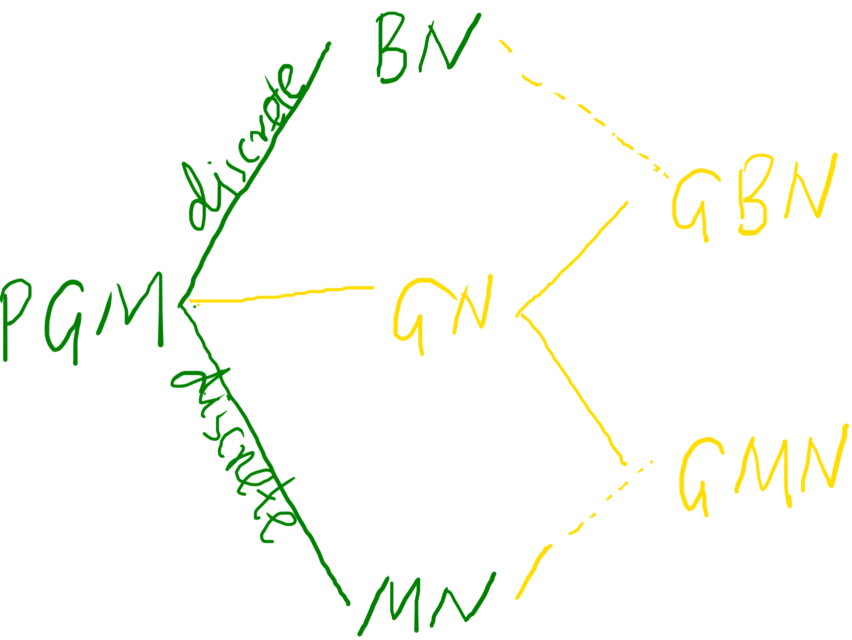
Gaussian Network
高斯网络（高斯图模型）



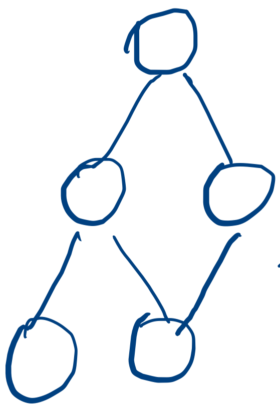
PGM: probabilistic graphical model

BN: bayesian network

MN: markov network

GN: Gaussian Network

Gaussian Network



$\Longleftrightarrow$

$x_i \sim N(\mu_i, \Sigma_i)$

$x = (x_1, x_2, \cdots, x_p)^T$

$$p(x) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}} \cdot \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

$\rightarrow \mu, \Sigma$

$$\Sigma = (\sigma_{ij}) = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \vdots & & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix}_{p\times p}$$

$x_i \perp x_j \Longleftrightarrow \sigma_{ij} = 0$

条件独立性：$x_A \perp x_B \mid x_C$

$$\Lambda = \Sigma^{-1} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1p} \\ \vdots & & & \vdots \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pp} \end{pmatrix}_{p\times p}$$

$x_i \perp x_j \mid -\{x_i, x_j\} \Longleftrightarrow \lambda_{ij} = 0$

precision matrix $\longrightarrow$ information matrix

Gaussian Network

连续型的PGM

↳ 有向：GBN

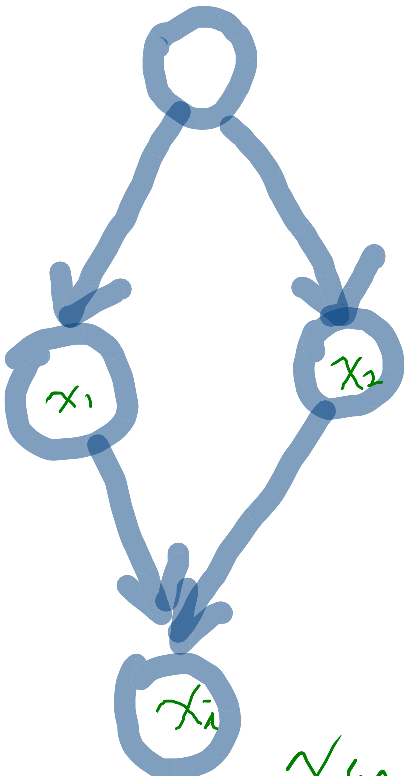
$$P(x)=\prod_{i=1}^{p} P(x_i|x_{pa(i)}) \longrightarrow \text{BN的因子分解}$$

$x_{pa(i)}$ ↳ 一个集合(父亲)

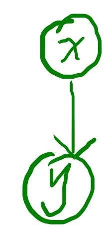
GBN is based on Linear Gaussian Model

GBN → global model

Linear Gaussian Model → local model



$$\begin{cases} P(x)=N(x|\mu_x,\Sigma_x) \\ P(y|x)=N(y|Ax+b,\Sigma_y) \end{cases}$$

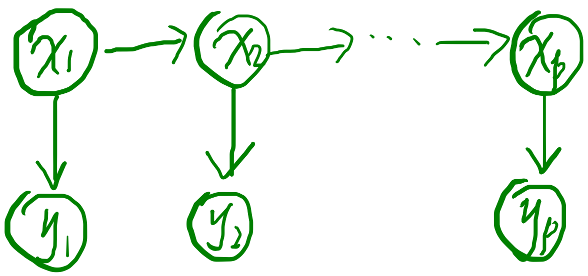


$x_{pa(i)} \to set=\{x_1,x_2\}$

Kalman Filter (HMM)

$\lambda = (\pi, A, B)$

$P(x_t \mid x_{t-1}) \quad P(y_t \mid x_t)$



$$\begin{cases} P(x_t \mid x_{t-1}) = N(x_t \mid A x_{t-1} + B,\ Q) \\ P(y_t \mid x_t) = N(y_t \mid C x_t + D,\ R) \end{cases}$$

$$\begin{cases} x_t = A x_{t-1} + B + \varepsilon, & \varepsilon \sim N(0, Q) \\ y_t = C x_t + D + \delta, & \delta \sim N(0, R) \end{cases}$$

GBN

Gaussian Bayesian Network

$x = (x_1, x_2, \cdots, x_p)^T$

$x \sim N(\mu, \Sigma)$ $\quad P(x) = \prod_{i=1}^{p} P(x_i \mid x_{pa(i)})$ $\quad x_{pa(i)} = (x_1, x_2, \cdots, x_k)^T$

$P(x_i \mid x_{pa(i)}) = N(x_i \mid \mu_i + W_i^T x_{pa(i)}, \sigma_i^2)$

x_i 是一维的

$x_i = \mu_i + \sum_{j \in x_{pa(i)}} W_{ij} \cdot (x_j - \mu_j) + \sigma_i \cdot \varepsilon_i$

ε_i is r.v.

$\mu = (\mu_1, \mu_2, \cdots, \mu_p)^T$

$x_i - \mu_i = \sum_{j \in x_{pa(i)}} W_{ij} (x_j - \mu_j) + \sigma_i \varepsilon_i$ $\quad \varepsilon_i \sim N(0,1)$

$W = [W_{ij}]$

$\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_p)^T$

$S = diag(\sigma_i)$

$x - \mu = W \cdot (x - \mu) + S \cdot \varepsilon$

$(I - W) \cdot (x - \mu) = S \cdot \varepsilon$

$x - \mu = (I - W)^{-1} S \cdot \varepsilon$

$\Sigma = cov(x) = cov(x - \mu) = cov((I - W)^{-1} S \cdot \varepsilon)$

Learning: 结构, 参数

Gaussian Markov Network

Λ: precision matrix

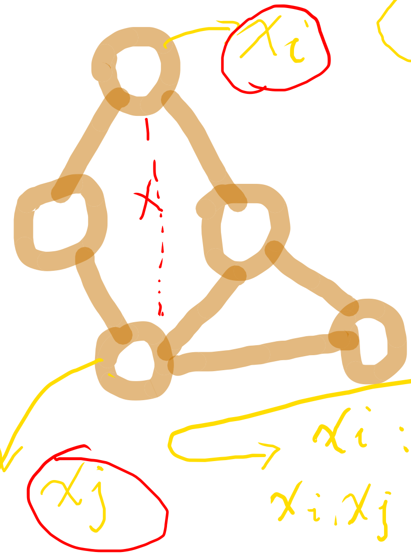
$\Lambda\mu$: potential vector $\to h = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_p \end{pmatrix}$

$$P(x) = \frac{1}{(2\pi)^{\frac{p}{2}} \cdot |\Sigma|^{\frac{1}{2}}} \cdot \exp\left\{-\frac{1}{2}(x-\mu)^T \cdot \Sigma^{-1}(x-\mu)\right\} \to \text{pdf}$$

$$P(x) = \frac{1}{Z} \prod_{i=1}^{p} \underbrace{\psi_i(x_i)}_{\text{node potential}} \cdot \prod_{i,j \in X} \underbrace{\psi_{i,j}(x_i, x_j)}_{\text{edge potential}}$$

$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ $\quad \Lambda = (\lambda_{ij})_{p\times p}$

$\Lambda = \Sigma^{-1} \to$ 对称 $\Rightarrow \Lambda^T = \Lambda$



$$P(x) \propto \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\}$$
$$= \exp\left\{-\frac{1}{2}(x^T\Lambda - \mu^T\Lambda)(x-\mu)\right\}$$
$$= \exp\left\{-\frac{1}{2}(x^T\Lambda x - x^T\Lambda\mu - \mu^T\Lambda x + \mu^T\Lambda\mu)\right\}$$
$$= \exp\left\{-\frac{1}{2}(x^T\Lambda x - 2\mu^T\Lambda x + \mu^T\Lambda\mu)\right\}$$
$$\propto \exp\left\{\underbrace{-\frac{1}{2}x^T\Lambda x}_{\text{二次}} + \underbrace{(\Lambda\mu)^T x}_{\text{一次}}\right\}$$

x_i: $-\frac{1}{2}x_i^2 \cdot \lambda_{ii} + h_i x_i$

x_i, x_j: $-\frac{1}{2}(\lambda_{ij} x_i x_j + \lambda_{ji} x_j x_i) = -\lambda_{ij} x_i x_j$

$x_i \perp x_j \mid -\{i,j\} \iff \lambda_{ij} = 0$

Schur complementary

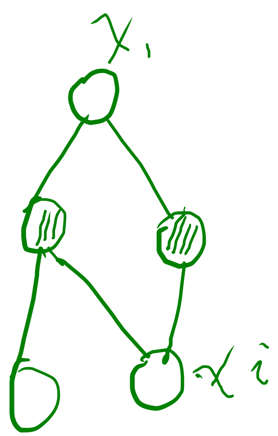
Gaussian Markov Network

① $x_i \perp x_j$, $(\Sigma = (\sigma_{ij}))$ $\Leftrightarrow$ $\sigma_{ij}=0$ $\longrightarrow$ marginal independent

② $x_i \perp x_j \mid -\{x_i, x_j\}$ $(\Lambda = \Sigma^{-1} = (\lambda_{ij}))$ $\Leftrightarrow$ $\lambda_{ij}=0$ $\longrightarrow$ 条件独立

③ $\forall x_i$, $\underbrace{x_i \mid -\{x_i\}}_{\text{条件概率分布}} \sim N\left(\sum_{j\neq i} \frac{\lambda_{ij}}{\lambda_{ii}} x_j, \frac{\lambda_{ii}^{-1}}{\sigma_{ii}}\right)$

x_i 可以看作跟它有连接的 x_j 的线性组合



$\lambda_{1i}=0$

$x = \begin{pmatrix} x_i \\ -\{x_i\} \end{pmatrix} = \begin{pmatrix} x_a \\ x_b \end{pmatrix}$

$\Lambda = \Sigma^{-1} = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}^{-1}$ 分块矩阵的求逆

$(A-BCD)^{-1} \longrightarrow$ woodbury formula