T1

已知 $X \in \mathbb{R}^{n\times n}$, $a, b \in \mathbb{R}^n$, 计算下列偏导数

$$\frac{\partial a^T X b}{\partial X}, \quad \frac{\partial a^T X^T b}{\partial X}$$

由于

$$a^T X b = \langle a, Xb\rangle = \sum_{i=1}^n a_i (Xb)_i = \sum_{i=1}^n a_i \sum_{j=1}^n X_{ij} b_j = \sum_{i=1}^n \sum_{j=1}^n a_i b_j X_{ij}$$

所以

$$\frac{\partial a^T X b}{\partial X} = \left(a_i b_j\right)_{i,j} = ab^T$$

同理,

$$a^T X^T b = \langle a, X^T b\rangle = \sum_{i=1}^n a_i \left(X^T b\right)_i = \sum_{i=1}^n a_i \sum_{j=1}^n X_{ji} b_j = \sum_{i=1}^n \sum_{j=1}^n a_j b_i X_{ij}$$

所以

$$\frac{\partial a^T X^T b}{\partial X} = \left(a_j b_i\right)_{i,j} = ba^T$$

.

T2

设 K 是一个锥, K^* 是其对偶锥 (i.e. $K^* = \{y : \langle x, y\rangle \geq 0, \forall x \in K\}$), 证明: K^* 也是一个锥 (即使 K 不是锥也成立).

设 $y_1, y_2 \in K^*$, $\alpha, \beta \geq 0$, 则对于任意 $x \in K$, 有

$$\langle x, \alpha y_1 + \beta y_2\rangle = \alpha\langle x, y_1\rangle + \beta\langle x, y_2\rangle \geq 0$$

所以 $\alpha y_1 + \beta y_2 \in K^*$, 因此 K^* 是一个锥.

T2

利用凸函数的二阶条件证明: log-sum-exp 函数 $f(x) = \log\left(\sum_{k=1}^{n} e^{x_k}\right)$ 是凸函数.

计算 $f(x)$ 的 Hessian 矩阵:

$$
\begin{aligned}
\left(\nabla^2 f(x)\right)_{ij} &= \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \\
&= \frac{\partial}{\partial x_j}\left(\frac{e^{x_i}}{\sum_{k=1}^{n} e^{x_k}}\right) \\
&= \frac{\delta_{ij} e^{x_i}\left(\sum_{k=1}^{n} e^{x_k}\right) - e^{x_i} e^{x_j}}{\left(\sum_{k=1}^{n} e^{x_k}\right)^2} \\
&= \frac{\delta_{ij}\left(\sum_{k=1}^{n} e^{x_k}\right) - e^{x_j}}{\left(\sum_{k=1}^{n} e^{x_k}\right)^2} e^{x_i}
\end{aligned}
$$

设 $S = \sum_{i=1}^{n} e^{x_i}$, 则对于任意 $z \in \mathbb{R}^n$, 有

$$
\begin{aligned}
\langle z, \nabla^2 f(x) z\rangle &= \sum_{i=1}^{n}\sum_{j=1}^{n} z_i z_j \left(\nabla^2 f(x)\right)_{ij} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} z_i z_j \frac{\delta_{ij} S - e^{x_j}}{S^2} e^{x_i} \\
&= \frac{1}{S^2}\sum_{i=1}^{n} z_i e^{x_i}\left[z_i S - \sum_{j=1}^{n} z_j e^{x_j}\right] \\
&= \frac{1}{S^2}\left[S\sum_{i=1}^{n} z_i^2 e^{x_i} - \left(\sum_{i=1}^{n} z_i e^{x_i}\right)^2\right]
\end{aligned}
$$

要证明 $\langle z, \nabla^2 f(x) z\rangle \geq 0$, 只需证明

$$
S\sum_{i=1}^{n} z_i^2 e^{x_i} \geq \left(\sum_{i=1}^{n} z_i e^{x_i}\right)^2
$$

利用 Cauchy 不等式, 有

$$\begin{aligned}\left(\sum_{i=1}^{n} z_i e^{x_i}\right)^2 &= \left(\sum_{i=1}^{n} z_i \sqrt{e^{x_i}} \sqrt{e^{x_i}}\right)^2 \\ &\leq \left(\sum_{i=1}^{n} z_i^2 e^{x_i}\right)\left(\sum_{i=1}^{n} e^{x_i}\right) = S \sum_{i=1}^{n} z_i^2 e^{x_i}\end{aligned}$$

因此 $\nabla^2 f(x)$ 是半正定的, 所以 $f(x)$ 是凸函数.