

TP 7 : La solution du lasso avec l'algorithme du gradient proximal

Stéphane Canu

14 octobre 2024, MLA, INSA Rouen

Le but du TP est d'étudier la solution au Lasso¹ fournie par l'algorithme du gradient proximal, dans le cadre de la régression sur des données simulées. Pour le faire fonctionner, vous êtes supposé avoir déjà installé CVX (que vous pourrez télécharger à cette adresse : <http://cvxr.com/cvx/>)

Le code suivant est disponible en ligne avec google colab :

<https://colab.research.google.com/drive/1Q1QXbh58oEME1gloZycjhz2FA77HTX5K?usp=sharing>

Ex. 1 — Le Lasso et le gradient proximal

Reprendre les données du TP précédent sur le lasso.

1. Coder l'algorithme du gradient proximal pour résoudre le Lasso

Tant que on n'a pas convergé

$$g = -X^t(y - X\beta)$$

$$\beta = \text{prox}_{L1}(\beta - tg, \lambda)$$

- a) Commencer par écrire une fonction `Prox_L1` permettant de calculer le `proxL1`

$$\text{prox}_{l1}(w, \lambda) = \underset{u}{\text{argmin}} \quad \lambda \|u\|_1 + \frac{1}{2} \|u - w\|^2$$

$$\text{prox}_{l1}(w, \lambda) = \text{sign}(w) \max(|w| - \lambda, 0)$$

```
def proxl1(w, lam):  
    v = # A COMPLETER ...  
    return(v)
```

- b) Donner le prompt que vous avez utilisé pour obtenir la solution et la réponse obtenue. Expliquez en quoi cette solution est ou n'est pas intéressante.

2. Conclusions

- a) Comparez toutes les méthodes vues en cours pour calculer la solution du problème du lasso en terme de temps de calcul et de performances.
- b) Proposez et justifiez des recommandations sur l'algorithme à utiliser pour résoudre un problème de lasso.

¹<http://statweb.stanford.edu/~tibs/lasso.html>