AA
(trab.)
6-11-2015

(4.3)

	margem
C = 2 → 27 support vectors	
C = 1 → 11 support vectors	
C = 0,8 → 10 support vectors	0,087
C = 0,9 → 10 support vectors →	0,06
C = 0,6 → 10	0,19 ~~0258~~ ← ✓
C = 0,7 → 10	0,13
C = 0,5 → 12	

Escolheu-se 0,6 pelo facto de dentro dos valores que tem menos support vectors ser o que tem maior margem.

(4.4)

↳ A forma da linha de classificação fica fortemente afectada pelos outliers.

↳ A margem de classificação é bastante menor porque os outliers estão muito próximos de clusters de elementos da outra classe

↳ O número de support vectors aumentou devido ao facto de a fronteira ser mais detalhada, necessitando assim de mais pontos para a definir.

↳ Todas estas alterações ocorrem devido ao facto de o classificador ser intolerante a erros de classificação.

(4.5)

↳ Para bound = inf, a classificação é semelhante à anterior pelo facto de o classificador continuar a ser intolerante a erros de classificação.

↳ bound = 10^3 → 1 training error → 13 support vectors → margem 0,058250

↳ bound = 10^2 → 1 → 15 → 0,19

↳ 10 → 2 → 17 → 0,34

↳ 1 → 2 → 59 → 0,3

↳ 0,1 → 18 → 90 → 1,41

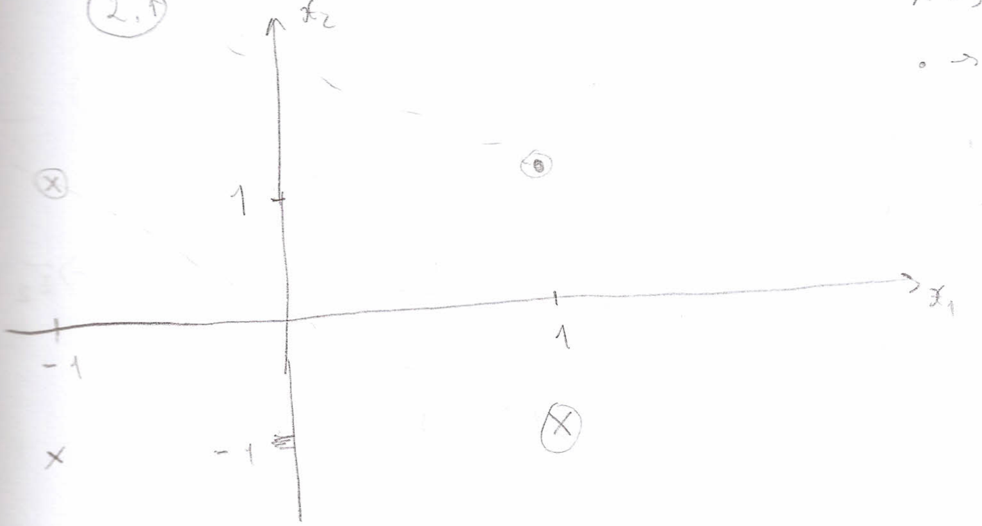
↳ 0,01 → 18 → 90 → 14,13

? just

②

x_1	x_2	d_{AND}	d_{XOR}
-1	-1	-1	-1
-1	1	-1	1
1	-1	-1	1
1	1	1	-1

2.1



$\times \to -1$ class

$\circ \to 1$ class

$(w \cdot x^s - b)\, d^s = 1$

$\Leftrightarrow$

$(x_1 + x_2 - 1)\, d^s = 1$

- $(-1, 1) \Rightarrow -1 \times d^s = 1 \Leftrightarrow 1 = 1$ P.V
- $(1, -1) \Rightarrow 1 = 1$ P.V
- $(1, 1) \Rightarrow d^s = 1 \Leftrightarrow 1 = 1$ P.V

$\bar{w}$ e b satisfazem a equação dada.

- Por inspeção, as margens serão m_b e m_A de tal forma que:

$m_b \Rightarrow x_2 = -x_1$

$m_A \Rightarrow x_2 = -x_1 + b \ \wedge\ 1 = -1 + b \Leftrightarrow b = 2 \Leftrightarrow m_A \Rightarrow x_2 = -x_1 + 2$

- a separação linear de classes será assim: $x_2 = -x_1 + 1 \Rightarrow m = $ boundary

- Os support vectors serão os pontos do seguinte conjunto: $S = \{(-1, 1), (1, -1), (1, 1)\}$

$$x_2 = -x_1 + 1 \Leftrightarrow x_1 + x_2 - 1 = 0 \Rightarrow \underbrace{[\,1\ 1\,]}_{\bar{w}^*} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\bar{X}} \underbrace{-1}_{b} = 0$$

$\bar{w} \cdot \bar{x} - b = 0 \to$ boundary

$\bar{w} \cdot \bar{x} - b > 0 \to$ classe 1

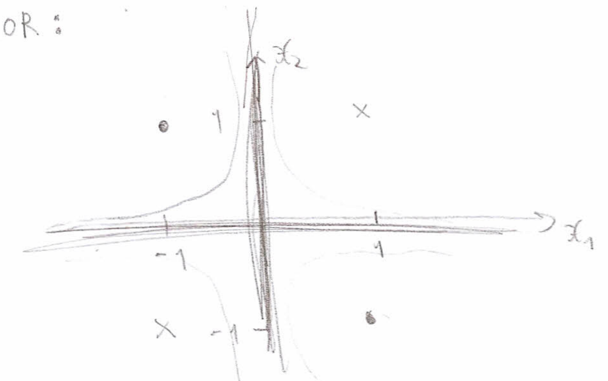
$\bar{w} \cdot \bar{x} - b < 0 \to$ classe 2

2.2

Para a função XOR:

X → classe -1

• → classe 1



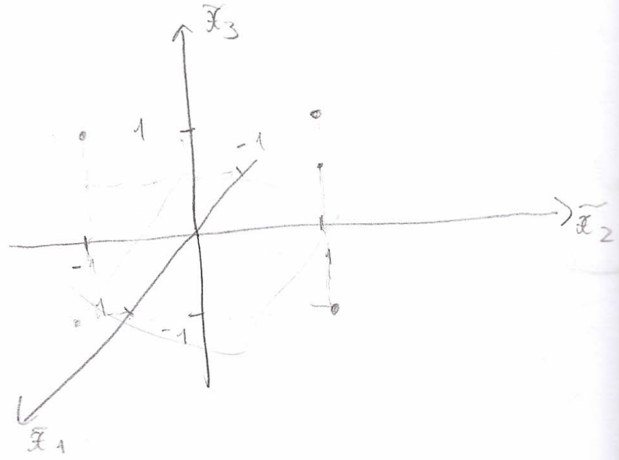
↳ Para a função XOR, não é possível realizar uma classificação linear pois não existe nenhuma recta capaz de separar as duas classes

↳ Non-linear mapping: $\tilde{x} = \varphi(x) = \begin{bmatrix} x_1 \\ x_2 \\ x_1 x_2 \end{bmatrix}$

$$K(\tilde{x},\tilde{y}) = \varphi(\tilde{x})\cdot\varphi(\tilde{y}) = (x_1, x_2, x_1x_2)\cdot(y_1, y_2, y_1y_2) = (x_1y_1, x_2y_2, x_1x_2y_1y_2)$$

2.3

x_1	x_2	$\boxed{\text{XOR}}$	$\tilde{x}_1$ ($= x_1$)	$\tilde{x}_2$ ($= x_2$)	$\tilde{x}_3$ ($= x_1x_2$)
-1	-1	-1	-1	-1	1
-1	1	1	-1	1	-1
1	-1	1	1	-1	-1
1	1	-1	1	1	1



boundary ⇒ $\tilde{x}_3 = 0$ → feature space

$$\underbrace{[0 \quad 0 \quad -1]}_{\tilde{w}} \underbrace{\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{bmatrix}}_{\tilde{x}} - \underbrace{0}_{b} = 0$$

$m_A \Rightarrow \tilde{x}_3 = 1$

$m_B \Rightarrow \tilde{x}_3 = -1$

↳ $(-1,-1,1) \rightarrow (-1)\times(-1) = 1$ ✓

↳ $(-1,1,-1) \rightarrow -1\times(-1) = 1$ ✓

↳ $(1,-1,-1) \rightarrow -1\times(-1) = 1$ ✓

↳ $(1,1,1) \rightarrow -1\times(1) = -1$ ✓

2.4

$-x_3 = 0$ → feature space

↓

$-x_1 x_2 = 0$ → input space

$\Leftrightarrow$

$x_1 = 0 \vee x_2 = 0$

ver 2.2 o sketch

$m_A \Rightarrow x_1 x_2 = 1 \Leftrightarrow x_2 = \frac{1}{x_1}$

$m_B \Rightarrow x_1 x_2 = -1 \qquad x_2 = -\frac{1}{x_2}$

2.5

$(x_1 < 0 \wedge x_2 > 0) \vee (x_1 < 0 \wedge x_2 > 0)$

$x_1 x_2 < 0$

3.1

$$K(x,y) = \phi(x) \cdot \phi(y) = ((x \cdot y) + a)^p - a^p = \left(\sum_{k=0}^{p} {}^{p}C_k \, (x \cdot y)^{p-k} a^k \right) - a^p =$$

$$= \underbrace{\left(\sum_{k=0}^{p} \frac{p!}{k!(p-k)!} (x \cdot y)^{p-k} a^k \right)}$$ $$- a^p$$

- Sendo x e y vectores, e n as suas dimensões
$$K(x,y) = \left(\sum_{k=0}^{p} \frac{p!}{k!(p-k)!} \left(\sum_{i=1}^{n} x_i y_i \right)^{p-k} a^k \right) - a^p$$

- Desprezando "$-a^p$", para $p=2$:
$$K(x,y) = \sum_{i=1}^{n} (x_i)^2 (y_i)^2 + \sum_{i=2}^{n} \sum_{j=1}^{i-1} (\sqrt{2}\, x_i x_j)(\sqrt{2}\, y_i y_j) + \sum_{i=1}^{n} (\sqrt{2a}\, x_i)(\sqrt{2a}\, y_i) + a^2$$

↳ 6 features n.º p/ para $p=2$ será então:

$$\phi(x) = (\underbrace{x_n^2, \dots, x_1^2}_{n}, \underbrace{\sqrt{2}\, x_n x_{n-1}, \dots, \sqrt{2}\, x_n x_1}_{n-1}, \sqrt{2}\, x_{n-1} x_{n-2}, \dots, \sqrt{2}\, x_{n-1} x_1, \dots, \sqrt{2}\, x_2 x_1, \sqrt{2a}\, x_n, \dots, \sqrt{2a}\, x_1, a)$$

Sendo $n=2$: $\phi(x) = (x_2^2, x_1^2, \sqrt{2}\, x_2 x_1, \sqrt{2a}\, x_2, \sqrt{2a}\, x_1, a)$

5 dimensões no feature space para $p=2$

$$\sum_{k=0}^{p} \frac{p!}{k!(p-k)!}(xy)^{p-k} a^k \text{, onde } xy = \sum_{i=0}^{2} x_i y_i = x_1 y_1 + x_2 y_2 =$$

• No caso geral:

$K(x,y) \Rightarrow p+1$ termos

$(xy)^{p-k} = (x_1y_1 + x_2y_2)^{p-k} = p-k+1$ termos

$p, p-1, p-2, \ldots, 0$

$$\sum_{k=0}^{p} \frac{p!}{k!(p-k)!}(xy)^{p-k} a^k;$$

$$= (xy)^{p} + p(xy)^{p-1} a + \frac{p!}{2(p-2)!}(xy)^{p-2}a^2 + \frac{p!}{3(p-3)!}(xy)^{p-3}a^3 + \ldots +$$

$$+ \frac{p!}{(p-1)!(p-p+1)!} xy\, a^{p-1} + a^p = (*)$$

//

$xy = x_1y_1 + x_2y_2$

$$(xy)^p = (x_1y_1 + x_2y_2)^p = (x_1y_1)^p + p(x_1y_1)^{p-1}(x_2y_2) + \frac{p!}{2(p-2)!}(x_1y_1)^{p-2}(x_2y_2)^2$$

//

$$(*) = (x_1y_1)^p + p(x_1y_1)^{p-1}(x_2y_2) + \frac{p!}{2(p-2)!}(x_1y_1)^{p-2}(x_2y_2)^2 + \ldots + \frac{p!}{(p-1)!} x_1y_1(x_2y_2)^{p-1} + (x_2y_2)^p +$$

$(xy)^p$

$p(xy)^{p-1}a$

$$+ p \cdot a \cdot \left[(x_1y_1)^{p-1} + (p-1)(x_1y_1)^{p-2}(x_2y_2) + \frac{(p-1)!}{2(p-3)!}(x_1y_1)^{p-3}(x_2y_2)^2 + \ldots + \frac{(p-1)!}{(p-2)!} x_1y_1(x_2y_2)^{p-2} + (x_2y_2)^{p-1}\right] +$$

$$+ \frac{p!}{2(p-2)!}a^2\left[(x_1y_1)^{p-2} + (p-2)(x_1y_1)^{p-3}(x_2y_2) + \frac{(p-2)!}{2(p-4)!}(x_1y_1)^{p-4}(x_2y_2)^2 + \ldots + \frac{(p-2)!}{(p-3)!}x_1y_1(x_2y_2)^{p-2} + (x_2y_2)^{p-2}\right] +$$

$$+ \frac{p!}{3(p-3)!}a^3\left[(x_1y_1)^{p-3} + (p-3)(x_1y_1)^{p-4}(x_2y_2) + \frac{(p-3)!}{2(p-5)!}(x_1y_1)^{p-4}(x_2y_2)^2 + \ldots + \frac{(p-3)!}{(p-4)!}x_1y_1(x_2y_2)^{p-3} + (x_2y_2)^{p-3}\right] +$$

$+ \ldots + p\, x_1y_1 a^{p-1} + p\, x_2y_2 a^{p-1} + \cancel{a^p}$ → termo que irá cortar com o termo "$-a^p$" até agora depois

$$f(x) = \left(x_1^p, x_2^p, \sqrt{p}\, x_1^{p-1} x_2, \sqrt{\frac{p!}{2(p-2)!}}\, x_1^{p-2} x_2^2, \ldots, \sqrt{p}\, x_1 x_2^{p-1}, \ldots,\right.$$

$$\ldots, \sqrt{pa}\, x_1^{p-1}, \sqrt{pa}\, x_2^{p-1}, \sqrt{pa}\sqrt{(p-1)}\, x_1^{p-2}x_2, \sqrt{pa}\sqrt{\frac{(p-1)!}{2(p-3)!}}\, x_1^{p-3}x_2^2, \ldots, \sqrt{pa}\sqrt{\frac{(p-1)!}{(p-3)!}}\, x_1 x_2^{p-2}, \ldots,$$

$$\ldots, \sqrt{\frac{p!\,a^2}{2(p-2)!}}\, x_1^{p-2}, \sqrt{\frac{p!\,a^2}{2(p-2)!}}\, x_2^{p-2}, \sqrt{\frac{p!\,a^2}{2(p-2)!}}\sqrt{p-2}\, x_1^{p-3}x_2, \sqrt{\frac{p!\,a^2}{2(p-2)!}}\sqrt{\frac{(p-2)!}{2(p-4)!}}\,(x_1)^{p-4}x_2^2 + \ldots +$$

$$\left., \sqrt{pa^{p-1}}\, x_1, \sqrt{pa^{p-1}}\, x_2\right)$$

dimensão: $(p+1)\times(p+1) = p^2+2p+1$?

3.2

Para $p=2$: $K(x,y) = x_2^2 y_2^2 + x_1^2 y_1^2 + 2 x_2^2 x_1^2 + 2a x_2^2 + 2a x_1^2 + \cancel{a^2} - \cancel{a^2}$

$\varphi(x) = (x_2^2, x_1^2, \sqrt{2} x_2 x_1, \sqrt{2a}\, x_2, \sqrt{2a}\, x_1)$

Us em 2.2 $\rightarrow \varphi(x) = (x_1, x_2, x_1 x_2)$ Temos mais variáveis mas afectadas de uma constante.

- 3.2 tem mais 2 dimensões

$\tilde{w}$?