Universidad de la República
Facultad de Ciencias Sociales
DEPARTAMENTO DE ECONOMIA

Notas Docentes

Notas de programación no lineal

Elvio Accinelli

Nota Docente No. 16

Contents

Notas de programación no lineal

Elvio Accinelli [*]

Prefacio

Las presentes notas no pretenden ni mucho menos sustituir la lectura de un buen libro de texto, necesidad impostergable para todo estudiante que aspire a entender con rigor y profundidad la teoría que se desarrolla en la clase. Son más bien una guía, que debe ser acompañada por lecturas más profundas y por la realización de ejercicios. Como tal son escuetas y descarnadas, pretenden mostrar el espíritu general del curso de matemática y una guía para el estudio.

La justificación de la matemática en la economía es a esta altura del desarrollo de ambas ciencias ociosa. No obstante a modo de justificación para este trabajo hacemos el siguiente comentario.

La teoría económica aparece como generadora de problemas desafiantes para la matemática cuya resolución supone una avance para la teoría económica. Ciertamente no parece sensato justificar el estudio de la matemática en la posible resolución de triviales ejercicios de matemática disfrazados de economía presentados en el apéndice de un libro de matemáticas para economistas. Estas triviales aplicaciones, decontextualizadas, oscurecen la verdadera relación entre ambas ciencias y son ciertamente desalentadoras para el buen estudiante de ciencias económicas.

El autor de las notas es totalmente contrario a presentar la matemática para los estudiantes de ciencias económicas como una matemática de segundo nivel, lejos de esto cree que debe ser una matemática de profundo contenido y rigor, enseñada como se les enseña a los estudiantes de matemática, que ayude a realizar aplicaciones verdaderas que ultrapasen el nivel del ejercicio simplificado sin connotación verdadera en la Teoría Económica, es totalmente justificable que los estudiantes de ciencias económicas estudien matemática en la facultad de ciencias con

[*] Quisiera agradecer a Ramón Garcia-Cobbian y a Loretta Gasco por haberme invitado a participar de las actividades de la Sección de Matematicas de La Universidad Católica de Lima, durante los meses de abril, mayo y junio de 2004, período este, en el cual estas notas fueron escritas en su primera aproximación. Este agradecimiento es extensivo a profesores y funcionarios de dicha sección por el apoyo y hospitalidad que me brindaron. Naturalmente que los errores que en este trabajo persisten son de responsabilidad exclusiva del autor, y toda sugerencia u observación tendiente a superarlos y en general a mejorarar el trabajo serán agradecidas por el autor. E-mail: elvio@cueyatl.uam.mx

matemáticos profesionales. Ciertamente el programa elegido debe comprender aquellos tópicos de la Matemática, de los que la Teoría Económica moderna se vale con mayor intensidad, pero dichos tópicos deben ser estudiados con rigor y comprendidos en su profundidad sea para desarrollar la Teoría Económica dentro de su esquema actual, o sea para criticarla y presentar una alternativa valedera. Encarada la matemática en cualquier carrera propia de las ciencias económicas de manera trivial, como una matemática de poco nivel, da lugar a la pregunta ingenua y tendenciosa: *¿ para qué sirve la matemática en las ciencias económicas?*

El rechazo de la matemática como herramienta y muchas veces como fuente del pensamiento económico por parte de algunos economistas, en base a la afirmación de que la realidad escapa a cualquier modelo no es de recibo como no lo es en ninguna ciencia, pues la trivialidad de la afirmación es clara y de valor universal y se convierte en una mutilación de la posibilidad del pensamiento creador. La alta formalización de la Teoría Económica moderna y la ductilidad de la Matemática moderna permiten una simbiosis rica y fecunda en descubrimientos e interacciones. Por otra parte, la fecundidad de la interdisciplinariedad se pierde tanto cuando se rechaza a la matemática como instrumento y fuente del pensamiento económico, como cuando se hace matemática primero y luego se buscan las posibles aplicaciones, esta manera de concebir la interdisciplinariedad encorceta la realidad económica en un marco fijado apriori y por lo tanto empobrecido. La matemática se aplica para la resolución de un problema originado en la realidad económica, y luego formalizado por el pensamiento del economista y del matemático, llegado a este nivel el problema es un desafío que puede o no ser resuelto de manera completa, en forma más o menos inmediata, pero que en todo caso pasa si es de interés, al cuerpo de la teoría y de manera más o menos indirecta al terreno de las aplicaciones.

¿ Qué matemática o qué rama de la matemática debe debe enseñarse a los estudiantes de ciencias económicas ? o ¿ con qué profundidad debe dictarse un curso para dichos estudiantes ? son preguntas de gran interés, con respuestas diferentes en su contenido. Para la primera no hay una respuesta clara, pues en tanto que herramienta del descubrimiento no puede fijarse a priori, en todo caso podemos ayudarnos conociendo el tipo actual de problemas que la ciencia económica se plantea, los que dan una pauta de los problemas del porvenir. Para la segunda la respuesta es clara: con toda la profundidad que la matemática requiere para ser bien utilizada.

En base a la forma actual de la Teoría Económica elegimos la optimización como tema para estas notas, la justificación de tal elección la encontramos en la presentación de la economía y la administración como ciencias que se refieren a la asignación de recursos escasos con eficiencia, por ejemplo importantes temas de la microeconomía: el Equilibrio General, la Teoría del Agente, la Teoría de la Firma, y la Teoría de los Contratos se refieren de manera directa a este tema. Elegimos estos no por ser los únicos posibles sino por entenderlos básicos. En definitiva la Economía es la ciencia que estudia la asignación óptima de recursos escasos.

1 Introducción

El objetivo final de este trabajo es presentar el teorema de Khum-Tucker mostrando en el camino la herramienta necesaria para su demostración y cabal comprensión. Prestaremos especial atención a aquellos puntos que entendemos son de mayor importancia para quienes pretenden entender la teoría económica moderna, como por ejemplo el teorema de separación de convexos.

Comenzaremos estas notas con unas breves consideraciones sobre la topología de R^n y sobre convexidad. Las mismas pueden encontrarse en cualquier libro de topología o Análisis en R^n, como por ejemplo [Lima, L.]. Luego haremos algunas consideraciones sobre convexidad en R^n y finalmente abordaremos el teorema de Kuhn-Tucker y alguna de sus posibles aplicaciones en economía. Las referencias básicas son [Takayama, A.] y para todo lo referente al estudio de la convexidad [Rockafellar, T.].

1.1 R^n como un espacio lineal.

Sea R el conjunto de los números reales. Consideremos conjuntos de $n-uplas$ de números reales $x = (x_1, x_2, .., x_n)$. Estos elementos serán llamados vectores o puntos de R^n. Notaremos como R^n al conjunto de todas las n-uplas de números reales. Entonces R^n es el producto cartesiano formado por n réplicas de R. El i-ésimo elemento de x se llama i-ésima coordenada y la representamos por x_i.

Definimos en R^n la suma de dos elementos $x = (x_1, x_2, .., x_n)$, $y = (y_1, x_2, .., y_n)$, cualesquiera como la $n-upla$ z cuya $i-esima$ coordenada, resulta de la suma de las correspondientes de x e y. Definimos también la multiplicación de un número α por un vector x al elemento w de R^n obtenido mediante la multiplicación de cada coordenada de x por el número α. Es decir

$$z = (z_1, z_2, ..., z_n);\ siendo\ z_i = x_i + y_i\ \ y\ \ w = (w_1, w_2, ..., w_n)\ donde\ w_i = \alpha x_i \quad \forall i = 1, 2, ..., n.$$

Con las operaciones definidas de esta forma R^n es un espacio vectorial sobre R.

Ejercicio 1.1 *Verificar que $\{R^n, R, +, \langle \cdot, \cdot \rangle\}$ donde $+$ representa la suma en R^n y $\langle \cdot, \cdot \rangle$ el producto de vector por escalar, es un espacio vectorial.*

Dados dos elementos x e y de R^n definimos el producto de x por y como el número real que se obtiene en la siguiente operación:

$$\langle x,y\rangle = xy = \sum_{i=1}^{n} x_i y_i.$$

Obsérvese que el producto interno así definido es una función $P : X \times X \to R$.

Este producto se llama producto interno o escalar.

Ejercicio 1.2 *Verificar que el producto interno satisface las siguientes propiedades:*

- *(i)* $\langle x,y\rangle = \langle y,x\rangle$.
- *(ii)* $\langle(\alpha x + \beta y), z\rangle = \alpha\langle x,z\rangle + \beta\langle y,z\rangle$.
- *(iii)* $\langle x,x\rangle \geq 0$ *y* $\langle x,x\rangle = 0$ *si y sólo si* $x = 0$.

En general dado un espacio vectorial X se llama producto interno a una función $f : X \times X \to R$ que verifica las propiedades antriormente mencionadas. El caso particular en el que: $f(x,y) = \langle xy\rangle$ se denomina producto euclideano.

1.2 R^n como espacio métrico

A continuación definiremos la **función distancia.**

Definición 1.3 *Sea X espacio vectorial. Decimos que la función $d : X \times X \to R$ es una distancia si verifica que:*

- (i) $d(x,y) \geq 0$, con igualdad si y solo si $x = y$.
- (ii) $d(x,z) \geq d(x,y) + d(y,z)$
- (iii) $d(x,y) = d(y,x)$.

Un espacio vectorial X en el que hay definido una función distancia se llama espacio métrico y se representa por (X,d).

Ejercicio 1.4 *Verifique que las siguientes funciones de $R^n \times R^n \to R$ son funciones distancia.*

1. $d_1(x,y) = \begin{cases} 1 & x \neq y \\ \\ 0 & x = y \end{cases}$
2. $d_2(x,y) = \sum_{i=1}^{n} |x_i - y_i|$.

3. $d_3(x, y) = sup_i\{|x_i - yi|\}$.

4. $d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$ *esta distancia se suele llamar distancia usual. Para la demostración necesitará de la desigualdad de Cauchy-Schwarz, es decir que dados dos puntos de* R^n x *e* y *se verifica que* $(\sum_{i=1}^{n} x_i y_i)^2 \leq \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2$ *equivalentemente que* $(\langle x, y\rangle)^2 \leq \langle x, x\rangle\langle y, y\rangle$.

   Demostración de la propiedad triangular: Definamos $u = z - x$; $v = y - z$; $w = y - x$. *Observe que* $\langle u, u\rangle = d^2(x, z)$; $\langle v, v\rangle = d^2(y, z)$ *y que* $\langle w, w\rangle = d^2(y, x)$. *Queremos probar entonces que*

$$(\langle w, w\rangle)^{\frac{1}{2}} \leq (\langle u,\ u\rangle)^{\frac{1}{2}} + (\langle v,\ v\rangle)^{\frac{1}{2}}.$$

   Elevando al cuadrado ambos miembros de la desigualdad obtenemos que la desigualdad a probar es:

$$\langle w,\ w\rangle \leq \langle u,\ u\rangle + \langle v,\ v\rangle + 2(\langle u\ u\rangle\langle v,\ v\rangle)^{\frac{1}{2}}.$$

   Siendo $w = u + v$ *se sigue que*

$$\langle w,\ w\rangle = \langle u + v,\ u + v\rangle = \langle u,\ u\rangle\ + \langle v,\ v\rangle + 2\langle u,\ v\rangle.$$

   Por la desigualdad de Cauchy-Schwarz, $\langle u\ v\rangle \leq (\langle u,\ u\rangle\langle v\ v\rangle)^{\frac{1}{2}}$. *De donde se sigue la desigualdad triangular.*

Decimos que la métrica o distancia proviene de un producto interno si $d(x, y) = \sqrt{< \langle x - y, x - y\rangle >}$. Definiremos ahora la **función norma.**

Definición 1.5 *Sea* X *espacio vectorial. Una función* $N : X \to R$ *es un norma en* X *desde que dicha función verifique las siguientes propiedades:*

1. $N(x) \geq 0$ con igualdad si y solamente si $x = 0$.

2. $N(x) + N(y) \geq N(x + y)$.

3. $N(\alpha x) = |\alpha|N(x)$.

Si en X se define una norma N entonces (X, N) se llama espacio normado.

Ejercicio 1.6 *Sea* X *un espacio vectorial,*

- *Muestre que todo espacio normado es un espació métrico,*

- *pero que el recíproco no es cierto.*

- *¿ qué propiedades debe verificar d para ser una norma?*

  1. *Muestre que con la distancia usual* $\|x\| = d(x, 0)$ *define una norma.*
  2. *Verifique que* $\|x\| = d(x, 0) = \sqrt{\langle x, x\rangle}$
  3. *Define la distancia* d_1 *del ejercicio (1.4) una norma?*

Sea $\{X, d\}$ un espacio métrico. Consideremos $x_0 \in R^n$ y definamos el conjunto $B_r(x_0)$ por

$$B_r(x_0) = \{x : x \in R^n, d(x, x_0) < r\},$$

donde r es algún número positivo y d se refiere al al distancia

Ejercicio 1.7 *Siendo d la distancia definida en el ejercicio (1.4), represente gráficamente* $B_r(0)$ *la correspondiente bola centrada en el cero de radio r.*

Decimos que un conjunto A en $\{R^n, d\}$ está acotado si existe $r \geq 0$ tal que $A \subseteq B_r(0)$. En general decimos que un subconjunto A de un espacio métrico o normada está acotado si existe $r > 0$ tal que $A \subset B_r(0)$.

1.3 R^n Como espacio topológico

Definición 1.8 Topología *Dado un conjunto cualquiera X decimos que una familia* $\mathcal{F}$ *de subconjuntos de X es una topología para X si*

1. $X \in \mathcal{F}$ *y* $\emptyset \in \mathcal{F}$
2. $V_i \in \mathcal{F}$ $i = 1, 2, ..., n$ *entonces* $\cap_{i=1}^{n} V_i \in \mathcal{F}$.
3. $V_\alpha \in \mathcal{F}$ $\forall \alpha \in A$ *entonces* $\cup_{\alpha \in A} V_\alpha \in \mathcal{F}$.

de Llamaremos a $\{X, \mathcal{F}\}$ *espacio topológico.* Observe que un subconjunto arbitrario de un espacio topológico no tiene por qué ser abierto o cerrado.

Como ejemplo de topologías extremas definibles en X, considere la formada solamente por el conjunto total y el vacío y por otra parte la formada por todos los subconjuntos posibles del conjunto X.

Los elementos de la familia $\mathcal{F}$ son llamados conjuntos abiertos, mientras que sus complementos se llaman conjuntos cerrados. Llamaremos entorno de un punto $x \in X$ a todo conjunto abierto que contenga a x. Diremos entonces que un conjunto A es un abierto si y solamente si es un entorno de cada uno de sus puntos. Obviamente en un mismo conjunto X pueden definirse distintas topologías. Decimos que la topología $\mathcal{F}'$ es más fina que la topología $\mathcal{F}$ si $\mathcal{F} \subset \mathcal{F}'$ No todas las topologías son comparables.

Definición 1.9 *Llamaremos* base de una topología $\mathcal{F}$ *en* X *a una subfamilia* $\mathcal{B}$ *de* $\mathcal{F}$ *tal que todo elemento* U *de* $\mathcal{F}$ *es unión de elementos de* $\mathcal{B}$. *Y llamaremos* subbase de una topología $\mathcal{F}$ *en* X *a una subfamilia* $\mathcal{S}$ *de* $\mathcal{F}$ *tal que la colección de todas las intersecciones finitas de elementos de* $\mathcal{S}$ *forman una base para la topología* $\mathcal{F}$.

Un espacio se dice satisfacer el primer axioma de numerabilidad si tiene una base numerable.

Definición 1.10 *Una base del sistema de entornos o base local en* x *es una familia de entornos de* x *tal que todo entorno del punto contiene un elemento de la familia.*

Un espacio se dice satisfacer el *primer axioma de numerabilidad* si tiene una base de entornos numerable. Mientras que si existe una base numerable se dirá que verifica el *segundo axioma de numerabilidad.*

Topologías definidas a partir de métricas y normas:

En un espacio métrico (X, d) es posible definir una *topología a partir de la métrica,* para esto procedemos de la siguiente manera. Consideremos los conjuntos $B_r(x_0) = \{x \in X : d(x, x_0) < r\}$ cada uno de estos será llamado bola abierta con centro x_0 y radio r. Definiremos la familia de abiertos de R^n como aquellos subconjuntos A de R^n tales que para todo $x \in A$ existe r tal que $B_r(x) \subseteq R^n$. Decimos que esta topología es generada por la métrica d.

Ejercicio 1.11 *Verifique que esta familia de conjuntos forma una topología.*

Es también posible definir una topología en un espacio normado (X, N) a partir de la norma N allí definida. Esta será la *topología definida por la norma.* Basta para esto considerar como abiertos en X aquellos subconjuntos A tales que para todo $x \in A$ existe $V_x = x + B_0(\epsilon)$ tal que $V_x \subseteq A$. Donde $B_0(\epsilon) = \{y \in X : N(y) < \epsilon\}$ es la bola abierta de centro cero y radio ϵ, definida por la norma N, de esta forma $V_x = \{z \in R^n : z = x + y,\ y \in B_0(\epsilon)\}$ define un entorno de x.

Se tiene entonces que la familia de las bolas abiertas $\mathcal{B}(x_0, r)\ \ \forall x_0 \in X,\ r \in R$ definen una base de la topología. Mientras que para cada $x \in X$ $\mathcal{B}(x_0, r)\ \ r \in R$ define una base loscal de x. Se concluye entonces que los espacios métircos y topológicos satisfacen el priemer axioma de numerabilidad.

Observación 1.12 *Todas las topologías en* R^n *generadas por normas definen los mismos abiertos*[1].

[1]Existe un teorema que generaliza esta afirmación, mismo dice que en R^n toda las topologías lineales que son Hausdorff definen los mismos abiertos Este teorema se encuentra en textos avanzados de topología general, no obstante daremos en el apéndice I de la sección (13.1) una demostración. Muestra que casi toda las topologías comunmente usadas en R^n son las mismas, lo que deja de ser válido en espacios de dimensión infinita.

Esta afirmación resulta del hecho de que en R^n todas las normas son equivalentes [2]. Es decir que, dadas dos normas N_1 y N_2 existen unas constantes A y B tales que $AN_1(x) \leq N_2(x) \leq BN_1(x)$.

Ejercicio 1.13 *Demuestre que las topologías generadas por normas equivalentes, definen los mismos abiertos.*

Ejercicio 1.14 *Verifique las siguientes afirmaciones:*

(i) El conjunto vacío y el total son cerrados.

(ii) Una unión finita de conjuntos cerrados es un conjunto cerrado.

(iii) La intersección arbitraria de conjuntos cerrados es cerrada.

(iv) Sea $\mathcal{F} = \{\emptyset, X\}$ entonces $\{X, \mathcal{F}\}$ es espacio topológico. La topología así definida se llama trivial.

(v) Considere $\mathcal{F}$ como el conjunto de todos los puntos de X entonces $\{X, \mathcal{F}\}$ es espacio topológico. La topología así definida se llama discreta.

Considere $\{R^n, d\}$ donde por d representaremos la distancia euclideana. Consideremos la topología formada por los siguientes abiertos: *un conjunto A es un abierto si y solamente si para cada punto $x \in A$ existe una bola abierta de radio ϵ totalmente contenida en A*. Decimos que esta es la *tpoología habitual*. Toda topología inducida por una norma en R^n es por la anteriormente visto equivalente a ésta. Por lo que todo lo dicho con respecto esta toplogía vale para cualquier otra que sea Hausdorf y lineal en R^n.

Ejercicio 1.15 *Probar que si Y es un subconjunto de un espacio topológico $\{X, \mathcal{F}\}$ la colección de $\mathcal{F}_\gamma$ de subconjuntos de Y definida por*

$$\mathcal{F}_\gamma = \{V \cap Y : V \in \mathcal{F}\}$$

es una topología en Y.

[2]Para ver la equivalencia de las normas consideremos N_2 como una norma cualquiera N, mientras que sea N_1 la norma del supremo $N_s(x) = max\{x_1, ..., x_n\}$. Observe que si $e_i, i = 1, .., n$ son los vectores de la base canónica de R^n la siguiente cadena de desigualdades se sigue: $N(x) \leq \sum_{i=1}^n |x_i| N(e_i) \leq BN_s(x)$, siendo $B = \sum_{i=1}^n N(e_i)$. Se sigue entonces que $|N(x) - N(y)| \leq N(x - y) \leq BN_s(x - y)$. Por lo tanto $N : R^n \to R$ es una función continua en (X, N_s) y por lo que alcanza su máximo y su mínimo en el conjunto compacto $V = \{x \in R^n : N_s(x) = 1\}$ es decir existen A y B tales que $A \leq N(x) \leq B$, $\forall x \in V$. Luego escribiendo $y = \frac{y}{N_s(y)} N_s(y)$ se sigue que $AN_s(y) \leq N(y) = N(\frac{y}{N_s(y)}) N_s(y) \leq BN_s(y)$.

La topología formada de esta manera se llama *topología relativa o inducida* por $\mathcal{F}$ en Y.

Definición 1.16 *Sea $f : X \to Y$ función continua entre los espacios topológicos (X, τ_x) y (Y, τ_y). Decimos que f es continua si y solamente si, la preimagen por f de todo abierto en (Y, τ_Y) es un abierto en (X, τ_x).*

Es importante recordar las siguientes definiciones:

Definición 1.17 *Sea (X, τ) espacio topológico y sea $K \subset X$, .*

1. *Decimos que $x \in X$ es un* punto de acumulación *de K si y solamente si en todo entorno suyo hay un punto de K distinto de x.*

2. *Llamaremos* clausura *K al conjunto formado por la intersección de todos los cerrados que lo contienen, notaremos a este conjunto como $\bar{K}$.*

3. *Decimos que* la sucesión $x_n \in X$ converge a $x \in X$ *si y solamente si x es el único punto de acumulación de la sucesión.*

Definición 1.18 *Un espacio vectorial, en el que hay definido una topología se dice ser un* **espacio vectorial topológico,** *abreviadamente e.v.t.*

1.4 Conjuntos compactos y conexos

A continuación haremos una breve introducción de los conceptos de compacidad y conexidad. Entendemos que el lector tiene cierta familiaridad con estos conceptos así como con el continuidad de funciones.

Definición 1.19 *Decimos que K es* compacto *si para todo cubrimiento por abiertos de K existe un subcubrimiento finito.*

En R^n compacto es equivalente a acotado y cerrado, esto es el teorema de Heine-Borel. Esta propiedad no se mantiene necesariamente en espacios más generales que R^n con la topología usual. No obstante para espacios métricos vale la siguiente definición, equivalente a la definición (1.19).

Definición 1.20 *Un conjunto K es compacto si y solamente si de toda sucesión de elementos de K tiene una subsucesión convergente.*

Definición 1.21 *Decimos que K es un conjunto* **conexo** *si y solamente si cada vez que $K \subset A \cup B$ siendo A y B abiertos en la topología relativa, y $A \cap B = \emptyset$ entonces $A = K$ y $B = \emptyset$ o bien $B = K$ y $A = \emptyset$.*

Obsérvese que X es conexo sii los únicos conjuntos de τ abiertos y cerrados a la vez, son el total y el vacío.

Puede probarse que en R los únicos conjuntos conexos son los intervalos[3].

Si bien no equivalente, a nuestra definición de conexidad, la noción de *conexidad por caminos* ayuda a entender la idea de conexión en general. Se dice que un conjunto X es *conexo por caminos* cuando, dados dos puntos a y b en X, existe un *camino*[4] en el conjunto que los une. Si bien no todo conjunto conexo de R^n es conexo por caminos, es cierto que en R^n un conjunto abierto es conexo si y solamente si es conexo por caminos. Ver [Lima, L.].

El siguiente concepto permitirá una caracterización muy útilizada de conjuntos compactos.

Definición 1.22 *Decimos que una familia $\mathcal{F}$ de subconjuntos de X tiene la* propiedad de intersección finita *(PIF) si y solamente si, se verifica que todo subconjunto finito de elementos de la familia $\mathcal{F}$ tiene intersección no vacía.*

Una importante aplicación de esta propiedad es el siguiente teorema:

Teorema 1.23 *Sea (K, τ) espacio topológico, K es compacto si y solamente si toda familia de cerrados con la PIF tiene intersección no vacía.*

Demostración: Sea K compacto y sea τ_α $\alpha \in A$ una familia de subconjuntos cerrados con la PIF. Supongamos que la intersección es vacía. Por lo tanto se verifica que $\bigcup_{\alpha \in A} \tau_\alpha^c = K$. Por ser K compacto y τ_α^c abiertos, existe una subfamilia F finita de elementos de A τ_β $\beta \in F$, tal que $\bigcup_{\alpha \in F} \tau_\alpha^c = K$. Luego $\bigcap_{\alpha \in F} \tau_\alpha = \emptyset$, lo que contradice que τ_α $\alpha \in A$ tiene la PIF. *Reciprocamente:* Supongamos que toda familia de cerrados en K con la PIF tiene intersección no vacía. Si K no es compacto, existe un cubrimiento de K, $A_\alpha; \alpha \in A$ de abiertos, tal que ninguna subfamilia finita $F \subset A$ es un subcubrimiento de K, es decir que $\bigcup_{\alpha \in F} A_\alpha^c \neq \emptyset$. Esto supone que $\forall F \subset A$, F finita, $(\bigcap_{\alpha \in F} A_\alpha)^c \neq \emptyset$. Como A_α^c es una familia de cerrados con la PIF se sigue que la interseccón

[3] Entendiendo por intervalos un subconjuntos I de R formado por todos los puntos de la recta comprendidos entre dos reales dados, el intervalo será abierto, cerrado o semiabiero, según contenga o no a los extremos a y b, o a uno de ellos solamente. Obviamente si un subconjunto C de R no es un intervalo, existen nmeros $a < b$ en C y $c \in C^c$ tales que $a < c < b$ luego el conjunto C puede escribirese como unión de intervalos abiertos, disjuntos, que no son el vacío ni todo x. Para el recíproco ver por ejemplo [Lima, L.].

[4] Se entiente acá por camino una función $f : I \to X$ donde I es un intervalo en R.

$\bigcap_{\alpha \in A} A_\alpha^c \neq \emptyset$, luego tomando complementos en ambos lados, se sigue que: $\bigcup_{\alpha \in A} A_\alpha \neq K$ lo que contradice la afirmación inicial.[$\cdot$]

A partir de esta caracterización de los conjuntos compactos es posible demostrar la existencia de la demanda considerando solamente consumidores con preferencias semicontinuas superiores, ver [Accinelli, E.].

1.5 Continuidad de funciones

Introduciremos en esta sección algunas consideraciones sobre las funciones continuas concepto que entendemos familiar al lector de estas notas. Dedicaremos un poco más de atención a a los conceptos de semicontinuidad superior e inferior, quizás menos conocidos por el lector. Los teoremas tratados en esta sección tienen importantes aplicaciones no sólo en el campo de la matemática, sino también en el de la ecponomía. El primer teorema que consideraremos afirma que la imagen de un conjunto compacto por una función continua es un compacto y el segundo lo análogo para conjuntos conexos. Ambos son la base para del teorema de Weierstrass, el que los economistas usan frecuentemente en muchas de sus aplicaciones en las que afirman la existencia de la demanda.

Definición 1.24 *Entendemos acá por* **función continua,** *a una función f entre dos espacios topológicos $\{X, \tau_X\}$ y $\{Y, \tau_Y\}$ tal que la preimagen por f de un conjunto A abierto en Y es un conjunto $B = f^{-1}(A)$ abierto en X.*

Observe que distintas funciones pueden ser continuas o dejar de serlo según las topologías definidas en los espacios involucrados. Afortundamente, las topologías generadas por normas equivalentes, son equivalentes, es decir, contienen los mismos abiertos. Esto hace que en R^n las funciones continuas no cambian con la elección de la norma.

Este concepto de continuidad es equivalente al quizás más conocido para el lector, usado corrientemente para funciones de R^n en R^n y generalizable para funciones entre espacios métricos. Descripta en términos de epsilones y deltas, esta definición es la siguiente:

Definición 1.25 *Sean (X, d_x) y (Y, d_y) espacios métricos. Una función $f : X \to Y$ es continua en $x_0 \in X$, si y solamente si, para todo $\epsilon > 0$ existe $\delta > 0$ tal que, para todo $x \in X$ que cumpla que $d_x(x, x_0) < \delta$ entonces se verifica que $d_y(f(x), f(x_0)) < \epsilon$. Luego f es continua en X si es continua para todo punto $x \in X$.*

Como anteriormente de la topología, la continuidad de una función depende de las métricas definidas en los espacios relacionados por la función. La definición de continuidad (1.25) es equivalente a la defención (1.24) si consideramos a X e Y como espacios topológicos con la topología inducida por la métrica.

Teorema 1.26 *Sean (X, τ_x) e (Y, τ_y) espacios topológicos. Sea $f : K \to Y$ siendo $K \subset X$ compacto y f continua. Entonces $f(K)$ es compacto.*

Teorema 1.27 *Sean (X, τ_x) e (Y, τ_y) espacios topológicos. Sea $f : K \to R$ siendo $K \subset X$ conexo f continua. Entonces $f(K)$ es conexo.*

Las demostraciones de estos teoremas se hacen por el absurdo, y las suponemos conocidas por el lector. Pueden verse en [Lima, L.].

Como consecuencia de los dos teoremas anteriormente mencionados se obtienen las siguientes propiedades para funciones reales continuas expresadas en los teoremas siguientes:

Teorema 1.28 (Teorema de Weierstrass) *Sean (X, τ_x) e (Y, τ_y) espacios topológicos. Toda función real continua definida en un conjunto compacto $K \subset X$ alcanza su máximo y su mínimo.*

Demostración: Siendo $f(K)$ compacto en R es acotado y cerrado. Considere entonces una sucesión $f(x_n)$ convergente al supremo (s) de f. Como K es compacto la sucesión $\{x_n\}$ tiene una subsucesión convergente. Sea esta $\{x_{n_i}\}$ y sea x su límite. Por la continuidad de f $f(x_{n_i})$ converge a $f(x) = s.[\cdot]$

Para demostrar que el ínfimo se alcanza repetimos la demostración anterior pero para $-f(x)$.

Teorema 1.29 (Teorema del valor intermedio) *Sea (X, τ_x) un espacio topológico. Sea $f : D \to Y$ función continua, $D \subset X$ conexo. Sean $y_1 = f(x_1)$ y $y_2 = f(x_2) = y_2$ donde $y_1 \leq y_2$ siendo x_1 y x_2 puntos en D. Entonces para todo $y : y_1 \leq y \leq y_2$ existe $x \in D$ tal que $f(x) = y$.*

Demostración: Es corolario del hecho de que funciones continuas transforman conjuntos conexos en conexos y que los únicos conjuntos conexos en R son los intervalos. $[\cdot]$

En muchos problemas importantes de la teoría económica es posible reemplazar el supuesto de un comportamiento modelado por funciones continuas por el más general representado por funciones semicontinuas. A continuación definiremos los conceptos de semicontinuidad superior e inferior sin entrar en detalles. No obstante mostraremos que la caracterización de los compactos por la PIF, permite extender el teorema de Weierstrass a funciones semicontinuas superiores (s.c.s) y semicontinuas inferiores (s.c.i).

Definición 1.30 *Sea $f : X \to R$, siendo X un espacio vectorial topológico.*

f es **semi continua superior** *(s.c.s.) en X si y solamente si, para todo $c \in R$, el conjunto $C = \{x \in X : f(x) \geq c\}$ es cerrado en X.*

f es **semi continua inferior** *(s.c.i.) en X si y solamente si para todo $c \in R$, el conjunto $C = \{x \in X : f(x) \leq c\}$ es cerrado en X.*

El siguiente teorema caracteriza a las funciones semicontinuas y puede encontrarse en cualquier texto de análisis. Recordamos que en términos de sucesiones se tienen las siguientes definiciones:

$$\liminf_{\alpha} f(x_\alpha) = \sup_{\alpha} inf_{\beta \geq \alpha} f(x_\beta)$$

mientras que:

$$\limsup_{\alpha} f(x_\alpha) = \inf_{\alpha} sup_{\beta \geq \alpha} f(x_\beta).$$

Equivalentemente en términos de entornos:

$$\lim \inf_{x \to x^*} f(x) = \sup_{\delta > 0} \left[inf \left\{ f(x) \right\} : x \in B_{x^*}(\delta) \cap X \right]$$

mientras que:

$$\lim \sup_{x \to x^*} f(x) = \inf_{\delta > 0} \left[sup\{f(x)\} : x \in B_{x^*}(\delta) \cap X \right].$$

Teorema 1.31 *Sea $f : X \to R$ X espacio vectorial topológico. Las siguientes afirmaciones son equivalentes:*

1. f es s.c.s. (resp. s.c.i).

2. lim $\sup_\alpha f(x_\alpha) \leq (resp. \geq) f(x)$, cuando $x_\alpha \to x$.

3. Para cada $\epsilon > 0$ existe $\delta > 0$ tal que si $x \in B_{x^*}(\delta) \cap X$ entonces $f(x) < f(x^*) + \epsilon$ (resp. $f(x) > f(x^*) - \epsilon$).

Este teorema admite una interpretación intuitiva: *una función s.c.s. no tiene saltos hacia abajo, mientras que una función s.c.i. no tiene saltos hacia arriba.* Esto debe ser interpretado cuidadosamente a la luz de diferentes ejemplos.

Demostración:

(3) $\to$ (1) Sea $c = f(x^*) + \epsilon$, y sea $L_c = \left\{ x \in X : f(x) < c \right\}$. Por (3) existe $B_{x^*}(\delta) \subset L_c$, luego $C = (L_c)^c$ es cerrado.

(1) $\to$ (3) Es inmediato.

(1) $\to$ (2) Sea $z = \lim\ sup_{x \to x^*} f(x)$, supongamos que no sea cierto que $\lim\ sup_{x \to x^*} f(x) \leq f(x^*)$. Existe entonces $\epsilon^* > 0$ tal que $f(x*) + 2\epsilon^* < z$:

- (i) Si $z = \infty$ entonces por la definición de límite superior, en toda bola de centro x^+ y radio δ existe x tal que $f(x) > f(x*) + 100$, luego no se cumpliría (3).
- (ii) Si $z \in R$ entonces en toda bola de centro x^* y radio δ existe x_δ tal que $f(x_\delta) > z - \epsilon^*$ lo que contradice (3).

(2) $\to$ (3) Como por (2) se verifica que $f(x^*) \geq \lim\ sup_{x \to x^*} f(x)$, por la definición de límite superior, se tiene que existe $\delta' > 0$ tal que para todo $0 < \delta < \delta'$ $\lim\ sup_{x \to x^*} f(x) = sup\{f(x) : x \in B_{x^*}(\delta) \cap X\} < f(x^*) + \epsilon$. Luego $f(x) < f(x^*) + \epsilon, \forall x \in B_{x^*}(\delta)$.[]

Ejercicio 1.32 *(i) Sea X el intervalo unidad en R. y sea $f : [0,1] \to R$*

$$f(x) = \begin{cases} 0 & 0 \leq x \leq 1 \\ 1 & x = 1. \end{cases}$$

Es s.c.s. pero no s.c.i. en $x = 1$.[]

(ii) Como puede verse fácilmente, la suma de funciones s.c.s. (resp. s.c.i.) es s.c.s. (resp. s.c.i.), pero no es cierta la afirmación análoga para el producto.

El siguiente teorema muestra la existencia de máximos (resp. mínimos) para funciones s.c.s. (resp. s.c.i.) definidas en compactos.

Teorema 1.33 *Sea $f : K \to R$, donde K es un subconjunto compacto de R^n, sea f s.c.s, (resp s.c.i) entonces el conjunto de puntos donde f alcanza el máximo (resp. mínimo) es no vacío y compacto.*

Demostración: Sea $c \in f(K)$, como f es s.c.s $F_c = \{x \in K : f(x) \geq c\}$ es un subconjunto de K cerrado y por lo tanto compacto, y no vacío. La familia F_c con $c \in f(K)$, tiene la PIF. Por lo tanto $\cap_{c \in f(K)} F_c \neq \emptyset$, y compacto. Obsérvese que si $y \in K$ es un máximo para f es decir $f(y) \geq f(x)\ \forall\ x \in X$ entonces $y \in \cap_{c \in f(K)} F_c$ y reciprocamente.[]

Una interesante aplicación a la economía es la demostración de que toda relación de preferencia $\succeq$ definida en un conjunto de consumo X compacto, s.c.s, tiene elemento maximal. Recordamos que $\succeq$ es s.c.s si y solamente si $\{a \in X : a \succeq b\}$ es cerrado para todo $b \in X$. La demostración de esta afirmación es análoga a la hecha para el teorema anterior. De esta forma se prueba que la demanda de un consumidor con una relación de preferencias s.c.s. en una restricción presupuestaria compacta existe y define un conjunto compacto.

2 Conjuntos convexos y algunas de sus propiedades

Muchas de las consideraciones que haremos a continuación son válidas en general en espacios vectoriales topológicos. Es decir en conjuntos a los que se les adjunta una estructura de espacio vectorial, necesaria para definir combinaciones lineales de sus elementos y una topología, necesaria para definir entornos de sus puntos. No obstante en esta presentación, los resultados más importantes estarán referidos siempre a espacios vectoriales de dimensión finita, más aun, nos remitiremos siempre a R^n y toda consideración sobre métrica será referida a la métrica euclidiana y toda consideración topológica a la generada por esta métrica. Si bien el tema central de esta sección es de de las propiedades de los conjuntos convexos, comenzaremos con algunas consideraciones sobre conjuntos afines y conos, necesarias para el análisis convexo. Un tratamiento desarrollado del tema puede encontrarse en [Rockafellar, T.].

La siguiente definición introduce notación necesaria para esta y las siguientes secciones.

Definición 2.1 *Se define la* **suma de dos conjuntos** A **y** B *de* R^n *como el conjunto*

$$C = A + B = \{z \in R^n : z = x + y : x \in A, y \in B\}$$

Siendo C *subconjunto de* R^n *y* $\alpha \in R$ *se define el conjunto* **producto**

$$\alpha C = \{z \in R^n : z = \alpha x, x \in C\}.$$

La consideración de las propiedades que definen a los conos y conjunto afines, son como veremos, de gran importancia para el estudio de la optimización convexa.

Definición 2.2 *Un conjunto* C *es un* **cono** *si para cada* $x \in C$ *y* $\lambda > 0$ *se tiene que* $\lambda x \in C$.

Definición 2.3 *Un conjunto* $A \subset R^n$ *es llamado* **afín** *si cada vez que* $x \in A$ *e* $y \in A$ *la linea recta que los une* $\lambda x + (1 - \lambda)y$, $\lambda \in R$ *pertenece a* A.

Puede verse facilmente que un subespacio vectorial es un conjunto afín cuya particularidad es que el cero pertenece al conjunto. Recíprocamente si A es un conjunto afín y $0 \in A$ entonces A es un subespacio vectorial

Diremos que dos conjuntos afines M y A son paralelos si existe $a \in A$ tal que $A + a = M$. Se verifica que si a y b pertenecen a A entonces $A + a$ y $A + b$ son paralelos. De esta forma puede definirse la clase de equivalencia de todos los conjuntos afines paralelos a uno dado.

Obsérvese que si A es un conjunto afín entonces

$$L = A - A = \{x - y : x \in A, y \in A\}$$

es un subespacio vectorial paralelo a A. A es paralelo a un único subespacio L, $A = x^* + L$, donde x^* es un elemento arbitrario de A. [5]

Definición 2.4 *Definimos la* **dimensión de un conjunto afín** *como la dimensión del único subespacio paralelo al conjunto. Un conjunto afín de R^n de dimensión $n-1$ es llamado un* **hiperplano.**

Todo hiperplano define dos semiespacios E_1 y E_2 que serán llamados cerrados o abiertos según contengan o no al hiperplano que los define.

Obsérvese que un hiperplano H en R^n queda determinado por un vector $\in R^n$ $p \neq 0$ y un real α siendo

$$H_{p,\alpha} = \{x \in R^n : \langle p, x \rangle = \alpha\}.$$

De esta manera quedarán determinados los semiespacios cerrados $E_1 = \{x \in R^n : \langle p, x \rangle \geq \alpha\}$ y $E_2 = \{x \in R^n : \langle p, x \rangle \leq \alpha\}$, los que serán llamados abiertos si las desigualdades se reemplazan por desigualdades estrictas. El hiperplano H es paralelo al subespacio vectorial

$$S = \{x \in R^n : \langle p, x \rangle = 0\}$$

es inmediato verificar que $H = x^* + S$ siendo $x^* = \frac{\alpha}{\langle p,p \rangle} p$.

Definición 2.5 *Un subconjunto C de un espacio vectorial X se dice convexo si*

$$\forall x \in C,\ \forall y \in C \ y \ \forall t \in [0,1] : tx + (1-t)y \in C.$$

Definición 2.6 *Decimos que x en una combinación convexa de elementos de C si*

$$x = \sum_{i=1}^{n} t_i x_i \ \ siendo \ \ \sum_{i=1}^{n} t_i = 1,$$

para $n \in N$ con $x_i \in C$, y $0 \leq t_i \leq 1; \forall i = 1, 2, ..., n$.

Sea X un espacio ventorial topológico.

Teorema 2.7 *Dado $C \subset X$. C es convexo si y sólo si contiene todas las combinaciones convexas de C.*

[5] En efecto, si H y K son subespacios afines, entonces $\alpha H + \beta K$ es un espacio afín, para cualesquiera α y β reales. En el caso $0 \in L$ por lo tanto L es un SEV. Sea $x \in x^* + L$ entonces $x = x^* + (y - z)$, $y, z \in A$. Como $\frac{1}{2}x^* + y \in A$ y además $\frac{1}{2}x^* - z \in A$ se sigue que $x \in A$. Por lo tanto $x^* + L \subset A$. Por otra parte, si $x \in A$ entonces de que $x - x^* \in L$ se sigue que $x = x * + (x - x^*) \in x^* + L$ es decir $A \subset x^* + L$[]

Demostración:

(i) La definición de conjunto convexo implica que si un conjunto contiene todas sus combinaciones convexas entonces es convexo, pues en particular contiene las de dos elementos.

(ii) Recíprocamente: Por ser C convexo entonces toda combinación convexa de dos elementos de C está en C. Razonando por inducción completa. Consideremos $x_i \in C; i = 1, 2, ..., (n+1)$, sea

$$x = \sum_{i=1}^{n+1} t_i x_i = t \sum_{i=1}^{n} \frac{t_i}{t} x_i + (1-t)x_{n+1}$$

donde $\sum_{i=1}^{n+1} t_i = 1$ y definamos $t = \sum_{i=1}^{n} t_i$. Luego por la hipótesis de inducción se sigue que $y = \sum_{i=1}^{n} \frac{t_i}{t} x_i \in C$. Finalmente, por la convexidad asumida $x = ty + (1-t)x_{n+1} \in C$. []

Definición 2.8 *Un conjunto C es un* **cono convexo** *si es un cono que además verifica que cada vez que $x, y \in C$ entonces $x + y \in C$.*

Importantes ejemplos de conjuntos convexos son:

1. Los semiespacios cerrados, es decir los conjuntos de la forma

$$\{x \in R^n : \langle x, b \rangle \leq \beta\}, \quad \{x \in R^n : \langle x, b \rangle \geq \beta\},$$

donde $b \in R^n$ es un vector no nulo y $\beta \in R$. Si las desigualdades son reemplazadas por desigualdades estrictas los semiespacios son abiertos, y obviamente también son conjuntos convexos.

2. Las conjuntos de la forma $B_r(x)$.

3. Las rectas, los segmentos de recta, los planos.

Decimos que un subconjunto C de un espacio vectorial topológico es estrictamente convexo, si para todo par de elementos x, e y de C y para todo $\lambda \in (0, 1)$ se verifica que, $\lambda x + (1-\lambda)y \in C^0$

Lema 2.9 *Sea X un espacio vectorial. Entonces las siguientes propiedades se verifican:*

1. Suma de dos conjuntos convexos es un conjunto convexo.

2. Multiplicando un conjunto convexo por un escalar se obtiene un conjunto convexo

3. C es convexo si y solamente si $\alpha C + \beta C = (\alpha + \beta)C$.

4. La calusura y el interior de un convexo, son conjuntos convexos.

Demostración: Probaremos solamente la segunda parte del último item. Las afirmaciones restantes quedan como ejercicios. Sea C^0 el interior de C. Considere $z \in \alpha C^0 + (1-\alpha)C^0$ por lo que $z = \alpha x + (1-\alpha)y$ con x e $y \in C^0$. Luego existe ϵ tal que $x + \epsilon B \in C$ y $y + \epsilon B \in C$. Siendo B la bola de radio 1 y centro 0, la convexidad de B escribir $\alpha\epsilon B + (1-\alpha)\epsilon B = \epsilon B$. Luego $(\alpha)[x+\epsilon B] + (1-\alpha)[y+\epsilon B] = z + \epsilon B \subset C$, por lo que $z \in C^0$. Se tiene que $\alpha C^0 + (1-\alpha)C^0 \subset C^0$ luego C^0 es convexo. []

Para la demostración de que la clausura de un convexo es convexa, ver ejercicio en la hoja de práctico 4[6]

Ejercicio 2.10 *1. Demuestre que si I es un conjunto de índices cualesquiera y $\{C_i\}$ una colección de conjuntos convexos de R^n entonces la intersección $\cap_{i\in I} C_i$ es un conjunto convexo.*

2. Sean C_1 y C_2 conjuntos convexos de R^n, entonces

$$C = \alpha C_1 + \beta C_2$$

es un conjunto convexo.

Definición 2.11 *Dado un conjunto $S \subset R^n$ no vacío, diremos que la* **cápsula convexa** *de S, la cual denotaremos como $co(S)$, es la intersección de todos los conjuntos convexos que contienen a S.*

Teorema 2.12 *La $co(S)$ es el conjunto de todas las combinaciones convexas de S.*

Demostración: Sea D el conjunto de todas las combinaciones convexas de S, entonces $S \subseteq D$, siendo D convexo, entonces $co(S) \subseteq D$. Recíprocamente, sea $y \in D$ por lo tanto y es una combinación convexa de elementos de S y por lo tanto $D \subseteq co(S)$. []

Teorema 2.13 *Para conjuntos convexos no vacíos $A_1, A_2, ..., A_n$ en un espacio vectorial topológico se verifica que*

$$co(\cup A_i) = \left\{ \sum_{i=1}^{n} \lambda_i x_i,\ \lambda_i \geq 0;\ x_i \in A_i,\ \sum_{i=1}^{n} \lambda_i = 1 \right\}.$$

En particular si cada A_i es compacto, entonces $co(\cup A_i)$ es compacto.

[6] Sean x e y elementos de la clausura de C. Luego para todo $\epsilon > 0$, existen $x' \in (x + \epsilon B) \cap C$ e $y' \in y + \epsilon B \cap C$ tales que por ser C convexo, su combinación convexa pertenece a C y además $\alpha x' + (1-\alpha)y' \in \alpha x + (1-\alpha)y + \epsilon B$. Luego $\alpha x + (1-\alpha)y \in \bar{C}$.[]

Demosración: Sea $x \in co(\cup A_i)$ entonces existen $x_i \in \cup_{i=1} A_i$ tales que x es una c.c. de estos elementos. Es decir

$$x = \sum_{i=1}^{m} \lambda_i x_i, \ \ \sum_{i=1}^{m} \lambda_i = 1, \ \lambda_i \geq 0, \ \ \forall \ i = 1, 2, ..., m.$$

Supongamos que la c.c. está compuesta por h_i elementos pertenecientes a A_i. Definamos t_i igual a la suma de los h_i coeficientes correspondientes a los elementos de A_i Escribamos x ubicando los elementos de cada A_i uno a continuación de otro comenzando con los corrsespondientes al A_i con menor i, sea este j. Llamemos a los elementos de j presentes en la c.c. $x_1, ..., x_{h_1}$, donde $h_1 = h_j$, y reescribmos a t_j como t_1. Sea k, el segundo menor i presente en la c.c. luego de reordenarlos si fuera preciso, tendremos: $x_{h_1+1}, ..., x_{h_2}$ donde $h_2 = h_1 + h_k$ renombramos como t_2 a t_k. Continuamos de esta manera. Así x puede escribirse entonces en la forma:

$$x = t_1 \sum_{i=1}^{h_1} \frac{\alpha_i}{t_1} + t_2 \sum_{i=h_1}^{h_2} \frac{\alpha_i}{t_2} + ... + t_n \sum_{i=h_{n-1}}^{h_n} \frac{\alpha_i}{t_n}$$

Obsérvese que $\sum_{i=h_{j-1}}^{h_j} \frac{\alpha_i}{t_j} \in A_j$ por ser c.c. de elementos de A_j y que $\sum_{j=1}^{n} t_j = 1, \ t_j \geq 0$.

Finalmente para probar la compacidad de $co(\cup_{i=1}^{n} A_i)$ definamos $f : A \times A_1 \times ... \times A_n \to co(\cup_{i=1}^{n} A_i)$ donde $A = \left\{\lambda \in R_+^n : \sum_{i=1}^{n} \lambda_i = 1\right\}$ tal que $f(\lambda, x_1, x_2, ..., x_n) = \sum_{i=1}^{n} \lambda_i x_i$. Es una función continua con dominio compacto y por lo tanto por lo que $co(\cup_{i=1}^{n} A_i)$ es compacto.[]

Definición 2.14 *Diremos que m es la* **dimensión de un conjunto convexo** $C \subset R^n$ *si esta es la dimensión del menor subconjunto afín de R^n que contiene a C, subconjunto al que denotaremos como $aff(C)$, afinidad de C. Obsérvese que la diemensión de $aff(C)$ es a lo sumo n.*

Definición 2.15 *Se define como* **interior relativo** *de un conjunto convexo $X \subseteq R^n$, que será denotado como riC, al interior de X en la topología relativa del menor subconjunto afín que contiene a X. Es decir $riX = \{x \in aff(X) : \exists \epsilon > 0, (x + B_0(\epsilon)) \cap (aff(X)) \subset X\}$.*

Si C es un subconjunto convexo de R^n entonces como facilmente puede comprobarse, la dimensión de C será n si y solamente si el interior de C coincide con el interior relativo de C, es decir si y solamente si $C^0 = riC$.

El siguiente teorema es el teorema fundamental en lo referente a resultados de dimensionalidad del análisis convexo. Dice que para conocer la cápsula convexa de un subconjunto $S \subset R^n$ es suficiente considerar combinaciones convexas de no más de $n + 1$ elementos de S.

Teorema 2.16 **(Teorema de Carathéodory)** *En un espacio vectorial de dimensión n, todo vector en la cápsula convexa de un conjunto no vacío, puede escribirse como la combinación convexa de a lo más $n + 1$ vectores de dicho conjunto.*

Demostración: Sea S subconjunto de R^n. Supongamos que $x \in co(S)$ se escribe como c.c. de menos de k elementos de S, siendo $k > n+1$ sea $x = \sum_{i=1}^{k} \alpha_i x_i$. De esta forma el conjunto $V = \{x_2 - x_1, x_3 - x_1, ..., x_k - x_1\}$ es linealmente dependiente. Por lo tanto existes $\lambda_i,\ i = 1, 2, ..., k$ no todos nulos tales que $0 = \sum_{i=2}^{k} \lambda_i (x_i - x_1)$. De donde se sigue que $0 = \sum_{i=2}^{k} \lambda_i x_i - x_1 \sum_{i=2}^{k} \lambda_i$. Sea $c_1 = -\sum_{i=2}^{k} \lambda_i$, y sea $c_i = \lambda_i\ \ i = 2, 3, ..., k$. Puede verse que $\sum_{i=1}^{k} c_i x_i = 0$ y que $\sum_{i=1}^{k} c_i = 0$.

Sea $c = min\{\frac{\alpha_i}{c_i} : c_i > 0\}$. Supongamos que $c = \alpha_m / c_m$. Se sigue que $\alpha_i - c c_i \geq 0,\ \forall i$ y además como $\alpha_m - c c_m = 0$ que $\sum_{i=1}^{n} \alpha_i - c c_i = 1$. Finalmente puede verse que $x = \sum_{i=1}^{n} (\alpha_i - c c_i) x_i$ donde al menos un elemento de la c.c. es cero, por lo que x puede escribirse con menos de k elementos. Luego cada vez que $k > n-1$ podemos obtener una c.c. que exprese a x como una c.c. con al menos un elemento menos.[]

2.1 Teoremas de separación

La siguiente sección estará dedicada al estudio de los teoremas de separación de conjuntos convexos. A partir de estos teoremas pueden realizarse un conjunto de importantes aplicaciones a la microeconomía. Estos teoremas son tan importantes que muy poco del actual desarrollo de la microeconomía puede obtenerse prescindiendo de los mismos. La extensiónde estos teoremas a espacios más abstractos que R^n puede hacerse, pero sólo bajo ciertos supuestos, en espacios de dimensión infnita los interiores de los conjuntos convexos pueden ser vacíos, en estos casos no puede hacerse la extensión de estos teoremas. Para convexos con interior relativo no vacío los teoremas de separación son resultado del teorema de Hahn-Banach. Muchas de las dificultades de la teoría económica modelada sobre espacios de dimesnión infinita, radican precisamente en la imposibilidad de extender a esots espacios los teoremas de separación de convexos.

Teorema 2.17 (Teorema de Separación (1)) *Sea X un conjunto convexo no vacío de R^n. Sea $x_0 \notin \bar{X}$. Entonces:*

i) Existe un punto $a \in \bar{X}$ tal que $d(x_0, a) \leq d(x_0, x)$ para todo $x \in \bar{X}$, con $d(x_0, a) > 0$. Aquí $d(\cdot, \cdot) : R^n \times R^n \to R_+$ representa la función distancia definida en R^n.

ii) Existe un $p \in R^n, p \neq 0, \|p\| < \infty$ y un $\alpha \in R$ tal que

$$\langle p, x \rangle \geq \alpha, \forall x \in \bar{X} \quad y \ ; \ \langle p, x_0 \rangle < \alpha.$$

en otras palabras, $\bar{X}$ y x_0 están separados por el hiperplano $H = \{x : \langle p, x \rangle = \alpha, x \in R^n\}$

Demostración:

(i) Sea $\bar{B}(x_0)$ una bola cerrada con centro en x_0 tal que la intersección con $\bar{X}$ es no vacía. Sea $A = \bar{B}(x_0) \cap \bar{X}$, este conjunto es no vacío, cerrado y acotado, por lo tanto compacto[7]. Sea $d(x_0, \cdot) : A \to R$ la función distancia. Es una función continua y por ser A compacto alcanza su mínimo en A (teorema de Weierstrass). Esto es, existe $a \in A$ tal que $d(x_0, a) \leq d(x_0, x)$ para todo $x \in A$ y por lo tanto para para todo $x \in \bar{X}$. Como $x_0 \notin \bar{X}$ entonces $d(x_0, a) > 0$.

(ii) Sea $p \equiv a - x_0$ y $\alpha = \langle p, a \rangle$. Obsérvese que $\langle p, x_0 \rangle = -\|p\|^2 + \alpha < \alpha$. Sea $x \in \bar{X}$ siendo $\bar{X}$ convexo se sigue que $x(t) = (1-t)a + tx \in \bar{X}$. Luego será $d(x_0, a) \leq d(x_0, x(t))$. Esto es

$$\|x_0 - a\|^2 \leq \|x_0 - [(1-t)a + tx]\|^2 = \|(1-t)(a - x_0) + t(x - x_0)\|^2,$$

siendo $\|(1-t)(a-x_0)+t(x-x_0)\|^2 = \langle (1-t)(a-x_0)+t(x-x_0)\ ,\ (1-t)(a-x_0)+t(x-x_0) \rangle$ se sigue que:

$$\|x_0 - a\|^2 \leq (1-t)^2\|a - x_0\|^2 + 2(1-t)t(a - x_0)(x - x_0) + t^2\|x - x_0\|^2.$$

Es decir $0 \leq (t-2)t\|a - x_0\|^2 + 2(1-t)t(a - x_0)(x - x_0) + t^2\|x - x_0\|^2$. Dividiendo por $t, (t > 0)$ y tomando en la expresión obtenida el límite para $t \to 0$ se obtiene que

$$0 \leq -2\|a - x_0\|^2 + 2(a - x_0)(x - x_0)$$

o equivalentemente $0 \leq \langle (a - x_0)\ ,\ (a - x_0) + (x - x_0) \rangle$. Finalmente operando y siendo $p = (a - x_0)$ obtenemos que $\langle p, a \rangle \leq \langle p, x \rangle$.[]

Corolario 2.18 *Sea X un conjunto convexo no vacío en R^n tal que $0 \notin \bar{X}$. Entonces existe $p \in R^n$ tal que $p \neq 0$, $\|p\| < \infty$, y un real α tal que:*

$$\langle p, x \rangle \geq \alpha > 0 \quad \forall x \in \bar{X}.$$

Teorema 2.19 (Teorema de Separación (2)) *Sea X un conjunto convexo no vacío en R^n y sea $x_0 \notin riX$. Entonces existe $p \in R^n$, $p \neq 0$, $\|p\| < \infty$, tal que*

$$\langle p, x \rangle \geq \langle p, x_0 \rangle \ \ \forall x \in X.$$

Demostración: Distinguimos dos casos.

(i) Si $x_0 \notin \bar{X}$ entonces el teorema anterior prueba la afirmación.

[7]Obsérvese que esta caraterización de los conjuntos compactos sólo vale en R^n.

(ii) Si $x_0 \in \bar{X}$, entonces para toda bola de radio r y centro en x_0 existe un punto que no pertenece a $\bar{X}$. Es decir existe una sucesión de puntos $x^q \notin \bar{X}$ convergente a x_0 y una sucesión de p^q tales que $\|p^q\| = 1$ para los que se verifica que $p^q x^q < p^q x$ para todo $x \in \bar{X}$. Como p^q pertenecen a la esfera unitario la que es un conjunto compacto, existe un punto límite, sea este p. La continuidad del producto interno[8] permite afirmar que $\langle p, x_0 \rangle \leq \langle p, x \rangle$ para todo $x \in \bar{X}$.[]

Teorema 2.20 Teorema de Separación (3) *Sean X e Y conjuntos convexos no vacíos en R^n tales que $riX \cap riY = \emptyset$. Entonces existe $p \in R^n$ $p \neq 0$, $\|p\| < \infty$, tal que*

$$\langle p, y \rangle \leq \alpha \leq \langle p, x \rangle; \ \forall y \in Y, \ \forall x \in X.$$

Demostración: Consideremos el conjunto convexo $S = X + (-Y)$. Entonces $0 \notin riS$ pues en otro caso deberían existir $x \in riX$ y $y \in -riY$ tales que $x = -y$ lo que no puede suceder pues la intersección de ambos interiores relativos es vacía. Escribamos $z = x - y$, por el teorema anterior existe $p \neq 0$ tal que $\langle p, z \rangle \geq \langle p, 0 \rangle$ para todo $z \in S$. Luego $\langle p, x \rangle \geq \langle p, y \rangle$ para todo $x \in X$ e $y \in Y$. Es decir $inf_{x \in X} \langle p, x \rangle \geq sup_{y \in Y} \langle p, y \rangle$.[]

Definición 2.21 *Decimos que un hiperplano H es un* **hiperplano soporte** *para un conjunto C si la clausura de C está contenida en unos de los semiespacios cerrados que H determina.*

Obsérvese entonces que el teorema de separción (2) establece que si $x^* \notin ri(C)$ siendo C un conjunto convexo, entonces existe un hiperplano soporte para C que contiene a x^* particularmente si x^*, es un elemento de la frontera de C.

2.2 Aplicaciones

Verifique las siguientes afirmaciones:

1. Un conjunto convexo, cerrado Ces laintersección de todos los semi-espacios que lo contienen. (Ayuda: use el teorema (1) de separación, luego la intersección de todos los semi-espacios que contien a C no podrá tener puntos qu eno sean de C)

2. Sea S un subconjunto cualquiera de R^n muestre que la capsula convexa de S es igual a la intersección de todos los semiespacios que contienen a S.

Tipos de separación

[8] Esta propiedad del producto interno no es necesaria en espacios de dimensión infinita, ver por ejemplo [Brézis, H.].

- Decimos que el hiperplano $H_{p,\alpha}$ **separa** dos conjuntos A y B, si
  $A \subseteq \{x \in R^n : \langle p, x\rangle \leq \alpha\}$ y $B \subseteq \{x \in R^n : \langle p, x\rangle \geq \alpha\}$
- Decimos que el hiperplano $H_{p,\alpha}$ **separa estrictamente** dos conjuntos A y B, si
  $A \subseteq \{x \in R^n : \langle p, x\rangle < \alpha\}$ y $B \subseteq \{x \in R^n : \langle p, x\rangle > \alpha\}$
- Decimos que el hiperplano $H_{p,\alpha}$ **separa fuertemente** dos conjuntos A y B, si
  $A \subseteq \{x \in R^n : \langle p, x\rangle \leq \alpha\}$ y $B \subseteq \{x \in R^n : \langle p, x\rangle \geq \alpha + \epsilon\}$

Ejemplo 2.22 *En el espacio vectorial R^2, definimos los siguientes conjuntos:* $A_1 = \{(x, y) : y > 0; o(y = 0, x > 0)\}$, *y* $B_1 = -A_1$, *así como* $A_2 = \left\{(x, y) : x > 0, y \geq \frac{1}{x}\right\}$ *y* $B_2 \left\{(x, y) : x > 0, y \leq -\frac{1}{x}\right\}$. *Entonces el hiperplano $y = 0$ separa A_1 y B_1 y separa estrictamente a A_2 y B_2. No obstnate los conjuntos A_1 y B_1 no pueden ser estricatmente separados, y A_2 no puede ser fuertemente separado de B_2.*

Definición 2.23 *Un* **poliedro** *en un espacio vectorial toplógico, es la intersección finita de semiespacios cerrados. Es decir un poliedro es la solución de un sistem finito de inecuaciones en X.*

3 Funciones cóncavas y cuasicóncavas

Dado que nuestro objetivo es el de resolver programas de optimización no lineales donde la función objetivo es cóncava o cuasi-cóncava, presentaremos en esta sección algunas de las principales características de este tipo de funciones. Muchas de las definiciones y propiedades que analizaremos en este capítulo pueden extenderse a espacios vectoriales topológicos como ser las definciones de funciones cóncavas o convexas. No obstante centraremos nuestra atención en R^n con la topología habitual, espacio este que representa un caso particular de tales e.v.t.

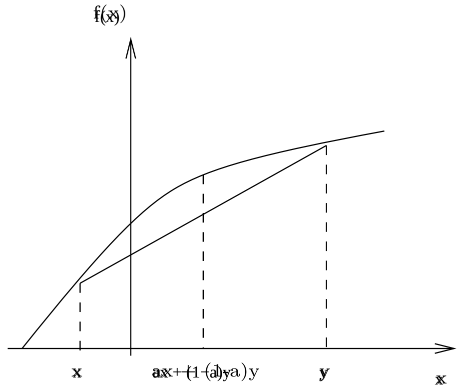
Definición 3.1 *Sea f una función real definida en X subconjunto convexo de R^n. La función f es llamada* **función cóncava** *si, para todo $x, y \in X$, y $0 \leq \theta \leq 1$,*

$$f[\theta x + (1-\theta)y] \geq \theta f(x) + (1-\theta) f(y)$$

Una función real definida en X subconjunto de R^n se dice **convexa** *si $-f$ es cóncava.*

Definición 3.2 *Una función* **afín** *es una función $f : X \to R$ finita, a la vez cóncava y convexa. Observe que f puede ser escrita como la suma de una función lineal $l : X \to R$ y una constante c, $f(x) = l(x) + c$.*

Si las desigualdades son estrictas, la función se dice que es **estrictamente cóncava.** Una función f es llamada **estricatmente convexa** si $-f$ es estrictamente cóncava.



Observación 3.3 *Intuitivamente f es cóncava si la cuerda que une dos puntos de su gráfica cae bajo la función.*

Definición 3.4 *Decimos que $f : X \to R$ siendo X un subconjunto convexo de un e.v.t. (en general de un espacio vectorial) es* **cuasicóncava** *si*

$$f(\alpha x + (1-\alpha)y) \geq \; min\,\{f(x), f(y)\}\,,\; 0 \leq \alpha \leq 1.$$

La función f es **cuasiconvexa** *si y solamente si $-f$ es cuasicóncava.*

Se sigue de la definición de cuasiconvexidad que f es cuasiconvexa si y solamente si $f(\alpha x + (1-\alpha)y) \leq \; max\,\{f(x), f(y)\}\,,\; 0 \leq \alpha \leq 1$.

Sea X un subconjunto convexo de R^n y $f : X \to R$. Algunas propiedades de las funciones cóncavas que se deducen de la definición en forma inmediata son:

Se verifica la siguiente desigualdad: $f(\sum_{i=1}^{n} \alpha_i x_i) \geq \sum_{i=1}^{n} \alpha_i f(x_i)$ es decir el valor que toma una función cóncava en una combianción convexa de elementos de X mayor o igual que la combinación convexa de los respectivos valores.

El subgrafo $S = \{(x, \lambda) \in X \times R : \lambda \leq f(x)\}$ de una función cóncava, es convexo en $R^n \times R$.

El epígrafo $S = \{(x, \lambda) \in X \times R : \lambda \geq f(x)\}$ de una función convexa, es convexo en $R^n \times R$.

Una combinación lineal no negativa, de funciones cóncavas, definidas en un conjunto convexo, es una función cóncava.

Dada una función $f : S \to R$ con $S \subset R^n$ se puede siempre prolongar a una finci'on $g : R^n \to \bar{R}$ donde $\bar{R}$ es el conjunt de los reales ampliado con los valores $\pm\infty$ defiida por:

$$g(x) = \begin{matrix} f(x) & si x \in S \\ \\ \infty & si x \notin S \end{matrix}$$

En muchos casos se define el concepto de función cóncava o convexa permitiendo que el recorrido de la función tome valores en el conjunto de los reales ampliado $\bar{R} = R \cup \{-\infty + \infty\}$. En este caso, funciones tales como

$$f(x) = \begin{cases} -\infty & |x| < 1, \\ \\ 0 & |x| = 1, \\ \\ \infty & |x| > 1. \end{cases}$$

representan funciones convexas, pues su epígrafo es convexo no obstante obsérvese que nuestra definición de convexidad tendría problemas técnicos. Funciones de este tipo son llamadas impropias. La definicón general de función cóncava que abarque también estos casos sería la siguiente:

Definición 3.5 *Una función $f : R^n \to \bar{R}$ es convexa, si y solamente si su epígrafo es convexo.*

Ejercicio 3.6 *Muestre que si una función verifica la defінción de convexidad dada anteriormente entonces verifica esta última.*

Definición 3.7 *Una función convexa se dice* **propia** *cuando su epígrafo es no vacío y no contine rectas verticales, esto es $f(x) < +\infty$ para al menos un x y $f(x) > -\infty$ para todo x.*

El subconjunto $\{x : f(x) < \infty\}$ es llamado dominio efectivo. De esta manera una función convexa definida en un conjunto C no vacío de un espacio vectorial topológico X, puede ser visto como una función convexa definida en todo $x \in X$, cuyo dominio efectivo es C. Las razones que llevan a considerar el recorrido de f como un subconjunto de $[-\infty, \infty]$ son de orden técnico y nosotros no entraremos en ellos aquí, por lo que el dominio de f coinsidirá siempre con su dominio efectivo, a menos que digamos o contrario. Para detalles ver [Rockafellar, T.].

Ejercicio 3.8 *Reescriba las anteriores definiciones de funciones convexas propias y dominio efectivo, para funciones cóncavas.*

Observación 3.9 *Considere las funciones cóncavas* $g_j : X \rightarrow, \ j = 1, 2, ..., m$. *Entonces el conjunto*

$$\mathcal{C} = \bigcap_{j=1}^{n} \{x : x \in X : g_j(x) \geq 0\}$$

es convexo. (La justificación queda a cargo del lector).

El lector familiarizado, realcionará esta observación con el conjunto factible, definido por funciones cóncavas, propios de problemas de maximización como los que presentaremos más adelante.

3.1 Algunas propiedades de las funciones cóncavas de variable real

Presentaremos a continuación algunas propiedades para funciones cóncavas reales de variable real.

Teorema 3.10 *Si* $f :\rightarrow R$ *es cóncava y si* $f(t) < \infty$ *para todo* $t \in I$ *donde* I *es un intervalo en* R, *entonces* f *es continua y existen las derivadas laterales en el interior de su dominio y se tiene*

$$f'_{-}(s) \geq f'_{+}(s) \geq \frac{f(t) - f(s)}{t - s} \geq f'_{-}(t) \geq f'_{+}(t)$$

para todo s *y para todo* t *en el interior del dominio de* f *tal que* $s \leq t$.

Demostración (intuitiva) de la continuidad: Consideremos $a < s < x < y < t < b$, representemos por X al punto $(x, f(x))$ análogamente en los demás casos. El punto X queda por encima de la recta que SY mientras que el punto Y queda por encima de la recta XT. Cuando $y \rightarrow x$ entonces $f(y) \rightarrow f(x)$. Con los límites por la izquierda procedemos de igual forma.[]

Demostración: (*Formal*). Considere $a < b < c$ pertenecientes al interior del dominio de f. Sea $h \in R$ tal que: $b = (1 - h)a + hc$. Usando la concavidad de f se obtiene que

$$\frac{f(b) - f(a)}{b - a} \geq \frac{f(c) - f(a)}{c - a}. \tag{1}$$

Análogamente, sea k tal que $b = ka + (1 - k)c$. Se obtienen entonces que

$$\frac{f(c) - f(a)}{c - a} \geq \frac{f(c) - f(b)}{c - b}. \tag{2}$$

Entonces vale la siguiente cadena de desigualdades:

$$\frac{f(b) - f(a)}{b - a} \geq \frac{f(c) - f(a)}{c - a} \geq \frac{f(c) - f(b)}{c - b}. \tag{3}$$

Sea $0 \leq \alpha \leq \beta$; entonces para t perteneciente al interior del dominio de f, $t - \beta < t - \alpha < t < t + \alpha < t + \beta$ y la cadena de desigualdades (2) implican que:

$$\frac{f(t) - f(t - \beta)}{\beta} \geq \frac{f(t) - f(t - \alpha)}{\alpha} \geq \frac{f(t + \alpha) - f(t)}{\alpha} \geq \frac{f(t + \beta) - f(t)}{\beta}.$$

Puede verse que la función $x \rightarrow \frac{f(t+x)-f(t)}{x}$ es una función no decreciente cuando $x \downarrow 0^+$ y que además está acotada por arriba, por lo que existe $f'_+(t)$.

Similarmente la función $x \rightarrow \frac{f(t)-f(t-x)}{x}$ es una función no creciente cuando $x \downarrow 0^+$ y que además está acotada por abajo, por lo que existe $f'_-(t)$.

Más aun, para $x > 0$ y t_1, t_2 en el interior del dominio de f se tiene que:

$$f'_+(t_1) \geq \frac{f(t_1+x)-f(t_1)}{x} \ \ y \ \ f'_-(t_2) \leq \frac{f(t_2)-f(t_2-x)}{x}.$$

Considerando: $x = t_2 - t_1$ se obtiene:

$$f'_+(t_1) \geq \frac{f(t_2)-f(t_1)}{x} \geq f'_-(t_2).[$$

Finalmente de la disigualdad $\frac{f(t)-f(t-x)}{x} \geq \frac{f(t+x)-f(t)}{x}$, $\forall\, x > 0$, tomando límites cuando $x \rightarrow 0$ se sigue que $f'_-(t) \geq f'_+(t)$.[]

Nota: Observe que la desigualdad última vale para $t_2 > t_1$ ambos en el interior del dominio de la función.

Teorema 3.11 *Si $f : I \rightarrow R$ diferenciable en I. Una condición necesaria y suficiente para que f sea cóncava en I es que f' sea decreciente en I.*

Demostración: El directo es una consecuencia inmediata del teorema anterior.
Recíprocamente: Sea $a \leq b \leq c$. Por el teorema del valor medio sabemos que existen $x_1 \in (a, b)$ y $x_2 \in (b, c)$ tales que:

$$f'(x_1) = \frac{f(b)-f(a)}{b-a} \ \ y \ \ f'(x_2) = \frac{f(c)-f(b)}{c-a}.$$

Por ser la derivada decreciente se concluye que:

$$\frac{f(b)-f(a)}{b-a} \geq \frac{f(c)-f(b)}{c-a} \quad (^*).$$

Luego haciendo $b = \alpha a + (1-\alpha)c$, sustituyendo en (*) y dado que $\alpha = \frac{a-b}{a-c}$ se llega a que $f(\alpha a + (1-\alpha)c) \geq \alpha f(a) + (1-\alpha) f(c)$.[]

Corolario 3.12 *Si $f : I \rightarrow R$ admite derivada segunda en I. Una condición necesaria y suficiente para que f sea cóncava en I es que $f''(x) \leq 0\ \forall x \in I$.*

Demostración: Ejercicio para el lector.[]

Observación 3.13 *Si f es tal que f'' es negativa (resp. positiva) en I entonces f es estrictamente cóncava (resp. convex) en I. No obstante puede un función ser estrictamente cóncava (resp. convexa) en un intervalo y la derivada anularse en algún punto del mismo. Caso $f(x) = x^4$.*

3.2 Propiedades para funciones cóncavas de n variables.

A continuación presentaremos algunas propiedades de continuidad verificadas por las funciones cóncavas y convexas.

Ejercicio 3.14 *Los siguientes dos teoremas serán enunciados y demostrados para funciones convexas, enuncie y demuestre los teoremas correspondientes para funciones cóncavas.*

Teorema 3.15 (De continuidad local) *Si una función f convexa está definida en X convexo y está acotada superiormente en un entorno $V_x \cap X$ de un punto $x \in ri(X)$, entonces es continua en el punto.*

Demostración: Sea $B_r(0)$ la bola con centro en cero y radio r tal que para $z \in B_r(0):\ x - z\ \in V_x$. Para $\delta \in [0,1],\ x + \delta z = (1-\delta)x + \delta(x+z)$, se tiene que $f(x+\delta z) \leq (1-\delta)f(x) + \delta f(x+z)$. Por lo tanto

$$f(x+\delta z) - f(x) \leq \delta[f(x+z) - f(x)]. \tag{4}$$

Cambiando z por $-z$ obtenemos:

$$f(x-\delta z) - f(x) \leq \delta[f(x-z) - f(x)]. \tag{5}$$

Como $x = \frac{1}{2}(x+\delta z) + \frac{1}{2}(x-\delta z)$ se sigue que $f(x) \leq \frac{1}{2}f(x+\delta z) - \frac{1}{2}f(x-\delta z)$. Multiplicando por 2, la expresión anterior, operando a partir de las ecuaciones (4) y (5) se obtiene:

$$f(x) - f(x+\delta z) \leq f(x-\delta z) - f(x) \leq \delta[f(x-z) - f(x)]. \tag{6}$$

A partir de (4) y (5)se obtiene:

$$|f(x+\delta z) - f(x)| \leq \delta max\left\{f(x-z) - f(x), f(x+z) - f(x)\right\}. \tag{7}$$

Luego, usando el hecho de que f está acotada por arriba en un entorno de x se sigue para δ suficientemente pequeño: $|f(x+\delta z) - f(x)| \leq \delta M$. Luego, para todo $y \in x + \delta B_r(0)$ para $\epsilon > 0$ arbitrario, eligiendo $\delta = \frac{\epsilon}{M}$ se sigue que $|f(y) - f(x)| < \epsilon$ para todo $y \in x + \delta B_r(0)$ desde que $\delta = \frac{\epsilon}{M}$.[]

Teorema 3.16 (De la continuidad global) *Sea $f : X \to R$ siendo X un subconjunto convexo de R^n, convexa. Las siguientes afirmaciones son equivalentes:*

1. *f es continua en $ri(X)$.*

2. *f es semicontinua superior en $ri(X)$.*

3. *f está acotada superiormente en algún entorno abierto $V_x \cap X$ de cada punto $x \in ri(X)$.*

4. *f es continua en algún $x \in ri(X)$.*

Demostración:

(1) $\to$ (2) es inmediato.

(2) $\to$ (3) Sea $x \in ri(X)$ como f es semicontinua superior se tiene que $\{y : f(y) < f(x) + 1\}$ es un entorno de x abierto .

(3) $\to$ (4) Es el teorema probado anteriormente.

(4) $\to$ (1) Supongamos que f es continua en $x \in ri(X)$. Queremos probar que entonces f es continua en cualquier otro punto $y \in ri(X)$. Sea V_x un entorno de x tal que $V_x \subset ri(X)$. Escribimos $V_x = x + B_0(\epsilon)$ donde $B_0(\epsilon)$ es la bola de centro cero y radio ϵ. Para cada $y \in ri(X)$ existe $z \in X$, tal que $y = \lambda x + (1 - \lambda)z$ [9]. Se verifica que $y + \lambda B_0(\epsilon) = \lambda(x + B_0(\epsilon)) + (1 - \lambda)z$. Sea $B_0' = \lambda B_0(\epsilon)$, entonces $V_y = y + B_0'$ es un entorno de y contenido en $ri(X)$. Sea $w \in V_y$ existen entonces $v \in B_0'$ y $u \in B_0(\epsilon)$ tales que $w = y + v = \lambda(x + u) + (1 - \lambda)z$. Como $f(w) = f(y+v)$ de la convexidad de f se sigue que $f(w) \leq \lambda f(x+u) + (1-\lambda)f(z)$. Como f es continua en x existe M tal que $f(x + u) \leq M, \ \ \forall u \in B_0(\epsilon)$ (ϵ suficientemente pequeño). Luego $f(w) \leq \lambda M + (1-\lambda)f(z)$ para todo $w \in V_y$ por lo tanto f está acotada en V_y entorno de y y por el teorema anterior es continua en y.[]

Obsérvese que los teoremas de continuidad anteriores son válidos en general en espacios vectoriales toplógicos cualesquiera. Su validez, en espacios vectoriales de dimensión finita como R^n es inmediata pues si una función convexa f, tiene por dominio un subconjunto convexo de R^n, entonces es finita en un entorno de cualquier punto interior.

En efecto, sea x interior a C, existen entonces a y b tales que $a < x < b$. Sea $[a, b] = \{y \in R^n : a \leq y \leq b\}$. Este es un entorno de x en C que forma la cápsula convexa de un conjunto finito de puntos (teorema de Charadtheodory) [10], luego f debe ser acotada por arriba en ese conjunto. La continuidad de f en C se sigue de los teoremas anteriores.

En general una función convexa definida en un espacio de dimensión infinita puede no ser continua, basta pensar en un funcional lineal discontinuo.

[9] Tal $z \in X$ existe, para probarlo considere V_y entorno abierto de y. Sea $\beta > 0$ suficientemente pequeño como para que $z = y + \beta(x - y) \in V_y$. Puede verse entonces que existe $0 < \lambda < 1$ tal que $y = \lambda x + (1 - \lambda)z$.

[10] Sea $h = b - a$ observe que los puntos de la forma $b - \delta h = (b_1 - \delta_1 h_1, ..., b_n - \delta_n h_n)$ siendo $\delta_i = 1 \ o \ 0 \ \ i = 1, 2, ..., n$ son puntos extremos. Entonces $[a, b]$ es la cápsula convexa de estos puntos.

Ejercicio 3.17 *(i) Muestre un ejemplo de función convexa semicontinua superior en su dominio que no sea continua en el mismo.*

(ii) Muestre un ejemplo de función cóncava semicontinua inferior en su dominio que no sea continua en el mismo.

(iii) Muestre que una función convexa no puede ser semicontinua inferior sin ser continua y que, análogamente, una función cóncava no puede ser semicontinua superior sin ser continua.

Ejercicio 3.18 *(i) Demuestre que si una función cóncava definida en X convexo, alcanza un máximo local en x^** [11] *entonces $f(x^*)$ es un máximo global.*

Para la demostración de este punto, considere $z(\alpha) = \alpha x^* + (1-\alpha)y \in X, \ \forall\ 0 \leq \alpha \leq 1$, observe que $z(\alpha)$ pertenece a cualquier entorno de de x^* si α está suficientemente próximo a 1. Utilice la definición de concavidad para contradecir entonces, que x^* es un máximo local.

(ii) Demuestre que si una función convexa definida en X convexo, alcanza un mínimo local en x^* entonces $f(x^*)$ es un mínimo global.

A continuación presentaremos algunas propiedades de las funciones convexas relacionadas con la derivabilidad. Notaremos por $\bigtriangledown f(x)$ al vector gradiente de f en x (es decir el vector de R^n formado por las derivadas parciales de f evaluadas en el punto x).

Teorema 3.19 *Sea C un conjunto convexo $C \subset R^n$ y $f : C \to R$ diferenciable. f es cóncava si y solamente si*

$$\langle \bigtriangledown f(y) - \bigtriangledown f(x),\ y - x \rangle \leq 0 \quad \forall x, y \in C. \tag{8}$$

Demostración: Para la demostracióndel teorema precisamos del siguiente lema:

Lema 3.20 *Sean x e y en C y $d = (y - x)$. Considere $F(t) = f(x + td)$. Entonces f es cóncava si y solamente si para todo $x, y \in X$ lo es F.*

Demostración del lema: Considere $F(t) = f(ty + (1-t)x)$, $0 \leq t \leq 1$, si F es cóncava se sigue que $F(t) \geq tF(1) + (1-t)F(0)$, como $F(1) = f(y)$ mientras que $F(0) = f(x)$ la concavidad de f se sigue. El recíproco queda a cargo del lector.

Directo: Sea f cóncava. Sea ahora $d = y - x$ entonces debido a la concavidad de F, se sigue que $F'(0) \geq F'(1)$ como $F'(0) = \bigtriangledown f(x), y - x$ y $F'(1) = \bigtriangledown f(y), y - x$ se tiene $\langle \bigtriangledown f(x), y - x \rangle \geq \langle \bigtriangledown f(y), y - x \rangle$

[11] Recuerde que una función alcanza en x^* su máximo local, si y solamente si, existe un entorno del punto donde $f(x^*) \geq f(y)$, para toda y en dicho entorno.

Recíproco: Si $\langle \bigtriangledown f(y) - \bigtriangledown f(x)\ y - x\rangle \leq 0 \quad \forall x, y \in C$ se deduce que F' es decreciente pues la expresión: $[F'(t) - F'(t')](t - t') = \langle \bigtriangledown f(x + td) - \bigtriangledown f(x + t'd), d(t - t')\rangle$ es negativa cada vez que $t \geq t'$. Por lo tanto si y solamente si F' es decreciente lo que implica que F es cóncava, luego f es también lo es.[]

Teorema 3.21 *Sea C un conjunto convexo $C \subset R^n$ y $f : C \to R$ dos veces diferenciable. f es cóncava si y solamente si $\bigtriangledown^2 f(x)$ es semidefinida negativa sobre C.*

Demostración: Considere $F(t) = f(x + td)$. Se ve que f es cóncava si y solamente si $F''(0) \leq 0, \ \forall x, d$ en consecuencia si y solamente si $\langle \bigtriangledown^2 f(x)d, d\rangle \leq 0 \ \forall x, d$. []

Ejercicio 3.22 *(i) Escriba un teorema análogo al anterior para funciones estrictamente cóncavas. Recuerde la observación (3.13).*

(ii) Verificar que si f es cóncava y $\lambda > 0$ entonces λf es cóncava.

(iii) Si f y g son funciones cóncavas entonces también lo es $f + g$.

(iv) Verificar que si f es cóncava y $K : R \to R$ es cóncava entonces $K(f)$ no es necesariamente cóncava.

(v) Probar que las siguientes funciones son cóncavas:

$$f(x) = x^{\frac{1}{2}}, \ \ f(x, y) = x^{\frac{1}{2}} - y^2, \ \ f(x, y) = -\frac{x^2}{y}.$$

(vi) Probar que las siguientes funciones son convexas

$$f(x) = x^2, \ \ f(x, y) = x^2 + y^2, \ \ f(x, y, z) = x^2 + 3y^4 + 2z^2$$

(vii) La siguiente función es cóncava o convexa? $f(x, y) = x^2 + y^3 - x^{\frac{1}{2}} y^{\frac{1}{2}}$. Si una u otra propiedad se verifica, indique las regiones del plano donde ello sucede.

Ejercicio 3.23 *Verifique que si $p = (p_1, ..., p_n)$ pertenece al simplex de R^n es decir que $0 \leq p_i \leq 1; \forall i$, y $\sum_{i=1}^n p_i = 1$, entonces para todo $x \in R^n$ se verifica que $\langle p, x\rangle \geq \Pi_{i=1}^n (x_i)^{p_i}$.*

Ayuda: *Tome logaritmos de $\langle p, x\rangle$ use entonces la concavidad de la función logaritmo y las propiedades conocidas de esta función.*

3.3 Derivadas direccionales y subgradientes

Una de las propiedades de gran utilidad de las funciones convexas es el poseer derivadas direccionales universalmente. Estas derivadas están relacionadas con los subgradientes, vectores que representan hiperplanos soportes del epígrafo de la función. Un interesante resumen de las principales propiedades de las funciones convexas pueden verse en [Coruzeix, J.P.; Ocaña, E; Sosa, W.]. Naturalmente que la referencia básica es [Rockafellar, T.].

Definición 3.24 *Sea $f : R^n \to [-\infty, \infty]$ finita en a. Se llama* **derivada direccional** *de f en a en la dirección de h, lo que se denota como $f'(a,h)$ al siguiente límite si éste existe:*

$$f'(a;h) = \lim_{\lambda \downarrow 0} \frac{f(a+\lambda h) - f(a)}{\lambda}$$

Teorema 3.25 *Sea f convexa, y sea x un punto donde f es finita. Para cada y el cociente en la definición de derivada direccional es una función no decreciente de $\lambda > 0$ y por lo tanto $f'(x;y)$ existe y es igual al*

$$inf \left\{ \frac{f(x+\lambda y) - f(x)}{\lambda}, \ \ \lambda > 0 \right\}.$$

Más aun, $f'(x;y)$ es una función convexa positivamente homogénea en y, con $f'(x,0) = 0$.

Demostracón: Debido a que la función

$$\phi(\lambda) = \frac{f(x+\lambda y) - f(x)}{\lambda}$$

es creciente, (decrece cuando λ decrece), entonces el límite de ϕ cuando $\lambda \to 0$ existe y es igual al ínfimo. La homogeneidad sigue de la identidad

$$\frac{f(x, \lambda(ky)) - f(x)}{\lambda} = k \frac{f(x, (k\lambda)y)) - f(x)}{k\lambda}$$

para todo $k > 0$. Luego tomamos límites en ambos miembros con $\lambda \downarrow 0$.[]

Ejercicio 3.26 *Enuncie y demuestre el teorema coorespondiente para funciones cóncavas.*

Definición 3.27 *Decimos que un vector $x^* \in R^n$ es un subgradiente*[12] *de una función convexa f en x si $f(z) \geq f(x) + \ < x^*, z - x > \ \forall z \in \ dom(f)$.*

[12] En general para el caso en el domino de f sea un subconjunto convexo de un espacio vectorial X podemos definir como sugradiente a un elemento del espacio dual de X para el que las propuedades enunciadas se verifican. En el caso particular de R^n su dual es el mismo R^n. En la sección (5) introduciremos el concepto de espacio dual.

El conjunto de subgradientes de f en x es llamado **subdiferencial** de f en x y es representado por $\partial f(x)$. El mapa multivaluado $x \to \partial f(x)$ es llamado subdiferencial de f en x. Obsérvese que el conjunto $\partial f(x)$ es convexo y cerrado. Si $\partial f(x) \neq \emptyset$ decimos que f es subdiferenciable en x.

Consecuetmente para una función cóncava g definimos **supergradientes** y **superdiferenciales.** De esta forma x^* es un supergradiente para g si:

$$g(z) \leq g(x) + \langle x^*, z - x \rangle, \ \ \forall z.$$

Como ejemplo considere la función $f(x) = \|x\|$ es diferenciable en $x \neq 0$, pero es subdiferenciable para todo $x \in R^n$.

$$\partial f(0) = \{x^* \in R^n : \|z\| \geq \langle x^*, z \rangle, \forall z\}.$$

mientras que $\partial f(x) = \frac{x}{\|x\|}$ para todo $x \neq 0$.

Teorema 3.28 *Sea f convexa y x un punto donde $f(x)$ es finito. Entonces x^* es un subgradiente si y solamente si, para todo $y \in R^n$ se verifica $f'(x; y) \geq \langle x^*, y \rangle$.*

Demostración: A partir de la desigualdad

$$\frac{f(x + \lambda y) - f(x)}{\lambda} \geq \langle x^*, y \rangle, \ \forall \ \lambda > 0$$

y teniendo en cuenta la definición de $f'(x, y)$.[$\cdot$]

Sea $f : R^n \to R$ finita en a. Se dice que f es Fréchet diferenciable en a si existe $x^* \in R^n$ tal que

$$\frac{f(a + h) - f(a) - \langle h, x^* \rangle}{\|h\|} \to 0 \tag{9}$$

cuando $h \to 0$. Observe que la definición no depende de la norma elegida. Si x^* existe y es único, escribimos entonces $x^* = \bigtriangledown f(a)$. Trivialmente se deduce que $\langle \bigtriangledown f(a), h \rangle = f'(a, h)$.

Teorema 3.29 *Sea f una función convexa propia, y sea $a \in dom(f)$. Entonces f es Fréchet diferenciable en a si y solamente si existe un único x^* en $\partial f(a)$. En este caso $x^* = \bigtriangledown f(a)$.*

Demostración: (Directo:) Sea f es Fréchet diferenciable, en a. Si $x^* \in \partial f$ se tiene que $\langle x^*, h \rangle \leq f'(x, h) = \langle \bigtriangledown f(a), h \rangle$. Entonces $\langle x^* - \bigtriangledown f(a), h \rangle \leq 0 \ \ \forall h$. Haciendo $h = x^* - \bigtriangledown f(a)$ se tine que $x^* = \bigtriangledown f(a)$.

(Recíprocamente:) Considere la función auxiliar $g(x) = f(x) - f(a) - \langle x - a, x^* \rangle$ puede verse que $g(a) = 0$ y que $g(a + h) \geq 0, \forall h$. Además como existe un único subgradiente de f en a, se verifica que $f'(a, h) = \langle h, x^* \rangle$ y por lo tanto $g'(a, h) = 0$ (*). Mostraremos luego que para todo

$h \in R^n$ es posible elegir unos vectores $d_i \in R^n$, $i = 1, 2, ..., 2n$ y $2n$ números reales t_i, tales que: $\sum_{i=1}^{2n} t_i = 1$, $t_i \geq 0$ $\forall i = 1, 2, ..., 2n$ de forma tal que se cumpla la igualdad $h = \|h\|_1 \sum_{i=1}^{2n} d_i t_i$. Por lo que $0 \leq g(a+h) = g(a+\sum_{i=1}^{2n} t_i \|h\|_1 d_i) \leq \sum_{i=1}^{2n} t_i g(a+\|h\|_1 d_i)$. La primera desigualdad se sigue de que $g(a) = 0$ y $g(a+h) \geq 0$ la segunda por la convexidad de g. Entonces $0 \leq \frac{g(a+h)-g(a)}{\|h\|_1} \leq \sum_{i=1}^{2n} t_i \frac{g(a+\|h\|_1 d_i)-g(a)}{\|h\|_1} \leq \sum_{i=1}^{2n} t_i \frac{g(a+\|h\|_1 d_i)-g(a)}{\|h\|_1}$. Donde $\|h\|_1 = \sum_{i=1}^{n} |h_i|$.

Tomando límites cuando $h \to 0$, y teniendo en cuenta (*) obtenemos que $\frac{g(a+h)-g(a)}{\|h\|_1} \to 0$ y por lo tanto $\bigtriangledown g(a) = 0$ y por lo tanto $\bigtriangledown f(a) = x^*$

Elección de los d_i y de los t_i : Sean $e_i, i = 1, 2, ..., n$ los elementos de la base canónica de R^n. Definamos $d_i = e_i$, *y* $d_{n+i} = -e_i$ $i = 1, 2, ..., n$. Luego $t_i = \frac{\xi_i}{\|h\|_1}$ con $\xi_i = 0$, *y* $\xi_{n+i} = -h_i$ si $h_i < 0$ mientras que si $h_i > 0$ hacemos $\xi_i = h_i$, *y* $\xi_{n+i} = 0$. Se sigue entonces que $h = \sum_{i=1}^{2n} d_i \xi_i$.[]

Veremos en el siguiente teorema que el grafo de una función convexa o cóncava puede ser considerado como la envolvente de funciones afines.

Teorema 3.30 *Sea $f : X \to \bar{R}$ una función convexa s.c.i. con dominio efectivo cerrado. Para cada x en el dominio efectivo definamos*

$$G(x) = sup\,\{g(x) : g \leq f, \forall g \text{ afín}\}$$

entonces $f(x) = G(x)$.

Demostración: Fijamos x en C dominio efectivo de f. Es suficiente probar que si $\alpha < f(x)$ entonces existe una función afín tal que $g < f$ y $g(x) = \alpha$. Como f es convexa y s.c.i. el epígrafo $E = \{(y, r) \in C \times R : r \geq f(x)\}$ es no vacío, convexo y cerrado en $R^n \times R$. Consideremos entonces $(x, \alpha) \notin E$. Por el teorema de separación de convexos (1) sabemos que existe un vector de $R^n \times R$, definido como (p, λ) tal que $\langle p, x\rangle + \lambda\alpha < \langle p, x\rangle + \lambda f(x)$ lo que implica $\lambda > 0$. Definamos ahora la función afín g como $g(y) = \frac{1}{\lambda}[-l(y) + l(x)] + \alpha$. Como puede verse $g(y) < f(y)$ $\forall y \in C$ y además $g(x) = \alpha$.[]

Ejercicio 3.31 *Reescriba el teorema anterior y su demostración para funciones cóncavas.*

3.4 Funciones cuasicóncavas y cuasiconvexas

En muchas aplicaciones a la teoría económica una generalización de los conceptos de función cóncava y convexa se hace necesaria. Los conceptos de función cuasicóncava y cuasiconvexa dan más generalidad a las conclusiones que la teoría económica puede obtener. Una muestra de su importancia es el hecho de que curvas de indiferencia convexas provienen de relaciones de preferencia representadas por funciones cuasicóncavas y generalmente las funciones de producción son también consideradas como cuasióncavas.

El teorema de Kuhn-Tucker se extiende a funciones cuasicóncavas. El teorema de Arrow-Enthoven de 1961, muestra esta posibilidad, el mismo será presentado más adelante en estas notas.

Notación En lo que sigue, a los efectos de simplificar la notación escribiremos cuando sea necesario xy para representar el producto interno $\langle x, y \rangle$.

La siguiente definición es como veremos, equivalente a la definición (3.4).

Definición 3.32 *Sea $f : X \to R$, donde X es convexo decimos que f es* **cuasicóncava** *si para todo $a \in R$ se verifica que $X_a = \{x \in X : f(x) \geq a\}$ es convexo. Diremos que f es cuasiconvexa sii $-f$ es cuasicóncava.*

Teorema 3.33 *Sea $f : X \to R$, siendo X es convexo las siguientes afirmaciones son equivalentes:*

1. f es cuasicónva.
2. Para todo $x, y \in X$ y $\theta \in [0, 1]$ se verifica que: $f(\theta x + (1-\theta)y) \geq min\{f(x), f(y)\}$.
3. Si $f(x) \geq f(y)$ entonces $f(\theta x + (1-\theta)y) \geq f(y)$.

Demostración:

(1) → (2) Supongamos que $f(y) = min\{f(x), f(y)\}$. Por ser f cuasicóncava el subconjunto de R^n $\{x \in X : f(x) \geq f(y)\}$ es convexo, luego $f(\theta x + (1-\theta)y) \geq f(y)$.

(2) → (3) Es inmediato.

(3) → (1) Sea $c \in R$ si existe $x \in X$ tal que $f(x) \geq c$ entonces sea $y \neq x \in X_c$. Sin pérdida de generalidad asumimos que $f(y) \geq f(x)$, por (3) $f(\theta x + (1-\theta)y) \geq f(x)$ por lo tanto $\theta x + (1-\theta)y \in X_c$. Es decir que X_c es conexo. En otro caso $X_c = \emptyset$ por lo tanto X_c es también convexo.[]

3.5 Conjuntos extremos

Estudiaremos en esta sección los conjuntos extremos, los que luego relazionaremos con los conjuntos ed maximizadores de funciones cuasicóncavas.

Definición 3.34 *Un* **subconjunto extremo** *de un conjunto C es un subconjunto no vacío E de C con la propiedad de que si $x = \alpha y + (1-\alpha)z \in E$, donde $0 < \alpha < 1$ y con $y, z \in C$ entonces $y, z \in E$. Los elementos de E cuando estos contienen un único elemento son llamados* **puntos extremos** *de C y los notaremos como $E(C)$.*

De esta forma un punto extremo de C no puede ser escrito como combinación convexa de dos puntos diferentes de C. Puede verse que un punto e de un conjunto convexo C es extremo si y solamente si $C \setminus \{e\}$ es convexo.

Ejemplos de conjuntos extremos son

- Las caras de un conjunto convexo.
- Vértices y aristas de un poliedro
- Cada uno de los puntos de la cáscara esférica que limita a una esfera, son puntos extremos.

Existen conjuntos convexos cuyos conjuntos extremos no tienen puntos extremos. Ejemplo de este es el subconjunto convexo C de R^2

$$C = \Big\{(x, y) \in R^2 : y \geq 0\Big\}.$$

El conjunto extremo es $E = \{(x, y) \in R^2 : y = 0\}$ el que claramente no tiene puntos extremos. No obstante, para conjuntos extremos compactos se tiene el siguiente lema:

Lema 3.35 *Sea C subconjunto de R^n entonces todo subconjunto E extremo y compacto de C contiene al menos un punto extremo.*

Demostración: Considere el conjunto de todos los subconjuntos compactos extremos de C incluídos en E ordenémoslos por inclusión. La intersección de los elementos de cada cadena es no vacía y compacta. Por el lema de Zorn, existe un conjunto G minimal extremo de C en F. Si este conjunto tuviera más de un elemento tendríamos una contradicción. Para ver esta contradicción, alcanza con suponer que existen dos elementos diferentes a y b en el minimal. Considere luego un funcional lineal f tal que $f(a) > f(b)$. El conjunto M de maximizadores de este funcional restringido a G es un conjunto extremo que no contiene a b, luego G no sería minimal.[]

Teorema 3.36 *Todo subconjunto K compacto y convexo de R^n está incluído en la cápsula convexa de sus puntos extremos* [13].

[13] Este teorema vale en espacios de Hausdorff localmente convexos. Esta generalización es conocida con el teorema de Krein-Milman.

Demostración: Suponga que existe $a \in K$ tal que $a \notin co(E(K))$. Luego por el teorema de separación (1), existe p funcional lineal en R^n tal que $p(a) > p(x), \ \forall \ x \in co(E(K))$. Sea A el conjunto de los maximizadores de p en K. $A \subset K \cap [co(E(K))]^c$ es un subconjunto extremo no vacío y compacto. Por el lema anterior contiene un punto extremo de K, por lo tanto $A \cap co(E(K)) \neq \emptyset$, lo que contradice la definición de A.[]

3.6 Conjuntos maximizadores para funciones cuasiconvexas y cuasicóncavas

Nos introduciremos en esta sección en la problematica de la maximizción de funciones cuasi cónvexass. Veremos propiedades de los ocnjuntos maximizdores de funciones cuasiconvexas y cuasicóncavas y algunas características sobre la forma en que crecen estas funciones.

Teorema 3.37 (Principio del máximo de Bauer) *Si C es un conjunto compacto y convexo, de R^n entonces toda función s.c.s, cuasiconvexa alcanza su máximo en un punto extremo.*

Idea de la demostración: La demostración es consecuencia de la extensión del teorema de Weierstrass para funciones s.c.s. y de que la cuasiconvexidad de f asegura que el conjunto de los maximizadores es un subconjunto extremo. La afirmación del teorema se obtiene ahora a partir del hecho de que todo subconjunto extremo tiene un punto extremo [14].[]

Ejercicio 3.38 *(i) Demuestre que el conjunto de maximizadores de una función cuasiconvexa es o bien vacío o bien un conjunto extremo y que*

(ii) el conjunto de minimizadores de una función cuasi-cóncava es o bien vacío o bien un conjunto extremo

(iii) Complete la demostración del teorema de Bauer.

(iv) Demuestre que el conjunto de minimizadores de una función cuasicóncava es un conjunto extremo.

(v) Demuestre que el conjunto de maximizadores de una función cuasicóncava no es necesariamente un subconjunto extremo.

Teorema 3.39 *Sea $X \subset R^n$, convexo. Si $f : X \to R$ es una función diferenciable con jacobiano en x^* definido por $f'(x^*) = \bigtriangledown f(x^*)$. Entonces una condición necesaria y suficiente para que f sea cuasicóncava en X es que la desigualdad*

[14] Este teorema vale en todo espacio de Hausdorff localmente convexo.

$$f(x) \geq f(x^*) \;\; implique \;\; \bigtriangledown f(x^*)(x-x^*) \geq 0 \;\; \forall\, x^*, x \in X. \tag{10}$$

Observación 3.40 *Este teorema afirma que una condición necesaria y suficiente para que una función real sea cuasicóncava, es que para todo $x^* \in X$ para el que existe $x \in X$ tal que $f(x) \geq f(x^*)$ entonces $f(y) \geq f(x^*)$ para todo y perteneciente al segmento de recta que une x con x^*. Es decir que la derivada direccional evaluada en cualquier punto de su dominio, sea positiva en la dirección del crecimiento* [15]

Observación 3.41 *Obsérvese que f es cóncava si y solamente si $f(y) - f(x) \leq f'(x)(y-x)$.*

Demostración: Siendo $f(x) \geq f(x^*)$ la cuasiconcavidad de f implica que $\frac{f(x^*+\theta(x-x^*))-f(x^*)}{\theta} \geq 0$; por lo tanto tomando límites cuando $\theta \to 0$ se obtine: $\bigtriangledown f(x^*)(x-x^*) \geq 0$ de donde se obtiene la afirmación.

(Recíprocamente) Sea $\phi(\lambda) = f(\lambda y + (1-\lambda)z)$ por lo tanto $\phi(1) = f(y)$; $\phi(0) = f(z)$. Supongamos que $f(y) \geq f(z)$ supongamos que se cumple (10) pero no obstante que f no es cuasicóncava. Esto es que existe $\bar{\lambda} \in (0,1)$ tal que $\phi(\bar{\lambda}) < \phi(0)$. En estas condiciones, existe $\lambda_0 \in (0,1)$ tal que $\phi(\lambda_0) < \phi(0)$ tal que $\phi'(\lambda_0) > 0$. Sea $x_0 = \lambda_0 y + (1-\lambda_0)z = z + \lambda_0(y-z)$, usando la regla de la cadena se sigue que:

$$\phi'(\lambda_0) = \bigtriangledown f(x_0)(y-z) > 0 \quad (^*).$$

Como $f(z) \geq f(x_0)$ se tiene que $\bigtriangledown f(x_0)(z-x_0) \geq 0$ $(^{**})$, y por ser $f(y) \geq f(x_0)$ se tiene que $\bigtriangledown f(x_0)(y-x_0) \geq 0$ $(^{***})$. Como además $z - x_0 = -\lambda_0(y-z)$ se sigue de $(^{**})$ que $-\lambda_0 \bigtriangledown f(x_0)(y-z) \geq 0$. Análogamente, como $y - x_0 = (1-\lambda_0)(y-z)$ se sigue de $(^{***})$ que $(1-\lambda_0)\bigtriangledown f(x_0)(y-z) \geq 0$. Por lo tanto $\bigtriangledown f(x_0)(y-z) = 0$ lo que contradice la desigualdad $(^*)$.[]

Definición 3.42 *Sea $f : C \to R$. Decimos que un vector $p \in E$ espacio vectorial toplógico define un soporte para en $x^* \in X$ si se verifica que $p(x-x^*) \geq 0$ para todo $x \in C_{x^*} = \{x \in C : f(x) \geq f(x^*)\}$.*

Definición 3.43 *Decimos que un semiespacio $E \subset R^n$ es un* **semiespacio soporte** *para X si X está totalmente contenido en E y existe al menos un punto $x^* \in \bar{X}$ en la frontera de E.*

De esta forma todo hiperplano soporte de X, define un semiespacio soporte de X. En otras palabras los hiperplanos soporte para X, son los hiperplanos que pueden ser representados en la

[15] Obsérvese que para $v = x - x^*$ se tiene que $\bigtriangledown f(x^*)(x-x^*) = \frac{\partial f}{\partial v}(x^*) = f'(x^*, v)$.

forma $H = \{x : \langle x, b \rangle = \beta, \}$ con $b \neq 0$ con $\langle x, b \rangle \leq \beta$ para todo $x \in X$ y $\langle x, b \rangle = \beta$ para al menos un punto en $fr(X)$. De acuerdo al teorema de separación (2), si C es un subconjunto convexo de R^n entonces para todo x^* en su frontera, existe un hiperplano soporte H para C que contiene a x^*. Definiendo a la vez este H un semiespacio soporte para C.

Teorema 3.44 *Sea $f : X \to R$ una función cuasicóncava diferenciable en x^*, siendo $X \subset R^n$ convexo, entonces el gradiente de f evaluado en x^*, $\bigtriangledown f(x^*)$ define un hiperplano soporte para x^*.*

Demostración: En efecto si $x \in C_{x^*}$ entonces por el teorema (3.39) se verifica que $\bigtriangledown f(x^*)(x - x^*) \geq 0$ y por lo tanto $\bigtriangledown f(x^*)$ es un soporte para x^*.[]

Corolario 3.45 *Sea $f : X \to R$ diferenciable en x^*, y $X \subset R^n$ convexo. Supongamos que x^* pueda ser considerado como c.c. de dos elementos de X y que $\bigtriangledown f(x^*) \neq 0$. Entonces f es cuasicóncava en X si y solamente si x^*, es la solución del siguiente programa de maximización no lineal:* $\max_{x \in X} \ f(x) \ \textit{s.a.} \ \ \bigtriangledown f(x^*)x \leq \bigtriangledown f(x^*)x^*$.

Demostración: Sea f cuasicóncava. Supongamos que existe $y \neq x^*$ que verifica las condiciones del programa no lineal y que a la vez verifica que $f(y) > f(x^*)$. Por el teorema anterior sabemos que $\bigtriangledown f(x^*)$ define un soporte para C_{x^*} por lo tanto debe verificarse la igualdad $\bigtriangledown f(x^*)(y - x^*) = 0$. Por la continuidad de f y el hecho de ser $f(y) > f(x^*)$ existe $\epsilon > 0$ tal que para todo $z \in B_y(\epsilon)$ se verifica que: $f(z) > f(x^*)$. A la vez, este z puede elegirse de forma tal que verifique $\bigtriangledown f(x^*)(z - x^*) < 0$ contradiciendo la cuasiconcavidad de f.
(Recíprocamente) Supongamos que x^* puede ser considerado como la combinación convexa de dos puntos $y, z \in C$. Entonces debe suceder que $\bigtriangledown f(x^*)x^* \geq \ min\{\bigtriangledown f(x^*)y, \bigtriangledown f(x^*)z\}$ Por lo tanto y o z pertenecerán a la restricción del programa de maximización. Se deduce entonces que $f(x^*) \geq \ min \ \{f(y), f(z)\}$.[]

Más adelante mostraremos algunas relaciones entre funciones cuasicóncavas y el hessiano orlado.

Muchas veces en economía se utilizan funciones que son cuasicóncavas y homogéneas de grado uno footnoteRecuerde que una función se dice homogénea de grado uno, cuando para todo $\lambda > 0$ verifica $f(\lambda x) = \lambda f(x)$ para todo x en el dominio de f. . Estas dos condiciones juntas implican en muchos casos la concavidad de la función. Se tiene el siguiente teorema que relaciona cuasiconcavidad y concavidad.

Teorema 3.46 *Sea $f : X \to R$ homogénea de grado uno, siendo X un subconjunto convexo de R^n, supongamos que para todo $x \in X; x \neq 0$ se verifica que $f(x) > 0$. Entonces f es cuasicóncava (respectivamente cuasiconvexa) si y solamente si es cóncava (resp. convexa).*

Demostración: La suficiencia es obvia. Sea ahora f cuasicóncava. Sean $x, y \in X - \{0\}$, y definamos $x^* = [1/f(x)]x$ e $y^* = [1/f(y)]y$. Entonces $f(x^*) = f(y^*) = 1$. La cuasiconcavidad de f hace que

$$f(\lambda x^* + (1-\lambda)y^*) \geq min\{f(x^*), f(y^*)\} = 1, \quad (*)$$

En particular esto vale para $\lambda = \frac{f(x)}{f(x)+f(y)}$ se tiene entonces, a partir de $(*)$ y la homogeneidad de f, que $f(x+y) \geq f(x) + f(y)$, es decir f es superaditiva. Por ser entonces f función super aditiva y homogénea de grado uno es cóncava.[]

Ejercicio 3.47 *Pruebe la última afirmación de la demostración del teorema anterior.*

4 Apéndice: Funciones mitcóncavas

Introduciremos en esta sección las llamadas funciones mitcóncavas, la mitconcavidad resulta a veces más sencilla de verificar que la convexidad, y si además es una función continua, entonces la mitconcavidad es equivalente a la concavidad de la función.

Definiremos mitconcavidad y diremos que una función f es mitconvexa si y solamente si $-f$ es mitcóncava.

Definición 4.1 *Sea $f : X \to R$ definida en $X \subset R^n$, convexo, decimos que f es mitcóncava si para todo $x, y \in X$ se verifica que*

$$f\left(\frac{x+y}{2}\right) \geq \frac{1}{2}\left[f(x) + f(y)\right].$$

Teorema 4.2 *Sea X un subconjunto convexo de R^n. Una función $f : X \to R$ es mitcóncava si y solamente si: para toda combinación convexa racional se verifica que:*

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \geq \sum_{i=1}^{n} \alpha_i f(x_i). \tag{11}$$

Nota: Aparece claro entonces que una función mitcóncava continua es cóncava. En efecto, la continuidad de la función obliga a que la desigualdad (11) se verifique también para α_i irracional. No obstante es posible construir, aun para el caso de variable real, funciones mitcóncavas no continuas. La condición de continuidad para mit cóncavas es la misma ya vista para funciones cóncavas, esto es la acotación de f en un entorno de algún punto de su dominio.

Demostración: Es claro que si una función es cóncava entonces es mitcóncava. Veamos entonces el recíproco. Facilmente puede probarse que si f es mitcóncava entonces, para $n = 2^k$ debe verificar que:

$$f\left(\frac{x_1 + x_2 + ... + x_n}{n}\right) \geq \frac{1}{n}\left[f(x_1) + f(x_2) + ... + f(x_n)\right].$$

Para $n = 3$ se tiene que:

$$f\left(\frac{x_1+x_2+x_3}{3}\right) = f\left[\frac{1}{2^2}(x_1+x_2+x_3) + (2^2-3)\left(\frac{x_1+x_2+x_3}{3}\right)\right]$$
$$\geq \frac{1}{2^2}\left[f(x_1)+f(x_2)+f(x_3)+(2^2-3)f\left(\frac{x_1+x_2+x_3}{3}\right)\right].$$

Entonces:

$$\left(1-\frac{2^3-3}{2^2}\right)f\left(\frac{x_1+x_2+x_3}{3}\right) \geq \frac{1}{2^2}\left[f(x_1)+f(x_2)+f(x_3)\right],$$

o equivalentemente,

$$\frac{3}{2^2}f\left(\frac{x_1+x_2+x_3}{3}\right) \geq \frac{1}{2^2}\left[f(x_1)+f(x_2)+f(x_3)\right].$$

Para n cualquier elegimos m tal que $2^{m-1} \leq n \leq 2^m$ y entonces escribimos:

$$f\left(\frac{x_1+x_2+...+x_n}{n}\right) = f\left[\frac{1}{2^m}(x_1+x_2+...+x_n) + (2^m-n)\frac{x_1+x_2+...+x_n)}{n})\right]$$
$$\geq \frac{1}{2^m}\left[f(x_1)+f(x_2)+...+f(x_n)+(2^m-n)f\left(\frac{x_1+x_2+...+x_n}{n}\right)\right].$$

Operando análogamente a lo hecho para el caso $n = 3$ obtenemos:

$$f\left(\frac{x_1+x_1+...+x_n}{n}\right) \geq \frac{1}{n}\left[f(x_1)+f(x_2)+...+f(x_n)\right].$$

Ahora para cualquier conjunto de n racionales $\alpha_i = p_i/q_i,\ i = 1,3...,n$ tales que $\sum_{i=1}^n \alpha_i = 1$, escribiendo ahora $\alpha_i = u_i/d$ donde d el denominador común de los α_i se tiene que $\sum_{i=1}^n u_i = d$ luego:

$$f\left(\frac{u_1}{d}x_1+\frac{u_2}{d}x_2+...+\frac{u_n}{d}x_n\right) = f\left(\frac{1}{d}[(x_1+...+x_1)+(x_2+...+x_2)+...+(x_n+...+x_n)]\right)$$

habiendo u_i sumandos en el pareéntesis que ocupa el lugar i. El resultado se obtiene apelando al caso ya probado.[]

5 Dualidad y optimización

En esta sseción introduciremos el concepto de funcional dual y daremos la demostración del teorema de dualidad de Fenchel.

Definición 5.1 *Se denomina* **funcional** *a una función real, con dominio en un espacio vectorial. Usaremos la siguiente notación:*

1. *Para f funcional convexo con dominio en C subconjunto convexo de un espacio vectorial X, usaremos la notación $[f, C]$ para referirnos al epígrafo de f. Es decir que*

$$[f, C] = \{(s, x) : x \in C : s \geq f(x)\}.$$

2. *Mientras que para un funcional cóncavo g definido en un conjunto convexo D entonces $[g, D]$ representa el subgrafo de g. Es decir definimos*

$$[g, D] = \{(r, x) : x \in D : r \leq f(x)\}.$$

La demostración de la siguiente proposición queda a cargo del lector.

Teorema 5.2 *El funcional f definido en un conjunto convexo C es convexo, si y solamente si, $[f, C]$ es convexo. Obsérvese que $[f, C] \subseteq R \times X$.*

5.1 Funcionales y conjuntos conjugados

Para un espacio vectorial X representaremos por X^* el espacio formado por todos los funcionales lineales acotados definidos en X. Un funcional lineal L en X se dice acotado si

$$\|L\| = \sup_{\|x\| \leq 1} |L(x)| < \infty.$$

Denominamos a este espacio, espacio dual normado de de X.

1. *Sea f un funcional convexo* definido en un subconjunto convexo C de un espacio normado X, el **conjunto conjugado** de C se define como:

$$C^* = \left\{ x^* \in X^* : \sup_{x \in C} [\langle x, x^* \rangle - f(x)] < \infty \right\}$$

   y el **funcional f^* con dominio es C^*** y definido como

$$f^*(x^*) = \sup_{x \in C} [\langle x, x^* \rangle - f(x)],$$

   se denomina **funcional conjugado de f en C.**

2. *Para g funcional cóncavo* definido en D suconjunto convexo de un espacio normado X, se define el **conjunto D^* conjugado** del subconjunto D como:

$$D^* = \left\{ x^* \in X^* : \inf_{x \in D} \langle x^*, x \rangle - g(x)] > -\infty \right\}$$

   y el **funcional g^* conjugado** de g como:

$$g^*(x^*) = \inf_{x \in D} [\langle x, x^* \rangle - g(x)]$$

Nótese que si $f = 0$ entonces f^* define un hiperplano soporte para C. Anlogamente para $g = 0$.

Teorema 5.3 *El funcional conjugado f^* de un funcional convexo, es convexo, el conjunto conjugado C^* es convexo y $[f^*, C^*]$ es un conjunto convexo y cerrado en $R \times X^*$. En el caso de ser g un funcional cóncavo, se verifica que su conjugado es cóncavo, D^* es convexo y $[g^*, D^*]$ es un conjunto convexo y cerrado en $R \times X^*$.*

Demostración: Probemos que f^* es convexo, las demás demostraciones quedan a cargo del lector. Conisdere x_1^* y x_2^* elementos duales. Sea $x^* = \alpha x_1^* + (1-\alpha)x_2^*$. Entonces

$$f^*(x^*) = f^*(\alpha x_1^* + (1-\alpha)x_2^*) = \sup_{x \in C} \{\alpha[\langle x, x_1^* \rangle - f(x)] + (1-\alpha)[\langle x, x_2^* \rangle - f(x)]\} \leq$$

$\leq \alpha \sup_{x \in C}[\langle x, x_1^* \rangle - f(x)] + (1-\alpha) \sup_{x \in C}[\langle x, x_2^* \rangle - f(x)] = \alpha f^*(x_1^*) + (1-\alpha) f^*(x_2^*).[]$

5.2 Problemas duales de optimización

Consideraremos una aplicación de la conjugación al problema de optimización, que consiste en resolver. $\inf_{C \cap D}[f(x) - (x)]$, donde f es un funcional convexo definido en C y g es cóncavo con dominio D ambos dominiso son subconjuntos convexos de un espacio ventorial X. En las aplicaciones más frecuentes, se considera $g = 0$.

Teorema 5.4 (de dualidad de Fenchel) *Sean f y g funcionales respectivamente, convexo y cóncavo definidos en C y D subconjuntos convexos de un espacio vectorial normado. Supongamos que $C \cap D$ continen puntos en el interior relativo y además que tanto $[f, C]$ como $[g, D]$ tienen interior no vacío, supongamos además que $\mu = \inf_{x \in C \cap D} \{f(x) - g(x)\}$ es finito. Entonces*

$$\mu = \inf_{x \in C \cap D} \{f(x) - g(x)\} = \max_{x^* \in C^* \cap D^*} \{g^*(x^*) - f^*(x^*)\},$$

siendo el máximo alcanzado por algún $x_0 \in C \cap D$, donde el máximo de la derecha es alcanzado por algún $x_0^* \in C^* \cap D^*$ Si el ínfimo a la izquierda de la igualdad anterior, es alcanzado por algń $x_0 \in C \cap D$ Entonces:

$$\max_{x \in C}[\langle x, x_0^* \rangle - f(x)] = [\langle x_0, x_0^* \rangle - f(x_0)]$$

y

$$\min_{x \in C}[\langle x, x_0^* \rangle - g(x)] = [\langle x_0, x_0^* \rangle - g(x_0)].$$

Demostración: Por definición, para todo $x^* \in C^* \cap D^*$ y $x \in C \cap D$ se obtiene:

$$\begin{aligned} f^*(x^*) &\geq \langle x, x^* \rangle - f(x) \\ g^*(x^*) &\leq \langle x, x^* \rangle - g(x) \end{aligned} \tag{12}$$

De donde:

$$f(x) - g(x) \geq g^*(x^*) - f^*(x^*) \tag{13}$$

Se obtiene que

$$\inf_{C\cap D}[f(x) - g(x)] \geq \sup_{C^*\cap D^*}[g^*(x^*) - f^*(x^*)] \tag{14}$$

Ahora la igualdad de la hipótesis puede obtenerse, encontrando $x_0^* \in C^* \cap D^*$ que verifique que: $\inf_{C\cap D}[f(x) - g(x)] = [g^*(x_0^*) - f^*(x_0^*)]$. Para esto considere el desplazamiento vertical de $[F, C]$ definido por $[f - \mu, C]$. Los conjuntos $[f - \mu, C]$ y $[g, D]$ están arbitrariamente próximos, pero sus interiores respectivos son no vacíos y tienen intersección vacía. Podemos entonces aplicar el teorema de separación de convexos para definir un hiperplano en $R \times X$ que los separa el que podemos representar por $(1, x_0^*, c) \in R \times X^* \times R$ Este hiperplano puede ser descripto como

$$H = \{(r, x) \in R \times X : \langle x, x_0^* \rangle - r = c\}.$$

Siendo que $[g, D]$ está por debajo de H pero arbitrariamente próximo, y $[f - \mu, C]$ por encima se tiene que

$$c = \inf_{x\in D}[\langle x, x_0^* \rangle - g(x)] = g^*(x_0^*)$$

y que

$$c = \sup_{x\in C}[\langle x, x_0^* \rangle - f(x) + \mu] = f^*(x_0^*) + \mu.$$

Luego $\mu = g^*(x_0^*) - f^*(x_0^*)$ basta entonces que el ínfimo μ sea alcanzado en alguún $x_0 \in C \cap D[\cdot]$

En las aplicaciones clásicas se trata, generalmente, de minimizar un funcional convexo o bien de maximizar un funcional cóncavo en un conjunto D convexo, definido como el conjunto de restricciones o factible. Para resolver cualquiera de los dos problemas mencionados de acuerdo al teorema de Fenchel elegimos $C = X$ y $g(x) = 0$, $\forall x \in D$. Por ser $g(x) \equiv 0$, se tiene que

$$g^*(x^*) = \inf_{x\in D}\{\langle x, x^* \rangle\}$$

De esta forma para el problema de minimización del funcional convexo, obtenemos el problema dual de maximización:

$$\inf_{x\in D}\{f(x)\} = \sup_{x^*\in D^*}\{\langle x, x^* \rangle - f^*(x^*)\}.$$

En el caso de maximizar el funcional cóncavo el teorema de dualidad toma la forma

$$\sup_{x\in D}\{f(x)\} = \inf_{x^*\in D^*}\{\langle x, x^* \rangle - f^*(x^*)\}.$$

El cálculo de f^* es en si mismo un problema de optimización, no obstante, sin restricciones, pues consideramos $X = C$. El cálculo de g^* es también un problema de optimización, en el que como elegimos $g = 0$, se transforma en un problema de optimización con objetivo lineal.

Para una aplicación del teorema de dualidad de Fenchel, ver por ejemplo [Accinelli, E.; Brida, G. Plata, L.; Puchet, M].

6 Introducción a la programación no lineal.

En esta sección haremos uso del teorema de separación de convexos en forma reiterada, y quedará en evidencia la importancia que el teorema tiene para la optimización y por lo tanto en la teoría económica. Si bien en estas notas, el teorema fue demostrado en R^n su validez es mucho más amplia. En espacios de dimensión infinita es un corolario del teorema de Hahn-Banach.

En muchos problemas de la economía y de la ingeniería como en general de la ciencia es necesario maximizar o minimizar ciertas funciones reales, $f(x)$ donde $x \in X$, subconjunto de R^n, sujeto a determinadas restricciones, como por ejemplo:

$$g_1(x) \geq 0, \;\; g_2(x) \geq 0, \;\; \dots \;\; , g_m(x) \geq 0$$

donde cada g_j es una función real. Por ejemplo en la teoría del consumidor f es la función de utilidad, x es una cesta de consumo con n bienes, X el ortante positivo de R^n, y las restricciones, las determinadas por la región presupuestaria, $M - \langle p, x \rangle \geq 0$, con $x \geq 0$. M es el nivel de ingresos del consumidor. Michas veces estas restricciones van acompa{\\a de la llamada restricción de no negatividad, esto es que $x_i \geq 0$ para todas las variables del problema. Si las funciones que intervienen en el problema son lineales, entonces estamos frente a un problema de programación lineal. Este problema fue formulado por el matemático ruso Kantorovich, y luego desarrollado en los Estados Unidos por Dantzig y otros.

En cuanto a la programación no lineal el trabajo fundamental en su desarrollo es el artículo de Kuhn y Tucker: Nonlinear Programing de 1951, *Proceeding of Second Berkeley Symposium.*

La relación funcional $g_j(x) = 0$, lineal o no lineal puede definir una superficie (o una curva) en el espacio (en el plano). Suponga que divide al espacio en dos regiones, una región donde $g_j(x) \geq 0$, otra donde $g_j(x) \leq 0$, y la frontera común donde $g_j(x) = 0$. El llamado **conjunto factible,** el que matemáticamente se representa como: C $= \{x : x \in X, \; g_j(x) \geq 0, j = 1, 2, \dots, m\}$, representa las restricciones impuestas al programa, donde X es un subconjunto del espacio vectorial en el que se plantea el problema. En principio, puede suceder que el conjunto C fuese vacío, esto es que no exista $x \in X$ que satisfaga a a la vez a todas las condiciones de factibilidad. Suponga que C $\neq \emptyset$, entonces el programa no lineal, consiste en maximizar la función objetivo: $f(x)$ sobre el conjunto no vacío C. Si el punto $\bar{x}$, que resuelve el programa pertenece al interior de la región $g_h(x) \geq 0$, decimos que la restricción h es **inactiva,** contrariamente las restricciones $g_i(x) \geq 0$ son **activas,** si se verifica que $g_i(\bar{x}) = 0$.

El conocido problema de Lagrange o Euler, que consiste en maximizar una cierta función real, con restricciones siempre activas $g_j(x) = 0, \ j = 1, 2, \ldots m$; no puede considerarse como un caso particular del programa no lineal expuesto sustituyendo las restricciones $g_j(x) = 0$ por $g_j(x) \geq 0$ y $-g_j(x) \geq 0$. Veremos que esto inhabilita la mayoría de las llamadas condiciones de cualificación. Las siguientes preguntas son las que aparecen naturalmente en un programa de maximización no lineal:

1. ¿ Es el conjunto $\mathcal{C}$ no vacío ? Esto es, ¿existe un punto factible ?
2. ¿Existe una solución $\bar{x}$? Es decir, ¿existe un punto en el conjunto factible que maximice a la función objetivo?
3. En caso de existir solución, ¿ es ésta única?
4. ¿ Existen algoritmos que permitan hallar la solución?

6.1 Programación cóncava. Punto silla

En esta sección caracterizaremos la solución del programa de maximización no lineal, cuando se trata de maximizar una función (objetivo) cóncava sujeto a restricciones que definen un conjunto convexo. Comenzaremos analizando el comportamiento de **punto silla,** el que caracterizará a la solución de nuestro problema. Su característica principal es la de definir un máximo relativo de una determinado función restringido a una determinada dirección y ser a la vez un mínimo de la misma función en otra dirección.

Definición 6.1 *Sea $\Phi(x, y)$ una función a valores reales definida en $X \times Y$ donde $x \in X$ e $y \in Y$. Un punto $(\bar{x}, \bar{y})$ en $X \times Y$ se llama* **punto de silla** *de $\Phi(x, y)$ si:*

$$\Phi(x, \bar{y}) \leq \Phi(\bar{x}, \bar{y}) \leq \Phi(\bar{x}, y), \ \forall x \in X, y \in Y$$

Es conocida la imagen que representa al punto silla a través del dibujo de una silla de montar, figura 1.2, pero no es esta la única posibilidad como se demuestra en la figura 1.3.

Ejemplo 6.2 *La siguiente descripción corresponde a la figura 1.3, donde el $(0, 0)$ es un punto de silla:*

$$\Phi(x, y) = 1 - x + y, \ \ 0 \leq x \leq 1, 0 \leq y \leq 1.$$

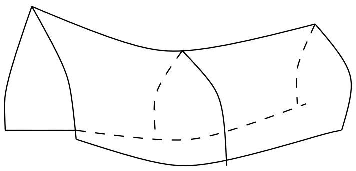
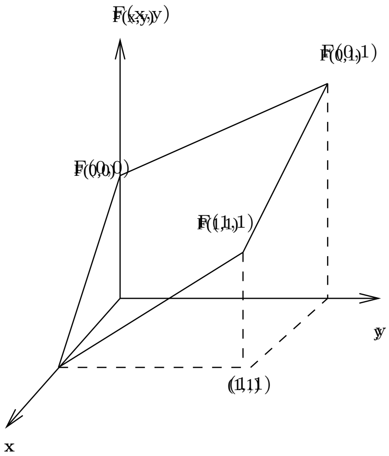


Figure 1: Una silla de montar, representa un punto silla.



7 El teorema de Kuhn-Tucker

El objetivo de esta sección es la demostarción rigurosa del teorema de Kuhn-Tucker, lo que nos permitirá su utilización rigurosa para resolver diferentes tipos de problemas de optimización. Comenzaremos considerando algunos teoremas previos, que nos permitirán demostrar el teorema de Kuhn-Tucker en forma clara.

7.1 Teoremas previos al de Kuhn-Tucker

Al siguiente teorema lo llamaremos *teorema fundamental* pues muchas de las técnicas que utilzaremos para resolver problemas de optimización que involucren funciones cóncavas o convexas se basan en él.

Teorema 7.1 (Teorema Fundamental) *Sea X un conjunto convexo en R^n y sean $f_1, f_2, ..., f_m$ funciones cóncavas, reales, definidas en X. Si el sistema:*

$$f_i(x) > 0, \; i = 1, 2, ..., m$$

no admite soluciones $x \in X$, entonces existen coeficientes $p_1, p_2, ...p_m$, con $p_i \geq 0$ que no se anulan simultaneamente tales que:

$$\sum_{i=1}^{m} p_i f_i(x) \leq 0 \quad \forall x \in X$$

Demostración: Dado un punto $x \in X$ definimos el conjunto Z_x como:

$$Z_x = \{(z_1, z_2, \dots, z_m) \in R^m;\ z_i < f_i(x)\ \ i = 1, 2, ..., m\}$$

Consideremos ahora el conjunto Z definido por:

$$Z = \cup_{x \in X} Z_x$$

El conjunto Z no contiene al cero y es convexo. En efecto, si contuviera al cero el sistema $f_i(x) > 0,\ i = 1, 2, ..., m$, tendría solución, lo que es una contradicción. Para probar la convexidad, suponga que $z_1 \in Z_{x_1}$ y $z_2 \in Z_{x_2}$ entonces

$$\lambda z_1 + (1 - \lambda) z_2 \leq \lambda f(x_1) + (1 - \lambda) f(x_2) \leq f(\lambda x_1 + (1 - \lambda) x_2).$$

Entonces el teorema de separación, obtenemos que existe $\bar{p}$ tal que $\langle \bar{p}, z \rangle \geq \langle \bar{p}, 0 \rangle = 0$ para todo $z \in Z$

Como z_j puede tomarse tan negativo como se quiera, obtenemos $\bar{p} \leq 0$. Consideramos a continuación $p = -\bar{p}$. Se sigue que $\langle p, z \rangle \leq 0$ para todo $z \in Z$. Finalmente consideremos $z_j = f_j(x) - \epsilon_j$. Obtenemos $\sum_{i=1}^{m} p_j [f_j(x) - \epsilon_j] \leq 0$, esto supone que $\sum_{i=1}^{m} p_j f_j(x) < \epsilon$, donde $\epsilon = \sum_{i=1}^{m} p_j \epsilon_j$. Como esta relación vale para cualquiera sea ϵ concluimos con que $\sum_{i=1}^{m} p_j f_j(x) \leq 0$. []

7.2 Teorema de Kuhn-Tucker

Teorema 7.2 *(Kuhn-Tucker, Uzawa): Sean $f, g_1, g_2, ..., g_m$ funciones reales y cóncavas definidas en un conjunto convexo $X \subset R^n$. Supongamos que en $\bar{x}, f$ alcanza su máximo, sujeto a las restricciones $g_i(x) \geq 0,\ i = 1, 2, ..., m$. Entonces, existen coeficientes, $\bar{p}_0, \bar{p}_1, \dots, \bar{p}_m$, ninguno de ellos negativo y no todos iguales a cero, tales que:*

$$\bar{p}_0 f(x) + \bar{p} g(x) \leq \bar{p}_0 f(\bar{x})\ \ \forall x \in X$$

donde $\bar{p} = (\bar{p}_0, \bar{p}_1, \dots, \bar{p}_n)$ y $g(x) = (g_1(x), g_2(x), ..., g_m(x))$.

Además $\bar{p} g(\bar{x}) = 0$. Pudiéndose elegir los coeficientes $\bar{p}_i$ tales que $\sum_{i=0}^{m} \bar{p}_i = 1$.

Demostración: La prueba de este teorema, sigue de aplicar el teorema fundamental al sistema:

$$\begin{array}{l} g_i(x) \geq 0,\ i = 1, 2, ..., m \\ f(x) - f(\bar{x}) > 0 \end{array}$$

para $\bar{x} \in X$ que maximiza a f. Este sistema no tiene solución, por lo tanto, tampoco lo tendrá el sistema que resulta del anterior sustituyendo desigualdades por desigualdades estrictas. La primera afirmación resulta entonces de aplicar ahora el teorema fundamental a este último sistema.

Para probar la segunda afirmación, considere la desigualdad:

$$\bar{p}_0[f(x) - f(\bar{x})] + \sum_{i=1}^{m} \bar{p}_i g_i(x) \leq 0, \quad \forall x \in X$$

o lo que es lo mismo:

$$\bar{p}_0 f(x) + \bar{p}g(x) \leq \bar{p}_0 f(\bar{x}), \ \forall x \in X$$

sustituya ahora x por $\bar{x}$ en la desigualdad anterior. La afirmación sigue del hecho de que $\bar{p} \geq 0$ y $g(\bar{x}) \geq 0$.[]

Corolario 7.3 *Supongamos que se satisfacen ademaś de las condiciones del teorema anterior las siguientes:*

(**S**) *Existe $\bar{x} \in X$ tal que $g_i(\bar{x}) > 0, \ i = 1, 2, ..., m$. Entonces tenemos que $\bar{p}_0 > 0$.*

(**SP**) *Si se satisface* (**S**) *así como las condiciones del teorema, entonces existen coeficientes $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_m)$ tales que: $\Phi(x, \bar{\lambda}) \leq \Phi(\bar{x}, \bar{\lambda}) \leq \Phi(\bar{x}, \lambda)$, para todo $x \in X$ y $\lambda \geq 0$, donde $\Phi(x, \lambda) = f(x) + \lambda g(x)$, con $\lambda = (\lambda_1, \lambda_2, ...\lambda_m)$. Siendo $\lambda_j = \frac{\bar{p}_j}{P_0}$.*

En otras palabras, $(\bar{x}, \bar{\lambda})$ es un punto de silla para $\Phi(x, \lambda)$.

Demostración: La demostración se hace por el absurdo, suponga que $\bar{p}_0 = 0$, obtenemos $\bar{p}g(x) \leq 0, \ \ \forall \ x \in X$, en particular $\bar{p}g(\bar{x}) \leq 0$, lo que obviamente contradice el supuesto de ser $\bar{x}$ un punto interior de la región factible.

La segunda parte sale inmediatamente de que: $\bar{p}g(\bar{x}) = 0, g(\bar{x}) \geq 0$ y $p \geq 0$. []

Observación 7.4 *La condición* **S** *es conocida como* **condición de Slater.** *Si esta condición no se verifica, entonces el corolario no necesariamente se cumple. Considere el siguiente problema: Maximizar $f(x) = x, x \in R$ sujeto a: $g(x) = -x^2 \geq 0$. Claramente $\bar{x} = 0$ es la solución, no obstante puede observarse que, $(0, \bar{\lambda})$ no es punto de silla de $\Phi(x, \lambda) = f(x) + \lambda g(x)$ para ningún λ no negativo. (Obsérvese que $\frac{\partial \Phi}{\partial x}$ evaluada en $x = 0$ es positiva, para cualquier valor de λ.)*

La función $\Phi(x, \lambda) = f(x) + \lambda g(x)$, o la no tan usual $\Phi(x, p) = p_0 f(x) + pg(x)$ definida en $X \times R^m_+$, es llamada **Lagrangiana.**

Se tiene el recíproco del corolario recién probado:

Teorema 7.5 *Sean $f, g_1, g_2, ..., g_m$ funciones reales definidas en $X \in R^n$. Si existe un punto $(\bar{x}, \bar{\lambda}) \in X \times R^m_+$ punto de silla para el Lagrangiano del problema de maximizar f, sujeto a $g_j(x) \geq 0, j = 1, 2, ...m$ entonces:*

i) El punto $\bar{x}$ maximiza f sujeto a las restricciones.

ii) $\bar{\lambda}g(\bar{x}) = 0$. Nótese que esta condición implica $\bar{\lambda}_i g_i(\bar{x}) = 0 \ \forall i \in (1, 2, ..., m)$.

Demostración: La desigualdad $\Phi(\bar{x}, \bar{\lambda}) \leq \Phi(\bar{x}, \lambda), \forall \lambda \in R^m_+$ implica $\bar{\lambda}g(\bar{x}) \leq \lambda g(\bar{x}), \forall \lambda \in R^m_+$, por lo tanto (basta elegir $\lambda = 0$) $\bar{\lambda}g(\bar{x}), \leq 0$, como $\lambda \in R^m_+$ y $\bar{x}$ verifica las restricciones, se sigue (ii).

Luego, $g_j(\bar{x}) \geq 0, \ j = 1, 2, ..., m$, tomando $\lambda = 0$ en la desigualdad de arriba, obtenemos $\bar{\lambda}g(\bar{x}) = 0$. Esto es (ii) se cumple.

Ahora de la desigualdad $\Phi(x, \bar{\lambda}) \leq \Phi(\bar{x}, \bar{\lambda}), \forall x$ y de (ii) se sigue que $f(\bar{x}) - f(x) \geq 0$ para toda $x \in X$ tal que $g_j(x) \geq 0, j = 1, 2, ..., m$. []

Observación 7.6 *Nótese que los multiplicadores de Kuhn-Tucker, correspondientes a restricciones inactivas, son ceros. La condiciones (ii) del teorema a, suelen llamarse* condiciones de complementariedad.

Combinando este teorema y el corolario anterior, obtenemos el siguiente resultado:

Teorema 7.7 *Sean $f, g_1, g_2, ..., g_m$ funciones cóncavas definidas sobre un conjunto convexo $X \in R^n$. Suponga que se cumple la condición de Slater (S). Entonces $\bar{x}$ maximiza a f sujeto a $g_j(x) \geq 0, j = 1, 2, ..., m$, si y sólo si existe $\bar{\lambda} \geq 0$ tal que $(\bar{x}, \bar{\lambda})$ que es un punto silla del Lagrangiano $\Phi(x, \lambda)$ para $x \in X$ y $\lambda \geq 0$ y tal que $\bar{\lambda}g(\bar{x}) = 0$.*

8 Caracterización por las condiciones de primer orden

En esta sección asumiremos que las funciones $f, g_1, g_2, ..., g_m$ reales, tienen dominio en un conjunto $X \subset R^n$, y que son diferenciables en X.

Discutiremos sobre las siguientes condiciones:

(M) *(Condición de Máximo)* Existe $\bar{x} \in X$ que maximiza a f sujeto a: $g_j(x) \geq 0, j = 1, 2, ..., m$; con $x \in X$.

(SP) *(Condición de Punto de Silla)* Existe $(\bar{x}, \bar{\lambda})$ en $X \times R^m_+$, punto de silla de $\Phi(x, \lambda) = f(x) + \lambda g(x)$.

(**CPO**) *(Condiciones de Primer Orden)* Existe un $(\bar{x}, \bar{\lambda})$ en $X \times \Omega^n$ tal que $\bar{f}_x + \bar{\lambda}\bar{g}_x = 0,\ g(\bar{x}) \geq 0$ y $\bar{\lambda}g(\bar{x}) = 0$.

Obsérvese que la condición $\bar{f}_x + \bar{\lambda}\bar{g}_x = 0$ puede escribirse como:

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \bar{\lambda}_j \frac{\partial g_j}{\partial x_i} = 0, i = 1, 2, ..., n$$

donde las derivadas parciales deben evaluarse en $x = \bar{x}$.

Teorema 8.1 *Sean $f, g_1, g_2, ..., g_m$ funciones reales y diferenciables en un subconjunto abierto $X \in R^n$.*

(i) La condicón (SP) implica las (CPO)

(ii) Si además las funciones son todas cóncavas, y X convexo entonces se cumple el recíproco, esto es (CPO) implica (SP).

Demostración:

(i) Por hipótesis $\Phi(x, \bar{\lambda}) \leq \Phi(\bar{x}, \bar{\lambda})$ para todo $x \in X$. Se sigue que $\Phi_x(\bar{x}, \bar{\lambda}) = \bar{f}_x + \bar{\lambda}.\bar{g}_x = 0$. Además por el teorema 1.4, se sigue que $\bar{\lambda}.g(\bar{x}) = 0$. Finalmente de $\Phi(\bar{x}, \bar{\lambda}) \leq \Phi_x(\bar{x}, \lambda)$, para todo $\lambda \geq 0$ se sigue que $g(\bar{x}) \geq 0$.

(ii) Por ser $\Phi(x, \lambda)$ combinación lineal no negativa de funciones cóncavas, ella misma es cóncava, luego como $\Phi_x(\bar{x}, \bar{\lambda}) = 0$ se sigue que $\Phi(x, \bar{\lambda}) \leq \Phi(\bar{x}, \bar{\lambda})$ para todo $x \in X$. Por último, por ser $\bar{\lambda}.g(\bar{x}) = 0$, obtenemos $\Phi(\bar{x}, \bar{\lambda}) \leq \Phi(\bar{x}, \lambda)$ para todo $\lambda \geq 0$, y $g(\bar{x}) \geq 0$.[]

Combinando este teorema con el corolario del teorema de Kuhn-Tucker Uzawa obtenemos el siguiente resultado:

Teorema 8.2 *Sean $f, g_1, g_2, ..., g_m$ funciones cóncavas y diferenciables en un conjunto convexo abierto, $X \in R^n$. Asuma que se cumpla la condición de Slater, entonces (M) si y sólo si (CPO).*

Observación 8.3 *En economía es frecuentemente necesario, usar la restricción $x \geq 0$, junto a otras como $g_j(x) \geq 0$, $j = 1, 2, ..., m$, en este caso la condición (CPO) se presenta de la siguiente manera:*

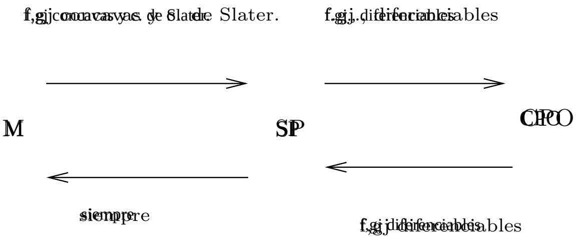


Figure 2: Relaciones entre, (M), (SP), (CPO

Existen $\bar{x}, \bar{\lambda}$ y $\bar{\mu}$ tales que:

$$\frac{\partial f(\bar{x})}{\partial x_i} + \bar{\lambda}\frac{\partial g(\bar{x})}{\partial x_i} + \mu = 0, \forall i = 1, 2, \ldots, n$$

$$\bar{\lambda} g(\bar{x}) + \bar{\mu}\bar{x} = 0$$

y además:

$$g(\bar{x}) \geq 0, \quad \bar{x} \geq 0$$

$$\bar{\lambda} \geq 0, \quad \bar{\mu} \geq 0$$

Esta condición puede transformarse en la expresión equivalente:

Existen $\bar{x} \geq 0, \bar{\lambda} \geq 0$ tales que:

$$\tfrac{\partial f(\bar{x})}{\partial x_i} + \bar{\lambda}\tfrac{\partial g(\bar{x})}{\partial x_i} \leq 0, \qquad [\tfrac{\partial f(\bar{x})}{\partial x_i} + \bar{\lambda}\tfrac{\partial g(\bar{x})}{\partial x_i}]\bar{x}_i = 0 \quad i = 1, 2, \ldots, n$$

$$\bar{\mu}_i \bar{x}_i = 0 \;\; \bar{\lambda}_i g_i(\bar{x}) = 0, \qquad\qquad g(\bar{x}) \geq 0$$

La figura 2.7 ilustra la condición anterior; se trata de maximizar una función cóncava f siendo $X = R$ con la restricción $x \geq 0$. El dibujo muestra que $\bar{x} = 0$ con $f_x(\bar{x}) < 0$, y $f_x(\bar{x})\bar{x} = 0$. (En este problema $f_x(\bar{x})$, deberá ser sustituído cuando corresponda por la derivada derecha, $f_x^+(\bar{x})$.)

8.1 Condiciones de cualificación

El teorema (8.2) analiza la relación entre la existencia de un máximo y las condiciones de primer orden pasando por la condición de Slater. El trabajo original de Kuhn y Tucker, relaciona la existencia del máximo y las condiciones de primer orden directamente. Se buscan condiciones necesarias para la optimalidad, es decir condiciones que aseguran que un óptimo verifica las llamadas CPO. Las condiciones que aseguran esto, son llamadas condiciones de cualificación. Sin estas condiciones no es necesariamente cierto que si un punto x^* resuelve un determinado programa de optimización las condiciones de primer orden (CPO) $f_x(x^*) + \lambda^* g(x^*) = 0$ $(*)$ $\lambda^* g(x^*) = 0,$ $\lambda^* \in$

R^n_+, se verifiquen. Es decir la afirmación *las CPO, son necesarias para un máximo restringido* no es necesariamente cierta. No obstante si es cierto que si modificamos la expresión (*) y consideramos en su lugar $\lambda_0^* f_x(x^*) + \lambda^* g(x^*) = 0$ con $(\lambda_0^*, \lambda^*) \in R^{n+1}$ admitiendo la posibilidad de que $\lambda_0^* = 0$ entonces la condición de máximo siempre implica estas CPO *modificadas*. En definitiva, las condiciones de calificación son las que garantizan que λ_0^* sea positiva, lo que permite elegir $\lambda_0 = 1$.. Obsérvese que basta la independencia lineal de $\bigtriangledown g_i(x^*)$ para que $\lambda_0^* \neq 0$.

Para seguir este camino precisamos definir **Punto Regular** o las llamadas condiciones de cualificación para un punto en el conjunto factible. Existen varios tipos de este tipo de condiciones. Introduciremos primeramente las definidas por Kuhn-Tucker.

Definición 8.4 Condición de Kuhn-Tucker *Sea C el conjunto factible,* $C = \{x : x \in X, g_j(x) \geq 0,\ j = 1, 2, \ldots, m\}$*. Un punto* $\bar{x} \in C$ *se dice* **regular***, si para cualquier* $\hat{x} \in X$ *tal* $g'_j(\bar{x})(\hat{x} - \bar{x}) \geq 0$ *para todo* $j \in E$*, donde* $E = \{j : j \in 1, 2, \ldots, m;\ g_j(\bar{x}) = 0\}$ *puede definirse una curva* $h(t);\ 0 \leq t \leq 1$ *con* $h(0) = \bar{x}$ *en el interior de* C*, tal que* $h'(0) = \alpha(\hat{x} - \bar{x})$ *para algún* $\alpha > 0$*.*

Esta condición es ilustrada en el ejemplo que sigue.

Considere las siguientes restricciones para un problema de maximización. $g_1(x) = x_2,\ g_2(x) = -x_2 + (x_1 - 1)^3$. Observe que el punto $x^* = (1, 0)$ no es regular, en este punto los gradientes de g_1 y g_2 apuntan en direcciones opuestas, se pierde la independencia lineal de los gradientes. Para este caso $\lambda_0^* = 0$. La condición de regularidad es una condición de independencia lineal, elimina a los puntos donde los límites de la región factible forman cúspides.

En esta sección haremos uso del llamado lema de Farkas.

Lema 8.5 Lema de Farkas-Minkowski *Sea* $a^1, a^2, ..., a^m$ *y* $b \neq 0$*, elementos de* R^n*. Supongamos que* $bx \geq 0$ *para todo* x *tal que* $a^i x \geq 0,\ i = 1, 2, ..., m$*. Entonces existe* $\bar{\lambda} \in R^m_+$ *que no se anulan simultaneamente y tales que* $b = \sum_{1=1}^n \bar{\lambda}_i a_i$*.*

Demostración: Sea $K = \left\{\sum_{i=1}^n \alpha_i a^i,\ \alpha_i \in R_+, i = 1, 2, ..., n\right\}$, el cono polihédrico generado por $a^1, a^2, ..., a^n$. Entonces K es un conjunto cerrado. Queremos ver que $b \in K$. Supongamos que no. Entonces existen p y β tales que $px \geq \beta$ *y* $pb < \beta$. Como K es un cono se tiene que si $x \in K$ entonces $\theta x \in K$ para todo $\theta \in R_+$ luego $p(\theta x) \geq \beta$ luego $px \geq \beta/\theta$ haciendo $\theta \to \infty$ se concluye que $px \geq 0$ para todo $x \in K$ y $pb < 0$. Esto contradice el supuesto. Luego $b \in K$.[]

Teorema 8.6 *Sean* $f, g_1, g_2, ..., g_m$ *funciones reales diferenciables en un subconjunto* X *de* R^n *abierto. Supongamos que* $\bar{x}$ *es regular, entonces si* $\bar{x}$ *verifica (M) se sigue la condición (SP).*

La demostración de este teorema puede verse en [Takayama, A.]. Daremos aquí una idea de la prueba.

Idea de la demostración: Sea $\bar{x}$ un máximo. La diferenciabilidad de f implica que:

$$f(x) - f(\bar{x}) = f'(\bar{x})(x - \bar{x}) + o(\|x - \bar{x}\|)$$

Definimos $E = \{j \in 1, 2, ..., m : g_j(\bar{x} = 0\}$, como es regular, existe $\hat{x}$ tal que $g_j(\bar{x})(\hat{x} - \bar{x}) \geq 0$ para todo $j \in E$ y la correspondiente función $h(t)$ tal que $x - \bar{x} = h(t) - h(0) = h'(0)t + o(|t|)$. Donde $o(|t|) = o(\|x - \bar{x}\|)$ Se sigue que $f(x) - f(\bar{x}) = f'(\bar{x})h'(0)t + o(|t|)$ Por lo tanto $f(x) - f(\bar{x}) = \alpha f'(\bar{x})(\hat{x}-\bar{x})t+o(|t|)$. Como $h(t) \in C$ para t suficientemente pequeño, $f(x)-f(\bar{x}) \leq 0$. Luego para t suficientemente pequeño $f'(\bar{x})(\hat{x}-\bar{x}) \leq 0$. Como a la vez se verifica $g_j'(\bar{x})(\hat{x}-\bar{x}),\ \forall J \in E$. Luego por el teorema de Farkas, existe $\bar{\lambda} \in R^e_+$, siendo $e =$ cardinal de E, tal que $-f'(x) = \sum_{i \in E} \bar{\lambda}_i g_i'(\bar{x})$.[]

Como puede verse a través de una serie de ejemplos, [Takayama, A.] y también entre los ejercicios del curso, si no se tiene la condición de regularidad no es posible en primera instancia, concluir que un punto que verifique la condición de ser un máximo para un programa de maximización sea un punto de silla del correspondiente lagrangiano.

En la búsqueda de condiciones de caulificación esto es de condiciones que impliquen que la solución de un programa de optimización verifique (CPO), la condición de regularidad de Kuhn-Tucker juega un papel importante. Pero no es particularmente fácil de verificar. Afortunadamente, no es la única que garantiza que en un máximo se satisfacen las condiciones de primer orden, obsérvese que el teorema (8.2) muestra que la concavidad de las funciones f y $g_j, j = 1, 2, \ldots, m$, y la condición de Slater, garantizan que en un máximo restringido se cumplen las condiciones de primer orden, esto es (M) implica (CPO). Demostraremos más adelante que la cuasiconcavaidad de las funciones $g_i, i = 1, 2, ..., m$ más la existencia de un punto interior al conjunto factible, sustituye a la condición de regularidad. Este es el teorema de Arrow-Enthoven. Condiciones alternativas pueden encontrarse en [Arrow, K.; Hurwicz, L,; Usawa, H.] y en [Takayama, A.].

Las siguientes condiciones introducidas en [Arrow, K.; Hurwicz, L,; Usawa, H.] dispensan la utilización de las condiciones de regularidad en Kuhn-Tucker. Las llamaremos condiciones fuertes (AHUf).

1. **Condición (CV).** Todas las $g_i(x), i = 1, 2, ..., m$, son convexas y existe $\bar{x}$ tal que $g_i(\bar{x}) < 0$.

2. **Condición (CO).** Todas las $g_i(x), i = 1, 2, ..., m$, son cóncavas y además se cumple la condición de Slater.

3. **Condición (CI).** El conjunto $C = \{x : g(x) \geq 0, j = 1, 2, ..., m, x \in R^m\}$ es convexo y posse un punto interior, y además $g_j'(x^*) \neq 0$ siendo para toda j tal que g_j sea efectiva.

4. **Condición (R).** La matriz formada por los gradientes de las restricciones efectivas evaluadas en el óptimo, tiene rango máximo.

La demostración de la afirmación puede encontrarse en la obra citada Si bien no daremos la demostración de estas condiciones el lector puede intentar demostrar que las condiciones anteriores implican la condición (AHU) que introduciremos a continuación la cual es suficiente para que M implique (CPO).

El siguiente teorema introduce una condición alternativa a la de regularidad dada por Kuhn-Tucker, es de más fácil verificación e implica a las cuatro condiciones dadas anteriormente.

Teorema 8.7 Arrow-Hurwicz-Uzawa *Sean $f, g_1, g_2, ..., g_m$ funciones reales diferenciables n definidas en $X \subset R^n$ abierto y convexo. Supongamos que la condición (AHU) se verifica. Entonces si en x^* se alcanza el máximo del programa de maximización, max $f(x)$; s.s. $g_i(x) \geq 0,\ i = 1, 2, ..., m;\ x \in X$ existe λ^* tal que (x^*, λ^*) las (CPO) se verifican.*

Definición 8.8 Condición AHU: *Sea $E = \{j \in \{1, 2, ..., m\} : g_j(x) = 0\}$. La condición AHU se verifica si existe un $h^* \in R^n$ tal que*

$$\bigtriangledown g_j(x^*)h^* \geq 0 \ \ \forall j \in J \ \ y \ \ \bigtriangledown g_j(x^*)h^* > 0 \ \ \forall j \in J'$$

Donde $J \subset E$ representa el conjunto de las restricciones activas convexas, mientras que $J' \subset E$ representa el conjunto de las restricciones activas no convexas.

Demostración:(Prueba de Hurwitz)

(i) Supongamos que $g_i(x) = 0,\ \forall i$. Entonces $\lambda_i = 0,\ \forall i$. Entonces $f'(x^*) = 0$ y CPO se verifica.

(ii) Admitamos entonces que $E \neq \emptyset$. Si g_j es convexa, entonces $g_j(x^*+th^*)-g_j(x^*) \geq g_j'(x^*)(th^*)$. Si g_j es no convexa entonces la diferenciabilidad permite escribir $g_j(x^* + th^*) - g_j(x^*) \geq g_j'(x^*)(th^*)$ para t suficientemente pequeño.

Siendo x^* un máximo para f se sigue que $f(x^* + h^*t) - f(x^*) \leq 0$. Luego $f'(x*)h^* \leq 0$. Ahora como h^* verifica (AHU), usando el teorema de Minkowsky-Farkas, se concluye en que existen $\lambda_j; j \in E$ tales que $-f'(x^*) = \sum_{j \in E} \lambda_j g_j(x^*)$. Luego eligiendo $\lambda_j = 0$ para todo $j \notin E$ se concluye el teorema.[]

Observación 8.9 *Algunas relaciones entre las diferentes condiciones de cualificación. Observe el lector que si la condición (CV) de las condiciones (AHUf) se verifica entonces basta considerar*

$h^* = 0$, y entonces (AHU) sigue. Si la condicón (CO) se verifica, la función f se comportará, en un entorno de x^* como una función cóncava, la demostración se obtiene usando el teorema fundamental. La convexidad de C en la condición (CI) implica que las funciones $g_i(x)$ sean cuasicóncavas, nuevamente el teorema fundamental termina la demostración. El hecho de que R implica (AHU) se encuentra más abajo en estas notas.

A los efectos de aclarar la relación entre las diferentes condiciones de regularidad, consideremos el siguiente ejemplo debido a Slater.

Ejemplo 8.10 *Considere el siguiente programa de optimización:*

$$\max_{x \in R^2} f(x) = x_2$$
$$s.a: \ -x_2 - x_1^2 \geq 0$$

Claramente $x^* = (0,0)$ es la solución de este problema, no obstante $x^* = (0,0)$ no es punto estacionario del Lagrangiano, no obstante es cierto que $\lambda_0^* f_x(0,0) + \lambda^* \bar{g}_x(0,0) = 0$ basta elegir $\bar{\lambda}^* = 0$.

El siguiente teorema muestra que la independencia lineal de los gradientes evaluados en el punto x^*, solución del programa de maximización, de las restricciones activas, es una condición suficiente para que x^* verifique las (CPO).

Teorema 8.11 *Suponga que x^* satisface las condiciones de máximo (M), para el problema de maximizar $f(x)$ con las condiciones $g_i(x) \geq 0, \ i = 1, 2, ..., m$. Suponga también que el rango de la matriz de m filas por n columnas, $\{\frac{\partial g_j}{\partial x_i}\}$ evaluada en $x = x^*$, es igual a m_e, estamos suponiendo $m_e \leq n$ donde m_e es el número de las restricciones activas. Entonces para $\bar{x}$ se satisfacen las condiciones de primer orden.*

Demostración: Sea $E = \{j \in \{1, 2, ..., m\} : g_j(x^*) = 0\}$. El hecho de que el rango de la matriz $\mathcal{R}$, cuyas columnas son los gradientes de las restricciones evaluados en x^*, sea máximo, equivale a que se cumple la condición AHU y por lo tanto la demostración se concluye a partir del teorema (8.7). Basta ver que si ubicamos en los primeros lugares de $\mathcal{R}$ las columnas correspondientes a las restricciones efectivas, obtenemos una submatriz $\mathcal{SR}$ invertible de rango m_e. Luego el sistema $\mathcal{SR}h = u$ donde u es un vector de unos tiene solución. Formamos ahora el h^* de la condición AHU, haciendo $h^* = (h, 0)$ es decir completamos las $n - m_e$ coordenadas con ceros. Luego se verifica trivialmente, para las restricciones efectivas las desigualdades $\bigtriangledown g_j(x^*) h^* > 0$.[]

Terminemos esta sección con el siguiente comentario. Como el lector ya habrá persivido las condiciones de cualificación son condiciones referidas a las restricciones unicamente, y no a las características de la función objetivo, de ahí que sean conocidas como condiciones de cualifiación de las restricciones.

9 Teorema de Lagrange

El problema clásico de Lagrange, hace referencia a restricciones de igualdad, esto es, se trata de hallar un $\bar{x} \in X$ abierto, que maximice a una cierta $f(x)$ sujeto a que $g_j(x) = 0, j = 1, 2, \ldots, m$. Discutiremos en esta sección la relación entre las (CPO) y la solución del problema de Lagrange. Este problema puede reducirse al anterior, considerando el cumplimiento simultaneo de las condiciones: $g_j(x) \geq 0$ y $-g_j(x) \leq 0, j = 1, 2, \ldots, m$, no obstante esto no resuelve el problema. Obsérvese que no pueden aplicarse las condiciones de regularidad, pues ningún punto que satisfaga las restricciones, será regular, tampoco estamos en las condiciones de la sección anterior, a menos que las restricciones de igualdad correspondan a funciones afines, pues si g_j es cóncava, $-g_j$ es convexa.Obsérvese también que aunque aparezca $2m$ restricciones, solamente existen m diferentes. Por otra parte las condiciones de cuaificación no pueden ser las mismas, pues el hecho de que un punto x^* resuelva el programa de maximización supone acá restricciones de igualdad, es decir maximiza en bajo condiciones más restrictivas que las de un programa con desigualdades.

El siguiente teorema muestra no obstante que, bajo condiciones de cuasiconcavidad, condiciones análogas a las (CPO) para el problema con restricciones de desigualdad, son también suficientes para la optimalidad del punto x^*. Estas condiciones están dadas por la existencia un vector $\lambda \in R^m_+$, siendo m es número de restricciones $g_i(x) = 0$ del programa de maximizar f tal que verifica $\bigtriangledown f(x^*) + \sum_{i=1}^m \lambda_i \bigtriangledown g_i(x^*) = 0$.

Teorema 9.1 *Sean $f, g_1, g_2, ..., g_m$ funciones cuasicóncavas y diferenciables e $X \in R^n$. Considere el siguiente programa de maximización, (P):*

$$max f(x) \;\; s.a. \;\; g_i(x) = 0 \; i = 1, 2, ..., m.$$

Suponga que x^* verifica las siguientes condiciones, existen escalares $\lambda_1^*, ..., \lambda_m^*$ no negativos, tales que $f'(x^*) + \lambda^* g(x^*) = 0$. y además que $g_i(x^*) = 0$ entonces x^* resuelve el problema (P).

Demostración: Para todo $x \in X$ tal que $g_i(x) \leq g_(x^*)$ se verifica que $\langle \bigtriangledown g(x^*), (x^* - x) \rangle \leq 0$. A partir de las condiciones de Kuhn-Tucker, se llega a:

$$\langle \bigtriangledown f(x^*), (x^* - x) \rangle = +\sum_{i=1}^m \lambda_i^* \langle \bigtriangledown g_i(x^*), (x^* - x) \rangle \leq 0.$$

Finalmente la cuasiconcavidad de f implica que $f(x^*) \geq f(x)$.[]

Observación 9.2 *Obsérvese que la condición requerida para que el teorema se aplique es que de la condición* $\langle \bigtriangledown f(x_0), (x - x_0) \rangle$ *puede obtenerse la condición* $f(x) \geq f(x_0)$. *Esta propiedad la tienen las funciones cuasicóncavas, y por lo tanto las cóncavas, pero hay una clase de funciones que no son cuasicóncavas que también la verifican, son las llamadas funciones pseudocóncavas. Las que se definen precisamente por satisfacer esta propiedad. Ejemplo de este tipo de funciones es* $f(x) = -x^3 - x$, *pues* $\langle (x_1 - x_2), -(3x_2^2 + 1) \rangle \geq 0$ *implica* $x_2 \geq x_1$ *y por lo tanto* $f(x_2) \geq f(x_1)$.

Un problema importante que aparece en la maximización con restricciones de igualdad es que de antemano no sabemos el signo de los multiplicadores correspondientes a estas restricciones en el lagrangiano. Considere el problema maximizar una determinada función sujeto a condiciones de igualdad, admitamos que estas condiciones estaán dada por funciones $g_i; i = 1, 2, ..., n$ afines Podemos entonces desdoblar las restricciones, reemplazando cada restricción $g_j(x) = 0$ por las siguientes dos condiciones de desigualdad: $g_j(x) \geq 0, \quad -g_j(x) \geq 0$. El lagrangiano para este problema es

$$\Phi(x, \lambda) = f(x) + \sum_{j=1}^{m} (\mu_j - \nu_j) g_j(x) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x)$$

donde $\lambda_j = \mu_j - \nu_j, j = i, 2, \ldots, m$ Aunque las condiciones de primer orden, conllevan la no negatividad de los multiplicadores μ_j y ν_j, no podemos afirmar nada sobre el signo de λ_j.

Ejemplo 9.3 *Considere el problema:*

$$\max_{x \in R^2} x_1 x_2 \ \ \textit{sujeto a}: \ x_1^2 + x_2^2 - 1 = 0$$

El Lagrangiano para este problema es: $\phi(x, \lambda) = x_1 x_2 + \lambda(x_1^2 + x_2^2 - 1)$, de donde se obtienen las (CPO):

$$\frac{\partial \Phi}{\partial x_1} = \bar{x}_2 + 2\bar{\lambda}\bar{x}_1 = 0, \quad \frac{\partial \Phi}{\partial x_2} = \bar{x}_1 + 2\bar{\lambda}\bar{x}_2 = 0 \ \ (x_1^2 + x_2^2 - 1) = 0$$

Resolviendo este sistema obtenemos la solución: $x_1 = x_2 =^{+}_{-} 1/\sqrt{2}$ y $\lambda = \frac{1}{2}$. Evaluada en este punto la matriz $\{\frac{\partial g_j}{\partial x_i}\} = (2\bar{x}_1, 2\bar{x}_2)$ tiene rango 1, por lo que la condición del teorema se cumple.

Como en la sección anterior, también aquí nos interesa encontrar condiciones que hagan que la solución de un problema de Lagrange verifique las (CPO), en este caso no ponemos condiciones en el signo de los multiplicadores.

Teorema 9.4 Condición de cualificación para el problema de Lagrange. *La independencia lineal de los gradientes de las restricciones es una condición suficiente para que si* x^* *resuelve el problema de maximizar* $f(x)$ *s.a* $h_i(x) = 0, \ i = 1, 2, ..., m.$ *verifique las (CPO) correspondientes.*

La siguiente demostración de este teorema es sencilla pero usa propiedades de las superficies regulares, una exposición de la teoría correspondiente y de este teorema puede verse en [Lima, L.]. No obstante daremos una breve demostración del teorema sin entrar mayormente en la discusión de los temas propios de la teoría de superficies.

Demostración Comenzamos con la definición de punto crítico de una función diferenciable. Acá M representa una superficie, y por f/M representamos la función restringida a M. Un **punto crítico** de la función $f : M \to R$ es un punto p tal que $df(p) = 0$ es decir tal que se verifica $grad\ f(p) \cdot v = 0$ para todo $v \in T_pM$ esto es que el gradiente de f en p es ortogonal al plano tangente en dicho punto. Entre los puntos críticos de $f : M \to R$ se encuentran los máximos y los mínimos locales de f pues si p es uno de esos puntos entonces para todo camino $\lambda : (-\epsilon, \epsilon) \to M$ con $\lambda(0) = p$ se debe verificar que $(f(\lambda))'(0) = 0$.

Supongamos ahora que la superficie $M \subset R^{m+n}$ está dada por la imagen inversa $M = \phi^{-1}(c)$ de un valor regular, donde $c = (c_1, c_2, ..., c_n)$ de la aplicación $\phi : U \to R^n$ de clase C^1 con $U \subset R^{m+n}$ si las funciones coordenadas son $\phi_1, \phi_2, ..., \phi_n$ entonces la condición de ser c **valor regular** significa que los vectores $\bigtriangledown\phi_1, \bigtriangledown\phi_2, ..., \bigtriangledown\phi_n$ son linealmente independientes, y por lo tanto constituyen una base del espacio vectorial perpendicular a T_pM. De aquí que el punto $p \in M$ es un punto crítico de f/M si y solamente si el vector $\bigtriangledown f(p)$ es una combinación lineal de los vectores $\bigtriangledown\phi_(p)$, es decir si y solamente si existen números reals $\lambda_1, ..., \lambda_n$ tales que

$$\bigtriangledown f(p) = \lambda_1 \bigtriangledown \phi_1(p) + \lambda_2 \bigtriangledown \phi_2(p) + ... + \lambda_n \bigtriangledown \phi_n(p).[]$$

Un caso interesante es aquel en que aparecen restricciones de igualdad y de desigualdad.

$$\max_{x \in X} f(x)$$

$$\text{sujeto a:} \quad g_j(x) \geq 0, \quad j = 1, 2, \ldots, m,$$

$$h_k(x) = 0, \quad k = 1, 2, \ldots, l.$$

Sea $\bar{x}$ una solución para este problema, y sea E el conjunto de índices tales que $g_j(\bar{x}) = 0$ (es decir las restricciones efectivas), sea m_e ese número. Para $m_e + l \leq n$ tenemos el siguiente resultado:

Teorema 9.5 *Sea $\bar{x} \in (M)$. Supongamos que el rango de la matriz $(m_e + l) \times n$*

$$\begin{bmatrix} \frac{\partial g_j}{\partial x_i} \\ \\ \frac{\partial h_k}{\partial x_i} \end{bmatrix} j \in E, k = 1, 2, \ldots, l$$

evaluada en $\bar{x}$ es $m_e + l$. Entonces se tiene que $\bar{x}$ es un punto estacionario para el Lagrangiano:

$$\Phi(x, \lambda, \mu,) = f(x) + \lambda g(x) + \mu h(x)$$

Es decir, en $\bar{x}$ se cumplen (CPO), con $\bar{\lambda} \geq 0$ mientras que $\bar{\mu}$ puede ser positivo o negativo.

Para una demostración de este teorema ver [Bazaraa, M.S.; Shetty, C.M.]. No obstante puede hacerse una demostración utilizando el teorema anterior. Basta considerar como restricciones de igualdad también a todas las restricciones efectivas, y repetir el razonamiento hecho en la demostración del teorema (9.4).[]

Presentaremos a continuación una aplicación del teorema.

Ejemplo 9.6 *Se trata de elegir $(x_1, x_2) \in R^2$ en el problema:*

$$\begin{array}{rll} \textit{Maximizar} & x_1 x_2 & \\ \textit{sujeto a:} & g(x) = x_1 + 8x_2 - 4 & \geq 0 \\ & h(x) = x^2 - 1 + x_2^2 - 1 & = 0 \end{array}$$

El Lagrangiano para este problema es: $\Phi(x, \lambda, \mu) = x_1 x_2 + \lambda(x_1 + 8x_2 - 4) + \mu(x^2 - 1 + x_2^2 - 1)$ y las (CPO):

$$\begin{aligned} \frac{\partial \bar{\Phi}}{\partial x_1} &= \bar{x}_2 + \bar{\lambda} + 2\bar{\mu}\bar{x}_1 = 0 \\ \frac{\partial \bar{\Phi}}{\partial x_2} &= \bar{x}_1 + 8\bar{\lambda} + 2\bar{\mu}\bar{x}_2 = 0 \\ \bar{x}_1^2 + \bar{x}_2^2 - 1 &= 0 \\ \bar{\lambda}(\bar{x}_1 + 8\bar{x}_2 - 4) &= 0 \\ \bar{\lambda} &\geq 0 \end{aligned}$$

Resolviendo este sistema obtenemos: $\bar{x}_1 = \bar{x}_2 = \frac{1}{\sqrt{2}}, \bar{\lambda} = 0$ y $\mu = -\frac{1}{2}$. Nótese que el multiplicador $\bar{\mu}$ que corresponde a la restricción de igualdad es negativo. Puede verse fácilmente que la restricción $g(x)$ está inactiva en $(\bar{x}_1, \bar{x}_2)$. Luego la condición sobre el rango se verifica en la matriz $(\frac{\partial h}{\partial x_1}, \frac{\partial h}{\partial x_1}) = (2x_1, 2x_2)$ que es igual a 1 en $\bar{x}_1 = \bar{x}_2 = \frac{1}{\sqrt{2}}$, por lo tanto es satisfecha.

Ejemplo 9.7 *(a) Observe que si bien el punto $(2, 0)$ resuelve el problema de*

$$\textit{max } x_1 \;\; \textit{s.a.}: \;\; x_1 \geq 0; \;\; x_2 = (2 - x_1)^2,$$

no es un punto de silla del correspondiente lagrangiano.

(b) Considere ahora el siguiente problema:

$$max\ x_1\ \ s.a.:\ \ x_1 \geq 0;\ \ 0 \leq x_2 \leq 1;\ \ x_2 = -1 + (2 - x_1)^2$$

Observe que para este caso el punto $(2, 0)$ *resuelve el problema y satisface la condición SP. ¿Cuál es la diferencia entre el problema (a) y el (b) ?.*

10 Progamación cuasicóncava

En la mayoría de los problemas de la teoría económica relacionados con programas de maximización, las funciones objetivos son cuasicóncavas y se agrega la condición de que el resultado debe pertenecer al ortante positivo del espacio en el que se modela la economía, por ejemplo R^n_+. Por ejemplo funciones de utilidad o beneficios, programas de minimización pueden presentarse como programas de maximización cambiando los signos de las respectivas funciones objetivos. Así el problema de minimizar el costo $C(y, w)$ para producir una determinada cantidad de producto y a precio de los factores w fijos, puede transformarse en el problema equivalente de maximizar $-C(y, w)$.

En estas sección probaremos el teorema de Arrow-Enthoven de 1961, que extiende las condiciones suficientes para que $x^* \in R^n$ definidas en el teorema de Kuhn Tucker al caso donde la función objetivo es cuasicóncava y restricciones de convexidad también definidas por funciones cuasicóncavas. Veremos también como corolario de este teorema una caracterización de la cuasiconcavidad para funciones reales definidas en subconjuntos convexos de R^n que hace uso del llamado hessiano orlado de f evaluado en x^*. Naturalmente que los resultado válidos para programas cuasicóncavos, lo son para cóncavos.

- Usaremos la siguiente notación:

  La función $G : X \to R^m$ queda definida por $G(x) = (g_1(x), g_2(x), ..., g_m(x))$ para todo $x \in X$ subconjunto convexo de R^n.

  Siendo λ un vector en R^n la notación $\lambda G(x)$ representa el producto interno de estos dos vectores.

  $G'(x) = (\bigtriangledown g_1(x), \bigtriangledown g_2(x), ..., \bigtriangledown g_m(x))$ y para $\lambda \in R^m$ escribiremos $\lambda G'(x) = \sum_{i=1}^m \lambda_i \bigtriangledown g_i(x)$.

  Sea $X \subset R^n$ diremos que $h \in R^n$ es una dirección admisible para $x \in X$ si existe $\epsilon > 0$ y $x + vh \in X$ para v $0 \leq v \leq \epsilon$.

Haremos uso en lo que sigue de los siguientes teoremas:

Teorema 10.1 *Sea f una función real diferenciable en $X \subset R^n$. Si f tiene un máximo en $x* \in X$ entonces la derivada de f en x^* se anula en toda dirección admisible.*

Demostración: Definimos $\phi : [0, \epsilon] \to R$ como $\phi(v) = f(x + vh)$. Esta es una función real de una variable real, que alcanza su extremo en $v = 0$, luego $\phi'(0) = \frac{d}{dv} f(x^* + vh)h|_{v=0} = 0$. Es decir $\frac{d}{dv} f(x^*)h = 0 \; \forall h$ admisible.[]

Teorema 10.2 *Supongamos ahora que $X^0 = \{x \in R^n : G(x) = 0\}$ donde $G : R^n \to R^m$ es una función dos veces diferenciable. Sea $L(x) = f(x) + \lambda G(x)$ donde $\lambda \in R^m$. Supongamos que existe $B_r(x^*)$ una bola de centro x^* y radio r, tal que*

$$f(x^*) > f(x) \;\; para \;\; todo \;\; x \in B_r(x^*) \cap X, \; x \neq x^*. \tag{15}$$

Entonces la matriz hessiana de $L(\cdot, \lambda^*) : R^n \to R^m$, evaluada en x^*, $L''_{xx}(x^*, \lambda^*)$ es definida negativa, para todo h tal que $\bigtriangledown L(x^*, \lambda^*)h = 0$.

Demostración: Observe que (15) es equivalente a

$$L(x^*, \lambda^*) > L(x, \lambda^*) \;\; para \;\; todo \;\; x \in B_r(x^*) \cap X, \; x \neq x^*.$$

Use ahora las condiciones suficientes para máximo local. (Condiciones de segundo orden).[]

El siguiente teorema define una condición suficiente para que x^* sea un máximo de un programa de maximización con función objetivo cuasicóncava y restricciones convexas. Diremos que la función $G : R^n \to R^m$ es cuasicóncava si y solamente si los son cada una de sus funciones coordenadas. Y notaremos como $G'(x)$ el jacobiano de G evaluado en x.

A continuación enunciaremos el teorema fundamental de la programación no lineal en el caso diferenciable. Primeramente en el caso general, luego en el caso en que nos restringimos a trabajar en el ortante positivo de R^n.

Consideremos el problema:

$$Maximizar \;\; f(x) \;\; sujeto \, a \;\; G(x) \geq 0. \tag{16}$$

Teorema 10.3 (Arrow-Enthoven, 1961). *Sea f una función real diferenciable en X subconjunto abierto y convexo de R^n. y sea $G : G : R^n \to R^m$ cuasicóncava en X, y supongamos además que se verifica la condicón de Slater. Entonces, si $x^* \in R^n$ es una solucón para el problema (16, existe $y^* \in R^m_+$ tal que*

$$\begin{aligned} L_x(x^*, y^*) &= f'(x^*) + y^* G'(x^*) = 0, && (i) \\ L_y(x^*, y^*) &= G(x^*) \geq 0, && (ii) \\ y^* L_y(x^*, y^*) &= y^* G(x^*) = 0, && (iii) \end{aligned} \tag{17}$$

Recíprocamente, *sea* $x^* \in X$ *e* $y^* \geq 0$ *tales que verifican las condiciones (i), (ii), y (iii) dadas en (16). Entonces si* f *es cuasicóncava,* x^* *resuelve (16).*

Demostración: Es inmediato que si x^* resuelve el problema de maximización, entonces la condición (ii) de (16) se verifica. Si además $f'(x^*) = 0$ entonces basta elegir $y^* = 0$ para que todas las condiciones de (16) se verifiquen. Podemos entonces asumir que $f'(x^*) \neq 0$. Siendo x^* optimal para f en $F = \{x : G(x) \geq 0\}$ se sigue de la diferenciabilidad de f que

$$f(x^* + \lambda(x - x^*)) - f(x^*) = f'(x^*)(\lambda(x - x^*)) + r(\lambda(x - x^*)) \leq 0$$

para todo $\lambda \in (0, 1]$. Haciendo ahora $\lambda \to 0$, se sigue que

$$f'(x^*)(\lambda(x - x^*)) \leq 0, \ \forall x \in F. \tag{18}$$

Sea $I = \{i : 1 \leq i \leq n : g_i(x^*) = 0\}$. Es fácil ver que $I \neq \emptyset$ pues en otro caso por la continuidad de de las g_i, sería posible elegir $\mu > 0$ y una bola de centro x^* y radio δ, $B(x^*, \delta)$ tal que $x^* \pm \mu e \in B(x^*, \delta)$ y por lo tanto, a partir de (18) se obtiene $f'(x^*) \leq 0$ y $f'(x^*) \geq 0$, por lo que $f'(x^*) = 0$ lo que contradice la afirmación de que $f'(x^*) \neq 0$. Entonces para todo $i \in I$ se verifica que $g_i(x) \geq g_i(x^*)$ por la cuasiconcavidad de cada g_i, $i \in I$ se sigue que

$$\bigtriangledown g_i(x^*)(x - x^*) \geq 0, \ \forall x \in F. \tag{19}$$

Entonces por el teorema fundamental (7.1) aplicado a las ecuaciones (18) y (19), existe $v_i \geq 0$, $i \in I$ tales que

$$\frac{\partial f}{\partial x_i}(x^*) + \sum_{i \in I} v_i \frac{\partial g}{\partial x_i}(x^*) = 0.$$

La igualdad ((i), 17) se concluye ahora eligiendo $y_i^* = v_i$ cada vez que $i \in I$ y cero en otro caso. Para estos mismos valores de y^* como facilmente puede verificarse, se tiene la igualdad ((ii)17).

Recíprocamente: A partir de (17) se sigue que $\bigtriangledown f(x^*)(x - x^*) \leq 0, \forall\ x \in F$. Entonces la cuasiconcavidad de f implica que $f(x^*) \geq f(x), x \in F$ lo que establece la optimalidad de x^* en F. Para ver esto, observe que $I \neq \emptyset$ elija como anteriormente y_i^*, para cada $i \in I$,la condición (iii) impone $y_i^* = 0\ \forall i \notin I$. A partir de (i) se sigue que:

$$\bigtriangledown f(x^*)(x - x^*) \leq -\sum_{i \in I} y_i^* \bigtriangledown g(x^*)(x - x^*) \leq 0.$$

Ahora la cuasiconcavidad de f termina la demostración.[]

Corolario 10.4 (Arrow-Enthoven, 1961). *Sea f una función real cuasicóncava definida en R^n_+ y sea $G : G : R^n \to R^m$ cuasicóncava Consideremos el problema:*

$$Maximizar \;\; f(x) \;\; sujeto\, a \quad G(x) \geq 0, \; x \geq 0. \tag{20}$$

Siendo $G(x) = (g_1(x), g_2(x), ..., g_m(x))$. Supongamos además que existe $\bar{x} \in R^n : \bar{x}_i > 0, i = 1, 2, ..., n$ y que además verifica $g_i(\bar{x}) > 0, \; \forall i = 1, 2, ..., m$. Entonces si $x^* \in R^n$ con $G(x^*) \geq 0$ es una solución para este problema existe $y^* \in R^m_+$ tal que

$$\begin{aligned} &f'(x^*) + y^* G'(x^*) \leq 0, && (i) \\ &y^* G(x^*) = 0, && (ii) \\ &[f'(x^*) + y^* G'(x^*)]\, x^* = 0, && (iii) \end{aligned} \tag{21}$$

Recíprocamente, sea $x^* \in X$ e $y^* \geq 0$ tales que verifican las condiciones (i), (ii), y (iii) dadas en (21). Entonces si f es cuasicóncava, y si una de las siguientes condiciones es satisfecha,

a) $f'_k(x^*) = \partial f(x^*)/\partial x_k < 0$ para al menos un k; y tal que existe $\bar{x} \in R^n_+$ tal que $\bar{x}_k > 0$ y $G(\bar{x}) > 0$.

b) $f'_j(x^*) = \partial f(x^*)/\partial x_j > 0$ para al menos un j tal que $x^*_j > 0$;

c) $f'(x^*) \neq 0$ y f es dos veces diferenciable en un entorno de x^*;

entonces x^* es optimal.

Demostración: Para ver que las condiciones $(i), (ii)$ y (iii) de (21) son necesarias, definimos $g_{m+k}(x) = x_k, \; k = 1, 2, ...n$. Considere entonces

$$F = \{x \in R^n : g_j(x) \geq 0; j = 1, 2, ..., m + n\}.$$

De esta forma

$$L(x, y) = f(x) + \sum_{i=1}^{m+n} y_i g_i(x) = f(x) + \sum_{i=1}^{m} y_i g_i(x) + \sum_{i=1}^{n} y_i x_i$$

donde $\mu_i = y_{m+i}$. Repitiendo el razonamiento anterior hecho en el teorema (10.3) se obtiene facilmente que $L_x(x^*, y^*) = 0$ y que por lo tanto $\bigtriangledown f(x^*) + \sum_{i \in I} y^*_i \bigtriangledown g_i(x^*) + \sum_{i \in E} \mu_i e_i \leq 0$, siendo: $I = \{i \in \{1, 2, ..., m\} : g_i(x^*) > 0\}$, $E = \left\{j \in \{1, 2, ..., n\} : x^*_j > 0\right\}$, (este conjunto puede ser vacío), e_i es el vector canónico en la i-ésima dirección, obsérvese que $\bigtriangledown x_i = e_i$. Finalmente del hecho de que $\mu_i = 0$ si $x^*_i > 0$, (esto por (ii) en (17)) y en otro caso $\mu_i \geq 0$ se tiene que

$\frac{\partial f}{\partial x_j}(x^*) + \sum_{i=1}^{m} \frac{\partial g_i}{\partial x_j}(x^*) \leq 0$, para todo $i = 1, 2, ..., m$. Si la desigualdad es estricta entonces debe ser $\mu_i > 0$ y por lo tanto $x_i = 0$. Por lo que la condición (iii) en (21) se cumple siempre.

Recíprocamente:

a) A partir de (21) (análogamente a lo hecho en la demostración del teorema (10.3)) se sigue que $y_i^* = 0$ para todo $i \notin I$, de donde resulta que tenemos que:

$$\bigtriangledown f(x^*)(x - x^*) \leq -\sum_{i \in I} y_i^* \bigtriangledown g_i(x - x^*) \leq 0 \tag{22}$$

para todo $x \in F$. Consideremos $x^1 = x^* + e^k$ donde $e^k = (0, ..., 0, e, 0, ..., 0)$ para e suficientemente pequeño, y cualquier $x \in F$ se tiene que $x(\lambda) = (1 - \lambda)x + \lambda x^1$, $0 \geq \lambda \geq 1$. Las siguientes desigualdades se verifican:

$$f'(x^*)(x^*(\lambda) - x^*) = (1 - \lambda)f'(x^*)(x - x^*) + \lambda \frac{\partial f}{\partial x_k}(x^*)e < 0$$

Haciendo ahora $\lambda \to 0, x(\lambda) \to x$ es decir que $f'(x^*)(x - x^*) < 0$ de donde por por la cuasiconcavidad se sigue que $f(x) \leq f(x^*)$.

b) Supongamos que no existe i que verifica (a). Por lo tanto $f'(x^*)x > 0$ para todo $x \in R_+^n$. Entonces $f'(x^*) \geq 0$. Razonando como en (22) obtenemos que $0 \leq f'(x^*)(x - x^*) \leq 0$. Escribiendo para $x \in R_+^n$ para el que $G(x) \geq 0$

$$x(\alpha) = (1 - \alpha)x, \;\; y, \;\; x^*(\alpha) = (1 - \alpha)x^*, \; 0 \leq \alpha \leq 1.$$

Se sigue que :

$$f'(x^*)(x^*(\alpha) - x^*) = -\alpha f'(x^*)x^* < 0, .$$

$$f'(x^*)(x(\alpha) - x^*(\alpha)) = (1 - \alpha)f'(x^*)(x - x^*) \leq 0$$

Sumando las desigualdades anteriores, luego haciendo $\alpha \to 0$ y usando la condición de cuasiconcavidad (10) se obtiene el resultado.

c) Para el caso restante, ver [Kemp, M.; Kimura, Y].

A continuación mostraremos algunas relaciones existentes entre funciones cuasicóncavas y sus hessianos orlados.

Definición 10.5 *Llamaremos* **hessiano orlado** *de* $f : X \to R$ *en* $x \in X$ *donde* $X \subset R^n$, *a la matriz*

$$[H(x)] = \begin{bmatrix} 0 & f'(x) \\ f'^t(x) & f''(x) \end{bmatrix}$$

donde $f'(x) = \bigtriangledown f(x)$ *y* $f''(x)$ *representa la matriz hessiana de* f *evaluada en* x

Teorema 10.6 (Arrow-Enthoven, 1961). *Para función f dos veces diferenciable definida en un subconjunto convexo X de R^n se verifica que:*

1. *Si f es cuasicóncava en X entonces $[H(x)]$ es semidefinida negativa, en cada $x \in X$*

2. *Es suficiente para que f sea cuasicóncava en X que $[H(x)]$ sea definida negativa en cada $x \in X$ para todo $h \in R^n$ tal que $f'(x)h = 0$.*

Demostración:

1. Si $f'(x) = 0$ para $x \in X$ entonces la condición necesaria se verifica trivialmente. Supongamos entonces que $f'(x_0) \neq 0$ para algún $x_0 \in X$. Consideremos entonces el problema de: $Maximizar\ f(x) \quad s.a. \quad f'(x_0)(x - x_0) \geq 0, \ x \geq 0$. Definamos entonces $G(x) = f'(x_0)(x_0 - x) \geq 0$ las condiciones definidas en ((21), (a), (b), (c)) con $y^* = 1$ se verifican para este caso. Por lo tanto x_0 es una solución para el problema de maximización. Luego $f(x_0)$ es un máximo bajo la restricción $f'(x_0)(x_0 - x) \geq 0$. Usando ahora el teorema (10.2) concluimos con que $H(x_0)$ es semidefinida negativa.

2. Sea $[H](x_0)$ definida negativa para todo h tal que $f'(x_0)h = 0$. Siendo f dos veces diferenciable en $x_0 \in X$ se tiene que $hf''(x_0)h < 0$. Para x suficientemente próximo a x_0 tal que $h = (x - x_0)$ se tiene que $f(x) < f(x_0)$ esto es $f(x_0)$ es un máximo local estricto para f sujeto a la condición $f'(x_0)(x - x_0) = 0$. Esto significa que para todo $\bar{x} \in X$ se verifica que $\bar{x}$ maximiza a $f(x)$ sujeto a la restricción $f'(\bar{x})(x - \bar{x}) = 0$.

   Supongamos haber demostrado que lo mismo ocurre si consideramos la restricción $f'(\bar{x})(x - \bar{x}) \leq 0$. Llamemos a este, supuesto (A). Sea entonces $x_1 \in X$ y sea x_2 una combinación convexa de x_0 y x_1. Por el supuesto (A), debe verificase que $f(x_2) \geq f(x)$ para todo x tal que $f'(x_2)(x - x_2) \leq 0$. Obsérvese que, por ser x_2 una combinación convexa de x_0 y x_1, no pueden verificarse simultáneamente siguientes desigualdades:

   $$f'(x_2)(x_0 - x_2) > 0, \ y \ f'(x_2)(x_1 - x_2) > 0;$$

   ni

   $$f'(x_2)(x_0 - x_2) < 0, \ y \ f'(x_2)(x_1 - x_2) < 0;$$

   Luego será cierto que $f(x_2) \geq f(x_0)$ o bien $f(x_2) \geq f(x_1)$ es decir $f(x_2) \geq min\{f(x_0), f(x_1)\}$ por lo que f es cuasicóncava.

Falta entonces probar el supuesto (A). Consideremos la restricción definida por el conjunto $X_0 = \{x \in X : f'(x_0)(x - x_0) \leq 0\}$. Queremos probar que $f(x_0)$ es el máximo de f restringido a X_0.

Supongamos que no lo sea. En este caso existe $x_1 \in X_0$ tal que $f(x_1) > f(x_0)$. Sea $x(\lambda) = (1 - \lambda)x_0 + \lambda x_1$ y $g(\lambda) = f((1-\lambda)x_0 + \lambda x_1)$, se verifica que $g(1) > g(0)$ y que $g'(\lambda) = \bigtriangledown f(x(\lambda))(x_1 - x_0)$. Se tiene además que $g'(0) < 0$, luego existe $\bar{\lambda} \in (0,1)$ que es un mínimo relativo para g para el que se cumple $g'(\bar{\lambda}) = 0$. Es posible elegir un h suficientemente pequeo tal que $\bar{\lambda} + h \in (0,1)$. Puede verificarse que $x(\bar{\lambda} + h) - x(\bar{\lambda}) = h(x_1 - x_0)$. Por lo tanto, $\bigtriangledown f(x(\bar{\lambda}))x(\bar{\lambda} + h) - x(\bar{\lambda}) = 0$ por o tanto $f(x(\bar{\lambda} + h)) < f(x(\bar{\lambda}))$ es decir $g(\bar{\lambda} + h) < g(\bar{\lambda})$. De donde se concluye el absurdo.[]

11 Aplicaciones

Las aplicaciones de la programación no lineal en economía son abundantes y muy importantes, mostraremos primeramente que loo multiplicadores de Kuhn-Tucker pueden ser interpretados heuristicamente como los *precios de equilibrio*, luego presentaremos dos aplicaciones, válidas tanto por su importancia en la teoría económica como por sus enseñanzas para la técnica que estamos estudiando.

Interpretación de los multiplicadores de Kuhn-Tucker. Supongamos que la función objetivo representa un costo que se intenta minimizar. Los argumentos de la función son los insumos necesarios para producir cierto producto. Las restricciones están definidas por la cantidad de recursos disponibles para la producción. El programa será un programa de nimización con restricciones convexas del tipo $g_j(x) \leq 0, \ j = 1, 2, ..., m$. Se ajusta al programa de maximización tratado en las secciones anteriores, cambiando en forma adecuada signos positivos por negativos y minimización de la función objetivo por maximización de su opuesta. Análogamente las opuestas de las restricciones, consideradas aquí dadas por funciones convexas, son funciones cóncavas.

Supongamos que es posible modificar las restricciones, adquiriendo una cantidad adicional de recursos en aquellos casos en que el óptimo x^* implica que la restricción es activa, $g_j(x^*) = 0, j \in E$, siendo E el conjunto de restricciones activas. Supongamos que el precio de la unidad adicional del recurso j tiene un precio $v_j^*, \ j \in E$. Sea entonces v^* el vector precios unitarios de los recursos $v_j, j \in E$. Tal modificación resulta interesante si se tiene la desigualdad $f(y^*) + \langle v^*, v \rangle \leq f(x^*)$. Esto es, el costo de producir bajo las nuevas condiciones $g_j(y^*) = v_j, \ j \in E$, el que es igual al costo $f(y^*)$, de producir y^*, donde y^* minimiza f con las condiciones $g_j(x) \leq v_j$ más el costo $\langle v^*, v \rangle$ de adquirir la cantidad adicional de recursos v, es menor que el costo de producir x^* en la situación anterior. La desigualdad anterior corresponde a la siguiente: $\min_{x \in C} \left\{ f(x) + \sum_{j \in E} v_j^* g_j(x) \right\} < f(x^*)$ siendo $C = \{x : g_j(x) = v_j\}$.

Supongamos ahora que $v_j^* = \lambda_j^*, \ j \in E$ es decir, que el precio unitario del recurso j, está dado por el correspondiente multiplicador de Kuhn-Tucker. En estas condiciones se tiene que

$$\left\{f(x)+\sum_{i=1}^{m}\lambda_i^* g_i(x)\right\} \geq f(x^*),$$

es decir que

$$\min_{x\in C}\left\{f(x)+\sum_{i=1}^{m}\lambda_i^* g_i(x)\right\} = f(x^*),$$

es decir que a precios λ^* la mejor opción es x^*, por lo tanto no hay incentivos a modificar la producción.

Ejemplo 11.1 *El primer ejemplo será el problema de minimizar costos (MC) por parte de una firma competitiva, que produce un único producto.*

La firma elegirá una cierta cantidad a producir $y \in R$, y sus correspondientes insumos, que serán representados por un vector $x \in R^n$, en forma acorde con la técnica de producción que posee, la que será representada por una función $f : R^n \to R$. La firma elegirá $x \in R^n$ tal que:

$$\begin{aligned} &\text{Minimizar:} \quad wx && (\text{Maximizar:} \quad -wx) \\ &\text{sujeto a:} \quad f(x) \geq y && \text{and} \quad x \geq 0 \end{aligned} \tag{23}$$

donde, w es el vector de precios (dado), admitimos $w_i > 0$ $i = 1, 2, \ldots n$; x el vector de insumos, $y(> 0)$ nivel del producto deseado, $f(x)$ función de producción. Las condiciones de primer orden (CPO), para este problema son:

$$-w_i + \bar{\lambda}\tfrac{\partial f(\bar{x})}{\partial x_i} \geq 0 \quad [-w_i + \bar{\lambda}\tfrac{\partial f(\bar{x})}{\partial x_i}]\bar{x}_i = 0 \quad i = 1, 2, ..., n$$

$$\bar{\lambda}[f(\bar{x}) - y] = 0 \quad f(\bar{x}) - y \geq 0$$

Estas condiciones determinan la función de demanda, $\bar{x} = x(w, y)$ tanto cuanto a $\bar{\lambda}(w, y)$

El conjunto factible para este problema es C $= \{x : x \geq 0, f(x) \geq y\}$. Suponga que existe $\bar{x} \geq 0$, tal que $f(\bar{x}) > y$ esto es que se cumple la condición de Slater, (S). Si la función de producción es cóncava, entonces la condición (1) es necesaria para la existencia de un óptimo local, siendo además la función objetivo lineal, y por lo tanto cóncava, la condición es también suficiente para un óptimo global.

Siendo f cóncava, y si se cumple la condición de Slater, entonces la condición (1) es *necesaria y suficiente* para la existencia de un óptimo global.

Obsérvese que es perfectamente posible obtener una solución de esquina, esto es una solución donde se cumpla la desigualdad estricta en (1); $\bar{\lambda}f_i(\bar{x}) < w_i$, en este caso $\bar{x}_i = 0$. Puede pensarse que para la tecnología existente es un factor muy caro y que por lo tanto no se usará.

En el caso en que se cumpla la igualdad

$$\bar{\lambda}\frac{\partial f(\bar{x})}{\partial x_i} = w_i, \;\; y = f(\bar{x}) \tag{24}$$

la función demanda $\bar{x}_i = x_i(w, y)$ así como $\bar{\lambda} = \lambda(w, y)$, se determinan a partir del teorema de la función implícita a partir del sistema (2).

Ejemplo 11.2 *Nuestro segundo ejemplo, es el problema de maximizar la función de utilidad supuesta cuasicóncava de un consumidor quien es un tomador de precios, y que tiene un ingreso fijo M.*

El problema se plantea de la siguiente forma:

$$\text{maximizar} \quad u(x)$$

$$\text{sujeto a:} \quad \langle p, x\rangle \leq M \quad x \geq 0$$

donde $u(x)$ es la función de utilidad, p el vector de precios, (se admite algún $p_i < 0$) y M el ingreso monetario (> 0).

Observe que la restricción presupuestaria define un conjunto de restricciones convexo, y que la condición (c) del teorema (10.4) se verifica, desde que el consumidor prefiera más a menos. También se cumplen la condiciones (e) y (b') del mencionado teorema, esta última pues $G'(x) = -(p_1, p_2, ..., p_n) \neq 0$. Las condiciones de primer orden para este problema son:

$$\begin{array}{ll} \frac{\partial u(\bar{x})}{\partial x_i} - \bar{\lambda}p_i \geq 0 & [\frac{\partial u(\bar{x})}{\partial x_i} - \bar{\lambda}p_i]\bar{x}_i = 0 \\ \bar{\lambda}(px - M) = 0, & p\bar{x} - M \leq 0 \end{array} \tag{25}$$

donde $\bar{x} \geq 0$ y $\bar{\lambda} \geq 0$ Como en el caso anterior, las condiciones (3) determinan la demanda y el valor del multiplicador, en función de los parámetros, p, M.

En el caso de una función de utilidad cóncava 8alcanza cuasicóncava) las condiciones (3) son necesarias y suficientes para la existencia del óptimo, siempre que se satisfaga la condición de Slater, la que en este caso por ser $M > 0$ se cumple siempre.

Obsérvese que si hay algún tipo de **no saciedad local** como por ejemplo que $\frac{\partial u(\bar{x})}{\partial x_i} > 0$ (esto implica $\bar{\lambda} > 0$), entonces el consumidor gastará todo su ingreso, luego, en la restricción presupuestaria obtendremos la igualdad. No obstante esto no alcanza para evitar las soluciones $\bar{x}$ donde se cumple $\frac{\partial u(\bar{x})}{\partial x_i} < \bar{\lambda}p_i$. En este caso, como se muestra en la figura 3.10, el vector de precios, no sería paralelo al gradiente de la curva de indiferencia del consumidor. En este caso, como se desprende de las (CPO), $\bar{x}_i = 0$, esto es, hay un bien demasiado caro, lo que haría que el agente no lo consuma.

Proponemos al lector analizar las posibilidades, $\frac{\partial u(\bar{x})}{\partial x_i} < 0$ y $p_i \geq 0$ o $p_i \leq 0$, así como los casos: $\frac{\partial u(\bar{x})}{\partial x_i} > 0$ con $p_i < 0$ y $\frac{\partial u(\bar{x})}{\partial x_i} = \bar{\lambda} p_i < 0$, e interpretar económicamente cada uno de estos casos.

no obstante en economía es frecuente imponer condiciones que aseguren que la solución sea interior, es decir $x_i > 0$, $\forall i = 1, 2, \ldots, n$. La figura 3.11, ilustra esta situación. En este caso nos encontramos con la solución familiar:

$$\frac{\partial u(\bar{x})}{\partial x_i} = \bar{\lambda} p_i, \; i = 1, 2, \ldots, n; \;\; p\bar{x} = M \tag{26}$$

Siendo posible en este caso obtener la función de demanda y el valor del multiplicador de Lagrange, a partir de (4), por el teorema de la función implícita, en función de los parámetros (p, M). Interpretando el multiplicador como la utilidad marginal de la moneda, llegamos a la conclusión de que si se verificase la desigualdad $\frac{\partial u(\bar{x})}{\partial x_i} = \bar{\lambda} p_i$, no habría demanda por el bien i puesto esto supone $\bar{x}_i = 0$.

Dedicaremos la próxima sección al análisis de sensibilidad, y al llamado *teorema de la envolvente.*

12 Análisis de Sensibilidad

Se analizará las modificaciones en la solución de un problema de programación no lineal, cuando se modifican los parámetros del problema, por ejemplo (w, y) en el caso de minimizar el costo o (p, M) en el caso de maximizar la función de utilidad. Para considerar este problema en forma general trabajaremos con las siguientes funciones: $f(x, \alpha)$ y $g_j(x, \alpha), j = 1, 2, \ldots, n$, reales, continuamente diferenciables, definidas en $R^(n + l)$ donde $x \in R^n, \alpha \in R^l$, consideremos el problema de fijado α hallar x tal que:

$$\text{maximizar:} \quad f(x, \alpha)$$

$$\text{sujeto a:} \quad g_j(x, \alpha) \geq 0, \quad j = 1, 2, \ldots, n \quad x \geq 0$$

Las condiciones de primer orden para este problema son:

$$f_x(\bar{x}, \alpha) + \bar{\lambda} g_x(\bar{x}, \alpha) \leq 0, \quad [f_x(\bar{x}, \alpha) + \bar{\lambda} g_x(\bar{x}, \alpha)]\bar{x} = 0$$

$$\bar{\lambda} g(\bar{x}, \alpha) = 0, \qquad\qquad g(\bar{x}, \alpha) \geq 0$$

Por $f_x(\bar{x}, \alpha)$ representamos el gradiente de f evaluado en $(\bar{x}, \bar{\alpha})$, análogamente para $g_x(\bar{x}, \alpha)$ en todas las restricciones. Con $\bar{x} \geq 0, \bar{\lambda} \geq 0$.

El análisis de sensibilidad se refiere a las modificaciones que aparecen en la solución óptima $\bar{x}(\alpha)$ y $\bar{\lambda}(\alpha)$ ante modificaciones del parámetro α.

Consideraremos el caso más sencillo en el que la solución es interior, esto es $\bar{x}_i > 0, \forall i$ y $\bar{\lambda}_j > 0, \forall j$. Por este motivo las (CPO) quedan restringidas a las igualdades:

$$f_x(\bar{x}, \alpha) + \bar{\lambda} g_x(\bar{x}, \alpha) = 0 \tag{27}$$

$$g(\bar{x}, \alpha) = 0 \tag{28}$$

Estas $(n+m)$ ecuaciones, se usarán para obtener $(n+m)$ incógnitas : $\bar{x}_i = x_i(\alpha), i = 1, 2, \ldots, n; \bar{\lambda}_j = \lambda_j(\alpha), j = 1, 2, \ldots, m$.

A continuación definimos la **función máximo valor:**

$$F(\alpha) = f[x(\alpha), \alpha]$$

$$\Psi(\alpha) = f[x(\alpha), \alpha] + \lambda(\alpha).g[x(\alpha), \alpha]$$

Tenemos el siguiente teorema:

Teorema 12.1 Teorema de la envolvente *Sean F y Ψ diferenciables, entonces tenemos que: $F_\alpha = \Psi_\alpha = \Phi_\alpha$ esto es:*

$$\frac{\partial F(\alpha)}{\partial \alpha_j} = \frac{\partial \Psi(\alpha)}{\partial \alpha_j} = \frac{\partial \Phi(x, \lambda, \alpha)}{\partial \alpha_j}$$

donde $\Phi(x, \lambda, \alpha) = f(x, \alpha) + \lambda g(x, \alpha)$, es el lagrangiano.

Demostración: La prueba sale de aplicar la regla de la cadena y aplicar las (CPO), definidas en (5) y (6):

$$\Psi_\alpha = \Phi_x x_\alpha + \Phi_\lambda \lambda_\alpha + \Phi_\alpha = \Phi_\alpha$$

$$F_\alpha = f_x x_\alpha + f_\alpha = -\lambda g_x + f_\alpha = \lambda g_\alpha + f_\alpha = \Phi_\alpha$$

para obtener la segunda igualdad se usó (5), para la tercera se usó (6) pues $g_x x_\alpha + g_\alpha = 0$.

Para justificar el nombre de este teorema considere el siguiente cambio de variables: $\beta = (\alpha_1, \ldots, \alpha_{l-1})$ y las siguientes ecuaciones:

$$\phi(F, \beta, \alpha_l) = F - F(\alpha) = 0, \phi(\beta, \alpha_l) = \partial F(\alpha)/\partial \alpha_l = 0$$

Dejando α_l fijo $\phi(F, \beta, \alpha_l) = 0$ define un superficie en el plano (β, α_l), la cual junto con $\phi(\beta, \alpha_l) = 0$ define la **envolvente** a la familia de superficies parametrizada por α_l.

12.1 Aplicaciones a la miecroeconomía

Ejemplo 12.2 *Consideremos el problema del consumidor, asumamos $p > 0$ y sea $x(p, M)$ la solución del problema, siendo $\lambda(p, M)$ el multiplicador de Lagrange. En este caso $\alpha = (p, M)$. La función de utilidad del consumidor será: u. Definimos :*

$$U(p, M) = u[x(p, M)] \text{ (\textbf{función de utilidad indirecta})}$$

La función lagrangiana será: $\Phi = u(x) + \lambda(M - px)$. Por el teorema anterior tendremos:

$$\partial U/\partial M = \lambda(p, M)$$

Esto es *multiplicador de Lagrange representa la utilidad marginal del ingreso.* Por otra parte obtenemos:

$$\partial U/\partial p_j = -\lambda x_j, j = 1, 2, \ldots, n.$$

Esto significa que un aumento en el precio disminuirá la satisfacción del consumidor.

Las igualdades obtenidas nos permiten obtener la **identidad de Roy**:

$$\partial U/\partial p_j + x_j \partial U/\partial M = 0, j = 1, 2, \ldots, n.$$

Ejemplo 12.3 Interpretación del multiplicador *Consideremos el problema de elegir $x \in R^n$ tal que:*

$$maximizar: \quad f(x)$$

$$sujeto\ a: \quad g_j(x) \leq b_j, \quad j = 1, 2, \ldots, m$$

El lagrangiano será: $\Phi(x, \lambda, b) = f(x) + \lambda[b - g(x)]$

Aplicando el teorema anterior con. $F(b) = f[x(b)]$ obtenemos:

$$\frac{\partial F(b)}{\partial b_j} = \lambda_j, j = 1, 2, \ldots m$$

Entonces el j-ésimo multiplicador de Lagrange, representa el variación marginal del valor del máximo $F(b)$ con respecto a la j-ésima restricción. Interpretando b_j como la cantidad existente del j-ésimo recurso, λ_j representa la pérdida en el valor máximo $F(b)$ por la pérdida de una unidad del mencionado recurso, por lo tanto es un precio sombra.

Sugerencia considere el problema de minimizar el costo, ya presentado, y haga las interpretaciones correspondientes. Verifique que el multiplicador de Lagrange representa el costo marginal por elevar en una unidad el producto. Observe también que en el óptimo, la demanda por el j-ésimo factor es igual a la variación marginal del costo respecto a su precio.

El lector interesado podrá verificar también el lema de Hotelling, esto, en un problema de maximizar beneficios, para una firma competitiva que produce m productos a partir de n factores, la variación marginal del beneficio $\pi(p,w) = py - wx$ respecto del precio del i-ésimo producto es igual a la producción óptima del mismo, y que el beneficio decrece con el aumento del precio de los factores.

También puede obtenerse, a partir de la verificación anterior, asumiendo que la función beneficio es dos veces diferenciable, la llamada **simetría de Hotelling,** esto es:

$$\partial x_j / \partial w_i = \partial x_i / \partial w_j$$

Estas relaciones pueden encontrarse en [Varian, H.].

12.2 Tres teoremas de la microeconomía y el teorema de separación de convexos

Terminemos esta sección mostrando que tres de los lemas más importantes de la microeconomía son en definitiva una aplicación directa del teorema de separación de convexos. El lector interesado en aplicaciones del análisis convexo a la producción económica podrá encontrar en [Uribe, P.] un excelente texto.

Recordamos que como corolario del teorema de separación de convexos, podemos asegurar que si un conjunto $K \subset R^n$ es convexo entonces para todo $\bar{x} \in Fr(k)$ existe $\bar{p} \in R^n$ tal que $\bar{p}\bar{x} \leq \bar{p}x, \forall\ x \in K$. (Este es el teorema de existencia del hiperplano soporte ya estudiado en la sección correspondiente al teorema de separción de convexos). El hiperplano determinado por el vector $\bar{p}$, como el lector recordará, es denominado hiperplano soporte de K en $\bar{x}$. Llamaremos a los vectores $\bar{p}$ vectores soporte.

Sea H_k el conjunto de vectores soporte del conjunto convexo K. Definamos a continuación la aplicación soporte de K $\mu_k : H \to K$ definida como

$$\mu_k(p) = \{\bar{y} \in K : p\bar{y} = \inf(py),\ \forall\ y \in K\}.$$

Lema 12.4 (Lema de dualidad:) *Supongamos que K es estrictamente convexo. Entonces el gradiente de μ_k evaluado en p $\bigtriangledown\mu_k(p) = \bar{y}$.*

Demostración: Es inmediato ver que en este caso existe un único $\bar{y} \in \mu_k(p)$ y además que pertenece a la frontera de K.Como puede facilmente verse la función μ_k es lineal. Definamos ahora la función $\xi(\bar{p}) = \mu_k(\bar{p}) - \bar{p}y \leq 0\quad \forall\ y \in K$. Dicha función alcanza su máximo valor en $\bar{y}$. Luego $\bigtriangledown\xi(\bar{p}) = \bigtriangledown\mu(\bar{p}) - \bar{y} = 0.[]$

1. Consideremos un conjunto de producción Y de vectores insumo-producto posibles dada la tecnología existente. Supongamos que Y es un conjunto convexo y cerrado (supuestos comunes en microeconomía) ver [Varian, H.] por ejemplo. Un vector $y \in Y$ se llama eficiente si $y \in Fr[Y]$. **Teorema 1:** *Un vector $\bar{y}$ es eficiente si y solamente si existe $\bar{p}$ vector con todas sus coordenadas positivas, tal que $\bar{p}\bar{y} \geq \bar{p}y \ \forall \ y \in Y$. Demostración:* La suficiencia es obvia. El recíproco: La existencia es una inmediata consecuencia del teorema de existencia del hiperplano soporte mencionado en el comienzo del epígrafe. Si alguna coordenada $\bar{p}_i$ fuese negativa entonces la coordenada $i-$ésima de de $v \in Y$ podría ser arbitrariamente grande. Lo que es un absurdo.

2. Sea $\Pi : R^n_+ \to R$, la función de beneficios de una empresa cuyo conjunto tecnológico Y es convexo, $\Pi(\bar{p}) = max_y \bar{p}y = \bar{p}\bar{y}$. La existencia de $\bar{y}$ se deduce de las condiciones habituales exigidas para el conjunto Y de producción [16]. **Teorema 2** (Teorema de Hotteling:)*Sea Y estricatmente convexo, entonces $\bigtriangledown\Pi(\bar{p}) = \bar{y}$. Es decir la oferta de productos y la demanda de insumos a precios $\bar{p}$ son iguales respectivamente a la derivada de la función de beneficios respecto a sus respectivos precios. Demostración :* La estricta convexidad de Ymuestra que Π es un función, el resto es una aplicación directa del lemma (12.4) []

3. Supongamos una empresa que produce un único producto a partir de $n-1$ insumos diferentes cuyos precios están dados por $w \in R^{n-1}_+$, con tecnología definida por una función f convexa de forma tal que $Y = \{(q, z_1, ..z_{n-1}) \in R^n : q - f(z) \leq 0.\}$. Los costos quedan definidos por la función $c(\bar{w}, q) = min_{z:q \leq f(z)} \bar{w}z = \bar{w}\bar{z}$. **Teorema 3** (Teorema de Shepard:)*Si la demanda por insumos $z(w, q)$ consite de un único punto entonces $\bigtriangledown_w c(w, q) = \bar{z}$. Demostración:* Si admitimos la estricta convexidad de la función f la unicidad de la demanda de insumos se deduce inmediatamente. Lo restante es nuevamente una aplicacion del lema de dualidad.[]

4. Finalmente veamos la relación entre la demanda hicksiana y la función de gasto. Supongamos una economía con n bienes diferentes cuyo espacio de consumo es $X \subset R^n$ y un agente cuya función de utilidad estdada por $u : R^n \to R$, cóncava. Recordamos que se define como demanda hicksiana la demada que minimiza el gasto que permite alcanzar un nivel de utilidad determinado. A precios p fijos el gasto corresponde al valor de la cesta de bienes elegida es decir px si dicha cesta es p. Definimos como función de gasto $e : R^n \times R \to$ definda como:

$$e(\bar{p}, \bar{u}) = min_x \ px$$
$$s.a. : u(x) \geq \bar{u}$$

[16] Por ejemplo la existencia de $a \in R^n$ tal que $y \leq a, \ \forall \ y \in Y$

La cesta de bienes $\bar{x}$ que resuelve este problema es precisamente la demanda hicksiana: $\bar{x} = h(\bar{p}, \bar{u})$. Obsérvese que siendo la funci
on de utilidad cóncava, entonces el conjunto $\{x \in X : u(x) \geq u\}$ es convexo. A partir del lema de dualidad podemos concluir ahora que, $\bigtriangledown_p e(\bar{p}, \bar{u}) = h(\bar{p}, \bar{u})$.

13 Apéndices

En los siguientes apéndices se presentan algunas demostraciones y cometarios no realizados en el texto, pero a los efectos de simplifacar la vida del lector interesado, las inclumos acá. Su lectura no es esencial para la compresnión de las notas presentadas, pero algunas de las demostraciones son encierta forma paradigmáticas.

13.1 Apéndice I

Mostraremos acá la equivalencia entre las diversas topologías lineales de Hausdorff en $R^{\cdot}$ Sea (X, τ) un espacio topológico decimos que la topología τ es Hausdorff si separa puntos, es decir si dados dos puntos diferentes , x e y existen abiertos disjuntos que los contienen. Puede demostrarse que si τ es Hausdorff, entonces todo conjunto compacto es cerrado.

Sea X espacio vectorial. Decimos que una topología *tau* en X es lineal si las funciones

- $f : X \times X \to X$ definida como $f(x, y) = x + y$
- y $g : R \times X \to X$ definida como $g(\alpha, X) = \alpha x$

son τ continuas.

Teorema *Topologas equivalentes en R^n* Todo espacio vectorial de dimensión finita admite una única topología de Hausdorff.

demostración Sea $X = R^n$ y τ_1 una topología lineal y Hausdorff en X y sea τ_1 la topología generada por la norma euclideana $\|x\|_2$. Mostremos primeramente que $\tau_1 \subset \tau$. Para esto considere la función identidad $I : (X, \tau) \to (X, \tau_1)$. Si demostramos que esta función en continua habremos demostrado que $\tau_1 \subseteq \tau$. Conisdere $\{x_\alpha\}$ una sucesión de R^n tal que $x_\alpha \to 0$ de acuerdo a $\|x\|_2$. Esti supone que cada sucesión coordenada $\{x_{\alpha i}\}$ converge a cero. Luego $x_\alpha = \sum_{i=1}^n x_i e_i$ siendo e_i los vectores canónicos, converge a cero sugún τ_1. Luego la identidad I por ser lineal es continua como función de (X, τ) en (X, τ_1). Mostraremos acontiniuación que $\tau \subseteq \tau_1$ Sea $B = \{x \in R^n : \|x\|_2 < 1\}$ y sea $B = \{x \in R^n : \|x\|_2 = 1\}$. Por ser S compacto en τ y como por la parte anterior sabemos que $\tau_1 \subset \tau$ se sigue que S es τ_1 compacto. Luego, como τ_1 Hausdorff, es cerrado entonces S es cerrado. Luego como $0 \notin S$ existe V τ_1, entorno de 0 y tal que $V \cap S = \emptyset$. Además $V \subseteq B$, pues si no lofuera, existiría $x \in V$ tal que $\|x\|_2 > 1$, por lo que $x/\|x\|_2 \in V \cap S$ lo que es absurdo. Luego $V \subseteq B$ por los que $\tau \subseteq \tau_1$.[·].

13.2 Apéndice II

Haremos acá algunos comsntarios sobre la continuidad de funciones. En primer lugar es interesante destacar que el hecho de ser una función entre dos espacios topológicos una función continua, depende de las topologías elegidas. Una topología con pocos abiertos en el domonio de las funciones, dará como resultados *pocas* funciones continuas. Cunato más rica en abiertos, la topología conisderada en el dominio, más funciones continuas. En el caso de ser el dominio un subconjunto de R^n las funciones continuas serán las mismas para toda topología Hausdorff y lineal. En segundo lugar, es importante destacar que la defindión de continuidad en espacios métricos realizadas en términos de proximidad, es equivalente a la definición, introducida en el texto en la sección(1.5), que sólo hace referencia a conjuntos abiertos, si consideramos la topología cuya base está formada por las bolas abiertas con la métrica considerada.

References

[Accinelli, E.] *Elementos de la topología y de la teoría de conjuntos en la teoría del equilibrio general* UAM(A), Eon-Sociales (2005).

[Accinelli, E.; Brida, G. Plata, L.; Puchet, M] "Bienestar social, óptimos de Pareto y equilibrios walrasianos". A aparecer en número especial del Triemstre Econ'omico dedicado a las VI Jolate.

[Arrow, K.; Enthoven, A.] *Quasi Concave Programing.* Econmetrica **29** pp 779-800 (1961).

[Arrow, K.; Hurwicz, L,; Usawa, H.] *Constraint Qualifications in Maximization Problems.* Naval Research Logistics Quarterly Vol 8/2 junio (1961).

[Bazaraa, M.S.; Shetty, C.M.] *Foundations of Optimization* Lecture Notes in Economics and Mathematical Systems, vol 122, (1976).

[Brézis, H.] *Analyse fonctionelle* Masson Editeur, de París 1983.

[Coruzeix, J.P.; Ocaña, E; Sosa, W.] *Análisis Convexo* Monografías de IMCA No. 33 (2003).

[Kemp, M.; Kimura, Y] *Introduction to Mathematical Economics* Springer -Verlag, (1978).

[Lima, L.] *Curso de Analise.* vol 2. Ed. Sociedad Brasilera de Matemáticas, Projeto Euclides.

[Rockafellar, T.] *Convex Analysis.* Princeton University Press (1970)

[Takayama, A.] *Mathematical Economics.* Cambridge University Press (1994)

[Uribe, P.] *Análisis de Actividades y Teoría del Capital.* Universidad de Guadalajara , México (1997).

[Varian, H.] *Micreconomic Analysis.* Ed. Norton, New York. Third edition.