作业七

1. 第一题

$$
\begin{aligned}
f'(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\big(1+e^{-x}\big)^{-1} = -\big(1+e^{-x}\big)^{-2}\cdot\frac{\mathrm{d}}{\mathrm{d}x}\big(1+e^{-x}\big)\\
&= -\big(1+e^{-x}\big)^{-2}\cdot(-e^{-x})\\
&= \frac{e^{-x}}{(1+e^{-x})^2} = f(x)\big(1-f(x)\big)\ .
\end{aligned}
$$

2. 第二题

1. 对 f 求导:

$$\frac{df}{dz} = -\frac{1}{2}f(z)\ .$$

维度: 1×1(标量)。

2. 对 z 求导:

$$\frac{\partial z}{\partial y} = \left(S^{-1}+S^{-\top}\right)y \text{ or } y^T\left(S^{-1}+S^{-\top}\right)\ ,$$

维度: $D\times 1$ 或者 $1\times D$(Jacobian)。

3. 对 y 求导:

$$\frac{\partial y}{\partial x} = I_D\ .$$

维度: $D\times D$。

4. 由链式法则可知:

$$\frac{df}{dx} = \frac{df}{dz}\cdot\frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial x} = -\frac{1}{2}f(z)\left(S^{-1}+S^{-\top}\right)(x-\mu)\ .$$

维度: $D\times 1$。

Jacobian 形式:

$$\frac{df}{dx} = -\frac{1}{2}f(z)\,(x-\mu)^\top\big(S^{-1}+S^{-\top}\big)$$

维度: $1\times D$。

3. 第三题

由迹的性质有,

$$\mathrm{tr}(xx^\top) = x^\top x, \qquad \mathrm{tr}(\sigma^2 I) = \sigma^2 D,$$

因此

$$f(x) = x^\top x + \sigma^2 D.$$

其中 $\sigma^2 D$ 为常数。有

$$\frac{df}{dx} = 2x^\top.$$

4. 第四题

函数 f 在点 x 处沿单位方向向量 u 的方向导数为 $D_u f(x)$. 设 f 在 x 可微，有

$$D_u f(x) = \nabla f(x) \cdot u = \|\nabla f(x)\| \, \|u\| \cos\theta \ .$$

其中 θ 为 u 与 $\nabla f(x)$ 的夹角。由于 $\|u\| = 1$,

$$D_u f(x) = \|\nabla f(x)\| \cos\theta \ .$$

$\theta = 0$ 时最大，即 u 与 $\nabla f(x)$ 同方向。因此函数变化最快的方向是梯度方向 $\nabla f(x)$。

5. 第五题

梯度:

$$\nabla f(x) = (A + A^T)x + b \ ,$$

Hessian 矩阵:

$$\nabla^2 f(x) = A + A^T \ .$$

6. 第六题

设

$$f(x) = x^T A x, \qquad A = \begin{pmatrix} 4 & 5 \\ 6 & 7 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

首先计算

$$f(x_0) = x_0^T A x_0 = 22.$$

梯度与 Hessian 为

$$\nabla f(x) = (A + A^T)x, \qquad \nabla^2 f(x) = A + A^T = \begin{pmatrix} 8 & 11 \\ 11 & 14 \end{pmatrix}.$$

因此二阶泰勒展开为

$$f(x) \approx 22 + \begin{pmatrix} 19 \\ 25 \end{pmatrix}^T (x - x_0) + \frac{1}{2}(x - x_0)^T \begin{pmatrix} 8 & 11 \\ 11 & 14 \end{pmatrix} (x - x_0).$$

7. 第七题

计算

$$\frac{d}{dx}\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] = -\frac{1}{2\sigma^2} \cdot 2(x-\mu) = -\frac{x-\mu}{\sigma^2}.$$

因此，有

$$f'(x) = f(x) \cdot \left(-\frac{x-\mu}{\sigma^2}\right) = -\frac{x-\mu}{\sigma^2} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right).$$