

Lab1 Algorithme du gradient proximal

El Mahdi Chayti

February 25, 2019

Objectif :

Résoudre le problème suivant

$$\min_{x \in \mathbf{R}^d} \frac{1}{2n} \|Ax - b\|_2 + \lambda \|x\|_1$$

Soit

$$g(x) = s \|x\|_1$$

Le prox de f est l'opérateur de "soft thresholding" donné par la formule suivante, pour $x = (x_1, \dots, x_n)$:

$$Prox_g(x)_i = \text{signe}(x_i) \max(|x_i| - s, 0)$$

pour tout $i \in \{1, \dots, n\}$.

Échantillon test : On génère aléatoirement une matrice $A \in \mathbf{R}^{n \times d}$ et un vecteur $\epsilon \in \mathbf{R}^n$ (le bruit) On choisit une cible (un vecteur x^* de \mathbf{R}^d) et on pose $b = Ax^* + \epsilon$. (c'est la fonction `simu_linreg` qui le fait dans mon code)

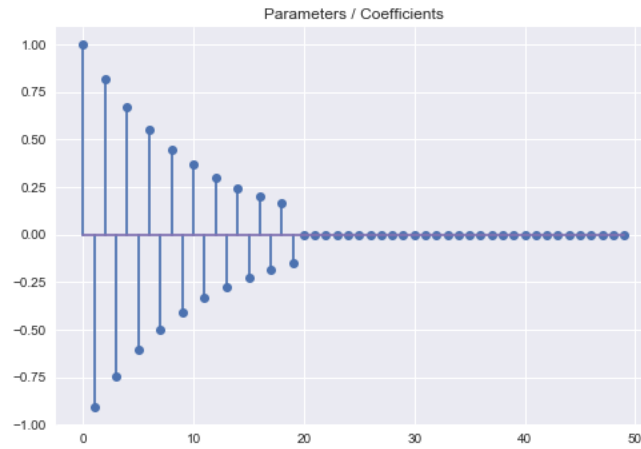


Figure 1: première cible

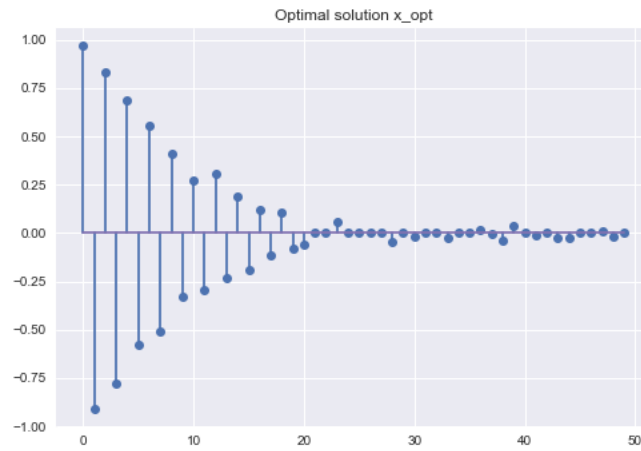


Figure 2: Solution pour la première cible

La figure 1 montre la première cible choisie.

On applique l'algorithme du gradient proximal sur l'échantillon généré pour ainsi avoir x_{opt} et on compare les deux.

On refait la même expérience sur un autre cible, et voici le résultat :

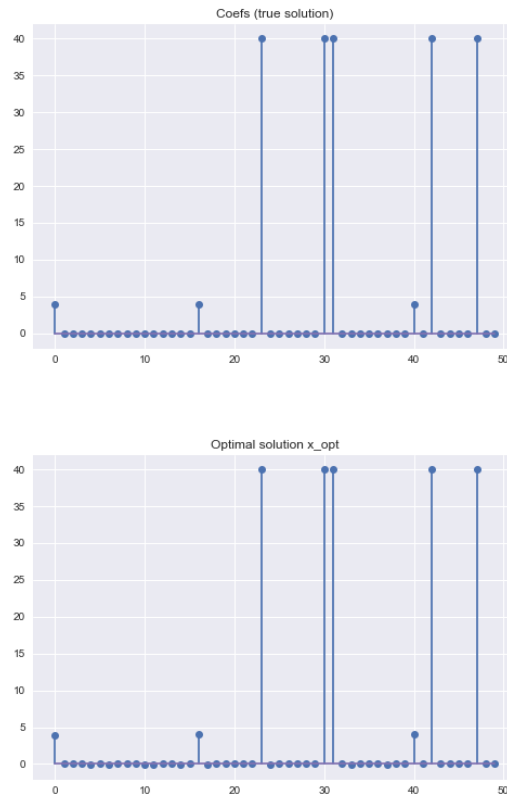


Figure 3: cible(en haut) et solution du PGD (en bas)

Comparaison de la vitesse de convergence de PGD (ou ISTA) et PGD accéléré (ou FISTA) :

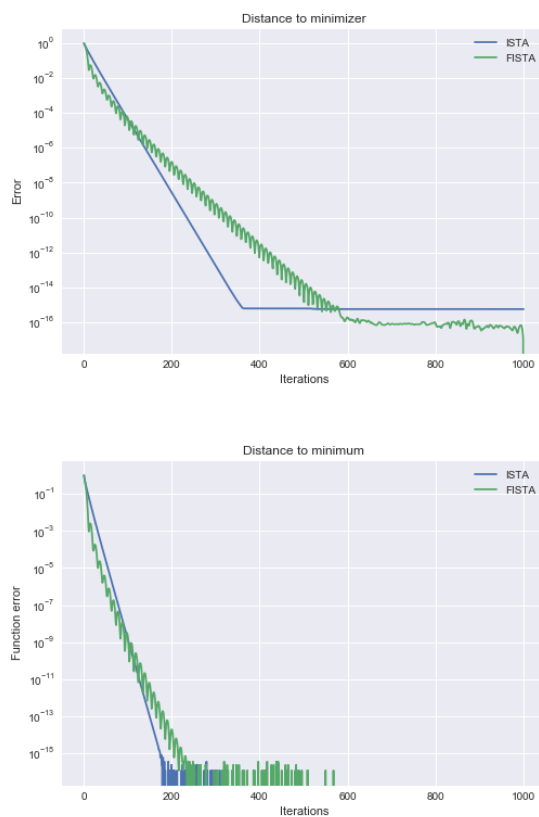


Figure 4: Vitesses de ISTA et FISTA

On remarque qu'au début FISTA est rapide par rapport à ISTA comme son l'indique, mais qu'après quelques itérations ceci s'inverse (bizarrement).

Effet de λ :

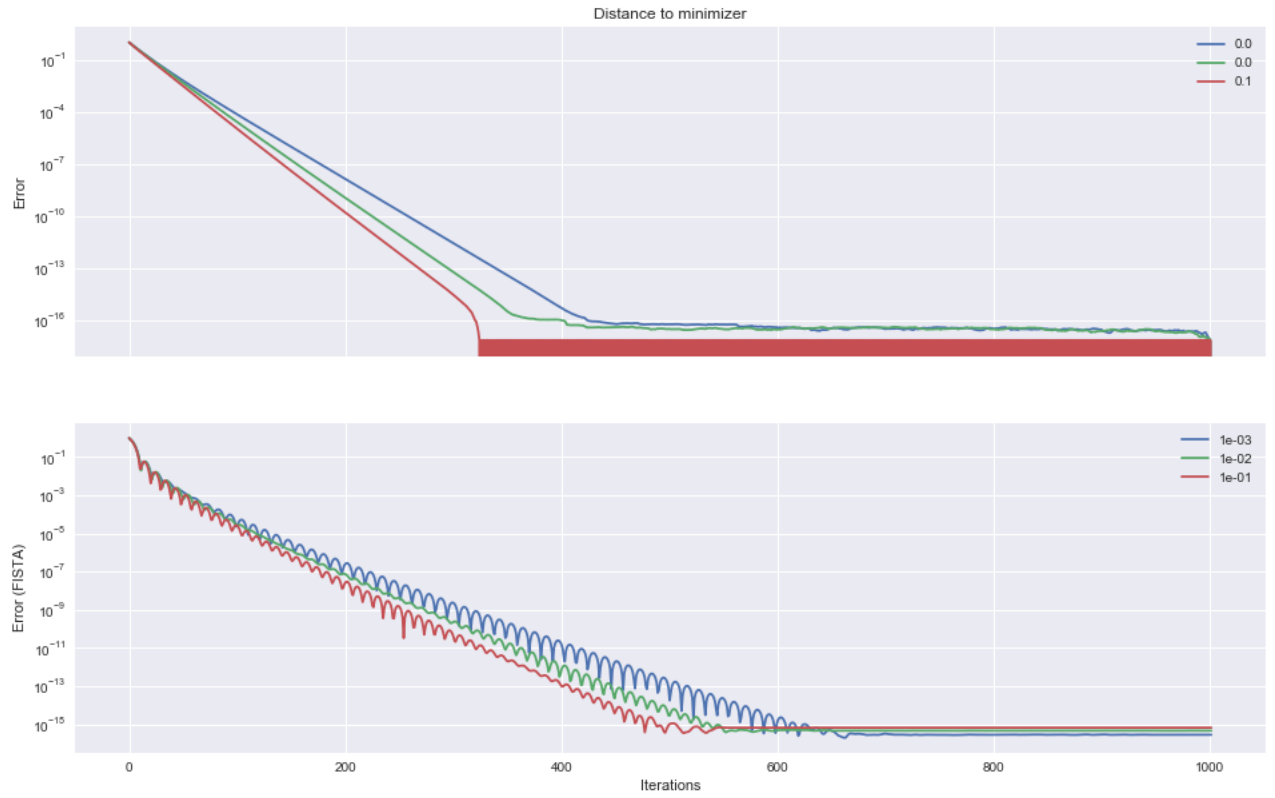


Figure 5: effet de λ sur ISTA et FISTA

L'algorithme converge d'autant plus rapidement que la constante de régularisation est grande, la raison est qu'elle est plus facile de minimiser le terme de lasso que de minimiser le terme quadratique.