CSE-5819 Written Assignment 1

Name: Kexin Chu
Studnet ID: 3145867

Q1. Linear Algebra and Probability

（1）

Answers:

For given vectors, can get:

$$x_2 - x_1 = [0.7, 0, -0.5, -0.2]$$
$$x_3 - x_1 = [0, 0.8, 0.5, 0]$$
$$x_4 - x_1 = [0.8, -0.1, 0.2, -0.2]$$

Then execute the result of different norm functions:

	$x_2 - x_1$	$x_3 - x_1$	$x_4 - x_1$	My final answer
(a) L_0	3	2	4	x_3
(b) L_1	1.4	1.3	1.3	x_3， x_4
(c) L_2	$\sqrt{0.78}$	$\sqrt{0.89}$	$\sqrt{0.73}$	x_4
(d) L_∞	0.7	0.8	0.8	x_2

（2）

Answers:

$$\mathrm{E}[(X - X')^2]$$
$$= E[X^2 - 2XX' - X'^2]$$
$$= E[X^2] - 2E[XX'] + E[X'^2]$$
$$= E[X^2] - 2E[X]E[X'] + E[X'^2]$$
$$= \mu_1^2 + \sigma^2 - 2\mu_1^2 + \mu_1^2 + \sigma^2$$
$$= 2\sigma^2$$

（3）

Answer:

Firstly, use w=2 to re-write the p(x) function:

$\mathrm{p}(x)=0$ if x < 0

$\mathrm{p}(x)=1-\frac{\mathrm{x}}{2}$ if 0 <= x <= 2

$\mathrm{p}(x)=0$ if x > 2

Then:

$$p(x=1 \mid w=2)=1-\frac{1}{2}=0.5$$

（4）

Answer:

Firstly, write the likelihood function

$$\mathrm{L}(\lambda \mid x)=\prod_{i=1}^{N} p(x_i \mid \lambda)=\prod_{i=1}^{N} \lambda e^{-\lambda x}=\lambda^{N} e^{-\lambda \sum_{i=1}^{N} x_i}$$

Secondly, execute the log

$$\ell(\lambda \mid x)=\log(L(\lambda \mid x))=\mathrm{N}\log(\lambda)-\lambda \sum_{i=1}^{N} x_i$$

Finally MLE:

$$\frac{\partial \ell(\lambda \mid x)}{\partial \lambda}=\frac{N}{\lambda}-\sum_{i=1}^{N} x_i=0$$

$$\lambda=\frac{\mathrm{N}}{\sum_{i=1}^{N} x_i}$$

Q2. Introduction to Optimization

(1)

Answer:

(a)

False

Reason: the definition of optimization problem is

$$\min_x f(x)$$
$$s.t. x \in F$$

So I think it is a minimization problem.

(b)

Here I use the $f''(x) > 0$? to judge whether f(x) is convex on $-\infty < x < c$

$$f'(x) = \frac{\partial f(x)}{\partial x} = \frac{2x(c-x)+x^2}{(c-x)^2} = \frac{2xc-x^2}{(c-x)^2}$$

$$f''(x) = \frac{\partial f'(x)}{\partial x}$$
$$= \frac{2(c-x)^3 - 2(2xc-x^2)(x-c)}{(c-x)^4}$$
$$= \frac{2(c^2-2cx+x^2+2xc-x^2)}{(c-x)^3}$$
$$= \frac{2c^2}{(c-x)^3}$$

When $-\infty < x < c$, The result for $f''(x) > 0$ is always True. So it is convex.

(2)

Answer:

When use Lagrange multipliers, we can get:

$$L(x, y, \lambda) = 6xy + \lambda(\frac{x^2}{9} + \frac{y^2}{16} - 1)$$
$$= 6xy - \frac{\lambda}{9}x^2 + \frac{\lambda}{16}y^2 - \lambda$$

Then, calculate:

$$\frac{\partial \mathrm{L}}{\partial x} = 6y + \frac{2\lambda}{9}x = 0$$

$$\frac{\partial \mathrm{L}}{\partial y} = 6x + \frac{\lambda}{8}y = 0$$

$$\frac{\partial \mathrm{L}}{\partial \lambda} = \frac{x^2}{9} + \frac{y^2}{16} - 1 = 0$$

Based on these equations,

$$\mathrm{y} = -\frac{\lambda}{27}x$$

$$\mathrm{y} = -\frac{48}{\lambda}x$$

$$\frac{x^2}{9} + \frac{y^2}{16} = 1$$

Then, use the first and second equation, we can get $\lambda^2 =（48*27）$

Finally, we can get:

$$x = \pm\sqrt{\frac{9}{2}}$$

$$y = \pm 2\sqrt{2}$$

Q3. Linear Regression

Answer:

(a)

$$L(w) = (y - Xw - \varepsilon)^T (y - Xw - \varepsilon)$$

Because the ε obeys the normalization distribution, It is a scalar, Then

$$L(w) = (y - Xw)^T (y - Xw)$$
$$= (y^T - w^T X^T)(y - Xw)$$
$$= y^T y - y^T Xw - w^T X^T y + w^T X^T Xw$$
$$= y^T y - 2y^T Xw + w^T X^T Xw$$

Then

$$\frac{\partial L(w)}{\partial w} = 0 - 2y^T X + 2w^T X^T X = 0$$
$$w^T X^T X = y^T X$$

Finally, transpose both sides of the equal sign, We can get the w based on X and y

$$X^T Xw = X^T y$$
$$w = (X^T X)^{-1}(X^T y)$$

(b)

$$X^T X = \{\begin{matrix} 6 & 8 \\ 8 & 14 \end{matrix}\}$$

$$X^T y = \{\begin{matrix} 13 \\ 17 \end{matrix}\}$$

=>

$$(X^T X)^{-1} = \{\begin{matrix} 0.7 & -0.4 \\ -0.4 & 0.3 \end{matrix}\}$$

$$w = \{\begin{matrix} 2.3 \\ -0.1 \end{matrix}\}$$