

Todo list

Проверить! Нет ли у β_1 особого положения?	1
на картинке три c : очень большое — дающее мнк решение, меньше — ненулевые β , маленькое — одна из β равна 0	1

1 Ridge/Lasso regression

LASSO — Least Absolute Shrinkage and Selection Operator. Метод построения регрессии, предложенный Robert Tibshirani в 1995 году.

Вспомним обычный МНК:

$$\min_{\beta} (y - X\beta)'(y - X\beta) \quad (1)$$

LASSO вместо исходной задачи решает задачу условного экстремума:

$$\min_{\beta} (y - X\beta)'(y - X\beta) \quad (2)$$

при ограничении $\sum_{j=1}^k |\beta_j| \leq c$.

Проверить! Нет ли у β_1 особого положения?

Естественно, при больших значениях c результат LASSO совпадает с МНК. Что происходит при малых c ?

Для наглядности рассмотрим задачу с двумя коэффициентами β : β_1 и β_2 . Линии уровня целевой функции — эллипсы. Допустимое множество имеет форму ромба с центром в начала координат.

на картинке три c : очень большое — дающее мнк решение, меньше — ненулевые β , маленькое — одна из β равна 0

То есть при малых c LASSO обратит ровно в ноль некоторые коэффициенты β .

Применим метод множителей Лагранжа для случая, когда ограничение $\sum_{j=1}^k |\beta_j| \leq c$ активно, то есть выполнено как равенство.

$$L(\beta, \lambda) = (y - X\beta)'(y - X\beta) + \lambda \left(\sum_{j=1}^k |\beta_j| - c \right) \quad (3)$$

Необходимым условием первого порядка является $\partial L / \partial \beta = 0$. Таким образом мы получили альтернативную формулировку метода LASSO:

$$\min_{\beta} (y - X\beta)'(y - X\beta) + \lambda \sum_{j=1}^k |\beta_j| \quad (4)$$

LASSO пытается минимизировать взвешенную сумму $RSS = (y - X\beta)'(y - X\beta)$ и «размера» коэффициентов $\sum_{j=1}^k |\beta_j|$.

Мы не будем вдаваться в численные алгоритмы, которые используются при решении этой задачи.

Ridge regression отличается от LASSO ограничением $\sum \beta_j^2 \leq c$. Также как и LASSO Ridge regression допускает альтернативную формулировку:

$$\min_{\beta} (y - X\beta)'(y - X\beta) + \lambda \sum_{j=1}^k \beta_j^2 \quad (5)$$

Также как и LASSO Ridge regression тоже приближает значения коэффициентов β_j к нулю. Принципиальное отличие LASSO и RR. В LASSO крайнее решение с несколькими коэффициентами равными нулю является типичной ситуацией. В RR коэффициент β_j может оказаться точно равным нулю только по чистой случайности. LASSO допускает байесовскую интерпретацию...