

Лекция по эконометрике № 6, 2 модуль

**Ridge (гребневые) и LASSO оценки
коэффициентов регрессии.**

**Демидова
Ольга Анатольевна
https://www.hse.ru/staff/demidova_olga
E-mail: demidova@hse.ru
30.11.2020**

План лекции № 6, 2 модуль

- 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, LASSO и т.п. оценки регрессии

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

$$OLS : RSS \rightarrow \min_{\hat{\beta}}$$

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

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

Ridge (гребневые) оценки регрессии

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

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

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

Ridge (гребневые) оценки регрессии

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

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

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

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

LASSO оценки регрессии

*LASSO – least absolut shrinkage
and selection operator*

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

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

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

Выбор λ

- CV (cross – validation)
- Adaptive LASSO

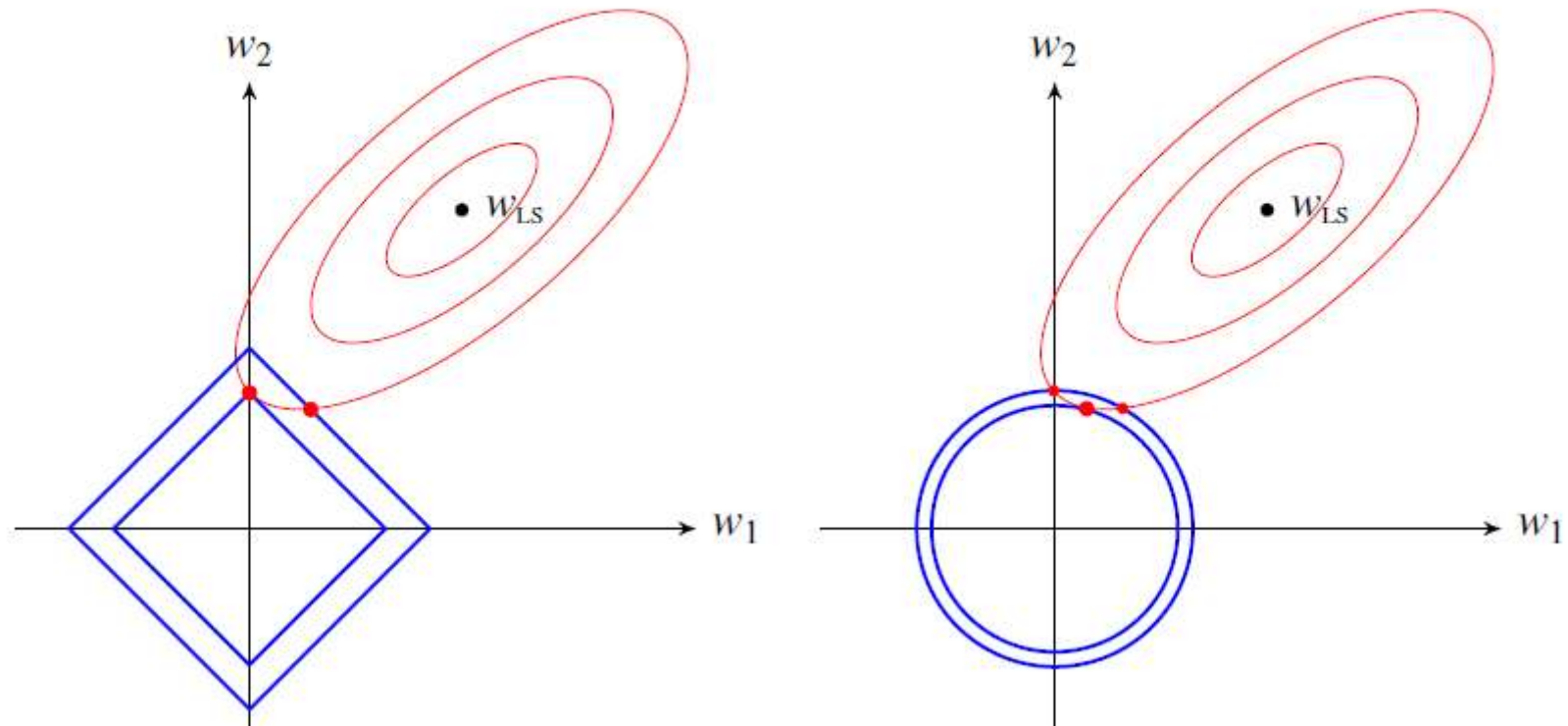
Cross-validation

- Все наблюдения делятся на K частей (folds)
- Для каждого $k = 1 \dots K$
- 1. Оцениваем модель по всем подмножествам, кроме k -го.
- 2. С помощью результатов оценивания заполняем пропуски в зависимой переменной.

Cross-validation

- Находим MSE (mean squared error).
- Усредняем
- Выбираем λ , где усредненная MSE наименьшая.

RIDGE REGRESSION VS LASSO



COEFFICIENT PROFILES: RR vs LASSO

