

Algorithmes de premier ordre pour la projection sur une contrainte de variation totale

Gabriel PEYRÉ¹, Jalal FADILI²

¹CEREMADE CNRS-Université Paris Dauphine

²GREYC CNRS-ENSICAEN-Université de Caen

`gabriel.peyre@ceremade.dauphine.fr, jalal.fadili@greyc.ensicaen.fr`

Résumé – Cet article propose un nouvel algorithme pour calculer la projection sur l’ensemble des images dont la variation totale est bornée par une constante. La projection est calculée à l’aide d’une formulation duale qui est résolue par des méthodes d’optimisation non-lisse du premier ordre. Ceci donne naissance à un algorithme calculant des seuillages doux itérés du champ de vecteurs dual. Cet algorithme de projection peut ensuite être utilisé comme un maillon pour la résolution d’un problème inverse sous contrainte de variation totale. Des résultats numériques montrent que notre algorithme est plus efficace que l’état de l’art pour résoudre les problèmes de débruitage, d’inpainting et de déconvolution par projection de variation totale.

Abstract – This article proposes a new algorithm to compute the projection on the set of images whose total variation is bounded by a constant. The projection is computed through a dual formulation that is solved by first order non-smooth optimization methods. This yields an iterative algorithm that computes iterative soft thresholding of the dual vector fields. This projection algorithm can then be used as a building block to solve inverse problems under a total variation constraint. Numerical results show that our algorithm competes favorably with state-of-the-art TV projection methods to solve denoising, inpainting and deconvolution problems.

1 Introduction

La variation totale est un a-priori très utilisé depuis son introduction en analyse d’image par Rudin, Osher et Fatemi [9]. Pour une image discrète $f \in \mathbb{R}^N$ de $N = n \times n$ pixels, la variation totale discrète s’écrit $\|f\|_{\text{TV}} = \|\nabla f\|_1$, où la norme ℓ^1 d’un champ de vecteur $u = (u_1, u_2) \in \mathbb{R}^{N \times 2}$ est

$$\|u\|_1 = \sum_{i,j} \sqrt{u_1[i,j]^2 + u_2[i,j]^2},$$

et le gradient discret

$$\nabla f[i,j] = (f[i+1,j] - f[i,j], f[i,j+1] - f[i,j]),$$

avec conditions aux bords de type Neumann. L’adjoint du gradient est $\nabla^* = -\text{div}$, où la divergence d’un champ de vecteurs $u \in \mathbb{R}^{N \times 2}$ est $\text{div}(u) = u_1[i,j] - u_1[i-1,j] + u_2[i,j] - u_2[i,j-1]$, avec conditions aux bords de type Neumann.

Le débruitage par régularisation utilisant la variation totale d’une image f_0 s’écrit

$$\min_{f \in \mathbb{R}^N} \frac{1}{2} \|f - f_0\|^2 + \lambda \|f\|_{\text{TV}}, \quad (1)$$

où le poids de régularisation λ doit être adapté selon le niveau de bruit ainsi que la régularité de f_0 . De nombreux algorithmes ont été développés pour résoudre ce problème, comme par exemple [2, 10].

2 Projection de variation totale

Projection pour le débruitage. Beaucoup moins de travaux se sont intéressés au débruitage d’une image par le calcul de sa projection sur une boule de variation totale de rayon $\tau < \|f_0\|_{\text{TV}}$, qui nécessite la résolution de

$$f^* = \underset{f \in \mathbb{R}^N, \|f\|_{\text{TV}} \leq \tau}{\text{argmin}} \|f - f_0\|. \quad (2)$$

Une telle formulation est préférable par rapport à (1) quand peu d’information est disponible sur le niveau du bruit, mais lorsque l’on dispose d’une estimée τ de la variation totale de l’image d’origine f_0 . Le calcul de la solution de (2) à l’aide d’un algorithme rapide est donc important comme expliqué par Combettes et Pesquet [4], qui proposent un algorithme de projection de sous-gradient. Dans ce papier, nous proposons une approche différente, qui exploite une formulation duale pénalisée (sans contrainte) de la projection primale (2).

Projection pour les problèmes inverses. Les problèmes inverses linéaires tels que l’inpainting et la déconvolution ont été l’objet d’une activité intense, et l’a priori de variation totale a été fréquemment utilisé pour régulariser l’inversion, voir par exemple [1].

La projection de variation totale peut également être utilisée pour résoudre des problèmes inverses [4]. La section 4 détaille l’utilisation de nos algorithmes de projection duale pour la résolution de problèmes inverses.

Projection duale. La proposition suivante, démontrée dans [6], montre que le problème d'optimisation sous contraintes (2) peut être transformé en un problème pénalisé.

Proposition 1. Pour tout $f \in \mathbb{R}^N$, la solution de (2) s'écrit $f^* = f_0 - \text{div}(u^*)$ où u^* est une solution du problème dual

$$\inf_{u \in \mathbb{R}^{N \times 2}} \frac{1}{2} \|f_0 - \text{div}(u)\|^2 + \tau \|u\|_\infty. \quad (3)$$

avec $\|u\|_\infty = \max_{i,j} \sqrt{u_1[i,j]^2 + u_2[i,j]^2}$.

3 Algorithmes de premier ordre

La résolution de (3) peut s'effectuer à l'aide d'algorithmes de premier ordre, qui sont des extensions à des fonctionnelles non-différentiables de la descente de gradient. Nous considérons ici l'itération explicite-implicite [5] (issue des schémas de décomposition d'opérateurs monotones maximaux), ainsi qu'un algorithme de Nesterov [8].

Opérateur proximal de la norme ℓ^∞ . Ces deux algorithmes utilisent la solution de la régularisation de Moreau-Yosida de la fonctionnelle $\kappa \|\cdot\|_\infty$, qui est l'unique solution du problème convexe

$$\text{prox}_{\kappa \|\cdot\|_\infty}(u) = \underset{v \in \mathbb{R}^{N \times 2}}{\text{argmin}} \frac{1}{2} \|u - v\|^2 + \kappa \|v\|_\infty. \quad (4)$$

La proposition suivante, démontrée dans [6], montre que cet opérateur proximal se calcule explicitement à l'aide d'un seuillage doux S_λ appliqué au champ de vecteur dual u . Pour obtenir cette valeur de λ , il est nécessaire de calculer $d[0] \leq d[1] \leq \dots \leq d[N-1]$ qui sont les valeurs ordonnées de l'ensemble des normes

$$\{d[t]\}_{t=0}^{N-1} = \{|u[i,j]|\}_{i,j=0}^{n-1}, \quad (5)$$

ainsi que les normes cumulées

$$D[s] = \sum_{t=s+1}^{N-1} d[t]. \quad (6)$$

Proposition 2. Pour $u \in \mathbb{R}^{N \times 2}$, avec d et D définis en (5)-(6), on a $\text{prox}_{\kappa \|\cdot\|_\infty}(u) = 0$ si $\|u\|_1 \leq \kappa$ et $\text{prox}_{\kappa \|\cdot\|_\infty}(u) = u - S_\lambda(u)$ dans le cas contraire, où

$$S_\lambda(u)[i,j] = \max\left(1 - \frac{\lambda}{|u[i,j]|}, 0\right) u[i,j] \quad (7)$$

et $\lambda > 0$ est donné par

$$\lambda = d[t] + (d[t+1] - d[t]) \frac{D[t+1] - \kappa}{D[t+1] - D[t]} \quad (8)$$

où t est tel que $D[t+1] \leq \kappa < D[t]$.

Projection par itération explicite-implicite. Une itération de l'algorithme de projection explicite-implicite pour résoudre (3) s'écrit

$$u^{(k+1)} = \text{prox}_{\mu\tau \|\cdot\|_\infty}\left(u^{(k)} + \mu \nabla\left(f_0 - \text{div}(u^{(k)})\right)\right) \quad (9)$$

L'opérateur proximal pour $\kappa = \mu\tau$ est donnée par la Proposition 2. Pour $\mu \in]0, 1/4[$, on peut montrer que $u^{(k)}$ converge vers u^* une solution de (3), avec une vitesse sur l'objectif dual en $O(1/k)$, voir par exemple [8]. Le théorème suivant, démontrée dans [6], montre que l'on a également une vitesse de convergence sur les itérés primaux vers l'unique solution de (2).

Theorem 1. Soit $\mu \in]0, 1/4[$ et $u^{(0)} \in \mathbb{R}^{N \times 2}$. Il existe $C > 0$ tel que

$$\|f^{(k)} - f^*\|^2 \leq C/k,$$

où $f^{(k)} = f_0 - \text{div}(u^{(k)})$ et $u^{(k)}$ est défini en (9).

Projection par algorithme de Nesterov. Y. Nesterov propose dans [8, 7] un schéma de gradient multi-pas pour optimiser la somme d'une fonctionnelle convexe régulière avec un gradient Lipschitz-continu et d'une fonctionnelle convexe non-différentiable dont la structure est simple ; c'est-à-dire dont l'opérateur proximal se calcule explicitement. Le problème dual (3) rentre dans ce cadre, et l'algorithme 1 détaille les étapes du schéma de Nesterov pour minimizer (3). Il est formulé à l'aide d'opérateurs proximaux, comme décrit dans [10].

Algorithm 1: Projection de variation totale de Nesterov .

Initialisation : $u^{(0)} \in \mathbb{R}^{N \times 2}$, $A_0 = 0$, $\xi^{(0)} = 0$, $\mu < 1/4$.

Itération principale :

while $\|u^{(k+1)} - u^{(k)}\| > \eta$ **do**

1. *Premier calcul proximal :*

$$v^{(k)} = \text{prox}_{A_k \tau \|\cdot\|_\infty}(u^{(0)} - \xi^{(k)}),$$

où l'opérateur proximal est calculé comme détaillé à la proposition 2 avec $\kappa = A_k \tau$.

2. Soit $a_k = (\mu + \sqrt{\mu^2 + 4\mu A_k})/2$ et

$$\omega^{(k)} = \frac{A_k u^{(k)} + a_k v^{(k)}}{A_k + a_k}.$$

3. *Deuxième calcul proximal :*

$$\tilde{\omega}^{(k)} = \omega^{(k)} - \frac{\mu}{2} \nabla\left(f_0 - \text{div}(\omega^{(k)})\right),$$

$$u^{(k+1)} = \text{prox}_{\mu\tau/2 \|\cdot\|_\infty}(\tilde{\omega}^{(k)}),$$

où l'opérateur proximal est calculé comme détaillé à la proposition 2 avec $\kappa = \mu\tau/2$.

4. Mettre à jour $A_{k+1} = A_k + a_k$ et $\xi^{(k+1)} = \xi^{(k)} + a_k \nabla(f_0 - \text{div}(u^{(k+1)}))$.

5. $k \leftarrow k + 1$, définir $f^{(k)} = f_0 - \text{div}(u^{(k)})$

Output : $f^* = f_0 - \text{div}(u^{(k+1)})$.

Pour $\mu \in]0, 1/4[$, Nesterov montre que $u^{(k)}$ converge vers u^* une solution de (3), avec une vitesse sur l'objectif dual en $O(1/k^2)$, voir [8, 7]. Le théorème suivant, démontrée dans [6], montre que l'on a également une vitesse de convergence sur les itérés primaux.

Theorem 2. Soit $\mu \in]0, 1/4[$ et $u^{(0)} \in \mathbb{R}^{N \times 2}$. Il existe $C > 0$ tel que la suite $f^{(k)}$ de l'algorithme 1 vérifie

$$\|f^{(k)} - f^*\|^2 \leq C/k^2.$$

Notons au passage qu'une distinction importante entre notre travail et celui dans [10, 7] est que la partie non-différentiable du problème dual n'est pas à domaine borné ici, alors que cette hypothèse est primordiale dans [10, 7]. De plus, nous exhibons la convergence sur les itérés primales et non pas sur l'objectif.

4 Problèmes inverses

Nos algorithmes de projection TV permettent également de résoudre un problème inverse mal posé où

$$y = \Phi f + w, \quad (10)$$

avec Φ un opérateur linéaire mal conditionné et w un bruit additif. On calcule alors

$$f^* = \operatorname{argmin}_{f \in \mathbb{R}^N} \|\Phi f - y\| \quad \text{avec} \quad \|f\|_{\text{TV}} \leq \tau. \quad (11)$$

On peut calculer une solution à ce problème par descente de gradient projetée

$$f^{(\ell+1)} = \operatorname{Proj}_{\| \cdot \|_{\text{TV}} \leq \tau}(\tilde{f}^{(\ell)}) \quad \text{où} \quad \tilde{f}^{(\ell)} = f^{(\ell)} + \nu \Phi^*(y - \Phi f^{(\ell)})$$

pour $0 < \nu < 2/\|\Phi\|^2$, où $\operatorname{Proj}_{\| \cdot \|_{\text{TV}} \leq \tau}$ est l'opérateur de projection sur la boule TV défini en (2). On peut démontrer que les itérés $f^{(\ell)}$ convergent vers une solution de (11), voir [6].

Le calcul de la projection $\operatorname{Proj}_{\| \cdot \|_{\text{TV}} \leq \tau}(\tilde{f}^{(\ell)})$ nécessite l'utilisation de l'algorithme de projection explicite-implicite (équation (9)) ou de l'algorithme de Nesterov 1 avec une tolérance $\eta = \eta^{(\ell)}$. Ce calcul approché engendre des erreurs à chaque itération, mais il est possible, comme expliqué dans [6], de contrôler la décroissance de la tolérance $\eta^{(\ell)}$ au cours des itérations pour que les projections approchées $f^{(\ell)}$ convergent encore vers la solution f^* .

Bien que le problème (11) puisse à son tour être résolu avec un schéma de Nesterov, on ne dispose pas à l'heure actuelle de preuve de stabilité d'un tel schéma lors de calculs approchés des itérés. C'est pour cette raison que nous considérons uniquement dans cet article une méthode de descente de gradient projeté.

5 Exemples numériques

Débruitage. La figure 1 montre des débruitages f^* calculés avec notre algorithme de projection TV dual, pour des valeurs décroissantes du rayon de la contrainte τ . L'image observée est $f_0 = \tilde{f}_0 + w$, avec $N = 512^2$ pixels contaminés par un bruit blanc Gaussien d'écart-type $0.06\|\tilde{f}_0\|_\infty$.

L'estimé $f^{(k)}$ de f^* calculé à l'itération k est $f^{(k)} = f_0 - \operatorname{div}(u^{(k+1)})$. La figure 2 compare les vitesses de convergence obtenues avec les algorithmes de projection explicite-implicite,

Nesterov, et la projection de sous-gradient proposée dans [3, 4]. Comme le coût d'une itération de l'algorithme de Nesterov est environ le double de celui des autres algorithmes, nous montrons en fait l'erreur générée par $f^{(k/2)}$ au lieu de $f^{(k)}$ pour la courbe de la méthode de Nesterov. Ceci montre que les vitesses observées coïncident avec les prédictions théoriques, et que l'algorithme de projection dual avec itérations de Nesterov est plus rapide que les deux autres méthodes.



FIGURE 1 – Exemples de débruitage par projections sur la boule TV pour différents τ calculées par notre algorithme.

Inpainting. L'inpainting tente de restaurer une image f_0 de laquelle un ensemble $\Omega \subset \{0, \dots, n-1\}^2$ de pixels est manquant. Ceci correspond au problème inverse mal posé (10) avec $(\Phi f)[i, j]$ égal à $f[i, j]$ si $(i, j) \notin \Omega$ et 0 sinon.

La figure 3, haut, montre une image endommagée y , avec $|\Omega|/N = 0.7$ pixels enlevé aléatoirement et un bruit blanc Gaussien additif tel que $\|w\| = 0.05\|f_0\|_\infty$ (PSNR=13 dB). L'image est restaurée en minimisant (11) avec une contrainte de variation totale $\tau = 0.6\|f_0\|_{\text{TV}}$.

Déconvolution. Un système optique introduit un flou modélisé par un opérateur Φ qui est une convolution circulaire avec un noyau φ .

La figure 3, bas, montre un exemple de restauration d'une image de $N = 512^2$ pixels convoluée par un noyau Gaussien φ d'écart-type 4 pixels, et bruitée par un bruit blanc Gaussien additif tel que $\|w\| = 0.02\|f_0\|_\infty$ (PSNR=17 dB). L'image déconvoluée est obtenue avec une projection sur la boule de variation totale de taille $\tau = 0.6\|f_0\|_{\text{TV}}$. La courbe de droite montre la décroissance en échelle logarithmique de l'erreur de conver-

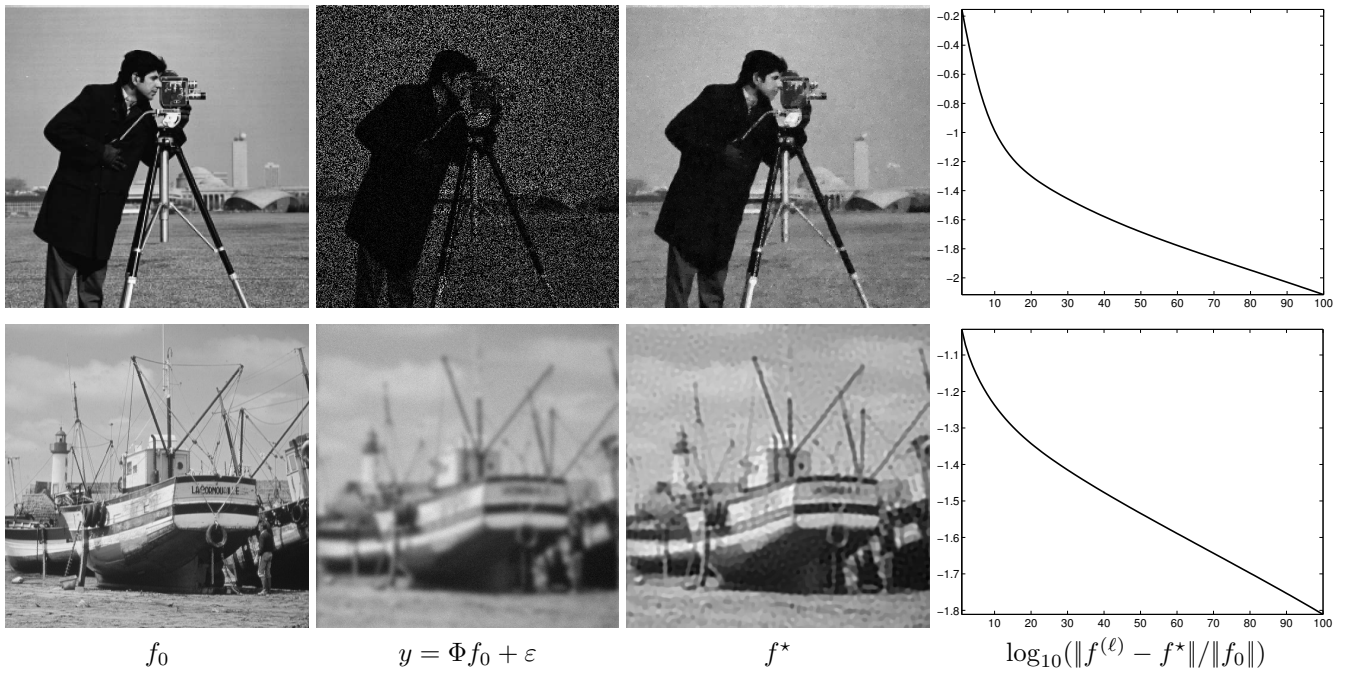


FIGURE 3 – Exemples d'inpainting (première ligne) et deconvolution (deuxième ligne) résolus par projection de variation totale. Les courbes de droite montre la décroissant de l'erreur.

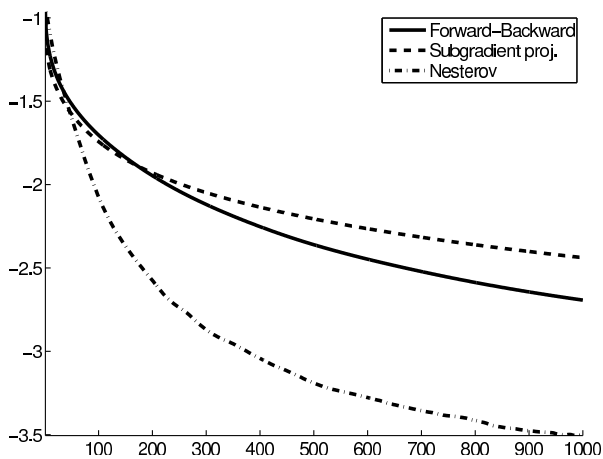


FIGURE 2 – Décroissance avec k de l'erreur $\log_{10}(\|f^{(k)} - f^*\|/\|f^*\|)$ pour l'algorithme dual avec itération explicite-implicite (trait plein), avec itérations de Nesterov (trait plein-pointillés) et par projection de sous-gradient [4] (trait en pointillés), avec $\tau = \|f_0\|_{TV}/4$.

gence.

Remerciements. Ce travail est soutenu par l'ANR NatImages, financement ANR-08-EMER-009. Nous remercions chaleureusement Pierre Weiss pour son aide et ses relectures lors de la préparation de ce manuscrit.

Références

- [1] J. Bect, L. Blanc Féraud, G. Aubert, and A. Chambolle. A ℓ_1 -unified variational framework for image restoration. In *Proc. of ECCV04*, pages Vol IV : 1–13. Springer-Verlag, 2004.
- [2] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20 :89–97, 2004.
- [3] P.-L. Combettes. A block-iterative surrogate constraint splitting method for quadratic signal recovery. *IEEE Trans. Signal Proc.*, 51(7) :1771–1782, 2003.
- [4] P.-L. Combettes and J.-C. Pesquet. Image restoration subject to a total variation constraint. *IEEE Trans. Image Proc.*, 13(9) :1213–1222, 2004.
- [5] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM Journal on Multiscale Modeling and Simulation*, 4(4), 2005.
- [6] M. J. Fadili and G. Peyré. Total variation projection with first order schemes. *Preprint Hal-00380491*, 2009. submitted.
- [7] Y. Nesterov. Smooth minimization of non-smooth functions. *Math. Program.*, 103(1, Ser. A) :127–152, 2005.
- [8] Y. Nesterov. Gradient methods for minimizing composite objective function. CORE Discussion Papers 2007076, Université catholique de Louvain, CORE, September 2007.
- [9] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4) :259–268, 1992.
- [10] P. Weiss, L. Blanc-Féraud, and G. Aubert. Efficient schemes for total variation minimization under constraints in image processing. *SIAM Journal on Scientific Computing*, 2009. to appear.