

Nov 05, 2018

1 Ordinary Least Square

Ordinary Least Square의 경우 오차의 평균이 0, 오차의 분산이 등분산, 오차가 서로 uncorrelated 일 경우 선형 모형 중에서 최적의 모델이다 (BLUE) 이때 오차의 분포가 정규분포라면 MLE와 동일한 결과를 보여준다

OLS가 아닌 다른 방법을 사용하는 이유는 변수의 개수 p 보다 관측치의 개수 n 이 충분히 많지 않은 경우 정확한 예측을 하지 못하는 문제가 발생하기 때문이다. 또한 실제 β 값중에서 0이 되어야 하는 β 대해서도 0에 가까운 값으로 예측 할 뿐 실제 0으로 예측하지 못하는 문제가 발생한다.

2 Ridge regression

$$\hat{\beta}^{ridge} \tag{1}$$