Part. 02

회귀분석

| 회귀계수 축소법 (Ridge)

FASTCAMPUS ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택

I 회귀계수 축소법

계수축소법의 종류

- 계수축소법은 기본적으로 다중선형회귀와 유사
- 다중선형회귀에서 잔차를 최소화했다면, 계수축소법에서는 잔차와 회귀계수를 최소화
- 계수축소법에는 크게 3 가지의 방법이 있음: Ridge 회귀, Lasso 회귀, Elastic-Net 회귀
- 아래 식은 다중선형회귀의 SSE이며, 다중선형회귀에서는 RSS가 최소화되는 회귀계수를 추정

$$minimize\ SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij}\right)^2$$

- 계수축소법에서는 위 식에 회귀계수를 축소하는 항을 추가

$$minimize\ SSE + f(\beta)$$

I 회귀계수 축소법

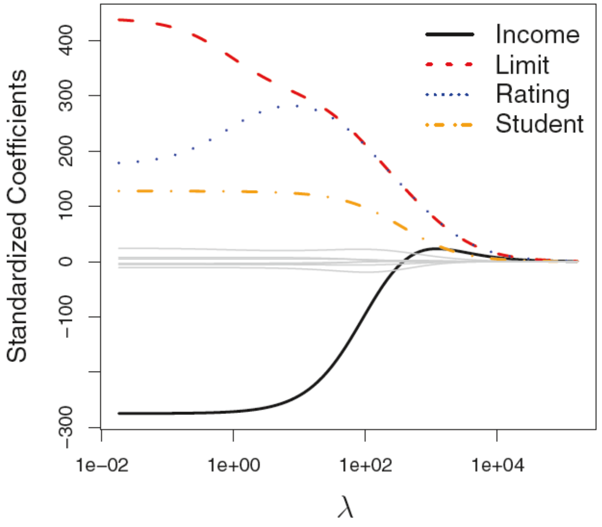
Ridge 회귀

- Ridge 회귀에서는 $f(\beta)$에 회귀계수의 제곱의 합을 대입
- λ는 tuning parameter로 크면 클 수록 보다 많은 회귀계수를 0으로 수렴

회귀계수의 제곱의 합

$$minimize \sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij}\right)^2 + \lambda\sum_{j=1}^{p}\beta_j^2$$

- 오른쪽 그림은 λ가 커질수록 입력변수인 Income, LimitRating, Student가 0으로 수렴하는 것을 표현
- 적절한 λ의 값은 데이터마다 달라지며, 현재는 e^4인 54.6의 값을 설정하였을 때 모든 입력 변수가 0으로 수렴함 (완전히 0이 되지는 않음)



James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. New York: Springer.

I 회귀계수 축소법

Ridge 회귀

다중회귀 β유도

$$\sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon'\varepsilon = (y - X\beta)'(y - X\beta)$$

$$= y'y + \beta'X'X\beta - 2\beta'X'y$$

$$\frac{\partial L}{\partial \beta} = 2X'X\beta - 2X'y = 0$$

$$\Rightarrow X'X\beta = X'y$$

$$\Rightarrow \beta = (X'X)^{-1}X'y$$

Ridge회귀 β유도

$$(y - X\beta)'(y - X\beta) + \lambda\beta'\beta$$

$$y'y + \beta'X'X\beta - 2\beta'X'y + \lambda\beta'\beta$$

$$\frac{\partial L}{\partial \beta} = 2X'X\beta - 2X'y + 2\lambda\beta = 0$$

$$\Rightarrow (X'X + \lambda I)\beta += X'y$$

$$\Rightarrow \beta = (X'X + \lambda I)^{-1}X'y$$

I 회귀계수 축소법

Ridge 회귀

Ridge회귀 β유도

$$(y - X\beta)'(y - X\beta) + \lambda\beta'\beta$$

$$y'y + \beta'X'X\beta - 2\beta'X'y + \lambda\beta'\beta$$

$$\frac{\partial L}{\partial \beta} = 2X'X\beta - 2X'y + 2\lambda\beta = 0$$

$$\Rightarrow \ (X'X + \lambda I)\beta += X'y$$

$$\Rightarrow \ \beta = (X'X + \lambda I)^{-1}X'y$$

$X'X$의 diagonal term에 λ가 추가된 형태

다중공선성은 X들간의 강한 선형관계가 있을때 발생

⬇

$X'X$의 역행렬을 구할 수가 없음

⬇

Ridge는 $X'X$의 역행렬을 구할 수 있도록
강제로 작은 값을 diagonal term에 추가한 것!

I 회귀계수 축소법

Ridge 회귀

$$X'X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$(X'X)^{-1} = \frac{1}{(ad - bc)} \begin{pmatrix} d & -b \\ -c & d \end{pmatrix}$$

$$X'X + \lambda I = \begin{pmatrix} a + \lambda & b \\ c & d + \lambda \end{pmatrix}$$

$$(X'X + \lambda I)^{-1} = \frac{1}{(\underline{(a + \lambda)(d + \lambda)} - bc)} \begin{pmatrix} d & -b \\ -c & d \end{pmatrix}$$

$ad = bc$가 되지 않게끔 하도록 함

다중공선성은 X들간의 강한 선형관계가 있을때 발생

$X'X$의 역행렬을 구할 수가 없음

Ridge는 $X'X$의 역행렬을 구할 수 있도록
강제로 작은 값을 diagonal term에 추가한 것!

Part. 02
회귀분석

| 회귀계수 축소법 (Lasso, ElasticNet)

FASTCAMPUS
ONLINE
머신러닝과 데이터분석 A-Z

강사. 이경택