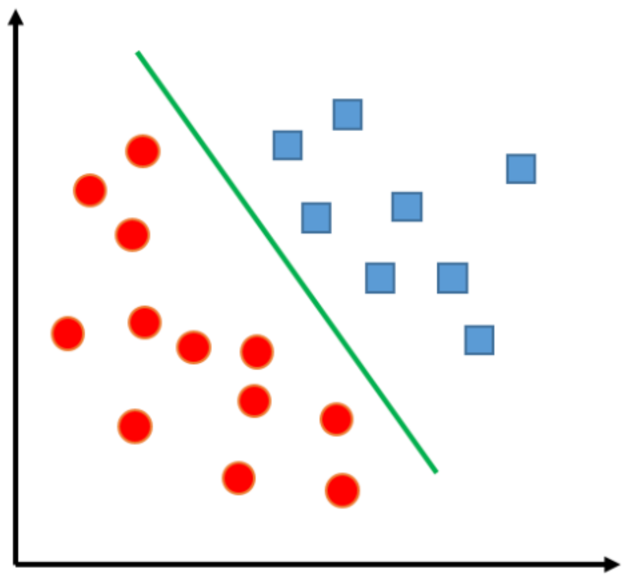
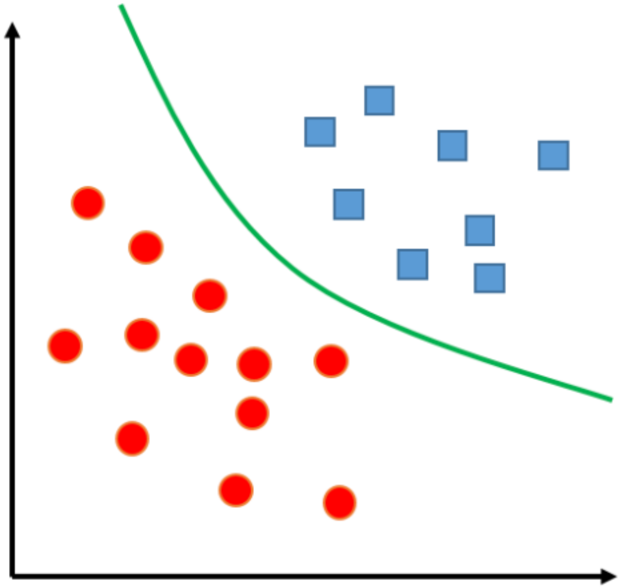
SVM

허석진

선형 분리





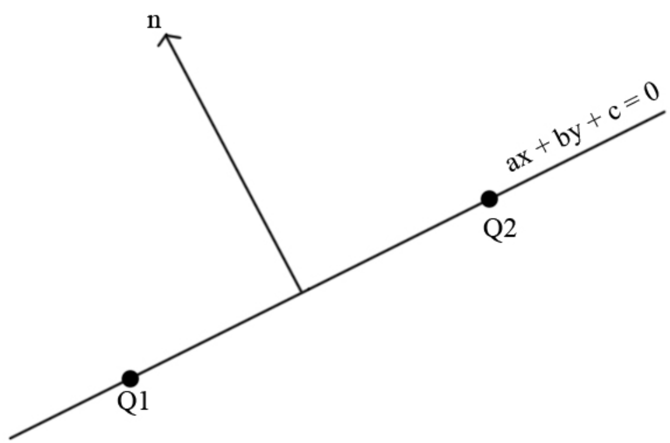
- 모수적 방법?

직선의 방정식

1. $y = m\,x + b$
2. $a\,x + by + c = 0$ → 일반적인 형식

- 내적을 이용한 식

$$\langle a, b\rangle \cdot \langle x, y\rangle + \ c = 0$$



- 점 (x_1, y_1) 까지의 거리

$$\frac{|\langle a,b\rangle\cdot\langle x,y\rangle + c|}{|\langle a,b\rangle|}$$

직선의 방정식

- 직선 중심의 위치 판정

$$\langle a, b \rangle \cdot \langle x, y \rangle + \ c > 0$$

$$\langle a, b \rangle \cdot \langle x, y \rangle + \ c = 0$$

$$\langle a, b \rangle \cdot \langle x, y \rangle + \ c < 0$$

평면의 방정식

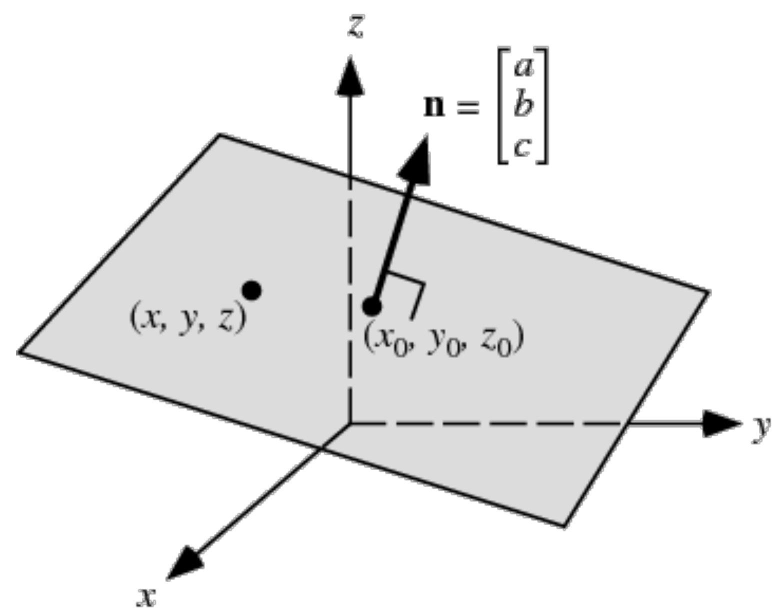
$$a\,x + by + cz + d = 0$$

- 내적을 이용한 식

$$\langle a, b, c \rangle \cdot \langle x, y, z \rangle + d = 0$$

- 점 (x_1, y_1, z_1) 까지의 거리

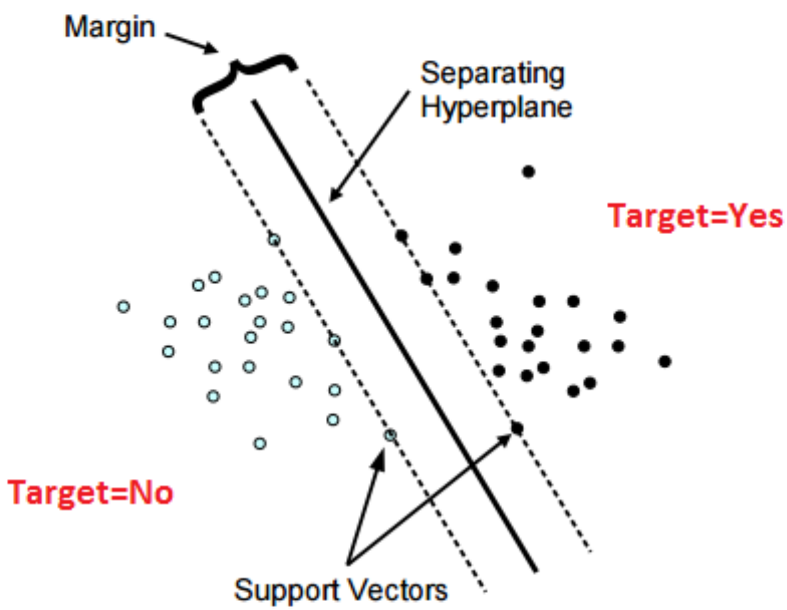
$$\frac{|\langle a,b,c\rangle \cdot \langle x,y,z\rangle + c|}{|\langle a,b,c\rangle|}$$



초평면(Hyperplane) 방정식

- 초평면: 전체 공간의 차원보다 하나 작은 차원의 공간
  - 2차원 평면에서는 직선
  - 3차원 공간에서는 평면
- 초평면 방정식은 $w \cdot x + b = 0$의 형태의 방정식
  - 방정식은 벡터들의 연산 (·은 내적)
- 초평면을 기준으로 한 분류는
  - $w \cdot x + b > 0$ 또는
  - $-w \cdot x + b > 0$ 에 따라 결정
  - 즉, $y(w \cdot x + b) > 0$ 에서 $y = 1$ 또는 $y = -1$에 의해 분류

최대 마진 분류기



Support Vector Machine

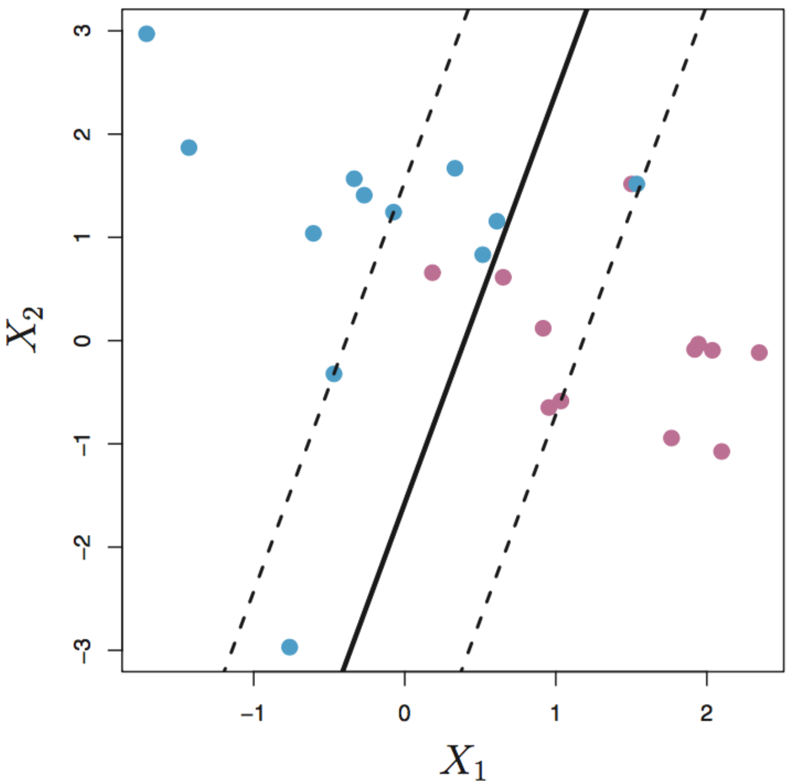
- Margin이 최대가 되도록 하는 분류기

최대 마진 분류기

- 수식에 의한 표현
  1. $|w| = 1$
  2. 모든 i 에 대해 $y_i\ (w \cdot x_i + b) \geq M$

  인 제한 조건 하에서 최대의 M(또는 최소의 $|w|/M$)을 찾음
- 여기서 $y_i = 1$ 또는 $y_i = -1$이 분류된 클래스에 대응
- 모든 관측값들이 직선의 올바른 쪽에 분류 가정

최대 마진 분류기



- 마진 안쪽으로 들어오거나
  분류 직선 기준으로 잘못 분류되는 관측값 허용?

서포트 벡터 분류기

- 일부 관측값이 올바르게 분류되지 않는 경우 허용

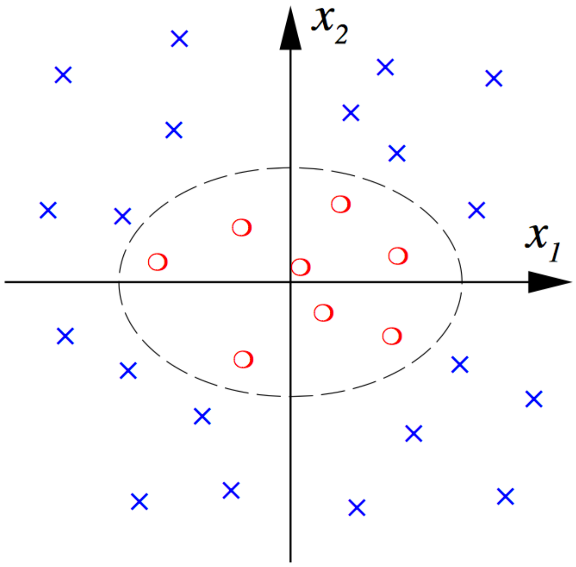
  1. $|w| = 1$
  2. 모든 i 에 대해 $y_i\ (w \cdot x_i + b) \geq M(1 - \varepsilon_i)$
  3. $\sum_i \varepsilon_i \leq C,\ \varepsilon_i \geq 0$

  인 제한 조건 하에서 최대의 M(또는 최소의 $|w|/M$)을 찾음

- 여기서 $y_i = 1$ 또는 $y_i = -1$이 분류된 클래스에 대응
- C 는 cost라고 부름

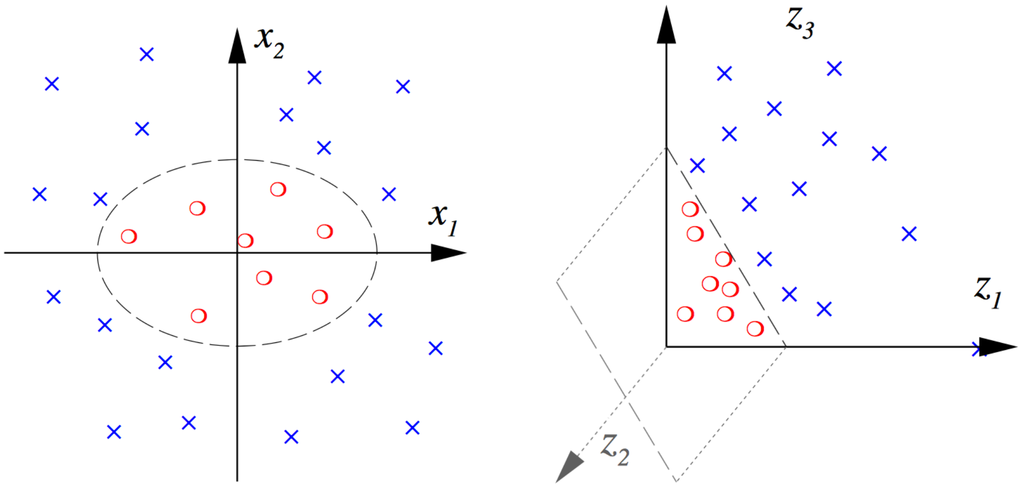
서포트 벡터 분류기

- 선형적으로 나뉘어지지 않는 관측값들은?



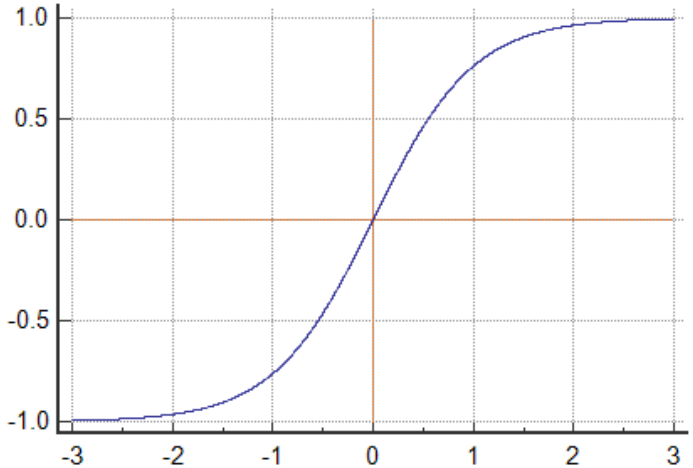
Kernel

- 선형적으로 나뉘어지지 않는 분류를 위해 내적을 재정의



Kernel

- Linear
  - $\mathrm{K}(x, y) = x \cdot y$
- Polynomial
  - $\mathrm{K}(x, y) = (c_0 + \gamma\, x \cdot y)^d$
- Radial
  - $\mathrm{K}(x, y) = \exp(-\gamma |x - y|^2)$
- Sigmoid
  - $\mathrm{K}(x, y) = \tanh(c_0 + \gamma\, x \cdot y)$



둘 보다 많은 분류

- 일대일 분류
  - 분류하려는 클래스의 개수가 K 일 때
  - K 들의 모든 조합에 대해 각각의 테스트 데이터 x_i를 SVM으로 분류
  - 각각의 x_i를 가장 많이 할당된 클래스로 분류
- 일대전부 분류
  - K 중 하나와 나머지 클래스로 나누고
  - 위의 이진 분류 데이터로 SVM으로 각각의 테스트 데이터 x_i를 분류
  - $y_i\,(w \cdot x_i + b)$가 최대가 되는 클래스로 분류