

TP 6 : Le chemin de régularisation du lasso (Homotopy et LARS)

Stéphane Canu

7 octobre 2024, MLA, INSA Rouen

Le but du TP est d'étudier le chemin de régularisation du Lasso¹, 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/>)

Ex. 1 — La mise œuvre du chemin de régularisation du lasso

1. Reprendre les données du TP précédent sur le lasso et calculer une solution du problème pour $\lambda = 10^{-3}n$.
2. Écrire les condition d'optimalité et vérifiez que cette solution est bien optimale
 - a) recherchez l'ensemble I_0 de toutes les composantes nulles de la solution
 - b) vérifiez que tous les gradient associés aux composantes non nulles de la solution sont bien égales à 0.
 - c) vérifiez que 0 appartient bien au sous gradient I_0 de toutes les composantes nulles de la solution
3. Faites un pas dans le chemin
 - a) Calculez analytiquement, à partir de la solution précédente la solution du lasso pour $\lambda = 10e - 3n + \varepsilon$, pour $\varepsilon = 0,01$.
 - b) vérifiez cette solution est correcte
4. Faites un autre pas dans le chemin
 - a) Calculez la plus petite valeur de λ telle que la solution du lasso soit le vecteur nul.
 - b) Calculez la plus petite valeur de λ telle que la solution du lasso soit le vecteur nul sauf une seule composante.
5. Comparez avec les méthodes de calcul du chemin de régularisation de sklearn. Vous préciserez les différences entre les méthodes `lars_path_gram` et `lasso_path`

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