わかりやすいパターン認識

第12章 ディリクレ過程混合モデルによるクラスタリグ

河野 和平

ディリクレ過程混合モデル

- 通常の混合モデル
  – 予め混合数(クラスタ数 c)を設定
- ディリクレ過程混合モデル
  – ディリクレ過程により決定

Step1. $G(\theta)|\alpha, G_0(\theta) \sim DP\left(\alpha, G_0(\theta)\right)$

Step2. $\theta^k \sim G(\theta) \quad (k = 1, \dots, n)$

Step3. $x_k|\theta \sim p(x|\theta^k) \quad (k = 1, \dots, n)$

CRPに基づくディリクレ過程混合モデル

- クラスタリング応用
  - 各パターンの所属クラスタを表す潜在変数 s_k

Step1. $s_1, \dots, s_n | \alpha \sim CRP(\alpha)$

Step2. $\theta_i | \beta \sim G_0(\theta) \quad (i = 1, \dots, \infty)$

Step3. $x_k | s_k, \theta \sim p(x|\theta_i) \quad (k = 1, \dots, n)$

$p(x|\theta_i)$ →正規分布
θ_i →平均ベクトルμ_i，共分散行列Σ_i

ギブスサンプリング(1)

- 生成モデル

$$p(\mathbf{s},\boldsymbol{\theta}|\mathbf{x}) = \frac{p(\mathbf{x}|\mathbf{s},\boldsymbol{\theta})P(\mathbf{s})p(\boldsymbol{\theta})}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\mathbf{s},\boldsymbol{\theta})P(\mathbf{s})p(\boldsymbol{\theta})}{\sum_{\mathbf{s}}\int p(\mathbf{x}|\mathbf{s},\boldsymbol{\theta})P(\mathbf{s})p(\boldsymbol{\theta})d\boldsymbol{\theta}} \quad (11\cdot4)$$

$$(\hat{\mathbf{s}},\hat{\boldsymbol{\theta}}) = \underset{\mathbf{s},\boldsymbol{\theta}}{\operatorname{argmax}}\{P(\mathbf{s},\boldsymbol{\theta}|\mathbf{x})\} \quad (11\cdot5)$$

$$P(\mathbf{s}|\mathbf{x}) = \frac{p(\mathbf{x}|\mathbf{s})P(\mathbf{s})}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\mathbf{s})P(\mathbf{s})}{\sum_{\mathbf{s}} p(\mathbf{x}|\mathbf{s})P(\mathbf{s})} \quad (11\cdot6)$$

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmax}}\{P(\mathbf{s}|\mathbf{x})\} \quad (11\cdot7)$$

  – 組み合わせ数:膨大

ギブスサンプリング(2)

- n 個のパターンのうち任意の1個をとる。
  - 最後に入力されたパターン x_n とみなす。
  - 所属クラスタを<u>確率的に決定</u>
  - いずれのパターンも最後のパターンとみなせる。
    - n 個すべてのパターンで繰り返す。

所属クラスタの確率的生成

- 事後確率

$$
\begin{aligned}
&P(s_k = \omega_i | \mathbf{x}_k, \mathbf{s}_{-k}, \mathbf{x}_{-k}, \theta) \\
&\quad = \frac{p(\mathbf{x}_k, s_k = \omega_i | \mathbf{s}_{-k}, \mathbf{x}_{-k}, \theta)}{p(\mathbf{x}_k | \mathbf{s}_{-k}, \mathbf{x}_{-k}, \theta)} && (12\cdot7) \\
&\quad \propto p(\mathbf{x}_k, s_k = \omega_i | \mathbf{s}_{-k}, \mathbf{x}_{-k}, \theta) && (12\cdot8) \\
&\quad = p(\mathbf{x}_k | s_k = \omega_i, \mathbf{x}_{-k}, \theta) P(s_k = \omega_i | \mathbf{s}_{-k}, \mathbf{x}_{-k}, \theta) && (12\cdot9) \\
&\quad = p(\mathbf{x}_k | s_k = \omega_i, \boldsymbol{\theta}) P(s_k = \omega_i | \mathbf{s}_{-k}) && (12\cdot10)
\end{aligned}
$$

所属クラスタとパラメータ(1)

- 既存クラスタ: ω_i
- 新規クラスタ: ω_{new}

$$P(s_k = \omega_i | x_k, s_{-k}, x_{-k}, \theta)$$
$$\propto p(x_k | s_k = \omega_i, \theta) P(s_k = \omega_i | s_{-k})$$

$$p(x_k | s_k = \omega_i, \theta) = \begin{cases} p(x_k | \theta_i) \\ \int p(x_k | \theta_{new}) G_0(\theta_{new})\, d\theta_{new} \end{cases}$$

$$p(s_k | s_{-k}) = \begin{cases} \dfrac{n_i'}{n - 1 + \alpha} \\ \dfrac{\alpha}{n - 1 + \alpha} \end{cases}$$

所属クラスタとパラメータ(2)

$$P(s_k|\mathbf{x}_k, \mathbf{s}_{-k}, \mathbf{x}_{-k}, \theta)$$
$$\propto \begin{cases} \dfrac{n'_i}{n-1+\alpha} \cdot p(\mathbf{x}_k|\boldsymbol{\theta}_i) \\ \qquad\qquad (s_k = \omega_i \text{ のとき}) \\ \dfrac{\alpha}{n-1+\alpha} \cdot \displaystyle\int p(\mathbf{x}_k|\boldsymbol{\theta}_{new}) G_0(\boldsymbol{\theta}_{new}) d\boldsymbol{\theta}_{new} \\ \qquad\qquad (s_k = \omega_{new} \text{ のとき}) \end{cases} \tag{12·14}$$

$$\begin{aligned} p\bigl(\boldsymbol{\theta}_i \big| \{\mathbf{x}_k;\ \mathbf{x}_k \in \omega_i\}\bigr) &= \frac{p\bigl(\{\mathbf{x}_k;\ \mathbf{x}_k \in \omega_i\} \big| \boldsymbol{\theta}_i\bigr) \cdot G_0(\boldsymbol{\theta}_i)}{p\bigl(\{\mathbf{x}_k;\ \mathbf{x}_k \in \omega_i\}\bigr)} \\ &= \frac{G_0(\boldsymbol{\theta}_i) \cdot \displaystyle\prod_{\mathbf{x}_k \in \omega_i} p(\mathbf{x}_k|\boldsymbol{\theta}_i)}{\displaystyle\int G_0(\boldsymbol{\theta}_i) \cdot \prod_{\mathbf{x}_k \in \omega_i} p(\mathbf{x}_k|\boldsymbol{\theta}_i)\, d\boldsymbol{\theta}_i} \end{aligned} \tag{12·15}$$

所属クラスタのみ

$$P(s_k|\mathbf{x}_k, \mathbf{s}_{-k}, \mathbf{x}_{-k}) \propto \begin{cases} \dfrac{n'_i}{n-1+\alpha}\displaystyle\int p(\mathbf{x}_k|\boldsymbol{\theta}_i)p(\boldsymbol{\theta}_i|\mathbf{x}_{-k})\,d\boldsymbol{\theta}_i \\ \qquad (s_k = \omega_i \text{ のとき}) \\ \dfrac{\alpha}{n-1+\alpha}\displaystyle\int p(\mathbf{x}_k|\boldsymbol{\theta}_{new})G_0(\boldsymbol{\theta}_{new})\,d\boldsymbol{\theta}_{new} \\ \qquad (s_k = \omega_{new} \text{ のとき}) \end{cases} \tag{12·23}$$

– θ が積分消去･･･崩壊型ギブスサンプリング

実験

- 正規分布

$$\mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}, \Sigma) = \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1})$$
$$= \frac{|\boldsymbol{\Lambda}|^{1/2}}{(2\pi)^{d/2}} \exp\left[-\frac{1}{2}(\mathbf{x}_k - \boldsymbol{\mu})^t \boldsymbol{\Lambda}(\mathbf{x}_k - \boldsymbol{\mu})\right]$$

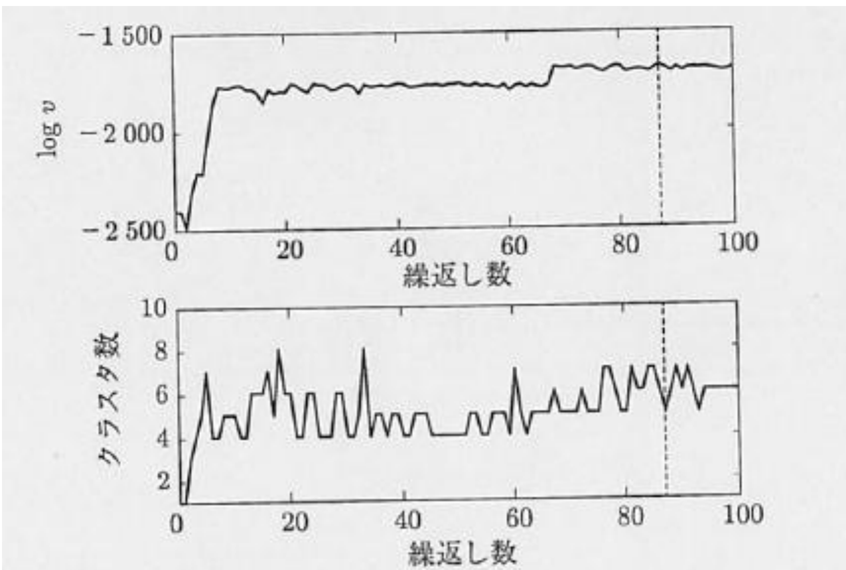
  – 基底分布
    - 精度行列　　，平均ベクトル　　が未知
    - 共役事前分布 : 正規-ウィーシャート分布

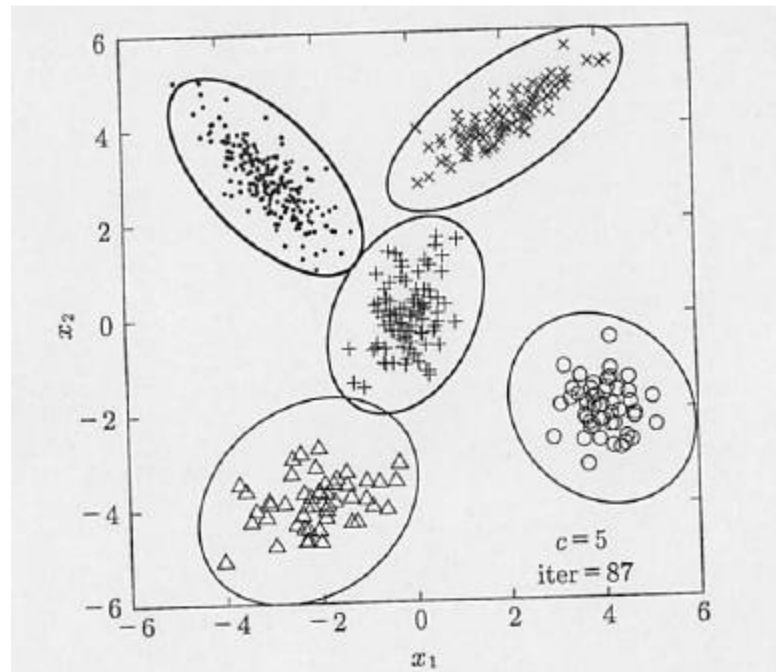
$$G_0(\theta) = G_0(\boldsymbol{\mu}, \boldsymbol{\Lambda}) = \mathcal{N}\left(\boldsymbol{\mu}; \boldsymbol{\mu}_0, (\beta\boldsymbol{\Lambda})^{-1}\right) \cdot \mathcal{W}(\boldsymbol{\Lambda}; \nu, \mathbf{S})$$

$$\alpha = 1.0$$
$$\boldsymbol{\mu}_0 = \boldsymbol{\mu}_{all}, \qquad \beta = 1/3, \qquad \nu = 15, \qquad \mathbf{S} = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$$

実験結果

- 所属クラスタとパラメータ





実験結果

- 所属クラスタのみ