

TP 6 : Le lasso composante par composantes (component wise)

Stéphane Canu

6 octobre 2025, MLA, INSA Rouen

Le but du TP est d'étudier différentes solutions au Lasso¹, dans le cadre de la régression sur des données simulées.

Le code suivant est disponible en ligne avec google colab :
<https://colab.research.google.com/drive/1Q1QXbh58oEME1gloZycjhz2FA77HTX5K?usp=sharing>

Ex. 1 — La mise œuvre du lasso composante par composantes (component wise)

1. Reprendre les données du TP 3 sur le lasso et vérifier que

$$\sum_{i=1}^n X_{ij}^2 = 1, \quad j = 1, \dots, p$$

2. Réécrire le problème d'optimisation du lasso comme une fonction d'une seule variable². Vous donnerez les prompts et le LLM que vous avez utilisé.
3. Ecrire en python la fonction `lasso_cost(X,y,beta,lamb)`. Vérifier que le cout du Lasso diminue après une seule itération de l'approche composante par composante : on commence en prenant $\beta = 0$, on calcule le cout de lasso de 0, puis on résout le problème d'optimisation du lasso pour une seule variable et on vérifie que le cout a bien diminué.
4. Coder l'algorithme de type Gauss Seidel - composant par composant (*component wise*) pour résoudre le Lasso
Pour $j = 1, p$
 $\beta^{(-j)} = (\beta_1, \dots, \beta_{j-1}, 0, \beta_{j+1}, \dots, \beta_p)$
 $r = y - X\beta^{(-j)}$
 $\hat{\beta}_j = \arg \min_{\beta \in \mathbb{R}} J^{(j)}(\beta, X_j, r, \lambda)$
fin de pour
où $J^{(j)}(\beta, X_j, r, \lambda)$ est la fonction de perte du Lasso de la seule variable β_j
5. Le Lasso pour sklearn³
 - a) Quels sont les différentes fonction proposée par sklearn pour résoudre le problème du Lasso ? Quelle est l'algorithme mis en œuvre ?
 - b) résoudre le problème à l'aide de la fonction Lasso de sklearn
 - c) résoudre le problème à l'aide de la fonction LassoCV de sklearn
6. Comparez toutes les solutions en terme de temps de calcul et de performances. Proposez et justifiez des recommandations sur la méthode à utiliser pour résoudre un problème de lasso.

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

²<https://projecteuclid.org/journals/annals-of-applied-statistics/volume-1/issue-2/Pathwise-coordinate-optimization/10.1214/07-AOAS131.full>

³scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso.html