Зміст

Вступ

Огляд літератури

Постановка задачі

Ми розглядаємо модель суміші зі змінними концентраціями. Об'єкти спостереження O_j, $j = 1, \ldots, n$ належать до m компонент з невідомими індексами $\kappa_j \in \{1, \ldots, m\}$, але наперед заданими ймовірностями $p^k_{j;n} = P\{\kappa_j = k\}$, які ще називають імовірностями змішення або концентраціями k-тої компоненти суміші під час j-ого спостереження.

Для об'єкта O аналізуємо модель лінійної регресії вигляду

$$y(O) = \sum_{i=1}^{d} \theta_{f,i}^{(\kappa(O))} x^i(O) + \varepsilon(O),$$

де $\xi(O) = \begin{pmatrix} y(O), & x_1(O), & \ldots, & x_d(O) \end{pmatrix}$, $x(O)$ вектор регресорів та $\varepsilon(O)$ – це помилка з нульовим середнім, що не залежить від $x(O)$. Дана вибірка $\{\xi_j = (x_j, y_j)\}$, задача оцінити параметри $\theta_{f,k}$ та θ_p, використовуючи функціонал якості

$$L = \sum_{k=1}^{m} \sum_{i=1}^{n} a^k_{i;n}(\theta_p) \left(y_i - f_k^{\theta_{f,k}}(x_i) \right)^2,$$

де $a^k_{i;n}(\theta_p) = \sum_{t=1}^{m} p^t_{i;m} \Gamma^{-1}_{t;m}$ – мінімаксні коефіцієнти, $\Gamma_n = P_n P_n^T$ та $P_n = (p^k_{i;n})_{i=1,k=1}^{n,m}$ [1].

Основні результати

Були знайдені такі обмеження L, які можуть бути диференційовані відносно θ_p, θ_f, щоб забезпечити оптимізацію.

$$|L| \leq \|\Gamma_n^{-1}\|_F \sum_{k=1}^{m} \sum_{i=1}^{n} \sqrt{\sum_{t=1}^{m} (p^t_{i;m})^2} \left(y_i - f_k^{\theta_{f,k}}(x_i) \right)^2,$$

$$|L| \leq \|\Gamma_n^{-1}\|_F \sqrt{\sum_{k=1}^{n} \sum_{t=1}^{m} \sum_{i=1}^{n} p^t_{i;n} \left(y_i - f_k^{\theta_{f,k}}(x_i) \right)^2}.$$

В особливому випадку $m = 2$ отримано:

$$|L| \leq \|\Gamma_n^{-1}\|_F \sqrt{\sum_{k=1}^{2} \left(\sum_{i=1}^{n} p^k_{i;n} \left(y_i - f_k^{\theta_{f,k}}(x_i) \right)^2 \right)^2}.$$

Доведення

Задана функція втрат:

$$L = \sum_{k=1}^{m} \sum_{i=1}^{n} a^k_{i;n}(\theta_p) \left(y_i - f_k^{\theta_{f,k}}(x_i) \right)^2.$$

де $a^k_{i;n}(\theta_p)$ визначені як: $a^k_{i;n}(\theta_p) = \sum_{t=1}^m p^t_{i;m} \cdot (\Gamma_n^{-1})_{t;k}$.

Перше обмеження

$$|L| \le \|\Gamma_n^{-1}\|_F \sum_{k=1}^m \sum_{i=1}^n \sqrt{\sum_{t=1}^m (p^t_{i;m})^2} \cdot \left(y_i - f_k^{\theta_{f,k}}(x_i)\right)^2$$

Доведення.

$$L = \sum_{k=1}^m \sum_{i=1}^n a^k_{i;n}(\theta_p) \left(y_i - f_k^{\theta_{f,k}}(x_i)\right)^2.$$

Застосуємо нерівність Коші–Буняковського до векторів

$$\vec{p}_{i;m} = \left(p^1_{i;m}, p^2_{i;m}, \ldots, p^m_{i;m}\right) \quad \text{та} \quad \vec{g}_k = \left((\Gamma_n^{-1})_{1k}, (\Gamma_n^{-1})_{2k}, \ldots, (\Gamma_n^{-1})_{mk}\right),$$

де перший вектор — це вектор ймовірностей для спостереження i, а другий — k-ий стовпець матриці Γ_n^{-1}. Тоді маємо:

$$\left|\sum_{t=1}^m p^t_{i;m} (\Gamma_n^{-1})_{t;k}\right| \le \sqrt{\sum_{t=1}^m (p^t_{i;m})^2} \cdot \sqrt{\sum_{t=1}^m \left((\Gamma_n^{-1})_{t;k}\right)^2}.$$

Зауважимо, що $\sqrt{\sum_{t=1}^m \left((\Gamma_n^{-1})_{t;k}\right)^2} \le \|\Gamma_n^{-1}\|_F$, де $\|\cdot\|_F$ норма Фробеніуса, оскільки $\left((\Gamma_n^{-1})_{t;k}\right)^2 = \sqrt{\sum_{k=1}^n \sum_{t=1}^m \left((\Gamma_n^{-1})_{t;k}\right)^2}$. Отже, справджується нерівність $a^k_{i;n}(\theta_p) \le \sqrt{\sum_{t=1}^m (p^t_{i;m})^2} \cdot \|\Gamma_n^{-1}\|_F$. Підставимо в обмеження :

$$|L| \le \|\Gamma_n^{-1}\|_F \sum_{k=1}^m \sum_{i=1}^n \sqrt{\sum_{t=1}^m (p^t_{i;m})^2} \cdot \left(y_i - f_k^{\theta_{f,k}}(x_i)\right)^2.$$

Матриця Грама Γ_n є симетричною і додатньо-визначеною, тому $\|\Gamma_n^{-1}\|_F$ можна виразити через власні числа Γ_n, позначені $\lambda_1, \ldots, \lambda_m$, як

$$\|\Gamma_n^{-1}\|_F = \sqrt{\operatorname{tr}\left((\Gamma_n^{-1})^\top \Gamma_n^{-1}\right)} = \sqrt{\sum_{i=1}^m \frac{1}{\lambda_i^2}},$$

де $tr(\cdot)$ – слід матриці, сума діагональних елеменів.

Друге обмеження

$$|L| \le \|\Gamma_n^{-1}\|_F \cdot \sqrt{\sum_{k=1}^m \sum_{t=1}^m \left(\sum_{i=1}^n p^t_{i;n} \cdot \left(y_i - f_k^{\theta_{f,k}}(x_i)\right)^2\right)^2}$$

Доведення. Змінемо порядок сум в функції L

$$L = \sum_{k=1}^m \sum_{t=1}^m \sum_{i=1}^n (\Gamma_n^{-1})_{k;t} \cdot p^t_{i;n} \cdot \left(y_i - f_k^{\theta_{f,k}}(x_i)\right)^2.$$

Для зручності позначимо $U_{t;k} := (\Gamma_n^{-1})_{t;k}$, $\quad V_{t;k} := \sum_{i=1}^n p_{i;n}^t \cdot \Big(y_i - f_k^{\theta_{f,k}}(x_i) \Big)^2$, тоді

$$L = \sum_{k=1}^m \sum_{t=1}^m U_{t;k} \cdot V_{t;k}.$$

Застосуємо нерівність Коші-Буняковського

$$\left| \sum_{k=1}^m \sum_{t=1}^m U_{t;k} V_{t;k} \right| \leq \sqrt{\sum_{k=1}^m \sum_{t=1}^m U_{t;k}^2} \cdot \sqrt{\sum_{k=1}^m \sum_{t=1}^m V_{t;k}^2}.$$

$$\sqrt{\sum_{k=1}^m \sum_{t=1}^m U_{t;k}^2} = \|\Gamma_n^{-1}\|_F.$$

$$\sum_{k=1}^m \sum_{t=1}^m V_{t;k}^2 = \sum_{k=1}^m \sum_{t=1}^m \left(\sum_{i=1}^n p_{i;n}^t \cdot \Big(y_i - f_k^{\theta_{f,k}}(x_i) \Big)^2 \right)^2.$$

Отримаємо обмеження

$$|L| \leq \|\Gamma_n^{-1}\|_F \cdot \sqrt{\sum_{k=1}^m \sum_{t=1}^m \left(\sum_{i=1}^n p_{i;n}^t \cdot \Big(y_i - f_k^{\theta_{f,k}}(x_i) \Big)^2 \right)^2}.$$

Третє обмеження

$$|L| \leq \|\Gamma_n^{-1}\|_F \cdot \sqrt{\sum_{k=1}^2 \left(\sum_{i=1}^n p_{i;n}^k \Big(y_i - f_k^{\theta_{f,k}}(x_i) \Big)^2 \right)^2}$$

Доведення. Розглянемо випадок $m = 2$. Нагадаємо, що розмір матриці Γ_n^{-1} – $m \times m$, в цьому випадку 2×2.

$$\Gamma_n^{-1} = \begin{pmatrix} (\Gamma_n^{-1})_{1;1} & (\Gamma_n^{-1})_{1;2} \\ (\Gamma_n^{-1})_{2;1} & (\Gamma_n^{-1})_{2;2} \end{pmatrix}, \quad \text{де } (\Gamma_n^{-1})_{1;2} = (\Gamma_n^{-1})_{2;1}.$$

Функція L набуває вигляду

$$\begin{aligned} L &= \sum_{i=1}^n \sum_{k=1}^2 \left(\sum_{t=1}^2 p_{i;n}^t \cdot (\Gamma_n^{-1})_{t;k} \right) \Big(y_i - f_k^{\theta_{f,k}}(x_i) \Big)^2 \\ &= \sum_{i=1}^n \Big[p_{i;n}^1 \cdot (\Gamma_n^{-1})_{1;1} \Big(y_i - f_1^{\theta_{f,1}}(x_i) \Big)^2 + p_{i;n}^2 \cdot (\Gamma_n^{-1})_{2;1} \Big(y_i - f_1^{\theta_{f,1}}(x_i) \Big)^2 \\ &\quad + p_{i;n}^1 \cdot (\Gamma_n^{-1})_{1;2} \Big(y_i - f_2^{\theta_{f,2}}(x_i) \Big)^2 + p_{i;n}^2 \cdot (\Gamma_n^{-1})_{2;2} \Big(y_i - f_2^{\theta_{f,2}}(x_i) \Big)^2 \Big]. \end{aligned}$$

Згрупуємо доданки $L = \sum_{i=1}^n \Big[\big(p_{i;n}^1 \cdot (\Gamma_n^{-1})_{1;1} + p_{i;n}^2 \cdot (\Gamma_n^{-1})_{2;1} \big) \Big(y_i - f_1^{\theta_{f,1}}(x_i) \Big)^2$
$+ \big(p_{i;n}^1 \cdot (\Gamma_n^{-1})_{1;2} + p_{i;n}^2 \cdot (\Gamma_n^{-1})_{2;2} \big) \Big(y_i - f_2^{\theta_{f,2}}(x_i) \Big)^2 \Big].$

Нехай $V_k = \sum_{i=1}^{n} p_{i;n}^k \left(y_i - f_k^{\theta_{f,k}}(x_i)\right)^2, \quad$ для $k = 1, 2$. Підставимо V_k і отримаємо

$$
\begin{aligned}
L &= (\Gamma_n^{-1})_{1;1}V_1 + (\Gamma_n^{-1})_{2;1}V_1 + (\Gamma_n^{-1})_{1;2}V_2 + (\Gamma_n^{-1})_{2;2}V_2 \\
&= V_1\left((\Gamma_n^{-1})_{1;1} + (\Gamma_n^{-1})_{2;1}\right) + V_2\left((\Gamma_n^{-1})_{1;2} + (\Gamma_n^{-1})_{2;2}\right).
\end{aligned}
$$

Елементи матриці Грама розміру 2×2, що розташовані на побічній діагоналі є недодатними $\left((\Gamma_n^{-1})_{1;2}, (\Gamma_n^{-1})_{2;1} \leq 0\right)$, тому справджується обмеження $L \leq (\Gamma_n^{-1})_{1;1}V_1 + (\Gamma_n^{-1})_{2;2}V_2$.

$$
V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}, \quad \Gamma = \begin{pmatrix} (\Gamma_n^{-1})_{1;1} \\ (\Gamma_n^{-1})_{2;2} \end{pmatrix}, \quad L \leq \langle \Gamma, V \rangle.
$$

Застосуємо нерівність Коші-Буняковського $|L| \leq \|\Gamma\|_F \cdot \|V\|_F$,

$$
\|V\|_F = \sqrt{V_1^2 + V_2^2}, \quad \|\Gamma\|_F = \sqrt{(\Gamma_n^{-1})_{1;1}^2 + (\Gamma_n^{-1})_{2;2}^2}.
$$

$\|\Gamma\|_F \leq \|\Gamma_n^{-1}\|_F \Rightarrow |L| \leq \|\Gamma_n^{-1}\|_F \cdot \|V\|_F$. Запишемо явно норму V і отримаємо

$$
|L| \leq \|\Gamma_n^{-1}\|_F \cdot \sqrt{\sum_{k=1}^{2} \left(\sum_{i=1}^{n} p_{i;n}^k \left(y_i - f_k^{\theta_{f,k}}(x_i)\right)^2 \right)^2},
$$

що і треба було довести.

Прикладна частина

Висновки

Список використаних джерел

[1] Miroshnichenko V., Maiboroda R. *Asymptotic normality of modified LS estimator for mixture of nonlinear regressions.* Modern Stochastics: Theory and Applications, Vol. 7, No. 4, pp. 435–448, 2020.