Última actualización: 16 de junio de 2022

Clase 43

La clase pasada conjeturamos que si deseamos hallar los valores máximo y mínimo de una función f sobre un conjunto de nivel de otra función, digamos $N_c(g)$ era suficiente con encontrar los puntos $\overline{x} \in N_c(g)$ en los cuales $\nabla f(\overline{x})$ y $\nabla g(\overline{x})$ son paralelos. Pusimos a prueba esta conjetura hallando la distancia de un punto fijo a un plano dado y funcionó.

Multiplicadores de Lagrange

Veamos, con un ejemplo, cómo podríamos generalizar nuestra conjetura.

Considere una recta $\mathscr{L}$ determinada por la intersección de los planos P_1 y P_2 que tienen como ecuaciones

$$A_1x + B_1y + C_1z + D_1 = 0 \qquad \text{y} \qquad A_2x + B_2y + C_2z + D_2 = 0,$$

respectivamente, con (A_1, B_1, C_1) y (A_2, B_2, C_2) vectores no cero y no paralelos, y $\overline{x}_0 \in \mathbb{R}^3$. Calcule la distancia de $\overline{x}_0$ a $\mathscr{L}$. Igual que en el ejemplo mencionado al inicio de este documento, hallaremos el punto $\overline{x}_{\mathscr{L}} \in \mathscr{L}$ que minimiza la distancia al cuadrado. Entonces, debemos hallar el punto $\overline{x}_{\mathscr{L}}$ del conjunto

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid g_1(x, y, z) = 0,\ g_2(x, y, z) = 0\}$$

tal que la función

$$f(x, y, z) = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$$

alcaza su valor mínimo sobre S, donde g_1 y g_2 están dadas por

$$g_1(x, y, z) = A_1x + B_1y + C_1z + D_1 \qquad \text{y} \qquad g_2(x, y, z) = A_2x + B_2y + C_2z + D_2.$$

Note que si $\overline{x}_{\mathscr{L}}$ es el punto de la recta $\mathscr{L}$ que está más cerca a $\overline{x}_0$, entonces $\overline{x}_{\mathscr{L}}$ es el único punto de $\mathscr{L}$ que pertenece al conjunto de nivel $c = \|\overline{x}_{\mathscr{L}} - \overline{x}_0\|^2$ de f (de hecho $N_c(f)$ es la esfera de radio $\|\overline{x}_{\mathscr{L}} - \overline{x}_0\|$ con centro en $\overline{x}_0$). Así, el vector $\overline{v} = \overline{x}_0 - \overline{x}_{\mathscr{L}}$ debe ser perpendicular a la recta $\mathscr{L}$, pues esta es tangente a $N_c(f)$. Por otro lado, como (A_1, B_1, C_1) y (A_2, B_2, C_2) son perpendiculares a los planos P_1 y P_2 respectivamente y $\mathscr{L}$ es la intersección de dichos plano, entonces (A_1, B_1, C_1) y (A_2, B_2, C_2) son perpendiculares a $\mathscr{L}$.

Tenemos entonces tres vectores distintos y perpendiculares a $\mathscr{L}$, además (A_1, B_1, C_1) y (A_2, B_2, C_2) son linealmente independientes, entonces $\overline{v}$ se puede escribir como una combinación lineal de (A_1, B_1, C_1) y (A_2, B_2, C_2).

Note también que $\nabla g_1(\overline{x}) = (A_1, B_1, C_1)$ y $\nabla g_2(\overline{x}) = (A_2, B_2, C_2)$ para todo $\overline{x} \in \mathbb{R}^3$, en particular $\nabla g_1(\overline{x}_{\mathscr{L}}) = (A_1, B_1, C_1)$ y $\nabla g_2(\overline{x}_{\mathscr{L}}) = (A_2, B_2, C_2)$, y que los vectores $\nabla f(\overline{x}_{\mathscr{L}})$ y $\overline{v}$ son paralelos, pues ambos son normales a $N_c(f)$.

En conclusión, $\nabla f(\overline{x}_{\mathscr{L}})$ se debe poder expresar como combinación lineal de $\nabla g_1(\overline{x}_{\mathscr{L}})$ y $\nabla g_2(\overline{x}_{\mathscr{L}})$, es decir, deben existir $\lambda_1, \lambda_2 \in \mathbb{R}$ tales que

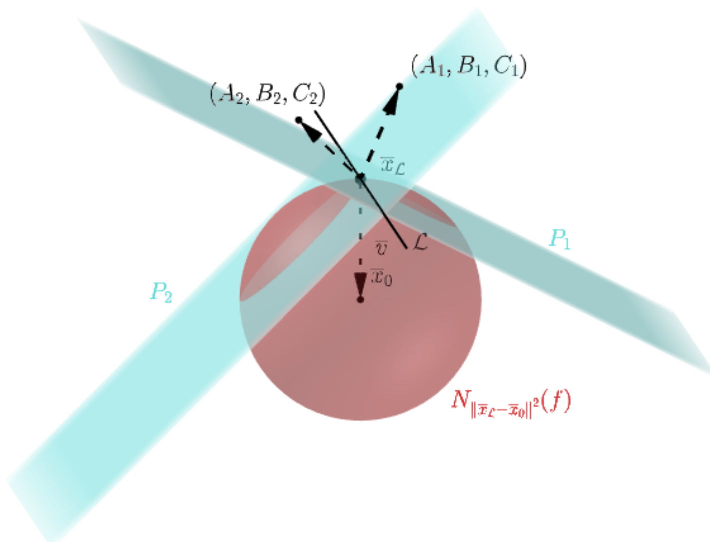


Figura 1: Dado que $\mathscr{L}$ es tangente a la esfera con centro en $\overline{x}_0$, los vectores $\overline{v}$, (A_1, B_1, C_1) y (A_2, B_2, C_2) son perpendiculares a $\mathscr{L}$, entonces pertenecen a un mismo plano.

$$\nabla f(\overline{x}_{\mathscr{L}}) = \lambda_1 \nabla g_1(\overline{x}_{\mathscr{L}}) + \lambda_2 \nabla g_2(\overline{x}_{\mathscr{L}}).$$

Las deducciones que hemos hecho hasta ahora son parte de algo más general, enunciado en el siguiente teorema.

Teorema 1 (de los multiplicadores de Lagrange) *Sean $U \subseteq \mathbb{R}^n$ y $f, g_1, ..., g_m : U \longrightarrow \mathbb{R}$, con $m < n$, tales que $f, g_1, ..., g_m \in C^1(U)$, $S = \{\overline{x} \in U \subseteq \mathbb{R}^n \mid g_1(\overline{x}) = 0, ..., g_m(\overline{x}) = 0\}$ y $\overline{x}_0 \in S$ tal que $\nabla g_1(\overline{x}_0), ..., \nabla g_m(\overline{x}_0)$ son linealmente independientes. Si f tiene un máximo o mínimo (global o solo local) en $\overline{x}_0$ sobre S, entonces existen $\lambda_1, ..., \lambda_m \in \mathbb{R}$ tales que*

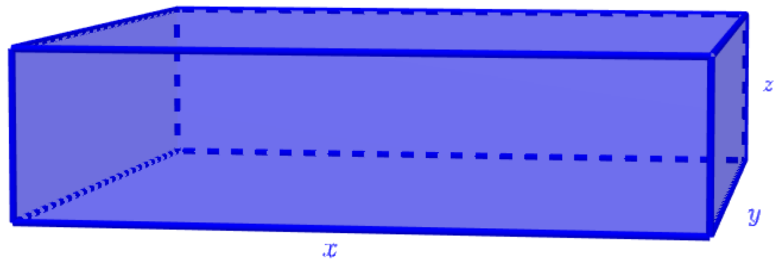
$$\nabla f(\overline{x}_0) = \lambda_1 \nabla g_1(\overline{x}_0) + \cdots + \lambda_m \nabla g_m(\overline{x}_0).$$

Ejemplo 2 *Una compañía planea fabricar cajas de cartón rectangulares cerradas de manera que el área de la superficie de cada una de ellas sea de $120cm^2$. Encuentre las dimensiones de la caja de forma que el volumen sea máximo.*

Solución. Representemos con las variables x, y y z las dimensiones de la caja. Se tiene que el volumen de la caja está dado por el producto xyz, donde todos ellos son positivos, pues queremos una caja de verdad. Consideremos entonces la función volumen $V : (0, \infty)^3 \longrightarrow \mathbb{R}$ dada por

$$V(x, y, z) = xyz.$$

Ahora, note que el área de la superficie está dada por la expresión $2(xy + yz + zx)$ y que esta debe ser igual a $120cm^2$, es decir, $2(xy + yz + zx) = 120$ o de manera equivalente



$$xy + yz + zx = 60.$$

De esta manera consideraremos la función $g : (0, \infty)^3 \longrightarrow \mathbb{R}$ dada por

$$g(x, y, z) = xy + yz + zx - 60.$$

Observe que nuestro problema se trata ahora de hallar el valor máximo de la función V sobre el conjunto

$$S = \{(x, y, z) \in (0, \infty)^3 \mid g(x, y, z) = 0\}.$$

Así que, por el Teorema de Multiplicadores de Lagrange, basta con hallar los puntos $(x, y, z) \in S$ tales que $\nabla V(x, y, z) = \lambda \nabla g(x, y, z)$ para algún $\lambda \in \mathbb{R}$. Como

$$\nabla V(x, y, z) = (yz, xz, xy) \qquad \text{y} \qquad \nabla g(x, y, z) = (y + z, x + z, y + x),$$

entonces debemos resolver el siguiente sistema de ecuaciones

$$\begin{aligned} yz &= \lambda(y + z) && (1)\\ xz &= \lambda(x + z) && (2)\\ xy &= \lambda(y + x) && (3)\\ 0 &= xy + yz + zx - 60. && (4) \end{aligned}$$

Multiplicando las primeras tres ecuaciones por x, y y z, en ese orden, tenemos el siguiente sistema de ecuaciones equivalente al anterior

$$\begin{aligned} xyz &= \lambda x(y + z) && (5)\\ xyz &= \lambda y(x + z) && (6)\\ xyz &= \lambda z(y + x) && (7)\\ 0 &= xy + yz + zx - 60. && (8) \end{aligned}$$

Igualando (5) y (6) y simplificando, tenemos que $\lambda xz = \lambda yz$ o bien que $\lambda(xz - yz) = 0$. Esta igualdad se da si $\lambda = 0$ o $xz = yz$, pero si $\lambda = 0$, entonces de (1) obtenemos que $y = 0$ o $z = 0$, lo cual no es posible. Así, tenemos que $xz = yz$ y como $z \neq 0$, entonces $x = y$. De manera similar, igualando (6) y (7) obtenemos que $y = z$. Por lo tanto $x = y = z$ y sustituyendo y simplificando esto último en (8), tenemos que $3x^2 = 60$, de donde $x = 2\sqrt{5}$. Por lo tanto, las dimensiones de la caja son $x = y = z = 2\sqrt{5}cm^2$, es decir, se trata de un cubo. ∎

En el ejemplo anterior obtuvimos solo una terna (x, y, z) que satisface la conclusión del Teorema de multiplicadores de Lagrange, pero ¿por qué el volumen es máximo en ese punto y no mínimo? Efectivamente, el volumen máximo se alcanza en $(2\sqrt{5}, 2\sqrt{5}, 2\sqrt{5})$ ¿puede justificarlo?
Continuemos con el siguiente ejemplo.

Ejemplo 3 *Demuestre, usando multiplicadores de Lagrange, la desigualdad de Cauchy-Schwarz: Para cualesquiera $\overline{x}, \overline{y} \in \mathbb{R}^n$ se cumple que*

$$|\overline{x} \cdot \overline{y}| \leq \|\overline{x}\|\|\overline{y}\|.$$

Solución. Primero, note que si $\overline{x} = \overline{0}$ o $\overline{y} = \overline{0}$ la desigualdad se satisface. Así que supongamos que $\overline{x} \neq \overline{0}$ y que $\overline{y} \neq \overline{0}$. Entonces la desigualdad deseada es equivalente a demostrar que

$$\left| \frac{\overline{x}}{\|\overline{x}\|} \cdot \frac{\overline{y}}{\|\overline{y}\|} \right| \leq 1$$

y esto a su vez es equivalente a demostrar que para cualesquiera $\overline{x}, \overline{y} \in \mathbb{R}^n$ tales que $\|\overline{x}\| = \|\overline{y}\| = 1$ se cumple que

$$|\overline{x} \cdot \overline{y}| \leq 1.$$

Y será esta última afirmación la que demostraremos. Para ello, consideremos las funciones $f, g_1, g_2 : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ dadas por

$$\begin{aligned} f(x_1, ..., x_n, y_1, ..., y_n) &= x_1y_1 + \cdots x_ny_n \\ g_1(x_1, ..., x_n, y_1, ..., y_n) &= x_1^2 + \cdots x_n^2 - 1 \\ g_2(x_1, ..., x_n, y_1, ..., y_n) &= y_1^2 + \cdots y_n^2 - 1 \end{aligned}$$

y el conjunto

$$S = \{(x_1, ..., x_n, y_1, ..., y_n) \in \mathbb{R}^n \times \mathbb{R}^n \mid g_1(x_1, ..., x_n, y_1, ..., y_n) = 0,\ g_2(x_1, ..., x_n, y_1, ..., y_n) = 0\}.$$

Note que el conjunto S es un conjunto compacto y como f es continua entonces queda garantizado que la función f alcanza su valor máximo y su valor mínimo sobre S, de hecho, usando el Teorema de los Multiplicadores de Lagrange deseamos demostrar que dichos valores son 1 y -1 respectivamente.

Sea $(\overline{x}_0, \overline{y}_0) = (x_{0,1}, ..., x_{0,n}, y_{0,1}, ..., y_{0,n}) \in S$ un punto donde f alcanza uno de sus valores extremos. Entonces existen $\lambda_1, \lambda_2 \in \mathbb{R}$ tales que

$$\nabla f(\overline{x}_0, \overline{y}_0) = \lambda_1 \nabla g_1(\overline{x}_0, \overline{y}_0) + \lambda_2 \nabla g_2(\overline{x}_0, \overline{y}_0),$$

es decir que

$$y_{0,i} = \frac{\partial f}{\partial x_i}(\overline{x}_0, \overline{y}_0) = \lambda_1 \frac{\partial g_1}{\partial x_i}(\overline{x}_0, \overline{y}_0) = 2\lambda_1 x_{0,i} \tag{9}$$

y

$$x_{0,i} = \frac{\partial f}{\partial y_i}(\overline{x}_0, \overline{y}_0) = \lambda_2 \frac{\partial g_2}{\partial y_i}(\overline{x}_0, \overline{y}_0) = 2\lambda_2 y_{0,i} \tag{10}$$

para cada $i \in \{1, ..., n\}$. Multiplicando (9) por $x_{0,i}$, luego sumando sobre el índice i y usando que $(\overline{x}_0, \overline{y}_0) \in S$, tenemos que

$$\begin{aligned} x_{0,1}y_{0,1} + \cdots + x_{0,n}y_{0,n} &= 2\lambda_1 \left(x_{0,1}^2 + \cdots + x_{0,n}^2\right) \\ &= 2\lambda_1(g_1(\overline{x}_0, \overline{y}_0) + 1) \\ &= 2\lambda_1. \end{aligned}$$

De manera análoga, multiplicando (10) por $y_{0,i}$, sumando sobre el índice i y usando que $(\overline{x}_0, \overline{y}_0) \in S$, entonces

$$\begin{aligned} x_{0,1}y_{0,1} + \cdots + x_{0,n}y_{0,n} &= 2\lambda_2 \left(y_{0,1}^2 + \cdots + y_{0,n}^2\right) \\ &= 2\lambda_2(g_2(\overline{x}_0, \overline{y}_0) + 1) \\ &= 2\lambda_2. \end{aligned}$$

Se sigue que $\lambda_1 = \lambda_2$, así que podemos escribir λ en (9) y (10) en vez de λ_1 y λ_2, y si además las combinamos, tenemos que

$$y_{0,i} = 2\lambda x_{0,i} = 2\lambda(2\lambda y_{0,i}) = 4\lambda^2 y_{0,i} \tag{11}$$

y

$$x_{0,i} = 2\lambda y_{0,i} = 2\lambda(2\lambda x_{0,i}) = 4\lambda^2 x_{0,i} \tag{12}$$

para cada $i \in \{1, ..., n\}$. Ahora, como $\|\overline{x}_0\| = 1$, entonces para algún $i \in \{1, ..., n\}$ se tiene que $x_{0,i} \neq 0$ y usando (12) tenemos que $\lambda = \pm\frac{1}{2}$. Luego, sustituyendo este valor en (9) o (10) tenemos que

$$y_{0,i} = \pm x_{0,i}$$

o

$$x_{0,i} = \pm y_{0,i}.$$

Así, tenemos

$$\begin{aligned} f(x_{0,1}, ..., x_{0,n}, y_{0,1}, ..., y_{0,n}) &= f(x_{0,1}, ..., x_{0,n}, \pm x_{0,1}, ..., \pm x_{0,n}) \\ &= x_{0,1}\left(\pm x_{0,1}\right) + \cdots + x_{0,n}\left(\pm x_{0,n}\right) \\ &= \pm\left((x_{0,1})^2 + \cdots + (x_{0,n})^2\right) \\ &= \pm 1 \end{aligned}$$

o

$$\begin{aligned} f(x_{0,1}, ..., x_{0,n}, y_{0,1}, ..., y_{0,n}) &= f(\pm y_{0,1}, ..., \pm y_{0,n}, y_{0,1}, ..., y_{0,n}) \\ &= y_{0,1}\left(\pm y_{0,1}\right) + \cdots + y_{0,n}\left(\pm y_{0,n}\right) \\ &= \pm\left((y_{0,1})^2 + \cdots + (y_{0,n})^2\right) \\ &= \pm 1. \end{aligned}$$

Por lo tanto, el valor mínimo de f sobre el conjunto S es -1 y el valor máximo es 1, es decir,

$$|\overline{x} \cdot \overline{y}| = |f(\overline{x}, \overline{y})| \leq 1$$

para todo $\overline{x}, \overline{y} \in \mathbb{R}^n$ tales que $\|\overline{x}\| = \|\overline{y}\| = 1$. ∎