作业 4

已知

Data

$$D = (x_i, y_i)_{i=1}^n, \quad x_i \in \mathbb{R}^d, \quad y_i \in 0,1$$

Model

$$P(x, y; \theta) = P(y; \theta)P(x|y; \theta)$$

用 0-1 分布来建模 y

y	0	1
$P$$	$1-p$	p

因此有:

$$P(y) = p^y(1-p)^{1-y}, \quad 0 < p < 1$$

题目

设x_i是 d维空间中的一个点，如果 x_i是离散的，我们可以做一个很强的假设，各个维度之间相互独立，每个维度的取值为 $0,1$。即 $x_{ij} \in 0,1$，此时可以用 0-1 分布来建模 x_i。假设 u_0, u_1为 d 维离散型随机向量，即:

$$u_i = \left(u_{i1} \; u_{i2} \; \vdots \; u_{id} \right), \quad i = 0,1$$

则有:

$$\begin{gathered} P(x_{ij}|y=0) = u_{0j}^{x_{ij}}(1-u_{0j})^{1-x_{ij}} \\ P(x_{ij}|y=1) = u_{1j}^{x_{ij}}(1-u_{1j})^{1-x_{ij}} \\ i = 1,2,\ldots,n; \quad j = 1,2,\ldots,d \end{gathered}$$

即:

$$P(x_i|y=0) = \prod_{j=1}^d u_{0j}^{x_{ij}}(1-u_{0j})^{1-x_{ij}}$$

$$P(x_i|y=1) = \prod_{j=1}^d u_{1j}^{x_{ij}}(1-u_{1j})^{1-x_{ij}}$$

任务：利用极大似然估计（MLE）求参数$\theta = p, u_0, u_1$

提示:

$$\ln P(D) = \ln \left(\prod_{i=1}^{n} P(x_i, y_i; \theta) \right) = \sum_{i=1}^{n} \ln P(x_i, y_i; \theta) = \sum_{i=1}^{n} \ln \big(P(y_i; \theta) P(x_i | y_i; \theta) \big)$$

定义:

$$L_D(\theta) \triangleq \ln P(D)$$

其中 $y_i = 0$或 1。

记:

$$n_1 = \sum_{i=1}^{n} y_i, \quad n_0 = \sum_{i=1}^{n} (1 - y_i)$$

求解

$$\begin{gathered}
\nabla_p L_D(\theta) = 0 \quad \Rightarrow \quad \frac{n_1}{p} - \frac{n_0}{1-p} = 0 \quad \Rightarrow \quad p = ? \\
\nabla_{\mu_0} L_D(\theta) = 0 \quad \Rightarrow \quad \mu_0 = ? \\
\nabla_{\mu_1} L_D(\theta) = 0 \quad \Rightarrow \quad \mu_1 = ?
\end{gathered}$$