RA-тесты для авторегрессионных моделей

Е.М. Ряднова

МГУ, МКСМ-10

5 июня 2025 г.

Мотивация-I

Рассмотрим стационарную авторегрессионную модель AR(1):

$$u_t = \beta u_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad |\beta| < 1,$$

где параметр β неизвестен, $\{\varepsilon_t\}$ – н.о.р.с.в. с неизвестной ф.р. G. Пусть наблюдаются $u_0, \dots, u_n$. Хотим проверить гипотезу

$$H_0 : \beta = \beta_0, \quad |\beta_0| < 1.$$

Хорошо известный тест наименьших квадратов имеет высокую чувствительность к выбросам. Альтернативой служит знаковый подход[1]. На основе $\{\mathsf{sign}(\varepsilon_t(\beta_0))\}$, где остатки $\varepsilon_t(\beta) = u_t - \beta u_{t-1}$, ЛНМ тестовая статистика против односторонних альтернатив

$$T_n(\beta_0) := \sum_{t=1}^{n-1} \beta_0^{t-1} \Gamma_{tn}(\beta_0), \quad \Gamma_{tn}(\beta_0) := \sum_{k=1}^{n-t} \mathsf{sign}(\varepsilon_k(\beta_0))\, \mathsf{sign}(\varepsilon_{k+t}(\beta_0)).$$

[1] Болдин, Симонова, Тюрин (1997). Знаковый статистический анализ линейных моделей. М.:Наука.

Мотивация-II

Величины Γ_{tn} — знаковые остаточные автоковариации (RA) и $T_n(\beta_0)$ – RA-статистика. Отметим, RA-оценки в ARMA моделях рассматривались ранее[2], для нашего случая оценка β есть решение уравнения

$$T_n^{\eta}(\beta) := \sum_{t=1}^{n-1} \beta^{t-1} \Gamma_{tn}(\beta) = 0, \quad \Gamma_{tn} := \sum_{k=1}^{n-t} \eta(\varepsilon_k(\beta), \varepsilon_{k+t}(\beta)).$$

Способы выбрать функцию η:

- Mallows type: $\eta(u, v) = \psi(u)\psi(v)$;
- Hampel type: $\eta(u, v) = \psi(uv)$;
- $\approx$ M-estimate: $\eta(u, v) = u\psi(v)$. В частности, $\approx$ LS: $\eta(u, v) = uv$.

Здесь функция ψ непрерывная и нечётная.

[2] Bustos, Yohai (1986) Robust Estimates for ARMA models, Journal of the American Statistical Association, 81:393, 155-168.

Постановка задачи

Рассмотрим стационарную авторегрессионную модель AR(p):

$$u_t = \beta_1 u_{t-1} + \dots + \beta_p u_{t-p} + \varepsilon_t, \quad t \in Z,$$

где ε_t — н.о.р.с.в. с неизвестной ф.р. G и $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ — неизвестные параметры такие, что корни характеристического уравнения $x^p = \beta_1 x^{p-1} + \dots + \beta_p$ лежат внутри единичного круга. Для проверки

$$H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$$

по наблюдениям $u_{1-p}, \dots, u_n$ введём остаточные автоковариации,

$$\Gamma_{tn}^{\varphi\psi}(\boldsymbol{\beta}) := \sum_{k=1}^{n-t} \varphi(\varepsilon_k(\boldsymbol{\beta}))\psi(\varepsilon_{k+t}(\boldsymbol{\beta}))$$

с некоторыми специальными функциями φ, ψ, где остатки $\varepsilon_t(\boldsymbol{\beta}) := u_t - \beta_1 u_{t-1} - \dots - \beta_p u_{t-p}$.

Постановка задачи

Определим последовательность $\delta_t(\boldsymbol{\beta})$ по рекуррентной формуле $\delta_t(\boldsymbol{\beta}) = \beta_1\delta_{t-1} + \dots + \beta_p\delta_{t-p}$, $t \in \mathbb{N}$, с начальными значениями $\delta_{1-p} = \dots = \delta_{-1} = 0$, $\delta_0 = 1$.
Тогда RA-тестовую статистику зададим p-вектором

$$T_n^{\varphi\psi}(\boldsymbol{\beta}_0) := \sum_{t=1}^{n-1} \Gamma_{tn}^{\varphi\psi}(\boldsymbol{\beta}_0)(\delta_{t-1}(\boldsymbol{\beta}_0), \dots, \delta_{t-p}(\boldsymbol{\beta}_0))'.$$

Наша цель:

1. Получить предельное распределение RA-статистики при H_0 и построить на её основе RA-тест.
2. Получить предельное распределение RA-статистики при локальных альтернативах и сравнить наши тесты с уже известными.
3. Исследовать робастные свойства предложенных тестов.

Предельное распределение при H_0

Теорема 1

Пусть $\varphi(\varepsilon_1), \psi(\varepsilon_1)$ имеют нулевые средние и конечные ненулевые дисперсии. Тогда при $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$,

$$\frac{1}{\sqrt{n}} T_n^{\varphi\psi}(\boldsymbol{\beta}_0) \xrightarrow{d} N^p(0, \mathsf{E}\,\varphi^2(\varepsilon_1)\,\mathsf{E}\,\psi^2(\varepsilon_1) K(\boldsymbol{\beta}_0)), \quad n \to \infty,$$

где $K(\boldsymbol{\beta}) := \sum_{t=1}^{\infty} (\delta_{t-1}(\boldsymbol{\beta}), \ldots, \delta_{t-p}(\boldsymbol{\beta}))'(\delta_{t-1}(\boldsymbol{\beta}), \ldots, \delta_{t-p}(\boldsymbol{\beta}))$.

Величины $\mathsf{E}\,\varphi^2(\varepsilon_1), \mathsf{E}\,\psi^2(\varepsilon_1)$ зависят вообще говоря от неизвестной $\varepsilon_1 \sim G$. Так что при построении критерия придётся использовать их состоятельные оценки.

Проверка гипотезы H_0

Следствие Теоремы1:

В условиях Теоремы 1, при гипотезе $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$,

$$A_n^{\varphi\psi} := \frac{n T_n^{\varphi\psi}(\boldsymbol{\beta}_0)' K^{-1}(\boldsymbol{\beta}_0) T_n^{\varphi\psi}(\boldsymbol{\beta}_0)}{\sum_{t=1}^{n} \varphi^2(\varepsilon_t(\boldsymbol{\beta}_0)) \sum_{t=1}^{n} \psi^2(\varepsilon_t(\boldsymbol{\beta}_0))} \xrightarrow{d} \chi^2(p), \quad n \to \infty.$$

RA-тест

Отвергаем гипотезу $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$ на асимптотическом уровне α, если

$$A_n^{\varphi\psi} > \chi^2_{1-\alpha,p},$$

где $\chi^2_{1-\alpha,p}$ есть $(1-\alpha)$-квантиль хи-квадрат распределения $\chi^2(p)$.

Предельное распределение при H_{1n}

Для нахождения асимптотической относительной эффективности наших тестов по сравнению с уже известными нужно предельное распределение тестовой статистики при локальных альтернативах.

Теорема 2

Пусть распределение $\varepsilon_1 \sim G$ имеет абсолютно непрерывную лебегову плотность $g(x) > 0$. Пусть $\varepsilon_1, g'/g(\varepsilon_1), \varphi(\varepsilon_1), \psi(\varepsilon_1)$ имеют нулевые средние и конечные ненулевые дисперсии. Тогда при локальных альтернативах $H_{1n} : \boldsymbol{\beta} = \boldsymbol{\beta}_n = \boldsymbol{\beta}_0 + n^{-1/2}\boldsymbol{\tau}$,

$$\frac{1}{\sqrt{n}} T_n^{\varphi\psi}(\boldsymbol{\beta}_0) \xrightarrow{d} N^p(-\mathsf{E}\,\varphi(\varepsilon_1)\varepsilon_1\,\mathsf{E}\,\psi(\varepsilon_1)\frac{g'}{g}(\varepsilon_1)K(\boldsymbol{\beta}_0)\boldsymbol{\tau},$$

$$\mathsf{E}\,\varphi^2(\varepsilon_1)\,\mathsf{E}\,\psi^2(\varepsilon_1)K(\boldsymbol{\beta}_0)), \quad n \to \infty.$$

Идея доказательства

Доказательство основано на третьей Лемме Ле Кама. Достаточные условия ЛАН для $\{u_t\}$ в авторегрессионных моделях получены Kreiss[3]. Как следствие методологии возникают дополнительные условия на конечность дисперсий ε_1 и $g'/g(\varepsilon_1)$.

Следствие Теоремы 2:

В условиях Теоремы 2 при локальных альтернативах H_{1n},

$$A_n^{\varphi\psi} \xrightarrow{d} \chi^2(p, \Delta^{\varphi\psi}(\boldsymbol{\beta}_0)), \quad n \to \infty,$$

$$\Delta^{\varphi\psi}(\boldsymbol{\beta}_0) = \frac{\mathsf{E}^2\,\varphi(\varepsilon_1)\varepsilon_1 \cdot \mathsf{E}^2\,\psi(\varepsilon_1)\frac{g'}{g}(\varepsilon_1)}{\mathsf{E}\,\varphi^2(\varepsilon_1)\cdot \mathsf{E}\,\psi^2(\varepsilon_1)}\boldsymbol{\tau}' K(\boldsymbol{\beta}_0)\boldsymbol{\tau}.$$

[3] Kreiss J. (1990) Testing Linear Hypothesis in Autoregressions, The Annals of Statistics, 18:3, 1470-1482.

Оптимальные φ, ψ

АОЭ по Питмену одного теста относительно другого с предельными хи-квадрат распределениями при локальных альтернативах есть отношение их параметров нецентральностей.

Наилучший выбор φ, ψ соответствует наибольшему параметру нецентральности:

$$\operatorname{argmax}_{\varphi(x)} \frac{\mathsf{E}^2\, \varphi(\varepsilon_1)\varepsilon_1}{\mathsf{E}\, \varphi^2(\varepsilon_1)} = x,$$

$$\operatorname{argmax}_{\psi(x)} \frac{\mathsf{E}^2\, \psi(\varepsilon_1)\frac{g'}{g}(\varepsilon_1)}{\mathsf{E}\, \psi^2(\varepsilon_1)} = \frac{g'}{g}(x).$$

Оптимальный выбор $\psi = g'/g$ зависит от $\varepsilon_1 \sim G$.
Для $N(0,1)$ наилучшие $\varphi(x) = \psi(x) = x$.
Для $\mathrm{Laplace}(0,1)$ наилучшие $\varphi(x) = x$, $\varphi(x) = \mathsf{sign}(x)$.

АОЭ RA-тестов относительно наименьших квадратов

$$\mathsf{АОЭ}_{\mathrm{RA,LS}} = \frac{\mathsf{E}^2\,\varphi(\varepsilon_1)\varepsilon_1 \cdot \mathsf{E}^2\,\psi(\varepsilon_1)\frac{g'}{g}(\varepsilon_1)}{\mathsf{E}\,\varphi^2(\varepsilon_1)\cdot \mathsf{E}\,\psi^2(\varepsilon_1)}.$$

φ	ψ	$N(0,1)$	$\mathrm{Lap}(0,1)$	$(1-\gamma)N(0,1)+\gamma N(0,\delta^2)$
x	x	1	1	1
x	$\operatorname{sign} x$	$\frac{2}{\pi}\approx 0.6$	2	$\frac{2}{\pi}(1-\gamma+\gamma\delta^2)(1-\gamma+\gamma/\delta)^2$
$\operatorname{sign} x$	$\operatorname{sign} x$	$\frac{4}{\pi^2}\approx 0.4$	1	$\frac{4}{\pi^2}(1-\gamma+\gamma\delta)^2(1-\gamma+\gamma/\delta)^2$

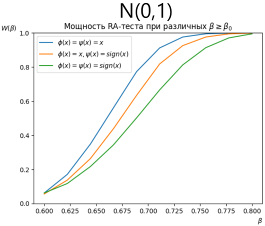
АОЭ RA-тестов относительно знаковых

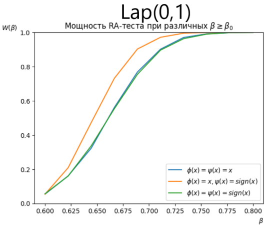
$$\text{АОЭ}_{\mathrm{RA,S}} = \frac{\mathsf{E}^2\,\varphi(\varepsilon_1)\varepsilon_1 \cdot \mathsf{E}^2\,\psi(\varepsilon_1)\frac{g'}{g}(\varepsilon_1)}{\mathsf{E}\,\varphi^2(\varepsilon_1)\cdot \mathsf{E}\,\psi^2(\varepsilon_1)} \times \frac{1}{(\mathsf{E}\,|\varepsilon_1|\cdot 2g(0))^2}.$$

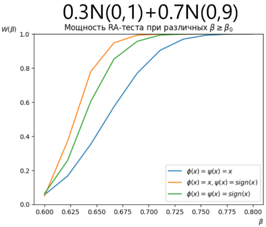
φ	ψ	$N(0,1)$	$Lap(0,1)$	$(1-\gamma)N(0,1)+\gamma N(0,\delta^2)$
x	x	$\frac{\pi^2}{4}\approx 2.5$	1	$\frac{\pi^2}{4}\frac{1}{(1-\gamma+\gamma\delta)^2(1-\gamma+\gamma/\delta)^2}$
x	$\mathrm{sign}\,x$	$\frac{\pi}{2}\approx 1.6$	2	$\frac{\pi}{2}\frac{(1-\gamma+\gamma\delta^2)}{(1-\gamma+\gamma\delta)^2}$

Сравнение мощностей RA-тестов

Смоделируем AR(1) длины $n = 400$. Проверяем $H_0 : \beta = \beta_0 = 0.6$.
Оценка функции мощности $W(\beta)$, $\beta \geqslant \beta_0$, по 2000 реализациям.

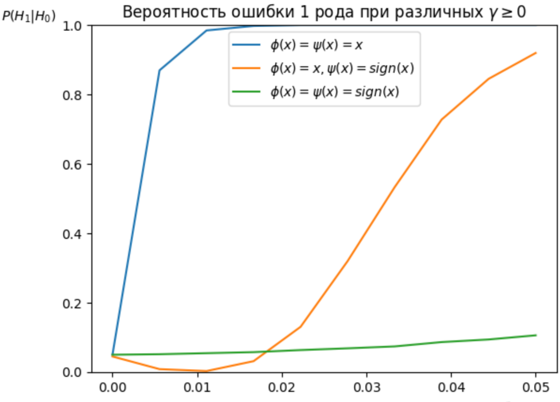






Робастность RA-тестов

Схема засорений Martin, Yohai[4]: вместо u_t наблюдаем $y_t = u_t + z_t^{\gamma}\xi_t$, где $z_t^{\gamma} \sim \mathrm{Bern}(\gamma)$ и $\xi_t \sim G_{\xi}$ — н.о.р.с.в. Моделируем $\varepsilon_t \sim N(0,1)$ и $\xi_t \sim N(0,100)$. Оцениваем $W(\beta_0)$ по 2000 реализаций в зависимости от степени засорения γ:



Литература

1. Болдин, Симонова, Тюрин (1997). Знаковый статистический анализ линейных моделей. М.:Наука.
2. Bustos, Yohai (1986) Robust Estimates for ARMA models, Journal of the American Statistical Association, 81:393, 155-168.
3. Kreiss (1990) Testing Linear Hypothesis in Autoresressions, The Annals of Statistics, 18:3, 1470-1482.
4. Очиров (2025) Дипломная работа под руководством М.В.Болдина.
5. Martin, Yohai (1986) Influence functionals for time series. The Annals of Statistics 14.3 p.781-818.