Gaussian Markov Random Fields

Sargur Srihari srihari@cedar.buffalo.edu

Topics

- Information Form of Gaussian
- Induced Pairwise Markov Network
 - Node and Edge potentials
- Pairwise MN as Gaussian

Representing Gaussian using undirected PGMs

- First show how a Gaussian distribution can be viewed as an MRF
- Derived almost immediately from the information form of Gaussian

Information Form of Gaussian

- Covariance Form is: $p(\mathbf{x}) = N(\mu, \Sigma) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp[(\mathbf{x} \mu)^t \Sigma^{-1} (\mathbf{x} \mu)]$
 - Using inverse covariance matrix precision $J=\Sigma^{-1}$ $-\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu) = -\frac{1}{2}(x-\mu)^t J(x-\mu) = -\frac{1}{2}\left[x^t J x 2x^t J \mu + \mu^t J \mu\right]$
 - Since last term is constant, we get information form $p(\mathbf{x}) \alpha \exp \left[-\frac{1}{2}\mathbf{x}^t J \mathbf{x} + (J \mathbf{\mu})^t \mathbf{x}\right]$
- We break-up exponent into two types of terms:
 - Involve single variable X_i , or involve pairs X_i , X_j
 - Terms that involve only the variable X_i are:
 - $\left| -\frac{1}{2}J_{ii}x_i^2 + h_ix_i \right|$ where the potential vector $\hat{h} = J\mu$
 - Terms that involve only the pair X_i , X_j are:

$$-\frac{1}{2} \Big[J_{ij} x_i \, x_j + J_{ji} x_i \, x_j \, \Big] = -J_{ji} x_i \, x_j$$

Induced pairwise Markov network

- Information form of Gaussian induces a pairwise Markov network
- Node potentials are derived from the potential vector h and the diagonal elements J_{ii} of the information matrix
- Edge potentials are derived from the offdiagonal entries of the information matrix
- When $J_{ij}=0$ there is no edge between X_i and X_j in the model
 - Corresponds directly to the independence assumption in the Markov network

Gaussian Markov Random Fields

Follows directly from information form

$$p(\mathbf{x}) \quad \alpha \quad \exp\left[-\frac{1}{2}\mathbf{x}^t J \mathbf{x} + \left(J \mathbf{\mu}\right)^t \mathbf{x}\right]$$

- which is obtained from covariance form with $J=\Sigma^{-1}$

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left[(\mathbf{x} - \boldsymbol{\mu})^t \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right]$$

- Break-up exponent into two types of terms
 - Using the potential vector $h=J\mu$
 - Terms involving single variable X_i

$$\boxed{-\frac{1}{2}J_{i,i}x_i^2 + h_ix_i}$$

Terms involving pairs of variables X_i , X_j

$$\left[-\frac{1}{2}\left[J_{i,j}x_{i}x_{j}+J_{j,i}x_{j}x_{i}\right]=-J_{i,j}x_{i}x_{j}\right]$$

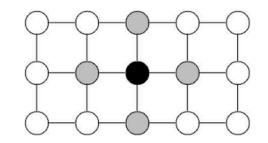
Called edge potentials (when $J_{i,j}=0$, there is no edge)

Converse: pairwise MN as Gaussian

 Consider any pairwise Markov network with quadratic node and edge potentials

We can write node and edge energy functions (log-potentials) as

$$\begin{aligned} & \boxed{ \boldsymbol{\varepsilon}_{i}(\boldsymbol{x}_{i}) = \boldsymbol{d}_{0}^{i} + \boldsymbol{d}_{1}^{i}\boldsymbol{x}_{i} + \boldsymbol{d}_{2}^{i}\boldsymbol{x}_{i}^{2} } \\ & \boldsymbol{\varepsilon}_{i,j}(\boldsymbol{x}_{i}, \boldsymbol{x}_{j}) = \boldsymbol{a}_{00}^{i,j} + \boldsymbol{a}_{01}^{i,j}\boldsymbol{x}_{i} + \boldsymbol{a}_{10}^{i,j}\boldsymbol{x}_{j} + \boldsymbol{a}_{11}^{i,j}\boldsymbol{x}_{i}\boldsymbol{x}_{j} + \boldsymbol{a}_{02}^{i,j}\boldsymbol{x}_{i}^{2} + \boldsymbol{a}_{20}^{i,j}\boldsymbol{x}_{j}^{2} \end{aligned} }$$



Charles

Debbie

Aggregating terms we can reformulate any set of potentials as

$$P(x) = \frac{1}{Z} \exp\left(-\frac{1}{2}x^T Jx + h^T x\right)$$
 where we assume that J is symmetric.

It defines a valid Gaussian iff J is a positive definite matrix If so J is a legal information matrix and we can take h to be a potential vector resulting in a Gaussian distribution in information matrix form

Ex: Gaussian to MN

• Given $p(X_1,X_2)=\mathcal{N}\left(\mu=\left(\begin{array}{cc}1\\3\end{array}\right);\Sigma=\left(\begin{array}{cc}2&1\\1&1\end{array}\right)\right)$

Build MN

$$J = \Sigma^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \qquad \qquad \boxed{\mathbf{X}_1 - \mathbf{X}_2}$$

$$P(X_1, X_2) = \frac{1}{Z} \exp\left[\phi_1(X_1) + \phi_2(X_2) + \phi_{1,2}(X_1, X_2)\right]$$

$$\phi_1(x_1) = -\frac{1}{2}J_{1,1}x_1^2 + J_{1,1}\mu_1x_1 = -\frac{1}{2} \cdot 1 \cdot x_1^2 + 1 \cdot 1 \cdot x_1 = -\frac{1}{2}x_1^2 + x_1$$

$$\phi_2(x_2) = -\frac{1}{2}J_{2,2}x_2^2 + J_{2,2}\mu_2x_2 = -\frac{1}{2} \cdot 2 \cdot x_2^2 + 2 \cdot 3 \cdot x_2 = -x_2^2 + 6x_2$$

$$\phi_{1,2}(x_1, x_2) = -J_{1,2}x_1x_2 = -(-1) \cdot x_1x_2 = x_1x_2$$