Асимптотика почти единичных корней*

Станислав Анатольев†

Российская экономическая школа, Москва, Россия

Николай Господинов‡

Университет Конкордия, Монреаль, Канада

В некоторых общепринятых асимптотических теориях предельное распределение претерпевает разрывы, или же асимптотическое распределение неточно приближает истинное конечномерное распределение. В подобных ситуациях оказывается полезен такой инструментарий, как дрейфующие параметризации, когда определенные параметры явным образом зависят от размера выборки. Дрейфующие параметризации используются среди прочего для анализа импульсных откликов и долгосрочного прогнозирования сильнозависимых процессов. Настоящее эссе представляет собой обзор подобных альтернативных асимптотических приближений в контексте временных рядов.

1 Введение

Известно, что когда авторегрессионный параметр в AR(1)-модели в точности равен единице, его МНК-оценка асимптотически ведет себя согласно распределению Дики–Фуллера. Это распределение несимметрично и, в частности, смещено и скошено. В то же время, если авторегрессионный параметр строго меньше единицы, асимптотическое распределение гауссово, то есть симметрично и, в частности, несмещено и нескошено. Истинное же распределение МНК-оценки при типичного размера выборке в случае, когда параметр близок к единице, не соответствует нормальному, а обладает всеми признаками распределения Дики–Фуллера. Это означает, что стандартная асимптотика работает в подобных случаях плохо, и хотелось бы иметь какое-то альтернативное, более адекватное, приближение реальности, некое расширение асимптотической теории Дики–Фуллера на случай корней не единичных, но близких к единичным. В дополнение к практическим соображениям, хотелось бы избавиться от скачка в асимптотическом распределении при переходе из зоны стационарности к ситуации с единичным корнем.

В данном эссе обсуждаются альтернативные асимптотические приближения, которые не рассматривают как фиксированный тот параметр, из-за которого наблюдается разрыв асимптотического распределения, а параметризуют его как дрейфующую последовательность, явным образом зависящую от размера выборки. По мере того, как выборка растет, последовательность сдвигается по направлению к границе множества значений параметра и в пределе стремится к ней. Такая искусственная статистическая конструкция гарантирует гладкий переход в асимптотической теории, что соответствует поведению и истинного конечновыборочного распределения. Подобные конструкции в современной эконометрике являются популярным инструментарием, используемым для моделей с сильно автокоррелированными

*Цитировать как: Анатольев, Станислав & Николай Господинов (2012) «Асимптотика почти единичных корней», Квантиль, №10, стр. 57–71. Citation: Anatolyev, Stanislav & Nikolay Gospodinov (2012) “Asymptotics of near unit roots,” Quantile, No.10, pp. 57–71.

†Адрес: 117418, г. Москва, Нахимовский проспект, 47, офис 1721(3). Электронная почта: *sanatoly@nes.ru*

‡Адрес: Department of Economics, Concordia University, 1455 de Maisonneuve Blvd. West, Montréal, Québec, H3G 1M8. Электронная почта: *gospodin@alcor.concordia.ca*

переменными. Недостатком такого асимптотического подхода является появление дополнительных состоятельно не оцениваемых шумовых параметров, что осложняет процедуру инференции.

Другой важной ситуацией, где введение дрейфующих последовательностей оказывается полезным, является построение долгосрочных прогнозов и функций импульсных откликов. Стандартная асимптотическая теория подразумевает, что горизонт прогнозирования фиксирован, и предел берется только по отношению к размеру выборки. В то же время, когда горизонт прогнозирования составляет соразмерную с длиной выборки величину, стандартная асимптотическая теория не дает разумных приближений для истинных конечновыборочных распределений. Поэтому выгодно параметризовать горизонт прогнозирования, заставив его расти с определенной скоростью по мере того, как размер выборки растет до бесконечности.

Замечательные обзоры некоторых тем, обсуждаемых в данном эссе, можно также найти в Stock (1994, 1997), Stock, Wright & Yogo (2002) и Stock & Watson (2008).

2 Параметризация локальности к единице

Рассмотрим AR(p)-процесс

$$y_t^* = \mu_1 + \mu_2 t + y_t, \quad \Gamma(L) y_t = e_t, \tag{1}$$

где $\Gamma(L) = 1 - \gamma_1 L - \gamma_2 L^2 - ... - \gamma_p L^p$ с корнями на или вне единичного круга, $y_{-p+1}, ..., y_0$ предполагаются фиксированными, а e_t является последовательностью мартингальных приращений с $\mathbb{E}[e_t^2] = \sigma^2$ и $\sup_t \mathbb{E}[e_t^{2+\xi}] < \infty$ при некотором $\xi > 0$. Полином $\Gamma(L)$ можно разложить как (Stock 1991)

$$\Gamma(L) = \psi(L)(1 - \phi L), \tag{2}$$

где ϕ обозначает наибольший корень AR-полинома, а $\psi(L) = 1 - \sum_{j=1}^{p-1} \psi_j L^j$ — полином от лага, описывающий краткосрочную динамику процесса без корней на или вблизи единичного круга.

Многие экономические временные ряды, включая процентные ставки, безработицу, реальные обменные курсы и подразумеваемую волатильность, характеризуются очень сильной связностью, но в то же время предположение о единичном корне плохо увязывается с экономической теорией. Скорее, речь идет о случае, когда ϕ близок к единице, но не равен в точности единице. Модель можно переписать в виде улучшенной регрессии Дики–Фуллера (ADF)

$$y_t = \rho y_{t-1} + \varphi_1 \triangle y_{t-1} + ... + \varphi_{p-1} \triangle y_{t-p+1} + e_t, \tag{3}$$

где $\rho = \gamma_1 + \gamma_2 + ... + \gamma_p$ и $\varphi_j = -(\gamma_{j+1} + ... + \gamma_p)$ для $j = 1, ..., p-1$. Такая форма часто используется на практике для оценивания, тестирования и прогнозирования. Модель (3) обычно оценивается с помощью МНК, а МНК-оценка состоятельна и асимптотически нормальна при $\rho < 1$.

Чтобы явно учесть сильную связность процесса, удобно использовать параметризацию локальности к единице (Chan & Wei 1987, Phillips 1987) наибольшего авторегрессионного корня

$$\phi_n = \exp\left(\frac{c}{n}\right) \approx 1 + \frac{c}{n},$$

где c — фиксированная константа, а n — размер выборки. Обычно предполагается, что $c \leq 0$, причем случай $c = 0$ соответствует точному единичному корню, но допускаются и слабо взрывные процессы ($c > 0$). В результате данная параметризация включает в себя целый набор значений ϕ, что особо полезно, когда неясно, каково точное значение наибольшего авторегрессионного корня. Из принятой параметризации следует, что в ADF-представлении процесса (3), $\rho_n = 1 + c\psi(1)/n$, где $\psi(1) = 1 - \psi_1 - ... - \psi_{p-1}$.

В рамках локальности к единице пространство параметров моделируется как окрестность единицы, сжимающаяся по мере того, как размер выборки растет. Такая (воображаемая!) статистическая конструкция устраняет разрыв в асимптотическом распределении и предоставляет отличное приближение для конечновыборочного распределения $\hat{\rho}$ в модели (3). Также она облегчает анализ непрерывных асимптотических пределов, т.к. $\lim_{n\to\infty}(\phi_n)^n = \exp(c)$. Наконец, стоит упомянуть еще одну важную черту принятой параметризации: в отличие от тестирования на единичные корни, где исследователю приходится в конечном счете выбирать между одним из двух типов приближений, асимптотика локальности к единице учитывает неопределенность в значении наибольшего авторегрессионного корня и дает возможность приближения во всем пространстве параметров.

Предельное распределение оценки $\hat{\rho}_n$ можно получить с помощью функциональной ЦПТ, а именно (опуская для простоты записи зависимость $\hat{\rho}_n$ от n), имеем (Phillips 1987)

$$n(\hat{\rho} - \rho_0) \Rightarrow \psi(1) \frac{\int_0^1 J_c^\tau(s) dW(s)}{\int_0^1 J_c^\tau(s)^2 ds},$$

где $\Rightarrow$ обозначает слабую сходимость, $J_c(r)$ — процесс Орнштейна–Уленбека, порождаемый стохастическим дифференциальным уравнением $dJ_c(r) = cJ_c(r)dr + dW(r)$, $\{W(r) : r \in [0,1]\}$ — стандартное броуновское движение, и $J_c^\tau(r) = J_c(r)$ если (1) не содержит детерминистических компонент ($\mu_1 = 0$, $\mu_2 = 0$), $J_c^\tau(r) = J_c(r) - \int_0^1 J_c(s)ds$ если (1) включает только константу ($\mu_1 \neq 0$, $\mu_2 = 0$) и $J_c^\tau(r) = J_c(r) - \int_0^1 (4 - 6s)J_c(s)ds - r\int_0^1 (12s - 6)J_c(s)ds$ если (1) включает и константу, и линейный тренд ($\mu_1 \neq 0$, $\mu_2 \neq 0$). Кроме того, нормализованная оценка (t-статистика) в пределе есть

$$t_\rho = \frac{\hat{\rho} - \rho_0}{\mathrm{se}(\hat{\rho})} \Rightarrow \frac{\int_0^1 J_c^\tau(s) dW(s)}{\sqrt{\int_0^1 J_c^\tau(s)^2 ds}}. \tag{4}$$

Предельные распределения для $\hat{\rho}$ и t_ρ — функции параметра локальности к единице, который состоятельно не оцениваем, так как $\hat{c} - c_0 = O_P(1)$. Данный факт осложняет инференцию в подобных моделях. В то же время параметризация локальности к единице предоставляет равномерное асимптотическое приближение для t_ρ для всех ρ в интервале $(0, 1]$ (Mikusheva 2007). В частности, если параметр локальности c уходит на $-\infty$, предельное распределение приближается к стандартному нормальному (Phillips 1987). Если $c = 0$, предельное распределение схлопывается в распределение Дики–Фуллера. Все промежуточные значения c непрерывным образом связывают стандартную асимптотику и асимптотику Дики–Фуллера.

Несмотря на то, что параметр локальности c состоятельно не оценить, существуют методы построения доверительных интервалов для c и ρ. Они основаны на обращении асимптотических (Stock 1991), бутстраповских (Hansen 1999) или Монте–Карловских (Andrews 1993) тестов на решетке значений c (или ρ). Mikusheva (2007) показывает, что упомянутые три метода дают асимптотически правильные доверительные интервалы для ρ.

Покажем, как построить доверительные интервалы для ρ. Следуя работе Stock (1991), обозначим за $q_c(\alpha)$ α-квантиль асимптотического распределения в (4). Тогда $100(1-\alpha)$%-ным доверительным интервалом для ρ будет $CI_\rho = \{\rho : q_c(\alpha/2) \leq t_\rho \leq q_c(1-\alpha/2)\}$. Асимптотические квантильные функции $q_c(\alpha/2)$ и $q_c(1-\alpha/2)$ обычно получают с помощью симуляций. Для конкретной длины выборки n эти квантильные функции завязаны на неявное значение ρ. Статистику t_ρ тогда можно вычислить для последовательности нулевых гипотез на решетке для ρ; пересечения тестовой статистики и квантильных функций дадут концы $100(1-\alpha)$% -го доверительного интервала. Методы в работах Andrews (1993) и Hansen

(1999) основаны на той же идее, но асимптотическая квантильные функции заменяются на их Монте–Карловские или бутстраповские аналоги.

Инструментарий локальности к единице также оказывается полезен при анализе последствий неверной спецификации процессов. Например, перепишем процесс (1)–(2) в виде

$$(1-\phi L)y_t = a(L)e_t,$$

где $a(L)=\psi^{-1}(L)=\sum_{i=0}^{\infty} a_i L^i$ и $a_0=1$, $\sum_{i=0}^{\infty} i|a_i|<\infty$, $\sum_{i=0}^{\infty} a_i \neq 0$, и предполагается отсутствие детерминистической компоненты.

Пусть $u_t=a(L)e_t$, тогда $u_t=\left(a(1)+(1-L)a^*(L)\right)e_t$ согласно разложению Бевериджа–Нельсона, где $a(1)=\sum_{i=0}^{\infty} a_i$, $a^*(L)=\sum_{i=1}^{\infty} a_i^* L^{i-1}$ и $a_i^*=-\sum_{j=i}^{\infty} a_j$. Обозначив $S_t=\sum_{j=1}^{t} e_j$ и используя рекурсивную подстановку и суммирование по частям, получим:

$$\begin{aligned}
y_t &= \sum_{j=1}^{t} \phi^{t-j} a(L) e_j \\
&= \left(a(1)+(1-L)a^*(L)\right) S_t + \sum_{j=1}^{t-1} \left(a(1)+(1-L)a^*(L)\right) S_{j-1}\left(\phi^{t-j+1}-\phi^{t-j}\right) \\
&= a(1)S_t + v_t + \sum_{j=1}^{t-1} a(1) S_{j-1}\left(\phi^{t-j+1}-\phi^{t-j}\right) + \sum_{j=1}^{t-1} v_{j-1}\left(\phi^{t-j+1}-\phi^{t-j}\right) \\
&= a(1)S_t + v_t + \frac{c}{n}\sum_{j=1}^{t-1} a(1)\phi^{t-j} S_{j-1} + \frac{c}{n}\sum_{j=1}^{t-1} \phi^{t-j} v_{j-1}, \qquad (5)
\end{aligned}$$

где $v_t=a^*(L)e_t$. Выражение (5) — алгебраическое разложение процесса y_t. Оно обобщает стандартное разложение Бевериджа–Нельсона для процессов с единичным корнем как частный случай ($c=0$). Интересно, что в то время как стандартное разложение Бевериджа–Нельсона задается компонентами $a(1)S_t$ (перманентной) и v_t (преходящей), разложение (5) содержит два дополнительных слагаемых. Четвертая часть

$$\frac{c}{n}\sum_{j=1}^{t-1} \phi^{t-j} v_{j-1}$$

асимптотически пренебрежимо мала, но

$$\frac{c}{n}\sum_{j=1}^{t-1} a(1)\phi^{t-j} S_{j-1}$$

нет, она порядка $O_P(n^{1/2})$, того же, что и $a(1)S_t$. Поэтому (5) содержит два слагаемых перманентной компоненты (т.е., слагаемые порядка $O_P(n^{1/2})$), так что упущение одной из них (например, неверное предположение точного единичного корня) приведет к смещенным оценкам и искаженным тестам, а также к смещенным выводам касательно стандартного разложения на тренд и цикл, основанного на предположении о точным единичном корне.

Будем обозначать как $\lfloor\cdot\rfloor$ взятие целой части и перейдем к непрерывному времени: $t=\lfloor rn\rfloor$ для $r\in[0,1]$. Тогда, деля обе части (5) на $\sqrt{n}$ и беря предел при $n\to\infty$, получаем:

$$\frac{1}{\sqrt{n}} y_{\lfloor rn\rfloor} \Rightarrow \sigma^2 a(1)\left(W(r)+c\int_0^r \exp((r-s)c)W(s)ds\right) \equiv J_c(r), \qquad (6)$$

где $J_c(r)$ — процесс Орнштейна–Уленбека, а $W(r)$ — стандартное броуновское движение. Предельный результат (6) подтверждает, что обе части (5) сходятся к одному и тому же процессу Орнштейна–Уленбека.

Полезно провести аналогичный анализ в многомерных моделях. Пусть теперь y_t — многомерный процесс

$$(I_m - \Phi L)y_t = A(L)e_t,$$

где $\Phi = I_m + C/n$, а $C = \operatorname{diag}\{c_1, c_2, ..., c_m\}$, и $c_1, c_2, ..., c_m$ — фиксированные постоянные. Используя ту же технологию, что и раньше, имеем

$$y_t = A(1)S_t + v_t + \frac{C}{n}A(1)\sum_{j=1}^{t-1}\Phi^{t-j}S_{j-1} + \frac{C}{n}\sum_{j=1}^{t-1}\Phi^{t-j}v_{j-1},$$

и наложение ограничений на $A(1)$ недостаточно, чтобы удалить перманентную компоненту, поскольку имеется другое $O_P(n^{1/2})$-слагаемое $Cn^{-1}\sum_{j=1}^{t-1}\Phi^{t-j}S_{j-1}$.

Обобщая далее, можно рассмотреть коинтеграционную модель

$$\begin{aligned} x_{1,t} &= \beta x_{2,t} + e_{1,t}, \\ x_{2,t} &= (1 + c/n)x_{2,t-1} + e_{2,t} \end{aligned}$$

рассмотренную в Elliott (1998), и получить похожее представление. В этом случае ограничение $\beta A(1) = 0$ не уничтожает перманентную компоненту. Как следствие, регрессии редуцированного ранга, которые налагают только это ограничение, неприемлемы и приводят к смещенным оценкам и искаженным тестам, как показал Elliott (1998). Более того, ограничения, наложенные на «неверную» перманентную компоненту, скорей всего негативно повлияют на импульсные отклики и разложения дисперсии, основанные на долгосрочных идентифицирующих ограничениях на $A(1)$.

3 Инференция и долгосрочное прогнозирование локальных к единице процессов

Отслеживание влияния структурных шоков на динамику эндогенных переменных является одним из главных инструментов анализа экономической политики. Часто интересно тестирование гипотез о величине и форме импульсных откликов при больших горизонтах, например $\theta_l \equiv \partial y_{t+l}/\partial e_t$ для горизонта l. Или же интересно измерить, сколько времени требуется невзрывному процессу, чтобы затухнуть на $100\omega\%$ от первоначального шока, т.е. $\sup_{l\in L}|\partial y_{t+l}/\partial e_t| \geq 1-\omega$ для некоторого фиксированного $\omega \in (0,1]$. В обоих случаях ограничения, накладываемые нулевой гипотезой, являются полиномами порядка l от параметров модели. В модели AR(1) структура затухания монотонна, и $l = \log(1-\omega)/\log(\rho)$ или $l = \lfloor \delta n \rfloor$, где $\delta = \log(1-\omega)/c$ — фиксированная положительная постоянная, если $c < 0$. Для AR-моделей более высокого порядка с корнем вблизи единицы Rossi (2005) показала, что $l = \lfloor \delta n \rfloor + o(1)$. Поэтому порядок полинома в ограничении увеличивается линейно с размером выборки по мере того, как процесс приближается к таковому с единичным корнем.

По этой причине удобно принять параметризацию $l = \lfloor \delta n \rfloor$ для некоторого фиксированного $\delta > 0$, что ранее использовалось в литературе для анализа импульсных откликов и долгосрочного прогнозирования почти нестационарных процессов (Stock 1996, Phillips 1998, Gospodinov 2002a, Rossi 2005). Данная конструкция позволяет сохранить в асимптотическом приближении неопределенность в оценивании параметров, свойственную конечноразмерному распределению. В случае же фиксированного горизонта планирования неопределенность в оценивании параметров асимптотически исчезает по мере того как n уходит в ∞.

Рассмотрим оценивание параметров модели (3) с $\rho_n = 1 + c/n$ и наложим ограничение, что θ_l равно некоторому конкретному значению $\theta_{0,l}$, где $l = \lfloor \delta n \rfloor$. Gospodinov (2004) показывает, что оценка ρ под ограничением сходится с большей скоростью ($n^{3/2}$), чем оценка без ограничений (n). В результате параметр локальности c можно оценить под ограничением. Данный результат — следствие параметризации горизонта импульсных откликов как функции размера выборки. Из-за этой параметризации параметр почти нестационарной компоненты решает сильно нелинейное полиномиальное ограничение, степень которого растет с размером выборки, что в свою очередь ускоряет сходимость оценки.

С целью иллюстрации этого результата рассмотрим модель AR(2)

$$y_t = \rho y_{t-1} + \varphi \triangle y_{t-1} + e_t.$$

Эту модель можно переписать в виде

$$\begin{pmatrix} y_t \\ \triangle y_t \end{pmatrix} = \begin{pmatrix} \rho & \varphi \\ \rho - 1 & \varphi \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \triangle y_{t-1} \end{pmatrix} + \begin{pmatrix} e_t \\ e_t \end{pmatrix}$$

или

$$Y_t = \Lambda Y_{t-1} + \bar{e}_t.$$

Тогда $Y_{t+l} = \Lambda^{l+1} Y_{t-1} + \Lambda^l \bar{e}_t + \Lambda^{l-1} \bar{e}_{t+1} + \ldots + \bar{e}_{t+l}$, а импульсный отклик при горизонте l равен

$$\theta_l \equiv \frac{\partial y_{t+l}}{\partial e_t} = (1,0) \Lambda^l \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Поскольку $\rho - 1 = c/n = o_P(1)$,

$$\begin{aligned}
&(1,0) \begin{pmatrix} \rho & \varphi \\ o_P(1) & \varphi \end{pmatrix}^l \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= (1,0) \begin{pmatrix} \rho^l + o_P(1) & \rho^{l-1}\varphi + \rho^{l-2}\varphi^2 + \ldots + \varphi^l + o_P(1) \\ o_P(1) & \varphi^l + o_P(1) \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \rho^l + \rho^{l-1}\varphi + \rho^{l-2}\varphi^2 + \ldots + \varphi^l + o_P(1) \\
&= \rho^l \left(1 + \frac{\varphi}{\rho} + \ldots + \left(\frac{\varphi}{\rho} \right)^l \right) + o_P(1).
\end{aligned}$$

Далее, если $l = \lfloor \delta n \rfloor$, по мере того как $n \to \infty$, имеем: $\rho^l = \exp(c\delta)$, $1 + \varphi/\rho + \ldots + (\varphi/\rho)^l = \rho/(\rho - \varphi) = (1-\varphi)^{-1} + o_P(1)$ и $\theta_l \to (1-\varphi)^{-1} \exp(c\delta)$ (Rossi 2004).

Обозначим как $\tilde{c}$ и $\tilde{\varphi}$ ограниченные МНК-оценки параметров c и φ при ограничении $\theta_l = \theta_{0,l}$. Из выражений выше следует, что ограниченная оценка $\tilde{c}$ является нелинейной функцией $\tilde{\varphi}$, а именно

$$\tilde{c} = \frac{\log\left(\theta_{0,l}(1-\tilde{\varphi})\right)}{\delta} = g(\tilde{\varphi}).$$

Разложение по Тейлору первого порядка $\tilde{c}$ вокруг c_0 дает

$$\sqrt{n}\,(\tilde{c} - c_0) = \frac{\partial g(\varphi)}{\partial \varphi} \sqrt{n}\,(\tilde{\varphi} - \varphi_0) + o_P(1),$$

т.е. оценка $\tilde{c}$ вынуждена сходиться с той же скоростью, что и оценка $\tilde{\varphi}$. Gospodinov (2004) развивает далее процедуру построения асимптотических поточечных доверительных коридоров для импульсных откликов на основе обращения $\mathcal{LR}$-теста.

Похожая технология используется в контексте долгосрочного условного прогнозирования локальных к единице процессов. Проиллюстрируем основную идею на примере модели AR(1) с нулевым средним

$$y_t = \rho y_{t-1} + e_t, \quad t = 1, 2, ..., n,$$

Положим, нам интересны прогнозы будущих значений y_{n+l}, условные на последнем наблюдении y_n, когда ρ находится вблизи единицы параметризовано как $\rho_n = 1+c/n$, а l сопоставимо с размером выборки и параметризовано как $l = \lfloor \lambda n \rfloor$. В авторегрессионных моделях зависимость между данными, используемыми для оценивания, и данными, используемыми для прогнозирования, порождает некоторые трудности в оценивании эффекта неопределенности в оценке параметров и их свойств на прогнозы (Phillips 1979). В отличие от стационарного случая, зависимость между $\hat{\rho}$ и y_n асимптотически не исчезает, если корни AR-процесса на или возле единичного круга. Это видно из совместной асимптотики МНК-оценки ρ и последнего наблюдения (Phillips 1987):

$$\left(n(\hat{\rho} - \rho_0), n^{-1/2} y_n\right) \Rightarrow \left(\frac{1}{2} \frac{J_c(1)^2 - J_c(0)^2 - 2c \int_0^1 J_c(s)^2 ds - 1}{\int_0^1 J_c(s)^2 ds}, J_c(1) \right).$$

Истинным значением будущего значения y через l-шагов будет

$$y_{n+l} = \rho_0^l y_n + \sum_{j=0}^{l-1} \rho_0^j e_{n+l-j},$$

истинным условным ожиданием — $\bar{y}_{n+l|n} = \mathbb{E}\left[y_{n+l}|y_n\right] = \rho_0^l y_n$, а доступным условным ожиданием — $\hat{y}_{n+l|n} = \hat{\rho}^l y_n$, где $\hat{\rho}$ — МНК-оценка. Целью является приближение распределения для $\hat{y}_{n+k|n} - \bar{y}_{n+k|n}$ и ошибок прогноза $\hat{y}_{n+k|n} - y_{n+k}$, условно на значении y_n. Поскольку y_n имеет порядок $O_P(n^{1/2})$ и расходится при $n \to \infty$, более разумно уславливать на масштабированном конечном значении $n^{-1/2} y_n = x$.

Условное асимптотическое представление нормализованной разности $\hat{y}_{n+k|n} - \bar{y}_{n+k|n}$ следующее (Gospodinov 2002a):

$$\left(n^{-1/2}\left(\hat{y}_{n+k|n} - \bar{y}_{n+k|n}\right)|x\right)$$
$$\Rightarrow \exp(\lambda c)\left(\exp\left(\frac{\lambda}{2}\frac{x^2 - J_c(0)^2 - 2c\int_0^1 J_{c|x}(s)^2 ds - 1}{\int_0^1 J_{c|x}(s)^2 ds}\right) - 1\right)x,$$

где $J_{c|x}(r)$ — траектория, согласующаяся с масштабированным конечным значением x. Условное асимптотическое представление нормализованных ошибок прогнозов

$$n^{-1/2}\left(\hat{y}_{n+k|n} - y_{n+k}\right)$$

включает дополнительную компоненту $N\left(0, \left(\exp(2c\lambda) - 1\right)/2c\right)$ (Gospodinov 2002a).

Предельные распределения зависят от состоятельно не оцениваемого c. Gospodinov (2002a) предлагает процедуру условного решетчатого бутстрапа и доказывает ее асимптотическую обоснованность при параметризациях локальности к единице и длинного горизонта.

4 Почти необратимое бегущее среднее и локальное к нулю отношение сигнала к шуму

Параметризации локальности к единице и к нулю также играют важную роль в моделировании почти необратимых процессов и процессов с почти постоянными локальными уровнями.

Чтобы понять основную идею, рассмотрим стохастический процесс $\{y_t\}_{t=1}^n$, порожденный моделью бегущего среднего первого порядка MA(1):

$$y_t = e_t - \theta e_{t-1},$$

где $|\theta| \leq 1$ и $e_t \sim iid(0, \sigma^2)$ с $\mathbb{E}[e_t^4] < \infty$.

Пусть нас интересует параметр θ, а за η обозначим (возможно бесконечномерный) вектор шумовых параметров, полностью описывающих распределение e. Если $|\theta|$ строго меньше единицы, процесс y_t обратим, и оценка максимального правдоподобия (ММП) параметра θ асимптотически нормально распределена со средним θ и дисперсией $(1-\theta^2)/n$. Когда θ находится вблизи единичного круга, гауссовское распределение является не очень качественным приближением истинного распределения ММП-оценки. Более того, когда истинный MA-параметр находится в окрестности единицы, в конечных выборках оценка принимает значения в точности на границе множества обратимости с положительной вероятностью.

Распределение ММП-оценки параметра θ в присутствии единичного MA-корня нестандартно (см. Davis & Dunsmuir 1996). Параметризовав MA-параметр как локальный к единице, т.е. в форме $\theta_n = 1 + c/n$, где $c \leq 0$ — фиксированная постоянная, мы получим конструкцию, полезную для анализа предельного поведения ММП-оценки и предоставляющую плавный переход от приближения нормальным распределением к приближению в случае необратимого MA-процесса.

Гибкая перепараметризация модели MA(1) задается процессом с постоянными локальными уровнями:

$$\begin{aligned} x_t &= \alpha_t + u_t, \\ \alpha_t &= \alpha_{t-1} + \tau \xi_t, \end{aligned}$$

где α_t ненаблюдаемый меняющийся во времени параметр, u_t и ξ_t взаимно некоррелированные белые шумы, а τ представляет из себя отношение сигнала к шуму. Взяв приращения и определив $\Delta x_t = y_t$, получим:

$$y_t = \tau \xi_t + \Delta u_t.$$

Легко показать, что здесь автокорреляционная структура та же, что и в ограниченной модели MA(1)

$$\triangle x_t = e_t - \theta e_{t-1}$$

с ограничением $0 \leq \theta \leq 1$. На самом деле существует взаимно однозначное отображение между параметрами двух представлений τ и θ, а именно, $\tau = \sqrt{(1-\theta)^2/\theta}$ и $\theta = (\tau^2 + 2 - \sqrt{\tau^4 + 4\tau^2})/2$, монотонные по θ и τ, соответственно. Отсюда следует, что тестирование на единичный MA-корень $\theta = 1$ эквивалентно тестированию на постоянство α, т.е. нулевой гипотезы $H_0 : \tau = 0$ против альтернативы $H_1 : \tau > 0$. Параметризация локальности к единице MA-параметра $\theta_n = 1 + c/n$ в ограниченной модели MA(1) соответствует параметризации локальности к нулю отношения сигнала к шуму $\tau_n = \lambda/n$ в модели с постоянными локальными уровнями. Gospodinov (2002b) предлагает бутстраповскую процедуру построения доверительного интервала и несмещенного (в смысле медианы) оценивания θ и τ и доказывает ее асимптотическую обоснованность при параметризациях локальности к единице и к нулю.

5 Почти коинтегрированные временные ряды с локальной к нулю дисперсией коинтеграционной ошибки

Введенные выше дрейфующие параметризации оказываются полезными и в анализе многомерных процессов. Рассмотрим треугольную систему

$$\begin{aligned} y_t &= \beta x_t + u_{y,t}, \\ x_t &= u_{x,t} \end{aligned} \tag{7}$$

и

$$\begin{pmatrix} (1-\rho L)u_{y,t} \\ (1-L)u_{x,t} \end{pmatrix} = \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix},$$

где $(\varepsilon_{y,t}, \varepsilon_{x,t})'$ — вектор iid-случайных величин с нулевым средним; для простоты мы предполагаем отсутствие динамики более высоких порядков и детерминистических компонент. Заметим, что коинтеграция соответствует случаю $\rho = 1$, а отсутствие таковой — случаю $-1 < \rho < 1$. Предположим теперь, что $\mathbb{V}(\varepsilon_{y,t}) \ll \mathbb{V}(\varepsilon_{x,t}) = 1$. Такую конструкцию можно мотивировать в контексте регрессий форвардных премий, которые описывают динамику спотовых и одномесячных форвардных обменных курсов и их разницы, которая и называется форвардной премией. Существует объемная литература по тестированию на коинтеграцию между спотовыми и форвардными обменными курсами. Из поведения их совместной динамики во времени видно, что два ряда очень тесно следуют друг за другом и не показывают намерения сильно отклоняться друг от друга. В то же время, сильная устойчивость, т.е. наличие почти единичного корня в поведении ошибок коинтеграции, приводит к тому, что статистические тесты часто отвергают гипотезу коинтеграции, что входит в противоречие с экономической интуицией. Анализ совместного процесса, однако, показывает, что компонента, связанная с единичным корнем, настолько маленькая, что не дает переменным отклоняться друг от друга в течение достаточно долгих периодов. Каким образом можно смоделировать и проанализировать такую систему? Чтобы отразить наблюдаемые свойства данных, вспомним и скомбинируем уже встречавшиеся дрейфующие параметризации.

Чтобы смоделировать сильную связность и малую дисперсию ошибок коинтеграции, т.е. форвардной премии, параметризуем $u_{y,t}$ как угасший процесс с почти единичным корнем. Это подразумевает двойную локализацию:

$$\rho_n = 1 + \frac{c}{n}$$

для некоторой фиксированной постоянной $c \leq 0$ и

$$\varepsilon_{y,t} = \tau v_t,\ \tau = \frac{\lambda}{\sqrt{n}}$$

для некоторых фиксированной постоянной $\lambda > 0$ и процесса $v_t \sim iid(0,1)$. Обозначим за $\{(V_1(r), V_2(r))' : r \in [0,1]\}$ двумерное броуновское движение с коэффициентом корреляции δ. Заметим, что хотя процесс $u_{y,t}$ почти интегрирован, он ограничен:

$$u_{y,t} = \frac{\lambda}{\sqrt{n}} \sum_{i=1}^{t} \left(1 + \frac{c}{n}\right)^{t-i} v_i \Rightarrow \lambda J_c(r),$$

где $J_c(r)$ — процесс Орнштейна–Уленбека, порожденный стохастическим дифференциальным уравнением $dJ_c(r) = cJ_c(r)dr + dV_2(r)$.

Теперь исследуем последствия такой параметризации для свойств оценок модели коинтеграции. Используя подход контролирующей переменной, получим, что эффективную оценку β можно получить из регрессии

$$y_t = \beta x_t + \omega \triangle x_t + e_t,$$

где ω — регрессионный коэффициент в уравнении $u_{y,t}$ на $\varepsilon_{x,t}$, а e_t — ошибки этой регрессии. МНК-оценка β в этой регрессии эквивалентна ММП-оценке (Phillips 1991). Оценка $\hat{\beta}$ асимптотически распределена как (Gospodinov 2009)

$$\sqrt{n}\left(\hat{\beta} - \beta_0\right) \Rightarrow \lambda \frac{\int_0^1 J_c(s)V_1(s)ds}{\int_0^1 V_1(s)^2 ds}.$$

Эта оценка состоятельна, но ее скорость сходимости ниже, чем n. Кроме того, общепринятая t-статистика для гипотезы $H_0 : \beta = \beta_0$ расходится со скоростью $n^{1/2}$, как в ложных регрессиях.

Векторное представление коррекции ошибки (VEC) модели (7) следующее:

$$\begin{pmatrix} \triangle y_t \\ \triangle x_t \end{pmatrix} = (\rho - 1) \begin{pmatrix} 1 & -\beta \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix},$$

а одиночное уравнение условной VEC-модели выглядит как

$$\triangle y_t = \gamma u_{y,t-1} + \bar{\omega} \triangle x_t + \varepsilon_{y,t}, \tag{8}$$

где $\gamma = \rho - 1$, а нулевая гипотеза коинтеграции — это $H_0 : \gamma = 0$. Обозначим через $\tilde{\gamma}$ МНК-оценку γ в (8). Взятие пределов выявляет, что асимптотическое поведение оценки $\tilde{\gamma}$ в условной VEC-модели такое же, как в Hansen (1995) и Zivot (2000):

$$n\left(\tilde{\gamma} - \gamma_0\right) \Rightarrow \delta \frac{\int_0^1 J_c(s)dW_1(s)}{\int_0^1 J_c(s)^2 ds} + (1-\delta)^{1/2} \frac{\int_0^1 J_c(s)dV_2(s)}{\int_0^1 J_c(s)^2 ds}$$

и

$$t_{\tilde{\gamma}} \Rightarrow \delta \bar{z} + \sqrt{1-\delta^2}\, \frac{\int_0^1 J_c(s)dV_2(s)}{\left(\int_0^1 J_c(s)^2 ds\right)^{1/2}}$$

где $\bar{z}$ — стандартная нормальная случайная величина, а W_1 — стандартное броуновское движение, независимое от V_2.

Часто VEC-модель определяется как (например, для прогнозирования)

$$\triangle y_t = \gamma u_{y,t-1} + \xi_t, \tag{9}$$

где $\gamma = \rho - 1$ и $\xi_t = \bar{\omega} \triangle x_t + \varepsilon_{y,t}$. Обозначим $\hat{\gamma}$ МНК-оценку γ в (9). В этом случае

$$\sqrt{n}\left(\hat{\gamma} - \gamma_0\right) \Rightarrow \frac{1}{\lambda} \left(\delta \frac{\int_0^1 J_c(s)dV_2(s)}{\int_0^1 J_c(s)^2 ds} + (1-\delta)^{1/2} \frac{\int_0^1 J_c(s)dW_1(s)}{\int_0^1 J_c(s)^2 ds} \right)$$

и

$$t_{\hat{\gamma}} \Rightarrow \delta \frac{\int_0^1 J_c(s)dV_2(s)}{\left(\int_0^1 J_c(s)^2 ds\right)^{1/2}} + \sqrt{1-\delta^2}\, \bar{z}.$$

Разница между моделью (9) и условной моделью (8) в более медленной сходимости оценки. Кроме того, в отличие от условной VEC-модели, параметр локальности λ из параметризации отношения сигнала к шуму входит в предельное распределение оценки. Поскольку λ фигурирует в знаменателе, его близкие к нулю значения приводят к сильной изменчивости оценки. Наконец, когда $\delta \to 0$, распределение t-статистики в безусловной VEC-модели асимптотически приближается к стандартному нормальному, в то время как в условной VEC-модели оно приближается к распределению типа Дики–Фуллера, и наоборот, когда $\delta \to 1$. Симуляции свидетельствуют, что смещение оценки β в условной VEC-модели может быть большим и отрицательным, когда корреляция между ошибками δ близка к единице, в то время как оценка в условной VEC-модели остается несмещенной. Более того, дисперсия оценки в безусловной модели может многократно превышать дисперсию оценки в условной (Gospodinov 2009).

6 Локальность к единице и почти идентификация

Существует тесная связь между локальностью к единице и слабой идентификацией в структурных моделях временных рядов. Рассмотрим двумерный векторный авторегрессионный процесс для $\tilde{y}_t = (y_{1,t}, y_{2,t})'$ порядка $p+1$

$$\Psi(L)(I_2 - \Phi L)\tilde{y}_t = u_t,$$

где матрица Φ содержит наибольшие корни системы и

$$\Psi(L) = I_2 - \sum_{i=1}^{p} \Psi_i L^i = \begin{bmatrix} \psi_{11}(L) & \psi_{12}(L) \\ \psi_{21}(L) & \psi_{22}(L) \end{bmatrix}$$

является полиномом p-го порядка с корнями вне единичного круга. Предполагается, что ошибки u_t являются двумерной мартингал-разностью, причем $\mathbb{E}\left[u_t u_t' | u_{t-1}, u_{t-2}, ...\right] = \Sigma > 0$ и $\sup_t \mathbb{E}[\|u_t\|^{2+\xi}] < \infty$ для некоторого $\xi > 0$, а стартовые значения предполагаются фиксированными.

Матрица наибольших корней Φ параметризуется как

$$\Phi_n = \begin{bmatrix} 1 & 0 \\ 0 & 1 + c/n \end{bmatrix} = I_2 + \frac{C}{n},$$

где

$$C = \begin{bmatrix} 0 & 0 \\ 0 & c \end{bmatrix},$$

и $c \leq 0$ — фиксированная постоянная. Внедиагональные элементы Φ_n приравниваются к нулю, чтобы исключить случай, когда какой-либо из процессов (почти) $I(2)$ (Elliott 1998, Phillips 1988). Структура этой матрицы также означает, что процессы некоинтегрированы.

Выгодно наложить наличие точного единичного корня на первую переменную, так что тогда $\triangle y_{1,t}$ будет стационарным процессом. Пусть $y_t = (\triangle y_{1,t}, y_{2,t})'$ и

$$D(L) = \Psi(L) \begin{bmatrix} 1 & 0 \\ 0 & 1 - (1 + c/n)L \end{bmatrix}.$$

Тогда приведенная форма VAR-модели выглядит как

$$D(L) y_t = u_t, \tag{10}$$

или

$$y_t = D_1 y_{t-1} + ... + D_{p+1} y_{t-p-1} + u_t.$$

Умножение слева обоих частей (10) на матрицу

$$B_0 = \left[\begin{array}{cc} 1 & -b_{12}^{(0)} \\ -b_{21}^{(0)} & 1 \end{array} \right]$$

дает

$$B(L)y_t = \varepsilon_t,$$

где $B(L) = B_0 D(L)$, а $\varepsilon_t = B_0 u_t$ обозначают структурные шоки $(\varepsilon_{1,t}, \varepsilon_{2,t})'$, которые согласно обычной практике предполагаются ортогональными с дисперсиями σ_1^2 и σ_2^2. Долгосрочное идентифицирующее ограничение об отсутствии долгосрочного влияния шоков ε_2 на y_1 накладывает нижнетреугольную структуру на матрицу бегущего среднего $D(1)^{-1}B_0^{-1}$. При этом идентифицирующем ограничении матрица долгосрочных мультипликаторов в структурной модели, $B(1) = B_0 D(1)$, также нижнетреугольна.

В демонстрационных целях упростим модель, сделав предположение $\Psi(L) = I_2$. Приведенную форму этой модели первого порядка можно переписать как

$$\begin{aligned} \triangle y_{1,t} &= u_{1,t}, \\ y_{2,t} &= \left(1 + \frac{c}{n}\right) y_{2,t-1} + u_{2,t}. \end{aligned} \tag{11}$$

Структурная форма модели выглядит так:

$$\begin{aligned} \triangle y_{1,t} &= b_{12}^{(0)} y_{2,t} + b_{12}^{(1)} y_{2,t-1} + \varepsilon_{1,t}, \\ y_{2,t} &= b_{21}^{(0)} \triangle y_{1,t} + b_{22}^{(1)} y_{2,t-1} + \varepsilon_{2,t}, \end{aligned} \tag{12}$$

где $b_{12}^{(1)} = -b_{12}^{(0)}(1 + c/n)$ и $b_{22}^{(1)} = 1 + c/n$.

Чтобы идентифицировать структурные параметры модели (12) из приведенной формы (11), мы налагаем долгосрочное ограничение о нижнетреугольной структуре $B(1)$. Точнее, добавляя и вычитая $b_{12}^{(0)} y_{2,t-1}$ из первого уравнения (12) и полагая долгосрочный мультипликатор $b_{12}^{(0)} + b_{12}^{(1)}$ нулю, мы преобразуем первое уравнение структурной модели к

$$\triangle y_{1,t} = b_{12}^{(0)} \triangle y_{2,t} + \varepsilon_{1,t}. \tag{13}$$

Поскольку $\triangle y_{2,t}$ эндогенно, неизвестный параметр $b_{12}^{(0)}$ можно оценить инструментальными переменными, используя $y_{2,t-1}$ в качестве инструмента. Соотношение между эндогенной переменной и инструментом задается вторым уравнением приведенной формы VAR (11)

$$\triangle y_{2,t} = \frac{c}{n} y_{2,t-1} + u_{2,t}. \tag{14}$$

Заметим что параметризация локальности к единице автоматически приводит к локальной к нулю корреляции между эндогенной переменной и инструментом, что напоминает аналитический инструментарий слабых инструментов (Staiger & Stock 1997, Паган 2007), но здесь корреляция между эндогенной переменной и инструментом сходится к нулю со скоростью n из-за почти нестационарности.

Подставляя (13) и (14) в инструментальную оценку $b_{12}^{(0)}$, равную

$$\hat{b}_{12}^{(0)} = \frac{\sum_{t=2}^{n} y_{2,t-1} \triangle y_{1,t}}{\sum_{t=2}^{n} y_{2,t-1} \triangle y_{2,t}},$$

получим

$$\hat{b}_{12}^{(0)} - b_{12}^{(0)} = \frac{n^{-1} \sum_{t=2}^{n} y_{2,t-1} \varepsilon_{1,t}}{c n^{-2} \sum_{t=2}^{n} y_{2,t-1}^2 + n^{-1} \sum_{t=2}^{n} y_{2,t-1} u_{2,t}}.$$

Предельное распределение инструментальной оценки $b_{12}^{(0)}$ задается как (Gospodinov 2010)

$$\hat{b}_{12}^{(0)} - b_{12}^{(0)} \Rightarrow \frac{d_2}{d_1} \frac{\delta \int_0^1 J_c(s)dV_1(s) + \sqrt{1-\delta^2} \int_0^1 J_c(s)dV_2(s)}{c\int_0^1 J_c(s)^2 ds + \int_0^1 J_c(s)dV_1(s)},$$

где $V_1(r)$ и $V_2(r)$ — независимые стандартные броуновские движения,

$$J_c(r) = \exp(cr) \int_0^r \exp(-cs)dV_1(s)$$

— процесс Орнштейна–Уленбека,

$$\delta^2 = \frac{(b_{21}^{(0)})^2 \sigma_1^2}{(b_{21}^{(0)})^2 \sigma_1^2 + \sigma_2^2}$$

— квадрат коэффициента корреляции между $\varepsilon_{1,t}$ и $u_{2,t}$, и d_2/d_1 — отношение стандартных отклонений $\varepsilon_{1,t}$ и $u_{2,t}$, соответственно.

Этот результат показывает, что инструментальная оценка $b_{12}^{(0)}$ несостоятельна. Причина несостоятельности в том, что $y_{2,t-1}$ является слабым инструментом в том смысле, что несет очень мало информации об эндогенной переменной $\triangle y_{2,t}$. В свою очередь, слабость инструмента возникает из-за персистентной природы процесса $y_{2,t}$, моделируемого как имеющего локальный к единице корень. Более того, предельное распределение $\hat{b}_{12}^{(0)} - b_{12}^{(0)}$ нестандартно. Во-первых, числитель предельного распределения — смесь гауссовой случайной величины и функционала от процесса Орнштейна–Уленбека с весами, определяющимися корреляцией между первым структурным шоком и шоком приведенной формы второго уравнения. Во-вторых, предельное распределение — отношение двух случайных величин, ибо знаменатель тоже случайная величина, содержащая функционалы процесса Орнштейна–Уленбека.

Полезно взглянуть на ситуацию, когда $y_{2,t}$ приближается к процессу с точным единичным корнем ($c \to 0$). В этом случае распределение $\hat{b}_{12}^{(0)} - b_{12}^{(0)}$ имеет предел

$$\frac{d_2}{d_1} \left(\delta + \sqrt{1-\delta^2} \frac{\int_0^1 V_1(s)dV_2(s)}{\int_0^1 V_1(s)dV_1(s)} \right).$$

Более того, если $\delta = 1$, предельное распределение $\hat{b}_{12}^{(0)} - b_{12}^{(0)}$ имеет точечную вероятностную массу в точке d_2/d_1. В случае другой крайности $\delta = 0$ распределение $\hat{b}_{12}^{(0)} - b_{12}^{(0)}$ сходится к отношению стандартного нормального распределения

$$\frac{\int_0^1 V_1(s)dV_2(s)}{\left(\int_0^1 V_1^2(s)ds\right)^{1/2}}$$

и распределения Дики–Фуллера

$$\frac{\int_0^1 V_1(s)dV_1(s)}{\left(\int_0^1 V_1^2(s)ds\right)^{1/2}}.$$

Асимптотические распределения оценок остальных структурных параметров и импульсные отклики можно получить аналогично (Gospodinov 2010). Оказывается, что проблема слабого инструмента, которая лежит в основе несостоятельности $\hat{b}_{12}^{(0)}$, также портит оценивание $b_{21}^{(0)}$ и импульсные отклики, делая их оценки несостоятельными. Подробности и улучшенную процедуру инференции можно найти в Gospodinov (2010).

Список литературы

Паган, А. (2007). Слабые инструменты. *Квантиль* 2, 71–81.

Andrews, D.W.K. (1993). Exactly median-unbiased estimation of first order autoregressive/unit root models. *Econometrica* 61, 139–165.

Chan, N.H. & C.Z. Wei (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* 15, 1050–1063.

Davis, R.A. & W.T.M. Dunsmuir (1996). Maximum likelihood estimation for MA(1) processes with a root on or near the unit circle. *Econometric Theory* 12, 1–29.

Elliott, G. (1998). On the robustness of cointegration methods when regressors almost have unit roots. *Econometrica* 66, 149–158.

Gospodinov, N. (2002a). Median unbiased forecasts for highly persistent autoregressive processes. *Journal of Econometrics* 111, 85–101.

Gospodinov, N. (2002b). Bootstrap-based inference in models with a nearly noninvertible moving average component. *Journal of Business & Economic Statistics* 20, 254–268.

Gospodinov, N. (2004). Asymptotic confidence intervals for impulse responses of near-integrated processes. *Econometrics Journal* 7, 505–527.

Gospodinov, N. (2009). A new look at the forward premium puzzle. *Journal of Financial Econometrics* 7, 312–338.

Gospodinov, N. (2010). Inference in nearly nonstationary SVAR models with long-run identifying restrictions. *Journal of Business & Economic Statistics* 28, 1–12.

Hansen, B.E. (1995). Rethinking the univariate approach to unit root tests: how to use covariates to increase power. *Econometric Theory* 11, 1148–1171.

Hansen, B.E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* 64, 413–430.

Hansen, B.E. (1999). The grid bootstrap and the autoregressive model. *Review of Economics and Statistics* 81, 594–607.

Mikusheva, A. (2007). Uniform inference in autoregressive models. *Econometrica* 75, 1411–1452.

Phillips, P.C.B. (1979). The sampling distribution of forecasts from a first-order autoregression. *Journal of Econometrics* 9, 241–261.

Phillips, P.C.B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika* 74, 535–547.

Phillips, P.C.B. (1988). Regression theory for near-integrated time series. *Econometrica* 56, 1021–1043.

Phillips, P.C.B. (1991). Optimal inference in cointegrated systems. *Econometrica* 59, 283–306.

Phillips, P.C.B. (1998). Impulse response and forecast error variance asymptotics in nonstationary VARs. *Journal of Econometrics* 83, 21–56.

Rossi, B. (2005). Confidence intervals for half-life deviations from purchasing power parity. *Journal of Business & Economic Statistics* 23, 432–442.

Stock, J.H. (1991). Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series. *Journal of Monetary Economics* 28, 435–459.

Stock, J.H. (1994). Unit roots, structural breaks and trends. Глава 46 в *Handbook of Econometrics* под редакцией Engle, R. & D. McFadden, том 4, стр. 2739–2841. North-Holland: Amsterdam.

Stock, J.H. (1996). VAR, error-correction and pretest forecasts at long horizons. *Oxford Bulletin of Economics and Statistics* 58, 685–701.

Stock, J. H. (1997). Cointegration, long-run comovements and long horizon forecasting. Глава в *Advances in Econometrics: Proceedings of the Seventh World Congress of the Econometric Society* под редакцией Kreps, D. & K.F. Willis, 34–60. Cambridge University Press: Cambridge.

Stock, J.H. & M.W. Watson (2008). What's new in econometrics – time series. *NBER Summer Institute.*

Stock, J.H, J.H. Wright & M. Yogo (2002). A Survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics* 20, 518–529.

Asymptotics of near unit roots

Stanislav Anatolyev

New Economic School, Moscow, Russia

Nikolay Gospodinov

Concordia University, Montréal, Canada

Sometimes the conventional asymptotic theory yields that the limiting distribution changes discontinuously, or that the asymptotic distribution does not approximate accurately the actual finite-sample distribution. In such situations one finds useful an asymptotic tool of drifting parameterizations where certain parameters are allowed to depend explicitly on the sample size. It proves useful, among other things, for impulse response analysis and forecasting of strongly dependent processes at long horizons. This essay provides a review of these alternative asymptotic approximations in the context of time series models.