} \in T(\mathbf{x}^*)$ de modo que existe una sucesión $\{\mathbf{x}_k\}$ que pertenece al conjunto factible y una sucesión de escalares $\{\lambda_k\} \subset \mathbb{R}_+$ tal que

$$\mathbf{d} = \lim_{k\to\infty} \lambda_k(\mathbf{x}_k - \mathbf{x}^*) \quad \text{y} \quad \mathbf{x}^* = \lim_{k\to\infty} \mathbf{x}_k$$

Debido a que f es diferenciable en $\mathbf{x}$, es posible construir la expansión de Taylor de primer orden alrededor del punto $\mathbf{x}_k$

$$f(\mathbf{x}_k) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}_k - \mathbf{x}^*) + R(\mathbf{x}_k - \mathbf{x}^*)$$

$$f(\mathbf{x}_k) - f(\mathbf{x}^*) = \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}_k - \mathbf{x}^*) + R(\mathbf{x}_k - \mathbf{x}^*)$$

Dado que $\mathbf{x}^*$ es un máximo local, entonces se verifica que $f(\mathbf{x}_k) \leq f(\mathbf{x}^*)$, de modo que se obtiene

$$\nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}_k - \mathbf{x}^*) + R(\mathbf{x}_k - \mathbf{x}^*) \leq 0$$

Multiplicando y dividiendo el último término de la inecuación por $(\mathbf{x}_k - \mathbf{x}^*)$

$$\nabla f(x^*)^T(x_k - x^*) + (x_k - x^*)\frac{R(x_k - x^*)}{(x_k - x^*)} \leq 0$$

Multiplicando por $\lambda_k > 0$ a toda la expresión y tomando el límite

$$\lim_{k\to\infty}\left[\lambda_k \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}_k - \mathbf{x}^*) + \lambda_k(\mathbf{x}_k - \mathbf{x}^*)\frac{R(\mathbf{x}_k - \mathbf{x}^*)}{(\mathbf{x}_k - \mathbf{x}^*)}\right] \leq 0$$

$$\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \mathbf{d} \cdot 0 \leq 0$$

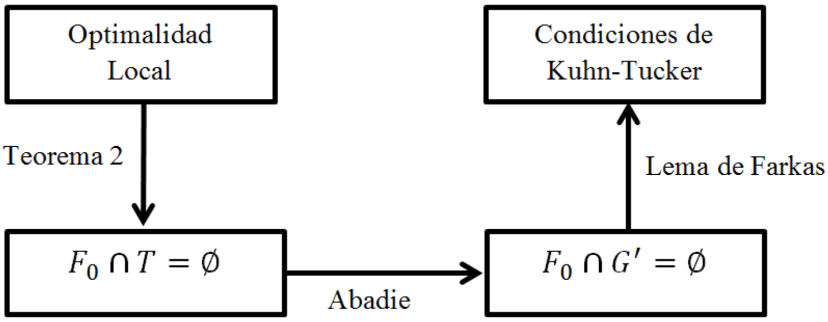
$$\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} \leq 0$$

Por lo tanto, se demuestra que $\mathbf{d} \notin F_0(\mathbf{x}^*)$. ■

2. TEOREMA DE KUHN-TUCKER

En esta sección, se demuestra el teorema de Kuhn-Tucker utilizando la condición necesaria de optimalidad enunciada en la sección anterior (véase Teorema 2) y la cualificación de restricciones de Abadie (1966). Bajo tales hipótesis, las condiciones de KT se derivan inmediatamente a través del Lema de Farkas (1902). El siguiente diagrama resume esquemáticamente las implicancias lógicas mencionadas

Gráfico 1. Hipótesis del teorema de Kuhn-Tucker



Fuente: Bazaraa *et al.*, 2013

Cualificación de restricciones de Abadie. Un vector $\mathbf{x} \in \mathcal{D}$ satisface la cualificación de restricciones de Abadie si $T(\mathbf{x}) = G'(\mathbf{x})$, donde $G'(\mathbf{x}) = \{\mathbf{d} : \nabla g_j(\mathbf{x})^{\mathrm{T}}\, \mathbf{d} \leq 0,\ \forall j \in A(\mathbf{x})\}$.

Lema de Farkas. Sea A una matriz de $m \times n$ y sea $\mathbf{b} \in \mathbb{R}^m$. Entonces uno y sólo uno de los siguientes sistemas tiene solución y el otro es inconsistente

$$A\mathbf{x} = \mathbf{b}, \qquad \mathbf{x} \geq \mathbf{0} \tag{3.1}$$

$$A^T\mathbf{x} \leq \mathbf{0}, \qquad \mathbf{b}^T\mathbf{y} > 0 \tag{3.2}$$

Demostración. Véase (Gale, 1989) o (Lauritzen, 2013).

Teorema de Kuhn-Tucker. Supóngase que $\mathbf{x}^* \in \mathcal{D}$ es un máximo local del problema (P) y verifica la cualificación de restricciones de Abadie. Entonces existe un vector $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ tal que se satisfacen las siguientes condiciones

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

$$\lambda_j^* \geq 0 \quad \text{y} \quad \lambda_j^* \left[b_j - g_j(\mathbf{x}^*)\right] = 0 \qquad j = 1, \ldots, m$$

Demostración. Dado que $\mathbf{x}^* \in \mathcal{D}$ es un máximo local, por el Teorema 2 se verifica $F_0(\mathbf{x}^*) \cap T(\mathbf{x}^*) = \emptyset$. A su vez, debido a que se satisface la cualificación de restricciones de Abadie, $T(\mathbf{x}^*) = G'(\mathbf{x}^*)$. De este modo, se sigue $F_0(\mathbf{x}^*) \cap G'(\mathbf{x}^*) = \emptyset$. Esto implica que no existe ninguna dirección $\mathbf{d} \in \mathbb{R}^n$ que sea solución del sistema

$$\nabla f(\mathbf{x}^*)^{\mathrm{T}}\, \mathbf{d} > 0, \qquad \nabla g_j(\mathbf{x}^*)^{\mathrm{T}}\, \mathbf{d} \leq 0 \quad \text{donde } j \in A(\mathbf{x}^*)$$

Por el Lema de Farkas, necesariamente existe una solución para el siguiente sistema

$$\sum_{j\in A(\mathbf{x}^*)} \lambda_j \nabla g_j(\mathbf{x}^*) = \nabla f(\mathbf{x}^*), \quad \lambda_j \geq 0 \quad \text{donde } \; j \in A(\mathbf{x}^*)$$

Luego, completando con ceros el vector de multiplicadores de Lagrange asociados a las restricciones inactivas

$$\lambda_k = 0 \qquad \forall\, k \notin A(\mathbf{x}^*)$$

Se demuestra la existencia de un $\boldsymbol{\lambda} \in \mathbb{R}^m$ que satisface las condiciones de Kuhn-Tucker. ■

Dado que la estructura lógica del teorema de Kuhn-Tucker es un condicional de la forma $A \Rightarrow B$, queda claro que la hipótesis de cualificación de restricciones resulta fundamental para asegurar la validez del mismo. Asumiendo que existe un máximo global, si las restricciones no cualifican, las condiciones de KT pueden no identificar candidato a óptimo alguno, conformando un sistema de ecuaciones inconsistente. En este sentido, para encontrar el conjunto de candidatos a óptimo, Sydsæter *et al.* (2008) proponen el siguiente procedimiento

i) Encontrar todos los puntos factibles que satisfacen las condiciones de Kuhn-Tucker.
ii) Encontrar todos los puntos factibles que no verifican la cualificación de restricciones.

La solución del problema (P) se encontrará entre estas dos clases de puntos candidatos.
No obstante, debido a que la condición de Abadie es una proposición demasiada abstracta, el paso (ii) resulta difícil de verificar operacionalmente. Por este motivo, en la sección que sigue a continuación se enuncian cualificaciones de restricciones alternativas que son más sencillas de comprobar en términos prácticos.

3. OTRAS CUALIFICACIONES DE RESTRICCIONES

En esta sección se presentan tres cualificaciones de restricciones. En virtud de los teoremas 5 y 6, todas éstas implican la cualificación de Abadie y, por transitividad, aseguran la validez del teorema de Kuhn-Tucker. El gráfico 2 describe sintéticamente el orden de implicancia e interdependencia que se establece entre las diferentes cualificaciones. Finalmente, se presentan dos ejemplos para ilustrar cómo verificar si las restricciones cualifican.

Cualificación de restricción de Mangasarian-Fromovitz. Un punto $\mathbf{x} \in \mathcal{D}$ satisface la *cualificación de restricción de Mangasarian-Fromovitz* (Mangasarian & Fromovitz, 1967) si existe un vector $\mathbf{d} \in \mathbb{R}^n$ tal que

$$\nabla g_j(\mathbf{x})^{\mathrm{T}}\, \mathbf{d} < 0 \qquad \forall j \in A(\mathbf{x})$$

Condición de Slater. El problema (P) satisface la *condición de Slater* (Slater, 1959) si cada función restricción $g_j\colon \mathbb{R}^n \to \mathbb{R}$ es convexa y, además, existe al menos un $\mathbf{x} \in \mathcal{D}$ tal que $g_j(\mathbf{x}) < 0$ para todo $j = 1, \dots m$.

Condición de regularidad. Un punto $\mathbf{x} \in \mathcal{D}$ es un *punto regular* o cumple la *condición de regularidad* si el conjunto de vectores gradientes de las restricciones activas en $\mathbf{x}$

$$\{\nabla g_j(\mathbf{x}) : j \in A(\mathbf{x})\}$$

conforman un sistema linealmente independiente (Fiacco & McCormick, 1990).

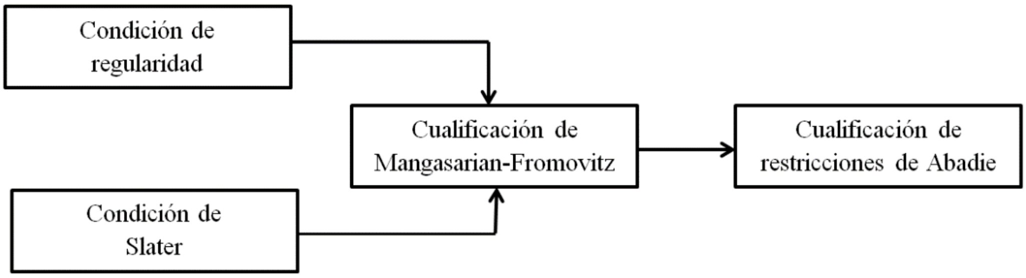
Teorema 5. Si el problema (P) satisface la condición de Slater, entonces se cumple la cualificación de Mangasarian-Fromovitz para todo $\mathbf{x} \in \mathcal{D}$.

Demostración. Véase (Andréasson *et al.*, 2005).

Teorema 6. Si un vector $\mathbf{x} \in \mathcal{D}$ satisface la condición de regularidad, entonces cumple la cualificación de Mangasarian-Fromovitz. A su vez, si $\mathbf{x} \in \mathcal{D}$ verifica la cualificación de Mangasarian-Fromovitz, entonces cumple la cualificación de Abadie.

Demostración. Véase (Andréasson *et al.*, 2005) o (Eustaquio *et al.*, 2008).

Gráfico 2. Interrelación entre cualificación de restricciones



Fuente. Elaboración propia.

Para finalizar este trabajo, se presentan dos ejemplos que ilustran el procedimiento para verificar la hipótesis de cualificación de restricciones. El primer ejemplo corresponde a (Sundaram, 1996) y tiene la particularidad que los candidatos a óptimo se obtienen resolviendo las condiciones de KT así como también identificando aquellos puntos donde las restricciones no cualifican. El segundo ejemplo se refiere al problema del consumidor, en el cual se observa que todo vector factible es un punto regular. De este modo, la canasta que maximiza la utilidad se deduce directamente de las condiciones de primer orden.

Ejemplo 2. $\max f(x) = 2x^3 - 3x^2 \quad$ s. a $\quad g(x) = -(3-x)^3 \leq 0$

Para encontrar todos los candidatos a óptimo, primero se calculan los puntos que satisfacen las condiciones de KT. Luego, se computan aquellos puntos que no satisfacen la condición de regularidad.

Se plantea el lagrangeano asociado al problema

$$L(x,\lambda) = 2x^3 - 3x^2 + \lambda(3-x)^3$$

Luego, se establecen las condiciones necesarias de optimalidad

$$\frac{\partial L}{\partial x} = 6x^2 - 6x - 3\lambda(3-x)^2 = 0 \tag{4.1}$$

$$\frac{\partial L}{\partial \lambda} = (3-x)^3 \geq 0,\ \ \lambda \geq 0,\ \ \lambda(3-x)^3 = 0 \tag{4.2}$$

Caso 1. Activa.

De la condición de holgura complementaria, se obtiene $(3-x)^3 = 0$. De esto, se sigue que $x = 3$. Reemplazando este valor en (4.1) se deduce una contradicción. Por lo tanto, no se identifican candidatos a óptimo.

Caso 2. Inactiva.

De la condición (4.1), se sigue $6x^2 - 6x = 0$. Resolviendo esta ecuación se obtienen dos candidatos a óptimo, $x_1 = 0$ y $x_2 = 1$.

Para identificar la totalidad de candidatos a óptimo, además se deben identificar aquellos puntos que no verifican la hipótesis de cualificación de restricciones. Para ello, se evalúa la condición de regularidad, la cual exige que los vectores gradientes de las restricciones activas sean linealmente independientes.

Caso 1. Activa

Dado que sólo una restricción está activa, se debe verificar que el vector $\nabla g = [3(3-x)^2]$ sea linealmente independiente[1]. El único punto que no satisface esta condición es $x = 3$. Por lo tanto, en este caso se identifica un candidato a óptimo adicional.

Caso 2. Inactiva

La cualificación de restricciones se satisface de manera trivial debido a que el conjunto de restricciones activas no contiene ningún elemento. Por lo tanto, no se identifican candidatos a óptimo en este caso.

En conclusión, para este ejemplo existen tres candidatos a óptimo $x_1 = 0, x_2 = 1$ y $x_3 = 3$.

Ejemplo 3. $\max u(\mathbf{x}) \quad \text{s. a} \quad \begin{cases} g_1(\mathbf{x}) = p_1x_1 + p_2x_2 \leq m \\ g_2(\mathbf{x}) = -x_1 \leq 0 \\ g_3(\mathbf{x}) = -x_2 \leq 0 \end{cases}$

donde $p_1, p_2, m > 0$. Además, se asume que $u: \mathbb{R}^n \to \mathbb{R}$ es una función cóncava, continuamente diferenciable y estrictamente creciente en x_1 y x_2. Debido a estos supuestos, por el teorema de Weierstrass se sigue que existe un máximo global. A continuación, se demuestra que la solución del problema necesariamente se deduce de las condiciones de primer orden.

Dado que existen tres restricciones, cada una de las cuales admite dos estados, se tienen 8 casos posibles. Sin embargo, debido a que u es una función estrictamente creciente, la restricción presupuestaria debe cumplirse con igualdad en el óptimo, por lo que se descartan los casos donde la primera restricción está inactiva. A su vez, es fácil comprobar que el caso donde todas las restricciones están activas implica una contradicción. En conclusión, el análisis queda acotado a 3 casos.

Caso 1. Activa-Inactiva-Activa.

La condición de regularidad se satisface sólo si los vectores gradiente de la primera y tercera restricción son linealmente independientes. Para ello, el determinante de la matriz conformada por los vectores fila ∇g_1 y ∇g_3 debe ser no nulo

$$\begin{vmatrix} p_1 & p_2 \\ 0 & -1 \end{vmatrix} \neq 0$$

Dado que p_1 y p_2 son estrictamente positivos, se verifica la condición de regularidad.

[1] Recuérdese que un vector $\mathbf{v} \in \mathbb{R}^n$ es linealmente independiente si y sólo si $\mathbf{v} \neq \mathbf{0}$.

Caso 2. Activa-Activa-Inactiva.

Operando de manera similar al caso anterior, se verifica la condición de regularidad.

Caso 3. Activa-Inactiva-Inactiva

Debido a que únicamente la primera restricción se encuentra activa, la condición de regularidad se satisface sólo si $\nabla g_1 = (p_1, p_2) \neq 0$. Una vez más, como los precios son estrictamente positivos, se cumple la cualificación de restricciones.

En resumen, se demuestra que todos los casos satisfacen la hipótesis de cualificación de restricciones. De este modo, la solución del problema del consumidor se deduce directamente de la resolución de las condiciones de primer orden.

CONCLUSIÓN

En este artículo se explicitaron las propiedades que deben satisfacer las restricciones para asegurar que las condiciones de Kuhn-Tucker sean condiciones necesarias de optimalidad local. Utilizando la cualificación de restricciones de Abadie, se demostró el teorema de KT y se concluyó que la cualificación de restricciones es una hipótesis fundamental para establecer la validez de las condiciones de primer orden. Además de la cualificación de Abadie, también se enunciaron cualificaciones de restricciones alternativas, tales como la cualificación de Mangasarian-Fromovitz, la condición de Slater y la condición de regularidad. Esta última destaca sobre el resto debido a que es fácilmente verificable en términos operacionales. Luego, se expuso mediante una representación diagramática el sentido de las relaciones de implicancia que se establecen entre las diferentes cualificaciones. Finalmente se presentaron dos ejemplos con el objetivo de ilustrar el procedimiento para determinar la cualificación de restricciones. Particularmente se demostró que, en el problema del consumidor, un ejemplo característico en la literatura económica, las restricciones cualifican, de modo tal que la canasta óptima se deduce directamente de las condiciones de primer orden.

REFERENCIAS BIBLIOGRÁFICAS

Abadie, J. (1966). *On the Kuhn-Tucker theorem*. Operations Research Department at the University of California.

Andréasson, N., Evgrafov, A., Patriksson, M., Gustavsson, E., & Önnheim, M. (2005). *An introduction to continuous optimization: foundations and fundamental algorithms*. Studentlitteratur Lund.

Arrow, K. J., Hurwicz, L., & Uzawa, H. (1961). Constraint qualifications in maximization problems. *Naval Research Logistics Quarterly*, *8*(2), 175-191.

Bazaraa, M. S., Goode, J. J., & Shetty, C. M. (1972). Constraint qualifications revisited. *Management Science*, *18*(9), 567–573.

Bazaraa, Mokhtar S, Sherali, H. D., & Shetty, C. M. (2013). *Nonlinear programming: theory and algorithms*. John Wiley & Sons.

Carter, M. (2001). *Foundations of mathematical economics*. MIT Press.

Eustaquio, R. G., Karas, E. W., & Ribeiro, A. A. (2008). Constraint qualifications for nonlinear programming. *Federal University of Parana*.

Farkas, J. (1902). Uber die Theorie der einfachen Ungeichungen. *J. Reine Angew. Math.*, *124*, 1-24.

Fiacco, A. V., & McCormick, G. P. (1990). *Nonlinear programming: sequential unconstrained minimization techniques* (Vol. 4). Siam.

Gale, D. (1989). *The theory of linear economic models*. University of Chicago press.

Gould, F., & Tolle, J. W. (1971). A necessary and sufficient qualification for constrained optimization. *SIAM Journal on Applied Mathematics*, *20*(2), 164-172.

Kuhn, H., & Tucker, A. (1951). Nonlinear Programming (pp. 481-492). Presentado en Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press.

Lauritzen, N. (2013). *Undergraduate convexity: from Fourier and Motzkin to Kuhn and Tucker*. World Scientific.

Mangasarian, O. L., & Fromovitz, S. (1967). The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. *Journal of Mathematical Analysis and applications*, *17*(1), 37-47.

Peterson, D. W. (1973). A review of constraint qualifications in finite-dimensional spaces. *SIAM Review*, *15*(3), 639-654.

Slater, M. (1959). *Lagrange Multipliers Revisited*. Cowles Foundation for Research in Economics, Yale University.

Sundaram, R. K. (1996). A first course in optimization theory. Cambridge university press.

Sydsæter, K., Hammond, P., Seierstad, A., & Strom, A. (2008). *Further mathematics for economic analysis*. Pearson education.