

SVM (Support Vector Machine)

- (선/면) 결정 경계, 마진, outlier 허용도, 커널 (decision boundary)

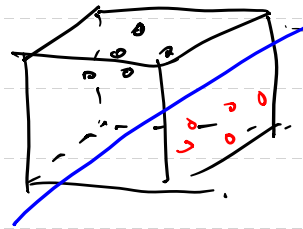
* 마진을 최대화 하는 결정 경계를 찾는 것이 SVM의 목표

* 결정 경계는 각 data 집합으로부터 등거리로 떨어져 있도록 한다.

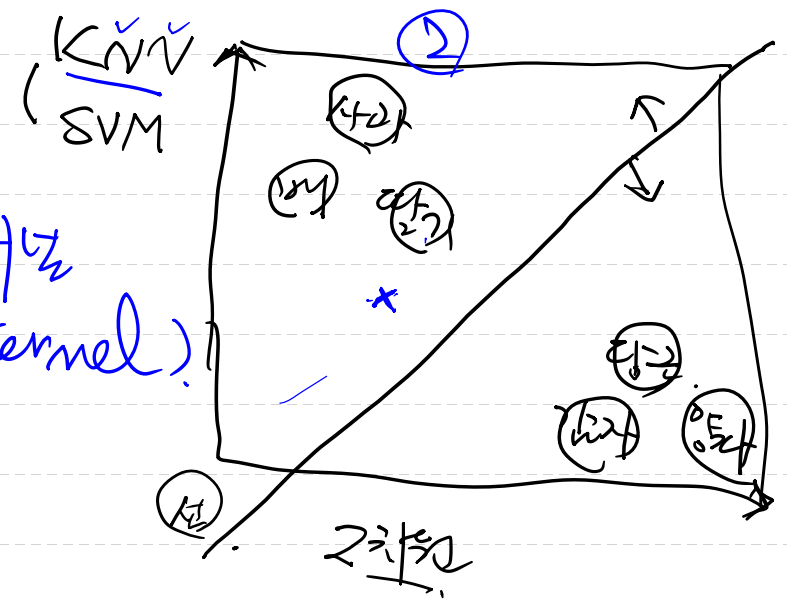
Support vector : 결정 경계에 가장 가까워 있는 데이터 점을 의미함.
(지지 벡터)

⇒ 지지 벡터들이 결정 경계를 결정함이다.

* n개의 특징을 가진 data 이면 최소 $n+1$ 개의 support 벡터가 존재해야 함.



커널 (kernel)



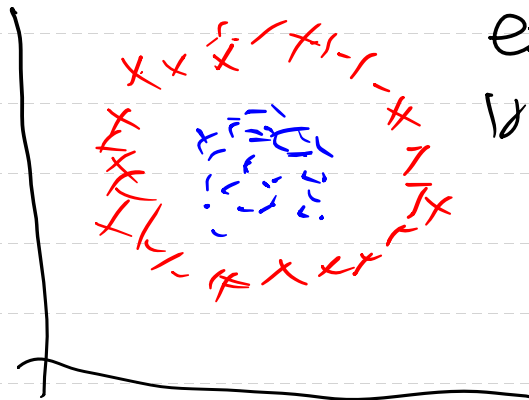
SVC 함수 유형.

① C : 기벌값 1

큰수쪽 \rightarrow 하드마진, 작은수쪽 \rightarrow 소프트마진

② 커널(kernel) : 결정 경계의 형식(type).

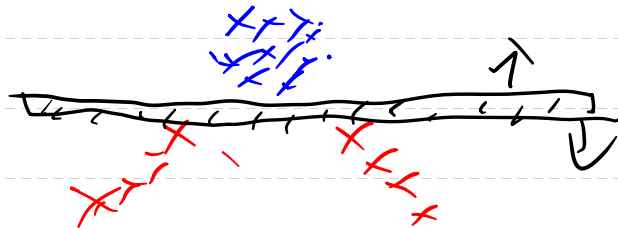
(선)

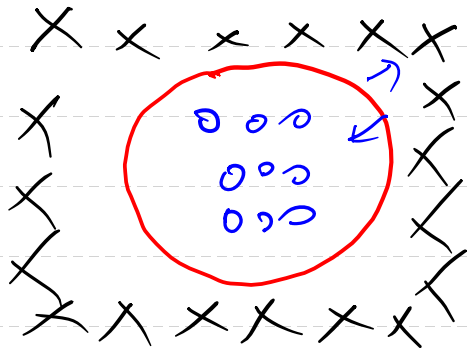


ex.

비선형 \rightarrow 분리되어 있는 data.

\Rightarrow 선분리





선형 SVM.

gamma : 결정 경계/2
유연하게

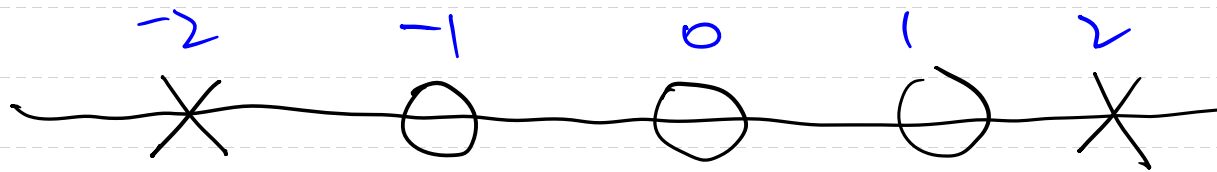
2리프를 확장한다.

rbf 커널 (가우시안 커널)

2차원 data \rightarrow R차원 space data로 변환.
(n 차원)

* 선형이지 않음:

gamma를 크게?
결정 경계/2이
많이 바뀌어짐
gamma를 작게?
결정 경계/2이
적당히 바뀌어짐



\times 2차원 벡터 공간의 기저를
 2차원 "으로 ~~늘~~린다. \Rightarrow 계층 정제를 계층

$y = x^2$ 함수를 1차원 \rightarrow 2차원.

Kernel - tan2k.

| 1차원(x) | 2차원 기저(x, y) |
|--------|--------------|
| -2 | $(-2, 4)$ |
| -1 | $(-1, 1)$ |
| 0 | $(0, 0)$ |
| 1 | $(1, 1)$ |
| 2 | $(2, 4)$ |

