DECOMPOSITION METHOD FOR NONSMOOTH CONVEX VARIATIONAL SIGNAL RECOVERY

Heinz H. Bauschke,[1] *Patrick L. Combettes,*[2] *and Jean-Christophe Pesquet*[3]

[1]Mathematics, University of British Columbia Okanagan, Kelowna, BC V1V 1V7, Canada
[2]Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie – Paris 6, 75005 Paris, France
[3]Institut Gaspard Monge and UMR–CNRS 8049, Université de Marne la Vallée, 77454 Marne la Vallée Cedex 2, France

ABSTRACT

Under consideration is the large body of signal recovery problems that can be formulated as the problem of minimizing the sum of two proper lower semicontinous convex (not necessarily smooth) functions in a real Hilbert space. A general analysis of this problem is given and a decomposition method is proposed to solve it. The convergence of the method, which is based on an extension of the standard Douglas-Rachford algorithm for monotone operators, is established under genral conditions. Various signal recovery applications are discussed and numerical results are provided.

1. INTRODUCTION

A wide array of methodological approaches have been proposed to solve signal recovery problems based on various physical, statistical, and numerical considerations, as well as on certain heuristic beliefs. Mathematically, though, the vast majority of signal recovery problems are posed as optimization problems and typically solved on a case-by-case basis by an *ad hoc* numerical algorithm. In [3], it was shown that a number of apparently unrelated problems fitted the following variational format in a real Hilbert space $\mathcal{H}$.

Problem 1 Let $f_1 \colon \mathcal{H} \to \;]-\infty, +\infty]$ and $f_2 \colon \mathcal{H} \to \mathbb{R}$ be two proper lower semicontinuous convex functions such that f_2 is differentiable on $\mathcal{H}$ with a $1/\beta$-Lipschitz continuous gradient, for some $\beta \in \;]0, +\infty[$. The objective is to minimize $f_1 + f_2$ over $\mathcal{H}$.

Problem 1 was shown to cover a variety of signal recovery formulations, including constrained least-squares problems, multiresolution sparse regularization problems, Fourier regularization problems, geometry/texture image decomposition problems, hard-constrained inconsistent feasibility problems, split feasibility problems, as well as certain maximum *a posteriori* problems [3]. Investigating this generic formulation therefore made it possible to derive existence, uniqueness, characterization, and stability results in a unified and standardized fashion for a large class of problems. Moreover, a relaxed forward-backward algorithm was proposed to solve Problem 1 and shown to capture, extend, and provide a simplified analysis for a variety of existing iterative methods, such as the projected Landweber method, alternating projections methods, or the iterative thresholding method recently proposed in [4] (see [3] for details).

Despite its relatively broad scope, Problem 1 fails to cover the important situations in which f_2 is differentiable with a non-Lipschitz gradient, or not finite everywhere, or, more importantly, when it is not differentiable at all. The latter situation includes for instance the problem of minimizing the total variation of a signal over a convex set, the problem of minimizing the sum of two set-distances, problems in which both functions are maxima of convex functions, or Tykhonov-like problems with L^1 norms. The objectives of the present paper are to extend Problem 1 by relaxing the assumptions on f_2 to a mere standard qualification condition, to analyze the properties of the resulting nonsmooth convex optimization problem, and to propose an iterative decomposition method to solve it based on recent developments [1] on the Douglas-Rachford method for monotone operator splitting [7]. In Section 3, we define our notation and provide the necessary mathematical background. The recovery problem is formulated, analyzed, and illustrated in Section 2. In Section 4, an algorithm is proposed to solve it. Finally, numerical applications to wavelet-based signal recovery are presented in Section 5.

2. NOTATION AND THEORETICAL TOOLS

Throughout this paper, $\mathcal{H}$ is a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$, norm $\| \cdot \|$, and distance d.

2.1. Convex analysis [9]

The *indicator function* of a subset C of $\mathcal{H}$ is

$$\iota_C \colon x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{if } x \notin C, \end{cases} \tag{1}$$

and the *distance* from a point $x \in \mathcal{H}$ to C is $d_C(x) = \inf \|x - C\|$; if C is also closed and convex then, for every $x \in \mathcal{H}$, there exists a unique point $P_C x \in C$ such that $\|x - P_C x\| = d_C(x)$; $P_C x$ is the *projection* of x onto C.

The *domain* of a function $f : \mathcal{H} \to \ ]-\infty, +\infty]$ is $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$. $\Gamma_0(\mathcal{H})$ is the class of all lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty, +\infty]$ which are proper in the sense that their domain is nonempty. Now let $f \in \Gamma_0(\mathcal{H})$. The *conjugate* of f is the function $f^* \in \Gamma_0(\mathcal{H})$ defined by

$$(\forall u \in \mathcal{H}) \quad f^*(u) = \sup_{x \in \mathcal{H}} \langle x \mid u \rangle - f(x) \qquad (2)$$

and $f^{**} = f$. The *subdifferential* of f at $x \in \mathcal{H}$ is the set

$$\partial f(x) = \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \ \langle y - x \mid u \rangle + f(x) \le f(y)\}. \qquad (3)$$

If f is (Gâteaux) differentiable at x with gradient $\nabla f(x)$, then $\partial f(x) = \{\nabla f(x)\}$.

Lemma 2 *Let f_1 and f_2 be two functions in $\Gamma_0(\mathcal{H})$ such that*

$$\bigcup_{\gamma > 0} \{\gamma(x_1 - x_2) \mid x_1 \in \operatorname{dom} f_1, x_2 \in \operatorname{dom} f_2\}$$

$$\textit{is a closed vector subspace}, \quad (4)$$

and let $x \in \mathcal{H}$. Then x minimizes $f_1 + f_2$ if and only if $0 \in \partial f_1(x) + \partial f_2(x)$.

Remark 3 Condition (4) is satisfied in each of the following cases.

  (i) f_1 or f_2 is finite.

  (ii) $\operatorname{dom} f_1 \cap \operatorname{int} \operatorname{dom} f_2 \ne \varnothing$ or $\operatorname{dom} f_2 \cap \operatorname{int} \operatorname{dom} f_1 \ne \varnothing$.

  (iii) $\dim \mathcal{H} < +\infty$ and the relative interior of $\operatorname{dom} f_1$ intersects with that of $\operatorname{dom} f_2$.

2.2. Proximity operators [3]

Let $f \in \Gamma_0(\mathcal{H})$. Then, for every $x \in \mathcal{H}$, the function $y \mapsto f(y) + \|x - y\|^2/2$ admits a unique minimizer denoted by $\operatorname{prox}_f x$. The operator

$$\operatorname{prox}_f : \mathcal{H} \to \mathcal{H} : x \mapsto \operatorname*{arg\,min}_{y \in \mathcal{H}} f(y) + \frac{1}{2}\|x - y\|^2 \quad (5)$$

is the *proximity operator* of f.

Lemma 4 *Let $f \in \Gamma_0(\mathcal{H})$. Then*

  (i) $(\forall \gamma \in \ ]0, +\infty[)(\forall x \in \mathcal{H})$

$$x = \operatorname{prox}_{\gamma f} x + \gamma \operatorname{prox}_{f^*/\gamma}(x/\gamma).$$

  (ii) $(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \ \|\operatorname{prox}_f x - \operatorname{prox}_f y\|^2$
$$\le \|x - y\|^2 - \|\operatorname{prox}_{f^*} x - \operatorname{prox}_{f^*} y\|^2.$$

  (iii) *If f is even, then prox_f is odd.*

2.3. Examples of proximity operators [3]

Example 5 Set $f = \iota_C$, where C is a nonempty closed convex subset of $\mathcal{H}$. Then $\operatorname{prox}_f = P_C$. Hence, proximity operators generalize the notion of a projection operator.

Example 6 Let C be a nonempty closed convex subset of $\mathcal{H}$, let $\gamma \in \ ]0, +\infty[$, and let $x \in \mathcal{H}$. Then

$$\operatorname{prox}_{\gamma d_C} x = \begin{cases} x + \dfrac{\gamma}{d_C(x)}(P_C x - x), & \text{if } d_C(x) > \gamma; \\ P_C x, & \text{if } d_C(x) \le \gamma. \end{cases}$$

Example 7 If $\mathcal{H} = \mathbb{R}$ and $C = \{0\}$ in Example 6, we obtain the soft-thresholding operation

$$\operatorname{prox}_{\gamma|\cdot|} x = \operatorname{sgn}(x) \max\{|x| - \gamma, 0\}. \qquad (6)$$

Example 8 Let $(e_k)_{k \in \mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$, let $(\phi_k)_{k \in \mathbb{N}}$ be functions in $\Gamma_0(\mathbb{R})$ such that

$$(\forall k \in \mathbb{N}) \ \phi_k \ge 0 \quad \text{and} \quad \phi_k(0) = 0, \qquad (7)$$

let $f : \mathcal{H} \to \ ]-\infty, +\infty] : x \mapsto \sum_{k \in \mathbb{N}} \phi_k(\langle x \mid e_k \rangle)$, and let $x \in \mathcal{H}$. Then $\psi \in \Gamma_0(\mathcal{H})$ and $\operatorname{prox}_f x = \sum_{k \in \mathbb{N}} \pi_k e_k$, where $\pi_k = \operatorname{prox}_{\phi_k} \langle x \mid e_k \rangle$ is the unique solution to the inclusion $\langle x \mid e_k \rangle - \pi_k \in \partial \phi_k(\pi_k)$.

Example 9 Let K be a (finite of infinite) subset of $\mathbb{N}$, let $(e_k)_{k \in K}$ be an orthonormal basis of $\mathcal{H}$, let $(p_k)_{k \in K}$ be a sequence in $[1, +\infty[$, let $(\omega_k)_{k \in K}$ be a sequence in $]0, +\infty[$, let $f : \mathcal{H} \to \ ]-\infty, +\infty] : x \mapsto \sum_{k \in K} \omega_k |\langle x \mid e_k \rangle|^{p_k}$, and let $x \in \mathcal{H}$. Then $\operatorname{prox}_f x = \sum_{k \in K} \pi_k e_k$ where, for every $k \in K$, π_k is the unique solution to the equation

$$p_k \omega_k |\pi_k|^{p_k - 1} \operatorname{sgn}(\pi_k) + \pi_k - \xi_k = 0,$$

where $\xi_k = \langle x \mid e_k \rangle$; in particular, π_k is given by

$$\begin{cases} \operatorname{sgn}(\xi_k) \max\{|\xi_k| - \omega_k, 0\}, & \text{if } p_k = 1; \\[2mm] \xi_k/(1 + 2\omega_k), & \text{if } p_k = 2; \\[2mm] \operatorname{sgn}(\xi_k)\left(|\xi_k| + \dfrac{9}{8}\omega_k^2\left(1 - \sqrt{1 + \dfrac{16|\xi_k|}{9\omega_k^2}}\right)\right), & \text{if } p_k = \dfrac{3}{2}; \\[2mm] \operatorname{sgn}(\xi_k)\dfrac{\sqrt{1 + 12\omega_k|\xi_k|} - 1}{6\omega_k}, & \text{if } p_k = 3. \end{cases}$$

Example 10 Let $\mathcal{H} = \mathbb{R}^m$, let $\gamma \in \ ]0, +\infty[$, and let f be the Burg entropy, i.e., $f : x = (\xi_k)_{1 \le k \le m} \mapsto \sum_{k=1}^m \phi(\xi_k)$, where

$$\phi : \xi \mapsto \begin{cases} -\ln(\xi), & \text{if } \xi > 0; \\ +\infty, & \text{if } \xi \le 0. \end{cases} \qquad (8)$$

Then

$$(\forall x \in \mathbb{R}^m) \ \operatorname{prox}_{\gamma f} x = \frac{1}{2}\left(\xi_k + \sqrt{\xi_k^2 + 4\gamma^2}\right)_{1 \le k \le m}.$$

3. PROBLEM FORMULATION

Problem 11 Let f_1 and f_2 be two functions in $\Gamma_0(\mathcal{H})$ which satisfy (4). The objective is to

$$\underset{x\in\mathcal{H}}{\text{minimize}}\ \ f_1(x) + f_2(x). \tag{9}$$

Proposition 12

(i) *Existence: Problem 11 possesses at least one solution if $\lim_{\|x\|\to+\infty} f_1(x) + f_2(x) = +\infty$.*

(ii) *Uniqueness: Problem 11 possesses at most one solution if $f_1 + f_2$ is strictly convex, as is the case when f_1 or f_2 is strictly convex.*

(iii) *Characterization: Let $x \in \mathcal{H}$ and $\gamma \in\ ]0, +\infty[$, and set $\text{rprox}_f = 2\text{prox}_f - \text{Id}$. Then the following statements are equivalent.*

    (a) *x solves Problem 11.*

    (b) *$x = \text{prox}_{\gamma f_2} y$, where $y = \text{rprox}_{\gamma f_1} \text{rprox}_{\gamma f_2} y$.*

Proof. (i)&(ii): [3, Proposition 3.1]. (iii): Combine [1, Lemma 2.6(iii)] and Lemma 2. □

Let us note that the only restriction imposed in Problem 11 is the relatively mild qualification condition (4). It therefore follows from Remark 3(i) that Problem 11 subsumes Problem 1. Therefore, all the examples of signal recovery problems discused in [3] are covered by Problem 11. We now list a few concrete scenarios which are covered by Problem 11 but not by Problem 1.

Example 13 $\mathcal{H}$ is either or $\mathbb{R}^N$ or $H^1(\Omega)$, where Ω is an open bounded domain of $\mathbb{R}^m$, f_1 is the L^1 norm or the indicator function of a nonempty closed convex set, and f_2 is the total variation. This follows from [2, Proposition 1].

Example 14 C_1 and C_2 are two nonempty closed convex sets, $f_1 = d_{C_1}$, and $f_2 = d_{C_2}$. Problem 11 is then an extension of the standard convex feasiblity problem, which corresponds to the case when $C_1 \cap C_2 \neq \varnothing$ [8].

Example 15 $\mathcal{H} = \text{L}^2(\Omega)$, where Ω is an open bounded domain of $\mathbb{R}^m$, and the observed data assume the form $z = L\overline{x} + w$, where L is a bounded linear operator from $\mathcal{H}$ to a Hilbert space $\mathcal{G}$ and $w \in \mathcal{G}$ is an additive noise component. Moreover, $f_1\colon x \mapsto \|Lx - z\|_{\text{L}^1}$ and $f_2 = \|\cdot\|_{\text{L}^1}$.

4. ALGORITHM

Theorem 16 *Let $\gamma \in\ ]0, +\infty[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0, 2[$, and let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be sequences in $\mathcal{H}$. Suppose that Problem 11 admits at least one solution,*

$\sum_{n\in\mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$, *and* $\sum_{n\in\mathbb{N}} \lambda_n(\|a_n\| + \|b_n\|) < +\infty$. *Take $x_0 \in \mathcal{H}$ and set, for every $n \in \mathbb{N}$,*

$$x_{n+\frac{1}{2}} = \text{prox}_{\gamma f_2} x_n + b_n \tag{10}$$

and

$$x_{n+1} = x_n + \lambda_n\left(\text{prox}_{\gamma f_1}\left(2x_{n+\frac{1}{2}} - x_n\right) + a_n - x_{n+\frac{1}{2}}\right). \tag{11}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to some point $y \in \mathcal{H}$ and $x = \text{prox}_{\gamma f_2} x$ is a solution to Problem 11.

Proof. Let A and B be the subdifferentials of f_1 and f_2, respectively, in [1, Corollary 5.2] and use Lemma 2. □

In the above algorithm, the minimization problem (9) is decomposed in two main steps: at iteration n, the current iterate x_n and the function f_2 is first utilized in (10) to compute $x_{n+\frac{1}{2}}$. The function f_1 is then utilized to produce the update x_{n+1} in (11). Note that the algorithms allows for the inexact implementation of these two proximal steps via the incorporation of errors b_n and a_n. Moreover, a variable relaxation parameter λ_n gives added flexibility. It is important to note tha the solution to Problem 11 is obtained as the image under $\text{prox}_{\gamma f_2}$ of the weak limit of $(x_n)_{n\in\mathbb{N}}$ and that in general little is known about the asymptotic behavior of $\left(\text{prox}_{\gamma f_2} x_n\right)_{n\in\mathbb{N}}$ unless $\text{prox}_{\gamma f_2}$ is weakly continuous. This is in particular the case when $\dim\mathcal{H} < +\infty$ (by continuity of $\text{prox}_{\gamma f_2}$, see Lemma 4(ii)) and we can then formulate a sharper convergence result, which is immadiately relevant in numerical computations with discretized data.

Corollary 17 *Suppose that $\dim\mathcal{H} < +\infty$ in Theorem 16. Then $\left(\text{prox}_{\gamma f_2} x_n\right)_{n\in\mathbb{N}}$ converges (strongly) to a solution to Problem 11.*

5. NUMERICAL RESULTS

We consider...

Fig. 1. Original image.

Fig. 2. Noisy image.
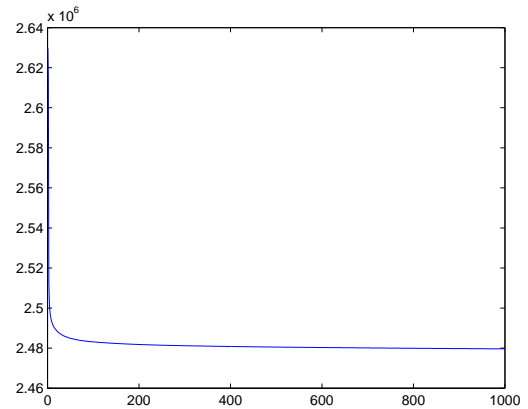

Fig. 3. Denoised image.

Fig. 4. Denoised image.

Fig. 5. Convergence profile.

6. REFERENCES

[1] P. L. Combettes, "Solving monotone inclusions via compositions of nonexpansive averaged operators," *Optimization,* vol. 53, pp. 475–504, 2004.

[2] P. L. Combettes and J. C. Pesquet, "Image restoration subject to a total variation constraint," *IEEE Trans. Image Proces.,* vol. 13, pp. 1213–1222, 2004.

[3] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Model. Simul.,* vol. 4, 2005.

[4] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Comm. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.

[5] C. de Mol and M. Defrise, "A note on wavelet-based inversion algorithms," *Contemp. Math.*, vol. 313, pp. 85–96, 2002.

[6] D. Dung, "Non-linear approximations using wavelet decompositions," *Vietnam J. Math.,* vol. 29, pp. 197–224, 2001.

[7] P. L. Lions and B. Mercier, "Splitting algorithms for the sum of two nonlinear operators," *SIAM J. Numer. Anal.,* vol. 16, pp. 964–979, 1979.

[8] D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1 – theory," *IEEE Trans. Medical Imaging,* vol. 1, pp. 81-94, Oct. 1982.

[9] C. Zălinescu, *Convex Analysis in General Vector Spaces.* River Edge, NJ: World Scientific, 2002.

```
  in long paper let us not forget  to discuss
Spingarn
```