Matemática: Optimización con Restricciones

Lara Sánchez Peña

Universidad Torcuato di Tella

10 de enero de 2022

Introducción - Geogebra

- En el curso hasta ahora, aprendimos a buscar mínimos y máximos de funciones en una o más variables, incluso para algunos casos sencillos donde se agregan restricciones adicionales.
- La función que queremos maximizar (o minimizar), $f(x,y)$, se llama **función objetivo** y cualquier función adicional que restrinja el problema $g(x,y)=c$ se conoce como **restricción**.
- Ejemplo 1: $\max\limits_{x,y} -x^2-y^2+4$ s.a. $x=1$
- Ejemplo 2: $\max\limits_{x,y} -x^2-y^2+4$ s.a. $y=x-1$

Podemos resolver estos problemas de dos maneras:

1. reemplazando la restricción en la función objetivo
2. usando el método de Lagrange.

Método de Lagrange

El método de Lagrange tiene la ventaja de que no siempre se puede despejar una de las variables respecto de las demás para reemplazarla en la función objetivo que queremos maximizar o minimizar.

Por lo tanto, $\min_{x,y} f(x,y)$ s.a. $g(x,y) = c$ se puede resolver usando el método de Lagrange:

$$\min_{x,y,\lambda} \mathcal{L}(x,y,\lambda) = \min_{x,y,\lambda} f(x,y) + \lambda(g(x,y) - c) \tag{1}$$

Análogamente $\max_{x,y} f(x,y)$ s.a. $g(x,y) = c$, se puede utilizar usando el método de Lagrange:

$$\max_{x,y,\lambda} \mathcal{L}(x,y,\lambda) = \max_{x,y,\lambda} f(x,y) - \lambda(g(x,y) - c) \tag{2}$$

Para entender por qué en un caso λ está acompañado por un $+$ o por un $-$, ver el apéndice.

Método de Lagrange

- Resolvamos las condiciones de primer orden del problema (2):

$$
\begin{aligned}
&(x) \quad & f'_x(x,y) - \lambda g'_x(x,y) = 0 \\
&(y) \quad & f'_y(x,y) - \lambda g'_y(x,y) = 0 \\
&(\lambda) \quad & g(x,y) - c = 0
\end{aligned}
$$

- Estas tres ecuaciones forman un sistema de tres ecuaciones con tres incógnitas. Bajo ciertas condiciones, dados valores de x e y resuelven el problema, el valor de λ deberá ser único.

Método de Lagrange

- Luego, se sigue que para que (x_0, y_0) sea un máximo del problema debe ser cierto que:

$$-\frac{f'_x(x_0, y_0)}{f'_y(x_0, y_0)} = -\frac{g'_x(x_0, y_0)}{g'_y(x_0, y_0)}$$

- La igualada anterior implica que el valor de la derivada implícita de y respecto de x en el punto (x_0, y_0) debe coincidir las funciones f y g.
- Por lo tanto, ∇f tiene la misma dirección y sentido que ∇g en (x_0, y_0), $\nabla f = \lambda \nabla g$.
- Entonces, aplicar el método de Lagrange es equivalente a buscar el (los) punto(s) (x_0, y_0) de la curva de nivel f tangente a la restricción g.

Método de Lagrange

Teorema (Lagrange)

Sean $f(x,y)$ y $g(x,y)$ dos funciones con derivadas parciales continuas y sea (x_0,y_0) un punto extremo del problema de optimización de f sujeto a $g(x,y)=c$. Si $g'_x(x,y)$ y $g'_y(x,y)$ no son simultáneamente 0, entonces existe un valor de λ tal que (x_0,y_0) sea la solución del Lagrangiano.

Intepretación de λ : En un problema de maximización, como en (2), sabemos que λ satisface que $\nabla f(x,y)=\lambda\nabla g(x,y)$. Intuitivamente, cuanto mayor sea el valor de λ "mayor" es la restricción impuesta por g, es decir, si hubiesemos podido mover la restricción $g(x,y)=c$ un poco en la dirección donde la función crece, el valor máximo que alcanza la función $f(x,y)$ hubiera sido mayor.

Método de Lagrange - Condiciones de Segundo Orden

- ¿Cómo sabemos si un punto (x_0, y_0) obtenido es un máximo o un mínimo? Chequeamos las condiciones de segundo orden para la función $\mathcal{L}$.
- Para saber si nos encontramos en presencia de un máximo o un mínimo simplemente debemos analizar al Hessiano del Lagrangiano $\mathcal{L}(x, y, \lambda)$, una matriz de 3×3.
- Como este Hessiano está construido en base a una función con una restricción u *orla*, se lo llama Hessiano **orlado** $(\overline{H})$. Específicamente, si (x_0, y_0, λ_0) es uno de los puntos críticos de $\mathcal{L}$, entonces:

$$\overline{H}(x_0, y_0, \lambda_0) = \begin{pmatrix} \mathcal{L}''_{xx}(x_0, y_0, \lambda_0) & \mathcal{L}''_{xy}(x_0, y_0, \lambda_0) & \mathcal{L}''_{x\lambda_0}(x_0, y_0, \lambda_0) \\ \mathcal{L}''_{yx}(x_0, y_0, \lambda_0) & \mathcal{L}''_{yy}(x_0, y_0, \lambda_0) & \mathcal{L}''_{y\lambda_0}(x_0, y_0, \lambda_0) \\ \mathcal{L}''_{\lambda_0 x}(x_0, y_0, \lambda_0) & \mathcal{L}''_{\lambda_0 y}(x_0, y_0, \lambda_0) & \mathcal{L}''_{\lambda_0\lambda_0}(x_0, y_0, \lambda_0) \end{pmatrix}$$

Método de Lagrange - Condiciones de Segundo Orden

- Si los **menores principales** del Hessiano Orlado **son todos positivos**, estaremos en presencia de un **mínimo**.
- Si los **menores principales** se **alternan en signo comenzando por un signo negativo**, entonces estaremos en presencia de un **máximo**.
- En la práctica, en economía, se suelen utilizar funciones cóncavas cuando se resuelven problemas de maximización y funciones convexas cuando se resuelven problemas de minimización. De este modo, no solemos chequear las condiciones de segundo orden (aunque deberíamos).

Método de Lagrange

- ¿Existe alguna forma de garantizar que un lagrangiano tiene solución? Bajo ciertas condiciones, sí.

Teorema (Weierstraß)

Sea $f(x, y)$ una función continua definida sobre un conjunto cerrado y acotado D. Entonces dicha función tiene al menos un máximo y un mínimo en D. Si además dicha función es estrictamente

- cóncava, el máximo es único.
- convexa, el mínimo es único.

- Este resultado implica que en la mayor parte de problemas económicos habrá una solución. Por ejemplo, para el problema de un individuo que elige comprar bienes con tal de maximizar su utilidad $u(x, y)$, como el **conjunto de consumo** es cerrado y acotado, sabremos que habrá una única solución a ese problema si la función $u(x, y)$ es una función estrictamente cóncava.

Ejercicios aplicados

Resuelva usando el método del Lagrangiano:

- $\max\limits_{x,y} u(x,y)$ s.a. $p_x x + p_y y = m$, donde $u(x,y) = xy$.
- $\min\limits_{L,K} wL + rK$ s.a. $f(L,K) = 36$, donde $f(L,K) = K^{0,5} L^{0,25}$.

Karesh-Kuhn-Tucker

Pasemos ahora a resolver problemas de optimización con **restricciones que no sean de igualdad**. Es decir, por ejemplo,

- $\min_{x,y} f(x,y)$ s.a. $g(x,y) \leq c$
- $\max_{x,y} f(x,y)$ s.a. $g(x,y) \leq c$

En este caso, el método de KKT, encuentra los puntos (x_0, y_0) que resuelvan dichos problemas teniendo en cuenta que:

1. la restricción **está operativa** $g(x,y) = c$
2. la restricción **no está operativa** $g(x,y) < c$

Karesh-Kuhn-Tucker

El método KKT utiliza la función $\mathcal{L}(x, y, \lambda)$, pero agrega una condición extra:

$$\min_{x,y,\lambda} \mathcal{L}(x, y, \lambda) = \min_{x,y,\lambda} f(x, y) + \lambda(g(x, y) - c) \tag{3}$$

$$\max_{x,y,\lambda} \mathcal{L}(x, y, \lambda) = \max_{x,y,\lambda} f(x, y) - \lambda(g(x, y) - c) \tag{4}$$

- Resolvamos las condiciones de primer orden del problema (4):

$$\begin{array}{ll} (x) & f'_x(x, y) - \lambda g'_x(x, y) = 0 \\ (y) & f'_y(x, y) - \lambda g'_y(x, y) = 0 \\ (\lambda) & \lambda \cdot (g(x, y) - c) = 0, \lambda \geq 0 \end{array}$$

- La tercera condición incorpora que la restricción puede o no ser operativa. Si $\lambda = 0$, la restricción **no es operativa** (hubiera sido equivalente a resolver un problema **sin restricciones**). Si $\lambda > 0$ la restricción es **operativa**, $g(x, y) = c$ (hubiera sido equivalente a usar el método de **Lagrange**).

Existencia de solución y CSO

- Para saber si existe al menos algún punto (x_0, y_0) que resuelva algún problema que se resuelve con KKT, necesitamos **condiciones necesarias**.
- Para garantizar que existe al menos algún punto (x_0, y_0) que resuelva algún problema que se resuelve con KKT, necesitamos **condiciones suficientes**.
- Para saber si la solución es única, es suficiente con pedir que la función objetivo $f(x, y)$ sea **estrictamente convexa para encontrar un mínimo** (estrictamente cóncava para encontrar un máximo)
- Al escribir restricciones, solemos pedir que:
  - si la restricción $g(x, y) \leq c$, $g(x, y)$ sea una función convexa.
  - si la restricción $g(x, y) \geq c$, $g(x, y)$ sea una función cóncava.
- Esto garantiza el conjunto de puntos (x, y) sobre el cual se busca un mínimo (máximo) sea un **conjunto convexo**. Eso quiere decir, que si dos puntos (a, b) y (c, d) están en el dominio del problema, entonces todo el segmento que une a esos puntos también lo estará.

Multiplicador de Lagrange

- Interpretemos nuevamente al multiplicador de Lagrange, λ para este tipo de problemas.
- Escribimos el Lagrangiano

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda (g(x, y) - c)$$

- Notemos que resulta inmediato que $\mathcal{L}'_c(\cdot) = \lambda$.
- Por el **teorema de la envolvente** vale que, en el óptimo vale que:

$$\frac{\partial f(x^*(c), y^*(c))}{\partial c} = \lambda$$

- Es decir, el multiplicador de Lagrange se puede interpretar como **el efecto de relajar marginalmente la restricción**, el valor máximo que puede tomar $g(x, y)$.

Aplicaciones - Repaso

- Un individuo debe elegir qué cantidad de dos bienes x e y consumir de tal modo de maximizar su "felicidad", bajo la restricción que dice que **el gasto en los bienes x e y no puede superar al ingreso** $p_x x + p_y y \leq m$, donde p_x es el precio del bien x, p_y el precio del bien y y m el ingreso disponible.

$$\max_{x,y} u(x,y) \text{ sujeto a } p_x x + p_y y \leq m$$

- Notar que la restricción presupuestaria es una función lineal en (x, y). Por lo tanto, es una función convexa.
- Una firma debe elegir cuánto trabajo L y cuanto capital K contratar para **producir por lo menos una cantidad** Q de bienes al menor costo posible, teniendo en cuenta que su producción está gobernada por la función $Q = f(L, K)$

$$\min_{L,K} wL + rK \text{ sujeto a } f(L,K) \geq Q$$

- Se suele pedir en economía que la función de producción $f(L, K)$ sea cóncava en (L, K).

Apéndice 1

- ¿Por qué se pone $+\lambda$ o $-\lambda$ en el lagrangiano?
- Si queremos buscar un mínímo (máximo) de $f(x, y)$ restringido a que $g(x, y) = 0$, queremos que f decrezca (crezca) a lo largo de g en la dirección de mínimo (máximo) crecimiento de f.
- Recordemos que $-\nabla f$ $(-\nabla f)$ es la dirección de mínimo (máximo) crecimiento de f. Demostración en la segunda parte del apéndice.
- Por lo tanto, queremos que

$$\begin{cases} \text{(mín)}: -\nabla f & = -\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \lambda \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) = \lambda \nabla g \\ \text{(máx)}: \nabla f & = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \lambda \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) = \lambda \nabla g \end{cases}$$

Apéndice 1

- Es decir que

$$\begin{cases} (\text{mín})\text{:}\nabla f + \lambda \nabla g = 0 \\ (\text{máx})\text{:}\nabla f - \lambda \nabla g = 0 \end{cases}$$

- Por lo tanto, cuando queremos
  - minimizar una función $f(x, y)$ sujeto a una restricción $g(x, y) = c$

$$\min_{x,y} f + \lambda g$$

  - maximizar una función $f(x, y)$ sujeto a una restricción $g(x, y) = c$

$$\max_{x,y} f - \lambda g$$

Apéndice 2

Dada una función $f(x, y)$ el crecimiento de una función f en una dirección (a, b) se mide a partir de la derivada direccional de f en la dirección (a, b) en un punto (x_0, y_0).

$$
\begin{aligned}
\frac{\partial f}{\partial \overrightarrow{(a, b)}} &= \lim_{(x,y)\to(x_0,y_0) \text{ dir. } (a,b)} \frac{F(x, y) - F(x_0, y_0)}{\|(x, y) - (x_0, y_0))} \\
&= \lim_{h\to 0} \frac{F(x_0 + ah, y_0 + bh) - F(x_0, y_0)}{h\|(a, b)\|} \\
&\underbrace{=}_{\text{Reg. de la cadena}} \frac{1}{\|(a, b)\|} \left(\frac{\partial f}{\partial x}(x_0, y_0) \cdot a + \frac{\partial f}{\partial y}(x_0, y_0) \cdot b \right)
\end{aligned}
$$

donde $\|(a, b)\| = \sqrt{a^2 + b^2}$.

Apéndice 2

Por lo tanto, podemos escribir al crecimiento de f en la dirección (a, b) en el punto (x_0, y_0), $\dfrac{\partial f}{\partial \overrightarrow{(a, b)}}(x_0, y_0)$, como:

$$\frac{\partial f}{\partial \overrightarrow{(a, b)}}(x_0, y_0) = \nabla^{\mathsf{T}} F(x_0, y_0) \cdot \frac{(a, b)}{\|(a, b)\|}$$

$$= \left\langle \frac{\nabla^{\mathsf{T}} F(x_0, y_0)}{\|\nabla^{\mathsf{T}} F(x_0, y_0)\|}, \frac{(a, b)}{\|(a, b)\|} \right\rangle \cdot \|\nabla^{\mathsf{T}} F(x_0, y_0)\|$$

Usando que $\left\langle \dfrac{\nabla^{\mathsf{T}} F(x_0, y_0)}{\|\nabla^{\mathsf{T}} F(x_0, y_0)\|}, \dfrac{(a, b)}{\|(a, b)\|} \right\rangle = cos(\theta)$, donde θ es el angúlo formado entre los vectores $\nabla^{\mathsf{T}} F(x_0, y_0)$ y (a, b).

Apéndice 2

Finalmente,

$$\frac{\partial f}{\partial \overrightarrow{(a,b)}}(x_0, y_0) = cos(\theta) \cdot \|\nabla^{\mathsf{T}} F(x_0, y_0)\|$$

La última expresión se

- maximiza cuando $cos(\theta) = 1$, cuando los vectores $\nabla^{\mathsf{T}} F(x_0, y_0)$ y (a, b) forman un ángulo de 0 grados. O sea, si (a, b) tiene la misma dirección y sentido que $\nabla^{\mathsf{T}} F(x_0, y_0)$.
- maximiza cuando $cos(\theta) = -1$, cuando los vectores $\nabla^{\mathsf{T}} F(x_0, y_0)$ y (a, b) forman un ángulo de 180 grados. O sea, si (a, b) tiene la misma dirección y sentido opuesto a $\nabla^{\mathsf{T}} F(x_0, y_0)$. Es decir cuando (a, b) tiene la misma dirección y sentido que $-\nabla^{\mathsf{T}} F(x_0, y_0)$.

Apéndice 3 - Teorema de la envolvente

El teorema de la envolvente dice que si la **función de valor** depende de parámetros de manera directa e indirecta, para los **valores óptimos** que maximizan (minimizan) **la función de valor solamente importan los efectos directos para ver cómo cambia la función de valor**. La demostración de este teorema es usar regla de la cadena.

1. $\max_{K,L} \pi(L, K) = \max_{K,L} pf(L, K) - wL - rK$, donde L, K son variables y p, w, r son parámetros.
2. $\min_{K,L} C(K, L) = \min_{K,L} wL + rK$ s.a. $f(L, K) = y$, donde L, K son variables y w, r son parámetros.

- En el problema 1, encontramos

$$pf(K(p,w,r),L(w,r,p)) - wL(w,r,p) - rK(p,w,r)$$

- Hay que usar la regla de la cadena y las CPO para mostrar que
  - $\frac{\partial \pi^*}{\partial p}(w,r,p) = f(L^*,K^*)$ **(Lema de Hotelling)**
  - $\frac{\partial \pi^*}{\partial w}(w,r,p) = -L^*$
  - $\frac{\partial \pi^*}{\partial r}(w,r,p) = -K^*$
- En el problema 2, encontramos

$$wL(w,r) + rK(w,r)$$

- Hay que usar la regla de la cadena y las CPO para mostrar que
  - $\frac{\partial C(K,L)^*}{\partial w}(w,r) = L^*$ **(Lema de Shephard)**
  - $\frac{\partial C(K,L)^*}{\partial r}(w,r,p) = K^*$ **(Lema de Shephard)**