# Principios de Optimización*

Profesor: Christian Belmar C.
Ayudante de investigación: Mauricio Vargas S.

Universidad de Chile
Facultad de Economía y Negocios

11 de julio de 2011

## Índice



*Material preparado para el curso de Introducción a la Microeconomía. En ese breve documento priman los conceptos teóricos y cómo aplicarlos por sobre la teoría y la formulación matemática. Este apunte se elaboro con el objeto de apoyar el transito desde Introd. a la Microeconomía a Microeconomía I, por tanto recopila los principios fundamentales de la optimización, esperamos que sea una ayuda a su proceso de aprendizaje.

# 1. Cálculo

La idea de esta sección es recordar algunos conceptos de cálculo y esbozar otras que no se ven en los primeros cursos. Para quienes deseen entender los detalles pueden consultar los puntos 1. y 4. de la bibliografía.

## 1.1. Funciones

Intuituivamente (y también de manera imprecisa), una función $y = f(x)$ se refiere a una "regla" de trasformación. Una función corresponde a una relación entre un conjunto $X$ (dominio) y un conjunto $Y$ (imagen) que a cada elemento de $X$, en caso de haber una asociación, asigna un único elemento de $Y$ a un elemento de $X$. Se denota

$$\begin{aligned} f : X &\to Y \\ x &\mapsto f(x) \end{aligned}$$

las funciones que nos interesan, en términos económicos, son las de la forma $f : \mathbb{R} \to \mathbb{R}$ y $f : \mathbb{R}^n \to \mathbb{R}$ principalmente, es decir las que transforman reales en reales y vectores de componentes reales en reales. Sólo en algunos casos (raros) son importantes las funciones $f : \mathbb{R}^n \to \mathbb{R}^m$ que transforman vectores en otros vectores con distinto número de componentes.

## 1.2. Funciones de una variable

Retomando los conceptos anteriores, para las funciones de una variable o de variable real, se hace la distinción de función inyectiva o sobreyectiva. *Inyectividad* se refiere a que dos elementos del dominio no tienen la misma imagen. *Sobreyectividad* se refiere a que todos los elementos del dominio tienen imagen.

No hay que confundir la inyectividad con la diferencia de función o no función. Para fijar ideas, si trazamos una recta vertical al gráfico de una función y efectivamente la recta y el gráfico de la función se intersectan en un único punto, entonces efectivamente se trata de una función. Por otra parte, si se traza una recta horizontal al gráfico de una función y esta intersecta al gráfico de la función en dos o más puntos, entonces la aplicación no es inyectiva.

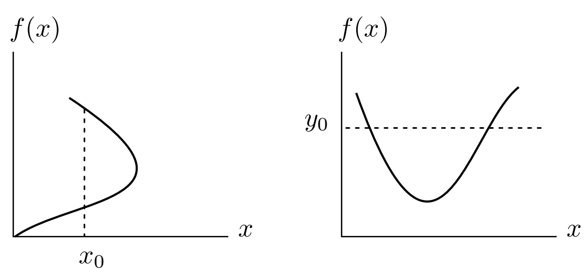


Figura 1: Un caso que no es función y otro de acuerdo al ejemplo.

Una función es *diferenciable* si es continua y a la vez es una función cuyo gráfico describe una curva "suave" (los libros en inglés se refieren al concepto de *smooth functions*), lo cual quiere decir que la función no presenta saltos ni puntas. Retomaremos esta idea tras dar algunos conceptos.

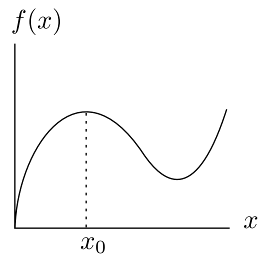


Figura 2: Una función diferenciable y otra que no lo es respectivamente.

La diferenciabilidad es una condición más estricta que la de continuidad. Toda función diferenciable es continua pero la inversa de esta afirmación no siempre es válida, trabajaremos con funciones diferenciables para poder trabajar los problemas con herramientas del cálculo básico.

*Continuidad* se refiere, intuitivamente, a que la función no presenta saltos. En $x_0$ una función $f : D \subset \mathbb{R} \to \mathbb{R}$ es continua si para todo $\varepsilon > 0$ $\exists \delta > 0$ tal que para todo $x \in dom(f)$ se cumple que

$$0 < |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$$

lo cual significa que

$$\lim_{\substack{x \to x_0 \\ x > x_0}} f(x) = \lim_{\substack{x \to x_0 \\ x < x_0}} f(x) = f(x_0)$$

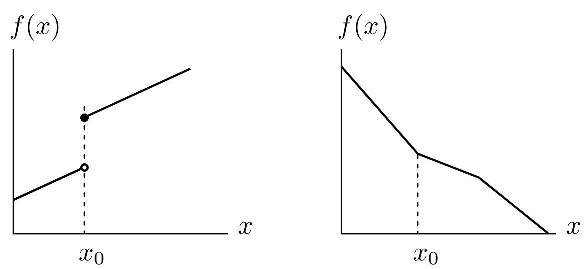


Figura 3: Caso discontinuo y caso continuo respectivamente.

La *derivada* de una función corresponde a la razón de cambio a lo largo de la “curva” que describe una función $f(x)$. La notación es la siguiente

$$\frac{dy}{dx} = f'(x)$$

que es una razón instantánea de cambio, es decir

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

y geométricamente es la recta tangente al gráfico de la función $f(x)$ en $x_0$.

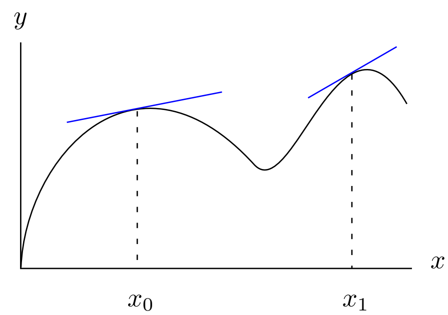


Figura 4: La derivada geométricamente.

si este límite queda bien definido cuando $h \to 0$ independientemente de que si $h > 0$ o $h < 0$ se tiene que $f$ es derivable en $x_0$.

Es importante hacer énfasis en esta idea, en otras palabras estamos diciendo que una función $f$ es diferenciable en $x_0$ si existe $a \in \mathbb{R}$ tal que

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = a$$

y más aún

$$\lim_{\substack{h\to 0\\ h>0}} \frac{f(x_0+h)-f(x_0)}{h} = a^+$$

$$\lim_{\substack{h\to 0\\ h<0}} \frac{f(x_0+h)-f(x_0)}{h} = a^-$$

de manera que $a^- = a^+$, la siguiente figura sirve para fijar ideas de lo que ocurre si la función tiene una "punta" en $x_0$

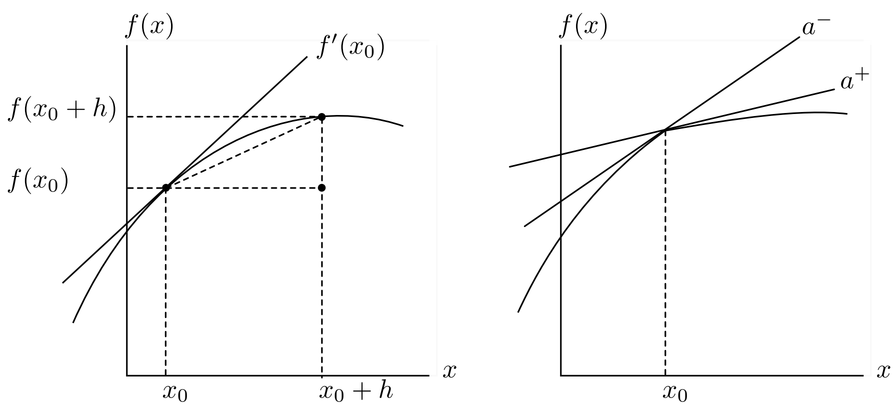


Figura 5: Una función diferenciable y otra que no lo es en $x_0$

Para las funciones de la forma $f : D \subset \mathbb{R} \to \mathbb{R}$ la diferenciabilidad se tiene cuando la tasa de crecimiento es calculable mediante derivadas, es decir derivable es lo mismo que diferenciable en tal caso. Una función de variable real es diferenciable si existe

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}$$

que implica

$$\lim_{x \to x_0} \frac{|f(x_0+h) - f(x_0) - f'(x_0)h|}{|h|} = 0$$

de manera alternativa, podemos decir que $f$ es diferenciable en $x_0$ si y sólo si existe $f'(x_0) \in \mathbb{R}$ tal que

$$f(x_0+h) = f(x_0) + f'(x_0)h + \theta(h)$$

donde $\theta(h) = f(x_0+h) - f(x_0) - f'(x_0)h$ que cumple

$$\lim_{h \to 0} \frac{|\theta(h)|}{|h|} = 0$$

Lo anterior da cuenta de que la función $\theta(h)$ en valor se acerca más rápido al valor 0 en comparación a la función $f(h) = h$. Recordemos algunas reglas básicas de derivadas:

1. $f(x) = c \Rightarrow f'(x) = 0$
2. $f(x) = ax^b \Rightarrow f'(x) = ba \cdot x^{b-1}$
3. $f(x) = e^x \Rightarrow f'(x) = e^x$
4. $f(x) = \ln(x) \Rightarrow f'(x) = 1/x$
5. $f(x) = a^x \Rightarrow f'(x) = a^x \cdot \ln(a)$
6. Si $f$ y $g$ son dos funciones diferenciables, entonces
   - $(f+g)'(x) = f'(x) + g'(x)$
   - $(f \cdot g)(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$
   - $\left(\frac{f}{g}\right)'(x) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{[g(x)]^2}$ (con $g(x) \neq 0$)
   - Regla de la cadena:
     $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$

**Ejemplo 1.1.** Si $C(x)$ es la función de costo de una empresa por producir $x$ unidades de un producto. Al aumentar la producción de $x_0$ a $x_1$ el costo adicional es

$$\Delta C = C(x_1) - C(x_0)$$

y la tasa promedio del cambio del costo es

$$\frac{\Delta C}{\Delta x} = \frac{C(x_1) - C(x_0)}{x_1 - x_0}$$

tomando $\Delta x = dx \approx 0$ se tiene la razón de cambio instantáneo del costo respecto a la cantidad producida, esto es el *costo marginal*. Supongamos que la función $f$ es diferenciable, entonces

$$CMa(x) = C'(x) \approx \frac{dC}{dx}$$

tomando $\Delta x = 1$ y $n \to \infty$, es razonable suponer que 1 es mucho más pequeño que $n$ y por lo tanto

$$C'(n) \approx \frac{C(n+1) - C(n)}{1} = C(n+1) - C(n)$$

en el caso en que asumimos $\Delta x = 1$ se tiene que $dC \approx C'(n)$ entonces el costo marginal para producir $n$ unidades es aproximadamente igual al costo para producir una unidad más (la unidad $n+1$). En el caso en que el aumento es distinto de 1 unidad la expresión encontrada para el costo marginal nos dice que el cambio en el costo está dado por

$$dC \approx C'(x)dx$$

Si la primera derivada de una función diferenciable es también una función diferenciable, se puede derivar nuevamente viendo la derivada como otra función, es decir

$$\frac{d^2y}{dx^2} = f''(x)$$

La segunda derivada, nos da entonces, la razón intantánea de cambio de la primera derivada, es decir, da cuenta de la razón de cambio de las primeras derivadas y esto nos dice si la función $f(x)$ es creciente o decreciente a tasa creciente o decreciente.

### Funciones cóncavas

Consideremos una función $f(x)$ diferenciable. La primera derivada está dada por

$$f'(x_0) = \frac{f(x) - f(x_0)}{x - x_0}$$

reordenando esto obtenemos

$$f(x) = f(x_0) + f'(x_0)(x - x_0)$$

De acuerdo a esto, una función diferenciable es *cóncava* si dada una recta tangente $l(x)$ a $f(x)$ se tiene que

$$l(x) = f(x_0) + f'(x_0)(x - x_0)$$

y además la recta tangente indica que la función no toma valores mayores a los de la recta tangente, es decir $l(x) \geq f(x)$, lo cual geométricamente indica que la función es creciente a tasa decreciente por lo que $f''(x) \leq 0$ (si $f''(x) < 0$ entonces la función es *estrictamente cóncava*).
Estos hechos nos llevan a que una función cóncava es tal que

$$f(x) \leq f(x_0) + f'(x_0)(x - x_0)$$

geométricamente se tiene lo siguiente

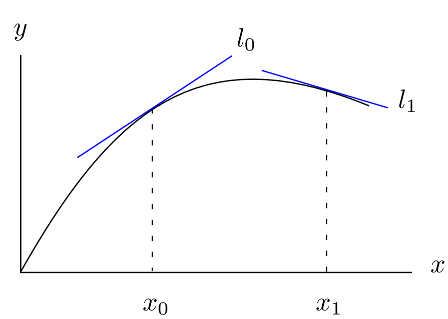


Figura 6: Concavidad (curvatura)

Cuando $f$ es diferenciable, en el caso de una variable basta con que $f''(x) \leq 0$ para determinar que $f$ es cóncava y si $f''(x) < 0$ entonces $f$ es estrictamente cóncava.

### Funciones convexas

De manera contraria al concepto anterior, una función diferenciable es *convexa* si dada una recta tangente $l(x)$ a $f(x)$ se tiene que

$$l(x) = f(x_0) + f'(x_0)(x - x_0)$$

y además la recta tangente indica que la función no toma valores menores a los de la recta tangente, es decir $l(x) \leq f(x)$, lo cual geométricamente indica que la función es creciente a tasa decreciente por lo que $f''(x) \geq 0$ (si $f''(x) > 0$ entonces la función es *estrictamente cóncava*).
Estos hechos nos llevan a que una función cóncava es tal que

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0)$$

geométricamente se tiene lo siguiente

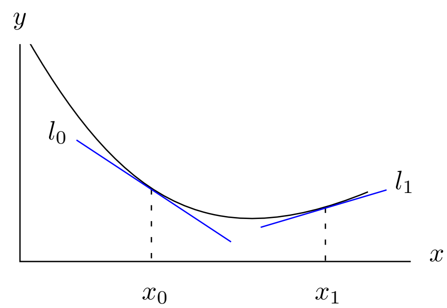


Figura 7: Convexidad (curvatura)

Cuando $f$ es diferenciable, en el caso de una variable basta con que $f''(x) \geq 0$ para determinar que $f$ es convexa y si $f''(x) > 0$ entonces $f$ es estrictamente convexa.

## 1.3. Funciones de varias variables

A lo largo del curso nos centraremos en el caso de las funciones de dos variables y, en ocasiones, en el caso de funciones de varias variables. Llamaremos a $f(x)$, con $x \in \mathbb{R}^n$ función a valores en $\mathbb{R}^m$ si $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$.
El argumento de $f(x)$ es un vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ y la imagen de $x$ es un vector de $\mathbb{R}^m$. Así $f(x) = (f_1(x), \cdots, f_m(x))$, donde las funciones $f_i : D \subseteq \mathbb{R} \to \mathbb{R}$, para cada $i$, se conocen como funciones coordenadas.

Una *bola abierta* de centro en $x_0$ y radio $r$ es un conjunto definido por

$$B(x_0, r) = \{x \in \mathbb{R}^n : \|x - x_0\| < r\}$$

donde $\| \cdot ||$ es la norma euclídea definida por

$$\|x\| = \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

A menos que se indique lo contrario, cuando nos refiramos a que las funciones transforman elementos de $D$ se asume que $D$ es un conjunto abierto para evitar problemas con las derivadas en los extremos del conjunto.

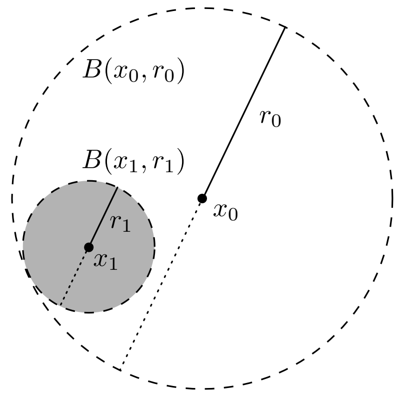


Figura 8: Idea geométrica de bola abierta.

En el caso de funciones de una variable teníamos dos direcciones para aproximarnos a $x_0$ en la medida que hacíamos tender un valor $h \to 0$. Para el caso de funciones de varias variables la idea de acercanos desde dos direcciones a un punto se desvanece. Con un caso de dos variables la idea de límite por la izquierda o por la derecha carece de sentido. De esta forma la definición de bola abierta nos ahorra el problema que surge al intentar acotar el límite de la función practicamente en infinitas direcciones, de manera más precisa a partir del concepto de bola abierta tenemos que si $f$ está definida en un conjunto abierto que contiene a $x_0$ nos ahorramos el problema señalado con las derivadas laterales.

Centrándonos en el caso de funciones de varias variables a valores reales, es decir funciones de la forma $D \subset \mathbb{R}^n \to \mathbb{R}$. Para $1 \leq j \leq n$ fijo, definimos la función:

$$\begin{aligned} f_j : \mathbb{R} &\to \mathbb{R} \\ h &\mapsto f(x + he_j) \end{aligned}$$

donde $e_j$ es la $j$-ésima componente canónica de $\mathbb{R}^n$. Hay que tener presente que

$$x + he_j = (x_1, \ldots, x_{j-1}, x_j + h, x_{j+1}, \ldots, x_n)$$

Luego, con este desarrollo podemos aplicar cualquier regla de diferenciabilidad a las funciones componentes $f_j(x)$ de $f(x)$ y las variables $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ se mantienen fijas para trabajar todo lo demás como constante.

Se define la *derivada parcial* de $f(x)$ con respecto a $x_j$ en $x_0 \in \mathbb{R}^n$ como

$$\frac{\partial f}{\partial x_j}(x_0) = \lim_{h \to 0} \frac{f(x_0 + he_j) - f(x_0)}{h}$$

Cuando la derivada parcial de $f$ con respecto a $x_j$ existe en todo punto de $D$, entonces define una función

$$\frac{\partial f}{\partial x_j} : D \subset \mathbb{R}^n \to \mathbb{R}$$

**Ejemplo 1.2.** Sea $f(x_1, x_2) = x_1^2 + 3x_1x_2 - x_2^2$. Calcular las derivadas parciales

Solución. Las derivadas parciales son

$$\frac{\partial f}{\partial x_1} = 2x_1 + 3x_2 \qquad \frac{\partial f}{\partial x_2} = 3x_1 - 2x_2$$

es decir, hemos considerado que sólo cambia un argumento y así podemos aplicar las reglas del cálculo de una variable. También es importante señalar que las derivadas parciales son funciones de $x_1$ y $x_2$ □

Las derivadas parciales dan cuenta de si la función es creciente o decreciente en el argumento $x_j$. De la misma forma, si las derivadas parciales definen funciones diferenciables, las segundas derivadas parciales de la función dan cuenta de que si las funciones son crecientes o decrecientes a tasa creciente o decreciente en $x_j$.

Para el caso de funciones de la forma $D \subset \mathbb{R}^n \to \mathbb{R}$ la idea de diferenciabilidad se refiere al concepto de mejor aproximación lineal de la función. Una función lineal $L : \mathbb{R}^n \to \mathbb{R}$ es toda función de la forma

$$L(x) = a \cdot x$$

donde $a = (a_1, \ldots, a_n)$ es un vector cuyas componentes están formadas por escalares y $x \in \mathbb{R}^n$. En la literatura se hace la distinción con la función lineal afín, que es de la forma

$$A(x) = a \cdot x + b$$

donde $b \in \mathbb{R}$ y así cuando $b = 0$ se tiene que la función lineal. De esta definición se concluye directamente que

$$b = A(x_0) - a \cdot x_0$$

y de esto se concluye

$$A(x) = A(x_0) + a \cdot (x - x_0)$$

Haciendo la analogía con el caso de funciones de variable real, la función lineal nos sirve para determinar que una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $x_0$ si y sólo si existe $f'(x_0) \in \mathbb{R}^n$ tal que

$$f(x_0 + h) = f(x_0) + f'(x_0) \cdot h + \theta(h)$$

donde $\theta(h) = f(x_0 + h) - f(x_0) - f'(x_0) \cdot h$ que cumple

$$\lim_{h \to 0} \frac{\|\theta(h)\|}{\|h\|} = 0$$

lo cual intuitivamente da la misma interpretación que en el caso de una función de variable real dado que $f$ tiene valores reales. Notemos que $f'(x_0)$ en este caso corresponde a un vector en $\mathbb{R}^n$ por como se define $f$, en caso de que este vector exista se denomina *gradiente* de $f$ en $x_0$ y es el vector derivada de la función y se denota

$$\nabla f(x_0) = \left( \frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0) \right)$$

El gradiente será de gran importancia cuando veamos problemas con restricciones y de maximización en general.

Por ahora, tendremos que dejar el caso de funciones a variable real y tomar el caso $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$. Pensemos en una función $F(x) = (F_1(x), \ldots, F_m(x))$. A partir del concepto de gradiente se tiene directamente el de *Matriz Jacobiana* que es una matriz definida por

$$DF(x_0) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(x_0) & \ldots & \frac{\partial F_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(x_0) & \ldots & \frac{\partial F_m}{\partial x_n}(x_0) \end{bmatrix}$$

con $i = \{1, \ldots, m\}$ y $j = \{1, \ldots, n\}$

Retomando las funciones a variable real, una buena aproximación lineal de $f$ está dada por la función lineal afín

$$A(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0)$$

Una función diferenciable es aquella cuyas derivadas parciales son continuas, pero esto es una condición suficiente y no necesaria. La continuidad de las derivadas parciales no significa que la función sea diferenciable. El concepto intuitivo de diferenciabilidad es que si existe una aproximación lineal de la función, dada por un plano, entonces la distancia del plano al gráfico de la función es una distancia pequeña.

Dado $x \in D$ y si existe $\varepsilon > 0$ podemos definir

$$g : (-\varepsilon, \varepsilon) \to \mathbb{R}$$
$$h \mapsto f(x + he_j) \in D$$

luego

$$\frac{\partial f(x)}{\partial x_j} = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = g'(0)$$

que corresponde a otra forma de expresar la derivada de una función componente.

Decimos que $f$ es *diferenciable* en $x_0$ si en $B(x_0, r)$ las derivadas parciales existen y son continuas, y además se tiene que

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - \nabla f(x_0) \cdot h\|}{\|h\|} = 0$$

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - \sum_{j=1}^{n} \frac{\partial f(x_0)}{\partial x_j}(x - x_0)\|}{\|h\|} = 0$$

Esto es, para que $f$ sea diferenciable su gráfico y el plano tangente al gráfico en el punto $x$ estén tan cerca que al dividir su distancia por $\|h\| = \|x - x_0\|$ la fracción todavía es pequeña.

La *derivada direccional* de $f$ está dada por

$$f'(x_0|v) = \frac{d}{dt} f(x_0 + tv)\Big|_{t=0} = \lim_{t \to 0} \frac{f(x_0 + tv - f(x_0)}{t}$$

intuitivamente este concepto se refiere a que en $x_0$ existe un vector que da cuenta del cambio en la función y apunta en una dirección $v$ (nótese que $v \in \mathbb{R}^n$), es decir, da cuenta de como cambia la función en la medida que nos movemos de $x_0$ a $v$.

Siguiendo el argumento, fijemos $x_0 \in \mathbb{R}^n$ y nos interesa saber como cambia el valor de $f(x)$ en la medida que nos movemos de $x_0$ en la dirección $v$ tal como en el caso de la derivada direccional. Definamos $g : \mathbb{R} \to \mathbb{R}$

$$g(t) = f(x_0 + tv)$$

con $t \in \mathbb{R}$, para $t = 0$ se tiene que $g(0) = f(x_0)$ y de esta forma si $g'(0)$ es creciente sabemos que $f(x_0)$ es creciente en la medida que nos movemos de $x_0$ a $v$. Como $g$ es una función de variable real, las reglas de derivadas vistas anteriormente nos sirven.

Si $t$ aumenta en una magnitud $\Delta t \to 0$, entonces estamos ante un problema de derivadas parciales cuando lo que nos interesa es ver como cambia $f$, de esta forma cambia cada una de las funciones componentes independientemente de las demás y así el cambio producido es la suma de todas las derivadas parciales de $f$ debido al manejo que se puede hacer con las funciones componentes. Esto proviene de

$$g(t) = f(x_0 + tv)$$

si derivamos con respecto a $t$

$$g'(t) = f'(x_0 + tv) \cdot v$$

evaluando en $t = 0$

$$\begin{aligned} g'(0) &= f'(x_0) \cdot v \\ g'(0) &= \nabla f(x_0) \cdot v \\ g'(0) &= \sum_{j=1}^{n} \frac{\partial f(x_0)}{\partial x_j} v_j \end{aligned}$$

y esto nos da un argumento para que el *diferencial*, que corresponde al cambio total de una función, se denote

$$df(x_0) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0) dx_j$$

**Ejemplo 1.3.** En el caso de que la utilidad asociada al consumo de $n$ productos es una función diferenciable $u : \mathbb{R}^n \to \mathbb{R}$, el cambio total ante cambios en el consumo de los distintos productos es

$$du = \frac{\partial u}{\partial x_1} dx_1 + \ldots + \frac{\partial u}{\partial x_2} dx_2 = \sum_{i=1}^{n} \frac{\partial u}{\partial x_i} dx_i$$

**Ejemplo 1.4.** Debido a los resultados del último torneo nacional de fútbol, el grandioso equipo mágico **Universidad de Chile** ha aumentado su popularidad y tiene nuevos seguidores (los que en verdad son hinchas apoyan al equipo siempre). Lo que no fue comentado por la prensa es el equipo rival, el cual no tuvo una buena actuación en el último torneo, tiene pocos hinchas de corazón y, lo que es sabido por todos, que muchos de los supuestos amantes del fútbol de este equipo en verdad son delincuentes que le dan mala fama. Ante esto los vecinos de Ñuñoa, preocupados de que haya seguridad en los estadios pero en su punto eficiente, han contratado a un profesor de microeconomía que determinó que el gasto en seguridad, en millones de pesos, está dado por

$$E = c(U, C) + g(U, C) + z(U, C)$$

donde $c$ = carabineros, $g$ = guanacos y $z$ = zorrillos corresponden a las funciones

$$\begin{aligned} c(U, C) &= 30C^{0,7} + 10U^{0,3} \\ g(U, C) &= 60C^{0,5} + 10U^{-0,5} \\ z(U, C) &= 100C^{0,4} \end{aligned}$$

Además, este profesor pidió a su ayudante que determinara cuánto cambió la cantidad de seguidores y se determinó que $dU = 3300$ y $dC = -2500$. Las estimaciones de público de este año indican que $U_0 = 20000$ y $C_0 = 17000$ en promedio.

Determine el cambio total en el gasto. ¿Qué había pasado si los cambios en cantidad hubieran sido al revés?

Solución. Calculamos las derivadas parciales

$$\begin{aligned}
\frac{\partial c(U,C)}{\partial C} &= 21C^{-0,3} & \frac{\partial c(U,C)}{\partial U} &= 30U^{-0,7}\\
\frac{\partial g(U,C)}{\partial C} &= 30C^{-0,5} & \frac{\partial c(U,C)}{\partial U} &= -5U^{-1,5}\\
\frac{\partial z(U,C)}{\partial C} &= 40C^{-0,6} & \frac{\partial c(U,C)}{\partial U} &= 0
\end{aligned}$$

Luego el diferencial está dado por

$$dE = \left(\frac{\partial c(U,C)}{\partial C} + \frac{\partial g(U,C)}{\partial C} + \frac{\partial z(U,C)}{\partial C}\right) dC + \left(\frac{\partial c(U,C)}{\partial U} + \frac{\partial g(U,C)}{\partial U} + \frac{\partial z(U,C)}{\partial U}\right) dU$$

debemos evaluar las derivadas parciales en $(C_0, U_0)$ y ya conocemos la magnitud de los cambios, entonces tenemos

$$\begin{aligned}
\frac{\partial c(U,C)}{\partial C} &= 1,1300 & \frac{\partial c(U,C)}{\partial U} &= 0,0029\\
\frac{\partial g(U,C)}{\partial C} &= 0,0728 & \frac{\partial c(U,C)}{\partial U} &= -1,7677 \cdot 10^{-6}\\
\frac{\partial z(U,C)}{\partial C} &= 0,1158 & \frac{\partial c(U,C)}{\partial U} &= 0
\end{aligned}$$

$$\begin{aligned}
\left(\frac{\partial c(U,C)}{\partial C} + \frac{\partial g(U,C)}{\partial C} + \frac{\partial z(U,C)}{\partial C}\right) dC &= -3296,4966\\
\left(\frac{\partial c(U,C)}{\partial U} + \frac{\partial g(U,C)}{\partial U} + \frac{\partial z(U,C)}{\partial U}\right) dU &= 9,6528
\end{aligned}$$

entonces

$$dE = -3296,4966 + 9,6528 = -3286,8438$$

el gasto disminuye pese a que van más hinchas de la U que del otro equipo. Esto puede explicarse por varias razones como que los hinchas de la U tienen un mejor comportamiento en el estadio y que además como los hinchas del otro equipo dejan de ir, de inmediato se reduce el gasto que ellos generan.

Si el cambio hubiera sido al revés tendríamos que $dE = 4344$. Hay que distinguir el cambio total en el gasto y el gasto que se atribuye a cada equipo, por ejemplo si asisten 3300 hinchas más de la U generan un gasto de $9,65$ millones de pesos pero si van 3300 hinchas más del otro equipo generan un gasto de $4351,38$ millones de pesos. □

Nota 1.1. Por el hecho de que el diferencial de $f$ proviene del concepto de derivada direccional, no es correcto calcular el diferencial de antemano si no sabemos si la función es diferenciable. Esta observación es importante porque pese a que en muchos casos se puede efectuar el desarrollo algebraico, se comente un error conceptual si se calcula el diferencial en un contexto en que las condiciones matemáticas invalidan el procedimiento.

**Ejemplo 1.5.** Consideremos la función Cobb-Douglas $f(x_1, x_2) = x_1^{1/2} x_2^{1/2}$. Queremos saber si es diferenciable en $(0,0)$.

Solución. La función es continua en el plano, tiene derivadas parciales en todos los puntos de $\mathbb{R}^2$, pero estas derivadas parciales son continuas en $(0,0)$ y aún así la función no es diferenciable en $(0,0)$.

Analizaremos por partes esta última afirmación.

La función $f$ es continua en $(0,0)$ si $f(x_{01}, x_{02}) = f(0,0)$ cuando $(x_{01}, x_{02}) \to (0,0)$, en efecto

$$\lim_{(x_{01},x_{02})\to(0,0)} f(x_{01}, x_{02}) = 0$$

$$|f(x_{01}, x_{02}) - f(0,0)| = |x_{01}^{1/2} x_{02}^{1/2}| \to 0$$

Las derivadas parciales de $f$ son continuas. Analizaremos sólo con respecto a $x_1$ pues el otro caso es idéntico

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \frac{1}{2}\left(\frac{x_2}{x_1}\right)^{1/2}$$

luego

$$\lim_{(x_{01},x_{02})\to(0,0)} \frac{1}{2}\left(\frac{x_{02}}{x_{01}}\right)^{1/2} = \frac{1}{2}$$

$$\left|\frac{\partial f}{\partial x_1}(x_{01}, x_{02}) - \frac{\partial f}{\partial x_1}(0,0)\right| \to 0$$

este cálculo se fundamenta en que $x_{01} \approx x_{02}$ en torno a $(0,0)$

Tenemos las siguientes propiedades, de fácil verificación, que no son conclusivas sobre la diferenciabilidad de la función $f$ en $(0,0)$.

1. Derivadas parciales continuas en $(0,0) \Rightarrow$ diferenciabilidad en $(0,0)$.

   No nos sirve para concluir diferenciabilidad, pues las derivadas parciales no son continuas. Tampoco nos sirve para mostrar que no es diferenciable, pues se trata de una condición suficiente para diferenciabilidad, no de una condición necesaria.

2. Diferenciabilidad en $(0,0) \Rightarrow$ las derivadas parciales existen en $(0,0)$.

   No nos sirve para probar que la función no es diferenciable, pues las derivadas parciales existen en todos los puntos. Tampoco sirve para probar diferenciabilidad, ya que es una propiedad que caracteriza una condición necesaria y no suficiente para diferenciabilidad.

Una propiedad conclusiva sobre la diferenciabilidad es la siguiente: Si $f$ es diferenciable en $(0,0)$ entonces $f$ tiene derivadas direccionales bien definidas en $(0,0)$.

Sea $v = (v_1, v_2)$ con $\|v\| = 1$, para $t \neq 0$ la derivada direccional está dada por

$$\frac{f[(0,0) + t(v_1, v_2)] - f(0,0)}{t} = v_1^{1/2} v_2^{1/2}$$

por lo tanto

$$f'(0,0) = v_1^{1/2} v_2^{1/2}$$

Evaluando en $v = (1,0)$

$$\frac{\partial f}{\partial x_1}(0,0) = 0$$

y evaluando en $v = (0,1)$

$$\frac{\partial f}{\partial x_2}(0,0) = 0$$

que es consistente con lo que habíamos determinado anteriormente.

Si $f$ fuera diferenciable en $(0,0)$ se tendría que $f'(0,0)=0$ dado cualquier $v$, pero dado $v=(1,1)$

$$f'(0,0)=1$$

en conclusión, $f$ no es diferenciable. En caso de intentar calcular el diferencial, siempre se obtendrá un valor cero para cualquier $v$.

Para fijar ideas, uno esperaría que el plano tangente a la función $f$ en el punto $(x_1,x_2,y)=(0,0,0)$ estuviera dado por la ecuación $y=0$, para ser consecuentes con el valor de las derivadas parciales de $f$ en el $(0,0)$. Sin embargo, este plano no puede ser tangente al gráfico de la función en $(0,0)$, pues sobre la recta $x_1=x_2$ el gráfico de $f$ tiene pendiente infinita en el origen. Veamos si la condición del límite se cumple para determinar si la función es diferenciable

$$\lim_{(x_{01},x_{02})\to(0,0)} = \frac{|(x_{01}x_{02})^{1/2}-\frac{1}{2}x_{0,1}^{-1/2}x_{02}^{1/2}\cdot x_{01}-\frac{1}{2}x_{01}^{1/2}x_{0,2}^{-1/2}\cdot x_{02}|}{\|(x_{01},x_{02})-(0,0)\|}$$

Para $(x_{0,1},x_{02})\neq(0,0)$ cada una de las funciones componentes es continua, entonces por álgebra de funciones continuas basta con analizar sólo una (en este caso para $x_1$)

$$\lim_{(x_1,x_2)\to(0,0)} = \frac{x_1^{2/3}-\frac{2}{3}x_1^{2/3}}{\sqrt{2}x_1} = \frac{\frac{1}{3}x_1^{2/3}}{\sqrt{2}x_1} = \frac{1}{3\sqrt{2}}\cdot\frac{1}{x_1^{1/3}}\neq 0$$

de esta forma lo que se obtiene es una demostración de que las derivadas parciales no son continuas en cero. Como se vió más arriba, esto no es conclusivo para diferenciabilidad. □

### Extendiendo el concepto de segunda derivada

Las segundas derivadas parciales de $f(x)$ corresponden a

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f(x_0)}{\partial x_j}\right)=\frac{\partial^2 f(x_0)}{\partial x_i\partial x_j}$$

es importante señalar que esta notación indica que primero se deriva parcialmente con respecto a $x_j$ y luego con respecto a $x_i$. Para $\frac{\partial f(x_0)}{\partial x_i}$ tenemos que el gradiente corresponde a

$$\nabla\left(\frac{\partial f}{\partial x_i}\right)(x_0)=\left(\frac{\partial^2 f(x_0)}{\partial x_i\partial x_1},\ldots,\frac{\partial^2 f(x_0)}{\partial x_i\partial x_n}\right)$$

El gradiente de $f$ consta de $n$ derivadas parciales y el gradiente de las derivadas parciales de $f$ consta de $n$ segundas derivadas parciales. Si reunimos los gradientes de todas las segundas derivadas parciales se obtiene

$$Hf(x)=\begin{bmatrix}\nabla\left(\frac{\partial f}{\partial x_1}\right)\\ \vdots\\ \nabla\left(\frac{\partial f}{\partial x_n}\right)\end{bmatrix}=\begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_n\partial x_1}\\ \vdots & \ddots & \vdots\\ \frac{\partial^2 f}{\partial x_1\partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{bmatrix}$$

esto se conoce como *matriz Hessiana* de $f$ y es un concepto similar al de la segunda derivada en el caso de funciones de una variable, ya que considera todos los efectos de las variables sobre el crecimiento o decrecimiento de la función. Retomaremos esta definición más adelante y nos será sumamente útil en problemas con y sin restricciones. Por ahora, diremos que esta matriz contiene una cantidad de derivadas parciales igual a $n^2$ y es una matriz simétrica debido al siguiente teorema:

**Teorema 1.1.** (Teorema de Schwarz)
Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ una función dos veces continua y diferenciable en $D$. Si las segundas derivadas parciales de $f$ son continuas en $x_0 \in D$ entonces

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) \qquad \forall i, j = 1, \cdots, n$$

**Ejemplo 1.6.** Sea $f(x_1, x_2) = x_1 x_2^2 + x_1 x_2$. Probar que la matriz Hessiana es simétrica.

Solución. Las derivadas parciales son

$$\frac{\partial f}{\partial x_1} = x_2^2 + x_2 \qquad \frac{\partial f}{\partial x_2} = 2x_1 x_2 + x_1$$

entonces, las segundas derivadas parciales son

$$\frac{\partial^2 f}{\partial x_1^2} = 0 \qquad \frac{\partial^f}{\partial x_2^2} = 2x_1 \qquad \frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial^2 f}{\partial x_1 \partial x_2} = 2x_2 + 1$$

de esta forma la matriz Hessiana corresponde a

$$Hf(x) = \begin{bmatrix} 0 & 2x_2 + 1 \\ 2x_2 + 1 & 2x_1 \end{bmatrix}$$

y queda en evidencia que $\dfrac{\partial^2 f}{\partial x_2 \partial x_1} = \dfrac{\partial^2 f}{\partial x_1 \partial x_2}$. □

#### Funciones cóncavas

Por definición una función $f$ es *cóncava* si se cumple que dado $\alpha \in [0, 1]$

$$f(\alpha x_1 + (1 - \alpha)x_2) \geq \alpha f(x_1) + (1 - \alpha)f(x_2)$$

Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$. Se dice que $f$ es cóncava si para todo $x \in D$ y todo $v \in \mathbb{R}^n$ tal que $v \neq$, la función $g : \mathbb{R} \to \mathbb{R}$ definida por

$$g(t) = f(x + tv)$$

es cóncava en $t \in \mathbb{R}$ tal que $x + tv \in D$. Para verificar esto supongamos que $g$ es cóncava, entonces

$$g(\alpha t_1 + (1 - \alpha)t_2) \geq \alpha g(t_1) + (1 - \alpha)g(t_2)$$

por definición

$$g(\alpha t_1 + (1 - \alpha)t_2) = f(x + [\alpha t_1 + (1 - \alpha)t_2]v)$$
$$g(\alpha t_1 + (1 - \alpha)t_2) = f[\alpha(x + t_1 v) + (1 - \alpha)(x + t_2 v)]$$
$$g(\alpha t_1 + (1 - \alpha)t_2) \geq \alpha f(x + t_1 v) + (1 - \alpha)f(x + t_2 v)$$

$$\alpha f(x+t_1v)+(1-\alpha)f(x+t_2v)=\alpha g(t_1)+(1-\alpha)g(t_2)$$

entonces,

$$g(\alpha t_1+(1-\alpha)t_2)\geq \alpha g(t_1)+(1-\alpha)g(t_2)$$

Asumiendo que $f$ es diferenciable. Un criterio más eficiente es tener presente que una matriz $A$ de $n\times n$ es semidefinida negativa si dado $v\in\mathbb{R}^n$ se cumple que

$$v^TAv\leq 0$$

y en base a esto una función $f$ es cóncava si su Hessiano es semi definido negativo, esto es

$$v^THf(x)v\leq 0$$

para determinar si el Hessiano es semi definido negativo nos servirá el siguiente criterio: $Hf(x)$ es semi definida negativa si los signos de los determinantes son alternados, comenzando con negativo, y se tiene $|H_1|\leq 0, |H_2|\geq 0,\dots$ tal que

$$(-1)^i|H_i|\geq 0 \qquad \forall i=\{1,\dots,n\}$$

Los determinantes de las submatrices de $Hf(x)$ corresponden a

$$H_1=\frac{\partial^2 f(x)}{\partial x_1^2}$$

$$H_2=\begin{vmatrix}\frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1\partial x_2}\\ \frac{\partial^2 f(x)}{\partial x_2\partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2}\end{vmatrix}$$

$$H_n=\begin{vmatrix}\frac{\partial^2 f(x)}{\partial x_1^2} & \dots & \frac{\partial^2 f(x)}{\partial x_n\partial x_1}\\ \vdots & \ddots & \vdots\\ \frac{\partial^n f(x)}{\partial x_1\partial x_n} & \dots & \frac{\partial^2 f(x)}{\partial x_n\partial x_n}\end{vmatrix}$$

Si una función es estrictamente cóncava, entonces su Hessiano es definido negativo, esto es

$$v^THf(x)v<0 \text{ y por ende } (-1)^i|H_i|>0$$

Para convencerse de esto tengamos presente que siendo $f$ cóncava, entonces $g''(t)\leq 0$ y $g(t)\leq g(t_0)+g'(t_0)(t-t_0)$. Se tiene que

$$g'(t)=\nabla f(x+tv)\cdot v$$

$$g'(t)=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x+tv)v_j$$

derivando $\frac{\partial f}{\partial x_j}(x+tv)$ con respecto a $t$ se obtiene

$$\frac{\partial^2 f}{\partial x_i\partial x_j}(x+tv)=\sum_{i=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(x+tv)v_i$$

para considerar el efecto de todas las variables se debe multiplicar por $v_j$ y agregar los $j$ términos

$$g''(t) = \sum_{j=1}^{n}\sum_{i=1}^{n} v_j v_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x+tv)v_i$$

$$g''(t) = v^T Hf(x+tv)v$$

adicionalmente existen $x, v$ tales que $0 = x + tv$, lo que permite concluir que

$$g''(0) \leq 0$$

$$v^T Hf(x+tv)v \leq 0$$

#### Funciones convexas

De manera similar a lo anterior, por definición una función $f$ es *convexa* si se cumple que dado $\alpha \in [0,1]$

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2)$$

Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$. Se dice que $f$ es convexa si para todo $x \in D$ y todo $v \in \mathbb{R}^n$ tal que $v \neq$, la función $g : \mathbb{R} \to \mathbb{R}$ definida por

$$g(t) = f(x+tv)$$

es convexa en $t \in \mathbb{R}$ tal que $x + tv \in D$.

Asumiendo que $f$ es diferenciable. Un criterio más eficiente es tener presente que una matriz $A$ de $n \times n$ es semidefinida positiva si dado $v \in \mathbb{R}^n$ se cumple que

$$v^T Av \geq 0$$

y en base a esto una función $f$ es cóncava si su Hessiano es semi definido positivo, esto es

$$v^T Hf(x)v \geq 0$$

para determinar si el Hessiano es semi definido positivo nos servirá el siguiente criterio: $Hf(x)$ es semi definida positiva si los signos de los determinantes son todos positivos y se tiene $|H_1| \geq 0, |H_2| \geq 0, \ldots$ tal que

$$|H_i| \geq 0 \qquad \forall i = \{1, \ldots, n\}$$

si una función es estrictamente cóncava, entonces su Hessiano es definido positivo, esto es

$$v^T Hf(x)v > 0 \text{ y por ende } |H_i| > 0$$

## 1.4. Función Implícita

Una *Función Implícita* es una expresión de la forma

$$F(x) = 0$$

la cual no siempre expresa una variable en función de otra dado que no siempre es posible “despejar” y expresar una variable en función de otra u otras.

Tengamos presente que muchas funciones implícitas no definen funciones tras despejar uan variable en función de otra. Por ejemplo, la circunferencia de centro $(h,k)$ y radio $r$ dada por

$$(x_1-h)^2+(x_2-k)^2-r^2=0$$

es tal que siempre existen $(x_{01},x_{02})$ que cumplen la ecuación pero si se obtiene $x_2(x_1)$ a algunos valores de $x_1$ les corresponden dos valores de $x_2$, de hecho a todos los que no se encuentren a una distancia $r$ del punto $(h,k)$. Veamos esto en detalle: Para $x_2>0$, en una bola $B(x_{01},r)$, es posible obtener

$$x_2(x_1)=\sqrt{r^2-(x_1-h)^2}+k$$

en el caso $x_2<0$, en una bola $B(x_{01},r)$, es posible obtener

$$x_2(x_1)=-\sqrt{r^2-(x_1-h)^2}-k$$

mientras que para el caso $x_2=0$ la ecuación se reduce a

$$(x_1-h)^2+k^2-r^2=0$$

entonces no es posible despejar $x_2(x_1)$ y en una bola $B(x_{01},r)$ se tendrán algunos puntos de la forma $(x_1,\sqrt{r^2-h^2}+k)$ y otros de la forma $(x_1,-\sqrt{r^2-h^2}-k)$, por lo tanto para cada $x_1$ se tienen dos imágenes $x_2(x_1)$.

**Teorema 1.2.** (Teorema de la función implícita) *Sea $F:D\subset\mathbb{R}^{n+1}\to\mathbb{R}$ una función continua y diferenciable que se puede escribir de la forma $F(x,y)=0$. Si existe $(a,b)\in\mathbb{R}^n\times\mathbb{R}$ tal que $F(a,b)=0$ y $\frac{\partial F}{\partial y}(a,b)\neq 0$. Entonces existen conjuntos abiertos $E=B(a,r)$ tal que $E\subset\mathbb{R}^n$ y $F=(-\varepsilon,\varepsilon)$ tal que $F\subset\mathbb{R}$ con $a\in E$ y $b\in F$ tales que para cada $x\in E$ existe un único $y\in F$ de forma que $F(x,y)=0$. Esto define una función $G:E\to F$ continua y diferenciable tal que $F(x,G(x))=0$ $\forall x\in F$ y*

$$DG(x)=-\frac{\frac{\partial F(x,G(x))}{\partial x_i}}{\frac{\partial F(x,G(x))}{\partial y}}\quad \forall x\in F,\ G(a)=b$$

Nota 1.2. De este teorema se tienen la siguiente consecuencia importante: Siempre y cuando $\frac{\partial G}{\partial y}(x_0,y_0)\neq 0$, utilizando la regla de la cadena la derivada parcial de $F$ con respecto a $x_i$ es

$$\frac{\partial F(x,G(x))}{\partial y}\frac{\partial G(x)}{\partial x_i}+\frac{\partial F(x,G(x))}{\partial x_i}=0$$

de lo cual es posible despejar de la siguiente forma

$$\frac{\partial G(x)}{\partial x_i}=\frac{\frac{\partial F(x,G(x))}{\partial x_i}}{\frac{\partial F(x,G(x))}{\partial y}}$$

**Ejemplo 1.7.** Por ejemplo, si la relación entre dos variables está dada por

$$x_1^2+x_1x_2+x_2^2=0$$

no es posible despejar las variables, entonces supongamos que $x_2(x_1)$ y diferenciando con respecto a $x_1$ se obtiene

$$2x_1+x_2+x_1\frac{dx_2}{dx_1}+2x_2\frac{dx_2}{dx_1}=0$$
$$2x_1+x_2+(x_1+2x_2)\frac{dx_2}{dx_1}=0$$

entonces

$$\frac{dx_2}{dx_1} = \frac{-(2x_1 + x_2)}{x_1 + 2x_2}$$

claramente este mismo procedimiento se puede realizar con otros problemas un poco más complejos y el teorema ahorra muchos problemas cuando no se puede factorizar convenientemente para resolver un problema (esto se retomará en problemas de optimización).

## 1.5. Funciones Homogéneas

Se denomina *cono* a todo conjunto $D \subset \mathbb{R}^n$ de la forma

$$tx \in D \; \forall x \in D \quad , t > 0$$

La estructura de cono de altura $h$ y radio $r$, que se tomará para fijar ideas, corresponde a una función implícita $f(x_1, x_2, x_3) = 0$ de la forma

$$\frac{x_1^2 + x_2^2}{(r/h)^2} - (x_3 - x_{03})^2 = 0$$

graficando el caso en que $r = 0$, $h = 2$ y $x_{03} = 0$ se obtiene un cono centrado en $(0, 0)$ y de la forma

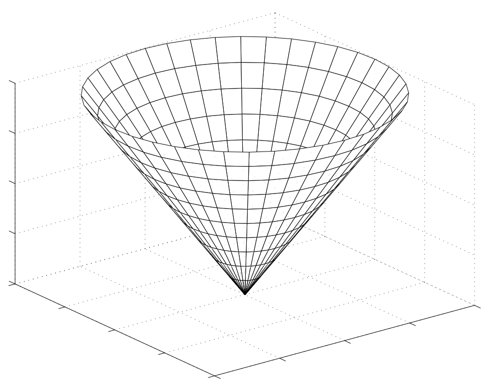

Figura 9: Idea geométrica de cono.

Se debe considerar el interior del cono en el contexto aludido, como aparecen términos cuadráticos en la forma implícita el gráfico completo incluye otro cono igual pero invertido hacia abajo. Consideraremos sólo la parte positiva del cono porque contiene los valores positivos, que son los que nos interesan en economía.

Además, se dice que un cono convexo $D$ define un espacio vectorial. Es decir si $x \in D$ e $y \in D$ entonces $\alpha x + \beta y \in D$ con $\alpha + \beta = 1$. La idea es que el cono se puede abrir y se puede formar, por ejemplo, el ortante positivo de $\mathbb{R}^2$.

Una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$, bajo la condición de que $D$ es un cono, se dice homogénea de grado $k$ si dado $t > 0 \in \mathbb{R}$ se tiene que

$$f(tx) = t^k f(x)$$

esto tiene relación con la proporcionalidad de los efectos que genera un cambio en las variables: si las variables aumentan todas en un %m, una función homogénea de grado 1 ($k = 1$) aumenta su valor en un $\%m$ mientras que si la función es homogénea de grado menor (mayor) a uno se tiene un aumento menos (más) que proporcional.

Se justifica que $t > 0$ porque económicamente nos interesan las transformaciones (o normalizaciones) que puedan resultar sobre las cantidades demandadas por productos, las cantidades utilizadas en la producción de un bien, etc. Adicionalmente, de forma matemática el hecho de que $t < 0$ lleva a problemas de inexistencia en los reales, por ejemplo, si $r = 0,5$ y $t < 0$ claramente $t^r \notin \mathbb{R}$ y esto carece de sentido económico.

Hay dos casos importantes de homogeneidad que son el de homogeneidad de grado 0 (la función no cambia de valor ante cambio en los argumentos) y homogeneidad de grado 1 que ya se explicó.

**Ejemplo 1.8.** Consideremos la función Cobb-Douglas $f(x_1, x_2) = Ax_1^\alpha x_2^\beta$. Determinar el grado de homogeneidad.

Solución. Por definición

$$\begin{aligned} f(tx_1, tx_2) &= A(tx_1)^\alpha (tx_2)^\beta \\ f(tx_1, tx_2) &= t^{\alpha+\beta} x_1^\alpha x_2^\beta \\ f(tx_1, tx_2) &= t^{\alpha+\beta} f(x_1, x_2) \end{aligned}$$

entonces la función es homegénea de grado $\alpha + \beta$ y dependiendo de la suma de los valores será homogénea de grado mayor, menor o igual a 1. □

**Teorema 1.3.** Si $f(x)$ es diferenciable y homogénea de grado $k$ sus derivadas parciales son homegéneas de grado $k - 1$

Demostración. Si $f$ es homegénea de grado $k$, entonces para $t > 0$

$$f(tx) = t^k f(x)$$

derivando con respecto a $x_j$

$$\begin{aligned} \frac{\partial[f(tx)]}{\partial x_j} &= t^k \frac{\partial f(x)}{\partial x_j} \\ \frac{\partial f(tx)}{\partial x_j} \frac{\partial(tx_j)}{\partial x_j} &= t^k \frac{\partial f(x)}{\partial x_j} \\ \frac{\partial f(tx)}{\partial x_j} t &= t^k \frac{\partial f(x)}{\partial x_j} \\ \frac{\partial f(tx)}{\partial x_j} &= t^{k-1} \frac{\partial f(x)}{\partial x_j} \end{aligned}$$

■

**Ejemplo 1.9.** Consideremos la función Cobb-Douglas $f(x_1, x_2) = Ax_1^\alpha x_2^\beta$ homogénea de grado 1 ($\alpha + \beta = 1$ a partir del ejemplo anterior). Probar que las derivadas parciales son homegéneas de grado 0.

Solución. Derivando con respecto a $x_1$

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \alpha A x_1^{\alpha-1} x_2^\beta$$

Por definición

$$\frac{\partial f(tx_1, tx_2)}{\partial x_1} = \alpha A(tx_1)^{\alpha-1}(tx_2)^{\beta}$$
$$\frac{\partial f(tx_1, tx_2)}{\partial x_1} = t^{\alpha+\beta-1}\alpha A x_1^{\alpha-1}x_2^{\beta}$$

del enunciado se tiene que $\alpha + \beta = 1$ entonces la derivada parcial con respecto a $x_1$ es homogénea de grado 0. Para el caso de $x_2$ es análogo. $\square$

**Teorema 1.4.** (Teorema de Euler)
Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ una función diferenciable, se dice homogénea de grado $k$ si y sólo si

$$kf(x) = \sum_{j=1}^{n} \frac{\partial f(tx)}{\partial x_j} x_j$$

Demostración. Definamos $g(t) = f(tx)$. Si derivamos con respecto a $t$ se obtiene

$$g'(t) = \frac{d}{dt} f(tx)$$
$$g'(t) = \sum_{j=1}^{n} \frac{\partial f(tx)}{\partial x_j} x_j$$

Si $f$ es homogénea, nuevamente derivando con respecto a $t$ se obtiene

$$\frac{d}{dt} f(tx) = \frac{d}{dt}[t^k f(x)]$$
$$\sum_{j=1}^{n} \frac{\partial f(tx)}{\partial x_j} x_j = kt^{k-1} f(x)$$

Del teorema anterior sabemos que $\dfrac{\partial f(tx)}{\partial x_j} = t^{k-1}\dfrac{\partial f(x)}{\partial x_j}$, reemplazando

$$\sum_{j=1}^{n} t^k \frac{\partial f(x)}{\partial x_j} x_j = kt^{k-1} f(x)$$
$$\sum_{j=1}^{n} t^{k-1} \frac{\partial f(x)}{\partial x_j} x_j = kt^{k-1} f(x)$$
$$\sum_{j=1}^{n} \frac{\partial f(x)}{\partial x_j} x_j = kf(x)$$

$\blacksquare$

**Ejemplo 1.10.** Consideremos la función Cobb-Douglas $f(x_1, x_2) = Ax_1^{\alpha}x_2^{\beta}$. Determinar si se cumple el teorema de Euler.

Solución. Derivando con respecto a $x_1$

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \alpha A x_1^{\alpha-1}x_2^{\beta}$$

derivando con respecto a $x_2$

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \beta A x_1^{\alpha} x_2^{\beta-1}$$

aplicando el teorema,

$$\begin{aligned}
\frac{\partial f(x_1, x_2)}{\partial x_1} x_1 + \frac{\partial f(tx_1, tx_2)}{\partial x_2} x_2 &= \alpha A x_1^{\alpha} x_2^{\beta} + \beta A x_1^{\alpha} x_2^{\beta} \\
\frac{\partial f(x_1, x_2)}{\partial x_1} x_1 + \frac{\partial f(tx_1, tx_2)}{\partial x_2} x_2 &= (\alpha + \beta) A x_1^{\alpha} x_2^{\beta} \\
\frac{\partial f(x_1, x_2)}{\partial x_1} x_1 + \frac{\partial f(tx_1, tx_2)}{\partial x_2} x_2 &= (\alpha + \beta) f(x_1, x_2)
\end{aligned}$$

del ejemplo anterior sabemos que las derivadas son homogéneas de grado $\alpha + \beta$ por lo que queda en evidencia que el teorema se cumple. $\square$

De lo anterior se tienen como consecuencia las siguientes propiedades:

1. Si $f, g : D \subset \mathbb{R}^n \to \mathbb{R}$ son funciones homogéneas de grado $k$, entonces

$$\begin{aligned}
(f+g)(tx) &= f(tx) + g(tx) = t^k (f+g)(x) \\
(\alpha f)(tx) &= \alpha f(tx) = t^k (\alpha f)(x) \\
(\alpha f + (1-\alpha) g)(tx) &= \alpha f(tx) + (1-\alpha) g(tx) = t^k (\alpha f + (1-\alpha) g)(x) \\
(f \cdot g)(tx) &= f(tx) \cdot g(tx) = t^k (f \cdot g)(x) \\
\left(\frac{f}{g}\right)(tx) &= \frac{f(tx)}{g(tx)} = \left(\frac{f}{g}\right)(x) \ , g(x) \neq 0
\end{aligned}$$

2. Si $g : E \subset \mathbb{R} \to \mathbb{R}$ es una función derivable y creciente tal que si $x_0 < x$ entonces $g(x_0) < g(x)$ y $f : D \subset \mathbb{R}^n \to \mathbb{R}$ es diferenciable y homogénea de grado 1, entonces

$$(g \circ f)(tx) = g(f(tx)) = g(t^k f(x)) = t^{k^2} g(f(x)) = t^{k^2} (g \circ f)(x)$$

3. Si $g : E \subset \mathbb{R} \to \mathbb{R}$ y $f : D \subset \mathbb{R}^n \to \mathbb{R}$ son homogéneas de grado $k$, entonces

$$\begin{aligned}
\frac{\frac{\partial f(tx)}{\partial x_i}}{\frac{\partial f(tx)}{\partial x_j}} &= \frac{\frac{\partial f(x)}{\partial x_i}}{\frac{\partial f(x)}{\partial x_j}} \\
\frac{\frac{\partial g[f(tx)]}{\partial x_i}}{\frac{\partial g[f(tx)]}{\partial x_j}} &= \frac{\frac{\partial g}{\partial f}\frac{\partial f}{\partial x_i}(tx)}{\frac{\partial g}{\partial f}\frac{\partial f}{\partial x_j}(tx)} = \frac{\frac{\partial f}{\partial x_i}(tx)}{\frac{\partial f}{\partial x_j}(tx)} = \frac{\frac{\partial f}{\partial x_i}(x)}{\frac{\partial f}{\partial x_j}(x)}
\end{aligned}$$

esto se conoce como *función homotética*, en términos económicos la $RTS$ de una función diferenciable no cambia bajo transformaciones tales como $\ln(\cdot)$ o alguna otra transformación creciente.

## 2. Optimización

### 2.1. Funciones de una variable

Consideremos una función $y = f(x)$ diferenciable. Diremos que $f$ tiene un *máximo local* en $x_0$ si $f(x_0) \geq f(x)$ para valores "cercanos" a $x_0$. Diremos que $f$ tiene un *mínimo local* en $x_0$ si $f(x_0) \leq f(x)$ para valores "cercanos" a $x_0$.

El caso que nos interesa es el de máximos y mínimos globales, es decir los casos en que en torno a $x_0$ se tiene que $f(x_0) > f(x)$ para un *máximo local* y $f(x_0) < f(x)$ para un *mínimo local*. Este es el caso que vamos a trabajar para poder utilizar convenientemente los criterios de convexidad y concavidad de la sección anterior.

Para fijar ideas, pensemos en una función $f : [0, 1] \to [0, 1]$ que, independiente de su forma funcional, su gráfico es el siguiente:

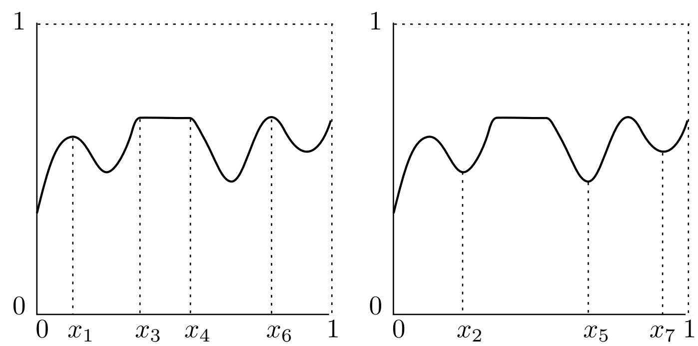


Figura 10: Caso en que hay varios máximos y mínimos.

En este caso la función alcanza máximos locales en $x_1$, $[x_3, x_4]$, $x_6$ y 1, estos valores corresponden a $f(x_1)$, $f(\alpha x_3 + (1-\alpha)x_4)$ con $\alpha \in [0, 1]$, $f(x_6)$ y $f(1)$. Los mínimos locales se alcanzan en 0, $x_2$, $x_5$ y $x_7$, estos valores corresponden a $f(0)$, $f(x_2)$, $f(x_5)$ y $f(x_7)$.

Los valores $f(0)$ y $f(1)$ corresponden a óptimos de esquina y los demás valores corresponden a óptimos interiores, la diferencia está en que los óptimos de esquina no se pueden encontrar directamente utilizando derivadas, la explicación la daremos de forma geométrica: Consideremos otra función $f : [0, 1] \to [0, 1]$ cuyo gráfico es el siguiente

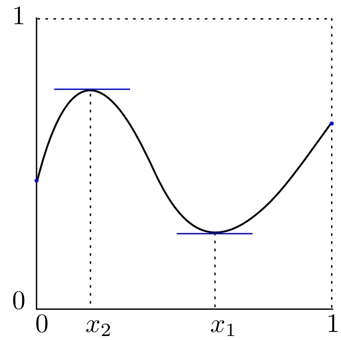


Figura 11: Óptimos interiores y de esquina.

En $f(x_1)$ la pendiente es única y está dada por $f'(x_1)$ que por tratarse de un máximo local se tiene que $f'(x_1) = 0$ (en este caso el máximo también es global). En $f(x_2)$ la pendiente es única y está dada por $f'(x_2)$ que por tratarse de un mínimo local se tiene que $f'(x_2) = 0$ (en este caso el mínimo también es global). Para $f(0)$ y $f(1)$ se tiene que la tangente en esos puntos del gráfico, como que la función es diferenciable en $[0, 1]$, toma valores bien definidos pero $f'(0) \neq 0$ y $f'(1) \neq 0$.

No hay que confundir los valores del eje $x$, llamados $\arg\max$, que son los argumentos que maximizan $f$, con los valores máximos.

Veamos ahora un caso en que una función de la forma $f : [0, 1] \to [0, 1]$ es continua pero no es diferenciable en todo su dominio

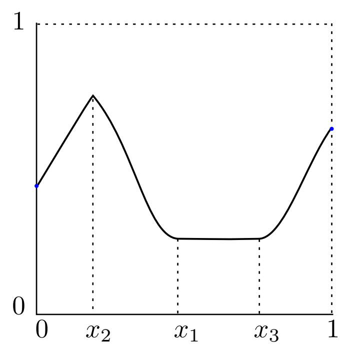


Figura 12: Caso en que la derivada en un punto no es única.

En $f(x_1)$ la pendiente no es única pero sin embargo $f(x_1)$ es el máximo valor de $f$ en su dominio y se tiene entonces un máximo global, pese a que no es posible aplicar el criterio de que $f'(x_1) = 0$. Un hecho que valida que $f(x_1)$ es máximo es que $f'(x_1 - \varepsilon) > 0$ y $f'(x_1 + \varepsilon) > 0$ con $\varepsilon \to 0$, entonces si $f$ fuera diferenciable en $x_1$ existiría $f'(x_1) = 0$. Existen herramientas más sofisticadas para la optimización no diferenciable pero no son de utilidad en el curso.

Para $f(\alpha x_2 + (1-\alpha)x_3)$ con $\alpha \in [0, 1]$ se tiene que $f'(\alpha x_2 + (1-\alpha)x_3) = 0$ y se tiene que cualquier punto de $[\alpha x_2 + (1-\alpha)x_3]$ es un mínimo local de la función.

Ahora estamos en condiciones de dar un criterio eficiente para la optimalidad de funciones: Supongamos que $f : D \subset \mathbb{R} \to \mathbb{R}$ es una función dos veces continua y diferenciable. Entonces $f$ tiene un óptimo en $x_0$ si se cumplen

1. En el caso de maximización

$$\text{Condición de primer orden: } f'(x_0) = 0$$

$$\text{Condición de segundo orden: } f''(x_0) \leq 0$$

2. En el caso de minimización

$$\text{Condición de primer orden: } f'(x_0) = 0$$

$$\text{Condición de segundo orden: } f''(x_0) \geq 0$$

## 2.2. Funciones de varias variables

Consideraremos el caso de las funciones dos veces continuas y diferenciables de la forma $f : D \subset \mathbb{R}^n \to \mathbb{R}$.

Conceptualmente diremos que $f$ alcanza un *máximo local* en $x_0$ si $f$ no puede aumentar su valor más allá del valor que alcanza en $x_0$. De manera más precisa, diremos que en torno a $x_0$ se tiene que $f(x_0 - he_j) \leq f(x_0)$ y $f(x_0 + he_j) \leq f(x_0)$ lo cual nos da la idea de que en una bola abierta de centro en $x_0$ y radio $r$, $f$ alcanza un máximo local en $x_0$ si $f(x_0) \geq f(x)\ \forall x \in B(x_0, r)$. Este máximo local es *único* si $f(x_0) > f(x)\ \forall x \in B(x_0, r)$. Se tiene que $f$ alcanza un *máximo global* en $x_0$ si $f(x_0) \geq f(x)\ \forall x \in D$. Además este máximo global es *único* si $f(x_0) > f(x)\ \forall x \in D$.

Para el caso en que $f$ tiene un *mínimo local* en $x_0$ si $f(x_0) \leq f(x)\ \forall x \in B(x_0, r)$. Este mínimo local es *único* si $f(x_0) < f(x)\ \forall x \in B(x_0, r)$. Se tiene que $f$ alcanza un *mínimo global* en $x_0$ si $f(x_0) \leq f(x)\ \forall x \in D$. Además este mínimo global es *único* si $f(x_0) < f(x)\ \forall x \in D$. Claramente, en todos los casos que hemos descrito $x_0 \neq x$.

Retomando el caso de funciones de una variable, si definimos $g(t) = f(x + tv)$ con lo que ya explicamos se tendrá que $f$ presenta un máximo global o local cuando $g'(t) = 0$, caso que se cumple cuando $t = 0$, es decir, $t = 0$ maximiza la función $g$ y en consecuencia $f$ tiene un máximo en $x_0$. Las condiciones de segundo orden para el caso de varias variables son un poco más complejas, por un lado se cumple la condición de segundo orden si $g''(t) \leq 0$ pero no hemos especificado la función $g$ por lo que se trata de un caso teórico. Un criterio eficiente en la práctica es el uso de la matriz Hessiana cuya descripción ya dimos en la sección anterior.

Una *condición de primer orden* para la optimalidad de $f$ es

$$\nabla f(x_0) = 0 \Leftrightarrow \frac{\partial f(x_0)}{\partial x_j} = 0 \quad \forall j = \{1, \ldots, n\}$$

Para convencernos de esto teníamos que $g(t) = f(x + tv)$ donde $v \in \mathbb{R}^n$ es una dirección hacia la cual nos podemos mover, en $t = 0$ $g(0) = f(x)$ y como ya se discutió $g'(0) = 0$. Derivando $g(t) = f(x_0 + tv)$ con respecto a $t$

$$g'(t) = \sum_{j=1}^{n} \frac{\partial f(x_0 + tv)}{\partial x_j} v_j$$

evaluando en $t = 0$

$$g'(0) = \sum_{j=1}^{n} \frac{\partial f(x_0)}{\partial x_j} v_j$$

$$g'(0) = \nabla f(x_0) \cdot v$$

como $g$ tiene un máximo en $t = 0$

$$g'(0) = 0$$

lo que quiere decir que

$$\nabla f(x_0) \cdot v = 0 \quad \forall v \in \mathbb{R}^n$$

que significa que cada una de las derivadas parciales de $f$ se anula independiente de $v$ y por lo tanto

$$\nabla f(x_0) = 0$$

**Ejemplo 2.1.** Dada la función $f(x, y) = y - 4x^2 + 3xy - y^2$, encontrar los puntos que maximizan o minimizan la función.

Solución. Aplicando la condición de primer orden

$$\frac{\partial f(x,y)}{\partial x} = -8x + 3y = 0 \qquad \frac{\partial f(x,y)}{\partial y} = 1 + 3x - 2y = 0$$

igualando ambas ecuaciones

$$-8x + 3y = 1 + 3x - 2y \Rightarrow -11x = 1 - 5y \Rightarrow x = \frac{5y - 1}{11}$$

reemplazando en la primera ecuación

$$\frac{-8(5y-1)}{11} + 3y = 0 \Rightarrow \frac{-40y + 8 + 33y}{11} = 0 \Rightarrow -7y + 8 = 0 \Rightarrow y = \frac{8}{7}$$

reemplazando en la segunda ecuación

$$1 + 3x - \frac{16}{7} = 0 \Rightarrow 3x - \frac{9}{7} = 0 \Rightarrow x = \frac{3}{7}$$

□

Una *condición de segundo orden* para la optimalidad de $f$ es la siguiente:

1. Para el caso de maximización si $f$ alcanza un máximo local interior en $x_0$, entonces $Hf(x_0)$ es semi definida negativa.

2. Para el caso de minimización si $f$ alcanza un mínimo local interior en $x_0$, entonces $Hf(x_0)$ es semi definida positiva.

Para convencernos de esto teníamos que $g(t) = f(x_0 + tv)$ y derivando $g(t) = f(x_0 + tv)$ con respecto a $t$ se obtiene

$$g'(t) = \sum_{j=1}^{n} \frac{\partial f(x_0 + tv)}{\partial x_j} v_j$$

derivando nuevamente con respecto a $t$

$$g''(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(x_0 + tv)}{\partial x_i x_j} v_j v_i$$

si $f$ alcanza un máximo en $x_0$

$$g''(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f(x_0+tv)}{\partial x_i x_j} v_j v_i \leq 0$$

cambiando la notación

$$g''(t) = v^T Hf(x_0)v \leq 0$$

si $f$ alcanza un mínimo en $x_0$

$$g''(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f(x_0+tv)}{\partial x_i x_j} v_j v_i \geq 0$$

cambiando la notación

$$g''(t) = v^T Hf(x_0)v \geq 0$$

En la sección anterior ya vimos el significado de semi definida negativa y semi definida positiva. Falta agregar que cuando $f$ alcanza un máximo o mínimo local en $x_0$ y este óptimo es único, entonces $Hf(x_0)$ es definida negativa o definida positiva respectivamente. Si $Hf(x_0)$ es definida negativa la función $f$ es cóncava en $B(x_0, r)$, mientras que si $Hf(x_0)$ es definida positiva la función $f$ es cónvexa en $B(x_0, r)$.

Cuando $f$ es estrictamente cóncava en todo su dominio, entonces $Hf(x_0)$ es definida negativa en $D$ y $x_0$ es un máximo global de $f$ que además es único. Para el caso en que $f$ es estrictamente convexa en todo su dominio, entonces $Hf(x_0)$ es definida positiva en $D$ y $x_0$ es un mínimo global de $f$ que además es único.

**Ejemplo 2.2.** Dada la función $f(x, y) = y - 4x^2 + 3xy - y^2$, determinar si las soluciones encontradas en el ejemplo anterior efectivamente maximizan $f$.

Solución. Aplicando la condición de segundo orden

$$\frac{\partial f(x,y)}{\partial x^2} = -8 \qquad \frac{\partial^2 f(x,y)}{\partial y^2} = -2$$

$$\frac{\partial^2 f(x,y)}{\partial xy} = 3 \qquad \frac{\partial^2 f(x,y)}{\partial yx} = 3$$

Luego,

$$Hf(x_0) = \begin{pmatrix} -8 & 3 \\ 3 & -2 \end{pmatrix}$$

Se tiene que las submatrices corresponden a

$$|H|_1 = -8 < 0$$

$$|H_2| = \begin{vmatrix} -8 & 3 \\ 3 & -2 \end{vmatrix} = 16 - 9 = 7 > 0$$

entonces, a partir de los signos de las submatrices, se tiene que estos son alternados y que $|H|_1$ es negativo por lo que la función es cóncava y presenta un máximo. Si el hessiano dependiera de $x$ e $y$ para ser definido negativo, entonces la función sería cóncava en $B(x_0, r)$ pero en el ejemplo es cóncava en $D$. $\square$

Para resumir lo presentado en esta sección tenemos los siguientes teoremas:

**Teorema 2.1.** Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ una función cóncava dos veces continua y diferenciable. Si $f$ alcanza un máximo interior en $x_0$ las siguientes propiedades son equivalentes

1. $\nabla f(x_0) = 0$
2. $f$ alcanza un máximo local en $x_0$
3. $f$ alcanza un máximo global en $x_0$

tal que $1 \Leftarrow 2$, $2 \Leftarrow 3$ y $1 \Leftarrow 3$.

Demostración.
($1 \Leftarrow 2$): Por definición si $f$ alcanza un máximo local en $x_0$, $f$ es cóncava en $B(x_0, r)$. Como $f$ es dos veces continua y diferenciable se tiene que $f(x_0) \geq f(x)$ y $g'(t) = 0$ dado $t = 0$, de lo cual llegamos a que $g'(0) = \nabla f(x_0) = 0$.
($2 \Leftarrow 3$): Por definición si $f$ alcanza un máximo global en $x_0$, $f$ es cóncava en $D$. Como $B(x_0, r) \subset D$ $f$ es cóncava en $B(x_0, r)$ y en torno a $x_0$ se tiene que $f(x_0) \geq f(x)$ para $x \in B(x_0, r)$.
($1 \Leftarrow 3$): Si se cumple la condición de primer orden, entonces $\nabla f(x_0) = 0$. Anteriormente vimos que para funciones de una variable $f(x) \leq f(x_0) + f'(x_0)(x - x_0)$ lo cual para el caso de varias variables corresponde a $f(x) \leq f(x_0) + \nabla f(x_0) \cdot (x - x_0)$ y esta ecuación, dado que $\nabla f(x_0) = 0$, se reduce a $f(x) \leq f(x_0)$ ■

**Teorema 2.2.** Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ una función convexa dos veces continua y diferenciable. Si $f$ alcanza un máximo interior en $x_0$ las siguientes propiedades son equivalentes

1. $\nabla f(x_0) = 0$
2. $f$ alcanza un mínimo local en $x_0$
3. $f$ alcanza un mínimo global en $x_0$

tal que $1 \Leftarrow 2$, $2 \Leftarrow 3$ y $1 \Leftarrow 3$.

Demostración. Es idéntica a la del teorema anterior, lo único que cambia es el sentido de la desigualdad. ■

En los últimos dos teoremas cabe la posibilidad de que los valores óptimos se alcancen en más de un punto similar a lo que se ilustra en la figura 12. Esta posibilidad se elimina si $f$ es *estrictamente cóncava* o *estrictamente convexa*.

**Teorema 2.3.** Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ una función estrictamente cóncava y dos veces continua y diferenciable. Si $f$ alcanza un máximo interior en $x_0$ entonces $f(x_0) > f(x)$ $\forall x \in D$ y $x \neq x_0$.

Demostración. Supongamos que $x_0$ no es el único argumento que maximiza $f$, digamos que existe $x' \neq x_0$ tal que $f(x') = f(x_0)$ y si esto fuera cierto por ser $f$ estrictamente cóncava si $x = \alpha x' + (1 - \alpha)x_0$ se tiene que

$$f(x) > \alpha f(x') + (1 - \alpha)f(x_0) \quad \forall \alpha \in (0, 1)$$

como $f(x') = f(x_0)$, entonces

$$\begin{aligned} f(x) &> \alpha f(x') + (1 - \alpha)f(x') \\ f(x) &> f(x') \end{aligned}$$

lo cual significa que ni $x_0$ ni $x'$ maximizan la función lo cual es una contradicción. Entonces, si $x_0$ maximiza $f$ que es estrictamente cóncava $f(x_0) > f(x)$ $\forall x \in D$. ■

**Teorema 2.4.** Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ una función estrictamente convexa y dos veces continua y diferenciable. Si $f$ alcanza un mínimo interior en $x_0$ entonces $f(x_0) < f(x)$ $\forall x \in D$ y $x \neq x_0$.

Demostración. Es idéntica a la del teorema anterior, lo único que cambia es el sentido de la desigualdad. ■

# 3. Optimización con restricciones

En economía muchas veces nos encontramos con que los problemas que debemos resolver presentan una o varias restricciones. Los consumidores cuentan con restricciones de tiempo y de presupuesto, las firmas tienen restricciones de capacidad, legales y tecnológicas por ilustrar sólo algunos casos a grandes rasgos. Por esto, tiene sentido que los agentes económicos busquen la mejor respuesta posible frente a las restricciones presentes. De esta forma, los conceptos que vimos anteriormente nos servirán como una base para lo que sigue pero no nos permiten resolver directamente un problema con restricciones.

## 3.1. Resolución implícita

Supongamos que tenemos una función diferenciable $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ y que enfrentamos un problema de la forma

$$\begin{array}{ll} \max\limits_{x_1,x_2} & f(x_1,x_2) \\ \text{s.a} & g(x_1,x_2)=0 \end{array} \qquad \text{ó} \qquad \begin{array}{ll} \min\limits_{x_1,x_2} & f(x_1,x_2) \\ \text{s.a} & g(x_1,x_2)=0 \end{array}$$

donde $g : D \subset \mathbb{R}^2 \to \mathbb{R}$ es diferenciable.

$f$ es nuestra *función objetivo* y $(x_1, x_2)$ son las *variables de decisión*. Cualquier valor $(x_1, x_2)$ tal que $g(x_1, x_2) = 0$ se dirá *factible* y si este valor maximiza o minimiza $f$ se dice que es un valor óptimo y factible.

Una forma eficiente de resolver, pero que presenta limitaciones, es dejar una variable en función de la otra, lo cual no siempre se puede hacer. Dependiendo de la forma funcional de $g(x_1, x_2) = 0$ en algunos casos es posible expresar $x_2 = g(x_1)$, entonces el problema que enfrentamos se reduce a

$$\max_{x_1} f(x_1, g(x_1)) \qquad \text{ó} \qquad \min_{x_1} f(x_1, g(x_1))$$

Derivando con respecto a $x_1$ es posible encontrar $x_{01}$ tal que

$$\frac{\partial f(x_0, g(x_{01}))}{\partial x_1} + \frac{\partial f(x_{01}, g(x_{01}))}{\partial x_2}\frac{\partial g(x_{01})}{\partial x_1} = 0$$

$$\frac{\frac{\partial f(x_{01},g(x_{01}))}{\partial x_1}}{\frac{\partial f(x_{01},g(x_{01}))}{\partial x_2}} = -\frac{g(x_{01})}{x_1}$$

se puede encontrar $x_{01}$ tras resolver directamente y luego reemplazar en $x_{02} = g(x_{01})$ para encontrar $(x_{01}, x_{02})$

La aplicación más común de esto en economía es con el problema de maximización de beneficios: Sean $p$ un precio dado por el mercado, $y$ una cantidad de producto, $(x_1, x_2)$ los factores productivos, $(w_1, w_2)$ los costos unitarios de cada factor y $f(x_1, x_2)$ una función diferenciable que resume la tecnología de la firma, el problema de maximización beneficios es

$$\begin{array}{ll} \max\limits_{x_1,x_2} & py - w_1x_1 - w_2x_2 \\ \text{s.a} & f(x_1,x_2) \geq y \end{array}$$

donde $f(x_1, x_2)$ en la restricción da cuenta de que con la tecnología de la firma debe producir a lo menos una cantidad mínima que está fija que puede atribuirse a contratos firmados con anterioridad, etc.

Si $f(x_1, x_2) = y$ podemos reemplazar en la función objetivo y se obtiene

$$\max_{x_1,x_2} pf(x_1, x_2) - w_1x_1 - w_2x_2$$

derivando con respecto a $x_1$

$$p\frac{\partial f(x_{01}, x_{02})}{\partial x_1} = w_1$$

derivando con respecto a $x_2$

$$p\frac{\partial f(x_{01}, x_{02})}{\partial x_2} = w_2$$

si dividimos las expresiones encontradas llegamos a

$$\frac{\frac{\partial f(x_{01}, x_{02})}{\partial x_1}}{\frac{\partial f(x_{01}, x_{02})}{\partial x_2}} = \frac{w_1}{w_2}$$

esto nos dice que la productividad marginal que aporta haber gastado $w_1$ en la última unidad utilizada del factor $x_1$ iguala a la productividad marginal que aporta haber gastado $w_2$ en la última unidad utilizada del factor $x_2$. Cuando se tiene esto, la combinación de factores es la mejor posible y cualquier cambio en la combinación de factores, con $(w_1, w_2)$ fijos, es desfavorable para la firma.

## 3.2. Multiplicadores de Lagrange

### Condiciones de primer orden

La idea de este método es llevar un problema con restricciones a uno equivalente pero sin restricciones. Una forma geométrica que da respuesta a esto es el siguiente teorema:

**Teorema 3.1.** (Teorema de los multiplicadores de Lagrange[1])
*Supongamos que tenemos una función diferenciable $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ y que enfrentamos un problema de la forma*

$$\begin{array}{llcll} \max\limits_{x_1,x_2} & f(x_1, x_2) & \quad ó \quad & \min\limits_{x_1,x_2} & f(x_1, x_2) \\ s.a & g(x_1, x_2) = c & & s.a & g(x_1, x_2) = c \end{array}$$

*donde $g : D \subset \mathbb{R}^2 \to \mathbb{R}$ es diferenciable. Sea $S$ el conjunto de todos los pares $(x_1, x_2)$ factibles, es decir*

$$S = \{(x_1, x_2) \in D : g(x_1, x_2) - c = 0\}$$

*Supongamos que $\nabla g(x_{01}, x_{02})) \neq 0$. Si $f(x_1^S, x_2^S)$, es decir considerando sólo pares factibles, tiene un mínimo o máximo local en $(x_{01}, x_{02}) \in S$, $(x_{01}, x_{02})$ es solución del problema que enfrentamos. Entonces existe un número real $\lambda$ tal que*

$$\nabla f(x_{01}, x_{02}) = \lambda \nabla g(x_{01}, x_{02})$$

[1]Joseph Louis Lagrange (1736-1813) fue un destacado matemático que desarrolló importantes trabajos en matemática pura con gran impacto en física y astronomía. Discutió y desarrollo el cálculo variacional con Euler. Lagrange nunca se preocupó de los aspectos aplicados de sus trabajos, contentándose sólo con la belleza teórica de sus resultados y fueron otros matemáticos de la época, como Laplace, los que desarrollaron las aplicaciones en física. El teorema mencionado lo desarrolló entre 1772 y 1778 y hasta nuestros días ha servido como base para métodos más eficientes o aplicables a casos más complejos en computación, astronomía, ingeniería, economía, etc.

Demostración. Supongamos que $(x_{01}, x_{02})$ es solución del problema de minimización (el otro caso es análogo) y sea

$$S = \{(x_1, x_2) \in S : g(x_1, x_2) - c = 0\}$$

Como $\nabla g(x_{01}, x_{02}) \neq 0$, entonces este vector es normal a la superficie $S$, por otro lado, el plano tangente a $S$ en $(x_{01}, x_{02})$ está definido por

$$\pi : \quad \nabla g(x_{01}, x_{02}) \cdot [(x_1, x_2) - (x_{01}, x_{02})] = 0$$

o equivalentemente, definiendo $\sigma : \mathbb{R} \to D$ (con la misma idea de cuando definiamos $g(t)$ en la sección anterior)

$$\pi = \{\sigma'(0) \mid \sigma : \mathbb{R} \to D,\ \sigma(t) \in S \quad \forall t,\ \sigma(0) = (x_{01}, x_{02})\}$$

Entonces, siendo $x_0$ un mínimo de $f$, la función $f(\sigma(t))$ tiene un mínimo en $t = 0$, por lo tanto

$$\nabla f(x_{01}, x_{02}) \cdot \sigma'(0) = 0$$

es decir, $\nabla f(x_{01}, x_{02})$ es ortogonal al plano tangente, es decir, paralelo al vector $\nabla g(x_{01}, x_{02})$. Por lo tanto, existe $\lambda \in \mathbb{R}$ tal que

$$\nabla f(x_{01}, x_{02}) = \lambda \nabla g(x_{01}, x_{02})$$

■

El escalar $\lambda$ se conoce como *multiplicador de Lagrange* y a la función de 3 variables

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda(g(x_1, x_2) - c)$$

se le conoce como *Lagrangeano*. Esta función es extensible a problemas de $n$ variables generando una función de $n + 1$, pero de momento vamos trabajar los problemas geométricamente.

Lo que el teorema nos dice es que se puede resolver un problema con restricciones por medio del siguiente sistema de ecuaciones (generalmente es un sistema no lineal) que en la literatura se conoce como *sistema Lagrangeano*

$$\begin{cases} \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_1} &= \dfrac{\partial f(x_{01}, x_{02})}{\partial x_1} - \lambda_0 \dfrac{\partial g(x_{01}, x_{02})}{\partial x_1} = 0 \\ \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_2} &= \dfrac{\partial f(x_{01}, x_{02})}{\partial x_2} - \lambda_0 \dfrac{\partial g(x_{01}, x_{02})}{\partial x_2} = 0 \\ g(x_{01}, x_{02}) - c &= 0 \end{cases}$$

Nota 3.1. **Importante:** Una pregunta muy frecuente es si debemos anotar

$$L = f(x_1, x_2) - \lambda(g(x_1, x_2) - c)$$

o se debe expresar

$$L = f(x_1, x_2) + \lambda(c - g(x_1, x_2))$$

Si $(x_{01}, x_{02})$ es un máximo de $f$ restringida a $S$, se cumple que

$$\nabla f(x_{01}, x_{02}) = \lambda \nabla g(x_{01}, x_{02})$$

Para la minimización tendríamos lo siguiente

$$-\nabla f(x_{01}, x_{02}) = \lambda \nabla g(x_{01}, x_{02})$$

Esto sugiere que si escribimos el Lagrangeano para el caso de la maximización como

$$L = f(x_1, x_2) + \lambda(c - g(x_1, x_2))$$

Para un problema de minimización sería

$$L = f(x_1, x_2) - \lambda(g(x_1, x_2) - c)$$

Sin embargo, esto es sólo por un tema conceptual de interpretación, al momento de resolver no hay diferencia. Más adelante daremos una interpretación al multiplicador, por ahoranos quedaremso con la idea de que para el caso de maximización da cuenta de cuanto puede aumentar el valor de la función objetivo si relajo la restricción y cuánto disminuye en un problema de minimización.

NOTA 3.2. El teorema de multiplicadores de Lagrange provee condiciones necesarias para la optimalidad. En algunos casos no es aplicable, por ejemplo, cuando todas las funciones involucradas son lineales. Existe un problema de sesgo con este método que es que en muchos casos hay multiples soluciones interiores y de esquina para las cuales el teorema no es aplicable.

Algunos problemas de optimización no se pueden resolver mediante este método, por ejemplo el de las funciones $\min\{\alpha x_1, \beta x_2\}$ u otras funciones no diferenciables. Surge un inconveniente adicional que es que algunos problemas los multiplicadores se anulan por lo que hay que ser cuidadosos y no pensar que este método nos provee una receta mágica.

Los siguientes ejemplos dan cuenta de la nota anterior:

**Ejemplo 3.1.** Consideremos el problema

$$\begin{array}{cc} \min\limits_{x_1,x_2} & \max\{x_1, x_2\} \\ \text{s.a} & 2x_1 - 3x_2 = 0 \end{array}$$

El inconveniente es con $\max\{x_1, x_2\}$ y esta función, además de que no es diferenciable, se minimiza cuando al menos uno de los dos argumentos tome el menor valor posible. Por ejemplo, si $x_2 > x_1$, entonces $\max\{x_1, x_2\} = x_2$ y cualquier par $(x_1, x_2)$ que cumpla $2x_1 - 3x_2 = 0$ es solución del problema. Por otra parte, se concluye que debido a que la función objetivo no es diferenciable entonces el teorema no es aplicable.

**Ejemplo 3.2.** Consideremos el problema

$$\begin{array}{cc} \min\limits_{x_1,x_2} & x_1 + x_2 \\ \text{s.a} & 2x_1 - 3x_2 = 0 \end{array}$$

Observemos que $\nabla f(x_1, x_2)|_{(0,0)} = (1, 1)$ y $\nabla g(x_1, x_2)|_{(0,0)} = (2, 3)$. Entonces, no existe $\lambda \in \mathbb{R}$ que haga sistema que definen las condiciones de primer orden del Lagrangeano tenga solución dado.

**Ejemplo 3.3.** Consideremos el problema

$$\begin{array}{cc} \min\limits_{x_1,x_2} & 2x_1 + 3x_2^2 \\ \text{s.a} & 2x_1 - 3x_2^2 = 0 \end{array}$$

Observemos que $\nabla f(x_1, x_2)|_{(0,0)} = (2, 0)$ y $\nabla g(x_1, x_2)|_{(0,0)} = (2, 0)$. Entonces, los gradientes son linealmente dependientes y no se cumple que $\nabla f(x_1, x_2)|_{(0,0)} - \lambda \nabla g(x_1, x_2)|_{(0,0)} = (0, 0)$ a menos que $\lambda = 1$ lo que hace que el sistema que definen las condiciones de primer orden del Lagrangeano no tenga solución dado que el par $(1, 0)$ no cumple la restricción del problema.

Para fijar ideas, la función lagrangeano resuelve un problema equivalente sin restricciones. Se evidencia esto si diferenciamos $L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda g(x_1, x_2)$

$$dL(x_1, x_2, \lambda) = \frac{\partial L(x_1, x_2, \lambda)}{\partial x_1} dx_1 + \frac{\partial L(x_1, x_2, \lambda)}{\partial x_2} dx_2 + \frac{\partial L(x_1, x_2, \lambda)}{\partial \lambda} d\lambda$$

reemplazando cada una de las derivadas parciales que aparecen en esto por las que aparecen en el sistema Lagrangeano se llega a

$$\begin{aligned} dL(x_{01}, x_{02}, \lambda_0) =& \frac{\partial f(x_{01}, x_{02})}{\partial x_1} dx_1 + \frac{\partial f(x_{01}, x_{02})}{\partial x_2} dx_2 - g(x_{01}, x_{02}) d\lambda \\ & - \lambda \left( \frac{\partial g(x_{01}, x_{02})}{\partial x_1} dx_1 + \frac{\partial g(x_{01}, x_{02})}{\partial x_2} dx_2 \right) \\ dL(x_{01}, x_{02}, \lambda_0) =& 0 \end{aligned}$$

como $g(x_{01}, x_{02}) = 0$

$$\begin{aligned} dL(x_{01}, x_{02}, \lambda_0) &= \frac{\partial f(x_{01}, x_{02})}{\partial x_1} dx_1 + \frac{\partial f(x_{01}, x_{02})}{\partial x_2} dx_2 - \lambda \left( \frac{\partial g(x_{01}, x_{02})}{\partial x_1} dx_1 + \frac{\partial g(x_{01}, x_{02})}{\partial x_2} dx_2 \right) \\ dL(x_{01}, x_{02}, \lambda_0) &= 0 \end{aligned}$$

a partir de $g(x_{01}, x_{02}) = 0$, se tiene que $dg(x_{01}, x_{02}) = \frac{\partial g(x_{01}, x_{02})}{\partial x_1} dx_1 + \frac{\partial g(x_{01}, x_{02})}{\partial x_2} dx_2 = 0$ y se puede simplificar

$$dL(x_{01}, x_{02}, \lambda_0) = \frac{\partial f(x_{01}, x_{02})}{\partial x_1} dx_1 + \frac{\partial f(x_{01}, x_{02})}{\partial x_2} dx_2 = 0 \quad \forall (x_{01}, x_{02}) \in S$$

y eso es la condición de primer orden aplicada a la función Lagrangeano.

De la última ecuación se obtiene

$$\frac{\frac{\partial f(x_{01}, x_{02})}{\partial x_1}}{\frac{\partial f(x_{01}, x_{02})}{\partial x_2}} = -\frac{dx_2}{dx_1}$$

lo cual es válido siempre y cuando $\frac{\partial f(x_{01}, x_{02})}{\partial x_2} \neq 0$ y $dx_1 \neq 0$, lo cual se consigue cuando la solución del problema es interior. La interpretación de esto es muy similar a la que dimos para el problema de maximización de beneficios, por medio de la función Lagrangeano se llega a una elección de variables tal que cualquier otra elección distinta es desfavorable o resulta más ineficiente.

Hemos detallado la geometrá de los multiplicadores de Lagrange, los siguientes gráficos nos servirán para fijar ideas:

De las condiciones de primer orden del Lagrangeano obtuvimos

$$\frac{\frac{\partial f(x_{01}, x_{02})}{\partial x_1}}{\frac{\partial f(x_{01}, x_{02})}{\partial x_2}} = -\frac{dx_2}{dx_1}$$

dado un valor fijo $y = f(x_1, x_2)$ se tiene que

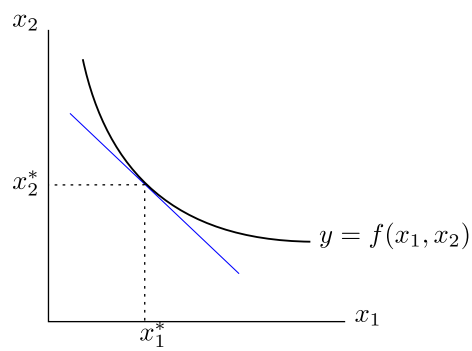


Figura 13: Curva de nivel de $f$.

en $(x_{01}, x_{02})$ la pendiente es $m = -\dfrac{dx_2}{dx_1}$

Diferenciando $g(x_1, x_2) = 0$ obtuvimos

$$\frac{\partial g(x_{01}, x_{02})}{\partial x_1} dx_1 + \frac{\partial g(x_{01}, x_{02})}{\partial x_2} dx_2 = 0$$

reordenando

$$\frac{\frac{\partial g(x_{01}, x_{02})}{\partial x_1}}{\frac{\partial g(x_{01}, x_{02})}{\partial x_2}} = -\frac{dx_2}{dx_1}$$

esto, dado un valor fijo $g(x_1, x_2) = 0$, corresponde a

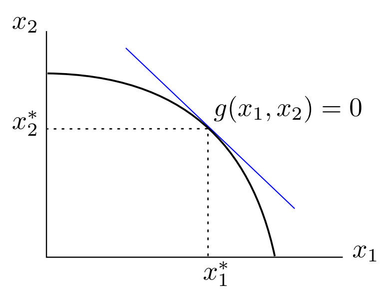


Figura 14: Curva de nivel de $g$.

en $(x_{01}, x_{02})$ la pendiente es $m = -\dfrac{dx_2}{dx_1}$

Combinando ambos gráficos se llega a lo siguiente

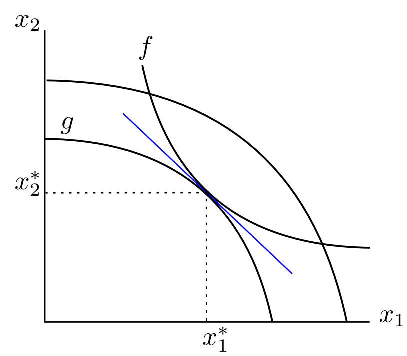


Figura 15: Condiciones de primer orden del Lagrangeano.

en $(x_{01}, x_{02})$ la pendiente es

$$m = -\frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}} = \frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} = -\frac{dx_2}{dx_1}$$

finalmente, concluiremos el análisis geométrico con el siguiente gráfico que se deduce a partir del anterior

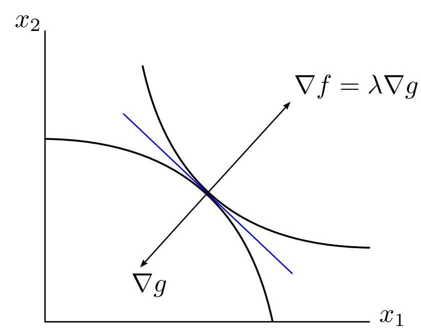


Figura 16: Relación entre el multiplicador y las curvas de nivel.

Este último gráfico nos da cuenta de que en el óptimo los gradientes de la función objetivo y la restricción son linealmente dependientes, es decir son proporcionales.

El teorema funciona con más restricciones y con $n$ variables. Ahora estamos en condiciones de presentar una forma general de la función Lagrangeano que corresponde a

$$L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g^j(x)$$

con $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$ y $1 \leq j \leq m$ pudiendo haber hasta $m < n$ cantidad de restricciones $g^1(x), \ldots, g^m(x)$.

Ahora daremos una aplicación al problema del consumidor: Considere que un individuo tiene preferencias representables por medio de una función de utilidad $U(x_{01}, x_{02})$ dos veces continua y diferenciable, donde $(x, y)$ corresponde a las cantidades consumidas de dos bienes. La restricción del consumidor es que cuenta con un ingreso $I$ finito, por tanto su gasto dados los precios $p_1, p_2$ debe ser $p_1 x_1 + p_2 x_2 \leq I$.

Luego el lagrangeano del problema es

$$L(x_1, x_2, \lambda) = U(x, y) - \lambda(p_1 x_1 + p_2 x_2 - I)$$

lo cual genera el siguiente sistema

$$\begin{cases} \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_1} & = \dfrac{\partial U(x_{01}, x_{02})}{\partial x_1} - \lambda p_1 = 0 \\ \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_2} & = \dfrac{\partial U(x_{01}, x_{02})}{\partial x_2} - \lambda p_2 = 0 \\ \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial \lambda} & = I - p_1 x_1 - p_2 x_2 = 0 \end{cases}$$

De las primeras dos ecuaciones se obtiene

$$\frac{\partial U(x_{01}, x_{02})}{\partial x_1} = \lambda p_1 \quad \text{y} \quad \frac{\partial U(x_{01}, x_{02})}{\partial x_2} = \lambda p_2$$

si la solución es interior, dividiendo ambas ecuaciones llegamos a

$$\frac{\frac{\partial U(x_{01}, x_{02})}{\partial x_1}}{\frac{\partial U(x_{01}, x_{02})}{\partial x_2}} = \frac{p_1}{p_2}$$

esto es una justificación del hecho de que, en el óptimo, la utilidad marginal del último peso gastado en un bien iguala a la utilidad marginal del último peso gastado en otro bien. Luego, la tercera ecuación del sistema es una condición de factibilidad y nos permite encontrar las cantidades $(x_{01}, x_{02})$ que son óptimas y factibles.

**Ejemplo 3.4.** Supongamos que una firma tiene una tecnología CES de la forma $f(x_1, x_2) = (x_1^\rho + x_2^\rho)^{1/\rho}$ y los costos unitarios de cada factor son $w_1$ y $w_2$ respectivamente. Luego su problema de minimizar costos está dado por:

$$\begin{array}{ll} \min\limits_{x_1, x_2} & w_1 x_1 - w_2 x_2 \\ \text{s.a} & f(x_1, x_2) = y \end{array}$$

Encuentre las cantidades $(x_{01}, x_{02})$ que minimizan costos y la forma de la función de costos en función de $(x_{01}, x_{02})$.

Solución. El lagrangeano del problema es

$$L(x_1, x_2, \lambda) = w_1 x_1 - w_2 x_2 - \lambda((x_1^\rho + x_2^\rho)^{1/\rho} - y)$$

lo cual genera el siguiente sistema

$$\begin{cases} \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_1} & = w_1 - \lambda \dfrac{1}{\rho}(x_{01}^\rho + x_{02}^\rho)^{(1-\rho)/\rho}(\rho x_{01}^{\rho-1}) = 0 \\ \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_2} & = w_2 - \lambda \dfrac{1}{\rho}(x_{01}^\rho + x_{02}^\rho)^{(1-\rho)/\rho}(\rho x_{02}^{\rho-1}) = 0 \\ \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial \lambda} & = (x_{01}^\rho + x_{02}^\rho)^{1/\rho} - y = 0 \end{cases}$$

De las primeras dos ecuaciones se obtiene

$$w_1 = \lambda \frac{1}{\rho}(x_{01}^\rho + x_{02}^\rho)^{(1-\rho)/\rho}(\rho x_{01}^{\rho-1}) \quad \text{y} \quad w_2 = \lambda \frac{1}{\rho}(x_{01}^\rho + x_{02}^\rho)^{(1-\rho)/\rho}(\rho x_{02}^{\rho-1})$$

si la solución es interior, dividiendo ambas ecuaciones llegamos a

$$\frac{w_1}{w_2} = \left(\frac{x_{01}}{x_{02}}\right)^{\rho-1}$$

despejando $x_2$ en la última ecuación y reemplazando en la restricción se obtiene

$$y = x_{01} w_1^{-1/(\rho-1)} (w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)})^{1/\rho}$$

y las demandas óptimas son

$$x_{01} = y w_1^{1/(\rho-1)} (w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)})^{-1/\rho}$$

$$x_{02} = y w_2^{1/(\rho-1)} (w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)})^{-1/\rho}$$

reemplazando estos valores en $w_1x_1 + w_2x_2$ se obtiene la función de costos en función de $(x_{01}, x_{02})$

$$C(x_{01}, x_{02}) = y(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)})^{(\rho-1)/\rho}$$

□

### Interpretación del multiplicador

Consideremos nuevamente el sistema

$$\begin{cases} \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_1} & = \dfrac{\partial f(x_{01}, x_{02})}{\partial x_1} - \lambda_0 \dfrac{\partial g(x_{01}, x_{02})}{\partial x_1} = 0 \\ \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_2} & = \dfrac{\partial f(x_{01}, x_{02})}{\partial x_2} - \lambda_0 \dfrac{\partial g(x_{01}, x_{02})}{\partial x_2} = 0 \\ \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial \lambda} & = g(x_{01}, x_{02}) - c = 0 \end{cases}$$

Viendo este sistema en conjunto como una función $L(x) = (L_1(x), L_2(x), L_3(x)) = 0$ el Jacobiano está dado por

$$DF(x_0) = \begin{bmatrix} -\frac{\partial g}{\partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_1^2}(x_0) - \lambda \frac{\partial^2 g}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) - \lambda \frac{\partial^2 g}{\partial x_2 \partial x_1}(x_0) \\ -\frac{\partial g}{\partial x_2}(x_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) - \lambda \frac{\partial^2 g}{\partial x_2 \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) - \lambda \frac{\partial^2 g}{\partial x_2^2}(x_0) \\ 0 & -\frac{\partial g}{\partial x_1}(x_0) & -\frac{\partial g}{\partial x_2}(x_0) \end{bmatrix}$$

y su determinante es $|DF(x_0)| \neq 0$. Dado esto el teorema de la función implícita nos dice que podemos expresar las demás variables en función de $c$,

$$\lambda_0 = \lambda(c) \quad \text{y} \quad (x_{01}, x_{02}) = (x_1(c), x_2(c))$$

entonces

$$L(x_{01}, x_{02}, \lambda_0) = f(x_{01}, x_{02}) - \lambda g(x_{01}, x_{02}, c)$$

diferenciando con respecto a $c$

$$\frac{dL}{dc} = \frac{\partial f(x_0)}{\partial x_1}\frac{dx_{01}}{dc} + \frac{\partial f(x_0)}{\partial x_2}\frac{dx_{02}}{dc} - [g(x_{01}, x_{02}) - c]\frac{d\lambda_0}{dc} - \lambda_0 \left[\frac{\partial g(x_0)}{\partial x_1}\frac{dx_{01}}{dc} + \frac{\partial g(x_0)}{\partial x_2}\frac{dx_{02}}{dc} - 1\right]$$

$$\frac{dL}{dc} = \left[\frac{\partial f(x_0)}{\partial x_1} - \lambda_0 \frac{\partial g(x_0)}{\partial x_1}\right]\frac{dx_{01}}{dc} + \left[\frac{\partial f(x_0)}{\partial x_2} - \lambda_0 \frac{\partial g(x_0)}{\partial x_2}\right]\frac{dx_{02}}{dc} - [g(x_{01}, x_{02}) - c]\frac{d\lambda_0}{dc} + \lambda_0$$

y debido a las condiciones de primer orden todos los terminos entre paréntesis se cancelan

$$\frac{dL}{dc} = \lambda_0$$

es decir, el **multiplicador** da cuenta de cómo cambia la solución óptima ante cambios en la restricción.

### Condiciones de segundo orden

A partir de las condiciones de primer orden de un problema con restricciones:

$$\begin{aligned} \nabla L(x) &= 0 \\ g^i(x) &= 0 \quad \forall i = \{1, \dots, m\} \end{aligned}$$

encontramos un punto $x_0$ que es candidato a máximo o mínimo local dependiendo del tipo de problema que enfrentamos. Lo que sigue a esto es preguntarse si efectivamente la solución encontrada corresponde a un óptimo del problema.

Ya habíamos dado una condición de segundo orden para el caso sin restricciones, que consistía determinar si la Matriz Hessiana de $f$ era definida negativa (caso maximización) o positivaen analizar (caso minimización).

Para el caso de un problema con restricciones, el criterio es similar pero no trabajaremos con la Matriz Hessiana de la función objetivo $f$ sino que debemos trabajar sobre la función Lagrangeano. El Hessiano del Lagrangeano es sólo con respecto a $x$ y está dado por

$$H_x L(x_0, \lambda) = \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2}(x_0, \lambda) & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n}(x_0, \lambda) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial x_n \partial x_1}(x_0, \lambda) & \cdots & \frac{\partial^2 L}{\partial x_n^2}(x_0, \lambda) \end{bmatrix}$$

No es necesario que el Hessiano sea definido positivo o negativo para cada dirección $v$, es necesario que el Hessiano, sólo con respecto a $x$, lo sea en un cierto conjunto que en la literatura se conoce como conjunto de direcciones críticas y lo definiremos como:

$$K(x) = \left\{ v \in \mathbb{R}^n \mid \nabla g^i(x_0) \cdot v = 0 \quad \forall i = \{1, \dots, m\} \ , \nabla f(x_0) \cdot v \geq 0 \right\}$$

para problemas de minimización y

$$K(x) = \left\{ v \in \mathbb{R}^n \mid \nabla g^i(x_0) \cdot v = 0 \quad \forall i = \{1, \dots, m\} \ , \nabla f(x_0) \cdot v \leq 0 \right\}$$

para problemas de maximización.

**Teorema 3.2.** *Sea $x_0 \in S = \left\{ x \mid g^i(x) = 0 \quad \forall i = \{1, \dots, m\} \right\}$. Supongamos que $\{\nabla g^1(x_0), \dots, \nabla g^m(x_0)\}$ es linealmente independiente, entonces existe $\lambda \in \mathbb{R}^m$ tal que $\nabla L(x_0, \lambda) = 0$. Si además se tiene que*

$$v^T H_x L(x_0, \lambda) v > 0 \quad \forall v \in K(x), \ v \neq 0$$

*entonces $x_0$ es un mínimo local del problema de optimización ya que el Hessiano es definido positivo. Mientras que si se tiene*

$$v^T H_x L(x_0, \lambda) v < 0 \quad \forall v \in K(x), \ v \neq 0$$

*entonces $x_0$ es un máximo local del problema de optimización ya que el Hessiano es definido negativo.*

**Ejemplo 3.5.** Considere el problema de minimo costo de una Cobb-Douglas dado por

$$\begin{array}{ll} \min\limits_{x_1,x_2} & 2x_1 + 3x_2 \\ \text{s.a} & x_1^{0,3} x_2^{0,3} = 100 \end{array}$$

Encuentre las cantidades $(x_{01}, x_{02})$ que minimizan costos y determine si se cumplen las condiciones de segundo orden.

Solución. El lagrangeano del problema es

$$L(x_1, x_2, \lambda) = 2x_1 + 3x_2 - \lambda(x_1^{0,3} x_2^{0,3} - 100)$$

lo cual genera el siguiente sistema

$$\begin{cases} \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_1} & = 2 - \frac{3\lambda}{10} x_{01}^{-0,7} x_{02}^{0,3} = 0 \\ \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial x_2} & = 3 - \frac{3\lambda}{10} x_{01}^{0,3} x_{02}^{-0,7} = 0 \\ \dfrac{\partial L(x_{01}, x_{02}, \lambda_0)}{\partial \lambda} & = x_1^{0,3} x_2^{0,3} - 100 = 0 \end{cases}$$

Si la solución es interior, dividiendo las primeras dos ecuaciones se obtiene

$$x_{02} = \frac{2x_{01}}{3}$$

reemplazando en la restricción

$$100 = x_{01}^{0,3} \left( \frac{2x_{01}}{3} \right)^{0,3}$$

y así se obtienen las demandas óptimas

$$x_{01} = 3076,42$$

$$x_{02} = 2050,95$$

despejando en cualquiera de las dos primeras ecuaciones se obtiene el valor del multiplicador

$$\lambda(x_{01}, x_{02}) = 15$$

Luego, las condiciones de segundo orden son

$$\begin{aligned} \frac{\partial^2 L(x_{01}, x_{02}, \lambda_0)}{\partial x_1^2} &= \frac{21\lambda}{100} x_{01}^{-1,7} x_{02}^{0,3} \\ \frac{\partial^2 L(x_{01}, x_{02}, \lambda_0)}{\partial x_2 \partial x_1} &= \frac{9\lambda}{100} x_{01}^{-0,7} x_{02}^{-0,7} \\ \frac{\partial^2 L(x_{01}, x_{02}, \lambda_0)}{\partial x_1 \partial x_2} &= \frac{9\lambda}{100} x_{01}^{-0,7} x_{02}^{-0,7} \\ \frac{\partial^2 L(x_{01}, x_{02}, \lambda_0)}{\partial x_2^2} &= \frac{21\lambda}{100} x_{01}^{0,3} x_{02}^{-1,7} \end{aligned}$$

y entonces el Hessiano corresponde a

$$H_x L(x_0, \lambda) = \frac{1}{100} \begin{bmatrix} 21\lambda x_{01}^{-1,7} x_{02}^{0,3} & 9\lambda x_{01}^{-0,7} x_{02}^{-0,7} \\ 9\lambda x_{01}^{-0,7} x_{02}^{-0,7} & 21\lambda x_{01}^{0,3} x_{02}^{-1,7} \end{bmatrix}$$

como el escalar que multiplica la matriz es positivo, nos bastará con determinar los determinantes de la matriz que aparece arriba

$$|H_1| = 21\lambda x_{01}^{-1,7} x_{02}^{0,3} = 0,00365 > 0$$
$$|H_2| = 360\lambda^2 x_{01}^{-1,4} x_{02}^{-1,4} = 0,000024 > 0$$

aunque los valores puedan parecernos extremadamente pequeños se cumple que la Matriz Hessiana es definida positiva y la solución encontrada efectivamente minimiza la función objetivo. □

Veamos ahora un ejemplo con "trampa":

**Ejemplo 3.6.** Consideremos el siguiente problema

$$\begin{array}{cl} \max\limits_{x,y,z} & xy + yz + xz \\ \text{s.a} & x + y + z = 3 \end{array}$$

Nos interesa saber si la solución (en caso de que exista) efectivamente corresponda a un máximo local estricto.

Solución. El Lagrangeano corresponde a

$$L(x, \lambda) = xy + yz + xz + \lambda(1 - x - y - z)$$

Aplicando el sistema Lagrangeano llegamos a lo siguiente

$$\begin{cases} y_0 + z_0 - \lambda = 0 \\ x_0 + z_0 - \lambda = 0 \\ x_0 + y_0 - \lambda = 0 \\ x_0 + y_0 + z_0 = 1 \end{cases} \implies \begin{array}{l} x_0 = y_0 \\ y_0 = z_0 \\ x_0 + y_0 + z_0 = 3 \end{array} \implies \begin{array}{l} x_0 = 1 \\ y_0 = 1 \\ z_0 = 1 \end{array}$$

Entonces $f(x_0, y_0, z_0) = 1$ y el valor del multiplicador de Lagrange es $\lambda = 2$.

Para las condiciones de segundo orden tenemos

$$Hf(x_0) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad Hg(x_0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \implies H_x L(x_0, \lambda) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Podríamos usar el criterio de las submatrices para determinar si la matriz resultante es definida positiva o negativa. Usando esto se obtiene $|H_1| = 0$, $|H_2| = -1$ y $|H_3| = 2$ lo cual bajo tal criterio nos dice que la matriz no es semi definida negativa ni tampoco semi definida positiva.

El teorema de condición de segundo orden establece una condición que pide que $H_x L(x_0, \lambda)$ sea semi definida positiva (caso minimización) o semi definida negativa (caso maximización) a lo menos en $K(x)$.

Definamos $h = (x, y, z)$ y entonces

$$h^T H_x L(x_0, \lambda) h = x(y + z) + y(x + z) + z(x + y) \tag{1}$$

Tengamos presente que en $K(x)$ se cumple que $\nabla g(x) \cdot h = 0$, entonces $x + y + z = 0$ por lo que podemos reescribir la última ecuación como

$$h^T H_x L(x_0, \lambda) h = -(x^2 + y^2 + z^2) < 0$$

observemos que $-(x^2 + y^2 + z^2) \leq 0$ pero restringido a $K(x)$ se tiene la desigualdad estricta y entonces la solución encontrada es un máximo local estricto. □

**Ejemplo 3.7.** La utilidad del agente representativo de una economía depende del consumo de $n$ productos y está dada por

$$u(x) = \prod_{i=1}^{n} x_i^{\alpha_i}$$

Con $\alpha_i > 0\ \forall i$ y $\sum_{i=1}^{n} \alpha_i = 1$. La restricción al consumo viene dada por un ingreso $I$ que gasta en su totalidad en productos perfectamente divisibles, los cuales se venden a un precio $p_i$ cada uno.

Plantee el problema de maximización y encuentre la demanda óptima por el producto j-ésimo que resuelve el problema. Finalmente obtenga una expresión para la función de utilidad en términos de las demandas óptimas.

Solución. El problema a resolver es

$$\max_{x} \prod_{i=1}^{n} x_i^{\alpha_i} \quad \text{s.a} \quad \sum_{i=1}^{n} p_i x_i = I$$

luego, el Lagrangeano corresponde a

$$L(x,\lambda) = \prod_{i=1}^{n} x_i^{\alpha_i} - \lambda \left( \sum_{i=1}^{n} p_i x_i - I \right)$$

y las condiciones de primer orden son

$$\frac{\partial L}{\partial x_i} = \frac{\alpha_i \prod_{i=1}^{n} x_i^{\alpha_i}}{x_i} - \lambda p_i = 0 \tag{*}$$

$$\frac{\partial L}{\partial \lambda} = I - \sum_{i=1}^{n} p_i x_i = 0 \tag{**}$$

Como $u(x) = \prod_{i=1}^{n} x_i^{\alpha_i}$, vamos a reemplazar en (*) y obtenemos

$$\alpha_i u = \lambda p_i x_i \tag{***}$$

De esto obtenemos $\lambda$ en términos de $(u, \alpha, p, x)$

$$\lambda = \frac{\alpha_i u}{p_i x_i}$$

Consideremos que $I = \sum_{i=1}^{n} p_i x_i$ y $\sum_{i=1}^{n} \alpha_i = 1$ por lo que en (***) vamos a agregar términos aplicando la sumatoria $\sum_{i=1}^{n} (\cdot)$

$$u = \lambda I$$

Si reemplazamos esto último en (***) obtenemos

$$\alpha_i I = p_i x_i$$

Sólo falta reordenar y obtenemos la demanda por el insumo j-ésimo (los índices son independientes) en términos de $(I, \alpha, p)$

$$x_j = \frac{I \alpha_j}{p_j}$$

Reemplazamos directamente el último resultado en $u(x) = \prod_{i=1}^{n} x_i^{\alpha_i}$ y obtenemos

$$u(x) = I \prod_{i=1}^{n} \left( \frac{\alpha_i}{p_i} \right)^{\alpha_i}$$

$\square$

## 3.3. Multiplicadores de Karush-Kuhn-Tucker

**Teorema 3.3.** (Teorema de Karush-Kuhn-Tucker[2], igualdades y desigualdades)
*Sean* $f, g_i, h_j : \mathbb{R}^n \to \mathbb{R}$ *funciones diferenciables* $\forall i = \{1, \dots, k\}$*,* $j = \{1, \dots, m\}$ *y* $f, h_j$ *son funciones cóncavas* $\forall j = \{1, \dots, m\}$ *y consideremos el siguiente problema de optimización:*

$$\begin{array}{ll} \max\limits_{x} & f(x) \\ s.a & g^i(x) = 0 \quad \forall i = \{1, \dots, k\} \\ & h^j(x) \geq 0 \quad \forall j = \{1, \dots, m\} \end{array}$$

*El problema se puede resolver mediante la función Lagrangeano*

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{k} \lambda_i g^i(x) + \sum_{j=1}^{m} \mu_j h^j(x)$$

*Si* $f$ *alcanza un máximo local en* $x_0 \in S$ *una condición necesaria para que* $x_0$ *sea solución del problema es que existan* $(\lambda, \mu) \in \mathbb{R}^{k+m}$*, que corresponde a una familia de multiplicadores de Karush-Kuhn-Tucker asociados con* $x_0$*, tales que*

$$\begin{aligned} \frac{\partial L}{\partial x_i}(x_0, \lambda, \mu) = \nabla f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{m} \mu_j \nabla h_j(x_0) &= 0 \\ g^i(x_0) = 0 \quad \forall i &= \{1, \dots, k\} \\ \mu_j h^j(x_0) = 0 \quad \forall j &= \{1, \dots, m\} \\ \mu_j \geq 0 \quad \forall j &= \{1, \dots, m\} \end{aligned}$$

*Además, si* $h^j(x_0) > 0$ *entonces* $\mu_j = 0$ *de forma tal que si* $h^j(x_0) = 0$ *entonces* $\mu_j > 0$ *por lo que en la tercera condición uno y sólo uno de los términos de la multiplicación se anula. Cuando* $f$ *es estrictamente cóncava las condiciones descritas son suficientes para que* $x_0$ *sea mínimo de* $f$ *en* $S$*.*

Nota 3.3. Este teorema no se demostrará porque la demostración requiere bastantes detalles. Un desarrollo completo se encuentra en el item 1. de la bibliografía. La importancia práctica es que permite resolver problemas con restricciones de desigualdad y problemas en que todas las funciones involucradas son lineales.

Nota 3.4. **Importante:** Hay que ser cuidadosos con el orden en que se colocan las restricciones y las constantes en las restricciones de desigualdad. Si el problema sólo tiene restricciones de igualdad, no hay restricciones sobre el signo del multiplicador porque no se puede decir algo respecto de su signo. En otro caso, el signo del multiplicador depende de si criterio es maximizar o minimizar y si las restricciones son de mayor o menor o igual. Por ejemplo si estamos minimizando con restricciones de menor o igual a cero, lo mismo que si estamos maximizando con restricciones de mayor o igual a cero, entonces los multiplicadores son todos mayores o iguales a cero y el signo de los multiplicadores en el Lagrangeano es positivo.

**Ejemplo 3.8.** Considere el problema del consumidor dado por

$$\begin{array}{ll} \max\limits_{x} & U(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha} \\ \text{s.a} & I = p_1 x_1 + p_2 x_2 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{array}$$

[2] William Karush (1917-1997) desarrolló estas condiciones en 1939 para su tesis de magíster. Se aprobó su tesis pero se le dijo que su trabajo no tenía gran relevancia en aspectos prácticos. Posteriormente en 1951, Harold Kuhn (1925 - ) y Albert Tucker (1905 - 1995) llegaron de forma independiente a estas condiciones pero bajo hipótesis más generales que son las que conocemos hoy. Tras redescubrirse el trabajo de Karush posterior a 1951, se rebautizaron las condiciones de Kuhn-Tucker como condiciones de Karush-Kuhn-Tucker

Encuentre las soluciones óptimas.

Solución. El lagrangeano está dado por

$$L(x,\lambda,\mu) = x_1^{\alpha}x_2^{1-\alpha} + \lambda(I - p_1x_1 + p_2x_2) + \mu_1(0 - x_1) + \mu_2(0 - x_2)$$

luego las condiciones de primer orden generan el siguiente sistema

$$\begin{cases} \alpha x_{01}^{\alpha}x_{02}^{1-\alpha} - \lambda p_1 + \mu_1 & = 0 \\ (1-\alpha)x_{01}^{\alpha}x_{02}^{-\alpha} - \lambda p_2 + \mu_2 & = 0 \\ p_1x_1 + p_2x_2 = I & , \lambda \text{ sin restricción de signo} \\ \mu_1 x_{01} = 0 & , \mu_1 \geq 0, \ x_{01} \geq 0 \\ \mu_2 x_{02} = 0 & , \mu_1 \geq 0, \ x_{02} \geq 0 \end{cases}$$

Si $\mu_1 > 0$ y $\mu_2 = 0$ entonces $x_{01} = 0$ (cuarta condición), $x_{02} = \frac{I}{p_2}$ (tercera condición) y $U(x_{01}, x_{02}) = 0$.

Si $\mu_1 = 0$ y $\mu_2 > 0$ entonces $x_{02} = 0$ (quinta condición), $x_{01} = \frac{I}{p_1}$ (tercera condición) y $U(x_{01}, x_{02}) = 0$.

Si $\mu_1 > 0$ y $\mu_2 > 0$ entonces $x_{01} = 0$ (cuarta condición), $x_{02} = 0$ (quinta condición) pero se tendría $I = 0$ y $U(x_{01}, x_{02}) = 0$.

Si $\mu_1 = 0$ y $\mu_2 = 0$ entonces de las primeras dos ecuaciones se obtiene

$$\lambda = \frac{\alpha x_{01}^{\alpha-1}x_{02}^{1-\alpha}}{p_1} = \frac{(1-\alpha)x_{01}^{\alpha}x_{02}^{-\alpha}}{p_2} \Longrightarrow x_{02} = \frac{(1-\alpha)p_1x_{01}}{\alpha p_2}$$

. Luego reemplazando en $I = p_1x_1 + p_2x_2$

$$x_{01} = \frac{\alpha I}{p_1} \qquad x_{02} = \frac{(1-\alpha)I}{p_2}$$

□

Nota 3.5. Este método nos permite resolver la maximización de una utilidad lineal o cuasilineal. Dejamos de ejercicio resolver el problema del consumidor para las siguientes funciones

$$\begin{aligned} U(x_1,x_2) &= x_1 + x_2 \\ U(x_1,x_2) &= x_1x_2 \\ U(x_1,x_2) &= x_1^{0,3} + x_2^{0,4} \\ U(x_1,x_2) &= x_1^{0,7}x_2^{0,3} + x_1 \end{aligned}$$

Daremos la solución del primer caso y en los demás sólo adelantaremos que se llega a soluciones de esquina que mediante multiplicadores de Lagrange no aparecen porque sólo nos permitían encontrar la solución interior.

**Ejemplo 3.9.** Considere el problema del consumidor dado por

$$\begin{array}{ll} \max\limits_{x} & U(x_1,x_2) = x_1 + x_2 \\ \text{s.a} & I = p_1x_1 + p_2x_2 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{array}$$

Encuentre las soluciones óptimas.

Solución. El lagrangeano está dado por

$$L(x,\lambda,\mu) = x_1 + x_2 + \lambda(I - p_1x_1 + p_2x_2) + \mu_1(0 - x_1) + \mu_2(0 - x_2)$$

luego las condiciones de primer orden generan el siguiente sistema

$$\begin{cases} 1 - \lambda p_1 + \mu_1 & = 0 \\ 1 - \lambda p_2 + \mu_2 & = 0 \\ p_1x_1 + p_2x_2 = I & , \ \lambda \text{ sin restricción de signo} \\ \mu_1 x_{01} = 0 & , \ \mu_1 \geq 0, \ x_{01} \geq 0 \\ \mu_2 x_{02} = 0 & , \ \mu_1 \geq 0, \ x_{02} \geq 0 \end{cases}$$

Si $\mu_1 > 0$ y $\mu_2 = 0$ entonces $x_{01} = 0$ (cuarta condición), $x_{02} = \frac{I}{p_2}$ (tercera condición) y $U(x_{01}, x_{02}) = x_2$.

Si $\mu_1 = 0$ y $\mu_2 > 0$ entonces $x_{02} = 0$ (quinta condición), $x_{01} = \frac{I}{p_1}$ (tercera condición) y $U(x_{01}, x_{02}) = x_1$.

Si $\mu_1 > 0$ y $\mu_2 > 0$ entonces $x_{01} = 0$ (cuarta condición), $x_{02} = 0$ (quinta condición) pero se tendría $I = 0$ y $U(x_{01}, x_{02}) = 0$.

Si $\mu_1 = 0$ y $\mu_2 = 0$ entonces de las primeras dos ecuaciones se obtiene

$$\frac{p_1}{p_2} = 1$$

y entonces la restricción coincide exactamente con la función de utilidad, por lo tanto cualquier combinación

$$\alpha x_{01} + (1-\alpha)x_{02} = \frac{\alpha I}{p_1} = \frac{\alpha I}{p_2}$$

con $\alpha \in [0,1]$ es una solución del problema. □

**Ejemplo 3.10.** Un inversionista que no estudío en la Facultad de Economía y Negocios de la Universidad de Chile tiene la posibilidad de invertir en $n$ activos $x_1, \ldots, x_n$ que ofrecen una tasa de retorno aleatoria $r_1, \ldots, r_n$ respectivamente y cada activo tiene una tasa de retorno promedio $\overline{r_i} = E(r_i)$ para $i = 1, \ldots, n$ y la covarianza del activo $i$ con el activo $j$ es $\sigma_{ij}$ para $j = 1, \ldots, n$. El portafolio $y = \sum_{i=1}^n r_i x_i$ formado por todos los activos tiene una tasa media de retorno dada por

$$E(y) = \sum_{i=1}^n \overline{r_i} x_i$$

mientras que la varianza de la inversión está dada por

$$\sigma^2 = E[(\sigma - \overline{\sigma})^2] = \sum_{i=1}^n \sum_{j=1}^n x_i \sigma_{ij} x_j$$

Lo que le interesa al inversionista es minimizar la volatilidad de la inversión y lo contrata a usted para que resuelva el problema. Plantee el problema y encuentre la solución.

Solución. El enunciado se traduce en

$$\begin{aligned} \min_x \quad & \sum_{i=1}^n \sum_{j=1}^n x_i \sigma_{ij} x_j \\ \text{s.a} \quad & \sum_{i=1}^n \overline{r_i} x_i = \overline{y}, \ \sum_{i=1}^n x_i = 1 \\ & x_i \geq 0 \ \forall i = 1, \ldots, n \end{aligned}$$

la última restricción nos sirve para normalizar las cantidades invertidas y si los pesos relativos suman uno, bastará con ponderar la solución óptima por algún escalar y se obtiene la cantidad pedida.

La función lagrangeano para este problema es

$$L(x,\lambda,\mu) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\sigma_{ij}x_j - \lambda_1\left(\sum_{i=1}^{n}\overline{r_i}x_i - \overline{y}\right) - \lambda_2\left(\sum_{i=1}^{n}x_i - 1\right) - \sum_{i=1}^{n}\mu_i x_i$$

Si la solución es interior, es decir $x_i > 0\ \forall i$ se puede derivar con respecto a $x_i$ para la condición de primer orden y llegamos a

$$\sum_{j=1}^{n}\sigma ij x_j - \lambda_1\overline{r_i} - \lambda_2 = 0$$

Es posible expresar esta condición considerando todos los activos, esto se obtiene con una expresión vectorial dada por

$$2Qx_0 - \lambda_1\cdot e - \lambda_2\cdot\overline{r} = 0$$

donde $Q$ denota la matriz de las covarianzas $\sigma_{ij}$, $e$ es el vector canónico de $\mathbb{R}^n$, es decir $e_i = (0,\ldots,1,\ldots,0)$, y $\overline{r} = (\overline{r_1},\ldots,\overline{r_n})$. Como la solución es interior todos los $\mu_i$ son cero. Luego, si $e$ y $\overline{r}$ son linealmente independientes, el teorema de Kuhn-Tucker es válido, en otro caso la validez no se pierde porque las restricciones son lineales (analice esto pero no por mucho tiempo, piense sólo en la idea de esta propiedad).

Si $Q$ es invertible, entonces

$$x_0 = \frac{1}{2}\left(Q^{-1}\lambda_1\cdot e + Q^{-1}\lambda_2\cdot\overline{r}\right)$$

y así se obtendrán las soluciones $(x_{01},\ldots,x_{0n})$

Por sustitución

$$x_0\cdot e = 1 \qquad x_0\cdot\overline{r} = \overline{y}$$

entonces

$$\begin{aligned}1 = x_0\cdot e &= \frac{1}{2}e\cdot(Q^{-1}e)\lambda_1 + \frac{1}{2}e\cdot(Q^{-1}\overline{r})\lambda_2\\ \overline{y} = x_0\cdot\overline{r} &= \frac{1}{2}\overline{r}\cdot(Q^{-1}e)\lambda_1 + \frac{1}{2}r\cdot(Q^{-1}\overline{r})\lambda_2\end{aligned}$$

como esto genera un sistema de ecuaciones se puede resolver para despejar $\lambda_1$ y $\lambda_2$ que son escalares de la forma $a + bx$ y se obtiene

$$\begin{aligned}\lambda_1 &= a_1 + b_1\overline{y}\\ \lambda_2 &= a_2 + b_2\overline{y}\end{aligned}$$

donde $a$ y $b$ son constantes que dependen del sistema anterior. Retomando la ecuación

$$x_0\cdot e = 1 \qquad x_0\cdot\overline{r} = \overline{y}$$

si reemplazamos $\lambda_1$ y $\lambda_2$ se llega a

$$x_0 = \overline{y}v + w$$

donde $v$ y $w$ son vectores de $\mathbb{R}^n$ que dependen de $Q$ y $\overline{r}$, luego la varianza de la inversión es

$$\sigma^2 = (\overline{y}v + w)\cdot[Q(\overline{y}v + w)] = (\alpha\overline{y} + \beta)^2 + \gamma$$

donde $\alpha$, $\beta$ y $\gamma$ dependen de $Q$ y $\bar{r}$.

En base a esto se puede construir una frontera de portafolios eficientes. Cada portafolio eficiente corresponde a un promedio ponderado de dos portafolios que se encuentren en la frontera de eficiencia como se ve en la siguiente figura:

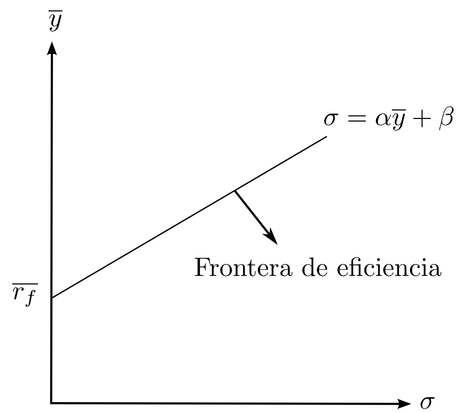


Figura 17: Frontera de eficiencia.

$\square$

Nota 3.6. El último ejemplo está muy lejos de ser abstracción. De hecho, es la base del modelo CAPM que se usa ampliamente en finanzas.

# 4. Equilibrio

Los teoremas que presentamos en esta parte son formas particulares de lo que se puede encontrar en la literatura. La idea es exponerlos nada más que para fijar ideas e introducir algunos conceptos que se utilizan en microeconomía y macroeconomía. Para más detalles y demostraciones de estos teoremas, el material de la bibliografía es suficiente y abordable.

Tengamos presente la siguiente definición: Una función $f : \mathbb{R}^n \to \mathbb{R}^n$ se dice *contractante* si existe $c \in (0, 1)$ tal que

$$\|f(x) - f(x_0)\| \leq c\|x - x_0\| \quad \forall x \in \mathbb{R}^n, x_0 \in \mathbb{R}^n$$

**Teorema 4.1.** (Teorema del punto fijo de Banach)
*Sea $F : \mathbb{R}^n \to \mathbb{R}^n$ una función contractante que representa un sistema de $n$ ecuaciones de la forma $f_i(x) = 0$ tales que $f_i : D \subset \mathbb{R}^n \to \mathbb{R}$ que se resumen en $F(x) = 0$, entonces existe un único $x \in \mathbb{R}^n$ tal que $f(x) = x$.*

**Teorema 4.2.** (Teorema del punto fijo de Brouwer)
*Sea $D = \overline{B}(x_0, r) \subset \mathbb{R}^n$ con $r$ finito tal que $D$ siempre es convexo además de cerrado, acotado y no vacío. Definamos $f : D \to D$. Si $f$ es continua entonces existe al menos un $x \in D$ tal que $f(x) = x$.*

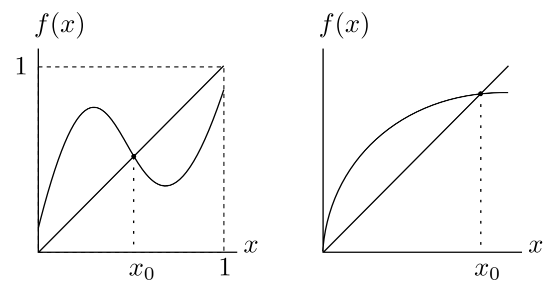


Figura 18: Idea geométrica del teorema del punto fijo de Brouwer y Banach respectivamente.

**Ejemplo 4.1.** Sean $Q \in \mathbb{R}^n$ y $p \in \mathbb{R}^n$ vectores de cantidades y precios. Sea $Q = D(p)$ la función de demanda por las cantidades $Q$. Sea $Q = S(p)$ la función de oferta de las cantidades. Si los oferentes a precios $p$ operan en el mercado su oferta está dada por $Q = S(p)$. Si la oferta en el mercado es $Q$ el vector de precios que prevalece es $p_0 = D^{1-}(Q) = D^{-1}[S(p)]$. Un vector de precios de equilibrio es tal que $p_0 = p$, es decir un vector de precios de equilibrio es $p = f(p)$ con $f(P) = D^{-1}[S(p)]$ que significa que no hay exceso de oferta ni demanda en ningún mercado.

Si el vector de precios $p$ es tal que ningún $p_i$ es menor a cero, ¿Se cumple la condición del teorema del punto fijo de Brouwer? Sí, siempre y cuando sea un vector que define un conjunto de precios que es convexo, cerrado, acotado y no vacío. Pensemos en un vector $p \in \mathbb{R}^3$, la condición del teorema se entiende facilmente si lo vemos geométricamente como que la bola $B(p, r)$ es similar a una pelota de fútbol tal que contiene a su propia frontera, no hay vacíos en su interior y su radio en ninguna dirección va para infinito.

En economía nos importan los precios relativos. Esto esto, si ponderamos todos los precios por un mismo escalar las cantidades ofertadas y demandadas no cambian, esto nos dice que las funciones de oferta y demanda son homogéneas de grado cero en precios. Entonces los precios se pueden normalizar, por ejemplo si definimos $p_i = 1$ entonces la suma de los precios es

$$\sum_{j=1}^{n} \frac{p_j}{p_i} = \frac{p_1}{p_i} + \ldots + 1 + \ldots + \frac{p_n}{p_i}$$

otra forma de normalizar es dividir todos los precios por $\sum_{j=1}^{n} p_j$, entonces

$$\frac{\sum_{i=1}^{n} p_i}{\sum_{j=1}^{n} p_j} = 1$$

para concluir podemos definir un vector de precios relativos

$$p' = \frac{p}{\|p\|} \quad \text{tal que} \quad \|p'\| = \frac{\|p\|}{\|p\|} = 1$$

Cualquier normalización de precios hace que el conjunto de precios cumpla las condiciones del teorema del punto fijo de Brouwer, la última normalización deja esto en evidencia.

**Ejemplo 4.2.** Consideremos una economía donde se transan $n$ productos a precios $p_i$ cada uno en que no todos los precios son cero, de lo contrario se demandaría infinito en todas las componentes del vector de cantidades y no habría problema de escasez como dice la economía. El precio relativo del producto $j$ está dado por

$$\overline{p_j} = \frac{p_j}{\sum_{i=1}^{n} p_i}$$

luego

$$\sum_{j=1}^{n} \overline{p_j} = \frac{\sum_{j=1}^{n} p_j}{\sum_{i=1}^{n} p_i} = 1$$

Recordemos que las demandas son homogeneas de grado cero en precios, entonces si definimos una función de exceso de demanda esta es homogenea de grado cero en precios. Luego, si en equilibrio existe exceso de oferta por algún bien este tendría precio cero (analizar esto). Definamos el exceso de demanda como

$$ED_i(P) = Q_i^d(P) - Q_i^s(P)$$

luego esta función tiene sentido económicamente hablando si

$$\begin{aligned} ED_i(P_0) &= 0 \quad \text{para} \quad p_{0i} > 0 \\ ED_i(P_0) &\leq 0 \quad \text{para} \quad p_{0i} = 0 \end{aligned}$$

Como sólo interesan los precios relativos, se puede cosntruir una función continua que transforme los precios relativos de equilibrio en otros precios relativos que también serán de equilibrio si los iniciales lo son, definamos entonces

$$F^i(P) = p_i + ED_i(P) \; \forall i$$

Analicemos la intuición detrás de esta función: Si al precio $p_i$ hay un exceso de demanda del producto $i$ entonces $ED_i(P) > 0$ y sube el precio $p_i$ hasta que $ED_i(P) \leq 0$. Si hay un exceso de oferta es análogo y el precio baja hasta el equilibrio. Dijimos de antemano que la función de exceso de demanda no se explicita la forma que tiene pero cualquier función continua nos sirve si razonamos de manera general y no de manera puntual. Entonces, $F^i$ es una función continua.

Hasta ahora todo se mueve en torno a los precios relativos. Sin embargo, hay un problema con la función $F^i$: Esta función perfectamente, independientemente de su continuidad, puede transformar precios relativos a precios relativos negativos lo cual es un detalle que intuitivamente es muy difícil de ver. Para evitar este problema se puede redefinir $F^i$ de la forma

$$F^i(P) = \text{máx}\{p_i + ED_i(P), 0\}$$

de esta forma siempre los precios, tras la transformación, serán $p_i \geq 0\ \forall i$ y $F^i$ sigue siendo continua.

Ahora aparece otro problema, tras la transformación no necesariamente la suma de todos los precios relativos es igual a uno. Por un argumento económico supongamos que a lo menos un precio es estrictamente positivo, es decir, $p_i + ED_i(P) > 0$ para a lo menos un $i$. Por el contrario, supongamos que

$$p_i + ED_i(P) \leq 0\ \forall i$$

multiplicando por $p_i$ y agregando todos los $p_i$ se tendrá que

$$\sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n} p_i ED_i(P) \leq 0\ \forall i$$

pero en este caso habría exceso de oferta para todos los productos de la economía, es decir $\sum_{i=1}^{n} p_i ED_i(P) = 0$ y por lo tanto

$$\sum_{i=1}^{n} p_i^2 \leq 0$$

que sólo se cumple con igualdad y así se tendría que $p_i = 0\ \forall i$, entonces el supuesto nos lleva nuevamente a que al menos un precio es estrictamente positivo en la economía.

Podemos entonces, normalizar la transformación de los precios, digamos que tras normalizar

$$\sum_{i=1}^{n} F^i(P) = 1$$

y esta normalización más allá del sentido económico cumple las condiciones del teorema del punto fijo de Brouwer. Como $F : D \to D$, es decir considerando un vector de precios, es una función de la forma $(F^1, \ldots, F^n)$ cumple las condiciones del teorema y existe $F^i(P_0)P_0$, en ese punto se cumple que

$$p_{0i} = \text{máx}\{p_i + ED_i(P), 0\}\ \forall i$$

entonces $P_0$ es un conjunto de precios de equilibrio para $p_{0i} > 0$, es decir

$$p_{0i} = p_{0i} + ED_i(P_0) \Rightarrow ED_i(P_0) = 0$$

mientras que para $p_{0i} = 0$

$$p_{0i} + ED_i(P) \leq 0 \Rightarrow ED_i(P) \leq 0$$

y se llega a que en equilibrio ningún precio es negativo.

Un hecho aparte es que, bajo condiciones particulares, el equilibrio de oferta y demanda podría no ser único o que existan fallas de mercado como monopolio. Lo anterior se puede entender de la siguiente forma: Bajo competencia perfecta, el sistema de precios opera de manera similar a la idea de la mano invisible ed Adam Smith.

Nota 4.1. La importancia de los teoremas de punto fijo recae en que si tenemos un problema que nos lleva a tener que resolver $f(x) = 0$, es lo mismo que resolver $f(x) + x = x$ y definiendo $g(x) = f(x) + x$ el problema es equivalente a encontrar un $x$ tal que $g(x) = x$. Se tiene entonces que si es posible encontrar un punto fijo entonces se puede resolver la ecuación. Adelantandonos un poco, estos resultados además de los que proporcionan otros teoremas de punto fijo, sirve para demostrar la existencia de equilibrio en una economía y se utiliza en problemas que dependen del tiempo en macroeconomía.

## 5. Bibliografía


1. Jofré, A. et al. **Apuntes de Clases: Cálculo en Varias Variables**. Departamento de Ingeniería Matemática, Universidad de Chile, 2011.

2. Krell, R., Rojas, E., Torres-Martínez, J. **Apuntes de Microeconomía I: Teoría del Consumidor**. Departamento de Economía, Universidad de Chile, 2009.

3. Luenberger, D., Ye, Y. **Linear and Nonlinear Programming**. Springer, 2008.

4. Piskunov, N. **Cálculo Diferencial e Integral, (vol. I y II)**. Editorial MIR, 1970.

5. Takayama, A. **Analytical Methods in Economics**. Editorial Harvester Wheatsheaf, 1994.