*УДК 519.254*

# ОБ ЭФФЕКТИВНОСТИ ОЦЕНОК ПРИ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ НЕЛИНЕЙНЫХ СТАТИЧЕСКИХ ОБЪЕКТОВ

©2004 О. А. Кацюба, М. Б. Линеенко

Самарская государственная академия путей сообщения

Рассматривается задача параметрического оценивания стохастических статических моделей, нелинейных относительно параметров. В качестве метода оценивания применяется метод квазиправдоподобных оценок; дано доказательство состоятельности и асимптотической нормальности оценок максимального квазиправдоподобия. Показано, что применение линейно-комбинированной оценки позволяет получить эффективность оценивания как угодно близкую к эффективности оценок максимального правдоподобия в условиях отсутствия априорной информации о законах распределения помех наблюдений.

Пусть имеет место стандартная задача несмещенного оценивания регрессионных функций; наблюдаемые величины $y_i$ $(i=\overline{1,n})$ представляется в виде $y_i = u(x_i, a_0) + \varepsilon(i)$, где $u(x_i, a_0)$ - функция, вид которой известен, $\dim a_0 = p$, $\varepsilon(i)$ - независимые случайные величины, удовлетворяющие условиям $M(\varepsilon(i)) = 0$, $M(\varepsilon^2(i)) \le \sigma_i^2$, ($M$ – оператор математического ожидания).

Требуется определить состоятельные оценки $\hat{a}(n)$, не зная априорно закона распределения шума

$$f_i(\varepsilon(i)) = f_i(y_i, a_0) = f_{\varepsilon_i}(y_i - u(x_i, a_0)).$$

В условиях априорной неопределенности существуют методы состоятельного оценивания (метод эмпирического риска, М-оценка и т.д.), однако очевидно, что эффективность этих оценок в общем случае далека от эффективности оценки максимального правдоподобия. В связи с этим возникает следующая задача: создание методов и алгоритмов определения таких состоятельных несмещенных оценок на основе множества оценок (линейно-комбинированных оценок), эффективность которых в широком смысле была бы в принципе достаточно близка к эффективности оценок максимального правдоподобия. Наиболее просто вопрос получения множества состоятельных оценок, из которых может быть сконструирована достаточно эффективная оценка, решается, если воспользоваться методом эмпирического риска, когда эмпирический функционал конструируется на основе некоторых законов распределения, и тогда указанный метод можно трактовать как непосредственное расширение метода максимального правдоподобия на случай априорной неопределенности при модификации некоторых его условий (такие оценки названы квазиправдоподобными), что позволяет получить более конструктивные условия состоятельности и асимптотической нормальности оценок, причем, этот метод распространен на случай неоднородных наблюдений, а также на задачу оценивания параметров линейных разностных уравнений при наличии помех в выходных переменных.

Отдельные вопросы состоятельности квазиправдоподобных оценок и теории линейно-комбинированных оценок рассмотрены в [1] и [2]. Предлагаемая статья дает возможность рассмотреть вопросы эффективности оценивания параметров нелинейных моделей в целом.

Пусть эксперименты $\{x^{(n)}, A_u^{(n)}, F_a^{(n)}\}$ порождаются наблюдениями $Y = (y_1,...,y_n)^T$ со значениями в $\{x, A_u\}$ и распределениями $\{F_a\}$, пусть семейство $F^{(n)}$ доминируется некоторой мерой $\upsilon$ и существует функция

$$\frac{\partial F_a^{(n)}(y_1,...,y_n;a)}{\partial\upsilon(y_1,...,y_n)} = \frac{\partial F_a^{(n)}(Y^T,a)}{\partial\upsilon} = f_{(n)}(Y^T,a).$$

*Определение 1.* Функцией квазиправдоподобия параметров $a$, отвечающей $\{x^{(n)}, A_u^{(n)}, F_a^{(n)}\}$ и $Y$, называется неотрицательная вещественная функция

$$\phi_{(n)}(Y^T, a) \neq f_{(n)}(Y^T, a),\ a \in A,$$

где

$$\phi_n(Y^T, a) = \phi^{(n)}_{\varepsilon_1,\dots\varepsilon_n}(y_1 - u(x_1, a), \dots, y_n - u(x_n, a)).$$

Оценкой максимального квазиправдоподобия (МКП) $a_0$ для заданной функции квазиправдоподобия $\phi_{(n)}(Y^T, a)$ по наблюдениям $Y$ называется $\hat{a}(n, \phi)$, определяемая из

$$\phi_{(n)}(Y^T, \hat{a}(n, \phi)) = \sup_{a \in \widetilde{A}} \phi_{(n)}(Y^T, a).$$

В частности, если

$$\phi_{(n)}(Y^T, a) = f_{(n)}(Y^T, a),$$

то оценка будет правдоподобной.

Оценку, при которой функция $\phi_{(n)}(Y^T, a)$ принимает абсолютно наибольшее значение на $\widetilde{A}$, назовем оценкой МКП в строгом смысле. Пусть функция $\phi_i(y_i, a)$ имеет частные производные по всем аргументам $a^{(j)}$ ($j = \overline{1, p}$), тогда функция

$$\ln \phi_{(n)}(Y^T, a) = \sum_{i=1}^{n} \ln \phi_i(y_i, a)$$

(в случае независимых случайных величин $y_i$) также имеет частные производные. Если для некоторого $a \in \widetilde{A}$ эта функция достигает максимума, то в этой точке

$$\sum_{i=1}^{n} \frac{\partial \ln \phi_i(y_i, a)}{\partial a^{(j)}} = 0. \quad (1)$$

Система уравнений (1) называется системой уравнения МКП. Ясно, что всякая оценка МКП в строгом смысле удовлетворяет системе уравнений МКП, поэтому любой корень системы уравнений (1), лежащий в $\widetilde{A}$, называется оценкой МКП в широком смысле. Из нижеизложенного следует, что оценка МКП, как в строгом смысле, так и в широком смысле не единственна, причем оценка МКП в широком смысле не обязательно тождественна оценке МКП в строгом смысле. В дальнейшем будем предполагать, что для всех $n$, за исключением конечного числа значений, существует самое большее одна оценка МКП.

Рассмотрим доказательство состоятельности квазиправдоподобных оценок в условиях независимых неоднородных наблюдений, причем докажем состоятельность оценок максимального квазиправдоподобия в строгом смысле.

Известная ограниченность доказательства состоятельности, основанного на “локальных” суждениях, не связанных с понятием точной верхней грани функции МКП заключается в том, что, во-первых, имеет место требование единственности корня, во-вторых, требуется выполнение довольно жестких ограничений на поведение функции квазиправдоподобия (существование производных и т.д.), поэтому далее рассмотрено доказательство состоятельности оценок МКП, не требующее существования производных и основанное на понятии точной верхней грани.

Введем следующие обозначения: $S_r(a)$ - замкнутый шар в $\overline{A}$ с центром в точке $a$ и радиусом $r$;

$$\phi_i^*(y_i, a, r) = \sup_{\bar{a} \in S_r(a)} \phi_i(y_i, \bar{a}),$$

$$\phi_i^{**}(y_i, \upsilon) = \sup_{\|\bar{a}\| > \upsilon} \phi_i(y_i, \bar{a}),\ \upsilon > 0;$$

$$\bar{\phi}_i^*(y_i, a, r) = \begin{cases} \phi_i^*(y_i, a, r), & \varphi_i^*(y_i, a, r) \geq 1 \\ 1, & \text{в противном случае} \end{cases}.$$

*Утверждение 1.* Пусть выполняются следующие условия:

$1^0$. При каждом $Y^T$ функция $\phi_{(n)}(Y^T, a)$ непрерывна относительно $a$ в $\overline{A}$ ($\overline{A}$ - замкнутое множество).

$2^0$. Вектора $x_i$ - неслучайные.

$3^0$. Для каждого малого $r > 0$ выполняются

$$M\left[\log \bar{\phi}_i^*(y_i, a, r)\right] \text{ - конечно} \quad (2)$$

и $\dfrac{1}{n}\sum_{i=1}^{n} M\left[\log \bar{\phi}_i^*(y_i, a, r)\right]$ сходится к $\lim\limits_{n\to\infty} \dfrac{1}{n}\sum_{i=1}^{n} M\left[\log \bar{\phi}_i^*(y_i, a, r)\right]$ равномерно относительно $r$.

$4^0$. Для каждого $a_0 \in \overline{A}$ существует $\lim\limits_{n\to\infty} \frac{1}{n}\sum\limits_{i=1}^{n} M\left[\log\phi_i(\varepsilon_i)\right]$.

$5^0$. Для всех $Y$ и каждого $a_0 \in \overline{A}$, $a \neq a_0$

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i(y_i,a)\right] <$$

$$< \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i(\varepsilon_i)\right];$$

$$\lim_{\upsilon\to\infty}\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i^{**}(y_i,\upsilon)\right] = -\infty.$$

$6^0$. При каждом $a \in \overline{A}$ и $r$, $\upsilon$ для дисперсии величины $\log\phi_i^{*}(y_i,a,r)$, $\log\phi_i(\varepsilon_i)$, $\log\phi_i^{**}(y_i,\upsilon)$ выполняются условия теоремы Маркова [3, с.209].

Тогда для любого замкнутого множества $A' \subset \overline{A}$ и $a_0 \notin A'$ и любого $\zeta > 0$, $\delta > 0$ имеет место

$$q_{a_0}\left(\left\| \frac{\sup\limits_{a\in A'} \phi_{(n)}(Y^T,a)}{\phi_{(n)}(Y^T,a_0)} \right\| < \zeta\right) > 1-\delta,$$

если $n > n(\zeta,\delta)$, где $q_{a_0}$ - значение вероятности.

*Доказательство утверждения 1.* Из предположения $1^0$ следует, что

$$\lim_{r\to 0} \log\overline{\phi}_i^{*}(y_i,a,r) = \log\overline{\phi}_i^{*}(y_i,a); \quad (3)$$

$$\overline{\phi}_i^{*}(y_i,a) = \begin{cases} \phi_i(y_i,a), \text{ если } \phi_i(y_i,a) > 1; \\ 1, \text{ в противном случае.} \end{cases}$$

Так как $\log\overline{\phi}_i^{*}(y_i,a,r)$ – невозрастающая функция $r$, тогда из (2) и (3) [4, с.202]

$$\lim_{r\to 0} M\left[\log\overline{\phi}_i^{*}(y_i,a,r)\right] = M\left[\log\overline{\phi}_i^{*}(y_i,a)\right].$$

Аналогично доказательство и для случая

$$\overline{\overline{\phi}}_i^{*}(y_i,a) = \begin{cases} 1, \text{ в противном случае;} \\ \phi_i(y_i,a), \text{ если } \phi_i(y_i,a) \le 1, \end{cases}$$

тогда

$$\lim_{r\to 0} M\left[\log\phi_i^{*}(y_i,a,r)\right] = M\left[\log\phi_i^{*}(y_i,a)\right]. \quad (4)$$

Из п. $3^0$ следует возможность почленного перехода к пределу [5] и, воспользовавшись (4), получаем

$$\lim_{r\to 0}\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i^{*}(y_i,a,r)\right] =$$

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i^{*}(y_i,a,r)\right] =$$

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i(y_i,a)\right]. \quad (5)$$

Далее выберем $\upsilon = \upsilon_0$ так, чтобы

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i^{**}(y_i,\upsilon_0)\right] <$$

$$< \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i(y_i,a)\right]. \quad (6)$$

Существование такого вытекает из $5^0$. Обозначим через $A'' \subset A'$ множество, состоящее из всех точек $a$ из $A'$, для которых выполняется неравенство $\|a\| \le \upsilon_0$, тогда в соответствии с (5), (6) для каждой точки $a \in A''$ поставим в соответствие значение $r_a$, что

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i^{*}(y_i,a,r)\right] <$$

$$< \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M\left[\log\phi_i(\varepsilon_i)\right]. \quad (7)$$

Множество $A''$ – компактно, поэтому [6, с.193] в $A''_\mu$ найдется конечное число точек $a_1$, …, $a_\mu$, что $A'' \subset \bigcup\limits_{k=1}^{\mu} S_{r_{a(k)}}(a_k)$. Очевидно, что для каждого $Y^T$ имеем

$$0 \le \sup_{a\in A'} \phi_{(n)}\left(Y^T,a\right) \le \sum_{k=1}^{\mu}\prod_{i=1}^{n} \phi_i^{*}\left(y_i,a_k,r_k\right) +$$

$$+ \prod_{i=1}^{n} \phi_i^{**}(y_i,\upsilon_0).$$

Следовательно, достаточно доказать, что для любых $\zeta'$, $\delta'$, $\zeta''$, $\delta''$

$$q_{a_0}\left(\left\| \frac{\prod\limits_{i=1}^{n} \phi_i^{*}(y_i,a_k,r_k)}{\phi_{(n)}(Y^T,a_0)} \right\| < \zeta'\right) > 1-\delta', \quad (8')$$

$n > n(\zeta', \delta', a_k)$, $k = \overline{1, \mu}$;

$$q_{a_0}\left(\left\|\frac{\prod_{i=1}^{n}\phi_i^{**}(y_i, \upsilon_0)}{\phi_{(n)}(Y^T, a_0)}\right\| < \zeta''\right) > 1 - \delta'', \quad (8")$$

$n > n(\zeta'', \delta'', \upsilon_0)$, так как $\mu$ не зависит от $n$.

Из условия теоремы и теоремы Слуцкого [7, с.282] следует, что

$$\frac{1}{n}\sum_{i=1}^{n}\log\phi_i^{*}(y_i, a_k, r_k) - \frac{1}{n}\sum_{i=1}^{n}\log\phi_i(y_i, a_0) \xrightarrow[n\to\infty]{п.в.}$$

$$\frac{1}{n}\sum_{i=1}^{n}M\left[\log\phi_i^{*}(y_i, a_k, r_k)\right] - \frac{1}{n}\sum_{i=1}^{n}M\left[\log\phi_i(y_i, a_0)\right];$$

$$\frac{1}{n}\sum_{i=1}^{n}\log\phi_i^{**}(y_i, \upsilon_0) - \frac{1}{n}\sum_{i=1}^{n}\log\phi_i(y_i, a_0) \xrightarrow[n\to\infty]{п.в.}$$

$$\frac{1}{n}\sum_{i=1}^{n}M\left[\log\phi_i^{**}(y_i, \upsilon_0)\right] - \frac{1}{n}\sum_{i=1}^{n}M\left[\log\phi_i(y_i, a_0)\right].$$

Однако по условиям (6), (7) эти величины отрицательны, поэтому

$$\sum_{i=1}^{n}\log\phi_i^{*}(y_i, a_k, r_k) - \sum_{i=1}^{n}\log\phi_i(y_i, a_0) \xrightarrow[n\to\infty]{п.в.} -\infty;$$

$$\sum_{i=1}^{n}\log\phi_i^{**}(y_i, \upsilon_0) - \sum_{i=1}^{n}\log\phi_i(y_i, a_0) \xrightarrow[n\to\infty]{п.в.} -\infty,$$

откуда

$$q_{a_0}\left(\left(\sum_{i=1}^{n}\log\phi_i^{*}(y_i, a_k, r_k) - \sum_{i=1}^{n}\log\phi_i(y_i, a_0)\right) < \log\zeta'\right) >$$

$$> 1 - \delta';$$

$$q_{a_0}\left(\left(\sum_{i=1}^{n}\log\phi_i^{*}(y_i, \upsilon_0) - \sum_{i=1}^{n}\log\phi_i(y_i, a_0)\right) < \log\zeta''\right) >$$

$$> 1 - \delta'',$$

что доказывает (8') и (8").

Аналогично доказывается неравенство для любых $k = \overline{1, \mu}$. Так как $\mu$ не зависит от $n$, то можно найти такое $n^*(\delta_1, \zeta_1)$, что одновременно выполняются все неравенства, из чего следует доказательство теоремы.

*Следствие 1.* Пусть выполняются все предположения теоремы и пусть для каждого $n \ge 1$ определена на $\chi^{(n)}$ измеримая функция $\hat{a}(n, \phi)$ такая, что для всех $Y^T \in \chi^{(n)}$ выполняется равенство

$$\sup_{a\in\overline{A}}\phi_{(n)}(Y^T, a) = \phi_{(n)}(Y^T, \hat{a}(n, \phi)).$$

Тогда $\{\hat{a}(n, \phi)\}$ – состоятельная последовательность оценивающих функций для $a \in \overline{A}$.

*Доказательство следствия 1.* По определению $\hat{a}(n, \phi)$

$$\phi_{(n)}(Y^T, \hat{a}(n, \phi)) \Big/ \phi_{(n)}(Y^T, a_0) \ge 1$$

для всех $n$ и всех $y_1, \ldots, y_n$. Предположим, что $\{\hat{a}(n, \phi)\}$ несостоятельная последовательность, тогда можно выбрать последовательность целых чисел $0 < n_1 < n_2 < \ldots$ и указать два положительных числа $\zeta_0$, $\delta_0$, таких, что $q_{a_0}\left(\left\|\hat{a}(n_i, \phi - a_0\right\| < \zeta_0\right) \le 1 - \delta_0$, $i = 1,2,\ldots$,

т.е. имеют место такие множества $\chi^{(n_i)}$, во всех точках $Y_{n_i}$ в которых с определенной вероятностью $\left\|\hat{a}(n, \phi) - a_0\right\| \ge \zeta_0$, откуда следует $Y_{n_i}^T \in \chi^{(n_i)}$:

$$\sup_{\|a-a_0\|\ge\varepsilon_0, a\in\overline{A}}\phi_{(n_i)}(Y_{n_i}^T, a) \ge \phi_{(n_i)}(Y_{n_i}^T, \hat{a}(n_i, \phi))$$

и

$$\sup_{\|a-a_0\|\ge\zeta_0, a\in\overline{A}}\frac{\phi_{(n_i)}(Y_{n_i}^T, a)}{\phi_{(n_i)}(Y_{n_i}^T, a_0)} \ge 1$$

для неограниченно больших $n_i$. Так как множество $\overline{A} \cap \left\{a : \left\|a - a_0\right\| \ge \zeta_0\right\}$ – замкнуто, то по утверждению 1 множества $\chi^{(n_i)}$ должны иметь произвольную малую $q_{a_0}$ меру, если $n_i$ достаточно велико, то это приводит к противоречию и доказательству утверждения. Тогда $\{\hat{a}(n, \phi)\}$ – состоятельная последовательность оценивающих функций для $a \in \overline{A}$.

Пусть имеется возможность получить последовательность состоятельных оценок с разной асимптотической эффективностью, а для сравнения эффективности этих оценок с эффективностью оценок максимального правдоподобия необходимо знать дисперсионные матрицы (хотя бы при $n \to \infty$).

Как известно, оценки максимального правдоподобия при определенных условиях регулярности [7] являются асимптотически эффективными оценками, поэтому, определив условия асимптотической нормальности оценки МКП, можно сравнить их эффективность с эффективностью оценки максимального правдоподобия.

*Утверждение 2.* Пусть выполняются следующие условия.

$1^0$. Все условия состоятельности (утверждение 1).

$2^0$. При всех $a \in A$ и любых $n$ существуют и единственны при $a = a_0$ первые и вторые производные $\ln \phi_i(y_i, a)$ и соответственно матрицы

$$H(Y^T, a) = \left| \frac{1}{n} \sum_{i=1}^{n} \frac{\partial \ln \phi_i(y_i, a)}{\partial a^{(j)}} \right|,$$

$$J(Y^T, a) = \left| \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \ln \phi_i(y_i, a)}{\partial a^{(m)} \partial a^{(j)}} \right|,$$

$$M\left\{ \sum_{i=1}^{n} H(y_i, a) H^T(y_i, a) \Big/ a = a_0 \right\} = W_n,$$

$j, m = \overline{1, p}$, $W_n$ - положительно определенная матрица.

$3^0$. Для любых $a \in A$ существует момент

$$M\left\{ \left\| \frac{\partial \ln \phi_i(y_i, a)}{\partial a} \Big/ a = a_0 \right\|^{2+\delta} \right\}$$

и

$$\lim_{n \to \infty} \sum_{i=1}^{n} M\left\{ \left\| [W_n]^{-1/2} \frac{\partial \ln \phi_i(y_i, a)}{\partial a} \Big/ a = a_0 \right\|^{2+\delta} \right\} = 0,$$

(условие Ляпунова), где $\delta$ – любое положительное число, тогда оценка МКП является асимптотически несмещенной и асимптотически нормальной с ковариационной матрицей (при $n \to \infty$)

$$\left\{ \sum_{i=1}^{n} M\left( \frac{\partial^2 \ln \phi_i^{(S)}(y_i, a)}{\partial a^{(m)} \partial a^{(j)}} \Big/ a = a_0 \right) \right\}^{-1} \times$$

$$\times W_n(a_0) \left\{ M\left( \sum_{i=1}^{n} \frac{\partial^2 \ln \phi_i^{(l)}(y_i, a)}{\partial a^{(m)} \partial a^{(j)}} \Big/ a = a_0 \right) \right\}^{-1}.$$

*Доказательство утверждения 2.* Пусть для всех $a \in A$ и каждого $y_i$ существует положительная дважды дифференцируемая функция $g(a)$ и такая функция $u_\phi(y_i)$, что

$$p \left\| \frac{\partial^2}{\partial a^{(d)} \partial a^{(m)}} \left( g(a) \frac{\partial \ln \phi_i(y_i, a)}{\partial a^{(j)}} \right) \right\| < u_\phi(y_i)$$

для всех $d$, $m$, $j$. Разложив функцию $\frac{1}{n} g(a) \sum_{i=1}^{n} \frac{\partial \ln \phi_i(y_i, a)}{\partial a}$ в ряд Тейлора относительно $a = a_0$, получаем

$$g(a) H(Y^T, a) = g(a_0) H(Y^T, a_0) +$$
$$+ g(b_0) J(Y^T, b_0) \widetilde{b} +$$
$$+ \left| \frac{1}{n} \sum_{i=1}^{n} \frac{\partial g(a_0)}{\partial a^{(m)}} \frac{\partial \ln \phi_i(y_i, a_0)}{\partial a^{(j)}} \right| \widetilde{a} +$$
$$+ \frac{1}{2} P(a, a_0, Y^T) \widetilde{a}^T \widetilde{a}, \tag{9}$$

где $P^{(j)}(a, a_0, Y^T) =$

$$= \frac{1}{n} \sum_{i=1}^{n} P^{(j)}(y_i, a, a_0) u_\phi(y_i) \left\| P^{(j)}(y_i, a, a_0) \right\| < 1.$$

Подставим $\hat{a}(n, \phi)$ в (9) и приравняем к 0:

$$g(a_0) H(Y^T, a_0) + g(a_0) J(Y^T, a_0) \times$$
$$\times (\hat{a}(n, \phi) - a_0) + \left| \frac{1}{n} \sum_{i=1}^{n} \frac{\partial g(a_0)}{\partial a^{(m)}} \frac{\partial \ln \phi_i(y_i, a)}{\partial a^{(j)}} \right| \times$$
$$\times (\hat{a}(n, \phi) - a_0) + \frac{1}{2} P(\hat{a}(n, \phi), a_0, Y^T) \times$$

$$(\hat a(n,\phi)-a_0)^T(\hat a(n,\phi)-a_0),$$

откуда

$$g(a_0)H(Y^T,a_0)=-\Big\{g(a_0)J(Y^T,a_0)+$$

$$+\left|\frac{1}{n}\sum_{i=1}^{n}\frac{\partial g(a_0)}{\partial a^{(m)}}\frac{\partial\ln\phi_i(y_i,a)}{\partial a^{(j)}}\right|\Big\}\times$$

$$\times(\hat a(n,\phi)-a_0)+\frac{1}{2}P(\hat a(n,\phi),a_0,Y^T)\times$$

$$\times(\hat a(n,\phi)-a_0)^T(\hat a(n,\phi)-a_0)=$$

$$=-\Big\{g(a_0)J(Y^T,a_0)+$$

$$+\left|\frac{1}{n}\sum_{i=1}^{n}\frac{\partial g(a_0)}{\partial a^{(m)}}\frac{\partial\ln\phi_i(y_i,a)}{\partial a^{(j)}}\right|+$$

$$+\frac{1}{2}P(\hat a(n,\phi),a_0,Y^T)(\hat a(n,\phi)-a_0)^T\Big\}\times$$

$$\times(\hat a(n,\phi)-a_0).$$

Окончательно

$$(\hat a(n,\phi)-a_0)=-\Big\{g(a_0)J(Y^T,a_0)+$$

$$+\left|\frac{1}{n}\sum_{i=1}^{n}\frac{\partial g(a_0)}{\partial a^{(m)}}\frac{\partial\ln\phi_i(y_i,a)}{\partial a^{(j)}}\right|+$$

$$+\frac{1}{2}P(\hat a(n,\phi),a_0,Y^T)(\hat a(n,\phi)-a_0)^T\Big\}^{-1}\times$$

$$\times g(a_0)H(Y^T,a_0). \qquad (10)$$

По теореме Слуцкого

$$\frac{1}{2}P(\hat a(n,\phi),a_0,Y^T)(\hat a(n,\phi)-a_0)^T\xrightarrow[n\to\infty]{п.в.}0$$

$(0:p\times p)$.

Как уже было показано выше,

$$\left|\frac{1}{n}\sum_{i=1}^{N}\frac{\partial g(a_0)}{\partial a^{(m)}}\frac{\partial\ln\phi_i(y_i,a)}{\partial a^{(j)}}\right|\xrightarrow[n\to\infty]{п.в.}0.$$

Таким образом, знаменатель выражения (10) сходится к некоторой постоянной матрице $-g(a_0)J(Y^T,a_0)$.

Исследуем числитель выражения (10): при выполнении условий Ляпунова случайный вектор $[W_n]^{-1/2}\sum_{i=1}^{n}\frac{\partial\ln\phi_i(y_i,a)}{\partial a}$ асимптотически нормален с нулевым математическим ожиданием и ковариационной матрицей в форме единичной матрицы, откуда $\sum_{i=1}^{n}\frac{\partial\ln\phi_i(y_i,a)}{\partial a}$ имеет нормальный закон распределения с нулевым математическим ожиданием и дисперсионной матрицей $W_n$, далее, $\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\ln\phi_i(y_i,a)}{\partial a}$ – нормальный закон распределения с дисперсионной матрицей $\frac{W_n}{n^2}$.

Таким образом, числитель выражения (10) распределен по $p$-нормальному закону с нулевым математическим ожиданием и дисперсионной матрицей $\frac{W_n}{n^2}$, откуда [7, с.281], получаем, что распределение величины (10) – $p$-нормальное с нулевым математическим ожиданием и дисперсионной матрицей

$$\left\{\sum_{i=1}^{n}M\left(\left.\frac{\partial^2\ln\phi_i^{(S)}(y_i,a)}{\partial a^{(m)}\partial a^{(j)}}\right/_{a=a_0}\right)\right\}^{-1}\times$$

$$\times W_n(a_0)\left\{M\left(\sum_{i=1}^{n}\left.\frac{\partial^2\ln\phi_i^{(l)}(y_i,a)}{\partial a^{(m)}\partial a^{(j)}}\right/_{a=a_0}\right)\right\}^{-1}.$$

Если же оценка правдоподобная, то эта матрица равна (при $n\to\infty$) $W_n^{-1}(a_0)$ (при условии регулярности в смысле вторых частных производных по $a$ для каждого "$i$").

Пусть имеется возможность получить последовательность состоятельных оценок с разной асимптотической эффективностью, при этом возникает задача повышения эффективности оценок путем комбинации их.

*Определение 2.* Любая линейная комбинация конечного числа квазиправдоподобных оценок $\hat a(n,\phi^{(l)})$ $(l=\overline{1,k})$ относительно разных функций квазиправдоподобия $\phi^{(l)}$ с определенными весами называется линейно-комбинированной квазиправдоподобной $\alpha_k$.

Покажем, что определенных условиях,

накладываемых на веса линейно-комбинированной оценки, имеет место следующее неравенство: $\operatorname{leff} \alpha_{k-1}/a_0 \le \operatorname{leff} \alpha_k/a_0$, причем, если в последовательности $\{\hat{a}(n,\phi^{(l)})\}$ имеет место $\hat{a}(n,\phi^{(l)}) \equiv \hat{a}(n,f)$ ($f(y_i,a)$ – закон распределения помех наблюдений), то $\alpha_k = \alpha_f$ (под значением $\operatorname{leff} \alpha_{k-1}/a_0$ понимается отношение значений обобщенной дисперсии оценки максимального правдоподобия и исследуемой оценки).

Для последовательности квазиправдоподобных оценок имеет место утверждение 3.

*Утверждение 3*. Пусть имеет место последовательность состоятельных, асимптотически несмещенных оценок $\hat{a}(n,\phi^{(1)})$ … $\hat{a}(n,\phi^{(k)})$, при всех $a \in A$ имеет место равенство

$$\int_{-\infty}^{\infty} \frac{\partial \ln \phi^{(l)}(y_i,a)}{\partial a^{(j)}} f(y_i,a)dy_i = 0$$

и выполняется условие дифференцирования по параметру $a$ под знаком интеграла (в частном случае, когда $\phi_i^{(l)} = f_i$ имеет место условие регулярности в смысле вторых производных по $a$), то линейно-комбинированная оценка

$$\alpha_k = \left[\begin{vmatrix} 1^T & 0^T & \cdots & 0^T \\ 0^T & 1^T & \cdots & 0^T \\ \vdots & \vdots & & \vdots \\ 0^T & 0^T & \cdots & 1^T \end{vmatrix} \Theta^{-1} \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix}\right]^{-1} \times$$

$$\times \begin{vmatrix} 1^T & 0^T & \cdots & 0^T \\ 0^T & 1^T & \cdots & 0^T \\ \vdots & \vdots & & \vdots \\ 0^T & 0^T & \cdots & 1^T \end{vmatrix} \Theta^{-1} \begin{vmatrix} \hat{a}_1[n,\phi^{(l)}] \\ \vdots \\ \hat{a}_p[n,\phi^{(l)}] \end{vmatrix}$$

будет состоятельной, асимптотически несмещенной с асимптотической эффективностью $\operatorname{leff} \alpha_{k-1}/a_0 \le \operatorname{leff} \alpha_k/a_0$, причем, если имеет место $\hat{a}(n,\phi^{(S)}) \equiv \hat{a}(n,f)$ ($S = \overline{1,k}$), то $\alpha_k \equiv \hat{a}(n,f)$, и $\operatorname{leff} \alpha_k/a_0 = 1$, где

$$\hat{a}_j[n,\phi^{(l)}] = \begin{vmatrix} \hat{a}^{(j)}(n,\phi^{(1)}) \\ \hat{a}^{(j)}(n,\phi^{(2)}) \\ \vdots \\ \hat{a}^{(j)}(n,\phi^{(k)}) \end{vmatrix};$$

$$\Theta = \begin{vmatrix} \operatorname{cov}(\hat{a}^{(1)}(n,\phi^{(l)}),\hat{a}^{(1)}(n,\phi^{(S)})) & \cdots \\ \operatorname{cov}(\hat{a}^{(2)}(n,\phi^{(l)}),\hat{a}^{(1)}(n,\phi^{(S)})) & \cdots \\ \vdots & \ddots \\ \operatorname{cov}(\hat{a}^{(p)}(n,\phi^{(l)}),\hat{a}^{(1)}(n,\phi^{(S)})) & \cdots \end{vmatrix} \to$$

$$\to \begin{vmatrix} \cdots & \operatorname{cov}(\hat{a}^{(1)}(n,\phi^{(l)}),\hat{a}^{(p)}(n,\phi^{(S)})) \\ \cdots & \operatorname{cov}(\hat{a}^{(2)}(n,\phi^{(l)}),\hat{a}^{(p)}(n,\phi^{(S)})) \\ \ddots & \vdots \\ \cdots & \operatorname{cov}(\hat{a}^{(p)}(n,\phi^{(l)}),\hat{a}^{(p)}(n,\phi^{(S)})) \end{vmatrix}$$

- ковариационная матрица квазиправдоподобных оценок $\hat{a}(n,\phi^{(l)})$ (положительно определенная, невырожденная, $pk \times pk$), порядок матрицы: $1: k \times 1$; $0: k \times 1$; $\hat{a}_j(n,\phi^{(l)}): k \times 1$.

*Следствие утверждения 3*. Ковариационная матрица линейно-комбинированных оценок вычисляется по формуле

$$D[\alpha_k] = \left[\begin{vmatrix} 1^T & 0^T & \cdots & 0^T \\ 0^T & 1^T & \cdots & 0^T \\ \vdots & \vdots & & \vdots \\ 0^T & 0^T & \cdots & 1^T \end{vmatrix} \Theta^{-1} \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix}\right]^{-1}.$$

Доказательство основных свойств полученных линейно-комбинированных оценок, исходя из свойств составляющих их квазиправдоподобных оценок, в том числе состоятельность, асимптотическая несмещенность, значение меры эффективности $\operatorname{leff} \alpha_k/a_0$ относительно наблюдений $n$ приведено в [8].

Достаточно сложным вопросом является выбор функции $\phi$ для получения квазиправдоподобных оценок, при этом функции $\phi$ должны обладать свойствами, описанными в утверждении 2. Было бы желательно, чтобы последовательность $\{\phi^{(l)}\}$ включала в себя наиболее распространенные критерии, соответствующие законам распределения Лапласа, Гаусса, которые обладают определенными экстремальными свойствами при робастном оценивании параметров.

Одним из возможных примеров такой последовательности $\{\phi^{(l)}\}$ является последовательность квазиправдоподобных оценок, соответствующая $l$-обобщенному нормальному закону

$$\phi(y,\lambda,l)=\frac{1}{2(2\lambda_l^2)_i^{1/l}\,\Gamma\left(\frac{l+1}{l}\right)}\exp\left(-\frac{|y-u(x,a)|^l}{(2\lambda_l^2)_i}\right),$$

где $\Gamma\left(l+1/l\right)$ – гамма-функция; $2\lambda_l^2$ – величина, являющаяся функцией дисперсии; $l$ – любое положительное число: при $l=1$ – закон Лапласа, при $l=2$ – закон Гаусса. Тогда в качестве формальных критериев при целых $l>0$ для получения квазиправдоподобных оценок примем последовательно

$$\min_{a\in R^p}\sum_{i=1}^{n}\frac{1}{(2\lambda_1^2)_i}|y_i-u(x_i,a)|,$$

$$\min_{a\in R^p}\sum_{i=1}^{n}\frac{1}{(2\lambda_2^2)_i}(y_i-u(x_i,a))^2, \ldots,$$

что соответствует методу наименьших модулей, квадратов и т.д., так как логарифмическая функция квазиправдоподобия имеет вид

$$\ln\phi^{(l)}(Y^T,\lambda,l)=\ln\left(\left[2\Gamma\left(\frac{l+1}{l}\right)\right]^n\left|\prod_{i=1}^{n}(2\lambda_l^2)_i^{1/l}\right|\right)^{-1}-$$

$$-\sum_{i=1}^{n}\frac{1}{(2\lambda_l^2)_i}|y_i-u(x_i,a)|^l.$$

На основе разработанных алгоритмов оценивания создано программное обеспечение, которое нашло применение при расчете прогноза концентрации вредных веществ в атмосфере.

## СПИСОК ЛИТЕРАТУРЫ


1. *Кацюба О. А.* О методе квазиправдоподобных оценок в задачах идентификации нелинейных объектов // Сиб. отд-ние АН СССР. Автометрия.1986. №6.
2. *Кацюба О. А., Хакимов Б. Б.* Алгоритм нелинейного параметрического оценивания в многомерных задачах статистической обработки // Сиб. отд-ние АН СССР. Автометрия. 1984. №2.
3. *Гнеденко Б. В.* Курс теории вероятностей. М., 2001.
4. *Ширяев А. Н.* Вероятность. М.: Наука. 1980.
5. *Будак Б. М., Фомин С. В.* Кратные интегралы и ряды. М.: Наука. 1965.
6. *Александров П. С.* Введение в теорию множеств и общую топологию. М.: Наука, 1977.
7. *Крамер Г.* Математические методы статистики. М.: Мир, 1975.
8. *Кацюба О. А., Линеенко М. Б.* Параметрическое оценивание нелинейных статических объектов в условиях априорной неопределенности // Труды III межд. конф. “Идентификация систем и задачи управления” SICPRO’04. М., 2004 г.


# ABOUT AN EFFICIENCY OF ESTIMATIONS WITH REFERENCE TO PARAMETRICAL IDENTIFICATION OF NONLINEAR STATIC OBJECTS

©2004 O.A. Katsyuba, M.B. Lineenko

Samara State Academy of Ways of Communications, Samara

It is considered the problem of a parametrical estimation of stochastic static models, nonlinear concerning parameters. The method of quasi-likelihood estimates is applied as a method of estimation; it is adduced proofs of a consistency and of an asymptotic normality of maximum quasi-likelihood estimates in conditions of independent heterogeneous observations. It is shown, that application of linearly-combined estimations allows receiving effectiveness of estimation somehow close to effectiveness of maximum likelihood estimations in conditions of lack of the information of distribution functions of parasites of observations.