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)
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

1. Setup
2. Bregman proximal mappings
3. The Bregman proximal average
4. When is the Bregman proximal average convex?
5. Duality via Combettes and Reyes’ anisotropic envelope
6. Relationships to arithmetic average and epi-average
7. Conclusion
The Euclidean space $\mathbb{R}^n$ has an inner product $\langle \cdot, \cdot \rangle$, and norm $\| \cdot \|$. Assume that

$$f, g : \mathbb{R}^n \rightarrow ]-\infty, +\infty]$$

are proper, lsc functions, and $\lambda > 0$.

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

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

$$\text{Prox}_\lambda f(x) = \arg\min_w \left\{ f(w) + \frac{1}{2\lambda} \| x - w \|^2 \right\}.$$
Define
\[ q = \frac{1}{2} \| \cdot \|_2^2, \]

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 \( \lambda \):
\[ \lambda \star f = \lambda f(\cdot/\lambda). \]
What is known?

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

Define the proximal average

\[
\rho_\mu(f, \lambda) = (\lambda_1(f_1^* \sqcap \mu \, q) + \cdots + \lambda_m(f_m^* \sqcap \mu \, q))^* - \mu^{-1} q
\]

\[
= (\lambda_1(f_1 + \mu^{-1} q)^* + \cdots + \lambda_m(f_m + \mu^{-1} q)^*)^* - \mu^{-1} q
\]
Moreau envelope, Fenchel conjugate and proximal mappings

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

1. $e_p \rho_p(f, \lambda) = \lambda_1 e_p f_1 + \cdots + \lambda_m e_p f_m.$
2. $(e_p \rho_p(f, \lambda))^* = \lambda_1 * (e_p f_1)^* \square \cdots \square \lambda_m * (e_p f_m)^*.$
3. $\text{Prox}_p \rho_p(f, \lambda) = \lambda_1 \text{Prox}_p f_1 + \cdots + \lambda_m \text{Prox}_p f_m.$

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

**Question:** What about Bregman proximal averages?
Outline

1. Setup
2. Bregman proximal mappings
3. The Bregman proximal average
4. When is the Bregman proximal average convex?
5. Duality via Combettes and Reyes’ anisotropic envelope
6. Relationships to arithmetic average and epi-average
7. Conclusion
Let
\[ \phi : \mathbb{R}^n \to [-\infty, +\infty) \] be convex and differentiable on \( U := \text{int dom } \phi \neq \emptyset. \)  

(3)

The Bregman distance associated with \( \phi : D_\phi : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty] : \)

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

(4)

Our standing assumptions on \( \phi \) are:

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

A2 \( \phi \) is 1-coercive, i.e., \( \lim_{\|x\| \to +\infty} \phi(x)/\|x\| = +\infty \). An equivalent requirement is \( \text{dom } \phi^* = \mathbb{R}^n \) (see [9, Theorem 11.8(d)]).
Let \( f : \mathbb{R}^n \to ]-\infty, +\infty] \) be proper and lower semicontinuous. We shall assume that \( \text{dom } f \cap \text{dom } \phi \neq \emptyset \).

**Definition 1**

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

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

and the *left Bregman proximal map* of \( f \) is

\[
\overleftarrow{\text{prox}}^\phi \lambda f : U \to U : y \mapsto \arg\min_x \left( f(x) + \frac{1}{\lambda} D_\phi(x, y) \right). \tag{6}
\]

The *right Bregman envelope* and *right Bregman proximal mapping* of \( f \) are defined analogously and denoted by \( \overrightarrow{\text{env}}^\phi \lambda f \) and \( \overrightarrow{\text{prox}}^\phi \lambda f \), respectively.
From the very definitions of $\hat{\text{env}}_\lambda^\phi f$, $\hat{\text{prox}}_\lambda^\phi f$, one can see that only the restriction of $f$ to $\text{dom} \phi$ matters.

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

**Example 2.1**

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

1. For $\phi(x) = |x|^3$, we have
   \[
   (\forall y > 0) \quad \hat{\text{env}}_1^\phi f(y) = 1/3 + 2y^3/3 - y^2, \quad (7)
   \]
   which is not convex on $(0, +\infty)$.

2. For $\phi(x) = -\ln x$ if $x > 0$ and $+\infty$ otherwise, we have
   \[
   (\forall y > 0) \quad \hat{\text{env}}_1^\phi f(y) = \ln y + 1/y - 1 \quad (8)
   \]
   which is not convex.
\(\phi\)-prox-bounded function

**Definition 2**

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

\[
\text{env}_\lambda^\phi f(x) > -\infty \quad \text{for some} \quad x \in \mathbb{R}^n.
\]

The supremum of all such \( \lambda \) is the threshold \( \lambda_f \) 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 $\ell_f = \lambda_f$. 

Definition 3

The $\lambda$-$\phi$-proximal hull (in short, $\lambda$-proximal hull) of $f$ is the function $h_\lambda f : \text{dom} \phi \rightarrow ]-\infty, +\infty]$ 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 $\phi = q$, the $\lambda$-$\phi$-proximal hull is the classical proximal hull.
Proposition 2.4

The following hold:

1. \( \underline{\text{hul}}_\lambda \phi f = - \underline{\text{env}}_\lambda \phi (- \underline{\text{env}}_\lambda \phi f) \), i.e.,
   \[
   \underline{\text{hul}}_\lambda \phi f(x) = \sup_{w \in U} \left( \underline{\text{env}}_\lambda \phi f(w) - \frac{1}{\lambda} D_\phi(x, w) \right).
   \]

   Moreover, \( \underline{\text{env}}_\lambda \phi (\underline{\text{hul}}_\lambda \phi f) = \underline{\text{env}}_\lambda \phi f \).

2. \( \underline{\text{hul}}_\lambda \phi f = (f + \frac{1}{\lambda} \phi)^{**} - \frac{1}{\lambda} \phi \), where we use the convention \( \infty - \infty = \infty \).

3. \( f \geq \underline{\text{hul}}_\lambda \phi f \geq \underline{\text{env}}_\lambda \phi f \) on \( U \).
Nice properties of $-\text{\text{env}}_{\lambda}^{\phi}f$

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 $\phi$ is second-order continuously differentiable. Then $-\text{\text{env}}_{\lambda}^{\phi}f$ is Clarke regular, and

\[
\hat{\partial}(-\text{\text{env}}_{\lambda}^{\phi}f)(x) = \partial(-\text{\text{env}}_{\lambda}^{\phi}f)(x) = \partial_{C}(-\text{\text{env}}_{\lambda}^{\phi}f)(x) = \frac{1}{\lambda} \nabla^{2} \phi(x)[\text{conv}(\text{prox}_{\lambda}^{\phi}f(x)) - x].
\]
Proposition 2.5

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

1. \( \overleftarrow{\text{prox}}_\lambda^\phi f \subseteq [\partial(\phi + \lambda f)]^{-1} \circ \nabla \phi. \) If
   \[
   \partial^\infty f(y) \cap -N_{\text{dom} \phi}(y) = \{0\} \text{ for every } y \in \text{dom} \phi, \tag{9}
   \]
   then \( \overleftarrow{\text{prox}}_\lambda^\phi f \subseteq (\nabla \phi + \lambda \partial f)^{-1} \circ \nabla \phi. \)

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

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

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

**Remark 2.6**

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

1. Setup
2. Bregman proximal mappings
3. The Bregman proximal average
4. When is the Bregman proximal average convex?
5. Duality via Combettes and Reyes’ anisotropic envelope
6. Relationships to arithmetic average and epi-average
7. Conclusion
Let $f_1, f_2$ be $\phi$-prox-bounded with thresholds $\lambda_{f_1}, \lambda_{f_2}$ respectively, and let 

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

For $0 < \lambda < \bar{\lambda}$, the Bregman proximal average of $f_1, f_2$ with respect to the Legendre function $\phi$ is defined by

$$P_{\phi}^\lambda(f_1, f_2, \alpha) = \left[\alpha \left( f_1 + \frac{1}{\lambda} \phi \right)^* + (1 - \alpha) \left( f_2 + \frac{1}{\lambda} \phi \right)^* \right]^* - \frac{1}{\lambda} \phi,$$  \hspace{1cm} (11)$$

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

Xianfu Wang (UBC, Canada)
As we shall see later that

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

so (11) means that

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

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

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

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

by the Fenchel-Moreau biconjugate theorem [1, Theorem 13.32].
Lemma 3.1

1. The function $P^\phi_\lambda(f_1, f_2, \alpha)$ is always lower semicontinuous on $U$.

2. If $\text{dom } \phi$ is closed, and $\phi$ is relatively continuous on $\text{dom } \phi$, then $P^\phi_\lambda(f_1, f_2, \alpha)$ is lower semicontinuous on $\mathbb{R}^n$. Suppose one of the following holds:
   - (a) $\text{dom } \phi$ is polyhedral.
   - (b) $\text{dom } \phi$ is locally simplicial.

Then $\phi$ is relatively continuous on $\text{dom } \phi$. 
Suppose that \( \phi \) 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. \( \text{dom} \, \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \alpha \text{conv}(\text{dom} \, f_1 \cap \text{dom} \, \phi) + (1 - \alpha) \text{conv}(\text{dom} \, f_2 \cap \text{dom} \, \phi) \subseteq \text{dom} \, \phi. \)

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 \( \phi \)-prox-bounded below with \( \lambda_f \geq \bar{\lambda}. \)

4. \( \overset{\sim}{\text{env}}_\lambda^\phi \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \alpha \overset{\sim}{\text{env}}_\lambda^\phi f_1 + (1 - \alpha) \overset{\sim}{\text{env}}_\lambda^\phi f_2. \)

5. \( (\forall x \in U) \) \( \overset{\sim}{\text{prox}}_\lambda^\phi \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha)(x) = \alpha \text{conv}(\overset{\sim}{\text{prox}}_\lambda^\phi f_1(x)) + (1 - \alpha) \text{conv}(\overset{\sim}{\text{prox}}_\lambda^\phi f_2(x)). \)

6. When \( \alpha = 0, \) \( \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \overset{\sim}{\text{hul}}_\lambda^\phi f_2; \) \text{ when } \( \alpha = 1, \) \( \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) = \overset{\sim}{\text{hul}}_\lambda^\phi f_1. \)
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 \((x_k)_{k \in \mathbb{N}}\) converging 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 \(C\).
Theorem 3.3 (epi-continuity I of Bregman proximal average)

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

2. As \( \alpha \uparrow 1 \), \( \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) \xrightarrow{\text{e}} \overrightarrow{\text{hul}}_\lambda^\phi f_1 \) 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

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

\( \alpha \mapsto \varphi^\alpha_\mu \) define a continuous (epi-topology) transform from \( h_\mu f \) to \( h_\mu g \).
Figure: $\alpha \mapsto \varphi^\alpha_{\mu}$ when $\alpha \in [0, 1]$. 

$\varphi^\alpha_{\frac{1}{2}, \frac{1}{2}}(f, g)(x)$
Corollary 3.5

Let the conditions in Theorem 3.2 hold and let $f_1, f_2 \in \Gamma_0(\mathbb{R}^n)$. Then

$$\forall x \in U, \partial \mathcal{P}_\phi^\lambda(f_1, f_2, \alpha)(x) = \hat{\partial} \mathcal{P}_\phi^\lambda(f_1, f_2, \alpha)(x) =$$

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

equivalently,

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

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

1. Setup
2. Bregman proximal mappings
3. The Bregman proximal average
4. When is the Bregman proximal average convex?
5. Duality via Combettes and Reyes’ anisotropic envelope
6. Relationships to arithmetic average and epi-average
7. Conclusion
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 nonexpansive 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 $S_\phi$ 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 \( g \in \Gamma_0(\mathbb{R}^n) \), \( \text{dom } g \subseteq \text{dom } \phi \), and \((\text{ri dom } g) \cap \text{int dom } \phi \neq \emptyset\). Then the following are equivalent:

1. \( g : \mathbb{R}^n \to ]-\infty, +\infty] \) is \( \phi \)-strongly convex, i.e., \( g = f + \phi \) for a convex function \( f \in \Gamma_0(\mathbb{R}^n) \).

2. \( g^* \) is a \( \phi^* \)-anisotropic envelope of \( f^* \) with \( f \in \Gamma_0(\mathbb{R}^n) \), i.e., \( g^* = f^* \Box \phi^* \).

3. \( g^* \) is differentiable with \( \nabla g^* \) being \( \nabla \phi \)-firmly nonexpansive on \( \mathbb{R}^n \).

4. \( (\phi^* - g^*) \circ \nabla \phi = \lambda \tilde{\text{env}}_{\lambda}^\phi f \) for a convex function \( f \in \Gamma_0(\mathbb{R}^n) \) and \( \lambda > 0 \).

5. \( g^* \) is differentiable with \( \nabla g^* \circ \nabla \phi = \tilde{\text{prox}}_{\phi}^f \) 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 $S_\phi$ be convex. Suppose that $f_1, f_2$ are convex, and that either $\text{ri dom } f_1 \cap U \neq \emptyset$ or $\text{ri dom } f_2 \cap U \neq \emptyset$. Then $P_\lambda^\phi(f_1, f_2, \alpha)$ is convex.

To make the notation simple in the proof, we use $f = P_\lambda^\phi(f_1, f_2, \alpha)$. 
Proof of Theorem 3.2

Recall that

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

Since \( f_i + \frac{1}{\lambda} \phi \) is \( \phi/\lambda \)-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

\[
\text{dom} \left[ \alpha \left( f_1 + \frac{1}{\lambda} \phi \right)^* + (1 - \alpha) \left( f_2 + \frac{1}{\lambda} \phi \right)^* \right]^* \\
= \alpha (\text{dom} f_1 \cap \text{dom} \phi) + (1 - \alpha) (\text{dom} f_2 \cap \text{dom} \phi),
\]

by the assumption, we have

\[ \text{ri} \left[ \alpha (\text{dom} f_1 \cap \text{dom} \phi) + (1 - \alpha) (\text{dom} f_2 \cap \text{dom} \phi) \right] \cap U \neq \emptyset. \]
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 $\phi/\lambda$-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

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

and $f$ is not convex.

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

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

and $f$ is not convex.
Outline

1. Setup
2. Bregman proximal mappings
3. The Bregman proximal average
4. When is the Bregman proximal average convex?
5. Duality via Combettes and Reyes’ anisotropic envelope
6. Relationships to arithmetic average and epi-average
7. Conclusion
Combettes and Reyes [4] defined anisotropic envelope and anisotropic proximity operator.

**Definition 5.1**

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

\[
f \square \phi : \mathbb{R}^n \rightarrow ]-\infty, +\infty] : x \mapsto \inf_{y \in \mathbb{R}^n} (f(y) + \phi(x - y)),
\]

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

\[
aprox_f^\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n : x \mapsto \text{argmin}_{y \in \mathbb{R}^n} (f(y) + \phi(x - y)).
\]
Theorem 5.2

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

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

Consequently,

$$\forall x \in U \quad \nabla \phi \left( \text{prox}^\lambda f(x) \right) + \lambda \text{aprox}_{f^*}^{1/\lambda \phi^*} \left( \nabla \phi(x)/\lambda \right) = \nabla \phi(x). \quad (18)$$
Proposition 5.3

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

1. $f_1 \square \phi = f_2 \square \phi$ if and only if $f_1 = f_2$.

2. $\text{env}_\phi f_1 = \text{env}_\phi f_2$ if and only if $f_1 = f_2$ on $\text{dom } \phi$. 
Proposition 5.4

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

1. $\text{dom } f \Box \phi = \text{dom } f + \text{dom } \phi$ and $f \Box \phi \in \Gamma_0(\mathbb{R}^n)$.

2. For every $x \in \text{dom } f + U$, $\text{aprox}_f^\phi(x)$ is single-valued, and $\text{dom } \text{aprox}_f^\phi \supset \text{dom } f + U$.

3. \[(\forall x \in \text{dom } f + U) \text{ aprox}_f^\phi(x) = (\text{Id} + \nabla \phi^* \circ \partial f)^{-1}(x).\]

4. $\text{argmin } f \cap U = \{ x \in U : \text{ aprox}_{f^*}^\phi(\nabla \phi(x)) = 0 \}$.

5. If $\phi$ is nonnegative, and $\phi(0) = 0$, then

\[
f \geq f \Box \phi, \quad \inf f = \inf(f \Box \phi), \text{ and } \tag{19}\]

\[
\text{argmin } f = \text{argmin } f \Box \phi. \tag{20}\]
Theorem 5.5

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

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

$$P^\phi_\lambda(f_1, f_2, \alpha)^* \Box (1/\lambda \ast \phi^*) = \alpha f_1^* \Box (1/\lambda \ast \phi^*) + (1 - \alpha) f_2^* \Box (1/\lambda \ast \phi^*),$$

(21)

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

$$\nabla \phi^* \left( \lambda(x^* - \text{aprox}_{P^\phi_\lambda(f_1, f_2, \alpha)^*}^{1/\lambda \ast \phi^*}(x^*)) \right) =$$

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

(22)
Suppose that $D_{\phi^*}$ is jointly convex. Then the anisotropic envelope and anisotropic proximal mapping of $\mathcal{P}_{1/\lambda}^{\phi^*}(f_1^*, f_2^*, \alpha)^*$ satisfy

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

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

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

$$\alpha \nabla \phi \left( (x - \text{aprox}_{f_1^*}(x))/\lambda \right) + (1 - \alpha) \nabla \phi \left( (x - \text{aprox}_{f_2^*}(x))/\lambda \right).$$
1. Setup
2. Bregman proximal mappings
3. The Bregman proximal average
4. When is the Bregman proximal average convex?
5. Duality via Combettes and Reyes’ anisotropic envelope
6. Relationships to arithmetic average and epi-average
7. Conclusion
\[ \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) \geq \left[ \alpha \text{conv } f_1 \left( \frac{\cdot}{\alpha} \right) \right] \square \left[ (1 - \alpha) \text{conv } f_2 \left( \frac{\cdot}{1 - \alpha} \right) \right]. \]

\[ \mathcal{P}_\lambda^\phi(f_1, f_2, \alpha) \leq \alpha f_1 + (1 - \alpha) f_2 \text{ on } \text{dom } \phi. \]
Suppose that $f_1, f_2$ are proper lower semicontinuous functions with prox-bound $\lambda_{f_1} > 0, \lambda_{f_2} > 0$ respectively. Define

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

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

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 \( \lambda \mapsto P_\lambda^\phi(f_1, f_2, \alpha)(x) \) is monotonically decreasing on \( ]0, \lambda[ \).

2. \( \lim_{\lambda \uparrow \lambda} P_\lambda^\phi(f_1, f_2, \alpha) = \left[ \alpha \ast \text{conv} \left( f_1 + \frac{1}{\lambda} \phi \right) \right] \square \left[ (1 - \alpha) \ast \text{conv} \left( f_2 + \frac{1}{\lambda} \phi \right) \right] - \frac{1}{\lambda} \phi \) pointwise. In particular, for \( \lambda = +\infty \) one has
   \( \lim_{\lambda \uparrow \infty} P_\lambda^\phi(f_1, f_2, \alpha) = \left[ \alpha \ast \text{conv} \tilde{f}_1 \right] \square \left[ (1 - \alpha) \ast \text{conv} \tilde{f}_2 \right] \) pointwise;
   consequently, \( P_\lambda^\phi(f_1, f_2, \alpha) \xrightarrow{\varepsilon} \text{cl} \left[ (\alpha \ast \text{conv} \tilde{f}_1) \square ((1 - \alpha) \ast \text{conv} \tilde{f}_2) \right] \) as \( \lambda \uparrow \infty \).

3. \( \lim_{\lambda \downarrow 0} P_\lambda^\phi(f_1, f_2, \alpha) = \alpha f_1 + (1 - \alpha)f_2 \) pointwise on \( U \). Consequently, when \( \text{dom } f_i \subseteq U \) for \( i = 1, 2 \), \( P_\lambda^\phi(f_1, f_2, \alpha) \xrightarrow{\varepsilon} \alpha f_1 + (1 - \alpha)f_2 \) as \( \lambda \downarrow 0 \).
Possible future works

1. Is $\alpha \text{Prox}_\mu f + (1 - \alpha) \text{Prox}_\mu g$ always a proximal mapping?

2. Extensions to Bui & Combettes’ warped resolvent $J^K_M = (K + M)^{-1} \circ K$?

3. Possible applications of Bregman proximal averages?
References


Thank you!