

Optimisation bi-convexe

Cecile Della Valle

March 12, 2021

Introduction

On introduit le probleme d'optimisation biconvex-concave de déconvolution aveugle d'images avec l' a priori de noyau de convolution régulier,

$$\min_{u \in U, K \in \Sigma} \sup_v \frac{1}{2} \|g - K * u\|^2 + \frac{1}{2} \alpha \|DK\|^2 + \langle \nabla u, v \rangle - \chi_{B, \mu}(v) .$$

L'espace Σ correspond à l'ensemble des simplexes (ensemble des noyaux à coefficient positif dont la somme des coefficients vaut 1). Par extension on notera Σ la fonction indicatrice de cet ensemble. L'espace U correspond à l'hypercube $[0, 1]^{N_x \times N_y}$, ou la boule unité pour la norme L^∞ . Sa fonction indicatrice correspond donc à $\chi_{B, 1}$. La fonction à minimiser peut alors se réécrire

$$\min_{u, K} \sup_v \frac{1}{2} \|g - K * u\|^2 + \chi_{B, 1}(u) + \frac{1}{2} \alpha \|DK\|^2 + \Sigma(K) + \langle \nabla u, v \rangle - \chi_{B, \mu}(v) .$$

On note

$$L(u, v, K) = \frac{1}{2} \|g - K * u\|^2 + \frac{1}{2} \alpha \|DK\|^2 + \langle \nabla u, v \rangle - \chi_{B, \mu}(v)$$

On souhaite comparer deux stratégies pour résoudre ce problème :

- Minimisation alternée : on scinde la fonction à minimiser en une partie qui dépend de l'image u et une partie qui dépend du noyau K , les deux fonctions sont minimisées jusqu'à convergence, alternativement.
- Minimisation par pas de descente alterné (PDA) : l'algorithme alterne les pas de descente entre l'image u , le noyau K et la variable auxiliaire (ou duale) v .

Pour chacune de ces deux méthodes, on effectue les trois tests suivants :

- Initialisation par le noyau exacte, bruit std = 0.05, $(\alpha, \mu) = (1, 5)$ (cas 1);
- Initialisation par un noyau approché, pas de bruit, $(\alpha, \mu) = (1, 5)$ (cas 2) ;
- Initialisation par un noyau approché, bruit std = 0.05, $(\alpha, \mu) = (1, 5)$ (cas 3) ;

On fixe les paramètres (α, μ) pour toutes les expériences, même s'ils ne sont pas optimaux en terme de reconstruction. Ce qui nous intéresse ici c'est seulement la convergence des algorithmes.

Remarque: J'ai retiré la projection sur Simplex finalement car les algorithmes ne marchaient pas bien. En fait le résultat de la projection donnait des noyaux irréguliers. A la place j'ai projeté sur un hypercube avec une valeur max et une valeur min, cela avait l'avantage d'être simple à implémenter, d'avoir un bon comportement et de rester cohérent avec le problème initial (?).

1 Minimisation alternée

On introduit les deux sous problèmes suivants :

$$E_1(K) = \frac{1}{2} \|g - K * u\|^2 + \frac{1}{2} \alpha \|DK\|^2 + \Sigma(K) \tag{1}$$

$$E_2(u) = \frac{1}{2} \|g - K * u\|^2 + \mu \text{TV}(u) \tag{2}$$

Dans la minimisation alternée, on considère d'abord la minimisation de E_2 pour reconstruire l'image u . La minimisation est obtenue par un algorithme de Chambolle-Pock. Puis on minimise E_1 pour reconstruire le noyau K avec un forward-backward. On alterne 5 fois entre chaque minimisation, chacune a 200 pas d'itérations, donc en fait on est pas complètement à convergence..

2 Pas de descente alterné (PDA)

Pour résoudre ce problème d'optimisation on se propose de tester l'algorithme suivant,

$$K_{k+1} = \text{prox}_{\Sigma} (K_k - \gamma \partial_K L(u_k, K_k, v_k)) \quad (3)$$

$$u_{k+1} = \text{proj}_{B_1} (u_k - \tau \partial_u L(u_k, K_{k+1}, v_k)) \quad (4)$$

$$v_{k+1} = \text{prox}_{B_{\mu}} (v_k + \sigma \partial_v L(\bar{u}_k, K_{k+1}, v_k)) \quad (5)$$

$$\bar{u}_{k+1} = u_{k+1} + \theta(u_{k+1} - u_k) \quad (\text{relaxation}) \quad (6)$$

On écrit alors les dérivées partielles de L ,

$$\partial_K L(u_k, K_k, v_k) = -(g - K_k * u_k) * \tilde{u}_k + \alpha D^T D K_k \quad (7)$$

$$\partial_u L(u_k, K_{k+1}, v_k) = \text{div} \cdot v_k - \tilde{K}_k * (g - K_k * u_k) \quad (8)$$

$$\partial_v L(\bar{u}_k, K_{k+1}, v_k) = \nabla \bar{u}_k \quad (9)$$

Pour les opérateurs ∇ et div j'ai repris les opérateurs discrets du document 120415-deconvolution que tu avais écrit.

On trace les énergies primales et duales au cours de la minimisation

$$Ep_K = \frac{1}{2} \|g - K * u\|^2 + \frac{1}{2} \alpha \|DK\|^2 \quad (10)$$

$$Ep_x = \frac{1}{2} \|g - K * u\|^2 + \mu \cdot \text{TV}(u) \quad (11)$$

$$Ed = -\langle \nabla u, v \rangle \quad (12)$$

L'algorithme est très sensible aux pas de descente choisis. Soit ℓ_1 le coefficient de Lipschitz de E_1 , le pas de descente pour le noyau K est choisi par

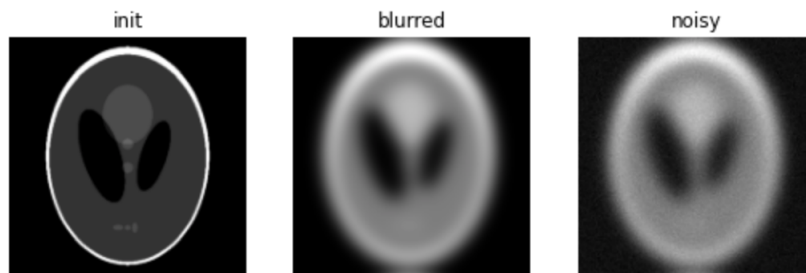
$$\tau = c \cdot \frac{2}{\ell_1} \leq c \cdot \frac{1}{\|u\|^2 + \alpha \|D\|^2} \quad 0 < c \leq 1.$$

En effet on ne veut pas dépasser la valeur limite du pas de descente pour assurer la convergence de la descente de gradient, mais on ne veut pas nécessairement effectuer un pas optimal non plus, par exemple si le noyau est initialisé de façon exacte. On choisit donc la stratégie suivante, à chaque pas de temps le coefficient c est multiplié par 0.98. La valeur initiale est ajustée à la main. Cela permet notamment de choisir un c supérieur à 1. En effet j'ai choisi cette stratégie quand le noyau est très peu modifié par rapport à l'image, cela améliore nettement la convergence. Pour le pas de descente de l'image et de la variable auxiliaire cette fois, on respecte la condition de convergence de Chambolle-Pock

$$\sigma_u \sigma_v < \frac{1}{L^2} \leq \left(\frac{1}{\sqrt{8} + \|k\|^2} \right)^2.$$

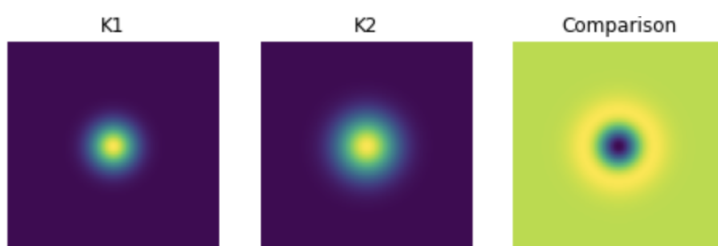
J'ai testé plusieurs stratégies, mais mettre le pas "optimal" marchait mieux jusqu'à présent.

3 Figures



Erreur blurred $|x_{\text{blurr}} - x_{\text{true}}|_2 : 0.6436$
 Erreur $|x_{\text{noisy}} - x_{\text{true}}|_2 : 0.6448$

Figure 1: Image utilisée pour la reconstruction, image convoluée et image bruitée convoluée.



Erreur $|K1 - K2| / |K2| : 0.3500$

Figure 2: Le noyau exact $K1$ et un noyau proche $K2$ pour l'initialisation de l'algorithme (de même centre mais un peu plus large). L'erreur initiale sur le noyau s'élève donc à environ 28%.

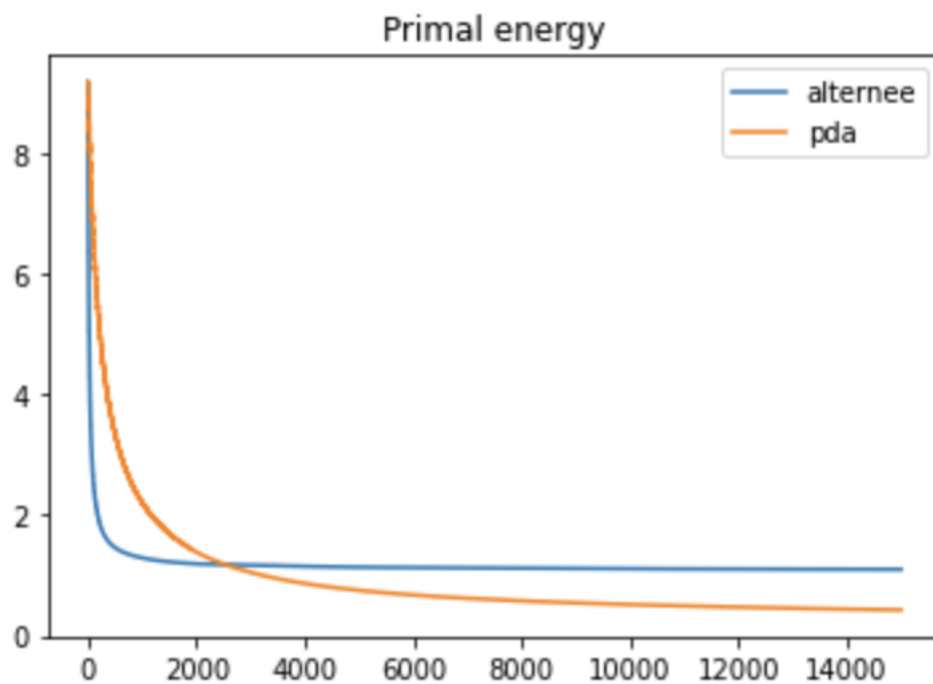


Figure 3: (cas1) initialisation du noyau K exact, bruit $\text{std} = 0.05$.

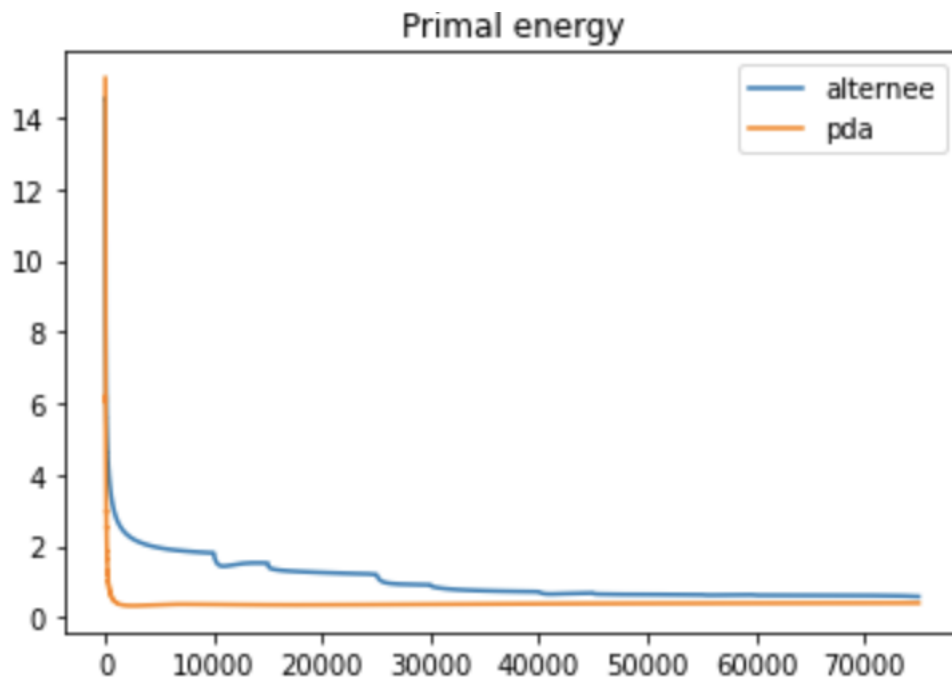


Figure 4: (cas2) initialisation du noyau approché, pas de bruit

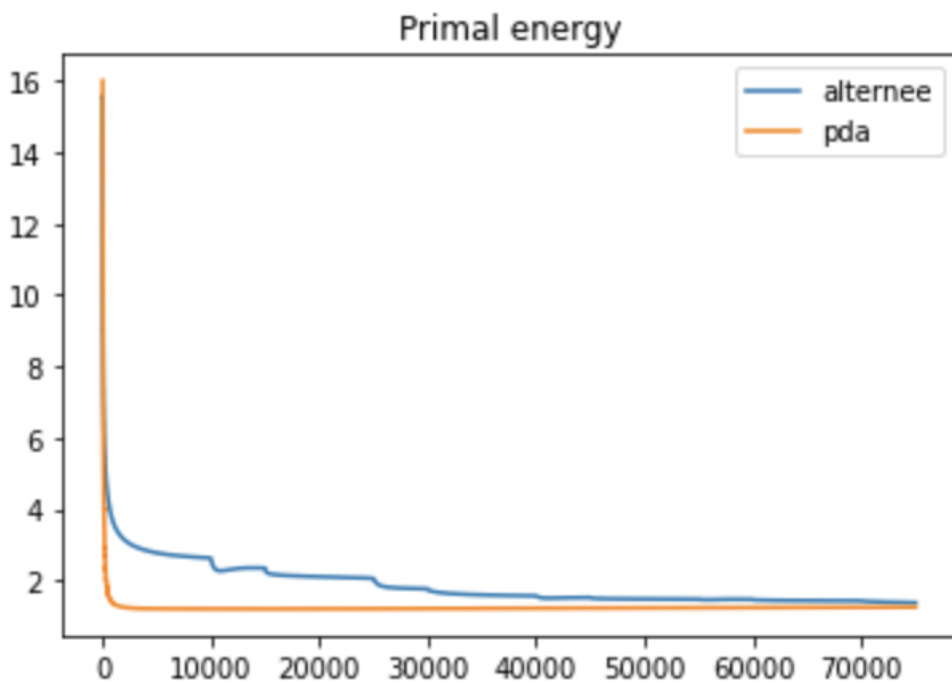
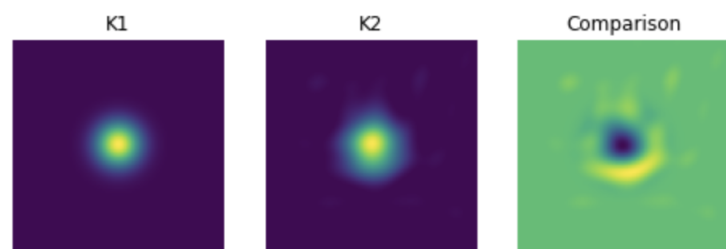
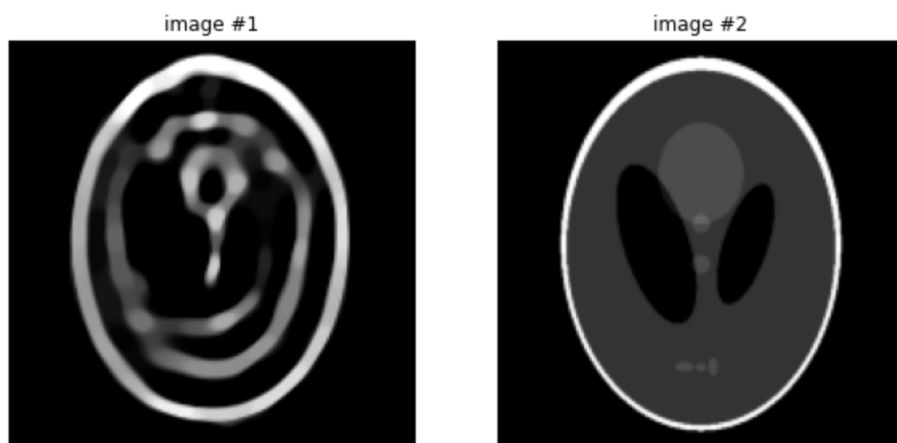


Figure 5: (cas3) initialisation du noyau approché, bruit std =0.05

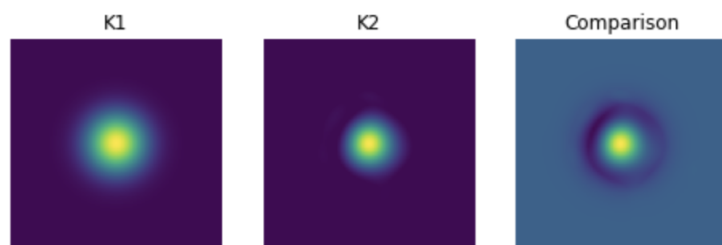


Erreur $|K1 - K2| / |K2| : 0.2267$

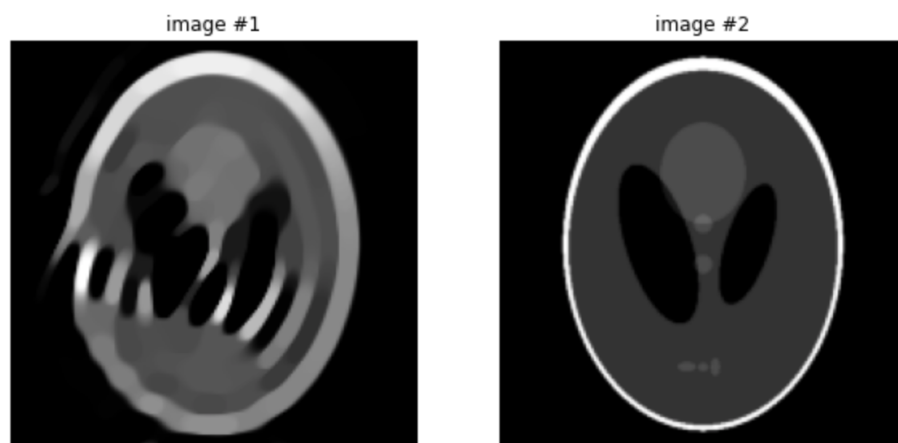


Erreur $|im1 - im2| / |im2| : 0.8603$

Figure 6: Reconstruction pour la minimisation alternée pour le cas 2.

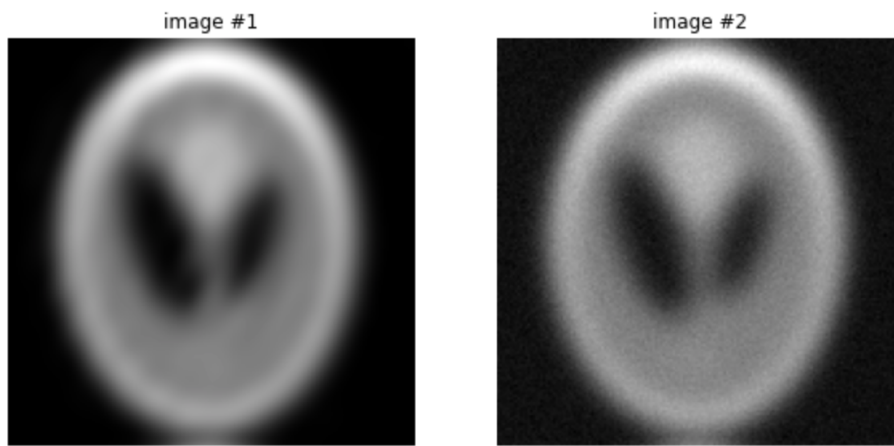


Erreur $|K1 - K2| / |K2| : 0.4890$



Erreur $|im1 - im2| / |im2| : 0.6048$

Figure 7: Reconstruction pour la minimisation par pas de descente alterné pour le cas 2.



Erreur $|im1 - im2|/|im2|$: 0.0518

Figure 8: Comparaison des images convoluées pour le cas 3 par pas de descente alternée. A gauche la convolution entre l'image et le noyau issus de la minimisation a droite l'image bruitée donnée en entrée de l'algorithme. La différence (ou erreur de reconstruction) est comparable au bruit.