

Reconstruction d’image en présence de bruit gaussien dépendant par un algorithme Explicite-Implicite à métrique variable.

Audrey REPETTI, Emilie CHOUZENOUX et Jean-Christophe PESQUET

Laboratoire d’Informatique Gaspard Monge, Université Paris-Est Marne la Vallée
Cité Descartes, 5 Boulevard Descartes, Champs sur Marne, 77454 Marne la Vallée Cedex 2, France
{audrey.repetti, emilie.chouzenoux, jean-christophe.pesquet}@univ-mlv.fr

Thème – Traitement et analyse

Problème traité – Reconstruction d’une image dégradée par un opérateur linéaire et perturbée par un bruit additif gaussien dont la variance dépend linéairement de l’image originale. La fonction d’attache aux données considérée (obtenue grâce à une estimation par maximum *a posteriori*) est non convexe et la fonction de régularisation est non lisse.

Originalité – Nous proposons de résoudre ce problème grâce à un algorithme Explicite-Implicite (*Forward-Backward*, FB) accéléré par l’introduction d’une métrique variable. Cette métrique est choisie suivant la théorie de la Majoration-Minimisation. L’algorithme présenté permet de traiter la somme d’une fonction (non nécessairement convexe) différentiable et d’une fonction (non nécessairement différentiable) convexe.

Résultats – Nous montrons la convergence de l’algorithme proposé en utilisant des résultats récents de l’analyse non lisse. Nous présentons des résultats expérimentaux illustrant l’accélération produite sur la vitesse de convergence du critère et celle des itérées. Les courbes de convergence obtenues montrent que notre algorithme se compare favorablement à l’algorithme FB classique et à ses formes accélérées telles que FISTA.

1 Introduction

On s’intéresse à l’estimation $\hat{x} \in \mathbb{R}^N$ d’une image originale $\bar{x} \in \mathbb{R}^N$ à partir d’observations $z \in \mathbb{R}^M$ de la forme $z = H\bar{x} + b$, où $H \in \mathbb{R}^{M \times N}$ est une matrice modélisant la dégradation subie par l’image (*e.g.* un opérateur de convolution) et $b \in \mathbb{R}^M$ est un bruit additif. Une stratégie efficace pour traiter ce problème est de définir \hat{x} comme étant un minimiseur d’un critère pénalisé $G = F + R$, où F est une fonction d’attache aux données et R est une fonction de régularisation qui favorise certaines propriétés *a priori* de la solution. Nous supposons que F est différentiable de gradient Lipschitz et R est convexe, propre, semi-continue inférieurement (sci). Dans le cas où R est lisse, des algorithmes efficaces pour résoudre ce problème de minimisation sont notamment les algorithmes de gradient préconditionné [1] et de Majoration-Minimisation (MM) [2]. Dans le cas où R est non lisse, l’algorithme Explicite-Implicite (*Forward-Backward*, FB) [3] peut être utilisé. Cependant, ce dernier peut souffrir de lenteurs de convergence. Deux approches ont été proposées dans la littérature pour remédier à ce problème. La première repose sur une accélération de type sous-espace qui utilise des informations des itérations précédentes pour construire la nouvelle itérée (par exemple, l’algorithme FISTA [4]). La seconde s’appuie sur l’emploi d’une métrique variable, c’est-à-dire d’une modification à chaque itération de la métrique sous-jacente. Cette stratégie est utilisée par l’algorithme Explicite-Implicite à Métrique Variable (*Variable Metric Forward-Backward*, VMFB) [5, 6] :

$$\begin{aligned} &x_0 \in \text{dom } R \\ &\text{Pour } k = 0, 1, \dots \\ &\left[\begin{array}{l} y_k = \text{prox}_{\gamma_k^{-1}A_k, R}(x_k - \gamma_k A_k^{-1} \nabla F(x_k)), \\ x_{k+1} = x_k + \lambda_k (y_k - x_k), \end{array} \right. \end{aligned} \tag{1}$$

où, pour tout $k \in \mathbb{N}$, $\nabla F(x_k)$ est le gradient de F en x_k , $A_k \in \mathbb{R}^{N \times N}$ est une matrice symétrique définie positive et $(\gamma_k, \lambda_k) \in]0, +\infty[^2$. De plus, pour tout $x \in \mathbb{R}^N$, $\text{prox}_{\gamma_k^{-1}A_k, R}(x)$ est l'opérateur proximal de R en x relatif à la métrique induite par $\gamma_k^{-1}A_k$ [5], i.e. l'unique minimiseur de $R + \frac{1}{2}\|\cdot - x\|_{\gamma_k^{-1}A_k}^2$ où, pour tout $y \in \mathbb{R}^N$, $\|y\|_{\gamma_k^{-1}A_k}^2 = y^\top (\gamma_k^{-1}A_k)y$. On remarque que, d'une part, si $R \equiv 0$, alors on retrouve un algorithme de gradient préconditionné. D'autre part, si, pour tout $k \in \mathbb{N}$, $A_k = I_N$, où I_N est la matrice identité de \mathbb{R}^N , alors on retrouve l'algorithme FB. En se limitant au cas où F est convexe, la convergence de l'algorithme (1) a été démontrée dans [5]. Cependant, aucune méthode constructive n'a été présentée pour choisir les matrices $(A_k)_{k \in \mathbb{N}}$.

Dans cet article, nous proposons une stratégie de construction des matrices $(A_k)_{k \in \mathbb{N}}$ basée sur une approche MM. Ce choix conduit à un nouveau résultat de convergence pour l'algorithme VMFB, qui reste valide dans le cas où la fonction F est non convexe. Nous illustrons les performances de l'approche proposée sur un exemple de reconstruction d'image.

2 Méthode proposée

Choix des métriques. Nous proposons de définir les matrices $(A_k)_{k \in \mathbb{N}}$, servant à définir la métrique variable de l'algorithme (1), en utilisant la théorie MM. Pour tout $k \in \mathbb{N}$, la matrice A_k est choisie de façon à ce que la fonction quadratique $Q(\cdot, x_k) = F(x_k) + (\cdot - x_k)^\top \nabla F(x_k) + \frac{1}{2}(\cdot - x_k)^\top A_k(\cdot - x_k)$ soit une fonction majorante de F en x_k sur le domaine de R , c'est-à-dire :

$$(\forall x \in \text{dom } R) \quad F(x) \leq Q(x, x_k) \quad \text{et} \quad F(x_k) = Q(x_k, x_k). \quad (2)$$

Par ailleurs, A_k est supposée symétrique et bornée au sens où il existe $(\underline{\nu}, \bar{\nu}) \in]0, +\infty[^2$, avec $0 < \underline{\nu} \leq \bar{\nu}$, tel que, pour tout $x \in \mathbb{R}^N$, $\underline{\nu}\|x\|^2 \leq \|x\|_{A_k}^2 \leq \bar{\nu}\|x\|^2$. Remarquons qu'il est toujours possible de construire de telles matrices. En particulier, (2) est vérifiée si $A_k = LI_N$, où L est la constante de Lipschitz de ∇F ([Prop. A.24][1]).

Théorème de convergence. La construction proposée pour les matrices $(A_k)_{k \in \mathbb{N}}$ conduit au résultat de convergence suivant décrivant le comportement asymptotique de la suite $(x_k)_{k \in \mathbb{N}}$ générée par l'algorithme (1) lorsque F n'est pas nécessairement convexe.

Théorème 1. *Supposons F de gradient L -Lipschitz sur $\text{dom } R$ et R convexe, propre, sci sur \mathbb{R}^N et continue sur son domaine. De plus, supposons que G est coercive et vérifie l'inégalité de Kurdyka-Łojasiewicz [7]. Si les conditions suivantes sont vérifiées :*

- (i) $0 < \inf_{l \in \mathbb{N}} (\lambda_l \gamma_l)$ et $\sup_{l \in \mathbb{N}} (\lambda_l \gamma_l) < 2$,
- (ii) pour tout $k \in \mathbb{N}$, $0 < \inf_{l \in \mathbb{N}} \lambda_l \leq \lambda_k \leq 1$,
- (iii) il existe $\underline{\eta} \in]0, 1]$ tel que, pour tout $k \in \mathbb{N}$, $G(x_{k+1}) \leq (1 - \underline{\eta})G(x_k) + \underline{\eta}G(y_k)$,

alors $\lim_{k \rightarrow +\infty} x_k = \hat{x}$, où \hat{x} est un point critique de G .

Remarquons que si, pour tout $k \in \mathbb{N}$, x_{k+1} est choisi tel que $G(x_{k+1}) \leq G(y_k)$ (c'est le cas en particulier pour $\lambda_k \equiv \underline{\eta} \equiv 1$), alors la condition (iii) est vérifiée. D'autre part, l'opérateur proximal n'étant pas toujours calculable de manière explicite ([6]), le Théorème 1 est démontré dans [8] pour une version inexacte de l'algorithme (1).

3 Application au problème de reconstruction d'image

Modèle. Afin d'illustrer les performances de l'algorithme VMFB, nous considérons un problème de reconstruction d'image en présence de bruit gaussien dépendant. Les observations sont de la forme :

$$(\forall m \in \{1, \dots, M\}) \quad z^{(m)} = [H\bar{x}]^{(m)} + \sqrt{\alpha^{(m)}[H\bar{x}]^{(m)} + \beta^{(m)}} w^{(m)}, \quad (3)$$

où $H \in [0, +\infty[^{M \times N}$ est une matrice de projection, $[Hx]^{(m)}$ est la m -ème composante de Hx , $(\alpha^{(m)})_{1 \leq m \leq M} \in [0, +\infty[^M$, $(\beta^{(m)})_{1 \leq m \leq M} \in [0, +\infty[^M$ et $(w^{(m)})_{1 \leq m \leq M}$ est une réalisation d'un vecteur aléatoire gaussien de moyenne nulle et de matrice de covariance I_M . Le terme d'attache aux données est défini, pour tout $x \in [0, +\infty[^N$, comme étant la log-vraisemblance des données : $F(x) = \frac{1}{2} \sum_{m=1}^M \left(\frac{([Hx]^{(m)} - z^{(m)})^2}{\alpha^{(m)}[Hx]^{(m)} + \beta^{(m)}} + \log(\alpha^{(m)}[Hx]^{(m)} + \beta^{(m)}) \right)$. Les matrices $(A_k)_{k \in \mathbb{N}}$ associées sont construites suivant une approche similaire à [9]. D'autre part, une régularisation hybride est considérée : $R = \iota_{[x_{\min}, x_{\max}]^N} + \sum_{j=1}^J \vartheta^{(j)} |[W \cdot]^{(j)}|$, où $\iota_{[x_{\min}, x_{\max}]^N}$ est la fonction indicatrice du pavé $[x_{\min}, x_{\max}]^N$, $W \in \mathbb{R}^{J \times N}$ et $(\vartheta^{(j)})_{1 \leq j \leq J} \in [0, +\infty[^J$.

Résultats. Nous présentons les résultats obtenus lorsque l'image originale \bar{x} correspond à une coupe du fantôme Zubal de dimension $N = 128 \times 128$, où $x_{\min} = 0$ et $x_{\max} = 1$, et H est une matrice de Radon modélisant $M = 16384$ projections parallèles à partir des acquisitions de 128 lignes et 128 angles. Pour tout $m \in \{1, \dots, M\}$, nous prenons $\alpha^{(m)} = 0.01$ et $\beta^{(m)} = 0.1$. En ce qui concerne W , nous utilisons une transformation en ondelettes redondantes de Daubechies d'ordre 8 sur 3 niveaux de résolution. Cette décomposition correspond à une trame à l'analyse ajustée (*tight frame*) de constante égale à 64. Les paramètres $(\vartheta^{(j)})_{1 \leq j \leq J}$ sont choisis de façon à maximiser le SNR entre \bar{x} et \hat{x} . Les figures 1(a)-(b) présentent les projections bruitées et l'image restaurée par la méthode proposée. De plus, les figures 1(c)-(d) illustrent les variations de $(G(x_k) - G(\hat{x}))_k$ et de $(\|x_k - \hat{x}\|)_k$ en fonction du temps de calcul pour les algorithmes VMFB, FB [7] et FISTA [4]. Pour VMFB, le paramètre de relaxation est fixé à $\lambda_k \equiv 1$. Les résultats de FB et de VMFB sont présentés pour deux valeurs du pas : $\gamma_k \equiv 1$ et $\gamma_k \equiv 1.9$. Nous observons que le second choix conduit à de meilleurs résultats en terme de vitesse de convergence. De plus, nous remarquons que l'introduction de la métrique variable permet d'accélérer la vitesse de convergence du critère ainsi que celle des itérées.

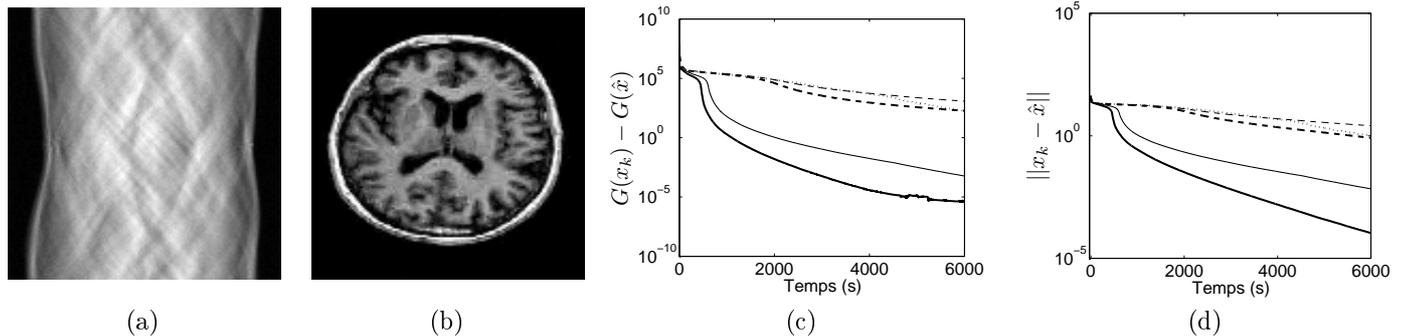


FIGURE 1 – Projections bruitées (a), image restaurée, SNR=18.9 dB (b) et comparaisons entre l'algorithme VMFB avec $\gamma_k \equiv 1$ (traits fins continus) et $\gamma_k \equiv 1.9$ (traits épais continus), l'algorithme FB avec $\gamma_k \equiv 1$ (traits fins discontinus) et $\gamma_k \equiv 1.9$ (traits épais discontinus) et FISTA (traits pointillés) en terme de convergence du critère (c) et des itérées (d).

Références

- [1] D. P. Bertsekas. *Nonlinear Programming*. 2nd edn. Athena Scientific, Belmont, MA, 1999.
- [2] M. Allain, J. Idier, et Y. Goussard. *On global and local convergence of half-quadratic algorithms*. IEEE Trans. Image Process., vol. 15, no. 5, pp. 1130-1142, Mai 2006.
- [3] H. H. Bauschke et P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.
- [4] A. Beck et M. Teboulle. *Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems*. IEEE Trans. Image Process., vol. 16, pp 2419-2434, Nov. 2009.
- [5] P. L. Combettes et B. C. Vũ. *Variable metric forward-backward splitting with applications to monotone inclusions in duality*. À paraître dans Optimization, 2013. Disponible sur <http://arxiv.org/abs/1206.6791>.
- [6] S. Becker et J. Fadili. *A quasi-Newton proximal splitting method*. Rap. Tech., 2012. Disponible sur <http://arxiv.org/abs/1206.1156>.
- [7] H. Attouch, J. Bolte et B. F. Svaiter. *Convergence of descent methods for semi-algebraic and tame problems : proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods*. Math. Program., vol. 137, pp. 91-129, Fév. 2011.
- [8] E. Chouzenoux, J.-C. Pesquet et A. Repetti. *Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function*. Rap. Tech., 2013. Disponible sur http://www.optimization-online.org/DB_HTML/2013/01/3749.html.
- [9] H. Erdogan et J. A. Fessler. *Monotonic algorithms for transmission tomography*. IEEE Trans. Med. Imag., vol. 18, no. 9, pp. 801-814, Sep. 1999.