Bregman Proximal Averages

Xianfu Wang

Department of Mathematics
University of British Columbia
Kelowna, BC, Canada

shawn.wang@ubc.ca

Wednesday, November 17, 2021, 16:40pm–17:10pm
Optimization and Operator Theory
Dedicated to Professor Lev Bregman 80th Birthday
Technion – Israel Institute of Technology

Joint work with H. Bauschke (UBC, Canada)

Abstract

We provide a proximal average with respect to a 1-coercive Legendre function. In the sense of Bregman distance,

- The Bregman envelope of the proximal average is the convex combination of Bregman envelope of individual functions.
- The Bregman proximal mapping of the average is the convex combination of convexified proximal mappings of individual functions.

Outline

Setup

The Euclidean space $\mathbb{R}^n$ has an inner product $\langle \cdot, \cdot \rangle$, and norm $\| \cdot \|$.

Assume that

$f, g : \mathbb{R}^n \to]-\infty, +\infty]$ are proper, lsc functions, and $\lambda > 0$.

Define the Moreau envelope and the proximal mapping $\operatorname{Prox}_\lambda f$:

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

$$\operatorname{Prox}_\lambda f(x) = \operatorname*{argmin}_w \left\{ f(w) + \frac{1}{2\lambda} \|x - w\|^2 \right\}.$$

Define

$$\mathfrak{q} = \frac{1}{2}\|\cdot\|^2,$$

Fenchel conjugate of f:

$$f^*(x) = \sup_w \{\langle x, w\rangle - f(w)\},$$

the infimal convolution of f, g:

$$f \Box g = \inf_w \{f(w) + g(x - w)\},$$

the epi-multiplication of f by λ:

$$\lambda \star f = \lambda f(\cdot/\lambda).$$

What is known?

Assume that $f_1, \ldots, f_m$ are proper, lsc and convex.

Define the proximal average

$$p_\mu(\boldsymbol{f}, \boldsymbol{\lambda}) = \left(\lambda_1(f_1^* \Box \mu\, \mathfrak{q}) + \cdots + \lambda_m(f_m^* \Box \mu\, \mathfrak{q})\right)^* - \mu^{-1}\, \mathfrak{q} \tag{1}$$

$$= \left(\lambda_1(f_1 + \mu^{-1}\, \mathfrak{q})^* + \cdots + \lambda_m(f_m + \mu^{-1}\, \mathfrak{q})^*\right)^* - \mu^{-1}\, \mathfrak{q} \tag{2}$$

Moreau envelope, Fenchel conjugate and proximal mappings

Fact 1.1 (Baushke, Goebel, Lucet and Wang, 2008)

1. $e_\mu p_\mu(\boldsymbol{f}, \boldsymbol{\lambda}) = \lambda_1 e_\mu f_1 + \cdots + \lambda_m e_\mu f_m$.
2. $\left(e_\mu p_\mu(\boldsymbol{f}, \boldsymbol{\lambda})\right)^* = \lambda_1 \star (e_\mu f_1)^* \Box \cdots \Box \lambda_m \star (e_\mu f_m)^*$.
3. $\operatorname{Prox}_\mu \left(p_\mu(\boldsymbol{f}, \boldsymbol{\lambda})\right) = \lambda_1 \operatorname{Prox}_\mu f_1 + \cdots + \lambda_m \operatorname{Prox}_\mu f_m$.

Hence, the set of proximal mappings of convex functions is a convex set!

Question: What about Bregman proximal averages?

Outline

Let

$\phi\colon \mathbb{R}^n \to \,]-\infty,+\infty]$ be convex and differentiable on $U := \operatorname{int}\operatorname{dom}\phi \neq \varnothing$. (3)

The *Bregman distance* associated with ϕ: $D_\phi\colon \mathbb{R}^n \times \mathbb{R}^n \to [0,+\infty]$:

$$(x,y) \mapsto \begin{cases} \phi(x) - \phi(y) - \langle \nabla\phi(y), x-y\rangle, & \text{if } \; y \in U; \\ +\infty, & \text{otherwise.} \end{cases} \tag{4}$$

Our standing assumptions on ϕ are:

A1 $\phi \in \Gamma_0(\mathbb{R}^n)$ is a convex function of Legendre type, i.e., ϕ is essentially smooth and essentially strictly convex in the sense of [8, Section 26].

A2 ϕ is 1-coercive, i.e., $\lim\limits_{\|x\|\to+\infty} \phi(x)/\|x\| = +\infty$. An equivalent requirement is $\operatorname{dom}\phi^* = \mathbb{R}^n$ (see [9, Theorem 11.8(d)]).

The Bregman envelope and proximal mapping

Let $f : \mathbb{R}^n \to \,]-\infty, +\infty]$ be proper and lower semicontinuous. We shall assume that $\operatorname{dom} f \cap \operatorname{dom} \phi \neq \varnothing$.

Definition 1

For $\lambda \in]0, +\infty[$, the *left Bregman envelope function* to f is defined by

$$\overleftarrow{\operatorname{env}}_{\lambda}^{\phi} f : \mathbb{R}^n \to \,]-\infty, +\infty] : y \mapsto \inf_{x} \left(f(x) + \frac{1}{\lambda} D_{\phi}(x, y) \right), \tag{5}$$

and the *left Bregman proximal map* of f is

$$\overleftarrow{\operatorname{prox}}_{\lambda}^{\phi} f \colon U \to U \colon y \mapsto \operatorname*{argmin}_{x} \left(f(x) + \frac{1}{\lambda} D_{\phi}(x, y) \right). \tag{6}$$

The *right Bregman envelope* and *right Bregman proximal mapping* of f are defined analogously and denoted by $\overrightarrow{\operatorname{env}}_{\lambda}^{\phi} f$ and $\overrightarrow{\operatorname{prox}}_{\lambda}^{\phi} f$, respectively.

From the very definitions of $\overleftarrow{\mathrm{env}}_\lambda^\phi f$, $\overleftarrow{\mathrm{prox}}_\lambda^\phi f$, one can see that only the restriction of f to $\operatorname{dom}\phi$ matters.

Even if f is convex, $\overleftarrow{\mathrm{env}}_\lambda^\phi f$ might not be convex.

Example 2.1

Let $\lambda = 1$, $f = \iota_{\{1\}}$ on $\mathbb{R}$.

1. *For $\phi(x) = |x|^3$, we have*

$$(\forall y > 0)\ \overleftarrow{\mathrm{env}}_1^\phi f(y) = 1/3 + 2y^3/3 - y^2, \tag{7}$$

   which is not convex on $(0, +\infty)$.

2. *For $\phi(x) = -\ln x$ if $x > 0$ and $+\infty$ otherwise, we have*

$$(\forall y > 0)\ \overleftarrow{\mathrm{env}}_1^\phi f(y) = \ln y + 1/y - 1 \tag{8}$$

   which is not convex.

ϕ-prox-bounded function

Definition 2

A function $f : \mathbb{R}^n \to]-\infty, +\infty]$ is ϕ-prox-bounded (in short, prox-bounded) if there exists $\lambda > 0$ such that

$$\overleftarrow{\mathrm{env}}^{\phi}_{\lambda} f(x) > -\infty \text{ for some } x \in \mathbb{R}^n.$$

The supremum of all such λ is the threshold λ_f of the prox-boundedness of f.

Proposition 2.2

The following hold:

1. *If f is prox-bounded with threshold* $\lambda_f > 0$*, then for every* $\lambda \in]0, \lambda_f[$ *the function* $f + \frac{1}{\lambda}\phi$ *is bounded below.*
2. *If there exists* $\ell > 0$ *such that for every* $\lambda \in]0, \ell[$ *the function* $f + \frac{1}{\lambda}\phi$ *is bounded below, then* $\lambda_f \geq \ell$*.*
3. *Define*

$$\ell_f = \sup \left\{ \ell > 0 : (\forall \lambda \in]0, \ell[) \ \inf \left(f + \frac{1}{\lambda}\phi \right) > -\infty \right\}.$$

Then $\ell_f = \lambda_f$.

The λ-ϕ-proximal hull

Definition 3

The λ-ϕ-proximal hull (in short,λ-proximal hull) of f is the function $h_\lambda f : \operatorname{dom}\phi \to \,]-\infty, +\infty]$ defined as the pointwise supremum of the collection of all the functions of the form

$$x \mapsto c - \frac{1}{2\lambda} D_\phi(x, w)$$

that are majorized by f, where $c \in \mathbb{R}$, $w \in U$.

Remark 2.3

When $\phi = \mathfrak{q}$, the λ-ϕ-proximal hull is the classical proximal hull.

Proposition 2.4

The following hold:

1. $\overleftarrow{\mathrm{hul}}_\lambda^\phi f = -\overrightarrow{\mathrm{env}}_\lambda^\phi(-\overleftarrow{\mathrm{env}}_\lambda^\phi f)$, *i.e.*,

$$\overleftarrow{\mathrm{hul}}_\lambda^\phi f(x) = \sup_{w\in U}\left(\overleftarrow{\mathrm{env}}_\lambda^\phi f(w) - \frac{1}{\lambda}D_\phi(x,w)\right).$$

   Moreover, $\overleftarrow{\mathrm{env}}_\lambda^\phi(\overleftarrow{\mathrm{hul}}_\lambda^\phi f) = \overleftarrow{\mathrm{env}}_\lambda^\phi f$.

2. $\overleftarrow{\mathrm{hul}}_\lambda^\phi f = \left(f + \frac{1}{\lambda}\phi\right)^{**} - \frac{1}{\lambda}\phi$, *where we use the convention* $\infty - \infty = \infty$.
3. $f \geq \overleftarrow{\mathrm{hul}}_\lambda^\phi f \geq \overleftarrow{\mathrm{env}}_\lambda^\phi f$ *on* U.

Nice properties of $-\overleftarrow{\operatorname{env}}_\lambda^\phi f$

Let $\hat{\partial}, \partial, \partial_C$ denote the Fréchet subdifferential, Mordukhovich limiting subdifferential, and Clarke subdifferential, respectively; see, e.g., [9, 6].

Fact 4 (Kan & Song 2012)

Suppose that $f : \mathbb{R}^n \to\ ]-\infty, +\infty]$ is proper lower semicontinuous and prox-bounded, and that ϕ is second-order continuously differentiable. Then $-\overleftarrow{\operatorname{env}}_\lambda^\phi f$ is Clarke regular, and

$$\begin{aligned}\hat{\partial}(-\overleftarrow{\operatorname{env}}_\lambda^\phi f)(x) = \partial(-\overleftarrow{\operatorname{env}}_\lambda^\phi f)(x) &= \partial_C(-\overleftarrow{\operatorname{env}}_\lambda^\phi f)(x)\\ &= \frac{1}{\lambda}\nabla^2\phi(x)[\operatorname{conv}(\overleftarrow{\operatorname{prox}}_\lambda^\phi f(x)) - x].\end{aligned}$$

The Bregman proximal mapping

Proposition 2.5

Suppose that $f : \mathbb{R}^n \to]-\infty, +\infty]$ is proper lower semicontinuous and ϕ-prox-bounded. Then the following hold:

1. $\overleftarrow{\operatorname{prox}}_\lambda^\phi f \subseteq [\partial(\phi + \lambda f)]^{-1} \circ \nabla\phi$. *If*

$$\partial^\infty f(y) \cap -N_{\operatorname{dom}\phi}(y) = \{0\} \text{ for every } y \in \operatorname{dom}\phi, \tag{9}$$

then $\overleftarrow{\operatorname{prox}}_\lambda^\phi f \subseteq (\nabla\phi + \lambda\partial f)^{-1} \circ \nabla\phi$.

2. *If $\lambda f + \phi$ is convex, then $(\forall x \in \mathbb{R}^n)$ $\overleftarrow{\operatorname{prox}}_\lambda^\phi f(x)$ is convex and closed, and $\overleftarrow{\operatorname{prox}}_\lambda^\phi f = [\partial(\phi + \lambda f)]^{-1} \circ \nabla\phi$. If, in addition, (9) holds and f is Clarke regular, then $\overleftarrow{\operatorname{prox}}_\lambda^\phi f = (\nabla\phi + \lambda\partial f)^{-1} \circ \nabla\phi$.*

3. If f is convex, and $\mathrm{dom}\, f \cap U \neq \varnothing$, then

$$\overleftarrow{\mathrm{prox}}_{\lambda}^{\phi} f = (\nabla\phi + \lambda\partial f)^{-1} \circ \nabla\phi = \left(\frac{1}{\lambda}\nabla\phi + \partial f\right)^{-1} \circ \left(\frac{1}{\lambda}\nabla\phi\right). \quad (10)$$

Moreover, $\overleftarrow{\mathrm{prox}}_{\lambda}^{\phi} f$ is continuous on U.

Remark 2.6

$\overleftarrow{\mathrm{prox}}_{\lambda}^{\phi} f$ is the warped proximity operator of $\lambda\partial f$ by Bui & Combettes [3].

Outline

Let f_1, f_2 be ϕ-prox-bounded with thresholds $\lambda_{f_1}, \lambda_{f_2}$ respectively, and let

$$\overline{\lambda} = \min\{\lambda_{f_1}, \lambda_{f_2}\}.$$

For $0 < \lambda < \overline{\lambda}$, the Bregman proximal average of f_1, f_2 with respect to the Legender function ϕ is defined by

$$\mathcal{P}_{\lambda}^{\phi}(f_1, f_2, \alpha) = \left[\alpha\left(f_1 + \frac{1}{\lambda}\phi\right)^* + (1-\alpha)\left(f_2 + \frac{1}{\lambda}\phi\right)^*\right]^* - \frac{1}{\lambda}\phi, \quad (11)$$

with the convention that $+\infty - (+\infty) = +\infty$, $+\infty - r = +\infty$ for every $r \in \mathbb{R}$.

As we shall see later that

$$\operatorname{dom}\left[\alpha\left(f_1+\frac{1}{\lambda}\phi\right)^* + (1-\alpha)\left(f_2+\frac{1}{\lambda}\phi\right)^*\right]^* \subseteq \operatorname{dom}\phi,$$

so (11) means that $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)(x) =$

$$\begin{cases} \left[\alpha\left(f_1+\frac{1}{\lambda}\phi\right)^* + (1-\alpha)\left(f_2+\frac{1}{\lambda}\phi\right)^*\right]^*(x) - \frac{1}{\lambda}\phi(x) & \text{if } x \in \operatorname{dom}\phi, \\ +\infty & \text{if } x \notin \operatorname{dom}\phi. \end{cases} \tag{12}$$

Hence, it is possible that $f(x) = +\infty$ when $x \in \operatorname{dom}\phi$. Moreover, if $f_1 = f_2$ is proper, lsc and convex with $\operatorname{dom} f_1 \cap \operatorname{dom}\phi \neq \varnothing$, then

$$\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \left(f_1+\frac{1}{\lambda}\phi\right)^{**} - \frac{1}{\lambda}\phi = f_1 + \iota_{\operatorname{dom}\phi}$$

by the Fenchel-Moreau biconjugate theorem [1, Theorem 13.32].

Lemma 3.1

1. *The function* $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)$ *is always lower semicontinuous on* U.
2. *If* $\operatorname{dom}\phi$ *is closed, and* ϕ *is relatively continuous on* $\operatorname{dom}\phi$*, then* $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)$ *is lower semicontinuous on* $\mathbb{R}^n$. *Suppose one of the following holds:*
   (a) $\operatorname{dom}\phi$ *is polyhedral.*
   (b) $\operatorname{dom}\phi$ *is locally simplicial.*

   Then ϕ *is relatively continuous on* $\operatorname{dom}\phi$.

Theorem 3.2

Suppose that ϕ is twice continuously differentiable on U and that for every $u \in U$, $\nabla^2\phi(u)$ is positive definite. Let $\lambda \in]0, \bar{\lambda}[$. Then the following hold:

1. $\operatorname{dom} \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) =$
   $\alpha \operatorname{conv}(\operatorname{dom} f_1 \cap \operatorname{dom} \phi) + (1-\alpha) \operatorname{conv}(\operatorname{dom} f_2 \cap \operatorname{dom} \phi) \subseteq \operatorname{dom} \phi$.
2. $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) + \lambda\phi \in \Gamma_0(\mathbb{R}^n)$.
3. *The function* $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)$ *is* ϕ*-prox-bounded below with* $\lambda_f \geq \bar{\lambda}$.
4. $\overleftarrow{\operatorname{env}}_\lambda^\phi \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \alpha \overleftarrow{\operatorname{env}}_\lambda^\phi f_1 + (1-\alpha) \overleftarrow{\operatorname{env}}_\lambda^\phi f_2$.
5. $(\forall x \in U)\ \overleftarrow{\operatorname{prox}}_\lambda^\phi \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)(x) =$
   $\alpha \operatorname{conv}(\overleftarrow{\operatorname{prox}}_\lambda^\phi f_1(x)) + (1-\alpha) \operatorname{conv}(\overleftarrow{\operatorname{prox}}_\lambda^\phi f_2(x))$.
6. *When* $\alpha = 0$, $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \overleftarrow{\operatorname{hul}}_\lambda^\phi f_2$; *when* $\alpha = 1$,
   $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \overleftarrow{\operatorname{hul}}_\lambda^\phi f_1$.

Definition 5

A sequence of functions $(f_k)_{k\in\mathbb{N}}$ from $\mathbb{R}^n \to]-\infty,+\infty]$ epi-converges to f at a point $x \in \mathbb{R}^n$ if both of the following conditions are satisfied:

1. whenever $(x_k)_{k\in\mathbb{N}}$ converges to x, we have $f(x) \leq \liminf_{k\to\infty} f_k(x_k)$;
2. there exists a sequence $(x_k)_{k\in\mathbb{N}}$ converges to x with $f(x) = \lim_{k\to\infty} f_k(x_k)$.

If $(f_k)_{k\in\mathbb{N}}$ epi-converges to f at every $x \in C \subseteq \mathbb{R}^n$, we say $(f_k)_{k\in\mathbb{N}}$ epi-converges to f on C.

Epi-continuity

Theorem 3.3 (epi-continuity I of Bregman proximal average)

1. *As* $\alpha \downarrow 0$, $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) \xrightarrow{e} \overleftarrow{\text{hul}}_\lambda^\phi f_2$ *on* U.
2. *As* $\alpha \uparrow 1$, $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) \xrightarrow{e} \overleftarrow{\text{hul}}_\lambda^\phi f_1$ *on* U.

Example 3.4

Define the W-shaped functions f and g by the following: for all $x \in \mathbb{R}$,

$$f(x) = \begin{cases} -2x - 12, & \text{if } x < -5, \\ 3x + 13, & \text{if } -5 \leq x < -1, \\ -x + 9, & \text{if } -1 \leq x < 7, \\ 3x - 19, & \text{if } x \geq 7, \end{cases}$$

and

$$g(x) = \begin{cases} -x, & \text{if } x < -4, \\ \frac{1}{2}x + 6, & \text{if } -4 \leq x < 4, \\ -\frac{1}{2}x + 10, & \text{if } 4 \leq x < 6, \\ x + 1, & \text{if } x \geq 6. \end{cases}$$

$\alpha \mapsto \varphi_\mu^\alpha$ define a continuous (epi-topology) transform from $h_\mu f$ to $h_\mu g$.

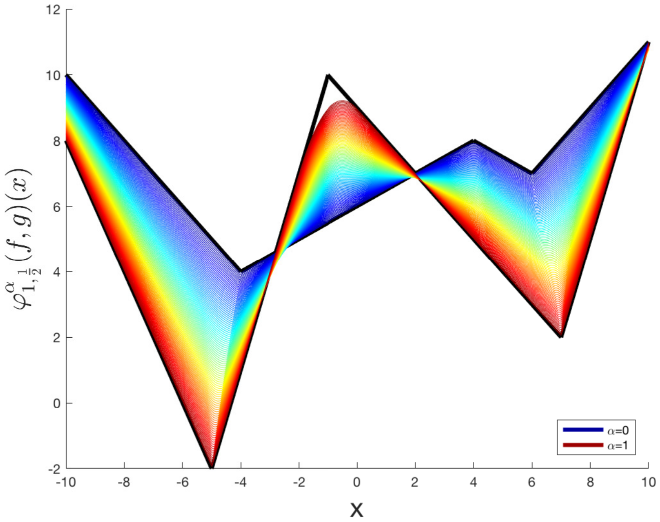

Figure: $\alpha \mapsto \varphi_{\mu}^{\alpha}$ when $\alpha \in [0, 1]$.

Corollary 3.5

Let the conditions in Theorem 3.2 hold and let $f_1, f_2 \in \Gamma_0(\mathbb{R}^n)$. Then $\forall x \in U$, $\partial \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)(x) = \hat{\partial} \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)(x) =$

$$\left[\alpha\left(\partial f_1 + \frac{1}{\lambda}\nabla\phi\right)^{-1} + (1-\alpha)\left(\partial f_2 + \frac{1}{\lambda}\nabla\phi\right)^{-1}\right]^{-1}(x) - \frac{1}{\lambda}\nabla\phi(x); \tag{13}$$

equivalently,

$$\left(\partial \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) + \frac{1}{\lambda}\nabla\phi\right)^{-1} = \alpha\left(\partial f_1 + \frac{1}{\lambda}\nabla\phi\right)^{-1} + (1-\alpha)\left(\partial f_2 + \frac{1}{\lambda}\nabla\phi\right)^{-1}.$$

Note that while ∂f_i is monotone, $\partial \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)$ may be not monotone.

Outline

$\nabla\phi$-firmly nonexpansive mappings

Define the symmetrized Bregman distance $S_\phi : U \times U \to \mathbb{R}$ by

$$S_\phi(x, y) = D_\phi(x, y) + D_\phi(y, x) = \langle \nabla\phi(x) - \nabla\phi(y), x - y \rangle .$$

We need $\nabla\phi$-firmly nonexpansive mappings.

Definition 6

Let $T : U \subset \mathbb{R}^n \to U$. We say that T is $\nabla\phi$-firmly nonexpanive on U if

$$\langle u - v, Tu - Tv \rangle \geq \langle \nabla\phi(Tu) - \nabla\phi(Tv), Tu - Tv \rangle = S_\phi(Tu, Tv)$$

$(\forall u \in U)(\forall v \in U)$.

Lemma 4.1

Let S_ϕ be convex. Suppose that T_1, T_2 are $\nabla\phi$-firmly nonexpansive on U. Then $\alpha T_1 + (1 - \alpha)T_2$ is $\nabla\phi$-firmly nonexpansive on U.

Lemma 4.2

Suppose that $g \in \Gamma_0(\mathbb{R}^n)$, $\operatorname{dom} g \subseteq \operatorname{dom} \phi$, *and* $(\operatorname{ri} \operatorname{dom} g) \cap \operatorname{int} \operatorname{dom} \phi \neq \varnothing$. *Then the following are equivalent:*

1. $g : \mathbb{R}^n \to]-\infty, +\infty]$ *is* ϕ*-strongly convex, i.e.,* $g = f + \phi$ *for a convex function* $f \in \Gamma_0(\mathbb{R}^n)$.
2. g^* *is a* ϕ^**-anisotropic envelope of* f^* *with* $f \in \Gamma_0(\mathbb{R}^n)$*, i.e.,* $g^* = f^* \Box \phi^*$.
3. g^* *is differentiable with* ∇g^* *being* $\nabla\phi$*-firmly nonexpansive on* $\mathbb{R}^n$.
4. $(\phi^* - g^*) \circ \nabla\phi = \lambda \overleftarrow{\operatorname{env}}_{\lambda}^{\phi} f$ *for a convex function* $f \in \Gamma_0(\mathbb{R}^n)$ *and* $\lambda > 0$.
5. g^* *is differentiable with* $\nabla g^* \circ \nabla\phi = \overleftarrow{\operatorname{prox}}_{1}^{\phi} f$ *for some* $f \in \Gamma_0(\mathbb{R}^n)$.

Remark 4.3

The above is an extended version of Baillon-Haddad Theorem.

When is the Bregman proximal average convex?

Theorem 4.4

Let S_ϕ be convex. Suppose that f_1, f_2 are convex, and that either $\operatorname{ri}\operatorname{dom} f_1 \cap U \neq \varnothing$ *or* $\operatorname{ri}\operatorname{dom} f_2 \cap U \neq \varnothing$. *Then* $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)$ *is convex.*

To make the notation simple in the proof, we use $f = \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)$.

Proof of Theorem 3.2

Recall that

$$f = \left[\alpha\left(f_1 + \frac{1}{\lambda}\phi\right)^* + (1-\alpha)\left(f_2 + \frac{1}{\lambda}\phi\right)^*\right]^* - \frac{1}{\lambda}\phi. \tag{14}$$

Since $f_i + \frac{1}{\lambda}\phi$ is ϕ/λ-strongly convex, by Lemma 4.2, each $T_i = \nabla\left(f_i + \frac{1}{\lambda}\phi\right)^*$ is $\nabla\phi/\lambda$-firmly nonexpansive. Lemma 4.1 implies

$$\alpha\nabla\left(f_1 + \frac{1}{\lambda}\phi\right)^* + (1-\alpha)\nabla\left(f_2 + \frac{1}{\lambda}\phi\right)^*$$

is $\nabla\phi/\lambda$-firmly nonexpansive. Because

$$\begin{aligned}&\operatorname{dom}\left[\alpha\left(f_1 + \frac{1}{\lambda}\phi\right)^* + (1-\alpha)\left(f_2 + \frac{1}{\lambda}\phi\right)^*\right]^* \\ &= \alpha(\operatorname{dom} f_1 \cap \operatorname{dom}\phi) + (1-\alpha)(\operatorname{dom} f_2 \cap \operatorname{dom}\phi),\end{aligned}$$

by the assumption, we have

$$\operatorname{ri}[\alpha(\operatorname{dom} f_1 \cap \operatorname{dom}\phi) + (1-\alpha)(\operatorname{dom} f_2 \cap \operatorname{dom}\phi)] \cap U \neq \varnothing.$$

Proof continued

Apply Lemma 4.2 again to obtain that

$$\left[\alpha\left(f_1 + \frac{1}{\lambda}\phi\right)^* + (1-\alpha)\left(f_2 + \frac{1}{\lambda}\phi\right)^*\right]^*$$

is ϕ/λ-strongly convex. Hence f is convex.

Example 4.5

Let $\lambda = 1$, and let $a > 0$, $f_1 = \iota_{\{a\}}$, $f_2 \equiv 0$ on $\mathbb{R}$.

1. *For $\phi(x) = |x|^3$, we have*

$$f(x) = \alpha|a|^3 + \frac{|x - \alpha a|^3}{(1-\alpha)^2} - |x|^3, \tag{15}$$

   and f is not convex.

2. *For $\phi(x) = -\ln x$ if $x > 0$ and $+\infty$ otherwise, we have*

$$f(x) = \begin{cases} -\alpha \ln a - (1-\alpha)\ln \frac{x-\alpha a}{1-\alpha} + \ln x & \text{if } x > \alpha a, \\ +\infty & \text{otherwise,} \end{cases} \tag{16}$$

   and f is not convex.

Outline

Combettes and Reyes' anisotropic envelope and proximity operator

Combettes and Reyes [4] defined anisotropic envelope and anisotropic proximity operator.

Definition 5.1

The anisotropic (or ϕ-anisotropic) envelope of f is defined by

$$f \Box \phi : \mathbb{R}^n \to \,]-\infty, +\infty] : x \mapsto \inf_{y \in \mathbb{R}^n} (f(y) + \phi(x - y)), \tag{17}$$

and the anisotropic (or ϕ-anisotropic) proximity operator of f is

$$\operatorname{aprox}_f^{\phi} : \mathbb{R}^n \to \mathbb{R}^n : x \rightrightarrows \operatorname*{argmin}_{y \in \mathbb{R}^n} (f(y) + \phi(x - y)).$$

Proximal mapping by anisotropic proximal mapping

Theorem 5.2

Suppose that $f \in \Gamma_0(\mathbb{R}^n)$ *and* $\operatorname{ri}\operatorname{dom} f \cap U \neq \varnothing$. *Then*

$$(\forall x \in U)\ \overleftarrow{\operatorname{prox}}_{\lambda}^{\phi} f(x) = \nabla\phi^*\left(\nabla\phi(x) - \lambda\ \operatorname{aprox}_{f^*}^{1/\lambda\star\phi^*}\left(\nabla\phi(x)/\lambda\right)\right).$$

Consequently,

$$(\forall x \in U)\ \nabla\phi\big(\overleftarrow{\operatorname{prox}}_{\lambda}^{\phi} f(x)\big) + \lambda\ \operatorname{aprox}_{f^*}^{1/\lambda\star\phi^*}\left(\nabla\phi(x)/\lambda\right) = \nabla\phi(x). \quad (18)$$

Proposition 5.3

Suppose that $f_1, f_2 \in \Gamma_0(\mathbb{R}^n)$ *with* $\operatorname{dom} f_i \cap \operatorname{dom}\phi \neq \varnothing$ *for* $i = 1, 2$. *Then the following hold:*

1. $f_1 \Box \phi = f_2 \Box \phi$ *if and only if* $f_1 = f_2$.
2. $\overleftarrow{\operatorname{env}}^{\phi}_{\lambda} f_1 = \overleftarrow{\operatorname{env}}^{\phi}_{\lambda} f_2$ *if and only if* $f_1 = f_2$ *on* $\operatorname{dom}\phi$.

Proposition 5.4

Let $f \in \Gamma_0(\mathbb{R}^n)$ *and* $\operatorname{dom} f \cap U \neq \varnothing$*. Then the following hold.*

1. $\operatorname{dom} f \Box \phi = \operatorname{dom} f + \operatorname{dom} \phi$ *and* $f \Box \phi \in \Gamma_0(\mathbb{R}^n)$*.*
2. *For every* $x \in \operatorname{dom} f + U$*,* $\operatorname{aprox}_f^{\phi}(x)$ *is single-valued, and* $\operatorname{dom} \operatorname{aprox}_f^{\phi} \supset \operatorname{dom} f + U$*.*
3. $$(\forall x \in \operatorname{dom} f + U) \ \operatorname{aprox}_f^{\phi}(x) = (\operatorname{Id} + \nabla \phi^* \circ \partial f)^{-1}(x).$$
4. $\operatorname{argmin} f \cap U = \{x \in U : \ \operatorname{aprox}_{f^*}^{\phi^*}(\nabla \phi(x)) = 0\}$.
5. *If* ϕ *is nonnegative, and* $\phi(0) = 0$*, then*
$$f \geq f \Box \phi, \quad \inf f = \inf(f \Box \phi), \ \textit{and} \tag{19}$$
$$\operatorname{argmin} f = \operatorname{argmin} f \Box \phi. \tag{20}$$

Theorem 5.5

Let $f_i \in \Gamma_0(\mathbb{R}^n)$ for $i = 1, 2$. Then the following hold:

1. *Suppose that $(\forall i)$ $\operatorname{ri} \operatorname{dom} f_i \cap U \neq \varnothing$, and that D_ϕ is jointly convex. Then the anisotropic envelope and anisotropic proximal mapping of $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)^*$ satisfy*

$$\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)^* \Box (1/\lambda \star \phi^*) = \alpha f_1^* \Box (1/\lambda \star \phi^*) + (1-\alpha) f_2^* \Box (1/\lambda \star \phi^*), \tag{21}$$

and $\forall x^* \in \mathbb{R}^n$,

$$\nabla \phi^* \left(\lambda (x^* - \operatorname{aprox}_{\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)^*}^{1/\lambda \star \phi^*}(x^*)) \right) = \tag{22}$$

$$\alpha \nabla \phi^* \left(\lambda (x^* - \operatorname{aprox}_{f_1^*}^{1/\lambda \star \phi^*}(x^*)) \right) + (1-\alpha) \nabla \phi^* \left(\lambda (x^* - \operatorname{aprox}_{f_2^*}^{1/\lambda \star \phi^*}(x^*)) \right).$$

2. Suppose that D_{ϕ^*} is jointly convex. Then the anisotropic envelope and anisotropic proximal mapping of $\mathcal{P}^{\phi^*}_{1/\lambda}(f_1^*, f_2^*, \alpha)^*$ satisfy

$$\mathcal{P}^{\phi^*}_{1/\lambda}(f_1^*, f_2^*, \alpha)^* \Box (\lambda \star \phi) = \alpha f_1 \Box (\lambda \star \phi) + (1-\alpha) f_2 \Box (\lambda \star \phi), \quad (23)$$

and $\forall x \in \mathbb{R}^n$,

$$\nabla \phi \left((x - \operatorname{aprox}^{\lambda \star \phi}_{\mathcal{P}^{\phi^*}_{1/\lambda}(f_1^*, f_2^*, \alpha)}(x))/\lambda \right) =$$
$$\alpha \nabla \phi \left((x - \operatorname{aprox}^{\lambda \star \phi}_{f_1}(x))/\lambda \right) + (1-\alpha) \nabla \phi \left((x - \operatorname{aprox}^{\lambda \star \phi}_{f_2}(x))/\lambda \right).$$

Outline

Epi-average

Theorem 6.1

The following holds:

1. $$\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) \geq \left[\alpha \operatorname{conv} f_1\left(\frac{\cdot}{\alpha}\right)\right] \Box \left[(1-\alpha) \operatorname{conv} f_2\left(\frac{\cdot}{1-\alpha}\right)\right].$$

2. $$\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) \leq \alpha f_1 + (1-\alpha) f_2 \text{ on } \operatorname{dom} \phi.$$

Arithmetic average

Suppose that f_1, f_2 are proper lower semicontinuous functions with prox-bound $\lambda_{f_1} > 0, \lambda_{f_2} > 0$ respectively. Define

$$\overline{\lambda} := \min\{\lambda_{f_1}, \lambda_{f_2}\},$$

$$\tilde{f}_i := f_i + \iota_{\mathrm{dom}\,\phi} \quad \text{for } i = 1, 2.$$

Let $\lambda \in]0, \overline{\lambda}[$.

Theorem 6.2 (epi-continuity II of Bregman proximal average)

The following hold:

1. *For every* $x \in \mathbb{R}^n$, *the function* $\lambda \mapsto \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)(x)$ *is monotonically decreasing on* $]0, \overline{\lambda}[$.
2. $\lim_{\lambda \uparrow \overline{\lambda}} \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \left[\alpha \star \operatorname{conv}\left(f_1 + \frac{1}{\overline{\lambda}}\phi\right)\right] \Box \left[(1-\alpha) \star \operatorname{conv}\left(f_2 + \frac{1}{\overline{\lambda}}\phi\right)\right] - \frac{1}{\overline{\lambda}}\phi$ *pointwise.* *In particular, for* $\overline{\lambda} = +\infty$ *one has* $\lim_{\lambda \uparrow \infty} \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \left[\alpha \star \operatorname{conv} \tilde{f}_1\right] \Box \left[(1-\alpha) \star \operatorname{conv} \tilde{f}_2\right]$ *pointwise; consequently,* $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) \xrightarrow{e} \operatorname{cl}\left[(\alpha \star \operatorname{conv} \tilde{f}_1) \Box ((1-\alpha) \star \operatorname{conv} \tilde{f}_2)\right]$ *as* $\lambda \uparrow \infty$.
3. $\lim_{\lambda \downarrow 0} \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \alpha f_1 + (1-\alpha) f_2$ *pointwise on* U*. Consequently, when* $\operatorname{dom} f_i \subseteq U$ *for* $i = 1, 2$, $\mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) \xrightarrow{e} \alpha f_1 + (1-\alpha) f_2$ *as* $\lambda \downarrow 0$.

Outline

Possible future works

1. Is $\alpha\,\mathrm{Prox}_\mu f + (1-\alpha)\,\mathrm{Prox}_\mu g$ always a proximal mapping?
2. Extensions to Bui & Combettes' warped resolvent $J_M^K = (K+M)^{-1} \circ K$?
3. Possible applications of Bregman proximal averages?

References

[1] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.

[2] H.H. Bauschke, R. Goebel, Y. Lucet, and X. Wang, The proximal average: basic theory, *SIAM J. Optim.* 19 (2008), 766–785.

[3] M.N. Bui and P.L. Combettes, Warped proximal iterations for monotone inclusions, *J. Math. Anal. Appl.* 491 (2020), 124315,

[4] P.L. Combettes and N.N. Reyes, Moreau's decomposition in Banach spaces, *Math. Program.* 139 (2013), Ser. B, 103–114.

[5] C. Kan and W. Song, The Moreau envelope function and proximal mapping in the sense of the Bregman distance, *Nonlinear Anal.* 75 (2012), 1385–1399.

[6] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation I*, Springer-Verlag, Berlin, 2006.

[7] S. Reich and S. Sabach, Two strong convergence theorems for a proximal method in reflexive Banach spaces, *Numer. Funct. Anal. Optim.* 31 (2010), 22–44.

[8] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[9] R.T. Rockafellar and R.J.-B. Wets, *Variational Analysis*, Springer-Verlag, 1998.

Thank you!