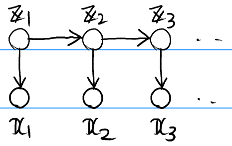
13.19

Webの解答見た



$\mathbf{Z} = \{z_1 \cdots z_N\}$

$\mathbf{X} = \{x_1 \cdots x_N\}$

欲しいのは $\mathbf{X}$ が観測値に固定されたとき $p(\mathbf{Z}, \mathbf{X})$ を最大化する $\mathbf{Z}$ である

これを $\mathbf{Z}^{max}$ とすると

$$\mathbf{Z}^{max} = \arg\max_{\mathbf{Z}} p(\mathbf{Z}, \mathbf{X})$$

である。ここで

$$p(\mathbf{Z}|\mathbf{X}) = \frac{p(\mathbf{Z}, \mathbf{X})}{p(\mathbf{X})}$$

なので、

$$\arg\max_{\mathbf{Z}} p(\mathbf{Z}|\mathbf{X}) = \arg\max_{\mathbf{Z}} \frac{p(\mathbf{Z}, \mathbf{X})}{p(\mathbf{X})} = \arg\max_{\mathbf{Z}} p(\mathbf{Z}, \mathbf{X}) = \mathbf{Z}^{max}$$

である。

またすべてのノードの分布が線形ガウス分布ならば

同時分布 $p(\mathbf{Z}, \mathbf{X})$ もガウス分布になる。(8.1.4節より)

$p(\mathbf{Z}, \mathbf{X})$ がガウス分布ならば $p(\mathbf{Z}|\mathbf{X})$ もガウス分布になる。(2.3.1節より)

$p(\mathbf{Z}|\mathbf{X})$ の平均を $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix}$ とすると

ガウス分布の最大を与える $\mathbf{Z}$ は分布の平均なので

$$\mathbf{Z}^{max} = \begin{pmatrix} z_1^{max} \\ \vdots \\ z_N^{max} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix}$$

となる。

また (2.98) より

$p(\mathbf{Z}|\mathbf{X})$ がガウス分布ならば周辺分布 $p(z_n|\mathbf{X})$ もガウス分布で

$p(z_n|\mathbf{X})$ の平均は μ_n に等しい。ガウス分布は平均で最大となるので

$$\therefore \arg\max_{z_n} p(z_n|\mathbf{X}) = \mu_n$$

である。

よって

$$\underset{\mathbb{Z}}{\operatorname{argmax}}\, p(\mathbb{Z}, \mathbb{X}) = \mathbb{Z}^{max} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix} = \begin{pmatrix} \underset{\mathbb{z}_1}{\operatorname{argmax}}\, p(\mathbb{z}_1 | \mathbb{X}) \\ \vdots \\ \underset{\mathbb{z}_N}{\operatorname{argmax}}\, p(\mathbb{z}_N | \mathbb{X}) \end{pmatrix}$$

を得る。