Сибирский математический журнал
Март—апрель, 2001. Том 42, № 2

УДК 519.237.5

АСИМПТОТИЧЕСКИ НОРМАЛЬНОЕ ОЦЕНИВАНИЕ МНОГОМЕРНОГО ПАРАМЕТРА В ЗАДАЧЕ ДРОБНО–ЛИНЕЙНОЙ РЕГРЕССИИ

Ю. Ю. Линке, А. И. Саханенко

Аннотация: Рассматривается задача оценивания неизвестного многомерного параметра для некоторой задачи так называемой дробно-линейной регрессии. Предлагается новый метод, позволяющий достаточно просто находить явные асимптотически нормальные оценки неизвестного параметра без использования метода наименьших квадратов. Библиогр. 5.

§ 1. Постановка задачи

Пусть в результате серии из N испытаний, $N \to \infty$, наблюдается последовательность случайных величин $Z_1, \dots, Z_N$, относительно которых предполагается, что они представимы в виде

$$Z_i = g_i(\theta) + \xi_i \equiv \frac{\alpha_i(\theta)}{\beta_i(\theta)} + \xi_i, \quad i = 1, \dots, N, \tag{1.1}$$

где

$$\alpha_i(\theta) \equiv a_{0i} + \sum_{j=1}^{m} a_{ji}\theta_j, \quad \beta_i(\theta) \equiv 1 + \sum_{j=1}^{m} b_{ji}\theta_j \tag{1.2}$$

— линейные функции, зависящие от неизвестного m-мерного параметра θ с координатами $\theta_1, \dots, \theta_m$, причем числа

$$b_{ji} \geq 0,\ a_{0i},\ a_{ji}, \quad i = 1, \dots, N,\ j = 1, \dots, m, \tag{1.3}$$

предполагаются известными. Случайные величины ξ_i, $i = 1, \dots, N$, в (1.1) — это ненаблюдаемые погрешности измерений. Ниже мы будем налагать ряд ограничений на предельное поведение распределений линейных комбинаций этих случайных величин.

В настоящей работе рассматривается задача оценивания неизвестного вектора θ с координатами $\theta_j > 0$, $j = 1, \dots, m$, по значениям случайных величин $Z_1, \dots, Z_N$. Авторы предлагают некоторый метод, позволяющий достаточно просто получать асимптотически нормальные оценки неизвестных параметров для модели дробно-линейной регрессии (1.1)–(1.3). В отличие от метода наименьших квадратов, который традиционно применяют при решении такого рода задач нелинейной регрессии, при реализации предлагаемого метода не нужно

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 99–01–00561) и INTAS (код проекта 98–1625).

использовать итерационные процедуры, порождающие, в свою очередь, проблемы выбора начального приближения, сходимости процесса и др., а также требующие использования вычислительной техники из-за большого числа итераций.

Основная цель данной работы — описать в общем виде предлагаемый метод построения оценок и схему изучения этих оценок, а также продемонстрировать применение ряда идей, которые могут быть использованы при изучении получаемых оценок. Полное математически строгое обоснование этого метода для простейшего одномерного случая задачи дробно-линейной регрессии изложено в [1]. В следующей работе авторов будет подробно изучен случай двух неизвестных параметров, включающий уравнение Михаэлиса — Ментен, играющее важную роль в биохимии.

В основе предлагаемого метода лежит замеченная авторами аналогия между задачей дробно-линейной регрессии (1.1) и классической задачей линейной регрессии (см. (2.2) и (2.3)). Поэтому имеется определенная аналогия между стандартным методом получения оценок параметров в классической задаче линейной регрессии (см., например, [2]) и предлагаемым в (2.5) способом получения оценок в более сложной задаче дробно-линейной регрессии (1.1). И там и здесь предлагается искать оценки параметров как решения систем линейных уравнений, в которых число уравнений совпадает с числом неизвестных параметров. В обоих случаях при получении этих оценок не нужно использовать трудоемкие итерационные процедуры для приближенного поиска оценок.

О структуре работы. В § 2 мы в общем случае опишем предлагаемый способ построения оценок в рассматриваемой задаче, а в § 3 приведем условия для состоятельности и асимптотической нормальности построенных оценок. В § 4 опишем еще более широкий класс оценок и найдем условия для их асимптотической нормальности. В § 5 получим необходимые условия для оптимальности введенных оценок и тем самым укажем возможный путь для нахождения таких оценок. В § 6, 7 мы рассмотрим несколько важных частных случаев. Ряд рекомендаций по практическому применению предлагаемых оценок можно найти в § 6–8.

Об обозначениях. Запись $\mathbf{A} = A_{m\times n}$ означает, что $\mathbf{A}$ — матрица, состоящая из m строк и n столбцов. Элемент на пересечении p-й строки и q-го столбца этой матрицы будем обозначать через $(\mathbf{A})_{pq}$. Далее, символ $^\top$ обозначает транспонирование вектора либо матрицы. Если t — вектор с координатами $t_1, \ldots, t_N$, то t — вектор-столбец, а $t^\top = (t_1, \ldots, t_N)$ — вектор-строка. Символы $\mathbf{I} = I_{n\times n}$ и $\mathbf{0} = 0_{n\times n}$ обозначают единичную и нулевую матрицы соответствующей размерности, а через $\operatorname{diag}\{h_1, \ldots, h_N\}$ будем обозначать диагональную матрицу размерности $N \times N$ с элементами $h_1, \ldots, h_N$ на главной диагонали. В случае симметричных неотрицательно определенных матриц $\mathbf{B}$ через $\mathbf{B}^{1/2}$ обозначим единственную симметричную неотрицательно определенную матрицу, удовлетворяющую равенству

$$(\mathbf{B}^{1/2})^2 = \mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{1/2\top}.$$

Всюду далее предполагается, что все пределы берутся при $N \to \infty$. Если $\mathbf{A}$ — случайная матрица или вектор (элементы которых могут зависеть от N), то под сходимостью $\mathbf{A} \xrightarrow{p} \mathbf{A}_0$ условимся понимать покоординатную сходимость по вероятности элементов этих матрицы или вектора (при $N \to \infty$). Сходимость

$$\mathbf{A} \Longrightarrow \Phi_m(0, \mathbf{I})$$

в этом случае означает, что распределение вектора $\mathbf{A}$ может зависеть от N и слабо сходится (при $N \to \infty$) к m-мерному стандартному нормальному распределению $\Phi_m(0, \mathbf{I})$.

Будем говорить, что некоторая m-мерная статистика $\tilde{\theta}^*$ является *асимптотически нормальной оценкой* m-мерного параметра θ с асимптотической ковариационной матрицей $\mathbf{K}\mathbf{K}^\top$, если распределение вектора $\mathbf{K}^{-1}(\tilde{\theta}^* - \theta)$ слабо сходится к m-мерному стандартному нормальному распределению $\Phi_m(0, \mathbf{I})$.

§ 2. Построение оценок неизвестного параметра

2.1. Прежде всего приведем уравнение (1.1) к виду, более удобному для нахождения оценок. С этой целью сначала домножим обе части уравнения (1.1) на знаменатель $\beta_i(\theta)$ из (1.2). В результате получим

$$Z_i + \sum_{j=1}^{m} b_{ji} Z_i \theta_j = a_{0i} + \sum_{j=1}^{m} a_{ji}\theta_j + \beta_i(\theta)\xi_i. \tag{2.1}$$

Введем обозначения

$$X_{ji} = a_{ji} - b_{ji} Z_i, \quad Y_i = Z_i - a_{0i}, \quad \eta_i = \beta_i(\theta)\xi_i. \tag{2.2}$$

Подставив в (2.1) обозначения из (2.2), перепишем уравнение (1.1) в следующем эквивалентном виде:

$$Y_i = \sum_{j=1}^{m} X_{ji}\theta_j + \eta_i, \quad i = 1, \dots, N. \tag{2.3}$$

2.2. Нетрудно заметить, что вид уравнений (2.3) аналогичен виду уравнений в классической задаче линейной регрессии (см., например [2]). В силу этой аналогии естественно попытаться найти оценки для неизвестных параметров $\theta_1, \dots, \theta_m$ точно так же, как это делается в задаче линейной регрессии. Домножим равенства (2.3) на некоторые константы c_{ki}, где $k = 1, \dots, m$, $i = 1, \dots, N$. Тогда, просуммировав по i, получаем

$$\sum_{i=1}^{N} c_{ki} Y_i - \sum_{i=1}^{N}\sum_{j=1}^{m} c_{ki} X_{ji}\theta_j = \sum_{i=1}^{N} c_{ki}\eta_i. \tag{2.4}$$

Мы вправе предполагать, что правая часть в тождестве (2.4), являющаяся взвешенной суммой погрешностей измерений, мала по сравнению с остальными слагаемыми, поэтому в (2.4) естественно отбросить правую часть, подставив в полученное тождество вместо неизвестного параметра θ оценку θ^*. Поэтому на первом шаге требуемые оценки $\theta_1^*, \dots, \theta_m^*$ всегда будем находить как решения следующей системы линейных уравнений:

$$\sum_{i=1}^{N}\sum_{j=1}^{m} c_{ki} X_{ji}\theta_j^* = \sum_{i=1}^{N} c_{ki} Y_i, \quad k = 1, \dots, m, \tag{2.5}$$

при соответствующих специально подобранных постоянных $\{c_{ki}\}$. Ниже, в § 6, можно найти рекомендации по выбору этих постоянных.

Отметим, что главное и очень серьезное отличие уравнений (2.3) от аналогичных уравнений линейной регрессии состоит в том, что $\{X_{ji}\}$ в (2.3) —

это не постоянные, а случайные величины специального вида (см. (2.2)). Указанное обстоятельство делает задачу исследования свойств оценок $\theta_1^*, \dots, \theta_m^*$ значительно более трудоемкой, чем в случае линейной регрессии.

2.3. При исследовании свойств построенных оценок нам будет удобнее перейти к матричным обозначениям. Введем в рассмотрение матрицы $\mathbf{C} = \mathbf{C}_{m\times N}$ и $\mathbf{X} = \mathbf{X}_{m\times N}$ с элементами

$$(\mathbf{C})_{ki} = c_{ki}, \quad (\mathbf{X})_{ji} = X_{ji} = a_{ji} - b_{ji}Z_i$$

и векторы

$$Y = (Y_1, \dots, Y_N)^\top, \quad \eta = (\eta_1, \dots, \eta_N)^\top, \quad \xi = (\xi_1, \dots, \xi_N)^\top,$$
$$\theta = (\theta_1, \dots, \theta_m)^\top, \quad \theta^* = (\theta_1^*, \dots, \theta_m^*)^\top.$$

В этих обозначениях уравнения (2.3) и (2.5) можно записать в следующем, более компактном виде:

$$Y = \mathbf{X}^\top\theta + \eta, \quad \mathbf{C}\mathbf{X}^\top\theta^* = \mathbf{C}Y. \tag{2.6}$$

В частности, из (2.6) вытекает равенство

$$\mathbf{C}\mathbf{X}^\top(\boldsymbol{\theta}^* - \boldsymbol{\theta}) = \mathbf{C}\eta. \tag{2.7}$$

Определим теперь матрицы $\mathbf{\Lambda} = \mathbf{\Lambda}(\boldsymbol{\theta})_{m\times N}$ и $\mathbf{\Psi} = \mathbf{\Psi}_{m\times N}$, полагая

$$(\mathbf{\Lambda})_{ji} = \lambda_{ji}(\theta) = a_{ji} - b_{ji}g_i(\theta), \quad (\mathbf{\Psi})_{ji} = \psi_{ji} = b_{ji}\xi_i.$$

В этом случае

$$\mathbf{X} = \mathbf{\Lambda} - \mathbf{\Psi} \tag{2.8}$$

и (2.7) можно переписать так:

$$(\mathbf{C}\mathbf{\Lambda}^\top - \mathbf{C}\mathbf{\Psi}^\top)(\theta^* - \theta) = \mathbf{C}\eta. \tag{2.9}$$

Равенства (2.7) и (2.9) будут полезны в следующем параграфе при изучении поведения разности $\theta^* - \theta$, которая определяет интересующую нас точность оценивания.

§ 3. Состоятельность и асимптотическая нормальность

3.1. Всюду далее считаем элементы матрицы $\mathbf{C}$ выбранными так, что матрица $\mathbf{C}\mathbf{\Lambda}^\top = (\mathbf{C}\mathbf{\Lambda}(\theta)^\top)_{m\times m}$ невырожденная. Домножая обе части уравнения (2.9) на $(\mathbf{C}\mathbf{\Lambda}^\top)^{-1}$, перепишем (2.9) в виде

$$(I - \mathbf{G}\mathbf{\Psi}^\top)(\boldsymbol{\theta}^* - \boldsymbol{\theta}) = \mathbf{G}\eta \quad \text{при } \mathbf{G} = (\mathbf{C}\mathbf{\Lambda}^\top)^{-1}\mathbf{C}. \tag{3.1}$$

Используя представление (3.1), нетрудно «угадать» вид приведенной ниже теоремы 1.

Теорема 1. *Пусть выполнены условия*

$$(\mathbf{C}\mathbf{\Lambda}^\top)^{-1}\mathbf{C}\mathbf{\Psi}^\top \xrightarrow{p} \mathbf{0}, \tag{3.2}$$

$$(\mathbf{C}\mathbf{\Lambda}^\top)^{-1}\mathbf{C}\eta \xrightarrow{p} \mathbf{0}. \tag{3.3}$$

Тогда оценка θ^* состоятельна.

Замечание 1. Ясно, что

$$\|\mathbf{G}\eta\|^2 = \eta^\top\mathbf{G}^\top\mathbf{G}\eta = \sum_{i,l=1}^{N}(\mathbf{G}^\top\mathbf{G})_{il}\beta_i(\theta)\beta_l(\theta)\xi_i\xi_l.$$

Таким образом, если $\mathbf{E}\xi_i^2 < \infty$ при всех $i = 1, 2, \ldots$, то для выполнения условия (3.3) достаточно, чтобы имела место сходимость

$$\mathbf{E}\|\mathbf{G}\eta\|^2 = \sum_{i,l=1}^{N} (\mathbf{G}^\top\mathbf{G})_{il}\beta_i(\theta)\beta_l(\theta)\mathbf{E}\xi_i\xi_l \xrightarrow{p} 0.$$

Аналогичным образом в этом случае нетрудно выписать условие, достаточное для сходимости (3.2):

$$\sum_{j,k=1}^{m} \mathbf{E}(\mathbf{G}\boldsymbol{\Psi}^\top)_{jk}^2 = \sum_{k=1}^{m}\sum_{i,l=1}^{N} (\mathbf{G}^\top\mathbf{G})_{il} b_{ki} b_{kl} \mathbf{E}\xi_i\xi_l \to 0.$$

3.2. Предположим теперь, что случайный вектор $\mathbf{C}\eta$ является асимптотически нормальным в том смысле, что найдется случайная или неслучайная невырожденная матрица $\mathbf{U}$ такая, что

$$\mathbf{UC}\eta \Longrightarrow \Phi_m(0, \mathbf{I}). \tag{3.4}$$

В этом случае из (2.7) вытекает равенство

$$\mathbf{UCX}^\top\mathbf{A}^{-1}\mathbf{A}(\theta^* - \theta) = \mathbf{UC}\eta \tag{3.5}$$

для любой невырожденной матрицы $\mathbf{A}$. Отсюда нетрудно «угадать» вид приведенной ниже теоремы 2.

Теорема 2. *Пусть выполнено условие* (3.4) *и*

$$\mathbf{UCX}^\top\mathbf{A}^{-1} \xrightarrow{p} \mathbf{I} \tag{3.6}$$

при некоторой невырожденной случайной или неслучайной матрице $\mathbf{A}$. *Тогда*

$$\mathbf{A}(\theta^* - \theta) \Longrightarrow \Phi_m(0, \mathbf{I}). \tag{3.7}$$

Следствие 1. *Пусть выполнено условие* (3.4) *и*

$$\mathbf{UC}\boldsymbol{\Lambda}^\top\mathbf{A}^{-1} \xrightarrow{p} \mathbf{I}, \tag{3.8}$$

$$\mathbf{UC}\boldsymbol{\Psi}^\top\mathbf{A}^{-1} \xrightarrow{p} \mathbf{0}. \tag{3.9}$$

Тогда справедливо утверждение (3.7) *теоремы* 2.

Следствие 2. *Пусть утверждение* (3.7) *теоремы* 2 *справедливо для некоторой неслучайной матрицы* $\mathbf{A}$. *Тогда*

$$(\mathbf{A}^\top\mathbf{A})^{1/2}(\theta^* - \theta) \Longrightarrow \Phi_m(0, \mathbf{I}).$$

Замечание 2. В приведенных выше утверждениях естественно положить $\mathbf{A} = \mathbf{UC}\boldsymbol{\Lambda}^\top$. В этом случае условие (3.8) следствия 1 будет выполнено автоматически. Именно так мы будем поступать ниже, в следствии 5 и теореме 6. В теореме 2 проще всего взять $\mathbf{A} = \mathbf{UCX}^\top$, поскольку в этом случае условие (3.6) проверяется элементарно (см. также замечание 13).

Замечание 3. Отметим, что оценка θ^* может удовлетворять предположениям теоремы 2, но не быть состоятельной. Соответствующий пример построен в [1, замечание 6].

В § 7 для случая независимых наблюдений будут приведены простые достаточные условия, гарантирующие выполнение всех указанных выше условий.

3.3. Перейдем к доказательствам сформулированных утверждений. Нам потребуется

Лемма 1. *Пусть m-мерный случайный вектор ζ_N удовлетворяет условию*

$$\zeta_N \Longrightarrow \Phi_m(0, \mathbf{I}), \tag{3.10}$$

а случайная $m \times m$-матрица $\mathbf{J}_N$ и m-мерный случайный вектор δ_N таковы, что

$$\mathbf{J}_N \xrightarrow{p} \mathbf{I}, \quad \delta_N \xrightarrow{p} \mathbf{0}.$$

Тогда

$$\mathbf{J}_N(\zeta_N + \delta_N) \Longrightarrow \Phi_m(0, \mathbf{I}) \quad \textit{и} \quad \mathbf{J}_N^{-1}(\zeta_N + \delta_N) \Longrightarrow \Phi_m(0, \mathbf{I}).$$

Это утверждение нетрудно извлечь из приема Крамера — Уолда, поэтому мы опускаем его вывод.

Доказательство теоремы 1. Перепишем (3.1) в виде

$$\theta^* - \theta = (\mathbf{I} - \mathbf{G}\mathbf{\Psi}^\top)^{-1}\mathbf{G}\eta,$$

где $\mathbf{G}\mathbf{\Psi}^\top \xrightarrow{p} 0$ в силу (3.2). Требуемое утверждение вытекает теперь из (3.3) и леммы 1 при $\mathbf{J}_N = \mathbf{I} - \mathbf{G}\mathbf{\Psi}^\top$.

Доказательство теоремы 2. Из представления (3.5) находим, что

$$\mathbf{A}(\theta^* - \theta) = (\mathbf{U}\mathbf{C}\mathbf{X}^\top\mathbf{A}^{-1})^{-1}\mathbf{U}\mathbf{C}\eta.$$

Утверждения теоремы нетрудно теперь получить из условий (3.4), (3.6) и леммы 1 при $\mathbf{J}_N = \mathbf{U}\mathbf{C}\mathbf{X}^\top\mathbf{A}^{-1}$.

Доказательство следствия 1. Заметим, что в силу (2.8)

$$\mathbf{U}\mathbf{C}\mathbf{X}\mathbf{A}^{-1} = \mathbf{U}\mathbf{C}\mathbf{\Lambda}^\top\mathbf{A}^{-1} - \mathbf{U}\mathbf{C}\mathbf{\Psi}^\top\mathbf{A}^{-1}. \tag{3.11}$$

Из этого представления и условий (3.8) и (3.9) немедленно следует предположение (3.6) теоремы 2.

3.4. Прежде чем перейти к доказательству следствия 2, приведем одно полезное утверждение, которое будем использовать и в дальнейшем.

Лемма 2. *Пусть выполнено условие* (3.10), *а $\mathbf{Q}_N$ — некоторая ортогональная матрица, т. е. $\mathbf{Q}_N\mathbf{Q}_N^\top = \mathbf{I}$. Тогда*

$$\mathbf{Q}_N\zeta_N \Longrightarrow \Phi_m(0, \mathbf{I}).$$

Доказательство. Используя обобщение теоремы о равномерной сходимости характеристических функций на случай многомерных случайных величин (см. [3, п. 13.3]), заключаем, что

$$\forall K < \infty \sup_{\|t\| \le K} |\mathbf{E}e^{i\langle\zeta_N, t\rangle} - e^{-\|t\|^2/2}| \to 0.$$

В силу ортогональности матрицы $\mathbf{Q}_N$ имеем $\|t\| = \|\mathbf{Q}_N t\|$, а потому

$$\sup_{\|t\| \le K} |\mathbf{E}e^{i\langle\mathbf{Q}_N\zeta_N, t\rangle} - e^{-\|t\|^2/2}| = \sup_{\|\mathbf{Q}_N^\top t\| \le K} |\mathbf{E}e^{i\langle\zeta_N, \mathbf{Q}_N^\top t\rangle} - e^{-\|\mathbf{Q}_N^\top t\|^2/2}| \to 0.$$

Полученная сходимость характеристических функций приводит к требуемому утверждению леммы 2.

Доказательство следствия 2. Положим $\mathbf{A}_0 = (\mathbf{A}^\top\mathbf{A})^{1/2}$. Тогда

$$(\mathbf{A}_0\mathbf{A}^{-1})(\mathbf{A}_0\mathbf{A}^{-1})^\top = \mathbf{A}_0\mathbf{A}^{-1}\mathbf{A}^{-1\top}{\mathbf{A}_0}^\top = \mathbf{A}_0(\mathbf{A}^\top\mathbf{A})^{-1}{\mathbf{A}_0}^\top = \mathbf{I}.$$

Значит, $\mathbf{Q}_N = \mathbf{A}_0\mathbf{A}^{-1}$ — ортогональная матрица. Таким образом, из теоремы 2 и леммы 2 немедленно вытекает требуемое утверждение:

$$(\mathbf{A}\mathbf{A}^\top)^{1/2}(\theta^* - \theta) = \mathbf{Q}_N\mathbf{A}(\theta^* - \theta) \Longrightarrow \Phi_m(0, \mathbf{I}).$$

§ 4. Улучшение оценок

4.1. Пусть на первом этапе оценивания для параметра θ построена некоторая оценка θ^* как решение системы уравнений (2.5). В этом случае на втором этапе мы можем ввести в рассмотрение оценки $\theta_1^{**}, \dots, \theta_N^{**}$ как решения системы уравнений

$$\sum_{i=1}^{N}\sum_{j=1}^{m}\gamma_{ki}(\theta^*)X_{ji}\theta_j^{**} = \sum_{i=1}^{N}\gamma_{ki}(\theta^*)Y_i, \quad k=1,\dots,m, \tag{4.1}$$

где $\gamma_{ki}(\theta)$ — некоторые специально подобранные функции, зависящие только от неизвестного параметра θ. Практические рекомендации по выбору функций $\gamma_{ki}(\theta)$ будут приведены в § 6.

Отметим, что система уравнений (4.1) отличается от системы (2.5) лишь заменой чисел c_{ki} статистиками $\gamma_{ki}(\theta^*)$. Подчеркнем, что использование статистик $\gamma_{ki}(\theta^*)$ в дополнение к числам c_{ki} существенно расширяет класс оценок. Тем самым при удачном выборе функций $\gamma_{ki}(\theta)$ появляется возможность вместо оценок θ_k^* использовать «улучшенные» оценки θ_k^{**}. В частности, в § 5, 6 будет показано, что при достаточно широких условиях можно так выбрать функции γ_{ki}, чтобы ковариационная матрица улучшенной оценки θ^{**} была минимальной.

4.2. Приступим теперь к изучению введенных оценок. По аналогии с (2.6) удобно переписать уравнения (4.1) в следующем эквивалентном матричном виде:

$$\mathbf{\Gamma}(\theta^*)\mathbf{X}^\top\theta^{**} = \mathbf{\Gamma}(\theta^*)Y, \tag{4.2}$$

где $\mathbf{\Gamma}(\boldsymbol{\theta}) = \mathbf{\Gamma}(\boldsymbol{\theta})_{m\times N}$ при $(\mathbf{\Gamma}(\theta))_{ki} = \gamma_{ki}(\theta)$. В частности, из (2.6) и (4.2) вытекает, что справедлив следующий аналог представления (2.7):

$$\mathbf{\Gamma}(\theta^*)\mathbf{X}^\top(\theta^{**}-\theta) = \mathbf{\Gamma}(\theta^*)\eta. \tag{4.3}$$

4.3. Предположим теперь, что случайный вектор $\mathbf{\Gamma}(\theta)\eta$ является асимптотически нормальным в том смысле, что найдется случайная или неслучайная невырожденная матрица $\mathbf{U}_\Gamma$ такая, что

$$\mathbf{U}_\Gamma\mathbf{\Gamma}(\theta)\eta \Longrightarrow \Phi_m(0,\mathbf{I}). \tag{4.4}$$

В этом случае справедливы следующие аналоги утверждений предыдущего параграфа.

Теорема 3. *Пусть выполнено условие* (4.4), *а также*

$$\delta_{0,N} \equiv \mathbf{U}_\Gamma(\mathbf{\Gamma}(\theta^*)-\mathbf{\Gamma}(\theta))\eta \xrightarrow{p} \mathbf{0}, \tag{4.5}$$

$$\mathbf{U}_\Gamma\mathbf{\Gamma}(\theta^*)\mathbf{X}^\top\mathbf{A}_\Gamma^{-1} \xrightarrow{p} \mathbf{I} \tag{4.6}$$

при некоторой случайной или неслучайной невырожденной матрице $\mathbf{A}_\Gamma$. *Тогда*

$$\mathbf{A}_\Gamma(\theta^{**}-\theta) \Longrightarrow \Phi_m(0,\mathbf{I}). \tag{4.7}$$

Следствие 3. *Пусть выполнены условия* (4.4), (4.5), *а также*

$$\mathbf{U}_\Gamma\mathbf{\Gamma}(\theta)\mathbf{\Lambda}^\top\mathbf{A}_\Gamma^{-1} \xrightarrow{p} \mathbf{I}, \tag{4.8}$$

$$\mathbf{\Delta}_{1,N} \equiv \mathbf{U}_\Gamma\mathbf{\Gamma}(\theta)\mathbf{\Psi}^\top\mathbf{A}_\Gamma^{-1} \xrightarrow{p} \mathbf{0}, \tag{4.9}$$

$$\mathbf{\Delta}_{2,N} \equiv \mathbf{U}_\Gamma(\mathbf{\Gamma}(\theta^*)-\mathbf{\Gamma}(\theta))\mathbf{\Lambda}^\top\mathbf{A}_\Gamma^{-1} \xrightarrow{p} \mathbf{0}, \tag{4.10}$$

$$\mathbf{\Delta}_{3,N} \equiv \mathbf{U}_\Gamma(\mathbf{\Gamma}(\theta^*)-\mathbf{\Gamma}(\theta))\mathbf{\Psi}^\top\mathbf{A}_\Gamma^{-1} \xrightarrow{p} \mathbf{0}. \tag{4.11}$$

Тогда справедливо утверждение (4.7) *теоремы* 3.

Следствие 4. *Пусть утверждение* (4.7) *теоремы* 3 *справедливо для некоторой неслучайной матрицы* $\mathbf{A}_\Gamma$. *Тогда*

$$(\mathbf{A}_\Gamma^\top \mathbf{A}_\Gamma)^{1/2}(\theta^* - \theta) \Longrightarrow \Phi_m(0, \mathbf{I}).$$

Замечание 4. В приведенных выше утверждениях проще всего взять $\mathbf{A}_\Gamma = \mathbf{U}_\Gamma \mathbf{\Gamma}(\boldsymbol{\theta})\mathbf{\Lambda}^\top$. В этом случае условие (4.8) следствия 3 будет выполнено автоматически. Именно так мы поступим в следствии 6. Если в теореме 3 положить $\mathbf{A}_\Gamma = \mathbf{U}_\Gamma \mathbf{\Gamma}(\theta^*)\mathbf{X}^\top$, то в этом случае условие (4.6) проверяется элементарно. Аналогичная идея предлагается также ниже, в замечании 13.

4.4. Остальная часть параграфа посвящена доказательствам приведенных утверждений.

Доказательство теоремы 3. Домножая равенство (4.3) слева на матрицу $\mathbf{U}_\Gamma$, получаем

$$\mathbf{U}_\Gamma \mathbf{\Gamma}(\theta^*)\mathbf{X}^\top \mathbf{A}_\Gamma^{-1}\mathbf{A}_\Gamma(\theta^{**} - \theta) = \mathbf{U}_\Gamma \mathbf{\Gamma}(\theta^*)\eta,$$

откуда, используя обозначения, введенные в формулировке теоремы, нетрудно извлечь следующее ключевое тождество:

$$\mathbf{A}_\Gamma(\theta^{**} - \theta) = \big(\mathbf{U}_\Gamma \mathbf{\Gamma}(\theta^*)\mathbf{X}^\top \mathbf{A}_\Gamma^{-1}\big)^{-1}(\mathbf{U}_\Gamma \mathbf{\Gamma}(\theta)\eta + \delta_{0,N}).$$

Для завершения доказательства теоремы остается с учетом леммы 1 применить к полученному тождеству условия теоремы.

Доказательство следствия 3. В силу (2.8)

$$\mathbf{\Gamma}(\theta^*)\mathbf{X}^\top = \mathbf{\Gamma}(\theta)(\mathbf{\Lambda} - \mathbf{\Psi}) + (\mathbf{\Gamma}(\theta^*) - \mathbf{\Gamma}(\theta))(\mathbf{\Lambda} - \mathbf{\Psi}).$$

Таким образом, используя обозначения (4.9)–(4.11), получаем

$$\mathbf{U}_\Gamma \mathbf{\Gamma}(\theta^*)\mathbf{X}^\top \mathbf{A}_\Gamma^{-1} = \mathbf{U}_\Gamma \mathbf{\Gamma}(\boldsymbol{\theta})\mathbf{\Lambda}^\top \mathbf{A}_\Gamma^{-1} - \mathbf{\Delta}_{1,N} + \mathbf{\Delta}_{2,N} - \mathbf{\Delta}_{3,N}.$$

Учитывая это тождество, легко убедиться, что условия (4.8)–(4.11) следствия 3 достаточны для справедливости предположения (4.6) теоремы 3.

Доказательство следствия 4. Повторяя рассуждения из доказательства следствия 2, нетрудно установить, что $\mathbf{Q}_N = \big(\mathbf{A}_\Gamma^\top \mathbf{A}_\Gamma\big)^{1/2}\mathbf{A}_\Gamma^{-1}$ — ортогональная матрица. Следовательно,

$$\big(\mathbf{A}_\Gamma^\top \mathbf{A}_\Gamma\big)^{1/2}(\theta^{**} - \theta) = \mathbf{Q}_N \mathbf{A}_\Gamma(\theta^{**} - \theta) \Longrightarrow \Phi_m(0, \mathbf{I}),$$

что немедленно вытекает из леммы 2 и теоремы 3.

§ 5. Оптимизация оценок

5.1. Далее мы будем считать, что погрешности измерений удовлетворяют следующим естественным предположениям:

$$\forall i\ \mathbf{E}\eta_i = \beta_i(\theta)\mathbf{E}\xi_i = 0,\ 0 < \mathbf{D}\eta_i = \beta_i^2(\theta)\mathbf{D}\xi_i < \infty.$$

В этом случае через $\mathbf{V}$ обозначим ковариационную матрицу случайного вектора η. Другими словами, далее мы предполагаем, что

$$\exists \mathbf{E}\eta = 0, \quad \exists \mathbf{V} = \mathbf{E}\eta\eta^\top, \quad \exists \mathbf{V}^{-1}.$$

Введем в рассмотрение класс матриц:

$$\mathscr{M}(\theta, \mathbf{V}) = \{\mathbf{\Gamma} = \mathbf{\Gamma}_{m\times N} : \exists(\mathbf{\Gamma}\mathbf{\Lambda}^\top)^{-1},\ \exists(\mathbf{\Gamma}\mathbf{V}\mathbf{\Gamma}^\top)^{-1}\}.$$

Для матриц $\mathbf{\Gamma} \in \mathscr{M}(\boldsymbol{\theta}, \mathbf{V})$ будем использовать обозначение

$$\mathbf{B}(\mathbf{\Gamma}, \theta, \mathbf{V}) \equiv (\mathbf{\Gamma}\mathbf{\Lambda}^\top)^{-1}\mathbf{\Gamma}\mathbf{V}\mathbf{\Gamma}^\top(\mathbf{\Gamma}\mathbf{\Lambda}^\top)^{-1\top}. \tag{5.1}$$

Всюду далее мы также предполагаем, что $\mathbf{C}, \mathbf{\Gamma}(\theta) \in \mathscr{M}(\theta, \mathbf{V})$.

5.2. Выберем в качестве матриц $\mathbf{U}$ и $\mathbf{U}_\Gamma$ любые неслучайные матрицы, удовлетворяющие условиям

$$\mathbf{U}^\top\mathbf{U} = (\mathbf{C}\mathbf{V}\mathbf{C}^\top)^{-1}, \quad \mathbf{U}_\Gamma^\top\mathbf{U}_\Gamma = (\mathbf{\Gamma}(\theta)\mathbf{V}\mathbf{\Gamma}^\top(\theta))^{-1}. \tag{5.2}$$

Далее будем полагать

$$\mathbf{A} = \mathbf{U}\mathbf{C}\mathbf{\Lambda}^\top, \quad \mathbf{A}_\Gamma = \mathbf{U}_\Gamma\mathbf{\Gamma}(\theta)\mathbf{\Lambda}^\top. \tag{5.3}$$

В этом случае предположения (3.8) и (4.8) для так определенных матриц выполняются автоматически, условия (3.4) и (4.4) означают справедливость классической центральной предельной теоремы, а следствия 2 и 4 примут следующий простой вид.

Следствие 5. *Пусть для матриц* $\mathbf{U}$ *и* $\mathbf{A}$, *введенных в* (5.2) *и* (5.3), *выполнены условия* (3.4) *и* (3.9). *Тогда*

$$\mathbf{B}_C^{-1/2}(\theta^* - \theta) \Longrightarrow \Phi_m(0, \mathbf{I}) \quad \textit{при} \quad \mathbf{B}_C = (\mathbf{A}^\top\mathbf{A})^{-1}.$$

Следствие 6. *Пусть для матриц* $\mathbf{U}_\Gamma$ *и* $\mathbf{A}_\Gamma$, *определенных в* (5.2) *и* (5.3), *выполнены условия* (4.4), (4.5), (4.9)–(4.11). *Тогда имеет место сходимость*

$$\mathbf{B}_\Gamma^{-1/2}(\theta^{**} - \theta) \Longrightarrow \Phi_m(0, \mathbf{I}) \quad \textit{при} \quad \mathbf{B}_\Gamma = ({\mathbf{A}_\Gamma}^\top\mathbf{A}_\Gamma)^{-1}.$$

Замечание 5. Нетрудно понять, что оценки θ^* и θ^{**} не изменятся, если уравнения (2.6) и (4.2), их определяющие, домножить слева на некоторые невырожденные матрицы. Иначе говоря, мы можем находить эти оценки как решения уравнений

$$\mathbf{R}_C\mathbf{C}\mathbf{X}^\top\theta^* = \mathbf{R}_C\mathbf{C}Y, \quad \mathbf{R}_\Gamma\mathbf{\Gamma}(\theta^*)\mathbf{X}^\top\theta^{**} = \mathbf{R}_\Gamma\mathbf{\Gamma}(\theta^*)Y$$

при произвольных невырожденных матрицах $\mathbf{R}_C = (\mathbf{R}_C)_{m\times N}$ и $\mathbf{R}_\Gamma = (\mathbf{R}_\Gamma)_{m\times N}$.

Нетрудно убедиться, что в этом случае асимптотические ковариационные матрицы оценок θ^* и θ^{**} не изменятся.

5.3. Таким образом, при выполнении условий следствий 5 и 6 оценки θ^* и θ^{**} асимптотически нормальны с асимптотическими ковариационными матрицами $\mathbf{B}_C$ и $\mathbf{B}_\Gamma$ соответственно. Ясно, что эти оценки тем точнее, чем меньше их асимптотические ковариационные матрицы. Заметим, однако, что

$$\mathbf{B}_C = \mathbf{B}(\mathbf{C}, \theta, \mathbf{V}), \quad \mathbf{B}_\Gamma = \mathbf{B}(\mathbf{\Gamma}(\theta), \theta, \mathbf{V}). \tag{5.4}$$

Поэтому естественно среди матриц вида $\mathbf{B}(\mathbf{\Gamma}, \boldsymbol{\theta}, \mathbf{V})$ попытаться найти в некотором смысле минимальные.

Далее, согласно [4] неравенство $\mathbf{B}_1 \geq \mathbf{B}_2$ между двумя симметричными неотрицательно определенными матрицами $\mathbf{B}_1$ и $\mathbf{B}_2$ означает, что матрица $\mathbf{B}_1 - \mathbf{B}_2$ неотрицательно определенная, т. е. справедливо неравенство $t^\top\mathbf{B}_1 t \geq t^\top\mathbf{B}_2 t$ между квадратичными формами для любого вектор-столбца $t = (t_1, \dots, t_N)^\top$.

Введем обозначения

$$\mathbf{B}^{opt}(\theta, \mathbf{V}) = (\mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{\Lambda}^\top)^{-1}, \quad \mathbf{\Gamma}^o(\theta, \mathbf{V}) = \mathbf{\Lambda}\mathbf{V}^{-1}. \tag{5.5}$$

Теорема 4. *Существует такая симметричная неотрицательно определенная матрица* $\mathbf{B}^o$, *что для всех матриц* $\mathbf{\Gamma} \in \mathscr{M}(\theta, \mathbf{V})$ *и всех невырожденных матриц* $\mathbf{R} = \mathbf{R}_{m\times m}$ *справедливо равенство*

$$(\mathbf{\Gamma}\mathbf{\Lambda}^\top)(\mathbf{B}(\mathbf{\Gamma}, \theta, \mathbf{V})) - \mathbf{B}^{opt}(\theta, \mathbf{V}))(\mathbf{\Gamma}\mathbf{\Lambda}^\top)^\top = (\mathbf{\Gamma} - \mathbf{R}\mathbf{\Gamma}^o)\mathbf{B}^o(\mathbf{\Gamma} - \mathbf{R}\mathbf{\Gamma}^o)^\top. \qquad (5.6)$$

Следствие 7. *Для всех матриц* $\mathbf{\Gamma} \in \mathscr{M}(\theta, \mathbf{V})$ *верно соотношение*

$$\mathbf{B}(\mathbf{\Gamma}, \theta, \mathbf{V}) \geq \mathbf{B}^{opt}(\theta, \mathbf{V}). \qquad (5.7)$$

При этом равенство

$$\mathbf{B}(\mathbf{\Gamma}^{opt}, \theta, \mathbf{V}) = \mathbf{B}^{opt}(\theta, \mathbf{V}) \qquad (5.8)$$

имеет место в том случае, когда

$$\mathbf{\Gamma}^{opt} = \mathbf{R}\mathbf{\Lambda}\mathbf{V}^{-1} \equiv \mathbf{R}\mathbf{\Gamma}^o(\theta, \mathbf{V}), \qquad (5.9)$$

где $\mathbf{R} = R_{m\times m}$ — *произвольная невырожденная матрица.*

Подчеркнем, что в общем случае матрица $\mathbf{\Gamma}^{opt}$ всегда зависит от неизвестного параметра θ, а также ковариационной матрицы $\mathbf{V}$. Нетрудно понять, что элементы матрицы $\mathbf{\Gamma}^{opt}$ являются постоянными только при очень специальных дополнительных ограничениях на матрицы $\mathbf{\Lambda}$ и $\mathbf{V}$ (см. примеры 3 и 5 в § 6).

5.4. Замечание 6. Тождество (5.6) можно интерпретировать следующим образом. Чем точнее матрица $\mathbf{C}$, используемая на первом шаге, или матрица $\mathbf{\Gamma}(\theta)$, используемая на втором шаге, приближаются некоторой матрицей $\mathbf{\Gamma}^{opt} = \mathbf{R}\mathbf{\Gamma}^o(\theta, \mathbf{V})$, тем ближе асимптотические ковариационные матрицы $\mathbf{B}_C$ и $\mathbf{B}_\Gamma$ оценок θ^* и θ^{**} к оптимальной матрице $\mathbf{B}^{opt}(\theta, \mathbf{V})$. При этом матрица, стоящая в правой части равенства (5.6), достаточно наглядно характеризует степень этой близости.

Замечание 7. Пусть погрешности ξ_i независимы, имеют стандартное нормальное распределение и дисперсии σ_i^2 не зависят от параметра θ. Тогда

$$\mathbf{B}^{opt}(\theta, \mathbf{V}) = \mathbf{I}_N^{-1}(\theta), \qquad (5.10)$$

где $\mathbf{I}_N(\theta)$ — информация Фишера для выборки $Z_1, \dots, Z_N$. Таким образом, по аналогии с неравенством Рао — Крамера следует ожидать неулучшаемости в некотором смысле оценок θ^*, когда выбрана $\mathbf{C} = \mathbf{\Gamma}^{opt}$, и оценок θ^{**}, если выбрана $\mathbf{\Gamma}(\theta) = \mathbf{\Gamma}^{opt}$.

Утверждение (5.10) следует из представления (5.7) для $\mathbf{B}^{opt}(\theta, \mathbf{V})$ и равенства

$$(\mathbf{I}_N(\theta))_{jk} = \sum \frac{\lambda_{ji}(\theta)\lambda_{ki}(\theta)}{\beta_i^2(\theta)\sigma_i^2} = (\mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{\Lambda}^\top)_{jk},$$

несложный вывод которого мы здесь опускаем.

5.5. Для доказательства сформулированных утверждений нам потребуется следующая лемма, которая уточняет утверждение теоремы 4.

Лемма 3. *Равенство* (5.6) *имеет место для матрицы*

$$\mathbf{B}^o \equiv \mathbf{V} - \mathbf{\Lambda}^\top \mathbf{B}^{opt} \mathbf{\Lambda}, \qquad (5.11)$$

которая удовлетворяет соотношению

$$\mathbf{B}^{o\top} = \mathbf{B}^o = \mathbf{B}^o \mathbf{V}^{-1} \mathbf{B}^{o\top} \geq 0. \qquad (5.12)$$

Доказательство. Из (5.1) и (5.7) для $\mathbf{B}(\mathbf{\Gamma}, \theta, \mathbf{V}))$ и $\mathbf{B}^{opt}(\theta, \mathbf{V})$ немедленно получаем, что левая часть в (5.6) совпадает с $\mathbf{\Gamma}^\top \mathbf{B}^o \mathbf{\Gamma}$ при $\mathbf{B}^o$, определенной в (5.11). Используя тождество

$$\mathbf{\Gamma}^o(\theta, \mathbf{V})\mathbf{B}^o = \mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{V} - \mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{\Lambda}^\top \mathbf{B}^{opt}\mathbf{\Lambda} = \mathbf{\Lambda} - (\mathbf{B}^{opt})^{-1}\mathbf{B}^{opt}\mathbf{\Lambda} = \mathbf{0},$$

нетрудно убедиться, что правая часть в (5.6) также совпадает с $\mathbf{\Gamma}^\top \mathbf{B}^o \mathbf{\Gamma}$.

Тем самым мы доказали первое утверждение леммы 3. Применяя теперь определения (5.5) и (5.11) для $\mathbf{B}^{opt}(\theta, \mathbf{V})$ и $\mathbf{B}^o$, прямой проверкой нетрудно убедиться в справедливости равенств в (5.12).

Таким образом, лемма 3 и теорема 4 полностью доказаны. Приступим теперь к выводу следствия 7. Соотношение (5.7) вытекает из того факта, что в силу тождества (5.6) $\mathbf{B}(\mathbf{\Gamma}, \theta, \mathbf{V})) - \mathbf{B}^{opt}(\theta, \mathbf{V})$ — неотрицательно определенная матрица. Если же выполнено условие (5.9), то равенство (5.8) является очевидным следствием все того же тождества (5.6).

§ 6. О выборе констант и функций, участвующих в определении оценок

6.1. Пример 1. Предположим, что ковариационная матрица $\mathbf{V}$ представима в виде

$$\mathbf{V} = \mathbf{W}(\theta)\sigma^2,$$

где $\mathbf{W}(\theta)$ — матрица, у которой все элементы являются известными функциями от θ, а σ^2 — некоторый неизвестный параметр. Положим

$$\mathbf{\Gamma}(\theta) = \mathbf{\Lambda}(\theta)\mathbf{W}^{-1}(\theta).$$

Если при так определенной матрице $\mathbf{\Gamma}(\theta)$ выполнены условия теоремы 3, то оценка θ^{**} асимптотически нормальна с оптимальной ковариационной матрицей $\mathbf{B}^{opt}(\theta, \mathbf{V})$.

Этот факт немедленно вытекает из следствия 8, поскольку в этом случае $\mathbf{\Gamma}(\theta) = \sigma^2 \mathbf{\Lambda}\mathbf{V}^{-1} \equiv \sigma^2 \mathbf{\Gamma}^o(\theta, \mathbf{V})$.

Замечание 8. В примере 1 в качестве чисел c_{ki} на первом шаге можно рекомендовать взять $c_{ki} = \gamma_{ki}(\theta_0) \equiv (\mathbf{\Gamma}(\theta_0))_{ki}$ при некотором заранее выбранном θ_0. Понятно, что в этом случае оценка θ^* будет тем точнее, чем ближе выбранное значение θ_0 будет к неизвестному истинному значению параметра θ.

Отметим, что рассмотренный в примере 1 достаточно общий случай включает в себя при $\mathbf{W}(\theta) \equiv \mathbf{I}$ и ситуацию классического регрессионного анализа, когда погрешности измерений $\xi_1, \ldots, \xi_N$ независимы, одинаково распределены, имеют нулевые средние, а дисперсии $\mathbf{D}\xi_i = \sigma^2$ одинаковые и неизвестные.

6.2. В примерах 2–5 считаем, что случайные погрешности $\xi_1, \ldots, \xi_N$ в (1.1) некоррелированы и удовлетворяют условиям

$$\forall i \ \mathbf{E}\xi_i = 0, \ 0 < \beta_i^2(\theta)\mathbf{D}\xi_i \equiv \mathbf{D}\eta_i = \sigma^2/w_i(\theta) < \infty, \tag{6.1}$$

где $w_i(\theta)$ — известная функция, а σ^2 — неизвестный параметр.

Пример 2. Если выполнено условие (6.1), то, как отмечено в примере 1, мы всегда можем положить

$$(\mathbf{\Gamma}^{opt})_{ki} = \lambda_{ki}(\theta)w_i(\theta) \equiv (a_{ki} - b_{ki}g_i(\theta))w_i(\theta). \tag{6.2}$$

Пример 3. Пусть

$$g_i(\theta) = a_{0i}/\beta_i(\theta) \tag{6.3}$$

и выполнено условие (6.1) при

$$w_i(\theta) = w_{0i}/g_i(\theta), \tag{6.4}$$

где w_{0i} — известные постоянные. Нетрудно убедиться, что в этом случае уже на первом шаге при построении оценки θ^* мы можем выбрать оптимальные константы c_{ki}, полагая

$$c_{ki} = b_{ki}w_{0i} \equiv (\mathbf{\Gamma}^{opt})_{ki}. \tag{6.5}$$

Пример 4. Пусть выполнено условие (6.1) и

$$g_i(\theta) = (c_0 a_{mi} + a_{mi}\theta_m)/\beta_i(\theta). \tag{6.6}$$

Как будет показано ниже, в лемме 4, в этом случае функции $(\mathbf{\Gamma}^{opt})_{ki}$ из следствия 8 можно определять по следующей, более простой, чем (6.2), формуле

$$(\mathbf{\Gamma}^{opt})_{mi} = (1 - c_0 b_{mi})g_i(\theta)w_i(\theta) \quad \text{и} \quad (\mathbf{\Gamma}^{opt})_{ki} = b_{ki}g_i(\theta)w_i(\theta) \text{ при } k < m. \tag{6.7}$$

Пример 5. Пусть выполнены условия (6.1), (6.4) и (6.6). При этом предположении определенные в (6.7) функции $(\mathbf{\Gamma}^{opt})_{ki}$ не зависят от θ. Значит, в этом случае уже на первом шаге мы можем выбрать оптимальные константы c_{ki}, полагая

$$c_{mi} = (1 - c_0 b_{mi})w_{0i} \equiv (\mathbf{\Gamma}^{opt})_{mi} \quad \text{и} \quad c_{ki} = b_{ki}w_{0i} \equiv (\mathbf{\Gamma}^{opt})_{ki} \text{ при } k < m. \tag{6.8}$$

6.3. Замечание 9. Если в примере 3 известно, что выполнено только условие (6.3), то можно рекомендовать использовать $c_{ki} = b_{ki}$ как самые простые c_{ki} из (6.5) и в случае, когда нам неизвестно, выполнено условие (6.4) или нет. Однако такой выбор c_{ki} уже не будет оптимальным.

Аналогично если в примере 5 нам известно только, что выполнено условие (6.4), то можно рекомендовать использовать c_{ki} из (6.8) при $w_{0i} \equiv 1$ также и в случае, когда условие (6.4) не выполнено или если нет информации о его справедливости.

Замечание 10. Если неизвестен точный вид ковариационной матрицы $\mathbf{V}$, то мы не сможем найти матрицу $\mathbf{\Gamma}^{opt}$ и построить оценку θ^{**} при $\mathbf{\Gamma}(\theta) = \Gamma^{opt}$. Тогда можно рекомендовать взять в качестве элементов $\mathbf{\Gamma}(\theta)$ функции, относительно которых можно предполагать, что они «не сильно отличаются» от неизвестных элементов матрицы $\mathbf{\Gamma}^{opt}$. Поэтому, как отмечалось в замечании 6, чем «лучше» выбранные элементы матрицы $\mathbf{\Gamma}(\theta)$ будут приближать элементы оптимальной матрицы $\mathbf{\Gamma}^{opt}$, тем меньше асимптотическая дисперсия полученной оценки будет отличаться от $\mathbf{B}^{opt}$.

Замечание 11. Если нет никакой информации о поведении ковариационной матрицы $\mathbf{V}$, то можно порекомендовать на первом шаге взять $c_{ki} = \lambda_{ki}(\theta_0)$ при некотором θ_0. В этом случае оценка θ^* будет тем точнее, чем ближе выбранное значение θ_0 будет к неизвестному истинному значению параметра θ и чем ближе неизвестная матрица $\mathbf{V}$ к матрице вида $\sigma^2 I$ при некотором σ.

6.4. Докажем теперь вспомогательное утверждение, которое существенно использовалось в примерах 4 и 5.

Лемма 4. *Пусть выполнены условия* (6.1) *и* (6.6), *а* $\mathbf{\Gamma}^{opt}$ — *матрица, элементы которой определены в* (6.7). *Тогда найдется матрица* $\mathbf{R}$ *такая, что*

$$\mathbf{\Gamma}^{opt} = \mathbf{R\Lambda V}^{-1}.$$

Доказательство. Пусть

$$(\mathbf{R})_{mj} = \begin{cases} \theta_j\sigma^2, & j < m, \\ (c_0+\theta_m)\sigma^2, & j = m, \end{cases} \quad \text{и} \quad (\mathbf{R})_{kj} = \begin{cases} -\sigma^2, & k = j < m, \\ 0, & m \ne k \ne j. \end{cases}$$

Отметим, что так определенная матрица $\mathbf{R}$ невырожденная. Поскольку

$$\lambda_{ji} = \begin{cases} -b_{ji}g_i(\theta), & j < m \\ a_{mi} - b_{mi}g_i(\theta), & j = m, \end{cases}$$

то при $k < m$

$$(\mathbf{R\Lambda})_{ki} = \sum_{j=1}^{m} (\mathbf{R})_{kj}(\mathbf{\Lambda})_{ji} = b_{ki}g_i(\theta)\sigma^2,$$

а при $k = m$

$$\begin{aligned}(\mathbf{R\Lambda})_{mi} &= -\sum_{j=1}^{m-1} b_{ji}\theta_j g_i(\theta)\sigma^2 + (c_0+\theta_m)(a_{mi} - b_{mi}g_i(\theta))\sigma^2 \\ &= g_i(\theta)\left(-\sum_{j=1}^{m} b_{ji}\theta_j + \beta_i(\theta) - c_0 b_{mi}\right)\sigma^2 = (1 - c_0 b_{mi})g_i(\theta)\sigma^2.\end{aligned}$$

В итоге получим требуемое утверждение:

$$(\mathbf{R\Lambda V}^{-1})_{ki} = \begin{cases} b_{ki}g_i(\theta)w_i(\theta), & k < m \\ (1 - c_0 b_{mi})g_i(\theta)w_i(\theta), & k = m. \end{cases}$$

§ 7. Следствия для независимых наблюдений

7.1. Предположим, что погрешности $\{\xi_i\}$ независимы и

$$\forall i\ \mathbf{E}\xi_i = 0,\ 0 < \mathbf{D}\xi_i = \sigma_i^2 < \infty. \tag{7.1}$$

В этом случае

$$\forall i\ \eta_i = \beta_i(\theta)\xi_i,\ \mathbf{D}\eta_i = \beta_i^2(\theta)\sigma^2,\ \mathbf{V} = \operatorname{diag}\{\mathbf{D}\eta_1, \dots, \mathbf{D}\eta_N\}. \tag{7.2}$$

Подчеркнем также, что далее в этом параграфе мы считаем выполненными все предположения из (5.2), (5.3) и будем использовать обозначение $\mathbf{B}_C$, введенное в следствии 5.

При проверке состоятельности полезным будет следующее утверждение.

Теорема 5. *Если погрешности измерений независимы, удовлетворяют* (7.1) *и выполнено условие*

$$\max_{j\le m}(\mathbf{B}_C)_{jj} \to 0, \tag{7.3}$$

то оценка θ^* *состоятельна.*

7.2. Предположим теперь, что погрешности $\{\xi_i\}$ представимы в виде

$$\forall i\ \xi_i = \sigma_i\varepsilon_i,\ \mathbf{E}\varepsilon_i = 0,\ \mathbf{D}\varepsilon_i = 1, \tag{7.4}$$

где $\varepsilon_1, \dots, \varepsilon_N$ — последовательность независимых одинаково распределенных случайных величин.

При проверке асимптотической нормальности будет полезна

Теорема 6. *Пусть независимые погрешности измерений представимы в виде* (7.4), *удовлетворяют условию* (7.3) *и*

$$\max_{i\le N}(\mathbf{V}^{1/2}\mathbf{C}^\top(\mathbf{C}\mathbf{V}\mathbf{C}^\top)^{-1}\mathbf{C}\mathbf{V}^{1/2})_{ii}\to 0. \tag{7.5}$$

Тогда сходимость

$$\mathbf{U}(\theta^*-\theta)\Longrightarrow\Phi_m(0,\mathbf{I})$$

имеет место для любой неслучайной матрицы $\mathbf{U}$, *которая удовлетворяет условию*

$$\mathbf{U}^\top\mathbf{U}=\mathbf{B}_C^{-1}. \tag{7.6}$$

Таким образом, при выполнении достаточно простых условий (7.3) и (7.5) оценка θ^* состоятельна и асимптотически нормальна.

Замечание 12. На первый взгляд, в (7.6) проще всего взять $\mathbf{U}=\mathbf{B}_C^{-1/2}$. Однако иногда оказывается легче, используя стандартную процедуру ортогонализации, найти треугольную матрицу $\mathbf{U}$, удовлетворяющую уравнению (7.6).

7.3. Замечание 13. Подчеркнем, что в теоремах 2, 3 и их следствиях разность $(\theta^*-\theta)$ нормируется матрицами $\mathbf{A}$ и $\mathbf{A}_\Gamma$, $\mathbf{B}_C$ и $\mathbf{B}_\Gamma$, которые существенно зависят от неизвестного параметра θ и от матрицы $\mathbf{V}$, которая также может быть неизвестной. По этой причине при построении доверительных интервалов и проверке гипотез могут оказаться полезными утверждения, в которых эти матрицы заменены известными.

Укажем возможный подход к решению этих задач в случае независимых наблюдений. Положим

$$\eta_i^*=\beta_i(\theta^*)Z_i-\alpha_i(\theta^*),\quad \mathbf{V}^*=\operatorname{diag}\{\eta_1^{*2},\dots,\eta_N^{*2}\},$$

$$\mathbf{A}^*=(\mathbf{C}\mathbf{V}^*\mathbf{C}^\top)^{-1/2}\mathbf{C}\mathbf{X}^\top\quad\text{и}\quad\mathbf{A}^{**}=(\mathbf{\Gamma}(\theta^*)\mathbf{V}^*\mathbf{\Gamma}^\top(\theta^*))^{-1/2}(\mathbf{\Gamma}(\theta^*)\mathbf{X}^\top.$$

Тогда по аналогии с [1] следует ожидать, что при выполнении достаточно жестких дополнительных предположений имеют место сходимости

$$\mathbf{A}^*(\theta^*-\theta)\Longrightarrow\Phi_m(0,\mathbf{I})\quad\text{и}\quad\mathbf{A}^{**}(\theta^{**}-\theta)\Longrightarrow\Phi_m(0,\mathbf{I}).$$

7.4. Приступим к доказательствам сформулированных утверждений. Докажем прежде несколько вспомогательных лемм.

Лемма 5. *Пусть независимые погрешности измерений удовлетворяют предположениям* (7.1) *и* (7.3). *Тогда выполнены условия* (3.2), (3.3) *и* (3.9).

Доказательство. Сравнивая обозначения, введенные в (3.1), (5.1) и (5.4), заключаем, что $\mathbf{B}_C=\mathbf{A}^{-1}\mathbf{A}^{-1\top}=\mathbf{G}\mathbf{V}\mathbf{G}^\top$. Учитывая этот факт и (7.2), находим, что условие (7.3) эквивалентно следующему:

$$(\mathbf{B}_C)_{jj}=\sum_{l=1}^m(\mathbf{A}^{-1})_{jl}^2=\sum_{i=1}^N(\mathbf{G})_{ji}^2\mathbf{D}\eta_i^2\to 0,\quad j=1,\dots,m. \tag{7.7}$$

Таким образом,

$$\mathbf{D}(\mathbf{G}\eta)_j=\mathbf{D}\left(\sum_{i=1}^N(\mathbf{G})_{ji}\eta_i\right)=\sum_{i=1}^N(\mathbf{G})_{ji}^2\mathbf{D}\eta_i=(\mathbf{B}_C)_{jj}\to 0. \tag{7.8}$$

Далее, в силу (1.3) числа b_{ji} неотрицательны, а потому

$$\mathbf{D}b_{qi}\xi_i = b_{qi}^2\sigma_i^2 \le \beta_i^2(\theta)\sigma_i^2/\theta_q^2 = \mathbf{D}\eta_i/\theta_q^2. \tag{7.9}$$

В частности, из (7.9) получаем

$$\mathbf{D}(\mathbf{G}\boldsymbol{\Psi}^\top)_{jk} = \mathbf{D}\left(\sum_{i=1}^N (\mathbf{G})_{ji} b_{ki}\xi_i\right) \le \sum_{i=1}^N (\mathbf{G})_{ji}^2 \mathbf{D}\eta_i/\theta_k^2 \to 0. \tag{7.10}$$

Опять применяя (7.9), находим

$$\begin{aligned}\mathbf{D}(\mathbf{U}\mathbf{C}\boldsymbol{\Psi}^\top\mathbf{A}^{-1})_{jk} &= \mathbf{D}\left(\sum_{i=1}^N\sum_{q=1}^m (\mathbf{U}\mathbf{C})_{ji} b_{qi}\xi_i(\mathbf{A}^{-1})_{qk}\right)\\ &= \sum_{i=1}^N\left(\sum_{q=1}^m (\mathbf{U}\mathbf{C})_{ji} b_{qi}(\mathbf{A}^{-1})_{qk}\right)^2\sigma_i^2 \le a_j\sum_{q=1}^m(\mathbf{A}^{-1})_{qk}^2/\theta_q^2, \end{aligned}\tag{7.11}$$

где

$$a_j = \sum_{i=1}^N (\mathbf{U}\mathbf{C})_{ji}^2 \mathbf{D}\eta_i.$$

Но последняя сумма равна диагональному элементу матрицы $(\mathbf{U}\mathbf{C})\mathbf{V}(\mathbf{U}\mathbf{C})^\top$. Однако

$$(\mathbf{U}\mathbf{C})\mathbf{V}(\mathbf{U}\mathbf{C})^\top = \mathbf{U}\mathbf{C}\mathbf{V}\mathbf{C}^\top\mathbf{U}^\top = \mathbf{U}(\mathbf{U}^\top\mathbf{U})^{-1}\mathbf{U}^\top = \mathbf{I}$$

ввиду (5.2). Следовательно, $a_j = 1$. Из этого факта, (7.7) и (7.11) имеем

$$\mathbf{D}(\mathbf{U}\mathbf{C}\boldsymbol{\Psi}^\top\mathbf{A}^{-1})_{jk} \le \sum_{q=1}^m (\mathbf{A}^{-1})_{qk}^2/\theta_q^2 \le \sum_{q=1}^m (\mathbf{B}_C)_{qq}/\theta_q^2 \to 0. \tag{7.12}$$

Но

$$\mathbf{E}(\mathbf{G}\eta)_j = \mathbf{E}(\mathbf{G}\boldsymbol{\Psi}^\top)_{jk} = \mathbf{E}(\mathbf{U}\mathbf{C}\boldsymbol{\Psi}^\top\mathbf{A}^{-1})_{jk} = 0.$$

Значит, в силу неравенства Чебышева из соотношений (7.8), (7.10) и (7.12) вытекают соответственно требуемые сходимости (3.3), (3.2) и (3.9).

Теорема 5 является очевидным частным случаем доказанной выше леммы 5. Чтобы убедиться в справедливости теоремы 6, нам осталось лишь проверить условие (3.4) теоремы 2, что будет сделано в следующей лемме.

Лемма 6. *Пусть независимые погрешности измерений представимы в виде* (7.4) *и удовлетворяют условию* (7.5). *Тогда выполнено условие* (3.4).

Доказательство. Положим

$$\widetilde{\mathbf{C}} = \mathbf{U}\mathbf{C}\mathbf{V}^{1/2}, \quad (\widetilde{\mathbf{C}})_{ij} = \tilde{c}_{ij}, \quad \varepsilon = (\varepsilon_1, \dots, \varepsilon_N)^\top. \tag{7.13}$$

В этом случае из предположения (5.2) вытекает, что

$$\widetilde{\mathbf{C}}\widetilde{\mathbf{C}}^\top = \mathbf{I}, \quad \text{т. е.} \quad \sum_{i=1}^N \tilde{c}_{ui}\tilde{c}_{vi} = \begin{cases} 0, & u \ne v,\\ 1, & u = v.\end{cases} \tag{7.14}$$

В силу определения (7.13) условие (3.4) можно переписать в виде

$$\mathbf{U}^{-1}\mathbf{C}\eta = \widetilde{\mathbf{C}}\varepsilon \Longrightarrow \Phi_m(0, \mathbf{I}). \tag{7.15}$$

Далее, согласно теореме Крамера — Уолда (см., например, [5]) случайные векторы $\zeta_n = (\zeta_{n1}, \dots, \zeta_{nm})$ в $\mathbf{R}^m$ сходятся по распределению к $\zeta = (\zeta_1, \dots, \zeta_m)$ тогда и только тогда, когда каждая линейная комбинация компонент ζ_n сходится по распределению к соответствующей линейной комбинации компонент ζ. Поэтому утверждение (7.15) эквивалентно сходимости для каждой точки $t = (t_1, \dots, t_m)$ из $\mathbf{R}^m$ распределений сумм $\sum\limits_{j=1}^{m} t_j \sum\limits_{i=1}^{N} \widetilde{\mathbf{C}}_{ji}\varepsilon_i$ к $\Phi_1\left(0, \sum\limits_{j=1}^{m} t_j^2\right)$. В силу (7.14) мы можем воспользоваться утверждением леммы 1 из [1]. Оказывается, достаточно показать, что

$$\max_{i\le N}\left(\sum_{j=1}^{m} t_j(\widetilde{\mathbf{C}})_{ji}\right)^2 \le \sum_{j=1}^{m} t_j^2 \max_{i\le N}\sum_{j=1}^{m}(\widetilde{\mathbf{C}})_{ji}^2 \to 0. \tag{7.16}$$

Однако, учитывая (5.2), нетрудно видеть, что последнюю сумму в (7.16) можно представить в виде

$$(\widetilde{\mathbf{C}}^\top\widetilde{\mathbf{C}})_{ii} = (\mathbf{V}^{1/2}\mathbf{C}^\top\mathbf{U}^\top\mathbf{U}\mathbf{C}\mathbf{V}^{1/2})_{ii} = (\mathbf{V}^{1/2}\mathbf{C}^\top(\mathbf{C}\mathbf{V}\mathbf{C}^\top)^{-1}\mathbf{C}\mathbf{V}^{1/2})_{ii}.$$

Следовательно, требуемая сходимость (7.16) вытекает из условия (7.5).

§ 8. Заключительные замечания

Замечание 14. Стоит отметить, что многие условия в утверждениях работы носят скорее методологический и концептуальный характер и такой путь изложения результатов представляется, по мнению авторов, наиболее наглядным, поскольку излишняя детализация может затуманить основные идеи. Так, например, условия теорем 1 и 2 — это соответственно многомерные ЗБЧ и ЦПТ для специальных схем серий, построенных по последовательности независимых одинаково распределенных случайных величин. Используя прием Крамера — Уолда, в каждом конкретном случае проверку условий (3.2)–(3.4) и (3.6) можно свести к проверке справедливости одномерных ЗБЧ и ЦПТ (и в случае независимых наблюдений это сделано в § 7). На первый взгляд, труднее всего проверять условия теоремы 3, которые даже при выполнении условия (6.1) прием Крамера — Уолда превратит в условия вида

$$\sum_{i=1}^{N}(f_{i,N}(\theta^*) - f_{i,N}(\theta))\xi_i \xrightarrow{p} 0,$$

но техника проверки такого рода условий развита авторами (см. [1, леммы 5–12]).

Замечание 15. Подчеркнем, что все утверждения работы остаются справедливыми также в ситуации, когда наблюдаемые случайные величины образуют схему серий, т. е. когда для i-го наблюдения справедливо представление

$$Z_i^{(N)} = \frac{a_{0i}^{(N)} + \sum\limits_{j=1}^{m} a_{ji}^{(N)}\theta_j}{1 + \sum\limits_{j=1}^{m} b_{ji}^{(N)}\theta_j} + \xi_i^{(N)},$$

где верхний индекс подчеркивает зависимость величин $\{Z_i\}$, $\{a_{ji}\}$, $\{b_{ji}\}$, $\{\xi_i\}$, а также элементов ковариационной матрицы $\mathbf{V}$ от числа наблюдений N. (Вид некоторых условий специально подбирался таким образом, чтобы было справедливо это замечание.) Есть единственное исключение: в теореме 6 надо требовать, чтобы распределения случайных величин $\varepsilon_i^{(N)}$ не зависели ни от i, ни от N.

Замечание 16. Всюду в работе θ — неизвестный параметр. Кроме того, неизвестными параметрами могут быть и элементы ковариационной матрицы ошибок. Таким образом, большинство условий во всех утверждениях работы — это ограничения на величины, содержащие неизвестные параметры. Понятно, что в случае практического применения этих утверждений мы должны проверять выполнение всех таких условий *при всех возможных значениях всех неизвестных параметров* (так же, как это делается, например, в [4]).

ЛИТЕРАТУРА

1. *Линке Ю. Ю., Саханенко А. И.* Асимптотически нормальное оценивание параметра в задаче дробно-линейной регрессии // Сиб. мат. журн. 2000. Т. 41, № 1. С. 150–163.
2. *Демиденко Е. З.* Линейная и нелинейная регрессии. М.: Финансы и статистика, 1981.
3. *Лоэв М.* Теория вероятностей. М.: Изд-во иностр. лит., 1962.
4. *Боровков А. А.* Математическая статистика. М.: Наука, 1984.
5. *Биллингсли П.* Сходимость вероятностных мер. М.: Наука, 1977.

Статья поступила 24 июля 2000 г.,
окончательный вариант — 30 ноября 2000 г.

Линке Юлиана Юрьевна
Институт математики им. С. Л. Соболева СО РАН, Новосибирск 630090
linke@math.nsc.ru

Саханенко Александр Иванович
Universidad Autonoma Metropolitana — Iztapalapa, Av. Michoacan y la Purisima s/n, Col. Vicentina, 09340 Mexico D.F., MEXICO
sakh@xanum.uam.mx