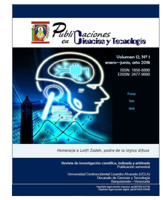

Publicaciones en Ciencias y Tecnología. Vol. 12 N^o 1, Ene-Jun (2018) 7–18

Artículo de Investigación

Método de punto proximal para sucesiones de funciones de Bregman convergentes puntualmente

Proximal point method for pointwise convergent sucessions of Bregman functions

Eibar Hernández[a], Raquel Silvana Quintana Carlone[a], Clavel María Quintana Carlone[a]

[a] *Universidad Centroccidental Lisandro Alvarado. Venezuela*

Recibido: 24-10-2017 Aceptado: 30-06-2018

Resumen

Se desarrolla una generalización del método de punto proximal clásico y el método de punto proximal con distancias de Bregman bajo condiciones de convexidad. Partiendo de una sucesión arbitraria de funciones de Bregman convergente puntualmente, el método propuesto permite generalizar los casos clásicos que han sido desarrollados para una función Bregman fija, considerando propiedades que regulan el comportamiento de la sucesión de distancias de Bregman. Como consecuencia, se obtiene un método que converge al minimizador de la función objetivo.

Palabras clave: Método de punto proximal, distancias de Bregman, sucesiones de funciones de Bregman.
Código UNESCO: 1207.11- Programación no lineal

Abstract

A generalization of the classical proximal point method and the method of proximal point with Bregman distances is developed under conditions of convexity. Starting from an arbitrary punctually convergent sucession of Bregman functions, our method allows both the generalization to the classic cases that have been developed for a fixed Bregman function and the addition of properties that regulate the behavior of the succession of Bregman distances. Thus, a method that converges to the minimizer of the objective function is obtained.

Key words: The proximal point method, Bregman distances, Bregman function successions.
UNESCO Code: 1207.11-Non-linear programming

ISSN 1856-8890, e-ISSN 2477-9660. Dep.Legal:pp200702LA2730. Dep.Legal:ppi201402LA4590.Venezuela.

Email addresses: ehernad@ucla.edu.ve (Eibar Hernández), raquel.quintana@ucla.edu.ve (Raquel Silvana Quintana Carlone), clavel.quintana@ucla.edu.ve (Clavel María Quintana Carlone)

1. Introducción

El problema de programación convexa es expresado comúnmente de la siguiente manera:

$$\widehat{f} = \inf\{f_0(x) \ : \ f_i(x) \leq 0 \quad i = 1, \ldots, m\}, \tag{P}$$

donde $f_i : \mathbb{R}^n \to \mathbb{R}$, para $i = 0, 1, \ldots, m$ son funciones convexas, propias y cerradas. Una forma de resolver el problema (P), es a través de un estudio dual-primal, en los que se utiliza el método de punto proximal, [1], [2],[3] . El método de punto proximal, en su forma clásica, consiste en generar una sucesión $\{x^k\}$ en $\mathbb{R}^n$ de la siguiente forma:

$$x^0 \in \mathbb{R}^n, \tag{1.1}$$

$$x^{k+1} = \arg\min\{f(x) + \lambda_k\|x - x^k\|^2 : x \in \mathbb{R}^n\}, \tag{1.2}$$

donde λ_k es un número real que satisface $0 < \lambda_k \leq \widetilde{\lambda}$, para $\widetilde{\lambda} > 0$ y $\|.\|$ es la norma Euclidiana. Es posible demostrar que bajo hipótesis de convexidad y diferenciabilidad de clase C^1, la sucesión generada por (1.1) y (1.2) converge a un minimizador de f, como lo desarrolla inicialmente Moreau [4], el cual propone la siguiente función:

$$F(y) = \arg\min\{f(x) + \frac{1}{2}\|x - x^k\|^2 : x \in \mathbb{R}^n\}, \tag{1.3}$$

que luego es modificada por Yosida, agregando el parámetro λ convirtiéndose en la regularización Moreau-Yosida [5], conocida por la expresión $F_\lambda(y) = \arg\min\{f(x) + \frac{1}{2}\lambda\|x - x^k\|^2 \ : \ x \in \mathbb{R}^n\}$ obsérvese que el método de punto proximal clásico sustituye el parámetro λ por una sucesión $\{\lambda_k\}$ acotada y de números positivos. Más tarde Martinet [6], introduce este método en el estudio de programación convexa, y posteriormente Rockafellar [7], generaliza la técnica para operadores monótonos maximales. En la década de los 90, surgen diferentes ideas para sustituir los núcleos cuadráticos por otros no cuadráticos. Attouch y Wets [8] consideran núcleos de la forma $k(z - x)$ donde $k : \mathbb{R}^n \to \mathbb{R}_+$ es una función estrictamente convexa y coerciva tal que $k(0) = 0$, recuperando con esta técnica la regularización Moreau-Yosida. Con la intención de continuar el desarrollo en esa dirección los métodos de punto proximal han ido evolucionando utilizando diversos tipos de casi-distancias, entre ellas se destacan las siguientes:

Las Distancias de Bregman: Estas son núcleos de la forma, $D_h(x, y) = h(x) - h(y) - (\nabla h(y))^t(x - y)$, donde $h : \overline{S} \to \mathbb{R}$ es una función estrictamente convexa sobre $\overline{S}$, y continuamente diferenciable sobre S, donde S es un conjunto abierto y convexo, además los conjuntos de subnivel de $D_h(x, y)$ son acotados. S es llamada la zona de h y $D_h : \overline{S} \times S \to \mathbb{R}$ es conocida como la distancia de Bregman asociada a h, [9]. El método de punto proximal con Distancias de Bregman está definido por: $x^0 \in S$, $x^{k+1} = \arg\min\{f(x) + \lambda_k D_h(x, x^k) : x \in S\}$, donde cada λ_k es un número real que satisface $0 < \lambda_k < \widetilde{\lambda}$ para algún $\widetilde{\lambda} > 0$.

Las ϕ-divergencias: Son otro tipo de casi-distancias con las cuales se desarrollan métodos de punto proximal. Por ejemplo la casi-distancias propuesta por Iusem-Svaiter-Teboulle, [2], denotada por; $d_\phi : \mathbb{R}^n_+ \times \mathbb{R}^n_{++} \to \mathbb{R}$, definida por; $d_\phi(x,y) = \sum_{i=1}^{n} y_i \phi\left(\frac{x_i}{y_i}\right)$, se dice que es una $\phi-$divergencia si cumple las siguientes propiedades: ϕ es de clase C^3, convexa $\phi(1) = \phi'(1) = 0$, $\phi''(1) > 0$, $\lim_{t\to 0^+} \phi(t) = +\infty$. El método proximal para este caso, [10], [11], genera una sucesión $\{x^k\} \subset \mathbb{R}^n$ dada por: $x^0 \in \mathbb{R}^n_{++}$, $x^{k+1} = \arg\min\{f(x) + \lambda_k d_\phi(x,x^k) : \mu \in \mathbb{R}^n_{++}\}$, donde cada λ_k es un número real que satisface $0 < \lambda_k < \widetilde{\lambda}$ para algún $\widetilde{\lambda} > 0$.

Los núcleos cuadráticos de segundo orden: Estas son casi-distancias cuya expresión es $d_\phi : \mathbb{R}^n_+ \times \mathbb{R}^n_{++} \to \mathbb{R}$, definida por; $d_\phi(x,y) = \sum_{i=1}^{n} y_i^2 \phi\left(\frac{x_i}{y_i}\right)$, donde, ϕ es de clase C^2, convexa, propia, cerrada con $dom(\phi) \subset [0,+\infty)$, $\phi(1) = \phi'(1) = 0$, $\phi''(1) > 0$, $\lim_{t\to 0^+} \phi^{'}(t) = -\infty$. Aunque similares a las ϕ-divergencias en su forma, son bastante diferentes en cuanto a la técnicas que se utilizan para probar la convergencia en los métodos que ambas generan. Esta casi-distancia fue propuesta por Auslender, Teboulle, Ben-Tiba, [1].

Minimización Proximal de Bregman (BPM): Esta es otra versión más general del método de punto proximal es presentado por Kiwiel, [12], cuyo método usa en su desarrollo una función de Bregman generalizada, la cual no es necesariamente diferenciable. Un método similar presenta Castillo-Gonzaga en [13], denominado método de multiplicadores basados en shifts. En la aplicación de este método se consideran dos tipos de penalidad, a saber las AL1 que son no coercivas y las AL2 que son coercivas. Una diferencia notable entre los métodos de Kiwiel y Castillo-Gonzaga es que Kiwiel no considera en su método de multiplicadores inexactos penalidades no coercivas, mientras Castillo-Gonzaga si las considera en su método de multiplicadores basados en shifts.

Casi-distancia generalizada: En [14] se considera una casi-distancia generalizada, la cual esta dada por,

$$d^p_{\theta^*}(s,\mu) = \sum_{i=1}^{m} \left[\frac{\mu_i^p}{c}\theta^*\left(\frac{cs_i}{\mu_i}\right) - \frac{\mu_i^p}{c}\theta^*(c) - \mu_i^{p-1}(\theta^*)'(c)(s_i - \mu_i) \right], \tag{1.4}$$

donde θ^* es la función conjugada de la función de penalidad θ la cual no pasa a través del origen y satisface que $\theta'(0) = \kappa, \kappa > 0$, $p \geq 0$ y c es tal que $(\theta^*)'(c) = \tilde{y} \in \mathbb{R}$. La casi-distancia generalizada contiene, como casos degenerados, ϕ-divergencias y casi-distancias con núcleos homogéneos de segundo orden. La motivación para conseguir esta casi-distancia proviene de estudiar funciones de penalidad con shift constante en el espacio primal que dan origen a dicha casi-distancia generalizada como núcleo para un método de punto proximal asociado al problema dual.

Enfoque unificado: Recientemente Castillo, Gonzaga, Karas y Matioli, en el 2014, [15], muestran un enfoque unificado tanto para el método de los multiplicadores como para métodos del punto proximal, usando la siguiente casi-distancia, $d^{(p,q)}_{\theta^*}(s,\mu) = \sum_{i=1}^{m} \left[\frac{\mu_i^{p-q}}{c}\theta^*\left(\frac{cs_i}{\mu_i^{1-q}}\right) - \frac{\mu_i^{p-q}}{c}\theta^*(c\mu_i^q) - \mu_i^{p-1}(\theta^*)'(c\mu_i^q)(s_i - \mu_i) \right]$, en los que se muestra una función de penalidad general en término de los parámetros $p, q, \widetilde{y} \in \mathbb{R}, r \in (0, 1]$ y $c > 0$ que incluyen como casos particulares casi todos los métodos de Lagrangeano aumentado conocidos hasta la fecha, cabe destacar que en [15] son presentadas solo las pruebas numéricas de este enfoque, que mas tarde Quintana, [16], presenta las pruebas de convergencia para un caso particular que incluye logaritmos.

2. Desarrollo

Esta sección se divide en dos partes donde en la primera se introduce el problema que se quiere resolver y se presenta una generalización de la definición de distancias de Bregman presentada en [9], en la segunda parte se presentan las pruebas de convergencia obteniendo que el método converge al minimizador de la función objetivo.

2.1. Planteamiento del problema

El problema de programación convexa es el siguiente:

$$\begin{aligned} &min \quad f(x) \\ &s.a \quad x \in S \end{aligned} \tag{2.1}$$

donde S es un conjunto abierto y convexo, f es continua y convexa sobre $\overline{S}$.

A continuación se presenta una generalización de la definición de distancias de Bregman, la cual será empleada para el desarrollo del método propuesto.

Definición 2.1. Sean $S \subset \mathbb{R}^n$ un conjunto abierto y convexo y $\overline{S}$ su clausura, sea $\{h_m\}$ con $m \in \mathbb{N}^* = \mathbb{N} - \{0\}$ una sucesión de funciones donde cada $h_m : \overline{S} \to \mathbb{R}$ es una función de Bregman que define $D_{h_m} : \overline{S} \times S \to \mathbb{R}$ dada por:

$$D_{h_m}(x, y) = h_m(x) - h_m(y) - \nabla h_m(y)^t (x - y),$$

donde para $m \geq 1$, se cumple que:

B1) h_m es $C^1(S)$.

B2) h_m es estrictamente convexa y continua en $\overline{S}$.

B3) Para todo $\delta \geq 0$

- $\Gamma_1^m(y, \delta) = \{x \in \overline{S} : D_{h_m}(x, y) \leq \delta\}$ para todo $y \in S$
- $\Gamma_2^m(x, \delta) = \{y \in S : D_{h_m}(x, y) \leq \delta\}$ para todo $x \in S$

son acotados.

B4) Si $\{y^k : k \in \mathbb{N}^*\} \subset S, y^k \to y^*$, entonces $D_{h_m}(y^k, y^*) \to 0$, cuando $k \to +\infty$.

B5) Si $\{x^k : k \in \mathbb{N}^*\} \subset \overline{S}$, $\{y^k : k \in \mathbb{N}^*\} \subset S$, $\{x^k\}$ es acotada, $\lim_{k\to\infty} y_k = y^*$ y $\lim_{k\to\infty} D_{h_m}(x^k, y^k) = 0$, entonces $\lim_{k\to\infty} x^k = y^*$.

B6) Para toda $m \in \mathbb{N}^*$, se dice que una función h_m de Bregman es coerciva en la frontera, si para $\{y^k\} \subset S$ tal que $\lim\limits_{k\to+\infty} y^k = y \in \partial S$ entonces

$$\lim_{k\to\infty} (\nabla h_m(y^k))^t (x - y^k) = -\infty \quad \text{para todo} \quad x \in S. \tag{2.2}$$

Dado $\{h_k\}_{k=1}^{\infty}$ una sucesión de funciones de Bregman, donde $h_k : \overline{S} \subset \mathbb{R}^n \to \mathbb{R}$, el método de punto proximal para sucesiones convergentes de funciones de Bregman, el cual se abreviara MPPSFBC, consiste en lo siguiente:

$$x^0 \in S \tag{2.3}$$

$$x^{k+1} = \arg\min\{f(x) + \lambda_k D_{h_k}(x, x^k) : x \in \overline{S}\}, \tag{2.4}$$

donde $D_{h_k} : \overline{S} \times S \to \mathbb{R}$ son la distancia de Bregman asociada a h_k para cada k, donde el parámetro λ_k, satisface la siguiente condición;

$$0 < \lambda_k < \overline{\lambda}. \tag{2.5}$$

2.2. Análisis de convergencia

Para el estudio de convergencia del MPPFSBC se requieren las siguientes suposiciones:

Suposición A. Sea $\{h_k\}$ una sucesión de funciones de Bregman convergente puntualmente a una función Bregman sobre S, donde para cada $k \in \mathbb{N}^*$, se tiene que h_k es coerciva en la frontera.

Suposición B. El conjunto optimal del problema denotado por X^* es no vacío.

Suposición C. La sucesión $\{\nabla h_k\}$ converge uniformemente, más aun, para cada $k \geq 1$ se tiene que ∇h_k es una función Lipschitz en S.

Suposición D. $D_{h_{k+1}}(x^*, y) \leq D_{h_k}(x^*, y)$, para todo $y \in S$ y $k \geq 1$.

Se destaca el hecho que la suposición **C** no fue considerada en [17], la cual es necesaria para la convergencia de la sucesión de las distancias de Bregman, y obtener así, una generalización al método de punto proximal clásico.

Teorema 2.1. Si se satisfacen las suposiciones **A,B** y **C**, entonces la sucesión generada por el MPPSFBC está bien definida.

Demostración:

Sea β una cota inferior de f en $\overline{S}$ y $f_k(x) = f(x) + \lambda_k D_{h_k}(x, x^k)$, entonces se cumple para todo $k \geq 1$ que, $f_k = f(x) + \lambda_k D_{h_k}(x, x^k) \geq \beta + \lambda_k D_{h_k}(x, x^k)$, por **B3** se tiene los conjuntos de subnivel de D_{h_k} son acotados, más aun, se obtiene que los conjuntos de subnivel de f_k también lo son. Además f_k es estrictamente convexa por la convexidad de f y porque $D_{h_k}(., x^k)$ estrictamente convexa. Luego podemos concluir entonces que x^{k+1} existe y es único. Solo falta ver que $x^{k+1} \in Int(S)$ en efecto, como $x^{k+1} = argmin\{f(x) + \lambda_k D_{h_k}(x, x^k) : x \in \overline{S}\}$, se tiena que $\lambda_k \nabla h_k(x^k) \in \partial(f + \lambda_k h_k)(x)$, ya que, $\lambda_k \nabla h_k(x^k) \in \partial f(x^{k+1}) + \lambda_k(\nabla h_k(x^{k+1}))$. Probemos bajo la hipótesis de que para todo $k \geq 1$ y h_k es coerciva en la frontera, que $\partial(f + \lambda_k h_k)(x) = \emptyset$ para todo x en el borde de S lo que permitiría concluir que $x_{k+1} \in Int(S)$. Supongamos por reducción al absurdo que existe w en el borde de S tal que $\partial(f + \lambda_k h_k)(w) \neq \emptyset$, tomemos $s^k \in \partial(f + \lambda_k h_k)(w)$, además $z \in S$ y definamos: $y^l = (1 - \epsilon_l)w + \epsilon_l z$, donde $\lim_{l \to \infty} \epsilon_l = 0$ y $\epsilon_l \neq 1$. Trabajando la ecuación anterior, obtenemos: $y^l = (1 - \epsilon_l)w + \epsilon_l z \Leftrightarrow y^l - w = \epsilon_l(z - w)$. Por otro lado, por la convexidad de f en $\overline{S}$ se tiene: $f(y^l) - f(w) \leq \epsilon_l(f(z) - f(w))$, luego, $\frac{1-\epsilon_l}{\lambda_k}((f(w) - f(z)) + (s^k)^t(z - w)) \leq \nabla h_k(y^l)^t(z - y^l)$.

Ahora bien, como $\lim_{l \to \infty} y^l = w$ y $w \in \partial S$ por la definición de coercividad en la frontera de h_k, se tiene que la parte derecha de la última desigualdad tiende a $-\infty$ cuando $l \to \infty$, pero la parte izquierda de la misma tiende a un número finito obteniendose una contradicción. Así $\partial(f + \lambda_k h_k)(x) = \emptyset$ para todo x en el borde de S y podemos concluir que $\{x^{k+1}\} \in Int(S)$.

Obsérvese que la prueba es similar a la que se sugiere en [9] para funciones de Bregman, solo que cabe la pena destacar que para cada k la función h_k no es necesariamente la misma.

Proposición 2.1. Si se satisfacen las suposiciones **A** y **B**, para todo k y toda solución $x^* \in X^*$ de (2.1), entonces se cumple que:

$$D_{h_k}(x^*, x^{k+1}) \leq D_{h_k}(x^*, x^k) - D_{h_k}(x^{k+1}, x^k) \tag{2.6}$$

Demostración:

Se tiene que, $0 \in \partial(f(.) + \lambda_k D_{h_k}(., x^k))(x^{k+1})$, así, $\lambda_k(\nabla h_k(x^k) - \nabla h_k(x^{k+1})) \in \partial f(x^{k+1})$, Se conoce además que el subdiferencial de f en x es el conjunto de vectores s que satisfacen la desigualdad $f(y) \geq f(x) + \langle s, y - x \rangle$ para todo $y \in \mathbb{R}^n$. Por tanto que para cada x^{k+1} y tomando $s^k = \lambda_k(\nabla h_k(x^k) - \nabla h_k(x^{k+1}))$ y $y = x^*, x = x^{k+1}$ tenemos que: $f(x^*) \geq f(x^{k+1}) + \langle s^k, x^* - x^{k+1} \rangle \Rightarrow \frac{f(x^{k+1}) - f(x^*)}{\lambda_k} \leq \frac{\langle s^k, x^{k+1} - x^* \rangle}{\lambda_k}$, utilizando la última desigualdad y la propiedad de distancias de Bregman, [9, proposición 9.1], con el hecho de que x^* es mínimo, tenemos que:

$$D_{h_k}(x^*, x^k) - D_{h_k}(x^*, x^{k+1}) - D_{h_k}(x^{k+1}, x^k) \geq \frac{f(x^{k+1}) - f(x^*)}{\lambda_k} \geq 0. \tag{2.7}$$

La demostración de la siguiente proposición fue inspirada por la presentada en [17] la cual define una función auxiliar con características particulares y parte de suponer que la sucesión $\{x^k\}$ no es acotada.

Proposición 2.2. Bajo las suposiciones A y B, se tiene que la sucesión $\{x^k\}$ generada por el MPPSFBC, es acotada.

Demostración:

Supongamos que la sucesión $\{x^k\} \in \mathbb{R}^n$ no es acotada, de donde existe un $s \in \{1, ..., n\}$ tal que $x_s^k \to +\infty$, ahí que existe $i \in \mathbb{N} \setminus x_s^* < x_s^i < x_s^{i+1}$, luego para este i, definamos $H_i : \mathbb{R}^n \to \mathbb{R}$ como: $H_i(x) =$

$h_i(x^i) + \nabla h_i(x^i)(x - x^i) - (h_i(x^{i+1}) + \nabla h_i(x^{i+1})(x - x^{i+1}))$, observemos que $\nabla H_i(x) = \nabla h_i(x^i) - \nabla h_i(x^{i+1})$ para todo x. Queremos probar que $H_i(x^*) > H_i(x^i) > 0$, para todo $x^* \in X^*$, se tiene que $H_i(x^i) > 0$, puesto que para cada i, h_i es convexa, así que, $H_i(x^i) = h_i(x^i) - h_i(x^{i+1}) - \nabla h_i(x^{i+1})(x^i - x^{i+1}) > 0$. Falta probar que $H_i(x^*) > H_i(x^i), \forall x^* \in X^*$. Supongamos por reducción al absurdo, que existe $x^* \in X^*$ tal que $H_i(x^*) \leq H_i(x^i)$, así, $(\nabla h_i(x^{i+1}) - \nabla h_i(x^i))(x^i - x^*) \leq 0$, esto es,

$$\nabla H_i(x)(x^* - x^i) = (\nabla h_i(x^i) - \nabla h_i(x^{i+1}))(x^* - x^i) \leq 0 \quad \forall x, \tag{2.8}$$

la desigualdad (2.8) es una contradicción ya que, $(\nabla h_i(x^i) - \nabla h_i(x^{i+1}))(x^* - x^i) \neq 0$, puesto que $i \in \mathbb{N}^*$ se tiene que $x_s^* < x_s^i < x_s^i$, así $(x^* - x^i) \neq \mathbf{0}$, además $\nabla h_i(x^{i+1}) \neq \nabla h_i(x^i)$ ya que h_i es estrictamente convexa, más aun, en particular para $x = x^* - x^i$ ya que (**??**) es para toda x, $D_{\nabla H_i(x^* - x^i)}(x^* - x^i) = \| \nabla H_i(x^* - x^i) \|^2 > 0$, de aquí se obtiene una contradicción, así que $H_i(x^*) > H_i(x^i) > 0$.

Por otro lado, puesto que, $H_i(x^*) > 0$, por lo que, $h_i(x^*) - h_i(x^{i+1}) - \nabla h_i(x^{i+1})(x^* - x^{i+1}) \geq h_i(x^*) - h_i(x^i) - \nabla h_i(x^i)(x^* - x^i)$;
esto es; $D_{h_i}(x^*, x^{i+1}) \geq D_{h_i}(x^*, x^i)$, obsérvese que la ecuación anterior es una contradicción con la proposición (2.4), así la sucesión $\{x^k\}$ es acotada.

Proposición 2.3. Suponga que se satisfacen las suposiciones **A, B, C** y **D**, entonces se cumple que:

1. Para todo $x^* \in X^*$ la sucesión $\{D_{h_k}(x^*, x^k)\}$ es convergente.
2. Si $\lim\limits_{j\to\infty} x^{k_j} = \overline{x}$ entonces, $\lim\limits_{j\to\infty} x^{k_j+1} = \overline{x}$.

Demostración:

Parte 1:

Por suposición D y de (2.6) se tiene que,

$$D_{h_{k+1}}(x^*, x^{k+1}) \leq D_{h_k}(x^*, x^{k+1}) \leq D_{h_k}(x^*, x^k). \tag{2.9}$$

Así sucesivamente se obtenemos que;

$$0 \leq D_{h_k}(x^*, x^k) \leq D_{h_1}(x^*, x^1). \tag{2.10}$$

Luego de (2.10) se tiene que la suceción $\{D_{h_k}(x^*, x^k)\}$ es acotada y de (2.9) se tiene que decrece, por lo tanto la sucesión $\{D_{h_k}(x^*, x^k)\}$ converge.

Parte 2:

De la proposición (2.2) se tiene que sucesión $\{x^k\}$ es acotada, es decir que, tiene al menos un punto de acumulación $\overline{x}$ y por el teorema de Weierstrass $\{x^k\}$ posee al menos una subsucesión que converge a ese punto límite, esto es, $\lim_{j\to\infty} x^{k_j} = \overline{x}$ donde x^{k_j} es subsucesión de $\{x_k\}$. Además x^{k_j} satisface que $D_{h_{k_j}}(\overline{x}, x^{k_j}) = h_{k_j}(\overline{x}) - h_{k_j}(x^{k_j}) - \langle \nabla h^t_{k_j}(\overline{x}), \overline{x} - x^{k_j} \rangle$, tomando el doble límite cuando $j \to \infty$ en la igualdad anterior, se tiene que,

$\lim_{j\to\infty} D_{h_{k_j}}(\overline{x}, x^{k_j}) = 0$. Por otra parte de (2.6) y la suposición **D**, obtenemos, $D_{h_{k_j}}(x^{k_j+1}, x^{k_j}) \leq D_{h_{k_j}}(x^*, x^{k_j}) - D_{h_{k_{j+1}}}(x^*, x^{k_j+1})$; luego de la parte 1 y puesto que $\{D_{h_{k_{j+1}}}(x^*, x^{k_j+1})\}$ es subsucesión de $\{D_{h_k}(x^*, x^k)\}$, se sigue que, $\lim_{j\to\infty} D_{h_{k_j}}(x^{k_j+1}, x^{k_j}) = 0$ Luego por propiedad (B5) $\lim\limits_{j\to\infty} x^{k_j+1} = \overline{x}$.

La siguiente proposición permite obtener en principio una convergencia débil ya que converge al punto limite de la sucesión, más aun, se hará uso del hecho que ∇h_k es una función Litchiptz para cada $k \geq 1$, esta condición no fue considerada por Quintana, [18], pero, esta hipótesis permite la generalización del método clásico cuadrático, ya que es considerado el hecho que λ_k es una constate distinta de cero.

Proposición 2.4. Bajo las suposiciones **A, B, C** y **D**, cualquier punto limite de la sucesión $\{x^k\}$ generada por el MPPSFBC con λ_k satisfaciendo (2.5) es solución óptima de (P).

Demostración:

De la proposición (2.2) se tiene que sucesión $\{x^k\}$ es acotada, es decir que, tiene al menos un punto de acumulación $\overline{x}$ y por el teorema de Weierstrass $\{x^k\}$ posee al menos una subsucesión que converge a ese punto límite, esto es: $\lim_{j\to\infty} x^{k_j} = \overline{x}$, donde x^{k_j} es subsucesión de $\{x_k\}$. Además x^{k_j} satisface que $-\lambda_{k_j}\nabla D_{h_{k_j}}(., x^{k_j})(x^{k_j+1}) \in \partial f(x^{k_j+1})$, por propiedad de las distancias de Bregman se tiene que, $\lambda_{k_j}[\nabla h_{k_j}(x^{k_j}) - \nabla h_{k_j}(x^{k_j+1}] \in \partial f(x^{k_j+1})$.

Sea x^* una solución optima de el problema primal, tomando $x = x^*$ y $y = x^{k_j+1}$ y aplicando la definición de subgradiente, se obtiene, $f(x^*) - f(x^{k_j+1}) \geq \langle y^{k_j}, x^* - x^{k_j+1}\rangle$, con $y^{k_j} = \lambda_{k_j}(\nabla h_{k_j}(x^{k_j}) - \nabla h_{k_j}(x^{k_j+1})$.

Por otra parte $|\langle y^{k_j}, x^* - x^{k_j+1}\rangle| \leq \|y^{k_j}\|\|x^* - x^{k_j+1}\|$, puesto que la sucesión $\{x^k\}$ es acotada, se tiene $f(x^*) - f(x^{k_j+1}) \geq -M\|y^{k_j}\|$. Finalmente se obtiene que $\lim_{j\to\text{ínf}} \|y^{k_j}\| = 0$, ya que $\lim_{j\to\text{ínf}} \|y^{k_j}\| = \lambda_{k_j}(\nabla h_{k_j}(x^{k_j}) - \nabla h_{k_j}(x^{k_j+1}) = 0$. Así, $0 \leq f(x^*) - f(\overline{x})$, luego la desigualdad anterior nos dice que todo punto limite es solución del problema (2.1).

Teorema 2.2. Bajo las suposiciones **A, B, C** y **D**, la sucesión $\{x^k\}$ generada por el MPPSFBC con λ_k satisfaciendo (2.5), converge a la solución optima de (P).

Demostración:

De la proposición (2.2) se tiene que la sucesión $\{x^k\}$ es acotada, así existe una subsucesión convergente, dígase $\{x^{k_j}\}$ una subsucesión convergente de $\{x^k\}$ tal que, $\lim\limits_{j\to\infty} x^{k_j} = \overline{x}$. Luego de la proposición (2.4) se tiene que $\lim\limits_{j\to\infty} x^{k_j} = \overline{x}$. De la definición (2.1) parte **B4** se tiene que, $\lim\limits_{j\to\infty} D_{h_{k_j}}(x^*, x^{k_j}) = 0$, ahora, del teorema (2.4), x^* es una solución óptima del problema (2.1), y de la proposición (2.3) se tiene que $\{D_{h_k}(x^*, x^k)\}$ es una sucesión convergente, con una subsucesión convergente a 0, así la subsucesión completa converge a cero, esto es, $\lim\limits_{j\to\infty} D_{h_k}(x^*, x^k) = 0$. Luego por la definición (2.1) parte **B5**, (para $y^k = x\prime lim_{k\to\infty} y^k = x\prime \{x^k\}$ es acotada) se tiene que $\lim\limits_{k\to\infty} x^k = x^* \in X^*$.

Esta trabajo es considerado como un primer paso para resolver un problema abierto que aparece en [14], en que el método de punto proximal desarrollado con la casi-distancia generalizada sugiere que deben estudiarse métodos de punto proximal asociados a familias de distancias de Bregman que dependan de un parámetro vectorial, el cual sera estudiado en investigaciones posteriores, este método contiene como casos particulares el método de punto proximal clásico y el método de punto proximal con distancias de Bregman, no solo en forma, si no, también en demostración.

2.3. Aportes computacionales

En este sección se consideran 10 problemas para minimizar, cuyo propósito es ilustrar el método que se propone, usando dos sucesiones de funciones de Bregman, el algoritmo fue desarrollado usando Octave 4.2 bajo un ambiente Windows en una computadora Intel Centrino con 2G de RAM y 1.6 Ghz.

Para las pruebas numéricas se usa BFGS con una búsqueda lineal monótona y no monótona basada en la presentada por Pan y Chen, [19].

En el cuadro (1) se presentan las sucesiones de funciones consideradas para el algoritmo, donde para cada $k \geq 1$ se tiene que el elemento de la sucesión es una función Bregman, $M \in \mathbb{R}^{n \times n}$ es una matriz simétrica y definida positiva y en el cuadro (2) se muestran las funciones a minimizar considerando funciones convexas, casi-convexas y la función de Rosenbrock, estas dos últimas con la finalidad de ilustrar futuras investigaciones.

En los cuadros (3) y (4), se observa que en el caso de las funciones convexas la sucesión SBM2 mostró un mejor desempeño en el desarrollo del algoritmo, más aún la búsqueda lineal de Armijo mejoró notablemente con el uso de la sucesión SBM2. En el caso de las funciones casiconvexas la búsqueda no monótona mostró una mejor aproximación a las soluciones (ver en el cuadro 4 el problema $f10$) ya que en este caso la búsqueda de Armijo se alejo de la solución con la sucesión SBM2.

	Sucesión Bregman	Referencia
SBM_1	$\{\frac{x^t M x}{k}\}_{k \geq 1}$, $M \in \mathbb{R}^{n \times n}$	[9, Ejemplo 9.1]
SBM_2	$\{(-log(1-x(1)) - log(1-x(2))) * (1/(k^2))\}_{k \geq 1}$	[2, Ejemplo 7.1 (2)]

Cuadro 1: Tabla de $\{h_k\}_{k \geq 1}$

f_i	Función
f_1	$x(1)^2 + x(2)^2$
f_2	$-1/(1 + x(1)^2 + x(2)^2) + 2$
f_3	$x(1)^2 + 1/2 * x(2)^2 + 3$
f_4	$x(1)^3 + x(2)^3 - 3 * x(1) - 3 * x(2) + 5$
f_5	$100 * (x(2) - x(1)^2)^2 + (1 - x(1))^2$
f_6	$-1/(1 + f1)$
f_7	$log(1 + f1)$
f_8	$2 + f1 + atan(f1)$
f_9	$f1 - cos(f1)$
f_{10}	$1 + \frac{x(1)^2}{x(2)^2}$

Cuadro 2: Funciones Objetivos, con $f1 = (1/2) * x^T * M * x$;

Búsqueda	**sucesión**	f_i	**óptimo**	**VOBJ**	**iter**	**tiempo**
Armijo	SBM1	$f1$	$(5{,}2037e^{-5}, 5{,}2037e^{-5})$	$5{,}4158e^{-9}$	100	6.71875
No monótona	SBM1	$f1$	$(3{,}8948e^{-5}, 3{,}8948e^{-5})$	$3{,}0339e^{-9}$	2	4.23435
Armijo	SBM2	$f1$	$(5{,}3958e^{-5}, 5{,}3958e^{-5})$	$5{,}8229e^{-9}$	2	12.9688
No monótona	SBM2	$f1$	$(2{,}8984e^{-5}, 2{,}8984e^{-5})$	$1{,}6802e^{-9}$	2	9.15625
Armijo	SBM1	$f2$	$(5{,}2221e^{-5}, 5{,}2221e^{-5})$	1	100	16.2188
No monótona	SBM1	$f2$	$(2{,}2421e^{-5}, 2{,}2421e^{-5})$	1	2	9.0937
Armijo	SBM2	$f2$	$(5{,}0704e^{-5}, 5{,}0704e^{-5})$	1	2	9.95312
No monótona	SBM2	$f2$	$(2{,}3531e^{-5}, 2{,}3531e^{-5})$	1	2	11,9375
Armijo	SBM1	$f3$	$(2{,}3123e^{-5}, 1{,}4564e^{-4})$	3	2	14.3438
No monótona	SBM1	$f3$	$(6{,}0691e^{-5}, 7{,}7448e^{-5})$	3	100	21.1094
Armijo	SBM2	$f3$	$(5{,}3983e^{-5}, 1{,}9670e^{-4})$	3	2	20.7188
No monótona	SBM2	$f3$	$(2{,}9064e^{-5}, 1{,}0313e^{-5})$	3	100	24.2344
Armijo	SBM1	$f4$	$(0{,}99, 0{,}99)$	1	2	7
No monótona	SBM1	$f4$	$(0{,}5, 0{,}5)$	2.25	100	532.156
Armijo	SBM2	$f4$	$(1{,}0041, 1{,}0041)$	1	2	21.2656
No monótona	SBM2	$f4$	$(0{,}5, 0{,}5)$	2.25	100	1096.31

Cuadro 3: Problemas

Armijo	SBM1	$f5$	$(0{,}9980, 0{,}9960)$	$3{,}9398e^{-8}$	100	386.562
No monótona	SBM1	$f5$	$(0{,}5, 0{,}5)$	6.5	100	561.281
Armijo	SBM2	$f5$	$(1{,}1010, 1{,}0200)$	$1{,}0336e^{-4}$	100	567.953
No monótona	SBM2	$f5$	$(0{,}5, 0{,}5)$	6.5	100	958.672
Armijo	SBM1	$f6$	$(1{,}0600e^{-4}, 1{,}0600e^{-4})$	-1	100	14.9219
No monótona	SBM1	$f6$	$(6{,}8494e^{-5}, 6{,}8494e^{-5})$	-1	2	10.1094
Armijo	SBM2	$f6$	$(1{,}0168e^{-4}, 1{,}0168e^{-4})$	-1	2	20.0625
No monótona	SBM2	$f6$	$(1{,}2520e^{-4}, 1{,}2520e^{-4})$	-1	2	16.1406
Armijo	SBM1	$f7$	$(1{,}0541e^{-4}, 1{,}0451e^{-4})$	$1{,}0922e^{-8}$	100	17.2344
No monótona	SBM1	$f7$	$(9{,}7310e^{-5}, 9{,}7310e^{-5})$	$9{,}4662e^{-8}$	2	8.6875
Armijo	SBM2	$f7$	$(1{,}2640e^{-4}, 1{,}2640e^{-4})$	$1{,}5978e^{-8}$	2	14.8125
No monótona	SBM2	$f7$	$(1{,}5391e^{-4}, 1{,}5391e^{-4})$	$2{,}3687e^{-8}$	2	12.6094
Armijo	SBM1	$f8$	$(5{,}2366e^{-5}, 5{,}2366e^{-5})$	2	100	11.6562
No monótona	SBM1	$f8$	$(3{,}9749e^{-5}, 3{,}9749e^{-5})$	2	2	7.4687
Armijo	SBM2	$f8$	$(5{,}9694e^{-5}, 5{,}9694e^{-5})$	2	2	12.3438
No monótona	SBM2	$f8$	$(3{,}1182e^{-5}, 3{,}1182e^{-5})$	2	2	12.1718
Armijo	SBM1	$f9$	$(1{,}0448e^{-4}, 1{,}0448e^{-4})$	-1	100	16.4219
No monótona	SBM1	$f9$	$(8{,}9195e^{-5}, 8{,}9195e^{-5})$	-1	2	7.8750
Armijo	SBM2	$f9$	$(1{,}4560e^{-4}, 1{,}4560e^{-4})$	-1	2	17.3906
No monótona	SBM2	$f9$	$(1{,}4500e^{-5}, 1{,}4500e^{-5})$	-1	2	10.500
Armijo	SBM1	$f10$	$(2{,}4156e^{-5}, 5{,}9565e^{-1})$	1	2	12.750
No monótona	SBM1	$f10$	$(-4{,}3503e^{-6}, 4{,}3101e^{-1})$	1	9	34.125
Armijo	SBM2	$f10$	$(-106, -47)$	5.9608	100	1494.88
No monótona	SBM2	$f10$	$(3{,}7487e^{-5}, 8{,}2335e^{-5})$	1.2073	100	1345.2

Cuadro 4: Problemas

3. Conclusiones

La presente investigación es considerada como un primer paso para resolver un problema abierto presentado por E.Hernandez, J.Campos y R. A. Castillo, [14], en que el método de punto proximal desarrollado con la casi-distancia generalizada sugiere que deben estudiarse métodos de punto proximal asociados a familias de distancias de Bregman que dependan de un parámetro, el método de punto proximal para sucesiones de funciones de Bregman convergente puntualmente presentado en esta investigación bajo condiciones de convexidad permite que la sucesión de funciones sea convergente uniformemente, este método contiene como casos particulares el método de punto proximal clásico y el método de punto proximal con distancias de Bregman, las pruebas de convergencia presentadas fueron inspiradas por [9].

Referencias

[1] A. Auslender; M. Teboulle; S. Ben-Tiba. Interior proximal and multiplier methods based on second order homogeneous kernels. *Mathematics of Operations Research*, 3:645–668, 1999.

[2] A. Iusem; M. Teboulle; B. Svaiter. Entropy-like proximal methods in convex programming. *Mathematics of Operations Research*, 4:790–814, 1994.

[3] J. Eckstein. Nonlinear proximal point algorithms using bregman functions with aplications to convex programming. *Mathematics of Operations Resaerchs*, 18:202–226, 1993.

[4] J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc.Math*, 93:97–116, 1976.

[5] C. Lemarechal; Hiriart-Urruty. Convex analysis and minimization algorithm ii. *Springer Verlag*, 1996.

[6] B. Martinet. Régularisation d'inéquations variationnelles par approximations successives. *Reue Francaise deInformatique et Recherche Operationelle*, 2:154–159, 1970.

[7] R.T. Rockafellar. Monotone operators and the proximal point algorithm in convex programming. *SIAM J. on Control and Optimization*, 14:877–898, 1976.

[8] H. Attouch; R. West. Ephigraphical analysis. *Ann. Ins. Poincaré: Anal. Nonlineair*, 1989.

[9] A. Iusem. Métodos de ponto proximal em optimização. *20 Coloquio Brasileiro de Matemática*, 1995.

[10] A. Iusem. Augmented lagrangian methods and proximal point methods for convex optimization. *Investigacion Operativa*, 8:11–50, 1997.

[11] C. Humes; J. Eckstein; P.J. Silva. Rescaling and stepsize selection in proximal methods using separable generalized distances. *SIAM Journal on Optimization*, 12:238–261, 2001.

[12] K. Kiwiel. Proximal minimizations methods with generalized bregman functions. *SIAM J. on Control and Optimization*, 35:1142–1168, 1997.

[13] C. Gonzaga; R. Castillo. Penalidades generalizadas e métodos de lagrangeano aumentado para promamacao nao linear. *Dsc. these, U.F.R.J.*, 1998.

[14] E.Hernandez; J.Campos; R. A. Castillo. A generalized like-distance in convex programming. *International Mathematical Forum*, 2:1811–1830, 2007.

[15] C. Gonzaga; E. Karas; L. Matioli; R. Castillo. An unified approach to multiplier and proximal methods. *Technical Report. Universidad Federal de Santa Catarina Brasil*, 2012.

[16] C. Quintana. Un estudio sobre métodos de multiplicadores y métodos de punto proximal. *Tesis Doctoral*, 2016.

[17] A. Larreal; E. Hernández. Método de punto proximal para sucesiones convergentes de funciones de bregman. caso convexo. *Tesis de Maestría UCLA, Barquisimeto*, 2014.

[18] R. Quintana. Un estudio sobre los métodos de punto proximal para familia de funciones de bregman dependientes de parámetros vectoriales. *Tesis Doctoral, UCLA*, page capitulo 6, 2018.

[19] S. Pan; J. Chen. Entropy-like proximal algorithms based on a second-order homogeneous distance function for quasi-conex programming. *Journal of Global Optimization*, 39:555–575, 2007.

Sobre los autores

Eibar Ramón Hernández: Doctor en Ciencias mención matemática. Profesor-Investigador en la Universidad Centroccidental Lisandro Alvarado, Venezuela. Correo: ehernand@ucla.edu.ve

Raquel Silvana Quintana Carlone: Magister en Matemática mención optimización. Profesora-Investigadora en la Universidad Centroccidental Lisandro Alvarado, Venezuela. Correo: raquel.quintana@ucla.edu.ve

Clavel María Quintana Carlone: Doctora en Matemática mención optimización. Profesora-Investigadora en la Universidad Centroccidental Lisandro Alvarado, Venezuela. Correo: clavel.quintana@ucla.edu.ve