République Islamique de Mauritanie
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université de Nouakchott
Faculté des Sciences et Techniques

Filière :
Statistiques et Sciences des Données

Projet SSD

Sous le thème :

Régression parcimonieuse avec ISTA : résolution du problème Lasso"(précis et académique)

Réalisé par :
Saad Bouh Aboubakar Hamar(C19871)

Encadrant :
Dr. EL BENANY MED MAHMOUD

Année Universitaire : 2025–2026

Table des matières

1 Introduction

Le Lasso (**Least Absolute Shrinkage and Selection Operator**) est une technique de régression linéaire régularisée qui favorise la parcimonie des coefficients. Dans ce TP, nous implémentons l'algorithme ISTA (*Iterative Soft-Thresholding Algorithm*) pour résoudre le problème de Lasso sur des données synthétiques. L'objectif est de minimiser la fonction coût :

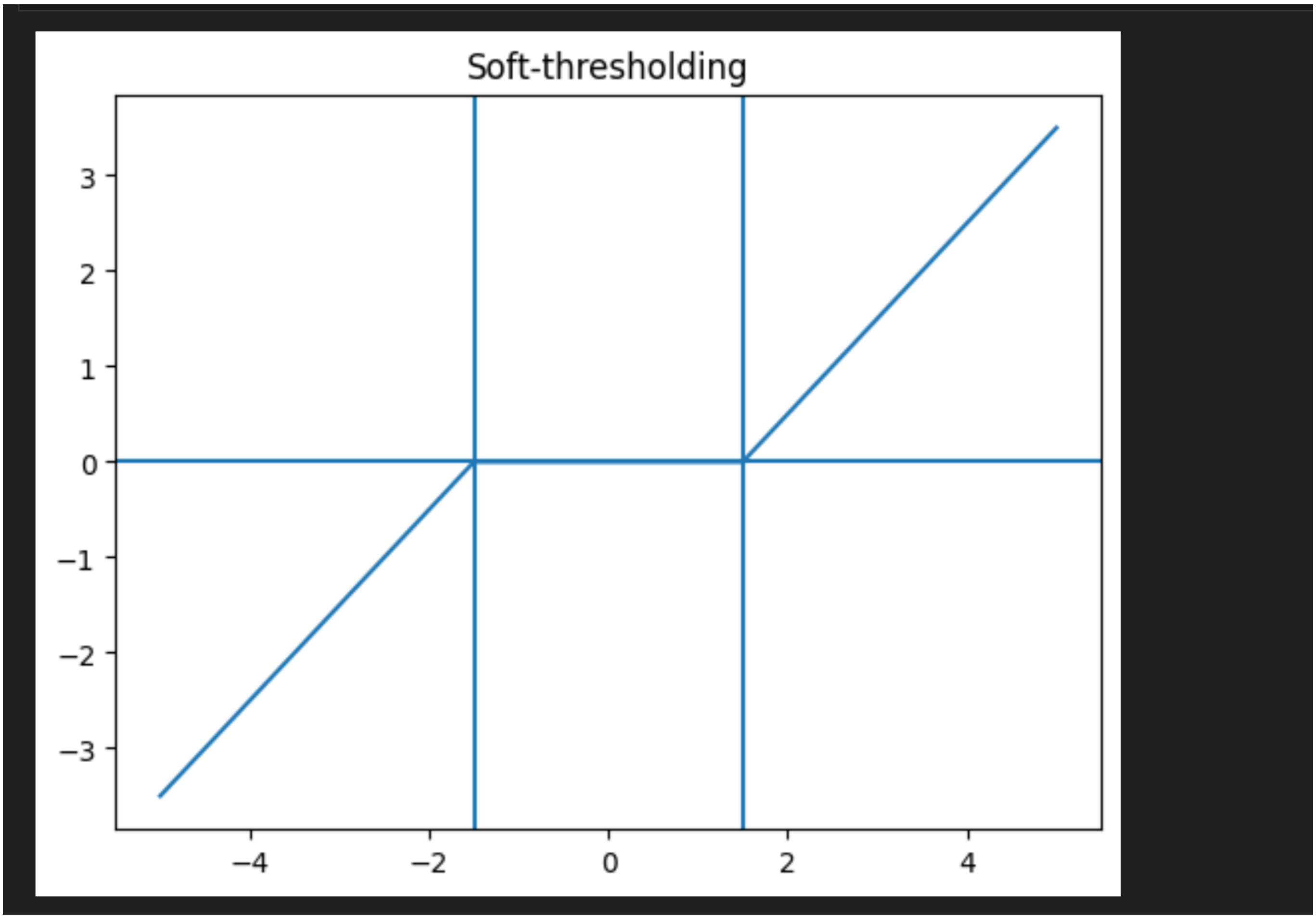

FIGURE 1 – Visualisation de l'opérateur de soft-thresholding

$$\min_w f(w) + \lambda \|w\|_1, \quad f(w) = \frac{1}{2n}\|Xw - y\|_2^2$$

où λ est le paramètre de régularisation favorisant la parcimonie.

2 Soft-thresholding

L'opérateur de **soft-thresholding** est le *proximal operator* de la norme ℓ_1 :

$$\text{SoftThreshold}(v, \gamma) = \text{sign}(v) \cdot \max(|v| - \gamma, 0)$$

Il permet de réduire les petites valeurs à zéro, en gardant les coefficients importants. La figure suivante illustre cet opérateur :

3 Données synthétiques

Pour tester ISTA, nous générons des données synthétiques :

- $n = 100$ observations et $d = 50$ variables.
- Vrai vecteur de coefficients w_{true} avec $k = 5$ coefficients non nuls.
- Bruit gaussien ajouté sur la sortie : $y = Xw_{\text{true}} + \epsilon$, $\epsilon \sim \mathcal{N}(0, 0.01)$

Ces données permettent de vérifier que l'algorithme ISTA retrouve correctement la parcimonie du vecteur w.

4 Méthodologie : ISTA

4.1 Gradient et constante de Lipschitz

La fonction de perte quadratique est :

$$f(w) = \frac{1}{2n}\|Xw - y\|_2^2$$

Le gradient est donné par :

$$\nabla f(w) = \frac{1}{n}X^\top(Xw - y)$$

La constante de Lipschitz L pour la descente de gradient est :

$$L = \frac{1}{n}\lambda_{\max}(X^\top X)$$

4.2 Algorithme ISTA

À chaque itération :

$$w^{(k+1)} = \text{SoftThreshold}\Big(w^{(k)} - \frac{1}{L}\nabla f(w^{(k)}), \frac{\lambda}{L}\Big)$$

La convergence est observée en suivant la valeur de la fonction objectif :

$$F(w) = f(w) + \lambda\|w\|_1$$

5 Implémentation

L'implémentation en Python comporte les étapes suivantes :

1. Génération de données synthétiques X, y et w_{true}
2. Calcul du gradient $\nabla f(w)$
3. Détermination de la constante de Lipschitz L
4. Boucle ISTA pour mettre à jour w
5. Calcul et suivi de la fonction objectif pour chaque itération
6. Visualisation de la convergence et de la parcimonie des coefficients

6 Résultats

6.1 Convergence

On observe que la fonction objectif décroît rapidement et converge vers une valeur stable après quelques dizaines d'itérations.

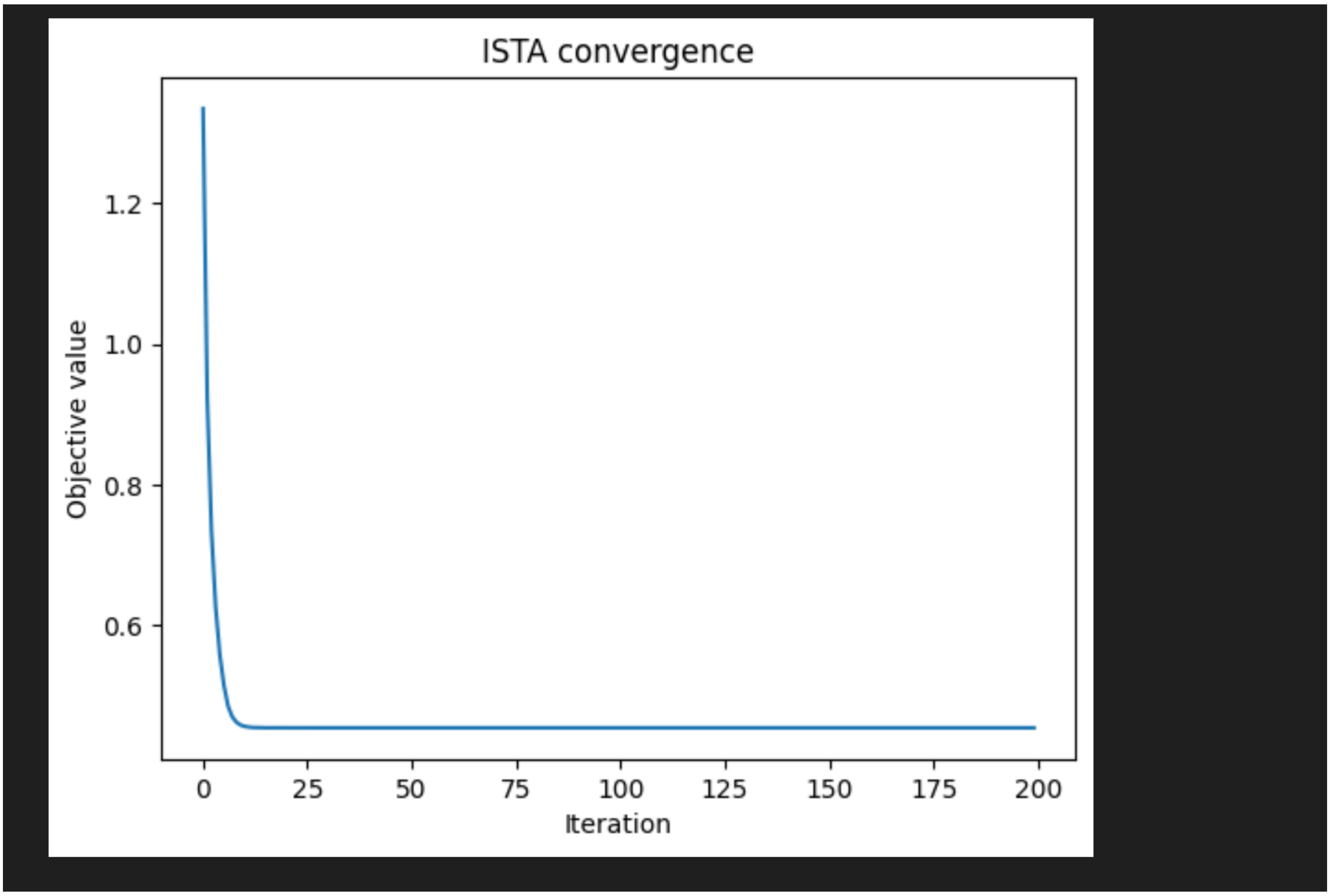

FIGURE 2 – Convergence de la fonction objectif au cours des itérations

6.2 Parcimonie

Le vecteur estimé w_{est} présente la propriété de parcimonie :

— Nombre de coefficients non nuls : $\sum_i \mathbf{1}_{w_i \neq 0}$

— Indices des coefficients non nuls : $\{i \mid w_i \neq 0\}$

Ces résultats confirment que ISTA est capable de retrouver les coefficients réellement actifs.

7 Analyse

— L'algorithme ISTA converge rapidement pour des données de petite dimension.

— Le paramètre λ contrôle la sparsité : une valeur trop élevée annule trop de coefficients, une valeur trop faible laisse apparaître du bruit.

— Le soft-thresholding est un outil simple et efficace pour imposer la parcimonie.

— Cette approche peut être généralisée à de grandes dimensions et à d'autres problèmes linéaires régularisés.

8 Conclusion

Nous avons implémenté ISTA pour résoudre le problème de Lasso sur des données synthétiques. L'algorithme montre une convergence rapide et est capable de récupérer la structure parcimonieuse du vecteur de coefficients. Cette expérience illustre l'intérêt des méthodes itératives avec régularisation ℓ_1 pour la régression et la sélection de variables.

9 Références

1. Beck, A., & Teboulle, M. (2009). A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. SIAM Journal on Imaging Sciences, 2(1), 183–202.
2. Hastie, T., Tibshirani, R., & Wainwright, M. (2015). Statistical Learning with Sparsity. CRC Press.