Обнаружение точки разладки при помощи обобщенного отношения правдоподобия

Бузун Назар postrealist@gmail.com

Оглавление

1 График отношения правдоподобия

1.1. Точка максимума правдоподобия

$$\alpha(\theta, \theta_0) = L(\theta) - L(\theta_0) - (\theta - \theta_0)^T \nabla L_G(\theta_0) + \frac{1}{2}\|D(\theta - \theta_0)\|^2$$

Пусть в локальной области $\Theta_0(r) = \{\|D(\theta - \theta^*)\| < r\}$, где $\theta, \theta_0 \in \Theta_0(r)$, справедливо неравенство с вероятностью $1 - e^{-x}$

$$\frac{|\alpha(\theta, \theta_0)|}{\|D(\theta - \theta_0)\|} \le 2\diamondsuit(r, x), \tag{a}$$

где $\diamondsuit(r, x) = (\delta(r) + 6 v_0 z_H(x)\omega) r$.

Разделим выборку на два отрезка по индексу, правдоподобие и параметр части выборки слева будем индексировать как $*_1$, справа – $*_2$. Выразим ОМП оценку на всей выборке через ОМП оценки слева ($\widehat{\theta}_1$) и справа ($\widehat{\theta}_2$):

$$L(\theta) = L_1(\theta) + L_2(\theta) = L_1(\widehat{\theta}_1) + L_2(\widehat{\theta}_2) -$$
$$-\frac{1}{2}(\theta - \widehat{\theta}_1)^T D_1^2 (\theta - \widehat{\theta}_1) - \frac{1}{2}(\theta - \widehat{\theta}_2)^T D_2^2 (\theta - \widehat{\theta}_2) + O(r\diamondsuit(r, x)),$$
$$\frac{1}{2}(\theta - \widehat{\theta}_1)^T D_1^2 (\theta - \widehat{\theta}_1) - \frac{1}{2}(\theta - \widehat{\theta}_2)^T D_2^2 (\theta - \widehat{\theta}_2) =$$
$$= \frac{1}{2}(\theta - \widehat{\theta})^T D^2 (\theta - \widehat{\theta}) + \mathrm{const}(\theta),$$
$$\widehat{\theta} = D^{-2}(D_1^2 \widehat{\theta}_1 + D_2^2 \widehat{\theta}_2), \quad D^2 = D_1^2 + D_2^2.$$

Лемма. Пусть $\tilde{\theta} = \arg\max L(\theta)$, $\widehat{\theta} \in \Theta_1(r) \cap \Theta_2(r)$, тогда с вероятностью $1 - 2e^{-x}$

$$\|D(\tilde{\theta} - \widehat{\theta})\|^2 \le O(r\diamondsuit(r, x)).$$

Замечание. Для выполнения условия $\widehat{\theta} \in \Theta_1(r) \cap \Theta_2(r)$ требуется ограничение на изменение параметра θ^* вида

$$\|D(\theta_1^* - \theta_2^*)\| \le r. \tag{L*}$$

1.2. Отношение правдоподобия

$$\triangle L = L_1(\widehat{\theta}_1) + L_2(\widehat{\theta}_2) - L(\widehat{\theta}) =$$
$$= \frac{1}{2}(\widehat{\theta} - \widehat{\theta}_1)^T D_1^2(\widehat{\theta} - \widehat{\theta}_1) + \frac{1}{2}(\widehat{\theta} - \widehat{\theta}_2)^T D_2^2(\widehat{\theta} - \widehat{\theta}_2) + O(r\diamondsuit(r,x)).$$
$$\widehat{\theta} - \widehat{\theta}_1 = D^{-2}(D_1^2\widehat{\theta}_1 + D_2^2\widehat{\theta}_2) - \widehat{\theta}_1 = D^{-2}D_2^2(\widehat{\theta}_2 - \widehat{\theta}_1),$$
$$\widehat{\theta} - \widehat{\theta}_2 = D^{-2}(D_1^2\widehat{\theta}_1 + D_2^2\widehat{\theta}_2) - \widehat{\theta}_2 = D^{-2}D_1^2(\widehat{\theta}_1 - \widehat{\theta}_2).$$
$$2\triangle L = (\widehat{\theta}_2 - \widehat{\theta}_1)^T \Sigma^2 (\widehat{\theta}_2 - \widehat{\theta}_1) + O(r\diamondsuit(r,x)),$$

где

$$\Sigma^2 = D_2^2 D^{-2} D_1^2 D^{-2} D_2^2 + D_1^2 D^{-2} D_2^2 D^{-2} D_1^2 = D_1^2 D^{-2} D_2^2 \approx \frac{1}{4}D^2. \qquad \text{(Sigma)}$$

Предположим, что выполнены неравенства с вероятностью $1 - 2e^{-x}$

$$\|D_1(\widehat{\theta}_1 - \theta_1^*) - \xi_1\| \le \diamondsuit(r,x),$$
$$\|D_2(\widehat{\theta}_2 - \theta_2^*) - \xi_2\| \le \diamondsuit(r,x).$$

Выполнив замены $\widehat{\theta}_2, \widehat{\theta}_1$ в выражении для $\triangle L$ при помощи неравенства

$$|\|a\|^2 - \|b\|^2| \le 2\|a\|\|a-b\| + \|a-b\|^2,$$

приходим к следующему результату.

Теорема 1. Пусть выполнено условие (L^*), а также условия квадратичной аппроксимации Лапласа и с вероятностью $1 - 2e^{-x}$ выполнено условие

$$\|\xi_i\| \le z(x), \quad z^2(x) = p + 6\lambda_{\max}x.$$

Тогда в локальной области с вероятностью $1 - 8e^{-x}$ справедливо равенство

$$2\triangle L = \|\triangle\xi_{12} + \triangle\theta_{12}^*\|^2 + O(\{r + z(x)\}f_D\diamondsuit(r,x)),$$

где

$$\triangle\xi_{12} = \Sigma(D_2^{-1}\xi_2 - D_1^{-1}\xi_1), \quad \triangle\theta_{12}^* = \Sigma(\theta_2^* - \theta_1^*),$$
$$\left\|\Sigma D_i^{-1}\right\|_\infty \le f_D, \quad i = \{1,2\}.$$

Замечание. При увеличении размера выборки $n \to \infty$ имеет место сходимость по распределению

$$\xi_{12} \to \mathcal{N}(0, I_p).$$

Замечание. Так как $\|\triangle\xi_{12}\|^2 \sim p$, то должно выполнятся

$$r\diamondsuit(r,x) \sim \frac{r^2}{\sqrt{n}} = o(p),$$

где n – длина выборки. Иначе Теорема 1 теряет смысл. Таким образом, из условия (L^*) получаем, что разность параметра $\|\theta_2^* - \theta_1^*\|$ не должна превышать величины

$$\frac{\sqrt{p}}{n^{1/4}},$$

так как $\|D\| = O(\sqrt{n})$. Если же пренебречь членом $\|\triangle\xi_{12}\|^2 \sim p$, то приходим к условию

$$\frac{r}{\sqrt{n}} = o(\sqrt{p}) \Rightarrow \|\theta_2^* - \theta_1^*\| = o(\sqrt{p}).$$

Установим аналогичный результат (теорема 1) для статистики $\sqrt{2\triangle L}$. Из условия (a) получаем с вероятностью не менее $1-2e^{-x}$

$$\left|\triangle L(\widehat{\theta}_1,\widehat{\theta}_2)-\frac{1}{2}\|\Sigma(\widehat{\theta}_2-\widehat{\theta}_1)\|^2\right| \le 2\|D_1(\widehat{\theta}_1-\widehat{\theta})\|\diamondsuit(r,x)+2\|D_2(\widehat{\theta}_2-\widehat{\theta})\|\diamondsuit(r,x) \le$$

$$\le 4\|\Sigma(\widehat{\theta}_2-\widehat{\theta}_1)\|\diamondsuit(r,x).$$

Воспользуемся неравенством $|a-b| \le |a^2-b^2|/b,\ b>0$.

$$\left|\sqrt{2\triangle L(\widehat{\theta}_1,\widehat{\theta}_2)}-\|\Sigma(\widehat{\theta}_2-\widehat{\theta}_1)\|\right| \le 8\diamondsuit(r,x).$$

Заменим $(\widehat{\theta}_1,\widehat{\theta}_2)$ на $(D_1^{-1}\xi_1+\theta_1^*,\ D_2^{-1}\xi_2+\theta_2^*)$

$$\Big|\|\Sigma(\widehat{\theta}_2-\widehat{\theta}_1)\|-\|\triangle\xi_{12}+\triangle\theta_{12}^*\|\Big| \le \|\Sigma(\widehat{\theta}_1-\theta_1^*)-\Sigma D_1^{-1}\xi_1\|+\|\Sigma(\widehat{\theta}_2-\theta_2^*)-\Sigma D_2^{-1}\xi_2\| \le 2f_D\diamondsuit(r,x).$$

Резюмируем результат

Теорема 2. Пусть выполнено условие (L^*), а также условия квадратичной аппроксимации Лапласа (a). Тогда с вероятностью не менее $1-4e^{-x}$ в локальной области $\Theta_1(r)\cap\Theta_2(r)$ выполнено неравенство

$$\left|\sqrt{2\triangle L(\widehat{\theta}_1,\widehat{\theta}_2)}-\|\triangle\xi_{12}+\triangle\theta_{12}^*\|\right| \le (8+2f_D)\diamondsuit(r,x).$$

где

$$\triangle\xi_{12}=\Sigma(D_2^{-1}\xi_2-D_1^{-1}\xi_1),\quad \triangle\theta_{12}^*=\Sigma(\theta_2^*-\theta_1^*),$$
$$\left\|\Sigma D_i^{-1}\right\|_\infty \le f_D,\quad i=\{1,2\}.$$

Замечание. Величина $f_D<1$ на погрешность погрешность в теоремах 1, 2 не оказывает большого влияния. Однако, константу при $\diamondsuit(r,x)$ можно уменьшить, раскладывая $L_1(\theta)$, $L_2(\theta)$ и $L(\theta)$ в окрестности точек θ_1^*, θ_2^* и θ^* вместо точек ОМП:

$$2\triangle L=-\|\xi\|^2+\|\xi_1\|^2+\|\xi_2\|^2-2\xi_1^TD_1D^{-2}D_2^2(\theta_2^*-\theta_1^*)+2\xi_2^TD_2D^{-2}D_1^2(\theta_2^*-\theta_1^*)+$$

$$+\|D_1D^{-2}D_2^2(\theta_2^*-\theta_1^*)\|^2+\|D_2D^{-2}D_1^2(\theta_2^*-\theta_1^*)\|^2\pm(2\diamondsuit(r,x)r+2\delta(r)r^2)=$$

$$=-\|\xi\|^2+\|\xi_1\|^2+\|\xi_2\|^2+2(D_2^{-1}\xi_2-D_1^{-1}\xi_1)^T\Sigma^2(\theta_2^*-\theta_1^*)+\|\Sigma(\theta_2^*-\theta_1^*)\|^2\pm(2\diamondsuit(r,x)r+2\delta(r)r^2).$$

Заменим $\|\xi\|^2$ на $\|D^{-1}(D_1\xi_1+D_2\xi_2)\|^2\pm 2\diamondsuit(r,x)z(x)$.

$$-\|\xi\|^2+\|\xi_1\|^2+\|\xi_2\|^2=\|\Sigma(D_2^{-1}\xi_2-D_1^{-1}\xi_1)\|^2\pm 2\diamondsuit(r,x)z(x).$$

Приходим к конечному выражению

$$\left|2\triangle L-\|\triangle\xi_{12}+\triangle\theta_{12}^*\|^2\right| \le (4\diamondsuit(r,x)r+2\delta(r)r^2).$$

2 Бутстреп

Лемма (Неравенство Хефдинга для векторов)**.** Пусть $\mathbb{E}X_i = 0$, $\|X_i\| < c_i$, тогда

$$\mathbb{P}\left(\left\|\sum_{i=1}^{n} X_i\right\| > t\right) \le \exp\left(-\frac{(t-v)^2}{2v^2}\right), \quad t \ge v,$$

$$v^2 = \frac{1}{4}\sum_{i=1}^{n} c_i^2.$$

Следствие.

$$\mathbb{P}\left(\left\|\sum_{i=1}^{n} X_i\right\| > v(1+\sqrt{2x})\right) \le e^{-x}.$$

Лемма (Неравенство для отклонений нормы эмпирического процесса)**.** Пусть выполнены ограничения при $\|\gamma_1\| = \|\gamma_2\| = 1$

$$\log \mathbb{E}\exp\left\{\frac{\lambda}{\omega}\gamma_1 D_0^{-1}\nabla^2\zeta_i(\theta)D_0^{-1}\gamma_2\right\} \le \frac{\lambda^2\nu_i^2}{2}, \quad \sum_i \nu_i^2 = \nu_0^2. \tag{ED}$$

Тогда при $\zeta = \sum_i \zeta_i$ в локальной области радиуса r с вероятностью $1 - e^{-x}$

$$\|D_0^{-1}(\nabla\zeta(\theta_2) - \nabla\zeta(\theta_1))\| \le 12\nu_0 z(x)\omega r.$$

Рассмотрим функционал

$$\alpha^o(\theta_2,\theta_1) = L^o(\theta_2) - L^o(\theta_1) - (\theta_2-\theta_1)^T\nabla L^o(\theta_1) + \frac{1}{2}\|D_0(\theta_2-\theta_1)\|^2$$

Будем подразумевать что дифференцирование производится по θ_2 для функций типа $(\theta_2,\theta_1) \to \mathbb{R}$. Оценим величину его среднего значения и отклонения от среднего

$$\|D_0^{-1}\nabla\mathbb{E}^0\alpha^o(\theta_2,\theta_1)\| = \|D_0^{-1}\nabla\alpha(\theta_2,\theta_1)\| \le 2\diamondsuit(r,x)$$

с вероятностью $1 - e^{-x}$.

$$S(\theta_2,\theta_1) = D_0^{-1}\{\nabla\alpha^o(\theta_2,\theta_1) - \mathbb{E}^o\nabla\alpha^o(\theta_2,\theta_1)\} = \sum_{i=1}^{n} D_0^{-1}(\nabla l_i(\theta_2) - \nabla l_i(\theta_1))(u_i - 1),$$

где $\mathbb{D}\, u_i = 1$, $\mathbb{E}u_i = 1$. Обозначим за $\overset{o}{S} = S(\theta_2,\theta_1) - \mathbb{E}S(\theta_2,\theta_1)$, тогда с вероятностью $1 - e^{-x}$

$$\|\overset{o}{S}\| \le 12\nu_0 z(x)\omega r\sqrt{\sum_{i=1}^{n}(\nu_i^2/\nu_0^2)(u_i-1)^2}.$$

Ограничим u_i:

$$\mathbb{E}e^{(u_i-1)^2/\sigma_u^2} \le e,$$

тогда вероятностью $1 - e^{-t}$

$$\sqrt{\sum_{i=1}^{n}(\nu_i^2/\nu_0^2)(u_i-1)^2} \le \sigma_u\sqrt{1+t}$$

Таким образом, с вероятностью $1 - 2e^{-x}$

$$\|\overset{o}{S}\| \le 12\nu_0 z(x)\omega r\sigma_u\sqrt{1+x}.$$

Наложим ограничение на $\mathbb{E}S(\theta_2, \theta_1)$ в виде

$$\|D_0^{-1} \nabla^2 \mathbb{E} l_i(\theta) D_0^{-1}\| \le \frac{C_i(r)}{n}, \tag{Lm}$$

в результате чего при помощи векторного неравенства Хефдинга получим с вероятностью $1 - e^{-x}$

$$\|\mathbb{E}S(\theta_2, \theta_1)\| \le \frac{1 + \sqrt{2x}}{2n} C(r) r c_u,$$

где $C(r) = \sqrt{\sum_i C_i^2(r)}$, $c_u = \max(u_i - 1)$. В итоге приходим к следующему результату

Теорема 3 (Бутстреп, Вилкс)**.** При выполнении условий (ED) и (Lm), (a), а также ограничений для весов u_i, в локальной области радиуса r имеет место неравенство с вероятностью $1 - 4e^{-x}$

$$|\alpha^o(\theta_2, \theta_1)| \le \|D_0(\theta_2 - \theta_1)\| \left(2\diamondsuit(r, x) + 12\nu_0 z(x) \omega r \sigma_u \sqrt{1 + x} + \frac{1 + \sqrt{2x}}{2n} C(r) r c_u \right) = \tag{a'}$$

$$= \|D_0(\theta_2 - \theta_1)\| \, 2\diamondsuit^o(r, x).$$

Для вывода теоремы Фишера для взвешенной функции правдоподобия в мире бутстрепа введем переменную

$$\chi^o(\theta_2, \theta_1) = D_0^{-1}(\nabla L^o(\theta_2) - \nabla L^o(\theta_1)) + D_0(\theta_2 - \theta_1),$$

из верхней оценки которой легко получить теорему Фишера. Заметим, что

$$\chi^o(\theta_2, \theta_1) = D_0^{-1} \nabla \alpha^o(\theta_2, \theta_1).$$

Тогда, проводя рассуждения аналогичные доказательству теоремы Вилкса, получаем верхнюю оценку $\chi^o(\theta_2, \theta_1)$.

Теорема 4 (Бутстреп, Фишер)**.** При выполнении условий (ED) и (Lm), (a), а также ограничений для весов u_i, в локальной области радиуса r имеет место неравенство с вероятностью $1 - 4e^{-x}$

$$\|\chi^o(\widehat{\theta}^o, \widehat{\theta})\| = \|D_0(\widehat{\theta}^o - \widehat{\theta}) - D_0^{-1} \nabla L^o(\widehat{\theta})\| \le 2\diamondsuit^o(r, x),$$

где $\widehat{\theta}^o$, $\widehat{\theta}$ – ОМП для взвешенной и невзвешенной функций правдоподобия.

Из теорем 4 и 3 можем получить аналог теоремы 2 только для взвешенной функции правдоподобия.

Теорема 5. Пусть выполнено условие (L^*), а также условия квадратичной аппроксимации Лапласа (a'). Тогда с вероятностью не менее $1 - 16e^{-x}$ в локальной области $\Theta_1(r) \cap \Theta_2(r)$ выполнено неравенство

$$\left| \sqrt{2\triangle L(\widehat{\theta}_1^o, \widehat{\theta}_2^o)} - \|\triangle \xi_{12}^o + \triangle \widehat{\theta}_{12}\| \right| \le (8 + 4f_D) \diamondsuit^o(r, x).$$

где

$$\triangle \xi_{12}^o = \Sigma(D_2^{-1} \xi_2^o - D_1^{-1} \xi_1^o), \quad \triangle \widehat{\theta}_{12} = \Sigma(\widehat{\theta}_2 - \widehat{\theta}_1),$$
$$\left\| \Sigma D_i^{-1} \right\|_\infty \le f_D, \quad i = \{1, 2\}.$$

Замечание. Для генерации нормы вектора ξ_{12}^o можно воспользоваться статистикой $T^o = \sqrt{2\triangle L(\widehat{\theta}_1^o, \widehat{\theta}_1^o + \triangle \widehat{\theta}_{12})}$, для которой также применима теорема 5, причем переменная $\triangle \widehat{\theta}_{12} = 0$ для данной статистики, т.е.

$$|T^o - \|\triangle \xi_{12}^o\|| \le (8 + 4f_D) \diamondsuit^o(r, x).$$

Найдем точность оценки распределения нормы вектора $\triangle\xi_{12}+b$ при помощи бутстрепа (аналога $\triangle\xi_{12}^o+b^o$, b, b^o неслучайные), приближая их распределения к нормальному. Последовательность сравнения распределений:

1. $\triangle\xi_{12}^o(\widehat{\theta}_1,\widehat{\theta}_2)\approx\triangle\xi_{12}^o(\theta_1^*,\theta_2^*)$;
2. $\|\triangle\xi_{12}^o(\theta_1^*,\theta_2^*)+b^o\|\approx\|\xi_{12}^o\|$, $\xi_{12}^o\sim\mathcal{N}(b^o,\Sigma^o)$;
3. $\|\triangle\xi_{12}(\theta_1^*,\theta_2^*)+b\|\approx\|\xi_{12}\|$.
4. $\|\xi_{12}^o\|\approx\|\xi_{12}\|$, $\xi_{12}\sim\mathcal{N}(b,\Sigma)$;

1) Рассчитаем отклонение $\triangle\xi_{12}^o(\widehat{\theta}_1,\widehat{\theta}_2)$ от $\triangle\xi_{12}^o(\theta_1^*,\theta_2^*)$. Поскольку

$$\xi_i^o(\theta)=D_0^{-1}\sum_{i=1}^n\nabla l_i(\theta)(u_i-1),$$

то, используя оценку для $S(\theta_2,\theta_1)$, получим неравенство в локальной области, с вероятностью $1-3e^x$, $i=\{1,2\}$

$$\|\xi_i^o(\theta_2)-\xi_i^o(\theta_1)\|\le\left(12\nu_0 z(x)\omega r\sigma_u\sqrt{1+x}+\frac{1+\sqrt{2x}}{2n}C(r)rc_u\right)=2\diamondsuit_\xi^o(r,x).$$

Запишем $\triangle\xi_{12}^o(\theta_1,\theta_2)$ в виде

$$\triangle\xi_{12}^o(\theta_1,\theta_2)=A\begin{pmatrix}\xi_1^o(\theta_1)\\ \xi_2^o(\theta_2)\end{pmatrix},\quad A=\Sigma\left(-D_1^{-1},D_2^{-1}\right).$$

Тогда справедливо утверждение

Лемма.

$$\|\triangle\xi_{12}^o(\widehat{\theta}_1,\widehat{\theta}_2)-\triangle\xi_{12}^o(\theta_1^*,\theta_2^*)\|\le4\diamondsuit_{\xi,F}^o(r,x),$$

где $F=A^TA$, $H_2=H_2(F)+4p$,

$$\diamondsuit_{\xi,F}^o(r,x)=\left(12\nu_0 z_F(x)\omega r\sigma_u\sqrt{1+x}+\frac{1+\sqrt{2x}}{2n}C(r)rc_u\right).$$

2),3) Найдем точность слабой аппроксимации $\|\triangle\xi_{12}^o(\theta_1^*,\theta_2^*)+b^o\|\approx\|\xi_{12}^o\|$, $\xi_{12}^o\sim\mathcal{N}(b^o,\Sigma^o)$.

Будем дополнительно считать, что $\triangle\xi_{12}$, $\triangle\xi_{12}^o$ зависят от координаты t. Выявим насколько хорошо распределение $\max_t\|\triangle\xi_{12}^o(\theta_1^*,\theta_2^*)(t)+b^o(t)\|$ приближается распределением нормы нормального вектора $\max_t\|\xi_{12}^o(t)\|$. Норму вектора S будем выражать через $\max_{\|\gamma\|=1}\gamma^TS$, а максимум через smooth max функцию:

$$h_\beta(x)=\beta^{-1}\log\left(\sum_i e^{\beta x_i}\right).$$

Функция $h_\beta(x)$ обладает свойством при $x\in\mathbb{R}^M$

$$\max_i(x_i)\le h_\beta(x)\le\max_i(x_i)+\frac{\log(M)}{\beta}.$$

Лемма. Для функции g_δ, аппроксимирующей индикатор (g_δ возрастает от 0 до одного на отрезке δ), имеет место неравенство при $\delta=\beta^{-1}\log(M)$

$$g_\delta h_\beta(x-\delta)\le\left[\max_{0\le i\le M}x_i>0\right]\le g_\delta h_\beta(x+\delta).$$

Следовательно, при условии $|\mathbb{E}g_\delta h_\beta(x)-\mathbb{E}g_\delta h_\beta(\widetilde{x})|\le C(\beta,M)\mu_n$

$$\left|\mathbb{P}\left(\max_{0\le i\le M}x_i>0\right)-\mathbb{P}\left(\max_{0\le i\le M}\widetilde{x}_i>\pm2\delta\right)\right|\le C(\delta,M)\mu_n. \tag{1}$$

Вынесем сдвиг $\pm 2\delta$ наружу при помощи леммы об анти-концентрации плотности распределения $\max_i \widetilde{x}_i$.

Лемма. Пусть $x \in \mathcal{N}(m, \Sigma) \in \mathbb{R}^M$, $\sigma_1 \leq \sqrt{\Sigma_{ii}} \leq \sigma_2$, $a_M = \max_i(\widetilde{x}_i - m_i)/\sqrt{\Sigma_{ii}}$, тогда $\forall c$

$$\mathbb{P}(|\max_i \widetilde{x}_i - c| \leq \varepsilon) \leq \frac{4\varepsilon}{\sigma_1}\left(\frac{\sigma_2}{\sigma_1}a_M + \left(\frac{\sigma_2}{\sigma_1} - 1\right)\sqrt{2\log\left(\frac{\sigma_1}{\varepsilon}\right)} + 2 - \frac{\sigma_1}{\sigma_2}\right) \approx 4\varepsilon\frac{\sigma_2}{\sigma_1^2}\sqrt{2\log\left(\frac{\sigma_1 M}{\varepsilon}\right)}.$$

Используя предыдущую лемму и выражение 1, получаем соотношение для разности вероятностных мер максимумов случайных векторов, один из которых имеет нормальное распределение

$$\left|\mathbb{P}\left(\max_{0\leq i\leq M} x_i > 0\right) - \mathbb{P}\left(\max_{0\leq i\leq M} \widetilde{x}_i > 0\right)\right| \leq C(\delta, M)\mu_n + \delta C_{\text{ak}}(M, \Sigma). \tag{2}$$

Проведем теперь расчет верхней границы в выражении

$$|\mathbb{E}g_\delta h_\beta(x) - \mathbb{E}g_\delta h_\beta(\widetilde{x})| \leq C(\beta, M)\mu_n.$$

Лемма. Для функции $f = g_\delta h_\beta$, где g_δ имеет ограниченные производные до третьего порядка, имеет место соотношение при $\delta = \beta^{-1}\log(M)$

$$\left|f(x+d) - f(x) - d^T f'(x) - d^T f''(x)d/2\right| \leq C(\delta, M)\|d\|_\infty^3,$$

где

$$C(\delta, M) = \frac{1}{6\delta^3}(|g'''| + \log(M)|g''| + \log^2(M)|g''|).$$

Пусть

$$\triangle\xi_{12}^o(\theta_1^*, \theta_2^*)(t) + b^o(t) = \xi(t) = \sum_{i=1}^n c_i(t)\varepsilon_i,$$

где $\mathbb{D}\sum_i c_i(t)\varepsilon_i = \sum_i c_i^2(t)V_i^2$.

При помощи приближения ε_i нормальными с.в. с такими же математическими ожиданиями и корреляционными матрицами получаем неравенство

$$|\mathbb{E}g_\delta h_\beta(\xi) - \mathbb{E}g_\delta h_\beta(\widetilde{\xi})| \leq C(\delta, M)\mu_n, \tag{3}$$

где

$$\mu_n = \sum_i (\|\varepsilon_i\|_\infty^3 + C\log^{3/2}(M)\sigma_i^3)\|c_i\|_\infty^3$$

Лемма. Пусть X, Y независимые нормальные вектора с параметрами m_X, m_Y, Σ_X, Σ_Y. Определим вспомогательные переменные

$$\triangle m = m_2 - m_1, \quad \triangle\Sigma = \Sigma_2 - \Sigma_1.$$

Справедливо неравенство

$$|\mathbb{E}g(\delta^{-1}h(X)) - \mathbb{E}g(\delta^{-1}h(Y))| \leq \left(\frac{\beta\|g'\|_\infty}{\delta} + \frac{\|g''\|_\infty}{2\delta^2}\right)\|\triangle\Sigma\|_\infty + \frac{\|g'\|_\infty}{2\delta}\|\triangle m\|_\infty.$$