

2020-02-07_sklapi

November 26, 2021

1 Régression Ridge, Lasso et nouvel estimateur

Ce notebook présente la régression Ridge, Lasso, et l'API de scikit-learn. Il explique plus en détail pourquoi la régression Lasso contraint les coefficients de la régression à s'annuler.

```
[1]: from jupyterhelper import add_notebook_menu
      add_notebook_menu()
```

```
[1]: <IPython.core.display.HTML object>
```

```
[2]: %matplotlib inline
```

1.1 Un jeu de données pour un problème de régression

```
[3]: from sklearn.datasets import load_boston
      data = load_boston()
      X, y = data.data, data.target
```

```
[4]: from sklearn.model_selection import train_test_split
      X_train, X_test, y_train, y_test = train_test_split(X, y)
```

On apprend la première régression linéaire.

```
[5]: from sklearn.linear_model import LinearRegression
      lin = LinearRegression()
      lin.fit(X_train, y_train)
      lin.coef_
```

```
[5]: array([-1.13205698e-01,  4.07774544e-02, -3.45089298e-03,  3.20617502e+00,
          -1.36532265e+01,  4.50004571e+00, -8.49372877e-03, -1.30815349e+00,
           2.91962099e-01, -1.28261282e-02, -8.74096357e-01,  9.76382670e-03,
          -4.48963490e-01])
```

```
[6]: from sklearn.metrics import r2_score
      r2_score(y_test, lin.predict(X_test))
```

```
[6]: 0.6595254960413792
```