

Тест 3 по курсу «Байесовский выбор моделей»

Время выполнения: 30 минут

Максимальный балл: 50 баллов

Задача (50 баллов). Пусть имеется НОР (i.i.d.) выборка x_1, \dots, x_n из смеси нормальных распределений с одинаковым мат. ожиданием и разными дисперсиями, то есть $x_i \sim \sum_{i=1}^K \pi_i \mathcal{N}(m, \sigma_i^2)$, где K – фиксированная постоянная. Введем априорные распределения на m и σ_i^2 вида

$$m \sim \mathcal{N}(m_0, \sigma_0^2), \frac{1}{\sigma_i^2} \sim \Gamma(\alpha_i, \beta_i),$$

где $\alpha_i, \beta_i, m_0, \sigma_0^2$ – известные гиперпараметры.

- Выписать совместное правдоподобие модели $p(\mathbf{x}, m, \sigma_1^2 | \boldsymbol{\alpha}, \boldsymbol{\beta}, m_0, \sigma_0, \boldsymbol{\pi})$ (2 балла);
- Выписать $p(m, \sigma^2 | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}, m_0, \sigma_0, \boldsymbol{\pi})$ с точностью до постоянной. Принадлежит ли оно известному параметрическому семейству (2 балла)?
- Ввести матрицу скрытых переменную \mathbf{Z} с $z_{jk} = 1$ дающим принадлежность объекта j к компоненте смеси k . Выписать совместное правдоподобие модели $p(\mathbf{x}, m, \sigma^2, \mathbf{Z} | \boldsymbol{\alpha}, \boldsymbol{\beta}, m_0, \sigma_0, \boldsymbol{\pi})$ со скрытой переменной (3 балла);
- Получить вариационное приближение $q(m, \sigma^2, \mathbf{Z}) = q(m)q(\mathbf{Z})q(\sigma^2)$ для полного апостериорного распределения $p(m, \sigma^2, \mathbf{Z} | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}, m_0, \sigma_0, \boldsymbol{\pi})$ при известном векторе весов компонент в смеси $\boldsymbol{\pi}$ (13 баллов). Что мешает успешно применить обычный EM-алгоритм (2 балла)?
- Найти наиболее обоснованные значения весов компонент смеси $\boldsymbol{\pi}$, решив задачу (20 баллов)

$$p(\mathbf{x} | \boldsymbol{\alpha}, \boldsymbol{\beta}, m_0, \sigma_0, \boldsymbol{\pi}) \rightarrow \max_{\boldsymbol{\pi}}.$$

Дает ли полученный $\boldsymbol{\pi}^*$ модель с наибольшей апостериорной плотностью вероятности (1 балл)? Как можно проинтерпретировать полученное выражение для $\boldsymbol{\pi}$ (2 балла)? Что нужно изменить в модели, чтобы «поощрить разреженность», то есть исключение избыточных компонент из смеси (5 баллов)?

Замечание: В формулировке выше векторы $\boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\beta}$ комбинируют соответствующие значения по компонентам, например, $\boldsymbol{\sigma} = [\sigma_1, \dots, \sigma_K]^T$.