CÁLCULO VECTORIAL

TEMA 5

EXTREMOS CONDICIONADOS MULTIPLICADORES DE LAGRANGE

El primer tema estudiado en lo que va de curso fue el cálculo de extremos absolutos y relativos de campos escalares. Dado un campo escalar $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ continuo y de clase C^3 en su interior, era posible determinar sus puntos críticos $\boldsymbol{x_0} \in U^\circ$ y clasificarlos si cierta condición de diferenciabilidad se daba, a saber, $\det(H_f(\boldsymbol{x_0})) \neq 0$. En la mayoría de los ejemplos vistos no nos interesaba cómo era el dominio U, pues no cumple un rol muy notorio en el criterio del diferencial de orden 2. Sin embargo, cuando queremos hallar extremos absolutos de f, en ciertos casos sí importa mucho cómo es su dominio. Por ejemplo, si U es un conjunto compacto, es posible determinar los extremos absolutos de f, que sabemos que existen por el Teorema de Weierstrass. Primero hallamos los puntos críticos en U° mediante el criterio del diferencial de orden 2, y luego los puntos críticos en la frontera ∂U.

Para la última situación que concierne a $\mathcal{C} = \partial U$, el cálculo de extremos relativos o absolutos de f alcanzados en puntos de $\mathcal{C}$ es lo que se conoce como *problema de extremos con condiciones*. En general son difíciles de resolver, pero para ciertos $\mathcal{C}$ existe un método para esto: el *teorema de multiplicadores de Lagrange*. Los subconjuntos $\mathcal{C}$ del dominio U, conocidos también como *condiciones*, consideraremos en estos problemas serán conjuntos de nivel de algún campo escalar diferenciable.

Estudiaremos además el problema más general de hallar extremos de f dado un número finito de condiciones, es decir, en el conjunto de puntos $\boldsymbol{x}$ del dominio de f que satisfacen $\boldsymbol{x} \in \mathcal{C}_1 \cap \cdots \cap \mathcal{C}_m$. Por ejemplo, se nos puede pedir determinar los extremos de un campo f restringido a la curva de intersección de dos superficies de nivel.

5.1 Extremos de campos escalares con una sola condición

Antes de dar el enunciado y demostración del teorema de multiplicadores de Lagrange con una condición, mencionemos algunas motivaciones:

1. **Topología:** Dada una superficie S que no pase por el origen, determinar los puntos de S que están más cerca del origen.

2. **Física:** Sea $t\colon \mathbb{R}^3 \to \mathbb{R}$ la función que mide la temperatura de cualquier punto en el espacio (en un instante fijo). Determinar las temperaturas máxima y mínima en una curva $\mathcal{C}$ de $\mathbb{R}^3$.

3. **Economía:** Sea $e\colon \mathbb{R}^2 \to \mathbb{R}$ una función que representa la ganancia obtenida de la venta de dos productos x e y. Acá, x representa la cantidad del primer producto, e y la cantidad del segundo producto. Seguramente nos interesa predecir la ganancia máxima de la venta de x e y, pero esto va a estar limitado por varias condiciones, como por ejemplo la cantidad de producto x e y del que dispongamos inicialmente.

Definición 5.1.1 (extremos condicionados). *Sean $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ y $g\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ campos escalares, y $\mathcal{C} = \{\boldsymbol{x} \in U : g(\boldsymbol{x}) = 0\}$. Se dice que f tiene en $\boldsymbol{x_0} \in \mathcal{C}$:*

1. *un **máximo condicionado** (resp., **mínimo condicionado**) **a $\mathcal{C}$** si*

$$f(\boldsymbol{x_0}) \geq f(\boldsymbol{x}) \text{ (resp., } f(\boldsymbol{x_0}) \leq f(\boldsymbol{x})\text{) para todo } \boldsymbol{x} \in \mathcal{C};$$

2. *un **máximo relativo condicionado** (resp., **mínimo relativo condicionado**) **a $\mathcal{C}$** si existe $\delta > 0$ tal que*

$$f(\boldsymbol{x_0}) \geq f(\boldsymbol{x}) \text{ (resp., } f(\boldsymbol{x_0}) \leq f(\boldsymbol{x})\text{) para todo } \boldsymbol{x} \in \mathcal{C} \cap B(\boldsymbol{x_0}, \delta).$$

Observación 5.1.2. *Si f alcanza un extremo en $\boldsymbol{x_0}$, entonces alcanza en $\boldsymbol{x_0}$ un extremo condicionado a $\mathcal{C}$ si $\boldsymbol{x_0} \in \mathcal{C}$. Sin embargo, un extremo condicionado de f no tiene por qué ser un extremo de f. Por ejemplo, si $f\colon \mathbb{R}^3 \to \mathbb{R}$ está dado por $f(\boldsymbol{x}) = ||\boldsymbol{x}||$ (la distancia de $\boldsymbol{x}$ al origen) y $\mathcal{C} = \{\boldsymbol{x} \in \mathbb{R}^3 : y = 1\}$ (es decir, $\mathcal{C}$ es el plano $y = 1$), entonces f tiene un mínimo condicionado al plano $y = 1$ en el punto $(0, 1, 0)$. Sin embargo, en este punto la función f no alcanza un mínimo.*

Teorema 5.1.3 (de multiplicadores de Lagrange para una condición). *Sean $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ y $g\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ campos escalares de clase C^1, y $\boldsymbol{x_0} \in U^\circ$ un punto tal que $g(\boldsymbol{x_0}) = 0$ y $\nabla g(\boldsymbol{x_0}) \neq 0$. Si $\boldsymbol{x_0} \in U$ es un extremo de f condicionado a $\mathcal{C} = \{\boldsymbol{x} \in U : g(\boldsymbol{x}) = 0\}$, entonces existe un escalar $\lambda \in \mathbb{R}$ tal que*

$$\nabla f(\boldsymbol{x_0}) = \lambda \nabla g(\boldsymbol{x_0}).$$

Observación 5.1.4.

1. *Al escalar λ del enunciado anterior se le conoce como* ***multiplicador de Lagrange****.*

2. *Geométricamente para el caso $n = 3$, el teorema dice que si f tiene un náximo o mínimo cuando está restringida a la superficie $\mathcal{C}$, entonces $\nabla f(\boldsymbol{x_0})$ es perpendicular a $\mathcal{C}$ en $\boldsymbol{x_0}$.*

Antes de demostrar el teorema de multiplicadores de Lagrange, es importante entender qué significa su enunciado geométricamente. Para esto lo mejor es tomar un ejemplo concreto. Consideremos el problema planteado al inicio en el punto 1. El campo $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ mide la distancia de (x, y, z) al origen. Por otro lado, sea $\mathcal{C}$ una superficie que no pasa por el origen y que viene dada por $g(x, y, z) = 0$ para algún campo escalar g (una esfera o un plano, por ejemplo). Queremos determinar los puntos de $\mathcal{C}$ que están más próximos al origen. Supongamos que esta distancia es r_0, y consideremos las superficies de nivel $f(x, y, z) = r \geq 0$. A partir de $r = 0$ (es decir, (x, y, z) es el origen) empezamos a incrementar este valor. Para $r > 0$ tenemos que (x, y, z) está en la superficie de nivel $f(x, y, z) = r$ (esfera de radio r centrada en el origen). Seguimos incrementando r, es decir, "expandiendo" esta esfera de nivel hasta que toca tangencialmente a $\mathcal{C}$ en uno o más puntos. Es en tales puntos (x_0, y_0, z_0) donde ocurre $f(x_0, y_0, z_0) = r_0$. Al tratarse de un punto de tangencia, tenemos que los planos tangentes de las superficies $f(x, y, z) = r_0$ y $g(x, y, z) = 0$ en el punto (x_0, y_0, z_0) coinciden, por lo que sus vectores normales $\nabla f(x_0, y_0, z_0)$ y $\nabla g(x_0, y_0, z_0)$ son paralelos, es decir, que existe $\lambda \in \mathbb{R}$ tal que $\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$.

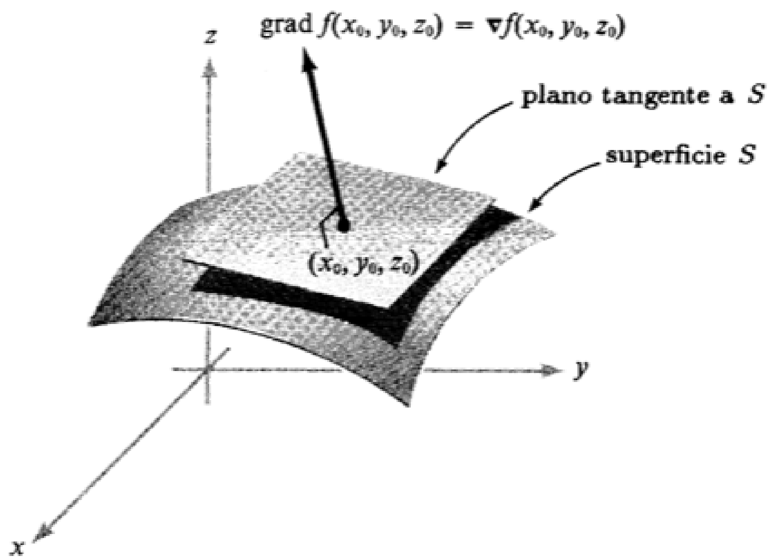

Figura 5.1: (Imagen tomada de Cálculo Vectorial – Marsden y Tromba)

Demostración del teorema de multiplicadores de Lagrange (una condición):
Tenemos un punto interior $\boldsymbol{x_0} \in U^\circ$ que está en la superficie $g(\boldsymbol{x}) = 0$, es decir, $g(\boldsymbol{x_0}) = 0$. Para simplificar un poco la demostración y darle una interpretación geométrica, supongamos sin pérdida de generalidad que $\boldsymbol{x_0} = \mathbf{0}$. La condición $\nabla g(\mathbf{0}) \neq 0$ implica que podemos considerar el plano tangente a esta superficie en el origen $\mathbf{0}$. A su vez, este plano tangente puede verse como el espacio vectorial $V = \{\boldsymbol{v} \in \mathbb{R}^n : \nabla g(\mathbf{0}) \cdot \boldsymbol{v} = 0\}$. Ahora, probar que $\nabla f(\mathbf{0})$ y $\nabla g(\mathbf{0})$ son paralelos significa demostrar que $\nabla f(\mathbf{0})$ pertenece al complemento ortogonal de V, es decir, al espacio vectorial $V^\perp = \{\boldsymbol{w} \in \mathbb{R}^n : \boldsymbol{w} \cdot \boldsymbol{v} = 0 \text{ para todo } \boldsymbol{v} \in V\}$. Note además que $V^\perp$ es el subespacio generado por $\nabla g(\mathbf{0})$, por lo cual $\nabla f(\mathbf{0}) \in V^\perp$ si, y solo si, existe $\lambda \in \mathbb{R}$ tal que $\nabla f(\mathbf{0}) = \lambda \nabla g(\mathbf{0})$.

Habiendo entendido a dónde queremos llegar, para cualquier $\boldsymbol{v} \in V$ queremos ver que $\nabla f(\mathbf{0}) \cdot \boldsymbol{v} = 0$. Para probar esto último aplicamos herramientas de cálculo en varias variables. Como $\boldsymbol{v}$ es un vector en el plano tangente a la superficie de nivel cero de g en $\mathbf{0}$, podemos considerar una curva $\boldsymbol{\sigma}\colon I \to \mathbb{R}^n$, con I un intervalo que contiene a 0, tal que $\boldsymbol{\sigma}(0) = \mathbf{0}$ y $\boldsymbol{\sigma}'(0) = \boldsymbol{v}$. Como $f|_C$ alcanza un extremo relativo en $\mathbf{0}$, la composición $f \circ \boldsymbol{\sigma}$ alcanza un extremo relativo en $t = 0$, de donde $(f \circ \boldsymbol{\sigma})'(0) = 0$. Por otro lado, por la regla de la cadena, se tiene que

$$(f \circ \boldsymbol{\sigma})'(0) = \nabla f(\boldsymbol{\sigma}(0)) \cdot \boldsymbol{\sigma}'(0) = \nabla f(\mathbf{0}) \cdot \boldsymbol{v}.$$

Por lo tanto, $\nabla f(\mathbf{0}) \cdot \boldsymbol{v} = 0$, como se quería demostrar. □

Ejemplo 5.1.5.

1. *Halle los valores extremos de la función $f(x, y) = xy$ sujeta a la condición $x + y = 1$.*

   La condición es la curva de nivel cero de la función $g(x, y) = x + y - 1$. Para hallar los posibles puntos en la recta $x + y = 1$ donde f alcanza sus valores extremos, debemos resolver el sistema

   $$\begin{cases} \nabla f(x,y) &= \lambda \nabla g(x,y), \\ g(x,y) &= 0. \end{cases}$$

   Por "resolver" nos referimos a hallar los puntos (x, y) en la recta $x + y = 1$ y los escalares $\lambda \in \mathbb{R}$ para los cuales la igualdad anterior se cumple. Tenemos que

   $$\begin{aligned} \nabla f(x,y) &= (y, x), \\ \nabla g(x,y) &= (1, 1), \\ (y, x) &= \lambda(1,1) = (\lambda, \lambda). \end{aligned}$$

   De lo anterior se sigue que $x = y = \lambda$. Por otro lado, recordemos que el punto (x, y) debe satisfacer la condición $x + y = 1$. Entonces, $2x = 1$, es decir, $x = y = \lambda = 1/2$. Tenemos así que $(x, y) = (1/2, 1/2)$ y $\lambda = 1/2$ son los puntos y escalares que cumplen la igualdad $\nabla f(x, y) = \lambda \nabla g(x, y)$. Ahora, esto no quiere decir aún que en $(1/2, 1/2)$ la *función f alcanza un extremo condicionado. Verifiquemos esto último. Debe ocurrir que $f(x,y) \geq f(1/2,1/2)$ o $f(x,y) \leq f(1/2,1/2)$ para todo (x,y) en la recta $x+y=1$, es decir, $f(x,1-x) \geq f(1/2,1/2)$ o $f(x,1-x) \leq f(1/2,1/2)$ para todo $x \in \mathbb{R}$. En este caso, va a ocurrir la segunda desigualdad. En efecto,*

$$\begin{aligned} f(x,1-x) \leq f(1/2,1/2) \quad &\iff \quad x - x^2 \leq \frac{1}{4} \\ &\iff \quad x^2 - x \geq -\frac{1}{4} \\ &\iff \quad x^2 - 2\cdot x \cdot \frac{1}{2} + \frac{1}{4} \geq 0 \\ &\iff \quad \left(x - \frac{1}{2}\right)^2 \geq 0. \end{aligned}$$

Como la última desigualdad en la equivalencia anterior siempre se cumple, tenemos que $f(x,y) \leq f(1/2,1/2)$ para todo (x,y) en la recta $x+y=1$. Por lo tanto, f alcanza en $(1/2,1/2)$ un máximo condicionado a $\mathcal{C} = \{(x,y) \in \mathbb{R}^2 : x+y=1\}$.

2. *Hagamos un ejemplo más geométrico. Supongamos que se nos pide calcular la distancia mínima entre el punto $(1,0)$ y la parábola $y^2 = 4x$.*

   La distancia de cualquier punto (x,y) del plano $\mathbb{R}^2$ al punto $(1,0)$ viene dada por

   $$d(x,y) = ||(x,y) - (1,0)|| = \sqrt{(x-1)^2 + y^2}.$$

   Queremos entonces hallar el o los puntos en la curva $y^2 = 4x$ cuya distancia a $(1,0)$ es mínima, y hallar dicha distancia. Debemos entonces resolver el sistema

   $$\begin{cases} \nabla d(x,y) &= \lambda \nabla g(x,y), \\ g(x,y) &= 0 \end{cases}$$

   donde $g\colon \mathbb{R}^2 \to \mathbb{R}$ es el campo escalar $g(x,y) = 4x - y^2$ que define la condición. Tenemos:

   $$\begin{aligned} \nabla d(x,y) &= \left(\frac{2(x-1)}{2\sqrt{(x-1)^2+y^2}}, \frac{2y}{\sqrt{2(x-1)^2+y^2}} \right) = \left(\frac{x-1}{d(x,y)}, \frac{y}{d(x,y)} \right), \\ \nabla g(x,y) &= (4,-2y), \\ \left(\frac{x-1}{d(x,y)}, \frac{y}{d(x,y)} \right) &= \lambda(4,-2y) = (4\lambda, -2\lambda y). \end{aligned}$$

   Lo anterior implica que

   $$x - 1 = 4\lambda d(x,y) \qquad e \qquad y = -2\lambda y d(x,y).$$

Si suponemos $y \neq 0$, entonces de la segunda igualdad se tiene que $1 = -2\lambda d(x, y)$, es decir, $\lambda d(x, y) = -1/2$. Sustituyendo en la primera igualdad, nos queda $x - 1 = 4(-1/2) = -2$, es decir, $x = -1$. Recordemos que (x, y) debe estar en la curva $y^2 = 4x$, de donde $y^2 = 4(-1) = -4$. Esto último es una contradicción, que viene de suponer que $y \neq 0$. Entonces, tiene que cumplirse que $y = 0$, y como $y^2 = 4x$, se tiene que $x = 0$. Finalmente, sustituyendo en $x - 1 = 4\lambda d(x, y)$ nos queda $-1 = 4\lambda d(0, 0)$, de donde $\lambda = -1/4$.

Falta verificar que d alcanza en $(0, 0)$ un mínimo condicionado a $y^2 = 4x$, es decir, que $d(x, y) \geq d(0, 0) = 1$ para todo (x, y) en la curva $y^2 = 4x$. Tenemos lo siguiente:

$$d\left(\frac{y^2}{4}, y\right) \geq 1 \iff \sqrt{\left(\frac{y^2}{4} - 1\right)^2 + y^2} \geq 1 \iff \frac{y^4}{16} + \frac{y^2}{2} + 1 \geq 1.$$

La última desigualdad de la equivalencia anterior siempre se cumple, por lo cual la distancia mínima de $(1, 0)$ a la parábola $y^2 = 4x$ es 1.

3. *En el ejemplo anterior, la curva de nivel o condición del problema era muy fácil de dibujar y el resultado era bastante predecible. Mas aún, tal resultado se pudo haber obtenido sin necesidad de multiplicadores de Lagrange. Efectivamente, pudimos haber planteado un problema de optimización en una variable, a saber, hallar el mínimo de la función $d(y^2/4, y)$. Lo anterior es posible porque toda la curva de condición se puede expresar como la gráfica de una función de variable y. La ventaja del uso de multiplicadores de Lagrange se nota más en ejemplos donde no resulta sencillo dibujar la curva ni parametrizarla completamente como una función de una variable.*

   Estudiemos entonces un ejemplo similar al anterior pero con una curva más complicada. Queremos hallar las distancias máxima y mínima del origen a la elipse de ecuación cartesiana $5x^2 + 6xy + 5y^2 = 8$. Debemos considerar entonces el campo escalar $f(x, y) = \sqrt{x^2 + y^2}$ del cual queremos sus extremos condicionados, y el campo escalar $g(x, y) = 5x^2 + 6xy + 5y^2 - 8$ que define la condición. Queremos resolver entonces

$$\begin{cases} \nabla f(x, y) &= \lambda \nabla g(x, y), \\ g(x, y) &= 0 \end{cases}$$

   Tenemos:

$$\begin{aligned} \nabla f(x, y) &= \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}}\right) = \left(\frac{x}{f(x, y)}, \frac{y}{f(x, y)}\right), \\ \nabla g(x, y) &= (10x + 6y, 10y + 6x), \\ \left(\frac{x}{f(x, y)}, \frac{y}{f(x, y)}\right) &= \lambda(10x + 6y, 10y + 6x), \end{aligned}$$

de donde nos queda el sistema:

$$\begin{cases} x &= \lambda(10x+6y)f(x,y), \\ y &= \lambda(6x+10y)f(x,y), \\ 5x^2+6xy+5y^2 &= 8. \end{cases}$$

Para poder simplificar, una posible opción es dividir la primera ecuación por la segunda (asumiendo $y \neq 0$):

$$\frac{x}{y} = \frac{\lambda(10x+6y)f(x,y)}{\lambda(6x+10y)f(x,y)} = \frac{10x+6y}{6x+10y},$$

entonces

$$\begin{aligned} x(6x+10y) &= y(10x+6y) \\ 6x^2+10xy &= 10xy+6y^2 \\ x^2 &= y^2 \\ x &= \pm y. \end{aligned}$$

Para $x = y$, sustituyendo en la tercera igualdad del sistema, se tiene que

$$\begin{aligned} 5x^2+6xy+5y^2 &= 8 \\ 5y^2+6y^2+5y^2 &= 8 \\ y &= \pm\frac{1}{\sqrt{2}} \end{aligned}$$

Para $x = -y$, de manera similar se tiene que $y = \pm\sqrt{2}$. Tenemos entonces que las soluciones son los puntos $(-\sqrt{2}, \sqrt{2})$, $(\sqrt{2}, -\sqrt{2})$, $(1/\sqrt{2}, 1/\sqrt{2})$ y $(-1/\sqrt{2}, -1/\sqrt{2})$. Notemos que el origen equidista de $(-\sqrt{2}, \sqrt{2})$ y de $(\sqrt{2}, -\sqrt{2})$, es decir, $f(-\sqrt{2}, \sqrt{2}) = f(\sqrt{2}, -\sqrt{2})$, por lo cual, si f alcanza un extremo condicionado en $(-\sqrt{2}, \sqrt{2})$, valdrá lo mismo que en $(\sqrt{2}, -\sqrt{2})$. Se puede aplicar también este análisis a $(1/\sqrt{2}, 1/\sqrt{2})$ y $(-1/\sqrt{2}, -1/\sqrt{2})$. Entonces, para el resto del problema es suficiente trabajar, por ejemplo, con los puntos $P_1 = (-\sqrt{2}, \sqrt{2})$ y $P_2 = (1/\sqrt{2}, 1/\sqrt{2})$. Para tales puntos tenemos que

$$\begin{aligned} f(P_1) &= \sqrt{(-\sqrt{2})^2+(\sqrt{2})^2} = \sqrt{2+2} = 2, \\ f(P_2) &= \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2+\left(\frac{1}{\sqrt{2}}\right)^2} = \sqrt{\frac{1}{2}+\frac{1}{2}} = 1. \end{aligned}$$

En principio, podríamos concluir que f alcanza en P_1 y en P_2 su máximo y mínimo condicionado a $\mathcal{C} = \{(x,y) \in \mathbb{R}^2 : 5x^2+6xy+5y^2 = 8\}$, respectivamente, pero falta analizar qué pasa si $y = 0$. Para esto último se tiene a partir de $5x^2+6xy+5y^2 = 8$ que $x = \pm\frac{2}{5}\sqrt{10}$. Por otro lado, $f\left(-\frac{2}{5}\sqrt{10}, 0\right) = f\left(\frac{2}{5}\sqrt{10}, 0\right) = \frac{2}{5}\sqrt{10}$, y es fácil verificar que $f(P_1) \geq f\left(-\frac{2}{5}\sqrt{10}, 0\right) = f\left(\frac{2}{5}\sqrt{10}, 0\right) \geq f(P_2)$. Ahora sí podemos afirmar que la *distancia máxima del origen a la elipse $5x^2 + 6xy + 5y^2 = 8$ es 2 (alcanzada en los puntos $(-\sqrt{2}, \sqrt{2})$ y $(\sqrt{2}, -\sqrt{2})$), y la mímima es 1 (alcanzada en los puntos $(1/\sqrt{2}, 1/\sqrt{2})$ y $(-1/\sqrt{2}, -1/\sqrt{2})$).*

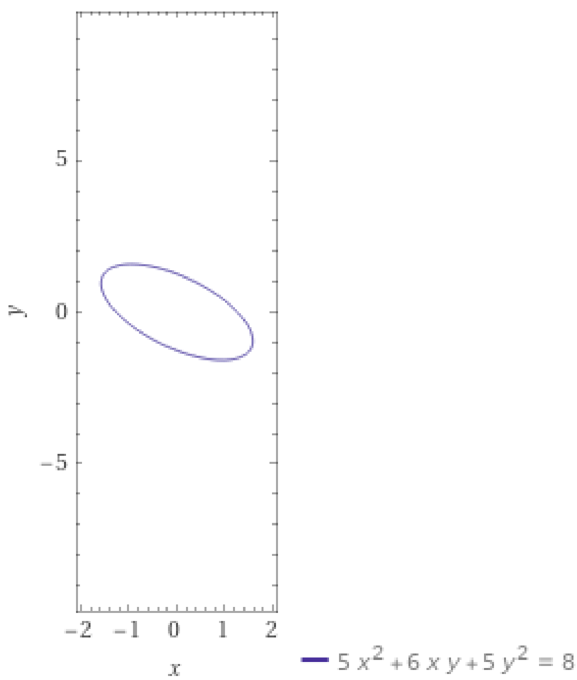

Figura 5.2: Normalmente, se habría tenido que probar que $f(P_1) \geq f(x, y) \geq f(P_2)$ para todo $(x, y) \in \mathcal{C}$, pero se nos dice en un principio que la curva que define la condición es una elipse, que a su vez es una curva cerrada, y por lo tanto un conjunto compacto en $\mathbb{R}^2$. Así, $f|_{\mathcal{C}}$ alcanza máximo y mínimo absoluto por el Teorema de Weierstrass.

5.2 Extremos de campos escalares con varias condiciones

En esta sección estudiaremos la generalización del teorema de multiplicadores de Lagrange a campos escalares con múltiples condiciones.

Teorema 5.2.1 (de multiplicadores de Lagrange para varias condiciones). *Sean $f, g_1, \ldots, g_m : U \subseteq \mathbb{R}^n \to \mathbb{R}$ campos escalares de clase C^1 con $m < n$. Sea $\boldsymbol{x_0} \in U$ un punto tal que $g_1(\boldsymbol{x_0}) = \cdots = g_m(\boldsymbol{x_0}) = 0$ y $\{\nabla g_1(\boldsymbol{x_0}), \ldots, \nabla g_m(\boldsymbol{x_0})\}$ es un conjunto linealmente independiente. Si f presenta en $\boldsymbol{x_0}$ un extremo condicionado a $\mathcal{C} = \{\boldsymbol{x} \in U : g_1(\boldsymbol{x}) = \cdots = g_m(\boldsymbol{x}) = 0\}$, entonces existen escalares $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ tales que*

$$\nabla f(\boldsymbol{x_0}) = \lambda_1 \nabla g_1(\boldsymbol{x_0}) + \cdots + \lambda_m \nabla g_m(\boldsymbol{x_0}) = \sum_{k=1}^{m} \lambda_k \nabla g_k(\boldsymbol{x_0}).$$

Demostración: Probemos solo el caso $m = 2$ ($n \geq 3$). Para presentar una demostración más geométrica, supongamos también que $n = 3$.[1] Sea $(x_0, y_0, z_0) \in \mathcal{C}$ un punto donde f presenta un extremo condicionado a $\mathcal{C}$. En este caso, note que $\mathcal{C}$ es la curva de intersección entre dos superficies de nivel, a saber, $g_1(x, y, z) = 0$ y $g_2(x, y, z) = 0$. Se asume además que el conjunto $\{\nabla g_1(x_0, y_0, z_0), \nabla g_2(x_0, y_0, z_0)\}$ es linealmente independiente. Queremos probar que $\nabla f(x_0, y_0, z_0)$ está en el subespacio generado por $\nabla g_1(x_0, y_0, z_0)$ y $\nabla g_2(x_0, y_0, z_0)$.

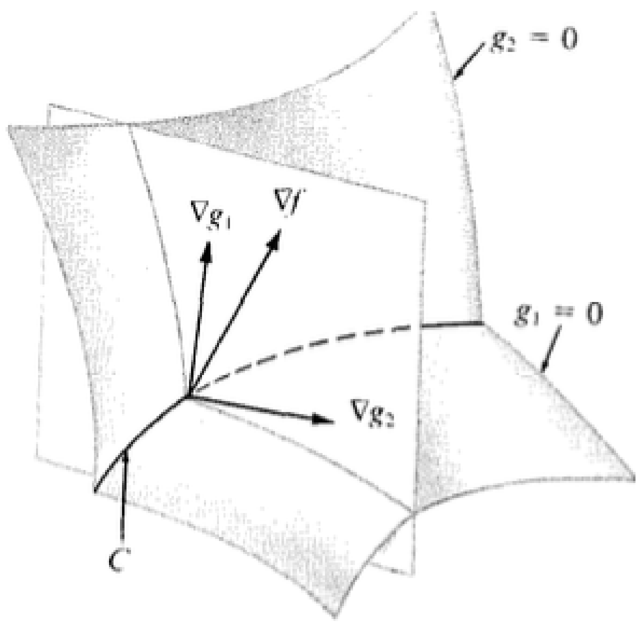

Figura 5.3: `(Imagen tomada de Calculus Vol. 2 - Apostol)`

Consideremos la recta tangente a $\mathcal{C}$ en (x_0, y_0, z_0), y sea $\boldsymbol{v}$ un vector de dirección en esa recta. Sea $I \subseteq \mathbb{R}$ un intervalo abierto que contiene a 0, y $\boldsymbol{\sigma} \colon I \to \mathcal{C}$ tal que $\boldsymbol{\sigma}(0) = (x_0, y_0, z_0)$ y $\boldsymbol{\sigma}'(0) = \boldsymbol{v}$. Es decir, podemos expresar $\mathcal{C}$ en un entorno de (x_0, y_0, z_0) como una curva paramétrica (esto lo justificaremos al final de la prueba). Entonces, tenemos que la composición $f \circ \boldsymbol{\sigma}(t)$ alcanza un extremo en $t = 0$, de donde $0 = (f \circ \boldsymbol{\sigma})'(0)$. Por otro lado, por la regla de la cadena tenemos que $0 = \nabla f(\boldsymbol{\sigma}(0)) \cdot \boldsymbol{\sigma}'(0) = \nabla f(x_0, y_0, z_0) \cdot \boldsymbol{v}$. Tenemos así que $\nabla f(x_0, y_0, z_0)$ pertenece al complemento ortogonal de la recta tangente a $\mathcal{C}$ en (x_0, y_0, z_0). Por otro lado, $\boldsymbol{v}$ pertenece al plano tangente a $g_1(x, y, z) = 0$ en (x_0, y_0, z_0), y al plano tangente a $g_2(x, y, z) = 0$ en (x_0, y_0, z_0), por lo cual $\nabla g_1(x_0, y_0, z_0) \cdot \boldsymbol{v} = 0$ y $\nabla g_2(x_0, y_0, z_0) \cdot \boldsymbol{v} = 0$. Entonces, $\nabla g_1(x_0, y_0, z_0)$ y $\nabla g_2(x_0, y_0, z_0)$ también están en el complemento ortogonal de la recta tangente a $\mathcal{C}$ en (x_0, y_0, z_0), que es un subespacio vectorial de dimensión dos. Entonces, al ser $\nabla g_1(x_0, y_0, z_0)$ y $\nabla g_2(x_0, y_0, z_0)$ linealmente independientes, tenemos que ellos generan a este subespacio. Por lo tanto, existen $\lambda_1, \lambda_2 \in \mathbb{R}$ tales que $\nabla f(x_0, y_0, z_0) = \lambda_1 \nabla g_1(x_0, y_0, z_0) + \lambda_2 g_2(x_0, y_0, z_0)$.

[1] La demostración general se puede hacer siguiendo los mismos argumentos.

Finalmente, vamos a justificar por qué se puede parametrizar $\mathcal{C}$ en un entorno de (x_0, y_0, z_0). Como $\nabla g_1(x_0, y_0, z_0)$ y $\nabla g_2(x_0, y_0, z_0)$ son linealmente independientes, su producto cruz $\nabla g_1(x_0, y_0, z_0) \times \nabla g_2(x_0, y_0, z_0)$ es diferente de cero. Entonces,

$$0 \neq \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial g_1}{\partial x}(x_0, y_0, z_0) & \frac{\partial g_1}{\partial y}(x_0, y_0, z_0) & \frac{\partial g_1}{\partial z}(x_0, y_0, z_0) \\ \frac{\partial g_2}{\partial x}(x_0, y_0, z_0) & \frac{\partial g_2}{\partial y}(x_0, y_0, z_0) & \frac{\partial g_2}{\partial z}(x_0, y_0, z_0) \end{vmatrix}.$$

Para que esto ocurra, alguno de los determinantes menores que involucra a g_1 y g_2 debe ser no nulo, es decir,

$$\begin{vmatrix} \frac{\partial g_1}{\partial x}(x_0, y_0, z_0) & \frac{\partial g_1}{\partial y}(x_0, y_0, z_0) \\ \frac{\partial g_2}{\partial x}(x_0, y_0, z_0) & \frac{\partial g_2}{\partial y}(x_0, y_0, z_0) \end{vmatrix} \neq 0,$$

$$\begin{vmatrix} \frac{\partial g_1}{\partial x}(x_0, y_0, z_0) & \frac{\partial g_1}{\partial z}(x_0, y_0, z_0) \\ \frac{\partial g_2}{\partial x}(x_0, y_0, z_0) & \frac{\partial g_2}{\partial z}(x_0, y_0, z_0) \end{vmatrix} \neq 0, \quad \text{o}$$

$$\begin{vmatrix} \frac{\partial g_1}{\partial y}(x_0, y_0, z_0) & \frac{\partial g_1}{\partial z}(x_0, y_0, z_0) \\ \frac{\partial g_2}{\partial y}(x_0, y_0, z_0) & \frac{\partial g_2}{\partial z}(x_0, y_0, z_0) \end{vmatrix} \neq 0.$$

En cualquier caso, se podrán despejar dos de las variables x, y o z en función de la otra variable restante en un entorno de (x_0, y_0, z_0), lo que da lugar a una curva paramétrica de la forma $\boldsymbol{\sigma}(t)$ especificada anteriormente. En principio, no tiene por qué ocurrir que $\boldsymbol{\sigma}(0) = \boldsymbol{x_0}$ (de hecho, $\boldsymbol{\sigma}$ podría no estar definida en 0), pero es posible reparametrizar $\boldsymbol{\sigma}$ mediante una traslación en el intervalo dominio para que esto suceda.[2] $\square$

Observación 5.2.2.

1. *En el enunciado del teorema se pide que el conjunto $\{\nabla g_1(\boldsymbol{x_0}), \ldots, \nabla g_m(\boldsymbol{x_0})\}$ esté formado por vectores linealmente independientes. Al ser éstos vectores en $\mathbb{R}^n$, no pueden ser más de n, lo cual implica que $m \leq n$. ¿Ahora, por qué de pide que m sea estrictamente menor que n? En el caso $n = m$, se tiene que $\{\nabla g_1(\boldsymbol{x_0}), \ldots, \nabla g_m(\boldsymbol{x_0})\}$ es linealmente independiente y la conclusión del teorema es trivial, y ni siquiera depende del hecho que f alcance en $\boldsymbol{x_0}$ un extremo condicionado a $\mathcal{C}$. Entonces, con $m < n$ y las hipótesis dadas, geométricamente hablando, el teorema nos dice que $\nabla f(\boldsymbol{x_0})$ pertenece al subespacio propio generado por $\{\nabla g_1(\boldsymbol{x_0}), \ldots, \nabla g_m(\boldsymbol{x_0})\}$.*

2. *Si $\{\nabla g_1(\boldsymbol{x_0}), \ldots, \nabla g_m(\boldsymbol{x_0})\}$ es linealmente dependiente el teorema falla. Como contraejemplo, analicemos el caso $n = 3$ y $m = 2$, con la función $f(x, y) = x^2 + y^2$. Supongamos que se nos pide hallar los extremos condicionados de f a la curva de intersección entre las superficies $g_1(x, y, z) = 0$ y $g_2(x, y, z) = 0$ donde $g_1(x, y, z) = z$ y $g_2(x, y, z) = z^2 - (y - 1)^3$. Las superficies de nivel cero correspondientes a g_1 y g_2 son un plano y un cilindro, respectivamente,*

[2] El tema de curvas parametrizadas y reparametrizaciones se estudiará en el siguiente tema de este curso de Cálculo Vectorial.

que se intersecan en la recta $\mathcal{C} = \{(x,y,z) \in \mathbb{R}^3 : x \in \mathbb{R},\ y = 1$ y $z = 0\}$. En este caso tenemos $\nabla g_1(x,y,z) = (0,0,1)$, $\nabla g_2(x,y,z) = (0,-3(y-1)^2, 2z)$. Por otro lado, para el punto $(0,1,0) \in \mathcal{C}$ se tiene que $f(0,1,0) = 1$, y además $f(x,y,z) = x^2+1 \geq f(0,1,0)$ para todo $(x,y,z) \in \mathcal{C}$. Es decir, f alcanza en $(0,1,0)$ un mínimo condicionado a $\mathcal{C}$. Sin embargo, $\nabla g_1(0,1,0) = (0,0,1)$ y $\nabla g_2(0,1,0) = (0,0,0)$ son linealmente dependientes.

Ejemplo 5.2.3. *Halle los puntos de la curva de intersección entre las superficies*

$$x^2 - xy + y^2 - z^2 = 1 \qquad \text{y} \qquad x^2 + y^2 = 1$$

que estén más próximos al origen.

En este caso, se nos pide minimizar el campo escalar $f(x,y,z) = \sqrt{x^2+y^2+z^2}$ condicionado a las curvas de nivel cero de los campos escalares

$$g_1(x,y,z) = x^2 - xy + y^2 - z^2 - 1 \qquad \text{y} \qquad g_2(x,y,z) = x^2 + y^2 - 1.$$

Para empezar, tenemos

$$\begin{aligned}
\nabla f(x,y,z) &= \left(\frac{x}{\sqrt{x^2+y^2+z^2}}, \frac{y}{\sqrt{x^2+y^2+z^2}}, \frac{z}{\sqrt{x^2+y^2+z^2}} \right) \\
&= \left(\frac{x}{f(x,y,z)}, \frac{y}{f(x,y,z)}, \frac{z}{f(x,y,z)} \right), \\
\nabla g_1(x,y,z) &= (2x-y, 2y-x, -2z), \\
\nabla g_2(x,y,z) &= (2x, 2y, 0).
\end{aligned}$$

Debemos resolver el sistema:

$$\left\{ \begin{array}{rl}
\nabla f(x,y,z) & = \lambda_1 \nabla g_1(x,y,z) + \lambda_2 \nabla g_2(x,y,z), \\
x^2 - xy + y^2 - z^2 & = 1, \\
x^2 + y^2 & = 1,
\end{array} \right.$$

es decir,

$$\left\{ \begin{array}{rl}
x & = (\lambda_1(2x-y) + 2\lambda_2 x) f(x,y,z), \\
y & = (\lambda_1(2y-x) + 2\lambda_2 y) f(x,y,z), \\
z & = -2\lambda_1 z f(x,y,z), \\
x^2 - xy + y^2 - z^2 & = 1, \\
x^2 + y^2 & = 1,
\end{array} \right.$$

Podemos empezar a resolver este sistema multiplicando la primera igualdad por y y la segunda por x, obteniendo:

$$\left\{ \begin{array}{rl}
xy & = (\lambda_1(2xy - y^2) + 2\lambda_2 xy) f(x,y,z), \\
xy & = (\lambda_1(2xy - x^2) + 2\lambda_2 xy) f(x,y,z).
\end{array} \right.$$

Luego,

$$\begin{aligned}(\lambda_1(2xy-y^2)+2\lambda_2xy)f(x,y,z) &= (\lambda_1(2xy-x^2)+2\lambda_2xy)f(x,y,z)\\ \lambda_1(2xy-y^2)+2\lambda_2xy &= \lambda_1(2xy-x^2)+2\lambda_2xy\\ 2\lambda_1xy-\lambda_1y^2+2\lambda_2xy &= 2\lambda_1xy-\lambda_1x^2+2\lambda_2xy\\ -\lambda_1y^2 &= -\lambda_1x^2\\ \lambda_1(x^2-y^2) &= 0.\end{aligned}$$

Note que hemos usado el hecho que $f(x,y,z)\neq 0$*, pues el punto* $(0,0,0)$ *no está en la intersección de las superficies de nivel consideradas.*

Ahora debemos considerar si $\lambda_1=0$ *o* $\lambda_1\neq 0$*. Supongamos primero que ocurre lo último. Entonces,* $x^2=y^2$*. Por otro lado, como* $x^2+y^2=1$*, tenemos que* $2x^2=1$*, es decir,* $x=\pm 1/\sqrt{2}$*. De esto se tiene que* $y=\pm 1/\sqrt{2}$*. Usamos ahora la igualdad* $x^2-xy+y^2-z^2=1$*, donde* $xy=\pm 1/2$*. Tenemos:*

$$\begin{aligned}x^2-xy+y^2-z^2 &= 1\\ 1-xy-z^2 &= 1\\ z^2 &= -xy,\end{aligned}$$

La última igualdad implica $xy=-1/2$*. Luego,* $z^2=1/2$*, de donde* $z=\pm 1/\sqrt{2}$*. Tenemos entonces que los puntos* $(\pm 1/\sqrt{2},\mp 1/\sqrt{2},\pm 1/\sqrt{2})$ *son solución del sistema anterior. Por otro lado,* $f(\pm 1/\sqrt{2},\mp 1/\sqrt{2},\pm 1/\sqrt{2})=\sqrt{1/2+1/2+1/2}=\sqrt{3/2}$*. Entonces, el valor* $\sqrt{3/2}$ *es un candidato a ser el mínimo de* f *condicionado a las superficies, pero debemos analizar el caso* $\lambda_1=0$ *para poder concluir.*

Supongamos ahora $\lambda_1=0$*. El sistema planteado al inicio se simplifica a:*

$$\begin{cases} x &= 2\lambda_2 x f(x,y,z),\\ y &= 2\lambda_2 y f(x,y,z),\\ z &= 0,\\ x^2-xy+y^2 &= 1,\\ x^2+y^2 &= 1.\end{cases}$$

De las dos últimas ecuaciones se tiene que $1-xy=1$*, de donde* $xy=0$*. Luego,* $x=0$ *o* $y=0$*. Para* $x=0$*, se tiene que* $0^2+y^2=1$*, de donde* $y=\pm 1$*. Tenemos entonces los puntos de solución*

$$P_1=(0,1,0) \qquad y \qquad P_2=(0,-1,0).$$

Para $y=0$*, se tiene que* $x=\pm 1$ *y así los puntos de solución*

$$P_3=(1,0,0) \qquad y \qquad P_4=(-1,0,0).$$

Notamos además que $f(P_k) = 1$ *para* $k = 1, 2, 3, 4$. *Como* $1 < \sqrt{3/2}$, *los puntos calculados en la parte anterior* $(\pm 1/\sqrt{2}, \mp 1/\sqrt{2}, \pm 1/\sqrt{2})$ *son descartados porque en ellos no se alcanza el valor mínimo condicionado. Por otro lado, para todo* $(x, y, z) \in \mathcal{C}$ *se tiene que*

$$f(x, y, z) = \sqrt{x^2 + y^2 + z^2} = \sqrt{1 + z^2} \geq 1 = f(P_k).$$

Por lo tanto, podemos concluir que la distancia mínima al origen de la curva de intersección entre $x^2 - xy + y^2 = 1$ *y* $x^2 + y^2 = 1$ *es 1, alcanzada en los puntos* $(0, \pm 1, 0)$ *y* $(\pm 1, 0, 0)$.

Escrito en LaTeX por **Marco A. Pérez**.

<u>Material consultado</u>:
- Cálculo Vectorial, de Jerrold E. Marsden y Anthony J. Tromba.
- Calculus Vol. 2, de Tom Apostol.

<u>Última actualización</u>: `14 de Septiembre de 2020`.