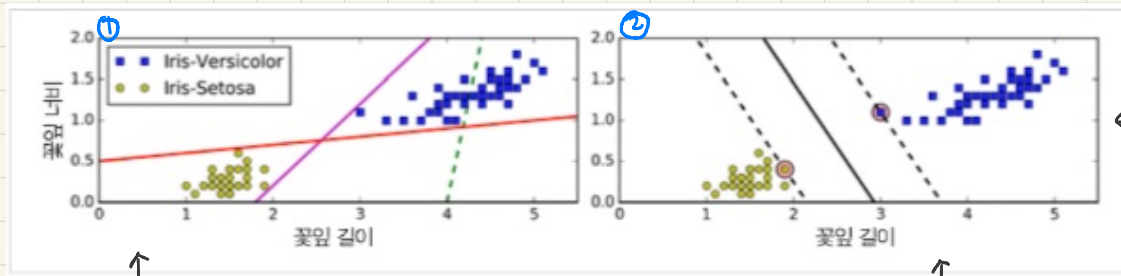


# 5 서포트 벡터 머신

: 다목적 머신러닝 모델. 선형/비선형 회귀, 이상치 탐색 사용가능. VERY STRONG

## 5.1 선형 SVM 분류



< 분류가 된 것 >

< 라지 마진 분류 >

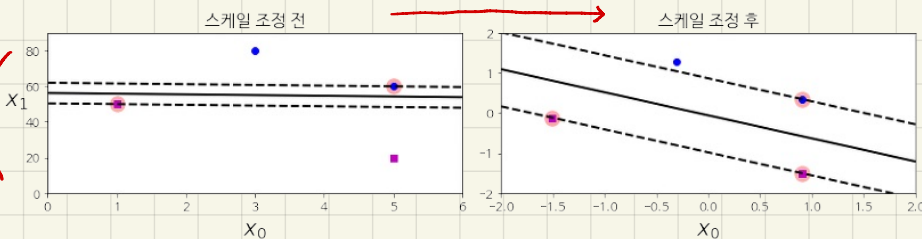
↑  
임의로 그린 것

결정 경계가 샘플에서 너무 가까워! 새로운 샘플에서는 잘 작동하지 않을 것.

↑  
가장 가까운 훈련 샘플로부터 최대한 멀리!

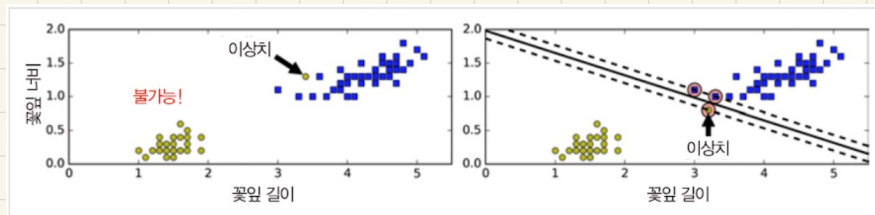
② 본 그래프에서 도중에 멀리 떨어진 곳에 샘플이 있더라도 결정 경계에는 영향 X  
도중 영역에 위치한 DATA SAMPLE 에 의해 전적으로 결정.

\* 특성에 민감한 SVM,  
이 스케일이 너무 커서  
거의 수평이 되어버림



### 5.1.1 소프트 마진 분류

[ 하드 마진 분류 ]



하드 마진 분류: 모든 샘플이 맞도록

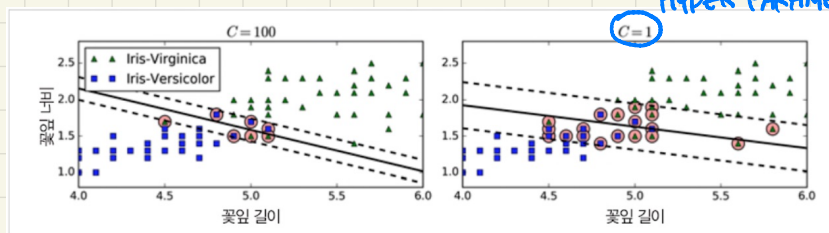
→ 이상치에 매우 민감

→ 선형분류되어야 잘 작동

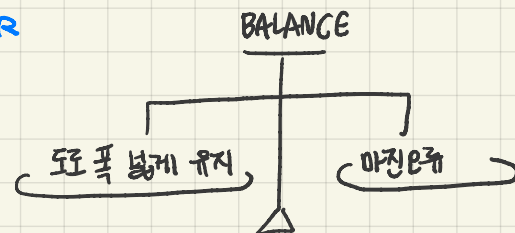
↑ 하드 마진 분류, 일반화가 잘 될 것 같지는 않네

→ 더 유연한 모델이 필요!

[ 소프트 마진 분류 ]



HYPER PARAMETER



- H.P ↑
- MARGIN ERROR ↓
- 일반화 ☺

- H.P ↓
- MARGIN ERROR ↑
- 일반화 ☹

모델이 과대 적합이라면 HP 강도시켜  
규제할 수 있다.