

scikit-learn ゼミ

1.1.3. Lasso

1.1.3.1. Setting regularization parameter

1.1.3.1.1. Using cross-validation

1.1.3.1.2. Information-criteria based model selection

國井慎也

lasso regression

- 過学習を抑えるために罰則項に係数の和を含めた目的関数

$$\min_w \frac{1}{2n_{sample}} \|Xw - y\|^2 + \alpha \|w\|$$

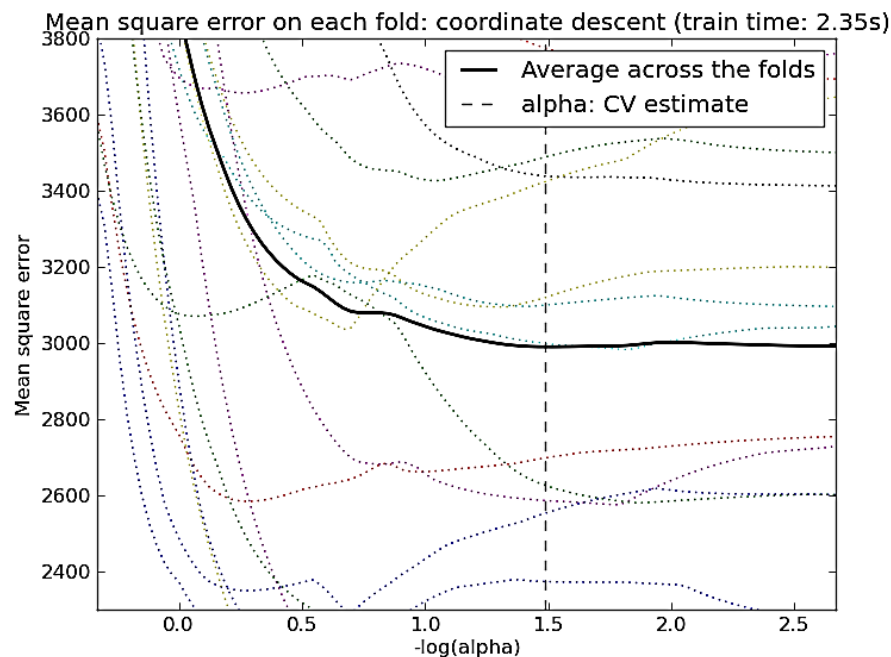
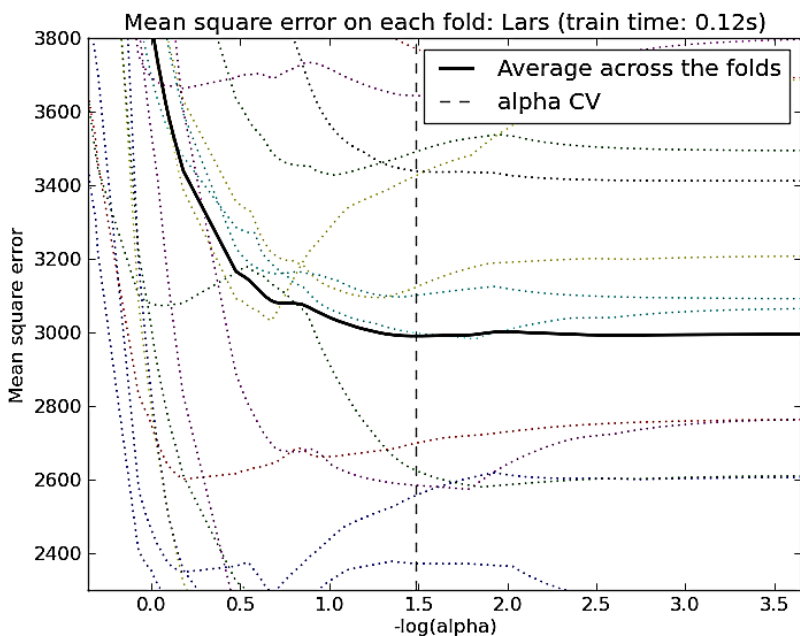
- Ridge Regression(リッジ回帰)
 - $\min_w \|Xw - y\|^2 + \alpha \|w\|^2$

実行例

```
>>> clf = linear_model.Lasso(alpha = 0.1)
>>> clf.fit([[0, 0], [1, 1]], [0, 1])
Lasso(alpha=0.1, copy_X=True, fit_intercept=True, max_iter=1000,
      normalize=False, positive=False, precompute='auto', tol=0.0001,
      warm_start=False)
>>> clf.predict([[1, 1]])
array([ 0.8])
```

罰則項の度合い α

- 罰則項の α
 - 求めた係数のスパースの度合いをコントロール
 - 適当に設定してきた
 - 交差検定を利用して α を求められる
 - LassoCVとLassoLarCV



情報化基準によるモデル選択

- LassoLarsIC

- AICまたはBICの値が少ないときの値を α として採用

- AIC

- The Akaike information criterion

- $AIC = -2\ln L + 2p$

- L : 尤度、 p : パラメータ数

- BIC

- The bayes information criterion

- $BIC = -2\ln L + p \log n$

- n : サンプル数

