Alternating Direction Method for image restoration: application to 3D biological imaging

Mikael Carlavan*, Pierre Weiss†, Laure Blanc-Féraud*

*Projet ARIANA, CNRS/I3S/INRIA
†Institut de Mathématiques de Toulouse, Université Paul Sabatier

August 9, 2010

Variational methods

Variational methods in image processing often result in solving:

$$\begin{array}{rl} \arg\min & f_1(Au) + \tau f_2(Bu) \\ \text{suject to} & u \in \mathbb{R}^n \end{array} \tag{1}$$

where:

- f_1 and f_2 are two closed convex functions.
- A and B are two linear transforms.

Variational methods

For example, considering a blurred and noisy acquisition z of an image u:

$$z = \Phi u + b \tag{2}$$

where:

- Φ is a blur function.
- b is an additive white gaussian noise.

Then, a restored image can be obtained by solving:

$$\begin{array}{rl} \arg\min & \|\Phi u - z\|_2^2 + \tau \|\nabla u\|_1, \\ \text{suject to} & u \in \mathbb{R}^n \end{array} \tag{3}$$

with τ being a regularizing parameter.

Motivations

Problem (1) is difficult to solve:

- It is not differentiable or its gradient is not Lipschitz.
- It deals with a significant amount of data.
- A and B are sources of difficulties.

We can split algorithms solving these problems in two categories:

- Ones which solve the problem in the space it has been set.
- Ones which add constraints or increase the dimensionality of the problem.

Minimization algorithms without transformation

- (Projected) sub-gradient descent [Polyak1987].
- Forward-backward splitting techniques [Combettes2005].
- Convergence acceleration techniques [Nesterov2009].
- Smoothing techniques [Weiss2009].
- Douglas-Rachford techniques [Combettes2007].

Minimization algorithms with transformation

- Douglas-Rachford splitting techniques.
- Dual formulation (for strongly convex problems) [Chambolle2004].
- Alternating direction techniques [Gabay1983].

Alternating direction technique

Framework:

$$\begin{array}{rll} \text{Find } (x^*, y^*) \in & \arg\min & f_1(x) + f_2(y) \\ & \text{suject to} & Ax + By = a \\ & & x \in \mathbb{R}^n, y \in \mathbb{R}^m \end{array} \quad (4)$$

where:

- $f_1 : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $f_2 : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ are the two closed convex functions.
- $A \in \mathbb{R}^{l \times n}$ and $B \in \mathbb{R}^{l \times m}$ are the two linear transforms.
- $a \in \mathbb{R}^l$ is a given vector.

Alternating direction technique

Augmented Lagrangian (with Lagrange multiplier $\lambda \in \mathbb{R}^l$) of this problem writes:

$$\mathcal{L}(x, y, \lambda) := f_1(x) + f_2(y) + \langle \lambda, Ax + By - a \rangle + \frac{\beta}{2} \|Ax + By - a\|_2^2 \quad (5)$$

where $\beta > 0$ is a parameter which favours the constraint.

Finding a saddle point of $\mathcal{L}(x, y, \lambda)$ gives one solution of the problem.
$\rightarrow$ Alternating minimization of (5) following x, y and λ converges to (x^*, y^*).

Alternating direction algorithm

Algorithm consists in three steps:

1. $x^{k+1} = \begin{array}{ll} \arg\min & \mathcal{L}(x, y^k, \lambda^k) \\ \text{suject to} & x \in \mathbb{R}^n \end{array}$
2. $y^{k+1} = \begin{array}{ll} \arg\min & \mathcal{L}(x^{k+1}, y, \lambda^k) \\ \text{suject to} & y \in \mathbb{R}^m \end{array}$
3. $\lambda^{k+1} = \lambda^k + \beta\gamma(Ax^{k+1} + By^{k+1} - a)$

γ is a relaxation parameter which has to be in $]0, \frac{\sqrt{5}+1}{2}[$ in order for (x^k, y^k) to converge to (x^*, y^*) [Glowinski1984].
Under some conditions it converges even if one of the steps is solved approximately [He2002].

Application to biological imaging

In biological imaging, image formation can be modeled by:

$$z = \mathcal{P}(Hu) \tag{6}$$

where:

- H is the PSF of the optical system.
- $\mathcal{P}$ is a Poisson random process.

Maximizing the posteriori probability with prior $p(u) = \alpha \exp\{-\tau \|Wu\|_1\}$ is equivalent to solving:

$$\text{Find } u^* = \begin{array}{rl} \arg\min & 1^T(Hu) - z^T \log(Hu) + \tau \|Wu\|_1 \\ \text{suject to} & u \in \mathbb{R}^n \\ & u \geq 0 \end{array} \tag{7}$$

where W is a linear transform.

Transformation of the problem

First, we can see that:

$$\text{Find } u^* = \begin{array}{rl} \arg\min & 1^T(Hu) - z^T \log(Hu) + \tau \|Wu\|_1 \\ \text{suject to} & u \in \mathbb{R}^n \\ & u \geq 0 \end{array} \quad (8)$$

is equivalent to:

$$\text{Find } u^* = \begin{array}{rl} \arg\min & 1^T v - z^T \log(v) + \tau \|w\|_1 \\ \text{suject to} & u \in \mathbb{R}^n, u \geq 0 \\ & v \in \mathbb{R}^n, v = Hu \\ & w \in \mathbb{R}^m, w = Wu \end{array} \quad (9)$$

Transformation of the problem

We set:

$$x = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m, \quad y = u \in \mathbb{R}^n, \tag{10}$$

$$A = -I, \quad B = \begin{bmatrix} I \\ H \\ W \end{bmatrix}, \quad a = 0 \tag{11}$$

such that:

$$Ax + By = a \tag{12}$$

gives:

$$\begin{cases} v = Hu \\ w = Wu \end{cases} \tag{13}$$

Transformation of the problem

Problem (7) becomes:

$$\begin{aligned} \text{Find } x^* = \quad \underset{\text{suject to}}{\arg\min} \quad & 1^T v - z^T \log(v) + \tau \|w\|_1 + \chi(u) \\ & Ax + By = a \\ & x \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m, \\ & y \in \mathbb{R}^n \end{aligned} \tag{14}$$

with χ being the indicator function on the convex set $\mathbb{R}^n_+$:

$$\chi(u) = \begin{cases} 0 & \text{if} \quad u \in \mathbb{R}^n, u \geq 0 \\ \infty & \text{otherwise} \end{cases} \tag{15}$$

Transformation of the problem

We see that we actually fit into the framework of the ADM method:

$$\begin{aligned} \text{Find } x^* = \quad & \arg\min \quad f_1(x) + f_2(y) \\ & \text{suject to} \quad Ax + By = a, \\ & \qquad\qquad x \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m, y \in \mathbb{R}^n \end{aligned} \tag{16}$$

with:

$$f_1(x) := 1^T v - z^T \log(v) + \tau \|w\|_1 + \chi(u), \quad f_2(y) := 0 \tag{17}$$

In this case, the augmented Lagrangian writes:

$$\mathcal{L}(x, y, \lambda) := f_1(x) + \langle \lambda, By - x \rangle + \frac{\beta}{2} \|By - x\|_2^2 \tag{18}$$

The algorithm in detail

First step of the algorithm is:

$$
\begin{aligned}
x^{k+1} = \quad & \arg\min \quad \mathcal{L}(x, y^k, \lambda^k) \\
& \text{suject to} \quad x \in \mathbb{R}^n \\
= \quad & \arg\min \quad f_1(x) + \langle \lambda^k, By^k - x \rangle + \frac{\beta}{2} \| By^k - x \|_2^2 \\
& \text{suject to} \quad x \in \mathbb{R}^n \\
= \quad & \arg\min \quad \frac{1}{\beta} f_1(x) + \frac{1}{2} \| By^k - x + \frac{\lambda^k}{\beta} \|_2^2 \\
& \text{suject to} \quad x \in \mathbb{R}^n \\
= & \operatorname{prox}_{\frac{1}{\beta} f_1} \left(By^k + \frac{\lambda^k}{\beta} \right)
\end{aligned}
\tag{19}
$$

where prox designates the proximal operator [Combettes2005]. It generalizes the notion of projection, so that $\forall x_0 \in \mathbb{R}^n$:

$$
\operatorname{prox}_{\gamma \Psi} (x_0) = \begin{array}{ll} \arg\min & \gamma \Psi(x) + \frac{1}{2} \| x_0 - x \|_2^2 \\ \text{suject to} & x \in \mathbb{R}^n \end{array}
\tag{20}
$$

The algorithm in detail

Some examples of closed-form proximal operation:

- $\Psi(x) = \|x\|_1$ then $\mathrm{prox}_{\gamma\Psi}\left(\ x_0\ \right)$ is the soft-thresholding operator of threshold γ given by:

$$\mathrm{prox}_{\gamma\Psi}\left(\ x_0\ \right) = \mathrm{sign}(x_0)\max(|x_0| - \gamma, 0) \tag{21}$$

- $\Psi(x) = 1^T x - z^T \log(x)$ then:

$$\mathrm{prox}_{\gamma\Psi}\left(\ x_0\ \right) = \frac{1}{2}\left(x_0 - \gamma + \sqrt{(x_0 - \gamma)^2 + 4\gamma z}\right) \tag{22}$$

- $\Psi(x) = \chi_C(x)$ is the indicator function on a convex set C, then:

$$\mathrm{prox}_{\gamma\Psi}\left(\ x_0\ \right) = \Pi_C(x_0) \tag{23}$$

is the orthogonal projection on this set.

The algorithm in detail

Referring back to (19), let $By^k + \frac{\lambda^k}{\beta} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$, then:

$$x^{k+1} = \begin{pmatrix} \max(u, 0) \\ \frac{1}{2}\left[v - \frac{1}{\beta} + \sqrt{\left(v - \frac{1}{\beta}\right)^2 + \frac{4z}{\beta}} \right] \\ \mathrm{sign}(w)\max(|w| - \frac{\tau}{\beta}, 0) \end{pmatrix} \tag{24}$$

The algorithm in detail

Second step of the algorithm is:

$$\begin{aligned} y^{k+1} = & \quad \arg\min \quad \mathcal{L}(x^{k+1}, y, \lambda^k) \\ & \quad \text{suject to} \quad y \in \mathbb{R}^m \\ = & \quad \arg\min \quad f_1(x^{k+1}) + \langle \lambda^k, By - x^{k+1} \rangle + \frac{\beta}{2} \| By - x^{k+1} \|_2^2 \\ & \quad \text{suject to} \quad x \in \mathbb{R}^n \\ = & \quad \arg\min \quad \| By - x^{k+1} + \frac{\lambda^k}{\beta} \|_2^2 \\ & \quad \text{suject to} \quad y \in \mathbb{R}^n \\ = & \; (H^* H + W^* W + I)^{-1} B^* \left(x^{k+1} - \frac{\lambda^k}{\beta} \right) \end{aligned} \tag{25}$$

which can be solved exactly in the Fourier domain (depending on W) or approximately with a Conjugate Gradient method.

Results

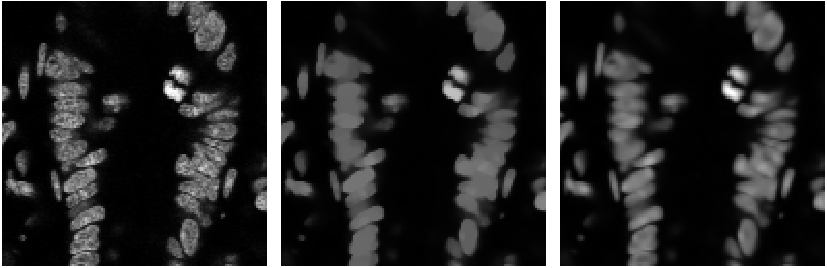

Figure 1: Restoration of a sample of mouse intestines. From left to right: original image, result obtained with a Total Variation prior, result obtained with a wavelet prior.

Conlusion

- Algorithm adapted to many problems including constrained problems.
- Fast convergence even if each step is solved approximately.
- Computing time of 25 minutes on a 256*256*64 voxels image.

- However, the algorithm involves significant memory ressources.