

519.233.2

Использование бутстрепа для построения равномерных доверительных интервалов в оценке плотности

А. С. Циглер

Московский физико-технический институт (государственный
университет)

Институт проблем передачи информации им. А. А. Харкевича РАН

План

- 1 Модель
- 2 Концентрация, теорема Вилкса
- 3 Гауссовская аппроксимация
- 4 Сравнение гауссовских распределений
- 5 Основной результат

Модель

$X_1, X_2, \dots, X_n \in \mathbb{R}^d$ - i.i.d. из плотности ρ .

Theorem (Loader)

Истинная функция плотности максимизирует функционал

$$F(\rho) = \mathbb{E}g(X_1) \log(\rho(X_1)) - \int g(x)\rho(x)dx \quad (1)$$

Параметрические предположения: $g(x) = K(\frac{x-\eta}{h})$,

$\log(\hat{\rho}(x, \theta)) = \Psi(\frac{x-\eta}{h})^T \theta$.

$$\mathbb{E}K(\frac{X_1 - \eta}{h})\Psi(\frac{X_1 - \eta}{h})^T \theta - \int K(\frac{x - \eta}{h}) \exp(\Psi(\frac{X_1 - \eta}{h})^T \theta) \rightarrow \max \quad (2)$$

Эмпирический функционал

Эмпирический функционал:

$$\frac{1}{n} \sum_{i=1}^n K\left(\frac{X_i - \eta}{h}\right) \Psi\left(\frac{X_i - \eta}{h}\right)^T \theta - \int K\left(\frac{x - \eta}{h}\right) \exp\left(\Psi\left(\frac{x - \eta}{h}\right)^T \theta\right) dx \rightarrow \max \quad (3)$$

Введем

$$S_\eta = \frac{1}{nh^d} \sum \Psi\left(\frac{X_i - \eta}{h}\right) K\left(\frac{X_i - \eta}{h}\right) \quad A(\theta) = \int K(z) \exp(\theta^T \Psi(z)) dz \quad (4)$$

$$L(\theta, \eta, h) = \theta^T S_\eta - A(\theta) \quad (5)$$

Основные обозначения

$$L(\theta, \eta, h) = \theta^T S_\eta - A(\theta) \quad (6)$$

$$\tilde{\theta}_\eta = \operatorname{argmax}_{\theta} L(\theta, \eta, h) \quad (7)$$

$$\theta_\eta^* = \operatorname{argmax}_{\theta} \mathbb{E} L(\theta, \eta, h) \quad (8)$$

$$D^2(\theta) = -\nabla_\theta^2 L(\theta, \eta, h) = \nabla_\theta^2 A(\theta) = \int K(z) \Psi(z) \Psi(z)^T \exp(\theta^T \Psi(z)) dz \quad (9)$$

Концентрация, теорема Вилкса

Обозначим $D_\eta^2 = D^2(\theta_\eta^*)$, $r_\eta = \|D_\eta(\tilde{\theta}_\eta - \theta_\eta^*)\|_2$,

$C_\eta = \sup_{z \in B_1(0)} \|D_\eta^{-1}\Psi(z)\|_2$.

Theorem (о концентрации)

$$\|D_\eta^{-1}(S_\eta - \mathbb{E}S_\eta)\|_2 \geq \sup_{r < r_\eta} \frac{1}{2} r \exp(-rC_\eta) \quad (10)$$

Theorem (Вилкс)

$$\begin{aligned} \left| 2L(\tilde{\theta}_\eta, \eta, h) - 2L(\theta_\eta^*, \eta, h) - \|D_\eta^{-1}(S_\eta - \mathbb{E}S_\eta)\|_2^2 \right| \leq \\ \leq r_\eta^2 \exp(r_\eta C_\eta)(\exp(r_\eta C_\eta) - 1) \quad (11) \end{aligned}$$

Дискретизация пространства

Пусть $\Xi = \{\eta_i\}_{i=1}^M$ - множество точек, в которых мы хотим построить доверительные интервалы.

Ранее вводились $S_\eta = \frac{1}{nh^d} \sum \Psi(\frac{X_i - \eta}{h}) K(\frac{X_i - \eta}{h})$, $D_\eta^2 = D^2(\theta_\eta^*)$.

Теперь можно считать $S \in \mathbb{R}^{pM}$, $D \in \text{Mat}_{pM \times pM}$.

Телескоп Линдеберга

Lemma

Пусть $\zeta_1, \dots, \zeta_n \in \mathbb{R}^k$ - i.i.d. независимы от $\tilde{\zeta}_1, \dots, \tilde{\zeta}_n \in \mathbb{R}^k$ - i.i.d. с теми же ожиданием и ковариацией, что и ζ_i .

Также, пусть $f : \mathbb{R}^k \rightarrow \mathbb{R}$ - трижды дифференцируемая функция .
Тогда

$$\begin{aligned} \mathbb{E} \left[f \left(\frac{1}{n} \sum_1^n \zeta_i \right) - f \left(\frac{1}{n} \sum_1^n \tilde{\zeta}_i \right) \right] = \\ = \mathbb{E} \frac{1}{6n^3} \sum_1^n \left((f^{(3)}(\bar{\zeta}_i), \zeta_i \otimes \zeta_i \otimes \zeta_i) - (f^{(3)}(\tilde{\zeta}_i), \tilde{\zeta}_i \otimes \tilde{\zeta}_i \otimes \tilde{\zeta}_i) \right) \quad (12) \end{aligned}$$

Гауссовская аппроксимация

Обозначим

$$S_{i,\eta} = \frac{1}{h^d} \Psi\left(\frac{X_i - \eta}{h}\right) K\left(\frac{X_i - \eta}{h}\right) \quad (13)$$

За S_i обозначим вектор $(S_{\eta_1}^T, \dots, S_{\eta_M}^T)^T$ (аналогично вектору S).

Введем также

$$\tilde{S}_i \in \mathcal{N}(\mathbb{E}S_i, \text{Var } S_i) \quad (14)$$

Мы хотим приблизить распределения $\max_i \|D_{\eta_i}^{-1}(S_{\eta_i} - \mathbb{E}S_{\eta_i})\|_2^2$ и $\|D_{\eta_i}^{-1}(\tilde{S}_{\eta_i} - \mathbb{E}S_{\eta_i})\|_2^2$.

То есть применить телескоп Линдеберга для функции

$$f(\zeta) \approx \mathbb{I}\{\max \|\zeta_{\eta_i}\|_2^2 - q < 0\}$$

Выбор функции

Возьмем

$$f = g \circ F_\beta, \quad (15)$$

где

$$F_\beta(x) = \beta^{-1} \log \left(\sum_j \exp(\beta(\|x_{\eta_j}\|_2^2 - q)) \right), \quad (16)$$

$g(x) = g_0(\phi^{-1}x)$, а $g_0 : \mathbb{R} \rightarrow \mathbb{R}$ - некоторая монотонная гладкая функция, такая что

$$g_0(x) = \begin{cases} 1, & x \leq 0, \\ 0, & x \geq 1. \end{cases} \quad (17)$$

Гауссовская аппроксимация - результат

Обозначим за G_1, G_2, G_3 соответственно максимумы первой, второй и третьей производных функции g_0 .

Обозначим $L = \sup_{x,j} \|D_{\eta_j}^{-1} \Psi(\frac{\eta_j - x}{h}) K(\frac{\eta_j - x}{h})\|_2$.

Введем

$$\begin{aligned} \chi_1(q) = \frac{4}{3(nh^d)^3} L^3 & \left(1 + 2\sqrt{\frac{2}{\pi}} \right) (\phi^{-1} G_1(24\beta q^{\frac{1}{2}} + 32\beta^2 q^{\frac{3}{2}} + 8\beta q^{\frac{3}{2}}) + \\ & + \phi^{-2} G_2(12q^{\frac{1}{2}} + 48\beta q^{\frac{3}{2}}) + \phi^{-3} G_3(8q^{\frac{3}{2}})). \end{aligned} \quad (18)$$

Тогда выполнено следующее утверждение:

$$\left| \mathbb{E} \left[f(D^{-1}(S - \mathbb{E}S)) - f(D^{-1}(\tilde{S} - \mathbb{E}S)) \right] \right| \leq \chi_1(q + \phi) \quad (19)$$

Процедура бутстрепа

Для бутстрепа мы будем генерировать векторы

$$S_{\eta_j}^b = \frac{1}{n} \sum_i S_i w_i^b = \frac{1}{nh^d} \sum_1^n \Psi\left(\frac{X_i - \eta_j}{h}\right) K\left(\frac{X_i - \eta_j}{h}\right) w_i^b, \quad (20)$$

где w_i - веса из $\mathcal{N}(1, 1)$.

Тогда

$$D^{-1}(\tilde{S} - \mathbb{E}S) \in \mathcal{N}\left(0, \frac{1}{n} D^{-1}(\mathbb{E}S_1 S_1^T - \mathbb{E}S_1 \mathbb{E}S_1^T) D^{-1}\right) \quad (21)$$

$$D^{-1}(S^b - S) \in \mathcal{N}\left(0, \frac{1}{n} D^{-1}\left(\frac{1}{n} \sum S_i S_i^T\right) D^{-1}\right) \quad (22)$$

Сравнение гауссовских распределений

Theorem

Пусть X и Y - два центрированных гауссовских вектора с ковариационными матрицами Σ_X и Σ_Y соответственно.

Пусть также $f(x)$ - гладкая функция.

Тогда существует такая матрица U , что для любой нормы $\|\cdot\|_*$ выполнено $\|U\|_* < \sup_x \|\nabla^2 f(x)\|_*$ и

$$|\mathbb{E}f(X) - \mathbb{E}f(Y)| \leq \frac{1}{2} |\text{tr}\{(\Sigma_X - \Sigma_Y)U\}| \quad (23)$$

Сравнение ковариационных матриц

Надо оценить $\text{tr}\{D^{-1}(\mathbb{E}S_1S_1^T - \sum S_iS_i^T - \mathbb{E}S_1\mathbb{E}S_1^T)D^{-1}U\}$

Мы уже ввели $L = \sup_{x,j} \|D_{\eta_j}^{-1}\Psi(\frac{\eta_j - x}{h})K(\frac{\eta_j - x}{h})\|_2$.

$\text{tr}\{D^{-1}(\mathbb{E}S_1S_1^T - \sum S_iS_i^T)D^{-1}U\}$ - Хёфдинг.

$\text{tr}\{D^{-1}\mathbb{E}S_1\mathbb{E}S_1^TD^{-1}U\}$ - имеет меньший порядок, чем предыдущий член.

Сравнение гауссовских распределений - результат

Введем

$$\begin{aligned}
 \chi_2(q) &= \\
 &= \frac{L^2}{2n} \left(\frac{2\sqrt{2}(\tau + 2 \log M)}{nh^{2d}} + \Upsilon_d^2 \max_x \rho(x)^2 \right) \left(G_1(2p^2 + 8\beta q) + 4G_2q \right),
 \end{aligned} \tag{24}$$

где Υ_d - объем единичного шара в d -мерном пространстве.

Верен следующий факт:

Для любого положительного τ с вероятностью $\geq 1 - 2 \exp(-\tau)$

$$\left| \mathbb{E} f \left(D^{-1}(\tilde{S} - \mathbb{E} S) \right) - \mathbb{E} f \left(D^{-1}(S^b - S) \right) \right| \leq \chi_2(q + \phi) \tag{25}$$

Основной результат работы

Theorem

Пусть $q < \frac{1}{eC}$, . Тогда для любых положительных ϕ, β выполнено следующее:

$$\begin{aligned} & \mathbb{P}\left\{\max_j [2L(\tilde{\theta}_\eta, \eta, h) - 2L(\theta_\eta^*, \eta, h)] \leq q^2(1 + 4\exp(3eqC)(\exp(eqC) - 1))\right\} \geq \\ & \geq \mathbb{P}\left\{\max_j [2L(\tilde{\theta}_{\eta_j}^b, \eta, h) - 2L(\tilde{\theta}_{\eta_j}, \eta, h)] \leq (q^2 - \phi - \beta^{-1}\log(M)) \times \right. \\ & \quad \times (1 - 4\exp(3eC\sqrt{q^2 - \phi - \beta^{-1}\log(M)})) \times \\ & \quad \left. \times (\exp(eC\sqrt{q^2 - \phi - \beta^{-1}\log(M)}) - 1)\right\} - \chi_2(q) - \chi_1(q) \quad (26) \end{aligned}$$

Список литературы

-  *Loader, C.:* Local Likelihood Density Estimation. Ann. Statist. Volume 24, Number 4, 1602-1618 (1996)
-  *Chernozhukov, V., Chetverikov, D. and Kato, K.:* Gaussian approximations and multiplier bootstrap for maxima of sums of highdimensional random vectors. Ann. Statist. 41 2786-2819 (2013).
-  *Spokoiny, V.:* Parametric estimation. Finite sample theory. Ann. Statist. Volume 40, Number 6, 2877-2909 (2012)
<http://arxiv.org/abs/1111.3029>

Спасибо за внимание!