

潜在変数モデルのベイズ推定・ディリクレ分布

樋口さぶろお <https://hig3.net>

龍谷大学大学院 先端理工学研究科 数理・情報科学コース

計算科学特論 L07(2026-06-01 Mon)

最終更新: Time-stamp: "2026-08-07 Fri 19:18 JST hig"

今日の目標

- カテゴリ分布, 多項分布, ディリクレ分布を説明できる
- 潜在変数モデルのベイズ推定ができる



L06-Q1

TA Prob and Sol: マルチレベルモデル

ある学年を対象に行われたテストで、各クラス i の平均点 $\mu^{(i)}$ は $N(\mu_0, 3^2)$ にしたがって、クラス i の各学生 j の点数 $x_j^{(i)}$ は $N(\mu^{(i)}, 2^2)$ にしたがうとする。

- クラス $i = 1$ の 2 名の学生の点数は、55, 59
- クラス $i = 2$ の 2 名の学生の点数は、61, 65

だった。学年平均点 μ_0 の事前分布を $N(50, 10^2)$ とする。
 μ_0 , $\mu^{(i)}$ の事後分布を求めよう。

略解

$$\propto e^{-\frac{(55-\mu^{(1)})^2}{2 \cdot 2^2}} e^{-\frac{(59-\mu^{(1)})^2}{2 \cdot 2^2}} e^{-\frac{(\mu^{(1)}-\mu_0)^2}{2 \cdot 3^2}} e^{-\frac{(61-\mu^{(2)})^2}{2 \cdot 2^2}} e^{-\frac{(65-\mu^{(2)})^2}{2 \cdot 2^2}} e^{-\frac{(\mu^{(2)}-\mu_0)^2}{2 \cdot 3^2}} e^{-\frac{(\mu_0-50)^2}{2 \cdot 10^2}}.$$

3次元正規分布である．MAP 推定すると，

$$2^{-2}(\mu^{(1)} - 55) + 2^{-2}(\mu^{(1)} - 57) + 3^{-2}(\mu^{(1)} - \mu_0) = 0$$

$$2^{-2}(\mu^{(2)} - 61) + 2^{-2}(\mu^{(2)} - 65) + 3^{-2}(\mu^{(2)} - \mu_0) = 0$$

$$3^{-2}(\mu_0 - \mu^{(1)}) + 3^{-2}(\mu_0 - \mu^{(2)}) + 10^{-2}(\mu_0 - 50) = 0$$

第3式から， $\mu_0 = \frac{3^{-2}\mu^{(1)} + 3^{-2}\mu^{(2)} + 10^{-2}50}{3^{-2} + 3^{-2} + 10^{-2}} = \frac{\mu^{(1)} + \mu^{(2)} + \frac{10^{-2}}{3^{-2}}50}{1 + 1 + \frac{10^{-2}}{3^{-2}}}$ ，ちょっと偏つ

た平均．これを第1,2式に代入すると，

$\mu^{(1)}$ は， $\frac{1}{2}(55 + 57)$ に近い値だが， $\mu^{(2)}$ ($\frac{1}{2}(61 + 65)$ に近い値) や， μ_0 に少し吸い寄せられる．

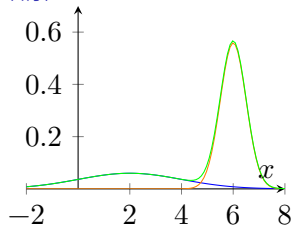
$\mu^{(1)}$ と $\mu^{(2)}$ は近づく．縮小 (shrinkage)．

L06-Q2

TA Prob and Sol: 混合ガウス分布の確率密度関数

X は混合ガウス分布 ($\pi_0 = \frac{3}{10}, \pi_1 = \frac{7}{10}, \mu_0 = 2, \mu_1 = 6, \sigma_0 = 2, \sigma_1 = \frac{1}{2}$) にしたがう. 確率密度関数 $f(x)$ のグラフの概形を描こう. 手で, または, `scipy.stats.norm` の浮動小数点計算で.

略解



L06-Q3

TA Prob and Sol: 混合ガウス分布の確率

X は混合ガウス分布 ($\pi_0 = \frac{3}{10}, \pi_1 = \frac{7}{10}, \mu_0 = 2, \mu_1 = 6, \sigma_0 = 2, \sigma_1 = \frac{1}{2}$) にしたがう.

- ① X の確率密度関数の $x = 4$ における値を, 標準正規分布の確率密度関数 $\phi(x)$ で表そう.
- ② 確率 $P(X \leq 4)$ を, 標準正規分布の累積分布関数 $\Phi(x)$ で表そう.

さらに `scipy.stats.norm` で浮動小数点数で求めてみると安心できるかも.

略解

$$\textcircled{1} \quad f(4) = \frac{3}{10} \frac{1}{\sqrt{2\pi \cdot 2^2}} e^{-\frac{(4-2)^2}{2 \cdot 2^2}} + \frac{7}{10} \frac{1}{\sqrt{2\pi \cdot (1/2)^2}} e^{-\frac{(4-6)^2}{2 \cdot (1/2)^2}} =$$

$$\frac{3}{20} \phi\left(\frac{4-2}{2}\right) + \frac{28}{20} \phi\left(\frac{4-6}{1/2}\right) = \frac{3}{20} \phi(1) + \frac{28}{20} \phi(+4).$$

$$\textcircled{2} \quad P(X \leq 4) = \frac{3}{10} \Phi\left(\frac{4-2}{2}\right) + \frac{7}{10} \Phi\left(\frac{4-6}{1/2}\right) = \frac{3}{10} \Phi(1) + 1 - \frac{7}{10} \Phi(4).$$

$\phi(x) = \phi(-x)$, $\Phi(x) + \Phi(-x) = 1$ なので表示は一意ではない.

ここまで来たよ

6 マルチレベルモデルのベイズ推定・混合ガウス分布

7 潜在変数モデルのベイズ推定・ディリクレ分布

- カテゴリ分布・多項分布・ディリクレ分布
- 潜在変数モデル

カテゴリ変数

量的変数の確率変数 ... 平均や分散に意味がある

- 離散型 確率関数=表 二項分布, ポアソン分布, ..., x は整数
- 連続型 確率密度関数 正規分布, カイ二乗分布, 指数分布, 一様分布, ..., x は実数

質的変数の確率変数 ... 平均や分散が考えられない

その中でも, 名義変数=カテゴリ (カル) 変数

順序や距離がなくぜんぶが対等. 例: 血液型, 性別, 携帯電話番号, チーム A 型, B 型などがカテゴリ

2 カテゴリなら, 0,1 のように番号を振って離散型と思える

3 カテゴリ以上なら, 順序や間隔によるので離散型には帰着できない.

あえてやるなら one-hot vector.

カテゴリ分布と多項分布 (multinomial distribution)

定義 (k カテゴリのカテゴリ分布 $C(p_1, \dots, p_k)$)

カテゴリ番号 $i = 1, \dots, k$ と, パラメタ $p_i \geq 0, p_1 + \dots + p_k = 1$. 略記 $\mathbf{p} = (p_1, \dots, p_k)$ に対して,

$$\text{確率関数 } p(i) = P(X = i) = p_i = \begin{cases} p_1 & (i = 1) \\ p_2 & (i = 2) \\ \vdots & \\ p_k & (i = k) \end{cases}$$

ベルヌイ分布 $B(1, p_1)$ は $k = 2$ の $C(1 - p, p)$.

定義 (k カテゴリの多項分布 $M(n, p_1, \dots, p_k)$)

確率変数は整数値ベクトル $(X_1, \dots, X_k)$. パラメタ $n, p_1, \dots, p_k, n$ ($p_1 + \dots + p_k = 1$).

$$\text{確率関数 } P(X_1 = x_1, \dots, X_k = x_k) = f(x_1, \dots, x_k) = \begin{cases} \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k} & (x_i \geq 0, x_1 + \dots + x_k = n) \\ 0 & (\text{他}) \end{cases}$$

カテゴリ分布にしたがう確率変数の n 個の実現値を得て, カテゴリ $1, \dots, k$ をそれぞれを $x_1, \dots, x_k$ 回得る確率.

	$n = 1$	一般の n	検定	共役事前分布
$k = 2$	ベルヌイ分布 $\text{Ber}(p)$	二項分布 $B(n, p)$	二項検定	ベータ分布
一般の k	カテゴリ分布 $C(p_1, \dots, p_k)$	多項分布 $M(n, p_1, \dots, p_k)$	適合度の検定	ディリクレ分布

L07-Q1

Quiz(多項分布の確率)

ある四面体サイコロで、 $i = 1, 2, 3, 4$ の目の出る確率がそれぞれ $p_i = 0.3, 0.2, 0.4, 0.1$ で与えられる.

- ① この四面体サイコロを 5 回ふったときに、 $i = 1, 2, 3, 4$ の目の出る回数 x_i がしたがう分布を答えよう.
- ② この四面体サイコロを 5 回ふったときに、
 $(x_1, x_2, x_3, x_4) = (3, 0, 1, 1)$ となる確率を求めよう.

多項分布の共役事前分布

定義 (ディリクレ分布)

多次元連続分布 $\text{Dir}(\alpha_1, \dots, \alpha_k)$. 確率変数 $X_1, \dots, X_k$, パラメタ $\alpha_1, \dots, \alpha_k > 0$.

$$f(x_1, \dots, x_k) = \begin{cases} \frac{1}{Z(\alpha_1, \dots)} \prod_{i=1}^k x_i^{\alpha_i-1} & (0 \leq x_i \leq 1, \sum_i x_i = 1) \\ 0 & (\text{他}) \end{cases}$$

L07-Q2

Quiz(ディリクレ分布の確率)

p_1, p_2, p_3 は, パラメタ $(1, 1, 1)$ のディリクレ分布にしたがう.

- ① 確率密度関数 $f(p_1, p_2, p_3)$ が 0 にならない $(p_1, p_2, p_3) \in \mathbb{R}^3$ は, $\mathbb{R}^3$ のどこにあるか.
- ② $P(p_1 \geq \frac{1}{2})$ を求めよう.

L07-Q3

Quiz(ディリクレ分布を事前分布とするベイズ推定)

ある四面体のサイコロで、 $i = 1, 2, 3, 4$ の目の出る確率を p_i とする。
 p_i の事前分布を、パラメタ α_i のディリクレ分布とする。
このサイコロを 5 回振った時に、目の出た回数は
 $(x_1, x_2, x_3, x_4) = (3, 0, 1, 1)$ だった。
事後分布を求めよう。

ここまで来たよ

6 マルチレベルモデルのベイズ推定・混合ガウス分布

7 潜在変数モデルのベイズ推定・ディリクレ分布

- カテゴリ分布・多項分布・ディリクレ分布
- 潜在変数モデル

ラベルが不可視であることによる問題

L07-Q4

Quiz(連続型の条件付き確率)

ある畑でとれる柿の実の $\frac{3}{10}$ は品種 A, $\frac{7}{10}$ は品種 B である.

品種 A の実の質量は, 母平均値 200 g, 母分散 20^2 g^2 ,

品種 B の実の質量は, 母平均値 180 g, 母分散 10^2 g^2 の, それぞれ正規分布にしたがう.

- ① 畑でとれる柿の実の質量の分布を求めよう (確率密度関数を書こう).
- ② 畑でとれたひとつの柿の実の質量が 190 g であるとき, 品種 A である確率を求めよう.

L07-Q5

Quiz(混合ガウス分布のナীবベイズ)

周辺分布 $f_X(x)$ が混合ガウス分布 (π_y, μ_y, σ_y) になる同時分布 $f(x, y)$ ($y = 0, 1$) を考える.

パラメタを $(\pi_0 = 3/10, \pi_1 = 7/10, \mu_0 = 2, \mu_1 = 6, \sigma_0 = 2, \sigma_1 = 1/2)$ とする.

$X = 5$ という条件のもとで $Y = 0$ である条件付き確率 $P(Y = 0|X = 5)$ を, 標準正規分布の確率密度関数 $\phi(\cdot)$ で表そう.

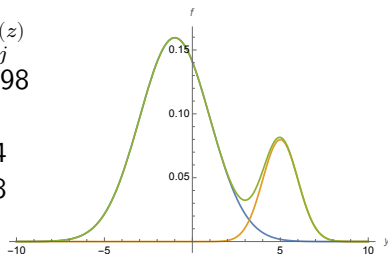
それが潜在変数 (latent variable)

潜在変数とは、観測できない（場合によっては想像上の）変数。
ノイズが加わった値すら観測できない（未知の母数は潜在変数ではない）
潜在変数の変数名としては、よく、 z が使われる。

例

- 5 科目テストデータが得られたときの、‘学力’ は潜在変数。
- ドーナツの重さデータが、製造機番号なしに与えられたときの、製造機番号は潜在変数。

製造機番号 z	ドーナツの重さ $x_j^{(z)}$
?	101, 103, 102, 99, 98
?	98, 100
?	103, 105, 104, 104
?	110, 115, 109, 108
?	102, 101, 100



二重苦：潜在変数モデル

- 潜在変数 2 クラスの場合, $Z = 0, 1, Z \sim B(1, p)$.
 $p_0 = 1 - p, p_1 = p$.
 - ▶ k クラスの場合 $\mathbf{p} = (p_1, \dots, p_k)$
- 観測値 $X^{(z)}$, 例えば, $X^{(z)} \sim N(\mu^{(z)}, (\sigma^{(z)})^2)$

$x^{(z)}$ の観測値が与えられるが, z がわからないときに, 母数 $(p_z, \mu^{(z)}, \sigma^{(z)})$ を推定せよ.

潜在変数モデルの階層ベイズモデリング

観測値 $x_j, j = 1, \dots, N$. 潜在変数 (ラベル) $i = z$. 母数 $(p_k, \mu^{(k)}, \sigma^{(k)})$.

- 事前分布 $f(\mathbf{p})f(\mu^{(0)}, \sigma^{(0)})f(\mu^{(1)}, \sigma^{(1)})$.
 - ▶ 必然ではないが、このように独立性が仮定されることが多い
- 尤度 $f(z_j|\mathbf{p})$
- 尤度 $f(x_j|z_j, \mu^{(0)}, \mu^{(1)}, \sigma^{(0)}, \sigma^{(1)})$
 - ▶ 例えば $p_{z_j} \frac{1}{\sqrt{2\pi(\sigma^{(z_j)})^2}} e^{-\frac{(x_j - \mu^{(z_j)})^2}{2(\sigma^{(z_j)})^2}}$.
- 事後分布

$$\begin{aligned}
 & f(\mathbf{p}, \mu^{(0)}, \mu^{(1)}, \sigma^{(0)}, \sigma^{(1)} | x_j) \\
 & \propto \prod_j \sum_{z_j} f(x_j | z_j, \mu^{(0)}, \mu^{(1)}, \sigma^{(0)}, \sigma^{(1)}) f(z_j | \mathbf{p}) \\
 & f(\mathbf{p}) f(\mu^{(0)}, \sigma^{(0)}) f(\mu^{(1)}, \sigma^{(1)})
 \end{aligned}$$