Message Passing Formula in Gaussian MRF

Xinhua Zhang

October 15, 2007

We basically explain (and point out a typo in) Eq. 2.15 to 2.20 in

Yair Weiss and William T. Freeman, "*Correctness of Belief Propagation in Gaussian Graphical Models of Arbitrary Topology*", Neural Computation, Vol 13, pp 2173–2200, 2001.

Let x_i be n dimensional and x_j be m dimensional. Then a is $n \times n$, b is $n \times m$, c is $m \times m$, P_{ii} is $n \times n$, V_{ij} is $(n+m) \times (n+m)$, P_{ij} is $m \times m$, P_0 is $n \times n$.

$$m_{ij} \propto \int_{\mathcal{R}^n} \psi_{ij}(x_i, x_j)\, \psi_{ii}(x_i, y_i) \prod_{k \in \mathcal{N}(i) \backslash j} m_{ki}(x_i) dx_i \tag{1}$$

where

$$\psi_{ij}(x_i, x_j) = \frac{1}{(2\pi)^{(n+m)/2} \left(\det V_{ij}^{-1}\right)^{1/2}} \exp\left(-\frac{1}{2}\left(x_i^\top, x_j^\top\right) \begin{pmatrix} a & b \\ b^\top & c \end{pmatrix} \begin{pmatrix} x_i \\ x_j \end{pmatrix}\right)$$

$$\psi_{ii}(x_i, y_i) = \frac{1}{(2\pi)^{n/2} \left(\det P_{ii}^{-1}\right)^{1/2}} \exp\left(-\frac{1}{2}(x_i - y_i)^\top P_{ii}(x_i - y_i)\right)$$

$$m_{ki}(x_i) = \frac{1}{(2\pi)^{n/2} \left(\det P_{ki}^{-1}\right)^{1/2}} \exp\left(-\frac{1}{2}(x_i - \mu_{ki})^\top P_{ki}(x_i - \mu_{ki})\right).$$

So (1) becomes:

$$m_{ij} \propto \int_{\mathcal{R}^n} \exp\left(\frac{-1}{2}\left(x_i^\top, x_j^\top\right) \begin{pmatrix} a & b \\ b^\top & c \end{pmatrix} \begin{pmatrix} x_i \\ x_j \end{pmatrix}\right) \exp\left(\frac{-1}{2}(x_i - y_i)^\top P_{ii}(x_i - y_i)\right)$$
$$\prod_{k \in \mathcal{N}(i) \backslash j} \exp\left(\frac{-1}{2}(x_i - \mu_{ki})^\top P_{ki}(x_i - \mu_{ki})\right) dx_i.$$

Denote the exponent in the integrand as J. Then

$$\begin{aligned} 2J &= -x_i^\top a x_i - 2x_i^\top b x_j - x_j^\top c x_j - x_i^\top P_{ii} x_i - \mu_{ii}^\top P_{ii} \mu_{ii} + 2\mu_{ii}^\top P_{ii} x_i \\ &\quad - \sum_{k \in \mathcal{N}(i) \backslash j} \left(x_i^\top P_{ki} x_i - 2\mu_{ki}^\top P_{ki} x_i + \mu_{ki}^\top P_{ki} \mu_{ki}\right) \\ &= -x_i^\top \left(a + P_{ii} + \sum_{k \in \mathcal{N}(i) \backslash j} P_{ki}\right) x_i + 2\left(-x_j^\top b^\top + \mu_{ii}^\top P_{ii} + \sum_{k \in \mathcal{N}(i) \backslash j} \mu_{ki}^\top P_{ki}\right) x_i \\ &\quad - x_j^\top c x_j + const(x_i, x_j), \end{aligned} \tag{2}$$

where $const\,(x_i, x_j)$ represents a function *independent* of x_i and x_j.

Now if we denote:

$$P_0 := P_{ii} + \sum_{k\in\mathcal{N}(i)\backslash j} P_{ki}, \qquad \mu_0 := P_0^{-1}\left(P_{ii}\mu_{ii} + \sum_{k\in\mathcal{N}(i)\backslash j} P_{ki}\mu_{ki}\right),$$

then (2) becomes:

$$2J = -x_i^\top (a + P_0)\, x_i + 2\left(-x_j^\top b^\top + \mu_0^\top P_0\right) x_i - x_j^\top c x_j + const\,(x_i, x_j) \tag{3}$$

$$= -(x_i - \mu)^\top (a + P_0)(x_i - \mu) + \mu^\top (a + P_0)\, \mu - x_j^\top c x_j + const\,(x_i, x_j), \tag{4}$$

where μ can be determined by comparing the linear terms (in x_i) in (3) and (4):

$$\mu^\top (a + P_0) = -x_j^\top b^\top + \mu_0^\top P_0,$$

which gives

$$\mu = (a + P_0)^{-1} (-b x_j + P_0 \mu_0).$$

So now (4) becomes

$$2J = -(x_i - \mu)^\top (a + P_0)(x_i - \mu) + (-b x_j + P_0\mu_0)^\top (a + P_0)^{-1} (-b x_j + P_0\mu_0)$$
$$- x_j^\top c x_j + const\,(x_i, x_j).$$

We know that

$$\int_{\mathcal{R}^n} -\frac{1}{2}(x_i - \mu)^\top (a + P_0)(x_i - \mu)\, dx_i = const\,(x_j),$$

regardless of the value of μ (which is a function of x_j). So

$$m_{ij} \propto \exp\left(\frac{1}{2}(-b x_j + P_0\mu_0)^\top (a + P_0)^{-1} (-b x_j + P_0\mu_0) - \frac{1}{2} x_j^\top c x_j\right)$$

$$= \exp\left(-\frac{1}{2} x_j^\top \left(c - b^\top (a + P_0)^{-1} b\right) x_j - \mu_0^\top P_0^\top (a + P_0)^{-1} b x_j\right) \tag{5}$$

$$\propto \exp\left(-\frac{1}{2}(x_j - \mu_{ij})^\top \left(c - b^\top (a + P_0)^{-1} b\right)(x_j - \mu_{ij})\right), \tag{6}$$

where μ_{ij} can be found by comparing the linear terms (in x_j) in (5) and (6):

$$\mu_{ij}^\top \left(c - b^\top (a + P_0)^{-1} b\right) = -\mu_0^\top P_0^\top (a + P_0)^{-1} b.$$

So

$$\mu_{ij} = -\left(c - b^\top (a + P_0)^{-1} b\right)^{-1} b^\top (a + P_0)^{-1} P_0 \mu_0.$$

In sum, the inverse covariance matrix of message m_{ij} is

$$P_{ij} \leftarrow c - b^\top (a + P_0)^{-1} b,$$

and the mean is

$$\mu_{ij} \leftarrow -P_{ij}^{-1} b^\top (a + P_0)^{-1} P_0 \mu_0.$$