NATIONAL RESEARCH UNIVERSITY

Лекция 10.12.21, 2 модуль

Ridge (гребневые) и LASSO оценки коэффициентов регрессии

Демидова О.А.
E-mail: demidova@hse.ru

План лекции

- Ridge (гребневые) оценки коэффициентов регрессии
- LASSO оценки коэффициентов регрессии

Методы борьбы с мультиколлинеарностью

- Переспецификация модели (функциональные преобразования переменных)
- Исключение одной или нескольких объясняющих переменных
- Метод главных компонент
- Использование ridge (гребневых), LASSO и т.п. оценок параметров

Ridge и LASSO оценки регрессии

$RSS = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \ldots - \hat{\beta}_k X_{ki})^2,$

$\mathrm{OLS} : \mathrm{RSS} \mapsto \min_{\hat{\beta}},$

$\mathrm{Ridge} : \mathrm{RSS} + \lambda \sum_{j=1}^{k} \hat{\beta}_j^2 \mapsto \min_{\hat{\beta}}$

$\mathrm{LASSO} : \mathrm{RSS} + \lambda \sum_{j=1}^{k} |\hat{\beta}_j| \mapsto \min_{\hat{\beta}},$

λ – параметр регуляризации

Ridge (гребневые) оценки регрессии

$Ridge\colon\ RSS + \lambda \sum_{j=1}^{k} \hat{\beta}_j^2 \rightarrow min_{\hat{\beta}}$

$\lambda \rightarrow 0, \quad \hat{\beta}_R \rightarrow \hat{\beta}_{OLS}$

$\lambda \rightarrow \infty, \quad \hat{\beta}_R \rightarrow 0$

$(Y - X\hat{\beta})' \left(Y - X\hat{\beta}\right) + \lambda\beta'\beta \rightarrow \ min_{\hat{\beta}}$

$\hat{\beta}_{Ridge} = (X'X + \lambda I)^{-1} X'Y$

$V(\hat{\beta}_{Ridge}) = \sigma_\varepsilon^2 (X'X + \lambda I)^{-1} X'X (X'X + \lambda I)^{-1}$

$V(\hat{\beta}_{OLS}) = \sigma_\varepsilon^2 (X'X)^{-1}$

LASSO оценки регрессии

$LASSO - \ least \ absolut \ shrinkage \ and \ selection \ operator$

$$LASSO{:}\ \ RSS + \lambda \sum_{j=1}^{k} \left|\hat{\beta}_j\right| \rightarrow \ min_{\hat{\beta}}$$

$$\lambda \rightarrow 0, \quad \hat{\beta}_{LASSO} \rightarrow \hat{\beta}_{OLS},$$

$$\lambda \rightarrow \infty, \quad \hat{\beta}_{LASSO} \rightarrow 0,$$

Переменные обычно предварительно центрируют и нормируют, константа после этого не включается в уравнение регрессии.

Выбор параметра регуляризации

CV (cross – validation)

Все наблюдения делятся на Q частей (foldes)

Для каждого q = 1,…,Q

1. Оцениваем модель по всем подмножествам, кроме q-го.

2. С помощью результатов оценивания заполняем пропуски в зависимой переменной.

3. Значения параметра λ берутся с некоторым шагом: $\lambda_1, \ldots, \lambda_N$.

4. Для каждого значения λ_j вычисляется ошибка прогноза:

$$RSS^q(\lambda_j) = \sum_{i=1}^{n}(y_i - \hat{\beta}_1^q(\lambda_j)\, x_{1i} - \ldots - \hat{\beta}_k^q(\lambda_j) x_{ki})^2$$

Выбор параметра регуляризации

CV (cross – validation)

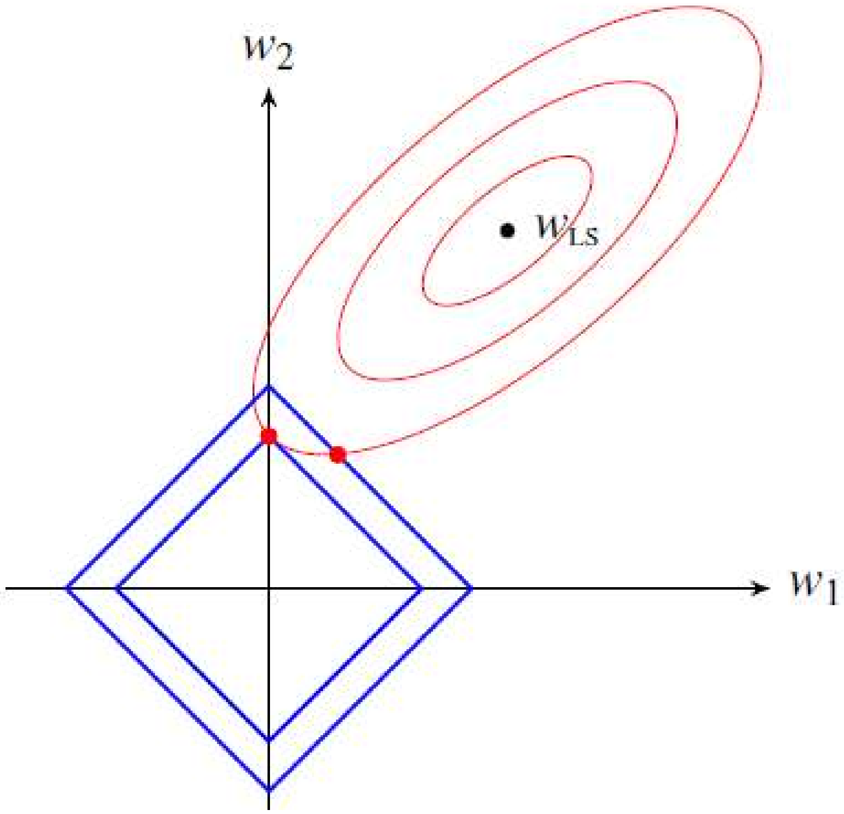
5. Вычисляется средняя ошибка прогноза по q блокам:

$$MSE\ (\lambda_j) = \frac{1}{Q}\sum_{q=1}^{Q} RSS^q(\lambda_j)$$

6. Выбирается λ_j, при котором $MSE\ (\lambda_j)$ является минимальным.

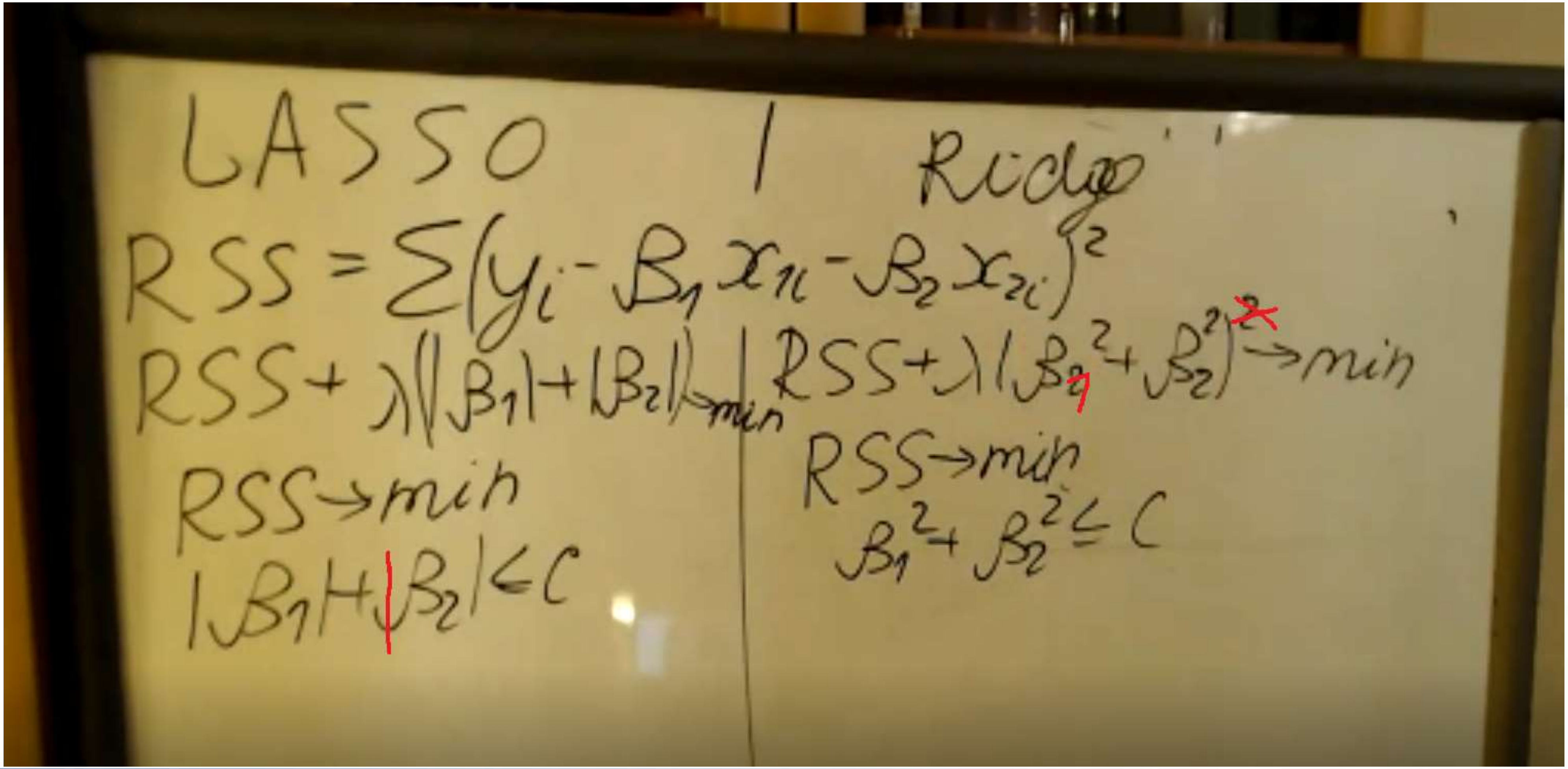
Особенности LASSO и Ridge оценок

Ridge Regression vs LASSO



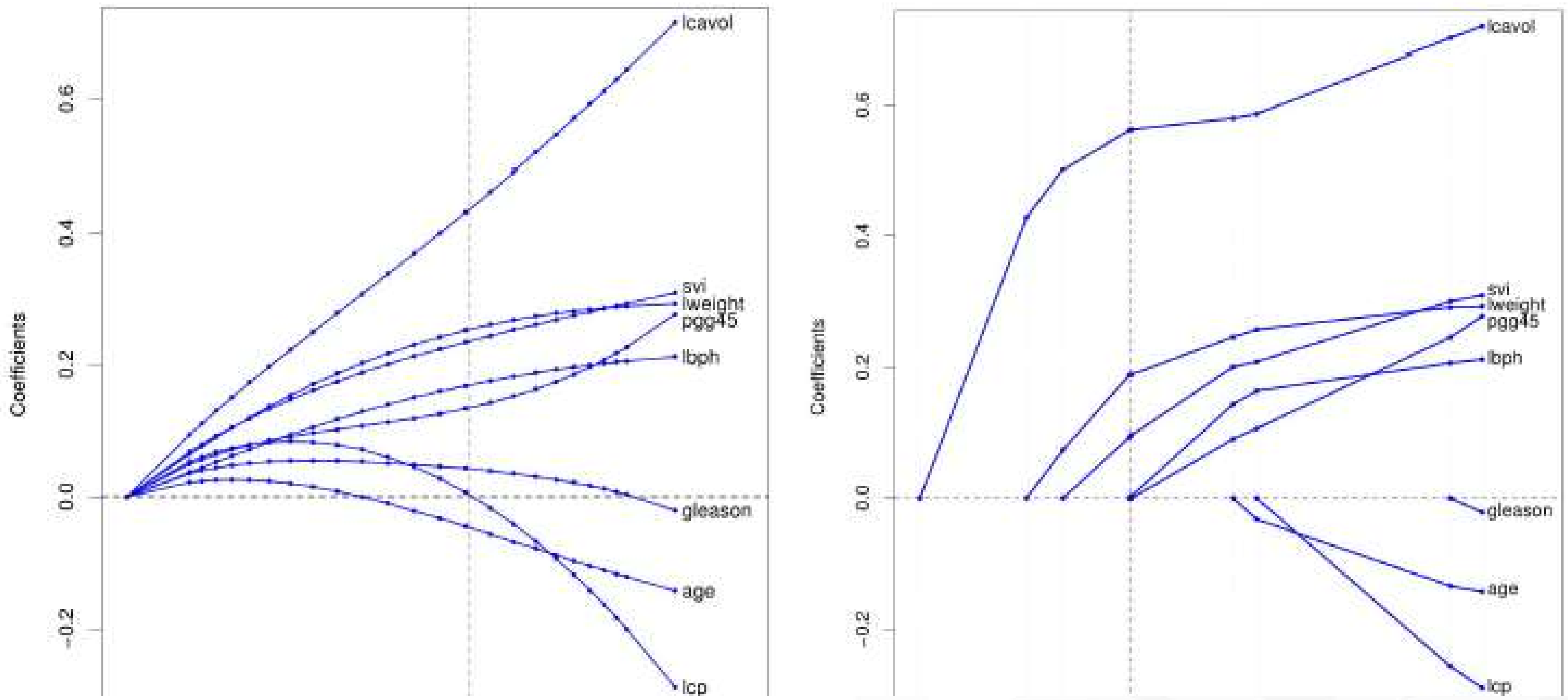
Исправленные опечатки в лекции

Дорогие студенты, в конце лекции в мои записи на доске вкрались опечатки, приношу глубочайшие извинения и исправляю.



Особенности LASSO и Ridge оценок

COEFFICIENT PROFILES: RR VS LASSO



Спасибо за внимание!

Демидова О.А.
E-mail: demidova@hse.ru