Departamento de Estadística y Matemática

Documento de Trabajo Nº 3

Facultad de Ciencias Económicas
Universidad Nacional de Córdoba

OPTIMIZACIÓN ESTÁTICA MULTIVARIANTE

Jorge Mauricio Oviedo[1]

[1] joviedo@eco.unc.edu.ar

Introducción

Las presentes notas de Optimización tienen por objeto brindar una guía intuitiva y práctica a la hora de resolver problemas de Maximización y Minimización de funciones en una y varias variables. La motivación de realizar este trabajo surge por la necesidad de suplir las deficiencias del plan de estudios para la Carrera de Licenciatura de Economía en virtud de que dichos conocimientos son necesarios en Materias como Microeconomía I y los alumnos no los poseen en esta etapa de la carrera.

En este sentido se destaca que las mismas no tienen por objeto realizar un análisis minucioso y detallado tanto de los teoremas que se involucran como a las demostraciones que los fundamentan, solamente se concentran en brindar simples “recetas prácticas” con fundamentaciones geométricas e intuitivas. Dicha aclaración es válida en el sentido de que de no hacerse este trabajo puede interpretarse como un estudio superficial que falta respeto al rigor matemático.

Clasificación de los Problemas de Optimización

En una primera gran división los problemas de Optimización pueden clasificarse:

- **Optimización Estática:** donde el objetivo es hallar el valor numérico de una o varias variables que optimizan el valor de una función en una o varias variables respectivamente. A su vez éstos problemas pueden subclasificarse en:
  - Optimización Libre
  - Optimización sujeto a restricciones de igualdad
  - Optimización sujeto a restricciones de desigualdad

- **Optimización Dinámica:** donde el objetivo es hallar una función o conjunto de funciones que maximicen o minimicen una integral definida. Estos problemas pueden clasificarse en tiempo discreto y continuo por un lado, en problemas determinísticos o aleatorios por otro y dentro de estos en variables aleatorias continuas o discretas. Este tipo de optimización exceden el alcance de este trabajo y se proveerá en un trabajo posterior.

Optimización Libre

Caso 2 variables:

En este problema el objetivo consiste en determinar los valores de x_1 y x_2 que hacen máximo o mínimo el valor de una función objetivo *f*. Formalmente el problema puede escribirse de la siguiente manera:

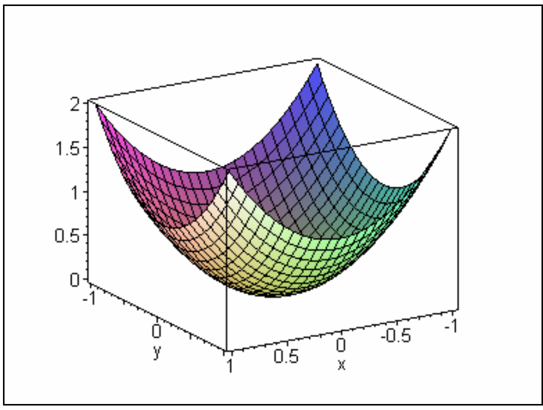
$$\boxed{\underset{x_1,x_2}{Max}\ f = f(x_1, x_2)}$$

Donde la condición de primer orden (condición necesaria) viene dada por el siguiente sistema de ecuaciones[2]:

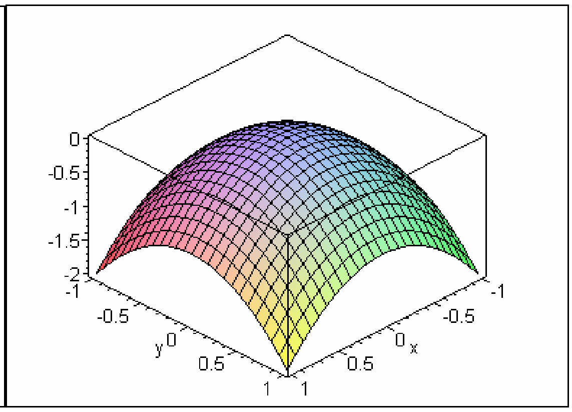
$$\begin{cases} f_{x_1}(x_1, x_2) = 0 \\ f_{x_2}(x_1, x_2) = 0 \end{cases} \Rightarrow (x_1^*, x_2^*)$$

Geométricamente la condición de primer orden establece que hay que buscar los valores de las variables, es decir el punto del dominio para el cual el plano tangente a la gráfica de *f* es horizontal. Dicho sistema por lo general es no lineal y puede tener varias soluciones, es decir pueden existir varios puntos para los cuales el plano tangente a la gráfica de *f* es horizontal. Dichos puntos son denominados puntos críticos o candidatos a óptimos de *f*. ya que los puntos donde *f* presenta plano tangente horizontal pueden ser máximos, mínimos o puntos de inflexión (puntos de silla) como se aprecian en las siguientes gráficas:

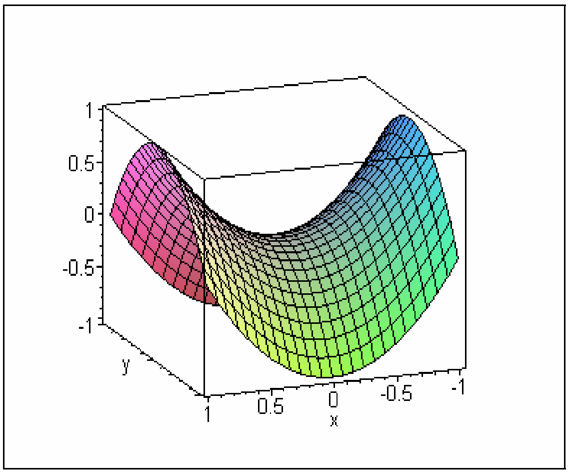
[2] Donde $f_{x_i}(x_1, x_2, ..., x_n) = \dfrac{\partial f(x_1, x_2, ..., x_n)}{\partial x_i} = \lim\limits_{h \to 0} \dfrac{f(x_1, x_2, ...x_i + h, ..., x_n) - f(x_1, x_2, ...x_i, ..., x_n)}{h}$



Mínimo en *(x,y)=(0,0)*



Máximo en *(x,y)=(0,0)*



Punto de silla en *(x,y)=(0,0)*

Para dicha distinción la condición de segundo orden (condición suficiente) hace referencia a la matriz hessiana (matriz de derivadas segundas) evaluada en cada punto crítico:

$$\mathbf{H}_f(x_1^*, x_2^*) = \begin{pmatrix} f_{x_1x_1}(x_1^*, x_2^*) & f_{x_1x_2}(x_1^*, x_2^*) \\ f_{x_2x_1}(x_1^*, x_2^*) & f_{x_2x_2}(x_1^*, x_2^*) \end{pmatrix}$$

Dicha matriz provee toda la información necesaria de *f* en las inmediaciones de cada punto por medio de una aproximación de Taylor de segundo orden en el sentido de que avisa si el entorno del punto es cóncavo o convexo. Algebraicamente, para discernir si cada punto crítico es efectivamente un punto que optimiza a *f* se hace referencia la concepto de menores principales de orden n asociado a una matriz cuadrada de orden m (m >= n). Un menor principal de orden n asociado a una matriz de orden m es el determinante que surge de considerar una submatriz de orden n conformada por las primeras n filas y las primeras n columnas. Con esta definición la condición de segundo orden se puede enunciar en términos de los menores principales de la matriz hessiana evaluada en cada punto crítico. En efecto:

$$
Mínimo \begin{cases} |\mathbf{H}_1| > 0 \\ |\mathbf{H}_2| > 0 \end{cases} \qquad Máximo \begin{cases} |\mathbf{H}_1| < 0 \\ |\mathbf{H}_2| > 0 \end{cases}
$$

Es decir para que efectivamente un punto crítico (que satisface la condición de primer orden) sea un máximo local de *f* es suficiente que los menores principales[3] asociados a la matriz hessiana evaluada en el punto crítico en cuestión alternen en signo empezando por signo negativo. Para el caso de un mínimo se requiere que todos sean positivos.

Caso n variables:

En esta situación el problema se presenta como:

$$
\boxed{\underset{x_1, x_2, \ldots, x_n}{Max} \quad f = f(x_1, x_2, \ldots, x_n)}
$$

donde ahora la condición de primer orden viene dada por es siguiente sistema de n ecuaciones:

$$
\begin{cases} f_{x_1}(x_1, x_2, \ldots, x_n) = 0 \\ f_{x_2}(x_1, x_2, \ldots, x_n) = 0 \\ \vdots \\ f_{x_n}(x_1, x_2, \ldots, x_n) = 0 \end{cases} \qquad \Rightarrow (x_1^*, x_2^*, \ldots, x_n^*)
$$

Donde al igual que en el caso anterior el sistema es no lineal y puede presentar múltiples soluciones. Cada una de ellas será un punto crítico y para decidir si dichos puntos son o no valores óptimos se recurre a analizar el signo de los menores principales de la matriz hessiana evaluada en cada punto critico:

$$
\mathbf{H}_f(x_1^*, x_2^*, \ldots, x_n^*) = \begin{pmatrix} f_{x_1 x_1}(x_1^*, x_2^*, \ldots, x_n^*) & \cdots & f_{x_1 x_n}(x_1^*, x_2^*, \ldots, x_n^*) \\ \vdots & \ddots & \vdots \\ f_{x_n x_1}(x_1^*, x_2^*, \ldots, x_n^*) & \cdots & f_{x_n x_n}(x_1^*, x_2^*, \ldots, x_n^*) \end{pmatrix}
$$

En esta situación, para el caso de mínimo, es suficiente que todos los menores principales sean positivos mientras que para el caso de máximo, se requiere que alternen en signo comenzando por signo negativo. Algebraicamente se tiene:

$$
Mínimo \begin{cases} |\mathbf{H}_1| > 0 \\ |\mathbf{H}_2| > 0 \\ \vdots \\ |\mathbf{H}_n| > 0 \end{cases} \qquad Máximo \begin{cases} |\mathbf{H}_1| < 0 \\ |\mathbf{H}_2| > 0 \\ \vdots \\ (-1)^n |\mathbf{H}_n| > 0 \end{cases}
$$

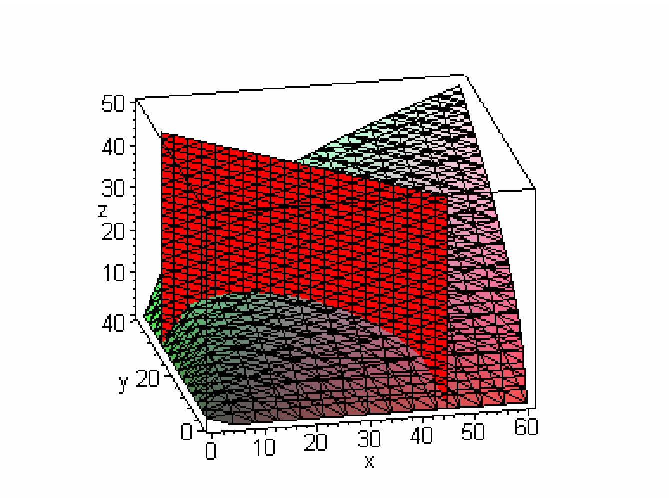
[3] Donde $|\mathbf{H}_n|$ denota al menor principal de orden *n* de la matriz $\mathbf{H}$

Optimización Restricciones de Igualdad

Caso 2 variables y una restricción:
Este tipo de problema consiste en hallar los valores de x_1 y x_2 que perteneciendo a una curva del dominio ($g(x_1, x_2) = m$) confieran a f un valor máximo o mínimo. Formalmente este problema puede escribirse de la siguiente manera:

$$\boxed{\begin{array}{l} \underset{x_1,x_2}{Max}\ f = f(x_1, x_2) \\ st: \quad g(x_1, x_2) = m \end{array}}$$

Geométricamente el problema puede visualizarse en la siguiente gráfica:



En dicha grafica se presenta una restricción del tipo lineal (*g(x,y)* es una relación lineal) sobre el dominio (plano x,y). Para entender el problema se procedió a cortar la gráfica de *f* con un plano vertical (en color rojo) que emerge sobre la curva de restricción (en este caso una recta). De la intersección de dicho plano con la superficie generada por *f* surge una curva que es la que hay que maximizar. Esa curva, que nace de la intersección de *f* con el plano de restricción, constituye los valores de *f* para los puntos que cumplen con la condición *g(x,y)=m*. En consecuencia, el problema consiste en hallar las coordenadas *(x,y)* que posadas sobre dicha curva confieran a *f* un valor máximo (o mínimo).

Para su resolución se hace uso de la función Lagrangeana (*L*) y de los multiplicadores de Lagrange (λ) como sigue:

$$\underset{x_1,x_2}{Max}\, L = f(x_1, x_2) + \lambda[m - g(x_1, x_2)]$$

La condición de primer orden ahora se escribe en términos de la función Lagrangeana y arroja el siguiente sistema de ecuaciones:

$$\begin{cases} L_{x_1}(x_1, x_2, \lambda) = 0 \\ L_{x_2}(x_1, x_2, \lambda) = 0 \Rightarrow (x_1^*, x_2^*, \lambda^*) \\ L_\lambda(x_1, x_2, \lambda) = 0 \end{cases}$$

Al igual que en los otros casos las ecuaciones pueden resultar ser no lineales y con soluciones múltiples. Cada una de dichas soluciones constituye un punto crítico que luego deberá verificar las condiciones de segundo orden. En este caso se recurre a una matriz hessiana de la función Lagrangeana orlada con ceros y evaluada en cada punto crítico. Luego sobre estas matrices se toman los menores principales:

$$\bar{\mathbf{H}}_L(x_1^*, x_2^*, \lambda^*) = \begin{pmatrix} 0 & g_{x_1}(x_1^*, x_2^*, \lambda^*) & g_{x_2}(x_1^*, x_2^*, \lambda^*) \\ g_{x_1}(x_1^*, x_2^*, \lambda^*) & L_{x_1x_1}(x_1^*, x_2^*, \lambda^*) & L_{x_1x_2}(x_1^*, x_2^*, \lambda^*) \\ g_{x_2}(x_1^*, x_2^*, \lambda^*) & L_{x_2x_1}(x_1^*, x_2^*, \lambda^*) & L_{x_2x_2}(x_1^*, x_2^*, \lambda^*) \end{pmatrix}$$

$$\textit{Mínimo: } |\mathbf{H}_3| < 0 \qquad \textit{Máximo: } |\mathbf{H}_3| > 0$$

Como se aprecia los menores principales se toman a partir del orden tres por lo que en definitiva solo hay que atender al signo del determinante de la matriz hessiana orlada evaluada en cada punto crítico.

Caso n variables y m restricciones (m < n)

Similarmente al caso planteado con anterioridad, el problema consiste en hallar un punto n dimensional (es decir los valores de $x_1, x_2, \ldots, x_n$) tales que satisfaciendo un conjunto de restricciones (siempre menor al número de incógnitas) confieran a f un valor máximo o mínimo. Formalmente el problema se traduce a:

$$\boxed{\begin{array}{l} \underset{x_1, x_2, \ldots, x_n}{Max} \quad f = f(x_1, x_2, \ldots, x_n) \\ st: \quad g_1(x_1, x_2, \ldots, x_n) = m_1 \\ \qquad\; g_2(x_1, x_2, \ldots, x_n) = m_2 \\ \qquad\; \vdots \\ \qquad\; g_m(x_1, x_2, \ldots, x_n) = m_m \end{array}}$$

Nuevamente se plantea la función Lagrangeana que ahora se generaliza de la siguiente manera:

$$\underset{x_1, x_2, \ldots, x_n}{Max} \; L = f(x_1, x_2, \ldots, x_n) + \lambda_1[m_1 - g_1(x_1, x_2, \ldots, x_n)] + \lambda_2[m_2 - g_2(x_1, x_2, \ldots, x_n)] + \cdots + \lambda_m[m_m - g_m(x_1, x_2, \ldots, x_n)]$$

Mientras que la condición de primer orden arroja el siguiente sistema de $n + m$ ecuaciones

$$\left\{\begin{array}{l} L_{x_1}(x_1, x_2, ..., x_n, \lambda_1, ...\lambda_m) = 0 \\ L_{x_2}(x_1, x_2, ..., x_n, \lambda_1, ...\lambda_m) = 0 \\ \vdots \\ L_{x_n}(x_1, x_2, ..., x_n, \lambda_1, ...\lambda_m) = 0 \\ L_{\lambda_1}(x_1, x_2, ..., x_n, \lambda_1, ...\lambda_m) = 0 \\ L_{\lambda_2}(x_1, x_2, ..., x_n, \lambda_1, ...\lambda_m) = 0 \\ \vdots \\ L_{\lambda_m}(x_1, x_2, ..., x_n, \lambda_1, ...\lambda_m) = 0 \end{array}\right. \Rightarrow (x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, \lambda_2^*, ..., \lambda_m^*)$$

Las soluciones para dicho sistema constituyen los puntos críticos candidatos a extremos relativos (máximos o mínimos) de la función sujeta a las restricciones. A continuación, se debe recurrir a la matriz hessiana orlada que para el caso de *m* restricciones adopta la siguiente forma en términos matriciales:

$$\bar{\mathbf{H}}_L(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*) = \begin{pmatrix} \mathbf{0}_{m\times m} & \mathbf{g_X}(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*)_{m\,x\,n} \\ \mathbf{g_X}^T(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*)_{m\,x\,n} & \mathbf{H}_L(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*)_{n\,x\,n} \end{pmatrix}$$

donde:

$\mathbf{0}_{m\,x\,m}$ es una matriz cuadrada de ceros de orden *m x m*

$$\mathbf{g_X}(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*)_{m\,x\,n} = \begin{pmatrix} g^1{}_{x_1}(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*) & \cdots & g^1{}_{x_n}(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*) \\ \vdots & \ddots & \vdots \\ g^m{}_{x_1}(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*) & \cdots & g^m{}_{x_n}(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*) \end{pmatrix}$$

$\mathbf{g_X}^T(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ...\lambda_m^*)_{m\,x\,n}$ es la transpuesta de la matriz anterior

Una vez obtenida la matriz hessiana orlada se debe atender al signo de los menores principales a partir del orden *2 m +1*. Para el caso de mínimo se requieren que todos tengan el signo de $(-1)^m$ y para máximo se requiere que los menores alternen en signo comenzando por el signo de $(-1)^{m+1}$.