EX 5 - Theory + SVM

207380528, 302673355

May 28, 2022

Q1 - Kernels and mapping functions

$K(x,y) = (x \cdot y + 1)^3$ function over $\mathbb{R}^2 \times \mathbb{R}^2$. Find Ψ for which K is a kernel

A

We'll expand the right hand side argument to find Ψ.

$$(x \cdot y + 1)^3 = (x \cdot y + 1)^2 (x \cdot y + 1) = ((x \cdot y)^2 + 2x \cdot y + 1)(x \cdot y + 1)$$

$$= (x^2 \cdot y^2 + 2x \cdot y + 1)(x \cdot y + 1) = x^3 \cdot y^3 + 2x^2 \cdot y^2 + x \cdot y + x^2 + y^2 + 2x \cdot y + 1$$

$$(x \cdot y)^3 + 3x \cdot y^2 + 3x \cdot y + 1$$

We now will apply $x \in \mathbb{R}^2$ to it and get:

$$(x_1y_1 + x_2y_2)^3 + 3(x_1y_1 + x_2y_2)^2 + 3(x_1y_1 + x_2y_2) + 1$$

And simplify it a bit more:

$$(x_1y_1+x_2y_2)^2(x_1y_1+x_2y_2)+3((x_1y_1)^2+2x_1y_1x_2y_2+(x_2y_2)^2)+3(x_1y_1+x_2y_2)+1$$

$$((x_1y_1)^2+2x_1y_1x_2y_2+(x_2y_2)^2)(x_1y_1+x_2y_2)+3((x_1y_1)^2+2x_1y_1x_2y_2+(x_2y_2)^2))+3(x_1y_1+x_2y_2)+1$$

$$(x_1y_1)^3+2(x_1y_1)^2x_2y_2+(x_2y_2)^2(x_1y_1)+(x_1y_1)^2(x_2y_2)+2(x_2y_2)^2x_1y_1+(x_2y_2)^3+3((x_1y_1)^2+2x_1y_1x_2y_2+(x_2y_2)^2))$$

$$(x_1y_1)^3+3(x_1y_1)^2(x_2y_2)+3(x_2y_2)^2(x_1y_1)+(x_2y_2)^3+3(x_1y_1)^2+6x_1y_1x_2y_2+3(x_2y_2)^2+3x_1y_1+3x_2y_2+1$$

Meaning, for $\Psi(x)$ to afford the kernel$K$ it means:

$$\Psi(x) = x_1^3 + \sqrt{3}x_1^2x_2 + \sqrt{3}x_2^2x_1 + x_2^3 + \sqrt{3}x_1^2 + \sqrt{6}x_1x_2 + \sqrt{3}x_2^2 + \sqrt{3}x_1 + \sqrt{3}x_2 + 1$$

$$\Psi(y) = y_1^3 + \sqrt{3}y_1^2y_2 + \sqrt{3}y_2^2y_1 + y_2^3 + \sqrt{3}y_1^2 + \sqrt{6}y_1y_2 + \sqrt{3}y_2^2 + \sqrt{3}y_1 + \sqrt{3}y_2 + 1$$

B

In class we called it “kernel for the full rational variaties mapping”.

C

The calculations themselves are done in $\mathbb{R}^2$ instead of in $\mathbb{R}^{10}$, so we'll save 8 multiplications per instance.

Q2 - Lagrange multipliers

A

$f(x,y) = 2x - y$. Find the min and max point for f under $g(x,y) = \frac{x^2}{4} + y^2 = 1$.

We can write the constraint as

$$\frac{x^2}{4} + y^2 - 1 = 0$$

And write the Lagrangian:

$$L(x,y) = 2x - y + \lambda(\frac{x^2}{4} + y^2 - 1)$$

Next, we'll derive by x, y and λ:

$$(1)\frac{\partial L}{\partial x} = 2 + \frac{2\lambda x}{4} = 0$$
$$\lambda = -4x$$

$$(2)\frac{\partial L}{\partial y} = -1 + 2\lambda y = 0$$

$$\lambda = \frac{1}{2y}$$

We'll put both (1) and (2) togehter and:

$$(3) - 4x = \frac{1}{2y}$$

$$x = -\frac{1}{8y}$$

$$y = -8x$$

$$\frac{\partial L}{\partial \lambda} = \frac{x^2}{4} + y^2 - 1 = 0$$

Plug in (3)

$$\frac{64y^2}{4} + y^2 - 1 = 0$$

$$17y^2 = 1$$

$$y = \pm\frac{1}{\sqrt{17}}$$

$$x = \mp\frac{8}{\sqrt{17}}$$

Plug it back in f:

$$f(-\frac{8}{\sqrt{17}}, \frac{1}{\sqrt{17}}) = -\sqrt{17}$$

$$f(\frac{8}{\sqrt{17}}, -\frac{1}{\sqrt{17}}) = \sqrt{17}$$

It seems that $\sqrt{17}$ is the max point and that $-\sqrt{17}$is the min point

Q3- PAC Learning

fa

fa
fas

Q4 - Confidence Intervals

We'll compute the confidence interval, according to the data provided. We'll take the upper bound of the interval, to find the maximum percentage that we can commit to in 95% confidence.

$$p \in [\hat{p} + 2se, \hat{p} - 2se]$$

$$se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

We know that $\hat{p} = 0.2$, $n = 1000$. We'll plug it in and get:

$$se = \sqrt{\frac{0.2(1-0.2)}{1000}} = 0.0126$$

$$p \in [0.175, 0.225]$$

The maximum error, with 95% confidnce, is 22.5%.

Q4 - SVM