PESQUIMAT 25(1): 22–35 (2022)
https://doi.org/10.15381/pesquimat.v25i1.23144

ISSN:1560-912X/ ISSN-E:1609-8439
Facultad de Ciencias Matemáticas – UNMSM

Método del Punto Proximal Inexacto Usando Cuasi-Distancias para Optimización de Funciones KL.

Erik A. Papa Quiroz [1] ***y Jose L. Huaman Ñaupa*** [2]

Resumen: Se introduce un algoritmo de punto proximal inexacto utilizando cuasi-distancias para dar solución a un problema de minimización en el espacio Euclideano. Este algoritmo ha sido motivado por el método proximal introducido por Attouch et al. [1] pero en este caso consideramos cuasi-distancias en vez de la distancia Euclidiana, funciones que satisfacen la desigualdad de Kurdyka-Lojasiewicz, errores vectoriales en el residual del punto crítico de los subproblemas proximales regularizados. Obtenemos bajo algunos supuestos adicionales la convergencia global de la sucesión generada por el algoritmo a un punto crítico del problema.

Palabras clave: Desigualdad de Kurdyka-Lojasewicz, cuasi-distancia, algoritmo de punto proximal.

Inexact Proximal Point Method Using Quasi-Distances for Optimization of KL Functions.

Abstract: An inexact proximal point algorithm using quasi-distances is introduced to give a solution of a minimization problem in the Euclidean space. This algorithm has been motivated by the proximal method introduced by Attouch, Bolte and Svaiter [1] but in this case we consider quasi-distance instead of the Euclidean distance, functions satisfying the Kurdyka-Lojasewicz inequality, vector errors in the critical point of the proximal subproblems. We obtain, under some additional assumptions, the global convergence of the sequence generated by the algorithm to a critical point of the problem.
Keywords: Kurdyka-Lojasiewicz inequality, quasi-distances, proximal point algorithm.

Recibido: 09/12/2021. *Aceptado:* 10/01/2022. *Publicado online:* 30/06/2022.

[1]Universidad Nacional Mayor de San Marcos, erikpapa@gmail.com
[2]Universidad Nacional Mayor de San Marcos, huamanjoselui@gmail.com

1. Introducción

El Algoritmo de Punto Proximal (APP) presentado en los años setenta por Martinet [10], y posteriormente muy estudiado por Rockafellar [15] con el objetivo de encontrar ceros de operadores monótonos con la propiedad de maximalidad, surgió inicialmente para encontrar una solución del problema

$$\min\{f(x) : x \in I\!R^n\}, \tag{1.1}$$

donde $f : I\!R^n \to I\!R$ es una función convexa. El APP genera una sucesión de puntos $\{x^k\}$ mediante el siguiente paso iterativo: sea $x^0 \in I\!R^n$ un punto inicial y dado x^k, se obtiene el siguiente punto x^{k+1} satisfaciendo la condición:

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

donde $\lambda_k > 0$ es el parámetro de regularización, $||.||$ denota la norma Euclidiana y la notación $\arg\min$ nos dice que x^{k+1} es un punto de mínimo global de $f(x) + \frac{\lambda_k}{2}||x - x^k||^2$ en todo el espacio $I\!R^n$. Si la sucesión $\{\lambda_k\}$ cumple la condición $\sum_{k=1}^{+\infty}(1/\lambda_k) = +\infty$, es posible demostrar que $\{f(x^k)\}$ converge al valor ínfimo de f; y si además el conjunto formado por todas las soluciones óptimas no es vacío, entonces $\{x^k\}$ converge a una solución óptima del problema. vea Güler [5].

Para aplicaciones del (APP) en Teoría de la Computación [16], Economia (elección del consumidor, funciones de utilidad) [6, 9] y Ciencias del Comportamiento (que incluye Inteligencia Artificial, Ciencias Administrativas, Procesos de Decisión, Filosofía, Psicología, Ciencias Políticas, Sociología, entre otras) es necesario considerar la no convexidad de la función f como también sustituir la norma Euclidiana $||x - x^{k-1}||$ en (1.2) por una cuasi-distancia $q(x, x^{k-1})$, ver por ejemplo [2, 4, 7, 8, 17]. Así, es de actual interés estudiar el APP usando cuasi-distancias:

$$x^{k+1} = \arg\min\{f(x) + \frac{\lambda_k}{2}q^2(x, x^k) : x \in I\!R^n\}. \tag{1.3}$$

La convergencia de la sucesión $\{x^k\}$ generada por la iteración (1.3) fue estudiada por la primera vez por Moreno et al. [11] y los autores probaron que si f satisface la condición de Kurdyka-Lojasiewicz (KL) sobre el conjunto de puntos críticos entonces toda sucesión acotada $\{x^k\}$, converge a un punto crítico del problema.

El 2015, los investigadores Bento y Soubeyran [3], motivados por los trabajos de Attouch [1] y Moreno [11] presentaron una versión inexacta de un algoritmo de suficiente descenso usando cuasi-distancias (donde el APP con cuasi-distancias es un caso particular) dado por la siguiente iteración: Para $k = 0, 1, 2, 3, \ldots$, dado el punto actual $x^k \in I\!R^n$, obtener un punto x^{k+1} y $\varphi^{k+1} \in I\!R^n$ satisfaciendo

$$f(x^{k+1}) \leq f(x^k) - \lambda_{k+1}(1 - \sigma)\Gamma(q(x^k, x^{k+1})) \tag{1.4}$$

$$\varphi^{k+1} \in \partial f(x^{k+1}) \tag{1.5}$$

$$\psi^{k+1} \in \partial q(x^k, .)(x^{k+1}) \tag{1.6}$$

$$||\varphi^{k+1}|| \leq b\Gamma'(q(x^k, x^{k+1}))||\psi^{k+1}||, \tag{1.7}$$

donde $\Gamma : [0, +\infty) \to [0, +\infty)$ satisface algunas condiciones adecuadas definidas en (2) y (3) de Bento y soubeyran [3], $\sigma \in (0, 1)$, ∂ denota la subdiferencial en el límite de f. Asumiendo que la sucesión $\{x^k\}$ es acotada, f satisface la condición de la desigualdad de Kurdyka-Lojasiewicz, los autores probaron que la sucesión converge a un punto crítico de f, estudiaron la tasa de convergencia de la sucesión e introducen una aplicación en Ciencias del Comportamiento.

En este trabajo introducimos una iteración inexacta del método del APP usando cuasi-distancias que considera errores vectoriales; para $k = 0, 1, 2, \ldots$, dado $x^k \in I\!R^n$, encontrar $x^{k+1} \in I\!R^n$ y $\varphi^{k+1} \in I\!R^n$ satisfaciendo

$$f(x^{k+1}) \leq f(x^k) - \tfrac{\lambda_{k+1}}{2} q(x^k, x^{k+1}) \tag{1.8}$$
$$\varphi^{k+1} \in \partial f(x^{k+1}) \tag{1.9}$$
$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1}), \tag{1.10}$$

donde

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}, \tag{1.11}$$

$\sigma > 0$ y

$$\psi^{k+1} \in \partial q(x^k, .)(x^{k+1}). \tag{1.12}$$

Para diferenciar nuestro trabajo del artículo de Bento y Soubeyran [3], realizaremos una comparación de los dos algoritmos. Observamos inicialmente que las condiciones (1.10) y (1.7) son diferentes, sin embargo, una condición suficiente y bien dificil de probar en cada iteración para que exista una relación estrecha entre ambas es que se cumpla que $\lambda_k ||\psi^{k+1}|| \geq 1$, $\sigma < 1$ and $\{\lambda_k\}$ sea acotado inferiormente. En efecto,

$$\begin{aligned} ||e^{k+1}||^2 = ||\varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}||^2 &\leq \sigma^2 q(x^k, x^{k+1})^2 \\ &\leq \sigma^2 (||\varphi^{k+1}||^2 + q(x^k, x^{k+1})^2) \\ &\leq \sigma^2 (||\varphi^{k+1}||^2 + ||\lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}||^2). \end{aligned}$$

Como $\sigma < 1$, entonces aplicando el Lemma 4.1 de Attouch y Bolte [1]

$$||\varphi^{k+1}|| \leq M \lambda_{k+1} q(x^k, x^{k+1}) ||\psi^{k+1}||.$$

Como $\{\lambda_k\}$ está acotado, obtenemos (1.7). Así, podemos concluir que los algoritmos inexactos propuestos por Bento y Soubeyran [3] y el algoritmo propuesto en este artículo son, en general, diferentes.

En este artículo estudiamos las propiedades de convergencia del APP inexacto dado por (1.8) - (1.12) y asumiendo propiedades sobre la función objetivo de ser propia, semicontinua inferior, coerciva y que se satisfaga la desigualdad de Kurdyka-Lojasiewicz, probamos que la sucesión generada por el algoritmo converge a un punto crítico de la función.

El principal aporte es la introducción de un nuevo error en métodos de punto proximal con cuasi-distancias con su respectivo análisis de convergencia para funciones que satisfacen la condición de Kurdyka-Lojasiewicz.

Este artículo está organizado de la siguiente manera: En la Sección 2 presentamos los elementos teóricos básicos necesarios para una lectura adecuada del artículo. En la Sección 3 el problema de optimización y sus hipótesis son introducidos, como también, el algoritmo proximal y sus propiedades teóricas de convergencia. En la Sección 4 presentamos las discusiones y en la Sección 5 las concluisones de la investigación.

2. Resultados Básicos

En este artículo $I\!R^n$ denota el espacio Euclideano, esto es, el espacio vectorial real n-dimensional dotado con la norma $\|.\|$ inducida por el producto interno canónico $\langle\, ,\, \rangle$, esto es, $\|x\| := \langle x, x\rangle^{1/2}$.

Definición 2.1 *Sea $C \subset I\!R^n$ no vacío. Se dice que C es un conjunto convexo si para cualquier $x, y \in C$ se tiene*

$$[x.y] := \{\alpha x + (1-\alpha)y;\ \alpha \in [0,1]\} \subset C.$$

Definición 2.2 *Dada una función $f : I\!R^n \to I\!R \cup \{\pm\infty\}$*

1. *El dominio efectivo (llamado también dominio) de f, es definido por* $\texttt{domf} := \{x \in I\!R^n : f(x) < +\infty\}$;
2. *f es propia si:*
   a) $\texttt{domf} \neq \emptyset$;
   b) $\forall\ x \in domf : f(x) > -\infty$
3. *f es semicontinua inferior en x^* si para toda sucesión $\{x^l\} \subset I\!R^n$ tal que $\lim\limits_{l\to+\infty} x^l = x^*$ se obtiene*

$$f(\bar{x}) \leq \liminf_{l\to+\infty} f(x^l).$$

La función f es semicontínua inferior (en $I\!R^n$) si es semicontínua inferior en todo punto de $I\!R^n$.

Definición 2.3 *Dado $\alpha \in I\!R$. El conjunto de nivel inferior de una función $f : I\!R^n \to I\!R \cup \{\pm\infty\}$, es dado por*

$$L_f(\alpha) = \{x \in I\!R^n; f(x) \leq \alpha\}.$$

Definición 2.4 *Decimos que una sucesión $\{x^k\}$ en $X \subset I\!R^n$ es crítica (con respecto al conjunto X) si:*

$$\lim_{k\to\infty} \| x^k \| = +\infty \quad o \quad \lim_{k\to\infty} x^k = x \in \overline{X} - X.$$

La función $f : X \subset I\!R^n \to I\!R \cup \{\pm\infty\}$ es coerciva en X si para toda sucesión crítica $\{x^k\}$ se tiene

$$\lim_{k\to\infty} supf(x^k) = +\infty,$$

donde

$$\lim_{k\to\infty} supf(x^k) = \inf_{n\in I\!N} \sup_{k\geq n} f(x^k).$$

Corolario 2.1 *([12], página 38). Dado un conjunto no vacio C y $f : C \subset I\!R^n \to I\!R$. Si f es coerciva y semicontinua inferior en C, entonces existe un punto de mínimo global de f en C.*

Teorema 2.1 *([14], página 9). Sea $f : I\!R^n \to \overline{I\!R}$ una función, los siguientes enunciados son equivalentes:*

i) *f es semicontinua inferior sobre $I\!R^n$;*

ii) *$epi(f)$ es cerrado en $I\!R^n \times I\!R$;*

iii) *$\forall \alpha \in I\!R$ el conjunto $L_f(\alpha)$ es cerrado en $I\!R^n$.*

Definición 2.5 *Sea $f : I\!R^n \to I\!R \cup \{+\infty\}$ una función propia semicontinua inferior y $x \in I\!R^n$*

(1) El subdiferencial de Fréchet de la función f en el punto x, denotado por $\widehat{\partial} f(x)$, es el conjunto de puntos $s \in I\!R^n$, tales que

$$f(y) \geq f(x) + \langle s, y-x \rangle + \varphi(||y-x||) \text{ donde } \lim_{\substack{x\neq y \\ y\to x}} \frac{\varphi(||y-x||)}{||y-x||} = 0$$

o equivalentemente

$$\widehat{\partial} f(x) = \begin{cases} \left\{ s \in I\!R^n : \liminf\limits_{\substack{x\neq y \\ y\to x}} \frac{f(y)-f(x)-\langle s,y-x\rangle}{||y-x||} \geq 0 \right\}, & si\ x \in domf \\ \emptyset, & si\ x \notin domf \end{cases}$$

(2) El subdiferencial en el limite de f en $x \in I\!R^n$, denotado por $\partial f(x)$, es definido como sigue

$$\partial f(x) := \left\{ s \in I\!R^n : \exists\, x^k \to x, f(x^k) \to f(x), s^k \in \widehat{\partial} f(x^k) \to s \right\}.$$

Observación 2.1 *Para una función $f : I\!R^n \to I\!R \cup \{+\infty\}$ y un punto $\bar{x}$ donde $f(\bar{x}) < +\infty$, los conjuntos $\partial f(\bar{x})$ y $\widehat{\partial} f(x)$ son cerrados, $\widehat{\partial} f(\bar{x})$ es convexo y $\widehat{\partial} f(\bar{x}) \subset \partial f(\bar{x})$.*

Proposición 2.1 *([14], ejercicio 10.10 página 431) Si f_1 es localmente Lipschitz continua en $\bar{x}$, f_2 es semicontinua inferior y propia con $f_2(\bar{x})$ finito, entonces*

$$\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x}).$$

Proposición 2.2 *Si $f : I\!R^n \to I\!R \cup \{+\infty\}$ tiene un mínimo local en $\bar{x} \in domf$ entonces $0 \in \widehat{\partial} f(\bar{x})$.*

Lema 2.1 *([13], Lema 2, página 44) Dadas $\{v^k\}, \{\lambda^k\}$ y $\{\beta^k\}$ tres sucesiones de valores reales no negativas satisfaciendo la desigualdad*

$$v^k \leq (1+\lambda^k) v^{k-1} + \beta^k,$$

donde $\sum_{k=1}^{+\infty} \beta^k < +\infty$ y $\sum_{k=1}^{+\infty} \lambda^k < +\infty$. Entonces, $\{v^k\}$ es convergente.

2.1. Cuasi-Distancia

En esta subsección se presenta el concepto de cuasi-distancia y algunas de sus principales propiedades.

Definición 2.6 *Sea X un conjunto no vacío del espacio Euclidiano $I\!R^n$. Una aplicación $q : X \times X \to I\!R_+$ es llamado una cuasi-distancia en X si para todo $x, y, z \in X$*

(1) $q(x,y) = q(y,x) = 0 \Leftrightarrow x = y$.

(2) $q(x,y) \leq q(x,z) + q(z,y)$.

En todo este artículo, consideremos la siguiente condición sobre la cuasi-distancia: existen constantes positivas α y β tal que

$$\alpha||x-y|| \leq q(x,y) \leq \beta||x-y||. \tag{2.13}$$

Proposición 2.3 *(Moreno [11], Proposición 3.6) Sea $q : I\!R^n \times I\!R^n \to I\!R_+$ una cuasi-distancia que verifica (2.13). Entonces para cada $\bar{z} \in I\!R^n$ las funciones $q(\bar{z}, .)$ y $q(., \bar{z})$ son Lipschitzianas.*

Proposición 2.4 *(Moreno [11], Proposición 3.7) Sea $\bar{z} \in I\!R^n$. Si q verifica (2.13) entonces $q^2(\bar{z}, .)$ y $q^2(., \bar{z})$ son funciones localmente Lipschitz continua sobre $I\!R^n$.*

Proposición 2.5 *(Moreno [11], Observación 5) Sea $\bar{z} \in I\!R^n$. Si q verifica (2.13) entonces $q^2(\bar{z}, .)$ y $q^2(., \bar{z})$ son funciones coercivas.*

Lema 2.2 *(Moreno [11], Lema 3.8) Sea $\bar{z} \in I\!R^n$ y $\lambda > 0$. Si $f : I\!R^n \to I\!R \cup \{+\infty\}$ es acotada inferiormente y $q^2(\bar{z}, .)$ es una función coerciva, entonces la función $f + \lambda q^2(\bar{z}, .) : I\!R^n \to I\!R \cup \{+\infty\}$ es coerciva.*

3. Algoritmo Proximal Inexacto

Consideremos el siguiente problema de optimización:

$$\text{mín}\{f(x) : x \in I\!R^n\}, \tag{3.14}$$

donde la función $f : I\!R^n \to I\!R \cup \{+\infty\}$ es extendida y satisface las siguientes condiciones:

($\mathbf{H_1}$) f es propia y semicontinua inferior.

($\mathbf{H_2}$) f es coerciva, esto es, $\lim\limits_{||x|| \to +\infty} f(x) = +\infty$

($\mathbf{H_3}$) $q : I\!R^n \times I\!R^n \to I\!R_+$ es una cuasi-distancia satisfaciendo: existe un $\alpha > 0$ y $\beta > 0$ tal que

$$\alpha ||x - y|| \leq q(x, y) \leq \beta ||x - y||. \tag{3.15}$$

3.1. Algoritmo Proximal Inexacto (API)

Sea $q(., .)$ una cuasi-distancia definida sobre $I\!R^n$ satisfaciendo la hipótesis (H_3) y f una función objetivo del problema (3.14) satisfaciendo las hipótesis (H_1) y (H_2).

Inicialización: Elegimos una sucesión acotada de parámetros positivos $\{\lambda_k\}$, un valor real $\sigma > 0$ y de forma arbitraria tomamos un punto de inicialización

$$x^0 \in I\!R^n. \tag{3.16}$$

Paso Primcipal: Para valores de $k = 0, 1, 2, ...,$ dado $x^k \in I\!R^n$, encontrar $(x^{k+1}, \varphi^{k+1}) \in I\!R^n \times I\!R^n$ tal que

$$f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1}) \tag{3.17}$$

$$\varphi^{k+1} \in \partial f(x^{k+1}) \tag{3.18}$$

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1}) \tag{3.19}$$

donde

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1} \tag{3.20}$$

y $\psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$.

Criterio de Finalización: Si $x^{k+1} = x^k$ o $0 \in \partial f(x^{k+1})$, entonces finalizar. Si no se satisface el criterio anterior, actualizar $k \leftarrow k + 1$ e ir al Paso Principal.

3.2. Existencia de los iterados

Teorema 3.1 *Bajo las hipótesis* $(\mathbf{H_1})$-$(\mathbf{H_3})$, *para cada* k, *existen* (x^{k+1}, φ^{k+1}) *satisfaciendo (3.17)-(3.20), esto es, (API) está bien definido.*

Demostración. Dado $x^k \in I\!R^n$ y considere el problema $\min\{f(x) + \frac{\lambda_{k+1}}{2} q^2(x^k, x) : x \in I\!R^n\}$. De la semicontinuidad inferior de f y $q^2(x^k, .)$, y la coercitividad de f tenemos que existe x^{k+1} tal que

$$f(x^{k+1}) + \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1}) \leq f(x) + \frac{\lambda_{k+1}}{2} q^2(x^k, x),$$

para todo $x \in I\!R^n$. Tomando $x = x^k$ tenemos (3.17). Por otro lado, de la condición de optimalidad, véase la Proposición 2.2, tenemos que.

$$0 \in \hat{\partial}\left(f(.) + \frac{\lambda_{k+1}}{2} q^2(x^k, .) \right)(x^{k+1}).$$

Por la Proposición 2.1 se obtiene (3.18), (3.19) y (3.20) con $e^{k+1} = 0$.

3.3. Propiedades de convergencia para el caso general

Proposición 3.1 *Bajo las hipótesis* $(\mathbf{H_1}), (\mathbf{H_2})$, $(\mathbf{H_3})$, $0 < \lambda_- < \lambda_{k+1} < \lambda_+$ *y si el algoritmo no finaliza tenemos que la sucesión* $\{x^k\}$ *generada por (API) satisface:*

(i) $\{f(x^k)\}$ *es una sucesión convergente.*

(ii) $q(x^k, x^{k+1})$ *converge a cero.*

(iii) $\{x^k\}$ *es una sucesión acotada.*

(iv) Si x^{k_j} *converge a* $\bar{x}$, *entonces* $\{x^{k_j+1}\}$ *converge a* $\bar{x}$.

Prueba.

(i) De (3.17) tenemos

$$f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1}).$$

Dividiendo por λ_{k+1} y asociando tenemos

$$(1/2)\, q^2(x^k, x^{k+1}) \leq \frac{1}{\lambda_{k+1}} f(x^k) - \frac{1}{\lambda_{k+1}} f(x^{k+1}). \tag{3.21}$$

Por lo tanto, de (3.21) tenemos:

$$f(x^{k+1}) < f(x^k). \tag{3.22}$$

Del Lema 2.1 obtenemos que $\{f(x^k)\}$ converge.

(ii) De la acotación de $\{\lambda_k\}$, la convergencia de $\{f(x^k)\}$ y de (3.21) obtenemos

$$\lim_{k\to\infty} q(x^k, x^{k+1}) = 0.$$

(iii) De (3.22) obtenemos:

$$f(x^k) < f(x^0),$$

para todo $k = 1, ...$ Así $x^k \in L_f(a_0)$ donde $a_0 = f(x^0)$. Como $L_f(a_0)$ es acotada pues f es una función coerciva, entonces se concluye que $\{x^k\}$ es acotada.

(iv) Usando la propiedad de la desigualdad triangular para la norma Euclidiana, tenemos

$$||x^{k_j+1} - \bar{x}|| \leq ||x^{k_j+1} - x^{k_j}|| + ||x^{k_j} - \bar{x}||.$$

Por la condición (3.15) se tiene que existe $\beta > 0$ tal que

$$||x^{k_j+1} - \bar{x}|| \leq \beta q(x^{k_j}, x^{k_j+1}) + ||x^{k_j} - \bar{x}||.$$

Como $q(x^{k_j}, x^{k_j+1})$ converge a cero por item(ii), se tiene $\{x^{k_j+1}\}$ converge a $\bar{x}$.

Para probar que todo punto de acumulación es un punto crítico, introduciremos la siguiente hipótesis

$(\mathbf{H_1})'$: f es propia, semicontinua inferior y continua sobre su dominio `domf`.

Teorema 3.2 *Asumiendo las hipótesis* $(\mathbf{H_1})', (\mathbf{H_2}), (\mathbf{H_3})$, *y que* $0 < \lambda_- < \lambda_k < \lambda_+$, *cada punto de acumulación de* $\{x^k\}$, *generada por el (API), es un punto crítico de (3.14).*

Proof. Consideremos $\bar{x} \in I\!R^n$ un punto de acumulación de la sucesión $\{x^k\}$, así existe $\{x^{k_j}\}$ tal que $x^{k_j} \to \bar{x}$ y por la Proposición 3.1, (iv), $x^{k_j+1} \to \bar{x}$. De (3.20) y la desigualdad triangular tenemos

$$||\varphi^{k_j+1}|| \leq ||e^{k_j+1}|| + \lambda_{k_j+1} q(x^{k_j}, x^{k_j+1})||\psi^{k_j+1}|| \leq \left(\sigma + \lambda_{k_j+1}||\psi^{k_j+1}||\right) q(x^{k_j}, x^{k_j+1}), \quad (3.23)$$

donde la última desigualdad viene de (3.19). Como $q(x^k, .)$ es localmente Lipschitz en $I\!R^n$ entonces $\partial q(x^k, .)$ es localmente acotado, ver Proposición 9,13 de Rockafellar y Wets [14], entonces del Lemma 5.1 de Moreno et al. [11], existe $M > 0$ tal que $||\psi^{k_j+1}|| \leq M, \forall j$. Así, la desigualdad anterior (3.23) implica que $||\varphi^{k_j+1}|| \leq (\sigma + \lambda_+ M)\, q(x^{k_j}, x^{k_j+1})$. Por lo tanto, existe $x^{k_j+1} \to \bar{x}$, with $f(x^{k_j+1}) \to f(\bar{x})$ y $\varphi^{k_j+1} \to 0$ con $\varphi^{k_j+1} \in \partial f(x^{k_j+1})$, esto implica que $0 \in \partial f(\bar{x})$. ∎

3.4. Convergencia para funciones que satisfacen la condición de Kurdyka-Lojasiewicz (KL)

En esta subsección probaremos la convergencia de la sucesión $\{x^k\}$ para funciones que satisfacen la propiedad de Kurdyka-Lojasiewicz. Nuestros resultados son motivados del paper de Moreno et al. [11], Attouch et al. [1] y Bento y Soubeyran [3].

Dado $\bar{r} \in (0, +\infty]$, definamos

$$K(0, \bar{r}) := \{\phi \in C[0, \bar{r}) \cap C^1(0, \bar{r}) : \phi(0) = 0 \text{ y } \phi'(r) > 0, \forall r \in (0, \bar{r})\},$$

donde $C[0, \bar{r})$ (respectivamente $C^1(0, \bar{r})$) denota el conjunto de funciones continuas sobre $[0, \bar{r})$ (respectivamente, funciones de clase C^1 sobre $(0, \bar{r})$). Para $\eta_1 < \eta_2 \leq +\infty$ definimos el siguiente conjunto

$$[\eta_1 < f < \eta_2] = \{x \in I\!R^n : \eta_1 < f(x) < \eta_2\}.$$

Definición 3.1 *La función f se dice que tiene la propiedad de Kurdyka-Lojasiewicz en $\bar{x} \in dom(\partial f)$ si existe un $\eta \in (0, +\infty]$, una vecindad U de $\bar{x}$ y una función cóncava $\varphi \in K(0,\eta)$, tal que para todo $x \in U \cap [f(\bar{x}) < f < f(\bar{x}) + \eta]$ la desigualdad de Kurdyka-Lojasiewicz cumple*

$$\varphi'(f(x) - f(\bar{x}))dist(0, \partial f(x)) \geq 1 \tag{3.24}$$

Es fácil probar que $\{f(x^k)\}$ converge a $\inf_{k\geq 0} f(x^k)$.
Dado $x \in I\!R^n$ y $\epsilon > 0$ fijo, denotamos por $B_q(x,\epsilon) = \{y \in I\!R^n; q(x,y) < \epsilon\}$. En particular si q es la distancia Euclideana, $B_q(x,\epsilon)$ es denotada como $B(x,\epsilon)$.

Teorema 3.3 *Sea $\{x^k\}$ la sucesión generada por el (API), f es una función que satisface $(\mathbf{H_1})'$ y $(\mathbf{H_2})$, q es una cuasi-distancia que satisface $(\mathbf{H_3})$, $\bar{x}$ es un punto de acumulación de la sucesión $\{x^k\}$ donde f satisface la desigualdad de Kurdyka-Lojasiewicz. Sea $U \subset I\!R^n$ una vecindad de $\bar{x}$, $\eta > 0$ y $\phi : [0,\eta) \to I\!R^+$ una función cóncava que satisfacen las condiciones de la definición de la condición de Kurdyka-Lojasiewicz. Dados $\delta > 0$ y $r \in (0,1)$ tal que $B(\bar{x}, \frac{\delta}{\alpha}) \subset U$ y $\overline{M} = 2\left(\frac{\sigma + M\lambda_+}{r(1-r)\lambda_-}\right)$ entonces existe un $N_0 \in I\!N$ satisfaciendo*

$$f(\bar{x}) < f(x^k) < f(\bar{x}) + \eta, \quad \textit{para todo} \ \ k \geq N_0 \tag{3.25}$$

$$q(\bar{x}, x^{N_0}) + \left(\frac{r}{1-r}\right)\sqrt{\frac{2}{\lambda_-}\left[f(x^{N_0-1}) - f(\bar{x})\right]} + \overline{M}\left[\phi(f(x^{N_0})) - \phi(f(\bar{x}))\right] + \sqrt{\frac{2}{\lambda_-}\left[f(x^{N_0}) - f(\bar{x})\right]} < \delta \tag{3.26}$$

y

$$x^{N_0+j} \in B_q(\bar{x}, \delta) \ \textit{con} \ j = 1, 2, \ldots \tag{3.27}$$

Además, la sucesión $\{x^k\}$ converge al punto crítico $\bar{x}$ y en particular

$$\sum_{j=1}^{+\infty} q(x^{N_0+j}, x^{N_0+j+1}) < +\infty. \tag{3.28}$$

Prueba 3.1 *Sin pérdida de generalidad suponemos que $f(x) := f(x) - \inf_{k\geq 0} f(x^k)$. Así, de la Proposición 3.1, (i), obtenemos que $\lim\limits_{k\to\infty} f(x^k) = 0$ y así, existe un $k_0 \in I\!N$ tal que $0 < f(x^k) < \eta$, para todo $k \geq k_0$.*
Definimos la sucesión real $\{b^k\}$ dado por

$$b^k = q(\bar{x}, x^k) + \left(\frac{1}{1-r}\right)\sqrt{\frac{2}{\lambda_-} f(x^k)} + \sqrt{\frac{2}{\lambda_-} f(x^{k-1})} + \overline{M}[\phi(f(x^{k+1}))].$$

0 es un punto de acumulación de la sucesión $\{b^k\}$, de aquí, existe $j_0 \in I\!N$ tal que para todo $j \geq j_0$ se tiene $b_{k_j} < \delta$. Definiendo $N_0 = max\{k_{j_0}, k_0\}$, las condiciones (3.25) y (3.26) se cumplen.
Observe que (3.26) implica que $x^{N_0} \in B_q(\bar{x}, \delta)$ y por la hipótesis $(\mathbf{H_3})$ esto a su vez implica que $x^{N_0} \in B(\bar{x}, \delta/\alpha) \subset U$. Por lo tanto de (3.25) tenemos que $x^{N_0} \in U \cap [0 < f < \eta]$ y por la condición de Kurdyka-Lojasiewicz se tiene

$$\phi'(f(x^{N_0}))\,||\varphi^{N_0}|| \geq 1, \ \textit{con} \ \varphi^{N_0} \in \partial f(x^{N_0}).$$

Así,

$$\phi'(f(x^{N_0})) \geq \frac{1}{||\varphi^{N_0}||}. \tag{3.29}$$

Ahora, de (3.18) y (3.20) tenemos que $\varphi^k \in \partial f(x^k)$ *y*

$$||\varphi^k|| \leq ||e^k|| + \lambda_k q(x^{k-1}, x^k)||\psi^k||, \tag{3.30}$$

donde $\psi^k \in \partial q(x^{k-1}, .)(x^k)$.
De (3.19) en (3.30) obtenemos

$$||\varphi^k|| \leq \sigma q(x^{k-1}, x^k) + \lambda_k q(x^{k-1}, x^k)||\psi^k|| \quad \textit{para } k = 1, 2, \ldots \tag{3.31}$$

Como $\lambda_- < \lambda_k < \lambda_+$ *y del Lema 5.1 de Moreno et al. [11] que afirma* $||\psi^k|| \leq M, k = 1, 2, \ldots,$ *se tiene en (3.31)*

$$||\varphi^k|| \leq (\sigma + \lambda_+ M)\, q(x^{k-1}, x^k) \quad \textit{para } k = 1, 2, \ldots \tag{3.32}$$

De (3.32) y (3.29) obtenemos para $k = N_0$

$$\phi'(f(x^{N_0})) \geq \frac{1}{(\sigma + \lambda_+ M)\, q(x^{N_0-1}, x^{N_0})}. \tag{3.33}$$

Usando la caracterización de la diferenciabilidad de la función cóncava $\phi(s)$ *para* $s, t \in (0, \eta)$ *tenemos*

$$\phi(t) - \phi(s) \geq \phi'(t)(t - s). \tag{3.34}$$

Tomando $s = f(x^{N_0+1})$ *y* $t = f(x^{N_0})$ *en (3.34) tenemos*

$$\phi(f(x^{N_0})) - \phi(f(x^{N_0+1})) \geq \phi'(f(x^{N_0}))[f(x^{N_0}) - f(x^{N_0+1})]. \tag{3.35}$$

De la condición (3.17) obtenemos:

$$\phi(f(x^{N_0})) - \phi(f(x^{N_0+1})) \geq \phi'(f(x^{N_0}))[\frac{1}{2}\lambda_{N_0+1} q^2(x^{N_0}, x^{N_0+1}))]. \tag{3.36}$$

Usando (3.33) en (3.36)

$$\phi(f(x^{N_0})) - \phi(f(x^{N_0+1})) \geq \frac{\frac{1}{2}\lambda_{N_0+1} q^2(x^{N_0}, x^{N_0+1})}{(\sigma + \lambda_+ M)\, q(x^{N_0-1}, x^{N_0})}.$$

Factorizando

$$\phi(f(x^{N_0})) - \phi(f(x^{N_0+1})) \geq \frac{q(x^{N_0}, x^{N_0+1})}{q(x^{N_0-1}, x^{N_0})} \left(\frac{\frac{1}{2}\lambda_{N_0+1} q(x^{N_0}, x^{N_0+1})}{\sigma + \lambda_+ M} \right). \tag{3.37}$$

Tomando $r \in (0, 1)$ *consideramos dos casos:*
Caso 1: Si $\frac{q(x^{N_0}, x^{N_0+1})}{q(x^{N_0-1}, x^{N_0})} \geq r$ *entonces tenemos que (3.37) implica*

$$q(x^{N_0}, x^{N_0+1}) \leq 2 \left(\frac{\sigma + \lambda_+ M}{r\lambda_{N_0+1}} \right) [\phi(f(x^{N_0})) - \phi(f(x^{N_0+1}))].$$

Caso 2: Ahora, si $\frac{q(x^{N_0}, x^{N_0+1})}{q(x^{N_0-1}, x^{N_0})} \leq r$, *entonces*

$$q(x^{N_0}, x^{N_0+1}) \leq r q(x^{N_0-1}, x^{N_0}).$$

De ambos casos obtenemos

$$q(x^{N_0}, x^{N_0+1}) \le rq(x^{N_0-1}, x^{N_0}) + 2\left(\frac{\sigma + \lambda_+ M}{r\lambda_{N_0+1}}\right)[\phi(f(x^{N_0})) - \phi(f(x^{N_0+1}))].$$

Finalmente como $\lambda_- < \lambda_k < \tilde{\lambda}$ *se obtiene*

$$q(x^{N_0}, x^{N_0+1}) \le rq(x^{N_0-1}, x^{N_0}) + 2\left(\frac{\sigma + \tilde{\lambda} M}{r\lambda_-}\right)[\phi(f(x^{N_0})) - \phi(f(x^{N_0+1}))] \tag{3.38}$$

Probaremos por inducción que $x^{N_0+j} \in B_q(\bar{x}, \delta)$ *para* $j \in I\!N$
Para $j = 1$. *Por la desigualdad triangular se tiene*

$$q(\bar{x}, x^{N_0+1}) \le q(\bar{x}, x^{N_0}) + q(x^{N_0}, x^{N_0+1}).$$

Usando la desigualdad (3,17) *obtenemos*

$$q(\bar{x}, x^{N_0+1}) \le q(\bar{x}, x^{N_0}) + \sqrt{\frac{2}{\lambda_-}(f(x^{N_0}) - f(x^{N_0+1})} < \delta,$$

donde la última desigualdad es debido a (3.26), lo cual establece que $x^{N_0+1} \in B_q(\bar{x}, \delta)$.
Supongamos que (3.27) cumple para $N_0 + 1, \dots, N_0 + j - 1$. *Probaremos que se cumple para* $N_0 + j$.
Observemos primero que de forma análoga a lo realizado para el punto x^{N_0}, *podemos verificar que (3.38) cumple para* $N_0 + 1, \dots, N_0 + j - 1$ *y por lo tanto*

$$\sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1}) \le r \sum_{k=N_0}^{N_0+j-1} q(x^{k-1}, x^k) + 2\left(\frac{\sigma + \lambda_+ M}{r\lambda_-}\right) \sum_{k=N_0}^{N_0+j-1} [\phi(f(x^k)) - \phi(f(x^{k+1}))].$$

Restando a ambas desigualdades el término $-r\sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1})$ *y usando la no negatividad de la cuasi-distancia obtenemos*

$$(1-r)\sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1}) \le rq(x^{N_0-1}, x^{N_0}) + 2\left(\frac{\sigma + \lambda_+ M}{r\lambda_-}\right)[\phi(f(x^{N_0})) - \phi(f(x^{N_0+j}))].$$

Dividiendo por $(1-r)$ *y denotando* $\overline{M} = 2\left(\frac{\sigma + M\lambda_+}{r(1-r)\lambda_-}\right)$.

$$\sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1}) \le \left(\frac{r}{1-r}\right) q(x^{N_0-1}, x^{N_0}) + \overline{M}[\phi(f(x^{N_0})) - \phi(f(x^{N_0+j}))] \tag{3.39}$$

Estamos ahora en condiciones de probar que la propiedadd es verdad para $N_0 + j$.
De la desigualdad triangular, tenemos

$$\begin{aligned} q(\bar{x}, x^{N_0+j}) &\le q(\bar{x}, x^{N_0}) + q(x^{N_0}, x^{N_0+1}) + \ldots + q(x^{N_0+j-1}, x^{N_0+j}) \\ &\le q(\bar{x}, x^{N_0}) + \sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1}) \\ &\le q(\bar{x}, x^{N_0}) + \left(\frac{r}{1-r}\right) q(x^{N_0}, x^{N_0+1}) + \overline{M}[\phi(f(x^{N_0})) - \phi(f(x^{N_0+j}))] \end{aligned}$$

donde la última desigualdad es por (3.39).
Usando el hecho que $-\phi(f(x^{N_0+j})) < 0$*, obtenemos*

$$q(\bar{x}, x^{N_0+j}) \leq q(\bar{x}, x^{N_0}) + \left(\frac{r}{1-r}\right)\sqrt{\frac{2}{\lambda_-}f(x^{N_0-1})} + \overline{M}[\phi(f(x^{N_0}))]$$

la cual, de (3.26), nos permite concluir la prueba por inducción.
Observe que de (3.39) cuando $j \to \infty$ *se obtiene (3.28). Por otro lado, por* $(\mathbf{H_3})$ *se tiene* $\alpha||x^k - x^{k+1}|| \leq q(x^k, x^{k+1})$*, combinado con la condición (3.28) se tiene*

$$\sum_{k=N_0}^{\infty} ||x^k - x^{k+1}|| < +\infty.$$

Concluimos que $\{x^k\}$ *es una sucesión de Cauchy y por lo tanto* $\{x^k\}$ *es convergente a* $\bar{x}$ *que es un punto crítico según Teorema 3.2.*

4. Discusiones

El presente artículo presenta contribuciones teóricas importantes con respecto a la convergencia del método del punto proximal sustituyendo la distancia Euclidiana por una cuasi-distancia para la minimización de funciones que satisfazen la desigualdad de Kurdyka-Lojasiewicz y complementa los estudios realizados por Bento y Soubeyran [3] ya que presentamos otro tipo de error y probamos la convergencia de la sucesión generada por el algoritmo propuesto a un punto crítico del problema. Como discutimos en la introducción estos dos tipos de errores son en general diferentes, sin embargo el error estudiado en este artículo es más natural y específico para métodos de punto proximal.

5. Conclusiones

- Se presenta resultados de convergencia del (API) para solucionar problemas de minimización donde la función objetivo satisface la condición de Kurdyka-Lojasiewicz y una condición de coercividad.
- Este trabajo motiva a investigar si el (API) se puede generalizar para funciones más generales.
- En un futuro trabajo se espera estudiar la velocidad de convergencia del algoritmo propuesto como también de generalizar los resultados para una clase mayor de distancias llamadas $w-$distancias, ver el artículo reciente de Soubeyran y Souza, [18].

6. Agradecimientos

Deseamos manifestar nuestra gratitud a los revisores del artículo y a los miembros del Instituto de Investigación de la FCM-UNMSM, en particular a su Director Dr. Edinson Montoro Alegre, por la gestión en el proceso de sumisión, revisión y publicación de este trabajo de investigación.

Referencias bibliográficas

[1] Attouch, H., Bolte, J. y Svaiter, B. (2013). Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. Mathematical Programming, Vol. 137, 1-2: 91-129.

[2] Bao, T., Mordukhovich, B. y Soubeyran, A. (2015). Variational Analysis in Psychological Modeling. Journal of Optimization Theory and Applications, 164(1), 290-315.

[3] Bento, G. y Soubeyran, A. (2015). A generalized inexact proximal point method for nonsmooth functions that satisfies Kurdyka Lojasiewicz inequality. Set-Valued and Variational Analysis, Vol. 23,3: 501-517.

[4] Da Cruz Neto, J., Oliveira, P., Soares, P.(2014). Proximal Point Method on Finslerian Manifolds and the .Effort-Accuracy"Trade-off, JOTA, vol 162 (3):873-891.

[5] Güler, O. (1991). On the convergence of the proximal point algorithm for convex minimization. SIAM Journal on Control and Optimization, Vol. 29, 2: 403-409.

[6] Hu, Ch., Lai, Sh. y Lai, Ch. (2020). Investigations to the price evolutions of goods exchange with CES utility functions. Physica, 549, 1-19.

[7] Janeiro, I.(2010). Motivational Dynamics in the Development of Career Attitudes Among Adolescents. Journal of Vocational Behavior. 76(2), 170-177.

[8] Hirschi, A. y Dauwalder, J. (2015). Dynamics in Career development. Personal and Organizational Perspective. In Book: Handbook of the Life Design paradigm: From practice to theory, from theory to practice. (pp 27-39) Boston, MA: Hogrefe.

[9] Levin, V.(1991). Some applications of set-valued mappings in mathematical economics. Journal of Mathematical Economics 20, 69-87.

[10] Martinet, B. (1970). Breve communication. régularisation d'inéquations variationnelles par approximations successives. Revue francaise. 4: 154-158.

[11] Moreno, F., Oliveira, P. y Soubeyran, A. (2012). A proximal algorithm with quasi distance Application to habit's formation. Optimization, Vol. 61, 12: 1383-1403.

[12] Papa Quiroz, E.A. (2019). Optimización matemática y computacional. Dhyblassoft E.I.R.L, Lima, Vol., 1.

[13] Polyak, B. (1987).Introduction to optimization, Series of Translations in Mathematics and Engineering. New York. Optimization Software Publications Division Inc.

[14] Rockafellar, R. y Wets,, J. (2009). Variational analysis, Springer Science & Business Media, Vol. 317.

[15] Rockafellar, R. (1976). Monotone operators and the proximal point algorithm. SIAM journal on control and optimization, Vol. 14, 5: 877-898.

[16] Romaguera y Sanchis (2003). Applications of utility functions defined on quasi-metric spaces. J. Math. Anal. Appl. 283, 219-235.

[17] Soubeyran, A. (2010). Variational Rationality and the unsatis?ed man: Routines and the course pursuit between aspirations, capabilities and beliefs. GREQAM, Aix Marseille University.

[18] Soubeyran, A. y Souza, J. (2020). General descent method using w-distance. application to emergence of habits following worthwhile moves. J. Nonlinear Var. Anal. 4 (2), 285-300.