HW 2　　2023-21408 정재원.

Problem 1.

1. $K(x, x') = \phi(x)^T \phi(x') = \dfrac{x^T x'}{\|x\| \|x'\|}$

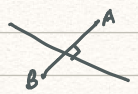
$\Rightarrow \phi(x) = \dfrac{x}{\|x\|}$

A		B	
X (x:y)	$\phi(x)$	X	$\phi(x)$
x:y = 1:2 or 2:1 →	(0.4, 0.9) (0.9, 0.4)	x:y = 1:1	(0.7, 0.7)

(linearly seperable!)

2. maximum-margin decision: $w_1 x_1 + w_2 x_2 + C = 0$

feature space에 mapping된 $\phi(A)$, $\phi(B)$ 값이용하여 계산.



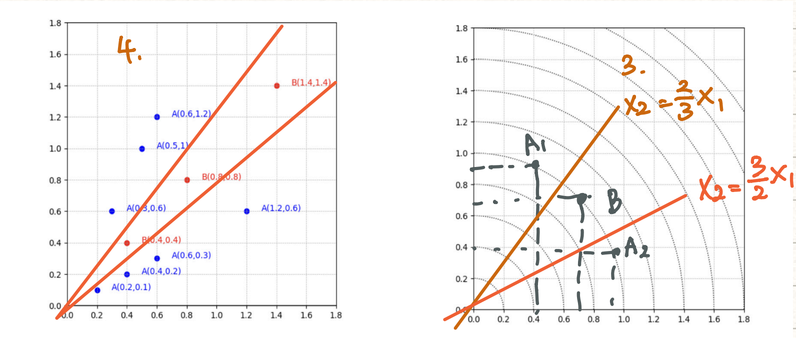
$(A-B)^T \cdot x = 0$

$(-0.3, 0.2)^T x = 0$

$\Rightarrow -0.3x_1 + 0.2x_2 + C = 0$ 이므로

ratio of $w_1, w_2 \Rightarrow \dfrac{w_2}{w_1} = -\dfrac{2}{3}$. (A1), $-\dfrac{3}{2}$ (A2)

3. & 4.



(a)　　(b)

↳ normalized linear kernel로 변은 space 역으로 돌림. (0,0)기45도 축.

Problem 2.

1.

(a)

$k(x, z) = \phi(x)^T \phi(z)$

$= (x_1^2, \sqrt{2}x_1x_2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2, 1)^T (z_1^2, \sqrt{2}z_1z_2, z_2^2, \sqrt{2}z_1, \sqrt{2}z_2, 1)$

$= x_1^2 z_1^2 + 2x_1x_2z_1z_2 + x_2^2 z_2^2 + 2x_1z_1 + 2x_2z_2 + 1$

$= (1 + x_1z_1 + x_2z_2)^2 = (1 + xz)^2$ ☑

(b) x: 각 element 제곱: 2, element 간 interaction (곱): 1

→ z도 same

⇒ 이후 6개의 term 모두 곱하고 더함.

∴ (2+1) × 2 + 6 + 6 = 18 ☑

↑ 곱　↑ 덧셈

(c) kernel 함수에 원래 데이터 (x z)의 내적값만 넣어주면 됨. $(1+xz)^2$

∴ 2+1 = 3. (2 elemt 각각 곱하고, 더함, $xz = x_1z_1 + x_2z_2$) ☑

2.

let) $k'(x,z) = k_1(x,z) + k_2(x,z)$

By Mercer's Condition,

$$\int_x \int_z k'(x,z) f(x) f(z) \, dx dz$$

$$= \int_x \int_z \underbrace{k_1(x,z) f(x) f(z) \, dx dz}_{\geq 0} + \int_x \int_z \underbrace{k_2(x,z) f(x) f(z) \, dx dz}_{\geq 0} \geq 0$$

⇒ $k_1(\cdot,\cdot)$, $k_2(\cdot,\cdot)$는 이미 valid kernel이므로 위 적분도 0 이상임.

양수인 수끼리 합이므로 위식도 0보다 크거나 같음 (무증명 생략)

∴ valid kernel ☑

3.

1. · Max depth of tree

: tree를 언제까지 growing 할 것인지 (너무 깊음 → overfitting / 너무 얕음 → under-fitting), default = 6.

· Learning rate.

: 이전 트리가 다음 트리에 끼치는 영향도 (overfitting 방지), default = 0.3

· Number of trees

: 생성할 tree수, 많을수록 복잡도↑ but 성능↑

· Subsample ratio: train data 에서 sampling 할 비율 (over fitting 방지), default = 1

· Regularization parameter

: [lambda : L2 regularization (주로사용), default = 1 / " = 0
alpha : L1 "] overfitting 방지