A FORMULA ON THE SUBDIFFERENTIAL OF THE DECONVOLUTION OF CONVEX FUNCTIONS

M. VOLLE

It is known that, under suitable assumptions, the subdifferential $\partial(f \square g)$ of the infimal convolution of two convex functions $f$ and $g$ coincides with the parallel sum of $\partial f$ and $\partial g$. We prove that a similar formula holds for the subdifferential of the deconvolution of two convex functions: under appropriate hypothesis it coincides with the parallel star-difference of the sub-differentials of the functions.

1. PRELIMINARIES

In what follows $(X, Y)$ is a couple of locally convex real topological spaces paired in separating duality by a bilinear form $\langle \cdot, \cdot \rangle$, and $\Gamma_0(X)$ (respectively $\Gamma_0(Y)$) is the class of convex, lower-semicontinuous proper functions defined on $X$ (respectively $Y$) with values in $\mathbb{R} \cup \{-\infty, +\infty\}$. Given $h, k : X \to \mathbb{R}$, the inf-convolution $h \square k$ is defined by

$$(h \square k)(x) = \inf_{u \in X} \{h(x - u) + k(u)\} \quad \text{for all } x \in X,$$

where $\hat{+}$ is the upper extension of the addition to $\overline{\mathbb{R}}$ (that is, $(+\infty)\hat{+}(-\infty) = (+\infty) - (+\infty) = +\infty$, see [10]).

The deconvolution, denoted by the symbol $\boxdot$, is a kind of inverse operation for the inf-convolution. It was introduced by Hiriart-Urruty and Mazure [5] in order to solve the inf-convolution equation

$$(1) \quad \text{find } \xi \in \overline{\mathbb{R}}^X \text{ such that } k \boxdot \xi = h.$$

It is known [9, Corollary 2] that a solution to (1) exists if and only if the function

$$(2) \quad z \mapsto (h \sqcup k)(z) = \sup_{u \in X} \{h(z + u) - k(u)\}$$

is one of them. The function defined by (2) is referred to as the deconvolution of $h$ and $k$. Here the symbol $\sqcup$ denotes the lower extension of the subtraction to $\overline{\mathbb{R}}$ (that is, $(+\infty) \sqcup (-\infty) = (+\infty) - (+\infty) = -\infty$, [10]).

Received 27th April, 1992
The operation $\Box$ has many interesting applications. For instance, taking the deconvolution of convex quadratic forms yields a variational formulation of the parallel subtraction of matrices and operators (for example [8, 13]).

The deconvolution operation is strongly linked to the star-difference of sets. Recall that the star-difference of two subsets $A$ and $B$ of a linear space $E$ is defined by

$$A^* - B = \{x \in E : x + B \subseteq A\}.$$ 

By setting

$$E(f) = \{(x, r) \in X \times \mathbb{R} : f(x) \leq r\}$$

for the epigraph of $f \in \mathbb{R}^E$ and $I_C$ for the indicator function of $C \subseteq X$ ($I_C(x) = 0$ if $x \in C$, $I_C(x) = +\infty$ if $x \in E \setminus C$), we have then [15, Proposition 6]

$$E(h \Box k) = E(h)^* - E(k)$$

for the epigraph of $f \in \mathbb{R}^E$ and $I_C$ for the indicator function of $C \subseteq X$.

In the context of epigraphical analysis [1], formula (3) suggests another terminology for the deconvolution operation, namely the epigraphical difference or, better, the epigraphical star-difference.

In connection with Fenchel's duality theory, the deconvolution operation enjoys some noteworthy properties. Recall that the Fenchel conjugate of $f \in \mathbb{R}^X$ is defined by

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

for all $y \in Y$. In a similar way one defines the conjugate of a function in $\mathbb{R}^Y$. A fundamental result concerning the conjugacy operation is that

$$f = f^{**}$$

for all $f \in \Gamma_0(X)$. Now, according to Hiriart-Urruty [4, Theorem 2.2], the formula

$$h^* - k^* = h \Box \hat{k}$$

holds for all $h, k \in \Gamma_0(X)$. It follows that

$$E(h \Box k) = (h^* - k^*)^*$$

for all $h, k \in \Gamma_0(X)$. If $h^* - k^*$ turns out to be in $\Gamma_0(Y)$, one can also write

$$(h \Box k)^* = h^* - k^*.$$
We end this section by recalling some facts and by introducing some notation. For any extended-real-valued function $\xi \in \mathbb{R}^Y$, the $\varepsilon$-subdifferential ($\varepsilon \geq 0$) of $\xi$ at $y \in \xi^{-1}(\mathbb{R})$ is the set

$$\partial_\varepsilon \xi(y) = \{x \in X : \forall v \in Y : \xi(v) - \xi(y) \geq (x,v - y) - \varepsilon\}.$$ 

For $\varepsilon = 0$ we set, as usual, $\partial \xi(y) = \partial_0 \xi(y)$. In connection with this concept, the following classical property will be used later on (see for example [6, p.351]):

**Lemma 1.** For any $\xi \in \mathbb{R}^Y$ and $y \in Y$ such that $\xi(y) = \xi^{**}(y) \in \mathbb{R}$ one has

$$\partial \xi(y) = \partial \xi^{**}(y).$$

**Proof:** As we always have $\xi^{**} \leq \xi$, the inclusion $\subseteq$ is obvious. Now, for any $x \in \partial \xi(y)$, the affine continuous form $(x, \cdot) - (x,y) + \xi^{**}(y)$ is smaller than $\xi$. As $\xi^{**}$ coincides with the upper hull of all affine continuous minorants of $\xi$, we have $(x,\cdot) - (x,y) + \xi^{**}(y) \leq \xi^{**}$, that is to say, $x \in \partial \xi^{**}(y)$. 

The directional derivative of a function $\xi \in \mathbb{R}^Y$ at a point $y \in \xi^{-1}(\mathbb{R})$ is defined, when it exists, by

$$\xi'(y,d) = \lim_{t \to 0^+} t^{-1}(\xi(y + td) - \xi(y))$$

for all $d \in Y$; the lower subdifferential of $\xi$ at $y$ is the set (for example [14])

$$\partial^-(\xi(y) = \{x \in X : (x, \cdot) \leq \xi'(y, \cdot)\}.$$ 

The above set obviously contains $\partial \xi(y)$ and coincides with $\partial \xi(y)$ when $\xi$ is convex; in that case ($\xi$ convex) it is well known (for example [6, p.354]) that $\xi'(y,\cdot)$ is a sublinear function whose Fenchel conjugate is the indicator function of $\partial \xi(y)$; in other words

$$\left(\xi'(y,\cdot)\right)^* = I_{\partial \xi(y)}.$$

If, moreover, $\xi$ is continuous at $y$, then $\xi'(y,\cdot)$ is finitely valued, continuous, and one has

$$\xi'(y,\cdot) = (I_{\partial \xi(y)})^*.$$  

2. **On the subdifferential of the difference of two convex functions**

Let $\varphi$ and $\psi$ in $\mathbb{R}^Y$ be two convex functions, finite at the point $y \in Y$. In connection with the subdifferentiability of the difference $\varphi - \psi$, there are two formulas worth mentioning:

$$\partial(\varphi - \psi)(y) = \bigcap_{\varepsilon > 0} \partial_\varepsilon \varphi(y) - \partial_\varepsilon \psi(y)$$

$$\partial^-(\varphi - \psi)(y) = \partial \varphi(y)^* - \partial \psi(y).$$
The first one is due to Martinez-Legaz and Seeger [7, Theorem 1] and applies to arbitrary functions \( \varphi \) and \( \psi \) in \( \Gamma_0(Y) \); the second one has been mentioned by Ellaia [3, p.94] for \( \varphi \) and \( \psi \) convex on \( \mathbb{R}^n \) and finitely valued. The next lemma extends the second formula to our general setting.

**Lemma 2.** Let \( \varphi, \psi \in \mathbb{R}^Y \) be convex functions, finite at \( y \in Y \), and assume that \( \psi \) is continuous at \( y \). Then

\[
\partial^-(\varphi - \psi)(y) = \partial^- \left( \varphi - \psi \right)(y) = \partial \varphi(y) - \partial \psi(y).
\]

**Proof:** Let us consider only the case \( - \). Each of the following lines is equivalent to \( x \in \partial \varphi(y) - \partial \psi(y) :\)

\[
\forall u \in \partial \psi(y) : x + u \in \partial \varphi(y)
\]

\[
\forall u \in \partial \psi(y) : (x + u,.) \leq \varphi^'(y,.) \hspace{1cm} (\text{as } \partial \varphi(y) = \partial^\varphi(y))
\]

\[
(x,.) + (I_{\partial \psi(y)})^* \leq \varphi^'(y,.) \hspace{1cm} (\text{by taking the supremum for } u \in \partial \psi(y))
\]

\[
(x,.) \leq \varphi^'(y,.) - \psi^'(y,.) \hspace{1cm} (\text{from (9)})
\]

\[
(x,.) \leq (\varphi - \psi)^'(y,.) \hspace{1cm} (\text{as } \psi^'(y,.) \text{ is finitely valued})
\]

\[
x \in \partial^- (\varphi - \psi)(y).
\]

Before passing to next section we record here a by-product of this lemma.

**Corollary.** Let \( \varphi, \psi \) be in \( \Gamma_0(Y) \); assume that \( \varphi - \psi \) is convex, finite at \( y \), and \( \psi \) is continuous at \( y \); then

\[
\bigcap_{\varepsilon > 0} \partial^\varepsilon \varphi(y) - \partial^\varepsilon \psi(y) = \partial \varphi(y) - \partial \psi(y).
\]

**Proof:** As \( \partial^- (\varphi - \psi) = \partial (\varphi - \psi) \), it suffices to apply formula (10) and Lemma 2.

3. **On the subdifferential of the deconvolution**

Let us present now the main results of this note. Given \( h, k \in \Gamma_0(X) \), the parallel sum of \( \partial h \) and \( \partial k \) is defined by (see for example [12])

\[
\partial h \square \partial k = \left( (\partial h)^{-1} + (\partial k)^{-1} \right)^{-1}.
\]
Here \((\partial h)^{-1}\) is set for the inverse of the multivalued operator \(\partial h\); in other words:
\[
y \in (\partial h \square \partial k)(z) \Leftrightarrow z \in (\partial h)^{-1}(y) + (\partial k)^{-1}(y)
\]
where \(+\) denotes the Minkowski vectorial addition. It turns out that, under appropriate constraint qualifications [11, 12], the formula \(\partial(h \square k) = \partial h \square \partial k\) holds. It is tempting to ask whether or not there is a similar formula for the deconvolution operator. To this end, let us introduce the notion of parallel star difference for subdifferentials.

**Definition:** Let \(h\) and \(k\) be in \(\Gamma_0(X)\). The parallel star difference of the subdifferentials \(\partial h\) and \(\partial k\) is the multivalued operator \(\partial h \square \partial k\) defined by
\[
\partial h \square \partial k = \left((\partial h)^{-1} - (\partial k)^{-1}\right)^{-1},
\]
that is, for any \((x, y) \in X \times Y,
\[
y \in (\partial h \square \partial k)(x) \Leftrightarrow x \in (\partial h)^{-1}(y) - (\partial k)^{-1}(y).
\]

In [5, Proposition 7] one finds a lower estimate for the subdifferential of two finitely valued convex functions \(f, g\) on \(\mathbb{R}^n\); namely it is shown that
\[
\partial(g \square f)(x) \supseteq \bigcup_{(x_1, x_2) \in A_x} \partial g(x_1) \cap \partial f(x_2),
\]
where \(A_x = \{(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n : z = x_1 - x_2, \ (g \square f)(x) = g(x_1) - f(x_2)\}\). As pointed out to me by A. Seeger, the condition \((x_1, x_2) \in A_x\) yields the inclusion \(\partial g(x_1) \subseteq \partial f(x_2)\). Moreover, the convexity assumption on \(f\) and \(g\) is superfluous:

**Proposition.** Let \(f, g\) be arbitrary extended real valued functions on \(X\). Then, for any \(x \in X\), we have
\[
\partial(g \square f)(x) \supseteq \bigcup_{v \in E(x)} \partial g(v),
\]
where \(E(x) = \{v \in X : g(v) - f(v - x) = (g \square f)(x) \in \mathbb{R}\}\).

**Proof:** Let \(v\) be in \(E(x)\) and \(y \in \partial g(v)\); then \(g(v)\) and \(f(v - x)\) are real numbers and we have,
\[
(g \square f)(z) \geq g(z + v - x) - f(v - x) \quad \text{for all } z \in X.
\]
Hence,
\[
(g \square f)(z) - (g \square f)(z) \geq g(z + v - x) - f(v - x) - g(v) + f(v - x) \quad \text{for all } z \in X,
\]
and finally
\[
(g \square f)(z) - (g \square f)(z) \geq g(z + v - x) - g(v) \geq (z - x, y) \quad \text{for all } z \in X.
\]
This shows that \( y \in \partial(g \sqcup f)(x) \).

The next result provides an upper estimate for the subdifferential of the deconvolution of two convex functions \( h, k \in \Gamma_0(X) \) in terms of the parallel star difference of \( \partial h \) and \( \partial k \); it involves the set

\[
C(h,k) = \{ y \in Y : k^{\ast} \text{ is finite and continuous at } y, \text{ and } (h^{\ast} - k^{\ast})(y) = (h^{\ast} - k^{\ast})^{\ast\ast}(y) \}
\]

**Theorem 1.** Let \( X, Y \) be locally convex spaces in separating duality, and let \( h, k \in \Gamma_0(X) \). Then, for all \( x \in X \), we have

\[
\partial(h \sqcup k)(x) \cap C(h,k) \subset (\partial h \sqcup \partial k)(x).
\]

**Proof:** Assume that \( y \in \partial(h \sqcup k)(x) \cap C(h,k) \). We have to show that \( x \in (\partial h)^{\ast\ast\ast}(y) - (\partial k)^{\ast\ast\ast}(y) \). As \( y \in \partial(h \sqcup k)(x) \) we have \( x \in \partial h \sqcup k^{\ast}(y) \). Now, from (6), \( (h \sqcup k)^{\ast} = (h^{\ast} - k^{\ast})^{\ast\ast\ast} \); then \( x \in \partial(h^{\ast} - k^{\ast})^{\ast\ast\ast}(y) \); as \( y \in C(h,k) \) it follows from Lemma 1 that \( x \in \partial(h^{\ast} - k^{\ast})(y) \) and, a fortiori, \( x \in \partial(h^{\ast} - k^{\ast})(y) \). So, by Lemma 2, we obtain \( x \in \partial h^{\ast}(y) - \partial k^{\ast}(y) \).

Let us give an example showing that Theorem 1 cannot be improved without additional assumptions. Take for \( X \) an Hilbert space with closed unit ball \( B \), \( h = \| \| \|, k = \left( \| \|^{\ast}\right)/2. \) We have then by (6) \( (h \sqcup k)^{\ast} = \left( I_B - (\| \|^{\ast})/2 \right)^{\ast\ast\ast} = I_B - 1/2 \) so that \( h \sqcup k = \| \| + 1/2 \). Note also that \( C(h,k) = \{ y \in X : \| y \| \geq 1 \} \). For the subdifferentials we have, on one hand,

\[
\partial(h \sqcup k)(x) = \begin{cases} x/\| x \| & \text{if } x \neq 0 \\ B & \text{if } x = 0 \end{cases}
\]

and, on the other hand, \( y \in (\partial h \sqcup \partial k)(x) \) if and only if

\[
x \in \partial I_B(y) - \partial \left( \| y \|^2/2 \right)(y) = \partial I_B(y) - y = \begin{cases} \emptyset & \text{if } \| y \| > 1 \\ [-1, +\infty[y] & \text{if } \| y \| = 1 \\ -y & \text{if } \| y \| < 1 \end{cases}
\]

In particular,

for \( \| x \| = 1 \) \( \partial(h \sqcup k)(x) = \{ x \} \subset (\partial h \sqcup \partial k)(x) = \{ x, -x \} \)

for \( x = 0 \) \( \partial(h \sqcup k)(0) = B \supset (\partial h \sqcup \partial k)(0) = \{ y : \| y \| = 1 \} \cup \{ 0 \} \).

With stronger assumptions it is possible to give an exact formula for \( \partial(h \sqcup k) \):
**Theorem 2.** Let $X,Y$ be locally convex spaces paired in separating duality, and let $h,k \in \Gamma_0(X)$. Assume that $k^*$ is finite and continuous over $Y$ and that $h^* - k^*$ is convex. Then,

$$\partial (h \Box k) = \partial h \Box \partial k.$$ 

**Proof:** Since for each $f \in \Gamma_0(X)$ one has $(\partial f)^{-1} = \partial f^*$, we easily obtain the equivalence between the assertions below:

1. $y \in \partial(h \Box k)(x)$
2. $x \in \partial(h \Box k^*)(y)$
3. $x \in \partial(h^* - k^*)(y)$ (by (5) as $h^* - k^*$ is convex proper lower semicontinuous)
4. $x \in \partial h^*(y) - \partial k^*(y)$ (from Lemma 2)
5. $x \in (\partial h)^{-1}(y) - (\partial k)^{-1}(y)$.

**Example:** Let us take for $X$ a Hilbert space, $h \in \Gamma_0(X)$, $k \in \Gamma_0(X)$. Assume that $k^*$ is finite over $X$ (hence continuous) and suppose that $h^* - k^*$ is strongly convex: there exists $t > 0$ and $f \in \Gamma_0(X)$ such that $h^* - k^* = f^* + \left( t \| \cdot \|^2 \right)/2$. We have then by (5)

$$h \Box k = \left( f^* + \frac{t \| \cdot \|^2}{2} \right)^* = f \Box \frac{\| \cdot \|^2}{2t}.$$ 

So, $h \Box k$ coincides with the Moreau-Yosida regularisation of $f \in \Gamma_0(X)$ (for example [2, p.195]). It follows that $h \Box k$ is continuously differentiable. As $h^*$ is also strongly convex, $h$ is continuously differentiable and we have, applying Theorem 2,

$$\nabla (h \Box k) = \nabla h \Box \partial k.$$ 

**References**


University of Avignon
33 rue L. Pasteur
84000 Avignon
France