HW5 - writeup

컴퓨터공학부 2017-18538 황선영

1 Theory Questions

1.1 SVM

(a)

"the positive examples are linearly separable from the negative examples"
가 참이라고 하자. 그러면 이를 만족하는 직선 $px+qy+r=0$ 이 존재할 것이다.
이 직선이 두 영역으로 분리할 때 두 점이 같은 영역에 있다면 직선의 방정식에 각 점을 대입한 값의 곱이 양수이고, 다른 영역에 있다면 음수일 것이다.
따라서

$$(a+b+c)(-a-b+c) > 0$$

$$(a+b+c)(a-b+c) < 0$$

$$(a+b+c)(-a+b+c) < 0$$

이다.

i) $a+b+c>0$ 인 경우

$$(a+b-c)<0,\ (a-b+c)<0,\ (-a+b+c)<0$$

이고 이 식을 모두 더하면 $a+b+c<0$ 이므로 모순이다.

ii) $a+b+c<0$ 인 경우

$$(a+b-c)>0,\ (a-b+c)>0,\ (-a+b+c)>0$$

이고 이 식을 모두 더하면 $a+b+c>0$ 이므로 모순이다.

따라서 세 부등식을 모두 만족할 수 없으므로, 주어진 가정은 거짓이다.

따라서 postive examples 와 negative exmples를 linearly separate 할 수 없다.

(b)

$\phi((1,1)) = [1,1,1,1]^T$, $\phi((-1,-1)) = [1,-1,-1,1]^T$

$\phi((1,-1)) = [1,1,-1,-1]^T$, $\phi((-1,1)) = [1,-1,1,-1]^T$

$X_+ = \{[1,1,1,1]^T, [1,-1,-1,1]^T\}$

$X_- = \{[1,1,-1,-1]^T, [1,-1,1,-1]^T\}$

(C)

$w = [a,b,c,d]^T$ 라 하자.

$\min_w \|w\| \text{ s.t. } y_i(w^T x_i + b) \geq 1 \text{ for all } i$ 이므로

$a+b+c+d \geq 1$ — ①

$a-b-c+d \geq 1$ — ②

$a+b-c-d \leq -1$ — ③

$a-b+c-d \leq -1$ — ④

이다.

① + ② : $2a + 2d \geq 2 \Rightarrow a + d \geq 1$

③ + ④ : $2a - 2d \leq -2 \Rightarrow a - d \leq -1$

a, d 모두 0 or 1 이므로

$a = 0, d = 1$ 이다.

① 에 대입하면 $b + c \geq 0$, ② 에 대입하면 $b + c \leq 0$ 이다.

따라서 $b + c = 0$ 이고, b, c 는 0 or 1 이므로

$b = c = 0$ 이다.

$\therefore w = [0, 0, 0, 1]^T$

1.2 K-means

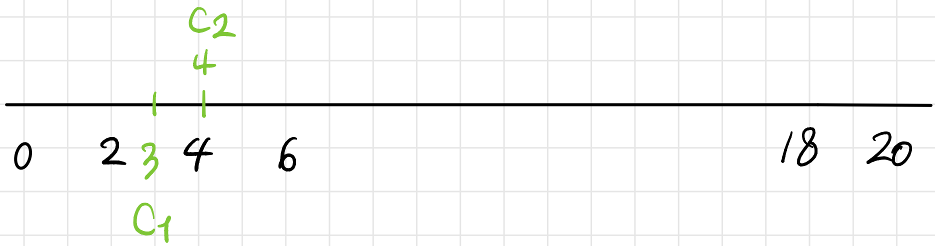
(a)

$d(x, c) = \sqrt{\sum_{i=1}^{N} \frac{(x_i - c_i)^2}{s_i^2}}$ 가 definition일 때 covariance matrix 는 diagonal 하다.

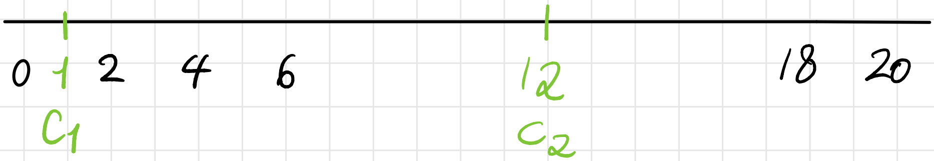
그리고 이 조건 하에서 resulting distance measure은 standardized Euclidean distance로 불린다. Mahalanobis distance가 Euclidean distance로 reduce 되기 때문이다.

(b)

initial



1st iter



$$C_1 = \frac{0+2}{2} = 1$$

$$C_2 = \frac{4+6+18+20}{4} = 12$$

2nd iter

$$C_1 = \frac{0+2+4+6}{4} = 3$$

$$C_2 = \frac{18+20}{2} = 19$$

(c) $C_1 = 4$일 때, 거리의 제곱의 최솟값 $D(x)$가 최소가 되게 하려면 first step에서 $C_2 = 19$이다.

$(0, 2, 4, 6)$, $(18, 20)$으로 나눔

second step에서 $c_3 = 1$일 때

거리가 최소이고,

(0, 2), (4, 6), (18, 20)으로 나뉜다.

따라서 initial $k=3$ center는

$c_1 = 4$, $c_2 = 19$, $c_3 = 1$이다.