

Model selection

내 모델의 문제점을 알기 위해선..

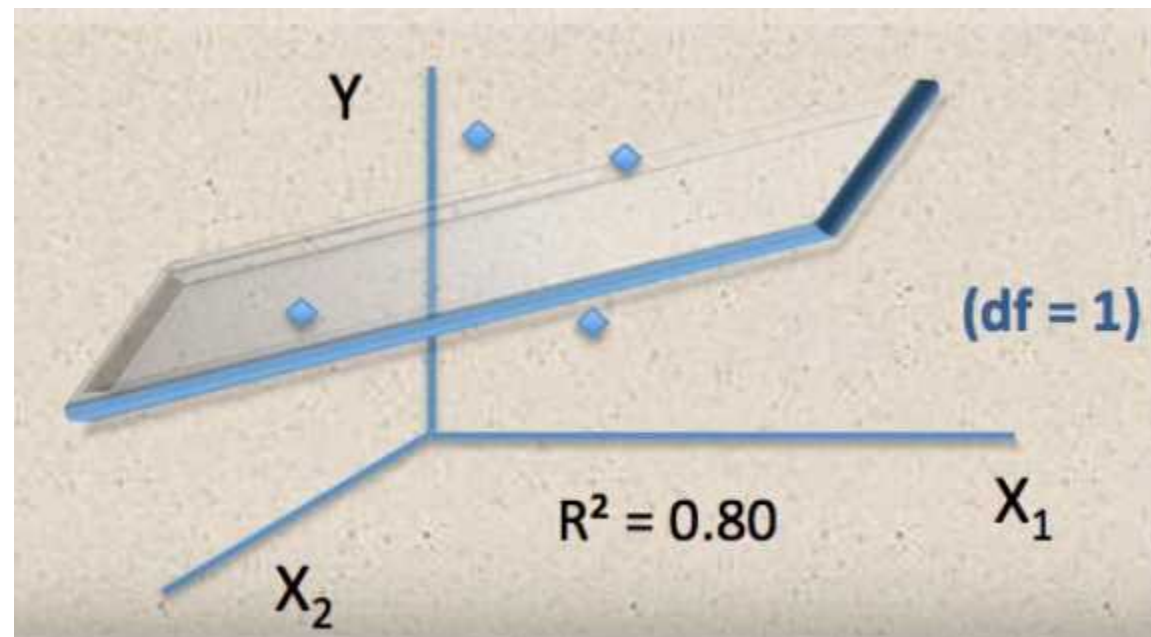
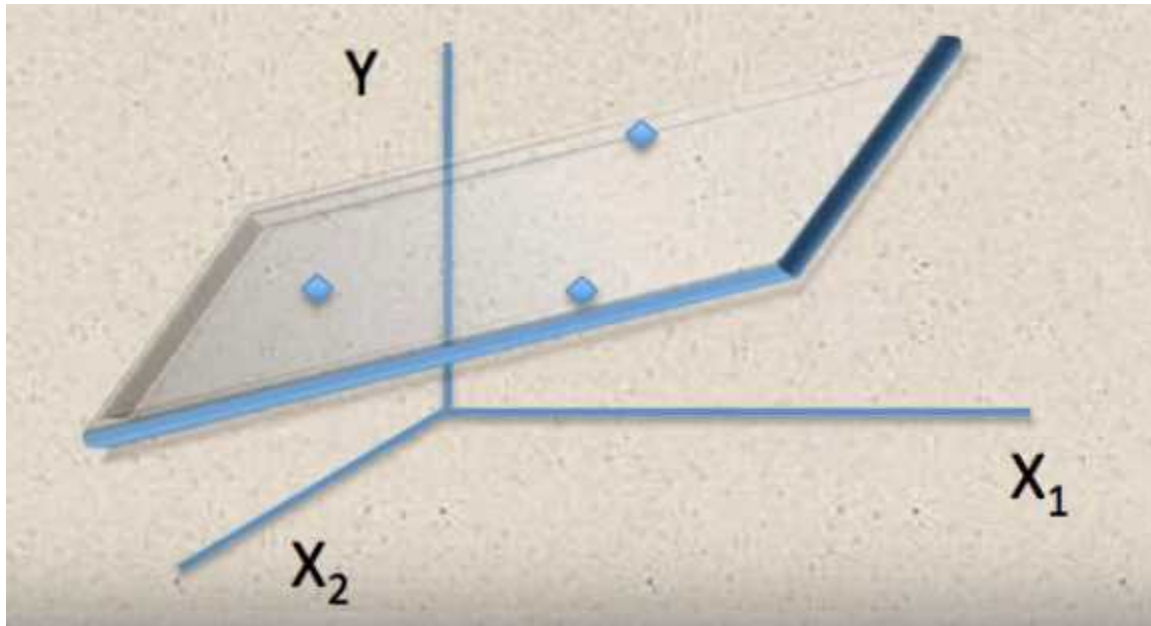
평가방법이 필요하다

그렇다면 모델을 평가하는 방법을 무엇일까??

방법

- Adjusted R square
- AIC(Akaike Information Criterion)
- BIC(Baysian Informaion Criterion)





보통 Residual의 총합 + 독립변수 개수에 따른 패널티로
만들어져 있다.

독립변수가 개수가 적은 심플한 모델을 선호한다고 한다.
하지만 심플할 수록 더 biased할 확률이 늘어나는 게 아
닌가?

모델의 biasedness는 피할수없는 현실적 문제라면
어떻게 해결해야 할까?

Correlation의 함정

$$X = [-2, -1, 0, 1, 2]$$

$$Y = [2, 1, 0, 1, 2]$$

선형적 관계가 없다는 것 뿐