第六章　习题

习题 1. 计算导数

$$f(\boldsymbol{x}) = \log(x^4)\sin(x^3)$$

习题 2. 计算导数

$$f(\boldsymbol{x}) = \frac{1}{1+\exp(-x)}$$

习题 3. 计算导数

$$f(\boldsymbol{x}) = \exp(-\frac{1}{2\boldsymbol{\sigma}^2}(\boldsymbol{x}-\boldsymbol{\mu})^2)$$

这里的 $\boldsymbol{\mu}$，$\boldsymbol{\sigma}$ 都是常量。

习题 4. 当 $\boldsymbol{x}_0 = 0$ 时计算泰勒多项式 T_n，$f(\boldsymbol{x}) = \sin(\boldsymbol{x}) + \cos(\boldsymbol{x})$，其中 $n = 0, \cdots, 5$。

习题 5. 有以下函数

$$f_1(\boldsymbol{x}) = \sin(\boldsymbol{x}_1)\cos(\boldsymbol{x}_2), \boldsymbol{x} \in \mathbb{R}^2$$

$$f_1(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x}^T\boldsymbol{y}, \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$$

$$f_1(\boldsymbol{x},\boldsymbol{y}) =, \boldsymbol{x} \in \mathbb{R}^n$$

1. $\frac{\partial f_t}{\partial \boldsymbol{x}}$ 的维数是多少？
2。计算雅克比矩阵。

习题 6. f 对 t 求导，g 对 $\boldsymbol{X}$ 求导，其中

$$f(\boldsymbol{t}) = \sin(\log(\boldsymbol{t}^T\boldsymbol{t})), \boldsymbol{t} \in \mathbb{R}^D$$

$$g(\boldsymbol{X}) = \mathrm{Tr}(\boldsymbol{AXB}), \boldsymbol{A} \in \mathbb{R}^{D*E}, \boldsymbol{X} \in \mathbb{R}^{E*F}, \boldsymbol{B} \in \mathbb{R}^{F*D}$$

Tr 表示迹。

习题 7. 用链式法则计算下列函数的导数 $\frac{df}{dx}$，给出每个偏导数的维数，详细描述你的步骤。
1.

$$f(z) = \log(1+z), z = \boldsymbol{x}^T\boldsymbol{x}, \boldsymbol{x} \in \mathbb{R}^D$$

2.

$$f(\boldsymbol{z}) = \sin(\boldsymbol{z}), \boldsymbol{z} = \boldsymbol{Ax} + b, \boldsymbol{A} \in \mathbb{R}^{E*D}, \boldsymbol{x} \in \mathbb{R}^D, b \in \mathbb{R}^E$$

其中 $\sin(\cdot)$ 作用于每个 $\boldsymbol{z}$ 元素。

习题 8. 计算下列函数的导数 $\frac{df}{dx}$，详细描述你的步骤。
1。使用链式法则，计算每个偏导数的维数。

$$f(z) = \exp(-\frac{1}{2}z)$$

$$z = g(\boldsymbol{y}) = \boldsymbol{y}^T\boldsymbol{S}^{-1}\boldsymbol{y}$$

$$\boldsymbol{y} = h(\boldsymbol{x}) = \boldsymbol{x} - \mu$$

其中 $\boldsymbol{x}, \mu \in \mathbb{R}^D, \boldsymbol{S} \in \mathbb{R}^{D*D}$。

2.

$$f(\boldsymbol{x}) = \mathrm{Tr}(\boldsymbol{x}\boldsymbol{x}^T + \sigma^2 \boldsymbol{I}), \boldsymbol{x} \in \mathbb{R}^D$$

这里 $\mathrm{Tr}(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的迹，即所有对角元素之和。提示: 需要明确写出外积。

3。使用链式法则。给出每个偏导数的维数。不需要明确地计算偏导数的乘积。

$$f = \tanh(\boldsymbol{z}) \in \mathbb{R}^M$$

$$\boldsymbol{z} = \boldsymbol{A}\boldsymbol{z} + b, \boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{A} \in \mathbb{R}^{M*N}, b \in \mathbb{R}^M$$

这里的 tanh 作用于 $\boldsymbol{z}$ 的每一个分量。

习题 9. 构建模型使得预测值与真实值的误差最小常用向量 l_2 范数度量，求解模型过程中需要计算梯度，求梯度：

- $f(\boldsymbol{A}) = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b} - \boldsymbol{y}\|_2^2$，求 $\frac{\partial f}{\partial \boldsymbol{A}}$。
- $f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b} - \boldsymbol{y}\|_2^2$，求 $\frac{\partial f}{\partial \boldsymbol{x}}$。

习题 10. 求 $\frac{\partial\, \mathrm{Tr}(\boldsymbol{W}^{-1})}{\partial \boldsymbol{W}}$，利用迹微分法求解。

习题 11. 二次型是数据分析中常用函数，求 $\frac{\partial \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}}{\partial \boldsymbol{x}}$，$\frac{\partial \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}}{\partial \boldsymbol{A}}$

习题 12. $f(\boldsymbol{z}) = \frac{\exp(\boldsymbol{z})}{\boldsymbol{1}^T \exp(\boldsymbol{z})}$ 称为 *softmax* 函数，$(\exp(\boldsymbol{z}))_i = \exp(\boldsymbol{z}_i)$，如果 $\boldsymbol{q} = \frac{\exp(\boldsymbol{z})}{\boldsymbol{1}^T \exp(\boldsymbol{z})}, J = -\boldsymbol{p}^T log(\boldsymbol{q})$，其中 $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{z} \in \mathbb{R}^n$，并且 $\boldsymbol{1}^{\mathrm{T}}\boldsymbol{p} = 1$，

- 证：$\frac{\partial J}{\partial \boldsymbol{z}} = \boldsymbol{q} - \boldsymbol{p}$
- 若 $\boldsymbol{z} = \boldsymbol{W}\boldsymbol{x}$，其中 $\boldsymbol{W} \in \mathbb{R}^{n\times m}, \boldsymbol{x} \in \mathbb{R}^m$，$\frac{\partial J}{\partial \boldsymbol{W}} = (\boldsymbol{q} - \boldsymbol{p})\boldsymbol{x}^T$ 是否成立。

习题 13. 以下内容是求解正态分布模型的关键步骤：$L = -\frac{Nd}{2} ln(2\pi) - \frac{N}{2} ln|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_t (\boldsymbol{x}_t - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_t - \boldsymbol{\mu})$。

1) 求 $\frac{\partial L}{\partial \boldsymbol{\mu}}$。

2) 当 $\mu = \frac{1}{N}\sum_t \boldsymbol{x}_t$ 时求 $\frac{\partial L}{\partial \boldsymbol{\Sigma}}$，求 $\boldsymbol{\Sigma}$ 使 $\frac{\partial L}{\partial \boldsymbol{\Sigma}} = 0, \boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{\Sigma} \in \mathbb{R}^{N\times N}$。

习题 14. 求 $\frac{\partial |\boldsymbol{X}^k|}{\partial \boldsymbol{X}}$。

习题 15. 求 $\frac{\partial\, \mathrm{Tr}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^T\boldsymbol{C})}{\partial \boldsymbol{X}}$。