Сибирский математический журнал
Март—апрель, 2009. Том 50, № 2

УДК 519.237.5

АСИМПТОТИЧЕСКИ ОПТИМАЛЬНОЕ ОЦЕНИВАНИЕ В ЗАДАЧЕ ЛИНЕЙНОЙ РЕГРЕССИИ ПРИ НЕВЫПОЛНЕНИИ НЕКОТОРЫХ КЛАССИЧЕСКИХ ПРЕДПОЛОЖЕНИЙ

Ю. Ю. Линке, А. И. Саханенко

Аннотация. Рассмотрена задача оценивания параметров линейной регрессии в случае зависимости дисперсий наблюдений от неизвестных параметров модели. Предложена двухшаговая процедура нахождения асимптотически линейных оценок. Найдены общие достаточные условия асимптотической нормальности введенных оценок, установлен явный вид наилучшей асимптотически линейной оценки. Детально изучено поведение оценок в случае одномерной по параметру модели регрессии.

Ключевые слова: линейная регрессия, двухшаговое оценивание, асимптотически нормальные оценки, наилучшая асимптотически линейная оценка.

§ 1. Введение

1.1. Пусть начиная с некоторого достаточно большого $n \geq m \geq 1$ даны наблюдения $Y_{n1}, Y_{n2}, \dots, Y_{nn}$, представимые в виде

$$Y_{ni} = \beta_{n1} X_{ni1} + \ldots + \beta_{nm} X_{nim} + \epsilon_{ni}, \quad i = 1, \dots, n, \tag{1}$$

где X_{nij}, $i = 1, \dots, n$, $j = 1, \dots, m$, — известные числа, $\beta_{n1}, \dots, \beta_{nm}$ — подлежащие оцениванию неизвестные параметры, а случайные величины $\epsilon_{n1}, \epsilon_{n2}, \dots, \epsilon_{nn}$ — ненаблюдаемые погрешности измерений. При некоторых дополнительных предположениях эта задача линейной регрессии достаточно широко известна (см., например, [1, 2]). Напомним [2], что при ее решении обычно используются *линейные* оценки $\boldsymbol{\beta}_n^* := (\beta_{n1}^*, \dots, \beta_{nm}^*)$ для m-мерного параметра $\boldsymbol{\beta}_n := (\beta_{n1}, \dots, \beta_{nm})$, которые определяются как решения системы линейных уравнений вида

$$\sum_{i=1}^{n} c_{nji} Y_{ni} = \beta_{n1}^* \sum_{i=1}^{n} c_{nji} X_{ni1} + \ldots + \beta_{nm}^* \sum_{i=1}^{n} c_{nji} X_{nim}, \quad j = 1, \dots, m, \tag{2}$$

при некоторых постоянных c_{nji} таких, что система (2) имеет единственное решение.

В частности, если погрешности измерений являются независимыми случайными величинами с нулевыми средними и равными дисперсиями, то рассматриваемая задача входит во многие руководства по математической статистике

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 08–01–00962), Совета по грантам президента РФ и государственной поддержке ведущих научных школ (код проекта НШ–3695.2008.1).

(см., например, опять же [1, 2]). В этом классическом случае существует наилучшая линейная несмещенная оценка $\boldsymbol{\beta}_n^*$ параметра $\boldsymbol{\beta}_n$, совпадающая с оценкой по методу наименьших квадратов, которую можно также определить как решение системы (2) при $c_{nji} = X_{nij}$.

Отметим, что более общий случай, когда $\mathbf{D}\epsilon_{ni} = \sigma_n^2/w_{ni} > 0$ при некоторых известных числах $w_{ni} > 0$, сводится к классическому домножением на $\sqrt{w_{ni}}$ обеих частей в (1). Следовательно, в этом случае наилучшей линейной несмещенной будет оценка $\boldsymbol{\beta}_n^*$, которая является решением системы (2) при $c_{nji} = X_{nij}w_{ni}$. После перехода к матричным обозначениям (см., например, [2]) удается найти явную формулу для наилучшей линейной несмещенной оценки и в случае, когда при некотором $\sigma_n > 0$ случайные величины $\epsilon_{n1}/\sigma_n, \ldots, \epsilon_{nn}/\sigma_n$ имеют нулевые средние и известную невырожденную матрицу ковариаций.

Подчеркнем, что во всех случаях, когда существует наилучшая линейная несмещенная оценка, она имеет матрицу ковариаций $\mathbf{B}_n^{opt}$, определяемую формулой

$$\mathbf{B}_n^{opt} := \left(\mathbf{X}_n^\top \mathbf{V}_n^{-1} \mathbf{X}_n\right)^{-1}, \tag{3}$$

где $\mathbf{X}_n$ — матрица регрессоров X_{nij} размера $n \times m$, а $\mathbf{V}_n$ — матрица ковариаций случайных погрешностей $\epsilon_{n1}, \ldots, \epsilon_{nn}$, у которых еще должны быть нулевые средние.

Отметим также, что в важном случае, когда погрешности измерений $\epsilon_{n1}, \ldots,$ ϵ_{nn} имеют совместное нормальное распределение, все упомянутые выше наилучшие линейные несмещенные оценки будут и эффективными оценками (см. [1, 2]).

1.2. Однако зачастую дисперсии ошибок ϵ_{ni} могут зависеть от неизвестного параметра (см., например, [3]) и в этом случае наилучшие линейные несмещенные оценки могут просто не существовать. Именно так обстоит дело, например, в случае, когда погрешности измерений независимы и

$$\mathbf{E}\epsilon_{ni} = 0 \quad \text{и} \quad \mathbf{D}\epsilon_{ni} = \sigma_n^2/w_{ni}(\boldsymbol{\beta}_n) > 0 \quad \text{при всех } i = 1, \ldots, n. \tag{4}$$

Как следует из сказанного в предыдущем пункте, в этом случае единственным претендентом на роль наилучшей линейной несмещенной оценки является вектор $\boldsymbol{\beta}_n^* := (\beta_{n1}^*, \ldots, \beta_{nm}^*)$, который должен удовлетворять (2) при $c_{nji} = X_{nij}w_{ni}(\boldsymbol{\beta}_n)$. Но такой вектор $\boldsymbol{\beta}_n^*$ не является статистикой при линейно независимых функциях $\{w_{ni}(\cdot)\}$.

Аналогичная проблема возникает и в более сложной ситуации, когда ковариационная матрица погрешностей $\{\epsilon_{ni}\}$ не диагональна, но также зависит от параметра $\boldsymbol{\beta}_n$.

1.3. В расчете на такого рода ситуации в данной работе мы по аналогии с нашими работами [4–7] вместо линейных оценок $\boldsymbol{\beta}_n^*$ предлагаем использовать несколько более сложно устроенные оценки $\boldsymbol{\beta}_n^{**} := (\beta_{n1}^{**}, \ldots, \beta_{nm}^{**})$, которые мы предлагаем находить как решение следующей системы уравнений:

$$\sum_{i=1}^n \gamma_{nji}(\boldsymbol{\beta}_n^*) Y_{ni} = \beta_{n1}^{**} \sum_{i=1}^n \gamma_{nji}(\boldsymbol{\beta}_n^*) X_{ni1} + \ldots + \beta_{nm}^{**} \sum_{i=1}^n \gamma_{nji}(\boldsymbol{\beta}_n^*) X_{nim}, \ j = 1, \ldots, m, \tag{5}$$

при некоторых специально подобранных функциях $\gamma_{nji}(\cdot)$ и оценках $\boldsymbol{\beta}_n^*$. Такие оценки $\boldsymbol{\beta}_n^{**}$ естественно назвать *асимптотически линейными.* В их определении (5) проще всего использовать линейные оценки $\boldsymbol{\beta}_n^*$, являющиеся решением системы (2) при некоторых специально подобранных константах c_{nji}.

В § 2, после введения матричных обозначений, будут получены общие достаточные условия для асимптотической нормальности введенных оценок $\boldsymbol{\beta}_n^{**}$. Кроме того, в теореме 3 и в следствии 3 в п. 2.4 в двух важных случаях будет установлен явный вид наилучшей асимптотически линейной оценки, т. е. будут найдены асимптотически линейные оценки $\boldsymbol{\beta}_n^{**}$, асимптотически нормальные с ковариационной матрицей $\mathbf{B}_n^{opt}$, определяемой все той же формулой (3), но теперь уже при более слабых ограничениях, чем в классическом случае линейных оценок.

В последнем пункте § 2 (см. теоремы 4 и 5) в случае одномерного неизвестного параметра ($m = 1$) проведем еще более детальное исследование поведения наших оценок. Отметим, что мы постарались так изложить эти одномерные результаты, чтобы их в принципе можно было читать сразу после введения, не вникая в технические проблемы, связанные с использованием матричных обозначений.

Доказательства всех результатов представлены в § 3.

Авторы благодарят рецензента за сделанные им полезные замечания.

§ 2. Основные результаты

2.1. При изучении многомерных параметров удобно перейти к матричным обозначениям. Условимся, что далее полужирные буквы, кроме $\mathbf{P}$, $\mathbf{E}$ и $\mathbf{D}$, всегда обозначают или матрицу, или вектор. Символ $\mathbf{A}^\top$ всюду в работе обозначает транспонирование вектора $\mathbf{A}$ либо матрицы $\mathbf{A}$. Если $\mathbf{t}$ — k-мерный вектор, то $\mathbf{t}$ — это всегда вектор-столбец: $\mathbf{t} = (t_1, \dots, t_k)^\top$. Символ $\mathbf{0}$ обозначает нулевой вектор размерности m или нулевую $(m \times m)$-матрицу, символ $\mathbf{I}$ — единичную $(m \times m)$-матрицу.

Далее в работе всюду предполагается, что все пределы берутся при $n \to \infty$. Если $\mathbf{Z}_n$ — случайная матрица или вектор, то под сходимостью $\mathbf{Z}_n \xrightarrow{p} \mathbf{A}$ условимся понимать покоординатную сходимость по вероятности элементов этой матрицы или вектора. Сходимость $\boldsymbol{\alpha}_n \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I})$ означает, что распределение m-мерного вектора $\boldsymbol{\alpha}_n$ слабо сходится к m-мерному стандартному нормальному распределению $\mathscr{N}_m(\mathbf{0}, \mathbf{I})$. В случае симметричных неотрицательно определенных матриц $\mathbf{B}$ через $\mathbf{B}^{1/2}$ будем обозначать единственную симметричную неотрицательно определенную матрицу, которая удовлетворяет равенству $(\mathbf{B}^{1/2})^2 = \mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{1/2\top}$. Если для некоторой оценки $\tilde{\boldsymbol{\beta}}_n^*$ параметра $\boldsymbol{\beta}_n$ имеет место сходимость $\mathbf{B}_n^{-1/2}(\tilde{\boldsymbol{\beta}}_n^* - \boldsymbol{\beta}_n) \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I})$, где матрица $\mathbf{B}_n$ неслучайна и неотрицательно определена, то матрицу $\mathbf{B}_n$ будем называть *асимптотической ковариационной матрицей* и говорить, что оценка $\tilde{\boldsymbol{\beta}}_n^*$ *асимптотически нормальна* с указанной асимптотической ковариационной матрицей.

При сделанных соглашениях уравнения (1) и (5) перепишутся в виде

$$\mathbf{Y}_n = \mathbf{X}_n\boldsymbol{\beta}_n + \boldsymbol{\epsilon}_n, \quad \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n\boldsymbol{\beta}_n^{**} = \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{Y}_n, \tag{6}$$

где $\mathbf{Y}_n := (Y_{n1}, Y_{n2}, \dots, Y_{nn})^\top$ — n-мерный вектор-столбец наблюдений; $\boldsymbol{\epsilon}_n := (\epsilon_{n1}, \epsilon_{n2}, \dots, \epsilon_{nn})^\top$ — ненаблюдаемый вектор погрешностей измерений; $\boldsymbol{\beta}_n = (\beta_{n1}, \dots, \beta_{nm})^\top$ — вектор неизвестных параметров; $\mathbf{X}_n$ — $(n \times m)$-матрица регрессоров, а $\boldsymbol{\Gamma}_n(\cdot)$ — $(m \times n)$-матрица, элементы которой — известные функции $\gamma_{nji}(\cdot)$.

Замечание 1. Если матрица $\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n$ невырожденна, то оценка $\boldsymbol{\beta}_n^{**}$, определенная в (6), существует, единственна и имеет вид

$$\boldsymbol{\beta}_n^{**} = (\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n)^{-1}\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{Y}_n.$$

Но если оценку $\boldsymbol{\beta}_n^{**}$ мы не можем однозначно определить, исходя из формулы (6), то условимся считать, что во всех таких случаях оценка доопределена произвольным измеримым образом. Подчеркнем, что условия в утверждениях работы подобраны так, что конкретный вид такого доопределения не играет никакой роли, поскольку каждый раз вероятность того, что мы использовали доопределение, будет стремиться к нулю.

2.2. Нам неоднократно потребуется следующее ограничение на оценки $\boldsymbol{\beta}_n^*$ и матрицу $\boldsymbol{\Gamma}_n(\cdot)$.

Предположение 1. *Существуют такие случайные или неслучайные матрицы $\mathbf{U}_n^*$ размера $m \times m$, что имеют место следующие сходимости:*

$$\boldsymbol{\alpha}_{n0} := \mathbf{U}_n^* \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n)\boldsymbol{\epsilon}_n \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}), \tag{7}$$

$$\boldsymbol{\Delta}_{n0}(\boldsymbol{\beta}_n^*) := \mathbf{U}_n^*(\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*) - \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n))\boldsymbol{\epsilon}_n \xrightarrow{p} \mathbf{0}. \tag{8}$$

Приведем сначала наиболее общие условия для асимптотической нормальности оценок $\boldsymbol{\beta}_n^{**}$.

Теорема 1. *Пусть выполнено предположение* 1, *а матрицы $\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n$ невырожденны с вероятностями, стремящимися к* 1. *В этом случае*

$$\boldsymbol{\alpha}_{n1} := \mathbf{U}_n^* \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}). \tag{9}$$

Следствие 1. *Пусть выполнено предположение* 1, *а матрицы $\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{X}_n$ невырожденны, начиная с некоторого n. Пусть, кроме того, существуют такие случайные или неслучайные невырожденные матрицы $\mathbf{U}_n$, что*

$$\mathbf{Q}_{n1}(\boldsymbol{\beta}_n^*) := \mathbf{U}_n^* \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n(\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{X}_n)^{-1}\mathbf{U}_n^{-1} \xrightarrow{p} \mathbf{I}. \tag{10}$$

Тогда

$$\boldsymbol{\alpha}_{n2} := \mathbf{U}_n \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{X}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}). \tag{11}$$

Рассмотрим теперь более классическую ситуацию, когда выполнено следующее

Предположение 2. *Для всех достаточно больших номеров n случайный вектор $\boldsymbol{\epsilon}_n$ имеет нулевое среднее и ковариационную матрицу $\mathbf{V}_n$, матрицы $\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{X}_n$ и $\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{V}_n\boldsymbol{\Gamma}_n^\top(\boldsymbol{\beta}_n)$ невырожденны, а предположение* 1 *и условие* (10) *справедливы при*

$$\mathbf{U}_n^* = \mathbf{U}_n = \mathbf{U}_{n0}, \quad \text{где} \;\; \mathbf{U}_{n0} := \big(\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{V}_n\boldsymbol{\Gamma}_n^\top(\boldsymbol{\beta}_n)\big)^{-1/2}. \tag{12}$$

Следствие 2. *Если верно предположение* 2, *то* (11) *имеет место при $\mathbf{U}_n = \mathbf{U}_{n0}$. Кроме того, в этом случае*

$$\boldsymbol{\alpha}_{n3} := \mathbb{B}_n^{-1/2}(\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n))(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}) \tag{13}$$

при

$$\mathbb{B}_n(\mathbf{G}) := (\mathbf{G}\mathbf{X}_n)^{-1}\mathbf{G}\mathbf{V}_n\mathbf{G}^\top\big(\mathbf{X}_n^\top\mathbf{G}^\top\big)^{-1} \equiv (\mathbf{G}\mathbf{X}_n)^{-1}\mathbf{G}\mathbf{V}_n((\mathbf{G}\mathbf{X}_n)^{-1}\mathbf{G})^\top. \tag{14}$$

2.3. Сходимость (13) означает, что матрицу $\mathbb{B}_n(\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n))$ мы можем интерпретировать как асимптотическую ковариационную матрицу вектора $\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n$.

Ясно, что оценка $\boldsymbol{\beta}_n^{**}$ тем точнее, чем меньше эта матрица. Поэтому естественным образом возникает вопрос о минимизации $\mathbb{B}_n(\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n))$. Но эта задача совпадает с аналогичной задачей для линейных оценок, решение которой хорошо известно (см., например, [1, 2]). Тем не менее ниже мы приведем свое решение этой задачи, которое нам представляется более простым.

Подчеркнем, что неравенство $\mathbf{B}_1 \geq \mathbf{B}_2$ между двумя квадратными матрицами $\mathbf{B}_1$ и $\mathbf{B}_2$ всюду в работе означает, что матрица $\mathbf{B}_1 - \mathbf{B}_2$ неотрицательно определена, т. е. справедливо неравенство $\mathbf{t}^\top \mathbf{B}_1 \mathbf{t} \geq \mathbf{t}^\top \mathbf{B}_2 \mathbf{t}$ между квадратичными формами для любого вектора-столбца $\mathbf{t}$ соответствующей размерности.

Теорема 2. *Пусть при некотором n матрица $\mathbf{V}_n$ невырожденна, а матрицы $\mathbf{X}_n$ и $\mathbf{G}_n$ таковы, что $\mathbf{G}_n\mathbf{X}_n$ невырожденна. В этом случае справедливо соотношение*

$$(\mathbf{G}_n\mathbf{X}_n)\big(\mathbb{B}_n(\mathbf{G}_n) - \mathbf{B}_n^{opt}\big)(\mathbf{G}_n\mathbf{X}_n)^\top = \big(\mathbf{G}_n\mathbf{B}_n^o\big)\mathbf{V}_n^{-1}\big(\mathbf{G}_n\mathbf{B}_n^o\big)^\top, \tag{15}$$

где матрица $\mathbf{B}_n^{opt} = \big(\mathbf{X}_n^\top\mathbf{V}_n^{-1}\mathbf{X}_n\big)^{-1}$ определена в (3) и

$$\mathbf{B}_n^o := \mathbf{V}_n - \mathbf{X}_n\mathbf{B}_n^{opt}\mathbf{X}_n^\top \equiv \mathbf{V}_n - \mathbf{X}_n\big(\mathbf{X}_n^\top\mathbf{V}_n^{-1}\mathbf{X}_n\big)^{-1}\mathbf{X}_n^\top. \tag{16}$$

В частности, $\mathbb{B}_n(\mathbf{G}_n) \geq \mathbf{B}_n^{opt}$, а равенство $\mathbb{B}_n(\mathbf{G}_n) = \mathbf{B}_n^{opt}$ имеет место тогда и только тогда, когда $\mathbf{G}_n = \mathbf{R}_n\mathbf{X}_n^\top\mathbf{V}_n^{-1}$ при некоторой невырожденной $(m \times m)$-матрице $\mathbf{R}_n$.

Замечание 2. Таким образом, из теоремы 2 вытекает, что

$$\mathbb{B}_n(\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n)) \geq \mathbf{B}_n^{opt} = \mathbb{B}_n\big(\mathbf{R}_n\mathbf{X}_n^\top\mathbf{V}_n^{-1}\big)$$

в случае, когда матрицы $\mathbf{R}_n$, $\mathbf{V}_n$ и $\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{X}_n$ невырожденны. Если к тому же удастся так подобрать $\mathbf{R}_n$, что матрица $\mathbf{R}_n\mathbf{X}_n^\top\mathbf{V}_n^{-1}$ будет функцией только от $\boldsymbol{\beta}_n$, то получим

$$\mathbf{B}_n^{opt} = \mathbb{B}\big(\boldsymbol{\Gamma}_n^{opt}(\boldsymbol{\beta}_n)\big) \quad \text{при } \boldsymbol{\Gamma}_n^{opt}(\boldsymbol{\beta}_n) = \mathbf{R}_n\mathbf{X}_n^\top\mathbf{V}_n^{-1}.$$

Тем самым имеем естественную рекомендацию по выбору матрицы $\boldsymbol{\Gamma}_n(\cdot)$.

2.4. Рассмотрим теперь наиболее важные случаи, когда удается указать явный вид оптимальных матриц $\boldsymbol{\Gamma}_n(\cdot)$ и тем самым построить оценки $\boldsymbol{\beta}_n^{**}$, для которых будет иметь место следующая сходимость:

$$\boldsymbol{\alpha}_{n4} := \big(\mathbf{B}_n^{opt}\big)^{-1/2}(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \equiv \big(\mathbf{X}_n^\top\mathbf{V}_n^{-1}\mathbf{X}_n\big)^{1/2}(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}). \tag{17}$$

Такую оценку $\boldsymbol{\beta}_n^{**}$ мы вправе называть *наилучшей асимптотически линейной оценкой*, поскольку она имеет асимптотическую ковариационную матрицу $\mathbf{B}_n^{opt}$.

Наиболее простой и важный класс таких оценок дает

Следствие 3. *Пусть ковариационная матрица $\mathbf{V}_n$ представима в виде*

$$\mathbf{V}_n = \sigma_n^2\mathbf{W}_n^{-1}(\boldsymbol{\beta}_n), \tag{18}$$

где все элементы матрицы $\mathbf{W}_n(\cdot)$ — некоторые известные функции. Пусть выполнены все условия предположения 2 при

$$\boldsymbol{\Gamma}_n(\cdot) = \mathbf{X}_n^\top\mathbf{W}_n(\cdot). \tag{19}$$

В этом случае имеет место сходимость (17).

Замечание 3. Если точное значение матрицы $\mathbf{W}_n(\boldsymbol{\beta}_n)$ известно, то для оценивания параметра $\boldsymbol{\beta}_n$ можно воспользоваться известными результатами из линейного регрессионного анализа, о чем мы уже упоминали во введении. В этом случае матрица $\mathbf{X}_n^\top \mathbf{W}_n(\boldsymbol{\beta}_n)$ состоит из известных постоянных и эти постоянные определяют по формуле (2) наилучшую линейную несмещенную оценку $\boldsymbol{\beta}_n^*$ с ковариационной матрицей $\mathbf{B}_n^{opt}$, которая определена в (3). Если к тому же вектор погрешностей $\boldsymbol{\epsilon}_n$ имеет нормальное распределение с нулевым средним, то эта оценка $\boldsymbol{\beta}_n^*$ еще и эффективна, т. е. является наилучшей оценкой параметра $\boldsymbol{\beta}_n$ в классе всех несмещенных оценок.

Но оценка $\boldsymbol{\beta}_n^{**}$ из следствия 2 имеет ту же самую асимптотическую ковариационную матрицу $\mathbf{B}_n^{opt}$, хотя наша оценка построена при значительно более слабых предположениях.

Следующее общее утверждение еще более расширяет класс ситуаций, когда можно построить такие наилучшие асимптотически линейные оценки $\boldsymbol{\beta}_n^{**}$.

Теорема 3. *Пусть все условия предположения* 2 *выполнены при*

$$\boldsymbol{\Gamma}_n(\cdot) = \mathbf{X}_n^\top \widetilde{\mathbf{W}}_n(\cdot), \tag{20}$$

и пусть, кроме того,

$$\mathbf{Q}_{n2} = \mathbf{U}_{n0} \mathbf{X}_n^\top \widetilde{\mathbf{W}}_n(\boldsymbol{\beta}_n) \mathbf{X}_n \left(\mathbf{B}_n^{opt}\right)^{1/2} \xrightarrow{p} \mathbf{I}. \tag{21}$$

В этом случае справедливо (17).

Замечание 4. Таким образом, если условие (18) выполнено с «достаточно хорошей» матрицей $\mathbf{W}_n(\cdot)$, то формула (19) дает нам явный вид оптимальных функций $\boldsymbol{\Gamma}_n(\cdot)$. Однако если условие (18) выполнено, но функции из (19) не удовлетворяют предположению 2, то имеет смысл вместо (19) использовать некие близкие к ним «хорошие» функции вида (20) (например, можно попытаться использовать сглаженные функции из (19)). Именно на такого типа ситуации и рассчитана теорема 3.

2.5. Во всех приведенных выше утверждениях разность $\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n$ нормируется матрицами, которые зависят от неизвестного параметра $\boldsymbol{\beta}_n$ и, возможно, от ковариационной матрицы вектора $\boldsymbol{\epsilon}_n$, которая также может быть неизвестной. Поэтому при построении доверительных интервалов и проверке гипотез могут оказаться полезными утверждения, в которых разность $\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n$ нормируется некоторыми матрицами, являющимися статистиками. Укажем возможные подходы к решению указанной задачи.

Положим

$$\boldsymbol{\epsilon}_n^{**} := \mathbf{Y}_n - \mathbf{X}_n \boldsymbol{\beta}_n^{**}, \quad \mathbf{V}_n^{**} = \mathbf{diag}\{(\boldsymbol{\epsilon}_n^{**})^2\}. \tag{22}$$

Таким образом, $\mathbf{V}_n^{**}$ — это диагональная матрица, у которой по диагонали стоят квадраты компонент вектора $\boldsymbol{\epsilon}_n^{**}$. Тем самым следующее утверждение полезно в случае, когда оцениваемая ковариационная матрица $\mathbf{V}_n$ либо сама диагональна, либо близка к диагональной.

Следствие 4. *Пусть выполнены все условия теоремы* 1 *и*

$$\mathbf{Q}_{n3}(\boldsymbol{\beta}_n^*) := \mathbf{U}_n^* \big(\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*) \mathbf{V}_n^{**} \boldsymbol{\Gamma}_n^\top(\boldsymbol{\beta}_n^*)\big)^{1/2} \xrightarrow{p} \mathbf{I}. \tag{23}$$

Тогда

$$\boldsymbol{\alpha}_{n5} := \big(\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*) \mathbf{V}_n^{**} \boldsymbol{\Gamma}_n^\top(\boldsymbol{\beta}_n^*)\big)^{-1/2} \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*) \mathbf{X}_n (\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}). \tag{24}$$

Если верно условие (18), то следующее утверждение может оказаться полезным, например, при $\widetilde{\mathbf{W}}_n = \mathbf{W}_n$, поскольку в этом случае вектор $\mathbf{W}_n^{1/2}(\boldsymbol{\beta}_n)\boldsymbol{\epsilon}_n/\sigma_n$ с необходимостью имеет единичную ковариационную матрицу.

Следствие 5. *Пусть верны все условия теоремы* 1 *и*

$$\mathbf{Q}_{n4}(\boldsymbol{\beta}_n^*) := \sigma_n^* \mathbf{U}_n^* \big(\mathbf{X}_n^\top \widetilde{\mathbf{W}}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n\big)^{1/2} \xrightarrow{p} \mathbf{I} \quad \text{при } (\sigma_n^*)^2 := (\boldsymbol{\epsilon}_n^{**})^\top \widetilde{\mathbf{W}}_n(\boldsymbol{\beta}_n^*)\boldsymbol{\epsilon}_n^{**}/n.$$

В этом случае

$$\boldsymbol{\alpha}_{n6} := \big(\mathbf{X}_n^\top \widetilde{\mathbf{W}}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n\big)^{1/2}(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n)/\sigma_n^* \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}).$$

2.6. Сделаем несколько общих замечаний.

Замечание 5. Всюду в работе $\boldsymbol{\beta}_n$ — это неизвестный параметр. Кроме того, неизвестными параметрами могут быть также и элементы ковариационной матрицы. 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 [1].)

Замечание 6. В классических задачах математической статистики естественным является предположение, что параметры не зависят от числа наблюдений. Однако при изучении вероятностей ошибок второго рода в случае сближающихся наблюдений приходится доказывать асимптотическую нормальность оценок в более трудной для изучения ситуации, когда параметры зависят от числа наблюдений. Аналогично обстоит дело и при получении некоторых равномерных по параметру результатов (см, например, [1, гл. 2 § 37]). Чтобы не исключать возможности применения результатов работы к таким задачам мы все время считаем, что все параметры, включая основной параметр $\boldsymbol{\beta}_n$, могут зависеть от числа наблюдений n.

Замечание 7. Условия теорем 1 и 3 и следствий 2–5 носят методологический и концептуальный характер и такой путь изложения результатов представляется, по мнению авторов, наиболее наглядным. Тем более что проверка некоторых из условий этих утверждений не содержит принципиальных проблем. Так, например, условие (7) — это многомерная ЦПТ для специальных схем серий. Используя прием Крамера — Уолда, в каждом конкретном случае проверку условия (7) можно свести к проверке справедливости одномерной ЦПТ. На первый взгляд труднее всего проверять условия (8) и (10), которые даже в случае независимых наблюдений прием Крамера — Уолда превратит в условия вида

$$\sum_{i=1}^n (f_{ni}(\boldsymbol{\beta}_n^*) - f_{ni}(\boldsymbol{\beta}_n))\epsilon_{ni} \xrightarrow{p} 0, \quad \sum_{i=1}^n (f_{ni}(\boldsymbol{\beta}_n^*) - f_{ni}(\boldsymbol{\beta}_n)) \xrightarrow{p} 0.$$

Но техника проверки такого рода условий развита авторами в [4, 7], а в настоящей работе мы продемонстрируем ее в доказательстве приводимой ниже теоремы 4.

Замечание 8. Как мы отмечали во введении, при выборе оценки $\boldsymbol{\beta}_n^*$ проще всего взять оценку, являющуюся решением системы (2) при $c_{nji} = X_{nij}$. В этом случае если матрица $\mathbf{X}_n^\top \mathbf{X}_n$ невырожденна, то $\boldsymbol{\beta}_n^* \equiv \big(\mathbf{X}_n^\top \mathbf{X}_n\big)^{-1}\mathbf{X}_n^\top \mathbf{Y}_n$ — это обычная оценка параметра $\boldsymbol{\beta}_n$ по методу наименьших квадратов.

Другая возможная рекомендация: положить в (2) $c_{nji} = \gamma_{nji}(\boldsymbol{\beta}_{no})$. Здесь $\boldsymbol{\beta}_{no}$ — это некоторое возможное значение параметра $\boldsymbol{\beta}_n$, которое должен выбрать статистик, желательно «поближе» к неизвестному истинному значению параметра $\boldsymbol{\beta}_n$. Такая рекомендация особенно разумна, когда выполнены условия следствия 3 или теоремы 3. Ниже, в замечании 13, будут приведены дополнительные аргументы в пользу такого выбора.

В замечании 10 показано, что неудачный выбор чисел $\{c_{nji}\}$ может сузить класс функций $\{\gamma_{nji}\}$, которые могут использоваться для построения улучшенных оценок $\boldsymbol{\beta}_n^{**}$.

Замечание 9. Если при изучении асимптотической нормальности оценки $\boldsymbol{\beta}_n^{**}$ получена сходимость вида $\mathbf{Q}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n)$ с некоторой матрицей $\mathbf{Q}_n$, то эту нормирующую матрицу $\mathbf{Q}_n$ можно несколько изменить (см. подробности в лемме 2 в п. 3.2).

2.7. В этом пункте в предположении независимости наблюдений мы, как неоднократно говорилось выше, проведем еще более детальное исследование поведения наших оценок в случае одномерного неизвестного параметра. Пусть далее

$$m = 1, \quad X_{ni} = X_{ni1}, \quad c_{ni} = c_{n1i}, \quad \gamma_{ni}(\cdot) = \gamma_{n1i}(\cdot), \quad \beta_n = \beta_{n1} = \boldsymbol{\beta}_n, \ \dots . \tag{25}$$

Другими словами, здесь предполагается, что при всех $n = 1, 2, \dots$ и каждом $i = 1, \dots, n$ наблюдаются случайная величина Y_{ni} и неслучайное число X_{ni}, о которых известно, что

$$Y_{ni} = \beta_n X_{ni} + \epsilon_{ni}, \quad \mathbf{E}\epsilon_{ni} = 0, \quad \mathbf{D}\epsilon_{ni} = \sigma_{ni}^2 < \infty, \tag{26}$$

где β_n — неизвестный параметр. При построении оценок для этого параметра будем использовать специально подобранные константы c_{ni} и функции $\gamma_{ni}(\cdot)$, вопрос о выборе которых более подробно обсудим ниже, в замечаниях 12 и 13.

Предположение 3. *Начиная с некоторого n*

$$A_{nc} := \sum c_{ni} X_{ni} \neq 0, \quad A_n := \sum \gamma_{ni}(\beta_n) X_{ni} \neq 0, \quad B_n^2 := \sum \gamma_{ni}^2(\beta_n)\sigma_{ni}^2 > 0, \tag{27}$$

случайные величины $\epsilon_{n1}, \dots, \epsilon_{nn}$ независимы между собой, а случайные величины $\{\gamma_{ni}(\beta_n)\epsilon_{ni}\}$ удовлетворяют условию Линдеберга. Кроме того,

$$d_{nc} \to 0 \quad \text{при } d_{nc}^2 := \sum c_{ni}^2 \sigma_{ni}^2 / A_{nc}^2. \tag{28}$$

Подчеркнем, что в (27) и далее символ $\sum$ без индексов будет использоваться только тогда, когда суммирование ведется по переменной i от 1 до n.

Для оценивания неизвестного параметра β_n в этом пункте будем использовать оценки β_n^* и β_n^{**}, которые имеют следующий вид:

$$\beta_n^* = \frac{\sum c_{ni} Y_{ni}}{\sum c_{ni} X_{ni}}, \quad \beta_n^{**} = \frac{\sum \gamma_{ni}(\beta^*) Y_{ni}}{\sum \gamma_{ni}(\beta^*) X_{ni}}. \tag{29}$$

Условимся доопределять их произвольным измеримым образом всякий раз, когда соответствующий знаменатель в (29) обращается в нуль. В силу предположения 3 оценка β_n^* определяется однозначно по формуле (29) начиная с некоторого n, когда $A_{nc} \neq 0$. Заметим еще, что из доказываемой далее асимптотической нормальности оценок β_n^{**} вытекает, что знаменатель в определении (29) оценки β_n^{**} может обращаться в нуль лишь с вероятностью, стремящейся к нулю. Поэтому способ доопределения этой оценки также не будет играть никакой роли.

Положим

$$K_{ni} := \sup\left\{ \frac{|\gamma_{ni}(t_2) - \gamma_{ni}(t_1)|}{|t_2 - t_1|} : \beta_n - 1 \le t_1 < t_2 \le \beta_n + 1 \right\}. \tag{30}$$

Подчеркнем, что используемое далее неявно условие $K_{ni} < \infty$ слабее, чем классическое условие Липшица.

Теорема 4. *Пусть выполнено предположение* 3 *и еще*

$$d_{nc}(\rho_{n1}+\rho_{n2})\to 0 \quad \text{при} \quad \rho_{n1}^2 := \sum K_{ni}^2\sigma_{ni}^2/B_n^2 \quad \text{и} \quad \rho_{n2} := \sum K_{ni}|X_{ni}|/|A_n|. \tag{31}$$

Тогда

$$\alpha_{n7} := \frac{\beta_n^{**}-\beta_n}{d_n(\gamma_\bullet)} \Longrightarrow \mathscr{N}(0,1) \quad \text{при} \quad d_n^2(\gamma_\bullet) := \frac{\sum \gamma_{ni}^2(\beta_n)\sigma_{ni}^2}{\left(\sum \gamma_{ni}(\beta_n)X_{ni}\right)^2} \equiv \frac{B_n^2}{A_n^2}. \tag{32}$$

При построении доверительных интервалов и проверке гипотез удобно иметь аналог утверждения (32), в котором вместо параметра $d_n(\gamma_\bullet)$ стоит какая-нибудь статистика. Приведем требуемый результат.

Следствие 6. *Пусть справедливы все условия теоремы* 4 *и*

$$\rho_{n3}^2 := \sum \gamma_{ni}^2(\beta_n)X_{ni}^2/A_n^2 \to 0. \tag{33}$$

Тогда

$$\alpha_{n8} := \frac{\beta_n^{**}-\beta_n}{d_n^{**}} \Longrightarrow \mathscr{N}(0,1) \quad \text{при } d_n^{**} := \frac{\left(\sum \gamma_{ni}^2(\beta_n^*)(Y_{ni}-\beta_n^{**}X_{ni})^2\right)^{1/2}}{\sum \gamma_{ni}(\beta_n^*)X_{ni}}. \tag{34}$$

Укажем теперь более наглядные достаточные условия для сходимостей (32) и (34) в случае, когда

$$\operatorname{sign}\gamma_{ni}(\beta_n) = \operatorname{sign}X_{ni} \quad \text{при всех } i=1,\dots,n. \tag{35}$$

Следствие 7. *Пусть выполнено предположение* 3 *и начиная с некоторого* n *справедливо условие* (35), *причем все функции* $\gamma_{ni}(\cdot)$ *имеют производные* $\gamma_{ni}'(\cdot)$. *Пусть, кроме того,*

$$d_{nc}K_n \to 0, \quad \text{где } K_n := \max_{\{i:X_{ni}\neq 0\}} \sup_{\beta_n-1\le t\le \beta_n+1} |\gamma_{ni}'(t)|/|\gamma_{ni}(\beta_n)|. \tag{36}$$

В этом случае имеет место сходимость (32). *Если еще*

$$\rho_{n4} := \max_{i\le n}\gamma_{ni}(\beta_n)X_{ni}/A_n \to 0, \tag{37}$$

то имеет место и сходимость (34).

Замечание 10. Из (29) и (28) немедленно следует, что

$$\beta_n^* = \beta_n + \sum c_{ni}\epsilon_{ni}/A_{nc} \quad \text{и} \quad \mathbf{E}(\beta_n^*-\beta_n)^2 = \mathbf{D}\beta_n^* = d_{nc}^2. \tag{38}$$

Таким образом, определенная в (28) величина d_{nc} характеризует точность, с которой оценка первого шага β_n^* оценивает неизвестный параметр β_n. С другой стороны, нетрудно видеть, что введенные в (30), (31) и (36) величины являются характеристиками гладкости функций $\gamma_{ni}(\cdot)$, которые определяют оценки второго шага β_n^{**}. Но условия (31) и (36) можно переписать в виде: $\rho_{n1}+\rho_{n2} = o\big(d_{nc}^{-1}\big)$ и $K_n = o\big(d_{nc}^{-1}\big)$ соответственно. Таким образом, чем точнее оценка первого шага β_n^* оценивает неизвестный параметр β_n, тем для большего класса функций $\gamma_{ni}(\cdot)$ справедливы условия (31) или (36) и, следовательно, тем для большего класса оценок второго шага β_n^{**} имеет место асимптотическая нормальность (32).

Рассмотрим более подробно случай, когда выполнено условие (4). Следующее утверждение усиливает результат теоремы 2.

Теорема 5. *Пусть при некотором n выполнены условия* (4) *и* (27). *В этом случае определена величина $d_n(\gamma_\bullet)$ и*

$$d_n^2(\gamma_\bullet) \geq d_n^2(\gamma_\bullet^{opt}) \equiv \sigma_n^2 / \sum X_{ni}^2 w_{ni}(\beta_n) > 0 \quad \text{при } \gamma_{ni}^{opt}(\cdot) := X_{ni} w_{ni}(\cdot). \tag{39}$$

Если к тому же верно условие (35), *то* $\operatorname{sign} \gamma_{ni}(\beta_n) = \operatorname{sign} \gamma_{ni}^{opt}(\beta_n)$ *при всех i и*

$$1 \leq \frac{d_n(\gamma_\bullet)}{d_n(\gamma_\bullet^{opt})} \leq 1 + \frac{(\sqrt{H_n/h_n} - 1)^2}{2\sqrt{H_n/h_n}} \leq \sqrt{\frac{H_n}{h_n}}, \tag{40}$$

где

$$h_n := \min_{\{i: X_{ni} \neq 0\}} \gamma_{ni}(\beta_n)/\gamma_{ni}^{opt}(\beta_n) > 0 \quad \text{и} \quad H_n := \max_{\{i: X_{ni} \neq 0\}} \gamma_{ni}(\beta_n)/\gamma_{ni}^{opt}(\beta_n) < \infty. \tag{41}$$

Замечание 11. Неравенство (40) можно интерпретировать как некоторое свойство устойчивости оценок β_n^{**} как функционалов, зависящих от набора функций $\gamma_{ni}(\cdot)$: чем лучше числа $\gamma_{ni}(\beta_n)$ приближают оптимальные $\gamma_{ni}^{opt}(\beta_n)$, тем меньше $d_n(\gamma_\bullet)$ отличается от $d_n(\gamma_\bullet^{opt})$.

Отметим, что неравенство (40) при малых значениях H_n/h_n может давать хорошие оценки. Так, из (40) при $H_n = 1/h_n = 2$ следует, что

$$1 \leq d_n(\gamma_\bullet)/d_n(\gamma_\bullet^{opt}) \leq 1.25, \quad \text{если } 1/2 \leq \gamma_{ni}(\beta_n)/\gamma_{ni}^{opt}(\beta_n) \leq 2,$$

т. е. мы не так много теряем, когда используем функции, повторяющие поведение оптимальных с точностью до константы.

Замечание 12. Если выполнено условие (4), то можно дать некоторые рекомендации по выбору функций $\gamma_{ni}(\cdot)$. Во-первых, в качестве $\gamma_{ni}(\cdot)$ лучше всего взять оптимальные функции $\gamma_{ni}^{opt}(\cdot) = X_{ni} w_{ni}$. Но это можно сделать только если функции $w_{ni}(\cdot)$ нам известны, а функции $\gamma_{ni}(\cdot) = X_{ni} w_{ni}(\cdot)$ удовлетворяют условиям (31) теоремы 4.

А если функции $w_{ni}(\cdot)$ известны, но $\gamma_{ni}(\cdot) = X_{ni} w_{ni}(\cdot)$ не удовлетворяют условиям (31), то можно использовать функции $\gamma_{ni}(\cdot) = X_{ni} \widetilde{w}_{ni}(\cdot)$, где $\widetilde{w}_{ni}(\cdot)$ получены в результате соответствующего сглаживания функций $w_{ni}(\cdot)$.

Если же функции $w_{ni}(\cdot)$ нам не известны, то также можно рекомендовать использовать функции вида $\gamma_{ni}(\cdot) = X_{ni} \widetilde{w}_{ni}(\cdot)$, где $\widetilde{w}_{ni}(\cdot)$ должны по-возможности «не очень сильно отличаться» от неизвестных функций $w_{ni}(\cdot)$.

Следствие 8. *Пусть при некотором n выполнены условия* (4) *и*

$$A_{nc} \neq 0 \quad \text{и} \quad \operatorname{sign} c_{ni} = \operatorname{sign} X_{ni} \quad \text{для всех } i = 1, \ldots, n.$$

В этом случае справедливо соотношение

$$1 \leq \frac{d_{nc}}{d_n(\gamma_\bullet^{opt})} \leq 1 + \frac{(\sqrt{H_{nc}/h_{nc}} - 1)^2}{2\sqrt{H_{nc}/h_{nc}}} \leq \sqrt{\frac{H_{nc}}{h_{nc}}},$$

где

$$h_{nc} := \min_{\{i: X_{ni} \neq 0\}} c_{ni}/\gamma_{ni}^{opt}(\beta_n) > 0 \quad \text{и} \quad H_{nc} := \max_{\{i: X_{ni} \neq 0\}} c_{ni}/\gamma_{ni}^{opt}(\beta_n) < \infty.$$

Замечание 13. В качестве констант c_{ni}, определяющих β_n^*, проще всего взять $c_{ni} = X_{ni}$: в этом случае $\beta_n^* = \sum Y_{ni} X_{ni} / \sum X_{ni}^2$ — это стандартная оценка по методу наименьших квадратов.

Другая возможная рекомендация: положить $c_{ni} = \gamma_{ni}(\beta_{no})$ при некотором значении β_{no}. Как следует из замечания 10, такой выбор будет наилучшим в случае, если нам удастся получить, что $d_{nc} = O\big(d_n(\gamma_\bullet^{opt})\big)$. В силу следствия 8 для этого достаточно добиться, например, выполнения условия $H_{nc}/h_{nc} = O(1)$.

§ 3. Доказательства

В п. 3.1 и частично в п. 3.5, где число n фиксировано, будем опускать нижний индекс n.

3.1. В этом пункте докажем теорему 2, поскольку при ее выводе потребуются лишь алгебраические свойства матриц и следующее элементарное утверждение.

Лемма 1. *Если* $\mathbf{B}_1 \geq \mathbf{B}_2$, *то* $\mathbf{A}\mathbf{B}_1\mathbf{A}^\top \geq \mathbf{A}\mathbf{B}_2\mathbf{A}^\top$ *для любой матрицы* $\mathbf{A}$ *соответствующей размерности.*

Доказательство теоремы. Поскольку матрица $\mathbf{G}_n\mathbf{X}_n$ невырожденна, то $\mathbf{X}_n$ имеет ранг m. Но, как отмечено в [2, § 3.6], в этом случае матрица $\mathbf{X}_n^\top \mathbf{V}_n^{-1}\mathbf{X}_n$ имеет обратную, а значит, существует матрица $\mathbf{B}_n^{opt}$.

Теперь из определения (16) нетрудно получить, что

$$\begin{gathered}\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{B}^o = \mathbf{X}^\top - (\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X})\mathbf{B}^{opt}\mathbf{X}^\top = \mathbf{0} = \mathbf{B}^o\mathbf{V}^{-1}\mathbf{X}, \\ \mathbf{B}^o = \mathbf{B}^{o\top} = \mathbf{B}^o\mathbf{V}^{-1}\mathbf{B}^{o\top}.\end{gathered} \tag{42}$$

В частности, из (42) вытекает, что верно равенство (15), поскольку обе части в нем равны $\mathbf{G}^\top\mathbf{B}^o\mathbf{G}$. Но из (15) и леммы 1 следует, что правая часть в (15) неотрицательно определена. Значит, $\mathbb{B}(\mathbf{G}) \geq \mathbf{B}^{opt}$.

Кроме того, из (15) и (42) вытекает, что $\mathbb{B}(\mathbf{R}\mathbf{X}^\top\mathbf{V}^{-1}) = \mathbb{B}(\mathbf{X}^\top\mathbf{V}^{-1}) = \mathbf{B}^{opt}$. Таким образом, осталось разобрать ситуацию, когда $\mathbb{B}(\mathbf{G}) = \mathbf{B}^{opt}$. В силу (15) в этом случае $\mathbf{G}\mathbf{B}^o = \mathbf{0}$. Значит, ввиду (16)

$$\mathbf{G}\mathbf{B}^o = \mathbf{0} = \mathbf{G}\mathbf{B}^o\mathbf{V}^{-1} = \mathbf{G} - \mathbf{G}\mathbf{X}\mathbf{B}^{opt}\mathbf{X}^\top\mathbf{V}^{-1}.$$

Следовательно, $\mathbf{G} = \mathbf{R}\mathbf{X}^\top\mathbf{V}^{-1}$ при $\mathbf{R} = \mathbf{G}\mathbf{X}\mathbf{B}^{opt}$. □

3.2. В этом пункте докажем остальные утверждения, сформулированные в пп. 2.2–2.4.

Лемма 2. *Пусть вектор* $\boldsymbol{\alpha}_n$ *и матрица* $\mathbf{Q}_n$ *таковы, что*

$$\mathbf{Q}_n\boldsymbol{\alpha}_n \Longrightarrow \mathscr{N}_m(\mathbf{0}, I), \tag{43}$$

а $\mathbf{Q}_n \xrightarrow{p} \mathbf{I}$. *Тогда*

$$\boldsymbol{\alpha}_n \Longrightarrow \mathscr{N}_m(\mathbf{0}, I). \tag{44}$$

Если же $\mathbf{Q}_n$ *неслучайна, то из* (43) *следует* (44) *для любой ортонормированной матрицы* $\mathbf{Q}_n$, *т. е. матрицы, удовлетворяющей условию* $\mathbf{Q}_n^\top\mathbf{Q}_n = \mathbf{I}$.

Доказательство первого утверждения леммы нетрудно извлечь из приема Крамера — Уолда, поэтому мы его опускаем. Доказательство (44) для неслучайной ортонормированной матрицы $\mathbf{Q}_n$ можно найти, например, в [5, лемма 2]. □

Доказательство теоремы 1. В силу (6) имеем

$$\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n\boldsymbol{\beta}_n^{**} = \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{Y}_n = \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n\boldsymbol{\beta}_n + \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\boldsymbol{\epsilon}_n,$$

поэтому $\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) = \boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\boldsymbol{\epsilon}_n$. Значит, ввиду (7) и (8)

$$\boldsymbol{\alpha}_{n1} \equiv \mathbf{U}_n^*\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) = \mathbf{U}_n^*\boldsymbol{\Gamma}_n(\boldsymbol{\beta}_n^*)\boldsymbol{\epsilon}_n = \boldsymbol{\alpha}_{n0} + \boldsymbol{\Delta}_{n0}(\boldsymbol{\beta}_n^*) \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}),$$

что доказывает (9). □

Доказательство следствия 1. Из условия (10) при невырожденных матрицах $\mathbf{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{X}_n$ вытекает, что матрицы $\mathbf{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n$ невырожденны с вероятностью, стремящейся к 1. Значит, выполнены все условия теоремы 1, и мы можем воспользоваться утверждением (9) этой теоремы. Учитывая обозначения, введенные в (10) и (11), имеем

$$\mathbf{Q}_{n1}(\boldsymbol{\beta}_n^*)\boldsymbol{\alpha}_{n2} = \mathbf{U}_n^*\mathbf{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \equiv \boldsymbol{\alpha}_{n1} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}).$$

Но теперь из леммы 2 находим, что $\boldsymbol{\alpha}_{n2} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I})$, когда $\mathbf{Q}_{n1}(\boldsymbol{\beta}_n^*) \xrightarrow{p} \mathbf{I}$. $\square$

Доказательство следствия 2. Сходимость (11) при $\mathbf{U}_n = \mathbf{U}_{n0}$ является частным случаем теоремы 1. Положим $\mathbf{Q}_n := \mathbf{U}_{n0}\mathbf{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{X}_n\mathbb{B}_n^{1/2}(\mathbf{\Gamma}_n(\boldsymbol{\beta}_n))$. В силу определения (14)

$$\mathbf{Q}_n\mathbf{Q}_n^\top = \mathbf{U}_{n0}\mathbf{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{V}_n\mathbf{\Gamma}_n^\top(\boldsymbol{\beta}_n)\mathbf{U}_{n0}^\top = \mathbf{U}_{n0}\mathbf{U}_{n0}^{-2}\mathbf{U}_{n0}^\top = \mathbf{I}.$$

Значит, $\mathbf{Q}_n$ — ортонормированная матрица. Далее, из уже доказанной сходимости (11) с учетом определения (13) следует, что

$$\mathbf{Q}_n\boldsymbol{\alpha}_{n3} = \mathbf{U}_{n0}\mathbf{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{X}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \equiv \boldsymbol{\alpha}_{n2} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}).$$

Таким образом, из леммы 2 находим, что $\boldsymbol{\alpha}_{n3} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I})$. $\square$

Доказательство теоремы 3. Используя следствие 2 и обозначения, введенные в (11), (17), (20) и (21), находим

$$\mathbf{Q}_{n2}\boldsymbol{\alpha}_{n4} = \mathbf{U}_{n0}\mathbf{\Gamma}_n(\boldsymbol{\beta}_n)\mathbf{X}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \equiv \boldsymbol{\alpha}_{n2} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}).$$

Значит, $\boldsymbol{\alpha}_{n4} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I})$ ввиду леммы 2. $\square$

Для доказательства следствия 3 достаточно воспользоваться утверждением теоремы 3 и заметить, что $\mathbf{Q}_{n2} = \mathbf{I}$ при выполнении условия (19).

Следствие 3 можно также извлечь из следствия 2 и замечания 2, поскольку в этом случае $\mathbf{\Gamma}_n(\boldsymbol{\beta}_n) = \sigma_n^2\mathbf{X}_n^\top\mathbf{V}_n^{-1} = \mathbf{R}_n\mathbf{X}_n^\top\mathbf{V}_n^{-1}$ при $\mathbf{R}_n = \sigma_n^2\mathbf{I}$.

Доказательство следствий 4, 5. Воспользуемся утверждением (9) теоремы 1. Учитывая обозначения, введенные в доказываемых следствиях, нетрудно получить следующие соотношения:

$$\mathbf{Q}_{n3}(\boldsymbol{\beta}_n^*)\boldsymbol{\alpha}_{n5} = \mathbf{U}_n^*\mathbf{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \equiv \boldsymbol{\alpha}_{n1} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}),$$

$$\mathbf{Q}_{n4}(\boldsymbol{\beta}_n^*)\boldsymbol{\alpha}_{n6} = \mathbf{U}_n^*\mathbf{\Gamma}_n(\boldsymbol{\beta}_n^*)\mathbf{X}_n(\boldsymbol{\beta}_n^{**} - \boldsymbol{\beta}_n) \equiv \boldsymbol{\alpha}_{n1} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I}).$$

Из этих фактов и леммы 2 немедленно следует, что $\boldsymbol{\alpha}_{n5} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I})$, когда $\mathbf{Q}_{n3}(\boldsymbol{\beta}_n^*) \xrightarrow{p} \mathbf{I}$ и $\boldsymbol{\alpha}_{n6} \Longrightarrow \mathscr{N}_m(\mathbf{0}, \mathbf{I})$, если $\mathbf{Q}_{n4}(\boldsymbol{\beta}_n^*) \xrightarrow{p} \mathbf{I}$. $\square$

3.3. В этом пункте докажем несколько вспомогательных утверждений, необходимых для вывода теоремы 4 и следствия 6. Положим

$$\tilde{\beta}_n = \beta_n + f(\delta_n), \quad \text{где } \delta_n := \beta_n^* - \beta_n \text{ и } f(x) = \begin{cases} x, & \text{если } |x| \le 1, \\ \operatorname{sign} x, & \text{если } |x| \ge 1. \end{cases} \tag{45}$$

Определенную таким образом величину $\tilde{\beta}_n$ будем интерпретировать как «срезку» случайной величины β_n^*.

Лемма 3. *Если справедливо предположение* 3, *то*

$$\mathbf{E}\Big|\sum \delta_{ni}(\tilde\beta_n)\epsilon_{ni}\Big| \le 8d_{nc}\Big(\sum K_{ni}^2\sigma_{ni}^2\Big)^{1/2} \quad \textit{при } \delta_{ni}(\cdot) := \gamma_{ni}(\cdot)-\gamma_{ni}(\beta_n). \tag{46}$$

Доказательство. Воспользуемся теоремой 5 из [7] при

$$u_i = c_{ni}\epsilon_{ni}/A_{nc}, \quad v_i = 0, \quad r = 1, \quad \theta = 4, \quad g_i(4+t) = \gamma_{ni}(t)\epsilon_{ni}, \quad \bar g_i = K_{ni}|\epsilon_{ni}|.$$

В этом случае утверждение (65) упомянутой теоремы из [7] превратится в требуемое неравенство (46). □

Замечание 14. Неравенство (46) является ключевым местом при выводе теоремы 4. Именно благодаря этому неравенству удается получить асимптотическую нормальность улучшенных оценок β_n^{**} при предположениях (31) или (36), в которых от функций $\{\gamma_{ni}(\cdot)\}$ требуется меньше, чем существование ограниченных первых производных. Техника получения такого рода результатов развита авторами в [4, 7].

Лемма 4. *Если выполнено предположение* 3, *то*

$$\mathbf{P}(\beta_n^* \ne \tilde\beta_n) \le \mathbf{E}\delta_n^2 = d_{nc}^2 \to 0.$$

Доказательство. Используя определение (45) и неравенство Чебышёва, имеем

$$\mathbf{P}(\beta_n^* \ne \tilde\beta_n) = \mathbf{P}(|\delta_n| > 1) \le \mathbf{E}\delta_n^2/1^2 = \mathbf{E}(\beta_n^* - \beta_n)^2.$$

Теперь из равенства (38) и условия (28) получаем утверждение леммы. □

Введем следующие обозначения:

$$\Delta_{n1}(\cdot) := \frac{1}{B_n}\sum \delta_{ni}(\cdot)\epsilon_{ni}, \quad \Delta_{n2}(\cdot) := \frac{1}{|A_n|}\sum |\delta_{ni}(\cdot)X_{ni}|, \qquad \Delta_{n3}^2(\cdot) := \frac{1}{B_n^2}\sum \delta_{ni}^2(\cdot)\epsilon_{ni}^2. \tag{47}$$

Лемма 5. *Если выполнены условия теоремы* 4, *то*

$$\Delta_{n1}(\beta_n^*) \xrightarrow{p} 0, \quad \Delta_{n2}(\beta_n^*) \xrightarrow{p} 0, \quad \Delta_{n3}(\beta_n^*) \xrightarrow{p} 0. \tag{48}$$

Доказательство. Прежде всего, докажем, что

$$\mathbf{E}|\Delta_{n1}(\tilde\beta_n)| \le 8d_{nc}\rho_{n1}, \quad \mathbf{E}\Delta_{n2}(\tilde\beta_n) \le d_{nc}\rho_{n2}, \quad \mathbf{E}\Delta_{n3}(\tilde\beta_n) \le d_{nc}\rho_{n1}, \tag{49}$$

где величины ρ_{nk} введены в (31). Заметим сразу, что, разделив обе части неравенства (52) на B_n, получим первое из неравенств в (49). Далее, из определений (30) и (45) имеем при всех n и i

$$|\delta_{ni}(\tilde\beta_n)| \equiv |\gamma_{ni}(\tilde\beta_n) - \gamma_{ni}(\beta_n)| \le K_{ni}|\tilde\beta_n - \beta_n| \equiv K_{ni}|f(\delta_n)| \le |\delta_n|K_{ni}, \tag{50}$$

так как $|f(x)| \le |x|$ при всех x ввиду определения (45). Но из (47) и (50) вытекает, что

$$\Delta_{n2}(\tilde\beta_n) \le |\delta_n|\sum K_{ni}|X_{ni}|/A_n \equiv |\delta_n|\rho_{n2} \quad \text{и} \quad \mathbf{E}\Delta_{n2}(\tilde\beta_n) \le \mathbf{E}|\delta_n|\rho_{n2} \le d_{nc}\rho_{n2}, \tag{51}$$

поскольку $\mathbf{E}|\delta_n| \le (\mathbf{E}\delta_n^2)^{1/2} = d_{nc}$ в силу леммы 4. Аналогично

$$\Delta_{n3}^2(\tilde\beta_n) \le \delta_n^2 S_n^2 \quad \text{при } S_n^2 := \sum K_{ni}^2\epsilon_{ni}^2/B_n^2.$$

Согласно неравенству Шварца

$$(\mathbf{E}\Delta_{n3}(\tilde\beta_n))^2 \le \mathbf{E}\delta_n^2\mathbf{E}S_n^2 = d_{nc}^2\sum K_{ni}^2\sigma_{ni}^2/B_n^2 \equiv d_{nc}^2\rho_{n1}^2, \qquad (52)$$

где использовано определение (31) величины ρ_{n1}.

Таким образом, из (51) и (52) вытекает, что доказаны все неравенства в (49). А этот факт при наличии условия (31) означает, что

$$\Delta_{n1}(\tilde\beta_n) \xrightarrow{p} 0, \quad \Delta_{n2}(\tilde\beta_n) \xrightarrow{p} 0, \quad \Delta_{n3}(\tilde\beta_n) \xrightarrow{p} 0.$$

Из полученных сходимостей немедленно следуют аналогичные сходимости из (48), потому что $\mathbf{P}(\beta_n^* \ne \tilde\beta_n) \to 0$ ввиду леммы 4. $\square$

Лемма 6. *Если случайные величины* $\{\gamma_{ni}(\beta_n)\epsilon_{ni}\}$ *удовлетворяют условию Линдеберга, то*

$$\alpha_n := \sum\gamma_{ni}(\beta_n)\epsilon_{ni}/B_n \Longrightarrow \mathscr{N}(0,1) \quad \text{и} \quad Q_n := \sum\gamma_{ni}^2(\beta_n)\epsilon_{ni}^2/B_n^2 \xrightarrow{p} 1. \qquad (53)$$

Указанные утверждения, а также достаточно подробные комментарии об условии Линдеберга можно найти, например, в [7, § 6].

3.4. В этом пункте докажем теорему 4 и следствие 6.

Лемма 7. *Если выполнены условия теоремы* 4, *то верно предположение* 2. *Кроме того, в этом случае*

$$\mathbf{\Gamma}_n(\boldsymbol\beta_n)\mathbf{X}_n = A_n, \quad \mathbf{U}_{n0} = 1/B_n, \quad \mathbb{B}_n(\mathbf{\Gamma}_n(\boldsymbol\beta_n)) = B_n^2/A_n^2 = d_n^2(\gamma_\bullet), \qquad (54)$$

$$\boldsymbol\alpha_{n0} = \alpha_n, \quad \mathbf{\Delta}_{n0}(\boldsymbol\beta_n^*) = \Delta_{n1}(\beta_n^*), \quad \mathbf{Q}_{n1}(\boldsymbol\beta_n^*) = \sum\gamma_{ni}(\beta_n^*)X_{ni}/A_n. \qquad (55)$$

Доказательство. Сравнивая обозначения, введенные в (12) и (14), с обозначениями из (27) и (32), нетрудно получить равенства (54). А если справедливы равенства из (12), то сравнение определений из (7), (8) и (10) с обозначениями из (53), (47) и (25) дает равенства (55).

Таким образом, осталось лишь проверить выполнение условий предположения 2, которые теперь имеют следующий вид:

$$\boldsymbol\alpha_{n0} \equiv \alpha_n \Longrightarrow \mathscr{N}(0,1), \quad \mathbf{\Delta}_{n0}(\boldsymbol\beta_n^*) \equiv \Delta_{n1}(\beta_n^*) \xrightarrow{p} 0, \quad \mathbf{Q}_{n1}(\boldsymbol\beta_n^*) \xrightarrow{p} 1. \qquad (56)$$

Первая сходимость в (56) установлена в лемме 6, а вторая — в лемме 5. Но последняя сходимость в (56) также следует из результатов леммы 6, поскольку из определений (46), (47) и (55) вытекает, что

$$\mathbf{Q}_{n1}(\boldsymbol\beta_n^*) - 1 = \sum\delta_{ni}(\beta_n^*)X_{ni}/A_n \quad \text{и} \quad |\mathbf{Q}_{n1}(\boldsymbol\beta_n^*) - 1| \le \Delta_{n2}(\beta_n^*) \xrightarrow{p} 0. \quad \square$$

Доказательство теоремы 4. Если выполнены условия теоремы 4, то ввиду леммы 7 верны также и условия следствия 2 из п. 2.2. Значит, справедливо и утверждение (13) этого следствия. Но из последнего равенства в (54) вытекает, что $\boldsymbol\alpha_{n3} = \alpha_{n7}$, т. е. доказанная сходимость (13) совпадает с точностью до обозначений с требуемой в теореме 4 сходимостью (32). $\square$

Лемма 8. *Если верны условия следствия* 6, *то при* $\mathbf{U}_n^* = 1/B_n$ *справедливы и предположения следствия* 4, *причем*

$$\boldsymbol{\alpha}_{n5} = \alpha_{n8} \quad \text{и} \quad \mathbf{Q}_{n3}^2(\boldsymbol{\beta}_n^*) = \sum \gamma_{ni}^2(\beta_n^*)\epsilon_{ni}^{**2}/B_n^2 \quad \text{при } \epsilon_{ni}^{**} = Y_{ni} - \beta_n^{**}X_{ni}. \tag{57}$$

Доказательство. Сравнивая обозначения из (22)–(24) с обозначениями из (34), нетрудно убедиться в справедливости обоих равенств в (57). Если выполнены все условия следствия 6, то справедливы и все предположения теоремы 4. Значит, в силу леммы 7 при $\mathbf{U}_n = 1/B_n$ верны и все условия предположения 1 как части условий предположения 2.

Таким образом, осталось проверить лишь условие (23) следствия 4, т. е. что $\mathbf{Q}_{n3}(\boldsymbol{\beta}_n^*) \xrightarrow{p} 1$. Из (26) и (57) нетрудно извлечь, что

$$\epsilon_{ni}^{**} - \epsilon_{ni} = -(\beta_n^{**} - \beta_n)X_{ni} = -\delta_n^{**}X_{ni}, \quad \text{где } \delta_n^{**} := \beta_n^{**} - \beta_n = \alpha_{n7}B_n/|A_n|. \tag{58}$$

При выводе последнего равенства использовано также (32). Поскольку $\gamma_{ni}(\beta_n^*) = \gamma_{ni}(\beta_n) + \delta_{ni}(\beta_n^*)$ ввиду (46), то

$$\gamma_{ni}(\beta_n^*)\epsilon_{ni}^{**} = \gamma_{ni}(\beta_n)\epsilon_{ni} + \delta_{ni}(\beta_n^*)\epsilon_{ni} - \delta_n^{**}\gamma_{ni}(\beta_n)X_{ni} - \delta_n^{**}\delta_{ni}(\beta_n^*)X_{ni}. \tag{59}$$

С целью сделать некоторые соотношения более обозримыми, условимся для наборов символов вида $a_{ni}(\cdot)b_{ni}$, $i = 1, \dots, n$, использовать следующее обозначение:

$$\|a_\bullet(\cdot)b_\bullet\|^2 := \sum a_{ni}^2(\cdot)b_{ni}^2. \tag{60}$$

Из (59) с учетом известных свойств введенной в (60) евклидовой нормы имеем

$$\big|\|\gamma_\bullet(\beta_n^*)\epsilon_\bullet^{**}\| - \|\gamma_\bullet(\beta_n)\epsilon_\bullet\|\big| \le \|\delta_\bullet(\beta_n^*)\epsilon_\bullet\| + |\delta_n^{**}|\|\gamma_\bullet(\beta_n)X_\bullet\| + |\delta_n^{**}|\|\delta_\bullet(\beta_n^*)X_\bullet\|.$$

Разделим это неравенство на B_n и воспользуемся представлением из (58) для δ_n^{**}. В итоге с учетом (57) и определений (33), (47) и (53) получаем

$$|\mathbf{Q}_{n3}(\boldsymbol{\beta}_n^*) - Q_n| \le \Delta_{n3}(\beta_n^*) + |\alpha_{n7}|\rho_{n3} + |\alpha_{n7}|\|\delta_\bullet(\beta_n^*)X_\bullet\|/|A_n|. \tag{61}$$

Но из (60) и (47) вытекает, что

$$\|\delta_\bullet(\beta_n^*)X_\bullet\|/|A_n| \le \sum |\delta_{ni}(\beta_n^*)X_{ni}|/|A_n| \equiv \Delta_{n2}(\beta_n^*) \xrightarrow{p} 0. \tag{62}$$

Используя сходимости (32), (33), (48) и (62), нетрудно заметить, что все три слагаемых в правой части формулы (61) сходятся к нулю по вероятности. Следовательно, $\mathbf{Q}_{n3}(\boldsymbol{\beta}_n^*) - Q_n \xrightarrow{p} 0$. Но $Q_n \xrightarrow{p} 1$ в силу леммы 6. Значит, $\mathbf{Q}_{n3}(\boldsymbol{\beta}_n^*) \xrightarrow{p} 1$. □

Доказательство следствия 6. Если выполнены условия следствия 6, то в силу леммы 8 справедливы также и условия следствия 4 при $\mathbf{U}_n^* = 1/B_n$. Тогда верно и утверждение (24) этого следствия. Из равенства $\boldsymbol{\alpha}_{n5} = \alpha_{n8}$ в (57) вытекает, что с точностью до обозначений доказанная в следствии 4 сходимость (24) совпадает с требуемой нам сходимостью (34). □

3.5. В этом пункте будут выведены остальные утверждения из п. 2.7.

Доказательство следствия. Из определения (30) и предположения (36) находим, что $K_{ni} \le K_n|\gamma_{ni}(\beta_n)|$. Подставляя эту оценку в (31) и учитывая условие (35), получаем, что $\rho_{n1}^2 \le K_n^2$ и $\rho_{n2} \le K_n$. Значит, предположение (36) следствия 7 влечет условие (31) теоремы 4, а потому из этой теоремы вытекает первое утверждение доказываемого следствия.

Используя опять условие (35) и сравнивая предположения (33) и (37), нетрудно заметить, что $\rho_{n3}^2 \le \rho_{n4}$. Теперь очевидно, что в рассматриваемом случае условие (37) сильнее, чем (33). Таким образом, требуемая сходимость (34) имеет место в силу следствия 6. $\square$

Доказательство теоремы 5. С целью упростить обозначения будем опускать нижний индекс n и положим

$$w_i := w_{ni}(\beta_n) > 0, \quad \gamma_i := \gamma_{ni}(\beta_n), \quad \gamma_i^{opt} := X_i w_i, \quad d_i := (\gamma_i - h\gamma_i^{opt})(\gamma_i - H\gamma_i^{opt}). \tag{63}$$

В этом случае определения (32) и (39) можно переписать в следующем виде:

$$d^2(\gamma_\bullet) = \frac{B^2}{A^2} = \frac{\sigma^2 S}{A^2} \quad \text{и} \quad d^2(\gamma_\bullet^{opt}) = \frac{\sigma^2}{S_0} \quad \text{при } S := \sum \frac{\gamma_i^2}{w_i} \text{ и } S_o := \sum \frac{(\gamma_i^{opt})^2}{w_i}. \tag{64}$$

Напомним, что величина $A = A_n = \sum \gamma_i X_i = \sum \gamma_i \gamma_i^{opt}/w_i$ введена в (27).

Из условий (4) и (35) вытекает, что при каждом i числа γ_i и γ_i^{opt} имеют одинаковый знак. Отсюда следуют неравенства (41). В частности, при каждом i число γ_i лежит между числами $h\gamma_i^{opt}$ и $H\gamma_i^{opt}$. Значит, для введенных в (63) чисел d_i справедливы неравенства $d_i \le 0$. Из этого факта с учетом обозначений из (64) получаем, что

$$S - (H+h)A + hHS_0 = \sum d_i/w_i \le 0.$$

Из полученного неравенства и (64) имеем

$$d^2(\gamma_\bullet) = \sigma^2 \frac{S}{A^2} \le \sigma^2 \frac{(H+h)A - hHS_0}{A^2} = \sigma^2 (H+h) f_1(A), \tag{65}$$

где, таким образом,

$$f_1(A) = (A - CS_0)/A^2 = A^{-1} - CS_0 A^{-2} \quad \text{при } C := Hh/(H+h).$$

Функция $f_1(A)$ достигает своего максимального значения при $A = 2CS_0$. Следовательно, из (65) находим, что

$$d^2(\gamma_\bullet) \le \sigma^2(H+h) f_1(A) \le \sigma^2 (H+h) f_1(2CS_0) = \frac{\sigma^2(H+h)}{4CS_0} = d^2(\gamma_\bullet^{opt}) \frac{(H+h)^2}{4Hh}.$$

Значит,

$$d(\gamma_\bullet)/d(\gamma_\bullet^{opt}) \le \frac{H+h}{2\sqrt{Hh}} = \frac{x+1}{2\sqrt{x}} = 1 + \frac{(\sqrt{x}-1)^2}{2\sqrt{x}}, \quad \text{где } x = H/h \ge 1.$$

Нетрудно убедиться, что эти соотношения эквивалентны (40). $\square$

Следствие 8 является очевидным частным случаем теоремы 5 при $\gamma_{ni}(\cdot) \equiv c_{ni}$.

Таким образом, доказаны все утверждения работы.

ЛИТЕРАТУРА

1. *Боровков А. А.* Математическая статистика. Новосибирск: Изд-во ИМ СО РАН; Наука, 1997.
2. *Себер Дж.* Линейный регрессионный анализ. М.: Мир, 1980.
3. *Kurata H., Kariya T.* Least upper bound for the covariance matrix of a generalized least squares estimator in regression with applications to a seemingly unrelated regression model and a heteroscedastic model // Ann. Statist. 1996. V. 24, N 4. P. 1547–1559.

4. *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.
5. *Линке Ю. Ю., Саханенко А. И.* Асимптотически нормальное оценивание многомерного параметра в задаче дробно-линейной регрессии // Сиб. мат. журн. 2001. Т. 42, № 2. С. 372–388.
6. *Линке Ю. Ю., Саханенко А. И.* Явное асимптотически нормальное оценивание параметров уравнения Михаэлиса — Ментен // Сиб. мат. журн. 2001. Т. 42, № 3. С. 610–633.
7. *Саханенко А. И., Линке Ю. Ю.* Асимптотически оптимальное оценивание в задаче дробно-линейной регрессии со случайными ошибками в коэффициентах // Сиб. мат. журн. 2006. Т. 47, № 6. С. 1372–1400.

Статья поступила 2 ноября 2007 г.

Линке Юлиана Юрьевна
Институт математики им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4, Новосибирск 630090
`linke@math.nsc.ru`

Саханенко Александр Иванович
Югорский гос. университет, ул. Чехова, 16, Ханты-Мансийск 628012
`aisakh@mail.ru`