EXTREMOS CONDICIONADOS SIN MULTIPLICADORES DE LAGRANGE

Salvador Gigena
Universidades Nacionales de Rosario y de Córdoba – ARGENTINA
sgigena@fceia.unr.edu.ar , sgigena@efn.uncor.edu , sgigena@yahoo.com

Campo de Investigación: Pensamiento matemático avanzado; Nivel Educativo: Superior

Resumen: En este trabajo exponemos un método que permite determinar, y clasificar, los extremos locales condicionados de funciones diferenciables reales que, a la vez, prescinde totalmente del uso de los clásicos Multiplicadores de Lagrange. Tal método se basa en el uso adecuado de las herramientas clásicas del Análisis Matemático (Cálculo) a saber: Teorema de la Función Implícita, Regla de la Cadena (cálculo de las derivadas de funciones compuestas) y, ocasionalmente, Teorema o Fórmula de Taylor.

Introducción histórica

El *método de los Multiplicadores de Lagrange* representa uno de los más importantes mojones del Análisis Matemático, no sólo por su significación intrínseca e histórica, tanto en cuestiones teóricas como prácticas, si no también porque ha sido de uso casi excluyente cuando se trata de determinar y analizar extremos locales condicionados: (Apóstol, 1967; Fleming, 1965; Spring, 1985; & Williamson et al., 1972).
F. Zizza ha descripto dos métodos alternativos que eliminan a los multiplicadores de la prueba (test) de las primeras derivadas, o sea de aquella parte del cálculo en que se determinan los puntos críticos, y lo logra usando formas diferenciales. Sin embargo, él no indica un método alternativo para analizar los puntos críticos en cuanto a que sean máximos, mínimos o puntos de ensilladura locales: (Zizza, 1998).
Con respecto a esto último, i.e., al análisis de puntos críticos, D. Spring ha descripto como clasificar los extremos condicionados usando fuertemente los Multiplicadores de Lagrange: (Spring, 1985); mientras que en S. Gigena, M. Binia y D. Abud se los clasifica sin usar los multiplicadores, pero sólo en lo concerniente a la prueba (test) de las segundas derivadas, ya que aún se supone su utilización para calcular los puntos críticos: (Gigena et al., 2001) .
Es el propósito de este artículo presentar un método comprensivo para determinar y analizar extremos locales condicionados que, al mismo tiempo, representa una alternativa diferente a los tres artículos mencionados, así como a todos los trabajos anteriores sobre el tema: se excluye en tal método todo y cualquier uso de los Multiplicadores de Lagrange y/o de formas diferenciales. Los requerimientos teóricos del método son bastante sencillos: Teorema de la Función Implícita, Regla de la Cadena y, en aquellos casos en que falla la prueba de las segundas derivadas, la Fórmula de Taylor. Esto permitirá, sin duda, su fácil implementación futura en el trabajo áulico por parte tanto de profesores, en la instrucción, como de alumnos en la práctica receptiva.

Planteo y solución propuesta para el Problema

Se trata de implementar un método para determinar y clasificar extremos condicionados, que sea válido para todos los casos dimensionales, y co-dimensionales, posibles. En general tendremos una función $y = f(x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m})$, cuyos extremos deben ser determinados en el caso en que la función f está sujeta a condiciones subsidiarias dadas por un conjunto de ecuaciones $G_i(x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}) = 0$, donde $i = 1, 2, ..., m$. Suponemos que las funciones dadas están definidas con suficiente diferenciabilidad en un subconjunto abierto $U \subset \mathbb{R}^{n+m}$: $f : U \to \mathbb{R}$, $G : U \to \mathbb{R}^m$. Luego, el problema es determinar y clasificar los puntos críticos de f restringida al conjunto $S = \{X = (x_1, x_2, ..., x_{n+m}) : G(X) = 0, rango(G'(X)) = m\} \subset \mathbb{R}^{n+m}$ que es, en lenguaje matemático, una *variedad diferenciable inmersa*. Siendo la función definida por $G = (G_1, ..., G_m) : U \to \mathbb{R}^m$ diferenciable, con su matriz Jacobiana definida por

$$G'(X) = \begin{bmatrix} \dfrac{\partial G_1}{\partial x_1} & \dfrac{\partial G_1}{\partial x_2} & \cdots & \dfrac{\partial G_1}{\partial x_{n+m}} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial G_m}{\partial x_1} & \dfrac{\partial G_m}{\partial x_2} & \cdots & \dfrac{\partial G_m}{\partial x_{n+m}} \end{bmatrix}, \qquad (1)$$

y cuyo rango, como indicado, se supone igual a m, i.e., $G'(X)$ tiene m columnas linealmente independientes. Sea $X_0 = (a_1, ..., a_n, a_{n+1}, ..., a_{n+m}) \in S$ y supongamos por conveniencia, en primer lugar y sin pérdida de generalidad, que las m columnas linealmente independientes en $G'(X_0)$ son las últimas (si no fuera así, las variables pueden ser re-ordenadas): $\dfrac{\partial G}{\partial x_{n+1}}, \dfrac{\partial G}{\partial x_{n+2}}, \ldots, \dfrac{\partial G}{\partial x_{n+m}}$. Llamando $U_0 = (a_1, ..., a_n)$ y $V_0 = (a_{n+1}, ..., a_{n+m})$, el Teorema de la Función Implícita establece que existe una función diferenciable $h = (h_1, ..., h_m) : N \to \mathbb{R}^m$ definida en un entorno abierto $N \subset \mathbb{R}^n$, de U_0, tal que $h(U_0) = V_0$ y $H(u) = (u, h(u)) \in S,\ \forall u \in N$. Por lo tanto, tenemos que

$$S \cap (N \times \mathbb{R}^m) = \text{Grafico}(h) = \text{Imagen}(H).$$

Esto es, la función $H : N \to \mathbb{R}^{n+m}$ efectúa una parametrización de la variedad diferenciable S en un entorno de X_0.

Podemos observar fácilmente que la función (restringida/condicionada)

$$f\big|_S : S \to \mathbb{R}$$

alcanza una extremidad local (máximo, mínimo o punto de ensilladura en X_0 sí, y sólo sí, la función (no condicionada)

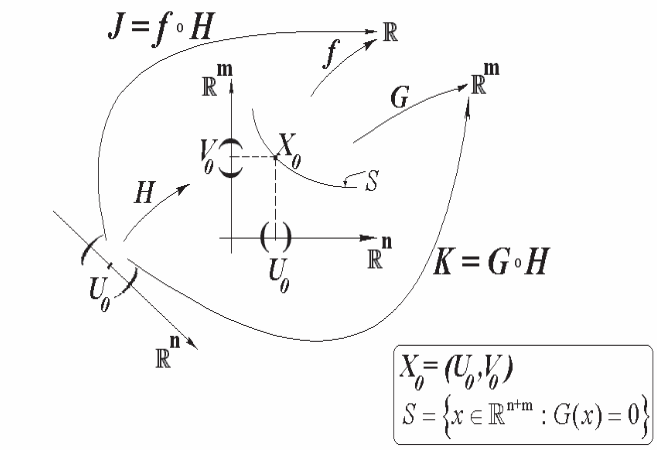
$$J := f \circ H : N \to \mathbb{R}$$

alcanza la misma clase de extremidad local en U_0.

En consecuencia, podemos reducir el problema original a la determinación y análisis de los posibles puntos críticos de la última función, siguiendo los pasos requeridos en el problema correspondiente para extremos libres, no condicionados.

Este último problema sería bastante fácil de resolver si conociéramos la expresión explícita de la función $H(u) = (u, h(u))$.. Sin embargo, como uno puede percibir en la mayoría de los ejemplos, resulta muy difícil, y a veces aún imposible, conseguir tal expresión. Así, en estos casos tenemos que computar las derivadas sucesivas de la función h implícitamente, usando el hecho de la función K obtenida por composición de G con H se anula idénticamente, i.e., $K := G \circ H = G \circ (Id, h) \equiv 0$

en un entorno abierto de $U_0 = (a_1,....,a_n)$, $N \subset \mathbb{R}^n$. Podemos representar toda la situación en el gráfico siguiente:



Recordemos que estamos suponiendo que tanto la función a valores reales f como la función a valores vectoriales G son suficientemente diferenciables. Por ejemplo, tenemos que requerir clase $C^{(2)}$ para desarrollar el test de las segundas derivadas. En tal caso, el gradiente y la matriz Hessiana de $J := f \circ H$ pueden ser calculados como se indica a seguir: denotamos $u = (u_1, u_2, \ldots, u_n)$ a las n primeras coordenadas del punto variable X que tiene coordenadas $X = (x_1,....,x_n, x_{n+1},....,x_{n+m})$ y sean $v = (v_1, v_2, \ldots, v_m)$ las últimas. Entonces, la función h está definida por : $h(u) = (v_1(u), v_2(u),..., v_m(u))$. Además, denotaremos con G'_v a la matriz cuya $i-e'sima$ fila consiste de las derivadas parciales de G_i respecto a las variables $v_1, v_2, ..., v_m$, i.e., en la expresión (1) de la matriz Jacobiana de G, tomamos la submatriz cuadrada formada por las últimas m columnas:

$$G'_v = \begin{bmatrix} \dfrac{\partial G_1}{\partial v_1} & \dfrac{\partial G_1}{\partial v_2} & \ldots & \dfrac{\partial G_1}{\partial v_m} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial G_m}{\partial v_1} & \dfrac{\partial G_m}{\partial v_2} & \ldots & \dfrac{\partial G_m}{\partial v_m} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial G_1}{\partial x_{n+1}} & \dfrac{\partial G_1}{\partial x_{n+2}} & \ldots & \dfrac{\partial G_1}{\partial x_{n+m}} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial G_m}{\partial x_{n+1}} & \dfrac{\partial G_m}{\partial x_{n+2}} & \ldots & \dfrac{\partial G_m}{\partial x_{n+m}} \end{bmatrix}. \qquad (2)$$

Así mismo, denotaremos con subíndices a las derivadas sucesivas de las funciones a valores reales y vectoriales, separando con una coma cuando necesario. Por ejemplo:

$$\frac{\partial f}{\partial x_i} := f_i \,;\ \frac{\partial^2 f}{\partial x_i \partial x_j} := f_{ij} \,;\ \frac{\partial G}{\partial x_i} := G_i \,;\ \frac{\partial^2 G}{\partial x_i \partial x_j} := G_{ij} \,;\ \frac{\partial G_{\alpha-n}}{\partial x_i} := G_{\alpha-n,i} \,;\ \frac{\partial^2 G_{\alpha-n}}{\partial x_i \partial x_j} := G_{\alpha-n,ij}$$

.

Los rangos de variación de los índices serán: letras latinas minúsculas indican variación del índice 1 al n, i.e., $1 \le i, j, k, l ... \le n$; letras griegas minúsculas van de $n+1$ a $n+m$, i.e., $n+1 \le \alpha, \beta, \gamma, \lambda, ... \le n+m$. Así, la matriz G'_v puede indistintamente ser escrita como

$$[G'_v] = \left(\frac{\partial G_{\alpha-n}}{\partial v_{\beta-n}} \right) = \left(\frac{\partial G_{\alpha-n}}{\partial x_{\beta}} \right) = \left(G_{\alpha-n,\beta} \right).$$

Y, puesto que es no-singular, denotaremos a su inversa por $\left[G'_v\right]^{-1} := \left(G^{\alpha\beta}\right)$.

Ahora, puesto que la función K se anula idénticamente en un conjunto abierto, i.e., $K(u) := G(u, h(u)) \equiv 0,\ \forall u \in N \subset \mathbb{R}^n$, también se anulan idénticamente todas las derivadas de esta función. En particular, para las primeras derivadas, respecto a u_i, obtenemos, usando también la Regla de la Cadena: $K_{\alpha-n,i} = G_{\alpha-n,i} + \sum_{\beta} G_{\alpha-n,\beta} h_{\beta-n,i} = 0$.

Entonces, obtenemos que: $h_{\gamma-n,i} = -\sum_{\alpha} G^{\gamma\alpha} G_{\alpha-n,i}$.

En forma similar, del hecho de que las segundas derivadas de K también se anulan :

$$K_{\alpha-n,ij} = G_{\alpha-n,ij} + \sum_{\mu} G_{\alpha-n,i\mu} h_{\mu-n,j} + \sum_{\beta} G_{\alpha-n,\beta j} h_{\beta-n,i} + \sum_{\beta,\nu} G_{\alpha-n,\beta\nu} h_{\beta-n,i} h_{\nu-n,j} + \sum_{\beta} G_{\alpha-n,\beta} h_{\beta-n,ij} = 0$$

también obtenemos

$$h_{\gamma-n,ij} = -\sum_{\alpha} G^{\gamma\alpha} G_{\alpha-n,ij} + \sum_{\alpha,\mu,\rho} G^{\gamma\alpha} G_{\alpha-n,i\mu} G^{\mu\rho} G_{\rho-n,j} + \sum_{\alpha,\beta,\rho} G^{\gamma\alpha} G_{\alpha-n,\beta j} G^{\beta\rho} G_{\rho-n,i} - \\ - \sum_{\alpha,\beta,\nu,\rho,\mu} G^{\gamma\alpha} G_{\alpha-n,\beta\nu} G^{\nu\rho} G_{\rho-n,j} G^{\beta\mu} G_{\mu-n,i} \ .$$

Así, para la función $J = f \circ H$ calculamos las primeras derivadas $\frac{\partial J}{\partial u_i}(u)$:

$$J_i = f_i + \sum_{\alpha} f_{\alpha} h_{\alpha-n,i} = f_i - \sum_{\alpha,\beta} f_{\alpha} G^{\alpha\beta} G_{\beta-n,i} \ .$$

A seguir, computamos las derivadas de esta última con respecto a la variable u_j, obteniendo a partir de las expresiones anteriores, y usando también nuevamente la Regla de la Cadena, la ecuación que representa cada una de las componentes de la Matriz Hessiana de la función J:

$$J_{ij} = f_{ij} - \sum_{\alpha,\beta} f_{i\alpha} G^{\alpha\beta} G_{\beta-n,j} - \sum_{\alpha,\beta} f_{\alpha j} G^{\alpha\beta} G_{\beta-n,i} + \sum_{\alpha,\beta,\mu,\nu} f_{\alpha,\beta} G^{\alpha\mu} G_{\mu-n,i} G^{\beta\nu} G_{\nu-n,j} + \\ + \sum_{\gamma} f_{\gamma} \left(\begin{array}{l} -\sum_{\alpha} G^{\gamma\alpha} G_{\alpha-n,ij} + \sum_{\alpha,\mu,\rho} G^{\gamma\alpha} G_{\alpha-n,i\mu} G^{\mu\rho} G_{\rho-n,j} + \sum_{\alpha,\beta,\rho} G^{\gamma\alpha} G_{\alpha-n,\beta j} G^{\beta\rho} G_{\rho-n,j} \\ - \sum_{\alpha,\beta,\nu,\rho,\mu} G^{\gamma\alpha} G_{\alpha-n,\beta\nu} G^{\nu\rho} G_{\rho-n,j} G^{\beta\mu} G_{\mu-n,i} \end{array} \right).$$

En consecuencia, para determinar los puntos críticos, tenemos que resolver el siguiente sistema de ecuaciones:

$$\left. \begin{array}{r} J_i = 0,\ i = 1, 2, \ldots, n \\ G_{\alpha-n} = 0,\ \alpha - n = 1, 2, \ldots, m \end{array} \right\}$$

Observemos que éste es un sistema de $n+m$ ecuaciones en las $n+m$ incógnitas representadas por las componentes del punto variable $X=(x_1,....,x_n,x_{n+1},....,x_{n+m})$.
A seguir, para clasificar los puntos críticos encontrados, podemos proceder a analizar la Hessiana $Hess(J):=\left(J_{ij}\right)$, cuyas componentes J_{ij} están descriptas más arriba en términos de las derivadas de las funciones dadas f y G, hasta el segundo orden
Finalmente, en caso de que el test de las segundas derivadas falle, i.e., que no permita obtener la clasificación total, al menos en algunos puntos, podemos proceder a calcular las derivadas parciales de orden superior de la función J y, usando además adecuadamente la Fórmula de Taylor, obtener la deseada clasificación de todos los puntos críticos.

Observaciones:

1) Al comienzo de nuestro argumento supusimos que las m columnas linealmente independientes de la matriz $G'(X_0)$ eran las últimas y que si así no fuera las podríamos re-ordenar. Puesto que puede existir en dicha matriz más de un conjunto de m columnas linealmente independientes será necesario, como es obvio, estudiar todos los otros posibles casos a fin de determinar puntos críticos <u>adicionales</u> de la función f.
2) Como hemos expuesto, el método presentado aquí prescinde del uso de los Multiplicadores de Lagrange para la determinación y clasificación de extremos condicionados. Sin embargo, para ciertas cuestiones de aplicación, como por ejemplo en Economía, puede ser deseable, y aún necesario, la obtención de los valores concretos de los mismos. Pues bien, vamos a demostrar a que los históricos multiplicadores, que no son ya parte del problema, también pueden ser obtenidos como parte de la solución del método aquí expuesto, sin prácticamente esfuerzo significativo adicional En efecto, supongamos haber calculado los puntos críticos y sea X_0 uno de ellos. Luego el Teorema de los Multiplicadores de Lagrange asegura que existen $\lambda_1,....,\lambda_m$ tales que $\left(f+\lambda_1 G_1+\cdots+\lambda_1 G_1\right)'(X_0)=(0,....,0)$. Puesto que las coordenadas de X_0 son ya conocidas, la última condición representa un sistema lineal de $n+m$ ecuaciones en las m incógnitas $\lambda_1,....,\lambda_m$. Sin embargo, por la suposición hecha en el desarrollo del método, con referencia a tomar las últimas m columnas de $G'(X_0)$ como linealmente independientes y formar la submatriz cuadrada representada en la ecuación (2), podemos a la vez reducir tal sistema, al de m ecuaciones con m incógnitas dado por:

$$\begin{bmatrix}\lambda_1 & \cdots & \lambda_m\end{bmatrix}\begin{bmatrix}\dfrac{\partial G_1}{\partial v_1} & \dfrac{\partial G_1}{\partial v_2} & \cdots & \dfrac{\partial G_1}{\partial v_m}\\ \vdots & \vdots & & \vdots\\ \dfrac{\partial G_m}{\partial v_1} & \dfrac{\partial G_m}{\partial v_2} & \cdots & \dfrac{\partial G_m}{\partial v_m}\end{bmatrix}=-\begin{bmatrix}\dfrac{\partial f}{\partial v_1} & \cdots & \dfrac{\partial f}{\partial v_m}\end{bmatrix},$$

donde todas las derivadas están evaluadas en X_0. Por lo tanto, como solución de tal sistema reducido, los multiplicadores se representan a través de la expresión

$$\lambda_{\beta-n}=-\sum_{\alpha} f_{\alpha-n}\left(X_0\right)G^{\alpha\beta}\left(X_0\right).$$

Comparación con los otros métodos

1) F. Zizza ha descrito dos métodos alternativos que eliminan a los multiplicadores de la prueba (test) de las primeras derivadas, o sea de aquella parte del cálculo en que se determinan los puntos críticos, y lo logra usando formas diferenciales. Sin embargo, él no indica un método alternativo para analizar los puntos críticos con respecto a la posibilidad de que sean máximos, mínimos o puntos de ensilladura locales: (Zizza, 1998).
2) Con respecto al análisis de puntos críticos, D. Spring ha descrito como clasificar los extremos condicionados usando fuertemente los Multiplicadores de Lagrange: (Spring, 1985); mientras que S. Gigena, M. Binia y D. Abud los clasifican sin usar los multiplicadores, pero sólo en lo concerniente a la prueba (test) de las segundas derivadas, ya que aún se supone su utilización para calcular los puntos críticos: (Gigena et al., 2001) .
3) En cuanto a los tiempos de computación F. Zizza ha determinado, experimentalmente usando Mathematika, que su método para el test de las primeras derivadas es más rápido que el que usa los multiplicadores (15 a 20 % de este último). El método aquí presentado es, en tal aspecto, prácticamente igual al de Zizza.
4) Además, y como dijéramos, en caso que el test de las segundas derivadas falle en lo concerniente a la clasificación, siempre el método nos permite recurrir a las derivadas superiores.

Referencias bibliográficas

Apostol, T. (1967). *Calculus. Volumen II*. Waltham, Massachussetts, USA: Blaisdell Publishing Company.

Fleming, W. (1965). *Functions of Several Variables*. Reading, Massachussetts, USA: Addison-Weslley Publishing Company.

Gigena, S., Binia, M., y Abud, D. (2001). Extremos Condicionados: Una propuesta metodológica para su resolución. *Revista de Educación Matemática 16* (3), pp. 31-53.

Spring, D. (1985). On the Second Derivative Test for Constrained Local Extrema. *American Mathematical Monthly 34*, pp. 631-643.

Williamson, R. Crowell, R. y Trotter, H. (1972). *Calculus of Vector Functions*, *3rd edition*. Englewood Cliffs, New Jersey, USA: Prentice Hall Inc..

Zizza, F. (1998). Differential Forms for Constrained Max-Min Problems: Eliminating Lagrange Multipliers. *The College Math Journal 29* (5), pp. 387-396.