

第 1 章 绪论

第 2 章 概率分布

第 3 章 线性回归模型

第 4 章 线性分类模型

第 5 章 神经网络

第 6 章 核方法

第 7 章 稀疏核机器

第 8 章 图模型

8.0.1 线性高斯模型

在前面的章节中，我们看到如何通过有向无环图对于一系列离散变量构建其联合概率分布。现在我们来展示多元高斯分布是如何表示为有向图的，而且该有向图的各个节点对应的是一个变量的线性高斯模型。这使得我们可以通过一般的高斯分布和协方差矩阵为对角矩阵的高斯分布构建不同的分布。概率主成分分析 (probabilistic principal component analysis)，因子分析 (factor analysis) 和线性动态系统 (linear dynamical system) 是比较常用的线性高斯模型 (Roweis and Ghahramani, 1999)。我们会在后续的章节中广泛应用本节中的一些结论。假设现在有一个包含有 D 个变量的有向无环图，其节点 i 表示一个一元连续随机变量 x_i ，且这个随机变量服从高斯分布。这个分布的均值被设置为节点 i 的父节点 pa_i 状态的线性组合

$$p(x_i|\text{pa}_i) = \mathcal{N}\left(x_i \mid \sum_{j \in \text{pa}_i} w_{ij}x_j + b_i, v_i\right) \quad (8.1)$$

其中 w_{ij} 和 b_i 为控制均值的参数， v_i 为 x_i 条件分布的方差。这个联合分布的对数可以写成图中所有节点所对应的条件分布乘积的对数，于是可以写成

$$\ln p(\mathbf{x}) = \sum_{i=1}^D \ln p(x_i|\text{pa}_i) \quad (8.2)$$

$$= - \sum_{i=1}^D \frac{1}{2v_i} \left(x_i - \sum_{j \in \text{pa}_i} w_{ij}x_j - b_j \right)^2 + \text{const} \quad (8.3)$$

其中 $\mathbf{x} = (x_1, \dots, x_D)^T$ ，const 项表示与 \mathbf{x} 相互独立的项。很明显这是一个关于 \mathbf{x} 的二次函数，所以联合分布 $p(\mathbf{x})$ 为多元高斯分布。

我们可以通过如下的递归方法确定联合分布的均值和协方差。对于每个变量 x_i ，它们都具有自己对应的条件高斯分布 (8.11)，当然是给定其父节点状态的条件下的分布，所

以

$$x_i = \sum_{j \in \text{pa}_i} w_{ij} x_j + b_i + \sqrt{v_i} \epsilon_i \quad (8.4)$$

其中 ϵ_i 表示的是均值为 0，方差为 1 的高斯随机变量，满足 $\mathbb{E}[\epsilon_i] = 0$ 和 $\mathbb{E}[\epsilon_i \epsilon_j] = I_{ij}$ ，其中 I_{ij} 表示单位矩阵中第 i 行，第 j 列的元素。对 (8.14) 求期望，可以得到

$$\mathbb{E}[x_i] = \sum_{j \in \text{pa}_i} w_{ij} \mathbb{E}[x_j] + b_i \quad (8.5)$$

这样一来