Condiciones Necesarias y Suficientes en Cálculo Variacional

Juliana Cervelati y Marko A. Rojas-Medar *

Departamento de Matemática Aplicada, Universidade Estadual de Campinas, CP 6065, 13083-970, Campinas-SP, Brasil

ju@ime.unicamp.br (J. Cervelati), marko@ime.unicamp.br (M.A. Rojas-Medar)

Resumen

En los últimos años surgió en la literatura asociada a la programación matemática la noción de función invex y sus generalizaciones, las cuales se han mostrado muy eficientes en la obtención de condiciones suficientes de optimalidad. Nuestro objetivo esencial es mostrar sus usos en problemas de cálculo variacional.

Palabras clave : *Cálculo variacional, funciones invexas, condiciones de optimalidad.*

1. Introducción

El Cálculo Variacional surgió alrededor de 1696 con el problema de la baquistrócrona. A partir de ahí, vários resultados fueron encontrados por algunos de los mayores matemáticos de los últimos 300 anõs, como por ejemplo: Euler, Lagrange, Legendre, Bolza, Hamilton, Bliss, Weierstrass y Jacobi. Su estudio fue muy importante no sólo para la Matemática, sino también para otras áreas del saber, como por ejemplo, Física, Ingeniería, Biología y Economía.

El Cálculo Variacional estudia los métodos que nos permiten encontrar los valores máximos y mínimos de funcionales.

En este trabajo introducimos algunos conceptos de invexidad y estudiamos condiciones necesarias y suficientes para problemas de cálculo variacional en el caso invex.

* Parcialmente financiado por el Ministerio De Ciencia y Tecnología, España, proyecto BFM 2003-06579 and CNPq-Brasil, proyecto 301354/03-0

2. Condiciones de optimalidad para problemas de Cálculo Variacional en el caso invexo

En esta sección haremos un análisis del problema general del Cálculo Variacional y presentaremos algunos resultados relativos a condiciones necesarias y suficientes de optimalidad para tales problemas en el caso invexo.

Primero daremos algunas definiciones preliminares.

Sea $I = [t_0, t_1]$ un intervalo en la recta, $f : \mathbb{R}^n \times \mathbb{R}^n \times I \to \mathbb{R}$ y $g : \mathbb{R}^n \times \mathbb{R}^n \times I \to \mathbb{R}^m$ funciones continuamente diferenciables.

Considere $f(x(t), \dot{x}(t), t)$ donde $x : I \to \mathbb{R}^n$ es diferenciable con derivada $\dot{x}(t)$, $X = C(I, \mathbb{R}^n)$ el espacio de las funciones $x : I \to \mathbb{R}^n$ suaves por partes, y el funcional $F : X \to \mathbb{R}$ definido por $F(x) = \displaystyle\int_{t_0}^{t_1} f(x(t), \dot{x}(t), t)dt$.

Con la finalidad de simplificar la notación usaremos $f(x, \dot{x}, t)$ en vez de $f(x(t), \dot{x}(t), t)$.

La definición de invexidad de una función f introducida por Hanson [12] es extendida para un funcional F por Mond y Hussain [14], y Mond y Smart [15].

Definición 1 *El funcional F se dice invex con relación a η si existe una función vetorial diferenciable $\eta(x, \bar{x}, t)$ con $\eta(x, x, t) = 0$ tal que para todo $x \in X$*

$$F(x) - F(\bar{x}) \geq \int_{t_0}^{t_1} \{\eta(x, \bar{x}, t) f_x(\bar{x}, \dot{\bar{x}}, t) + (\frac{d\eta}{dt}(x, \bar{x}, t)) f_{\dot{x}}(\bar{x}, \dot{\bar{x}}, t)\} dt$$

Definición 2 *El funcional F se dice pseudoinvexo con relación a η si existe una función vectorial diferenciable $\eta(x, \bar{x}, t)$ con $\eta(x, x, t) = 0$ tal que para todo $x \in X$*

$$\int_{t_0}^{t_1} \{\eta(x, \bar{x}, t) f_x(\bar{x}, \dot{\bar{x}}, t) + (\frac{d\eta}{dt}(x, \bar{x}, t)) f_{\dot{x}}(\bar{x}, \dot{\bar{x}}, t)\} dt \geq 0 \Rightarrow F(x) - F(\bar{x}) \geq 0$$

o equivalentemente

$$F(x) - F(\bar{x}) < 0 \Rightarrow \int_{t_0}^{t_1} \{\eta(x, \bar{x}, t) f_x(\bar{x}, \dot{\bar{x}}, t) + (\frac{d\eta}{dt}(x, \bar{x}, t)) f_{\dot{x}}(\bar{x}, \dot{\bar{x}}, t)\} dt < 0$$

Donde $\dfrac{d\eta}{dt}$ es un vector cuya i-ésima componente viene dada por $(\dfrac{d}{dt})\eta_i(x, \bar{x}, t)$.

Es fácil verificar que si f no depende de t entonces las definiciones anteriores se reducen a las definiciones de función invex y pseudoinvex respectivamente como en [5] y [12]. Vale la pena destacar que invexidad y pseudoinvexidad son la generalización de los conceptos de convexidad y pseudoconvexidad respectivamente.

Considere el problema de Cálculo Variacional siguiente:

$$\left.\begin{array}{ll} \text{Minimizar} & F(x) = \displaystyle\int_{t_0}^{t_1} f(x,\dot{x},t)dt \\ & \\ \text{sujeto a} & g(x,\dot{x},t) \leq 0 \quad t \in [t_0,t_1] \\ & x(t_0) = x_0, \;\; x(t_1) = x_1 \\ & x \in X \end{array}\right\} \quad (PV)$$

Sea H el conjunto de las soluciones factibles del problema anterior, es decir:

$$H = \{x \in X \mid g(x,\dot{x},t) \leq 0, t \in [t_0,t_1] \;\; y \;\; x(t_0) = x_0, \;\; x(t_1) = x_1\}$$

La función Lagrangiana asociada a (PV) es

$$L(x,\dot{x},\lambda,y,t) = \lambda F + \sum_{i=1}^{m} y_i(t) g_i(x,\dot{x},t)$$

o en notación vectorial

$$L(x,\dot{x},\lambda,y,t) = \lambda F + y(t)^T g(x,\dot{x},t).$$

Cuando $\lambda = 0$ decimos que el problema es anormal y cuando $\lambda \neq 0$ normal (vea [4]). En el caso normal es suficiente considerar $\lambda = 1$. Resaltamos que en el caso anormal las condiciones usuales de primer orden tienen poca utilidad, pues en este caso el funcional objectivo, F, no interfiere. Son necesarias otras herramientas para estudiar este caso (vea [2]). Cuando el problema es normal, escribimos

$$L(x,\dot{x},\lambda,y,t) \equiv L(x,y).$$

Definición 3 *$\bar{x} \in H$ es una solución óptima, o mínimo global de (PV) si*

$$F(\bar{x}) \leq F(x) \quad \forall x \in H$$

.

Ahora será enunciado un resultado que nos da una condición necesaria de optimalidad para problemas de Cálculo Variacional.

Definición 4 *$x \in H$ es llamado Punto Crítico tipo Kuhn-Tucker si existe una función suave por partes, $y : I \to \mathbb{R}^n$, tal que:*

$$\begin{aligned} f_x(x,\dot{x},t) + y(t)^T g_x(x,\dot{x},t) &= \frac{d}{dt}\{f_{\dot{x}}(x,\dot{x},t) + y(t)^T g_{\dot{x}}(x,\dot{x},t)\} && (1) \\ y(t)^T g(x,\dot{x},t) &= 0 && (2) \\ y(t) &\geq 0 \quad t \in [t_0,t_1] && (3) \end{aligned}$$

En otras palabras un punto crítico tipo Kuhn-Tucker es un punto que satisface las ecuaciones de Euler-Lagrange asociada a la Lagrangiana para problemas normales.

En el problema de Cálculo Variacional con restricciones, la condición de Punto Crítico Tipo Kuhn-Tucker en el punto $\bar{x}$ es una condición necesaria para que $\bar{x}$ sea un óptimo global. El siguiente resultado puede encontrarse en [8], [11].

Teorema 1 *Si $\bar{x}$ es una solución óptima normal (es decir, estamos tratando un problema normal) para (PV), entonces $\bar{x}$ es un Punto Crítico Tipo Kuhn-Tucker.*

Es conveniente recordar que en la teoría clásica de optimización, las condiciones de Kuhn-Tucker son también suficientes, cuando las funciones involucradas son convexas.

Desde el surguimiento de la noción de invexidad en 1980 como una generalización del concepto de convexidad, muchos resultados han sido obtenidos con la finalidad de aumentar la clase de funciones para las cuales estas condiciones sean también suficientes, como podemos constatar en [1], [5], [12], [14], [15], [16] entre otros.

En [14] Mond y Hussain enunciarón un resultado que nos daba condiciones suficientes de optimalidad para el problema del Cálculo Variacional, en el cual solamente la función Lagrangeana asociada al problema tendría que ser pseudoinvex.

Teorema 2 *Sea $\bar{x} \in H$. Si existe una función suave por partes $\bar{y} : I \to \mathbb{R}^n$ tal que $(\bar{x}(t), \bar{y}(t))$ sea un Punto Crítico Tipo Kuhn-Tucker, y si la función Lagrangiana*

$$L(x, \bar{y}) = \int_{t_0}^{t_1} f(x, \dot{x}, t) + \bar{y}(t) g(x, \dot{x}, t) dt$$

es pseudoinvex en $\bar{x}$ con relación al η, entonces $\bar{x}$ es una solución óptima para (PV).

Arana-Jiménez y alg. [1], definierón L-KT pseudoinvexidad, y, a partir de ahíencontraron resultados que además de necesarios son también suficientes.

Considere el problema (PV) y las funciones f, g, donde $F(x) = \int_{t_0}^{t_1} f(x, \dot{x}, t) dt$ y $G(x) = \int_{t_0}^{t_1} g(x, \dot{x}, t) dt$.

Definición 5 *El par (F, G) es L-KT-pseudoinvex en $\bar{x} \in X$, si para todo $x \in X$, $\bar{y} : I \to \mathbb{R}^m$ es una función suave por partes, que verifica la Condición de Kuhn-Tucker (1)-(3), entonces existe una función vectorial diferenciable $\eta(x, \bar{x}, \bar{y}, t)$, con $\eta(x, \bar{x}, \bar{y}, t_0) = 0 = \eta(x, \bar{x}, \bar{y}, t_1)$, tal que*

$$F(x) - F(\bar{x}) < 0 \Rightarrow \int_{t_0}^{t_1} \{(f_x(x, \dot{\bar{x}}, t) + \bar{y}(t)^T g_x(x, \dot{\bar{x}}, t))\eta(x, \bar{x}, \bar{y}, t) +$$

$$+ (f_{\dot{x}}(x, \dot{\bar{x}}, t) + \bar{y}(t)^T g_{\dot{x}}(x, \dot{\bar{x}}, t)) \frac{d\eta}{dt}(x, \bar{x}, \bar{y}, t)\} dt < 0$$

o equivalentemente

$$\int_{t_0}^{t_1} \{(f_x(x,\dot{\bar{x}},t) + \bar{y}(t)^T g_x(x,\dot{\bar{x}},t))\eta(x,\bar{x},\bar{y},t) +$$

$$+(f_{\dot{x}}(x,\dot{\bar{x}},t) + \bar{y}(t)^T g_{\dot{x}}(x,\dot{\bar{x}},t))\frac{d\eta}{dt}(x,\bar{x},\bar{y},t)\}dt \geq 0 \Rightarrow F(x) - F(\bar{x}) \geq 0$$

El problema (PV) es L-KT-pseudoinvex si lo es para $x, \bar{x} \in H$.

Note que dados $x, \bar{x} \in X$, existe $\bar{y}$ que verifica la Condición de Kuhn-Tucker (1)-(3). Por ejemplo, considere $\bar{y}(t) = 0, t \in I$.

Definición 6 *El Problema Variacional con restricciones (PV) se dice L-KT pseudoinvex si la Definición 2 es verificada para $x, \bar{x} \in H$.*

Proposición 3 *Si (F,G) es pseudoinvex en $x, \bar{x} \in H$, entonces F es pseudoinvex en $\bar{x}$.*

En el mismo trabajo, Arana-Jiménez y alg. reenunciarón el Teorema 2 para funciones del tipo L-KT pseudoinvexas.

Teorema 4 *Sea $\bar{x} \in H$. Si para todo $\bar{y} : I \to \mathbb{R}^n$, existe una función suave por partes tal que $(\bar{x}, \bar{y})$ verifique (1)-(3), la función Lagrangiana $L(x, \bar{y}) = \int_{t_0}^{t_1} f(x,\dot{x},t) + \bar{y}(t)g(x,\dot{x},t)dt$ com $x \in H$, es pseudoinvex en $\bar{x}$ con relación a η, entonces (PV) es L-KT pseudoinvex.*

A partir de este resultado, se probó que la L-KT pseudoinvexidad de (PV) es una condición necesaria y suficiente para optimalidad de (PV). Primero recordaremos el siguiente lema.

Lema 5 *Sean w_1, w_2 funciones vectoriales continuas por partes tal que $\int_a^b (w_1^T(t)\psi(t) + w_2^T(t)\dot{\psi}(t))dt = 0$, $\forall \psi$ diferenciable, $\psi(a) = 0 = \psi(b)$, entonces w_2 es suave por partes y $Dw_2 = w_1$ en I.*

Teorema 6 *Si todos los Puntos Críticos Tipo Kuhn-Tucker son soluciones óptimas para (PV), entonces (PV) es L-KT pseudoinvex.*

Demostración. Considere $x, \bar{x} \in H$, $(\bar{x}, \bar{y})$ verificando (2) y (3), tal que $F(x) - F(\bar{x}) < 0$. Debemos encontrar $\eta(x,\bar{x},\bar{y},t)$ diferenciable, con $\eta(x,\bar{x},\bar{y},t_0) = 0 = \eta(x,\bar{x},\bar{y},t_1)$, tal que

$$P(\eta(x,\bar{x},\bar{y},\cdot)) \equiv \int_{t_0}^{t_1} \{\ (f_x(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_x(t,\bar{x},\dot{\bar{x}}))\eta(x,\bar{x},\bar{y},t)$$

$$+(f_{\dot{x}}(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_{\dot{x}}(t,\bar{x},\dot{\bar{x}}))\frac{d}{dt}\eta(x,\bar{x},\bar{y},t)\ \}\,dt < 0$$

Supongamos que $P(\eta(x,\bar{x},\bar{y},\cdot)) < 0$ no tenga solución $\eta(x,\bar{x},\bar{y},t)$ y entonces $P(\eta(x,\bar{x},\bar{y},\cdot)) > 0$ tampoco tiene solución, dado que podríamos considerar $-\eta(x,\bar{x},\bar{y},t)$. Además,

$$P(\eta(x,\bar{x},\bar{y},\cdot)) = \int_{t_0}^{t_1} \{ \; (f_x(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_x(t,\bar{x},\dot{\bar{x}}))\eta(x,\bar{x},\bar{y},t)$$
$$+(f_{\dot{x}}(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_{\dot{x}}(t,\bar{x},\dot{\bar{x}}))\frac{d}{dt}\eta(x,\bar{x},\bar{y},t) \; \} \, dt = 0$$

$\forall \eta(x,\bar{x},\bar{y},t)$ diferenciable, con $\eta(x,\bar{x},\bar{y},t_0) = 0 = \eta(x,\bar{x},\bar{y},t_1)$. El Lema 5, implica que

$$f_{\dot{x}}(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_{\dot{x}}(t,\bar{x},\dot{\bar{x}})$$

es suave por partes y

$$f_x(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_x(t,\bar{x},\dot{\bar{x}}) = \frac{d}{dt}\left\{ f_{\dot{x}}(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_{\dot{x}}(t,\bar{x},\dot{\bar{x}}) \right\}$$

y además, $(\bar{x},\bar{y})$ verifica (1), (2) and (3), $\bar{x}$ es un punto crítico de Kuhn-Tucker, y entonces $\bar{x}$ es una solución óptima para (PV), lo cual es contradictorio $F(x) - F(\bar{x}) < 0$. Así, existe $\eta(x,\bar{x},\bar{y},t)$ diferenciable, con $\eta(x,\bar{x},\bar{y},t_0) = 0 = \eta(x,\bar{x},\bar{y},t_1)$, tal que $P(\eta(x,\bar{x},\bar{y},\cdot)) < 0$, por lo tanto, (PV) es L-KT-pseudoinvex. $\square$

Teorema 7 *Si (PV) es L-KT pseudoinvex, entonces todos los Puntos Críticos Tipo Kuhn-Tucker son soluciones óptimas para (PV).*

Demostración. Sea $\bar{x} \in H$ un punto crítico de Kuhn-Tucker, es decir, existe una función suave por partes $\bar{y}$ tal que $(\bar{x},\bar{y})$ satisface (1), (2) y (3). Sea $x \in H$. dado que (PV) es L-KT-pseudoinvex, existe $\eta(x,\bar{x},\bar{y},t)$ diferenciable, con $\eta(x,\bar{x},\bar{y},t_0) = 0 = \eta(x,\bar{x},\bar{y},t_1)$ verificando la definición. Tenemos,

$$\begin{aligned}
&\int_{t_0}^{t_1} \{(f_x(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_x(t,\bar{x},\dot{\bar{x}}))\eta(x,\bar{x},\bar{y},t) \\
&\qquad + (f_{\dot{x}}(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_{\dot{x}}(t,\bar{x},\dot{\bar{x}}))\frac{d}{dt}\eta(x,\bar{x},\bar{y},t)\}dt \\
&\quad = \int_{t_0}^{t_1} \left\{(f_x(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_x(t,\bar{x},\dot{\bar{x}}))\eta(x,\bar{x},\bar{y},t)\right. \\
&\qquad \left. -\frac{d}{dt}\left(f_{\dot{x}}(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_{\dot{x}}(t,\bar{x},\dot{\bar{x}})\right)\eta(x,\bar{x},\bar{y},t)\right\} dt \\
&\qquad + (f_x(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_x(t,\bar{x},\dot{\bar{x}}))\eta(x,\bar{x},\bar{y},t)|_{t=t_0}^{t=t_1} \\
&\qquad \text{(por integración por partes)} \\
&\quad = \int_{t_0}^{t_1} \left\{f_x(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_x(t,\bar{x},\dot{\bar{x}})\right. \\
&\qquad \left. -\frac{d}{dt}\left(f_{\dot{x}}(t,\bar{x},\dot{\bar{x}}) + \bar{y}(t)^T g_{\dot{x}}(t,\bar{x},\dot{\bar{x}})\right)\right\} \eta(x,\bar{x},\bar{y},t)dt = 0 \\
&\qquad \text{(by (1))}
\end{aligned}$$

Como (PV) es L-KT-pseudoinvex deducimos que $F(x) - F(\bar{x}) \geq 0$, es decir, $F(x) \geq F(\bar{x})$, $\forall x \in H$. Por lo tanto, $\bar{x}$ es una solución óptima de (PV). $\square$

Nota 1 *En este trabajo hemos dado una caracterización completa para la determinación de soluciones óptimas de un problema de cálculo variacional. Vimos que la KT-pseudoinvexidad es una condición necesaria y suficiente para saber si un punto es solución óptima. Hemos extendido estos resultados a problemas de control óptimo, pero esto será dado en otro trabajo.*

Referencias

[1] Arana-Jiménez, M., Osuna-Gómez, R., Ruiz-Garzón, G., Rojas-Medar, M.A., Characterization of solutions in variational problems. Duality. (sometido a publicación)

[2] Avakov, E. R., First-order necessary conditions for abnormal problems in the calculus of variations, Differential Equations, 27, (1991), 495-500.

[3] Barbolla, R., Cerdá, E., Sanz, P., Optimización. Cuestiones, ejercicios y aplicaciones a la economía. Prentice Hall, Madrid, 2000.

[4] Bliss, G. A., Lectures on the Calculus of variations, Univ. Chicago Press, Chicago, 1946.

[5] Brandão, A.J., Rojas-Medar, M.A., Silva, G.N., Uma introdução às funções invexas diferenciáveis com aplicações em otimização. Bol. Soc. Paran. de Matemática, Vol.19, n.1-2, (1999) pp.51-65.

[6] Ben-Israel, A., Mond, B., What´s Invexity?, J. Austral. Math. Soc. Ser. B, n.28, (1986) pp.1-9.

[7] Cesari, L., Optimization - Theory and Applications, Problems with ordinary differential equations. Springer-Verlag, New York, 1983.

[8] Cerdá, E., Optimización dinámica. Prentice Hall, Madrid, 2001.

[9] Elsgoltz, L., Ecuaciones diferenciales y cálculo variacional. Mir-Moscú, 1969.

[10] Giorgi, G., A note on the relashionships between convexity and invexity. J. Austral. Math. Soc. Ser. B, n.32, (1990) pp.97-99.

[11] Gregory, J., Lin, C., Constrained optimization in the calculus os variations and optimal control theory. Chapman & Hall, London, 1996.

[12] Hanson, M.A., On sufficiency of the Kuhn-Tucker conditions. J. Math. Anal. Appl., n.80, (1981) pp.545-550.

[13] Mangasarian, O.L., Nonlinear Programming. Classics in Applied Mathematics, SIAM, n.10, 1994.

[14] Mond, B., Husain, I., Sufficient optimality criteria and duality for variational problems with generalised invexity. J. Austral. Math. Soc. Ser. B, n.31, (1989) pp.108-121.

[15] Mond, B., Smart, I., Duality and sufficiency in control problems with invexity. J. Math. Anal. Appl., n.136, (1988) pp.325-333.

[16] Moretti, A.C., Rojas-Medar, M.A., Condiciones suficientes de optimalidad em programación no lineal. Cubo Matem. Educac., Vol.3, n.2, (2001) pp.129-146.