Tsinghua-Berkeley Shenzhen Institute
Inference and Information
Fall 2017

Homework 7

赵丰 March 15, 2018

- **Acknowledgments:** This coursework referes to wikipedia: `https://en.wikipedia.org`.
- **Collaborators:** I finish this coursework by myself.

I use `enumerate` to generate answers for each question:

7.1. (a) $D(q||p) \geq q(M)\log\frac{q(M)}{p(M)} = \infty$ 所以 $D(q||p) = \infty$

(b) 因为 $\lim\limits_{x\to 0+} x\log(x) = 0$，$D(p||q) = \displaystyle\sum_{y=0}^{M-1} p(y)\log\frac{p(y)}{q(y)} < \infty$

(c) 运用 Lagrange 乘子法可以求出 $p(k) = tq(k), k = 0,\dots,M-1$, 通过归一化条件求出 $t = \frac{1}{Q(M-1)}$, 另外由函数的凸性可得局部最小为全局最小，从而

$$p^*(k) = \frac{q(k)}{Q(M-1)}, k = 0,\dots,M-1 \text{ and } D(p^*||q) = \log\frac{1}{Q(M-1)}$$

另解：

$$\begin{aligned} D(p||q) &= \sum_{y=0}^{M-1} p(y)\log\frac{p(y)}{\frac{q(y)}{Q(M-1)}Q(M-1)} \\ &= \sum_{y=0}^{M-1} D\left(p||\frac{q(y)}{Q(M-1)}1_{y<M}\right) + \log\frac{1}{Q(M-1)} \end{aligned}$$

(d) 对

$$f(p) = \sum_{m=0}^{\infty} p_m\log\frac{p_m}{q_m} - \lambda(\sum_{m=0}^{\infty} p_m - 1) - \mu(\sum_{m=M}^{\infty} p_m - \epsilon)$$

运用 Lagrange 乘子法求出

$$\begin{aligned} p_k^* &= tq_k, m \leq M-1 \\ p_k^* &= t'q_k, m \geq M \end{aligned}$$

由归一化条件和 $\mathcal{P}_\epsilon$ 的性质可得

$$\begin{aligned} t =& \frac{1-\epsilon}{Q(M-1)} \\ t' =& \frac{\epsilon}{1-Q(M-1)} \end{aligned}$$

从而得到 $D(p_\epsilon^*||q) = (1-\epsilon)\log t + \epsilon\log t'$ 并且有：

$$\lim_{\epsilon\to 0+} D(p_\epsilon^*||q) = D(p^*||q)$$

(e) $t(y) = 1(y \geq M), c = \epsilon$

(f) $p_\epsilon^*(y) = q(y)\exp(xt(y) - \alpha(x)) \Rightarrow e^x = \frac{t'}{t}$ 所以

$$x = \log \frac{\epsilon Q(M-1)}{(1-\epsilon)(1-Q(M-1))}$$

7.2. (a) 我们证明 $D(p_{\mathsf{x,y}}||q_{\mathsf{x}}q_{\mathsf{y}}) \geq D(p_{\mathsf{x,y}}||p_{\mathsf{x}}p_{\mathsf{y}})$

$$\begin{aligned}
D(p_{\mathsf{x,y}}||q_{\mathsf{x}}q_{\mathsf{y}}) - D(p_{\mathsf{x,y}}||p_{\mathsf{x}}p_{\mathsf{y}}) &= \sum_{x,y} p_{\mathsf{x,y}}(x,y) \log \frac{p_{\mathsf{x}}(x)p_{\mathsf{y}}(y)}{q_{\mathsf{x}}(x)q_{\mathsf{y}}(y)} \\
&= \sum_{x,y} p_{\mathsf{x,y}}(x,y) \log \frac{p_{\mathsf{x}}(x)}{q_{\mathsf{x}}(x)} + \sum_{\mathsf{x,y}} p_{\mathsf{x,y}}(x,y) \log \frac{p_{\mathsf{y}}(y)}{q_{\mathsf{y}}(y)} \\
&= \sum_{x} p_{\mathsf{x}}(x) \log \frac{p_{\mathsf{x}}(x)}{q_{\mathsf{x}}(x)} + \sum_{y} p_{\mathsf{y}}(y) \log \frac{p_{\mathsf{y}}(y)}{q_{\mathsf{y}}(y)} \\
&= D(p_{\mathsf{x}}||q_{\mathsf{x}}) + D(p_{\mathsf{y}}||q_{\mathsf{y}}) \geq 0
\end{aligned}$$

当 $q_x = p_x, q_y = p_y$ 时取等号，因此 $\min D(p_{\mathsf{x,y}}||q_{\mathsf{x}}q_{\mathsf{y}}) = I(\mathsf{x},\mathsf{y})$

(b) 注意到 $p = 0 \Rightarrow q = 0$, 否则 $D(q||p_{\mathsf{x,y}}) = \infty$。所以 $q_{\mathsf{x}}(4)q_{\mathsf{y}}(k) = 0, k = 1,2,3$, 若 $q_{\mathsf{x}}(4) \neq 0$, 则 $q_{\mathsf{y}} = (0,0,0,1), q_{\mathsf{x}} = (0,0,0,1) \Rightarrow D(q||p_{\mathsf{x,y}}) = \log 4$; 若 $q_{\mathsf{x}}(4) = 0 \Rightarrow q_{\mathsf{y}}(4) = 0$, 因为 $q_{\mathsf{x}}(3)q_{\mathsf{y}}(k) = 0, k = 1,2$，若 $q_{\mathsf{x}}(3) \neq 0$，则 $q_{\mathsf{y}} = (0,0,1,0), q_{\mathsf{x}} = (0,0,1,0) \Rightarrow D(q||p_{\mathsf{x,y}}) = \log 4$; 若 $q_{\mathsf{x}}(3) = 0 \Rightarrow q_{\mathsf{y}}(3) = 0$。为记号简便，设 $q_{\mathsf{x}}(1) = a, q_{\mathsf{y}}(1) = b$，则只需极小化下面的函数

$$f(a,b) = ab\log(8ab) + (1-a)(1-b)\log(8(1-a)(1-b)) + a(1-b)\log(8a(1-b)) + b(1-a)\log(8b(1-a))$$

化简得：

$$f(a,b) = \log 8 + a\log a + (1-a)\log(1-a) + b\log b + (1-b)\log(1-b)$$

当 $a = b = \frac{1}{2}$ 时，$f(a,b)$ 取得最小值 $\log 2$ 而 (a) 中的解使得 $D(q||p_{\mathsf{x,y}}) = \infty$。