TP 7 : Des variantes du Lasso

Stéphane Canu

16 novembre 2026, MLA, INSA Rouen

Le but du TP est d'étudier la solution à différentes variantes du Lasso. Pour le faire fonctionner, vous êtes supposé avoir déjà installé CVX (que vous pourrez télécharger à cette adresse : `http://cvxr.com/cvx/`)

Ex. 1 — des variantes du Lasso

1. Reprendre les données du TP précédent sur le lasso.

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

url_data = "https://web.stanford.edu/~hastie/ElemStatLearn/datasets/prostate.data"
df = pd.read_csv(url_data, delimiter='\t')
variables = df.columns[1:9]
df[variables] = df[variables].apply(lambda x: (x - x.mean()) / x.std())

# Get the training and test sets
Y_train = df.loc[df["train"]=="T", 'lpsa'].to_numpy()
X_train = df.loc[df["train"]=="T", variables].to_numpy()
print("Training set : n = {} samples and p = {} dimensions".format(X_train.shape[0],
    X_train.shape[1]))

Y_test = df.loc[df["train"]=="F", 'lpsa'].to_numpy()
X_test = df.loc[df["train"]=="F", variables].to_numpy()
print("Test set : n = {} samples and p = {} dimensions".format(X_test.shape[0], X_test.
    shape[1]))
nt = X_test.shape[0]
X_test_1 = np.hstack((np.ones((nt,1)),X_test))
```

2. Recherchez la solution du lasso telle qu'elle à été publiée dans le livre (comme dans le TP1)
3. Recherchez la solution de l'elastic net

$$J_{\text{el}}(\beta) \;=\; \tfrac{1}{2}\|X\beta - y\|^2 + \lambda\|\beta\|_1 + \tfrac{\gamma}{2}\|\beta\|_2^2 \,,$$

   a) comme un QP (avec CVX) et $\lambda = 6$ et $\gamma = 0.005$. Vérifiez que vous obtenez de mailleurs résultats que ceux du livre
   b) comme un problème de lasso
4. Recherchez la solution du lasso adaptatif

$$J_a(\beta) \;=\; \tfrac{1}{2}\|X\beta - y\|^2 + \lambda \sum_{j=1}^{p} w_j|\beta_j| \,,$$

   a) comme un QP (avec CVX) et $\lambda = 3$. Vérifiez que vous obtenez de mailleurs résultats que ceux du livre
   b) comme un problème de lasso
5. Donner les prompts que vous avez utilisé pour obtenir les solutions et la réponse obtenue de votre IA préféré. Expliquez en quoi cette solution est ou n'est pas intéressante.
6. The fused lasso en 1 d
   a) Générez des données 1d bruitées

```
x = np.arange(-1, 1, .1)
n = len(x)
stdnoise = 0.3
y = np.sign(x) + np.random.randn(n)*stdnoise
```

b) Ecrire un programme CVX qui résoud le fused lasso 1d suivant

$$J_f(\beta) \;=\; \tfrac{1}{2}\|\beta - y\|^2 + \mu \sum_{j=1}^{n-1} |\beta_j - \beta_{j+1}| \;,$$

pour $\mu = \frac{3}{2}$.

c) Visualisez les données (comme la figure du cours)