

Мультимоделирование как универсальный способ описания выборки общего вида

Качанов В. В., Стрелкова Е. С., Адуенко А. А.

Strelkova.ES@phystech.edu

¹Организация; ²Организация

В данной работе рассматривается метод инкрементального уточнения структуры мультимодели при появлении новых объектов. Для повышения точности мультимодели необходимо разработать и сравнить разные алгоритмы обновления структуры, и в результате построить оптимальную схему уточнения мультимодели в зависимости от объёма выборки. Для решения данной задачи предлагается использовать Байесовский метод на основании обоснованности, проанализировать связь обоснованности и статистической значимости.

Введение

Решается задача инкрементального уточнения структуры мультимодели.

Постановка задачи

Пусть имеется K моделей, $k \in [1, K]$, $\{\bar{x}_i, y_i\}_{i=1}^m$, - выборка

$$y_i = \bar{w}_k^T \bar{x}_i + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \beta^{-1})$$

Априорное распределение на $\bar{\pi}$: $p(\bar{\pi}|\mu) = Dir(\bar{\pi}|\mu)$ Априорное распределение моделей: $p(\bar{w}_k) = \mathcal{N}(\bar{w}_k|0, A_k)$

Совместное правдоподобие:

$$p(\bar{y}, \bar{W}, \bar{\pi}|X, A, \beta, \mu) = Dir(\bar{\pi}|\mu) \prod_{k=1}^K \mathcal{N}(\bar{w}_k|0, A_k) \prod_{i=1}^m \sum_{j=1}^K \pi_k \mathcal{N}(y_i|\bar{w}_k^T \bar{x}_i \bar{w}_k, \beta^{-1})$$

Апостериорное распределение пропорционально:

$$p(\bar{W}, \bar{\pi}|X, \bar{y}, A, \beta, \mu) \sim \prod_{i=1}^m \left(\sum_{j=1}^K \pi_k \exp \left(-\frac{\beta}{2} (y_i - \bar{w}_j^T \bar{x}_i)^2 \right) \right) * \\ * \prod_{k=1}^K \pi_k^{\mu-1} \exp(-0.5 \bar{w}_k^T A_k \bar{w}_k)$$

Для решения задачи воспользуемся вариационным ЕМ-алгоритмом со скрытой переменной $Z = ||z_{ik}||$, тогда совместное правдоподобие переписывается в виде

$$p(\bar{y}, \bar{W}, \bar{\pi}, Z|X, A, \beta, \mu) = Dir(\bar{\pi}|\mu) \prod_{k=1}^K \mathcal{N}(\bar{w}_k|0, A_k) \prod_{i=1}^m * \\ * \left(\prod_{j=1}^K \pi_j \mathcal{N}(y_i|\bar{w}_k^T \bar{x}_i \bar{w}_k, \beta^{-1}) \right)^{z_{ij}}$$

Воспользовавшись вариационным приближением: $q(\bar{\pi}, Z, W) = q(\bar{\pi}) q(Z) q(W)$

$$\begin{aligned}
 \log q(\bar{\pi}) &= \sum_{k=1}^K \log \pi_k \left(\sum_{i=1}^m \mathbb{E} z_{ik} + \mu - 1 \right) \\
 &\Rightarrow q(\bar{\pi}) = Dir(\bar{\pi} | \mu + \bar{\alpha}), \quad \alpha_k = \sum_{i=1}^m z_{ik} \\
 \log q(W) &\sim \sum_{i=1}^m -\frac{1}{2} \bar{w}_k^T A_k \bar{w}_k + \sum_{i=1}^m \sum_{l=1}^K \mathbb{E} z_{il} \frac{\beta}{2} (\bar{w}_l^T x_i x_i^T A_k \bar{w}_l - 2 y_i \bar{w}_l^T x_i) \\
 q(\bar{w}_k) &= \mathcal{N}(\bar{w}_k | m_k, \Sigma_k^{-1}) \\
 \log q(Z) &\sim \sum_{k=1}^K \sum_{i=1}^m z_{ik} \left(\mathbb{E} \log \pi_k - \frac{\beta}{2} (y_i - \bar{w}_k^T \bar{x}_i)^2 \right) \Rightarrow \\
 &\Rightarrow p(z_{ik} = 1) = C \exp \left(\mathbb{E} \log \pi_k - \frac{\beta}{2} (y_i - \bar{w}_k^T \bar{x}_i)^2 \right) \\
 \mathbb{E}_q \log p(\bar{y}, \bar{p}, W, Z | X, A, \beta, \mu) &= \mathcal{F}(A, \beta) \propto \\
 \sum_{k=1}^K ((\mu + 2\alpha_k - 1) \mathbb{E} \log \pi_k - \frac{1}{2} \mathbb{E} \bar{w}_k^T A_k^{-1} \bar{w}_k + \frac{1}{2} \log \det A_k^{-1} + \\
 \sum_{i=1}^m \mathbb{E} z_{ik} (\log \beta - \frac{\beta}{2} \mathbb{E} (y_i - \bar{w}_k^T \bar{x}_i)^2)) \\
 \frac{\partial \mathcal{F}}{\partial A_k^{-1}} = 0 &\Rightarrow \tilde{A}_k = Diag(\mathbb{E}(w_k^i)^2) \\
 \frac{\partial \mathcal{F}}{\partial \beta} = 0 &\Rightarrow \tilde{\beta} = \frac{\sum_{k=1}^K \sum_{i=1}^m \frac{1}{2} \mathbb{E} z_{ik} (y_i - \bar{w}_k^T \bar{x}_i)^2}{\sum_{k=1}^K \sum_{i=1}^m \mathbb{E} z_{ik}}
 \end{aligned}$$

ЛитератураReferences

- [1] Адуенко А. А. Выбор мультимodelей в задачах классификации, 2017.
- [2] *Bishop C.M.* Pattern recognition and machine learning *Berlin: Springer*, 2006.
- [3] *MacKay D.J.C* The evidence framework applied to classification networks *Neural computation* 4.5, 1992. Pp.720–736.
- [4] *Gelman A.* Bayesian data analysis *Florida: Chapman and Hall/CRC*, 2013.
- [5] *Motrenko A., Strijov V., Weber G.-W.* Sample size determination for logistic regression. *Journal of Computational and Applied Mathematics*, 2014. Vol.255, Pp.743-752.