T1

计算负熵的共轭函数

$$\mathcal{H}(\boldsymbol{x}) = \sum_{i=1}^{n} x_i \log x_i, \quad \boldsymbol{x} \in \mathbb{R}^n$$

根据 Fenchel 共轭的定义, 我们有

$$\mathcal{H}^*(\boldsymbol{y}) = \sup_{\boldsymbol{x} \in \text{ dom } \mathcal{H}} \{\langle \boldsymbol{y}, \boldsymbol{x} \rangle - \mathcal{H}(\boldsymbol{x})\}$$

考虑广义函数的情形: $\mathcal{H} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, 那么其定义域就是 dom $\mathcal{H} = \{\boldsymbol{x} \in \mathbb{R}^n : x_i \geq 0, i = 1, 2, ..., n\}$, 这里我们约定 $0 \log 0 = 0$.

考虑到我们要优化的目标函数

$$\langle \boldsymbol{y}, \boldsymbol{x} \rangle - \mathcal{H}(\boldsymbol{x}) = \sum_{i=1}^{n} (y_i x_i - x_i \log x_i)$$

是关于 $\boldsymbol{x}$ 的可分函数, 因此我们可以将其拆分为 n 个独立的子问题:

$$\sup_{x_i \geq 0} \{y_i x_i - x_i \log x_i\}, \quad i = 1, 2, ..., n$$

对 x_i 求导并令导数为零, 我们得到

$$\frac{\mathrm{d}}{\mathrm{d}x_i}(y_i x_i - x_i \log x_i) = y_i - \log x_i - 1 = 0 \Rightarrow x_i = \exp(y_i - 1)$$

由于 $x_i \geq 0$ 恒成立, 因此我们不需要考虑约束条件. 将 $x_i = \exp(y_i - 1)$ 代入目标函数, 我们得到

$$\mathcal{H}^*(\boldsymbol{y}) = \sum_{i=1}^{n} \exp(y_i - 1).$$

T2

证明: 任意多个凸集的交还是凸集.

回忆凸集的定义: 设 $\mathcal{C}$ 是一个实向量空间 V 的子集, 如果对于任意的 $x_1, x_2 \in \mathcal{C}$ 和任意的 $\theta \in [0,1]$, 都有 $\theta x_1 + (1-\theta)x_2 \in C$, 则称 $\mathcal{C}$ 是凸集.

现在假设 $\mathcal{C}$ 和 $\mathcal{D}$ 都为凸集, 考虑 $x_1, x_2 \in \mathcal{C} \cap \mathcal{D}$:

1. 如果 $x_1, x_2 \in \mathcal{C}$, 则根据 $\mathcal{C}$ 的凸性, 对于任意的 $\theta \in [0,1]$, 都有 $\theta x_1 + (1-\theta)x_2 \in \mathcal{C}$.
2. 同理, 如果 $x_1, x_2 \in \mathcal{D}$, 则根据 $\mathcal{D}$ 的凸性, 对于任意的 $\theta \in [0,1]$, 都有 $\theta x_1 + (1-\theta)x_2 \in \mathcal{D}$.
3. 结合 (1) 和 (2), 由于 $x_1, x_2 \in \mathcal{C} \cap \mathcal{D}$, 那么我们既有 $x_1, x_2 \in \mathcal{C}$, 又有 $x_1, x_2 \in \mathcal{D}$, 因此对于任意的 $\theta \in [0,1]$, 都有 $\theta x_1 + (1-\theta)x_2 \in \mathcal{C}$ 且 $\theta x_1 + (1-\theta)x_2 \in \mathcal{D}$, 也就是说 $\theta x_1 + (1-\theta)x_2 \in \mathcal{C} \cap \mathcal{D}$. 这就证明了 $\mathcal{C} \cap \mathcal{D}$ 是凸集.

T3

证明: 仿射集一定是凸集, 但是凸集不一定是仿射集.

1. 因为仿射集要求对于任意的 $x_1, x_2 \in \mathcal{A}$ 和任意的 $\theta \in \mathbb{R}$, 都有 $\theta x_1 + (1-\theta)x_2 \in \mathcal{A}$, 这显然包含了凸集的定义中 $\theta \in [0,1]$ 的情况, 因此仿射集一定是凸集.
2. 但是凸集不一定是仿射集. 例如, 在二维平面上, 单位圆盘 $\mathcal{C} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ 是一个凸集, 但是它不是一个仿射集. 因为对于 $x_1 = (1,0)$ 和 $x_2 = (0,1)$, 当 $\theta = 2$ 时, 我们有 $\theta x_1 + (1-\theta)x_2 = (2,-1)$, 这个点不在单位圆盘内, 因此单位圆盘不是仿射集.

T4

证明: 凸函数 $f(x)$ 的下水平集 $\mathcal{C} = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$ 是凸集, 但是任意下水平集为凸集的函数不一定是凸函数.

1. 设 $x_1, x_2 \in \mathcal{C}$, 则根据下水平集的定义, 我们有 $f(x_1) \leq \alpha$ 且 $f(x_2) \leq \alpha$. 因为 $f(x)$ 是凸函数, 根据凸函数的定义, 对于任意的 $\theta \in [0,1]$, 我们有

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta)f(x_2) \leq \theta\alpha + (1-\theta)\alpha = \alpha$$

因此 $\theta x_1 + (1-\theta)x_2 \in \mathcal{C}$, 这就证明了 $\mathcal{C}$ 是凸集.

2. 但是任意下水平集为凸集的函数不一定是凸函数. 考虑函数 $f(x) = \sqrt{|x|}$, 当 $\alpha \geq 0$ 时, 其下水平集为 $\mathcal{C} = \left\{x \in \mathbb{R} : \sqrt{|x|} \leq \alpha\right\} = [-\alpha^2, \alpha^2]$, 闭区间显然为凸集; 当 $\alpha < 0$ 时, 下水平集为空集, 空集也被认为是凸集. 因此 $f(x)$ 的任意下水平集都是凸集. 但是 $f(x)$ 不是凸函数, 因为对于 $x_1 = 0$ 和 $x_2 = 1$, 当 $\theta = 0.5$ 时, 我们有 $f(0.5x_1 + 0.5x_2) = f(0.5) = \sqrt{0.5} \approx 0.707$, 而 $0.5f(x_1) + 0.5f(x_2) = 0.5 \times 0 + 0.5 \times 1 = 0.5$, 显然 $f(0.5) > 0.5f(x_1) + 0.5f(x_2)$, 因此 $f(x)$ 不是凸函数.

T5

证明: 任意函数的共轭函数都是凸函数.

设 $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, 根据 Fenchel 共轭的定义, 我们有

$$f^*(y) = \sup_{x \in \text{ dom } f} \{\langle y, x\rangle - f(x)\}$$

考虑 $y_1, y_2 \in \mathbb{R}^n$ 和 $\theta \in [0, 1]$, 计算

$$\begin{aligned}
&f^*(\theta y_1 + (1-\theta)y_2) \\
&\qquad = \sup_{x \in \text{ dom } f} \{\langle \theta y_1 + (1-\theta)y_2, x\rangle - f(x)\} \\
&\qquad = \sup_{x \in \text{ dom } f} \{\theta\langle y_1, x\rangle + (1-\theta)\langle y_2, x\rangle - f(x)\} \\
&\qquad = \sup_{x \in \text{ dom } f} \{\theta[\langle y_1, x\rangle - f(x)] + (1-\theta)[\langle y_2, x\rangle - f(x)]\} \\
&\qquad \leq \theta \sup_{x \in \text{ dom } f} \{\langle y_1, x\rangle - f(x)\} + (1-\theta) \sup_{x \in \text{ dom } f} \{\langle y_2, x\rangle - f(x)\} \\
&\qquad = \theta f^*(y_1) + (1-\theta)f^*(y_2)
\end{aligned}$$

因此 $f^*(y)$ 是凸函数.