UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

INTRODUCCIÓN A LA TEORÍA DE LA OPTIMIZACIÓN. MATERIAL PARA LA ENSEÑANZA DE LA ECONOMÍA MATEMÁTICA

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICO

P R E S E N T A:

FRANCISCO VALDEZ NÚÑEZ

TUTOR
DR. SERGIO HERNÁNDEZ CASTAÑEDA

2007

A la memoria de mi abuelita Eva

Y a todas aquellas personas que
hicieron posible este gran logro, gracias.
En especial, a mi mamá Luz María
y a Alfredo quienes han sido un gran ejemplo;
a mis tías Lety y Lulú por su apoyo incondicional.
Y a los dos grandes amores de mi vida,
mi esposa Alejandra y mi hija Emily Sofía
quien es ya un nuevo motor en mi vida.

Índice general

Introducción

La optimización, se puede considerar como la búsqueda de la mejor solución entre las posibles a un problema determinado. De hecho, este ejercicio mental lo realizamos cuando cotidianamente elegimos, entre diferentes opciones, la más adecuada.

La teoría que nos proporciona los resultados y herramientas precisos para estudiar este tipo de problemas es la "optimización matemática". El desarrollo de esta teoría no es reciente, ya que, aunque las aportaciones más importantes se produjeron en los años cuarenta y cincuenta del siglo XX, muchos de los resultados se conocían ya en el siglo XVIII.

Dada la diversidad de áreas y materias en las que se plantean problemas de optimización, como son la Ingeniería, Biología, Física, Investigación de Operaciones, entre otras, la presente tesis ha sido concebida como una introducción a las técnicas de optimización y su aplicación a la teoría económica.

Sin embargo, cabe aclarar que no se abordan todos los aspectos de la optimización matemática, la intención de este trabajo es, presentar la teoría de programación lineal y la teoría de la programación cóncava. Por otra parte, se dan por conocidos conceptos de álgebra lineal y análisis matemático, esto último viene motivado porque la tesis ha sido pensada para que pueda ser accesible a alumnos que, de alguna manera ya poseen estos conocimientos, en particular está dirigida a los alumnos de quinto semestre de las carreras de matemáticas y actuaría de la Facultad de Ciencias, interesados en abordar la problemática económica. No obstante, también puede resultar útil a estudiantes y profesores de economía interesados en la economía matemática, pues resultaría apropiado como complemento de algunos cursos.

Es preciso señalar que el contenido teórico de este trabajo, así como los ejemplos de la función de utilidad casi Bernoulli y Bernoulli están basados en las notas del profesor Sergio Hernández Castañeda. Me propuse enriquecer dichas notas con algunos ejemplos de aplicación, así como con ideas geométricas; además de reescribir las demostraciones con un mayor desarrollo analítico y haciendo explícitos varios pasos para su mejor comprensión. Lo anterior, con el fin de que los alumnos interesados en profundizar en un tratamiento más riguroso de la parte teórica, puedan acceder a las demostraciones de los principales teoremas que fundamentan dicha teoría, apoyándose en los gráficos y en los problemas de aplicación. Otra de las ventajas es que la notación usada es la del cálculo diferencial e integral y la del álgebra lineal que a mi modo de ver es más conocida por los estudiantes.

Así pues, este trabajo se distribuye en seis capítulos cuyo orden y clasificación obedece a un aumento gradual de su dificultad, siendo la estructura de cada uno de ellos la siguiente:

El primer capítulo es introductorio y de motivación en donde se plantean, sin solución, algunos de los diferentes tipos de problemas de optimización que aparecen en la teoría económica. En particular se aborda el planteamiento del problema del consumidor con funciones de utilidad

Bernoulli. Dichos problemas se van resolviendo conforme se va adquieriendo la herramienta teórica necesaria para ello.

En los capítulos 2, 3, y 5, se exponen los aspectos teóricos de los métodos modernos de optimización matemática, ya que éstos constituyen una parte importante de los instrumentos fundamentales para abordar los ejemplos económicos del presente trabajo. Para presentarlos, se sigue a grandes rasgos el proceso cronológico del surgimiento de las ideas. Así, en el capítulo 2 partiremos de la problemática inicial manejada por Leibnitz. En el capítulo 3, esbozaremos cómo fue desarrollada esa problemática inicial por Lagrange y, por último en el capítulo 5, se plantea la problemática moderna desarrollada por Kuhn, Tucker y Uzawa, además se incluyen los aspectos teóricos para resolver problemas de programación cóncava con restricciones de no negatividad.

En los capítulos 4 y 6 se trata el problema del consumidor en forma analítica.

En el capítulo 4, se supone que la función de utilidad es de tipo casi Bernoulli o, equivalentemente, de tipo Cobb-Douglas. Aquí, el instrumento principal es el clásico teorema de Lagrange.

En el capítulo 6, generalizando la situación anterior, se supone que la función de utilidad puede expresarse como la suma de ciertas funciones estrictamente cóncavas. Para esto, ahora el instrumento analítico fundamental viene a ser el teorema de Kuhn Tucker.

Toda está temática no fue fácil de abordar, pero durante el desarrollo de este trabajo tuve la gran fortuna de recibir valiosos comentarios y sugerencias del grupo de Economía Matemática y Teoría de Juegos de la Facultad de Ciencias de esta institución, cuyos integrantes enlisto en órden alfabético, ya que para mí todas sus contribuciones han sido igual de importantes. Alejandra Calvo, Salvador Ferrer, Sergio Hernández, Enrique Hueda, Alma Jiménez, María Esther Linares, Gabriela Marmolejo, César Sousa, Claudia Villegas y Paloma Zapata; también quiero agradecer al profesor Oscar Palmas, pues aunque no pertenece al grupo, me hizo comentarios y observaciones de gran utilidad para mejorar la redacción de éste trabajo y por supuesto a mi esposa Alejandra, quien me asesoró en la captura y uso de Scientific Work Place, así como en la elaboración de las gráficas. Naturalmente cualquier error o deficiencia son míos.

Y en especial, no puedo dejar de mostrar mi reconocimiento y más profundo agradecimiento al profesor Sergio Hernández Castañeda por el material que me facilitó y todas las horas que me dedicó para poder realizar este trabajo. A todos ellos nuevamente ¡GRACIAS!.

FVN 2007.

Capítulo 1

Algunos problemas de optimización en la teoría económica

Introducción

En la teoría de la optimización, muchos de los problemas que se plantean tienen una estructura lógica común, ya que la idea que subyace es encontrar la mejor solución posible en determinadas circunstancias. Dicha solución permitirá decidir las medidas a tomar para alcanzar los objetivos que se persiguen. Este tipo de problemas aparece en distintas disciplinas, en particular en economía, como el caso de los problemas del consumidor y de la empresa.

Por otro lado, parece lógico plantearse, en primer lugar, sobre qué variables podemos decidir. En el caso del consumidor, se trata de decidir la cantidad de cada uno de los bienes y servicios disponibles que el consumidor debe adquirir, dados los precios que rigen en el mercado, de tal forma que maximice su utilidad o satisfacción. En cambio, la empresa tiene que elegir las cantidades de insumos a utilizar y de bienes a producir de tal forma que obtenga los máximos beneficios o en su caso los costos mínimos.

En este capítulo únicamente expondremos el planteamiento de algunos problemas de optimización aplicados a la economía, con el objetivo de mostrar cómo quedarían expresados en términos matemáticos y así poder ver la aplicación de algunos de los teoremas contenidos en la presente tesis.

Ejemplos de optimización sin restricciones

Maximizar los beneficios de una empresa

Supongamos una empresa que produce un solo tipo de bien, en una cantidad Q a la cual se le puede estimar una condición de demanda del mercado mediante la siguiente expresión:

$$4P + Q - 16 = 0 \tag{1.1}$$

donde la constante positiva P denota el precio del bien en el mercado. Además supongamos que la empresa también puede estimar el costo promedio de producción de una unidad del bien y lo puede expresar en términos de la cantidad que produce por medio de la siguiente función de costos:

$$C(Q) = \frac{4}{Q} + 2 - \frac{3}{10}Q + \frac{1}{20}Q^2$$

entonces, la empresa se plantea el problema de encontrar la producción Q adecuada para obtener el máximo ingreso y el máximo beneficio dadas las condiciones anteriores.

Para empezar se construye la función de ingreso total $I_T(Q)$ y la función de costo total $C_T(Q)$, donde Q es la variable de decisión.

Definimos los ingresos totales como el precio multiplicado por la cantidad producida, en este ejemplo de la ecuación 1.1 podemos despejar a P y escribir la función de ingresos mediante la siguiente expresión:

$$\begin{aligned} I_T(Q) &= P * Q \\ &= \left(\frac{16 - Q}{4}\right) Q \\ &= 4Q - \frac{1}{4}Q^2; \end{aligned} \tag{1.2}$$

por consiguiente el problema al que se enfrenta la empresa para maximizar sus ingresos se define como:

$$\begin{aligned} \text{máx } I_T(Q) &= 4Q - \frac{1}{4}Q^2 \\ \text{s.a. } Q &\geq 0. \end{aligned}$$

Definimos los beneficios de la empresa como los ingresos totales menos los costos totales. Así, definimos la función de costos totales $C_T(Q)$ como el producto de la función de costos por la cantidad producida, es decir:

$$\begin{aligned} C_T(Q) &= C(Q) * Q \\ &= \left(\frac{4}{Q} + 2 - \frac{3}{10}Q + \frac{1}{20}Q^2\right) Q \\ &= 4 + 2Q - \frac{3}{10}Q^2 + \frac{1}{20}Q^3 \end{aligned} \tag{1.3}$$

Luego de las ecuaciones 1.2 y 1.3 la función de beneficios queda definida por:

$$\begin{aligned} B(Q) &= I_T(Q) - C_T(Q) \\ &= 4Q - \frac{1}{4}Q^2 - \left(4 + 2Q - \frac{3}{10}Q^2 + \frac{1}{20}Q^3\right) \\ &= -4 + 2Q + \frac{1}{20}Q^2 - \frac{1}{20}Q^3 \end{aligned}$$

Por lo tanto, el problema de maximización de los beneficios de la empresa se plantea como:

$$\begin{aligned} \text{máx } B(Q) &= -4 + 2Q + \frac{1}{20}Q^2 - \frac{1}{20}Q^3 \\ \text{s.a. } Q &\geq 0. \end{aligned}$$

Obsérvese que la condición de no negatividad, dada por $Q \geq 0$, que aparece en ambos problemas de optimización que se plantea la empresa, es considerada únicamente al final del problema, pues de los candidatos a ser máximo se descartan los que no cumplan con esta restricción. Por lo que el problema se puede resolver como un problema de optimización sin restricciones, que se ve representado en la siguiente gráfica:

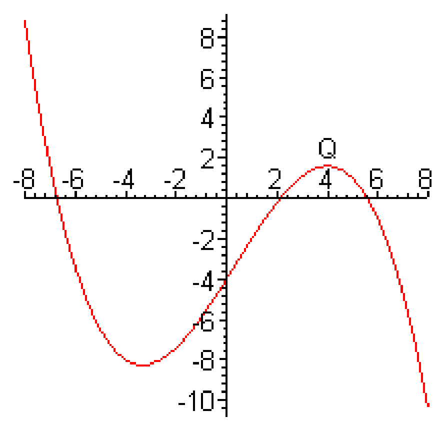


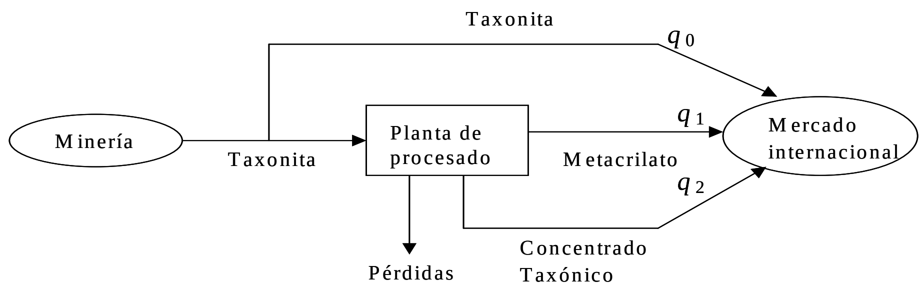
Figura 1-1

Una política óptima de producción

El gobierno de cierto país vendía en bruto el mineral de taxonita extraído de sus minas, sin ningún procesado previo del mismo. Para generar un valor añadido a su producción minera, se crea la empresa nacional Taxonita, S.A. (TAXSA) que proyecta instalar una planta de procesado de taxonita, de la que se piensa obtener dos productos diferentes:

- El metracrilato de taxonita, líquido a temperatura ambiente.
- El concentrado taxónico, residuo sólido de alta densidad.

Parte de la producción del mineral de taxonita se exportará directamente, y parte será procesado para obtener dichos productos, los cuales también se exportarán. La política anterior se expresa en el siguiente diagrama:



Donde q_0, q_1 y q_2 son las cantidades de los diferentes productos del mineral de taxonita, medidas en millones de toneladas.

Supondremos que el procesado de taxonita tiene un rendimiento de $\alpha\,\%$ en metacrilato, $\beta\,\%$ en concentrado taxónico y $\gamma\,\%$ en pérdidas, donde se verifica la relación $\alpha+\beta+\gamma=1$, y además que los precios de venta de dichos productos disminuyen con la cantidad ofertada en el mercado, según las funciones de demanda:

$$\begin{aligned} P_1 &= (4-q_1)10^3 \\ P_2 &= (3-q_2)10^3 \end{aligned} \tag{1.4}$$

dadas en unidades monetarias por tonelada, donde P_1 y P_2 son los precios del metacrilato y concentrado taxónico respectivamente.

También se supone que el costo de extracción del mineral de taxonita es de 1000 unidades monetarias por tonelada, y que el costo de operación de la planta de procesado es de 200 unidades monetarias por tonelada de taxonita tratada.

El problema que se plantea el gobierno es determinar la capacidad óptima de la planta de procesado de taxonita que debe instalar TAXSA para maximizar su beneficio de explotación; suponiendo que los costos de la inversión son asumidos totalmente por el estado. Para plantear el problema, se debe construír la función de beneficios de la empresa TAXSA, la cual se define como:

$$\text{Beneficio} = \text{Ingreso total - Costo total.}$$

Donde el ingreso total de la empresa es el precio de cada uno de los dos productos en el mercado, multiplicado por la cantidad que se produce en la planta, es decir:

$$Ingreso\ Total = P_1q_1 + P_2q_2$$

y los costos totales de la empresa se obtienen mediante la siguiente expresión:

$$Costo\ Total = 1000q + 200q$$

donde el primer sumando representa los costos de extracción del mineral y el segundo representa los costos de procesado, siendo q la cantidad total de mineral destinada a la planta.

Así pues los beneficios de la empresa están dados por:

$$B = P_1q_1 + P_2q_2 - (1000q + 200q)$$

sujeto a que la empresa debe satisfacier las restricciones de demanda dadas por las ecuaciones 1.4, las restricciones de rendimiento en el procesado de la planta,

$$\begin{aligned} q_1 &= \alpha q \\ q_2 &= \beta q \end{aligned}$$

y la restricción de no negatividad:

$$q \geq 0.$$

Por consiguiente, el problema queda definido en términos matemáticos como sigue:

$$\begin{aligned} \text{máx } B &= P_1 q_1 + P_2 q_2 - (1000q + 200q) \\ \text{s.a.} P_1 &= (4 - q_1)10^3 && (1.5) \\ P_2 &= (3 - q_2)10^3 && (1.6) \\ q_1 &= \alpha q && (1.7) \\ q_2 &= \beta q && (1.8) \\ q &\geq 0 && (1.9) \end{aligned}$$

para simplificar se sustituyen las restricciones 1.7 y 1.8 en 1.5 y 1.6, de modo que se obtiene lo siguiente:

$$\begin{aligned} P_1 &= (4 - \alpha q)10^3 \\ P_2 &= (3 - \beta q)10^3 \end{aligned}$$

así, sustituyendo en la función de beneficios obtenemos la siguiente expresión:

$$B = (4 - \alpha q)10^3 \alpha q + (3 - \beta q)10^3 \beta q - 1200q$$

simplificando términos y dividiendo entre 10^3 se tiene que el problema queda definido como sigue:

$$\begin{aligned} \text{máx } B(q) &= (4\alpha + 3\beta - \frac{5}{6})q - (\alpha^2 + \beta^2)q^2 \\ \text{s.a. } q &\geq 0. \end{aligned}$$

Un monopolio

Supóngase que una empresa monopolista[1] vende dos productos 1 y 2. Ahora bien, la empresa ha podido estimar por algún método las demandas de cada uno de estos mercados, mediante las siguientes ecuaciones:

$$\begin{aligned} \frac{1}{10}p_1 - \frac{6}{5} + \frac{1}{5}x_1 &= 0 \\ 10p_2 - 320 + 40x_2 &= 0 \end{aligned}$$

[1] Los mercados se pueden clasificar en relación al número de vendedores de un tipo de bien. Un mercado con un solo vendedor es un monopolio, con dos es un duopolio y con un número reducido pero superior a dos es un oligopolio.

En donde p_1, p_2, x_1 y x_2 son respectivamente los precios y las cantidades de los productos 1 y 2. Supongamos además que la empresa ha estimado su función de costos totales $C(x_1, x_2)$ por medio de la siguiente expresión:

$$C(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2 \tag{1.10}$$

Entonces la empresa se plantea el problema de cómo deben ser los precios del bien en cada mercado a fin de maximizar su beneficio.

Dado que el monopolista se plantea un problema cuya solución permitirá conocer los precios que debe imponer en cada mercado, entonces las variables de decisión del problema serán:

$$\begin{aligned} p_1 &= \text{precio a fijar por el monopolista en el mercado I} \\ p_2 &= \text{precio a fijar por el monopolista en el mercado II} \end{aligned}$$

y la función de beneficios B deberá estar en términos de las variables de decisión, así como en el ejemplo anterior, la función de beneficios queda definida por:

$$B(x_1, x_2) = \text{Ingresos totales} - \text{Costos totales.}$$

Luego, los ingresos totales estarán definidos como los ingresos del mercado I más los ingresos del mercado II, así se tiene:

$$\text{Ingresos Totales} = p_1x_1 + p_2x_2. \tag{1.11}$$

Utilizando las ecuaciones 1.10 y 1.11, podemos escribir una expresión para los beneficios dada por:

$$B(x_1, x_2) = p_1x_1 + p_2x_2 - \left(x_1^2 + 2x_1x_2 + x_2^2\right) \tag{1.12}$$

y finalmente para obtener la función de beneficios en términos de los precios, despejamos p_1 y p_2 de las funciones de demanda del problema, así se tiene que:

$$\begin{aligned} p_1 &= 12 - 2x_1 \\ p_2 &= 32 - 4x_2 \end{aligned}$$

y sustituyendo estos valores en la función de beneficios 1.12 podemos escribir la siguiente expresión:

$$\begin{aligned} B(x_1, x_2) &= (12 - 2x_1)\, x_1 + (32 - 4x_2)\, x_2 - \left(x_1^2 + 2x_1x_2 + x_2^2\right) \\ &= 12x_1 - 3x_1^2 + 32x_2 - 5x_2^2 - 2x_1x_2 \end{aligned}$$

luego el problema que enfrenta la empresa monopolista es:

$$\begin{aligned} \text{máx}\, B(x_1, x_2) &= 12x_1 - 3x_1^2 + 32x_2 - 5x_2^2 - 2x_1x_2 \\ \text{s.a. } x_1 &\geq 0 \\ x_2 &\geq 0. \end{aligned}$$

y su gráfica se puede ver en la figura 1-2.

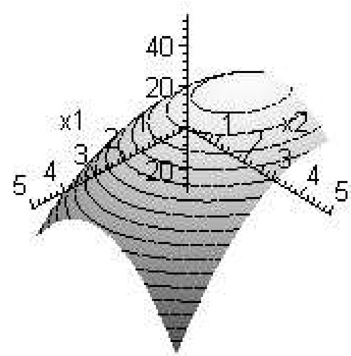


Figura 1-2

Ejemplos de optimización con restricciones de igualdad

Los siguientes ejemplos son problemas de optimización con restricciones de igualdad. Es importante señalar que algunos problemas que aparentan ser de este tipo se pueden transformar en problemas de optimización sin restricciones despejando alguna variable de decisión en las restricciones para después sustituirla en la función objetivo, como es el caso del ejemplo 1 en la página 9. No obstante, los siguientes ejemplos son problemas en los que esto no sucede, por lo tanto sólo pueden ser resueltos haciendo uso del teorema de Lagrange.

Programa de admisión universitario

Considérese el problema al que se enfrenta un administrador universitario que desea maximizar el ingreso de las cuotas universitarias. La Universidad está obligada por el gobierno a cobrar una cuota específica a los estudiantes nacionales, pero puede cobrar lo que desee a estudiantes extranjeros. Pareciera entonces, que el objetivo de la Universidad es inscribir a tantos estudiantes extranjeros como se pueda, pero si inscribe demasiados el gobierno impondrá castigos financieros a la Universidad. El administrador tiene que hacer un balance entre el mayor número de estudiantes extranjeros sin incurrir en ningún castigo. Precisamente, su trabajo consiste en evaluar tres variables de decisión:

- El número de estudiantes nacionales.
- El número de estudiantes extranjeros.
- Las cuotas que debe cobrar a los estudiantes extranjeros.

De tal forma que maximice sus ingresos netos. Estas variables, de cualquier manera, no pueden variar independientemente pues deben satisfacer ciertas condiciones. Por ejemplo, el número total de estudiantes está sujeto a los recursos docentes de la Universidad y el número de estudiantes extranjeros que desea entrar depende de las cuotas que les cobren.

Para modelar matemáticamente este problema, supóngase que el número de estudiantes nacionales admitidos se denota por x_1, el número de estudiantes extranjeros por x_2, y la cuota que se cobrará a los estudiantes extranjeros por x_3. Supóngase también que el gobierno requiere 900 de un total de 1000 lugares disponibles para estudiantes nacionales y fija el castigo por $\$5(\ x_1-900)^2$, sí el número de estudiantes extranjeros sobrepasa estas cantidades, entonces el ingreso neto para la Universidad es:

$$f(x_1,x_2,x_3) = R - 5(x_1-900)^2 + x_2x_3 \tag{1.13}$$

donde R es la cantidad que aporta el gobierno para sus estudiantes nacionales. Se supone también que hay suficientes estudiantes que desean entrar a dicha Universidad para ocupar todos estos lugares, es decir:

$$x_1 + x_2 = 1000 \tag{1.14}$$

y la relación entre el número de estudiantes extranjeros y la cuota que se les cobrará se aproxima a la relación lineal dada por:

$$x_2 = 250 - \frac{x_3}{20}. \tag{1.15}$$

El problema es entonces maximizar 1.13, sujeto a 1.14 y 1.15. Entonces se tiene:

$$\begin{aligned} \text{máx}\ \ f(x_1,x_2,x_3) &= R - 5(x_1-900)^2 + x_2x_3 \\ \text{s.a.} \qquad\quad x_2 &= 250 - \frac{x_3}{20} \\ x_1 + x_2 &= 1000. \end{aligned}$$

Negociar los precios del trigo

Después de varias negociaciones el departamento de adquisiciones de una compañía procesadora de alimentos ha firmado contratos con tres ciudades productoras de trigo C_1, C_2 y C_3, para satisfacer la demanda de trigo por 12 meses. Hubo mucho regateo para establecer los precios de contrato, pero eventualmente se acordó que el precio pactado sería el precio local en el momento de entrega del trigo. Sí el precio de cada ciudad estuviera establecido, claramente la compañía hubiera firmado contrato únicamente con la ciudad del precio más bajo. Como, desde luego, sólo puede pronosticar el precio (p_1, p_2, p_3 por tonelada respectivamente) la compañía ha decidido poner todas sus apuestas a comprarle a las tres ciudades, con excepción de que uno o más de los pronósticos fuera realmente malo. Para calcular la proporción x_i que la compañía comprará a la ciudad C_i se ha decidido presupuestar el precio promedio $\widehat{p}$ para llevar a cabo su plan. Las proporciones deben satisfacer:

$$x_1p_1 + x_2p_2 + x_3p_3 = \widehat{p}$$

junto con la igualdad

$$x_1 + x_2 + x_3 = 1.$$

Pensando conservadoramente la compañía desea escoger x_1, x_2, x_3 que minimicen la incertidumbre de que el precio promedio $\widehat{p}$ se salga salvajemente del precio de mercado de los últimos doce meses.

Una medida obvia del riesgo es la varianza del precio pronosticado, que se puede expresar como:

$$\sum_{i=1}^{3} x_i^2 \sigma_i^2$$

donde σ_i^2 es la varianza del precio pronóstico p_i, si asumimos que las condiciones de clima y otros factores afectan la cosecha en las tres ciudades. Entonces el problema de la compañía se reduce a:

$$\begin{aligned} \text{mín}\, f(x_i) &= \sum_{i=1}^{3} x_i^2 \sigma_i^2 \\ \text{s.a.} \sum_{i=1}^{3} x_i p_i &= \widehat{p} \\ \sum_{i=1}^{3} x_i &= 1. \end{aligned}$$

Un caso particular sería considerar:

$$\begin{aligned} (p_1, p_2, p_3) &= (10, 15, 20) \\ (\sigma_1^2, \sigma_2^2, \sigma_3^2) &= (15, 10, 5) \end{aligned}$$

y entonces el problema a resolver es:

$$\begin{aligned} \text{mín}\, f(x_1, x_2, x_3) &= 15\ x_1^2 + 10x_2^2 + 5x_3^2 \\ \text{s.a.}\ 10x_1 + 15x_2 + 20x_3 &= \widehat{p} \\ x_1 + x_2 + x_3 &= 1. \end{aligned}$$

Un problema geométrico

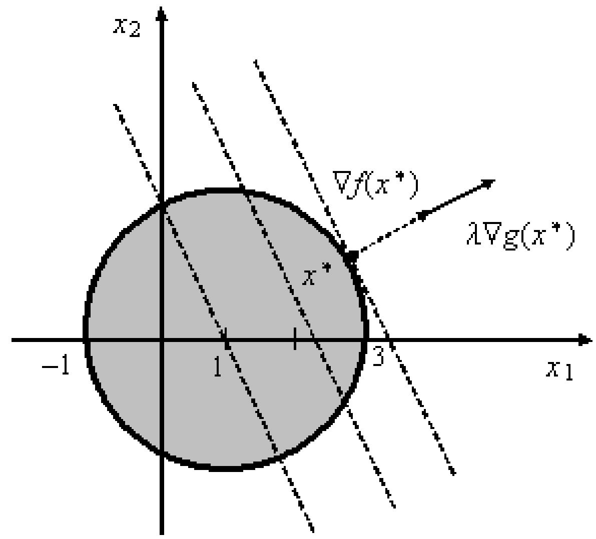
El siguiente ejemplo, aunque es sencillo, nos permitira ver cómo es la interpretación geométrica del teorema de Lagrange. La idea que subyace en dicho teorema es que el gradiente[2] de la función objetivo se puede expresar como combinación lineal de los multiplicadores de Lagrange y los gradientes de las restricciones del problema. Para este caso particular, se tiene una sola restricción por lo que las condiciones del teorema se reducen a que el gradiente de la función objetivo es igual a un escalar multiplicado por el gradiente de la restricción. Lo cual también demostraremos analíticamente una vez que se obtenga la solución del problema.

Ahora bien, se plantea el siguiente problema:

$$\begin{aligned} \text{máx}\, f(x_1, x_2) &= 2x_1 + x_2 \\ \text{sujeto a}\ \ (x_1 - 1)^2 + x_2^2 &= 4 \end{aligned}$$

[2]Si $f : U \subset R^n \to R$ es diferenciable, el gradiente de f en $(x_1, x_2, ..., x_n)$ es el vector en el espacio R^n dado por: $\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}\right)$.

Como ya se había mencionado el problema cuenta con una sola restricción la cual es una circunferencia con radio igual a 2 y centro en las coordenadas $(1, 0)$. Además se observa que las curvas de nivel de la función objetivo corresponden a líneas rectas con pendiente igual a -2 y surgen de despejar x_2 en la ecuación $2x_1 + x_2 = k$. Ahora analizaremos su gráfica, en donde el conjunto de soluciones factibles es el conjunto de puntos que están en el interior de la circunferencia y en su frontera además las curvas de nivel son las rectas punteadas.



La función de producción Cobb-Douglas

Paul Douglas y Charles Cobb[3] publicaron un estudio en el que modelaron el crecimiento de la economía estadounidense durante el periodo de 1899 y 1922, ellos consideraron una simplificación de la economía en la que la producción está determinada por la cantidad de mano de obra empleada y la cantidad de capital invertido. Si bien hay numerosos factores que afectan el comportamiento de la economía, el modelo resulto ser bastante preciso. La función que utilizaron para hacer este modelo de producción tenía la siguiente forma:

$$P(L, K) = bL^{\alpha}K^{1-\alpha} \tag{1.16}$$

donde:

P es la producción total de un año medida en valor monetario,

L es la cantidad de mano de obra medida en horas-hombre el mismo año,

K es la cantidad de capital invertido.[4]

Para entender cómo se obtiene el modelo planteado por Cobb y Douglas en la ecuación 1.16, emplearemos algunos supuestos sobre la economía que ellos hicieron.

Si la función se denota por $P(L, K)$, entonces la derivada parcial $\frac{\partial P}{\partial L}$ es la razón de cambio de la producción con respecto a la cantidad de mano de obra la cual es conocida como productividad

[3] Paul Douglas fue un economista del siglo XX, profesor de la Universidad de Chicago y más tarde senador de E.U. Charles Cobb fue matemático y profesor del Amherst College. La forma funcional Cobb-Douglas se utilizó inicialmente para estudiar la producción.

[4] El capital invertido representa el valor monetario de toda la maquinaria, equipos y edificios de las empresas.

marginal de la mano de obra. Del mismo modo la derivada parcial $\frac{\partial P}{\partial L}$ es la razón de cambio de la producción con respecto al capital, la cual se le conoce como productividad marginal del capital.

Cobb y Douglas hacen los siguientes supuestos:

i) Si desaparece la mano de obra o el capital, desaparece la producción, esto es si $L = 0$ o $K = 0$ entonces:

$$P(L, K) = 0$$

ii) La productividad marginal de la mano de obra es proporcional a la cantidad de producción por unidad de mano de obra, es decir:

$$\frac{\partial P}{\partial L} = \alpha \frac{P}{L} \tag{1.17}$$

para cierta constante α. Si suponemos K constante es decir $K = K_0$, entonces la ecuación diferencial parcial 1.17 se convierte en una ecuación diferencial ordinaria:

$$\frac{dP}{dL} = \alpha \frac{P}{L}. \tag{1.18}$$

Para resolver esta ecuación diferencial usamos el método de variables separables, por lo que de la ecuación 1.18 se tiene:

$$\frac{1}{p} dp = \frac{\alpha}{L} dL;$$

integrando de ambos lados se obtiene:

$$\int \frac{1}{p} dp = \int \frac{\alpha}{L} dL$$

y resolviendo

$$\ln(p) = \alpha \ln(L) + C_1.$$

Componiendo de ambos lados con la exponencial se tiene:

$$\begin{aligned} e^{\ln(p)} &= e^{\alpha \ln(L) + C_1} \\ p &= e^{\alpha \ln(L)} * e^{C_1} \\ p &= C_1 L^{\alpha}. \end{aligned}$$

Utilizando la hipótesis de que el capital es constante, podemos escribir la siguiente expresión:

$$P(L, K_0) = C_1 K_0 L^{\alpha}. \tag{1.19}$$

Análogamente, si se supone que también la productividd marginal del capital es proporcional a la cantidad de producción por unidad de capital, se tiene que:

$$\frac{dP}{dK} = \beta \frac{P}{K}.$$

Si suponemos ahora que L es constante, es decir $L = L_0$ y haciendo un razonamiento análogo, se tiene que:

$$P(L_0, K) = C_2 L_0 K^{\beta} \tag{1.20}$$

comparando las ecuaciones podemos deducir que la producción se puede expresar como una función en términos del capital y del trabajo como sigue:

$$P(L, K) = bL^{\alpha}K^{\beta} \tag{1.21}$$

donde b es una constante independiente de L y de K.

Observación 1 *Se sabe por el supuesto i) que $\alpha > 0$, también se observa en la ecuación 1.21 que si la mano de obra y el capital aumentan en un factor m se tiene:*

$$\begin{aligned} P(mL, mK) &= b\,(mL)^{\alpha}\,(mK)^{\beta} \\ &= m^{\alpha+\beta} bL^{\alpha}K^{\beta} \\ &= m^{\alpha+\beta} P(L, K); \end{aligned}$$

si $\alpha + \beta = 1$ entonces

$$P(mL, mK) = mP(L, K)$$

lo cual indica que la producción tiene rendimientos constantes a escala, por esta razón Cobb y Douglas supusieron que $\alpha + \beta = 1$ por lo que la ecuación 1.21 se puede escribier como:

$$P(L, K) = bL^{\alpha}K^{1-\alpha} \tag{1.22}$$

donde la ecuación anterior denota un modelo de producción simple.

Cabe señalar que Cobb y Douglas utilizaron datos económicos publicados por el gobierno de E.U. y construyeron una tabla con los datos L, K y P para los años de 1899 a1922 y con el método de mínimos cuadrados ajustaron los datos a la función:

$$P(L, K) = 1{,}01 L^{0{,}75} K^{0{,}25}$$

la cual resultó bastante precisa para el modelo de producción de E.U. en dichos años. Su gráfica está representada por la siguiente figura

Si suponemos que el costo de una unidad de mano de obra es n y el costo de una unidad de capital es m y el sector productivo dispone de w cantidad de dinero como presupuesto total entonces podemos plantear el problema de maximizar la producción P sujeta a la restricción presupuestal

$$nL + mK = w$$

es decir el problema queda como sigue:

$$\begin{aligned} \text{máx } P(L, K) &= bL^{\alpha}K^{1-\alpha} \\ \text{s.a. } nL + mK &= w \\ L, K &> 0 \\ m, n, \alpha, w &> 0 \end{aligned}$$

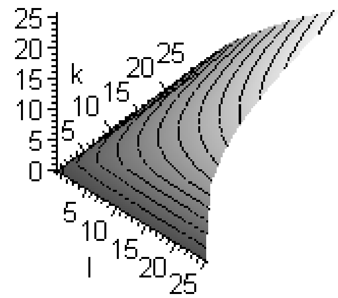


Figura 1-3

El problema anterior puede ser planteado por cualquier tipo de empresa.

Otra variante para el problema de la empresa, supone que la producción

$$bL^{\alpha}K^{1-\alpha} = Q$$

está fija, es decir Q es constante. Entonces el problema ahora consiste en encontrar los valores L y K que minimizan la función de costos

$$C(L, K) = mL + nK,$$

el cual queda planteado de la siguiente manera:

$$\begin{aligned} \text{mín}\, C(L,K) &= mL + nK \\ \text{s.a.}\ \ bL^{\alpha}K^{1-\alpha} &= Q \\ L, K &> 0 \\ m, n, \alpha, w &> 0. \end{aligned}$$

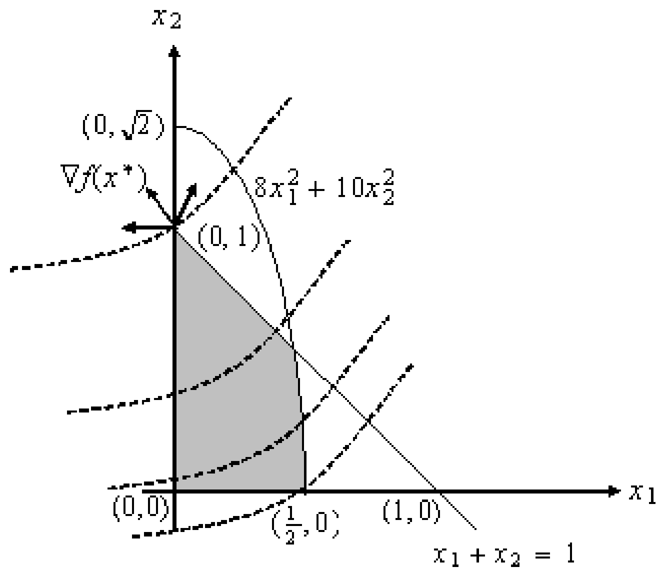
Ejemplos de optimización con restricciones de desigualdad

Un problema geométrico

Convencidos de que una idea geométrica siempre es una buena ruta para entender un problema, se incluye este ejemplo, en el que se puede ver como las condiciones de Kuhn-Tucker pueden caracterizar la solución a un problema de programación cóncava con restricciones de no negatividad. La idea geométrica que se encuentra detrás de las condiciones de no negatividad, es que, en una solución de contorno la dirrección del gradiente de la función objetivo debe ser una combinación lineal no negativa de las normales a los planos tangentes de las restricciones y de los multiplicadores de Lagrange en el punto que es solución del problema. Consideremos entonces el siguiente problema de programación no lineal:

$$\begin{aligned} \text{máx } f(x_1, x_2) &= -8x_1^2 - 10x_2^2 + 12x_1x_2 - 50x_1 + 80x_2 \\ \text{s.a. } x_1 + x_2 &\leqq 1 \\ 8x_1^2 + 10x_2^2 &\leqq 2 \\ x_1, x_2 &\geqq 0 \end{aligned}$$

cuya gráfica es la siguiente:



La parte sombreada corresponde al conjunto de soluciones factibles y las curvas punteadas corresponden a las curvas de nivel de la función objetivo.

Para resolver este problema de forma analítica es necesario llevarlo a la forma estándar planteada en el problema PPCNN en la página 111, por lo que el problema estándar se convierte en:

$$\begin{aligned} \text{máx } f(x_1, x_2) &= -8x_1^2 - 10x_2^2 + 12x_1x_2 - 50x_1 + 80x_2 \\ \text{s.a. } 1 - x_1 - x_2 &\geqq 0 \\ 2 - 8x_1^2 - 10x_2^2 &\geqq 0 \\ x_1, x_2 &\geqq 0 \end{aligned}$$

Maximizar de las ventas de una empresa

Este ejemplo tiene una sola variable de decisión, pues es conveniente resaltar que cuando existen más de dos variables de decisión la complejidad para resolver este tipo de problemas aumenta de manera considerable. Es por esto que, para la solución del ejemplo anterior hicimos uso de su gráfica.

En el análisis convencional, la empresa suele suponer que su objetivo es maximizar los beneficios, sin embargo cuando en la empresa la propiedad y la gestión están separadas, puede ser muy racional para la gestión perseguir el objetivo alternativo de maximizar las ventas es decir los ingresos, ya que éste resulta ser un indicador importante de la posición competitiva de la empresa dentro de un mercado. Y por otra parte también la gestión se preocupa de que el nivel de beneficios Π_0 no esté por debajo de un mínimo prefijado, esto asociado al hecho de que el ingreso marginal debe ser igual al costo marginal, es decir $IM = CM$.

Definamos entonces a: I

$Q > 0$ la cantidad producida por la empresa,

$I(Q)$ la función de ingresos (ventas),

$C(Q)$ la función de costos,

Π_0 el nivel mínimo de beneficios requeridos por la empresa.

Entonces $\Pi(Q) = I(Q) - C(Q)$ son los beneficios de la empresa definidos como la diferencia de los ingresos totales menos los costos totales.

La empresa se plantea el siguiente problema de optimización:

$$\begin{aligned} & \text{máx}\, I(Q) \\ \text{s.a. } \Pi(Q) &= I(Q) - C(Q) > \Pi_0 \\ Q &\geqq 0. \end{aligned}$$

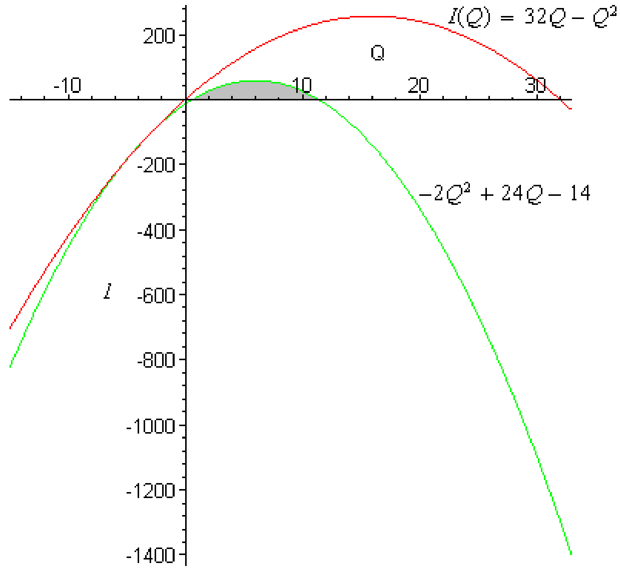
Para tener una idea más concreta, supóngase que:

$$\begin{aligned} I(Q) &= 32Q - Q^2 \\ C(Q) &= Q^2 + 8Q - 4 \\ \Pi_0 &= 18 \end{aligned}$$

entonces el problema se convierte en:

$$\begin{aligned} \text{máx}\, I(Q) &= 32Q - Q^2 \\ \text{s.a. } -2Q^2 + 24Q - 14 &\geqq 0 \\ Q &\geqq 0 \end{aligned}$$

que tiene la forma estándar del problema PPCNN definido en la página 111. Geométricamente, lo que se busca es un punto en la región sombreada (región factible) el cual maximice la función $I(Q) = 32Q - Q^2$.



Problema del consumidor

El siguiente ejemplo cuenta con cuatro variables de decisión, y en este caso la forma de resolverlo se vuelve más sencilla, ya que la función objetivo es estrictamente cóncava y se tiene una sola restricción lineal, lo que nos lleva a que la solución es única. Por lo tanto, la forma que toman las condiciones de Kuhn Tucker no es tan compleja, inclusive este problema se resuelve muy detalladamente para n bienes en el capítulo siete.

Considérese el problema de un consumidor, que desea obtener la máxima utilidad o satisfacción al consumir una canasta de consumo con cuatro tipos diferentes de bienes, en un período de tiempo determinado.

Ahora bien, se sabe que el consumidor tiene la siguiente función de utilidad:

$$U\left(x_1, x_2, x_3, x_4\right) = \sum_{i=1}^{4} a_i \log\left(x_i + b_i\right).$$

Donde a_i y b_i son constantes positivas definidas entre $0 < a_i < 1$ y $0 < b_i < 1$. Las variables de decisión son las x_i que se definen como las cantidades que el individuo consume de cada uno de los cuatro tipos de bienes. Además sean p_i los precios unitarios de cada uno de estos bienes, para toda $i = 1, 2, 3, 4$. Se sabe también que el consumidor dispone de una cantidad de dinero para gastar, que denotamos como M, la cual es una constante positiva.

Entonces el consumidor se enfrenta al problema de máximizar su función de utilidad, sujeto a la cantidad de dinero que tiene disponible para gastar.

La restricción del problema se expresa matemáticamente como sigue:

$$p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4 \leqq M$$

la cual es conocida como la restricción presupuestaria del consumidor.

Así pues el problema que se desea resolver, queda expresado matemáticamente de la siguiente forma:

$$
\begin{aligned}
\text{máx } U(x_1, x_2, x_{3,} x_4) &= \sum_{1=1}^{4} a_i \log\left(x_i + b_i\right) \\
\text{s.a. } p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_4 &\leqq M \\
\text{con } x_1, x_2, x_{3,} x_4 &\geqq 0
\end{aligned}
$$

Capítulo 2

Optimización sin restricciones

Introducción

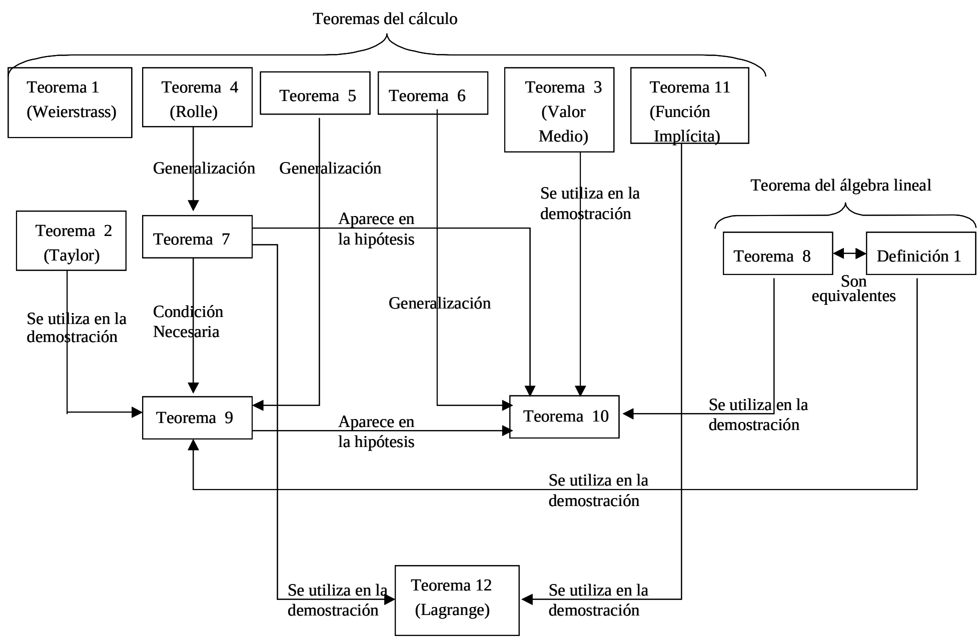
En los problemas de optimización sin restricciones las variables de decisión, en principio, pueden tomar cualquier valor. Esto podría sugerir que el estudio de este tipo de problemas carece de interés pues normalmente las variables han de cumplir con alguna restricción. Sin embargo, conviene señalar que no sólo razones pedagógicas y teóricas justifican su estudio pues, en ocasiones, es posible abordar la solución de un problema con restricciones mediante las técnicas de optimización sin restricciones.

Iniciamos este capítulo enunciando el teorema de Weierstrass, que nos dice cuándo una función alcanza un máximo o mínimo global, posteriormente se hace la demostración rigurosa del teorema de Taylor y se enuncia el conocido teorema del valor medio. Luego se estudian los teoremas que abordan las condiciones de primer orden y segundo orden las cuales caracterizan a los óptimos, para después concluir con las generalizaciones de dichos teoremas para funciones de varias variables, cuyas demostraciones estarán basadas en el teorema de Taylor.

Usaremos frecuentemente los resultados conocidos en relación a las propiedades de R^n como espacio vectorial, a las de las funciones lineales $f : R^n \to R$ y las propiedades de R^n como espacio topológico, es decir, nos referimos a las propiedades conocidas de los subconjuntos abiertos y cerrados de R^n, a las propiedades de las sucesiones de puntos de R^n, a las de los subconjuntos compactos de R^n, a las de las funciones continuas $f : \Omega \to R$ con $\Omega \subset R^n$, con las propiedades conocidas.

En el siguiente diagrama se muestra la relación entre los teoremas que aparecen en los capítulos dos y tres, hasta concluír con el teorema de Lagrange. Los teoremas de la raíz del diagrama corresponden a teoremas conocidos del cálculo que se enuncian sin demostración, así mismo se enuncia un teorema del álgebra lineal y como se puede ver todo el diagrama nos lleva al teorema principal de está sección que es el teorema de Lagrange.

Veamos un ejemplo que nos ayudará a entender el diagrama. El teroema 4 es el conocido teorema de Rolle y se encuentra en la raíz del diagrama, de ahí parte una flecha que nos lleva hacia el teorema /, dicho teorema es una generalización del teorema de Rolle. Así del teorema 7 salen tres flechas; una hacia el teorema 10 pues el teorema 7 aparece en la hipótesis del teroema 10; la otra flecha va hacia el teorema 9 la cual indica que es una condición necesaria para demostrar el teorema 9; y por último sale una flecha hacia el teorema 12 ya que el teorema 7 se utiliza en la demostración del teorema de Lagrange.



El teorema de Weierstrass y el teorema de Taylor

Antes de enunciar el teorema de Weierstrass, el cual consideramos que es fundamental para el presente trabajo en sus primeras secciones y en general para la teoría de la optimización matemática, enunciaremos algunas definiciones.Al mismo tiempo nos referiremos frecuentemente a las propiedades de las funciones $f : \Omega \to R$ continuamente diferenciables en donde $\Omega \subset R^n$ abierto y también a las funciones de clase $C^r(\Omega)$ con $r \geqq 1$ entero.

Definición 1 *Sea Ω un subconjunto de $R^n, (\Omega \subset R^n)$. Se dice que Ω es un conjunto acotado si existe un número real $M > 0$ tal que para todo $x \in \Omega$ se cumple que:*

$$\|x\| = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} \leq M$$

Definición 2 *Decimos que un subconjunto* Ω *de* R^n *es cerrado cuando su complemento* $R^n - \Omega$ *es abierto. Es decir, para todo* $x \notin \Omega$, *existe* $\varepsilon > 0$ *tal que* $B(x, \varepsilon) \subset R^n - \Omega$. *Donde* $B(x, \varepsilon)$ *es la vecindad de radio* $\varepsilon > 0$ *y centro en* x.

Definición 3 *Sea* Ω *un subconjunto de* R^n, *decimos que* Ω *es un conjunto compacto, si es un conjunto cerrado y acotado.*

Teorema 1 (De Weierstrass) *Sea* $f : \Omega \to R$ *continua con* $\Omega \subset R^n$ *compacto. Entonces existen* x^* *y* $x^{**} \in \Omega$ *tales que:*

$$\begin{aligned} f(x) &\leqq f(x^*)\ \forall x \in \Omega \\ f(x^{**}) &\leqq f(x)\ \forall x \in \Omega \end{aligned}$$

Dicho en otras palabras, f *alcanza un máximo global y un mínimo global en* Ω.

Para entender dicho teorema, veamos el siguiente problema de optimización y analicemos gráficamente sus hipótesis.

$$\begin{aligned} \textit{optimizar } f(x_1, x_2) &= 3x_1 + x_2 \\ \textit{sujeto a } x_1^2 + x_2^2 &\leq 5 \\ x_1 - x_2 &\leq 1 \\ x_1 &\geq 0 \end{aligned}$$

En este ejemplo se puede ver que la función $f(x_1, x_2) = 3x_1 + x_2$, claramente es una función continua, por ser una función polinomial. La parte sombreada en la figura 2-1 corresponde a la región factible que generan las restricciones del problema, la cual es un conjunto compacto. Por lo tanto el teorema de Weierstrass nos asegura la existencia de un máximo y mínimo globales.

Las curvas de nivel de la función objetivo $f(x_1, x_2) = 3x_1 + x_2$ constituyen una familia de rectas de ecuación

$$3x_1 + x_2 = k$$

es decir, como $x_2 = 3x_1 + k$, tenemos las rectas de pendiente -3 y ordenada al origen k. Al mover las curvas de nivel de la función objetivo en dirección del gradiente $\bigtriangledown f(x_1, x_2) = (3, 1)$, es decir, según aumenta el valor de k. Este desplazamiento lo podemos ver en las rectas punteadas de la figura 2-1.

Aunque el teorema no nos indica como encontrar los puntos óptimos, en este caso podemos ver graficamente que el máximo se encuentra en el primer cuadrante, en donde se intersectan la circunferencia $x_1^2 + x_2^2 \leq 5$ y la recta $x_1 - x_2 \leq 1$ y para encontrarlo analíticamente se resuelve el sistema

$$\begin{aligned} x_1^2 + x_2^2 &= 5 \\ x_1 - x_2 &= 1 \end{aligned}$$

de donde

$$x_2^2 + x_2 - 2 = 0$$

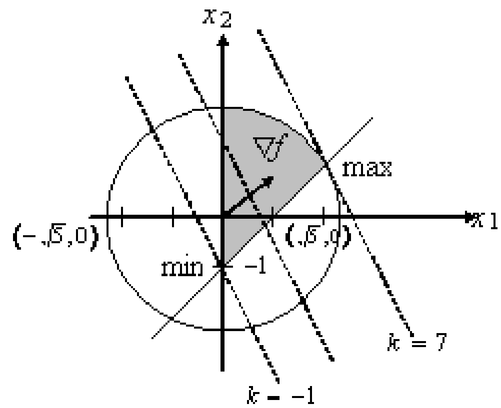


Figura 2-1

así pues, x_2 puede tomar los valores -2 y 1 y x_1 es igual a -1 y 2 de donde se generan los puntos $(2,1)$ y $(-1,-2)$. Como el punto $(-1,-2)$ no es factible, el máximo se alcanza en el punto $(2,1)$, siendo $f(2,1)=7$. El mínimo se alcanza en el punto más bajo de la parte sombreada que es el punto $(0,-1)$, y el valor mínimo de la función se encuentra en $f(0,-1)=-1$.

Habitualmente los conjuntos de soluciones factibles de los problemas que aparecen en optimización son cerrados y acotados. Y cuando la región factible no es acotada, ocasiona que no exista máximo o mínimo global o ninguno de los anteriores, al existir la posibilidad de desplazar las curvas de nivel de la función objetivo indefinidamente en las direcciones correspondientes sin abandonar el conjunto de soluciones factibles.

Es preciso señalar que el teorema de Weierstras nos proporciona las condiciones suficientes para saber si el problema alcanza su máximo y su mínimo, sin embargo es posible encontrar un ejemplo de un problema de optimización que no verifique las condiciones del teorema pero que posea máximo y mínimo globales, como lo es el siguiente:

$$\begin{array}{rcl} \textit{optimizar } f(x_1,x_2) & = & x_2 \\ \textit{sujeto a } \quad -x_1+x_2 & \leq & 1 \\ x_2 & \leq & 3 \\ x_1 & \geq & 0, x_2 \geq 0 \end{array}$$

Cuya solución gráfica se puede ver en la figura 2-2.

De lo anterior vemos que el mínimo global se alcanza en todos los puntos del conjunto

$$\left\{(x_1,x_2)\in R^2 \text{ tales que } x_2=0,\ x_1\geq 0\right\},$$

siendo el valor mínimo de la función objetivo igual a cero. El máximo global se alcanza en todos los puntos del conjunto

$$\left\{(x_1,x_2)\in R^2 \text{ tales que } x_2=3,\ x_1\geq 2\right\},$$

y el valor máximo de la función objetivo es tres. Observemos además que la función

$$f(x_1,x_2)=x_2$$

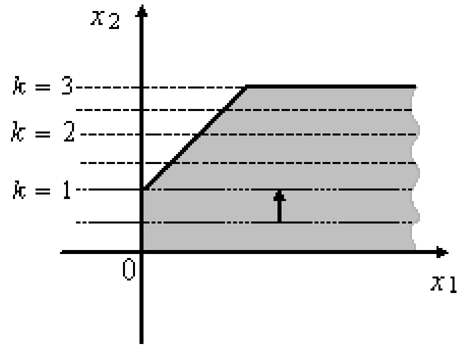


Figura 2-2

cumple con la hipótesis de ser continua en el espacio de R^2 y el conjunto de soluciones factibles

$$F = \left\{(x_1, x_2) \in R^2 \text{ tales que } -x_1 + x_2 \leq 1, x_2 \leq 3,\ x_1 \geq 0 \text{ y } x_2 \geq 0\right\}$$

es un conjunto cerrado, pero no esta acotado por lo que no cumple con las hipótesis del teorema.

Ahora demostraremos el teorema de Taylor para funciones $f : R \to R$ que formularemos del siguiente modo:

Teorema 2 (Taylor) *Sea $f : (a, b) \to R$ de clase $C^r(a, b)$ y sean $x \in (a, b)$ y $h \in R$ tales que $x + h \in (a, b)$, entonces se tiene que*

$$f(x+h) = f(x) + f^{(1)}(x)h + \frac{1}{2!}f^{(2)}(x)h^2 + ... + \frac{1}{r!}f^{(r)}(x)h^r + w(h)$$

en donde $\dfrac{w(h)}{h^r} \to 0$ cuando $h \to 0$.

Demostración. *Demostraremos nuestro teorema por inducción sobre r.*

Si $r = 1$ entonces se tiene

$$\begin{aligned} f(x+h) &= f(x) + f^{(1)}(x)h + w(h) \\ f(x+h) - f(x) &= f^{(1)}(x)h + w(h) \end{aligned}$$

dividiendo la expresión anterior entre h se tiene:

$$\frac{f(x+h) - f(x)}{h} = f^{(1)}(x) + \frac{w(h)}{h},$$

tomando el límite de ambos lados cuando $h \to 0$ se tiene:

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} f^{(1)}(x) + \lim_{h \to 0} \frac{w(h)}{h}$$

se puede ver que del lado izquierdo de la ecuación se tiene la derivada de f por definición y como $f'(x) = f^{(1)}(x)$ entonces:

$$f^{(1)}(x) = f^{(1)}(x) + \lim_{h \to 0} \frac{w(h)}{h}$$

de donde se deduce $\lim_{h\to 0} \dfrac{w(h)}{h} = 0$*, por lo tanto nuestro teorema es claro para* $r = 1$ *.*

Ahora si suponemos que es válido para r *y* $f : (a, b) \to R$ *es de clase* $C^{r+1}(a, b)$*, entonces* $f^{(1)}$ *es de clase* $C^r(a, b)$ *y por la hipótesis de inducción, tendremos que:*

$$f^{(1)}(x + h) = f^{(1)}(x) + f^{(2)}(x)h + \frac{1}{2!}f^{(3)}(x)h^2 + ... + \frac{1}{r!}f^{(r+1)}(x)h^r + \overline{w}(h)$$

en donde $\overline{w}(h)$ *está definida en* $(a - x, b - x)$ *y es continua, es decir,* $h \in (a - x, b - x)$ *para que estemos en el intervalo* (a, b)*, donde la expresión se puede escribir de manera compacta como:*

$$f^{(1)}(x + h) = \sum_{i=0}^{r} \frac{1}{i!} f^{(i+1)}(x)(h)^i + \overline{w}(h). \tag{2.1}$$

Definamos ahora la siguiente función $F : [0, 1] \to R$ *como*

$$F(t) = f(x + th)$$

obsérvese que por la regla de la cadena

$$F'(t) = f'(x + th)h$$

y por el teorema fundamental del cálculo se tiene lo siguiente:

$$\begin{aligned} f(x + h) - f(x) &= f(x + 1(h)) - f(x + 0(h)) \\ &= F(1) - F(0) \\ &= \int_0^1 f'(x + th)hdt \end{aligned} \tag{2.2}$$

por definición de la ecuación 2.1 y tomando en cuenta que

$$f'(x + th) = f^{(1)}(x + h)$$

se tiene lo siguiente:

$$\begin{aligned} \int_0^1 f^{(1)}(x + th)hdt &= \int_0^1 \left(\sum_{i=0}^{r} \frac{1}{i!} f^{(i+1)}(x)(th)^i + \overline{w}(th) \right) hdt \\ &= \int_0^1 \left(\sum_{i=0}^{r} \frac{1}{i!} f^{(i+1)}(x)t^i h^{i+1} \right) dt + \int_0^1 \overline{w}(th)hdt \\ &= \sum_{i=0}^{r} \frac{1}{i!} f^{(i+1)}(x)h^{i+1} \left. \frac{t^{i+1}}{i+1} \right|_0^1 + \int_0^1 \overline{w}(th)hdt \\ &= \sum_{i=0}^{r} \frac{1}{i!} f^{(i)}(x)h^i + \int_0^1 \overline{w}(th)hdt. \end{aligned}$$

Si hacemos correr el índice desde $i = 1$ hasta $r + 1$ en la expresión anterior tenemos que:

$$= \sum_{i=1}^{r+1} \frac{1}{i!} f^{(i)}(x) h^i + \int_0^1 \overline{w}(th) h dt. \tag{2.3}$$

Nuevamente de la ecuación 2.2 y de la ecuación 2.3 se tiene que

$$\begin{aligned} f(x+h) - f(x) &= \sum_{i=0}^{r+1} \frac{1}{i!} f^{(i)}(x) h^i + \int_0^1 \overline{w}(th) h dt \\ f(x+h) &= f(x) + \sum_{i=0}^{r+1} \frac{1}{i!} f^{(i)}(x) h^i + \int_0^1 \overline{w}(th) h dt \end{aligned}$$

por lo tanto

$$f(x+h) = \sum_{i=0}^{r+1} \frac{1}{i!} f^{(i)}(x) h^i + w(h)$$

en donde

$$w(h) = \int_0^1 \overline{w}(th) h dt = h \int_0^1 \overline{w}(th) dt.$$

Ahora demostraremos que

$$\lim_{h \to 0} \frac{w(h)}{h^{r+1}} = \lim_{h \to 0} \frac{h \int_0^1 \overline{w}(th) dt}{h^{r+1}} \to 0,$$

obsérvese que

$$\begin{aligned} \frac{w(h)}{h^{r+1}} &= \frac{h \int_0^1 \overline{w}(th) dt}{h^{r+1}} \\ &= \frac{\int_0^1 \overline{w}(th) dt}{h^r}. \end{aligned} \tag{2.4}$$

y si suponemos $h \neq 0$ *y definimos para* $t \in [0, 1]$ *a* $g(t) = \dfrac{\overline{w}(th)}{t^r h^r}$, *donde*

$$g(t) = \begin{cases} \dfrac{\overline{w}(th)}{t^r h^r} & si\ t \neq 0 \\ 0 & si\ t = 0 \end{cases},$$

ahora si suponemos que $t \neq 0$ *y* $g(0) = 0$, *entonces:*

$$g(t) = \frac{\overline{w}(th)}{t^r h^r} \iff t^r g(t) = \frac{\overline{w}(th)}{h^r}$$

de lo anterior y de la ecuación 2.4 podemos escribir la siguiente igualdad

$$\frac{w(h)}{h^{r+1}} = \int_0^1 \frac{\overline{w}(th)}{h^r} dt = \int_0^1 t^r g(t) dt.$$

Y si utilizamos nuevamente nuestra hipótesis de inducción tenemos que:

$$f^{(1)}(x+h) = f^{(1)}(x) + f^{(2)}(x)h + \frac{1}{2!} f^{(3)}(x)h^2 + ... + \frac{1}{r!} f^{(r+1)}(x)h^r + \overline{w}(h)$$

donde $\lim_{h \to 0} \dfrac{\overline{w}(h)}{h^{r+1}} = 0$ *y utilizando la definición de límite se tiene que para todo* $\varepsilon > 0$ *existe* $\delta > 0$ *tal que* $0 < |h - 0| < \delta$ *entonces* $\left| \dfrac{\overline{w}(h)}{h^r} - 0 \right| < \frac{\varepsilon}{2}$, *es decir para todo* $\varepsilon > 0$ *existe* $\delta > 0$ *tal que*

$$0 < |h| < \delta \Longrightarrow \left| \frac{\overline{w}(h)}{h^r} \right| < \frac{\varepsilon}{2}.$$

Supongamos que $0 < |h| < \delta$, *para* $t \in [0, 1]$, *y obsérvese que* $|g(t)| < \varepsilon$ *donde*

$$|g(t)| = \left| \frac{\overline{w}(h)}{t^r h^r} \right| < \varepsilon$$

y se tiene:

$$\begin{aligned} \left| \frac{w(h)}{h^{r+1}} \right| &= \left| \frac{h \int_0^1 \overline{w}(th) dt}{h^{r+1}} \right| = \left| \frac{\int_0^1 \overline{w}(th) dt}{h^r} \right| = \left| \int_0^1 t^r g(t) dt \right| \\ &\leqq \int_0^1 |t^r g(t) dt| \leqq \int_0^1 |t^r| \, |g(t)| \, dt \leqq \int_0^1 |g(t)| \, dt \leqq \int_0^1 \frac{\varepsilon}{2} dt = \frac{\varepsilon}{2} < \varepsilon \end{aligned}$$

Por lo tanto $\lim_{h \to 0} \dfrac{w(h)}{h^{r+1}} \to 0$ *y el teorema queda probado.* ■

Y por último, enunciaremos también el conocido teorema del valor medio del cálculo diferencial que usaremos posteriormente.

Teorema 3 (Valor medio) *Si* $f : [a, b] \to R$ *es continua y diferenciable en el intervalo abierto* (a, b), *entonces existe un punto* c *en* (a, b) *tal que*

$$f(b) - f(a) = f'(c)(b - a).$$

Teoremas relacionados con las condiciones de primer y segundo orden

Los siguientes teoremas que enunciamos son conocidos y aparecen en cualquier libro de cálculo diferencial:

Teorema 4 *Sea* $f:(a,b)\to R$ *de clase* $C^{(1)}(a,b)$ *y* $x^*\in(a,b)$ *un máximo local de* f, *es decir, existe una vecindad*

$$B(x^*;\varepsilon)=\{x\in R \text{ tal que } |x-x^*|<\varepsilon\}$$

con centro en x^**tal que* $f(x^*)\geqq f(x)$ *para toda* $x\in B(x^*;\varepsilon)\cap(a,b)$. *Entonces* $f'(x^*)=0$.

Demostración. *Supongamos que* $f'(x^*)\neq 0$ *entonces se desprenden dos casos* $f'(x^*)>0$ *o* $f'(x^*)<0$.

Si $f'(x^*)>0$, *entonces se tiene*

$$\lim_{x\to x^*}\frac{f(x)-f(x^*)}{x-x^*}>0,$$

y se sabe que $f(x)$ *es continua y de clase* $C^{(1)}(a,b)$. *Entonces existe* $x\in B(x^*;\varepsilon)$ *con* $x\neq x^*$ *tal que se cumple*

$$\frac{f(x)-f(x^*)}{x-x^*}>0,$$

lo que nos lleva a

$$f(x)-f(x^*)>0,$$

es decir se cumple

$$f(x)>f(x^*).$$

Lo cual es una contradicción, ya que x^* *es un máximo, análogamente llegamos a una contradicción si suponemos que* $f'(x^*)<0$, *por lo que se deduce que* $f'(x^*)=0$. ■

Teorema 5 *Sea* $f:(a,b)\to R$ *de clase* $C^{(2)}(a,b)$ *y* $x^*\in(a,b)$ *un máximo local de* f *. Entonces* $f''(x^*)\leqq 0$ *.*

Teorema 6 *Sea* $f:(a,b)\to R$ *de clase* $C^{(2)}(a,b)$ *y* $x^*\in(a,b)$ *tal que:*

i) $f'(x^*)=0$ *y*

ii) $f''(x^*)<0$. *Entonces* x^* *es un máximo local de* f *.*

Estos teoremas constituyen el fundamento de los métodos de optimización que datan de Leibnitz. Las afirmaciones de los teoremas 4 y 5 son conocidas como las condiciones necesarias de primer orden y de segundo orden respectivamente, en tanto que las hipótesis (i) y (ii) del teorema 6 son conocidas como las partes de primer orden y de segundo orden respectivamente de la condición suficiente, por tanto, están directamente vinculadas a nuestro tema.

Generalización de las condiciones de primer y segundo orden

Probaremos aquí las generalizaciones de estos teoremas para funciones de varias variables, aún cuando dichas generalizaciones aparecen en cualquier libro de cálculo.

Teorema 7 *Sea $f : \Omega \to R$ de clase $C^{(1)}(\Omega)$, con $\Omega \subset R^n$ un conjunto abierto y $x^* \in \Omega$ un máximo local de f, es decir, existe una vecindad*

$$B(x^*; \varepsilon) = \{x \in R^n \text{ tal que } |x - x^*| < \varepsilon\}$$

de x^* tal que $f(x^*) \geqq f(x)$ para toda $x \in B(x^*; \varepsilon) \cap \Omega$. Entonces

$$\frac{\partial f}{\partial x_i}(x^*) = 0 \qquad \forall i = 1, 2, ..., n$$

Demostración. *Considérese i fija, y sea e_i el $i-$ésimo vector de la base canónica de R^n, es decir que tiene en todos sus componentes cero y un uno en la $i-$ésima entrada dado por*

$$e_i = (0, 0, ..., 1, ..., 0),$$

y sea $h \in R$ con $h > 0$ suficientemente pequeña tal que

$$x + he_i \in B(x^*; \varepsilon) \cap \Omega,$$

entonces se tiene que

$$f(x^* + he_i) - f(x^*) \leqq 0$$

y dividiendo entre h se tiene que

$$\frac{f(x^* + he_i) - f(x^*)}{h} \leqq 0.$$

Por lo tanto cuando tomamos el límite y hacemos tender h a cero se tiene:

$$\frac{\partial f}{\partial x_i}(x^*) \leqq 0,$$

análogamente si tomamos ahora $h \in R$ con $h < 0$, suficientemente pequeña tal que

$$x^* + he_i \in B(x^*; \varepsilon) \cap \Omega$$

y con e_i como lo definimos anteriormente se tiene que:

$$f(x^* + he_i) - f(x^*) \leqq 0$$

dividiendo la expresión anterior entre $h < 0$

$$\frac{f(x^* + he_i) - f(x^*)}{h} \geqq 0,$$

donde si nuevamente tomamos el límite cuando h tiende a cero tenemos que:

$$\frac{\partial f}{\partial x_i}(x^*) \geqq 0$$

lo que implica que

$$\frac{\partial f}{\partial x_i}(x^*) = 0 \ .$$

Por lo que queda demostrado el teorema. ■

A continuación citaremos algunos conceptos del álgebra lineal, que utilizaremos posteriormente.

Definición 4 *Sea A una matriz cuadrada dada por $A = (a_{ij})$, decimos que es positiva o negativa definida si:*

i) Para toda $h = (h_1, h_2, ..., h_n) \in R^n$, se tiene lo siguiente:

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}h_ih_j &\geqq 0 \text{ es positiva definida}\\ \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}h_ih_j &\leqq 0 \text{ es negativa definida}\end{aligned}$$

ii)

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}h_ih_j = 0 \text{ si solo si } h = \overline{0}$$

En el caso en que $A = a_{ij}$ cumple únicamente con la condición (i), diremos que es semipositiva o seminegativa definida.

Ejemplo 1 *Sea* $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ *verificar que es positiva definida.*

Obsérvese que la matriz A se puede descomponer como el producto de las siguientes matrices:

$$\begin{aligned} A &= \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & \frac{4}{3} \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{2.5}$$

Sea $h \in R^3$ entonces se quiere demostrar

$$hAh^t = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}h_ih_j \geqq 0,$$

desarrollando y escribiendo A como en 2.5 se tiene:

$$
\begin{aligned}
&(h_1, h_2, h_3)\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & \frac{4}{3} \end{pmatrix}\begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} \\
= &\left(h_1 - \frac{1}{2}h_2, h_2 - \frac{2}{3}h_3, h_3\right)\begin{pmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & \frac{4}{3} \end{pmatrix}\begin{pmatrix} h_1 - \frac{1}{2}h_2 \\ h_2 - \frac{2}{3}h_3 \\ h_3 \end{pmatrix} \\
= &\left(2\left(h_1 - \frac{1}{2}h_2\right), \frac{3}{2}\left(h_2 - \frac{2}{3}h_3\right), \frac{4}{3}h_3\right)\begin{pmatrix} h_1 - \frac{1}{2}h_2 \\ h_2 - \frac{2}{3}h_3 \\ h_3 \end{pmatrix} \\
= &\ 2\left(h_1 - \frac{1}{2}h_2\right)^2 + \frac{3}{2}\left(h_2 - \frac{2}{3}h_3\right)^2 + \frac{4}{3}(h_3)^2 \geqq 0
\end{aligned}
$$

Lo anterior se cumple para todo $h \in R^3$ por lo tanto A es positiva definida.

En general a la expresión $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}h_ih_j$ se le conoce como forma cuadratica [1]. Ahora enunciaremos dos proposiciones importantes de esta teoría.

Proposición 1 *Sea una A matriz cuadrada entoces se cumple*

$$A = \left(\frac{A + A^t}{2}\right) + \left(\frac{A - A^t}{2}\right)$$

Demostración. *Demostraremos la proposición anterior para el caso de una matriz cuadrada de orden dos. Sean las siguientes matrices*

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad y \quad A^t = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$$

definamos ahora $\left(\frac{A+A^t}{2}\right) + \left(\frac{A-A^t}{2}\right)$ *entonces:*

$$\left(\frac{A + A^t}{2}\right) + \left(\frac{A - A^t}{2}\right) =$$

[1] Definimos una forma como una expresión polinomial la cual puede tener una o más variables, cada una con un exponente entero no negativo y además cumple con que cada término es de grado uniforme, es decir, la suma de los exponentes de cada término es la misma. Por ejemplo, en el polinomio $4x^2 - xy + 3y^2$ se observa que cada uno de sus términos es de segundo grado, es decir, la suma de los exponentes enteros es igual a 2 por lo que constituye una forma cuadrática de dos variables. También se pueden encontrar formas cuadráticas de trés variables, tales como $x^2 + 2xy - yw + 7w^2$ y en general de n variables.

$$\begin{aligned}
&= \frac{1}{2}\left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}\right] + \frac{1}{2}\left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}\right] \\
&= \begin{pmatrix} \frac{a_{11}+a_{11}}{2} & \frac{a_{12}+a_{21}}{2} \\ \frac{a_{21}+a_{12}}{2} & \frac{a_{22}+a_{22}}{2} \end{pmatrix} + \begin{pmatrix} \frac{a_{11}-a_{11}}{2} & \frac{a_{12}-a_{21}}{2} \\ \frac{a_{21}-a_{12}}{2} & \frac{a_{22}-a_{22}}{2} \end{pmatrix} \\
&= \begin{pmatrix} a_{11} & \frac{a_{12}+a_{21}}{2} \\ \frac{a_{21}+a_{12}}{2} & a_{22} \end{pmatrix} + \begin{pmatrix} 0 & \frac{a_{12}-a_{21}}{2} \\ \frac{a_{21}-a_{12}}{2} & 0 \end{pmatrix} \\
&= \begin{pmatrix} a_{11} & \frac{a_{12}+a_{21}+a_{12}-a_{21}}{2} \\ \frac{a_{21}+a_{12}+a_{21}-a_{12}}{2} & a_{22} \end{pmatrix} \\
&= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}
\end{aligned}$$

■

Proposición 2 *Sea A una matriz cuadrada y $h \in R^n$, entonces*

$$h\left(\left(\frac{A+A^t}{2}\right)+\left(\frac{A-A^t}{2}\right)\right)h^t$$

es una forma cuadratica.

Demostración. *De igual manera, demostraremos esta proposición para el caso de dos variables.*

Sean A y A^t como se definieron anteriormente y definamos a $h = (h_1, h_2)$ y $h^t = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$.

Por lo que definimos

$$h\left(\left(\frac{A+A^t}{2}\right)+\left(\frac{A-A^t}{2}\right)\right)h^t =$$

$$\begin{aligned}
&= h\left(\frac{A+A^t}{2}\right)h^t + h\left(\frac{A-A^t}{2}\right)h^t \\
&= (h_1,h_2)\begin{pmatrix} a_{11} & \frac{a_{12}+a_{21}}{2} \\ \frac{a_{21}+a_{12}}{2} & a_{22} \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + (h_1,h_2)\begin{pmatrix} 0 & \frac{a_{12}-a_{21}}{2} \\ \frac{a_{21}-a_{12}}{2} & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \\
&= \left(a_{11}h_1 + \left(\frac{a_{21}+a_{12}}{2}\right)h_2, \left(\frac{a_{12}+a_{21}}{2}\right)h_1 + a_{22}h_2\right)\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \\
&\quad + \left(\left(\frac{a_{21}-a_{12}}{2}\right)h_2, \left(\frac{a_{12}-a_{21}}{2}\right)h_1\right)\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \\
&= a_{11}h_1^2 + \left(\frac{a_{21}+a_{12}}{2}\right)h_1h_2 + \left(\frac{a_{12}+a_{21}}{2}\right)h_1h_2 + a_{22}h_2^2 + \left(\frac{a_{21}-a_{12}}{2}\right)h_1h_2 + \left(\frac{a_{12}-a_{21}}{2}\right)h_1h_2 \\
&= a_{11}h_1^2 + \left(\frac{a_{12}+a_{21}}{2}\right)h_1h_2 + \left(\frac{a_{12}+a_{21}}{2}\right)h_1h_2 + a_{22}h_2^2 - \left(\frac{a_{12}-a_{21}}{2}\right)h_1h_2 + \left(\frac{a_{12}-a_{21}}{2}\right)h_1h_2 \\
&= a_{11}h_1^2 + 2\left(\frac{a_{12}+a_{21}}{2}\right)h_1h_2 + a_{22}h_2^2 \\
&= a_{11}h_1^2 + (a_{12}+a_{21})\,h_1h_2 + a_{22}h_2^2
\end{aligned}$$

La cual es una forma cuadrática, además se puede observar que en el caso de que $A = A^t$, la forma cuadrática que resulta es

$$a_{11}h_1^2 + 2a_{12}h_1h_2 + a_{22}h_2^2$$

■

Todo lo anterior se puede generalizar a matrices de orden mayor que dos, es decir a cualquier matriz cuadrada se le puede asociar una forma cuadratica. Y utilizando este último resultado, podemos enunciar de otra forma la definición 4.

Proposición 3 *Sea A una matriz cuadrada dada por $A = (a_{ij})$, decimos que es positiva (negativa) definida si la forma cuadrática asociada a la matriz A es positiva (negativa).*

Una caracterización útil, para ver cuando una matriz es positiva o negativa definida, se da en términos de los menores principales de la matriz A. El menor principal de orden $k = 1, 2, ..., n$ denotado $M_k(A)$ de la matriz A está definido como

$$M_k(A) = \begin{vmatrix} a_{11} & a_{12} & .. & a_{1k} \\ a_{21} & a_{22} & .. & a_{2k} \\ .. & .. & .. & .. \\ a_{k1} & a_{k2} & .. & a_{kk} \end{vmatrix}$$

Ahora enunciaremos el siguiente teorema del álgebra lineal[2].

Teorema 8 *Sea $A = (a_{ij})$ una matriz cuadrada de orden $n \times n$ y tal que $A^t = A$ entonces*

(i) A es positiva definida si y sólo si

$$M_1(A) > 0, M_2(A) > 0, ..., M_n(A) > 0$$

(ii) A es negativa definida si y sólo si

$$M_1(A) < 0, M_2(A) > 0, ..., (-1)^n M_n(A) > 0$$

Para entender la relación que hay entre el teorema enterior, y la proposición 3, consideremos una matriz A que sea definida positiva (negativa). Y que además cumpla con las hipótesis del teorema 8, es decir que la matriz A sea igual a su transpuesta, nos fijaremos entonces, en el caso cuando la matriz es de dimensión dos, por lo que de la proposición 3 se tiene la forma cuadratica asociada a la matriz A,debe cumplir:

$$a_{11}h_1^2 + 2a_{12}h_1h_2 + a_{22}h_2^2 > 0.$$

De la ecuación anterior sumando y restando el término $\frac{a_{12}^2}{a_{11}}h_2^2$ del lado izquierdo de la desigualdad, se tiene:

$$a_{11}h_1^2 + 2a_{12}h_1h_2 + \frac{a_{12}^2}{a_{11}}h_2^2 + a_{22}h_2^2 - \frac{a_{12}^2}{a_{11}}h_2^2 > 0$$

[2]Una demostración se puede ver en la sección 6.2, en el teorema 6B del libro Álgebra lineal y sus aplicaciones, de Gilbert Strang.

lo cual se puede reescribir como

$$a_{11}\left(h_1^2+\frac{2a_{12}}{a_{11}}h_1h_2+\frac{a_{12}^2}{a_{11}^2}h_2^2\right)+\left(a_{22}-\frac{a_{12}^2}{a_{11}}\right)h_2^2>0$$

agrupando de manera conveniente, esta última expresión se obtiene

$$a_{11}\left(h_1+\frac{a_{12}}{a_{11}}h_2\right)^2+\left(\frac{a_{11}a_{22}-a_{12}^2}{a_{11}}\right)h_2^2>0.$$

Observemos que a_{11}es el menor principal de orden uno de la matriz A que se denota como $M_1\left(A\right)$ y que $a_{11}a_{22}-a_{12}^2$ es el menor principal de orden dos de la matriz, que se denoto como $M_2\left(A\right)$, por lo que condición de que la matriz A sea positiva (negativa) definida, la cual se ve reflejada en la última desigualdad, se puede expresar como:

$$M_1\left(A\right)\left(h_1+\frac{a_{12}}{a_{11}}h_2\right)^2+\frac{M_2\left(A\right)}{M_1\left(A\right)}h_2^2>0$$

Por lo tanto se deduce, que A es positiva definida si y sólo si $M_1\left(A\right)>0$ y $M_2\left(A\right)>0$, observemos también que para el caso en que A es negativa definida, se debe cumplir que $M_1\left(A\right)<0$ y $M_2\left(A\right)>0$.Lo cual en el caso más general, es decir para matrices cuadradas de dimensión mayor a dos, las condiciones que nos permiten ver cuando una matriz es positiva o negativa definida, quedan establecidas en el teorema 8.

Ejemplo 2 *Verificar que la matriz A del ejemplo 1 es positiva definida ahora utilizando el teorema 8.*

Claramente se ve que $A^t=\begin{pmatrix}2&-1&0\\-1&2&-1\\0&-1&2\end{pmatrix}=A$; ahora verificaremos que cada uno de los menores es positivo.

$$\begin{aligned}M_1(A)&=|2|>0,\\M_2(A)&=\begin{vmatrix}2&-1\\-1&2\end{vmatrix}=4-1=3>0\end{aligned}$$

$$\begin{aligned}M_3(A)&=\begin{vmatrix}2&-1&0\\-1&2&-1\\0&-1&2\end{vmatrix}\\&=2\begin{vmatrix}2&-1\\-1&2\end{vmatrix}+1\begin{vmatrix}-1&-1\\0&2\end{vmatrix}\\&=2(3)+1(-2)=4>0.\end{aligned}$$

Por lo tanto A es definida positiva.

Teorema 9 *Sea $f:\Omega\to R$ de clase $C^2(\Omega)$ con $\Omega\subset R^n$ un conjunto abierto y $x^*\in\Omega$ un máximo local de f. Entonces la matriz Hessiana*[3] *$H(x^*)$ es seminegativa definida, en donde $H(\ x^*)$ es la matriz cuyas entradas son:*

$$a_{ij}=\frac{\partial^2 f}{\partial x_j\partial x_i}(x^*).$$

Demostración. Sea $h\in R^n$ arbitraria fija. Por ser Ω un conjunto abierto, existe $\eta>0$ tal que $x^*+th\in\Omega$ para toda $t\in[0,\eta]$, definamos ahora $F:[0,\eta]\to R$ como $F(t)=f(x^*+th)$ entonces, por la regla de la cadena, se tiene:

$$\begin{aligned} F'(t) &= \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x^*+th)h_i \quad \text{y} \\ F''(t) &= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_j\partial x_i}(x^*+th)h_ih_j. \end{aligned}$$

Nótese que como x^* y h son vectores en R^n y $t\in[0,\eta]$, entonces x^*+th se puede escribir en forma desarrollada como sigue:

$$x^*+th=(x_1^*,x_2^*,...,x_n^*)+t(h_1,h_2,...,h_n)$$

por lo que

$$\begin{aligned} F'(t) &= \frac{\partial f}{\partial x_1}(x^*+th)h_1+\frac{\partial f}{\partial x_2}(x^*+th)h_2+...+\frac{\partial f}{\partial x_n}(x^*+th)h_n \\ &= \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x^*+th)h_i \end{aligned}$$

y

$$F''(t)=$$

$$\begin{array}{l} \frac{\partial^2 f}{\partial x_1^2}(x^*+th)h_1^2+\frac{\partial^2 f}{\partial x_1\partial x_2}(x^*+th)h_1h_2+...+\frac{\partial^2 f}{\partial x_1\partial x_n}(x^*+th)h_1h_n \\ \frac{\partial^2 f}{\partial x_2\partial x_1}(x^*+th)h_2h_1+\frac{\partial^2 f}{\partial x_2^2}(x^*+th)h_2^2+...+\frac{\partial^2 f}{\partial x_2\partial x_n}(x^*+th)h_2h_n \\ \vdots \hspace{30em} \vdots \\ \frac{\partial^2 f}{\partial x_n\partial x_1}(x^*+th)h_nh_1+\frac{\partial^2 f}{\partial x_n\partial x_2}(x^*+th)h_nh_2+...+\frac{\partial^2 f}{\partial x_n^2}(x^*+th)h_n^2 \end{array}$$

$$F''(t)=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_j\partial x_i}(x^*+th)h_ih_j.$$

[3] La matriz Hessiana es la matriz de segundas derivadas parciales, es decir, es una matriz cuadrada cuyas entradas son $a_{ij}=\frac{\partial^2 f}{\partial x_j\partial x_i}(x)$.

Por otro lado sea $h \in R^n$ arbitraria, dada por el vector $h = (h_1, h_2, ..., h_n)$ y sea

$$H(x^*) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x^*) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x^*) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x^*) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x^*) & \frac{\partial^2 f}{\partial x_2^2}(x^*) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x^*) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x^*) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x^*) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x^*) \end{pmatrix}$$

de donde

$$hH(x^*)h^t = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x^*) h_i h_j.$$

Supongamos que $t > 0$, por definición tenemos que:

$$F(t) = f(x^* + th)$$

y por lo tanto

$$F(0) = f(x^* + 0h) = f(x^*), \tag{2.6}$$

así mismo, se tiene que $0 + t \in [0, \eta]$ y además que $F : [0, \eta] \to R$ es de clase $C^2(0, \eta)$, entonces aplicando el teorema de Taylor 2 en cero, se cumple que:

$$\begin{aligned} F(0+t) &= F(0) + F^{(1)}(0)t + \frac{1}{2!}F^{(2)}(0)t^2 + w(t) \\ F(t) &= F(0) + F^{(1)}(0)t + \frac{1}{2!}F^{(2)}(0)t^2 + w(t) \end{aligned}$$

donde

$$F(t) - F(0) = F^{(1)}(0)t + \frac{1}{2!}F^{(2)}(0)t^2 + w(t) \tag{2.7}$$

y de 2.6 se tiene que

$$F(t) - F(0) = f(x^* + th) - f(x^*) \leqq 0 \tag{2.8}$$

ya que por hipótesis x^* es máximo local de f.

Ahora de 2.7 y de 2.8 se obtiene lo siguiente:

$$F^{(1)}(0)t + \frac{1}{2!}F^{(2)}(0)t^2 + w(t) = f(x^* + th) - f(x^*) \leqq 0 \tag{2.9}$$

y sabemos que $f(x^*) = F(0)$, entonces $f'(x^*) = F'(0)$ y en virtud del teorema 7 por ser x^* es máximo local de f se tiene:

$$f'(x^*) = 0 = F'(0),$$

entonces podemos escribir la ecuación 2.9 como sigue:

$$\frac{1}{2!}F^{(2)}(0)t^2 + w(t) = f(x^* + th) - f(x^*) \leqq 0$$

y dividiendo la expresión anterior entre t^2, obtenemos

$$\frac{1}{2!}F^{(2)}(0) + \frac{w(t)}{t^2} = \frac{f(x^* + th) - f(x^*)}{t^2} \leqq 0$$

de donde, haciendo tender t a cero, obtenemos que

$$\frac{1}{2!}F^{(2)}(0) \leqq 0$$

es decir $F^{(2)}(0) \leqq 0$ y como se definió $F(t)$, se tiene que:

$$\begin{aligned} F''(0) &= F^{(2)}(0) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x^* + 0h)h_i h_j \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x^*)h_i h_j \end{aligned}$$

lo cual implica que:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x^*)h_i h_j \leqq 0.$$

Obsérvese que lo anterior se cumple para $h \in R^n$ arbitraria, por lo que la matriz Hessiana es seminegativa definida. Por lo tanto queda demostrado nuestro teorema. ■

Estableceremos finalmente la generalización del teorema 6.

Teorema 10 *Sea $f : \Omega \to R$ de clase $C^2(\Omega)$ con $\Omega \subset R^n$ un conjunto abierto y $x^* \in \Omega$ tal que:*

(i) Para todo $i = 1, 2, ..., n$ se tiene $\frac{\partial f}{\partial x_i}(x^*) = 0$ y

(ii) La matriz hessiana $H(x^*)$ es negativa definida, entonces x^* es un máximo local de f.

Demostración. Para cada $k = 1, 2, ..., n$ se tiene que $M_k(H(x))$ es una función continua de $x \in \Omega$, por ser $f : \Omega \to R$ de clase $C^2(\Omega)$ y por definición sabemos que $H(x^*)$ es negativa definida, entonces

$$(-1)^k M_k(H(x^*)) > 0$$

por lo que existe $\delta_k > 0$ tal que $\|x - x^*\| < \delta_k$, de aquí, si definimos a $\delta = \text{mín}_k(\delta_k)$ y $\|x - x^*\| < \delta$, entonces la matriz $H(x)$ es negativa definida.

Probaremos ahora que si $\|x - x^*\| < \delta$, entonces $f(x) \leqq f(x^*)$. Para ello supongamos que $f(x) > f(x^*)$ y definamos como en el teorema 9 una función $F : [0, 1] \to R$, como:

$$F(t) = f(x^* + t(x - x^*)),$$

entonces, si hacemos $h = x - x^*$ se tiene:

$$F(t) = f(x^* + t(h)),$$

donde la primera y la segunda derivada de F se obtienen como en la demostración del teorema 9, es decir:

$$\begin{aligned} F'(t) &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^* + th)h_i \quad \text{con } h = (h_1, h_2, ..., h_n) \\ F''(t) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x^* + th)h_i h_j. \end{aligned}$$

Ahora bien, como

$$\begin{aligned} F(t) &= f(x^* + t(h)) \\ &= f(x^* + t(x - x^*)) \end{aligned}$$

de donde observamos que

$$\begin{aligned} F(1) &= f(x^* + x - x^*) = f(x) \\ F(0) &= f(x^* + 0(x - x^*)) = f(x^*) \end{aligned}$$

y usando el teorema del valor medio se tiene que:

$$\frac{F(1) - F(0)}{1 - 0} = F'(\eta_1) \text{ para alguna } \eta_1 \in (0, 1),$$

es decir,

$$0 < f(x) - f(x^*) = F'(\eta_1) \;\text{ con }\; \eta_1 \in (0, 1).$$

Pero por la hipotesis (i), se tiene $F'(0) = \dfrac{\partial f}{\partial x_i}(x^*) = 0$ para toda $i = 1, ..., n$ por lo que aplicando de nuevo el teorema del valor medio se tiene:

$$\frac{F'(\eta_1) - F'(0)}{\eta_1 - 0} = F''(\eta_2) \text{ para alguna } \eta_2 \in (0, \eta_1)$$

es decir,

$$0 < F'(\eta_1) - F'(0) = F''(\eta_2)\eta_1 \text{ con } \eta_2 \in (0, \eta_1) \text{ y } F''(\eta_2)\eta_1 > 0.$$

Entonces se tiene:

$$F''(\eta_2) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x^* + \eta_2 h)h_i h_j > 0,$$

y la matriz Hessiana $H(x^* + \eta_2 h)$ no es negativa definida. Esto es un absurdo, pues claramente $x^* + \eta_2 h$ dista de x^* menos que δ, de donde el teorema queda probado. ■

Soluciones a los problemas planteados en el capítulo 1

Maximizar los beneficios de una empresa

Este ejemplo se encuentra en la página 8 y plantea dos problemas de optimización, el primero es maximizar los ingresos de la empresa y el segundo es maximizar los beneficios de la misma. En esta ocasión sólo resolveremos el segundo caso debido a que el otro problema se resuelve de una manera similar, con la diferencia de que en este último la función objetivo es de mayor grado. Luego el problema quedó definido por:

$$\text{máx}\, B(Q) = -4 + 2Q + 0{,}05Q^2 - 0{,}05Q^3$$

entonces la condición de primer órden es:

$$B'(Q) = \frac{dB}{dQ} = 2 + 0{,}1Q - 0{,}15Q^2$$

que equivale a resolver la siguiente ecuación de segundo grado:

$$2 + 0{,}1Q - 0{,}15Q^2 = 0$$

y usando la fórmula general se tiene:

$$\begin{aligned} Q_1 &= \frac{-0{,}1 + 1{,}1}{-0{,}3} = -3.\overline{3} \\ Q_2 &= \frac{-0{,}1 - 1{,}1}{-0{,}3} = 4 \end{aligned}$$

Claramente la solución que nos sirve es el valor positivo $Q_2 = 4$ pues Q representa la cantidad del bien que se desea producir, por lo que no puede tomar valores negativos. Así pues:

$$Q^* = 4.$$

Ahora utilizando el teorema 5 que caracteriza la solución del problema para asegurar que $Q^* = 4$ es máximo se tiene que:

$$B''(Q) = \frac{d^2B}{dQ} = 0{,}1 - 0{,}3Q$$

y aplicando el teorema 5:

$$\begin{aligned} B''(Q^*) &= 0{,}1 - 0{,}3\,(4) \\ &= 1{,}1 < 0 \end{aligned}$$

Por lo tanto, podemos asegurar que $Q^* = 4$ es un máximo. Entonces $Q^* = 4$ es la cantidad del bien que es necesario producir para maximizar los beneficios.

Una política óptima de producción

Este ejemplo que se encuentra en la página 9 también se resuelve usando el teorema 5.

$$\text{máx } B(q) = (4\alpha + 3\beta - 1{,}2)q - (\alpha^2 + \beta^2)q^2$$

Por lo tanto:

$$B'(q) = 4\alpha + 3\beta - 1{,}2 - 2(\alpha^2 + \beta^2)q = 0$$

entonces despejando q, se tiene:

$$q^* = \frac{4\alpha + 3\beta - 1{,}2}{2(\alpha^2 + \beta^2)}$$

donde $0 < \alpha < 1$, $0 < \beta < 1$ y ahora para demostrar que q^* es máximo, calculamos la condición que establece el teorema 5:

$$B''(q^*) = -2(\alpha^2 + \beta^2)q < 0$$

por lo tanto q^* es la cantidad de mineral que se debe producir para que la empresa TAXSA maximice sus beneficios.

Un monopolio

Para resolver el problema planteado en la página 1 empezamos por obtener las condiciones de primer orden:

$$\begin{aligned} \frac{\partial \pi}{\partial x_1} &= 12 - 6x_1 - 2x_2 = 0 \\ \frac{\partial \pi}{\partial x_2} &= 32 - 10x_2 - 2x_1 = 0 \end{aligned}$$

de donde dichas condiciones forman el sistema de ecuaciones siguiente:

$$\begin{aligned} 6x_1 + 2x_2 &= 12 \\ 2x_1 - 10x_2 &= 32 \end{aligned}$$

y resolviéndolo se tiene que $x_1 = 1$ y $x_2 = 3$. Por lo que se concluye que el punto óptimo $x^* = (x_1^*, x_2^*) = (1, 3)$ y para demostrar que x^* es máximo, hacemos uso del teorema 10 antes enunciado.

La matriz Hessiana del problema está dada por:

$$H(x^*) = \begin{pmatrix} -6 & -2 \\ -2 & -10 \end{pmatrix}$$

de donde los menores principales están dados por:

$$\begin{aligned} M_1(A) &= -6 < 0 \\ M_2(A) &= \begin{vmatrix} -6 & -2 \\ -2 & -10 \end{vmatrix} = 60 - 4 = 56 > 0. \end{aligned}$$

Por lo que $H(x^*)$ es definida negativa y podemos asegurar que $x^* = (1, 3)$ es un máximo. Por consiguiente los precios óptimos se obtienen sustituyendo $\widehat{x}_1 = 1$ y $\widehat{x}_2 = 3$ en p_1 y p_2 respectivamente, es decir:

$$\begin{aligned} p_1^* &= 12 - 2 = 10 \\ p_2^* &= 32 - 12 = 20 \end{aligned}$$

Capítulo 3

Optimización con restricciones de igualdad

Introducción

En este capítulo abordaremos problemas de optimización con restricciones de igualdad. En ellos el conjunto factible esta formado por las soluciones del sistema de ecuaciones construído por las restricciones. Dado que el método de sustitución no siempre es aplicable para resolver dicho sistema de ecuaciones, es preciso disponer de criterios más generales mediante los cuales sea posible hallar, en caso de existir, la solución óptima. El teorema de Lagrange nos proporciona la herramienta necesaria para resolver este tipo de problemas, por lo que este capítulo está dedicado a su demostración rigurosa.

El problema estándar

El problema de Lagrange de programación no lineal consiste en maximizar f en el conjunto:

$$F = \{x \in \Omega \ \text{ tal que } g_j(x) = 0 \ \text{ para } j = 1, 2, ..., n\}$$

llamado conjunto factible del problema. A fin de simplificar llamaremos al "problema de programación no lineal" como PPL. El cual definimos a continuación:

Problema 1 (PPL) *Sea $\Omega \subset R^n$ un conjunto abierto y sean*

$$f, g_1, g_2, ..., g_m : \Omega \to R$$

de clase C^1 en Ω, con $m < n$

$$\begin{aligned} & \text{máx}\, f(x) \\ s.a\ g_j(x) &= 0 \\ con\ j &= 1, 2, ..., m. \end{aligned}$$

Que en forma desarrollada se puede escribir como:

$$\begin{aligned} \text{máx}\, f(x_1, ..., x_n) & \\ sujeto\ a\ \ g_1(x_1, ..., x_n) &= 0 \\ g_2(x_1, ..., x_n) &= 0 \\ &\vdots \\ g_m(x_1, ..., x_n) &= 0 \end{aligned}$$

El teorema de Lagrange

Establecemos ahora el resultado fundamental de esta sección que es el teorema de Lagrange, cuya demostración está basada en la siguiente afirmación:

Teorema 11 (Función implícita) *Sea $f : \Omega \to R$ con $\Omega \subset R^{n+1}$ una función continua de clase C^1en Ω, denotamos los puntos de Ω como (x, z), donde $x \in R^n$ y $z \in R$, y suponemos que $(x_0, z_0) \in \Omega$ satisface*

$$F(x_0, z_0) = 0$$

y

$$\frac{\partial f}{\partial z}(x_0, z_0) \neq 0.$$

Entonces existe una bola U que contiene a x_0 en R^n y una vecindad V de z_0 en R tal que existe una función única $z = g(x)$ definida para x en U y z en V que satisface

$$f(x, g(x)) = 0.$$

En el caso de tener m ecuaciones y $n + m$ variables es decir:

$$\begin{aligned} & f_1 (x_1, x_2, ..., x_n, z_1, z_2, ..., z_m) = 0 \\ & f_2 (x_1, x_2, ..., x_n, z_1, z_2, ..., z_m) = 0 \\ & \vdots \\ & f_m (x_1, x_2, ..., x_n, z_1, z_2, ..., z_m) = 0. \end{aligned} \tag{3.1}$$

el teorema se puede generalizar de la siguiente forma, la condición $\frac{\partial f}{\partial z}(x_0, z_0) \neq 0$ se convierte en la condición $\Delta \neq 0$, donde Δ es el determinante de la matriz evaluado en el punto $(x_0, z_0) \in R^{n+m}$, es decir:

$$\Delta(x_0, z_0) = \begin{vmatrix} \frac{\partial f_1}{\partial z_1} & \frac{\partial f_1}{\partial z_2} & \cdots & \frac{\partial f_1}{\partial z_m} \\ \frac{\partial f_2}{\partial z_1} & \frac{\partial f_2}{\partial z_2} & \cdots & \frac{\partial f_2}{\partial z_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_m}{\partial z_1} & \frac{\partial f_m}{\partial z_2} & \cdots & \frac{\partial f_m}{\partial z_m} \end{vmatrix} (x_0, z_0) \neq 0$$

en donde la ecuación 3.1 define de manera única funciones continuas y diferenciables

$$z_i = k_i\left(x_1, x_2, ..., x_n\right) \text{ con } i = 1, 2, ..., m$$

es decir

$$f_i\left(x_1, x_2, ..., x_n, k_1\left(x_1, x_2, ..., x_n\right), k_2\left(x_1, x_2, ..., x_n\right), ..., k_m\left(x_1, x_2, ..., x_n\right)\right)$$
$$\text{con } i = 1, 2, ..., m.$$

Continuamos ahora con el teorema de Lagrange y su demostración.

Teorema 12 (Lagrange) *Sea un problema del tipo PPL, sea* $x^* \in \Omega$ *solución del problema y supongamos que la matriz*

$$G_{mxn} = \left(\frac{\partial g_j}{\partial x_i}(x^*)\right) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

es de rango m*; es decir que tiene* m *renglones (columnas) que son linealmente independientes, entonces existe* $\lambda^* \in R^m$ *tal que*

$$\frac{\partial f}{\partial x_i}(x^*) = \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) \quad \textit{para} \quad i = 1, 2, ..., n;$$

que en forma desarrollada se ve como sigue

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x^*) &= \lambda_1^* \left(\frac{\partial g_1}{\partial x_1}, \frac{\partial g_1}{\partial x_2}, \cdots, \frac{\partial g_1}{\partial x_n}\right) + \lambda_2^* \left(\frac{\partial g_2}{\partial x_1}, \frac{\partial g_2}{\partial x_2}, \cdots, \frac{\partial g_2}{\partial x_n}\right) + \cdots \\ & ... + \lambda_m^* \left(\frac{\partial g_m}{\partial x_1}, \frac{\partial g_m}{\partial x_2}, \cdots \frac{\partial g_m}{\partial x_n}\right) \end{aligned}$$

y que a su vez se escribe como

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x^*) &= \lambda_1^* \frac{\partial g_1}{\partial x_1} + \lambda_2^* \frac{\partial g_2}{\partial x_1} + \cdots + \lambda_m^* \frac{\partial g_m}{\partial x_1} \\ &\vdots \\ \frac{\partial f}{\partial x_n}(x^*) &= \lambda_1^* \frac{\partial g_1}{\partial x_n} + \lambda_2^* \frac{\partial g_2}{\partial x_n} + \cdots + \lambda_m^* \frac{\partial g_m}{\partial x_n}. \end{aligned}$$

Donde los números $\lambda_1^*, \lambda_2^*, ..., \lambda_m^*$ *, componentes de* λ^* *, son llamados los multiplicadores de Lagrange del problema en la solución* x^**.*

Demostración. Puesto que G es de rango m, existe una submatriz cuadrada de orden $m \times m$ de G, cuyo determinante es no nulo. No se pierde generalidad si suponemos que dicha submatriz es la matriz

$$A_{m\times m} = \left(\frac{\partial g_j}{\partial x_i}(x^*)\right) \text{ con } i = 1, 2, .., m \text{ y } j = 1, 2, .., m.$$

Sea

$$B_{m\times(n-m)} = \left(\frac{\partial g_j}{\partial x_i}(x^*)\right) \text{con } i = m+1, m+2, .., n \text{ y } j = 1, 2, .., m$$

la submatriz de orden $m \times (n-m)$ que queda al eliminar las entradas de la matriz A de la matriz G

$$\begin{aligned} G_{m\times n} &= \begin{pmatrix} A_{m\times m} & B_{m\times(n-m)} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_m} & \frac{\partial g_1}{\partial x_{m+1}} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_m} & \frac{\partial g_2}{\partial x_{m+1}} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_m} & \frac{\partial g_m}{\partial x_{m+1}} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}. \end{aligned}$$

Denotemos a cada vector $x = (x_1, x_2, ..., x_n) \in R^n$ como $x = (v, w)$ en donde $v = (x_1, x_2, ..., x_m)$ y $w = (x_{m+1}, x_{m+2}, ..., x_n)$. Entonces sea $x^* = (v^*, w^*)$ solución del problema, por hipótesis

$$g_j(x^*) = g_j(v^*, w^*)$$

y

$$G_{mxn} = \left(\frac{\partial g_j}{\partial x_i}(x^*)\right).$$

Ahora, por el teorema de la función implícita, existe una vecindad $W \subset R^{n-m}$ de w^* y una función $\gamma : W \to R^m$ de clase C^1 tal que para cada $w \in W$ se tiene $(\gamma(w), w) \in \Omega$,

$$g_j\left(\gamma(w), w\right) = 0 \quad \text{para} \quad j = 1, 2, ..., m \tag{3.2}$$

y además tal que

$$\gamma(w^*) = v^*.$$

Definamos $F : W \to R$ como

$$F(w) = f(\gamma(w), w)$$

entonces w^* es un máximo de F en el abierto W y por el teorema 7 se tiene

$$\frac{\partial F}{\partial x_{m+r}}(w^*) = 0 \text{ para } r = 1, 2, ..., n-m.$$

Pero por la regla de la cadena, tenemos para $r = 1, 2, ..., n-m$ que

$$\frac{\partial F}{\partial x_{m+r}}(w^*) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^*)\frac{\partial \gamma_i}{\partial x_{m+r}}(w^*) + \sum_{s=1}^{n-m} \frac{\partial f}{\partial x_{m+r}}(x^*)\delta_{m+r}^{m+s},$$

donde δ es la delta de Kronecker, es decir

$$\frac{\partial F}{\partial x_{m+r}}(w^*) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^*) \frac{\partial \gamma_i}{\partial x_{m+r}}(w^*) + \frac{\partial f}{\partial x_{m+r}}(x^*). \tag{3.3}$$

Ahora bien, derivando la ecuación 3.2, con respecto a x_{m+r} se obtiene que

$$\sum_{k=1}^{m} \frac{\partial g_j}{\partial x_k}(x^*) \frac{\partial \gamma_k}{\partial x_{m+r}}(w^*) + \frac{\partial g_j}{\partial x_{m+r}}(x^*) = 0 \text{ para } \quad j = 1, 2, ..., m$$

lo cual se expresa en forma matricial como

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & .. & \frac{\partial g_1}{\partial x_m} \\ \vdots & .. & \vdots \\ \frac{\partial g_m}{\partial x_1} & .. & \frac{\partial g_m}{\partial x_m} \end{pmatrix} (x^*) \begin{pmatrix} \frac{\partial \gamma_1}{\partial x_{m+1}} & .. & \frac{\partial \gamma_1}{\partial x_n} \\ \vdots & .. & \vdots \\ \frac{\partial \gamma_m}{\partial x_{m+1}} & .. & \frac{\partial \gamma_m}{\partial x_n} \end{pmatrix} (w^*) + \begin{pmatrix} \frac{\partial g_1}{\partial x_{m+1}} & .. & \frac{\partial g_1}{\partial x_n} \\ \vdots & .. & \vdots \\ \frac{\partial g_m}{\partial x_{m+1}} & .. & \frac{\partial g_m}{\partial x_n} \end{pmatrix} (x^*) = \overline{0}$$

de lo anterior y de la definición de la matriz A y B, se tiene

$$A \begin{pmatrix} \frac{\partial \gamma_1}{\partial x_{m+1}} & \cdots & \frac{\partial \gamma_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \gamma_m}{\partial x_{m+1}} & \cdots & \frac{\partial \gamma_m}{\partial x_n} \end{pmatrix} (w^*) = -B;$$

por otro lado, sabemos que A es no singular, es decir existe A^{-1}.

Multiplicando por la izquierda la expresión anterior obtenemos

$$\begin{pmatrix} \frac{\partial \gamma_1}{\partial x_{m+1}} & \cdots & \frac{\partial \gamma_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \gamma_m}{\partial x_{m+1}} & \cdots & \frac{\partial \gamma_m}{\partial x_n} \end{pmatrix} (w^*) = -A^{-1}B. \tag{3.4}$$

Al mismo tiempo, expresando también la ecuación 3.3 en forma matricial tenemos

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, .., \frac{\partial f}{\partial x_m}\right)(x^*) \begin{pmatrix} \frac{\partial \gamma_1}{\partial x_{m+1}} & .. & \frac{\partial \gamma_1}{\partial x_n} \\ \vdots & .. & \vdots \\ \frac{\partial \gamma_m}{\partial x_{m+1}} & .. & \frac{\partial \gamma_m}{\partial x_n} \end{pmatrix} (w^*) + \left(\frac{\partial f}{\partial x_{m+1}}, ..., \frac{\partial f}{\partial x_n}\right)(x^*) = \overline{0}$$

y sustituyendo la ecuación 3.4 con la expresión anterior se obtiene

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_m}\right)(x^*)\left(-A^{-1}B\right) + \left(\frac{\partial f}{\partial x_{m+1}}, \ldots, \frac{\partial f}{\partial x_n}\right)(x^*) = \overline{0}, \tag{3.5}$$

pero, por otro lado se cumple la siguiente igualdad

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_m}\right)(x^*)\left(-A^{-1}A\right) + \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_m}\right)(x^*) = \overline{0} \tag{3.6}$$

pues $-A^{-1}A = -I_{mxm}$, así que, utilizando las ecuaciones 3.5 y 3.6, obtenemos

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, .., \frac{\partial f}{\partial x_m}\right)(x^*)(-A^{-1})(AB) + \left(\frac{\partial f}{\partial x_1}, .., \frac{\partial f}{\partial x_m}, .., \frac{\partial f}{\partial x_n}\right)(x^*) = \overline{0}$$

donde $G = (AB)$ y $\left(\frac{\partial f}{\partial x_1}, .., \frac{\partial f}{\partial x_m}, .., \frac{\partial f}{\partial x_n}\right)(x^*)$ es el gradiente de $f(\gamma(w), w)$

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, .., \frac{\partial f}{\partial x_m}\right)(x^*)(-A^{-1})G + \left(\frac{\partial f}{\partial x_1}, .., \frac{\partial f}{\partial x_m}, .., \frac{\partial f}{\partial x_n}\right)(x^*) = \overline{0}.$$

Ahora si definimos

$$(\widehat{y}_1, \widehat{y}_2, ..., \widehat{y}_m) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, .., \frac{\partial f}{\partial x_m}\right)(x^*)(+A^{-1})$$

se tiene que

$$(\widehat{y}_1, \widehat{y}_2, ..., \widehat{y}_m)\, G = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, .., \frac{\partial f}{\partial x_n}\right)(x^*)$$

es decir

$$\frac{\partial f}{\partial x_j}(x^*) = \sum_{i=1}^{m} \widehat{y}_i \frac{\partial g_i}{\partial x_j}(x^*) \quad \text{para} \quad j = 1, 2, ..., n.$$

Por lo que el teorema queda probado. ■

Soluciones a los problemas planteados en el capítulo 1

Programa de admisión universitario.

El problema está planteado en la página 13 y tiene la siguiente forma:

$$\begin{aligned} \text{máx } f(x_1, x_2, x_3) &= R - 5(x_1 - 900)^2 + x_2 x_3 \\ \text{s.a.} \qquad x_2 &= 250 - \frac{x_3}{20} \\ x_1 + x_2 &= 1000. \end{aligned}$$

de modo que

$$\begin{aligned} f(x_1, x_2, x_3) &= R - 5(x_1 - 900)^2 + x_2 x_3 \\ g_1(x_1, x_2, x_3) &= x_2 + \frac{x_3}{20} - 250 \\ g_2(x_1, x_2, x_3) &= x_1 + x_2 - 1000. \end{aligned}$$

Ahora bien, aplicando el teorema de Lagrange con $n = 3$ y $m = 2$ se tiene que existe

$$\lambda^* = (\lambda_1^*, \lambda_2^*) \in R^2$$

tal que:

$$\frac{\partial f}{\partial x_i}(x^*) = \sum_{j=1}^{2} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) \text{ para } i = 1, 2, 3.$$

En nuestro caso particular la forma desarrollada se ve como sigue:

$$\left(-10\left(x_1-900\right), x_3, x_2\right)=\lambda_1\left(0,1,\frac{1}{20}\right)+\lambda_2(1,1,0),$$

que a su vez nos lleva a las siguientes ecuaciones:

$$\begin{aligned}
-10\left(x_1-900\right) &= \lambda_2 \\
x_3 &= \lambda_1+\lambda_2 \\
x_2 &= \frac{\lambda_1}{20}.
\end{aligned}$$

Así pues, se genera el sistema de ecuaciones que a continuación se enuncia:

$$-10\left(x_1-900\right) = \lambda_2 \tag{3.7}$$

$$x_3 = \lambda_1+\lambda_2 \tag{3.8}$$

$$x_2 = \frac{\lambda_1}{20} \tag{3.9}$$

$$x_2+\frac{x_3}{20} = 250 \tag{3.10}$$

$$x_1+x_2 = 1000. \tag{3.11}$$

Ahora bien, de las ecuaciones 3.7, 3.8 y 3.9 se tiene:

$$x_1 = 900-\frac{\lambda_2}{10} \tag{3.12}$$

$$x_2 = \frac{\lambda_1}{20} \tag{3.13}$$

$$x_3 = \lambda_1+\lambda_2 \tag{3.14}$$

las cuales se sustituyen en las ecuaciones 3.10 y 3.11 generando así, el sistema de ecuaciones en términos solamente de λ_1 y λ_2, esto es:

$$\begin{aligned}
\frac{\lambda_1}{20}+\frac{\lambda_1+\lambda_2}{20} &= 250 \\
900-\frac{\lambda_2}{10}+\frac{\lambda_1}{20} &= 1000;
\end{aligned}$$

simplificando

$$\begin{aligned}
2\lambda_1+\lambda_2 &= 5000 \\
\lambda_1-2\lambda_2 &= 2000
\end{aligned}$$

cuyas soluciones son $\lambda_1=2400$ y $\lambda_2=200$. Sustituyendo estos valores en las ecuaciones 3.12, 3.13 y 3.14, la solución a nuestro problema esta dada por:

$$\begin{aligned}
x_1^* &= 880 \\
x_2^* &= 120 \\
x_3^* &= 2600;
\end{aligned}$$

es decir, la política de la universidad debe ser aceptar a 880 estudiantes nacionales, 120 extranjeros y con una cuota de 2600 unidades monetarias.

Negociando los precios del trigo

Una vez establecido el teorema de Lagrange podemos resolver el problema 1 planteado en el capítulo 1 en la página 14.

$$\begin{aligned} \text{mín}\, f(x_1, x_2, x_3) &= 15\ x_1^2 + 10x_2^2 + 5x_3^2 \\ \text{s.a. } 10x_1 + 15x_2 + 20x_3 &= \widehat{p} \\ x_1 + x_2 + x_3 &= 1. \end{aligned}$$

Obsérvese que el problema es minimizar una función objetivo sujeta a dos restricciones de igualdad. Para poder aplicar el teorema de Lagrange en la versión del teorema 12, debemos plantear el problema como un problema de maximización es decir, como el problema estándar 1 de la página 45 lo cual equivale a escribirlo de la siguiente forma:

$$\begin{aligned} \text{máx}\, f(x_1, x_2, x_3) &= -\ \left(15\ x_1^2 + 10x_2^2 + 5x_3^2\right) \\ \text{s.a. } 10x_1 + 15x_2 + 20x_3 &= \widehat{p} \\ x_1 + x_2 + x_3 &= 1. \end{aligned}$$

donde

$$\begin{aligned} f(x_1, x_2, x_3) &= -15\ x_1^2 - 10x_2^2 - 5x_3^2 \\ g_1(x_1, x_2, x_3) &= 10x_1 + 15x_2 + 20x_3 - \widehat{p} \\ g_2(x_1, x_2, x_3) &= x_1 + x_2 + x_3 - 1. \end{aligned}$$

Como se tienen dos restricciones j corre desde uno hasta dos y se tienen tres variables. Por lo tanto $i = 1, 2, 3$; por el teorema de Lagrange se tiene la siguiente igualdad:

$$\frac{\partial f}{\partial x_i}(x^*) = \sum_{j=1}^{2} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) \text{ para } i = 1, 2, 3$$

que en forma desarrollada puede escribirse como sigue:

$$(-30x_1, -20x_2, -10x_3) = \lambda_1\,(10, 15, 20) + \lambda_2\,(1, 1, 1)$$

que a su vez genera las siguientes ecuaciones:

$$\begin{aligned} -30x_1 &= 10\lambda_1 + \lambda_2 \\ -20x_2 &= 15\lambda_1 + \lambda_2 \\ -10x_3 &= 20\lambda_1 + \lambda_2 \end{aligned}$$

y aunadas a las restricciones mismas del problema nos plantean el siguiente sistema de ecuaciones:

$$\begin{aligned} -30x_1 &= 10\lambda_1 + \lambda_2 && (3.15) \\ -20x_2 &= 15\lambda_1 + \lambda_2 && (3.16) \\ -20x_3 &= 20\lambda_1 + \lambda_2 && (3.17) \\ 10x_1 + 15x_2 + 20x_3 &= \widehat{p} && (3.18) \\ x_1 + x_2 + x_3 &= 1. && (3.19) \end{aligned}$$

Para resolverlo, primero despejamos λ_2 de 3.15, 3.16 y 3.17, de donde se tiene:

$$\begin{aligned} \lambda_2 &= -30x_1 - 10\lambda_1 && (3.20)\\ \lambda_2 &= -20x_2 - 15\lambda_1 && (3.21)\\ \lambda_2 &= -20x_3 - 20\lambda_1. && (3.22) \end{aligned}$$

Igualando ahora 3.20 con 3.21 y a su vez despejando λ_1 se tiene:

$$\begin{aligned} -30x_1 - 10\lambda_1 &= -20x_2 - 15\lambda_1\\ \lambda_1 &= 6x_1 - 4x_2. && (3.23) \end{aligned}$$

Análogamente con las ecuaciones 3.21 y 3.22 de donde se tiene:

$$\begin{aligned} -20x_2 - 15\lambda_1 &= -10x_3 - 20\lambda_1\\ \lambda_1 &= 4x_2 - 2x_3. && (3.24) \end{aligned}$$

Igualando las ecuaciones obtenidas 3.23 y 3.24 se tiene:

$$\begin{aligned} 6x_1 - 4x_2 &= 4x_2 - 2x_3\\ x_2 &= \frac{3}{4}x_1 + \frac{1}{4}x_3 && (3.25) \end{aligned}$$

por otro lado, sustituyendo la ecuación 3.25 en 3.19 se tiene:

$$\begin{aligned} x_1 + \left(\frac{3}{4}x_1 + \frac{1}{4}x_3\right) + x_3 &= 1\\ \frac{7}{4}x_1 + \frac{5}{4}x_3 &= 1 \end{aligned}$$

de donde:

$$x_1 = \frac{4}{7} - \frac{5}{7}x_3 \qquad (3.26)$$

y para escribir a x_2 en términos de x_3 se sustituye 3.26 en 3.25 por tanto:

$$\begin{aligned} x_2 &= \frac{3}{4}\left(\frac{4}{7} - \frac{5}{7}x_3\right) + \frac{1}{4}x_3\\ x_2 &= \frac{4}{7} - \frac{5}{7}x_3. && (3.27) \end{aligned}$$

Así, sustituyendo a 3.26 y a 3.27 en la ecuación 3.18, se tiene todo en términos de x_3:

$$10\left(\frac{4}{7} - \frac{5}{7}x_3\right) + 15\left(\frac{4}{7} - \frac{5}{7}x_3\right) + 20x_3 = \widehat{p}.$$

Desarrollando paréntesis y despejando x_3 se tiene que:

$$x_3 = \frac{7}{60}\widehat{p} - \frac{17}{12};$$

por lo tanto:

$$x_3^* = \frac{7}{60}\widehat{p} - \frac{17}{12}$$

donde su valor aproximado en decimales es:

$$x_3^* \approx 0{,}12\widehat{p} - 1{,}42.$$

Para obtener el valor de x_1^* y x_2^* se sustituye x_3^* en las ecuaciones 3.26 y 3.27 respectivamente, obteniendo así que:

$$\begin{aligned} x_1^* &= \frac{4}{7} - \frac{5}{7}\left(\frac{7}{60}\widehat{p} - \frac{17}{12}\right) \\ &= -\frac{1}{12}\widehat{p} + \frac{19}{12} \end{aligned}$$

$$x_1^* \approx -0{,}08\widehat{p} + 1{,}58$$

y para:

$$\begin{aligned} x_2^* &= \frac{3}{7} - \frac{2}{7}\left(\frac{7}{60}\widehat{p} - \frac{17}{12}\right) \\ &= -\frac{1}{30}\widehat{p} + \frac{5}{6} \end{aligned}$$

$$x_2^* \approx -0{,}03\widehat{p} + 0{,}83$$

Por último se calculan los valores de los multiplicadores de Lagrange λ_1^* y λ_2^*. Para λ_1^* utilizaremos la ecuación 3.24 sustituyendo a x_2^* y x_3^* se obtiene:

$$\begin{aligned} \lambda_1^* &= 4\left(-\frac{1}{30}\widehat{p} + \frac{5}{6}\right) - 2\left(\frac{7}{60}\widehat{p} - \frac{17}{12}\right) \\ &= -\frac{11}{30}\widehat{p} + \frac{37}{6} \end{aligned}$$

$$\lambda_1^* \approx -0{,}37\widehat{p} + 6{,}17.$$

Ahora para λ_2^* se sustituye a x_2^* y λ_1^* en la ecuación 3.21 y se obtiene:

$$\begin{aligned} \lambda_2^* &= -20\left(-\frac{1}{30}\widehat{p} + \frac{5}{6}\right) - 15\left(-\frac{11}{30}\widehat{p} + \frac{37}{6}\right) \\ &= \frac{37}{6}\widehat{p} - \frac{655}{6} \end{aligned}$$

$$\lambda_2^* \approx 6{,}17\widehat{p} - 109{,}17$$

Recordemos que el valor de los multiplicadores de Lagrange están afectados por un signo negativo, ya que transformamos el problema original de minimización en un problema de maximización, por lo que las soluciones a nuestro problema original estan dadas por:

$$\begin{aligned} x_1^* &\approx -0{,}08\widehat{p}+1{,}58 \\ x_2^* &\approx -0{,}03\widehat{p}+0{,}83 \\ x_3^* &\approx 0{,}12\widehat{p}-1{,}42 \\ \lambda_1^* &\approx -0{,}37\widehat{p}+6{,}17 \\ \lambda_2^* &\approx -6{,}17\widehat{p}+109{,}17 \end{aligned}$$

donde $\widehat{p}$ es el precio presupuestado por la compañía.

Un problema geométrico

Para la solución analítica del problema planteado en la página 15, obsérvese que las derivadas parciales de f y g estan dadas por:

$$\begin{aligned} \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right)(x) &= (2,1) \\ \left(\frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}\right)(x) &= (2(x_1-1), 2x_2) \end{aligned}$$

por lo tanto las condiciones de primer orden son las siguientes:

$$\lambda 2(x_1-1) = 2 \tag{3.28}$$
$$\lambda 2x_2 = 1 \tag{3.29}$$
$$(x_1-1)^2+x_2^2 = 4. \tag{3.30}$$

Despejando λ de 3.28 y 3.29 se tiene que:

$$\lambda = \frac{2}{2(x_1-1)} \tag{3.31}$$
$$\lambda = \frac{1}{2x_2} \tag{3.32}$$

igualando ambas y despejando x_2, se tiene que:

$$x_2 = \frac{x_1-1}{2} \tag{3.33}$$

sustituyendo el valor de x_2 en 3.30, se tiene que:

$$\begin{aligned} (x_1-1)^2+\left(\frac{x_1-1}{2}\right)^2 &= 4 \\ 4(x_1-1)^2+(x_1-1)^2 &= 16 \\ 5\,(x_1-1)^2 &= 16 \\ (x_1-1)^2 &= \frac{16}{5}. \end{aligned}$$

lo que nos lleva a una ecuación de segundo grado en x_1, donde las soluciones son:

$$x_1 = 1 \pm \frac{4}{\sqrt{5}}.$$

Ahora para encontrar el valor de x_2, sustituimos los valores de x_1 en 3.33, por lo tanto:

$$x_2 = \frac{1 \pm \frac{4}{\sqrt{5}} - 1}{2} = \pm\frac{2}{\sqrt{5}}$$

De donde se tiene que las soluciones del problema son:

$$x^* = \left(1 + \frac{4}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$$

y

$$x^{**} = \left(1 - \frac{4}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right).$$

Y por último para decidir cuál de los dos puntos críticos es máximo, se verifica cuál de las siguientes desigualdades se cumple ya sea

$$f(x^*) > f(x^{**}) \text{ o } f(x^*) < f(x^{**})$$

de acuerdo con el teorema 7.

$$\begin{aligned} f(x^*) &= 2\left(1 + \frac{4}{\sqrt{5}}\right) + \frac{2}{\sqrt{5}} \\ &= 2 + \frac{10}{\sqrt{5}} \end{aligned}$$

$$\begin{aligned} f(x^{**}) &= 2\left(1 - \frac{4}{\sqrt{5}}\right) - \frac{2}{\sqrt{5}} \\ &= 2 - \frac{10}{\sqrt{5}} \end{aligned}$$

como $f(x^*) > f(x^{**})$ entonces x^* es máximo.

Para encontrar λ^* se sustituye x^* en cualquiera de las ecuaciones ya sea la 3.31 o 3.32, en este caso, para facilitar los calculos se seleccionó 3.32, por lo tanto se tiene que:

$$\lambda^* = \frac{1}{2\left(\frac{2}{\sqrt{5}}\right)} = \frac{\sqrt{5}}{4}.$$

Con la solución del problema $x^* = \left(1 + \frac{4}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ y $\lambda^* = \frac{\sqrt{5}}{4}$, se cumple la implicación del teorema de Lagrange, es decir, dado x^* existe λ^* tal que:

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x^*) &= \lambda^* \frac{\partial g}{\partial x_1}(x^*) \\ \frac{\partial f}{\partial x_2}(x^*) &= \lambda^* \frac{\partial g}{\partial x_2}(x^*) \end{aligned}$$

pues sustituyendo los valores se tiene que:

$$\begin{aligned} 2 &= \frac{\sqrt{5}}{4}\left(2\left(1+\frac{4}{\sqrt{5}}-1\right)\right) \\ 1 &= \frac{\sqrt{5}}{4}\left(2\left(\frac{2}{\sqrt{5}}\right)\right). \end{aligned}$$

La función de producción Cobb-Douglas

Para resolver el problema planteado en la página 16, donde la función de producción es de tipo Cobb-Douglas, nuevamente haremos uso del teorema de Lagrange. El problema está dado por:

$$\begin{aligned} \text{máx } P(L,K) &= bL^{\alpha}K^{1-\alpha} \\ \text{s.a. } nL+mK &= w \\ L,K &\geqq 0 \end{aligned}$$

En este caso contamos únicamente con la restricción presupuestal, por lo que aplicando el teorema 12 de Lagrange se tienen las siguientes condiciones que caracterizan la solución:

$$\begin{aligned} \left(\alpha bL^{\alpha-1}K^{1-\alpha},(1-\alpha)bL^{\alpha}K^{-\alpha}\right) &= \lambda\,(n,m) \\ mL+nK &= w \end{aligned}$$

lo cual equivale a resolver el siguiente sistema de ecuaciones:

$$\alpha bL^{\alpha-1}K^{1-\alpha} = \lambda n \tag{3.34}$$

$$(1-\alpha)bL^{\alpha}K^{-\alpha} = \lambda m \tag{3.35}$$

$$mL+nK = w \tag{3.36}$$

para resolver el sistema anterior, primero despejamos λ de las ecuaciones 3.34 y 3.35, es decir:

$$\begin{aligned} \lambda &= \frac{\alpha bL^{\alpha-1}K^{1-\alpha}}{n} \\ \lambda &= \frac{(1-\alpha)bL^{\alpha}K^{-\alpha}}{m} \end{aligned}$$

de donde se deduce la siguiente igualdad:

$$\frac{\alpha bL^{\alpha-1}K^{1-\alpha}}{n}=\frac{(1-\alpha)bL^{\alpha}K^{-\alpha}}{m}$$

de esta ecuación podemos despejar alguna de las variables de decisión, en este caso lo haremos con K por lo tanto:

$$\frac{K^{1-\alpha}}{K^{-\alpha}} = \frac{(1-\alpha)nL^{\alpha}}{\alpha mL^{\alpha-1}}$$

$$K = \frac{(1-\alpha)nL}{\alpha m} \tag{3.37}$$

ahora sustituyendo el valor de K en la ecuación 3.36, se tiene:

$$nL + m\left(\frac{(1-\alpha)nL}{\alpha m}\right) = w$$

despejando L,se tiene que:

$$L^* = \frac{\alpha w}{n}$$

L^* es solución del problema, ya que se expresa en términos de datos que son conocidos pues α se puede estimar y w, n son variables exógenas[1] del modelo. Por último, para encontrar el valor de K^* se sustituye L^* en la ecuación 3.37, entonces:

$$\begin{aligned} K^* &= \frac{(1-\alpha)n\left(\frac{\alpha w}{n}\right)}{\alpha m} \\ &= \frac{(1-\alpha)w}{m} \end{aligned}$$

por lo tanto la producción máxima para el sector productivo en este modelo está dada por:

$$P^*(L^*, K^*) = b\left(\frac{\alpha w}{n}\right)^{\alpha}\left(\frac{(1-\alpha)w}{m}\right)^{1-\alpha}.$$

Recordemos que otra variante del problema de la empresa, supone que la producción es fija. Dicho problema quedó planteado de la siguiente manera:

$$\begin{aligned} \text{mín}\, C(L,K) &= mL + nK \\ \text{s.a.}\;\; bL^{\alpha}K^{1-\alpha} &= Q \end{aligned}$$

y para plantearlo como un problema de maximización se multiplica por un signo menos a la función objetivo, es decir:

$$\begin{aligned} \text{máx}\, -C(L,K) &= -(mL + nK) \\ \text{s.a.}\;\; bL^{\alpha}K^{1-\alpha} &= Q \end{aligned}$$

entonces las condiciones que caracterizan la solución del problema anterior son:

$$\begin{aligned} (-n,-m) &= \lambda\left(\alpha bL^{\alpha-1}K^{1-\alpha}, (1-\alpha)bL^{\alpha}K^{-\alpha}\right) \\ bL^{\alpha}K^{1-\alpha} &= Q \end{aligned}$$

lo que equivale al siguiente sistema de ecuaciones:

$$-n = \lambda\alpha bL^{\alpha-1}K^{1-\alpha} \tag{3.38}$$
$$-m = \lambda(1-\alpha)bL^{\alpha}K^{-\alpha} \tag{3.39}$$
$$Q = bL^{\alpha}K^{1-\alpha} \tag{3.40}$$

[1] Una variable exógena es aquella que se supone determinada por fuerzas o causas externas al modelo y cuyos valores se aceptan sólo como datos dados. Así mismo se conocen como variables endógenas a aquellas cuyos valores solución se encuentran a partir del modelo.

despejando λ de 3.38 y 3.39, se tiene:

$$\begin{aligned}\lambda &= \frac{-n}{\alpha b L^{\alpha-1}K^{1-\alpha}}\\ \lambda &= \frac{-m}{(1-\alpha)bL^{\alpha}K^{-\alpha}}\end{aligned}$$

lo que nos lleva a establecer la siguiente igualdad

$$\frac{-n}{\alpha L^{\alpha-1}K^{1-\alpha}} = \frac{-m}{(1-\alpha)L^{\alpha}K^{-\alpha}}$$

que equivale a

$$\frac{\alpha L^{\alpha-1}K^{1-\alpha}}{n} = \frac{(1-\alpha)L^{\alpha}K^{-\alpha}}{m}$$

y despejando K se obtiene que

$$K = \frac{(1-\alpha)nL}{\alpha m}. \tag{3.41}$$

Esta expresión es análoga a la obtenida en la ecuación 3.37 en la variante anterior del problema en donde se maximizó la producción.

Sustituyendo el valor de K en la ecuación 3.40 se obtiene

$$\begin{aligned}Q &= bL^{\alpha}\left(\frac{(1-\alpha)nL}{\alpha m}\right)^{1-\alpha}\\ &= bL\left(\frac{(1-\alpha)n}{\alpha m}\right)^{1-\alpha}\end{aligned}$$

por lo que despejando se tiene el valor de L^* que es solución

$$L^* = \frac{Q}{b}\left(\frac{\alpha m}{(1-\alpha)n}\right)^{1-\alpha}$$

finalmente para obtener K^* se sustituye L^* en la ecuación 3.41, lo que nos lleva a

$$K^* = \frac{Q}{b}\left(\frac{(1-\alpha)n}{\alpha m}\right)^{\alpha}$$

por lo tanto L^* y K^* son los valores que minimizan los costos de la empresa.

Capítulo 4

El problema del consumidor cuando la función de utilidad es casi Bernoulli

Motivación

La interpretación que hace Daniel Bernoulli es la siguiente:

Sea $U(x) = k$ la utilidad que le reporta a un individo tener x cantidad de oro y sea

$$U(x+\varepsilon) - U(x) = a$$

la utilidad que le reporta a un individuo, tener una cantidad adicional ε de oro, bajo la hipótesis de que ya tenia x cantidad de oro.

La idea de Bernoulli es que la utilidad que le produce tener una cantidad adicional por unidad de oro es inversamente proporcional a la cantidad por unidad de oro que ya tenía, es decir:

$$\frac{U(x+\varepsilon) - U(x)}{\varepsilon} = \frac{a}{x}$$

que en términos de la derivada de la función de utilidad, podemos escribir como:

$$U'(x) = \frac{a}{x}$$

en donde $a > 0$ es la constante de proporcionalidad. De ahí se deduce que:

$$U(x) = a\log(x) + c.$$

Para generalizar la idea de Bernoulli, supondremos que las elecciones posibles de consumo de un individuo en cierta economía se pueden representar mediante la elección de un punto particular en un espacio de tipos de bienes. Si suponemos que existe un número finito n de tipos de bienes, las cantidades de cada uno de estos tipos de bienes adquiridos por el individuo, se representan por un vector de consumo dado por:

$$x = (x_1, x_2, ..., x_n),$$

donde x_i es la cantidad del bien del tipo i, con $i = 1, 2, ..., n$, en el supuesto de que se puede adquirir cualquier cantidad no negativa de él. Los vectores de consumo que cumplen esta condición forman el espacio de consumo dado por:

$$C = \{x = (x_1, x_2, ..., x_n) / \ x_i \geq 0, \ i = 1, 2, ..., n\}.$$

Es decir el espacio de consumo es el ortante no negativo en el espacio euclidiano de dimensión n. Este conjunto se denota como R^n_+ y tiene las propiedades de ser un conjunto convexo[1], y cerrado[2].

Sea $x = (x_{1,}x_2, .., x_n) \geqq \overline{0}$, un vector de bienes de consumo no negativo; entonces la utilidad generada para cada uno de los n bienes está dada por:

$$\begin{aligned} U(x_1) &= a_1 \log(x_1) + c \\ U(x_2) &= a_2 \log(x_2) + c \\ &\cdots \\ U(x_n) &= a_n \log(x_n) + c \end{aligned}$$

por lo que la utilidad total que le reporta al individuo cuando consume de todos los bienes, se obtiene sumando cada componente, es decir:

$$U(x) = \sum_{i=1}^{n} a_i \log(x_i) + c. \tag{4.1}$$

Así pues, la interpretación de las a_i en la función de utilidad, corresponde a la proporción fija de la renta que el individuo gasta en el $i - \acute{e}simo$ bien.

Definición de la función casi Bernoulli

Con esta motivación procederemos a formalizar primero la definición de la función de utilidad casi Bernoulli. Cabe destacar que buscando simplificar el problema se hace $c = 0$ en la expresión 4.1.

Definición 5 *Sea $U : R^n_+ \to \overline{R}$ donde $\overline{R}$ denota el sistema de los reales extendidos es decir $R \cup \{+\infty, -\infty\}$. La función de utilidad casi Bernoulli está dada por:*

$$U(x) = \begin{cases} \displaystyle\sum_{i=1}^{n} a_i \log(x_i) & si \quad x > \overline{0} \\ -\infty & si \quad x \ngtr \overline{0} \end{cases}.$$

[1] Definición de convexo, ver página 74 definición 6.

[2] Un subconjunto Ω de R^n es cerrado cuando su complemento $R^n - \Omega$ es abierto. Es decir, para todo $x \notin \Omega$, existe $\varepsilon > 0$ tal que $B(x; \varepsilon)$ está contenida en $R^n - \Omega$.

Donde $a_i > 0 \ \ \forall \ i = 1, 2, ..., n$ *y además* $\sum_{i=1}^{n} a_i = 1$.

Planteamiento y comprensión del problema

Para construir el problema del consumidor definiremos el concepto de restricción presupuestaria, que es la que restringe al consumidor a un subconjunto del espacio de consumo. Ésta establece que el gasto total monetario en todos los bienes no puede exceder del ingreso monetario del individo o consumidor. Supongamos entonces un vector de precios representado por:

$$p = (p_1, p_2, ..., p_n),$$

siendo p_i el precio unitario del bien del tipo i. Se define al ingreso monetario del individuo en un determinado periodo como r, ambos parámetros positivos dados. La restricción presupuestaria nos dice que el gasto no puede exceder al ingreso, lo cual matemáticamente puede formularse como el producto punto del vector de precios p y un vector de consumo x, así se tiene que:

$$p \cdot x \leqq r$$

es decir

$$\sum_{i=1}^{n} p_i x_i \leqq r$$

que en forma desarrollada

$$p_1 x_1 + p_2 x_2 + ... + p_n x_n \leqq r$$

siendo $p_i x_i$ el gasto realizado en el bien o artículo i.

Lo que nos lleva a decir que el conjunto factible o de oportunidades para el consumidor es

$$F = \{x \in C \ / \ p \cdot x \leqq r\},$$

el cual es un subconjunto convexo, compacto[3] y no vacío definido en el espacio de consumo C.

Si suponemos el caso de solamente dos bienes, la región está comprendida entre las siguientes rectas

$$\begin{aligned} p_1 x_1 + p_2 x_2 &\leqq r \\ x_1 &\geqq 0 \\ x_2 &\geqq 0 \end{aligned}$$

en donde el conjunto factible satisface las tres restricciones que corresponden a conjuntos que son cerrados, en consecuencia, el conjunto que satisface las tres condiciones es una intersección de cerrados. Por lo tanto, la región factible es cerrada y está dada por la figura 4-1.

En el caso de tener tres bienes, la región factible también es un conjunto cerrado y está dada por la figura 4-2.

[3] Un conjunto compacto es un conjunto cerrado y acotado.

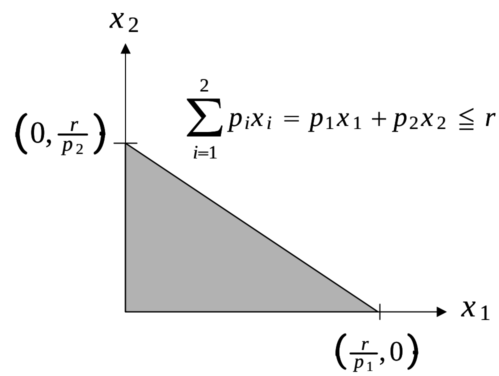


Figura 4-1 Región factible para dos bienes

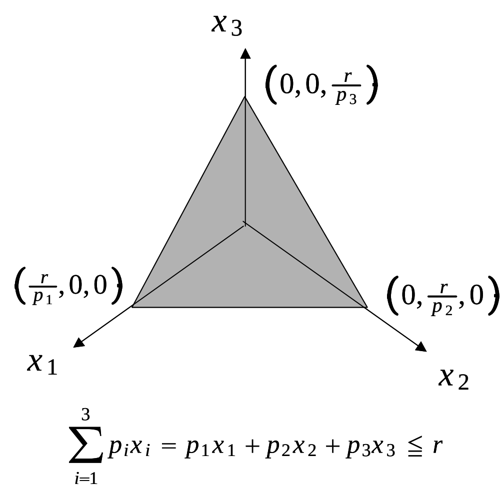


Figura 4-2 Región factible para tres bienes

En el caso de n bienes la figura es un simplejo[4] y es un conjunto cerrado que queda contenido en el espacio n-dimensional .

Para establecer el problema del consumidor, empecemos suponiendo una función de utilidad casi Bernoulli y una economía con n diferentes tipos de bienes, en la cual rige el vector de precios $p = (p_1, p_2, .., p_n) > \overline{0}$. Además, el individuo dispone de una cantidad $r > 0$. Entonces, el individuo se plantea el problema de demandar cierta cantidad de cada uno de los bienes, de tal forma que maximice su satisfacción o utilidad, sujeto a lo que pueda gastar, o dicho de otra forma:

Problema 2

[4]Un simplejo canónico es el conjunto $S = \left\{ x \in R^n_+, \text{ tal que } \sum_{i=1}^{n} x_i = 1 \right\}$.

$$
\begin{aligned}
& & \text{máx} \quad U(x) \\
\text{s.a.} \ \sum_{i=1}^{n} p_i x_i & \leqq & r \\
\text{con } x_i & \geqq & 0 \\
\text{para toda } \ i & = & 1, 2, ..n.
\end{aligned}
$$

Con el objetivo de comprender mejor el problema, se considera el siguiente caso particular, en el cual esbozaremos su gráfica.

Ejemplo 3 *Se tiene una economía con dos bienes* (x_1, x_2), *con el vector de precios* $(25, 20)$, *y suponemos que el consumidor dispone de la cantidad* $r = 100$, *y la función de utilidad casi Bernoulli dada por*

$$
U(x_1, x_2) = \begin{cases} 0{,}5 \log(x_1) + 0{,}5 \log(x_2) & si \ (x_1, x_2) > \overline{0} \\ -\infty & si \ (x_1, x_2) \ngtr \overline{0} \end{cases};
$$

es decir, si suponemos que consume de ambos bienes, entonces el problema que se plantea es el siguiente

$$
\begin{aligned}
\text{máx } \ U(x_1, x_2) & = & 0{,}5 \log(x_1) + 0{,}5 \log(x_2) \\
s.a \ \ 25x_1 + 20x_2 & \leqq & 100 \\
con \ \ x_1, x_2 & > & 0 \\
para\ toda\ i & = & 1, 2.
\end{aligned}
$$

Empezaremos por graficar la función de utilidad

$$
U(x_1, x_2) = 0{,}5 \log(x_1) + 0{,}5 \log(x_2),
$$

la cual está contenida en el primer ortante del espacio tridimensional. Ver figura **??**.Obsérvese que la función es monótona creciente, en particular es una función cóncava.

Por otro lado, si hacemos una proyección al plano x_1, x_2, nos permitirá ver las curvas de nivel de la función de utilidad cuyas ecuaciones son de la forma

$$
x_2 = \frac{k}{x_1},
$$

que se obtienen al igualar la función de utilidad casi Bernoulli a una constante y despejar a x_2.

Una curva de nivel de la función de utilidad es el conjunto de puntos del espacio euclidiano de dimensión n, para la cual el valor de la función de utilidad es constante, de modo que diferentes constantes dan lugar a curvas de nivel diferentes. En el caso de funciones monótonas crecientes, a mayor valor de la constante mayor será la utilidad que le reporta al consumidor. En términos económicos, dichas curvas son comúnmente conocidas como las curvas de indiferencia.

Si graficamos la restricción presupuestaria y las curvas de nivel en el mismo plano se obtiene la figura 4-4.

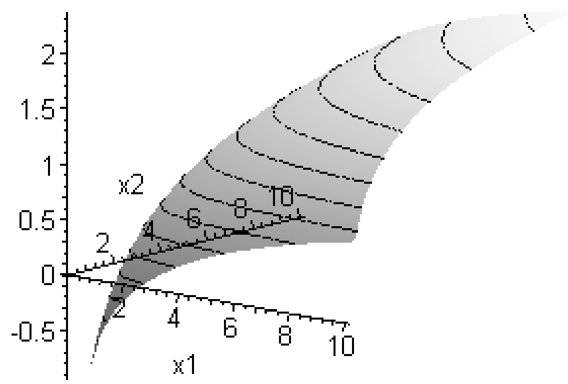


Figura 4-3 Función de Utilidad Bernoulli

Donde se observa que la parte sombreada corresponde al conjunto factible F.

A la vista de la representación, vemos que el individuo maximiza su utilidad en el punto donde alguna de las curvas de nivel $U(x_1, x_2)$ es tangente a la recta presupuestaria. Las tres curvas de nivel representadas aquí son:

$$\begin{aligned}
U_1(x_1, x_2) &= 0{,}5 \log(x_1) + 0{,}5 \log(x_2) = 0{,}3010 \\
U_2(x_1, x_2) &= 0{,}5 \log(x_1) + 0{,}5 \log(x_2) = 0{,}3494 \\
U_3(x_1, x_2) &= 0{,}5 \log(x_1) + 0{,}5 \log(x_2) = 0{,}3890.
\end{aligned}$$

Más adelante resolveremos este problema de forma analítica.

Para darnos una idea de cómo sería la gráfica de la función de utilidad casi Bernoulli con diferentes parámetros se incluye la figura 4-5, donde se grafica la ecuación

$$U(x_1, x_2) = \frac{1}{9} \log(x_1) + \frac{8}{9} \log(x_2).$$

Se puede ver que la función sigue siendo monótona creciente y más recargada hacia el eje x_1.

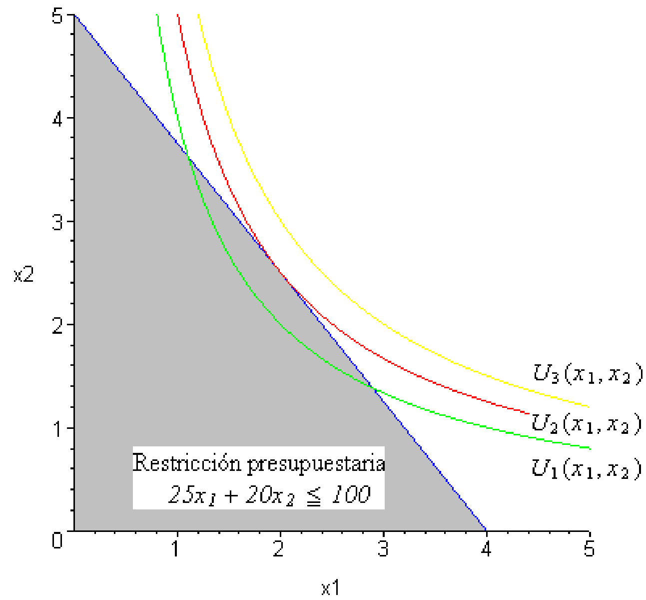


Figura 4-4

Existencia de solución

Una vez expresado el problema que pretendemos resolver nos preguntamos si tiene solución y en caso afirmativo, veremos cuál es ésta.

Como se vio en la sección anterior, cuando tenemos un problema con dos variables su solución se puede ver gráficamente. Pero dada la limitación que presenta el método gráfico cuando el número de variables es mayor que dos, conviene disponer de métodos generales que garanticen la existencia de la sulución óptima, como es el caso del teorema de Weierstrass, ver teorema 1 en la página 25.

Obsérvese que en las hipótesis del teorema se pide que el dominio de la función sea un conjunto compacto, lo cual ha sido analizado, además se pide que la función sea continua en el dominio. Para el caso de la función de utilidad casi Bernoulli este hecho no está claro, pues cuando alguno de los consumos vale cero la función no está bien definida. Para solucionar este problema, utilizaremos el hecho de que cualquier transformación monótona de la función de utilidad representa exactamente las mismas preferencias de consumo del individuo, es decir, si $U(x)$ es una función de utilidad, entonces lo será también $\varphi[U(x)]$, donde φ es una función estrictamente creciente. Por otro lado sabemos que la función exponencial es una función estrictamente creciente por lo que $e^{U(x)}$ seguirá siendo una función de utilidad.

Para nuestro caso, sí suponemos que el individuo consume de todos los bienes, se tiene que la función casi Bernoulli esta dada por

$$U(x) = \sum_{i=1}^{n} a_i \log(x_i).$$

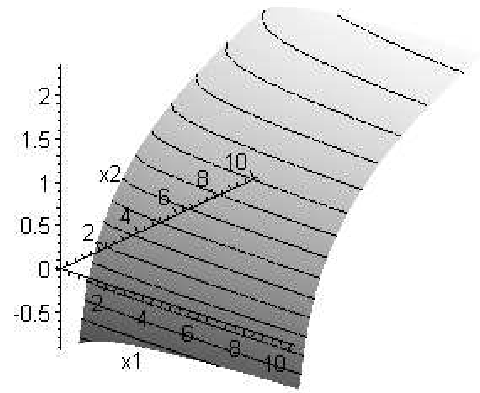


Figura 4-5 $U(x_1, x_2) = \frac{1}{9}\log(x_1) + \frac{8}{9}\log(x_2)$.

Componiendo de ambos lados con la función exponencial se tiene

$$\begin{aligned} e^{\sum_{i=1}^{n} a_i \log(x_i)} &= e^{a_1 \log(x_1) + a_2 \log(x_2) + \ldots + a_n \log(x_n)} \\ &= e^{\log(x_1)^{a_1}} e^{\log(x_2)^{a_2}} \cdots e^{\log(x_n)^{a_n}} \\ &= x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}, \end{aligned}$$

donde la función obtenida es un producto de funciones del tipo $x_i^{a_i}$ en la que cada una es continua y siempre está bien definida, ya que el producto de continuas es continua, y al igual que la función casi Bernoulli también cumple con ser monótona creciente. En consecuencia ambas funciones están fuertemente emparentadas.[5]

Si aplicamos el teorema de Weierstrass a la nueva función podemos asegurar que existe x^* solución del problema del consumidor.

Solución del problema

En la presente sección resolveremos de manera analítica el problema del consumidor dado que sabemos que existe solución al problema.

[5] La función $x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ es comúnmente conocida en teoría económica como la función de tipo Cobb-Douglas (ver ejemplo 1 en la página 16).

Para ello, en una primera aproximación demostraremos que la solución se encuentra sobre el hiperplano de la restricción; es decir, el problema se puede plantear como un problema de optimización con restricciones de igualdad. Se sabe que para resolver este tipo de problemas, el teorema de Lagrange nos proporciona las condiciones que nos permiten encontrar los puntos óptimos. En el caso de la función de utilidad casi Bernoulli, se desea encontrar el vector de consumo que la maximiza.

Reducción al problema de Lagrange

La solución del problema del consumidor está sobre el hiperplano de la restricción. Para demostrarlo y para tener una idea geométrica consideremos solamente el caso de dos bienes. Definimos el siguiente conjunto:

$$F = \left\{ x \in R_+^2 \text{ tal que } \sum_{i=1}^{2} p_i x_i \leqq r \right\}.$$

Se observa que F es un conjunto compacto, donde geométricamente tendríamos la región para dos bienes dada por la figura 4-1 de la página 64.

De esto se deduce que x^* podría estar dentro de los siguientes dos casos:

- Que la solución esté dentro de la región sombreada, es decir que $p_1x_1 + p_2x_2 < r$.
- Que la región este sobre el hiperplano, es decir sobre la recta $p_1x_1 + p_2x_2 = r$.

Caso 1 (Dentro del conjunto factible) *Supongamos que $x^* = (x_1^*, x_2^*)$ solución del problema, está dentro de la región, entonces se tiene:*

$$p_1x_1^* + p_2x_2^* < r,$$

definamos a un nuevo vector de consumo dado por:

$$\overline{x} = (\overline{x}_1, \overline{x}_2) = (x_1^* + \Delta, x_2^*) \;\; \textit{con } \Delta > 0.$$

Por demostrar que $\overline{x} \in F$.

Demostración. Por construcción $\overline{x} \in R_+^2$ ya que es un vector no negativo, definamos

$$\Delta = \frac{r - p_1x_1^* - p_2x_2^*}{p_1} > 0.$$

Por hipótesis $p_1x_1^* + p_2x_2^* < r$, considerese ahora $p_1\overline{x}_1 + p_2\overline{x}_2$ y como se definió $\overline{x}$, se tiene lo siguiente:

$$\begin{aligned} p_1\overline{x}_1 + p_2\overline{x}_2 &= p_1\left(x_1^* + \Delta\right) + p_2x_2^* \\ &= p_1x_1^* + p_1\Delta + p_2x_2^* \end{aligned}$$

donde

$$p_1x_1^* + p_1\Delta + p_2x_2^* = p_1x_1^* + p_1\left(\frac{r - p_1x_1^* - p_2x_2^*}{p_1}\right) + p_2x_2^* = r,$$

por lo que se demuestra que $\overline{x} \in F$, se tiene entonces que pertenece a la región factible de nuestro problema. ■

Ahora, utilizando el hecho de que la función de utilidad Cobb-Douglas es monótona creciente, demostraremos la siguiente afirmación:

$$U(\overline{x}) > U(x^*)$$

Demostración. Por hipótesis se tiene que:

$$x_1^* + \Delta > x_1^*$$

de donde se cumple que

$$\begin{aligned}
(x_1^* + \Delta)^{a_1} &> (x_1^*)^{a_1} \text{ con } 0 < a_1 < 1 \\
(x_1^* + \Delta)^{a_1}(x_2^*)^{a_2} &> (x_1^*)^{a_1}(x_2^*)^{a_2} \text{ con } 0 < a_2 < 1 \text{ y } x_2^* \geqq 0 \\
U(x_1^* + \Delta, x_2^*) &> U(x_1^*, x_2^*),
\end{aligned}$$

por lo que podemos decir que se cumple la desigualdad

$$U(\overline{x}) > U(x^*).$$

Lo cual contradice la hipótesis de que x^* sea solución del problema. ■

De aquí, se deduce que x^* debe estar sobre la restricción del problema. Obsérvese que no puede estar sobre los ejes ya que en estos puntos la función vale cero, lo cual es una solución trivial del problema en el caso de la función Cobb-Douglas y en el caso de la función casi Bernoulli, dicha función no está definida para el valor cero. Es decir, para el caso de dos bienes la solución está sobre la recta excluyendo las orillas, y dicha región es un conjunto abierto, ver figura 4-6.

De manera análoga se puede demostrar que para el caso de n bienes, también se cumple que la solución está sobre el hiperplano de la restricción.

Solución del problema utilizando el teorema de Lagrange

Basándonos en que la solución está sobre el hiperplano de la restricción, podemos redefinir el problema del consumidor, ahora como un problema de optimización con restricciones de igualdad, es decir:

$$\begin{aligned}
& \text{máx} \quad U(x) \\
\text{s.a. } \sum_{i=1}^{n} p_i x_i - r &= 0 \\
\text{con } x_i &> 0 \\
\text{para toda } i &= 1, 2, \ldots, n.
\end{aligned}$$

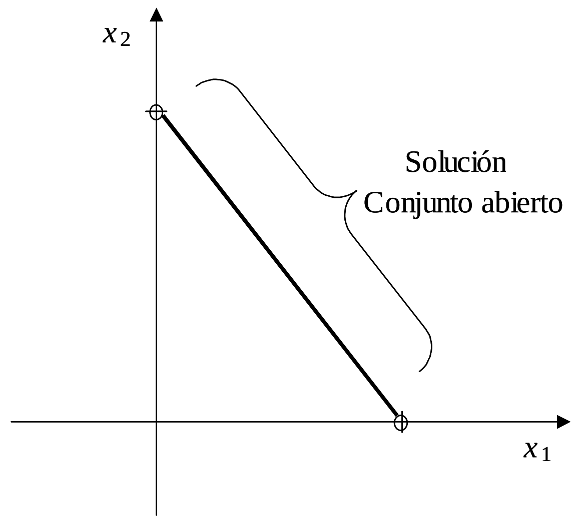


Figura 4-6 Conjunto solución.

El problema cumple con las hipótesis del teorema 12 de Lagrange, pues está definido en un conjunto abierto, y en este caso se reduce a una sola restricción.

Establecemos las $n+1$ condiciones para la solución del problema.

Las primeras n están dadas por

$$\frac{\partial U}{\partial x_i}(x^*) = \lambda \frac{\partial g}{\partial x_i}(x^*) \text{ para toda } i = 1, 2, \ldots, n$$

que en nuestro caso particular serían

$$a_i \frac{1}{x_i} = \lambda p_i \text{ para toda } i = 1, 2, .., n \tag{4.2}$$

y además tenemos la restricción misma del problema como una condición más dada por

$$\sum_{i=1}^{n} p_i x_i = r. \tag{4.3}$$

De las primeras n condiciones dadas en (4.2) se cumple que:

$$a_i \frac{1}{\lambda} = p_i \; x_i \quad \text{para toda } i = 1, 2, .., n$$

ahora sustituyendo lo anterior en la restricción (4.3), se tiene que:

$$\begin{aligned}\sum_{i=1}^{n} p_i x_i &= \sum_{i=1}^{n} a_i \frac{1}{\lambda} \\ &= \frac{1}{\lambda} \sum_{i=1}^{n} a_i \\ &= \frac{1}{\lambda}\end{aligned}$$

y por otro lado

$$\begin{aligned}\sum_{i=1}^{n} p_i x_i &= r \\ \frac{1}{\lambda} &= r\end{aligned}$$

por lo que

$$\lambda = \frac{1}{r}.$$

Nuevamente de las primeras n condiciones dadas por (4.2), obtenemos

$$\begin{aligned}x_i &= a_i \frac{1}{\lambda p_i} \\ &= a_i \frac{1}{\frac{1}{r} p_i} \\ &= \frac{r a_i}{p_i} \quad \text{para toda } i = 1, 2, .., n.\end{aligned}$$

Por lo tanto

$$x^* = (x_1^*, x_2^*, ..., x_n^*) = \left(\frac{r a_1}{p_1}, \frac{r a_2}{p_2}, \ldots, \frac{r a_n}{p_n} \right) \tag{4.4}$$

es la solución para el problema del consumidor en el caso de la función de utilidad casi Bernoulli, donde x_i^* es la demanda del $i-$ésimo bien.

Retomando el ejemplo 3 de la página 65, en el cual hemos analizado detalladamente su gráfica, daremos la solución analítica a este caso en particular.

Sustituyendo los valores $a_1 = 0{,}5$, $a_2 = 0{,}5$, $p_1 = 25$, $p_2 = 20$ y $r = 100$ en la expresión 4.4, se tiene que el vector de demanda está dado por:

$$\begin{aligned}x^* &= (x_1^*, x_2^*) \\ &= \left(\frac{r a_1}{p_1}, \frac{r a_2}{p_2} \right) \\ &= \left(\frac{100 * 0{,}5}{25}, \frac{100 * 0{,}5}{20} \right) \\ &= \left(2, \frac{5}{2} \right),\end{aligned}$$

lo que significa que el individuo debe consumir sólo 2 unidades del bien tipo 1 y 2,5 unidades del bien tipo 2. Obsérvese que las a_i son iguales, por lo que se puede inferir que la propensión al consumo para ambos bienes es la misma.

Capítulo 5

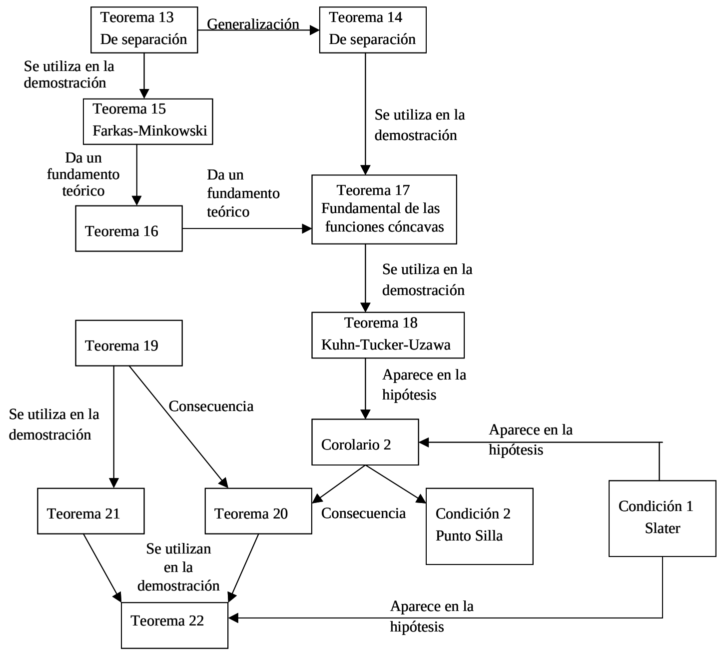
Introducción a la programación cóncava y casi cóncava

Introducción

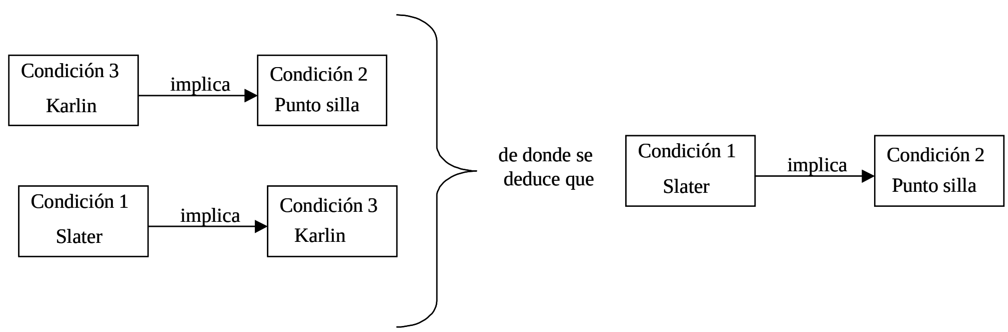
En ocasiones es importante determinar la solución para problemas de optimización, donde no necesariamente aparezcan restricciones de igualdad, como en el caso del teorema de Lagrange. Para ello, se requiere de la teoría de la programación cóncava. En este capítulo principalmente abordaremos los teoremas que nos permiten caracterizar las soluciones para un problema de programación no lineal con funciones cóncavas.

Primero se demuestran el teorema de separación y el teorema fundamental de las funciones cóncavas, los cuales dan el sustento teórico para la demostración del teorema de Kuhn Tucker Uzawa y su corolario 2, que a su vez proporcionan una caracterización a la solución del "problema estándar de programación no linial con funciones cóncavas" al que denotamos como PPC, resumiendo lo anterior en los teoremas 19 y 20. Cabe destacar que en dichos teoremas la solución al problema PPC queda caracterizada en términos de la condición punto silla y para relacionarlo con el cálculo diferencial se definen las condiciones casi punto silla. En el teorema 21 se demuestra la relación entre la condición punto silla y la condición casi punto silla, así en el teorema 22 caracterizamos nuevamente la solución al problema PPC pero ahora en términos de las condiciones casi punto silla, con lo cual podremos calcular la solución al problema pero ahora utilizando como herramienta el cálculo diferencial. Finalmente se enuncia la definición de funciones casi cóncavas la cual es una caracterización más débil de las funciones cóncavas y se demuestran algunos resultados sobre la relación entre ellas.

El siguiente diagrama muestra cómo están relacionados los diferentes teoremas que aparecen en el presente capítulo.



Por otro lado, se tiene un segundo diagrama que se deriva de las proposiciones 5 y 6.



Teoremas de separación

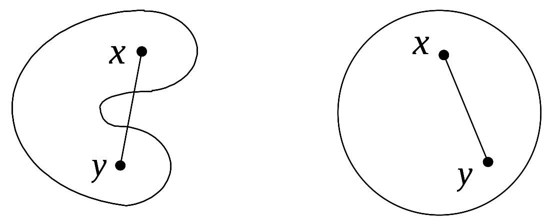
Los diversos resultados que veremos en esta sección son los preámbulos necesarios para abordar las siguientes tres secciones en donde se generalizan los resultados vistos anteriormente.

Empezamos por recordar el concepto de conjunto convexo quizá ya familiar por su aparición frecuente en el estudio de R^n.

Definición 6 *Un conjunto $X \subset R^n$ es* ***convexo*** *si y sólo si para todo par de puntos $x, y \in X$ se tiene que para todo $t \in [0,1]$ entonces*

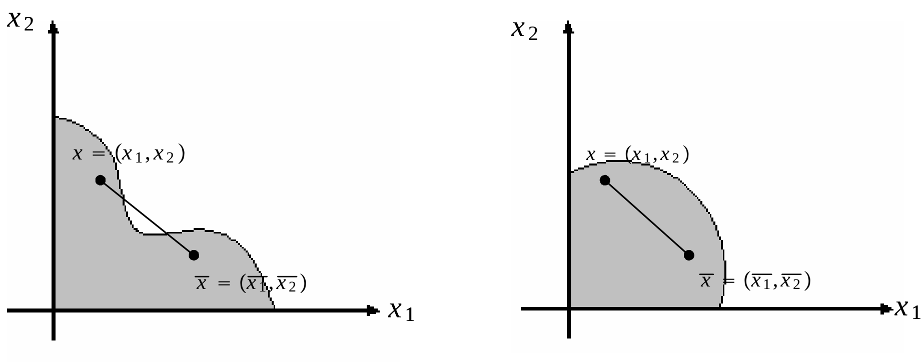
$$ty + (1-t)x \in X.$$

Para dar una idea intuitiva de la definición anterior se muestran las siguientes figuras.



(a) (b)

En el inciso (a) tenemos un conjunto no convexo, ya que existen puntos x e y de dicho conjunto tales que el segmento que los une no está contenido en él; sin embargo esto no sucede en el conjunto del inciso (b) por lo que el conjunto (b) es un ejemplo de conjunto convexo. Otros ejemplos vistos en el plano cartesiano serían los siguientes.



(a) (b)

A continuación se presenta un ejemplo más detallado con su demostración.

Ejemplo 4 *Demostrar que el conjunto* $S = \{(x_1, x_2) \in R^2 \mid x_2 \geqslant x_1\}$ *es un conjunto convexo.*

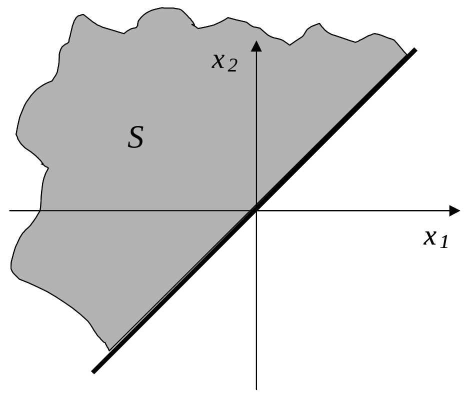
Sean $x = (x_1, x_2)$ e $y = (y_1, y_2)$ dos elementos de S. Probaremos que para todo $t \in [0, 1]$

$$\begin{aligned} z &= (z_1, z_2) \\ &= ty + (1-t)x \\ &= (ty_1 + (1-t)x_1, ty_2 + (1-t)x_2) \end{aligned}$$

también pertenece a S. Por ser x e y elementos de S se verifica que $x_2 \geqslant x_1$ e $y_2 \geqslant y_1$ y como $t \geqslant 0$ y $1 - t \geqslant 0$ se tiene que $ty_2 \geqslant ty_1$ y $(1-t)x_2 \geqslant (1-t)x_1$. Por tanto

$$z_2 = ty_2 + (1-t)x_2 \geqslant ty_1 + (1-t)x_1 = z_1,$$

lo que prueba que S es convexo ya que $z \in S$. Gráficamente el conjunto es el siguiente:



Definición 7 *Sea $P \in R^n$ no nulo y $\alpha \in R$. El conjunto*

$$H = \{x \in R^n \text{ tal que } P \cdot x = \alpha\}$$

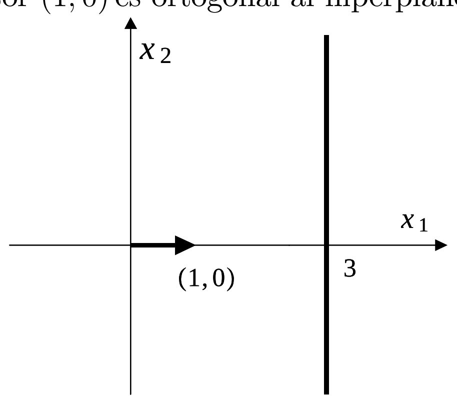
*es llamado un **hiperplano** en R^n con normal P.*

Ejemplo 5 *Para ilustrar esta definición se muestran las gráficas de un hiperplano en R^2 y un hiperplano en R^3 respectivamente.*

Sea $P = (1, 0)$ y $\alpha = 3$ entonces

$$\begin{aligned} P \cdot x &= \alpha \\ (1,0)(x_1, x_2) &= 3 \\ x_1 &= 3. \end{aligned}$$

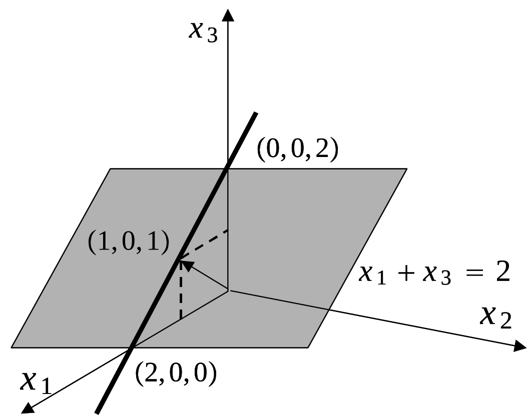
Gráficamente tenemos que el vector $(1, 0)$ es ortogonal al hiperplano $x_1 = 3$.



Sea $P = (1, 0, 1)$ y $\alpha = 2$ entonces

$$\begin{aligned} P \cdot x &= \alpha \\ (1,0,1)(x_1, x_2, x_3) &= 2 \\ x_1 + x_3 &= 2. \end{aligned}$$

Gráficamente tenemos que el vector $(1, 0, 1)$ es ortogonal al hiperplano $x_1 + x_3 = 2$.



Teorema 13 ***De separación entre un conjunto convexo y cerrado y un punto fuera de él.*** *Sea $X \neq \varnothing$ un subconjunto convexo y cerrado en R^n y sea $x_0 \notin X$; entonces:*

i) $\exists\ a \in X$ *tal que* $d(x_0, a) \leqq d(x_0, x)\ \forall\ x \in X$ *además,* $d(x_0, a) > 0$.

ii) $\exists\ p \in R^n$ *no nulo y* $\alpha \in R$ *tal que* $p \cdot x = \alpha, \forall x \in X$ *y* $p \cdot x_0 < \alpha$

Es decir, X *y* x_0 *quedan separados por el hiperplano*

$$H = \{x \in R^n \text{ tal que } p \cdot x = \alpha\}.$$

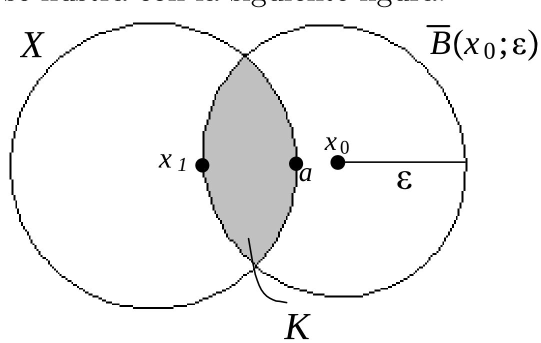
Demostración. Por hipótesis $X \neq \varnothing$. Sea $x_1 \in X$ y consideremos la bola cerrada

$$B(x_0; \varepsilon) = \{x \in R^n \text{ tales que } \|x_o - x\| \leqq \varepsilon\}$$

en donde $\varepsilon = \|x_o - x_1\| > 0$. Es decir que

$$K = B(x_0; \varepsilon) \cap X \neq \varnothing$$

es un conjunto compacto, que se ilustra con la siguiente figura:



Sea $f : K \to R$ definida como

$$f(x) = d(x, x_0)$$

de donde f es una función continua y por lo tanto alcanza un valor mínimo. Sea $a \in K$ mínimo; es decir,

$$d(a, x_0) \leqq d(x, x_0) \forall x \in K.$$

Si $x \in X - K$ entonces $x \in \overline{B}(x_0; \varepsilon)$ y $d(x, x_0) > \varepsilon \geqq d(a, x_0)$ entonces

$$d(x, a) \leqq d(x_0, x) \forall x \in X$$

y además, puesto que $x_0 \in X$, entonces $a \neq x_0$ de donde $d(a, x_0) > 0$, por lo tanto el inciso (i) queda probado.

Para probar (ii) : Sea $P = (a - x_0) \neq \overline{0}$ y sea $P \cdot a = \alpha$, de donde primeramente se tiene los siguiente

$$\begin{aligned} 0 &< P^2 \\ &= P(a - x_0) \\ &= P \cdot a - P \cdot x_0 \end{aligned}$$

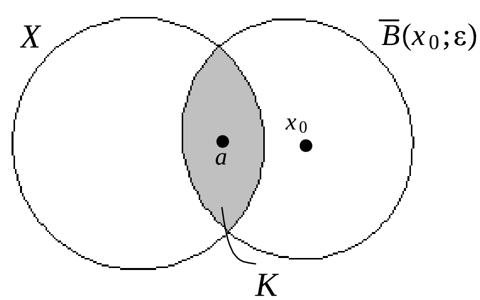
entonces

$$\begin{aligned} P \cdot a - P \cdot x_0 &> 0 \\ \alpha - P \cdot x_0 &> 0 \end{aligned}$$

por lo tanto

$$\alpha > P \cdot x_0$$

véase figura



Consideremos ahora $x \in X$ con $x \neq a$ y supongamos que $P \cdot x < \alpha$. Entonces sea

$$t = \frac{(a - x) \cdot P}{(x - a)^2} = \frac{P \cdot a - P \cdot x}{(x - a)^2} = \frac{\alpha - P \cdot x}{(x - a)^2} > 0$$

ya que por hipótesis $P \cdot x < \alpha$, entonces $\alpha - P \cdot x > 0$.

Pero al mismo tiempo se tiene

$$\begin{aligned}
1-t &= 1-\frac{P\cdot a \ -P\cdot x}{(x-a)^2}\\
&= \frac{(a-x)^2-(P\cdot a-P\cdot x)}{(x-a)^2}\\
&= \frac{(a-x)^2-P\cdot(a-x)}{(x-a)^2}\\
&= \frac{(a-x)\left[(a-x)-P\right]}{(x-a)^2}\\
&= \frac{(a-x)(a-x-a+x_0)}{(x-a)^2}\\
&= \frac{(a-x)(x_0-x)}{(x-a)^2};
\end{aligned}$$

sumando $-x_0+x_0$ se tiene

$$= \frac{(a-x_0+x_0-x)(x_0-x)}{(x-a)^2}$$

y resolviendo el producto

$$\begin{aligned}
&= \frac{(ax_0-ax-x_0^2+x_0x)+(x_0^2-x_0x-xx_0+x^2)}{(x-a)^2}\\
&= \frac{x_0(a-x_0)-x(a-x_0)+(x-x_0)^2}{(x-a)^2}\\
&= \frac{(a-x_0)\,(x_0-x)+(x-x_0)^2}{(x-a)^2}\\
&= \frac{P\cdot(x_0-x)+(x-x_0)^2}{(x-a)^2};
\end{aligned}$$

una vez más sumando $-P^2+P^2$

$$\begin{aligned}
&= \frac{(x-x_0)^2-P^2+P^2+P\cdot(x_0-x)}{(x-a)^2}\\
&= \frac{(x-x_0)^2-P^2+P\cdot(P+x_0-x)}{(x-a)^2}\\
&= \frac{(x-x_0)^2-P^2+P\cdot(a-x_0+x_0-x)}{(x-a)^2}\\
&= \frac{(x-x_0)^2-P^2+P\cdot(a-x)}{(x-a)^2}\\
&= \frac{(x-x_0)^2-P^2+P\cdot a-P\cdot x}{(x-a)^2}
\end{aligned}$$

y como $P\cdot a=\alpha$

$$= \frac{(x-x_0)^2-P^2+\alpha-P\cdot x}{(x-a)^2} > 0$$

y es mayor que cero pues

$$(x-x_0)^2 - P^2 = (x-x_0)^2 - (a-x_0)^2 > 0$$

ya que $(x-x_0)^2 > P^2 = (a-x_0)^2$ y por hipótesis

$$P \cdot x < \alpha \Leftrightarrow \alpha - P \cdot x > 0.$$

En virtud de que $x \in X$ por lo tanto

$$x = (1-t)a + tx \in X$$

ya que X es un conjunto convexo y además

$$\begin{aligned}
(\overline{x} - x_0)^2 &= [(1-t)a + tx - x_0]^2 \\
&= [a - at + tx - x_0]^2 \\
&= [t(x-a) + (a-x_0)]^2 \\
&= \left[(a-x_0) + \frac{(a-x)\cdot P}{(a-x)^2}(x-a)\right]^2 \\
&= (a-x_0)^2 + 2\left[\frac{(a-x)P(x-a)}{(x-a)^2}\cdot(a-x_0)\right] + \frac{[(a-x)\cdot P]^2 (x-a)^2}{(x-a)^4} \\
&= (a-x_0)^2 + 2\frac{(a-x)P(x-a)\cdot P}{(x-a)^2} + \frac{[(a-x)\cdot P]^2}{(x-a)^2} \\
&= (a-x_0)^2 - 2\frac{[(a-x)\cdot P]^2}{(x-a)^2} + \frac{[(a-x)\cdot P]^2}{(x-a)^2} \\
&= (a-x_0)^2 - \frac{((a-x)\cdot P)^2}{(x-a)^2} < (a-x_0)^2
\end{aligned}$$

Absurdo pues a es de todos los puntos de X el que dista menos, es decir

$$d(a, x_0) \leqq d(x, x_0) \forall x \in X.$$

Por lo que nuestro teorema queda probado. ■

Establecemos ahora nuestro segundo teorema de separación que generaliza al anterior.

Teorema 14 ***De separación entre un conjunto convexo y un punto fuera de él .*** *Sean $X \neq \varnothing$ subconjunto convexo de R^n no necesariamente cerrado y $x_0 \notin X$. Entonces existe $P \in R^n$ no nulo tal que se cumple*

$$P \cdot x \geqq P \cdot x_0 \quad \forall x \in X.$$

Demostración. Consideremos $\overline{X}$ la cerradura de X que también es convexo. Nos referimos a la cerradura como el conjunto de todos los puntos de acumulación unión el conjunto.

Si $x_0 \notin \overline{X}$ entonces nuestra afirmación se desprende inmediatamente del teorema 13. Ahora si $x_0 \in \overline{X}$, entonces, puesto que X es convexo sucede que

$$x_0 \notin int(\overline{X}) = int(X).$$

Por lo tanto, $x_0 \in Fr(\overline{X})$ donde $Fr(\overline{X})$ es el conjunto de puntos de la frontera de $\overline{X}$, y existe una sucesión $\{x^k\} \subset C(\overline{X})$, donde $C(\overline{X})$ denota a la cerradura del conjunto $\overline{X}$, tal que $x^k \to x_0$ cuando $k \to \infty$. Entonces, en virtud del teorema 13 se tiene que para todo $k = 1, 2, ...$ existe $\overline{P}^k \neq \overline{0}$ en R^n tal que

$$\overline{P}^k \cdot x > \overline{P}^k \cdot x^k \; \forall x \in \overline{X}.$$

Sea

$$P^k = \frac{\overline{P}^k}{\left|\overline{P}^k\right|} \text{ entonces } P^k \cdot x > P^k \cdot x^k \quad \forall x \in \overline{X}$$

y además

$$P^k \in S = \{P \in R^n \text{ tales que } |P| = 1\}.$$

Como S es un conjunto compacto, toda sucesión de sus elementos contiene una subsucesión convergente a un elemento del conjunto, es decir existe una subsucesión $\{P^{\varphi_k}\}$ de $\{P^K\}$ que converge a $P \in S$. Entonces se tiene que para todo $x \in \overline{X}$ se cumple

$$P^{\varphi_k} \cdot x \geqq P^{\varphi_k} \cdot x^{\varphi_k}$$

y haciendo tender k a ∞ tenemos que

$$P \cdot x \geqq P \cdot x_0.$$

Por lo que nuestro teorema queda probado. ■

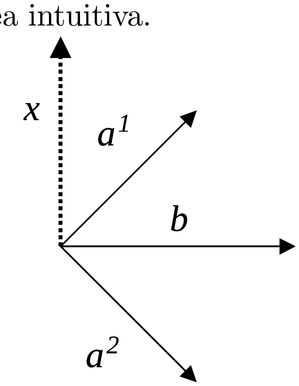
Finalmente, establecemos el útil teorema conocido como el lema de Farkas-Minkowski clave en la teoría de la programación lineal, en la teoría de juegos y para el teorema de Kuhn-Tucker así como para el teorema de Arrow-Hurwitz-Uzawa.

Teorema 15 (Farkas-Minkowski) *Sean $a^1, a^2, ..., a^m$ y $b \neq \overline{0}$ puntos de R^n. Supongamos que para todo $x \in R^n$ tales que $a^i \cdot x \geqq 0$ para $i = 1, 2, ..., m$ se tiene que $b \cdot x \geqq 0$. Entonces existen coficientes $\lambda_1, \lambda_2, ..., \lambda_m \geqq 0$ no todos nulos tales que*

$$b = \sum_{i=1}^{m} \lambda_i a^i$$

es decir b es combinación lineal de los a^i.

Demostración. Antes de empezar con la demostración, representaremos geométricamente el caso para cuando $m = 2$ y así tener una idea intuitiva.



Es decir b no puede salirse del abanico generado por las a^i.

Regresando a la demostración se define el siguiente conjunto

$$K = \left\{ x \in R^n \text{ tal que } x = \sum_{i=1}^{m} \lambda_i a^i \text{ con } \lambda_i \geqq 0 \text{ para } i = 1, 2, ..m \right\}$$

y probaremos que K es un conjunto convexo y cerrado.

Por simple inspección se puede ver que K es cerrado, sólo falta demostrar que K es convexo. Sean

$$\begin{aligned} x_1 &= \sum_{i=1}^{m} \lambda_i a^i \\ x_2 &= \sum_{i=1}^{m} \alpha_i a^i \end{aligned}$$

con $\lambda_i, \alpha_i \geqq 0$ y $t \in [0, 1]$ entonces

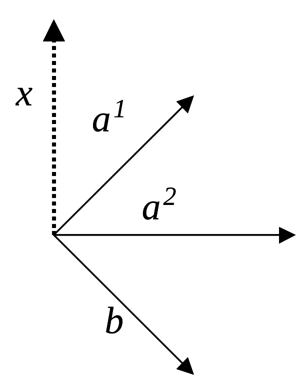
$$\begin{aligned} (1-t)x_1 + tx_2 &= (1-t)\sum_{i=1}^{m} \lambda_i a^i + t\sum_{i=1}^{m} \alpha_i a^i \\ &= \sum_{i=1}^{m} \left[(1-t)\lambda_i + t\alpha_i\right] a^i \end{aligned}$$

ahora si $\sigma_i = (1-t)\lambda_i + t\alpha_i$ donde $\sigma_i \geqq 0$ ya que λ_i, $\alpha_i \geqq 0$ para toda $i = 1, 2, ..., m$ y $t \in [0, 1]$ entonces

$$\sum_{i=1}^{m} \left[(1-t)\lambda_i + t\alpha_i\right] a^i = \sum_{i=1}^{m} \sigma_i a^i$$

por lo que $(1-t)x_1 + tx_2 \in K$ para todo $t \in [0, 1]$ es decir K es convexo.

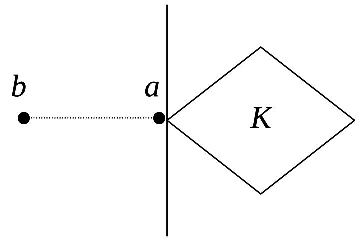
Continuando con la demostración, supongamos que $b \notin \dot{K}$; geométricamente para el caso $m = 2$ se tiene,



es decir b no está en el abanico formado por a^1 y a^2. Ahora por el teorema 13, sea $a \in K$ tal que

$$d(b, a) \leqq d(b, x) \text{ para todo } x \in K$$

donde gráficamente se tiene



Ahora sea $P = a - b$. Por construcción, como $b \notin K$ se tiene que $a \neq b$, entonces $P \neq \overline{0}$ y con lo anterior se define

$$P^2 = P \cdot (a - b) \text{ donde } P^2 > 0$$

es decir

$$\begin{aligned} P \cdot a - P \cdot b &> 0 \\ P \cdot a &> P \cdot b \end{aligned}$$

nuevamente utilizando el teorema 13 y la desigualdad anterior se tiene que $P \in R^n$ tales que

$$P \cdot x \geqq P \cdot a \quad \forall x \in K \tag{5.1}$$

y

$$P \cdot b < P \cdot a$$

donde $P \cdot a$ es el mínimo, y puesto que $\overline{0} \in K$ lo podemos escribir de la forma $\overline{0} = \sum_{i=1}^{m} \lambda_i a^i$, donde $\lambda_i = 0$ para todo $i = 1, 2, ..., m$, de lo anterior se tiene que

$$\overline{0} = P \cdot \overline{0} \geqq P \cdot a.$$

Supongamos que $P \cdot a < 0$ y sea $t > 1$; lo que nos lleva a decir que $t \cdot a \in K$ pues

$$a = \sum_{i=1}^{m} \lambda_i a^i$$

entonces

$$ta = \sum_{i=1}^{m} t\lambda_i a^i \text{ donde } t\lambda_i \geqq 0$$

por lo tanto

$$P \cdot ta = tP \cdot a < P \cdot a$$

lo cual es un absurdo, ya que $P \cdot a$ es el mínimo, de donde se deduce que $P \cdot a = 0$. Ahora como $a^i \in K \ \ \forall i = 1, 2, .., m$ con $P \cdot a = 0$ y por la condición 5.1 se cumple que

$$P \cdot a^i \geqq 0 \ \ \forall i = 1, 2, ..., m$$

pero

$$P \cdot b < 0.$$

Ésto contradice la hipótesis. Luego $b \in K$ por lo tanto nuestro teorema queda probado. ■

Programación Cóncava

Empezaremos por considerar el "problema estándar de programación no lineal con funciones cóncavas", al cual nos referiremos como PPC. Nuestro objetivo es en este caso caracterizar a las posibles soluciones del problema, así como demostrar los teoremas más importantes de la programación cóncava, pero antes tendremos que establecer algunas definiciones y hechos preliminares.

Procedemos ahora a definir el problema central de está sección:

Problema 3 (PPC)

$$\begin{aligned} & \text{máx} \;\; f(x) \\ \textit{sujeto a } g_j(x) & \geqq 0 \\ \textit{para } j & = 1, 2, .., m \end{aligned}$$

En donde $f, g_1, g_2, ..., g_m : X \to R$ *son funciones cóncavas definidas en el convexo* $X \subset R^n$.

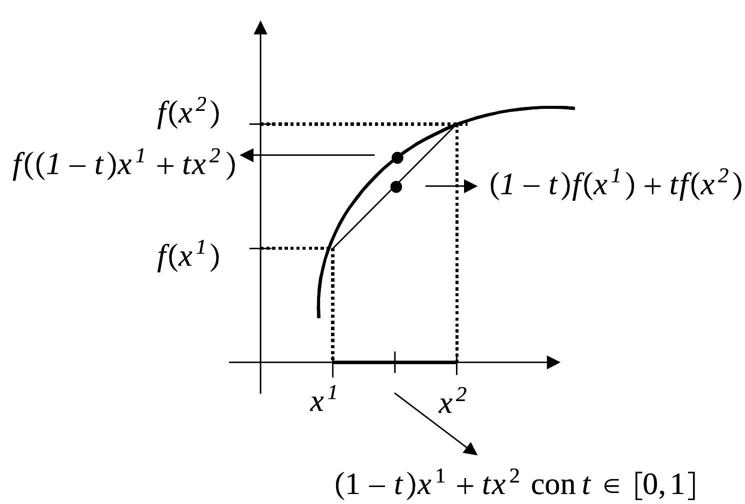
Decimos que una función es cóncava si sucede lo siguiente:

Definición 8 *Sea* $f : X \to R$ *en donde* $X \subset R^n$ *es un conjunto convexo, entonces decimos que* f *es una función cóncava si y sólo si para todo* $x^1, x^2 \in X$ *y* $t \in [0,1]$ *se cumple que*

$$f\left((1-t)\, x^1 + tx^2\right) \geqq (1-t)\, f(x^1) + tf(x^2)$$

En el caso en que la desigualdad estricta valga para $t \in (0,1)$ *con* $x^1, x^2 \in X$ *arbitrarias, decimos que* f *es estrictamente cóncava.*

Para tener una idea geométrica analizaremos el caso de una función cuya gráfica está contenida en R^2. Sea $f : X \to R$ con $X = [x_1, x_2]$ un convexo de R.



Si consideramos un ejemplo numérico de la definición anterior podríamos usar la función raíz cuadrada. Sea $f : R_+ \to R$ dada por $f(x) = \sqrt{x}$ con $x^1, x^2 \geqslant 0$ y $t \in [0,1]$, ya que se cumple la siguiente desigualdad

$$\begin{aligned} f\left((1-t)\,x^1 + tx^2\right) &\geqslant (1-t)\,f(x^1) + tf(x^2) \\ \sqrt{(1-t)\,x^1 + tx^2} &\geqslant (1-t)\sqrt{x^1} + t\sqrt{x^2} \end{aligned}$$

puesto que

$$t\,(1-t)\left[\sqrt{x^1} + t\sqrt{x^2}\right]^2 \geqslant 0 \text{ para todo } t \in [0,1]\,.$$

A continuación demostraremos la desigualdad anterior. Si

$$\begin{aligned} \sqrt{(1-t)\,x^1 + tx^2} &\geqslant (1-t)\sqrt{x^1} + t\sqrt{x^2} \\ \left(\sqrt{(1-t)\,x^1 + tx^2}\right)^2 &\geqslant \left((1-t)\sqrt{x^1} + t\sqrt{x^2}\right)^2 \\ (1-t)\,x^1 + tx^2 &\geqslant (1-t)^2\,x^1 + 2t(1-t)\sqrt{x^1x^2} + t^2x^2. \end{aligned}$$

Desarrollando la expresión anterior se tiene

$$x^1 - tx^1 + tx^2 \geqslant x^1 - 2tx^1 + t^2x^1 + 2t\sqrt{x^1x^2} - 2t^2\sqrt{x^1x^2} + t^2x^2;$$

es decir,

$$x^1 - tx^1 + tx^2 - x^1 + 2tx^1 - t^2x^1 - 2t\sqrt{x^1x^2} + 2t^2\sqrt{x^1x^2} - t^2x^2 \geqslant 0.$$

Eliminando y agrupando términos semejantes

$$tx^1 - t^2x^1 + tx^2 - t^2x^2 - 2t\sqrt{x^1x^2} + t^2\sqrt{x^1x^2} \geqslant 0$$

factorizando $t - t^2$

$$\begin{aligned} \left(t-t^2\right)x^1 + \left(t-t^2\right)x^2 - \left(t-t^2\right)2\sqrt{x^1x^2} &\geqslant 0 \\ \left(t-t^2\right)\left(x^1 - 2\sqrt{x^1x^2} + x^2\right) &\geqslant 0 \end{aligned}$$

por lo tanto

$$t\,(1-t)\left[\sqrt{x^1} + t\sqrt{x^2}\right]^2 \geqslant 0.$$

Teorema 16 **(i)** *Sea* $X \subset R^n$ *un conjunto convexo y* $f : X \to R$ *una función cóncava. Entonces para todo* $y \in R$ *se tiene que el conjunto*

$$S_y = \{x \in X \ \ \textit{tal que } f(x) \geqq y\} \textit{ es convexo.}$$

(ii) *Si* $f_1, f_2, ..., f_m : X \to R$ *son cóncavas y* $\lambda_1, \lambda_2, ..., \lambda_m$ *son reales no negativos, entonces*

$$f = \sum_{i=1}^{m} \lambda_i\, f_i \quad \textit{es también cóncava.}$$

(iii) *Si* $f : X \to R$ *es una función cóncava y* $x \in Int(X)$, *entonces* f *es continua en* x.

Demostración. (i) Sea $y \in R$ y sean $x^1, x^2 \in S_y$. Consideremos $t \in [0,1]$ y por hipótesis f es cóncava, es decir

$$f\ \left((1-t)\,x^1 + tx^2\right) \geqq (1-t)\,f\ (x^1) + tf\ (x^2) \geqq (1-t)y + ty = y$$

lo anterior por estar x^1 y x^2 en S_y, de donde se deduce

$$f\ \left((1-t)\,x^1 + tx^2\right) \geqq y$$

por tanto

$$(1-t)\,x^1 + tx^2 \in S_y$$

de donde se observa que S_y es convexo. ■

Demostración. (ii) Sean $x^1, x^2 \in X$ y $t \in [0,1]$ entonces se tiene

$$f\ \left[(1-t)\,x^1 + tx^2\right] = \sum_{i=1}^{m} \lambda_i\ f_i\left[(1-t)\,x^1 + tx^2\right] \tag{5.2}$$

ahora por hipótesis se tiene que por ser $f_1, f_2, ..., f_m$ funciones cóncavas

$$\sum_{i=1}^{m} \lambda_i\ f_i\left[(1-t)\,x^1 + tx^2\right] \geqq \sum_{i=1}^{m} \lambda_i\left[(1-t)\,f_i(x^1) + tf_i(x^2)\right]$$

donde

$$\begin{aligned} \sum_{i=1}^{m} \lambda_i\left[(1-t)\,f_i(x^1) + tf_i(x^2)\right] &= (1-t)\sum_{i=1}^{m} \lambda_i\ f_i(x^1) + t\sum_{i=1}^{m} \lambda_i\ f_i(x^2) \\ &= (1-t)f\ (x^1) + tf\ (x^2) \end{aligned} \tag{5.3}$$

y de las ecuaciones anteriores 5.2 y 5.3 se deduce que

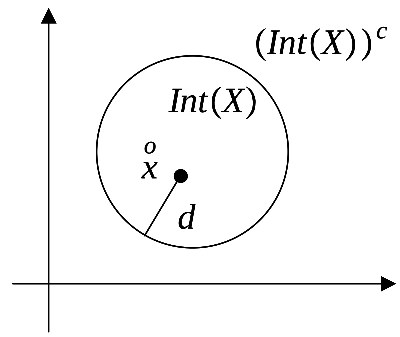
$$f\ \left((1-t)\,x^1 + tx^2\right) \geqq (1-t)f\ (x^1) + tf\ (x^2).$$

Por lo tanto queda demostrado que si para alguna i, se tiene que $\lambda_i > 0$ y f_i es *estrictamente cóncava*, entonces claramente, f es también *estrictamente cóncava.* ■

Demostración. (iii) Sea $\overset{o}{x} \in Int(X)$ donde $Int(X)$ es el abierto de X y definimos

$$d = \text{distancia de } \overset{o}{x} \text{ a } \left(Int(X)\right)^c, \tag{5.4}$$

donde $\left(Int(X)\right)^c$ es el complemento del conjunto $Int(X)$ el cual es un conjunto cerrado, véase la siguiente gráfica

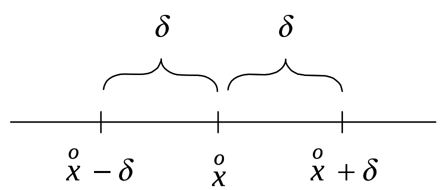


obsérvese que si $Int(X) = R^n = X$, entonces $d = \infty$. Sí $0 < \delta < \frac{d}{\sqrt{n}}$, con δ un número positivo arbitrario, hagamos

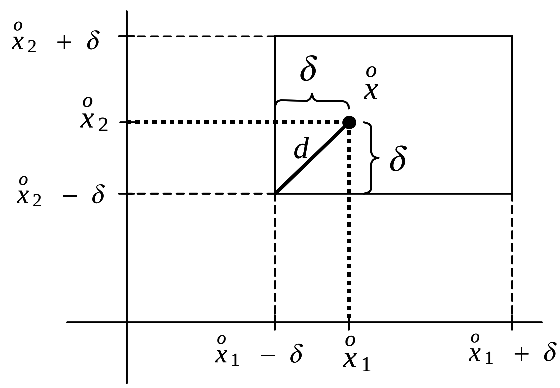
$$C = \prod_{i=1}^{n} \left[\overset{o}{x}_i - \delta, \overset{o}{x}_i + \delta \right]$$

donde C es el producto cartesiano.

Si $i = 1$



Si $i = 1, 2$



Obsérvese de la gráfica anterior que la distancia de $\overset{o}{x}$ a cualquier vértice esta dada por

$$\sqrt{2\delta^2} = \delta\sqrt{2} < \frac{d\sqrt{2}}{\sqrt{2}} = d.$$

Entonces, sea V el conjunto de vértices de C.Puesto que V es finito, podemos definir

$$M = \max_{x \in V} \{-f(x)\}$$

y sea

$$X_M = \{x \in X \text{ tal que } -f(x) \leqq M\}$$

donde X_M es convexo en virtud del inciso (i) del presente teorema, es decir $f(x) \geqq -M$ con $y = -M$.

Puesto que C es la cerradura convexa de V, esto es el menor convexo de V y $V \subset X_M$, entonces $C \subset X_M$. Consideremos ahora $x \in X$ tal que

$$0 < \left| x - \overset{o}{x} \right| < \delta$$

de esta expresión se tiene que

$$0 < \frac{\left\| x - \overset{o}{x} \right\|}{\delta} < 1$$

ya que por definición

$$\begin{aligned} \left| x - \overset{o}{x} \right| &= \left\| x - \overset{o}{x} \right\| \\ &= \sqrt{\left(x - \overset{o}{x} \right) \left(x - \overset{o}{x} \right)} \\ &= \sqrt{\left(x_1 - \overset{o}{x}_1 \right)^2 + \left(x_2 - \overset{o}{x}_2 \right)^2 + \ldots + \left(x_n - \overset{o}{x}_n \right)^2}. \end{aligned}$$

Hagamos ahora

$$\mu = \frac{x - \overset{o}{x}}{\left| x - \overset{o}{x} \right|} \delta, \tag{5.5}$$

despejando x de la ecuación 5.5 se tiene

$$x = \frac{\left\| x - \overset{o}{x} \right\| \mu}{\delta} + \overset{o}{x},$$

sumando un cero de manera conveniente del lado derecho de la última expresión, podemos escribir

$$\begin{aligned} x &= \frac{\left\| x - \overset{o}{x} \right\|}{\delta} \mu + \overset{o}{x} + \frac{\left\| x - \overset{o}{x} \right\|}{\delta} \overset{o}{x} - \frac{\left\| x - \overset{o}{x} \right\|}{\delta} \overset{o}{x} \\ &= \frac{\left\| x - \overset{o}{x} \right\|}{\delta} \mu + \frac{\left\| x - \overset{o}{x} \right\|}{\delta} \overset{o}{x} + \overset{o}{x} - \frac{\left\| x - \overset{o}{x} \right\|}{\delta} \overset{o}{x} \\ &= \frac{\left\| x - \overset{o}{x} \right\|}{\delta} (\mu + \overset{o}{x}) + (1 - \frac{\left\| x - \overset{o}{x} \right\|}{\delta}) \overset{o}{x}; \end{aligned} \tag{5.6}$$

por otro lado despejando a $\overset{o}{x}$ de 5.5 se obtiene

$$\overset{o}{x} = -\frac{\left\| x - \overset{o}{x} \right\|}{\delta} \mu + x,$$

sumando de ambos lados $\frac{\left\|x-\overset{o}{x}\right\|}{\delta}\overset{o}{x}$

$$\frac{\left\|x-\overset{o}{x}\right\|}{\delta}\overset{o}{x}+\overset{o}{x}=\frac{\left\|x-\overset{o}{x}\right\|}{\delta}-\frac{\left\|x-\overset{o}{x}\right\|}{\delta}\mu+x.$$

Lo anterior lo podemos expresar como

$$\left(\frac{\left\|x-\overset{o}{x}\right\|}{\delta}+1\right)\overset{o}{x}=\frac{\left\|x-\overset{o}{x}\right\|}{\delta}(\overset{o}{x}-\mu)+x,$$

despejando a $\overset{o}{x}$ se tiene que

$$\overset{o}{x}=\frac{\frac{\left\|x-\overset{o}{x}\right\|}{\delta}}{\frac{\left\|x-\overset{o}{x}\right\|}{\delta}+1}(\overset{o}{x}-\mu)+\frac{1}{\frac{\left\|x-\overset{o}{x}\right\|}{\delta}}x. \tag{5.7}$$

Luego mediante la ecuación 5.5, x ha sido expresado como combinación convexa de $\overset{o}{x}-\mu$ y $\overset{o}{x}$ y mediante la ecuación 5.6 $\overset{o}{x}$ ha sido expresado como combinación convexa de $\overset{o}{x}-\mu$ y x.Ahora puesto que por hipótesis f es cóncava se tiene

$$f(x)\geqq tf(\overset{o}{x}-\mu)+(1-t)f(\overset{o}{x}) \tag{5.8}$$

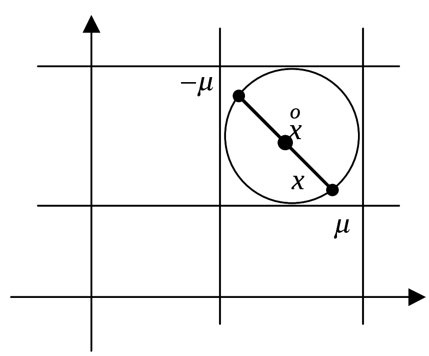
y

$$f(\overset{o}{x})\geqq\frac{t}{1+t}f(\overset{o}{x}-\mu)+\frac{1}{1+t}f(x)=\tau f(\overset{o}{x}-\mu)+(1-\tau)f(x) \tag{5.9}$$

en donde

$$t=\frac{\left\|x-\overset{o}{x}\right\|}{\delta}\in[0,1]\ \text{ y } \tau=\frac{t}{1+t}.$$

Gráficamente se puede ver



donde x se expresa como combinación convexa de $\overset{o}{x}$ y μ, y $\overset{o}{x}$ se expresa como combinación convexa de $-\mu$ y x.

Ahora de la ecuación 5.8 se tiene

$$\begin{aligned}-f(x) &\leqq -t\left(f\left(\overset{o}{x}+\mu\right)\right)-(1-t)\left(f(\overset{o}{x})\right)\\ &\leqq t\left(-f\left(\overset{o}{x}+\mu\right)\right)+(1-t)\left(-f(\overset{o}{x})\right),\end{aligned}$$

en donde se cumple la siguiente desigualdad por la definición de M y por estar x en X_M

$$t\left(-f\left(\overset{o}{x}+\mu\right)\right)+(1-t)\left(-f(\overset{o}{x})\right)\leqq tM+(1-t)(-f(\overset{o}{x}))$$

lo cual implica

$$-f(x)\leq tM-f(\overset{o}{x})+tf(\overset{o}{x}),$$

es decir

$$f(\overset{o}{x})-f(x)\leq t\left(M+f(\overset{o}{x})\right). \tag{5.10}$$

Ahora de la ecuación 5.9 se tiene

$$-f(\overset{o}{x})\leqq -\frac{t}{1+t}f\left(\overset{o}{x}-\mu\right)-\frac{1}{1+t}f(x),$$

de donde de manera análoga por estar $\overset{o}{x}$ en X_M y por la definición de M se cumple

$$-\frac{t}{1+t}f\left(\overset{o}{x}-\mu\right)-\frac{1}{1+t}f(x)\leqq\frac{t}{1-t}M-\frac{1}{1+t}f(x);$$

es decir se tiene de las dos últimas ecuaciónes

$$-f(\overset{o}{x})\leqq\frac{t}{1-t}M-\frac{1}{1+t}f(x)$$

y multiplicando por menos uno

$$f(\overset{o}{x})\geqq-\frac{t}{1-t}M+\frac{1}{1+t}f(x)$$

restando $f(x)$ de ambos lados de la desigualdad y agrupando de manera conveniente se obtiene

$$\begin{aligned} f(\overset{o}{x})-f(x) &\geqq -\frac{t}{1-t}M+\left(\frac{1}{1+t}-1\right)f(x)\\ &\geqq -\frac{t}{1-t}M-\frac{t}{1-t}f(x)\\ &= -\frac{t}{1-t}\left(M+f(x)\right), \end{aligned}$$

de donde se tiene

$$\begin{aligned} (1-t)\left(f(\overset{o}{x})-f(x)\right) &\geqq -t\left(M+f(x)\right)\\ f(\overset{o}{x})-f(x)+tf(\overset{o}{x})-tf(x) &\geqq -t\left(M+f(x)\right); \end{aligned}$$

por lo tanto

$$f(\overset{o}{x})-f(x)\geqq -t\left(M+f(x)\right) \tag{5.11}$$

y utilizando las ecuaciones 5.10 y 5.11 se deduce

$$\left|f(\overset{o}{x})-f(x)\right|\leqq t\left(M+f(x)\right)=\frac{M+f(\overset{o}{x})}{\delta}\left\|x-\overset{o}{x}\right\|.$$

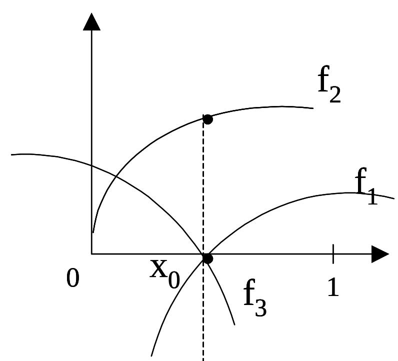
Esta desigualdad prueba que f es continua en $\overset{o}{x}$ por lo tanto el teorema queda probado. ■

Teorema 17 (Fundamental de las funciones cóncavas) *Sea $X \subset R^n$ un conjunto convexo no vacío y sean $f_1, f_2, ..., f_m : X \to R$ funciones cóncavas. Si el sistema $f_j(x) > 0$ para $j = 1, 2, ..., m$ no admite solución para $x \in X$, entonces existen $p_1, p_2, ..., p_m$ reales no negativos y que no se anulan simultáneamente tales que*

$$\sum_{j=1}^{m} p_j \, f_j(x) \leqq 0 \qquad \forall x \in X.$$

Observación 2 *Si el sistema $f_j(x) > 0$ para $\forall\ j = 1, 2, .., m$ no admite solución para $x \in X$, entonces tenemos que para todo x en el dominio, existe $f_j(x)$, tal que $f_j(x) \leqq 0$.*

Para dar una idea de la hipótesis en el teorema anterior, supóngase que $X = [0, 1]$, $n = 1$ y $m = 3$ entonces se tiene que $f_1,\ f_2,\ f_3 : X \to R$ son funciones cóncavas. Si $f_j(x) > 0$ para $j = 1, 2, 3$ no admite solución para $x \in X$, es decir no existe un punto del dominio x_0 tal que al evaluarlo, cumpla que $f_1(x_0) > 0$, $f_2(x_0) > 0$ y $f_3(x_0) > 0$, como se muestra en la siguiente gráfica.

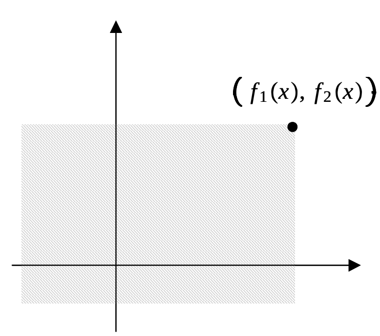


Demostración. Para todo $x \in X$, hagamos

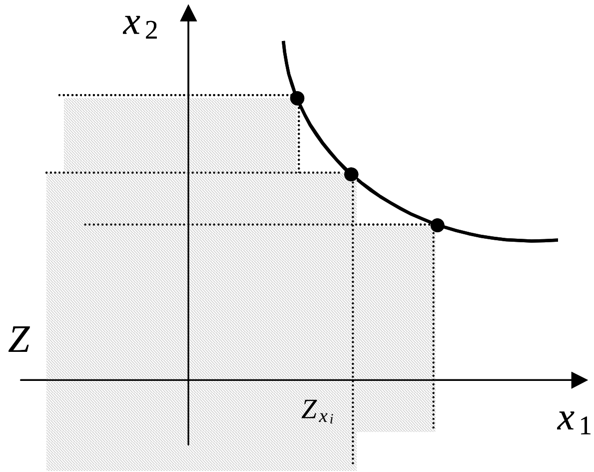
$$Z_x = \{(z_1, z_2, ..., z_m) \in R^m \quad \text{tales que} \quad z_j < f_j(x)\ \forall j = 1, 2, ...m\}.$$

Esbozaremos una idea geométrica para Z_x en el caso en que $m = 2$ en donde se tiene

$$Z_x = \left\{(z_1, z_2) \in R^2 \quad \text{tales que} \quad z_j < f_j(x)\ \forall j = 1, 2\right\}$$



y para la función gráficamente se tiene:



Definamos entonces

$$Z = \bigcup_{x \in X} Z_x.$$

Sabemos que $\overline{0} \notin Z$ en virtud de nuestra hipótesis. Supongamos que $\overline{0} \in Z$; esto implica que existe $x_0 \in X$ tal que $\overline{0} \in Z_{x_0} \Longrightarrow 0 < f_j(x_0) \, \forall j = 1, 2, ..., m,$ pero esto contradice la hipótesis. Sean z y $z' \in Z$ y además sea $t \in [0,1]$. Supongamos entonces que $z \in Z_x$ y que $z' \in Z_{x'}$. Consideremos ahora

$$\begin{aligned} z'' &= (1-t)z + tz' \quad \text{y} \\ x'' &= (1-t)x + tx' \end{aligned}$$

es decir z'' es una combinación convexa de z y z', tomemos ahora la coordenada j-ésima de z'' entonces

$$\begin{aligned} z''_j &= (1-t)z_j + tz'_j \\ &< (1-t)f_j(x) + tf_j(x') \\ &\leqq f_j\left((1-t)\,x + tx'\right) \end{aligned}$$

como se definió el conjunto y por ser f una función cóncava, por lo tanto

$$z''_j \in Z_{[(1-t)x+tx']} \subset Z$$

ya que

$$(1-t)\,x + tx' \in X.$$

Hemos probado entonces que Z es convexo. Entonces en virtud del teorema 14 y aplicado a Z y $\overline{0} \notin Z$ se tiene que existe $\widetilde{P} \in R^m$ con $\widetilde{P} \neq \overline{0}$ tal que

$$\widetilde{P} \cdot z \geqq \widetilde{P} \cdot \overline{0} = 0 \qquad \forall z \in Z$$

es decir $\widetilde{P} \cdot z \geqq 0$,esto es

$$\sum_{j=1}^{m} \widetilde{p}_j z_j \geqq 0$$

Sea $j_o = 1, 2, ..., m$ fijo y $z \in Z$ y suponemos que $\widetilde{p}_{j_0} > 0$, si tomamos

$$Z(M) = (z_1, z_2, ..., z_{j_0} - M, ..., z_m).$$

Para M real positivo arbitrario, se cumplirá que $Z(M) \in Z$ ya que

$$Z(M) < Z_j < f_j(x) \ \forall j = 1, 2, ..., m$$

y por el teorema 14, tenemos que

$$0 \leq \widetilde{P} \cdot Z(M) = \sum_{j=1}^{m} \widetilde{p}_j z_j(M) = \sum_{j=1}^{m} \widetilde{p}_j z_j - \widetilde{p}_{j_0} M$$

por lo tanto

$$\sum_{j=1}^{m} \widetilde{p}_j z_j \geqq \widetilde{p}_{j_0} M$$

para todo M positivo arbitrario.

Esto nos lleva a un absurdo, de suponer que $\widetilde{p}_{j_0} > 0$. Entonces $\widetilde{p}_{j_0} \leqq 0 \ \forall \ j_0 = 1, 2, ..., m$. Hagamos ahora $P = -\widetilde{P} \geqq \overline{0}$. Luego se tiene que

$$P \cdot z \leqq 0 \quad \forall z \in Z \tag{5.12}$$

Ahora sea $\varepsilon > 0$ y $x \in X$ definimos

$$Z_\varepsilon = (f_1(x) - \varepsilon \ , \ f_2(x) - \varepsilon \ , ..., \ f_m(x) - \varepsilon) \in Z.$$

Ya que $f_j(x) - \varepsilon < f_j(x) \quad \forall j = 1, 2, ..., m$. Ahora utilizando nuevamente el teorema 14 y la desigualdad 5.12, sustituyendo Z_ε se tiene que

$$P \cdot Z_\varepsilon = \sum_{j=1}^{m} p_j \left(f_j(x) - \varepsilon \right) \leqq 0$$

entonces

$$\sum_{j=1}^{m} p_j \ f_j(x) - \varepsilon \sum_{j=1}^{m} p_j \leqq 0$$

de donde

$$\sum_{j=1}^{m} p_j \ f_j(x) \leqq \varepsilon \sum_{j=1}^{m} p_j$$

luego, haciendo $\varepsilon \to 0$ tenemos que

$$\sum_{j=1}^{m} p_j \ f_j(x) \leqq 0 \quad \forall x \in X.$$

Por lo tanto el teorema queda demostrado. ■

Suponiendo ahora que las funciones son convexas, tenemos un resultado análogo.

Corolario 1 *Sea $X \subset R^n$ un conjunto convexo y sean $f_1, f_2, ..., f_m : X \to R$ funciones convexas entonces o bien el sistema*

$$f_j(x) < 0 \quad para \ \ j = 1, 2, ..., m$$

admite solución en $x \in X$ o existen $p_1, p_2, ..., p_m$ no negativos y que no se anulan simultáneamente tales que

$$\sum_{j=1}^{m} p_j \ f_j(x) \geqq 0 \quad \forall x \in X.$$

Teorema 18 (Kuhn-Tucker-Uzawa) *Sean $f, g_1, g_2, ..., g_m : X \to R$ funciones cóncavas en $X \subset R^n$ un conjunto convexo. Supongamos que f alcanza un máximo en x^* sobre X sujeto a $g_j(x) \geqq 0 \ \ \forall \ j = 1, 2, ..., m; \ \ x \in X$. Entonces existen coeficientes $p_0^*, p_1^*, ..., p_m^*$ no negativos y que no se anulan simultáneamente, tal que cumplen lo siguiente:*

$$p_0^* \ f(x) + \sum_{j=1}^{m} p_j \ g_j(x) \leqq p_0^* \ f(x^*) \qquad \forall x \in X.$$

Además

$$\sum_{j=1}^{m} p_j^* \ g_j(x^*) = 0.$$

Demostración. Por hipótesis, el siguiente sistema no admite solución en X

$$\begin{aligned} g_j(x) &\geqq 0 \quad \text{para} \ \ j = 1, 2, ..., m \\ f(x) - f(x^*) &> 0 \end{aligned}$$

ya que por hipótesis x^* es máximo de f entonces se tiene que

$$f(x) \leqq f(x^*) \ \ \forall x \in X$$

es decir

$$f(x) - f(x^*) \leqq 0.$$

Ahora bien, por lo anterior el sistema

$$\begin{aligned} g_j(x) &> 0 \quad \text{para} \ \ j = 1, 2, ..., m \\ f(x) - f(x^*) &> 0 \end{aligned}$$

tampoco admite solución en X. Por el argumento anterior y por el teorema 17 existen números $p_0^*, p_1^*, p_2^*, ..., p_m^*$ reales no negativos y que no se anulan simultáneamente tales que

$$p_0^* \left(f(x) - f(x^*)\right) + \sum_{j=1}^{m} p_j^* \ g_j(x) \leqq 0 \ \ \forall x \in X,$$

es decir

$$p_0^* f(x) - p_0^* f(x^*) + \sum_{j=1}^{m} p_j^* g_j(x) \leqq 0 \ \ \forall x \in X,$$

de donde

$$p_0^* \ f(x) + \sum_{j=1}^{m} p_j^* \ g_j(x) \leqq p_0^* \ f(x^*) \qquad \forall x \in X$$

por lo que se demuestra la primera afirmación.

Ahora la desigualdad anterior se cumple para todo $x \in X$, en particular dado que $x^* \in X$ se tiene

$$p_0^* \ f(x^*) + \sum_{j=1}^{m} p_j^* \ g_j(x^*) \leqq p_0^* \ f(x^*)$$

de donde

$$\sum_{j=1}^{m} p_j^* \ g_j(x^*) \leqq p_0^* \ f(x^*) - p_0^* \ f(x^*)$$

entonces se tiene

$$\sum_{j=1}^{m} p_j^* \ g_j(x^*) \leqq 0,$$

pero por hipótesis $g_j(x) \geqq 0 \ \ \forall j = 1, 2, ..., m$; en particular para x^* se tiene $g_j(x^*) \geqq 0 \ \forall$ $j = 1, 2, ..., m$, además por el teorema 17 sabemos que $p_j^* \geqq 0 \ \ \forall j = 1, 2, ..., m$ que no se anulan simultáneamente, entonces la única posibilidad es que

$$\sum_{j=1}^{m} p_j^* \ g_j(x^*) = 0.$$

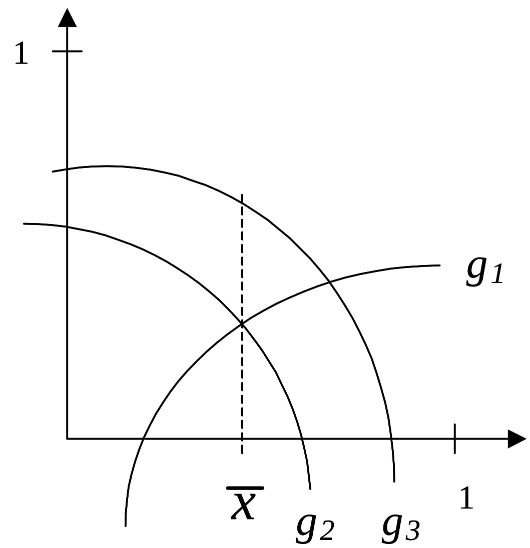
Por lo tanto nuestro teorema queda probado. ■

Condición 1 (Slater) *Sean $g_1, g_2, ..., g_m : X \to R$ cóncavas en donde $X \subset R^n$, entonces existe $\overline{x} \in X$ tal que $g_j(\overline{x}) > 0$ para todo $j = 1, 2, .., m$.*

Ejemplo 6 *El siguiente es un ejemplo gráfico para el caso en el que $m = 3$ y $n = 1$. Sea*

$$X = (0, 1) \subset R$$

y $g_j(x) > 0$ con $j = 1, 2, 3$ cóncavas.



Condición 2 (Punto Silla) *Sea L la función Lagrangiana del problema estándar de programación no lineal, dada por:*

$$L(x,\lambda) = f(x) + \sum_{j=1}^{m} \lambda_j \; g_j(x) \quad \textit{para} \quad x \in X, \lambda \in R_+^m.$$

Se dice que se cumple la condición punto silla si existen

$$(x^*, \lambda^*) \in X \times R_+^m$$

tales que

$$L(x,\lambda^*) \leqq L(x^*,\lambda^*) \leqq L(x^*,\lambda) \qquad \forall \;\; (x,\lambda) \in X \times R_+^m$$

Corolario 2 (Del Teorema de Kuhn-Tucker-Uzawa) *Si además de las hipótesis del teorema de Kuhn-Tucker-Uzawa se cumple la condición de Slater, entonces se cumple la condición punto silla.*

Demostración. Sea x^* máximo de f como en la demostración del teorema 18 y sea $\overline{x}$ como en la condición 1. Entonces

$$p_0^* \; f(\overline{x}) + \sum_{j=1}^{m} p_j^* \; g_j(\overline{x}) \leqq p_0^* \; f(x^*) \qquad \forall x \in X$$

si suponemos que $p_0^* = 0$

$$\sum_{j=1}^{m} p_j^* g_j(\overline{x}) \leqq 0. \tag{5.13}$$

Pero existe al menos un $j = 1, 2, .., m$ tal que $p_j^* > 0$ y además $g_j(\overline{x}) > 0$ para todo $j = 1, 2, .., m$ entonces la ecuación 5.13, es un absurdo. Luego $p_o^* > 0$.

Hagamos ahora para $j = 1, 2, ..., m$

$$\lambda_j^* = \frac{p_j^*}{p_o^*}$$

donde

$$\lambda^* = (\lambda_1^*, \lambda_2^*, ..., \lambda_m^*) \geqq 0$$

entonces $\lambda^* \in R_+^m$, además sea $x \in X$ y x^* como en el teorema de Kunh-Tucker-Uzawa se tiene

$$\begin{aligned} L(x, \lambda^*) &= f(x) + \sum_{j=1}^{m} \lambda_j^* g_j(x) \\ &\leqq f(x^*) + \sum_{j=1}^{m} \lambda_j^* g_j(x^*) \\ &= L(x^*, \lambda^*). \end{aligned}$$

Ahora por el teorema 18 se tiene

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

y sabemos que $\lambda_j \geqq 0$ y que $g_j(x^*) \geqq 0$ lo que implica $\sum_{j=1}^{m} \lambda_j g_j(x^*) \geqq 0$ por lo que se deduce

$$\sum_{j=1}^{m} \lambda_j^* g_j(x^*) \leqq \sum_{j=1}^{m} \lambda_j g_j(x^*) \quad \text{para todo } \lambda \in R_+^m$$

entonces sumando de ambos lados $f(x^*)$ se tiene

$$f(x^*) + \sum_{j=1}^{m} \lambda_j^* g_j(x^*) \leqq f(x^*) + \sum_{j=1}^{m} \lambda_j g_j(x^*) \quad \text{para todo } \lambda \in R_+^m$$

es decir

$$L(x^*, \lambda^*) \leqq L(x^*, \lambda) \quad \text{para todo } \lambda \in R_+^m.$$

Por lo tanto esto completa la prueba. ■

Recordemos el problema estándar de programación no lineal, que definimos al principio del capítulo (ver página 84).

$$\begin{aligned} & \text{máx } f(x) \\ \text{sujeto a } g_i(x) &\geqq 0 \\ \text{para } i &= 1, 2, .., m \end{aligned}$$

En donde $f, g_1, g_2, ..., g_m : X \to R$ son funciones cóncavas definidas en el convexo $X \subset R^n$.

Condición 3 (Karlin) *Decimos que el problema 3 cumple la condición de Karlin, si y sólo si para cada* $P \geq \overline{0}$, *con* $P \in R^m$, *existe* $\overline{x} \in X$, *tal que*

$$\sum_{j=1}^{m} p_j g_j(\overline{x}) > 0.$$

Proposición 4 *La condición de Karlin implica que* $p_0^* > 0$ *en el teorema 18.*

Demostración. Para demostrar este hecho supongamos que $p_0^* = 0$ y tomemos en particular $P^* = (p_1^*, p_2^*, ..., p_m^*) \geq \overline{0}$ con $P^* \in R^m$. Como el problema cumple la condición de Karlin se tiene que existe $\overline{x}$ tal que

$$\sum_{j=1}^{m} p_j^* g_j(\overline{x}) > 0,$$

entonces en el teorema 18

$$p_0^* \ f(\overline{x}) + \sum_{j=1}^{m} p_j^* \ g_j(\overline{x}) \leqq p_0^* \ f(x^*)$$

por hipótesis $p_0^* = 0$ es decir

$$\sum_{j=1}^{m} p_j^* \ g_j(\overline{x}) \leqq 0$$

lo cual es un absurdo ya que también por hipótesis

$$\sum_{j=1}^{m} p_j^* g_j(\overline{x}) > 0;$$

por lo tanto $p_0^* > 0$. ■

A continuación demostraremos un resultado relevante en términos de la condición punto silla y de la condición de Karlin antes mencionadas.

Proposición 5 *La condición de Karlin implica la condición punto silla.*

Demostración. De la anterior demostración sabemos que $p_0^* > 0$ en el teorema de Kuhn-Tucker-Uzawa. Ahora consideremos $\lambda^* \geq \overline{0}$ con $\lambda \in R_+^m$ y $\lambda_j^* = \dfrac{p_j^*}{p_0}$ para $j = 1, 2, ..., m$ por lo que si construimos la función Lagrangiana de nuestro problema estándar tenemos

$$L\left(x, \lambda^*\right) = f\left(x\right) + \sum_{j=1}^{m} \lambda_j^* g_j(x) = f\left(x\right) + \sum_{j=1}^{m} \frac{p_j^*}{p_0} g_j(x).$$

Ahora del teorema 18 se cumple

$$p_0^* f(x) + \sum_{j=1}^{m} p_j^* g_j(x) \leqq p_0^* f(x^*) \text{ para toda } x \in X,$$

dividiendo la desigualdad anterior entre p_0^* se tiene

$$f(x) + \sum_{j=1}^{m} \frac{p_j^*}{p_0^*} g_j(x) \leqq f(x^*).$$

Entonces regresando a la función Lagrangiana del problema se tiene

$$L(x, \lambda^*) = f(x) + \sum_{j=1}^{m} \lambda_j^* g_j(x) = f(x) + \sum_{j=1}^{m} \frac{p_j^*}{p_0} g_j(x) \leqq f(x^*), \tag{5.14}$$

ahora nuevamente ocupando el teorema 18

$$\sum_{j=1}^{m} p_j^* g_j(x^*) = 0,$$

dividiendo ambos lados entre p_0^*

$$\sum_{j=1}^{m} \frac{p_j^*}{p_0^*} g_j(x^*) = 0,$$

utilizando la última expresión podemos escribir

$$f(x^*) + 0 = f(x^*) + \sum_{j=1}^{m} \frac{p_j^*}{p_0} g_j(x^*) \tag{5.15}$$

y por lo tanto de las ecuaciones 5.14 y 5.15 se tiene que

$$L(x, \lambda^*) \leqq f(x^*) + \sum_{j=1}^{m} \frac{p_j^*}{p_0} g_j(x^*) = L(x^*, \lambda^*), \tag{5.16}$$

lo cual implica la primera desigualdad de la condición punto silla.

Para demostrar la otra desigualdad sabemos

$$L(x^*, \lambda^*) = f(x^*) \leqq f(x^*) + \sum_{j=1}^{m} \lambda_j^* g_j(x) = L(x^*, \lambda), \tag{5.17}$$

lo anterior se sabe ya que $\sum_{j=1}^{m} \lambda_j^* g_j(x)$ es una suma de un producto de no negativos por la condición de Karlin, por lo que resumiendo las ecuaciones 5.16 y 5.17 llegamos a la siguiente expresión:

$$L(x, \lambda^*) \leqq L(x^*, \lambda^*) \leqq L(x^*, \lambda).$$

Por lo que queda demostrada nuestra proposición. ■

Citaremos nuevamente la condición de Slater (página 95), la cual tiene una implicación importante con lo demostrado ya que la condición de Slater es más débil que la condición de Karlin. En otras palabras se tiene la siguiente proposición.

Proposición 6 *La condición de Slater implica la condición de Karlin.*

Demostración. Supongamos que existe $\overline{x} \in X$ tal que $g_j(\overline{x}) > 0$ y sea $P \in R^m_+$ tal que $P \geq \overline{0}$, consideremos

$$\sum_{j=1}^{m} p_j g_j(\overline{x}) = p_1 g_1(\overline{x}) + p_2 g_2(\overline{x}) + ... + p_m g_m(\overline{x}),$$

que es una suma de términos no negativos y más aún podemos asegurar que alguno de los términos es estrictamente positivo ya que por hipótesis $g_j(\overline{x}) > 0$ y además $p_j > 0$ para toda $j = 1, 2, ..., m$ donde no todos los p_j se anulan simultaneamente, de donde podemos asegurar que

$$\sum_{j=1}^{m} p_j g_j(\overline{x}) > 0,$$

lo que demuestra nuestra proposición. ■

Por lo demostrado en las proposiciones 5 y 6 se deduce que la condición de Slater implica la condición punto silla, (ver segundo diagrama en la introducción del presente capítulo).

Ahora enunciaremos el siguiente teorema y su demostración, el cual nos lleva a caracterizar las soluciones del problema estándar 3 (ver página 84).

Teorema 19 *Sean $f, g_1, g_2, ..., g_m : X \to R$ funciones cóncavas con $X \subset R^n$ un conjunto convexo. Si existe $(x^*, \lambda^*) \in X \times R^m_+$ tal que*

$$L(x, \lambda^*) \leqq L(x^*, \lambda^*) \leqq L(x^*, \lambda) \qquad \forall (x, \lambda) \in X \times R^m_+.$$

Entonces:

(i) x^* máximiza a f sujeto a $g_j(x) \geqq 0$ con $j = 1, 2, ..., m$ y $x \in X$

(ii) $\lambda^* \cdot g(x^*) = 0$. Esta condición implica que los vectores

$$\lambda^* = (\lambda^*_1, \lambda^*_2, ..., \lambda^*_m)$$

y

$$(g_1(x^*), g_2(x^*), ..., g_m(x^*))$$

son perpendiculares.

Demostración. Definamos

$$L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) \text{ para toda } (x, \lambda) \in X \times R^m_+$$

la función Lagrangiana del problema.

Ahora si utilizamos la segunda desigualdad de la hipótesis de nuestro teorema,

$$f(x^*) + \sum_{j=1}^{m} \lambda^*_j g_j(x^*) \leqq f(x^*) + \sum_{j=1}^{m} \lambda_j g_j(x^*) \text{ para toda } (x, \lambda) \in X \times R^m_+$$

es decir se tiene que:

$$\sum_{j=1}^{m} \lambda_j^* g_j(x^*) \leqq \sum_{j=1}^{m} \lambda_j g_j(x^*) \text{ para toda } \lambda \in R_+^m \tag{5.18}$$

observemos que tanto x^* como λ^* estan fijos, por lo que tenemos un escalar del lado izquierdo de la desigualdad, lo que nos dice que el conjunto $\lambda \cdot g_j(x^*)$ está acotado inferiormente, ya que la ecuación 5.18 se cumple para todo $\lambda \in R_+^m$.

Utilizando lo anterior demostraremos que

$$\sum_{j=1}^{m} \lambda_j g_j(x^*) \geqq 0 \text{ para toda } \lambda \in R_+^m$$

lo cual haremos por reducción al absurdo. Supongamos entonces que existe $\lambda^* \in R_+^m$ tal que cumple con

$$\sum_{j=1}^{m} \lambda_j^* g_j(x^*) < 0 \text{ para toda } \lambda \in R_+^m$$

y sea $\rho > 0$ un escalar, definamos $\lambda^*(\rho) = \rho\lambda^* = (\rho\lambda_1^*, \rho\lambda_2^*, ..., \rho\lambda_m^*) \in R_+^m$ entonces

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

lo cual contradice nuestra hipótesis de que el conjunto $\lambda \cdot g_j(x^*)$ está acotado inferiormente pues el producto $\rho \sum_{j=1}^{m} \lambda_j^* g_j(x^*)$ es un escalar negativo tan grande como se quiera, por lo que se cumple

$$\sum_{j=1}^{m} \lambda_j g_j(x^*) \geqq 0 \text{ para toda } \lambda \in R_+^m.$$

Sin pérdida de generalidad sea $\lambda = (1, 0, 0, ..., 0) \in R_+^m$. Calculemos ahora

$$\sum_{j=1}^{m} \lambda_j g_j(x^*) = g_1(x^*) + 0 + 0 + ... + 0 = g_1(x^*) \geqq 0$$

análogamente podemos hacerlo para todo $j \in R_+^m$, de tal forma que $g_j(x^*) \geqq 0$ para todo $j = 1, 2, ..., m$ por lo que se deduce que x^* satisface las condiciones del problema, por lo tanto es factible.

Como $g_j(x^*) \geqq 0$ y λ_j^* para todo $j = 1, 2, ..m$ son no negativos y sabemos que

$$0 \leqq \sum_{j=1}^{m} \lambda_j^* g_j(x^*) \tag{5.19}$$

por otro lado, si hacemos $\lambda = \overline{0} \in R_+^m$ y sustituyendo λ en la ecuación 5.18 tenemos que

$$\sum_{j=1}^{m} \lambda_j^* g_j(x^*) \leqq \sum_{j=1}^{m} 0 \cdot g_j(x^*) = 0,$$

es decir

$$\sum_{j=1}^{m} \lambda_j^* g_j(x^*) \leqq 0. \tag{5.20}$$

Por lo que de las ecuaciones 5.19 y 5.20 podemos decir que

$$\sum_{j=1}^{m} \lambda_j^* g_j(x^*) = \lambda^* g_j(x^*) = 0. \tag{5.21}$$

Esto demuestra el inciso (ii) del teorema y sólo falta demostrar que efectivamente x^* maximiza a f.

Para ello nuevamente utilizaremos la primera desigualdad de nuestra hipótesis

$$L\left(x, \lambda^*\right) \leqq L\left(x^*, \lambda^*\right),$$

es decir

$$f(x) + \sum_{j=1}^{m} \lambda_j^* g_j(x) \leqq f(x^*) + \sum_{j=1}^{m} \lambda_j^* g_j(x^*) \text{ para toda } (x, \lambda) \in X \times R_+^m$$

y por la ecuación 5.21

$$f(x) + \sum_{j=1}^{m} \lambda_j^* g_j(x) \leqq f(x^*),$$

lo cual podemos expresar como

$$\sum_{j=1}^{m} \lambda_j^* g_j(x) \leqq f(x^*) - f(x) \text{ para todo} (x, \lambda) \in X \times R_+^m. \tag{5.22}$$

Observemos que si x satisface las restricciones del problema se cumple que $g_j(x) \geqq 0$ para $j = 1, 2, ..., m$ y por ser $\lambda^* \geqq \overline{0}$, sabemos que

$$0 \leqq \sum_{j=1}^{m} \lambda_j^* g_j(x)$$

entonces de la ecuación 5.22 se tiene que

$$f(x^*) - f(x) \geqq 0,$$

es decir $f(x^*) \geqq f(x)$ para todo $x \in X$ por lo que x^* es máximo de f. Por lo que queda demostrado el inciso (i). ■

Una consecuencia del corolario 2 y el teorema 19 es el siguiente teorema.

Teorema 20 *Sean $f, g_1, g_2, .., g_m : X \to R$ funciones reales y cóncavas definidas sobre un convexo $X \subset R^n$. Supongamos que se satisface la condición de Slater. Entonces existe una solución x^* al problema estándar 3 y $x \in X$ si y sólo si existe un punto silla para $L(x, \lambda)$ con $(x, \lambda) \in X \times R_+^m$, en este caso existe un punto silla de la forma (x^*, λ^*) con $\lambda^* \in R_+^m$.*

Demostración. $(\to)$ Corolario 2 $\Longrightarrow$ Condición puto-silla 2 ■
Demostración. $(\leftarrow)$ Condición punto-silla 2 $\Longrightarrow$ Teorema 19. ■

La caracterización del casi-punto-silla

En la presente sección se expondrán los siguientes conceptos:

- La condición de maximalidad
- La condición punto-silla
- La condición casi-punto-silla

Los cuales nos permiten caracterizar las posibles soluciones de un problema de programación no lineal ahora con las siguientes condiciones:

Problema 4 *Sea $X \subset R^n$ un conjunto abierto y convexo*

$$\begin{array}{rcl} & \text{máx} & f(x) \\ s.\ a.\ g_j(x) & \geqq & 0 \\ para\ j & = & 1, 2, ..., m \\ con\ \ x & \in & X \end{array}$$

Condición 4 (Casi-Punto-Silla) *El problema 4, satisface la condición casi-punto-silla si y sólo si existe $(x^*, \lambda^*) \in X \times R_+^m$ tal que*

(i) $f'(x^*) + \lambda^* \cdot g'(x^*) = \dfrac{\partial f}{\partial x_i}(x^*) + \displaystyle\sum_{j=1}^{m} \lambda_j^* \dfrac{\partial g_j}{\partial x_i}(x^*) = 0$ para toda $i = 1, 2, ..., n$

(ii) $g(x^*) \geqq \overline{0}$

(iii) $\lambda^* \cdot g(x^*) = \displaystyle\sum_{j=1}^{m} \lambda_j^* g_j(x^*) = 0$

Continuemos ahora con el siguiente teorema.

Teorema 21 *Sean $f, g_1, g_2, .., g_m : X \to R$ diferenciables en $X \subset R^n$ un conjunto abierto.*

i) La condición punto-silla implica la condición casi-punto-silla

ii) Si además las funciones son cóncavas y X es un conjunto convexo, entonces la condición casi-punto-silla implica la condición punto-silla.

Demostración. **(i)** Sea

$$F(x) = L(x, \lambda^*) = f(x) + \sum_{j=1}^{m} \lambda_j^* g_j(x).$$

Por hipótesis se cumple que

$$F(x) = L(x, \lambda^*) \leqq L(x^*, \lambda^*) = F(x^*) \qquad \forall x \in X$$

por lo cual $F(x)$ alcanza su máximo en x^*, lo que implica que

$$F'(x^*) = 0$$

es decir

$$\frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) = 0,$$

de manera análoga, de la hipótesis y del teorema 19 se tiene que

$$\sum_{j=1}^{m} \lambda_j^* g_j(x^*) = 0,$$

y nuevamente utilizando la hipótesis sabemos que

$$L(x^*, \lambda^*) \leqq L(x^*, \lambda) \ \ \forall \lambda \in R_+^m,$$

entonces se tiene que

$$f(x^*) + \sum_{j=1}^{m} \lambda_j^* g_j(x^*) \leqq f(x^*) + \sum_{j=1}^{m} \lambda_j g_j(x^*),$$

es decir

$$\sum_{j=1}^{m} \lambda_j^* g_j(x^*) \leqq \sum_{j=1}^{m} \lambda_j g_j(x^*) \ \ \forall \lambda \in R_+^m.$$

Y por lo anterior sabemos que

$$0 \leqq \sum_{j=1}^{m} \lambda_j g_j(x^*) \forall \lambda \in R_+^m,$$

donde

$$\lambda \cdot g(x^*) = \sum_{j=1}^{m} \lambda_j \ g_j(x^*) \ \ \geqq 0 \ \text{ para todo } \lambda \in R_+^m.$$

Sin pérdida de generalidad, sea $\lambda = (1, 0, 0, ..., 0) \in R_+^m$. Calculemos ahora

$$\begin{aligned} \sum_{j=1}^{m} \lambda_j g_j(x^*) &= g_1(x^*) + 0 + 0 + ... + 0 \\ &= g_1(x^*) \geqq 0, \end{aligned}$$

análogamente podemos hacerlo para todo $j \in R_+^m$, de tal forma que $g_j(x^*) \geqq 0$ para todo $j = 1, 2, ..., m$ ya que x^* satisface las condiciones del problema, es decir $g(x^*) \geqq \overline{0}$. Por lo que quedan demostradas las condiciones casi-punto-silla. ■

Demostración. (ii) Para λ^* se tiene que

$$\begin{aligned} F(x) &= L(x,\lambda^*) \\ &= f(x) + \sum_{j=1}^{m} \lambda_j^* g_j(x), \end{aligned}$$

lo anterior nos dice que $F(x)$ es un combinación lineal no negativa de funciones cóncavas, es decir $F(x)$ es una función cóncava. Ahora utilizando la hipótesis de la condición casi-punto-silla se cumple que

$$\frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*)$$

es decir $F'(x^*) = 0$,de donde sabemos que $F(x)$ alcanza un máximo global en x^*, o dicho de otro modo se cumple que $F(x) \leqq F(x^*)$ esto es

$$L(x,\lambda^*) \leqq L(x^*,\lambda^*) \quad \text{para todo } x \in X.$$

Ahora nuevamente de la hipótesis tenemos que

$$\begin{aligned} L(x^*,\lambda^*) &= f(x^*) + \sum_{j=1}^{m} \lambda_j^* g_j(x^*) \\ &= f(x^*) + 0 \\ &= f(x^*); \end{aligned}$$

utilizando lo anterior y de saber que $\sum_{j=1}^{m} \lambda_j g_j(x)$ es una combinación lineal no negativa de funciones cóncavas, podemos plantear la siguiente desigualdad

$$f(x^*) \leqq f(x^*) + \sum_{j=1}^{m} \lambda_j g_j(x) = L(x^*,\lambda),$$

lo cual nos lleva a que

$$L(x^*,\lambda^*) \leqq L(x^*,\lambda) \text{ para todo } x \in X.$$

Por último, resumiendo todo lo anterior se deduce que

$$L(x,\lambda^*) \leqq L(x^*,\lambda^*) \leqq L(x^*,\lambda) \text{ para todo } \lambda \in R_+^m.$$

La cual es la condición punto silla que se quería demostrar en el inciso (ii). ■

Como un corolario del teorema 20 y del teorema 21 se presenta el siguiente teorema.

Teorema 22 *Sean $f, g_1, g_2, .., g_m : X \to R$ funciones reales, cóncavas y diferenciables en el conjunto $X \subset R^n$ abierto y convexo. Supongamos que se cumple la condición de Slater, entonces existe x^* que maximiza a $f(x)$ sujeto a las restricciones $g_j(x) \geqq 0$ y $x \in X$ si y sólo si se cumplen las condiciones casi-punto-silla.*

Demostración. $(\Rightarrow)$ Sean $f, g_1, g_2, .., g_m : X \to R$ funciones reales, cóncavas y diferenciables en el conjunto $X \subset R^n$ abierto y convexo. Supongamos que se cumple la condición de Slater, entonces existe x^* solución del problema 4, lo que implica por el teorema 20 que se cumple la condición punto-silla. Y por el inciso (i) del teorema 21 se cumple la condición casi-punto-silla. ■

Demostración. $(\Leftarrow)$ Sean $f, g_1, g_2, .., g_m : X \to R$ funciones reales, cóncavas y diferenciables en el conjunto $X \subset R^n$ abierto y convexo. Supongamos que cumple la condición casi-punto-silla, lo que implica por el inciso (ii) del teorema 21 que se cumple condición punto-silla; entonces por el regreso del teorema 20 existe x^* solución del problema 4. ■

Algunos conceptos de la programación casi cóncava

Muchos de los hechos demostrados anteriormente para el problema de programación no lineal en el caso de funciones cóncavas, se pueden generalizar para el caso de funciones casi cóncavas diferenciables, siempre que estemos restringidos al problema moderno de programación no lineal con condiciones de no negatividad. En está sección sólo se abordarán algunos conceptos elementales y algunas demostraciones relativas a la programación casi cóncava.

Definición 9 *Sea $f : X \to R$ donde $X \subset R^n$ convexo, es llamada una función casi-cóncava si $f(x_1) \geqq f(x_2)$, entonces*

$$f(tx_2 + (1-t)x_1) \geqq f(x_2)$$

para todo $x_1, x_2 \in X$ y $t \in [0,1]$.

Proposición 7 *Sea $f : X \to R$ donde $X \subset R^n$ convexo, entonces f es casi-cóncava si y sólo si para todo $y \in R$, el conjunto*

$$S_y = \{x \in X \text{ tal que } f(x) \geqq y\}$$

es convexo.

Demostración. $(\rightarrow)$ Supongamos que f es casi-cóncava y que $y \in R$, además que $x_1, x_2 \in S_y$ con $t \in [0,1]$. Se tienen dos casos, ya sea que

$$f(x_1) \geqq f(x_2) \text{ ó } f(x_1) < f(x_2).$$

Sin pérdida de generalidad, supongamos que ocurre el primer caso ya que el segundo sería totalmente análago. Entonces por ser X un conjunto convexo y de suponer que $f(x_1) \geqq f(x_2)$ tenemos que

$$f(tx_2 + (1-t)x_1) \geqq f(x_2),$$

pero por hipótesis $x_2 \in S_y$, es decir $f(x_2) \geqq y$ por lo que de la ecuación anterior se cumple

$$f(tx_2 + (1-t)x_1) \geqq y,$$

y

$$tx_2 + (1-t)\,x_1 \in S_y,$$

por lo tanto S_y es convexo. ∎

Demostración. $(\longleftarrow)$ Supongamos ahora que para todo $y \in S_y$, se tiene que S_y es convexo y sean $x_1, x_2 \in X$, tales que $f(x_1) \geqq f(x_2)$ y además $t \in [0,1]$. Entonces si $y = f(x_2)$, S_y es convexo y

$$tx_2 + (1-t)\,x_1 \in S_y$$

es decir, se cumple que

$$f(tx_2 + (1-t)\,x_1) \geqq y.$$

Pero por otro lado tenemos que $y = f(x_2)$ para alguna y, por lo que

$$f(tx_2 + (1-t)\,x_1) \geqq f(x_2),$$

es decir f es una función casi-cóncava. ∎

Definición 10 *Sea $f : X \to R$ donde $X \subset R^n$ convexo, decimos que f es estrictamente casi-cóncava, si para todo par $x_1, x_2 \in X$ tales que $f(x_1) \geqq f(x_2)$ y para todo real $t \in (0,1)$ se tiene que*

$$f(tx_2 + (1-t)\,x_1) > f(x_2).$$

Es obvio que toda función estrictamente casi-cóncava es casi-cóncava, pero el recíproco no es válido. Para ilustrar lo anterior se pueden ver los gráficos 5-1 y 5-2.

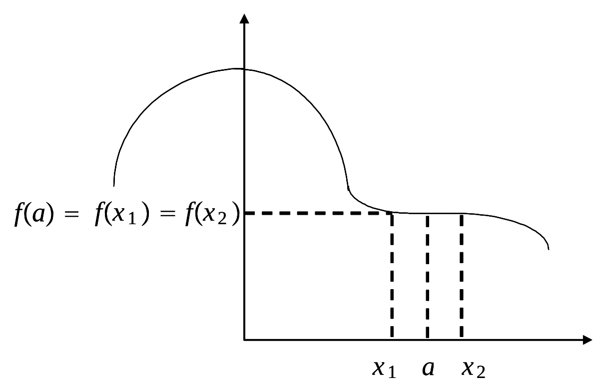


Figura 5-1 Gráfica de una función casi-cóncava no estricta.

En la figura 5-1 se tiene que a es un punto del intervalo (x_1, x_2), es decir $a = tx_1 + (1-t)\,x_2$ para alguna $t \in (0,1)$. Del dibujo se observa que $f(a) = f(x_2)$, por lo tanto la función no cumple con la definición de ser estrictamente cóncava. Un contraejemplo de que el recíproco no es válido es la gráfica 5-2.

La siguiente proposición, relaciona el concepto de funciones cóncavas con las funciones casi cónvavas.

Proposición 8

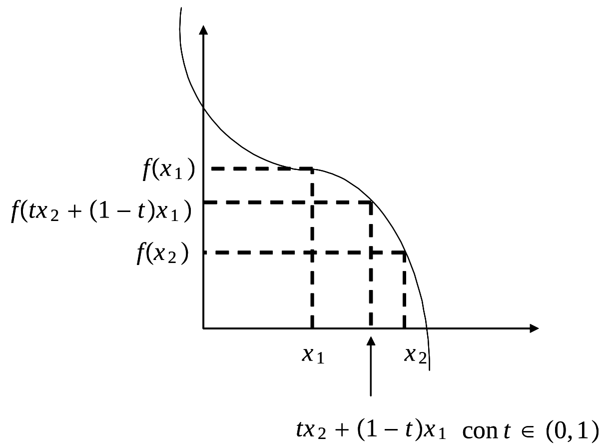


Figura 5-2 Gráfica de una función casi-cóncava estricta.

(i) Cada función cóncava es casi-cóncava, pero el reciproco no es válido

(ii) Si $X \subset R$ convexo, cada $f : X \to R$ monótona es casi-cóncava

Demostración. (i) Suponemos que para todo $x_1, x_2 \in X$ se tiene que $f(x_1) \geqq f(x_2)$ y sea $t \in [0,1]$ entonces

$$\begin{aligned} tf(x_1) &\geqq tf(x_2) \\ tf(x_1) + f(x_2) &\geqq tf(x_2) + f(x_2) \\ tf(x_1) + f(x_2) - tf(x_2) &\geqq f(x_2), \end{aligned}$$

por hipótesis f es cóncava, por lo que se cumple

$$f\left(tx_1 + (1-t)\,x_2\right) \geqq tf\left(x_1\right) + (1-t)\,f\left(x_2\right),$$

de donde se deduce que

$$f\left(tx_1 + (1-t)\,x_2\right) \geqq f(x_2).$$

Por lo tanto f es casi-cóncava. ■

Demostración. (ii) Sean $x_1, x_2 \in X$ un conjunto convexo y suponemos que $f(x_1) \geqq f(x_2)$, lo que implica que $x_1 \geqq x_2$. Ya que si suponemos que $x_1 < x_2$, se tiene que $f(x_1) < f(x_2)$ por ser f una función monótona, lo cual nos lleva a contradecir nuestra suposición.

Por lo tanto se tiene que

$$x_1 \geqq x_2$$

multiplicando por t de ambos lados se tiene

$$tx_1 \geqq tx_2$$

y sumando de ambos lados $(1-t)\,x_2$ se tiene

$$tx_1 + (1-t)\,x_2 \geqq tx_2 + (1-t)\,x_2,$$

es decir

$$tx_1 + (1-t)\,x_2 \geqq x_2.$$

Como f es una función monótona se cumple la siguiente desigualdad

$$f\,(tx_1 + (1-t)\,x_2) \geqq f\,(x_2)$$

es decir f es casi-cóncava. Por lo que queda demostrada la afirmación. ■

Condiciones de Kuhn Tucker

Frecuentemente, en las aplicaciones económicas aparecen problemas con restricciónes de desigualdad no lineales, donde se pide la no negatividad en las variables de elección; a este tipo de problemas les llamaremos "problema de programación cóncava con restricciones de no negatividad", para referirnos a ellos utilizaremos la siguiente simplificación PPCNN. Los cuales se plantean de la siguiente forma:

Problema 5 (PPCNN) *Sea $X \subset R^n$ un conjunto convexo y abierto, sean*

$$f, g_1, g_2, ..., g_m : X \to R$$

funciones cóncavas, entonces el problema de programación cóncava está definido como sigue:

$$\begin{array}{rcl} & & \text{máx } f(x) \\ g_j(x) & \geqq & 0 \\ x_i & \geq & 0 \\ \textit{para toda } j & = & 1,2,...,m \\ \textit{y para } i & = & 1,2,...,n, \end{array}$$

que en forma desarrollada se puede escribir como:

$$\begin{array}{rcl} & & \text{máx } f(x) \\ s.a.\ g_1(x) & \geq & 0 \\ g_2(x) & \geq & 0 \\ & & \vdots \\ g_m(x) & \geq & 0 \\ x_1 & \geq & 0 \\ x_2 & \geq & 0 \\ & & \vdots \\ x_n & \geq & 0. \end{array}$$

Definición 11 *Considérese un problema de programación cóncava con restricciones de no negatividad como el planteado anteriormente, diremos que satisface la condición de Slater, si existe* $x \in X$ *tal que:*

$$\begin{aligned} g_j(x) &> 0 \text{ para } j = 1, 2, ..., m \\ x_i &> 0 \text{ para } i = 1, 2, ..., n. \end{aligned}$$

Ahora bien, si se considera el problema de programación cóncava con restricciones de no negatividad y, se supone además, que las funciones $f, g_1, g_2, ..., g_m : X \to R$ son continuamente diferenciables en $X \subset R^n$ un conjunto convexo y abierto. Si el problema satisface la condición de Slater y además existen

$$x^* \in X \text{ y } \lambda^* = (\lambda_1, \lambda_2, ..., \lambda_m) \in R^m_+$$

tales que, se cumplen las siguientes condiciones para $i = 1, 2, ..., n$ y $j = 1, 2, ..., m$

$$\frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) \leqq 0 \tag{5.23}$$

$$\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) \right) x_i^* = 0 \tag{5.24}$$

$$\sum_{j=1}^{m} \lambda_j^* \, g_j(x^*) = 0 \tag{5.25}$$

$$g_j(x^*) \geqq 0 \tag{5.26}$$

$$x_i^* \geqq 0 \tag{5.27}$$

$$\lambda_j^* \geqq 0, \tag{5.28}$$

entonces se dice que x^* es solución del problema.

A estas condiciones las llamaremos condiciones de Kuhn Tucker para un problema de programación cóncava con restricciones de no negatividad.

Una idea intuitiva de las condiciones de Kuhn Tucker

Una función con una sola variable y la condición de no negatividad

A fin de que, se tenga un mejor entendimiento de las condiciones de Kuhn Tucker, suponemos el caso más simple. Una función de una sola variable y sin restricciones, esto es, cuando $n = 1$ y $m = 0$.

Sea $X \subset R$ un conjunto convexo y abierto con una función $f : X \to R$ estrictamente cóncava y la restricción de no negatividad para la variable de elección, entonces el problema a resolver es

$$\begin{aligned} &\text{máx } f(x) \\ \text{s.a. } x &\geq 0. \end{aligned}$$

Analizaremos la gráfica de los posibles casos, de donde podremos deducir intuitivamente las condiciones para la solución. Dadas las hipótesis sabemos que su gráfica está contenida en R^2, y por las condiciones de no negatividad podemos restringir el problema al primer cuadrante; así pues, se tendran los siguientes dos casos:

Caso 1 Para este primer caso suponemos que la solución está en el interior de X, es decir decir $x^* > 0$, lo cual corresponde a la gráfica 5-3.

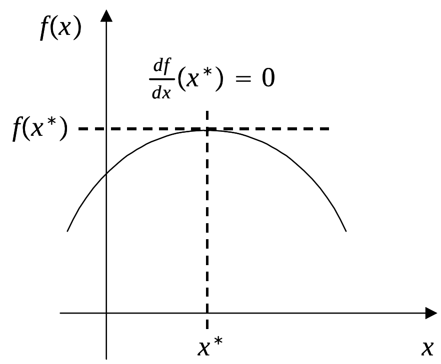


Figura 5-3

La solución es estrictamente positiva y al evaluar su derivada vale cero como se muestra en la figura 5-3. Se observa entonces que si $x^* \in X$ es solución interior del problema debe satisfacer las siguientes dos condiciones:

$$\begin{aligned} \frac{df}{dx}(x^*) &= 0 \\ x^* &> 0. \end{aligned}$$

Caso 2 Soluciones frontera o de contorno. En este caso se supone que la solución es igual a cero, es decir, $x^* = 0$. De aquí se deducen 2 posibles subcasos. Uno de ellos es que el máximo global de la función se alcance en $x^* = 0$, si esto sucede al evaluar la derivada, está vale cero, ver figura 5-4 inciso (a). El otro subcaso es que el máximo global este fuera de la región factible, así pues, por la hipótesis de concavidad la solución esta nuevamente en $x^* = 0$, pero al evaluar la derivada está es menor que cero, ver figura 5-4 inciso (b).

Así, de la figura 5-4 se observa que, sí x^* es solución frontera o de contorno del problema, debe satisfacer las siguientes tres condiciones:

$$\begin{aligned} \frac{df}{dx}(x^*) &< 0 \\ \frac{df}{dx}(x^*) &= 0 \\ x^* &= 0. \end{aligned}$$

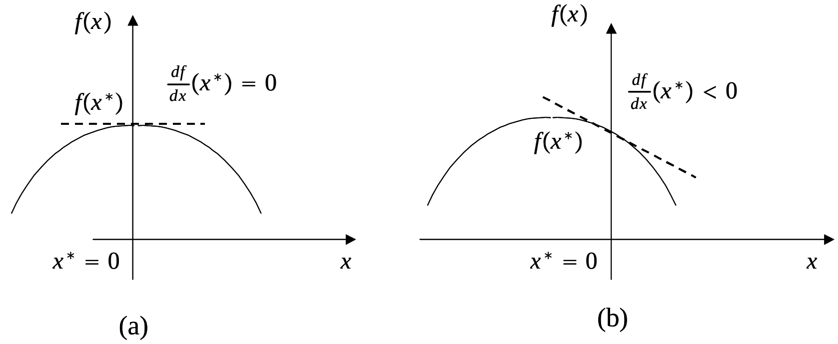


Figura 5-4

Es conveniente aclarar que estamos ante un problema con una función de una sola variable, donde la notación no requiere del uso de derivadas parciales, no obstante, para poder hacer la analogía con las condiciones de Kuhn Tucker en el caso más general, diremos que:

$$\frac{df}{dx}(x^*) = \frac{\partial f}{\partial x}(x^*).$$

Entonces, de las gráficas 5-3 y 5-4 se puede inferir el siguiente resumen de las condiciones necesarias para el PPCNN cuando $n = 1$ y $m = 0$.

Si x^* es solución, entonces:

$$\frac{\partial f}{\partial x}(x^*) \leq 0 \tag{5.29}$$

$$\frac{\partial f}{\partial x}(x^*)\, x^* = 0 \tag{5.30}$$

$$x^* \geq 0 \tag{5.31}$$

Finalmente se tiene que las analogías con las condiciones de Kuhn Tucker son :

- La condición 5.23 con la condición del caso particular 5.29,
- La condición 5.24 con la condición del caso particular 5.30 y
- La condición 5.27 con la condición del caso particular 5.31.

Una función con varias variables y la condición de no negatividad

En está subsección mostraremos un ejemplo con una función vectorial y la condición de no negatividad para las variables de decisión, esto es $n > 1$ y $m = 0$.

Sea $X \subset R^n$ un conjunto convexo y abierto, con una función dos veces diferenciable $f : X \to R$, estrictamente cóncava y la restricción de no negatividad para la variable de elección, entonces el problema a resolver es

$$\begin{aligned} & \text{máx } f(x) \\ \text{s.a. } x_i &\geq 0 \\ \text{para } i &= 1, 2, .., n. \end{aligned}$$

Supongamos que existe un $x^* \in X$ solución del problema, entonces se tiene que para todos los puntos vecinos $x^* + \Delta x$ se cumple la siguiente desigualdad:

$$f(x^*) \geq f(x^* + h\Delta x) \tag{5.32}$$

donde Δx es una dirección de movimiento en $X \subset R^n$ y h es un número real positivo arbitrariamente pequeño. Por ser f dos veces diferenciable utilizaremos el desarrollo de Taylor en el segundo miembro de la ecuación 5.32 alrededor de x^*, obteniendo lo siguiente

$$f(x^* + h\Delta x) = f(x^*) + \frac{\partial f}{\partial x}(x^*)\, h\Delta x + \frac{1}{2!}\frac{\partial^2 f}{\partial x^2}(x^* + \theta(h\Delta x))(h\Delta x)^2 \tag{5.33}$$

con $0 < \theta < 1$ y $0 < h < 1$.

De las expresiones 5.32 y 5.33 se obtiene la siguiente desigualdad

$$f(x^*) \geq f(x^*) + \frac{\partial f}{\partial x}(x^*)\, h\Delta x + \frac{1}{2!}\frac{\partial^2 f}{\partial x^2}(x^* + \theta(h\Delta x))(h\Delta x)^2,$$

eliminando $f(x^*)$ de ambos lados se tiene

$$\frac{\partial f}{\partial x}(x^*)\, h\Delta x + \frac{1}{2!}\frac{\partial^2 f}{\partial x^2}(x^* + \theta(h\Delta x))(h\Delta x)^2 \leq 0 \tag{5.34}$$

la cual es una condición necesaria para un máximo local en x^*.

Nuevamente estamos ante una situación con dos casos. Un caso es que x^* sea un punto interior de la región factible y el otro que x^* sea un punto frontera o de contorno de la región factible.

Caso1 Si x^* es un punto interior entonces $x^* > \overline{0}$, y por ser x^* máximo del problema se tiene que

$$\frac{\partial f}{\partial x_i}(x^*) = 0 \text{ para toda } i = 1, 2, ..., n,$$

lo que implica que la ecuación 5.34 puede expresarse como

$$\frac{1}{2!}\frac{\partial^2 f}{\partial x^2}(x^* + \theta(h\Delta x))(h\Delta x)^2 \leq 0. \tag{5.35}$$

Obsérvese que

$$\frac{\partial^2 f}{\partial x^2}(x^* + \theta(h\Delta x))$$

es la matriz de segundas derivadas parciales de f, y también que la expresión del lado izquierdo de la desigualdad 5.35 es una forma cuadratica. Por otro lado, sabemos que cuando la función f es cóncava la forma cuadratica es semidefinida negativa sin importar que forma tienen las h, lo que nos lleva a las mismas condiciones de la programación clasica; por lo tanto, se puede escribir la ecuación 5.35 en forma desarrollada como sigue

$$\frac{1}{2!}(h_1\Delta x_{1,}.., h_n\Delta x_{n,})\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x^*) & \cdots & \frac{\partial^2 f}{\partial x_1 x_n}(x^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1}(x^*) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x^*) \end{pmatrix}\begin{pmatrix} h_1\Delta x_1 \\ \vdots \\ h_n\Delta x_n \end{pmatrix}$$

de tal forma que la expresión 5.35, también se puede escribir como

$$\frac{1}{2!}(h\Delta x)^t \frac{\partial^2 f}{\partial x^2}(x^* + \theta(h\Delta x))(h\Delta x) \le 0. \tag{5.36}$$

Lo cual es una condición necesaria para un máximo local en x^*.

Caso 2 Supongamos ahora que x^* es una solución de contorno o frontera de la región factible del problema, esto es:

$$x^* = \left(x_1^*, x_2^*, .., x_{k-1}^*, x_k^*, x_{k+1}^*, .., x_n^*\right)$$

con $x_k^* = 0$ para alguna k.

Para visualizarlo supongamos $n = 3$, y que x^*es una solución de la forma $(0, x_2^*, x_3^*)$, graficamente tendríamos la figura 5-5.

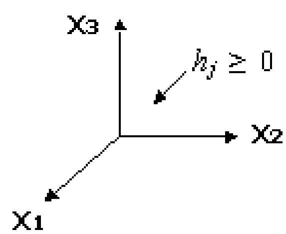


Figura 5-5 Un punto frontera

Obsérvese que la única dirección posible para la cordenada x_1^* es hacia el interior del octante positivo, puesto que las condiciones de no negatividad restringen el problema a dicha región factible. Por lo tanto $h\Delta x_1^* > 0$.

Sin perdida de generalidad, para el caso de n variables sea $x_k^* = 0$ la $k-$ésima entrada de x^*, y sea $x_i^* > 0$ para toda $i \neq k$ con $i = 1, 2, ..., n$, entonces $h\Delta x_k^* > 0$. De la ecuación 5.36 y desarrollando en x_k^* se tiene que

$$\frac{\partial f}{\partial x}(x_k^*)\, h\Delta x_k + \frac{1}{2!}\frac{\partial^2 f}{\partial x^2}(x^* + \theta(h\Delta x))(h\Delta x)^2 \le 0,$$

de donde por ser $h > 0$ podemos dividir la ecuación anterior entre h sin alterar la desigualdad. Después haciendo tender h a cero se tiene

$$\frac{\partial f}{\partial x}(x_k^*)\, \Delta x_k \le 0$$

y sabemos que $\Delta x_k > 0$ por ser x_k^* un punto frontera o de contorno, por lo que se deduce que:

$$\frac{\partial f}{\partial x}(x_k^*) \le 0.$$

Resumiendo, en ambos casos se tiene que la primera derivada con respecto a x^* se anula necesariamente en una solución interior, esto es, cuando $x^* > \overline{0}$; y en el caso de una solución frontera o de contorno, es decir $x^* \geq \overline{0}$, la primera derivada es necesariamente menor o igual a cero.

De lo anterior podemos inferir que una condición necesaria es

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x}(x^*)\, x_i^* = 0,$$

y en general las condiciones necesarias para que x^* sea solución del PPCNN están determinadas por

$$\begin{aligned} \frac{\partial f}{\partial x}(x^*) &\leq \overline{0} \\ \sum_{i=1}^{n} \frac{\partial f}{\partial x}(x^*)\, x_i^* &= 0 \\ x_i^* &\geq 0 \\ \text{para toda } i &= 1, 2, ..., n. \end{aligned}$$

Teoremas que establecen las condiciones de Kuhn Tucker en el caso más general

En esta sección, enunciaremos los teoremas que nos llevan a establecer las condiciones que debe cumplir la solución para un problema de programación cóncava con condiciones de no negatividad, a partir de los resultados de la programación cóncava analizados en el capítulo 4.

Recordemos el problema de programación cóncava PPC

$$\begin{aligned} & \text{máx } f(x) \\ \text{sujeto a } g_j(x) &\geqq 0 \\ \text{para } \ j &= 1, 2, .., m, \end{aligned}$$

el cual definimos con anterioridad en la página 84, y recordemos también que el problema PPC cumple la condición de Slater definida en la página 95. Dadas $g_1, g_2, ..., g_m : X \to R$ cóncavas con $X \subset R^n$, existe $\overline{x} \in X$ tales que $g_j(\overline{x}) > 0$ para todo $j = 1, 2, ..., m$.

Utilizando el resultado del teorema 22 (ver página 105) establecemos los siguientes teoremas.

Teorema 23 (A) *Sea $X \subset R^n$ un conjunto convexo y abierto. Sean*

$$f, g_1, g_2, ..., g_m : X \to R$$

funciones cóncavas y contiuamente diferenciables, si x^* es solución de PPC y si además se satisface la condición de Slater, entonces existe $\lambda^* \in R_+^m$ tal que

(i) $\dfrac{\partial f}{\partial x}(x^*) + \sum_{j=1}^{m} \lambda_j^* \dfrac{\partial g_j(x)}{\partial x}(x^*) = 0$

(ii) $g_j(x) \geqq 0$

(iii) $\sum_{j=1}^{m} \lambda_j^* (g_j(x^*)) = 0$

para toda $i = 1, 2, ..., n$ y $j = 1, 2, ..., m$.

Teorema 24 (B) *Sea $X \subset R^n$ un conjunto convexo y abierto. Sean*

$$f, g_1, g_2, ..., g_m : X \to R$$

funciones cóncavas y contiuamente diferenciables. Si $x^* \in X$ es factible para el problema PPC, y además existe $\lambda^* \in R_+^m$ tal que se cumple

(i) $\dfrac{\partial f}{\partial x}(x^*) + \sum_{j=1}^{m} \lambda_j^* \dfrac{\partial g_j(x)}{\partial x}(x^*) = 0$

(ii) $\sum_{j=1}^{m} \lambda_j^* (g_j(x^*)) = 0$

para toda $i = 1, 2, .., n$ y $j = 1, 2, .., m$. Entonces x^* es solución para el problema PPC.

Obsérvese que las condiciones (i) y (iii) del teorema 23 corresponden a las condiciones (i) y (ii) en la hipótesis del teorema 24.

Finalmente definiremos los teoremas que resuelven el problema que definimos al principio de esté capítulo al que llamamos PPCNN. Éstos teoremas se desprenden de los dos teoremas anteriores y son la base para obtener las condiciones de Kuhn Tucker, que caracterizan la solución al problema PPCNN.

Teorema 25 (A') *Considérese un problema del tipo PPCNN, entonces si x^* es solución y se cumple la condición de Slater, entonces existe $\lambda^* \in R_+^m$ tal que:*

(i)' $\dfrac{\partial f}{\partial x_i}(x_i^*) + \sum_{j=1}^{m} \lambda_j^* \dfrac{\partial g_j(x)}{\partial x_i}(x^*) \leqq 0$

(ii)' $\sum_{i=1}^{n} \left(\dfrac{\partial f}{\partial x_i}(x_i^*) + \sum_{j=1}^{m} \lambda_j^* \dfrac{\partial g_j(x)}{\partial x_i}(x^*) \right) x_i^* \leqq 0$

(iii)' $\sum_{j=1}^{m} \lambda_j^* (g_j(x^*)) = 0$

(iv)' $g_j(x^*) \geqq 0$

para toda $i = 1, 2, .., n$ y $j = 1, 2, .., m$.

Teorema 26 (B') *Considérese un problema del tipo PPCNN y sean*

$$f, g_1, g_2, ..., g_m : X \to R$$

funciones cóncavas y contiuamente diferenciables. Si x^* cumple las restricciones del problema, y existe $\lambda^* \in R_+^m$ tal que

(i)' $\dfrac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j^* \dfrac{\partial g_j(x)}{\partial x_i}(x^*) \leqq 0$

(ii)' $\sum_{i=1}^{n} \left(\dfrac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j \dfrac{\partial g_j(x)}{\partial x_i}(x^*) \right) x_i^* = 0$

(iii)' $\sum_{j=1}^{m} \lambda_j^* (g_j(x^*)) = 0$

(iv)' $g_j(x^*) \geqq 0$

para toda $i = 1, 2, .., n$ y $j = 1, 2, .., m$. Entonces x^* es solución del problema PPCNN.

Por lo que las condiciones de Kuhn Tucker quedan establecidas por las cuatro condiciones del teorema 26, y por la condición de no negatividad, las cuales se resumen en las siguientes condiciones para $i = 1, 2, ..., n$ y $j = 1, 2, ..., m$

$$\begin{aligned}
\frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) &\leqq 0 \\
\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) \right) x_i^* &= 0 \\
\sum_{j=1}^{m} \lambda_j^* \; g_j(x^*) &= 0 \\
g_j(x^*) &\geqq 0 \\
x_i^* &\geqq 0.
\end{aligned}$$

Soluciones a los problemas planteados en el capítulo 1

Un problema geométrico

El problema planteado en la página 19 es el siguiente

$$
\begin{aligned}
\text{máx } f(x_1,x_2) &= -8x_1^2-10x_2^2+12x_1x_2-50x_1+80x_2\\
\text{s.a. } 1-x_1-x_2 &\geqq 0\\
2-8x_1^2-10x_2^2 &\geqq 0\\
x_1,x_2 &\geqq 0
\end{aligned}
$$

Se puede ver que tanto la función objetivo como las restricciones son funciones cóncavas y continuamente diferenciables, por lo tanto podemos usar las condiciones de Kuhn-Tucker (ver página 129). Sea $n=2$ y $m=2$, entonces las condiciones de Kuhn-Tucker son las siguientes

Condición i)

$$
\begin{aligned}
\frac{\partial f}{\partial x_i}(x^*)+\sum_{j=1}^{2}\lambda_j^*\frac{\partial g_j}{\partial x_i}(x^*) &\leqq 0\\
\text{para todo } i,j &= 1,2
\end{aligned}
$$

y resolviendo para nuestro problema particular se tiene que

$$(-16x_1+12x_2-50,12x_1-20x_2+80)+$$

$$+\lambda_1(-1,-1)+\lambda_2(-16x_1,-2x_2)\leqq(0,0)$$

por lo que la expresión anterior genera las dos condiciones siguientes

$$
\begin{aligned}
-16x_1+12x_2-50-\lambda_1-16x_1\lambda_2 &\leqq 0\\
12x_1-20x_2+80-\lambda_1-2x_1\lambda_2 &\leqq 0.
\end{aligned}
$$

Condición ii)

$$\sum_{i=1}^{2}\left(\frac{\partial f}{\partial x_i}(x^*)+\sum_{j=1}^{2}\lambda_j^*\frac{\partial g_j}{\partial x_i}(x^*)\right)x_i^*=0$$

por lo tanto

$$(-16x_1+12x_2-50-\lambda_1-16\lambda_2x_1)\,x_1+$$

$$+\,(12x_1-20x_2+80-\lambda_1-2\lambda_2x_2)\,x_2=0.$$

Condición iii)

$$\sum_{j=1}^{2} \lambda_j^* \, g_j(x^*) = 0$$

es decir

$$\lambda_1 (1 - x_1 - x_2) + \lambda_2 \left(2 - 8x_1^2 - x_2^2\right) = 0.$$

Condición iv)

$$g_j(x^*) \geqq 0 \text{ para todo } j = 1, 2$$

es decir

$$\begin{aligned} 1 - x_1 - x_2 &\geqq 0 \\ 2 - 8x_1^2 - x_2^2 &\geqq 0. \end{aligned}$$

Condición v)

$$x_i^* \geqq 0 \text{ para todo } i = 1, 2$$

es decir

$$\begin{aligned} x_1 &\geqq 0 \\ x_2 &\geqq 0. \end{aligned}$$

Condición vi)

$$\lambda_j^* \geqq 0 \text{ para todo } j = 1, 2$$

es decir:

$$\begin{aligned} \lambda_1 &\geqq 0 \\ \lambda_2 &\geqq 0. \end{aligned}$$

Si bien, estas condiciones caracterizan a la solución de nuestro problema, no es fácil deducir de ellas donde se encuentra, es decir las condiciones de Kuhn Tuker en la practica no son muy útiles para calcular dicha solución.

En este caso, si nos fijamos en la gráfica de la página 20 se tiene que

$$x^* = (x_1^*, x_2^*) = (0, 1)$$

resulta ser la solución, y dado x^* se obtiene que

$$\lambda^* = (\lambda_1^*, \lambda_2^*) = (60, 0)$$

de donde se observa que x^* y λ^* satisfacen las condiciones de Kuhn-Tuker.

Maximización de las ventas de una empresa

Resolveremos el problema planteado en la página 20 cuya forma estándar esta dada por

$$\begin{aligned} \text{máx}\, I(Q) &= 32Q - Q^2 \\ \text{s.a.} \; -2Q^2 + 24Q - 14 &\geqq 0 \\ Q &\geqq 0, \end{aligned}$$

en donde las condiciones de Kuhn-Tucker son:

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{1} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) &\leqq 0 \\ \sum_{i=1}^{1} \left(\frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{1} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) \right) x_i^* &= 0 \\ \sum_{j=1}^{1} \lambda_j^* \; g_j(x^*) &= 0 \\ g_j(x^*) &\geqq 0 \\ x_i^* &\geqq 0 \\ \lambda_j^* &\geqq 0 \end{aligned}$$

con $i, j = 1$, y aplicadas a nuestro caso particular dichas condiciones quedan establecidas por

$$32 - 2Q + \lambda(-4Q + 24) \leqq 0 \tag{5.37}$$

$$(32 - 2Q + \lambda(-4Q + 24))\, Q = 0 \tag{5.38}$$

$$\lambda\left(-2Q^2 + 24Q - 14\right) = 0 \tag{5.39}$$

$$-2Q^2 + 24Q - 14 \geqq 0 \tag{5.40}$$

$$Q \geqq 0 \tag{5.41}$$

$$\lambda \geqq 0 \tag{5.42}$$

Para encontrar la solución, suponemos que $Q = 0$, entonces de la ecuación 5.37 se tiene que $\lambda \leqq -\frac{4}{3}$, lo cual nos lleva a una contradicción en la ecuación 5.42, de donde se deduce que

$$Q > 0.$$

Por el resultado anterior se debe cumplir que

$$32 - 2Q + \lambda(-4Q + 24) = 0, \tag{5.43}$$

lo que nos lleva a una ecuación en dos variables.

Por otro lado sabemos que $Q > 0$, entonces $\lambda > 0$ o $\lambda = 0$.

Si suponemos que $\lambda = 0$, entonces $Q = 16$, pero si utilizamos este valor en la ecuación 5.40 llegamos a una contradicción pues

$$-2(16)^2 + 24(16) - 14 = -142 \ngeqq 0,$$

por lo tanto $\lambda > 0$, de donde se deduce que

$$-2Q^2 + 24Q - 14 = 0,$$

resolviendo esta ecuación de segundo grado se tiene que

$$\begin{aligned} Q_1 &= 1 \\ Q_2 &= 11. \end{aligned}$$

Si suponemos que $Q^* = 1$, no se satisface la ecuación 5.43 pues hace que $\lambda = -\frac{3}{2}$, por lo tanto

$$Q^* = 11,$$

de donde se obtiene que

$$\lambda^* = \frac{1}{2}$$

entonces $(Q^*, \lambda^*) = \left(11, \frac{1}{2}\right)$es la solución a nuestro problema ya que satisface todas las condiciones de Kuhn-Tucker. Por lo tanto la producción maximiza las ventas con $(Q^*, \lambda^*) = \left(11, \frac{1}{2}\right)$.

Problema del consumidor

Contamos ahora con la herramienta necesaria para poder resolver el problema del consumidor planteado en el capítulo 1 en la página 22, el cual propone que el consumidor desea obtener la máxima utilidad al consumir una canasta con cuatro tipos de bienes, y donde su función de utilidad es del tipo Bernoulli (ver definición en la pag 127), es decir:

$$\begin{aligned} \text{máx } \; U(x_1, x_2, x_{3,} x_4) &= \sum_{1=1}^{4} a_i \log (x_i + b_i) \\ \text{s.a. } p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_4 &\leqq M \\ \text{con } x_1, x_2, x_{3,} x_4 &\geqq 0. \end{aligned}$$

Se tiene que si x^* es solución del problema y satisface las condiciones de Kuhn Tucker entonces sabemos que existe λ^* un real no negativo tal que cumple con

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) &\leqq 0 \\ g(x^*) &\geqq 0 \\ x_i^* &\geqq 0 \\ \lambda^* g(x^*) &= 0 \\ \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_k}(x^*) + \lambda^* \frac{\partial g_j}{\partial x_i}(x^*) \right) x_i^* &= 0 \\ \lambda^* &\geqq 0 \end{aligned}$$

para todo $i = 1, 2, .., n$ y $j = 1, 2, .., m$.

Aplicando las condiciones de Kuhn Tucker al problema se tiene

$$\begin{aligned}
\frac{a_k}{x_k^* + b_k} - \lambda^* p_k &\leqq 0 && (5.44)\\
M - \sum_{i=1}^{4} p_i x_i^* &\geqq 0 && (5.45)\\
x_k^* &\geqq 0 && (5.46)\\
\lambda^* \left(M - \sum_{i=1}^{4} p_i x_i^* \right) &= 0 && (5.47)\\
\sum_{k=1}^{4} \left(\frac{a_k}{x_k^* + b_k} - \lambda^* p_k \right) x_k^* &= 0 && (5.48)\\
\lambda^* &\geqq 0 && (5.49)
\end{aligned}$$

para todo $k = 1, 2, 3, 4$.

Entonces de la ecuación 5.44 se deduce que

$$\frac{a_k}{x_k^* + b_k} \leqq \lambda^* p_k \text{ para todo } k = 1, 2, 3, 4$$

lo cual se puede expresar como

$$\left(\frac{a_k}{x_k^* + b_k} \right) \frac{1}{p_k} \leqq \lambda^* \text{ para todo } k = 1, 2, 3, 4.$$

Por la hipótesis del problema y de la definición función de utilidad Bernoulli se sabe que los valores $a_k > 0, b_k > 0$ y $p_k > 0$, de donde se deduce que λ^* es estrictamente positivo, de este hecho y utilizando la condición 5.47, se establece la siguiente igualdad

$$M - \sum_{i=1}^{4} p_i x_i^* = 0$$

por lo que la solución del problema está en el hiperplano

$$\sum_{i=1}^{4} p_i x_i^* = M \tag{5.50}$$

además utilizando la condición 5.44 se tiene que para todo $k = 1, 2, 3, 4$

$$\frac{a_k}{x_k^* + b_k} \leqq \lambda^* p_k$$

es decir

$$\frac{a_k}{\lambda^* p_k} \leqq (x_k^* + b_k)$$

lo que podemos expresar como

$$\frac{a_k}{\lambda^* p_k} - b_k \leqq x_k^*.$$

Obsérvese que x_k^* nos indica cuál debe ser el consumo del individuo de cada uno de los cuatro bienes que existen en el mercado, demostraremos que ese consumo esta dado por la siguiente cantidad

$$x_k^* \; = \text{máx}\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right) \text{para todo } k = 1, 2, 3, 4.$$

Para demostrarlo supondremos dos casos, primero suponemos a $x_k^* > 0$ y después que $x_k^* = 0$, basados en las condiciones de no negatividad del problma.

Caso $x_k^* > 0$ De las condiciones 5.44 y 5.48, se cumple para todo $k = 1, 2, 3, 4$

$$\left(\frac{a_k}{x_k^* + b_k} - \lambda^* p_k\right) x_k^* \; = 0$$

y por hipótesis $x_k^* > 0$, por lo que se deduce

$$\frac{a_k}{x_k^* + b_k} - \lambda^* p_k \; = 0$$

es decir

$$\frac{a_k}{\lambda^* p_k} - b_k = x_k^* > 0$$

entonces

$$x_k^* \; = \text{máx}\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right) \text{para todo } k = 1, 2, 3, 4.$$

Caso $x_k^* = 0$ Nuevamente de la condición 5.44 se tiene que:

$$\frac{a_k}{\lambda^* p_k} - b_k \leqq x_k^* = 0 \text{ para todo } k = 1, 2, 3, 4,$$

lo que contradice la condición de no negatividad en la ecuación 5.46, de donde se deduce que

$$x_k^* \; = \text{máx}\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right) = 0 \text{ para todo } k = 1, 2, 3, 4,$$

ya que el individuo no puede consumir cantidades negativas de ninguno de los bienes.

Entonces de ambos casos podemos concluír que la solución del problema esta dada por

$$x_k^* \; = \text{máx}\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right) \text{para todo } k = 1, 2, 3, 4,$$

obsérvese que x_k^* esta en términos de λ^*, la cual es una variable positiva. Ahora sabemos que la solución debe satisfacer la ecuación 5.50, es decir

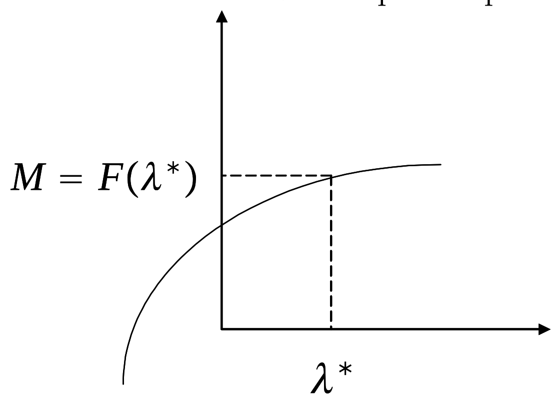
$$\sum_{i=1}^{4} p_i \, \text{máx}\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right) = M. \tag{5.51}$$

Entonces debemos buscar el valor de λ^*, de tal forma que se cumpla la ecuación 5.51.

Para poder analizar como debe ser λ^*, definamos la función $F : R_{++} \to R$ dada por

$$F(\lambda) = \sum_{i=1}^{4} p_i \max\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right), \tag{5.52}$$

en donde la solución gráfica consiste en encontrar λ^* tal que cumpla con $M = F(\lambda^*)$, es decir:



y para simplificar la ecuación 5.52, hagamos el siguiente cambio de variable

$$\mu = \frac{1}{\lambda},$$

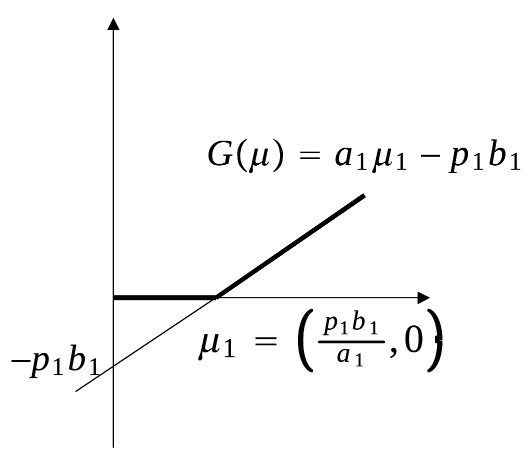
entonces la función queda definida por:

$$G(\mu) = \sum_{i=1}^{4} p_i \max\left(0, \frac{a_k}{p_k}\mu - b_k\right). \tag{5.53}$$

Sabemos por hipótesis que a_k, b_k y p_k son constantes positivas, supongamos que el individuo consume únicamente del bien tipo 1, entonces la función en la ecuación 5.53 toma la forma:

$$\begin{aligned} G(\mu) &= p_1\left(\frac{a_1}{p_1}\mu - b_1\right) \\ &= a_1\mu - p_1 b_1, \end{aligned} \tag{5.54}$$

en donde su gráfica corresponde a una línea recta con pendiente positiva a_1 y ordenada al origen $-p_1 b_1$, es decir

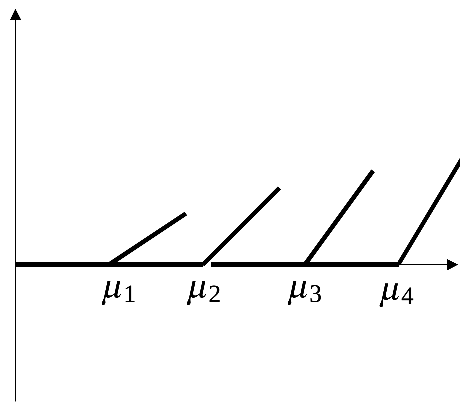


La solución al problema es el conjunto de puntos de la función que quedan dentro del primer cuadrante donde μ_1 es el valor μ que hace que la función $G(\mu)$ sea igual a cero.

Análogamente si suponemos que ahora consume de los 2 primeros tipos de bienes se tiene que

$$G(\mu) = (a_1 + a_2)\,\mu - (p_1 b_1 + p_2 b_2) \tag{5.55}$$

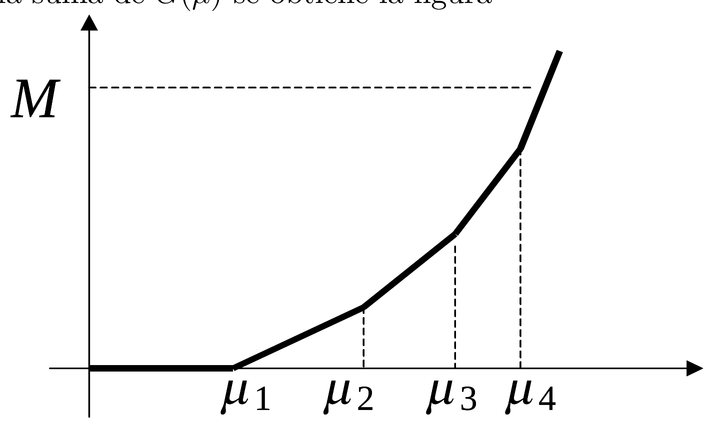
lo cual representa como en el caso anterior la ecuación de una recta, pero ahora con pendiente mayor. Y $\mu_2 = \frac{p_1 b_1 + p_2 b_2}{a_1 + a_2}$ es el valor que hace cero a la función $G(\mu)$, la gráfica es muy parecida al caso anterior. Haciendo un razonamiento análogo para los otros dos tipos de bienes y dibujando sus gráficas en el mismo plano, se tiene la figura



donde se cumple la condición

$$0 < \mu_1 \leqq \mu_2 \leqq \mu_3 \leqq \mu_4.$$

Ahora, si consideramos la suma de $G(\mu)$ se obtiene la figura



por lo que podemos decir que la demanda del individuo dependerá de la renta, es decir si la renta es muy pequeña solo demandará el primer tipo de bien, si la renta es mayor podrá demandar de los otros tipos de mercancias, entonces la demanda óptima de cada tipo de bien está dada por:

$$x_k^* \;= \text{máx}\left(0, \frac{a_k}{p_k}\mu - b_k\right) \text{para todo } k = 1, 2, 3, 4.$$

Donde para calcular el valor de μ, solamente necesitamos conocer el valor de la renta M.

Capítulo 6

El problema del consumidor cuando la función de utilidad es Bernoulli

Planteamiento del problema

Definamos ahora la función de utilidad Bernoulli.

Definición 12 *Sea* $U : R_+^n \to R$ *la función de utilidad Bernoulli dada por:*

$$U(x) = \sum_{i=1}^{n} a_i \log(x_i + b_i).$$

Donde $a_i, b_i > 0 \ \ \forall i = 1, 2, .., n$ *y además* $\sum_{i=1}^{n} a_i = 1.$

Y con ella, definamos ahora un nuevo problema, que está dado por:

Problema 6

$$\begin{aligned} & \text{máx} \quad U(x) \\ s.a. \ \sum_{i=1}^{n} p_i x_i &\leqq r \\ con \ \ x_i &> 0 \\ para\ toda\ i &= 0, 1, 2, ..n \end{aligned}$$

El cual podemos transformar en el siguiente problema de programación cóncava con restricciones de no negatividad.

Problema 7

$$\begin{aligned} & \text{máx } U \ (x) \\ s.a. \ \ r - \sum_{i=1}^{n} p_i x_i & \geqq 0 \\ con \ \ x_i & \geqq 0 \\ para\ toda\ i & = 0, 1, 2, ..n \end{aligned}$$

Este nuevo problema es muy parecido al problema PPCNN 5 que se definió en el capítulo 6 (ver página 109). Y ahora podemos enunciar las condiciones de Kuhn Tucker, las cuales nos caracterizan las soluciones.

Supóngase que cumple la condición de Slater (ver página 95). Además por la hipótesis del problema 7 se tiene que $U(x)$ es una función cóncava y la única restricción del problema

$$g(x) = r - \sum_{i=1}^{n} p_i x_i$$

también es una función cóncava. Si además tomamos precios positivos $p_i > 0 \ \forall i = 1, 2, .., n$ y una renta $r > 0$, nos queda que el conjunto factible de la función es un conjunto compacto, en general es un simplejo, como en el problema 2.

Y análogamente por el teorema de Weierstrass que nos dice que toda función continua definida en un compacto alcanza su máximo y su mínimo, podemos asegurar que existe x^*solución del problema 6. Por otro lado, sabemos que el problema satisface la condición de Slater.

Solución del problema

Sea x^*solución del problema 7, por las condiciones de Kuhn Tucker entonces sabemos que existe λ^* un real no negativo tal que cumple con

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) & \leqq 0 \\ g(x^*) & \geqq 0 \\ x_i^* & \geqq 0 \\ \lambda^* g(x^*) & = 0 \\ \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_k}(x^*) + \lambda^* \frac{\partial g_j}{\partial x_i}(x^*) \right) x_i^* & = 0 \\ \lambda^* & \geqq 0 \end{aligned}$$

para todo $i = 1, 2, .., n$ y $j = 1, 2, .., m$.

Aplicando las condiciones anteriores al problema 7, con la función de utilidad Bernoulli se tiene:

$$\begin{aligned}
\frac{a_k}{x_k^* + b_k} - \lambda^* p_k &\leqq 0 && (6.1)\\
r - \sum_{i=1}^{n} p_i x_i^* &\geqq 0 && (6.2)\\
x_k^* &\geqq 0 && (6.3)\\
\lambda^* \left(r - \sum_{i=1}^{n} p_i x_i^* \right) &= 0 && (6.4)\\
\sum_{k=1}^{n} \left(\frac{a_k}{x_k^* + b_k} - \lambda^* p_k \right) x_k^* &= 0 && (6.5)\\
\lambda^* &\geqq 0 && (6.6)
\end{aligned}$$

para todo $k = 1, 2, .., n$.

Entonces de la condición 6.1 se deduce que:

$$\frac{a_k}{x_k^* + b_k} \leqq \lambda^* p_k$$

es decir

$$\left(\frac{a_k}{x_k^* + b_k} \right) \frac{1}{p_k} \leqq \lambda^*$$

y por hipótesis $a_k, b_k, p_k > 0$, lo que nos lleva a afirmar que λ^* es estrictamente positivo.

Por otro lado, utilizando lo anterior y de la condición 6.4 se deduce que:

$$r - \sum_{k=1}^{n} p_k x_k^* = 0 \quad \text{ya que } \lambda^* > 0$$

es decir la solución del problema esta en el hiperplano

$$\sum_{k=1}^{n} p_k x_k^* = r,$$

además de la condición 6.1, se obtiene que:

$$\begin{aligned}
\frac{a_k}{x_k^* + b_k} &\leqq \lambda^* p_k\\
\frac{a_k}{p_k \lambda^*} &\leqq (x_k^* + b_k)
\end{aligned}$$

de donde se obtiene la siguiente desigualdad:

$$\frac{a_k}{\lambda^* p_k} - b_k \leqq x_k^*.$$

Obsérvese que x_k^* nos indica cual deberá ser el consumo por el individuo del k-esimo bien. Ahora solo falta demostrar que:

$$x_k^* = \max\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right)$$

Para ello supondremos dos casos, primero suponemos $x_k^* > 0$ y después suponemos que $x_k^* = 0$, lo anterior por las condiciones de no negatividad del problema.

Caso 2 *Sea $x_k^* > 0$ entonces de las condiciones 6.1, 6.5 y de la condición de factibilidad 6.3 se deduce que:*

$$\left(\frac{a_i}{x_i^* + b_i} - \lambda^* p_i\right) x_i^* = 0 \qquad \forall i = 1, 2, .., n$$

en particular la ecuación anterior se cumple para $i = k$, de donde se obtiene

$$\left(\frac{a_k}{x_k^* + b_k} - \lambda^* p_k\right) x_k^* = 0$$

por hipótesis tenemos que $x_k^* > 0$ por lo que se deduce que

$$\frac{a_k}{x_k^* + b_k} - \lambda^* p_k = 0$$

de donde se tiene

$$\frac{a_k}{\lambda^* p_k} - b_k = x_k^* > 0$$

es decir

$$x_k^* = \max\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right)$$

Caso 3 *Sea $x_k^* = 0$, entonces de la condición 6.1 se tiene que:*

$$\frac{a_k}{\lambda^* p_k} - b_k \leqq x_k^* = 0$$

$$\text{es decir } \frac{a_k}{\lambda^* p_k} - b_k \leqq 0$$

de donde se deduce que

$$x_k^* = \max\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right) = 0$$

por lo tanto en cualquiera de los dos casos se tiene que:

$$x_k^* = \max\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right).$$

De lo anterior por estar x_k^* en términos de λ,entonces sabemos que la solución al problema esta sobre el hiperplano

$$\sum_{k=1}^{n} p_k x_k^* = r$$

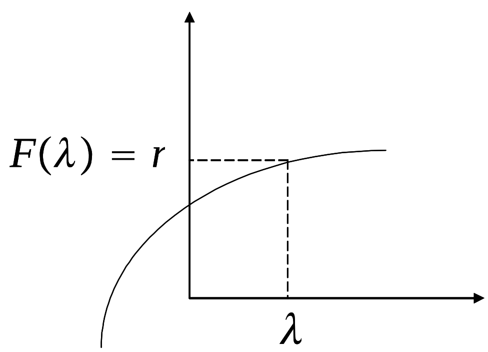
por lo tanto tenemos que buscar $\lambda > 0$ tal que cumpla

$$\begin{aligned} r &= \sum_{k=1}^{n} p_k x_k^* \\ &= \sum_{k=1}^{n} p_k \left(\frac{a_k}{\lambda^* p_k} - b_k \right) \end{aligned}$$

definamos una función $F : R_+ \to R$ dada por

$$F(\lambda) = \sum_{k=1}^{n} p_k \max\left(0, \frac{a_k}{\lambda^* p_k} - b_k\right)$$

gráficamente tendríamos que el problema se reduce a buscar una λ tal que cumpla



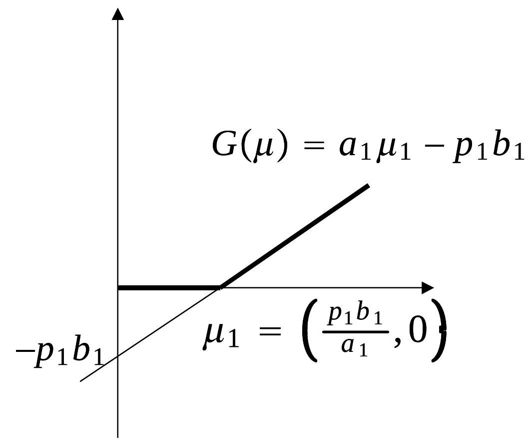
y para simplificar la expresión hagamos el siguiente cambio de variable sea $\mu = \dfrac{1}{\lambda}$ entonces definimos:

$$G(\mu) = \sum_{k=1}^{n} p_k \max\left(0, \frac{a_k}{p_k}\mu - b_k\right).$$

Obsérvese que por hipótesis se tiene a_k , p_k ,$b_k > 0$ de donde la gráfica de $G(\mu)$, por ejemplo para $k = 1$ sería

$$G(\mu) = p_1 \max\left(0, \frac{a_1}{p_1}\mu - b_1\right)$$

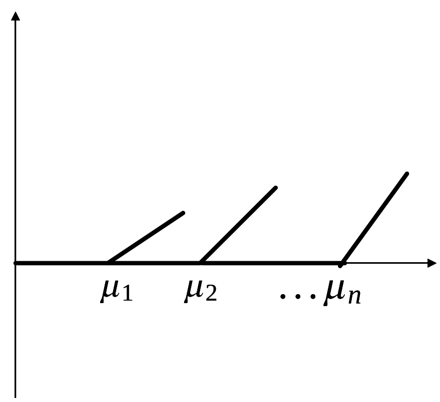
donde tenemos una recta con pendiente positiva a_1 y ordenada al origen $-p_1 b_1$ graficamente se tendría lo siguiente:



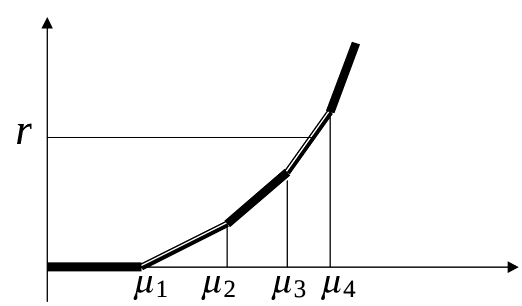
obsérvese también que μ_1 es el punto donde la función se hace cero, es decir sin pérdida de generalidad se tiene que para $k = i$ su cumple $a_i\mu_i - p_ib_i = 0$ de donde $a_i\mu_i = p_ib_i$ por lo que $\mu_i = \dfrac{p_ib_i}{a_i}$ donde μ_j cumple que:

$$0 < \mu_1 \leqq \mu_2 \leqq ..... \leqq \mu_n.$$

Graficamente por separado para cada μ_i con $i = 1, 2, .., n$ se tendría:



se cumple que $0 < \mu_1 \leqq \mu_2 \leqq \mu_3 \leqq ..... \leqq \mu_n$ ya que es la suma de términos no negativos, ahora si se considera la suma $G(\mu)$ graficamente se vería de la siguiente forma:



donde r es la renta del consumidor, por lo que la demanda de éste dependerá de la renta para encontrar el valor de λ, por lo que el punto óptimo sería:

$$x_i^* = \text{máx}\left(0, \frac{a_i}{p_i}\mu - b_i\right).$$

Si la renta es muy pequeña solo demandará de un solo tipo de mercancia, si la renta es mayor entonces podrá demandar más mercancias donde la variable

$$\mu_i = \frac{p_i b_i}{a_i}$$

$p_i =$ precio de la mercancia i
$b_i =$ pondera la satisfacción del consumidor
$a_i =$ pondera las preferencias del consumidor para adquirir la mercancia del tipo i.

Capítulo 7

Conclusiones

En el grupo de trabajo de economía matemática y teoría de juegos de la Facultad de Ciencias, y en particular, en el seminario de tesistas, hacemos un esfuerzo por generar material y sugerir ideas didácticas a los profesores, para la enseñanza de temas de matemáticas con aplicaciones a la economía.

Con ese epíritu, en un principio este material surge con la idea de estar dirigido a estudiantes de economía, sin embargo la primera reflexión que hago con este trabajo, es percatarme de la gran dificultad que existe para diseñar material que lleve de la mano a dichos estudiantes, a abordar temáticas tan complejas como el teorema de Kuhn Tucker.

Por este motivo reduje mis expectativas iniciales, a un trabajo de un nivel intermedio que resulta de utilidad a los estudiantes de las carreras de actuaría y de matemáticas interesados en la problemática económica. Lo anterior lo he constatado a través de mi experiencia como ayudante en la Facultad de Ciencias, pues una buena parte de este trabajo ha sido utilizado en los cursos de economía I y II, así como en seminarios de matemáticas aplicadas a la economía.

Dando como resultado una mejor comprensión de la problemática económica, en particular me refiero a los capítulos 4 y 6, en donde se resuelve el problema del consumidor, pues al ser abordados en clase permiten que los alumnos entiendan cómo se forman las demandas para los diferentes tipos de bienes, y posteriormente al utilizar estos resultados en temáticas como la teoría del equilibrio económico general, los alumnos logran moverse con soltura.

Como es sabido, introducir el aspecto geométrico ayuda al estudiante a aprovechar su intuición y lograr una mayor destreza para comprender y resolver problemas. Con base en ello, uno de mis mayores propósitos fue hacer mucho hincapié en el aspecto geométrico y por este motivo, me dí a la tarea de complementar en la medida de lo posible tanto los ejemplos como la teoría con gráficas.

Con respecto a la colección de ejemplos que se incluyen en el capítulo de motivación, la gran mayoría se plantean sobre problemáticas económicas, además se buscaron de tal forma que no puedan resolverse por sustitución; así los estudiantes se ven obligados a plantear problemas donde aparezcan restricciones. Por otro lado, sin descuidar el propósito de ser pedagógicos, primero se plantean ejemplos sencillos en los que se puede ver su solución geométrica, y posteriormente una vez que se cuenta con la teoría, retomarlos y resolverlos de forma analítica, verificando así la solución. Cabe señalar que no sólo nos remitimos a ejemplos sencillos, pues también se incluyen otros ejemplos más complicados, que sólo pueden ser resueltos de manera analítica, con el objeto de resaltar la utilidad de la teoría que aquí se demuestra.

En el camino, también me encontré con una diversidad muy grande de formas de abordar esta problemática; simplemente al plantear el problema estándar algunos autores lo consideran como maximizar y otros como minimizar, por supuesto todas sus demostraciones varían según el planteamiento del problema. A mi parecer esto lejos de ayudar, complica la comprensión de estos temas; a raíz de ello y para aprovechar que los alumnos conocen la notación usual del cálculo diferencial y el álgebra lineal, todas las demostraciones de los capítulos 2, 3, y 5 están tratadas con dicha notación. Es preciso señalar que los teoremas de programación cóncava usualmente no se incluyen en los cursos de la Facultad de Economía, por lo que está tesis proporciona una alternativa a los profesores que quieran incluír está temática.

Por último quiero hacer énfasis en lo siguiente, después de haber realizado está tesis, me queda la espinita de seguir trabajando para generar material a los economistas, pues considero que si bien las matemáticas no son la única herramienta para los economistas, son de gran utilidad para ayudar a modelar y comprender fenómenos sociales.

Capítulo 8

Referencias

[1] Barbolla, Rosa., Cerdá, Emilio., y Sanz, Paloma., "Optimización. Cuestiones, ejercicios y aplicaciones a la economía", Prentice Hall. España, 2005.

[2] Baumol, William., "Teoría económica y análisis de operaciones", Prentice Hall Internacional. España, 1996.

[3] Courant, Richard., y Fritz, John., "Introducción al cálculo y al análisis matemático vol 1", Limusa Noriega Editores. México, 2003.

[4] Courant, Richard., y Fritz, John., "Introducción al cálculo y al análisis matemático vol 2", Limusa Noriega Editores. México, 2004.

[5] Fryer, MJ and Greenman, JV., "Optimisation Theory. Applications in or and Economics", Edward Arnold. Great Britain,1987.

[6] Henderson, James., Quandt, Richard., "Teoría microeconómica", Editorial Ariel. España, 1979.

[7] Intriligator, Michael D., "Optimización matemática y teoría económica", Prentice Hall Internacional. España, 1973.

[8] Intriligator, Michael D., "Mathematical Programming with Applications to Economics", Handbook of Mathematical Economics, vol. I, Amsterdam, North-Holland, 1981.

[9] Stewart, James., "Calculo Multivariable", Thomson Learning. Colombia, 2004.

[10] Strang, Gilbert., "Álgebra lineal y sus aplicaciones", Addison-Wesley Iberoamericana. México,1986.

[11] Takayama, Akira., "Mathematical Economics", Hinsdale, Illinois, Dryden Press, 1974.

[12] Varian, Hal., "Microeconomía intermedia. Un enfoque actual." Antony Bosh editor. España, 1999.