ベイズ本 4.3.2 節

ここでは，例として，ポアソン混合モデルの事後分布からパラメータと潜在変数をサンプルする．導出は手書きの資料に載せているので，ここでは省略する．

■Notation

以下の記号を用いる．

- $\boldsymbol{X} = \{x_1, \ldots, x_N\}$: データ
- $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_K\}$: ポアソン分布のパラメータの集合
- $\boldsymbol{S} = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$: 潜在変数の集合
  - $\boldsymbol{s}_n$: k 次元ベクトル
  - $s_{n,k} = 1 \Longleftrightarrow k$ 番目のクラスタが指定された
- $\boldsymbol{\pi}$: 混合比率
  - $\sum_{k=1}^{K} \pi_k = 1$

■モデル

まず，以下の確率分布を定義する．

- クラスタ k に対する観測モデル:

$$p(x_n|\lambda_k) = \mathrm{Poi}(x_n|\lambda_k), \quad k = 1, \ldots, K. \tag{1}$$

- $\boldsymbol{s}_n$ をサンプルするための分布:

$$p(\boldsymbol{s}_n|\boldsymbol{\pi}) = \mathrm{Cat}(\boldsymbol{s}_n|\boldsymbol{\pi}). \tag{2}$$

- 混合分布における条件付き分布:

$$p(x_n|\boldsymbol{s}_n, \boldsymbol{\lambda}) = \sum_{k=1}^{K} \mathrm{Poi}(x_n|\lambda_k)^{s_{n,k}}. \tag{3}$$

- ポアソン分布のパラメータ $\boldsymbol{\lambda}$ に対する事前分布:

$$p(\lambda_k) = \mathrm{Gam}(\lambda_k|a, b), \quad k = 1, \ldots, K. \tag{4}$$

  - a, b: ハイパーパラメータ．

- 混合比率 $\boldsymbol{\pi}$ に対する事前分布:

$$p(\boldsymbol{\pi}) = \mathrm{Dir}(\boldsymbol{\pi}|\boldsymbol{\alpha}). \tag{5}$$

  - $\boldsymbol{\alpha}$: ハイパーパラメータ．K 次元ベクトル．

これらを用いると，同時分布は

$$\begin{aligned} p(\boldsymbol{X}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{\pi}) &= p(\boldsymbol{X}|\boldsymbol{S}, \boldsymbol{\lambda})p(\boldsymbol{S}|\boldsymbol{\pi})p(\boldsymbol{\lambda})p(\boldsymbol{\pi}) \\ &= \left(\prod_{n=1}^{N} p(x_n|\boldsymbol{s}_n, \boldsymbol{\lambda})p(\boldsymbol{s}_n|\boldsymbol{\pi})\right)\left(\prod_{k=1}^{K} p(\lambda_k)\right)p(\boldsymbol{\pi}) \end{aligned} \tag{6}$$

と表される [1].

■ギブスサンプリング

データ $\boldsymbol{X}$ が観測されたあと，パラメータ $\boldsymbol{\lambda}$, $\boldsymbol{\pi}$ と潜在変数 $\boldsymbol{S}$ をサンプルする．潜在変数とパラメータを分けてサンプルすることを考えると，

$$\boldsymbol{S} \sim p(\boldsymbol{S}|\boldsymbol{X}, \boldsymbol{\lambda}, \boldsymbol{\pi}) \tag{7}$$

$$\boldsymbol{\lambda}, \boldsymbol{\pi} \sim p(\boldsymbol{\lambda}, \boldsymbol{\pi}|\boldsymbol{X}, \boldsymbol{S}) \tag{8}$$

となる．

潜在変数 $\boldsymbol{S}$ について．条件付き独立性が示され，各 $\boldsymbol{s}_n$ に対して，

$$\boldsymbol{s}_n \sim \mathrm{Cat}(\boldsymbol{s}_n|\boldsymbol{\eta}_n), \tag{9}$$

ただし，

$$\eta_{n,k} \propto \exp\left(x_n \ln \lambda_k - \lambda_k + \ln \pi_k\right), \quad \text{s.t.} \quad \sum_{k=1}^{K} \eta_{n,k} = 1 \tag{10}$$

である [1].

$\boldsymbol{\lambda}$, $\boldsymbol{\pi}$ について．$\boldsymbol{X}$, $\boldsymbol{S}$ が与えられた下では，2 つの分布が独立であることが示される．$\boldsymbol{\lambda}$ の分布は，各 k に対して

$$\lambda_k \sim \mathrm{Gam}(\lambda_k|\hat{a}_k, \hat{b}_k), \tag{11}$$

ただし，

$$\hat{a}_k = \sum_{n=1}^{N} s_{n,k} x_n + a \tag{12}$$

$$\hat{b}_k = \sum_{n=1}^{N} s_{n,k} + b \tag{13}$$

である [1].

また，$\boldsymbol{\pi}$ の分布は

$$\boldsymbol{\pi} \sim \mathrm{Dir}(\boldsymbol{\pi}|\hat{\boldsymbol{\alpha}}), \tag{14}$$

ただし，

$$\hat{\alpha}_k = \sum_{n=1}^{N} s_{n,k} + \alpha_k \tag{15}$$

である．

以上を用いると，ポアソン混合モデルの事後分布に対するギブスサンプリングは Algorithm 1 で与えられる．

Algorithm 1 Gibbs Sampling for Poisson mixture model.

Input: MAXITER: the number of iteration, N: the number of data, K: the number of clusters.

Output: $\boldsymbol{S}_{\text{sample}}, \boldsymbol{\lambda}_{\text{sample}}, \boldsymbol{\pi}_{\text{sample}}$: The set of sampled parameters and the set of hidden variables.

Initialisation :

$\boldsymbol{S}_{\text{sample}} \leftarrow \emptyset,\ \boldsymbol{\lambda}_{\text{sample}} \leftarrow \emptyset,\ \boldsymbol{\pi}_{\text{sample}} \leftarrow \emptyset$. Set $\boldsymbol{\lambda}_{\text{init}}$ and $\boldsymbol{\pi}_{\text{init}}$.

LOOP Process

for $i = 1$ to MAXITER **do**
  for $n = 1$ to N **do**
    Sample $\boldsymbol{s}_n$ using (9).
  end for
  $\boldsymbol{S}_i \leftarrow \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$, $\boldsymbol{S}_{\text{sample}} \leftarrow \boldsymbol{S}_{\text{sample}} \cup \{\boldsymbol{S}_i\}$.
  for $k = 1$ to K **do**
    Sample λ_k using (11).
  end for
  $\boldsymbol{\lambda}_i \leftarrow \{\lambda_1, \ldots, \lambda_K\}$, $\boldsymbol{\lambda}_{\text{sample}} \leftarrow \boldsymbol{\lambda}_{\text{sample}} \cup \{\boldsymbol{\lambda}_i\}$.
  Sample $\boldsymbol{\pi}$ using (14).
  $\boldsymbol{\pi}_i \leftarrow \boldsymbol{\pi}$, $\boldsymbol{\pi}_{\text{sample}} \leftarrow \boldsymbol{\pi}_{\text{sample}} \cup \{\boldsymbol{\pi}_i\}$.
end for
return $\boldsymbol{S}_{\text{sample}}, \boldsymbol{\lambda}_{\text{sample}}, \boldsymbol{\pi}_{\text{sample}}$

参考文献

[1] A. Suyama, ベイズ推論による機械学習 入門, 講談社, 2017.