УДК 62.50;517.972

©2001 г. О.Н.Граничин, канд. физ.–мат. наук
(Санкт-Петербургский государственный университет)

ОЦЕНИВАНИЕ ПАРАМЕТРОВ ЛИНЕЙНОЙ РЕГРЕССИИ ПРИ ПРОИЗВОЛЬНЫХ ПОМЕХАХ

Рассматривается задача оценки параметров линейной регрессии с произвольными помехами, т.е. такими, среднее значение которых может быть неизвестно и отлично от нуля, либо такими, которые представляют собой реализацию коррелированного случайного процесса, либо такими, которые могут задаваться неизвестной, но ограниченной детерминированной функцией.

1. Введение

Обычное предположение о помехах в задачах линейной регрессии заключается в том, что их считают реализацией некоторой последовательности независимых случайных величин с нулевым средним значением. Однако в приложениях это допущение часто нарушается, что может сильно сказываться на работе стандартных оценочных процедур. Поэтому важно исследовать возможность оценки параметров регрессии при минимальных предположениях о статистических свойствах помех. На первый взгляд, это кажется удивительным, но параметры регрессии могут быть эффективно оценены в случае нецентрированных, коррелированных и даже неслучайных помех (см. [1–4]). В этой работе будет показано, как этого можно достичь при определенном условии, когда входы случайны.

Будем рассматривать модель линейной регрессии

$$y_n \;=\; \varphi_n^{\mathrm{T}}\theta^\star \;+\; v_n$$

с вектором параметров $\theta^\star \in \mathbb{R}^r$, который должен быть оценен по наблюдениям $y_n,\ \varphi_n,\ n = 1,2,\ldots$. Допустим, что входы $\{\varphi_n\}_{n\geq 1}$ независимы с помехами $\{v_n\}_{n\geq 1}$. Это предположение будет гарантировать хорошие свойства оценок при чрезвычайно умеренных ограничениях на помехи. Как ни странно, долгое время не замечали того факта, что алгоритмы поиска с последовательным изменением оценки $\{\theta_n\}$ в направлении по оси некоторого случайного центрированного (среднее значение равно нулю) вектора Δ_n, называемого пробным одновременным возмущением,

$$\theta_n = \theta_{n-1} - \Delta_n Y_n,$$

могут сходиться к истинному вектору параметров $\theta^\star$ и в условиях, когда наблюдения Y_n производятся в некоторой точке, определяемой предыдущей оценкой θ_{n-1} и вектором пробного возмущения Δ_n, на фоне почти произвольных помех (см., например, [1–3]). Алгоритмы такого типа будем называть *рандомизированными*, так как обоснование их сходимости при почти произвольных помехах существенно использует вероятностную (стохастическую) природу пробного возмущения. При оценивании неизвестных параметров линейной регрессии роль пробного возмущения могут играть центрированные входы $\Delta_n = \varphi_n - \mathrm{E}\{\varphi_n\}$. (Здесь и далее $\mathrm{E}\{\cdot\}$ — обозначение для среднего значения).

Идея использования случайных входных сигналов для устранения эффекта смещения была выдвинута Фишером [5] в виде рандомизированного принципа планирования эксперимента. Помимо задачи планирования эксперимента, в которой регрессоры могут быть рандомизированны экспериментатором, случайные входы возникают во многих задачах идентификации, фильтрации, распознавания и т.д. (см., например, [6,7]). Имея в виду эти приложения, будем использовать термины *входы, выходы* вместо названий, используемых в регрессионном анализе (подобных термину *регрессор*).

В [1], наверное, впервые был предложен состоятельный алгоритм оценки параметров линейной регрессии при почти произвольных помехах, при этом рассматривалась более общая по сравнению со стандартной постановка задачи. Предполагалось, что вектор неизвестных параметров может меняться со временем, и в алгоритме оценивалось его среднее значение. Такое же обобщение постановки задачи рассматривалось в виде примера в [2]. Использование при решении этой обобщенной задачи методов исследования сходимости при почти произвольных помехах, разработанных в [3], позволяет получить алгоритмы, достигающие оптимальной в средне-квадратичном смысле скорости сходимости.

2. Постановка задачи и основные предположения

Рассмотрим модель линейной регрессии

$$(1) \qquad y_n \;=\; \varphi_n^{\mathrm{T}} \theta_n^\star \;+\; v_n, \quad \theta_n^\star = \theta^\star + w_n, \quad n \;=\; 1, 2, \ldots,$$

с выходами (наблюдениями) $y_n \in \mathbb{R}^1$, входами $\varphi_n \in \mathbb{R}^r$ и помехами $v_n \in \mathbb{R}^1,\ w_n \in \mathbb{R}^r$. требуется оценить значение $\theta^\star$, базируясь на наблюдениях $y_n,\ \varphi_n,\ n \;=\; 1, 2, \ldots.$

Пусть $\mathcal{F}_n$ — σ-алгебра вероятностных событий, порожденная $\{\varphi_1, \ldots, \varphi_n, w_1, \ldots, w_n, v_1, \ldots, v_n\}$, $\hat{\mathcal{F}}_{n-1}$ — σ-алгебра, порожденная $\{\varphi_1, \ldots, \varphi_{n-1}, w_1, \ldots, w_{n-1}, v_1, \ldots, v_n\}$, и $\tilde{\mathcal{F}}_{n-1}$ — σ-алгебра, порожденная $\{\varphi_1, \ldots, \varphi_{n-1}, w_1, \ldots, w_n, v_1, \ldots, v_n\}$, $\mathcal{F}_{n-1} \subset \hat{\mathcal{F}}_{n-1} \subset \tilde{\mathcal{F}}_{n-1} \subset \mathcal{F}_n$.

Сделаем следующие предположения.

(A) Входы $\{\varphi_n\}_{n\geq 1}$ представляют собой последовательность независимых случайных векторов с известными ограниченными математическими ожиданиями $\|\mathrm{E}\{\varphi_n\}\| \leq M_\varphi < \infty$, при каждом n φ_n независимы от $\tilde{\mathcal{F}}_{n-1}$. Случайные величины $\Delta_n = \varphi_n - \mathrm{E}\{\varphi_n\}$ имеют симметричные функции распределения $\mathrm{P}_n(\cdot)$, т.е. $\mathrm{P}_n(\Omega) = \mathrm{P}_n(-\Omega)$ для любого борелевского множества $\Omega \subset \mathbb{R}^r$, и матрицы ковариаций $\mathrm{E}\{\Delta_n \Delta_n^{\mathrm{T}}\} \;=\; \mathrm{B_n} \;>\; 0,\ \|\mathrm{B}_n\| \;\leq\; \sigma_\Delta^2 \;<\; \infty$, ($\mathrm{B} > 0$ означает, что B — положительно-определенная матрица, $\|\cdot\|$ обозначение для евклидовой нормы в $\mathbb{R}^r$).

(B) При каждом n случайные векторы w_n независимы от $\hat{\mathcal{F}}_{n-1}$ и $\mathrm{E}\{w_n\} = 0$. Последовательности помех $\{v_n\}_{n\geq 1}$ и $\{w_n\}_{n\geq 1}$ удовлетворяют одному из условий:

$$(i) \qquad \mathrm{E}\{\, v_n^2 | \mathcal{F}_{n-1}\} \leq \sigma_v^2 < \infty,\ \mathrm{E}\{\|w_n\|^2 |\, \mathcal{F}_{n-1}\} \leq \sigma_w^2 < \infty \text{ a. s.;}$$

$$(ii) \qquad \mathrm{E}\{v_n^2\} \;\leq\; \sigma_v^2 \;<\; \infty, \mathrm{E}\{w_n w_n^{\mathrm{T}}\} \;\leq\; Q_w \;\leq\; \sigma_w^2 \mathrm{I} \;<\; \infty;$$

$$(iii) \qquad |v_n| \;\leq\; C_v \;<\; \infty, \qquad \|w_n\| \;\leq\; C_w \;<\; \infty \text{ a. s.,}$$

где σ_v, σ_w, C_v, C_w — некоторые постоянные и I — матрица с единицами на диагонали и остальными элементами равными нулю.

Заметим, что стандартные предположения в задаче оценивания параметров линейной регрессии со случайными входными сигналами несколько различны (см., например, [8]). Это выражается, в частности, в отсутствии условия $\mathrm{E}\{v_n\} = 0$ и предположения о том, что помехи $\{v_n\}_{n\geq 1}$ представляют собой последовательность одинаково распределенных случайных величин независимых между собой и независимых от $\{\varphi_n\}_{n\geq 1}$.

3. Оценивание по методу стохастической аппроксимации

Исследуем сначала рандомизированный алгоритм типа стохастической аппроксимации (РСА) для рассматриваемой модели наблюдений

$$\theta_n = \theta_{n-1} - \alpha_n \Gamma \Delta_n (\varphi_n^{\mathrm{T}} \theta_{n-1} - y_n), \quad n = 1, 2, \ldots, \tag{2}$$

в котором $\alpha_n \geq 0$ — неслучайная последовательность, определяющая шаг алгоритма, и Γ — некоторая положительно-определенная матрица (см. [6–8]). Предположим, что начальное значение θ_0 — произвольный неслучайный вектор из $\mathbb{R}^r$.

Т е о р е м а 1 . *Пусть для входов модели выполнено допущение* **(A)** *и* $\sum_{n=1}^{\infty} \alpha_n = \infty$, $\alpha_n \to 0$, $\alpha_n \mathrm{E}\{\|\Delta_n\|^4\} \to 0$ *при* $n \to \infty$.

Если для помех выполнено условие **(Bi)** *и* $\sum_{n=1}^{\infty} \alpha_n^2 (1+\mathrm{E}\{\|\Delta_n\|^4\}) < \infty$, **тогда** *последовательность оценок, доставляемых алгоритмом (2), сильносостоятельная, т.е.* $\theta_n \to \theta^\star$ *при* $n \to \infty$ *с вероятностью единица.*

Если для помех выполнено условие **(Bii)**, **тогда** *алгоритм (2) доставляет состоятельные оценки и* $\mathrm{E}\{(\theta_n - \theta^\star)(\theta_n - \theta^\star)^{\mathrm{T}}\} \to 0$ *при* $n \to \infty$.

Доказательство теоремы 1 и двух последующих приведены в приложении.

Следующая теорема устанавливает среднеквадратичную скорость сходимости последовательности оценок, доставляемых алгоритмом (2).

Т е о р е м а 2 . *Если* $\alpha_n = n^{-1}$, *входы модели удовлетворяют условию* **(A)** *и* $n^{-1}\mathrm{E}\{\|\Delta_n\|^4\} \to 0$ *при* $n \to \infty$, *для помех выполнено условие* **(Bii)**, *существуют матрицы* U *и* $\mathrm{B} > 0$ *такие, что* $\mathrm{E}\{\Delta_n \Delta_n^{\mathrm{T}} Q_w \Delta_n \Delta_n^{\mathrm{T}}\} \leq \mathrm{U} + \mathcal{O}(n^{-1})$, $\|\mathrm{U}\| < \infty$ *и* $\|\mathrm{B}_n - \mathrm{B}\| = \mathcal{O}(n^{-1})$, $-\Gamma\,\mathrm{B} + \frac{1}{2}\mathrm{I}$ — *устойчивая матрица, т.е. все её собственные значения лежат в левой полуплоскости,* **тогда**

$$\mathrm{E}\,(\theta_n - \theta^\star)(\theta_n - \theta^\star)^{\mathrm{T}} \leq n^{-1}\,\mathrm{S} + o(n^{-1}),$$

где матрица S *является решением матричного уравнения*

$$\Gamma\,\mathrm{B}\,\mathrm{S} \;+\; \mathrm{S}\,\mathrm{B}\,\Gamma \;-\; \mathrm{S} \;=\; \Gamma\mathrm{R}\Gamma,$$

в котором $\mathrm{R} \;=\; (\sigma_v^2(1+M_\varphi^2\rho)+M_\varphi^2\sigma_w^2)\mathrm{B} \;+\; \mathrm{U}$ *с любым* $\rho>0$.

Если $\Gamma \;=\; \mathrm{B}^{-1}$, тогда последнее уравнение для матрицы S может быть легко решено

$$\mathrm{S} \;=\; \mathrm{B}^{-1}\mathrm{R}\mathrm{B}^{-1}.$$

При этом для алгоритма (2), принимающего вид

$$\theta_n \;=\; \theta_{n-1} \;-\; (n\mathrm{B})^{-1}\Delta_n(\varphi_n^{\mathrm{T}}\theta_{n-1}-y_n), \quad n=1,2,\ldots,$$

при $\sigma_w=0$ имеем

$$\mathrm{E}\,(\theta_n-\theta^\star)(\theta_n-\theta^\star)^{\mathrm{T}} \le n^{-1}\sigma_v^2(1+M_\varphi^2\rho)\mathrm{B}^{-1} \;+\; o(n^{-1}).$$

Для последнего алгоритма получили почти ту же самую скорость сходимости, что и наилучшая в случае, когда помехи v_n являются независимыми случайными величинами с нулевыми средними, см. [8]. Кроме того, как показано в [8], этот выбор α_n и Γ является оптимальным для алгоритма (2) при $M_\varphi=0$.

З а м е ч а н и е 1. Утверждения теорем 1 и 2 также выполняются, если предположить равенство нулю третьего центрального момента φ_n вместо симметричности распределения Δ_n.

З а м е ч а н и е 2. В случае равенства в предположении (Bii) в утверждении теоремы 2 неравенство в оценках скорости сходимости также можно заменить на равенство.

З а м е ч а н и е 3. В [3] предлагалось рассматривать несколько отличающийся от (2) алгоритм стохастической аппроксимации. Сравнение соответствующих результатов показывает, что в случае $M_\varphi\neq 0$ алгоритм (2) дает лучшие оценки, чем получаемые в [3].

4. Оценки по методу наименьших квадратов

Теперь рассмотрим для той же регрессионной модели наблюдений (1) оценки по рандомизированному методу наименьших квадратов (РМНК) вида

(3*a*) $$\theta_n = \theta_{n-1} - \Gamma_n\Delta_n(\varphi_n^{\mathrm{T}}\theta_{n-1}-y_n),$$

$$\Gamma_n = \Gamma_{n-1} - \frac{\Gamma_{n-1}\Delta_n\Delta_n^{\mathrm{T}}\Gamma_{n-1}}{1+\Delta_n^{\mathrm{T}}\Gamma_{n-1}\Delta_n},\ \Gamma_0 = \rho^{-1}\mathrm{I}, \tag{3b}$$

где $\rho > 0$ является малым положительным числом (параметром регуляризации, см. [7]). Как и выше, примем в качестве начального значения θ_0 произвольный случайный вектор из $\mathbb{R}^r$.

Т е о р е м а 3 . *Пусть выполнены предположения о входах* **(A)** *и они одинаково распределены при всех* n*,* $\mathrm{B}_n = \mathrm{B} > 0$*,* $\mathrm{E}\{\|\Delta_n\|^4\} < \infty$.

Если выполнено условие **(Bi)** *для помех,* **тогда** *для алгоритма (3)* $\theta_n \to \theta^\star$ *при* $n \to \infty$ *с вероятностью единица.*

Если выполнено условие **(Biii)** *для помех и пробное возмущение равномерно ограничено с вероятностью единица:* $\|\Delta_n\| \leq C_\Delta < \infty,\ n = 1, 2, \ldots,$ **тогда** *для алгоритма (3) при* $n \to \infty$ $\mathrm{E}\{(\theta_n - \theta^\star)(\theta_n - \theta^\star)^{\mathrm{T}}\} \to 0$.

5. Пример

Рассмотрим задачу обнаружения (детектирования) скалярного сигнала $\{\varphi_n\}$, который может быть попадает, а может быть и нет в зашумленный канал наблюдения (измерения производятся с помехами). В задачах обнаружения сигнала оцениваемая величина $\theta^\star$ обычно принимает конечное число значений и часто представляет из себя характеристику типа "да— нет". Будем считать, что событие $\{\theta^\star = 1\}$ соответствует наличию сигнала в приемнике, а $\{\theta^\star = 0\}$ — его отсутствию. С учетом вышесказанного, наблюдаемые величины $\{y_n\}$ можно представить в виде (1) с $w_n = 0$. Предположим, что помехи задаются неизвестной ограниченной детерминированной функцией: $|v_n| \leq C_v = 2,\ n = 1, 2, \ldots$.

Пусть полезный сигнал имеет статистическую природу, представляя из себя последовательность независимых между собой наблюдаемых случайных величин, одинаково равномерно распределенных на интервале $[0,5; 1,5]$. Выполнены все условия теорем 1–3, и для построения последовательности состоятельных оценок $\{\theta_n\}$ величины $\theta^\star$ можно воспользоваться любым из алгоритмов РСА (2) или РМНК (3), выбрав в качестве начального приближения $\theta_0 = 0$. В качестве "решающего" правила в момент времени n при выборе гипотезы о наличии полезного сигнала в канале наблюдения или об его отсутствии обычно выбирают операцию сравнения величины текущей оценки с задаваемым пороговым значением $\delta > 0$ (например $\delta = 0,5$). Если $\theta_n < \delta$, то принимается гипотеза "сигнала нет", в противном случае — "сигнал есть".

Для иллюстрации работоспособности предложенных выше рандомизированных алгоритмов оценивания и сравнения при "плохих" помехах с оценками обыкновенного регуляризованного метода наименьших квадратов (МНК) была проведена серия экспериментов на ЭВМ. Типичный результат сравнительного моделирования работы алгоритмов (2) с $\alpha_n \Gamma = 1/(\sigma_\varphi^2 n)$, (3) с $\rho = 0,99$ и МНК

$$\theta_n = \frac{\sum_{k=1}^n \varphi_k y_k}{\sum_{k=1}^n \varphi_k^2 + (0,99)^{-1}}. \tag{4}$$

приведен на рис. 1. В этом эксперименте значения помехи v_n из интервала $[-2; 2]$ вычислялись по детерминированной функции и имели среднее значение больше единицы. Анализ поведения типичных траекторий последовательного изменения оценок для трех алгоритмов показывает, что уровень помехи настолько высок, что оценки обыкновенного МНК почти всегда превышают уровень принятия решения вне зависимости от наличия или отсутствия сигнала, в то время как после 50 итераций рандомизированные алгоритмы дают правильные ответы.

6. Заключение

При случайных входных сигналах предложенные в работе алгоритмы для обоснования сходимости требуют выполнения очень умеренных условий на помехи. В частности, помеха может задаваться неизвестной, но ограниченной детерминированной функцией. По этой обнадеживающей причине эти алгоритмы могут быть полезны во многих приложениях. Численное моделирование продемонстрировало эффективность алгоритмов при разнообразных помехах v_n. В частности, в скалярном случае эксперименты были выполнены с неслучайной константой, нецентрированной случайной переменной и различными неслучайными последовательностями помех. В этих экспериментах исследовались траектории оценок стохастической аппроксимации и оценок рандомизированного метода наименьших квадратов. Поведение типичных траекторий рандомизированных алгоритмов при высоком уровне нерегулярной помехи существенно лучше поведения траекторий обычных алгоритмов.

ПРИЛОЖЕНИЕ

Доказательство теоремы 1. Обозначив $\eta_n = \theta_{n-1} - \theta_n^\star$, $\xi_n = v_n - \mathrm{E}\{\varphi_n\}^{\mathrm{T}}\eta_n$, из уравнений для модели наблюдений и алгоритма (2) получаем

$$\theta_n - \theta^\star = \theta_{n-1} - \theta^\star - \alpha_n \Gamma \Delta_n \Delta_n^{\mathrm{T}} \eta_n + \alpha_n \Gamma \Delta_n \xi_n,$$

и, следовательно,

$$\|\theta_n - \theta^\star\|^2 = ((\theta_{n-1} - \theta^\star)^{\mathrm{T}} - \eta_n^{\mathrm{T}} \alpha_n \Delta_n \Delta_n^{\mathrm{T}} \Gamma)((\theta_{n-1} - \theta^\star) - \alpha_n \Gamma \Delta_n \Delta_n^{\mathrm{T}} \eta_n) +$$
$$+ \alpha_n^2 \xi_n^2 \Delta_n^{\mathrm{T}} \Gamma^2 \Delta_n + ((\theta_{n-1} - \theta^\star)^{\mathrm{T}} - \eta_n^{\mathrm{T}} \alpha_n \Delta_n \Delta_n^{\mathrm{T}} \Gamma) \alpha_n \xi_n \Gamma \Delta_n +$$
$$+ \alpha_n \xi_n \Delta_n^{\mathrm{T}} \Gamma ((\theta_{n-1} - \theta^\star) - \alpha_n \Gamma \Delta_n \Delta_n^{\mathrm{T}} \eta_n).$$

Рассмотрим условные по отношению к σ-алгебре $\tilde{\mathcal{F}}_{n-1}$ математические ожидания от обеих частей последней формулы. Предварительно заметим, что в силу предположения **(A)** выполняются соотношения

$$\mathrm{E}\{\eta_n^{\mathrm{T}} \alpha_n \xi_n \Gamma \Delta_n | \tilde{\mathcal{F}}_{n-1}\} = \eta_{n-1}^{\mathrm{T}} \alpha_n v_n \Gamma \mathrm{E}\{\Delta_n | \tilde{\mathcal{F}}_{n-1}\} = \eta_{n-1}^{\mathrm{T}} \alpha_n \xi_n \Gamma \mathrm{E}\{\Delta_n\} = 0$$

и $\mathrm{E}\{(\theta_{n-1} - \theta^\star)^{\mathrm{T}} \alpha_n \xi_n \Gamma \Delta_n \,|\, \tilde{\mathcal{F}}_{n-1}\} = 0$, а также из-за симметричности распределения $\mathrm{P}_n(\cdot)$ $\mathrm{E}\{\eta_n^{\mathrm{T}} \alpha_n^2 \Delta_n \Delta_n^{\mathrm{T}} \Gamma^2 \Delta_n \xi_n \,|\, \tilde{\mathcal{F}}_{n-1}\} = 0$. С учетом последних соотношений имеем

$$\mathrm{E}\{\|\theta_n - \theta^\star\|^2 \,|\, \hat{\mathcal{F}}_{n-1}\} = \|\theta_{n-1} - \theta^\star\|^2 - \eta_n^{\mathrm{T}} \alpha_n \mathrm{E}\{\Delta_n \Delta_n^{\mathrm{T}}\} \Gamma (\theta_{n-1} - \theta^\star) -$$
$$- (\theta_{n-1} - \theta^\star)^{\mathrm{T}} \alpha_n \Gamma \mathrm{E}\{\Delta_n \Delta_n^{\mathrm{T}}\} \eta_n + \alpha_n^2 \eta_n^{\mathrm{T}} \mathrm{E}\{\Delta_n \Delta_n^{\mathrm{T}} \Gamma^2 \Delta_n \Delta_n^{\mathrm{T}}\} \eta_n +$$
$$+ \alpha_n^2 \mathrm{E}\{\xi_n^2 | \tilde{\mathcal{F}}_{n-1}\} \mathrm{E}\{\Delta_n^{\mathrm{T}} \Gamma^2 \Delta_n\} \le \|\theta_{n-1} - \theta^\star\|^2 + \alpha_n^2 \|\eta_n\|^2 \mathrm{E}\{\mathrm{Tr}[\Delta_n \Delta_n^{\mathrm{T}} \Gamma^2 \Delta_n \Delta_n^{\mathrm{T}}]\} -$$
$$- \eta_n^{\mathrm{T}} \alpha_n \mathrm{B_n} \Gamma (\theta_{n-1} - \theta^\star) - (\theta_{n-1} - \theta^\star)^{\mathrm{T}} \alpha_n \Gamma \mathrm{B_n} \eta_n + \alpha_n^2 \xi_n^2 \mathrm{Tr}[\Gamma\, \mathrm{B_n} \Gamma].$$

Усредняя по w_n, получаем

$$\mathrm{E}\{\|\theta_n - \theta^\star\|^2 \,|\, \hat{\mathcal{F}}_{n-1}\} \le \|\theta_{n-1} - \theta^\star\|^2 (\mathrm{I} + \alpha_n^2 (\|\Gamma\|^2 \mathrm{E}\{\|\Delta_n\|^4\} + 2 M_\varphi^2 \mathrm{Tr}[\Gamma \mathrm{B_n} \Gamma])) -$$
$$- \alpha_n (\theta_{n-1} - \theta^\star)^{\mathrm{T}} (\mathrm{B_n} \Gamma + \Gamma \mathrm{B_n})(\theta_{n-1} - \theta^\star) + \alpha_n^2 r \sigma_w^2 \|\Gamma\|^2 \mathrm{E}\{\|\Delta_n\|^4\} +$$
$$+ \alpha_n^2 (2 v_n^2 + r M_\varphi^2 \sigma_w^2) \mathrm{Tr}[\Gamma \mathrm{B_n} \Gamma].$$

Далее, усредняя по v_n, имеем

$$\mathrm{E}\{\|\theta_n-\theta^\star\|^2 \,|\, \mathcal{F}_{n-1}\} \le \|\theta_{n-1}-\theta^\star\|^2(\mathrm{I}+\alpha_n^2(\|\Gamma\|^2\mathrm{E}\{\|\Delta_n\|^4\}+2M_\varphi^2\mathrm{Tr}[\Gamma \mathrm{B_n}\Gamma])) -$$
$$-\alpha_n(\theta_{n-1}-\theta^\star)^{\mathrm{T}}(\mathrm{B_n}\Gamma+\Gamma\mathrm{B_n})(\theta_{n-1}-\theta^\star)+$$
$$+\alpha_n^2(r\sigma_w^2\|\Gamma\|^2\mathrm{E}\{\|\Delta_n\|^4\}+2\sigma_v^2+rM_\varphi^2\sigma_w^2)\mathrm{Tr}[\Gamma\mathrm{B_n}\Gamma]).$$

Так как $\sum_{n=1}^\infty \alpha_n^2\mathrm{E}\{\|\Delta_n\|^4\} < \infty$ и $\sum_{n=1}^\infty \alpha_n^2\mathrm{Tr}[\Gamma\,\mathrm{B_n}\,\Gamma]) < \infty$, то, применяя лемму Роббинса-Сигмунда [9], получаем существование конечного предела $\lim_{n\to\infty}\|\theta_n-\theta^\star\|^2$ и, более того, сходимость ряда

$$\sum_{n=1}^{\infty}(\theta_{n-1}-\theta^\star)^{\mathrm{T}}\,\alpha_n\,(\mathrm{B_n}\,\Gamma\,+\,\Gamma\mathrm{B_n})\,(\theta_{n-1}-\theta^\star)<\infty\,.$$

Из того, что $\sum\alpha_n=\infty$ и $\|\mathrm{B_n}\|\le\sigma_\Delta^2<\infty$, следует $\|\theta_n-\theta^\star\|^2\to 0$ с вероятностью единица. Первое утверждение теоремы доказано.

Докажем второе утверждение теоремы. Обозначим $D_n=(\theta_n-\theta^\star)(\theta_n-\theta^\star)^{\mathrm{T}}$. Как и выше, из уравнений для модели наблюдений и алгоритма (2) можем получить

$$D_n\,=\,D_{n-1}\,-\,\alpha_n\Gamma\Delta_n\Delta_n^{\mathrm{T}}\eta_n\,(\theta_{n-1}-\theta^\star)^{\mathrm{T}}\,-\,\alpha_n(\theta_{n-1}-\theta^\star)\eta_n^{\mathrm{T}}\Delta_n\Delta_n^{\mathrm{T}}\Gamma+$$
$$+\,\alpha_n^2\Gamma\Delta_n\Delta_n^{\mathrm{T}}\eta_n\eta_n^{\mathrm{T}}\Delta_n\Delta_n^{\mathrm{T}}\Gamma\,+\,\alpha_n^2\xi_n^2\Gamma\Delta_n\Delta_n^{\mathrm{T}}\Gamma+$$
$$+\,(\theta_{n-1}-\theta^\star-\alpha_n\Gamma\Delta_n\Delta_n^{\mathrm{T}}\eta_n)\alpha_n\xi_n\Delta_n^{\mathrm{T}}\Gamma+\alpha_n\xi_n\Gamma\Delta_n(\theta_{n-1}-\theta^\star-\alpha_n\Gamma\Delta_n\Delta_n^{\mathrm{T}}\eta_n)^{\mathrm{T}}.$$

Далее, усреднив обе части последнего соотношения по σ-алгебре $\tilde{\mathcal{F}}_{n-1}$, в силу допущения (**A**) заключаем, что

$$\mathrm{E}\{D_n|\tilde{\mathcal{F}}_{n-1}\}\,=\,D_{n-1}-\alpha_n\,\Gamma\,\mathrm{B}_n\eta_{n-1}\,(\theta_{n-1}-\theta^\star)^{\mathrm{T}}-\alpha_n\,(\theta_{n-1}-\theta^\star)\eta_{n-1}^{\mathrm{T}}\,\mathrm{B}_n\Gamma+$$
$$+\,\alpha_n^2\Gamma\mathrm{E}\{\Delta_n\Delta_n^{\mathrm{T}}\eta_n\eta_n^{\mathrm{T}}\Delta_n\Delta_n^{\mathrm{T}}|\tilde{\mathcal{F}}_{n-1}\}\Gamma\,+\,\alpha_n^2\xi_n^2\Gamma\mathrm{B}_n\Gamma\,\le$$
$$\le\,D_{n-1}\,-\,\alpha_n\,\Gamma\,\mathrm{B}_n\eta_{n-1}\,(\theta_{n-1}-\theta^\star)^{\mathrm{T}}\,-\,\alpha_n\,(\theta_{n-1}-\theta^\star)\eta_{n-1}^{\mathrm{T}}\,\mathrm{B}_n\Gamma+$$
$$+\alpha_n^2\Gamma\mathrm{E}\{\Delta_n\Delta_n^{\mathrm{T}}(D_{n-1}+Q_w)\Delta_n\Delta_n^{\mathrm{T}}\}\Gamma+$$
$$+\,\alpha_n^2(v_n^2(1+M_\varphi^2\rho)+\|D_n\|\rho^{-1}+M_\varphi^2\|w_n\|^2)\Gamma\,\mathrm{B}_n\Gamma.$$

Теперь произведем усреднение по σ-алгебре $\hat{\mathcal{F}}_{n-1}$, используя вторую часть (Bii),

$$\mathrm{E}\{D_n|\hat{\mathcal{F}}_{n-1}\}\,\le\,D_{n-1}-\alpha_n(\Gamma\,\mathrm{B}_nD_{n-1}+D_{n-1}^{\mathrm{T}}\,\mathrm{B}_n\Gamma)+\alpha_n^2(\|D_n\|\mathrm{E}\{\|\Delta_n\|^4\}\Gamma^2+$$

$$+\Gamma((v_n^2(1+M_\varphi^2\rho)+\|D_n\|\rho^{-1}+rM_\varphi^2\sigma_w^2)\mathrm{B}_n+\mathrm{E}\{\Delta_n\Delta_n^{\mathrm{T}}Q_w\Delta_n\Delta_n^{\mathrm{T}}\})\Gamma).$$

В заключение, взяв безусловное математическое ожидание и использовав первую часть допущения (Bii), получаем для матриц $\mathbf{V}_n := \mathrm{E}\{D_n\}$

$$\mathbf{V}_n \le \mathbf{V}_{n-1} - \alpha_n(\Gamma\mathrm{B}_n\mathbf{V}_{n-1} + \mathbf{V}_{n-1}\mathrm{B}_n\Gamma) +$$
$$+\alpha_n^2\Gamma((\sigma_v^2(1+M_\varphi^2\rho)+M_\varphi^2\sigma_w^2)\mathrm{B}_n+\mathrm{U})\Gamma+\alpha_n\beta_n\,\mathcal{O}(\|\,\mathbf{V}_{n-1}\,\|)$$

с некоторой числовой последовательностью $\{\beta_n\}$, $\beta_n \to 0$ при $n \to \infty$. Применяя к последнему неравенству лемму 3 из [10], имеем $\mathbf{V}_n \to 0$ при $n \to \infty$.

Доказательство теоремы 2. Для начала покажем, что матричное уравнение из утверждения теоремы имеет решение. Перепишем его в виде

$$(\Gamma\,\mathrm{B} - \frac{1}{2}\mathrm{I})\mathrm{S} + \mathrm{S}(\mathrm{B}\,\Gamma - \frac{1}{2}\mathrm{I}) = \Gamma\mathrm{R}\Gamma.$$

Так как $-\Gamma + \frac{1}{2}\mathrm{I}$ является гурвицевой, то по лемме Ляпунова существует положительно определенная матрица S, которая является решением соответствующего матричного уравнения.

Вернемся к последнему неравенству из доказательства предыдущей теоремы. В силу $\|\mathrm{B}_n - \mathrm{B}\| = \mathcal{O}(n^{-1})$ получаем

$$\mathbf{V}_n \le \mathbf{V}_{n-1} - n^{-1}(\Gamma\mathrm{B}\mathbf{V}_{n-1} + \mathbf{V}_{n-1}\mathrm{B}\Gamma) + n^{-2}\Gamma\mathrm{R}\Gamma + n^{-1}\beta_n\mathcal{O}(\|\,\mathbf{V}_{n-1}\,\|).$$

Пусть $\mathbf{W}_n = n\,\mathbf{V}_n - \mathrm{S}$. Тогда в условиях теоремы имеем

$$\mathbf{W}_n \le \mathbf{W}_{n-1} - (n-1)^{-1}((\Gamma\mathrm{B} - \frac{1}{2}\mathrm{I})\mathbf{W}_{n-1} + \mathbf{W}_{n-1}(\mathrm{B}\,\Gamma - \frac{1}{2}\mathrm{I})) + n^{-1}\beta_n\mathcal{O}(\|\mathbf{W}_{n-1}\|).$$

Следовательно, применяя опять лемму 3 из [10], получаем $\mathbf{W}_n \to 0$ при $n \to \infty$ и, таким образом, теорема 2 доказана.

Доказательство теоремы 3. Обозначим $\eta_n = \theta_n - \theta_n^\star$, $\xi_n = v_n - \mathrm{E}\{\varphi_n\}^{\mathrm{T}}\eta_n$. Следующие вспомогательные результаты будут использоваться в доказательстве теоремы 3.

При выполнении условий теоремы 3

(a) $\mathrm{E}\{\eta_n^{\mathrm{T}}\Gamma_n\,\xi_n\,\Delta_n\,|\tilde{\mathcal{F}}_{n-1}\} = 0$ и $\mathrm{E}\{\eta_n^{\mathrm{T}}\Delta_n\,\Delta_n^{\mathrm{T}}\,\Gamma_n^2\,\Delta_n\,\xi_n|\tilde{\mathcal{F}}_{n-1}\} = 0$,

(b) $\sum\limits_{n=1}^{\infty}\Delta_n^{\mathrm{T}}\,\Gamma_n^2\,\Delta_n < \infty$ с вероятностью единица,

(c) $\sum\limits_{n=1}^{\infty} \Delta_n^{\mathrm{T}} \Gamma_n \Delta_n = \infty$ с вероятностью единица,

(d) $\sum_{n=1}^{\infty} \|\Delta_n\|^4 \lambda_{max}^2(\Gamma_n) < \infty$ с вероятностью единица.

Доказательство (a). Из (3b) имеем

$$\Gamma_n \Delta_n = \Gamma_{n-1} \Delta_n - \frac{\Gamma_{n-1} \Delta_n \Delta_n^{\mathrm{T}} \Gamma_{n-1} \Delta_n}{1 + \Delta_n^{\mathrm{T}} \Gamma_{n-1} \Delta_n} = \frac{\Gamma_{n-1} \Delta_n}{1 + \Delta_n^{\mathrm{T}} \Gamma_{n-1} \Delta_n},$$

$$\mathrm{E}\{\eta_n^{\mathrm{T}} \Gamma_n \xi_n \Delta_n | \tilde{\mathcal{F}}_{n-1}\} = \eta_n^{\mathrm{T}} \xi_n \mathrm{E}\{\frac{\Gamma_{n-1} \Delta_n}{1 + \Delta_n^{\mathrm{T}} \Gamma_{n-1} \Delta_n} \Big| \tilde{\mathcal{F}}_{n-1}\}$$

$$= \eta_n^{\mathrm{T}} \xi_n \Gamma_{n-1} \int \frac{x}{1 + x^{\mathrm{T}} \Gamma_{n-1} x} \mathrm{P_n}(dx) = 0,$$

так как интеграл от нечетной функции аргумента x по симметричному распределению $\mathrm{P_n}(\cdot)$ равен нулю. Далее,

$$\Delta_n \Delta_n^{\mathrm{T}} \Gamma_n^2 \Delta_n = \Delta_n \frac{\Delta_n^{\mathrm{T}} \Gamma_{n-1}^2 \Delta_n}{(1 + \Delta_n^{\mathrm{T}} \Gamma_{n-1} \Delta_n)^2}$$

и $\mathrm{E}\{\Delta_n \Delta_n^{\mathrm{T}} \Gamma_n^2 \Delta_n \xi_n | \tilde{\mathcal{F}}_{n-1}\} = 0$ по тем же причинам.

Доказательство (b). Из (3b) также имеем

$$\mathrm{Tr}[\Gamma_n] = \mathrm{Tr}[\Gamma_{n-1}] - \frac{\mathrm{Tr}[\Gamma_{n-1} \Delta_n \Delta_n^{\mathrm{T}} \Gamma_{n-1}]}{1 + \Delta_n^{\mathrm{T}} \Gamma_{n-1} \Delta_n} = \mathrm{Tr}[\Gamma_{n-1}] - \frac{\Delta_n^{\mathrm{T}} \Gamma_{n-1}^2 \Delta_n}{1 + \Delta_n^{\mathrm{T}} \Gamma_{n-1} \Delta_n}.$$

Так как $\mathrm{Tr}[\Gamma_n] \leq \mathrm{Tr}[\Gamma_{n-1}] \leq \ldots \leq \mathrm{Tr}[\Gamma_0]$, то в силу того, что Γ_n — положительно-определенные матрицы для всех n, существует предел $\lim_{n\to\infty} \mathrm{Tr}[\Gamma_n]$. Далее получаем

$$\mathrm{Tr}[\Gamma_n] = \mathrm{Tr}[\Gamma_0] - \sum_{k=1}^{n} \frac{\Delta_k^{\mathrm{T}} \Gamma_{k-1}^2 \Delta_k}{1 + \Delta_k^{\mathrm{T}} \Gamma_{k-1} \Delta_k} \leq \mathrm{Tr}[\Gamma_0] - \sum_{k=1}^{n} \frac{\Delta_k^{\mathrm{T}} \Gamma_{k-1}^2 \Delta_k}{(1 + \Delta_k^{\mathrm{T}} \Gamma_{k-1} \Delta_k)^2}.$$

Устремляя $n \to \infty$, в силу того, что

$$\Delta_k^{\mathrm{T}} \Gamma_{k-1}^2 \Delta_k = \frac{\Delta_k^{\mathrm{T}} \Gamma_{k-1}^2 \Delta_k}{(1 + \Delta_k^{\mathrm{T}} \Gamma_{k-1} \Delta_k)^2},$$

получаем доказательство (b).

Доказательство (с). По свойству собственных значений матрицы получаем

$$\sum_{n=1}^{\infty} \Delta_n^{\mathrm{T}}\, \Gamma_n\, \Delta_n \;\geq \sum_{n=1}^{\infty}\; \|\Delta_n\|^2\, \lambda_{min}(\Gamma_n) \;=\; \sum_{n=1}^{\infty}\; \|\Delta_n\|^2\,/\lambda_{max}(\Gamma_n^{-1}) \;=$$

$$= \sum_{n=1}^{\infty} \frac{\|\Delta_n\|^2\, \mathrm{Tr}[\Gamma_n^{-1}]}{(\rho + \sum_{k=1}^n \|\Delta_k\|^2)\, \lambda_{max}(\Gamma_n^{-1})}.$$

Но $\sum_{n=1}^{\infty} \|\Delta_n\|^2/(\rho + \sum_{k=1}^n \|\Delta_k\|^2) = \infty$ с вероятностью единица в силу леммы Абеля–Дини (см. [11]) и $\mathrm{Tr}[\Gamma_n^{-1}]/\lambda_{max}(\Gamma_n^{-1}) = \mathcal{O}(1)$. Следовательно, утверждение (c) верно.

Доказательство (d). Нетрудно убедиться, что $\sum_{n=1}^{\infty} \|\Delta_n\|^4 \lambda_{max}^2(\Gamma_n) = \sum_{n=1}^{\infty} \|\Delta_n\|^4 \lambda_{min}^{-2}(\Gamma_n^{-1})$. В соответствии с усиленным законом больших чисел Колмогорова, $n^{-1}\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;\to\; \mathrm{B}$ с вероятностью единица при $n \to \infty$. Следовательно, $(n\mathrm{B})^{-1}\Gamma_n^{-1} \;\to\; \mathrm{I}$ с вероятностью единица и $n^{-1}\lambda_{min}(\Gamma_n^{-1}) \;\to$ г $>\; 0$ с вероятностью единица. Сходимость с вероятностью единица $\sum_{n=1}^{\infty}(\Delta_n^{\mathrm{T}}\Gamma_n\Delta_n)^2$ будет обеспечена, если с вероятностью единица сходится $\sum_{n=1}^{\infty} n^{-2}\|\Delta_n\|^4$. Для доказательства сходимости этого ряда используем суммирование частями

$$\sum_{k=1}^{n} k^{-2}\|\Delta_k\|^4 \;=\; n^{-1}(n^{-1}\sum_{k=1}^{n}\|\Delta_k\|^4) + \sum_{k=1}^{n-1}(k^{-2} - (k+1)^{-2})\sum_{t=1}^{k}\|\Delta_t\|^4 \;\leq$$

$$\leq\; n^{-1}(n^{-1}\sum_{k=1}^{n}\|\Delta_k\|^4) + \sum_{k=1}^{n-1} 2k^{-2}\frac{1}{k}\sum_{t=1}^{k}\|\Delta_t\|^4.$$

Но $\epsilon_k := k^{-1}\sum_{t=1}^k \|\Delta_t\|^4$ сходится с вероятностью единица при $k \to \infty$, так как $\mathrm{E}\{\|\Delta_n\|^4\} < \infty$, следовательно, $\sum_{k=1}^{\infty} 2k^{-2}\epsilon_k < \infty$ с вероятностью единица. Более того, $n^{-2}\sum_{k=1}^n \|\Delta_k\|^4 \;\to\; 0$ при $n \;\to\; \infty$. Сходимость $\sum_{n=1}^{\infty} n^{-2}\|\Delta_n\|^4$ доказана, а значит, закончено доказательство (d).

Вернемся к доказательству теоремы 3. В силу алгоритма оценивания (3) имеем $\theta_n - \theta^\star = \theta_{n-1} - \theta^\star - \Gamma_n\Delta_n\Delta_n^{\mathrm{T}}\eta_n + \Gamma_n\Delta_n\xi_n$ и, следовательно,

$$\|\theta_n - \theta^\star\|^2 \;=\; ((\theta_{n-1} - \theta^\star)^{\mathrm{T}} - \eta_n^{\mathrm{T}}\, \Delta_n\, \Delta_n^{\mathrm{T}}\, \Gamma_n)((\theta_{n-1} - \theta^\star) - \Gamma_n\, \Delta_n\, \Delta_n^{\mathrm{T}}\eta_n) +$$

$$+\, \xi_n^2\, \Delta_n^{\mathrm{T}}\, \Gamma_n^2\, \Delta_n \;+\; ((\theta_{n-1} \;-\; \theta^\star)^{\mathrm{T}} \;-\; \eta_n^{\mathrm{T}}\Delta_n\, \Delta_n^{\mathrm{T}}\, \Gamma_n)\xi_n\, \Gamma_n\, \Delta_n +$$

$$+\,\xi_n\,\Delta_n^{\mathrm{T}}\;\Gamma_n\,((\theta_{n-1}\,-\,\theta^\star)\,-\,\Gamma_n\,\Delta_n\Delta_n^{\mathrm{T}}\eta_n).$$

Учитывая полученные выше соотношения (a), рассмотрим условные по отношению к σ-алгебре $\tilde{\mathcal{F}}_{n-1}$ математические ожидания от обеих частей последнего соотношения

$$\mathrm{E}\{\|\theta_n-\theta^\star\|^2\,|\,\tilde{\mathcal{F}}_{n-1}\}\;=\;\|\theta_{n-1}\,-\,\theta^\star\|^2\,-\eta_n^{\mathrm{T}}\mathrm{E}\{\Delta_n\Delta_n^{\mathrm{T}}\Gamma_n\,|\,\tilde{\mathcal{F}}_{n-1}\}(\theta_{n-1}-\theta^\star)-$$
$$-\,(\theta_{n-1}-\theta^\star)^{\mathrm{T}}\mathrm{E}\{\Gamma_n\Delta_n\Delta_n^{\mathrm{T}}\,|\,\tilde{\mathcal{F}}_{n-1}\}\eta_n\,+\,\eta_n^{\mathrm{T}}\mathrm{E}\{\Delta_n\Delta_n^{\mathrm{T}}\Gamma_n^2\Delta_n\Delta_n^{\mathrm{T}}\,|\,\tilde{\mathcal{F}}_{n-1}\}\eta_n+$$
$$+\,\mathrm{E}\{(v_n-\mathrm{E}\{\varphi_n\}^{\mathrm{T}}\eta_n)^2\,|\tilde{\mathcal{F}}_{n-1}\}\mathrm{E}\{\Delta_n^{\mathrm{T}}\Gamma_n^2\Delta_n\,|\,\tilde{\mathcal{F}}_{n-1}\}.$$

Усреднив по отношению к σ-алгебре $\hat{\mathcal{F}}_{n-1}$, имеем

$$\mathrm{E}\{\|\theta_n-\theta^\star\|^2|\hat{\mathcal{F}}_{n-1}\}\;\le\;\|\theta_{n-1}-\theta^\star\|^2\,+\,\mathrm{E}\{\|\Delta_n\|^4\,\lambda_{max}^2(\Gamma_n)\|\eta_n\|^2|\hat{\mathcal{F}}_{n-1}\}+$$
$$+\,(2v_n^2+2M_\varphi^2\|\theta_{n-1}-\theta^\star\|^2+rM_\varphi^2\sigma_w^2)\mathrm{E}\{\Delta_n^{\mathrm{T}}\Gamma_n^2\Delta_n|\hat{\mathcal{F}}_{n-1}\}-$$
$$-\,(\theta_{n-1}-\theta^\star)^{\mathrm{T}}\mathrm{E}\{(\Gamma_n\Delta_n\Delta_n^{\mathrm{T}}\,+\,\Delta_n\Delta_n^{\mathrm{T}}\Gamma_n)|\hat{\mathcal{F}}_{n-1}\}(\theta_{n-1}-\theta^\star).$$

Далее, усредняя по v_n, на основании допущения (**Bi**) заключаем

$$\mathrm{E}\{\|\theta_n-\theta^\star\|^2|\mathcal{F}_{n-1}\}\;\le\;\|\theta_{n-1}-\theta^\star\|^2(1+\mathrm{E}\{2M_\varphi^2\Delta_n^{\mathrm{T}}\Gamma_n^2\Delta_n+\|\Delta_n\|^4\,\lambda_{max}^2(\Gamma_n)|\mathcal{F}_{n-1}\})+$$
$$+\,r\sigma_w^2\mathrm{E}\{\|\Delta_n\|^4\,\lambda_{max}^2(\Gamma_n)|\mathcal{F}_{n-1}\}\,+\,(2\sigma_v^2+rM_\varphi^2\sigma_w^2)\mathrm{E}\{\Delta_n^{\mathrm{T}}\Gamma_n^2\Delta_n|\mathcal{F}_{n-1}\}-$$
$$-\,(\theta_{n-1}-\theta^\star)^{\mathrm{T}}\mathrm{E}\{(\Gamma_n\Delta_n\Delta_n^{\mathrm{T}}\,+\,\Delta_n\Delta_n^{\mathrm{T}}\Gamma_n)|\mathcal{F}_{n-1}\}(\theta_{n-1}-\theta^\star).$$

Применяя лемму Роббинса–Сигмунда из [9] к последнему соотношению, в силу утверждений (b) и (d) заключаем, что последовательность $\{\|\theta_n-\theta^\star\|^2\}$ с вероятностью единица имеет конечный предел и

$$\sum_{n=1}^{\infty}(\theta_{n-1}-\theta^\star)^{\mathrm{T}}\,\mathrm{E}\{\Gamma_n\,\Delta_n\,\Delta_n^{\mathrm{T}}\,+\,\Delta_n\,\Delta_n^{\mathrm{T}}\,\Gamma_n|\hat{\mathcal{F}}_{n-1}\}(\theta_{n-1}-\theta^\star)\;<\;\infty.$$

В силу утверждения (c) $\sum_{n=1}^{\infty}\,\mathrm{E}\{\Gamma_n\,\Delta_n\,\Delta_n^{\mathrm{T}}|\hat{\mathcal{F}}_{n-1}\}\;=\;\infty$, следовательно, $\|\theta_n-\theta^\star\|^2\;\to\;0$ с вероятностью единица при $n\;\to\;\infty$. Первое утверждение теоремы доказано.

Докажем второе утверждение теоремы. Из логики построения оценок РМНК следует, что

$$\theta_n\;=\;\Gamma_n\,\sum_{k=1}^{n}\,\Delta_k(y_k-\mathrm{E}\{\varphi_k^{\mathrm{T}}\}\theta_{n-1}),\qquad\Gamma_n\;=\;(\sum_{k=1}^{n}\,\Delta_k\Delta_k^{\mathrm{T}}\,+\,\rho\mathrm{I})^{-1}$$

и, значит,

$$\theta_n - \theta^\star \;=\; \Big(\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;+\; \rho \mathrm{I}\Big)^{-1} \Big(\sum_{k=1}^n \Delta_k(\xi_k + \Delta_k^{\mathrm{T}}(\theta_k^\star - \theta^\star)) - \rho\theta^\star\Big).$$

Обозначив $\epsilon_k = \xi_k + \Delta_k^{\mathrm{T}}(\theta_k^\star - \theta^\star)$, имеем

$$(\theta_n - \theta^\star)(\theta_n - \theta^\star)^{\mathrm{T}} \;=\; \Big(\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;+\; \rho \mathrm{I}\Big)^{-1} \Big(\rho^2\theta^\star(\theta^\star)^{\mathrm{T}} + \sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}}\, \epsilon_k^2 -$$

$$- \rho\theta^\star \sum_{i=1}^n \Delta_i^{\mathrm{T}}\, \epsilon_i \;-\; \rho \sum_{i=1}^n \Delta_i\, \epsilon_i\, (\theta^\star)^{\mathrm{T}} + \sum_{\substack{i,j=1 \\ i\neq j}}^n \Delta_i\, \Delta_j^{\mathrm{T}}\, \epsilon_i\, \epsilon_j\Big) \Big(\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;+\; \rho \mathrm{I}\Big)^{-1}.$$

Сначала покажем, что для любого $i \neq j,\ i,j \;=\; 1, \ldots, n$

$$(e) \qquad \mathrm{E}\{\Big(\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;+\; \rho \mathrm{I}\Big)^{-1} \Delta_i\, \Delta_j^{\mathrm{T}}\, \epsilon_i\, \epsilon_j \Big(\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;+\; \rho \mathrm{I}\Big)^{-1}\} \;=\; 0.$$

Пусть для определенности $i > j$ и $\mathcal{F}_n^i$ — σ-алгебра, порождаемая случайными величинами $\{\Delta_1, \ldots, \Delta_{i-1}, \Delta_{i+1}, \ldots, \Delta_n, \theta_1^\star, \ldots, \theta_{i-1}^\star, \theta_{i+1}^\star, \ldots, \theta_n^\star, \xi_1, \ldots, \xi_n\}$. Учитывая независимость случайной величины Δ_i от $\theta_i^\star$, рассмотрим условное математическое ожидание:

$$\mathrm{E}\{\Big(\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;+\; \rho \mathrm{I}\Big)^{-1} \Delta_i\, \Delta_j^{\mathrm{T}}\, \epsilon_i\, \epsilon_j \Big(\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;+\; \rho \mathrm{I}\Big)^{-1} |\mathcal{F}_n^i\} \;=$$

$$=\; \Delta_j^{\mathrm{T}}\, \epsilon_j \mathrm{E}\{\Gamma_n(\Delta_i\xi_i + \Delta_i\Delta_i^{\mathrm{T}} w_i)\Gamma_n|\mathcal{F}_n^i\} \;=$$

$$=\; \Delta_j^{\mathrm{T}}\, \epsilon_j \mathrm{E}\{\Gamma_n\Delta_i\xi_i\Gamma_n|\mathcal{F}_n^i\} + \Delta_j^{\mathrm{T}}\, \epsilon_j \mathrm{E}\{\Gamma_n\Delta_i\Delta_i^{\mathrm{T}}\}\mathrm{E}\{w_i\Gamma_n\} \;=$$

$$=\; \Delta_j^{\mathrm{T}}\, \epsilon_j \mathrm{E}\{\Big(\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;+\; \rho \mathrm{I}\Big)^{-1} \Delta_i\, \xi_i \Big(\sum_{k=1}^n \Delta_k\, \Delta_k^{\mathrm{T}} \;+\; \rho \mathrm{I}\Big)^{-1} |\mathcal{F}_n^i\} \;= 0,$$

так как в силу допущения $(\mathbf{A})$ случайная величина Δ_i не зависима от $\{\Delta_1, \ldots, \Delta_{i-1}, \Delta_{i+1}, \ldots, \Delta_n, \xi_i\}$, и выражение под знаком условного математического ожидания является нечетной функцией от Δ_i, интегрируемой по симметричному распределению $\mathrm{P_i}(\cdot)$. Следовательно, (e) выполняется.

Аналогично можно показать, что для любого $i = 1, \ldots, n$

$$\mathrm{E}\Big\{\Big(\sum_{k=1}^{n} \Delta_k\, \Delta_k^{\mathrm{T}} \,+\, \rho \mathrm{I}\Big)^{-1} \rho \theta^{\star} \Delta_i^{\mathrm{T}}\, \epsilon_i \Big(\sum_{k=1}^{n} \Delta_k\, \Delta_k^{\mathrm{T}} \,+\, \rho \mathrm{I}\Big)^{-1}\Big\} \,=\, 0.$$

Теперь получаем в соответствии с предположением об ограниченности Δ_n и выполнением условия (**Biii**) при достаточно больших n

$$\mathrm{E}\{(\theta_n - \theta^{\star})(\theta_n - \theta^{\star})^{\mathrm{T}}\} \,=\, \mathrm{E}\Big\{\Big(\sum_{k=1}^{n} \Delta_k\, \Delta_k^{\mathrm{T}} \,+\, \rho \theta^{\star} (\theta^{\star})^{\mathrm{T}}\Big)^{-1} \times$$

$$\times (\rho^2 \theta^{\star} (\theta^{\star})^{\mathrm{T}} + \sum_{k=1}^{n} \Delta_k\, \Delta_k^{\mathrm{T}}\, \epsilon_k^2) \Big(\sum_{k=1}^{n} \Delta_k\, \Delta_k^{\mathrm{T}} \,+\, \rho \mathrm{I}\Big)^{-1}\Big\} \le \hat{C} \mathrm{E}\{\Gamma_n\}$$

с некоторой постоянной $\hat{C}$. Так как $\sum_{k=1}^{n} \Delta_k\, \Delta_k^{\mathrm{T}} \,\to\, \infty$ при $n \,\to\, \infty$ с вероятностью единица и $\|\Gamma_n\| \le \rho^{-1}$, то по теореме Лебега о доминирующей последовательности получаем $\mathrm{E}\{(\theta_n - \theta^{\star})(\theta_n - \theta^{\star})^{\mathrm{T}}\} \,\to\, 0$ при $n \,\to\, \infty$.

Доказательство теоремы 3 закончено.

СПИСОК ЛИТЕРАТУРЫ

1. *Граничин О.Н.*, "Алгоритм стохастической аппроксимации в возмущением на входе для идентификации статического нестационарного дискретного объекта" // Вестн. Ленингр. ун-та., сер. 1., 1988. Вып. 3. С. 92–93.

2. *Граничин О.Н.*, "Оценивание точки минимума неизвестной функции, наблюдаемой на фоне зависимых помех" // ППИ. 1992. No. 2. С. 16–20.

3. *Goldenshluger A.V., Polyak B.T.*, "Estimation of regression parameters with arbitrary noise" // Math. Methods of Statistics. 1993. V. 2. No. 1. P. 18–29.

4. *Ljung L., Guo L.*, "The Role of Model Validation for Assessing the Size of the Unmodeled Dynamics" // IEEE Trans. Automat. Control. 1997. V. 42. No. 9. P. 1230–1239.

5. *Fisher R.A.*, "The Design of Experiments." Oliver and Boyd. Edinburgh. 1935.

6. *Цыпкин Я.З.*, "Информационная теория идентификации." М.: Наука, 1995.

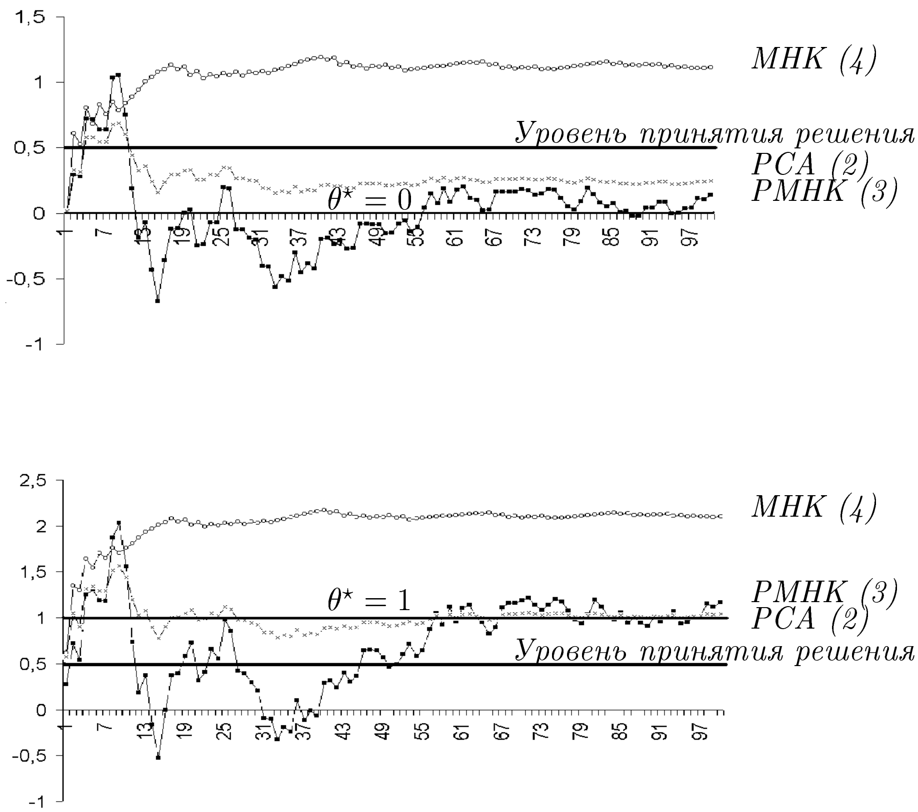
7. *Ljung L., Söderström T.*, "Theory and Practice of Recursive Identification." MIT Press, Cambridge, MA - London, 1983.

8. *Поляк Б.Т., Цыпкин Я.З.*, "Адаптивные алгоритмы оценивания (сходимость, оптимальность, устойчивость)" // АиТ. 1979. No. 3. С. 71–84.

9. *Robbins H., Siegmuud D.*, "A convergence theorem for nonnegative almost super–marlingales and some applications", Optimizing Methods in Statistics, J.S.Rustagi ed., Academic Press, NY. 1971. P. 233–257.

10. *Поляк Б.Т.*, "Сходимость и скорость сходимости итеративных стохастических алгоритмов. 2.Линейный случай" // АиТ. 1977. No. 4. С. 101–107.

11. *Poznyak A. S.*, "Estimating the parameters of autoregressive processes by the method of least squares" //Int. J. Syst. Sci. 1980. V. 11, P. 577–588.



МНК (4)
Уровень принятия решения
PCA (2)
РМНК (3)
$\theta^\star = 0$
МНК (4)
РМНК (3)
$\theta^\star = 1$
PCA (2)
Уровень принятия решения

Рис. 1.