高斯网络

高斯图模型（高斯网络）是一种随机变量为连续的有向或者无向图。有向图版本的高斯图是高斯贝叶斯网络，无向版本的叫高斯马尔可夫网络。

高斯网络的每一个节点都是高斯分布：$\mathcal{N}(\mu_i,\Sigma_i)$，于是所有节点的联合分布就是一个高斯分布，均值为 μ，方差为 Σ。

对于边缘概率，我们有下面三个结论：

1. 对于方差矩阵，可以得到独立性条件：$x_i\perp x_j\Leftrightarrow\sigma_{ij}=0$，这个叫做全局独立性。
2. 我们看方差矩阵的逆（精度矩阵或信息矩阵）：$\Lambda=\Sigma^{-1}=(\lambda_{ij})_{pp}$，有定理：

   > $x_i\perp x_j|(X-\{x_i,x_j\})\Leftrightarrow\lambda_{ij}=0$

   因此，我们使用精度矩阵来表示条件独立性。
3. 对于任意一个无向图中的节点 x_i，$x_i|(X-x_i)\sim\mathcal{N}(\sum\limits_{j\ne i}\frac{\lambda_{ij}}{\lambda_{ii}}x_j,\lambda_{ii}^{-1})$

   也就是其他所有分量的线性组合，即所有与它有链接的分量的线性组合。

高斯贝叶斯网络 GBN

高斯贝叶斯网络可以看成是 LDS 的一个推广，LDS 的假设是相邻时刻的变量之间的依赖关系，因此是一个局域模型，而高斯贝叶斯网络，每一个节点的父亲节点不一定只有一个，因此可以看成是一个全局的模型。根据有向图的因子分解：

$$
p(x)=\prod\limits_{i=1}^pp(x_i|x_{Parents(i)}) \tag{1}
$$

对里面每一项，假设每一个特征是一维的，可以写成线性组合：

$$
p(x_i|x_{Parents(i)})=\mathcal{N}(x_i|\mu_i+W_i^Tx_{Parents(i)},\sigma^2_i) \tag{2}
$$

将随机变量写成：

$$
x_i=\mu_i+\sum\limits_{j\in x_{Parents(i)}}w_{ij}(x_j-\mu_j)+\sigma_i\varepsilon_i,\varepsilon_i\sim\mathcal{N}(0,1) \tag{3}
$$

写成矩阵形式，并且对 w 进行扩展：

$$
x-\mu=W(x-\mu)+S\varepsilon \tag{4}
$$

其中，$S=diag(\sigma_i)$。所以有：$x-\mu=(\mathbb{I}-W)^{-1}S\varepsilon$

由于：

$$
Cov(x)=Cov(x-\mu) \tag{5}
$$

可以得到协方差矩阵。

高斯马尔可夫网络 GMN

对于无向图版本的高斯网络，可以写成：

$$p(x)=\frac{1}{Z}\prod_{i=1}^{p}\phi_i(x_i)\prod_{i,j\in X}\phi_{i,j}(x_i,x_j) \tag{6}$$

为了将高斯分布和这个式子结合，我们写出高斯分布和变量相关的部分：

$$\begin{aligned}p(x)&\propto\exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))\\&=\exp(-\frac{1}{2}(x^T\Lambda x-2\mu^T\Lambda x+\mu^T\Lambda\mu))\\&=\exp(-\frac{1}{2}x^T\Lambda x+(\Lambda\mu)^Tx)\end{aligned} \tag{7}$$

可以看到，这个式子与无向图分解中的两个部分对应，我们记 $h=\Lambda\mu$为 Potential Vector。其中和 x_i 相关的为：$x_i:-\frac{1}{2}\lambda_{ii}x_i^2+h_ix_i$，与 x_i,x_j 相关的是：$x_i,x_j:-\lambda_{ij}x_ix_j$，这里利用了精度矩阵为对称矩阵的性质。我们看到，这里也可以看出，x_i,x_j 构成的一个势函数，只和 λ_{ij} 有关，于是 $x_i\perp x_j|(X-\{x_i,x_j\})\Leftrightarrow\lambda_{ij}=0$。