UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Ciencias y Humanidades

INTRODUCCION A LA PROGRAMACION NO-LINEAL Y SU APLICACION EN UN MODELO ECONOMICO

DAVID RIDELMAN TREJOS

Trabajo de investigación presentado para optar al Grado Académico de Licenciado en Matemática

Guatemala

1977

Vo.Bo. del asesor:

Dcotor, Eduardo Suger Cofiño

TRIBUNAL:

Doctor, Eduardo Suger Cofiño

Licenciado, Leonel Morales

Licenciado, Jacinto Quan

Fecha de aprobación: de octubre de 1977.

A mis padres y
hermanos.

Agradezco a mis profesores que con toda dedicación y esmero me transmitieron sus conocimientos, y a todas aquellas personas que me brindaron su amistad durante mis estudios en Guatemala, en especial a la familia Alcahé Behar.

INDICE

1. Introducción.

El presente trabajo no pretende ser original en ninguno de los temas tratados, pues ha sido mi interés principal el de exponer de manera sencilla y corta la teoría de Programación No-Lineal (primera parte), aproveché la oportunidad para exponer en la segunda parte una axiomatización de la Teoría del Consumidor, la cual se presta a ser un magnífico ejemplo para la aplicación de la teoría expuesta en la primera parte.

El nivel matemático necesario no va más allá de un primer curso de álgebra lineal y cálculo diferencial en varias variables. En los momentos en que omito una demostración o abandono un tema cito, el libro se puede hallar la parte correspondiente.

Me hubiese gustado mucho penetrar más en los campos tratados en el presente estudio, pero por ser esta la primera vez que lo hago en Modelos Económicos y Programación No-Lineal, me veo obligado a dejar la profundización para el futuro.

PARTE I

PROGRAMACION NO LINEAL

2. Concavidad de funciones.

Definición 2.1. Diremos que el conjunto A es "convexo" si x, y están en A, entonces

$$\theta x + (1-\theta) y \quad \text{está en A}, \quad \forall \theta, \quad 0 \leq \theta \leq 1$$

Se deduce de está definición que la intersección de conjuntos convexos es convexa.

Definición 2.2. Diremos que una función de $D \subset E^n$ en E, con D convexo es "convexa" si

$$f(\theta x + (1-\theta)x') \leq \theta f(x) + (1-\theta) f(x')$$

para todo $0 \leq \theta \leq 1$ y x, x' en D. Además, si -f es convexa, diremos que f es "cóncava".

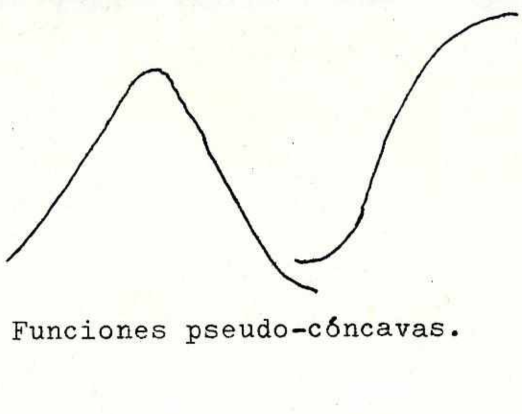
Definición 2.3. Diremos que f: $E^n \longrightarrow E$, diferenciable, es "pseudo-cóncava" si

$$\nabla f(x)' (y-x) \leq 0 \implies f(y) \leq f(x) .$$

Además, si -f es pseudo-cóncava , diremos que f es "pseudo-convexa".

Definición 2.4. Diremos que f:$E^n \longrightarrow E$, es "cuasi-cóncava", si para todo γ en R, el conjunto

$$F_\gamma = \{ x / f(x) \geq \gamma \} \quad \text{es convexo.}$$



Funciones pseudo-cóncavas; intuitivamente significa que si la derivada direccional indica una disminución, la función continua disminuyendo en esa dirección.

Funciones pseudo-cóncavas.

Si f es diferenciable tenemos una definición alternativa para concavidad.

Lema 2.1. Sea f diferencable en E^n , entonces f es cóncava ssi

$$f(y) \leq f(x) + \nabla f(x)'(y-x), \text{para todo } x,\ y.$$

Demostración. Si f es cóncava y $0 \leq \theta \leq 1$

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

para $\theta > 0$

$$\frac{f(\theta y + (1-\theta)x) - f(x)}{\theta} \geq f(y) - f(x)$$

tomando limite

$$\lim_{\theta \to 0^+} \frac{f(\theta y + (1-\theta)x) - f(x)}{\theta} = \lim_{\theta \to 0^+} \frac{f(x+\theta(y-x)) - f(x)}{\theta} \geq f(y) - f(x)$$

o sea

$$\nabla f(x)'(y-x) \geq f(y) - f(x).$$

Para demostrar la conversa asumimos

$$f(y) \leq f(x) + \nabla f(x)'(y-x)$$

tomando $x = \theta x' + (1-\theta)x''$ & $y = x'$ obtenemos

$$f(x') \leq f(\theta x' + (1-\theta)x'') + \nabla f(\theta x' + (1-\theta x'')' [x' - (\theta x' + (1-\theta)x'')] \tag{2.1}$$

y si hacemos $y = x''$

$$f(x'') \leq f(\theta x' + (1-\theta)x'') + \nabla f(\theta x' + (1-\theta)x'')[x'' - (\theta x' + (1-\theta)x'')] \tag{2.2}$$

como

$$x' - (\theta x' + (1-\theta)x'') = (1-\theta)(x' - x'')$$

$$x'' - (\theta x' + (1-\theta)x'') = (-\theta)(x' - x'')$$

haciendo $\omega = \theta x' + (1-\theta)x''$

las ecuaciones (2.1) y (2.2) se vuelven

$$f(x') \leq f(w) + \nabla f(w)'(1-\theta)(x'-x'') \tag{2.3}$$

$$f(x'') \leq f(w) - \nabla f(w)'\theta(x'-x'') \tag{2.4}$$

multiplicando (2.3)por $\theta > 0$ y (2.4) por $(1-\theta) \geq 0$ y luego sumando obtenemos

$$\theta f(x') + (1-\theta)f(x'') \leq f(\theta x' + (1-\theta)x'').$$

q.e.d.

Lema 2.2. Si f tiene segundas derivadas parciales continuas, entonces f es cóncava ssi la matriz Hessiana es negativa **semi-definida.**

Demostración. La expanción de f en serie de Taylor es

$$f(y) = f(x) + \nabla f(x)'(y-x) + \frac{1}{2}(y-x)'H(x+\theta(y-x))(y-x)$$

claramente

$$(y-x)'H(x+\theta(y-x))(y-x) \leq 0 \tag{2.5}$$

si y solo si

$$f(y) \leq f(x) + \nabla f(x)'(y-x) \,. \tag{2.6}$$

Si el Hessiano es negativo semidefinido, se cumple (2.5) y por lo tanto (2.6), pero (2.6) implica que f es cóncava por el lema (2.1).

Si f es cóncava pero el Hessiano no es negativo semi-definido para alguna x, entonces existe y tal que

$$(y-x)'H(x)(y-x) > 0$$

por continuidad del Hessiano, se tiene para un y cerca de x y para todo θ, $0 \leq \theta \leq 1$

$$(y-x)'H(x+\theta(y-x))(y-x) > 0.$$

pero entonces por la ecuación (2.5), nos damos cuenta (2.6) sería falso. La contradicción viene del lema anterior.

Observación. Los lemas (2.1) y (2.2) tienen análogos para los casos en que se trata con funciones convexas, con lo que obtendríamos que, si f tiene segundas derivadadas parciales continuas, entonces f es convexa si y solo si la matriz Hessiana es positiva semidefinida. Las demostraciones se siguen de igual manera a las ya hechas, solo que invirtiendo los signos de desigualdad.

Si definimos una función como "estrictamente cóncava"si la desigualdad de la definición de concavidad se da en sentido estricto y $0<\theta<1$, obtenemos el resultado que una función es estrictamente cóncava si y solo si su Hessiana es negativo definido y en este caso por los teoremas (a.2) y (A.3) del apéndice B, el máximo será absoluto.

Similarmente ocurre con la definición de estrictamente convexa, y la conclusión sería que una función es etrictamente convexa si y solo si su Hessiano es positivo definido y, por lo tanto, su mínimo será absoluto.

3. Formulación del problema de programación no lineal (NLP).

Sea f un campo escalar. El problema de programación no lineal NLP aparece cuando tenemos que maximizar (o minimizar) una función f que puede ser lineal o no lineal, sujeto a restricciones que limitan nuestro dominio de busqueda, estas restricciones vienen dadas por medio de funciones lineales o no lineales en la forma $g_i(x) \geq 0$. A la función f se la llama " función objetivo" y a las funciones g_i "restricciones".

Para tratar el problema de manera general, tomaré como dominio de f al espacio euclideo n-dimensional E^n .Un punto x^o que satisfaga las restricciones se dice que es "factible", y al conjunto de estos puntos se le llama "región factible". Si x^o es factible y f alcanza su máximo ,se dice que es un "óptimo".

El problema descrito anteriormente se escribirá en su forma estandard

$$\text{Max} \quad f(x)$$
$$\text{sujeto a} \quad g_i(x) \geq 0, \qquad i= 1, \ldots, n.$$

Si el caso es $g(x)=0$, sabemos que esto es equivalente a dos desigualdades del tipo $g(x)\geq 0$ y $-g(x) \geq 0$.

He puesto maximizar f,pues si el problema real indica buscar mínimo de f; esto se logra al maximizar -f; más aún si se tiene $af(x)+ b$, con $a>0$, el problema de maximización no se ha variado, por lo tanto, el sumar constantes o multiplicar constantes positivas no altera la solución del problema. Supondremos,por facilidad, que todas las funciones son continuas y diferenciables, a menos que se diga lo contrario.

3.1 Condiciones necesarias de Kuhn-Tucker (K-T).

En esta sección me dedicaré a estudiar criterios que facilitan la identificación de puntos óptimos para el problema NLP.

Una "dirección" será un vector columna "d" en E^n .Entonces $x + td$ con t en R, será una recta que pasa por x en dirección d.

Teorema 3.1. Si f es difenciable en x y si d es una dirección tal que $\nabla f(x)'d > 0$, entonces existe $s > 0$ tal que para todo t con

$$s \geq t > 0 \quad , \quad f(x+ td) > f(x).$$

Demostración. Sabemos que

$$\lim_{t \to 0} \frac{f(x+ td) - f(x)}{t} = \nabla f(x)'d$$

entonces

$$\lim_{t \to 0} \frac{f(x+ td) - f(x)}{t} > 0 \quad ;$$

por propiedad de límite tenemos que existe $s > 0$ tal que para todo $t \neq 0$

$$\frac{f(x+ td) - f(x)}{t} > 0, \quad \text{si} \quad s > t > -s.$$

q.e.d.

Intuitivamente el teorema es claro, pues el gradiente marca la dirección de máximo incremento; por lo tanto, si $\nabla f(x)'d > 0$, entonces d va en el mismo sentido del gradiente y un desplazamiento em esa dirección nos dará un mayor valor de f.

Si el problema NLP no tiene restricciones tenemos el siguiente corolario al teorema anterior.

Corolario 3.1.1. Sea f diferenciable . Si x° maximiza a f en E , entonces $\nabla f(x^{\circ}) = 0$.

Demostración. Supongamos que $\nabla f(x^{\circ})$ es diferente de cero. Hagamos $d = \nabla f(x^{\circ})$; entonces

$$\nabla f(x^{\circ})'d = \nabla f(x^{\circ})'\nabla f(x^{\circ}) > 0 \; .$$

Por el teorema 3.1 habría un punto en el cual f sería mayor que $f(x^{\circ})$. De esta contradicción tenemos el corolario.

El corolario 3.1.1 da una condición necesaria para que un punto x^o sea máximo de una función f que no está sujeta a restricciones. Esta fue la característica que utilizamos en el corolario, pues escogimos como dirección d al gradiente, sabiendo de antemano que todos los puntos en esta dirección eran factibles.

Si agregamos restricciones puede ser que en dirección del gradiente todos los puntos violen al menos una restricción, y la prueba no estaría correcta.

En lo que sigue de la sección me dedicaré a resolver este problema, por medio de las condiciones de Kuhn-Tucker, que son una generalización del corolario 3.1.1.

<u>Definición 3.1</u>. Sea F el conjunto de puntos factibles, entonces

$$D(x') = \{ d / \exists\, s > 0 ,\ s \geq t > 0 \Rightarrow x' + td \in F \} .$$

<u>Corolario 3.1.2</u>. Si x^o es optimal para el problema NLP, entonces

$$\nabla f(x^o)'d \leq 0 \quad , \text{para toda } d \text{ en } D(x^o).$$

La demostración se sigue inmediatamente del teorema 3.1. El corolario nos dice que en el punto óptimo, cualquier dirección factible nos indica una disminución en el valor de la función.

Antes de seguir adelante, recordaré que un conjunto cerrado es el que contiene a todos sus puntos de acumulación. La cerradura de A se denota $\overline{A}$.

Corolario 3.1.3. Si x^o es optimal para el NLP, entonces $\nabla f(x^o)'d \leq 0$, para todo d en $\overline{D}(x^o)$.

Demostración. Si $d^{\infty} = \lim_{k\to\infty} d^k$ con d^k en $D(x^o)$, entonces d^{∞} está en $\overline{D}(x^o)$.

$$\nabla f(x^o)'d^k \leq 0, \quad \text{para toda } K;$$

por lo tanto $\nabla f(x^o)'d^{\infty} = \lim \nabla f(x^o)'d^k \leq 0.$

Trataré ahora de dar una expresión de $\overline{D}(x^o)$ en término de las restricciones $g_i(x)$; las únicas que interezan son las restricciones donde $g_i(x) = 0$, pues si es mayor que cero, esto nos dice que nos podemos mover en cualquier dirección y por continuidad de la restricción, si el movimiento es pequeño, no nos salimos de la región factible.

Llamaremos "restricciones activas en x" a aquellas donde $g_i(x) = 0$; el conjunto de indices de las restricciones activas lo denotaremos por $\mathcal{A}(x)$, es decir:

$$g_i(x) = 0 \quad \text{si} \quad i \in \mathcal{A}(x),$$
$$g_i(x) > 0 \quad \text{si} \quad i \notin \mathcal{A}(x).$$

Definición 3.2.

$$\mathcal{D}(x) = \{d / \nabla g_i(x)'d \geq 0 \ , \ \forall\, i \in \mathcal{A}(x)\} \ .$$

Observamos que $\mathcal{D}(x)$ es cerrado, y la demostración de este hecho es análoga a la del corolario 3.1.3. Por otra parte, notamos que en la definición de $\mathcal{D}(x)$ lo que nos intereza es el ángulo entre las restricciones activas (el gradiente de éstas) y la dirección d; no nos intereza la magnitud de ninguna cantidad o vector.

Lema 3.1.

$$\overline{D}(x) \subset \mathcal{D}(x).$$

Demostración. Si d está en D(x) y $g_i(x) = 0$, entonces $\nabla g_i(x)'d < 0$ implica por el teorema 3.1 que para todo t suficientemente pequeño

$$g_i(x + td) < g_i(x) = 0 \implies d \notin D(x);$$

por lo tanto, $\nabla g_i(x)'d \geq 0$ implica que $D(x) \subset \mathcal{D}(x)$. Pero $\mathcal{D}(x)$ es cerrado, de donde $\overline{D}(x) \subset \mathcal{D}(x)$.

Observación. Existen casos en que $d \in \mathcal{D}(x)$ y $d \notin \overline{D}(x)$, pero estas corresponden generalmente a situaciones que no se dan em la práctica y que son de fabricación matemática. Por lo tanto, con poca pérdida de generalidad asumiré que $\mathcal{D}(x) \subset \overline{D}(x)$, es decir

$$\overline{D}(x) = \mathcal{D}(x).$$

LLamaremos a esta suposición " la característica de restricción". Con la característica de restricción nuestro objetivo de expresar el conjunto de direcciones factibles en función de las restircciones está logrado, y podemos resumir diciendo que si x° es óptimo, entonces

$$\overline{D}(x^\circ) = \mathcal{D}(x^\circ) = \left\{ d / \nabla g_i(x^\circ)'d \geq 0, \ \forall\, i \in A(x^\circ) \right\} .$$

Con esto hemos establecido el siguiente lema.

Lema 3.2. Si x° es óptimo para el NLP, entonces bajo la característica de restricción

$$\nabla f(x^\circ)'d \leq 0, \ \forall\, d \in \mathcal{D}(x^\circ) = \left\{ d / \nabla g_i(x^\circ)'d \geq 0, \ \forall\, i \in A(x) \right\} . \qquad (3.1)$$

Haciendo uso del Lema de Farkas (ver apéndice C), tenemos que la expresión (3.1) es equivalente a la existencia de multiplicadores $\lambda_i > 0$, $i \in A(x^0)$, tal que

$$\nabla f(x^0) + \sum_{i \in A(x^0)} \lambda_i \nabla g_i(x^0) = 0 \qquad (3.2)$$

Se deduce inmediatamente lo deseado.

<u>Teorema 3.2</u>. (<u>Condiciones necesarias de Kuhn-Tucker</u>). Sea el problema NLP

$$\text{Max } f(x)$$
$$\text{sujeto a } g_i(x) \geq 0, \quad i = 1, \ldots, n.$$

donde todas las funciones son diferenciables. Sea x^0 optimal, y asumamos la característica de restricción. Entonces, tenemos las tres condiciones siguientes

(1) x^0 es factible;

(2) existen multiplicadores $\lambda_i \geq 0$, con $i = 1, \ldots, n$; tales que

$$\lambda_i g_i(x^0) = 0 \ , \ i = 1, \ldots, n \ ;$$

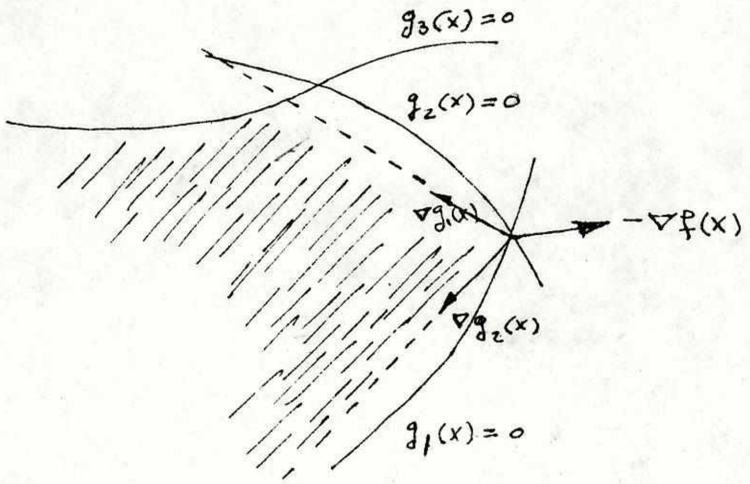
(3) $$\nabla f(x^0) + \sum_{i=1}^{n} \lambda_i \nabla g_i(x^0) = 0$$

Demostación. La condición (1) es trivial; por otra parte si $i \notin A(x^0)$, entonces $g_i(x^0) > 0$ y $\lambda_i = 0$ por la condición (2), y esto hace que la condición tres sea precisamente la ecuación (3.2).

q.e.d.

Las condiciones (1), (2) y (3) son llamadas " las Condiciones de Kuhn-Tucker". Notemos que las condiciones de Kuhn-Tucker son equivalentes a decir $\nabla f(x^0)' d \leq 0$ para toda d en $\mathcal{D}(x^0)$. Su interpretación geométrica nos dice que en cualquier dirección factible f debe decrecer.

En el punto optimal x°, el gradiente de f está expresado como una combinación lineal de los gradientes de las restricciones activas, es decir $-\nabla f(x^\circ)$ está dentro del cono formado por los gradientes de las restricciones activas.



Condiciones de K-T para restricciones de igualdad. Escribiendo el problema NLP cuando aparecen restricciones de igualdad, la formulación queda

$$\begin{array}{ll} \text{Maximizar} & f(x) \\ \text{sujeto a} & g_i(x) \geq 0, \quad i = 1, \ldots, m' \\ & g_i(x) = 0, \quad i = m'+1, \ldots, m. \end{array}$$

Las condiciones necesarias de Kuhn-Tucker son para este caso:

(1) x° es factible;

(2) existen multiplicadores $\lambda_i \geq 0$, $i = 1, \ldots, m'$ y multiplicadores $\lambda_i \in \mathbb{R}$, $i = m'+1, \ldots, m$, tal que $\lambda_i g_i(x^\circ) = 0$, $i = 1, \ldots, m'$;

(3) $\nabla f(x^\circ) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^\circ) = 0.$

Demostración. Todo lo que hay que hacer es expresar las condiciones de igualdad en forma de desigualdades, conforme a lo dicho al principio de la sección; y el teorema se reduce al anterior.

Observación. Las condiciones de K-T son necesarias pero no suficientes, o sea que nos sirven únicamente para identificar puntos optimales pero no para encontrarlos. Lamentablemente se tendrán que dar unas restricciones sobre las funciones de restricción y la función objetivo.

3.2 Condiciones de suficiencia.

Teorema 3.3. Supongamos que tenemos un problema NLP, donde f es diferenciable y pseudo-cóncava y las restricciones g_i son diferenciables y cuasi-cóncavas, entonces las condiciones de K-T son suficientes para optimalidad (x° satisface las condiciones K-T, entonces x° es óptimo).

Demostración. Tomemos un punto factible "y" en F. Como g_i son cuasi-cóncavas, entonces F es convexo (def. 2.4), y por lo tanto la dirección $d^\circ = y - x^\circ$ es factible. Por el lema 2.1 $d^\circ \in \mathcal{D}(x^\circ)$, y de esto tenemos

$$\nabla g_i(x^\circ)' d^\circ \geq 0 , \qquad i \in \mathcal{A}(x^\circ).$$

Por otra parte, las condiciones de K-T nos dicen

$$0 \geq \nabla f(x^\circ)' d, \quad \forall\, d \in \mathcal{D}(x^\circ);$$

como $d^\circ \in \mathcal{D}(x^\circ)$ implica que $0 \geq \nabla f(x^\circ)' d^\circ = \nabla f(x^\circ)'(y - x^\circ)$ y por pseudo- concavidad de f, $f(y) \leq f(x)$.

Observación. La interpretación geométrica del teorema anterior es clara, al recordar que las condiciones K-T nos dicen que en cualquier dirección factible d, la derivada direccional de f en x^{o} nos indica una disminución de f. Al pedir pseudo-concavidad de f es para que decrezca en cualquier dirección factible, y como las restricciones son cuasi-cóncavas, el conjunto factible es convexo; por esto, $d^{o} = y - x^{o}$ es factible. De aqui que la derivada direccional nos de $f(y) \leq f(x^{o})$.

3.3 La característica de restricción.

Quisiera en este punto profundizar en el significado geométrico de la característica de restricción, sobre la cual no he hablado hasta el presente. Ya que el objetivo señalado me obliga aintroducir cierto material que para el desarrollo posterior es dispensable, el lector puede omitir esta sección y continuar en la sección 3.4 con Teoría Lagrangiana.

Una curva $x(t)$ en una superficie S es una función continua con dominio un intervalo de la forma $[a, b]$ y comtradominio S, me interezarán las curvas cuya primera y segunda derivadas existan. Denotaremos a estas por $\dot{x}(t)$ y $\ddot{x}(t)$, respectivamente. Si $x(t^{o}) = x^{o}$, decimos que la curva pasa por x^{o}, donde $a \leq t^{o} \leq b$.

Si x^{o} está en S, el conjunto de todas las derivads de todas las curvas que pasan por x^{o} se llama "Plano Tangente en x^{o}" (T); nótese que T es subespacio de E^{n} y que solo me interzan las curvas diferenciables.

En la sección 3.1 supuse que el conjunto de direcciones factibles en x^o era igual al conjunto $\mathcal{D}(x^o)$; como en $\mathcal{D}(x^o)$ interezan los puntos frontera de la región factible S, la suposición quiere decir que para cada punto frontera x^o donde se cumpla la característica de restricción , debe existir una curva suave (diferenciable) que pase por x^o y que esté en S. Además, si x^o es máximo, f(x) no puede decrecer a lo largo de ninguna curva de este tipo. Para el caso especial en que f(x) es cóncava y g (x) son convexas, lo anterior quiere decir que existe un punto en el conjunto factible que satisface todas las desigualdades en sentido estricto, es decir, existe x^o tal que $g_i(x^o) > 0$ para toda i.

Empezaré la demostración de los hechos citados con el concepto clave de punto regular; suponiendo que mi problema NLP solo tiene igualdades h_i . Denotaremos h(x) al vector $(h_1(x), \ldots, h_m(x))$.

<u>Definición 3.3</u>. Un punto x^o que satisface las restricciones $h(x^o)= 0$, se dice "regular" si los vectores gradientes $\nabla h_i(x^o)$ son linealmente independientes para toda i.

<u>Teorema 3.4</u> . En un punto regular x^o de S definido por h(x) = 0, el plano tangente T es igual a

$$M = \{ y / \nabla h(x^o)'y = 0 \} .$$

Demostración. Es obvio que $T \subset M$, no importando si x^o es regular o no, por lo tanto, solo queda la conversa.

Para tener $M \subset T$, debemos tener que si y M, entonces existe una curva en S que pasa por x^o con derivada y. Para construir la curva, Consideremos la ecuación

$$h(x^\circ + ty + \nabla h(x^\circ)'u(t)) = 0 \tag{3.4}$$

donde para un t fijo consideramos a $u(t) \in E^n$ es la incógnita. Este es un sistema no lineal de m ecuaciones con m incógnitas, con un parámetro continuo t, y para $t = 0$, $u(0) = 0$. La matriz Jacobiana es respecto a $u(t)$ en t° la matriz

$$\nabla h(x^\circ)\ \nabla h(x^\circ)'$$

que es regular por ser x° regular. Por el teorema de la función implícita, existe una solución continua $u(t)$ en una región adecuada $(-a < t < a)$.

La curva $x(t) = x^\circ + ty + \nabla h(x^\circ)'u(t)$ es una curva en S por construcción. Si diferenciamos el sistema(3.4) con respecto a t en $t = 0$, obtenemos

$$0 = \frac{d}{dt} h(x(t))\Big|_{t=0} = \nabla h(x^\circ)'y + \nabla h(x^\circ)\nabla h(x^\circ)'\ \dot{u}(0).$$

Por definición de y se tiene que $\nabla h(x^\circ)'y = 0$, y como $\nabla h(x^\circ)\nabla h(x^\circ)$ es no singular; por lo tanto $\dot{u}(0) = 0$; com lo que concluimos

$$x(0) = y + \nabla h(x^\circ)'\dot{u}(0) = y\ ,$$

habiendo construido una curva con derivada y en x°.

q.e.d.

Es importante reconocer que la condición de ser punto regular no es una condición en la superficie de restricción, sino más bien en la representación de un punto en términos de $h_i(x)$. El plano tangente se define independientemente de las restricciones, mientras que M no.

En el caso que hayan restricciones del tipo $g_i \geq 0$, las que nos interezarán son las activas en x^o, pues no me podré mover en cualquier dirección arbitraria; si $g_i(x^o) > 0$, por continuidad de g_i existe un entorno en x^o en el cual me puedo desplazar libremente sin violar la restricción.

Concluimos de lo anterior que bajo ciertas advertencias, puedo considerar mi problema NLP en su forma estandard, como si realmente estuviera formado nada más por igualdades como restricciones.

Como una observación más al teorema 3.4, tenemos pues que la suposición de la característica de restricción válida en x^o se puede interpretar como la suposición de que x^o es punto regular. Talvez el lector se pregunte que significado tiene en la definición de $\mathcal{D}(x^o)$, la expresión

$$\nabla g_i(x^o)'d \geq 0 \qquad (3.5)$$

y como la adaptamos a la teoría de esta sección. El problema se resuelve al recordar que la expresión (3.5) nos implica, por el teorema 3.1, que existen puntos en dirección d donde g_i alcanza un valor mas alto que en x^o; por lo tanto estos puntos no violan la restricción, y son factibles. Hemos obtenido pues, una generalización del teorema 3.4, pero antes de enunciarlo daré una definición.

Definicion 3.4. Sea x^o un punto de S tales que el conjunto de los gradientes de las condiciones activas en x^o es linealmente independiente; entonces, decimos que x^o es "punto regular de las restricciones activas".

Teorema 3.5. Si x° es punto regular de las restricciones activas, entonces

$$\bar{D}(x^\circ) = \mathcal{D}(x^\circ) .$$

3.4 Teoría Lagrangiana.

LLamaremos "función lagrangiana" o "Lagrangiano" del problema NLP a

$$L(x,\lambda) = f(x) + \lambda' g(x) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x),$$

donde $\lambda = (\lambda_1, \ldots, \lambda_m)$ y $g(x) = (g_1(x), \ldots, g_m(x))'$.

Si $x \in X \subset E^n$ y $\lambda \in Y \subset E^m$, decimos de (x°, λ°) que es un "punto de silla de L" o que "L tiene un punto silla en (x°, λ°)" si

$$L(x, \lambda^\circ) \leq L(x^\circ, \lambda^\circ) \leq L(x^\circ, \lambda), \text{ para toda } x \in X, \text{ y } \lambda \in Y.$$

Para simplicidad de estudio supondré que tiene sentido hablar de máximos y mínimos de L.

Definimos dos funciones siguientes:

$$L^*(\lambda) = \max_{x \in X} L(x,\lambda) \quad \text{y} \quad L_*(x) = \min_{\lambda \in Y} L(x,\lambda).$$

Una propiedad importante de estas dos finciones es

$$L^*(\lambda) \geq L(x,\lambda) \geq L_*(x), \tag{3.6}$$

asi que

$$L^*(\lambda) \geq L_*(x) \text{ para cualquier } (x,\lambda) \text{ en } X \times Y. \tag{3.7}$$

Daremos las definiciones de punto silla en función de L^* y L_*.

Lema 3.3. Las siguientes tres propiedades son equivalentes:

I) $\min_{\lambda\in Y} \max_{x\in X} L(x,\lambda) = \max_{x\in X} \min_{\lambda\in Y} L(x,\lambda)$;

II) $L^*(\lambda^\circ) = L_*(x^\circ)$, $\lambda^\circ \in Y$, $x^\circ \in X$;

III) $L(x,\lambda^\circ) \leq L(x^\circ,\lambda^\circ) \leq L(x^\circ,\lambda)$, para todo $x\in X$, $\lambda\in Y$.

Demostración. La expresión (i) se puede escribir como

$$\min_{\lambda\in Y} L^*(\lambda) = \max_{x\in X} L_*(x) \tag{3.8}$$

Si λ° minimiza $L^*(\lambda)$ y x° maximiza $L_*(x)$, entonces (ii) es cierto. por ecuación (3.7), (ii) implica (3.8) y de aqui (i).Por lo tanto, debido a (3.6), (ii) es equivalente a (i).

$$L^*(\lambda^\circ) = L(x^\circ,\lambda^\circ) = L_*(x^\circ); \tag{3.9}$$

reescribiendo (3.9) en términos de la definición se tiene

$$\max_{x\in X} L(x,\lambda^\circ) = L(x^\circ,\lambda^\circ) = \min_{\lambda\in Y} L(x^\circ,\lambda) .$$

Debido a (3.6) se tiene que esta última ecuación es equivalente a (iii); es decir (iii) es una reformulación de (ii).

De esta manera tenemos varias maneras de identificar un punto silla. Hay que tomar en cuenta que existen funciones que no tienen puntos silla pues $\min_{\lambda\in Y} L^*(\lambda) > \max_{x\in X} L_*(x)$.

Antes de seguir adelante definiré la "función primal" como

$$L_*(x) = \min_{\lambda\in Y} L(x,\lambda)$$

y la "función dual" como

$$L^*(\lambda) = \max_{x\in X} L(x,\lambda).$$

El siguiente problema es el determinar un punto x° llamado "optimal" tal que

$$L_*(x^\circ) = \max_{x\in X} L_*(x),$$

y el de hallar λ° tal que $L^*(\lambda^\circ) = \min_{\lambda\in Y} L^*(\lambda)$; a la primer busqueda se le llama "problema primal" y a la segunda "problema dual".

<u>Problemas</u> <u>primal</u> <u>y</u> <u>dual</u>. Si nos damos cuenta,

$$L_*(x) = \min_{\lambda\in Y} \left\{ f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \right\} = \begin{cases} f(x) & \text{si } g_i(x) \geqq 0; \forall i \\ -\infty & \text{si } \exists i \text{ con } g_i(x) < 0 \end{cases} \qquad (3.10)$$

Esto último pues $\lambda_i \geqslant 0$, $\forall i$. Por lo tanto, si $g_i(x) \geqq 0$, $\lambda_i = 0$ minimizan. Sin embargo, si $g_i(x) < 0$, L es mínimo cuando $\lambda_i \to +\infty$.

Si trabajamos con el problema NLP, el caso que nos intereza es el primero, de esto se llega a que el problema primal queda dado por

$$\max_{x\in X} L_*(x) = \max f(x) \text{ sobre } x \in X, \text{ tal que } g_i(x) \geqq 0 \text{ para toda } i \qquad (3.11)$$

es decir el problema primal se nos vuelve un problema NLP si $X = E^n$; por esta razón a veces se le llama de esta manera al NLP.

Por otra parte, una pareja (x', λ') se dice que es "factible para el problema dual" si

$$L^*(\lambda') = L(x', \lambda').$$

Si λ° es óptimo para el dual y (x°, λ°) es factible, entonces la pareja (x°, λ°) se llama "optimal".

Lema 3.4. Si x' es factible para elNLP, entonces $L_*(x') = f(x')$; más aún, si (x'', λ'') es factible para el dual,

$$L_*(x') = f(x') \leq f(x'') + \sum_{i=1}^{m} \lambda_i'' g(x'') = L(x'', \lambda'') = L^*(\lambda''). \qquad (3.12)$$

Demostración. Si x' es factible, por (3.10) $L_*(x') = f(x')$, y por ecuación (3.7), $L_*(x') \leq L^*(\lambda'')$.

Teorema 3.6. Sea (x°, λ°) un punto silla del Lagrangiano, esto es

$$\max_{x \in X} L(x, \lambda^\circ) = L(x^\circ, \lambda^\circ) = \min_{\lambda \in Y} L(x^\circ, \lambda);$$

entonces, (x°, λ°) resuelve el prpblema primal Y el dual (NLP si $X \equiv E^n$).

Demostración. Se deduce inmediatamente del Lema 3.3.

Teorema. 3.7. El punto x' es optimal para el problema primal si se cumple una de las siguientes condiciones. Más aún, si $X = E^n$, una de las siguientes condiciones es suficiente para que x' sea optimal del problema NLP.

i) $L_*(x') = L^*(\lambda'')$ para alguna $x' \in X$ y $\lambda'' \in Y$;

ii) x'es factible para el problema (3.11), (x'', λ'') es factible para el dual y $f(x') = L(x'', \lambda'')$;

iii) 1. x' es factible para el problema (3.11);

2. $\sum_{i=1}^{m} \lambda_i' g_i(x') = \breve{0}$ donde $\lambda_i' \geq 0$

3. $L(x', \lambda') = \max_{x \in X} L(x, \lambda')$;

iv) 1. x' satisface las condiciones K-T para el problema NLP y $x' \in X$;

2. si $\lambda' = (\lambda_i')$ son los correspondientes multiplicadores de K-T

$$\nabla L(x', \lambda') = 0 \Rightarrow L(x', \lambda') = \max_{x \in X} L(x, \lambda');$$

v) las funciones f y g_i para toda i son cóncavas , $x' \in X$, x' satisface las condiciones K-T.

Demostraciones. Todas las condiciones se deducen del teorema 3.6 despues de verificar que un punto silla existe.

i) del lema 3.3, esta proposición es equivalente a la existencia de un punto silla.

ii) usando el lema 3.4, $f(x') = L_*(x') = L^*(\lambda'')$, pero entonces estamos en la parte (i) de nuevo;

iii) de (iii-2) y (iii-3),

$$f(x') = f(x')+\sum_{i=1}^{m} \lambda_i' g_i(x') = L(x',\lambda') = \max_{x\in X} L(x,\lambda') = L^*(\lambda'),$$

de donde (x',λ') es factible para el dual, y estamos en (ii) de nuevo.

iv) Como las condiciones K-T se dan, x' es factible y $\sum_{i=1}^{m} \lambda_i' g_i(x) = 0$; más aún, la tercera condición K-T es $\nabla L(x', \lambda') = 0$ y hemos caído en el caso (iii);

v) por concavidad de $g_i(x)$, $f(x)$, $\lambda_i' \geqq 0$, el lema 3. , tenemos que

$$L(x, \lambda') = f(x)+\sum_{i=1}^{m} \lambda_i' g_i(x)$$

es una función cóncava en x; por lo tanto

$$\nabla L(x', \lambda') = 0 \Rightarrow L(x', \lambda') = \max_{x\in E^n} L(x, \lambda') = \max_{x\in X} L(x, \lambda'),$$

como $X \subset E^n$, estamos en el caso (iv).

q.e.d.

De esta manera, con los dos últimos teoremas hemos dado una serie de condiciones suficientes para un punto optimal.

PARTE II

MODELO ECONOMICO.

4. Teoría del consumidor racional.

El problema que trataré es el relacionado con la decisión que debe hacer un individuo -el consumidor- al tener que escoger entre varios bienes con el fin de satisfacer sus necesidades. Se tomará como "bien" cualquier cosa que el consumidor desea tener más. Este problema surge al reconocer que las necesidades de un individuo son ilimitadas, y por el contrario, sus recursos son limitados.

He señalado que el consumidor es un individuo, pero tambien entenderemos como consumidor a la familia como unidad, o cualquier grupo pequeño de personas que tomen decisiones en conjunto debido a sus intereses comunes.

Las decisiones son, por ejemplo, qué debo comprar, cúanto he de pagar, debo ahorrar, si o no, y en qué forma, cúando debo hacer una compra, etc. Las respuestas a estas preguntas están sujetas a la información que en un momento dado se posee, al tiempo que se tiene para escoger, a la duración que va a tener el efecto de la decisión, etc.

Empezaré, por lo tanto, haciendo una serie de suposiciones: (1) el consumidor posee una cantidad fija de dinero, llamada "ingreso", para hacer sus gastos, los cuales son hechos en un periodo determinado (un día, una semana, un mes, etc.); despues de este periodo el consumidor recibirá otro ingreso; (2) el consumidor estáperfectamente informado de los precios de todos los bienes y conoce las circunstancias relevantes para hacer su decisión; (3) estas circunstancias y otras, llamadas también condiciones calificantes o parámetros, se mantienen constantes en el intervalo de tiempo establecido; por ejemplo, si A es preferido a B en t, pudiera ser que en t' ya no se verificara, sino que en t' se tuviera que B es preferido a A. Esta situación queda, por lo tanto, excluída.

Para finalizar, quiero hacer notar que los axiomas de la teoría de la elección que se darán más tarde pueden dejar de funcionar; el proceso de selcción mental puede ser llevado a cabo de manera diferente, dependiendo de la situación de elección, o se podría actuar aleatoriamente y romper así con las condiciones propuestas por los axiomas (transitividad, por ejemplo). Y por último, si la selección es a nivel individual, no se puede dar una regla arbitraria para llegar a la selección de grupo, ya que se puede llegar a contradecir algún axioma; por ejemplo, si tenemos tres individuos y sus preferencias sobre A, B y C son:

A-P-B-P-C (A es preferido a B y B preferido a C)
B-P-C-P-A (B es preferido a C y C preferido a A)
C-P-A-B-B (C es preferido a A y A preferido a B),

La regla de la mayoría da entonces A-P-B, B-P-C y C-P-A, lo cual nos lleva a una contradicción con la transitividad de la relación de preferencia que más adelante se estudiará.

Como no estoy haciendo teoría de la decisión, me veo forzado a dejar este interesante tema, y paso enseguida a establecer los axiomas de la teoría.

Como dije anteriormente, el consumidor se encuentra ante un conjunto de varias alternativas, llamado $S = \{ x , x', \ldots \}$ y tiene que escoger una de ellas. Cada alternativa le da a la persona un grado de utilidad (satisfacción); entonces, qué alternativa escogerá?. Lógicamente, la que le de mayor utilidad; pero lamentablemente no siempre se escoge la que más se prefiere, y esto es debido a las restricciones que aparecen en el momento de decisión.

Las alternativas son conjuntos de bienes; así, la alternativa x tiene una cantidad x del bien 1, una cantidad x del bien dos, y así sucesivamente hasta agotar los bienes con que estamos trabajando; x' tiene cantidades x'_1 , x'_2 , ..., x'_n del bien uno, dos, hasta N, respectivamente.

Lo primero que tiene que hacer el consumidor para resolver su problema es ordenar su conjunto de alternativas de acuerdo a sus preferencias. Por lo tanto, por ahí comenzará nuestra teoría, dando unas propiedades que se deben cumplir en esta categorización, o sea que la relación de preferencia (según grado de utilidad) de las alternativas tendrá una serie de propiedades que serán, en última instancia, nuestros axiomas.

Así, decimos que el consumidor siempre puede definir una relación R, y diremos que x es considerado tan bueno como x' si x está relacionado con x' (x R x').

Antes de seguir adelante, tendremos que hacer una idealización bastante difícil de aceptar, pero nos traerá resultados muy fructíferos;

supondremos que los bienes son infinitamente divisibles, esto es, podemos adquirir cualquier cantidad del bien que nosotros deseemos; expresamos esto escribiendo

$$S=\{ x =(x_1, \ldots,x_n) \ / \ x_i \geq 0 \ y \ i = 1, \ldots, n\}$$

<u>Axioma I (Completitud)</u>

Existe una relación R en S, tal que x R x' ó x' R x ó ambas.

El axioma nos dice que por lo menos se debe tener una de las alternativas anteriores, o sea que dos alternativas siempre se podrán relacionar; si se dan ambas diremos que "x es indiferente a x'" y escribiremos x I x'. Por otra parte, si tenemos x R x' pero no se cumple x' R x, diremos que "x es preferido a x'" y se denota x P x'. En resumen entre dos alternativas siempre tendremos que una es preferida a la otra, o que es indiferente a la otra.

<u>Corolario</u>. Para toda x, x' en S se tiene

i) una y solo una de las siguientes se cumple: x P x', x' P x, x I x';

ii) x I x;

iii) x I x' entonces x' I x.

El significado intituitivo del siguiente axioma nos dice que si x se considera tan bueno como x' y x' tan bueno como x'', entonces, x es considerado tan bueno como x''.

Axioma II (Transitividad).

Para toda x, x', x'' en S, se tiene que si: x R x' Y x' R x'', entonces, x R x''.

Corolario.

i) $x\ P\ x'$ y $x'\ P\ x''$,entonces $x\ P\ x''$;

ii) $x\ I\ x'$ y $x'\ I\ x''$,entonces $x\ I\ x''$;

iii) $x\ R\ x'$ y $x'\ P\ x''$,entonces $x\ P\ x''$;

iv) $x\ P\ x'$ y $x'\ R\ x''$,entonces $x\ P\ x''$;

v) $x\ R\ x'$ y $x'\ I\ x''$,entonces $x\ R\ x''$;

vi) $x\ I\ x'$ y $x'\ R\ x''$,entonces $x\ R\ x''$;

vii) $x\ P\ x'$ y $x'\ I\ x''$,entonces $x\ P\ x''$;

La relación de indiferencia I es de equivalencia y por lo tanto, nos induce a una partición en S. A cada clase de la partición la llamaremos "clase de indiferencia" (el consumidor es indiferente entre dos elementos de la misma clase). Al mismo tiempo,las clases pueden ser ordenadas de acuerdo a la relación de preferencia P. Más tarde, se darán unos axiomas que van a determinar a las clases de indiferencia a ser funciones contínuas, además de tener otras propiedades.

Axioma III (Continuidad).

Para cualquier x°, el conjunto $U = \{x/\ x\ R\ x^\circ\}$ y el conjunto $L = \{x/\ x^\circ\ R\ x\}$ son cerrados.

Se deduce del axioma anterior que el conjunto $\{x/\ x\ I\ x^\circ\}$ que es la intersección de U y L es cerrado.

Axioma IV (Convexidad).

Si x' P x'' y $0 \leq \alpha < 1$, entonces $(1-\alpha)x' + \alpha x''$ P x'' .

Axioma V (No saturación).

No existe x^{o} tal que x^{o} P x, para todo x en S.

El Axioma de Continuidad garantiza que la relación no cambia bruscamente, esto es, la relación entre dos alternativas se mantiene si nos movemos un entorno muy pequeño alrededor de estas, por ejemplo; si x' P x'', entonces, existen vecindades N' y N'' de x' y x'' respectivamente, talque x P x'', para todo x en N' y x' P x para todo x en N''.

Lema I.

(a) Para toda x' en S, existe x'' en S arbitrariamente cercano a x' para el cual x'' P x'.

(b) Para todo x^{o}, el conjunto $\{x / x \text{ P } x^{o}\}$ es convexo.

Demostración. (a) Por A. V existe x''' P x'. Por A. IV, $x''=(1-\alpha)x' + \alpha x'''$ P x', $0 \leq \alpha < 1$, si hacemos α tender a uno, x'' queda arbitrariamente cerca de x'. (b) Sean x' y x'' pertenecientes a $\{x / x R x^{o}\}$ y sea $x= (1-\alpha)x' + \alpha x''$, $0 \leq \alpha \leq 1$. Probaremos que x R x^{o}. Si $\alpha = 0$ ó $\alpha = 1$ es trivial. Si $0 < \alpha < 1$, entonces de A. I podemos suponer que x' R x''. Si x^{*} está arbitrariamente cercano a x' y x^{*} P x'. Entonces x^{*} P x'' por transitividad. Si $x^{**}=(1-\alpha)x^{*}+\alpha x''$. Por A. IV x^{**}P x'' y por transitividad x^{**}P x^{o}, en particular, $x^{**} \in \{x / x \text{ R } x^{o}\}$. Como x^{*} Está arbitrariamente cercano a x', x^{**} está arbitrariamente cercano a x. Por lo tanto x es punto de acumulación de $\{x / x \text{ R } x^{o}\}$ y por el axioma de continuidad pertenece a este conjunto.

Nota: Del lema anterior se desprenden resultados muy interesantes, en la parte (a) se probó que no hay saturación localmente. Por otra parte, sabemos que el conjunto de indiferencia para una alternativa es una curva en R^n, en la parte (b) se probó que esta curva es convexa. Se puede probar que de los axiomas anteriores y el lema se deduce la continuidad de la curva de indiferencia, pero me limitaré nadamás a hacerlo notar.

5. Función Utilidad.

Definición 1. Una función U de S en R se llama "función de utilidad" si cumple con la propiedad: $x'\ R\ x''$ si y solo si $U(x') \geq U(x'')$.

Se deduce de lo anterior que si $x'\ I\ x''$ entonces $U(x') = U(x'')$ igualmente $x'\ P\ x''$ si y solo si $U(x') > U(x'')$. En la siguiente sección se verá que las preferencias que un consumidor tiene, no solo se pueden ordenar por una relación de tipo cualitativo, sino tambien por medio de una de tipo cuantitativo; quiero decir con esto que por medio de los axiomas anteriores se demuestra la existencia de una función de utilidad.

5.1 Existencia de la función utilidad.

Primero, tomemos un punto x° en S y sea $C(x^\circ)=\{x/\ x\ R\ x^\circ\}$, la función $\rho(x) = |x - x^\circ|$ es continua y como $C(x^\circ)$ es cerrado y acotada por abajo $\rho(x)$ alcanza su mínimo. Definimos:

$$U(x) = \operatorname*{Min}_{x' \in C(x)} \rho(x') \qquad (1)$$

$$M(x) = C(x) \cap \{x'/ \rho(x') = U(x)\} . \qquad (2)$$

Como $M(x) \subset C(x)$, entonces x' R x para todo x' en $M(x)$. En realidad tendremos x I x', pues si x' P x, por continuidad se tiene que en una vecindad de x', x'' P x, para todo x'' en la vecindad de x'. Escojamos $\alpha > 0$ tal que $(1-\alpha)x' + \alpha x^\circ$ R x, en particular si $\alpha < 1$; pero:

$$\rho[(1-\alpha)x' + \alpha x^\circ] = |(1-\alpha)x' - \alpha x^\circ - x^\circ| = (1-\alpha)|x' - x^\circ|$$
$$= (1-\alpha)\, \rho(x') = (1-\alpha)U(x) .$$

Por (1) sabemos que $U(x) \leq \rho[(1-\alpha)x' + \alpha x^\circ]$, asi que $(1-\alpha)U(x) \geq U(x)$, pero entonces, $U(x) \leq 0$, es decir $U(x) = 0$ por ser U una función no negativa. Por lo tanto $\rho(x') = 0$ ó $x' = x^\circ$, como supusimos x' R x°, llegamos a la conclusión que es falso x' P x°. Resumiendo:

$$\text{Si } x' \in M(x), \text{ entonces } x' \text{ I } x. \qquad (3)$$

Para probar que U es en efecto una función de utilidad hay que probar que x' R x'' implica $U(x') \geq U(x'')$ y que x' P x'' implica que $U(x') > U(x'')$ Lo primero se sigue inmediatamente de (1) y de la transitividad A. II que asevera que $C(x') \subset C(x'')$.

Por otra parte, supongamos que x' P x''; entonces $M(x') \subset C(x') \subset C(x'')$. Si $U(x') = U(x'')$, entonces $M(x') \subset M(x'')$, por (2). Si x''' está en $M(x')$, entonces por (3), $x' \sim x'''$ y tambien $x''' \sim x''$, asi que $x' \sim x''$, que es una contradicción.* Además, se tiene $U(x') \geq U(x'')$ y se probó que $U(x') \neq U(x'')$, es decir $U(x') > U(x'')$. Por lo tanto U es una función de utilidad.

Es suficiente probar que $\{x/ U(x) \geq u\}$ y $\{x/ U(x) \leq u\}$ son conjuntos cerrados, para probar continuidad de U.

(*) $\sim \equiv$ I

Sea $\{x^{\nu}\}$ una sucesión tal que $x^{\nu} \to x$ y $U(x^{\nu}) = u$. Sea $x' \in M(x)$ por el lema 1podemos tomar x'' cercano a x' y x'' P x', además $U(x'') \leq \varrho(x'')$. Por otra parte como x'' P x' y x' I x, entonces x'' P x. Por A. III, x'' R x^{ν} ,para ν suficientemente grande, $U(x'') \geq U(x^{\nu}) \geq u$, asi que $U(x'') \geq u$; pero x'' está arbitrariamente cerca de x' y como $\varrho(x)$ es continua, podemos que $u \leq \varrho(x') = U(x)$, con lo que probamos que $\{x/ \ U(x) \geq u\}$ es cerrado.

Si tenemos la sucesión $\{x^{\nu}\}$ tal que $U(x^{\nu}) \leq u$, para todo ν, y $x^{\nu} \to x$. Para cada ν escojamos x'^{ν} en $M(x^{\nu})$. Como $\varrho(x'^{\nu}) = U(x^{\nu})$, tenemos que $\varrho(x'^{\nu}) \leq u$ para todo ν, asi que la sucesión $\{x'^{\nu}\}$ es acotada y tiene límite x', con $\varrho(x') \leq u$. Como x'^{ν} I x^{ν} , por (3) y $x^{\nu} \to x$, se sigue x' I x donde hay que hacer uso de A. III. Entonces por definición de $U(x)$ y por ser función de utilidad

$$U(x) = U(x') \leq \varrho(x') \leq u \quad ,$$

asi que $\{x/ \ U(x) \leq u\}$ es cerrado y por lo tanto la continuidad de la función de utilidad ha sido establecida sobre el conjunto $\{x/ \ x \text{ R } x^{\circ}\}$.

Si hacemos que el punto x° coincida con el origen, la función de utilidad queda definida y es continua para todo S.

Finalmente, en lo que sigue supondré que la función de utilidad no solo es concava, sino que tiene sus segundas derivadas parciales continuas. Para esto hay que agregar unos axiomas que omitiré para no desviarme del tema principal.

La teoría expuesta se puede aplicar para el caso en que el conjunto de alternativas no sea todo S; como sucede en general. Si se desea profundizar en este tema el lector puede recurrir al librode Arrow y Hahn citado en la bibliografía.

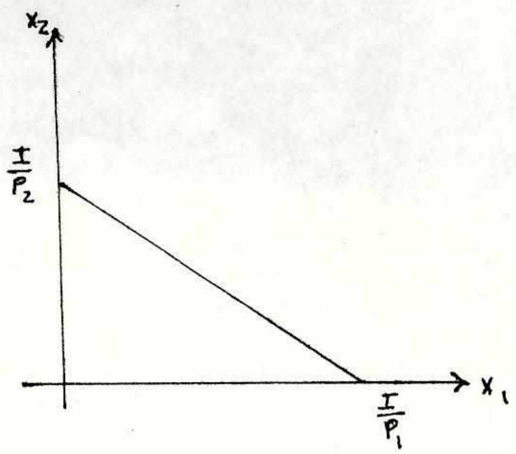
5.2 Maximización de la utilidad.

En las secciones anteriores se vió que el grado de preferencia de un individuo se puede representar perfectamente por medio de un campo escalar U, que madia el grado de satisfacción.

Nuestro objetivo será ahora maximizar el grado de satisfacción del consumidor, sabiendo su función de utilidad y las restricciones que pesan sobre el; en este caso será el ingreso I. Si p_i es el precio del bien i-ésimo la restricción queda

$$p_1 x_1 + \ldots + p_n x_n \leq I$$

donde x_i es la cantidad del bien i que se va a comprar. Esta restricción es un plano, pero en el caso de dos bienes es una recta. Como la función objetivo es cóncava y las restricciones nos dan un conjunto factible convexo; por el teorema 2.3 tenemos que las condiciones K-T son suficientes para garantizar optimalidad.



Escribiendo el problema NLP tenemos

$$\text{Maximizar:} \quad U = U(x_1, x_2, \ldots, x_n)$$

$$\text{sujeto a :} \quad p_1 x_1 + \ldots + p_n x_n \leq I$$

$$x_i \geq 0 \quad , \text{para toda } i = 1, \ldots, n.$$

Utilizando notación vectorial el Lagrangiano del problema es:

$$L(x, \lambda) = U(x) + \lambda(I - px)$$

las condiciones de K-T son

$$\frac{\partial L}{\partial x} = \frac{\partial U}{\partial x} - \lambda p = 0$$

$$\frac{\partial L}{\partial \lambda} = I - px = 0 \tag{4.1}$$

En dos dimensiones esto es

$$\frac{\partial U}{\partial x_1}(x_1, x_2) - \lambda p_1 = 0$$

$$\frac{\partial U}{\partial x_2}(x_1, x_2) - \lambda p_2 = 0 \tag{4.2}$$

$$I - p_1 x_1 - p_2 x_2 = 0$$

Por otra parte, si tomamos una curva de indifernncia particular, U es constante, por lo tanto dU= 0; i. e.

$$\frac{\partial U}{\partial x_1} dx_1 + \frac{\partial U}{\partial x_2} dx_2 = 0$$

esto implica que

$$\frac{\frac{\partial U}{\partial X_1}}{\frac{\partial U}{\partial X_2}} = -\frac{dX_2}{dX_1} \qquad (4.3)$$

de (4.2) resulta

$$\frac{\frac{\partial U}{\partial X_1}}{\frac{\partial U}{\partial X_2}} = \frac{P_1}{P_2} \qquad (4.4)$$

de esto último y de (4.3) podemos concluir

$$\frac{dX_2}{dX_1} = -\frac{P_2}{P_1}$$

En el punto óptimo la pendiente de la recta de ingreso y la pendiente de la curva de indiferencia son iguales.

Si llamamos a $\frac{\partial U}{\partial X_i}$ "utilidad marginal del bien i" (lo denotaremos U_i), de (4.4) tenemos

$$\frac{U_1}{p_1} = \frac{U_2}{p_2}$$

esto no es más que enel óptimo, las utilidades marginales son proporcionales a los precios. Esto era de esperar, ya que de otra manera, si por ejemplo $\frac{U_2}{P_2} > \frac{U_1}{P_1}$, entonces nosotros podríamos lograr aún un grado de mayor utilidad sacrificando cantidades del bien uno y comprando bienes del tipo dos.

De la ecuación (4.1) obtenemos en general

$$\frac{U_1}{p_1} = \ldots = \frac{U_n}{p_n} = \lambda$$

Para entender mejor el significado de λ derivemos el Lagrangiano con respecto a I y valuamos en el punto óptimo

$$\frac{\partial L}{\partial I} = \frac{\partial U}{\partial I} + \lambda = 0$$

de donde

$$\frac{\partial U}{\partial I} = \lambda$$

en la ecuación anterior se observa que λ indica una medida del promedio en el cual la utilidad cambia al variar el ingreso. Por esto se le llama a λ "utilidad marginal del dinero".

Por el teorema 2.5 maximizar U equivale a maximizar L y por las condiciones de segundo orden se tiene que la siguiente matriz es negativa definida

$$\begin{bmatrix} 0 & -p \\ -p' & H \end{bmatrix} = \begin{bmatrix} 0 & -p_1 & -p_2 & \cdots & -p_n \\ -p_1 & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \vdots & & & & \vdots \\ -p_n & \frac{\partial^2 L}{\partial x_n \partial x_n} & \frac{\partial^2 L}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{bmatrix}, \qquad (4.5)$$

donde H es el Hessiano de U y por lo tanto negativa definida; i.e.

$$p\, H\, p' < 0 \quad , \text{ para todo } p \neq 0$$

Es fácil probar que

$$\begin{bmatrix} 0 & -p \\ -p' & H \end{bmatrix}^{-1} = \begin{bmatrix} \mu & \mu p H^{-1} \\ \mu H^{-1} p' & \mu H^{-1} p' p H^{-1} + H^{-1} \end{bmatrix} \qquad (4.6)$$

donde

$$\mu = \frac{-1}{p\, H^{-1} p'} > 0$$

Podemos volver a formular las condiciones de K-T (4.1)

$$\begin{aligned} \varphi_1(\lambda, x, p, I) &= I - px = 0 \\ \varphi_2(\lambda, x, p, I) &= \frac{\partial U(x)}{\partial X} - \lambda p = 0 \end{aligned} \qquad (4.7)$$

Este sistema tiene solución (por el teorema de la función implícita), si el deteminante del Jacobiano es diferente de cero. pero el Jacobiano del sistema es

$$\begin{bmatrix} \frac{\partial \varphi_1}{\partial \lambda} & \frac{\partial \varphi_1}{\partial X} \\ \frac{\partial \varphi_2}{\partial \lambda} & \frac{\partial \varphi_2}{\partial X} \end{bmatrix} = \begin{bmatrix} 0 & -p \\ -p' & H \end{bmatrix}$$

y como vimos en la ecuación (4,6) el inverso de la matriz de la derecha existe, por lo tanto el determinante es distinto de cero. De donde encontramos como soluciones a las siguientes ecuaciones (funciones) :

$$\begin{aligned} x^* &= x^*(p, I) \\ \lambda^* &= \lambda^*(p, I) \end{aligned} \qquad (4.8)$$

A las n funciones x^* se les llama "funciones de demanda"; el teorema de la función implícita tambien nos garantiza que las funciones (4.8) tienen derivadas parciales continuas en una región de la solución de (4.7).

6. Funciones de demanda.

Una propiedad importante de las funciones de demanda es la homogeneidad. Decimos que una función $f(x)$ es homogenea de grado k si $f(\alpha x) = \alpha^k f(x)$.

Si suponemos que todos los precios se afectan por medio de un factor multiplicativo k, y el ingreso tambien aumenta en la misma proporción, entonces nuestras funciones de demanda no varían, por lo tanto son homogeneas de grado cero; la prueba sigue asi:

El Lagrangiano del problema se torna

$$L(x,\lambda) = U(x) + \lambda(kI - kpx)$$

las condiciones K-T son entonces

$$\frac{\partial L}{\partial x} = \frac{\partial U}{\partial x} - \lambda kp = 0 \tag{5.1}$$

$$\frac{\partial L}{\partial \lambda} = kI - kpx = 0 \tag{5.2}$$

La ecuación (5.1) da como resultado la (4.4)

$$\frac{U_i}{p_i} = \lambda \qquad , \text{ para todo } i = 1, \ldots, n$$

y la (5.2) nos da el resultado previamente obtenido de

$$I - px = 0$$

con esto hemos llegado al sistema inicial (4.1) y resumimos el resultado en un teorema

Teorema 1.

La función de demanda para cualquier comodidad es una función univaluada que depende de los precios y del ingreso; además son funciones homogeneas de grado cero.

$$x_i = x_i(p_1, p_2, \ldots, p_n; I)$$

Definimos "elasticidad dela función x_i con respecto al precio p_j " (ε_{ij}), como la derivada parcial de x_i con respecto a p_j , por $\frac{p_j}{x_i}$

$$\varepsilon_{ij} = \frac{p_j}{x_i}\frac{\partial x_i}{\partial p_j}$$

A veces se le llama "elasticidad propia" si i es igual a j y "elasticidad cruzada" en caso contrario.

Si en vez de derivar x_i con respecto a p_j , lo hacemos con respecto a I, obtenemos la elasticidad de la función demanda con respecto al ingreso (M):

$$\varepsilon_{i,M} = \frac{M}{x_i}\frac{\partial x_i}{\partial M}$$

Teorema 2.

$$\frac{p_1}{x_i}\frac{\partial x_i}{\partial p_1} + \ldots + \frac{p_n}{x_i}\frac{\partial x_i}{\partial p_n} + \frac{I}{x_i}\frac{\partial x_i}{\partial I} = 0$$

i.e.

$$\sum_{j=1}^{n} \varepsilon_{ij} + \varepsilon_{i,M} = 0 \qquad \text{, para toda } i = 1, \ldots, n.$$

La demostración se basa en el teorema de Euler, el cual establece que si $f_i(p_1, \ldots, p_n; I)$ es una función de grado r, entonces:

$$p_1 \frac{\partial f_i}{\partial p_1} + \cdots + M \frac{\partial f_i}{\partial M} = r\, f_i(p_1, \ldots, p_n ; M)$$

en nuestro caso f_i es la función de demanda x_i que sabemos que es de grado cero y de aquí la conclusión del teorema.

Para demostrar el teorema de Euler basta con derivar con respecto a k la siguiente ecuación

$$f_i\,(kp_1, \ldots, kp_n ; kI) = k^r f_i\,(p_1, \ldots, p_n ; I)$$

lo cual nos da

$$\frac{\partial f_i}{\partial kp_1} \frac{\partial kp_1}{\partial k} + \cdots + \frac{\partial f_i}{\partial kM} \frac{\partial kM}{\partial k} = r k^{r-1} \; f_i\,(p_1, \ldots, p_n ; M) \quad ,$$

haciendo K = 1 llegamos al resultado.

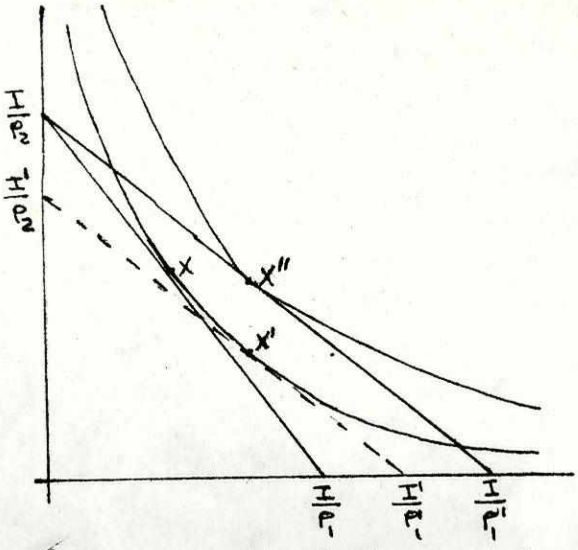
7. Análisis de sensitividad y ecuación de Slutsky.

Supongamos que el consumidor ha escogido el punto x como su mejor alternativa, esto quiere decir que para todos los puntos x' de la curva de indiferencia en que está situado x, se tiene la siguiente relación

$$px < px' \qquad (6.1)$$

Si el precio cambia de p a p', entonces la recta de ingreso varía. En la figura 6.1 se está trabajando con dos bienes y se supuso que el precio bajó para el bien uno y se mantuvo constante el precio del bien dos.

Fig. 6.1



Como el consumidor tiene una nueva recta de ingreso el punto óptimo será un x'' situado en una curva de indiferencia de mayor utilidad. Nosotros podemos imaginarnos una reducción en el ingreso del consumidor de I a I' de manera que contrarreste la baja de precio y mantenga al consumidor en la misma curva de indiferencia, en un punto nuevo x'; em este punto se mantiene la relación

$$p'x' < p'x \tag{6.2}$$

De (6.1) y (6.2) llegamos a la ecuación

$$(p' - p)(x' - x) < 0 \tag{6.3}$$

En el caso de n bienes, se supone que solo se hace variar el precio del bien i, p_i , con lo cual obtenemos

$$(p_i' - p_i)(x_i' - x_i) < 0$$

y se deduce que x_i' es distinto de x_i siempre que las curvas de indiferencia sean estrictamente convexas. Por lo tanto dividiendo ambos lados por $(x_i' - x_i)^2$;

$$\frac{\Delta p_i}{\Delta x_i} < 0 \tag{6.4}$$

con $\Delta p_i = p_i' - p_i$,y $\Delta x_i = x_i' - x_i$.

Hasta el momento hemos maximizado la función de utilidad U, y como consecuencia se obtuvo n+l funciones que corresponden a las condiciones de primer orden. Estas n+l funciones son las de demanda y la de utilidad marginal del dinero que tiene como parámetros a los precios y al ingreso, (p_i, I).

Me preocuparé de estudiar como varían estas funciones al variar los parámetros, a este estudio se le llama análisis de sensitividad y ya hicimos algo de esto en el comienso de esta sección.

Para comenzar sustituimos nuestra solución al problema de optimización en las condiciones de primer orden; esto es sustituir las ecuaciones (4.8) en las (4.1) y por el teorema 2.2,

$$\begin{aligned} &I - px^*(p,I) = 0 \\ &\frac{\partial U}{\partial x}(x^*(p,I)) - \lambda^*(p,I)p = 0 \end{aligned} \tag{6.5}$$

Consideremos los efectos de un cambio en el ingreso, diferenciando (6.5) con respecto a I

$$\begin{aligned} &1 - \sum_{j=1}^{n} p_j \frac{\partial x_j^*}{\partial I} = 0 \\ &\sum_{k=1}^{n} \frac{\partial^2 U}{\partial x_j \partial x_k} \frac{\partial x_k^*}{\partial I} - p_j \frac{\partial \lambda^*}{\partial I} = 0 \qquad 1 \leq j \leq n, \end{aligned} \tag{6.6}$$

si hacemos $\quad \frac{\partial x^*}{\partial I} = \left(\frac{\partial x_1^*}{\partial I}, \ldots, \frac{\partial x_n^*}{\partial I}\right)$

la ecuación (6.6) se nos vuelve

$$-p \frac{\partial x^*}{\partial I} = -1$$

$$-p' \frac{\partial \lambda^*}{\partial I} + H \frac{\partial x^*}{\partial I} = 0 \tag{6.6.a}$$

o, equivalente en notación vectorial (matricial);

$$\begin{bmatrix} 0 & -p \\ -p' & H \end{bmatrix} \begin{bmatrix} \dfrac{\partial \lambda^*}{\partial I} \\ \dfrac{\partial x^*}{\partial I} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \tag{6.6.b}$$

Como segundo paso consideremos un cambio en los precios y sus efectos, diferenciando con respecto a p_ℓ (6.5)

$$-x_\ell^* - \sum_{j=1}^{n} p_j \frac{\partial x_j^*}{\partial p_\ell} = 0$$

$$\sum_{k=1}^{n} \frac{\partial^2 U}{\partial x_j \partial x_k} \frac{\partial x_k^*}{\partial p_\ell} - p_j \frac{\partial \lambda^*}{\partial p_\ell} - \lambda^* \delta_{j\ell} = 0 \quad ; j = 1, \ldots, n . \tag{6.7}$$

donde $\delta_{j\ell}$ es la delta de Kroenecker (vale 1 si j es igual l y 0 si j es distinto de l). Si hacemos

$$\frac{\partial x^*}{\partial p} = \begin{bmatrix} \dfrac{\partial x_1^*}{\partial p_1} & \dfrac{\partial x_1^*}{\partial p_2} & \cdots & \dfrac{\partial x_1^*}{\partial p_n} \\ \vdots & & & \vdots \\ \dfrac{\partial x_n^*}{\partial p_1} & \dfrac{\partial x_n^*}{\partial p_2} & \cdots & \dfrac{\partial x_n^*}{\partial p_n} \end{bmatrix} \tag{6.8}$$

y $$\frac{\partial \lambda^*}{\partial p} = \left(\frac{\partial \lambda^*}{\partial p_1}, \ldots, \frac{\partial \lambda^*}{\partial p_n}\right) \qquad (6.8')$$

la ecuación (6.7) se puede escribir

$$\begin{aligned} -p \frac{\partial x^*}{\partial p} &= x^{*\prime} \\ -p' \frac{\partial \lambda^*}{\partial p} + H \frac{\partial x^*}{\partial p} &= \lambda^* I_n \end{aligned} \qquad (6.7.a)$$

o, equivalentemente

$$\begin{bmatrix} 0 & -p \\ -p & H \end{bmatrix} \begin{bmatrix} \dfrac{\partial \lambda^*}{\partial p} \\ \\ \dfrac{\partial x^*}{\partial p} \end{bmatrix} = \begin{bmatrix} x^{*\prime} \\ \\ \cancel{\gamma}^* I_n \end{bmatrix} \qquad (6.7.b)$$

λ^*

Finalmente consideremos como lo hicimos al principio de la sección un cambio en el precio, donde el ingreso se varía de manera tal que la utilidad se mantenga constante

$$\begin{aligned} dU &= \frac{\partial U}{\partial x} dx = \lambda\, p(dx) \\ dI &= p(dx) + (dp)x \end{aligned} \qquad (6.9)$$

como dU es igual a cero, entonces p(dx) es cero; por lo tanto para obtener U constante hay que asegurarse de que dI sea igual a (dp)x

Si derivamos las ecuaciones (6.5) con respecto a p_ℓ , pero hacemos al mismo tiempo $dI = (dp_\ell)x_\ell$, nos da

$$-\sum_{j=1}^{n} p_j \frac{\partial x_j^*}{\partial p_\ell} = 0$$

(6.10)

$$\sum_{k=1}^{n} \frac{\partial^2 U}{\partial x_j \partial x_k} \frac{\partial x_k^*}{\partial p_\ell} - p_j \frac{\partial \lambda^*}{\partial p_\ell} - \lambda^* \delta_{j\ell} = 0 \quad ; \quad 1 \leq j \leq n$$

estas ecuaciones pueden ser escritas

$$-p\left(\frac{\partial x^*}{\partial p}\right)_{comp} = 0$$

(6.10.a)

$$-p'\left(\frac{\partial \lambda^*}{\partial p}\right)_{comp} + H\left(\frac{\partial x^*}{\partial p}\right)_{comp} = \lambda^* I_n$$

donde $\left(\frac{\partial \lambda^*}{\partial p}\right)_{comp}$ y $\left(\frac{\partial x}{\partial p}\right)_{comp}$ son las derivadas parciales de (6.8) y (6.8'), pero con el ingreso compensado para mantener la utilidad constante. O, equivalentemente;

$$\begin{bmatrix} 0 & -p \\ -p' & H \end{bmatrix} \begin{bmatrix} \left(\frac{\partial \lambda^*}{\partial p}\right)_{comp} \\ \left(\frac{\partial x^*}{\partial p}\right)_{comp} \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda^* I_n \end{bmatrix} \qquad (6.10.b)$$

Resumiendo las ecuaciones (.6.b), (6.7.b) y (6.10.b) en una sola matriz es

$$\begin{bmatrix} 0 & -p \\ -p' & H \end{bmatrix} \begin{bmatrix} \frac{\partial \lambda^*}{\partial I} & \frac{\partial \lambda^*}{\partial p} & \left(\frac{\partial \lambda^*}{\partial p}\right)_{comp} \\ \frac{\partial x^*}{\partial I} & \frac{\partial x^*}{\partial p} & \left(\frac{\partial x^*}{\partial p}\right)_{comp} \end{bmatrix} = \begin{bmatrix} -1 & x^{*\prime} & 0 \\ 0 & \lambda^* I_n & \lambda^* I_n \end{bmatrix} \qquad (6.11)$$

Este resultado es llamado "Ecuación Matricial Fundamental de la Teoría del consumidor".

Recordamos que nuestro objetivo es recordar como varían nuestras funciones x^* y λ^* con respecto a los parámetros, por lo tanto despejando en (6.11)llegamos a

$$\begin{bmatrix} \frac{\partial \lambda^*}{\partial I} & \frac{\partial \lambda^*}{\partial p} & \left(\frac{\partial \lambda^*}{\partial p}\right)_{comp} \\ \frac{\partial x^*}{\partial I} & \frac{\partial x^*}{\partial p} & \left(\frac{\partial x^*}{\partial p}\right)_{comp} \end{bmatrix} = \begin{bmatrix} 0 & -p \\ -p' & H \end{bmatrix}^{-1} \begin{bmatrix} -1 & x^{*\prime} & 0 \\ 0 & \lambda^* I_n & \lambda^* I_n \end{bmatrix} \quad (6.12)$$

sustituyendo (4.6) en (6.12) y efectuando la multiplicación,el resultado para los cambios de la funciones de demanda son

$$\frac{\partial x^*}{\partial I} = -\mu H^{-1} p' \quad (6.13)$$

$$\frac{\partial x^*}{\partial p} = \mu H^{-1} p' x^{*\prime} + \mu H^{-1} p' p H^{-1} \lambda^* + H^{-1} \lambda^* \quad (6.14)$$

$$\left(\frac{\partial x^*}{\partial p}\right)_{comp} = \mu H^{-1} p' p H \lambda^* + H^{-1} \lambda^* \quad (6.15)$$

en particular estas ecuaciones se pueden combinar para dar la "Ecuación de Slutsky"

$$\frac{\partial x_j^*}{\partial p_\ell} = \left(\frac{\partial x_j^*}{\partial p_\ell}\right)_{comp} - \frac{\partial x_j^*}{\partial I} x_\ell^* \quad ; \; j,l = 1,\ldots,n. \quad (6.16)$$

en notación vectorial es

$$\frac{\partial x^*}{\partial p} = \left(\frac{\partial x^*}{\partial p}\right)_{comp} - \frac{\partial x^*}{\partial I} x^* \quad (6.16.a)$$

al término $\frac{\partial x_j^*}{\partial p_\ell}$ se le llama "efecto total", al término $\left(\frac{\partial x_j^*}{\partial p_\ell}\right)_{comp}$ "efecto sustitución" y al último $\left(-\frac{\partial x_j^*}{\partial I}\right) x_\ell$ el"efecto ingreso".

Volviendo nuevamente al inicio de la sección, notamos que el resultado (6.4)es el cambio compensado para variaciones de la función de demanda del bien i con respecto a variaciones del bien j; haciendo que las variaciones sean infinitesimales (6.4) se torna:

$$\left(\frac{\partial x_i}{\partial p_i}\right)_{comp} < 0 \quad , \qquad i = 1, \ldots, n. \qquad (6.17)$$

Además haciendo una comparación de la ecuación (6.16) para el caso de dos bienes y la figura 6.1 vemos que el efecto total es el cambio de x a x", el efecto sustitución es el cambio de x' a x".

En la fig. 6.1 que x" está a la izquierda de x', esto quiere decir que al aumentar el ingreso de I a I', el consumidor deja de comprar el bien tipo uno, a esta clase de bienes cuyo consumo disminuye al aumentar el ingreso se les llama "bienes inferiores" , $\left(\frac{\partial x_i^*}{\partial I} < 0\right)$, contrario sería si aumentara su consumo al aumentar el ingreso, a estos se les llama "bienes superiores" $\left(\frac{\partial x_j^*}{\partial I} > 0\right)$ .Como x" no puede ser menor que x', deducimos que en general tiene que haber un bien superior.

Notamos tambien que la demanda del bien uno disminuyó al disminuir el precio, a este tipo se le llama "bien Giffen", en el caso conrio se le llama "bien normal".

Por último, en el caso de n biemes puede suceder que una reducción compensada del bien i, baje la demanda del bien j, entonces decimos que el bien i y el bien j son "sustitutos"; en caso contrario, si un cambio compensado en el precio del bien i induce a una reducción del bien j, decmos que el bien i y el bien j son "complementarios".

En resumen, hemos hecho la serie de categorizaciones siguientes:

$$\left\{\begin{matrix}\text{normal}\\ \text{Giffen}\end{matrix}\right\} \text{ si } \frac{\partial x_j^*}{\partial p_j} \left\{\begin{matrix}<\\ >\end{matrix}\right\} 0$$

$$\left\{\begin{matrix}\text{superior}\\ \text{inferior}\end{matrix}\right\} \text{ si } \frac{\partial x_j^*}{\partial I} \left\{\begin{matrix}>\\ <\end{matrix}\right\} 0$$

$$\left\{\begin{matrix}\text{sustitutos}\\ \text{complementerios}\end{matrix}\right\} \text{ si } \left(\frac{\partial x_i^*}{\partial p_j}\right)_{comp} \left\{\begin{matrix}>\\ <\end{matrix}\right\} 0 \qquad i \neq j$$

Con esta nueva terminología podemos seguir adelante. Si multiplicamos (6.15) por p' nos da

$$\left(\frac{\partial x^*}{\partial p}\right)_{comp} p' = 0 \tag{6.18}$$

en notación de suma esto es:

$$\sum_{i=1}^{n} \left(\frac{\partial x_j^*}{\partial p_i}\right)_{comp} p_i = 0 \qquad , j= 1,\ldots, n. \tag{6.18.a}$$

Por la ecuación (6.17) sabemos que los elementos de la diagonal son negativos, es necesario que para alguna i suceda

$$\left(\frac{\partial x_j}{\partial p_i}\right)_{comp} > 0 \qquad , j= 1,\ldots, n. \tag{6.19}$$

queda demostrado el teorema siguiente

Teorema 6.1

Para cada bien existe un sustituto. En particular, si solo hay dos bienes estos son sustitutos.

Antes de seguir adelante, podemos verificar el teorema 2 de **la** sección 5 si sustituimos la ecuación (6.18) en la ecuación de Slutsky (6.16.a)

$$\left(\frac{\partial x^*}{\partial p}\right)p' + \left(\frac{\partial x^*}{\partial I}\right)I = 0 \qquad (6.20)$$

además multiplicando por la izquierda la ecuación (6.13) por p, entonces

$$p\left(\frac{\partial x^*}{\partial I}\right) = 1 \qquad (6.21)$$

en notación matricial, pasándolo a sumatoria esto es

$$\sum_{i=1}^{n} p_i\left(\frac{\partial x_i^*}{\partial I}\right) = 1 \qquad (6.21.a)$$

lo que nos dice este resultado es que por lo menos un bien es superior, pues uno de los términos de la sumatoria tiene que ser positivo para que el resultado sea uno, enunciamos la propiedad;

Teorema 2.

Al menos existe un bien superior.

Se deduce de la ecuación de Slutsky (ó (6.20));

Teorema 3.

Todo bien superior es normal

Y para finalizar, citaremos la condición de sumatoria de "Cournot".

Teorema 4.

La cantidad demandada del bien j es el negativo de la combinación lineal de los precios de cada bien i por el cambio en la demanda del bien i con respecto al precio del bien j, esto es

$$x_j^* = - \sum_{i=1}^{n} p_i \frac{\partial x_i^*}{\partial p_j} \qquad j= 1,\dots, n.$$

Para la comprobación, multiplicaremos la ecuación de Slutsky (6.16) por p y utilizamos los resultados (6.21) y (6.18), llegando a :

$$p\left(\frac{\partial x^{*}}{\partial p}\right) + x^{*\prime} = 0 \ .$$

8. Bibliografía.

Apostol, T. Calculus. Vol. II. New York, Blaisdell Publishing Co., 1962.

Arrow, K. y F. Hahn. General competitive analysis. San Francisco, Holden-Day, Inc.

Green, J. Consumer theory. Middlesex, Penguin Books Ltd., 1971.

Henderson, J. M. y R. F. Quandt. Microeconomic theory: a mathematical approach. Tokio, McGraw-Hill Kaga Kusha Ltd., 1971.

Intriligator, M. D. Mathematical optimization and economic theory. New Jersey, Prentice Hall, 1971.

Luenberger, D. Introduction to linear and non linear programming. Filipinas, Wesley Publishing Co. Inc., 1973.

Zangwill, W. I. Nonlinear programming: a unified approach. New Jersey, Prentice Hall, 1969.

Apéndice

A. Conceptos de álgebra lineal.

En todo mi estudio se toma al espacio vectorial $V = R^n = E^n$. Recordamos que un producto interior definido en V es una función f de VxV en R con las propiedades siguientes:

i) $f(x, y) = f(y, x)$

ii) $f(x, y+z) = f(x, y) + f(x, z)$

iii) $f(cx, y) = cf(x, y)$

iv) $f(x, x) = 0$ si $x \neq 0$, donde x es un vector columna nx1, $c \in R$.

Un espacio vectorial con producto interior se llama "espacio euclideo". Si tenemos un operador lineal T simétrico definido en un espacio euclideo tal que $(T(x), y) = (x, T(y))$ para todo x, y en V. Este operador induce una forma cuadrática Q,

$$Q : V \longrightarrow R \quad \text{con } Q(x) = (T(x), x).$$

Se puede probar que si A es la matriz nxn asociada con el operador T, entonces

$$Q(x) = x' A x.$$

Si tenemos una matriz A cualquiera nxn, se puede probar que

$$x' A x = x' B x, \quad \text{donde } B = 1/2(A + A'),$$

es simétrica, y A' indica la transpuesta de A.

En resumen, toda matriz induce una forma cuadrática.

Por último, si A es una matriz simétrica real,

$$Q(y) = y' A y = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} y_i y_j .$$

Si se tiene

$Q(y) > 0$, se dice que Q es una forma cuadrática positiva definida;

$Q(y) \geqslant 0$, se dice que Q es una forma cuadrática positiva semidefinida;

$Q(y) < 0$, se dice que Q es una forma cuadrática negativa definida;

$Q(y) \leqslant 0$, se dice que Q es una forma cuadrática negativa semidefinida.

B. Máximos, mínimos y puntos silla.

Definición. Llamaremos "campo escalar" a una función que tiene como dominio a un subconjunto de R^n y como contradominio R.

Definición. Un campo escalar tiene un "máximo" (absoluto o global) en el punto x° de un conjunto S en R si

$$f(x^\circ) \geqslant f(x) \quad , \text{ para todo } x \text{ en } S.$$

Si la relación anterior se comple para todo x en una vecindad de x°, decimos que x° es un "máximo relativo".

Nota: las definiciones para mínimo (absoluto o global) y mínimo relativo son análogas, sólo se cambia el sentido de la desigualdad en la definición.

Definición. Llamaremos "punto extremo" de f a un punto que es máximo o mínimo relativo.

Definición. Sea f diferenciable. x° es "punto estacionario de f" si $\nabla f(x^\circ) = 0$.

Definición. Un punto estacionario x° es "punto silla" se para toda vecindad de x° existen x & y tal que

$$f(x^\circ) > f(x) \quad y \quad f(x^\circ) < f(y).$$

Fórmula de Taylor para campos escalares.

Si f es un campo escalar diferenciable y "a" un punto del dominio, la fórmula de Taylor nos da para una aproximación de primer orden:

$$f(a + y) - f(a) = \nabla f(a)'y + \|y\|E(a, y),$$

donde E(a, y) tiende a 0 si y tiende a 0.

Si "a" es un punto estacionario, $\nabla f(a) = 0$ y tenemos que

$$f(a+y) - f(a) = \|y\| E(a, y).$$

Para tener más información del punto estacionario "a", hay que analizar el miembro derecho de la ecuación. El siguiente teorema garantiza que si las segundas derivadas parciales de f son contínuas en "a" se da

$$\|y\| E(a, y) = (1/2) \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij} f(a) y_i y_j + \text{términos de orden}$$

menor que $\|y\|^2$, donde $D_{ij} f(a) = [D_i(D_j f)](a)$.

Llamaremos "matriz Hessiana de f" a

$$H(x) = [D_{ij} f(a)]_{i,j=1}^{n}$$

Podemos enunciar ahora el teorema para la fórmula de Taylor de segundo orden para campos escalares.

Teorema A.1. Sea f un campo escalar con segundas derivadas parciales contínuas en una vecindad de a. Entonces, para toda y en $\dot{R}$ tal que a+y esté en la vecindad de a, se tiene

$$f(a+y) - f(a) = \nabla f(a)'y + (1/2)y'H(a+ y)y, \quad 0 < \lambda < 1;$$

ésto último se puede escribir también

$$f(a+y) - f(a) = \nabla f(a)'y + (1/2)y'H(a)y + \|y\|^2 E_2(a, y),$$

$E (a, y) \rightarrow 0$ si $y \rightarrow 0$.

Naturaleza de puntos estacionarios.

Si "a" es un punto estacionario, $\nabla f(a) = 0$ y

$$f(a+y) - f(a) = (1/2)y'H(a)y + \|y\|^2 E_2(a, y).$$

Como el término $E_2(a, y)$ va a cero más rápidamente que $\|y\|^2$, podemos esperar que el signo de f(a+y) - f(a) esté dado por el de la forma cuadrática (1/2)y'H(a)y.

Teorema A.2. Si A es una matriz real simétrica y Q su forma cuadrática asociada, entonces

i) $Q(y) > 0$ para todo $y \neq 0$ ssi todos los valores propios son positivos;

ii) $Q(y) < 0$ para todo $y \neq 0$ ssi todos los valores propios son negativos.

Teorema A.3. Sea f un campo escalar con segundas derivadas parciales contínuas en una vecindad de un punto "a". Entonces

i) Si los valores propios de H(a) son positivos, a es mínimo relativo de f;

ii) Si los valores propios de H(a) son negativos, a es máximo relativo de f;

iii) Si los valores propios de H(a) son positivos y negativos, a es punto silla de f.

Las demostraciones a los anteriores teoremas se pueden encontrar en Apostol "Calculus", Vol. II, N. Y., Blaisdell Publishing Co., 1962, que es el libro que he tomado como base para esta sección.

C. Programación lineal.

El problema siguiente:

$$\text{(A.1)} \qquad \begin{aligned} &\text{maximizar } q'x \\ &\text{sujeto a } Ax = b \\ &\qquad\qquad x \geqslant 0, \end{aligned}$$

donde A es una matriz mxn, es el problema típico de la programación lineal (PL). El problema dual es

$$\text{(A.2)} \qquad \begin{aligned} &\text{minimizar } b'u \\ &\text{sujeto a } A'u \geqslant q \\ &u \text{ sin restricciones.} \end{aligned}$$

Si cualquiera de los problemas (A.1) y (A.2) tiene solución óptima dualidad, (A.5) se satisface con valor de la función objetivo igual a cero. Pero esto implica que (A.3) es cierto.

finita, entonces el otro lo tiene y el valor de las funciones objetivo es el mismo. Al par de problemas se le denomina "problemas duales".

<u>Teorema</u> (Lema de Farkas). La expresión

(A.3) $\quad q'x \leq 0$ para todo x, tal que $Ax \geq 0$

es equivalente a la expresión

(A.4) $\quad$ Existe $u \geq 0$ tal que $q + A'u = 0$.

Demostración. Consideremos el problema

$$\text{(A.5)} \qquad \begin{array}{l} \text{maximizar } q'x \\ \text{sujeto a } Ax \geq 0. \end{array}$$

El problema dual es

$$\text{(A.6)} \qquad \begin{array}{l} \text{minimizar } 0 \\ \text{sujeto a } -A'u = q \\ \qquad\qquad\quad u \geq 0. \end{array}$$

Si la expresión (A.3) es cierta, haciendo $x = 0$ obtenemos la solución óptima finita para (A.5), con valor de la función objetivo igual a cero. Por dualidad tenemos que (A.6) tiene solución, y esto implica que (A.4) se verifica.

Conversamente, si (A.4) se verifica, entonces (A.6) tiene solución óptima finita, pues ocupamos el valor de u dado por (A.4) para satisfacer (A.6) y que hace que la función objetivo valga cero. Por