Continuous Primal-Dual Methods in Image Processing

Michael Goldman

CMAP, Polytechnique

August 2012

Introduction

Maximal monotone operators

Application to the initial problem

Numerical illustration

Introduction

Many problems in image processing can be solved by minimizing

$$J(u) = \int_\Omega |Du| + G(u)$$

where G is a convex lsc function on L^2.

Example : the denoising using the ROF model corresponds to $G(u) = \frac{\lambda}{2}|u - f|^2$

can be used for zooming, deblurring, inpainting etc...

Our approach extends to :

- more general convex functionals with at least linear growth

$$J(u) = \int_\Omega F(x, Du) + G(u)$$

  where F is convex in p and $F(x,p) \geq C|p|^\alpha$ with $\alpha \geq 1$,
- problems with boundary conditions.

Idea of the method

Reminder : The total variation is defined as

$$\int_\Omega |Du| = \sup_{\substack{\xi \in \mathcal{C}^1_c(\Omega) \\ |\xi|_\infty \leq 1}} \int_\Omega u \,\mathrm{div}\, \xi$$

The minimization problem then reads

$$\min_{u \in BV} J(u) = \min_{u \in BV} \sup_{\substack{\xi \in \mathcal{C}^1_c(\Omega) \\ |\xi|_\infty \leq 1}} - \int_\Omega u \,\mathrm{div}\, \xi + G(u)$$

$\Rightarrow$ It can thus be recasted as a saddle point problem

The Arrow-Hurwicz Method

For a function K, the Arrow-Hurwicz method reads

$$\begin{cases} \frac{\partial u}{\partial t} = -\nabla_u K(u, \xi) \\ \frac{\partial \xi}{\partial t} = \nabla_\xi K(u, \xi) \end{cases}$$

It is a gradient descent in the Primal variable u and a gradient ascent in the Dual variable ξ.

If $K(u,\xi) = -\int_\Omega u \operatorname{div} \xi + G(u)$ then

$$\begin{cases} \nabla_u K = -\operatorname{div} \xi + \partial G(u) \\ \nabla_\xi K = Du \end{cases}$$

If $K(u,\xi) = -\int_\Omega u \operatorname{div} \xi + G(u)$ then

$$\begin{cases} \nabla_u K = -\operatorname{div} \xi + \partial G(u) \\ \nabla_\xi K = Du \end{cases}$$

which formally leads to :

$$\begin{cases} \frac{\partial u}{\partial t} = \operatorname{div} \xi - \partial G(u) \\ \frac{\partial \xi}{\partial t} = Du \qquad |\xi|_\infty \leq 1 \end{cases}$$

This is exactly the method proposed by Appleton and Talbot. It corresponds to the continuous analogue of the method proposed by Chan and Zhu.

Theorem

The Cauchy problem associated with the previous system as a unique solution.

Moreover, if $G(u) = \frac{\lambda}{2}|u - f|^2$ *then this solution converges toward the minimizer* $\overline{u}$ *of* J *and we have the a posteriori estimate*

$$|u - \overline{u}| \leq \frac{1}{2}\left(\frac{|\partial_t u|}{\lambda} + \sqrt{\frac{|\partial_t u|^2}{\lambda^2} + \frac{8|\Omega|^{\frac{1}{2}}}{\lambda}|\partial_t \xi|}\right)$$

Crucial observation

Formally we have :

$$\begin{pmatrix} -\operatorname{div}\xi + \partial G(u) \\ -Du \end{pmatrix} \cdot \begin{pmatrix} u \\ \xi \end{pmatrix} = \partial G(u) \cdot u \geq 0$$

Thus the operator defining the system is monotone.

Maximal monotone operators

Definition

Let H be a Hilbert space. An operator A on H is monotone if :

$$\forall x_1, x_2 \in D(A), \qquad (A(x_1) - A(x_2), x_1 - x_2) \geq 0.$$

Definition

It is called maximal monotone if it is maximal in the set of monotone operators.

Proposition

Let φ be a convex lsc function on H then $\partial\varphi$ is maximal monotone.

Reminder : $p \in \partial\varphi(x)$ if for every y

$$\varphi(y) - \varphi(x) \geq p \cdot (y - x).$$

Theorem

For every $u_0 \in D(A)$*, there exists a unique function* $u(t)$ *from* $[0, +\infty[$ *in* H *such that*

- $u(t) \in D(A)$ *for every* $t > 0$
- $u(t)$ *is Lipschitz on* $[0, +\infty[$*, i.e* $\frac{du}{dt} \in L^\infty(0, +\infty; H)$*.*
- $-\dfrac{du}{dt} \in A(u(t))$ *for a.e.* t*.*
- $u(0) = u_0$*.*
- *if* u *and* $\hat{u}$ *are two solutions then* $|u(t) - \hat{u}(t)| \leq |u(0) - \hat{u}(0)|$*.*

Application to finding saddle points

Theorem (Rockafellar 68)

Let K be a proper saddle function. Assume that K is lsc in y and usc in z then the associated Arrow-Hurwicz operator T is maximal monotone.

Idea of the proof

Let

$$H(y, z^*) = \sup_z z^* \cdot z + K(y, z)$$

We then have :

Lemma

H is a lsc convex function and

$$(y^*, z^*) \in T(y, z) \Leftrightarrow (y^*, z) \in \partial H(y, z^*)$$

Idea of the proof

Let

$$H(y, z^*) = \sup_z z^* \cdot z + K(y, z)$$

We then have :

Lemma

H is a lsc convex function and

$$(y^*, z^*) \in T(y, z) \Leftrightarrow (y^*, z) \in \partial H(y, z^*)$$

Unfortunately, this theorem doesn't apply directly !

Application to the initial problem

We remind that we look for a saddle point of

$$K(u,\xi) = -\int_\Omega u \,\mathrm{div}\, \xi + G(u)$$

We then let

$$H(u,\xi^*) = \sup_{|\xi|_\infty \le 1} \langle \xi, \xi^* \rangle - \int_\Omega u \,\mathrm{div}\, \xi + G(u)$$
$$= \int_\Omega |Du + \xi^*| + G(u)$$

Application to the initial problem

We remind that we look for a saddle point of

$$K(u,\xi) = -\int_\Omega u \operatorname{div} \xi + G(u)$$

We then let

$$H(u,\xi^*) = \sup_{|\xi|_\infty \le 1} \langle \xi, \xi^* \rangle - \int_\Omega u \operatorname{div} \xi + G(u)$$
$$= \int_\Omega |Du + \xi^*| + G(u)$$

<u>H **is a lsc convex function !**</u>

We can thus define the maximal monotone T by

$$(u^*, \xi^*) \in T(u, \xi) \Leftrightarrow (u^*, \xi) \in \partial H(u, \xi^*)$$

Problem : compute T.

Characterization of T

Proposition

$(u^*, \xi^*) \in T(u, \xi)$ if and only if :

- $u \in BV \cap L^2$ *and* $\xi \in H_0^1(\mathrm{div})$ *with* $|\xi|_\infty \leq 1$.
- $u^* + \mathrm{div}\,\xi \in \partial G$
- $\displaystyle\int_\Omega |\xi^* + Du| = \langle \xi^*, \xi \rangle + \int_\Omega [\xi, Du]$

About convergence...

Proposition

For the denoising problem, there is convergence towards the minimizer $\overline{u}$.

Key idea of the proof :

It rests on the simple estimate

$$\frac{d}{dt}\left(|u-\overline{u}|^2+|\xi-\overline{\xi}|^2\right)\leq -C|u-\overline{u}|^2$$

and the *a posteriori* estimates

Proposition

There holds the following a posteriori estimates

$$|u-\overline{u}| \leq \frac{1}{2}\left(\frac{|\partial_t u|}{\lambda} + \sqrt{\frac{|\partial_t u|^2}{\lambda^2} + \frac{8|\Omega|^{\frac{1}{2}}}{\lambda}|\partial_t \xi|}\right)$$

Idea of the proof :

We start from

$$\begin{aligned} u &= f + \frac{1}{\lambda}(\operatorname{div}\xi - \partial_t u) \\ \bar{u} &= f + \frac{1}{\lambda}\operatorname{div}\bar{\xi} \end{aligned}$$

To obtain

$$|u-\bar{u}|^2 = \frac{1}{\lambda}\langle \operatorname{div}(\xi - \bar{\xi}) - \partial_t u, u - \bar{u}\rangle$$

Numerical illustration

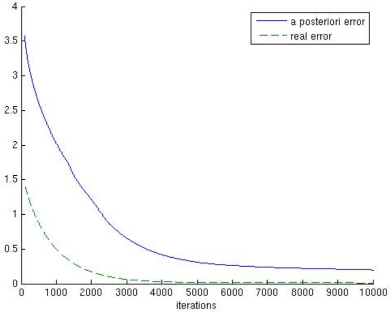
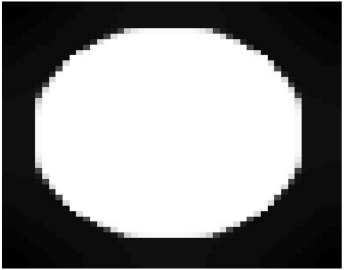
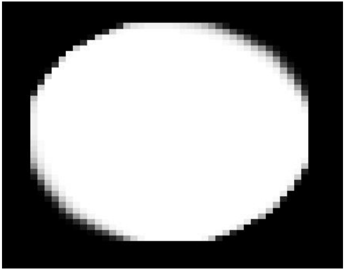


Illustration of the stopping criterion

Interest of the continuous approach

- Leads to a better understanding of the discrete model
- Gives answers that were still unknown in the discrete model
- Gives rise to less anisotropical algorithms

Restauration by AT on the left and CZ on the right

Zoom on the top right corner.