Cálculo Vectorial
Lectura 4: Multiplicadores de Lagrange

Ing. Aldo Jiménez Arteaga

Enero de 2020

1. Optimización de Funciones

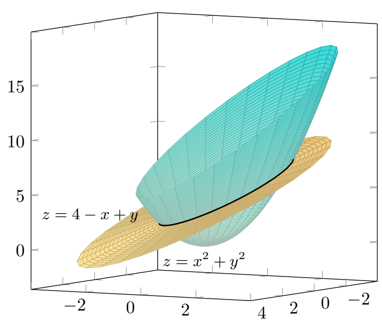

Figura 1. La curva de intersección entre las superficies contiene dos puntos extremos: uno minimiza la distancia al origen y el otro la maximiza.

Una vez que se ha estudiado la manera de analizar los extremos de funciones multivariable hay que preguntarse lo siguiente: ¿qué sucede si se necesita encontrar un extremo sujeto a ciertas condiciones? Esta pregunta plantea el análisis de una función que tiene extremos pero que está condicionada a una o varias restricciones.

Por ejemplo, en la figura 1 se muestran un paraboloide y un plano cuyas ecuaciones cartesianas son

$$z = x^2 + y^2 \tag{1}$$
$$z = 4 - x + y \tag{2}$$

Al intersecarse se obtiene una curva, donde cada uno de sus puntos se encuentra a una distancia determinada del origen. De entre todas las distancias mencionadas existen dos que son interesantes: la máxima y la mínima distancia. Este problema involucra el cálculo de máximos y mínimos condicionados.

Para resolver los extremos condicionados se debe identificar las funciones involucradas; existen dos tipos de estas funciones: la función objetivo, y la(s) función(es) restricción. La función objetivo es aquélla que contiene los puntos extremos condicionados; es la llamada función a optimizar. Las funciones restricción son aquéllas que limitarán a la función objetivo para determinar los extremos. En otras palabras se busca optimizar la función objetivo a partir de las restricciones. Para el caso de estudio la función objetivo es

$$d\,(x, y, z) = x^2 + y^2 + z^2 \tag{3}$$

y las restricciones son las superficies (1) y (2) reescritas en la forma $g\,(\mathbf{x}) = 0$:

$$g_1\,(x, y, z) = x^2 + y^2 - z \tag{4}$$
$$g_2\,(x, y, z) = 4 - x + y - z \tag{5}$$

Una vez establecidos los dos tipos de funciones se requieren dos conceptos importantes para optimizar el objetivo: el paralelismo entre vectores mediante combinación lineal, y la relación entre las curvas de nivel y el gradiente de una superficie.

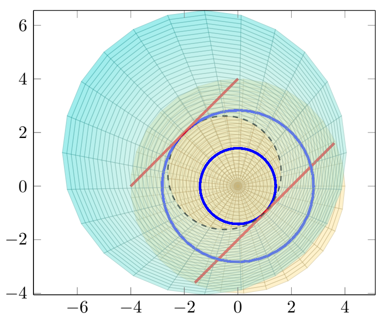

Figura 2. Las curvas de nivel del paraboloide (en azul) y del plano (en rojo) son tangentes en los lugares donde existen extremos condicionados.

Para obtener la optimización se determinará el lugar geométrico en el cual las restricciones y el objetivo coinciden. Realizar este trabajo puede ser largo si se busca la intersección geométrica. Sin embargo, el proceso se vuelve eficiente si se analizan las curvas de nivel de las funciones involucradas, ya que se puede analizar los puntos de tangencia entre dichas curvas; es decir, el lugar donde las restricciones y el objetivo coinciden es donde sus respectivas curvas de nivel son tangentes. La figura 2 muestra las curvas de nivel del paraboloide y del plano que coinciden con la curva de intersección.

1.1. Ecuación de Lagrange

Una vez que se sabe que los extremos condicionados están relacionados con la tangencia de curvas de nivel, se requiere de una herramienta que indique en qué momento exista dicha tangencia. En este punto entra el gradiente en un punto P, ya que su significado geométrico es un vector ortogonal a una curva de nivel que pasa por dicho punto.

Considerando la función objetivo (3) y las restricciones (4) y (5), sus respectivas curvas de nivel son tangentes en algún punto y por lo tanto sus respectivos gradientes son paralelos entre sí.

En otras palabras uno de los gradientes es combinación lineal de los otros; es decir,

$$\nabla d = \alpha_1 \nabla g_1 + \alpha_2 \nabla g_2 \qquad (6)$$

La figura 3 muestra la geometría de los gradientes y las respectivas curvas de nivel.

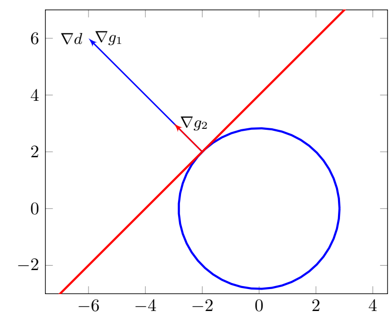

Figura 3. Relación geométrica de los gradientes de las funciones (3), (4) y (5). g_1 y d poseen la misma curva de nivel y el mismo gradiente.

En conjunto con las restricciones g_1 y g_2, la ecuación (6) permite resolver el problema de los extremos condicionados.

La combinación lineal en la ecuación (6) introduce los escalares α_1 y α_2 asociados a los gradientes de las restricciones. Dichos escalares se llaman multiplicadores de Lagrange.

Sea $f(\mathbf{x})$ una función cuyos estremos están condicionados por el conjunto de restricciones $G = \{g_1(\mathbf{x}), g_2(\mathbf{x}), \cdots, g_n(\mathbf{x})\}$. Los puntos extremos de f se obtienen al resolver la ecuación de Lagrange:

$$\nabla \left(f(\mathbf{x}) - \sum_{i=1}^{n} \alpha_i g_i(\mathbf{x}) \right) = 0$$

Los escalares α_i son los multiplicadores de Lagrange y la función

$$\mathcal{L}(\mathbf{x}, \alpha_1, \cdots, \alpha_n) = f(\mathbf{x}) - \sum_{i=1}^{n} \alpha_i g_i(\mathbf{x})$$

es la función de Lagrange.

Hay que aclarar que al plantear la ecuación de Lagrange los multiplicadores son variables, lo cual indica que son susceptibles de derivarlos. Partiendo del gradiente de la función de Lagrange

$$\begin{aligned}\nabla\mathcal{L}\left(\mathbf{x},\alpha_1,\cdots,\alpha_n\right)&=0\\ \nabla\left(f\left(\mathbf{x}\right)-\sum_{i=1}^{n}\alpha_i g_i\left(\mathbf{x}\right)\right)&=\end{aligned}$$

se debe aplicar la derivada parcial respecto de cada variable involucrada: $\mathbf{x}=\begin{bmatrix}x_1\\ \vdots\\ x_m\end{bmatrix}$, $\alpha_1,\alpha_2,\cdots,\alpha_n$. Por lo tanto, cada derivada parcial de $\mathcal{L}$ se lista a continuación:

$$\begin{aligned}\nabla f\left(\mathbf{x}\right)-\sum_{i=1}^{n}\alpha_i\nabla g_i\left(\mathbf{x}\right)&=\begin{cases}\mathcal{L}_{x_1}=f_{x_1}-\sum_{i=1}^{n}\alpha_i\dfrac{\partial g_i}{\partial x_1}\\ \vdots\\ \mathcal{L}_{x_m}=f_{x_m}-\sum_{i=1}^{n}\alpha_i\dfrac{\partial g_i}{\partial x_m}\end{cases}\\ \mathcal{L}_{\alpha_1}&=g_1\left(\mathbf{x}\right)\\ &\vdots\\ \mathcal{L}_{\alpha_n}&=g_n\left(\mathbf{x}\right)\end{aligned}$$

Así, las ecuaciones a resolver en un problema de extremos condicionados son

$$\begin{aligned}\nabla f\left(\mathbf{x}\right)-\sum_{i=1}^{n}\alpha_i\nabla g_i\left(\mathbf{x}\right)&=0\\ g_1\left(\mathbf{x}\right)&=0\\ &\vdots\\ g_n\left(\mathbf{x}\right)&=0\end{aligned}$$

Ejemplo. Obtenga los puntos extremos de la función $d\left(x,y,z\right)=x^2+y^2+z^2$, restringidos por las condiciones $z=x^2+y^2$ y $z=4-x+y$.

Para comenzar las restricciones se reescriben como $g_1=x^2+y^2-z$ y $g_2=4-x+y-z$. Planteando la ecuación $\nabla d-\alpha_1\nabla g_1-\alpha_2\nabla g_2=0$, componente a componente se obtiene

$$\begin{aligned}2x-\alpha_1\left(2x\right)-\alpha_2\left(-1\right)&=0 &&(7)\\ 2y-\alpha_1\left(2y\right)-\alpha_2\left(1\right)&=0 &&(8)\\ 2z-\alpha_1\left(-1\right)-\alpha_2\left(-1\right)&=0 &&(9)\end{aligned}$$

Las restricciones originales son

$$\begin{aligned}x^2+y^2-z&=0 &&(10)\\ 4-x+y-z&=0 &&(11)\end{aligned}$$

Para resolver el sistema de las ecuaciones (7) a (11), primero se suman término a término las ecuaciones (7) y (8).

$$\begin{aligned}\left(2x-2x\alpha_1+\alpha_2\right)+\left(2y-2y\alpha_1-\alpha_2\right)&=0\\ 2x\left(1-\alpha_1\right)+2y\left(1-\alpha_1\right)&=\\ 2\left(x+y\right)\left(1-\alpha_1\right)&= &&(12)\end{aligned}$$

La ecuación (12) se satisfface con $x=-y$ ó $\alpha_1=1$, sin embargo con ésta última solución se obtendrían valores complejos, por lo cual se descarta. Sustituyendo $x=-y$ en las ecuación (10),

$$\begin{aligned}x^2+y^2-z&=0\\ \left(-y\right)^2+y^2-z&=\\ 2y^2-z&=0 &&(13)\end{aligned}$$

Y ahora, sustituyendo $x=-y$ en la ecuación 11,

$$
\begin{aligned}
4 - x + y - z &= 0 \\
4 - (-y) + y - z &= \\
4 + 2y - z &= 0 \qquad (14)
\end{aligned}
$$

Al igualar las ecuaciones (13) y (14),

$$
\begin{aligned}
2y^2 - z &= 4 + 2y - z \\
2y^2 - 2y - 4 &= 0 \\
y^2 - y - 2 &= \\
(y - 2)(y + 1) &= 0
\end{aligned}
$$

se obtienen los valores $y = -1$ y $y = 2$, que al sustituirlos en x se obtiene $x = 1$ y $x = -2$, respectivamente. Finalmente, para el valor de z se utiliza la ecuación (14).

$$
\begin{aligned}
4 + 2y - z &= 0 \\
4 + 2(-1) - z &= 0 \\
4 - 2 - z &= 0 \\
z &= 2
\end{aligned}
$$

$$
\begin{aligned}
4 + 2y - z &= 0 \\
4 + 2(2) - z &= 0 \\
4 + 4 - z &= 0 \\
z &= 8
\end{aligned}
$$

Para conocer el máximo y el mínimo se evalúan los valores calculados en la función objetivo:

$$
d(1, -1, 2) = 6, \qquad d(-2, 2, 8) = 72
$$

Hay un mínimo en $P(1, -1, 2, 6)$ y un máximo en $Q(-2, 2, 8, 72)$.