2013年度　後期　情報論的学習理論/情報論的機械学習　レポート問題
2014年1月14日

以下の8問のうち3問を選択して答えよ。
レポート期限：2/18，提出先：山西のポスト

問題 1) dを与えられた正整数として、$\mathcal{X}_i$を$\mathcal{X}_i = m_i$ $(i = 1, \ldots, d)$とする有限集合とし、d次元ベクトル$\mathbf{x} = (x^{(1)}, \ldots x^{(d)}) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_d$の同時分布の分解を考える。例えば、

$$P(x^{(1)}, x^{(2)}, x^{(3)}) = P(x^{(2)}|x^{(1)})P(x^{(3)}|x^{(1)})P(x^{(1)})$$

などがある。このような分解は有向グラフで表現できる。これを**ベイジアンネットワーク**と呼ぶ。一般には、G_iを$x^{(i)}$につながる親ノード集合として、

$$P(x^{(1)}, \ldots x^{(d)}) = \prod_{i=1}^{d} P(x^{(i)}|G_i)$$

の形で表現できる。G_iの与え方がベイジアンネットワークの構造に依存する。ここで、パラメータは

$$\theta_{ijk} = P(x^{(i)} = k|G_i\text{のパタン} = j)$$
$$(i = 1, \ldots, d,\ j = 1, \ldots, G_i\text{のパタンの数}, k = 1, \ldots, m_i)$$

をパラメータとして持つ。このとき、n個の独立なデータ列$\mathbf{x}^n = \mathbf{x}_1 \ldots \mathbf{x}_n$からベイジアンネットワークのグラフ構造$G$とパラメータ$\{\theta_{ijk}\}$を確率的コンプレキシティに基づいて推定する方式を与えよ。

問題 2) kを与えられた正整数として、$\mathcal{X} = \{0, 1, \ldots, k\}$を有限アルファベットとする．$P(X = i; \theta) = \theta_i$ $(i = 0, \ldots, k)$, $\sum_{i=0}^{k} \theta_i = 1$と書き、$\theta = (\theta_0, \ldots, \theta_k)$とおく．$\Theta_k$をそのような$\theta$の全体とする．このとき、$\mathcal{P}_k = \{P(X; \theta) : \theta \in \Theta_k\}$を$k$**次元離散分布**のクラスと呼ぶ.

Chap. 3 で
1次元で示した内容を
多次元に！

一方で、モデルクラス$\mathcal{P} = \{P(X; \theta)\}$を用いる**逐次的正規化最尤予測アルゴリズム**とは、データ列$x_1, x_2, ..$が逐次的に与えられる下で、各時刻tに対して、逐次型正規化最尤予測分布：

$$P(X|x^{t-1}) = \frac{P(X \cdot x^{t-1}; \hat{\theta}(X \cdot x^{t-1}))}{\sum_{X'} P(X' \cdot x^{t-1}; \hat{\theta}(X' \cdot x^{t-1}))}$$