Tsinghua-Berkeley Shenzhen Institute
LEARNING FROM DATA
Fall 2021

Homework 0

Issued: Sunday 12th September, 2021 **Due:** Friday 17th September, 2021

Tips: It is not a formal homework and will not be graded. The goal is to help you recall the basic mathematics you have learnt before.

Calculus & Linear Algebra

0.1. (Inner product)

(a) If $\boldsymbol{x} = [1, -1]^{\mathrm{T}}, y = [1, 1]^{\mathrm{T}}$, calculate $\|\boldsymbol{x}\|$, $\|\boldsymbol{y}\|$ and $\|\boldsymbol{x} + \boldsymbol{y}\|$;

(b) If $\boldsymbol{x} \in \mathbb{R}^n$ is orthogonal to $\boldsymbol{y} \in \mathbb{R}^n$, please show that

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.$$

0.2. (Orthogonal Matrix)

(a) For matrix $\mathbf{Q} = \begin{bmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{bmatrix}, \boldsymbol{x} \in \mathbb{R}^2$, please show that $\|\mathbf{Q}\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$.

(b) For orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}, \boldsymbol{x} \in \mathbb{R}^n$, please show that $\|\mathbf{Q}\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$.

0.3. (Trace) For any matrices $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}$, please show that

(a) $\mathrm{trace}(\mathbf{AB}) = \mathrm{trace}(\mathbf{BA})$

(b) $\mathrm{trace}(\mathbf{ABC}) = \mathrm{trace}(\mathbf{CAB}) = \mathrm{trace}(\mathbf{BCA})$

(c) $\nabla_{\mathbf{A}} \mathrm{tr}(\mathbf{AB}) = \mathbf{B}^{\mathrm{T}}$.

0.4. (Chain rule) When $x \in \mathbb{R}$ is a scalar, we have

$$y = ax + b$$
$$z = \frac{1}{1 + e^{-y}}$$

Please calculate $\frac{\partial z}{\partial x}$.

0.5. (Matrix derivative) For vectors $\boldsymbol{x}, \boldsymbol{w} \in \mathbb{R}^n$, and matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, we have $f : \mathbb{R}^n \to \mathbb{R}$ as

$$f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\mathbf{A}\boldsymbol{x} + \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$$

Please calculate $\nabla_{\boldsymbol{x}} f(\boldsymbol{x})$.

0.6. (Matrix derivative) For vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and matrix $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$, we have $\ell : \mathbb{R}^n \to \mathbb{R}$ as

$$\ell(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{y})^{\mathrm{T}}(\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{y}).$$

Please calculate $\nabla_{\boldsymbol{x}} \ell(\boldsymbol{x})$.

Probability Theory

0.7. (Conditional Probability) Explain that $\mathbb{E}\left[\mathbb{E}\left[X|Y\right]\right] = \mathbb{E}[X]$

0.8. (Bayes) A city has a 50% chance to rain everyday and the weather report has a 90% chance to correctly forecast.
You will take an umbrella when the report says it will rain and you have a 50% chance to take an umbrella when the report says it will not rain.
Compute

(a) the probability of raining when you don't take an umbrella;

(b) the probability of not raining when you take an umbrella.

0.9. (Joint Distribution) Random Variables X and Y have a joint distribution with joint probability density function

$$f(x,y) = \begin{cases} Ce^{-(2x+y)} & x > 0, y > 0 \\ 0 & ow. \end{cases}$$

Please find C by

$$\int_0^\infty \int_0^\infty f(x,y)\mathrm{d}x\mathrm{d}y = 1.$$

(Hint: The integral of a probability density function is 1.)

0.10. (Covariance) We have a joint probability density function

$$f(x,y) = \begin{cases} 4xy & 0 < x < 1, 0 < y < 1 \\ 0 & ow. \end{cases}$$

Please show that the covariance of X and Y is 0.

0.11. (Uncorrelated and independent RVs) We have a uniform distribution of X and Y on a disk. The probability density function is

$$f(x,y) = \frac{1}{\pi} \qquad x^2 + y^2 \leq 1$$

(a) Calculate $\mathbb{E}[X], \mathbb{E}[Y]$ and $\mathbb{E}[XY]$.

(b) Please show that X and Y are uncorrelated but not independent.

0.12. (Gaussian Distribution) There is a famous integral here

$$\int_{-\infty}^\infty e^{-x^2}\mathrm{d}x = \sqrt{\pi}$$

It is called Gaussian Integral. Based on it, please find some results of the Normal Distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad -\infty < x < \infty$$

(a) Prove that it is a probability density function ($\sigma > 0$).

(b) Compute the expectation $\mathbb{E}[X]$ and variance $\mathrm{Var}(X)$.