TECNIA, Vol.13 Nº 1, págs. 83 - 90, 2003
Universidad Nacional de Ingeniería
Lima - Perú

UNA PRUEBA GENERAL DE LA BUENA DEFINICIÓN DEL MÉTODO LAGRANGEANO AUMENTADO

Yna Consuelo Rezza Espinoza
Facultad de Ingeniería Económica y Ciencias Sociales
Escuela Profesional de Ingeniería Estadística
E-mail: yna_c@yahoo.com

RESUMEN

En este trabajo presentamos una demostración de la buena definición del Método Lagrangeano Aumentado con Penalidades $P_i \in \mathcal{P}$ la cual incluye casi todos los casos existentes en la literatura. Las penalidades que describimos consideran dos subfamilias; una, continuamente diferenciable y estrictamente convexa (entre otras características) y la otra, continuamente diferenciable, estrictamente convexa en [-b, +∞) para algún b>0 y constante en (- ∞, -b].Las pruebas de buena definición realizadas por Rockafellar [10], Bertsekas [1], Polyak y Teboulle [7] y Gonzaga y Castillo [2] vendrían a ser casos particulares.

ABSTRACT

We present a proof of the well definition of the Aumented Langrangian Method with Penalties $P_i \in P$ that includes almost all cases existing in the literature. The penalties we describe consider two subfamilies: one, continuously differentiable and strictly convex (among other characteristics) and the other one, continuously differentiable, strictly convex in [-b, +∞) for some b>0 and constant in (-∞, -b]. The tests of good definition made by Rockafellar [10], Bertsekas [1], Polyak and Teboulle [7] and Gonzaga y Castillo [2] would be particular cases.

INTRODUCCIÓN

Problemas no lineales de optimización surgen espontáneamente en muchos campos de aplicación. Uno de los más estudiados es el siguiente:

$$\begin{array}{ll} \text{minimizar } f(x) & \\ \text{s.a. } g_i \leq 0, & i = 1, \ldots, m \\ x \in \mathbb{R}^n & \end{array}$$

donde $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, $g_i : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ para $i = 1, \ldots, m$ son funciones convexas y cerradas.

Para resolver este problema de optimización convexa apareció a fines de la década de los 70 un método que se ha venido aplicando con éxito hasta el día de hoy; estamos refiriéndonos al Método Lagrangeano Aumentado.

Este método como muchos otros métodos, se basa en la sustitución del problema inicial por una secuencia de subproblemas, el primer objetivo a alcanzar será garantizar la buena definición de la sucesión formada por alguna solución de cada subproblema. Rockafellar [10], demostró esto mediante el uso de penalidades estándar, Bertsekas [1] la garantizó para penalidades de tipo exponencial.

Otras pruebas de la buena definición del Método Lagrangeano Aumentado se encuentran en [7] y [2].

La prueba de buena definición que presentamos a continuación, incluye todos estos casos.

Nosotros definimos las penalidades $P_i \in \mathcal{P}$ considerando dos subfamilias, una continuamente diferenciable y estrictamente convexa y la otra, continuamente diferenciable, estrictamente convexa en [-b, +∞) para algún b>0 y constante en (- ∞, -b].

LA BUENA DEFINICIÓN DEL MÉTODO LAGRANGEANO AUMENTADO

Definición 2.1.- Se dice que la función $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ no idénticamente $+\infty$ es convexa si para todo $(x, x') \in \mathbb{R}^n \times \mathbb{R}^n$ y todo $\alpha \in (0, 1)$ tenemos:

$$f(\alpha x + (1-\alpha)x') \leq \alpha f(x) + (1-\alpha) f(x')$$

considerando esta desigualdad en $\mathbb{R} \cup \{+\infty\}$.
Esta clase de funciones será denotada así $f \varepsilon$ Conv $\mathbb{R}^n$.

Definición 2.2.- La función $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ se dice cerrada, si para cada $x \in \mathbb{R}^n$ se verifica $\lim_{y \to x} \inf f(y) \geq f(x)$.

Si $f \in$ Conv $\mathbb{R}^n$ y además f es cerrada, denotaremos $f \in \overline{\text{Conv}}\ \mathbb{R}^n$.

Consideremos el siguiente problema de optimización convexa:

$$\begin{array}{ll} \text{minimizar } f(x) & \\ \text{s.a. } g_i \leq 0,\ i = 1, \ldots, m & \qquad (P) \end{array}$$

donde $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $g_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ para $i = 1, \ldots, m$ son funciones convexas y cerradas.

Hipótesis 2.3.- Supongamos que el conjunto de soluciones óptimas del problema (P) es no vacío y limitado.

Hipótesis 2.4.- (Condición de Calificación de Slater). Supongamos que existe $\bar{x} \in \mathbb{R}^n$ tal que

$$g_i(\bar{x}) < 0 \text{ para todo } i = 1, \ldots, m.$$

Definición 2.5.- Dado $b > 0$ (posiblemente $+\infty$), denotemos por $\mathcal{P}$ a la familia de funciones $P: \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}$ o $P: \mathbb{R} \times \mathbb{R}_{+} \to \mathbb{R}$ que satisface lo siguiente:

Si $b = +\infty$ (respectivamente $0 < b < +\infty$)

1.- $P(\cdot, u)$ es continuamente diferenciable en $\mathbb{R}$, estrictamente convexa en $[-b, +\infty)$ y constante en $(-\infty, -b]$.

2.- $P(0, u) = 0$, $\frac{\partial P}{\partial t}(0, u) = u$

3.- $\lim_{t \to -\infty} \frac{\partial P}{\partial t}(t, u) = 0$ y

4.- $\lim_{t \to +\infty} \frac{\partial P}{\partial t}(t, u) = +\infty$

para cualquier $u > 0$ y para todo $t \in \mathbb{R}$ respectivamente para cualquier $u \geq 0$ y para todo $t \in \mathbb{R}$).

Por comodidad de aquí en adelante escribiremos $P'(\cdot, u)$ en lugar de $\frac{\partial P}{\partial t}(\cdot, u)$, para cualquier $u \geq 0$.

Las funciones Lagrangeano y Lagrangeano Aumentado

La función Lagrangeano asociada al problema (P) es definida por

$$x \in \mathbb{R}^n, \mu \in \mathbb{R}^m \to l(x, \mu) = f(x) + \sum_{i=1}^{m} \mu_i g_i(x),$$

aparece en la teoría de la optimización al tratar de las condiciones necesarias y suficientes que una solución del problema (P) verifica.

Como veremos más adelante, estas condiciones (y la función lagrangeano clásico) serán parte fundamental en la construcción del método Lagrangeano Aumentado en estudio. De manera similar definimos, asociada al problema (P), L, la función Lagrangeano aumentado (con penalidades $P_i \in \mathcal{P}$) dada por:

$$x \in \mathbb{R}^n, \mu \in \mathbb{R}^m, \lambda > 0 \to L(x, \mu, \lambda) = f(x) + \lambda \quad (2.1)$$

donde las m funciones $P_i :: \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}$ (o las m funciones $P_i : \mathbb{R} \times \mathbb{R}_{+} \to \mathbb{R}$) llamadas **penalidades** pertenecerán a la familia $\mathcal{P}$ definida antes.

La primera función Lagrangeano aumentado introducida tenía aplicación a problemas con restricciones de igualdad (ver [5] y [8]).

Rockafellar [10] generalizó la metodología para el caso donde las restricciones son de desigualdad. Esta primera función L.A. es dada por:

$$x \in \mathbb{R}^n, \mu \in \mathbb{R}^m, \lambda > 0 \to L(x, \mu, \lambda) = f(x)$$

$$+\sum_{i=1}^{m}\frac{1}{2}\left[(\max\{0,\mu_i+\frac{g_i(x)}{\lambda}\})^2-\mu_i^2\right] \quad (2.2)$$

Resultados preliminares a la descripción del método

Definición 2.6.- Sea $C \subset \mathbb{R}^n$ un conjunto convexo. La aplicación $F : C \to \mathbb{R}^n$ se dice monótona (respectivamente estrictamente monótona) en C si para todo $(x, x') \in C \times C$, $(F(x)-F(x'))^t(x-x') \geq 0$ (respectivamente $(F(x)-F(x'))^t(x-x') > 0$, cuando $x \neq x'$).

Teorema 2.7.- Sea f una función diferenciable en un abierto $\Omega \subset \mathbb{R}^n$ y sea $C \subset \Omega$ un conjunto convexo. Luego f es convexa (respectivamente estrictamente convexa) en C si y solo si ∇f es monótono (respectivamente estrictamente monótono) en C.

Prueba.-
Ver Teorema 4.1.4, pág. 185 en [3].

Observación 2.8.- Sea $f : \mathbb{R} \to \mathbb{R}$ convexa (respectivamente estrictamente convexa) y diferenciable.
Para $\mu \geq 0$ (respectivamente $\mu > 0$), si $f'(0) = \mu$ y $\lim_{t\to-\infty} f'(t) = 0$ entonces $f'(t) \geq 0$ (respectivamente $f'(t) > 0$) para todo $t \in \mathbb{R}$.

Prueba.-

a) Para f convexa. Dado que $f'(t) = \mu \geq 0$ para $t = 0$, analizaremos dos casos.

i) Si $t \geq 0$ entonces por el Teorema 2.7 $f'(t) \geq f(0)$. Así, $f'(t) \geq \mu \geq 0$.

ii) Caso $t < 0$. Supongamos que existe $\bar{t} < 0$ tal que $f'(\bar{t}) < 0$ entonces, para todo $t < \bar{t}$ tenemos $f'(t) \leq f'(\bar{t}) < 0$ por el mismo teorema.

Haciendo tender $t \to -\infty$ obtenemos,

$$\lim_{t\to-\infty} f'(t) \leq f'(\bar{t}) < 0$$

lo que contradice la hipótesis. Luego, para todo $t < 0$, $f'(t) \geq 0$.

Por tanto, $f'(t) \geq 0$ para todo $t \in \mathbb{R}$.

b) Para f estrictamente convexa. Supongamos que existe $\bar{t} \in \mathbb{R}$ tal que $f'(\bar{t}) = 0$. Entonces, para do $t < \bar{t}$, nuevamente por el Teorema 2.7, tenemos $f'(t) < f'(\bar{t})$. Así $f'(t) < f'(\bar{t}) = 0$.

Por el resultado anterior (caso f convexa) sabemos que $f'(t) \geq 0$ para todo $t \in \mathbb{R}$. Luego, $0 \leq f'(t) < 0$, lo que es una contradicción. Por tanto, no existe $t \in \mathbb{R}$ tal que $f'(t) = 0$, o sea, retiramos la igualdad en la conclusión de la Observación.

Proposición 2.9.- Sea $f \in \mathrm{Conv}\, \mathbb{R}^n$ (respectivamente $\overline{\mathrm{Conv}}\, \mathbb{R}^n$) y sea $g \in \mathrm{Conv}\, \mathbb{R}$ (respectivamente $\overline{\mathrm{Conv}}\, \mathbb{R}^n$) creciente.

Supongamos que existe $x^o \in \mathbb{R}^n$ tal que $f(x^o) \in \mathrm{dom}\, g$. Luego la función com-puesta $x \to g(f(x))$ está en $\mathrm{Conv}\, \mathbb{R}^n$ (respectivamente $\overline{\mathrm{Conv}}\, \mathbb{R}^n$).

Prueba.-
Ver Proposición 2.1.8, pág. 160, en [3].

Observación 2.10.- Para cada $u \geq 0$ y $\lambda > 0$ fijos. Considere $P \in \mathrm{P}$ y $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ una función convexa y cerrada. Entonces:

i) $\frac{\partial P}{\partial t}(t, u) = P_u'(t) \geq 0$ para todo $t \in \mathbb{R}^n$. O sea $P_u = P(\cdot, u)$ es creciente.

ii) $P(\frac{g_i(\cdot)}{\lambda}, u) = P_u \circ \frac{g}{\lambda}$ es convexa y cerrada.

Además, si $P \in \mathrm{P}$ es tal que para cualquier $u > 0$, $P(\cdot, u)$ es estrictamente convexa (el caso en que $b = +\infty$), entonces $P_u'(t) > 0$, para todo $t \in \mathbb{R}^n$. O sea, P_u será estrictamente creciente.

Prueba.-

i) Como $P \in \mathrm{P}$, para $u \geq 0$ fijo $P(\cdot, u) = P_u(\cdot)$ es convexa y diferenciable en $\mathbb{R}$. Además de eso P_u verifica las condiciones de la observación 2.8 entonces $P_u'(t) \geq 0$ para todo $t \in \mathbb{R}^n$ luego, P_u será creciente..

ii) Dado que g es convexa y cerrada, para $\lambda > 0$ fijo, g / λ también lo será.

Como para $u \geq 0$ fijo, P_u es convexa y creciente, tenemos por la proposición 2.9 que la compuesta

$P_u \circ \frac{g}{\lambda}$ también estará en $\overline{C\,onv}$ $\mathbb{R}^n$.

La última afirmación se obtiene del caso alternativo en la observación 2.8.

Observación 2.11.- Para $\mu \geq 0$ en $\mathbb{R}^m$ y $\lambda > 0$ fijos, consideremos la función Lagrangeano aumentado con penalidades $P_i \in \mathrm{P}$ definida en (2.1).

Entonces:

$$x \in \mathbb{R}^n \rightarrow L(x, \mu, \lambda) = L\mu_{,\lambda}(x) = f(x) + \lambda \sum_{i=1}^{m} P_i\left(\frac{g_i(x)}{\lambda}, \mu_i\right)$$

es convexa y cerrada.

Prueba.-

Para $\mu \geq 0$ en $\mathbb{R}^m$ fijo, tenemos $\mu_i \geq 0$ para todo i = 1 ,..., m. Entonces, como cada $P_i \in \mathrm{P}$ y g_i son convexas y cerradas, por la Proposición 2.9 (i), $P_{i\mu_i} \circ \frac{g_i}{\lambda}$ estará en $\overline{Conv}$ $\mathbb{R}^n$, para $\lambda > 0$ fijo.

Finalmente, como $L\mu_{,\lambda}$ es una combinación positiva de funciones convexas y cerradas, $L\mu_{,\lambda}$ también estará en $\overline{C\,onv}$ $\mathbb{R}^n$.

Proposición 2.12.- Para $f \in$ Conv $\mathbb{R}^n$, se tiene:

$x \in \mathbb{R}^n$ minimiza f si y solo si $0 \in \partial f(x)$.

Prueba.-
Consecuencia del Corolario 1.4.4., pag. 48 dado en [4].

Corolario 2.13.- Sean f_1, f_2, dos funciones convexas y cerradas. Asumamos que ri dom $f_1 \cap$ ri dom $f_2 \neq 0$. Luego:

$\partial(f_1+f_2)(x) = \partial f_1(x)+\partial f_2(x)$ para todo $x \in$ dom $f_1 \cap$ dom f_2.

Prueba.-
Ver Corolario 3.1.2, pag. 114 en [4].

Teorema 2.14.- Sean $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, convexa y $g \in \overline{Conv}$ $\mathbb{R}^n$ creciente. Asumamos que $f(\mathbb{R}^n) \cap$ int dom $g \neq 0$.

Luego, para todo x tal que:

$f(x) \in$ dom g $s \in \partial(g \circ f)(x)$ si y solo si $\exists\, \alpha \geq 0$ tal que $s \in \partial(\alpha f)(x)$, $\alpha \in \partial g(f(x))$.

Si además, g es diferenciable en f(x), la relación anterior puede expresarse por:

$$\partial(g \circ f)(x) = \partial g(f(x))\, \partial f(x).$$

Prueba.-
Consecuencia de los Teoremas 3.6.1, págs 125 y 126, en [4] y 4.3.1., pág. 265 en [3].

Definición 2.15.- Un punto $\bar{x} \in \mathbb{R}^n$ es llamado un punto de Karush Kuhn Tucker (K.K.T) si existe algún punto $\bar{u} \in \mathbb{R}^n$ tal que se verifica

a) $0 \in \partial f(\bar{x}) + \sum_{i=1}^{m} u_i g_i(\bar{x}) = \partial l(\bar{x}\ \bar{u})$

b) $\bar{u} \geq 0$

c) $g_i(\bar{x}) = 0$ para todo i = 1 ,..., m

d) $\bar{u}_i\ g_i(\bar{x}) = 0$ para todo i = 1 ,..., m

Si se verifica apenas a) diremos que $(\bar{x}\,\bar{u})$ cumplen la condición de Lagrange.

Definición 2.16.- Una función definida de la siguiente manera: $x \in \mathbb{R}^n \rightarrow \langle s, x \rangle + k$, para algún $s \in \mathbb{R}^n$ y algún $k \in \mathbb{R}^n$, es llamada función a fín.

Sea $X = \{ x \in \mathbb{R}^n : g_i(x) \leq 0\}$ el conjunto de puntos viables para el problema (P). Denotemos por J_a al conjunto de índices correspondientes a las restricciones g_i afines:

$$J_a = \{ i = 1 ,..., m: g_i \text{ es una función a fin} \}$$

Definición 2.17.- Decimos que X satisface la hipótesis débil de Slater si existe un punto $x^o \in \mathbb{R}^n$ en el cual:

i) $g_i(x^o) \leq 0$ para $i \in J_a$

ii) $g_i(x^o) < 0$ para $i \notin J_a$

Al asumir la Hipótesis 2.4, se estará satisfaciendo la hipótesis débil de "Slater".

Teorema 2.18.- Supongamos que se verifica la hipótesis débil de "Slater". Entonces, para cualquier $\bar{x} \in \mathbb{R}^n$, el hecho de que $\bar{x}$ sea un punto de Karush Kuhn Tucker es suficiente y necesario para que $\bar{x}$ sea una solución del problema (P).

Prueba.-
Ver Teorema 2.2.5, pag 310, en [3].

Descripción del Método Lagrangeano Aumentado

Consideremos la función Lagrangeano Aumentado con penalidades $P_i \in$ P asociada al problema (P) y $\{\lambda_k\} \subset \mathbb{R}$ limitada de números reales positivos. El Método Lagrangeano Aumentado para la optimización convexa intenta aproximar un punto de K.K.T. (que bajo la Hipótesis 2.3 y por el Teorema 2.18, será también una solución del problema (P)) mediante la secuencia $\{x^k\} \subset \mathbb{R}^n$ generada al resolver iteractivamente, partiendo de algún punto $\mu^0 \in \mathbb{R}^m_+$ ($\mu^0 \in \mathbb{R}^m_{++}$) dado, problemas del tipo:

$$\text{Minimizar}_{\ x\in\mathbb{R}^n}\ L(x, \mu^k, \lambda_k) = f(x) + \lambda_k \sum_{i=1}^{m} P_i\left(\frac{g_i(x)}{\lambda_k}, \mu_i^k\right) \qquad (\hat{P}_k)$$

donde la actualización de μ^k se realiza por la fórmula:

$$\mu_i^{k+1} = \frac{\partial P_i}{\partial y_i}(y_i^{k+1}, \mu_i^k),\ \ _i^{k+1} = \frac{g_i(x^{k+1})}{\lambda_k} \quad \text{para todo} = 1, \dots, m \qquad (2.3)$$

(x^{k+1} denota una solución del problema $(\hat{P}_k)$).

La idea es procurar verificar las condiciones dadas por la Definición 2.15.

Primero imponemos la verificación de la condición de Lagrange para todos los iterados. Expliquemos esto.

Para x^{k+1} una solución del problema $(\hat{P}_k)$ (que por ahora supondremos que existe), tenemos por la Observación 2.11 y la Proposición 2.12.

$$0 \in \partial L\mu^k, \lambda_k (x^{k+1}) \qquad (2.4)$$

Evaluando la relación (2.1) en (μ^k, λ_k) obtenemos:

$$L\mu^k, \lambda_k(x) = f(x) + \lambda_k \sum_{i=1}^{m} P_{i\mu_i^k}\left(\frac{g_i(x)}{\lambda_k}\right),$$

Usando ahora el Corolario 2.13:

$$\partial L\mu^k, \lambda_k(x) = \partial f(x) + \lambda_k \sum_{i=1}^{m} \partial\left[P_{i\mu_i^k}\left(\frac{g_i(x)}{\lambda_k}\right)\right] \qquad (2.5)$$

Como $P_{i\mu_i^k}$ es creciente (Observación 2.10 (i)) y diferenciable en $\mathbb{R}$, por el Teorema 2.14,

$$\partial\left(P_{i\mu_i^k} \circ \frac{g_i(x)}{\lambda_k}\right)(x) = \frac{\partial P_{i\mu_i^k}}{\partial y_i}\left(\frac{g_i(x)}{\lambda_k}\right).\frac{\partial g_i(x)}{\lambda_k},\ y_i = \frac{g_i(x)}{\lambda_k}$$

Sustituyendo esto en (2.5), evaluando en x^{k+1} y retomando (2.4).

$$0 \in \partial L\mu^k, \lambda_k(x^{k+1}) = \partial f^{k+1}) + \sum_{i=1}^{m} \underbrace{\left(\frac{\partial P_{i\mu_i^k}}{\partial y_i}\left(\frac{g_i(x^{k+1})}{\lambda_k}\right)\right)}_{\mu_i^{k+1}} \partial g_i(x^{k+1}) \qquad (2.6)$$

Así, (x^{k+1}, μ^{k+1}) verifica la condición de Lagrange, lo que garantiza que un punto límite de la secuencia $\{x^k, \mu^k\}$ (si existe) también satisfacerá dicha condición.

Por otro lado, usando la Observación 2.10 (i) (con $\mu = \mu^k \geq 0$, $\lambda = \lambda_k > 0$ y $\bar{x} = x^{k+1}$,para un $k \in \mathbb{N}$ dado) y la fórmula (2.3) obtenemos $\mu^{k+1} \geq 0$ (respectivamente $\mu^{k+1} > 0$ para todo $k \in \mathbb{N}$, para penalidades P_{i} (., μ_i^k estrictamente convexas).

Luego, la viabilidad dual se verifica para un punto límite $\bar{u}$ de la secuencia $\{\mu^k\}$. Viabilidad primal y complementaridad (items c y d de la Definición 2.15 respectivamente), no se satisfacen a lo largo de los iterados. No en tanto deberán verificarse en un punto límite $(\bar{x}, \bar{u})$ (esto acontece por ejemplo en el caso de las penalidades de Polyak y Teboulle [7]).

Para conseguir estos objetivos, así como para garantizar la existencia de soluciones del problema $(\hat{P}_k)$ y la convergencia de las secuencias {xk} y {μk}, deberemos asumir algunas hipótesis, pero ese no es nuestro objetivo.

Nosotros estamos interesados específicamente en probar la buena definición de la secuencia $\{xk\} \subset$ IRn generada por el Método Lagrangeano Aumentado con penalidades $P_i \in$ P.

Esta prueba incluye casi todos los casos existentes en la literatura, tales como los que se encuentran en [10] (donde se usa la penalidad Estándar), en [1] (que usa la penalidad Exponencial), en [3] (con penalidades estrictamente Convexas) y en [2] que considera penalidades con Cambio de Variable. Comenzaremos presentando algunas definiciones y resultados que serán de utilidad.

Definiciones y Resultados Preliminares

Para f: $\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ y algún $r \in \mathbb{R}$ denotemos por $S_r(f)$ a los conjuntos de nivel de f dados por: $S_r(f) = \{x \in \mathbb{R}^n : f(x) \leq r\}$.

Sea h: $\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ una función convexa cerrada, recordemos su función de recesión $h'_\infty(d)$:

$$d \in \mathbb{R}^n \to h'_\infty(d) = \lim_{\lambda \to +\infty} \frac{h(x+\lambda d) - h(x)}{\lambda},$$

para $x \in$ dom h.

Proposición 2.19.- Para $f \in \overline{\text{Conv}}\, \mathbb{R}^n$, los siguientes items son equivalentes:

i) existe $r \in \mathbb{R}$ para el que $S_r(f)$ es no vacío y compacto.

ii) todos los conjuntos de nivel de f son compactos,

iii) $f'_\infty(d) > 0$ para todo $d \neq 0$ en $\mathbb{R}^n$.

Prueba.-
Ver Proposición 3.2.5, pág. 180 en [3].

Definición 2.20.- Una dirección $d \neq 0$ en $\mathbb{R}^n$ es una dirección de recesión de h si $h'_\infty(d) \leq 0$, donde $h'_\infty(.)$ denota la función de recesión de h.

Observación 2.21.- Si el conjunto de soluciones óptimas del problema (P) es no vacío y limitado (Hipótesis 2.3), entonces las funciones f, $g_1, \ldots\ldots, g_m$ no tienen ninguna dirección de recesión común.

Prueba.-
Por contradicción. Supongamos que h es una dirección de recesión común.

Entonces, para x^*, una solución óptima, tenemos que para todo $\lambda \geq 0$, $x^* + \lambda h$ será viable y $f(x^* + \lambda h) \leq f(x^*)$, lo que es una contradicción pues el conjunto óptimo es limitado. Luego, no existe dirección de recesión común.

Proposición 2.22.- Sean $f_1, \ldots\ldots, f_m$ m funciones en $\overline{\text{Conv}}\, \mathbb{R}^n$ y $t_1, \ldots\ldots, t_m$ m números positivos. Asumamos que existe $x^0 \in \mathbb{R}^n$ en el que f_j es finito para todo j = 1 ,..., m. Luego para $f = \sum_{j=1}^{m} t_j f_j$

tenemos: $$f'_\infty = \sum_{j=1}^{m} t_j f'_{j\infty}.$$

Prueba.-
Ver Proposición 3.2.9, pág 182 en [3].

Observación 2.23.- Sea f: $\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ una función convexa y cerrada. Luego para $\bar{x} \in \mathbb{R}^n$ y para todo $h \in \mathbb{R}^n$ tal que $h \neq 0$, $f'_\infty(h) = \lim_{\lambda \to +\infty} f'(\bar{x} + \lambda d, h)$ donde $f'(x, h)$ denota la derivada direccional de f en el punto x, en la dirección h.

Prueba.-
Para todo $x \in \mathbb{R}^n$ y para cualquier $h \in \mathbb{R}^n$ tal que $h \neq 0$, por la definición de $f'(x, h)$ tenemos $f(x+\lambda h) \geq f(x) + \lambda f'(x,h)$ para todo $\lambda > 0$.

Entonces:

$$\lim_{\lambda \to +\infty} \frac{f(x+\lambda h) - f(x)}{\lambda} \geq f'(x, h),$$

O sea, $f'_\infty(h) \geq f'(x, h)$. Haciendo $x = \bar{x} + \lambda h$ para $\in \mathbb{R}^n$ fijo y para todo $\lambda > 0$ tenemos: $f'_\infty(h) \geq f'(\bar{x} + \lambda h, h)$ para todo $\lambda > 0$.

Tomando límite cuando $\lambda \to +\infty$ $f'_\infty(h) \geq \lim_{\lambda \to +\infty} f'(\bar{x} + \lambda h, h)$ para todo $\lambda > 0$.

Para $y = \bar{x} + \lambda h \in \mathbb{R}^n$ para todo $\lambda > 0$ entonces $f(x) \geq f(y) + \lambda f'(y, -h)$. Sustituyendo y, $f \geq f(\bar{x} + \lambda h) + \lambda f'(y, -h)$. Pero, como f es convexa se verifica también $f'(y, h) \geq -f'(y, -h)$.

Sustituyendo esto en la relación anterior:

$f(\bar{x} + \lambda h) - f(x) = \lambda f'(\bar{x} + \lambda h, h)$ para todo $\lambda > 0$.

Tomando límite cuando $\lambda \to +\infty$:

$$\lim_{\lambda\to+\infty} \frac{f(x+\lambda h)-f(x)}{\lambda} = \lim_{\lambda\to+\infty} f'(\bar{x}+\lambda h, h)$$

o sea

$$f'_{\infty}(h) = \lim_{\lambda\to+} f'(\bar{x}+\lambda h, h).$$

Luego tenemos probada la igualdad.

Observación 2.24.- Sea $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ una función convexa y cerrada, $h \in \mathbb{R}^n$ y $x \in \mathbb{R}^n$. Entonces:

$$f'(x, h) = \max_{s\in\partial f(x)} \{\langle s, h\rangle\}$$

donde: $\partial f(x)$ es no vacío compacto y convexo.

Prueba.-
Consecuencia del resultado (2.0.1), pág 102, en [4].

Observación 2.25.- Para $P: \mathbb{R} \to \mathbb{R}$ una función convexa diferenciable y $g: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ una función convexa cerrada, definamos la compuesta $x \in \mathbb{R}^n \to f(x) = P(g(x))$.
Si $P'(t) \geq 0$ para todo $t\in \mathbb{R}$ entonces $f'(x, h) = P'(g(x))\, g'(x, h)$ para todo $h\in \mathbb{R}^n$, $x \in \mathbb{R}^n$.

Prueba.-
Para $x \in \mathbb{R}^n$ arbitrario, $h\in \mathbb{R}^n$ fijo, de la Observación 2.24 tenemos:

$$\begin{aligned} f'(x, h) &= \max \{\langle s,h\rangle, s\in\partial f(x)\} \\ &= \max \{s^t h : s = P'(g(x))\gamma, \gamma\in\partial g(x)\} \\ &= \max \{P'(g(x))\gamma^t h : \gamma\in\partial g(x)\} \\ &= P'(g(x)) \max \{\gamma^t h : \gamma\in\partial g(x)\} \end{aligned}$$

usando nuevamente la Observación 2.24 ahora con g: $f'(x, h) = P'(g(x))\, g'(x, h)$, que es lo que queríamos probar.

La buena definición del Algoritmo Lagrangeano Aumentado con penalidades $P_i \in \mathrm{P}$

Consideremos el problema (P) bajo la Hipótesis 2.3. Sea $\{x^k\} \subset \mathbb{R}^n$ una secuencia generada por el algoritmo Lagrangeano Aumentado con penalidades $P_i \in \mathrm{P}$ siguiente:

$$\mu^0 \in \mathbb{R}^m_+ \tag{2.7}$$

$$x^{k+1} = \operatorname{argmin}_{x\in\mathbb{R}^n} \{L(x,\mu^k,\lambda_k)\} \tag{2.8}$$

$$\mu_i^{k+1} = \frac{\partial P_i}{\partial y_i}(y_i^{k+1}, \mu_i^k) \text{ para todo } i = 1, \ldots, m \text{ y para todo } k \in \mathbb{N} \tag{2.9}$$

donde $\{\lambda_k\}$ es una secuencia limitada de números reales positivos, L es la función Lagrangeano Aumentado con penalidades:

$$P_i \in \mathrm{P} \text{ y } y_i^{k+1} = \frac{g_i(x^{k+1})}{\lambda_k} \text{ para } i = 1, \ldots, m.$$

Afirmación 2.26.- La secuencia $\{x^k\}$ está bien definida.

Prueba.-
Dados $\lambda_k > 0$ y $\mu^k \geq 0$ ($k \in \mathbb{N}$ arbitrario), probaremos que L_{μ^k,λ_k} no tiene dirección de recesión (lo que según la Proposición 2.19 será equivalente a afirmar que los subconjuntos de nivel de L_{μ^k,λ_k} son compactos. Luego L_{μ^k,λ_k} alcanza su mínimo en $\mathbb{R}^n$).

De la definición de L_{μ^k,λ_k}, como f y $(P_{i\mu_i^k} \circ \frac{g_i}{\lambda})$ están en $\overline{\mathrm{Conv}}\ \mathbb{R}^n$ por la Proposición 2.2 tenemos:

$$L'_{\infty}(z, \mu^k, \lambda_k) = f'_{\infty}(z) + \lambda_k \sum_{i=1}^{m} h'_{i\infty}(z) \text{ para todo } z\in \mathbb{R}^n$$

$$\text{donde } h_i(x) = P_i\left(\frac{g_i(x)}{\lambda_k}, \mu_i^k\right) \text{ para todo } i = 1,\ldots, m$$

o también usando la Observación 2.23:

$$\lim_{\lambda\to+\infty} L'_{\mu^k,\lambda_k}(x+\lambda z, z) = \lim_{\lambda\to+\infty} f'(x+\lambda z, z) + \lambda_k \sum_{i=1}^{m} \lim_{\lambda\to+\infty} h_i(x+\lambda z, z) \tag{2.10}$$

Por otro lado, para $z \neq 0$, en $\mathbb{R}^n$, tenemos dos alternativas para una restricción g_i arbitraria: z es dirección de recesión de g_i o no es.

Supongamos que $g'_{i\,\infty} \leq 0$.

Entonces, para $\bar{x}$ un punto viable del problema (P) y $\lambda > 0$ tenemos, por la Observación 2.25,

$$h_i'(\bar{x}+\lambda z, z) = P_i'\left(\frac{g_i(\bar{x}+\lambda z)}{\lambda_k}, \mu_i^k\right)\left[\frac{g_i'(\bar{x}+\lambda z, z)}{\lambda_k}\right]$$

Cuando $\lambda \longrightarrow +\infty$ tenemos dos alternativas:

i) $g_i(\bar{x}+\lambda z) \to -\infty$ .Entonces $P_i'\left(\frac{g_i(\bar{x}+\lambda z)}{\lambda_k}, \mu_i^k\right)$

$\to 0$ y $g_i'(\bar{x}+\lambda z, z)$ es limitada, o

ii) $g_i(\bar{x}+\lambda z)$ es limitada. Luego

$P_i'\left(\frac{g_i(\bar{x}+\lambda z)}{\lambda_k}, \mu_i^k\right)$ es limitada y como g_i es

convexa $g_i'(\bar{x}+\lambda z, z) \to 0$.

En ambos casos,

$$h_i'(\bar{x}+\lambda z, z) \to 0 \text{ cuando } \lambda \longrightarrow +\infty \tag{2.11}$$

iii) Si $g_{i\infty}'(z) > 0$, $g_i(\bar{x}+\lambda z)$ $+\infty$ y

$P \longrightarrow_i'\left(\frac{g_i(\bar{x}+\lambda z)}{\lambda_k}, \mu_i^k\right)$ $+\infty$ cuando $\lambda \longrightarrow +\infty$.

$$\text{O sea, } h_i'(\bar{x}+\lambda z, z) \to +\infty. \tag{2.12}$$

Sustituyendo estos resultados en (2.10) y usando la Observación 2.23, tenemos:

$L_\infty'(z, \mu^k, \lambda_k) = \{f_\infty'(z)$ si z es dirección de recesión de las funciones $g_{i,\infty}$ caso contrario. (2.13)

Finalmente, supongamos que existe $z^0 \neq 0$ dirección de recesión de L μ^k λ_k luego, de (2.13) :$f_\infty'(z^0) \leq 0$ (o sea, z^0 es dirección de recesión de las funciones g_i y de f), lo que contradice la Observación 2.21.

Por tanto,

$L_\infty'(z, \mu^k, \lambda_k) > 0$ para todo $z \neq 0$ en $\mathbb{R}^n$.

SIMBOLOGÍA

int S Interior del conjunto S.
ri S Interior relativo del conjunto S.
$\mathbb{R}_+$ Conjunto de números reales no negativos
$\mathbb{R}_{++}$ Conjunto de números reales positivos.
$\mathbb{R}_+^m$ Conjunto de puntos de $\mathbb{R}^m$ cuyos componentes son no negativas.
$\mathbb{R}_{++}^m$ Conjunto de puntos de $\mathbb{R}^m$ cuyos componentes son positivas $\partial f(x)$. Conjunto de subgradientes de la función de f en el punto x.

BIBLIOGRAFÍA

1. **Bertsekas, D. P.,** "Constrained Optimization and Lagrange Multiplier Method", Academic Press, New York, 1982.
2. **Gonzaga, C. y Castillo, R.**, ″Métodos de Lagrangeano Aumentado usando Penalidades Generalizadas para Programação não linear" Tesis, COPPE, UFRJ, 1998.
3. **Hiriart Urruty, J., Baptiste, Lemaréchal, C.** "Convex Analysis and Minimization Algorithms I, 1 ed". New York, Springer-Verlag, 1993.
4. **Hiriart Urruty, J., Baptiste, Lemaréchal, C.** "Convex Analysis and Minimization Algorithms II, 1 ed". New York, Springer-Verlag, 1993.
5. **Hestenes, M.**, ″Multiplier and Gradient Methods", Jota, vol 4, pp.303-320, 1969.
6. **Iusem, A.**, ″Lagrangian methods and proximal point methods for convex optimization" Minicurso I workshop en Optimización, Florianópolis, Brasil, Diciembre, 1997.
7. **Polyak, R., Teboulle, M.**, ″Nonlinear rescaling and proximal-like methods in convex programming", Departament of Operations Research and Applied Statistics School of Information Technology and Engineering, George Mason University, 1995.
8. **Powell, M.**, ″A method for nonlinear constraints in minimizations problems", Ed., Academic Press, N.Y, pp.283-298, 1969.
9. **Rockafellar, R. T.**, "Convex Analysis, 1 ed". Princeton University Press, Princeton, N.J., 1970.
10. **Rockafellar R. T.**, ″Augmented Lagrangians and applications of the proximal point algorithm inconvex programming", Mathematics of Operations Research, vol 1, pp. 97-116, 1976.