

MEDIA PROXIMAL Y LAS ENVOLTURAS DE MOREAU Y GOEBEL

Marlo Carranza ¹, Tomas Nuñez ²

Resumen: Se tiene una función $\mathcal{P}(f, \lambda)$ llamada el promedio o media proximal de dos o más funciones, el cual que tiene la ventaja que dichas funciones pueden tener sus dominios disjuntos o no. Entre sus aplicaciones se destaca que algunos esquemas de aproximación, se pueden expresar en términos de ella y que a partir de ella, también se pueden lograr otros esquema de aproximación, que además, pueden ser autoduales y en este caso es sencillo probar este hecho.

Palabras clave : Análisis convexo, media proximal, conjugada de Fenchel, envoltura de Moreau, envoltura de Goebel.

Abstract: It has a function $\mathcal{P}(f, \lambda)$ call proximal half the average of two or more functions, which has the advantage that these functions may be disjoint or not their domains. Applications include highlights some approximation schemes, can be expressed in terms of it and from it, you can also achieve other approximation scheme, which also can be autoduales and in this case it is easy to prove this fact.

Keywords: : Convex analysis, proximal average, Fenchel conjugate, Moreau envelope, Goebel envelope.

1. Introducción

Dadas dos funciones f_0 y f_1 podemos construir una función $P(x, \lambda)$, de tal manera que $P(x; 0) = f_0(x)$ y $P(x; 1) = f_1(x)$, el método común suele ser el promedio aritmético es decir

$$\lambda_1 f_0(x) + \lambda_2 f_2(x) \quad (1)$$

tal que λ_1 y λ_2 son reales no negativos y $\lambda_1 + \lambda_2 = 1$, un motivo que genera esta necesidad sería el problema de hallar el $\min\{f_0(x); f_1(x)\}$. Es decir, para tratar este tipo de problemas un camino sería el promedio de funciones, y naturalmente el punto de inicio es la media aritmética, pero surge la dificultad que cuando los dominios de las funciones no se intersectan entonces el resultado de (1) es la función que nos da infinito. Para tratar esta dificultad en el año 2007, Bauschke, Lucet [1] propusieron el promedio o media proximal, $\mathcal{P}_\mu(x, \lambda)$ que supera la dificultad de calcular un promedio de ambas funciones aunque sus dominios sean disjuntos; vamos a ver una prueba sencilla de la semi dualidad del promedio proximal, es decir $(\mathcal{P}_\mu(f, \lambda))^* = \mathcal{P}_{\mu^{-1}}(f^*, \lambda)$ esto nos va a permitir comprobar que algunos esquemas de regularización que se pueden expresar en términos del promedio proximal tienen la propiedad de ser autoduales, por ejemplo la envoltura de Goebel [3], [4] es auto dual es decir $(G_\lambda f)^* = G_\lambda f^*$ lo cual no sucede con la envoltura de Moreau, es decir $(e_\beta f)^* \neq e_\beta f^*$, vamos a considerar $f_1, \dots, f_n \in \Gamma_0(\mathbb{X})$, y $\lambda_1, \dots, \lambda_n$ como números reales no negativos tales que $\sum_{i=1}^n \lambda_i = 1$.

2. Preliminares

2.1. Análisis convexo

Denotaremos por \mathbb{X} un subespacio de Hilbert de \mathbb{R}^n con producto interno $\langle \cdot, \cdot \rangle$ y norma $\| \cdot \|$, con estas notaciones revisaremos aspectos fundamentales del análisis convexo.

¹Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemática, E-mail: marlocpx5@hotmail.com

²Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemática E-mail: tnunez1@gmail.com

Definición 2.1. Argumento mínimo

Dada la función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, se define

$$\operatorname{argmin} f = \begin{cases} \{x \in X / f(x) = \inf_X f\} & \text{si } \inf_X f < +\infty \\ \phi & \text{si } \inf_X f = +\infty \end{cases}$$

Definición 2.2.

El conjunto de las funciones convexas propias y semicontinuas inferiormente sobre \mathbb{X} será denotado por $\Gamma_0(\mathbb{X})$.

Definición 2.3. Conjugada de Fenchel

La conjugada de Fenchel de la función $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ es definida como

$$f^*(x^*) = \sup_{x \in \mathbb{X}} \{\langle x, x^* \rangle - f(x)\}$$

Teorema 2.1

Si $f_1 \leq f_2$, entonces $f_1^* \geq f_2^*$

Demostración

Sea $y \in \mathbb{X}$, como $f_1 \leq f_2$, entonces $\langle x, y \rangle - f_1(x) \geq \langle x, y \rangle - f_2(x)$ para todo $x \in \mathbb{X}$. Tomando el supremo se tiene

$$\sup_{x \in \mathbb{X}} \{\langle x, y \rangle - f_1(x)\} \geq \sup_{x \in \mathbb{X}} \{\langle x, y \rangle - f_2(x)\}$$

Por lo tanto se tiene que $f_1^* \geq f_2^*$. ■

Teorema 2.2. Aproximación por funciones lineales afines

Una función $f : X \rightarrow \overline{\mathbb{R}}$ propia, convexa y cerrada es el supremo de la familia afín de funciones h tal que $h \leq f$

La demostración se encuentra en [2], Teorema 2.4.4.

Teorema 2.3.

Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, una función propia, se cumplen:

- (i) Desigualdad de Young Fenchel $f(x) + f^*(p) \geq \langle p, x \rangle$, para cada $p \in \mathbb{R}^n$
- (ii) $f^{**}(x) \leq f(x)$, para cada x , $f^{**} = f$ si y solo si f es convexa y cerrada.

Demostración

- (i) De la definición de conjugada convexa $f^*(p) = \sup_x \{\langle p, x \rangle - f(x)\}$ se tiene

$$f^*(p) \geq \langle p, x \rangle - f(x) \text{ si y solo si } f(x) + f^*(p) \geq \langle p, x \rangle \quad (2)$$

- (ii) Para esta demostración, consideramos la conjugada de la función f^* así tenemos $f^{**}(x) = \sup_y \{\langle y, x \rangle - f^*(y)\}$ pero de la ecuación (2) tenemos $\langle y, x \rangle - f^*(y) \leq f(x)$ de aquí se tiene

$$\sup_y \{\langle y, x \rangle - f^*(y)\} \leq f(x)$$

Por lo tanto

$$f^{**} \leq f$$

Ahora probaremos $f^{**} \geq f$, para ello, consideramos el conjunto Q formado por todas las funciones afines mayoradas por f , para cada $h \in Q$ tenemos $h(x) = \langle p, x \rangle - \alpha$, para cada x por lo tanto

$$\alpha \geq \sup_x \{\langle p, x \rangle - f(x)\} = f^*(p)$$

$h(x) = \langle p, x \rangle - \alpha \leq f(x)$, para cada x , $\langle p, x \rangle - f(x) \leq \alpha$, para cada x

$$\sup_x \{\langle p, x \rangle - f(x)\} \leq \alpha$$

$f^*(p) \leq \alpha$ si y solo si $-\alpha \leq -f^*(p)$ y consecuentemente

$$\langle p, x \rangle - \alpha \leq \langle p, x \rangle - f^*(p)$$

así que $h(x) \leq \langle p, x \rangle - f^*(p)$, para cada x por lo tanto

$$\sup\{h / h \in Q\} \leq \sup_p \{\langle p, x \rangle - f^*(p)\} = f^{**} \quad (3)$$

Si f es convexa y cerrada, entonces es propia; por el Teorema , es justamente igual a la función del lado izquierdo de (3) , es decir $f \leq f^{**}$ y como $f^{**} \leq f$ se tiene que $f^{**} = f$. Recíprocamente si $f = f^{**}$ entonces f es la conjugada de f^* por lo tanto es convexa y cerrada. ■

Ejemplo 2.1.

Sea \mathbb{X} , un espacio de Hilbert; la conjugada convexa de $q(x) = \frac{1}{2}\|x\|^2$ es

$$q^*(x) = \frac{1}{2}\|x\|^2 = q(x)$$

Además es la única función tal que $f^* = f$.

En efecto $f^*(p) = \sup_x \{\langle p, x \rangle - f(x)\} = \sup_x \{\langle p, x \rangle - \frac{1}{2}\|x\|^2\}$, definimos

$$h(x) = \langle p, x \rangle - \frac{1}{2}\|x\|^2$$

Si $\nabla h(x) = 0$ entonces $x = p$ este valor maximiza la función h , $f^*(p) = \frac{\|p\|^2}{2}$.

Por otro lado, supongamos que f es una función convexa tal que $f^* = f$ entonces es una función propia y usando la desigualdad de Young - Fenchel se tiene $\langle x, x \rangle \leq f(x) + f^*(x) = 2f(x)$ es decir tenemos $q(x) \leq f(x)$. entonces por el teorema se tiene $f^* \leq q^* = q$.

Como $f^* = f$ entonces tenemos $f \leq q$ y por lo tanto se tiene $q \leq f \leq q$, así concluimos $f = q$.

Teorema 2.4.

Sea una función $f : \mathbb{X} \rightarrow]-\infty, +\infty]$, $\alpha, \beta \in \mathbb{R}$, se cumple que

(i) si $\alpha > 0$ entonces $(\alpha f)^*(x^*) = \alpha f^*(\alpha^{-1}x^*)$ para cada $x^* \in \mathbb{X}$.

(ii) si $\beta \neq 0$ entonces $(f(\beta \cdot))^*(x^*) = f^*(\beta^{-1}x^*)$ para cada $x^* \in \mathbb{X}$.

La demostración se encuentra en [7], Teorema 2.3.1

Definición 2.4. inf-convolución o epi-adición

Sean $f; g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$.

La inf-convolución (o epi-adición), $f \sharp g$ está definida como

$$(f \sharp g)(x) = \inf_u \{f(u) + g(x - u)\}$$

0

$$(f_1 \# f_2 \# \cdots \# f_n)(x) = \inf_{x_1 + x_2 + \cdots + x_n = x} \{f_1(x_1) + \cdots + f_n(x_n)\}$$

Definición 2.5. epi- multiplicación

Sea la función, $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$.

Para un escalar μ la epi-multiplicación $\mu \star f$ es dada por

$$\mu \star f = \begin{cases} \mu f(\mu^{-1} \cdot) & \text{si } \mu > 0, \\ i_{\{0\}} & \text{si } \mu = 0. \end{cases}$$

Teorema 2.5.

Sea $\alpha \geq 0$, entonces se cumple que

1. $(\alpha f)^* = \alpha \star f^*$
2. $(\alpha \star f)^* = \alpha f^*$
3. $(f_1 \# f_2 \# \cdots \# f_n)^* = f_1^* + f_2^* + \cdots + f_n^*$
4. Si $f \in \Gamma_0(\mathbb{X})$ se cumple que $f \# q$ y $f^* + q$ son conjugados uno del otro además se cumple $(f \# q) + (f^* \# q) = q$

Demostración.

1. Si $\alpha > 0$, tenemos

$$(\alpha f)^*(x^*) = \sup_x \{\langle x, x^* \rangle - \alpha f(x)\} = \alpha \sup_x \{\langle x, \frac{x^*}{\alpha} \rangle - f(x)\} = \alpha f^*\left(\frac{x^*}{\alpha}\right).$$

Si $\alpha = 0$

$$(\alpha f)^* = \sup_x \{\langle x, x^* \rangle - 0\} = \begin{cases} 0 & \text{si } x^* = 0, \\ +\infty & \text{otro caso.} \end{cases} = i_{\{0\}}(x^*) = (\alpha \star f^*)(x^*)$$

tenemos $(\alpha f)^*(x) = (\alpha \star f^*)(x^*)$

2.

$$(\alpha \star f)^* = \begin{cases} \sup_x \{\langle x, x^* \rangle - \alpha f(x/\alpha)\} & \text{si } \alpha > 0 \\ \sup_x \{\langle x, x^* \rangle - i_{\{0\}}(x)\} & \text{si } \alpha = 0 \end{cases} = \alpha f^*(x^*)$$

3.

$$\begin{aligned} (f_1 \# f_2 \# \cdots \# f_n)^* &= \sup_x \left\{ \langle x, x^* \rangle - \inf_{x_1 + \cdots + x_n = x} \{f_1(x_1) + \cdots + f_n(x_n)\} \right\} \\ &= \sup_{x_1} \{\langle x_1, x^* \rangle - f_1(x_1)\} + \cdots + \sup_{x_n} \{\langle x_n, x^* \rangle - f_n(x_n)\} \\ &= (f_1^* + \cdots + f_n^*)(x^*) \end{aligned}$$

4. Se tiene

$(f \# q)^* = f^* + q$ tomando conjugada nuevamente se tiene $(f \# q)^{**} = (f^* + q)^*$ de aquí $(f \# q) = (f^* + q)^*$.

Así $f \# q$ y $f^* + q$ son conjugados uno del otro, ahora haciendo los cálculos

$$\begin{aligned} (f \# q)(w) &= (f^* + q)^*(w) \\ &= \sup_x \{\langle x, w \rangle - f^*(x) - q(x)\} \\ &= \sup_x \{\langle x, w \rangle - f^*(x) - q(x - w) + q(w) - \langle x, w \rangle\} \\ &= \sup_x \{q(w) - f^*(x) - q(x - w)\} \\ &= q(w) - (f^* \# q)(w) \end{aligned}$$

Es decir $(f \# q) + (f^* \# q) = q$, así queda probado el teorema. ■

Teorema 2.6.

Se cumplen las siguientes proposiciones

1. Si $\text{int dom } f_1 \cap \dots \cap \text{int dom } f_n \neq \emptyset$, entonces $(f_1 + \dots + f_n)^* = f_1^* \# \dots \# f_n^*$
2. Si $\text{int dom } f_1^* \cap \dots \cap \text{int dom } f_n^* \neq \emptyset$, entonces $(f_1 \# \dots \# f_n)^*$ y $\text{epi}(f_1 \# \dots \# f_n) = \text{epi}(f_1) + \dots + \text{epi}(f_n)$

La demostración se encuentra en [7], Teorema 2.8.7

Lema 2.7.

$$(\lambda_1 \star (f_1 + q) \# \dots \# \lambda_n \star (f_n + q))^* = \lambda_1(f_1^* \# q) + \dots + \lambda_n(f_n^* \# q)$$

Demostración.

Aplicando el Teorema 2.5 (iii) y (ii), Teorema 2.6 y el Ejemplo 2.1, tenemos

$$\begin{aligned} (\lambda_1 \star (f_1 + q) \# \dots \# \lambda_n \star (f_n + q))^* &= (\lambda_1 \star (f_1 + q))^* + \dots + (\lambda_n \star (f_n + q))^* \\ &= \lambda_1(f_1 + q)^* + \dots + \lambda_n(f_n + q)^* \\ &= \lambda_1(f_1^* \# q^*) + \dots + \lambda_n(f_n^* \# q^*) \\ &= \lambda_1(f_1^* \# q) + \dots + \lambda_n(f_n^* \# q) \end{aligned}$$

Teorema 2.8.

Sea $g \in \Gamma_0(\mathbb{X})$ se cumple que

$$\text{para todo } x^* \in \mathbb{X}, (g - \mu \star q)^*(x^*) = \mu(q - \mu^{-1}g^*)^* - \mu^{-1} \star q$$

Demostración.

Usando el Teorema 2.5 tenemos

$$\begin{aligned} (g - \mu \star q)^*(x^*) &= \sup_{y^* \in \mathbb{X}} (g^*(y^*) - \mu q(y^* - x^*)) \\ &= \sup_{y^* \in \mathbb{X}} (g^*(y^*) - \mu q(y^*) - \mu q(x^*) + \mu \langle y^*, x^* \rangle) \\ &= -\mu q(x^*) + \sup_{y^* \in \mathbb{X}} (\mu \langle y^*, x^* \rangle - (\mu q(y^*) - q^*(y^*))) \\ &= -\mu q(x^*) + \mu \sup_{y^* \in \mathbb{X}} (\langle y^*, x^* \rangle - (q(y^*) - \mu^{-1}q^*(y^*))) \\ &= -(\mu^{-1} \star q)(x^*) + \mu(q - \mu^{-1}g^*)^*(x^*). \end{aligned}$$

3. Media proximal

Definición 3.1 Media proximal

Sean $f = (f_1, \dots, f_n)$, y $\lambda = (\lambda_1, \dots, \lambda_n)$. La media proximal de peso λ con parámetro μ de n funciones f_i , $1 \leq i \leq n$ se define como

$$\mathcal{P}_\mu(f, \lambda) = \lambda_1 \star (f_1 + \mu \star q) \# \dots \# \lambda_n \star (f_n + \mu \star q) - \mu \star q \quad (4)$$

Sea $x \in \mathbb{X}$ tenemos

$$\mathcal{P}_\mu(f, \lambda)(x) = \lambda_1(f_1 + \mu \star q)\left(\frac{x_1}{\lambda_1}\right) \sharp \cdots \sharp \lambda_n(f_n + \mu \star q)\left(\frac{x_n}{\lambda_n}\right) - \mu \star q\left(\frac{x}{\mu}\right)$$

$$\mathcal{P}_\mu(f, \lambda)(x) = \lambda_1\left(f_1\left(\frac{x_1}{\lambda_1}\right) + \frac{1}{\mu}q\left(\frac{x_1}{\lambda_1}\right)\right)\left(\frac{x_1}{\lambda_1}\right) \sharp \cdots \sharp \lambda_n\left(f_n\left(\frac{x_n}{\lambda_n}\right) + \frac{1}{\mu}q\left(\frac{x_n}{\lambda_n}\right)\right)\left(\frac{x_n}{\lambda_n}\right) - \frac{1}{\mu}q(x)$$

y de aquí se tiene por definición de la epi suma

$$\mathcal{P}_\mu(f, \lambda)(x) = \frac{1}{\mu} \left(\inf_{\sum_{i=1}^n x_i = x} \left\{ \lambda_1(\mu f_1 + q)\left(\frac{x_1}{\lambda_1}\right) + \cdots + \lambda_n(\mu f_n + q)\left(\frac{x_n}{\lambda_n}\right) \right\} - q(x) \right)$$

esto es para todo $x \in \mathbb{X}$, se tiene

$$\mathcal{P}_\mu(f, \lambda)(x) = \frac{1}{\mu} \left(\inf_{\sum_{i=1}^n x_i = x} \sum_{i=1}^n \left\{ \lambda_i \left[\mu f_i\left(\frac{x_i}{\lambda_i}\right) + q\left(\frac{x_i}{\lambda_i}\right) \right] \right\} - q(x) \right) \quad (5)$$

Observación 3.1.

Cuando en la media proximal $\mu = 1$ escribiremos $\mathcal{P}(f, \lambda)(x)$ y no $\mathcal{P}_1(f, \lambda)(x)$.

Teorema 3.1.

La media proximal cumple las siguientes proposiciones

1. $\mathcal{P}_\mu(f, \lambda)(x) = f_i(x)$, cuando $\lambda_i = 1$
2. $\mathcal{P}_\mu(f, \lambda)(x) = \frac{1}{\mu} \left(\inf_{\sum \lambda_i y_i = x} \left\{ \sum_i \lambda_i \mu f_i(y_i) + \sum_i \lambda_i q(y_i) \right\} - q(x) \right)$
3. $\mathcal{P}_\mu(f, \lambda) = (\lambda_1(f_1 + \mu^{-1}q)^* + \cdots + \lambda_n(f_n + \mu^{-1}q)^*)^* - \mu^{-1}q$

Demostración.

1. Como $\lambda_k \geq 0$ para $k = 1, \dots, n$ y $\sum_{k=1}^n \lambda_k = 1$ basta tomar $\lambda_i = 1$ en la definición de media proximal, donde $i = 1, \dots, n$ y tenemos que $\mathcal{P}_\mu(f, \lambda)(x) = f_i(x)$, cuando $\lambda_i = 1$
2. Usamos el cambio de variables $y_i = \frac{x_i}{\lambda_i}$ en la relación (5) de la definición 3.1 y tenemos inmediatamente concluida la prueba.
3. Para esta demostración hay que tener en cuenta que $\text{dom } q = \mathbb{X}$ y para todo $i \in \mathbb{N}$, $\text{dom}(f_i \sharp q) = (\text{dom } f_i) + (\text{dom } q) = \mathbb{X}$ entonces usando el Teorema 2.6 (1), el Teorema 2.5 (1) y el Teorema 2.3, tenemos

$$\begin{aligned} & (\lambda_1(f_1 + \mu^{-1}q)^* + \cdots + \lambda_n(f_n + \mu^{-1}q)^*)^* = \\ & = (\lambda_1(f_1 + \mu^{-1}q)^*)^* \sharp \cdots \sharp (\lambda_n(f_n + \mu^{-1}q)^*)^* \\ & = \lambda_1 \star (f_1 + \mu^{-1}q)^{**} \sharp \cdots \sharp \lambda_n \star (f_n + \mu^{-1}q)^{**} \\ & = \lambda_1 \star (f_1 + \mu \star q) \sharp \cdots \sharp \lambda_n \star (f_n + \mu \star q). \end{aligned}$$

Teorema 3.2.

Se cumple $\text{dom } \mathcal{P}_\mu(f, \lambda) = \lambda_1 \text{dom } f_1 + \cdots + \lambda_n \text{dom } f_n$

Demostración.

Se tiene

$$\begin{aligned}
\text{dom } \mathcal{P}_\mu(f, \lambda) &= \text{dom}(\lambda_1 \star (f_1 + \mu \star q) \# \cdots \# \lambda_n \star (f_n + \mu \star q)) \\
&= \text{dom}(\lambda_1 \star (f_1 + \mu \star q)) + \cdots + \text{dom}(\lambda_n \star (f_n + \mu \star q)) \\
&= \lambda_1 \text{dom}(f_1 + \mu^{-1} q(\frac{\cdot}{\mu})) + \cdots + \lambda_n \text{dom}(f_n + \mu^{-1} q(\frac{\cdot}{\mu})) \\
&= \lambda_1 \text{dom}(f_1) + \cdots + \lambda_n \text{dom}(f_n)
\end{aligned}$$

Teorema 3.3. Conjugada de Fenchel

$$\text{Se cumple } (\mathcal{P}_\mu(f, \lambda))^* = \mathcal{P}_{\mu^{-1}}(f^*, \lambda)$$

Demostración.

Tenemos

$$\begin{aligned}
\mathcal{P}_\mu(f, \lambda) &= \lambda_1 \star (f_1 + \mu \star q) \# \cdots \# \lambda_n \star (f_n + \mu \star q) - \mu \star q \\
\mathcal{P}_\mu(f, \lambda) + \mu \star q &= \lambda_1 \star (f_1 + \mu \star q) \# \cdots \# \lambda_n \star (f_n + \mu \star q) \\
g = (\mathcal{P}_\mu(f, \lambda) + \mu \star q)^* &= \lambda_1 (f_1^* \# \mu q) + \cdots + \lambda_n (f_n^* \# \mu q)
\end{aligned}$$

Ahora usando el Teorema 2.6 y 2.5, tomando conjugada y aplicando el Lema 2.7 tenemos

$$\begin{aligned}
(q - \mu^{-1} g)^* &= (\lambda_1 (\mu^{-1} \star f_1 \# q) + \cdots + \lambda_n (\mu^{-1} \star f_n \# q))^* \\
&= (\lambda_1 (\mu^{-1} \star f_1 \# q))^* \# \cdots \# (\lambda_n (\mu^{-1} \star f_n \# q))^* \\
&= \lambda_1 \star (\mu^{-1} \star f_1 \# q)^* \# \cdots \# \lambda_n \star (\mu^{-1} \star f_n \# q)^* \\
&= \lambda_1 \star ((\mu^{-1} \star f_1)^* + q^*) \# \cdots \# \lambda_n \star ((\mu^{-1} \star f_n)^* + q^*) \\
&= \lambda_1 \star (\mu^{-1} f_1^* + q) \# \cdots \# \lambda_n \star (\mu^{-1} f_n^* + q) \\
\mu(q - \mu^{-1} g)^* &= \mu(\lambda_1 \star (\mu^{-1} f_1^* + q) \# \cdots \# \lambda_n \star (\mu^{-1} f_n^* + q)) \\
&= \lambda_1 \star (f_1^* + \mu q) \# \cdots \# \lambda_n \star (f_n^* + \mu q) \\
&= \lambda_1 \star (f_1^* + \mu^{-1} \star q) \# \cdots \# \lambda_n \star (f_n^* + \mu^{-1} \star q)
\end{aligned}$$

usando la definición de g y el Teorema 2.8 se tiene

$$\begin{aligned}
(\mathcal{P}_\mu(f, \lambda))^* &= (\lambda_1 \star (f_1 + \mu \star q) \# \cdots \# \lambda_n \star (f_n + \mu \star q) - \mu \star q)^* \\
&= \mu(q - \mu^{-1}((\lambda_1 \star (f_1 + \mu \star q) \# \cdots \# \lambda_n \star (f_n + \mu \star q))^*))^* - \mu^{-1} \star q \\
&= \lambda_1 \star (f_1^* + \mu^{-1} \star q) \# \cdots \# \lambda_n \star (f_n^* + \mu^{-1} \star q) - \mu^{-1} \star q \\
(\mathcal{P}_\mu(f, \lambda))^* &= \mathcal{P}_{\mu^{-1}}(f^*, \lambda)
\end{aligned}$$

Teorema 3.4.

Para $\mu = 1$ se cumple

$$(\mathcal{P}(f, \lambda))^* = \mathcal{P}(f^*, \lambda)$$

Ejemplo 3.1.

Sea $F = (f, f^*)$ y $\lambda = (\frac{1}{2}, \frac{1}{2})$, entonces $\mathcal{P}(F, \lambda) = q$.

En efecto, por el Teorema 3.4 tenemos que $(\mathcal{P}(F, \lambda))^* = \mathcal{P}(F^*, \lambda)$.

Como $F^* = (f^*, f^{**}) = (f^*, f)$ entonces tenemos que $\mathcal{P}(F^*, \lambda) = \mathcal{P}(F, \lambda)$.

Por lo tanto, usando el Ejemplo 2.1 tenemos que $(\mathcal{P}(F, \lambda))^* = q$.

Corolario 3.5.

$\mathcal{P}(f, \lambda)$ es una función convexa, semicontinua inferiormente y propia

Demostración.

Por el Teorema 3.4 tenemos que

$$(\mathcal{P}(f, \lambda))^{**} = (\mathcal{P}(f^*, \lambda))^* = \mathcal{P}(f, \lambda)$$

Por lo tanto usando el Teorema 2.3 $\mathcal{P}(f, \lambda) \in \Gamma_0(\mathbb{X})$ ■

Definición 3.2. Envoltura de Moreau

Para funciones propias, sci. de valores extendidos f y $\lambda > 0$, se tiene la envoltura o sobre de Moreau de la función f , que es denotado por $e_\lambda f$ y está definido como

$$e_\lambda f(x) = \inf_u \left\{ f(u) + \frac{1}{2} \|x - u\|^2 \right\}$$

la cual luego fue modificada por Yosida, agregando el parámetro λ convirtiéndose en la regularización Moreau-Yosida o simplemente Moreau, conocida por la expresión siguiente

$$e_{\lambda} f(x) = \inf_u \left\{ f(u) + \frac{1}{2\lambda} \|x - u\|^2 \right\}$$

Por lo tanto $e_{\lambda} f$ es la *inf*-convolución de la función f con la función

$$q(x) = \frac{\|x\|^2}{2}$$

entonces con nuestra notación tenemos $e_{\lambda} f(x) = f \sharp (\lambda \star q(x))$.

El mapeo proximal $P_{\lambda} f(x)$ es un conjunto de mapeos dado por

$$P_{\lambda} f(x) = \operatorname{argmin}_u \left\{ f(u) + \frac{1}{2\lambda} \|x - u\|^2 \right\}$$

Observe que el método de punto proximal clásico dado como

$$P_{\lambda_k} f(x) = \operatorname{argmin}_u \left\{ f(u) + \frac{1}{2\lambda_k} \|x - u\|^2 \right\}$$

sustituye el parámetro λ por una sucesión $\{\lambda^k\}$ acotada y de números positivos. Martinet, en 1970, introduce este método en el estudio de la programación convexa, y posteriormente Rockafellar [5], [6], generaliza la técnica para operadores monótonos maximales.

Este método de punto proximal, debe ser visto primeramente como un método teórico, ya que en su aspecto práctico nos lleva a la resolución de una sucesión de subproblemas relativamente similares al original, que tendrán que ser resueltos con el apoyo de otros métodos; sin embargo, hay situaciones en las que es conveniente reemplazar un problema de minimización por una sucesión de subproblemas, como es el caso en que se tiene un problema original con restricciones y se convierte en una sucesión de subproblemas irrestrictos, situación que se presenta en el método de Lagrangeano aumentado, usando la teoría de dualidad de Fenchel.

Definición 3.3. Función de proximidad acotada

Una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ es Proximidad acotada (Prox-acotada) si existe $\lambda > 0$ tal que $e_{\lambda} f(x) > -\infty$ para $x \in \mathbb{R}^n$.

Para una función f de proximidad acotada, el supremo de todos los λ es el umbral λ_f de proximidad acotada para f .

Teorema 3.6

Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ una función propia, sci y de proximidad acotada por el umbral λ_f entonces, para $0 < \lambda \leq \mu < \lambda_f$, se cumplen las siguientes afirmaciones

1. $e_\mu f \leq e_\lambda f \leq f$. Por lo tanto, como $\lambda \rightarrow 0^+$, las envolturas $e_\lambda f$ forman una familia creciente de funciones acotadas por f y

$$\sup_{\lambda > 0} e_\lambda f(x) = f(x) \quad \text{para todo } x$$

2. $e_\lambda(e_\mu f) = e_{\lambda+\mu} f$
 3. $\inf e_\lambda f = \inf f$ y $\operatorname{argmin} e_\lambda f = \operatorname{argmin} f$

Demostración.

1. Si $0 < \lambda \leq \mu$, entonces $f(u) + \frac{1}{2\mu}\|x - u\|^2 \leq f(u) + \frac{1}{2\lambda}\|x - u\|^2$, para todo u por lo tanto tomando el ínfimo respecto de u se tiene $e_\mu f \leq e_\lambda f$. La otra desigualdad sigue de tomar $u = x$ en la definición de la envoltura de Moreau, tenemos $e_\lambda f \leq f$
2. Con la notación $j_\lambda(u) = \frac{1}{2\lambda}|u|^2$ tenemos que $e_\mu f(x) = f \# j_\lambda(x)$

$$\begin{aligned} e_\lambda(e_\mu f) &= e_\mu f \# j_\lambda(x) \\ &= f \# (j_\mu(x) \# j_\lambda(x)) \\ &= f \# (j_{\lambda+\mu}(x)) \\ &= e_{\lambda+\mu} f \end{aligned}$$

3. Se deduce de (1) que $\inf e_\mu f \leq \inf f$. Por otro lado, tenemos que $f(u) \leq f(u) + \frac{1}{2\lambda}\|x - u\|^2$, para cada x y cada u , por lo que tomando el ínfimo en ambos lados, tenemos $\inf f \leq e_\lambda f$, esto concluye la primera parte $\inf e_\lambda f = \inf f$, ahora para la segunda igualdad podemos proceder así
 $\operatorname{argmin} e_\lambda f = \partial(e_\mu f)^*(0) = \partial(f^* + \mu q)(0) = (\partial f^* + \mu Id)(0) = \partial f^*(0) = \operatorname{argmin} e_\lambda f = \operatorname{argmin} f$

■

Teorema 3.7.

Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ una función propia, sci y proximidad acotada con umbral positivo t_f . Entonces para cada $\lambda \in \langle 0, 1 \rangle$ el conjunto P_λ es no vacío y compacto, donde el valor $e_\lambda f(x)$ es finito y depende continuamente en $\langle \lambda, x \rangle$, con

$$\lim_{\lambda \rightarrow 0^+} e_\lambda f(x) = \sup_{\lambda > 0} e_\lambda f(x) = f(x) \quad \text{para todo } x$$

Además si $x_m \rightarrow x$ y $\lambda_m \rightarrow 0^+$ en $\langle 0, t_f \rangle$ de tal manera que la sucesión $\left\{ \frac{\|x_m - x\|}{\lambda_m} \right\}_{m \in \mathbb{N}}$ es acotada, entonces $e_{\lambda_m} f$ converge continuamente a f , es decir

$$e_{\lambda_m} f(x_m) \rightarrow f(x)$$

Demostración.

La demostración está en [6], Teorema 1.25

3.1. Aproximación Regular

En esta sección se introduce una regularización que se puede aplicar a un amplio espectro de funciones. Su definición de basa en el célebre envoltorio de Moreau. Una propiedad notable de esta regularización es su auto-dualidad con respecto a la conjugación de Fenchel, cuando se está restringido a la clase de funciones propias, convexas y sci. Por otra parte, una extensión de esta

regularización conserva las mismas características cuando se aplica a la clase de funciones silla de montar, propias y cerradas.

Definición 3.4. Envoltura de Goebel

Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ propia, sci y de proximidad acotada con umbral λ_f . Para cada $\lambda \in \langle 0, \lambda_f \rangle$, se define la aproximación regular $G_\lambda f$ como

$$G_\lambda f(x) = (1 - \lambda^2)e_\lambda f(x) + \frac{\lambda}{2}\|x\|^2 \quad (6)$$

Teorema 3.8.

Cuando $\lambda \rightarrow 0^+$, se cumple: $G_\lambda f \rightarrow f$

Demostración.

La demostración es directa a partir del Teorema 3.7.

Observación 3.2.

Es importante señalar, que G_λ no conserva otras características de la envoltura de Moreau. Ninguno de los resultados del Teorema se cumple si se sustituye $e_\lambda f$ por $G_\lambda f$

Ejemplo 3.2.

Sea la función $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x$, se tiene que $e_\lambda f(x) = x - \frac{\lambda}{2}$ y $G_\lambda f(x) = \frac{\lambda x^2}{2} + (1 - \lambda^2)x - \frac{\lambda(1 - \lambda^2)}{2}$.

En efecto, sea $f(x) = x$ se tiene que $e_\lambda f(x) = \inf_u \{f(u) + \frac{1}{2\lambda}|x - u|^2\}$, definimos

$$h(x) = f(u) + \frac{1}{2\lambda}|x - u|^2$$

Tenemos $h'(u) = 1 + \frac{u-x}{\lambda} = 0$ de aquí $u = x - \lambda$ por lo tanto

$$e_\lambda f(x) = x - \lambda + \frac{(-\lambda)^2}{2\lambda} = x - \lambda + \frac{\lambda}{2} = x - \frac{\lambda}{2}$$

Por otro lado tenemos que

$G_\lambda f(x) = (1 - \lambda^2)e_\lambda f(x) + \frac{\lambda}{2}|x|^2 = \frac{\lambda}{2}x^2 + (1 - \lambda^2)x - \frac{\lambda}{2}(1 - \lambda^2)$, en este caso por verificación directa se muestra que (1) ni (2) del Teorema se cumplen.

Teorema 3.9.

Para $0 < \beta \leq 1$, tenemos

$$\begin{aligned} G_\beta f &= (1 + \beta)\mathcal{P}(f, 1 - \beta, 0, \beta) + \beta q \\ G_\beta f(x) &= (1 + \beta)^2 \mathcal{P}\left(f, \frac{1 - \beta}{1 + \beta}, q, \frac{2\beta}{1 + \beta}\right) \left(\frac{x}{1 + \beta}\right) \end{aligned} \quad (7)$$

Demostración.

Por la definición de la envoltura de Moreau

$$\begin{aligned} G_\beta f(x) &= (1 - \beta^2)e_\beta f(x) + \beta q(x) \\ &= (1 - \beta^2) \inf_u \left(f\left(\frac{u}{1 - \beta}\right) + \frac{1}{\beta} q\left(x - \frac{u}{1 - \beta}\right) + \frac{\beta}{1 - \beta^2} q(x) \right) \end{aligned} \quad (8)$$

usando el hecho que $\frac{1}{(1-\beta)^2} + \frac{1}{\beta(1-\beta)} = \frac{1}{\beta(1-\beta)^2}$ y $\frac{1}{\beta(1-\beta)} - \frac{1}{(1-\beta)^2} = \frac{1}{\beta(1-\beta)^2}$ tenemos

$$\begin{aligned} \frac{1}{\beta}q\left(x - \frac{u}{1-\beta}\right) + \frac{\beta}{1-\beta^2}q(x) &= \\ &= \frac{1}{2\beta}\|x - \frac{u}{1-\beta}\|^2 + \frac{\beta}{1-\beta^2}\frac{\|u\|^2}{2} \\ &= q\left(\frac{u}{1-\beta}\right) + \frac{1}{\beta(1-\beta)}q(x-u) - \frac{1}{1-\beta^2}q(x) \end{aligned}$$

Reemplazamos esto en (34), obtenemos

$$G_{\beta}f(x) = (1+\beta)\inf_u \left[(1-\beta)(f+q)\left(\frac{u}{1-\beta}\right) + \frac{1}{\beta}q(x-u) \right] - q(x)$$

Usando la Definición

$$\begin{aligned} G_{\beta}f(x) &= (1+\beta) \left[\left\{ (1+\beta)(f+q)\left(\frac{\cdot}{1-\beta}\right) \right\} \# \frac{1}{\beta}q \right] (x) - q(x) \\ &= (1+\beta) [(1+\beta)(f+q)^* + \beta q]^*(x) - q(x) \end{aligned}$$

A continuación, aplicamos el Teorema 2.6 (1) y el Teorema 2.3, Ejemplo 2.1 y el Teorema 2.5 (1) obtenemos, la primera igualdad en (7)

$$\begin{aligned} G_{\beta}f(x) &= (1+\beta) [(1-\beta)(f+q)^* + \beta q]^*(x) - q(x) + \beta q(x) \\ &= (1+\beta)^2 \left[\left(\frac{1-\beta}{1+\beta}(f+q)^* + \frac{2\beta}{1+\beta}(q+q)^* \right) \left(\frac{x}{1+\beta} \right) - \frac{1}{2} \frac{\|x\|^2}{(1+\beta)^2} \right] \end{aligned}$$

Para cada x , por la definición de media proximal, por el Teorema 2.5 (1) y el Ejemplo 2.1 se tiene la segunda parte de la igualdad de 7.

$$G_{\beta}f(x) = (1+\beta)^2 \mathcal{P} \left(f, \frac{1-\beta}{1+\beta}, q, \frac{2\beta}{1+\beta} \right) \left(\frac{x}{1+\beta} \right)$$

Teorema 3.10.

La aproximación $G_{\beta}f$ es auto dual, es decir

$$(G_{\beta}f)^* = G_{\beta}f^*$$

Demostración.

Tenemos

$$(G_{\beta}f(x))^* = \left((1+\beta)^2 \mathcal{P} \left(f, \frac{1-\beta}{1+\beta}, q, \frac{2\beta}{1+\beta} \right) \left(\frac{x}{1+\beta} \right) \right)^*$$

De los teoremas 2.4 y 2.5 (1) tenemos

$$\begin{aligned} (G_{\beta}f)^* &= \left((1+\beta)^2 \mathcal{P} \left(f, \frac{1-\beta}{1+\beta}, q, \frac{2\beta}{1+\beta} \right) \right)^* ((1+\beta)\cdot) \\ &= (1+\beta)^2 \left(\mathcal{P} \left(f, \frac{1-\beta}{1+\beta}, q, \frac{2\beta}{1+\beta} \right) \right)^* \left(\frac{(1+\beta)\cdot}{(1+\beta)^2} \right) \\ &= (1+\beta)^2 \mathcal{P} \left(f^*, \frac{1-\beta}{1+\beta}, q, \frac{2\beta}{1+\beta} \right) \left(\frac{\cdot}{(1+\beta)} \right) \\ (G_{\beta}f)^* &= G_{\beta}f^* \end{aligned}$$

4. Conclusiones

Se prueba que la media proximal es autodual. La importancia de este hecho radica en que algunos de los esquemas de regulacización se pueden expresar en términos de la media proximal y al ser autoduales las pruebas se simplifican como en el caso de la envoltura de Goebel.

REFERENCIAS BIBLIOGRÁFICAS

- [1] H. Bauschke, Y. Lucet **How to transform one convex continuously into another.** SIAM 2007
- [2] J. M. Borwein and J. Vanderwerff, **Convex Functions: Constructions, Characterizations and Counterexamples.** , CMS Books in Mathematics, 2nd edition, Springer-Verlag, New York, 2008.
- [3] R. Goebel. **Convexity, Convergence and Feedback in Optimal Control** PhD thesis, University of Washington, Seattle, 2000.
- [4] R. Goebel. **Convex Optimal Control Problems with Smooth Hamiltonians** SIAM J. control optim. 43(5):1781-1811 ,2005.
- [5] Rockafellar, R.T., **Convex Analysis**, Princeton University Press,1970.
- [6] Rockafellar, J.B Wets, **Variational Analysis**,Book to be published by Springer-Verlag in the series ,1997.
- [7] C. Zalinescu, **Convex analysis in general vector spaces** , World Scientific Publishing Co. Pte. Ltd.2002.