Статьи: эконометрическая теория

Тестирование предсказательной способности при наличии структурных сдвигов*

Павел Яськов†

Московский государственный университет, Москва, Россия

В настоящей статье предложен новый подход к тестированию предсказательной способности при наличии структурных сдвигов в данных. Подход расширяет результаты работ West (1996) и West & McCracken (1998) и является альтернативой методам, предложенным в Giacomini & White (2006) и Giacomini & Rossi (2010).
Ключевые слова: тестирование предсказательной способности, оценивание качества прогнозов, структурные сдвиги
Классификация JEL: C22, C52, C53

1 Введение

В настоящей работе решается задача сравнения качества прогнозов двух конкурирующих моделей для заданного временного ряда. Первый строгий подход к ее решению дан в статье West (1996). Дальнейшее развитие этого подхода можно найти в West & McCracken (1998), McCracken (2000), White (2000), Clark & McCracken (2001), Corradi & Swanson (2005).

Во всех указанных работах рассматриваются ряды, имеющие стационарную структуру. Новый метод тестирования предсказательной способности, пригодный при наличии структурных сдвигов в данных, разработан Giacomini & White (2006). Он радикально отличается от предшествующих, поскольку не предполагает существование популяционных значений параметров, участвующих в моделях прогноза. При этом, вообще говоря, сравниваются не модели прогноза, а конкретные схемы, по которым строятся такие прогнозы (см. подробнее Giacomini & White, 2006). Другие варианты данного подхода рассмотрены в Giacomini & Rossi (2010).

В данной работе предлагается альтернативный к Giacomini & White (2006) способ тестирования предсказательной способности, продолжающий линию West (1996) и применимый для данных с небольшим числом структурных сдвигов. Наша работа отчасти связана со статьей Китов (2009), где выводится асимптотическое смещение второго порядка для статистик, рассмотренных в West & McCracken (1998).

2 Постановка задачи

Пусть имеется временной ряд $\{(x_t, y_t)\}_{t=1}^{\infty}$ и две модели прогноза величин y_t, здесь $y_t \in \mathbb{R}$ – зависимая переменная, а $x_t \in \mathbb{R}^d$ – набор объясняющих факторов. Допустим также, что i-ая модель ($i = 1, 2$) имеет вектор параметров θ_i и случайные функции прогноза $p_i(x_t, \theta_i)$. При этом оптимальное значение θ_{it} параметров i-ой модели в момент t определяется из экстремальной задачи

$$\min_{\theta_i} \mathbb{E}[\phi_i(x_t, y_t, \theta_i)],$$

*Цитировать как: Яськов, Павел (2010). «Тестирование предсказательной способности при наличии структурных сдвигов», Квантиль, №8, стр. 127–135. Citation: Yaskov, Pavel (2010). "Testing for predictive ability in the presence of structural breaks," Quantile, No.8, pp. 127–135.

†Адрес: 119992, Россия, г. Москва, Воробьевы горы. Электронная почта: *pavel.yaskov@gmail.com*

где ϕ_i – некоторая функция. Для заданной функции потерь L положим

$$f_t(\theta) = L(y_t, p_1(x_t, \theta_1)) - L(y_t, p_2(x_t, \theta_2)), \quad \theta = (\theta_1', \theta_2')'.$$

Определим нулевую и альтернативную гипотезы как

$$H_0 : \limsup_n \frac{1}{n} \sum_{t=1}^{n} \mathbb{E} f_t \le 0, \quad H_A : \limsup_n \frac{1}{n} \sum_{t=1}^{n} \mathbb{E} f_t > 0,$$

здесь и далее

$$f_t = f_t(\theta_t), \quad \theta_t = (\theta_{1t}', \theta_{2t}')'.$$

Таким образом, при H_0 прогноз у первой модели в среднем не хуже, чем у второй.

Как и в West & McCracken (1998), будем рассматривать вневыборочную схему получения оценок со скользящим окном. Другими словами, для любого $m < n$ и $t \in [m, n)$ определим

$$\hat\theta_{it} = \arg\min_{\theta_i} \sum_{s=t-m+1}^{t} \phi_i(x_t, y_t, \theta_i), \quad \hat\theta_t = (\hat\theta_{1t}', \hat\theta_{2t}')',$$

и рассмотрим статистику $\hat S_{mn}$, заданную формулой

$$\hat S_{mn} = \frac{1}{\sqrt{n}} \sum_{t=m}^{n-1} f_{t+1}(\hat\theta_t).$$

В данном случае m является длиной окна, по которому в каждый момент $t = m, \dots, n-1$ переоцениваются параметры моделей 1 и 2. Таким образом имитируется ситуация, когда в момент t мы знаем лишь значения величин x_s и y_s вплоть до момента $s = t$ и можем использовать только эту информацию для построения будущих прогнозов. Параметр m, вообще говоря, подлежит выбору исследователя. Мы предполагаем, что этот выбор уже сделан, так что m фиксировано.

При определенных условиях (см. следующий раздел) статистика $\hat S_{mn}$ асимптотически нормальна, а тест H_0 подходящего размера α имеет вид

$$\hat S_{mn} - \hat\mu_{mn} > \hat\sigma_{mn} \cdot z_\alpha, \tag{1}$$

где $\hat\mu_{mn}, \hat\sigma_{mn}$ – оценки асимптотических смещения и дисперсии, а z_α – квантиль уровня α распределения $\mathcal{N}(0,1)$.

3 Предположения

Для простоты изложения будем считать, что $\theta_i \in \mathbb{R}$. Результаты и предположения в общем случае по существу не меняются. Введем условия:
(A1) $\theta = (\theta_1, \theta_2)$ лежит в некотором компакте $\Theta \subset \mathbb{R}^d$;
(A2) ϕ_1, ϕ_2, f_t почти всюду трижды дифференцируемы по θ;
(A3) абсолютные моменты порядка p величин

$$\sup_\theta |f_t(\theta)|^{1/p}, \quad f_t(\theta_t), \quad \frac{\partial^k f_t}{\partial \theta_i^k}(\theta_t), \quad \sup_\theta \left| \frac{\partial^3 f_t}{\partial \theta_i^3}(\theta) \right|, \quad \sup_{\theta_i} \left| \frac{\partial^k \phi_i}{\partial \theta_i^k}(x_t, y_t, \theta_i) \right|$$

равномерно ограничены по t для всех $i = 1, 2$, $k = 0, 1, 2$ и некоторого $p > 3$;
(A4) для достаточно малых $\varepsilon > 0$ при всех t и i

$$\sup_{\|\theta_i - \theta_{it}\| > \varepsilon} \mathbb{E}\phi_i(x_t, y_t, \theta_i) - \mathbb{E}\phi_i(x_t, y_t, \theta_{it}) > 0;$$

(A5) функция

$$\psi_i(\varepsilon) = \max_{k=0,1} \sup_{t,q_i} \Big| \mathbb{E} \sup_{\|\theta_i - q_i\| \le \varepsilon} \frac{\partial^k \phi_i}{\partial \theta_i^k}(x_t, y_t, \theta_i) - \mathbb{E} \inf_{\|\theta_i - q_i\| \le \varepsilon} \frac{\partial^k \phi_i}{\partial \theta_i^k}(x_t, y_t, \theta_i) \Big|$$

убывает к 0, когда $\varepsilon \to 0$;
(A6) число различных временных отрезков вида $[a, b]$, на которых ряд $\{(x_t, y_t)\}_{t=1}^n$ является стационарным в узком смысле, равно $l = l(n)$, и m удовлетворяет

$$l = o(\sqrt{n}/m), \quad m, n \to \infty;$$

(A7) $\{(x_t, y_t)\}_{t=1}^\infty$ удовлетворяет условию сильного перемешивания и имеет коэффициенты перемешивания α_n порядка $o(n^{3p/(2p-6)})$.
Замечание: Условие стационарности в узком смысле ряда $\{(x_t, y_t)\}_{t=1}^n$ можно ослабить, предполагая стационарность в широком смысле векторов из величин вида $h(x_t, y_t, \theta_t)$, где в качестве h взято $f, \partial f/\partial\theta, \phi_i, \partial\phi_i/\partial\theta$ и т.п.

4 Приближенное распределение основной статистики

Введем следующие обозначения:

$$g_t = g_t(\theta_t), \quad \partial g = \frac{\partial g}{\partial \theta'}, \quad \partial^1 (g_1, g_2)' = \begin{pmatrix} \dfrac{\partial g_1}{\partial \theta_1} \\ \dfrac{\partial g_2}{\partial \theta_2} \end{pmatrix}, \quad \partial^i (g_1, g_2)' = \begin{pmatrix} \dfrac{\partial^i g_1}{\partial \theta_1^i} & 0 \\ 0 & \dfrac{\partial^i g_2}{\partial \theta_2^i} \end{pmatrix}, \quad i = 2, 3,$$

$$\phi_t(\theta) = \begin{pmatrix} \phi_1(x_t, y_t, \theta_1) \\ \phi_2(x_t, y_t, \theta_2) \end{pmatrix}, \quad \begin{pmatrix} a \\ b \end{pmatrix}^* = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad u_t = \mathbb{E}[\partial^2 \phi_t]^{-1} \partial^1 \phi_t, \quad w_t = \mathbb{E}[\partial^2 f_{t+1}] \cdot u_t,$$

$$v_t = \mathbb{E}[\partial f_{t+1}] \mathbb{E}[\partial^2 \phi_t]^{-1} \partial^2 \phi_t, \quad z_t = \mathbb{E}[\partial f_{t+1}] \mathbb{E}[\partial^2 \phi_t]^{-1} u_t^* \, \mathbb{E}[\partial^3 \phi_t], \quad h_t = \mathbb{E}[\partial f_{t+1}] \cdot u_t.$$

Положим также

$$S_{mn} = \frac{1}{\sqrt{n}} \sum_{t=m}^{n-1} (f_{t+1}(\hat\theta_t) - \mathbb{E} f_{t+1}).$$

Теорема 1: Допустим, что выполнены условия (A1)–(A7). Если

$$\liminf_{m,n} \sigma_{mn}^2 > 0$$

(σ_{mn} определено ниже), то для некоторого $\gamma = \gamma(p) < 3$

$$\sup_{x \in \mathbb{R}} \big| \mathbb{P}\{S_{mn} - \mu_{mn} \le x \sigma_{mn}\} - \Phi(x) \big| = o(1), \quad m, n \to \infty, \quad n = O(m^\gamma),$$

$\Phi(x)$ – функция распределения закона $\mathcal{N}(0, 1)$,

$$\mu_{mn} = \frac{1}{m\sqrt{n}} \sum_{k=-m}^{m} \left(1 - \frac{|k|}{m}\right) \sum_{t=m \vee (m-k)}^{(n-1)\wedge(n+k-1)} \left(\mathbb{C}(v_t, u'_{t-k}) + \frac{\mathbb{C}(w'_t, u'_{t-k}) - \mathbb{C}(z_t, u'_{t-k})}{2} \right)$$

$$-\frac{1}{m\sqrt{n}} \sum_{k=0}^{m} \sum_{t=m \vee (m-k)}^{(n-1)\wedge(n+k-1)} \mathbb{C}(\partial f_{t+1}, u'_{t-k}),$$

$$\sigma^2_{mn} = \frac{1}{n}\sum_{s,t=m}^{n-1}\mathbb{C}(f_{t+1}-h_t, f_{s+1}-h_s) + \frac{1}{n}\sum_{t,s=1}^{m-1}\frac{ts}{m^2}\mathbb{C}(h_t,h_s)$$

$$+\frac{1}{n}\sum_{t,s=1}^{m-1}\frac{ts}{m^2}\mathbb{C}(h_{n-t},h_{n-s}) - \frac{1}{n}\sum_{t,s=1}^{m-1}\frac{s}{m}\mathbb{C}(f_{n-t+1}-h_{n-t},h_{n-s}).$$

Если верна H_0, то при $m = o(n)$

$$\limsup_{m,n}\frac{1}{n}\sum_{t=m}^{n-1}\mathbb{E}f_{t+1} = \limsup_{n}\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}f_{t+1} \le 0,$$

и по теореме 1 тест подходящего уровня значимости имеет вид (1). При этом, однако, встает вопрос о том, как состоятельно оценить μ_{mn} и σ_{mn}, т.е. подобрать такие $\hat{\mu}_{mn}$ и $\hat{\sigma}_{mn}$, что

$$\hat{\mu}_{mn} - \mu_{mn} \xrightarrow{p} 0, \quad \hat{\sigma}_{mn} - \sigma_{mn} \xrightarrow{p} 0, \quad m,n \to \infty.$$

Возможна следующая схема оценивания: (a) разбить ряд $\{(x_t,y_t)\}_{t=1}^n$ на непересекающиеся части вида $\{(x_t,y_t)\}_{t=a}^b$, $a<b$, не имеющие структурных сдвигов; (b) оценить на этих частях параметры θ_t, а также все величины $u_t, v_t, ...$, участвующие в μ_{mn} и σ_{mn}; (c) использовать стандартные методы HAC оценивания (см. Andrews, 1991) для построения $\hat{\mu}_{mn}$ и $\hat{\sigma}_{mn}$, предполагая нулевую корреляцию у величин из различных отрезков $[a,b]$. Некоторые подходы к решению задачи из пункта (a) описаны в Perron (2006), Andreou & Ghysels (2009) и Boldeay & Hall (2010).

Замечание 1: Если вместо статистики S_{mn} рассматривать

$$\bar{S}_{mn} = S_{mn} - \frac{1}{\sqrt{n}}\sum_{t\in U_{mn}}(f_{t+1}(\hat{\theta}_t) - \mathbb{E}f_{t+1}),$$

где $U_{mn} \subset [m,n]$ таково, что $|U_{mn}| = o(n)$ и

$$\mathbb{P}\{[\tau,\tau+m-1] \subset U_{mn} \text{ для всех } \tau, \text{ в которые ряд } (x_t,y_t) \text{ меняет структуру}\} \to 0,$$

то условия (A3) и (A6) можно ослабить, не предполагая, что $\mathbb{E}\sup_\theta|f_t(\theta)| < \infty$, и меняя соотношение $l = o(\sqrt{n}/m)$ на $m = o(n)$. При этом результат теоремы 1 остается в силе при замене S_{mn} на $\bar{S}_{mn}$.

Замечание 2: Альтернативный способ вычисления σ^2_{mn} рассмотрен Giacomini & White(2006):

$$\sigma^2_{mn} = \frac{1}{n}\sum_{s,t=m}^{n-1}\mathbb{C}(f_{t+1}(\hat{\theta}_t), f_{s+1}(\hat{\theta}_s)).$$

Таким образом, при построении $\hat{\sigma}^2_{mn}$ допустимо использовать величины $f_{t+1}(\hat{\theta}_t)$. Формально здесь необходимо обосновать возможность применения нормального приближения S_{mn}, фигурирующего в Giacomini & White (2006), наряду с приближением из теоремы 1. Это можно сделать при некоторых ограничениях на величины $f_{t+1}(\hat{\theta}_t)$, параметр m и порядок убывания коэффициентов перемешивания α_n из условия (A7).

5 Компьютерное моделирование

В качестве примера реального применения теоремы 1 был рассмотрен ряд $\{y_t\}_{t=0}^n$, заданный следующим образом:

$$y_0 = \frac{2e_0}{\sqrt{3}}, \quad y_t = \begin{cases} y_{t-1}/2 + e_t, & t \le n/2, \\ 1/\sqrt{3} + e_t, & t > n/2, \end{cases}$$

где $n = 900$, e_t образуют последовательность независимых стандартных гауссовских величин. Нетрудно проверить, что

$$\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}f_t(\theta_{1t},\theta_{2t}) = 0,$$

если $f_t(\theta_1,\theta_2) = (y_t - \theta_1 y_{t-1})^2 - (y_t - \theta_2)^2$, а оптимальные значения параметров θ_{1t}, θ_{2t} для всех моментов t определяются методом наименьших квадратов. В рассматриваемом случае $\mathbb{E}\partial f_{t+1} = 0$. Поэтому

$$\mu_{mn} = \frac{1}{m\sqrt{n}}\sum_{k=-m}^{m}\left(1-\frac{|k|}{m}\right)\sum_{t=m\vee(m-k)}^{(n-1)\wedge(n+k-1)}\frac{\mathbb{E}w_t' u_{t-k}}{2} - \frac{1}{m\sqrt{n}}\sum_{k=0}^{m}\sum_{t=m\vee(m-k)}^{(n-1)\wedge(n+k-1)}\mathbb{E}\partial f_{t+1}u_{t-k},$$

$$\sigma^2_{mn} = \frac{1}{n}\sum_{s,t=m}^{n-1}\mathbb{C}(f_{t+1}, f_{s+1}).$$

Для простоты предполагалось, что момент структурного сдвига $\tau = n/2$ известен, и вместо $\hat{S}_{mn}$ использовалась статистика

$$\bar{S}_{mn} = \hat{S}_{mn} - \frac{1}{\sqrt{n}}\sum_{t=\tau}^{\tau+m-1} f_{t+1}(\hat{\theta}_t).$$

В качестве $\hat{\mu}_{mn}, \hat{\sigma}^2_{mn}$ были взяты HAC-оценки с ядром Парзена. При этом количество лагов оцениваемых ковариаций (для каждого из промежутков $[m, n/2]$ и $[n/2+m, n]$) вычислялось по алгоритму, предложенному в Hirukawa (2010). Кроме того, как и в Giacomini & White (2006), вместо $f_{t+1}(\theta_{t+1})$ при построении $\hat{\sigma}^2_{mn}$ брались $f_{t+1}(\hat{\theta}_t)$. Последняя замена показала лучший результат по сравнению с использованием величин $f_t(\hat{\theta})$, $t \in [m, n/2]$, и $f_t(\tilde{\theta})$, $t \in [n/2+m, n]$, где $\hat{\theta}$, $\tilde{\theta}$ – оценки θ по первой и второй части выборки соответственно.

Результаты 1000 симуляций ряда $\{y_t\}_{t=0}^{n}$ представлены в таблицах ниже.

m	$< -1,96$	$< -1,645$	$> 1,645$	$> 1,96$
α	0,025	0,050	0,050	0,025
20	0,025	0,060	0,058	0,039
30	0,020	0,057	0,051	0,026

Таблица 1: Доля значений $(\bar{S}_{mn} - \hat{\mu}_{mn})/\hat{\sigma}_{mn}$, попавших в критические множества вида $(-\infty, -z_\alpha)$ и (z_α, ∞) для различных α, где z_α – α-квантиль распределения $\mathcal{N}(0,1)$.

m	$< -1,96$	$< -1,645$	$> 1,645$	$> 1,96$
α	0,025	0,050	0,050	0,025
20	0,007	0,024	0,089	0,052
30	0,007	0,028	0,068	0,038

Таблица 2: Доля значений $\bar{S}_{mn}/\hat{\sigma}_{mn}$, попавших в критические множества вида $(-\infty, -z_\alpha)$ и (z_α, ∞) для различных α, где z_α – α-квантиль распределения $\mathcal{N}(0,1)$.

Несмотря на малость истинного значения μ_{mn} (величина порядка десятых), из таблиц видно, что учет асимптотического смещения $\hat{S}_{mn}$ делает процедуру тестирования более точной в конечных выборках. Использование μ_{mn} является отличительной особенностью нашей схемы тестирования по сранению с работами West & McCracken (1998) и Giacomini & White (2006) (см. также Китов, 2009).

6 Заключение

В настоящей работе предложен новый подход к тестированию предсказательной способности для данных с небольшим числом структурных сдвигов. С помощью компьютерных симуляций показана полезность данного подхода. Несмотря на это, остается ряд нерешенных вопросов: выбор длины окна m, построение эффективных оценок $\hat{\mu}_{mn}$ и $\hat{\sigma}_{mn}$, подходящее разбиение данных на непересекающиеся части без структурных сдвигов. Все эти вопросы – темы будущих исследований.

Благодарности

Автор благодарен своему научному руководителю профессору С.А. Анатольеву за постановку задачи и плодотворное обсуждение работы. Хотелось бы также выразить признательность М. Маккракену за конструктивные замечания, способствовавшие улучшению статьи.

Список литературы

Китов, В. (2009). Парные тесты на одинаковую точность прогнозов. *Квантиль* 6, 77–91.

Яськов, П.А. (2010). Некоторые оценки норм дискретных стохастических интегралов. *ДАН: Математика* 432, 322–325.

Andreou E. & E. Ghysels (2009). Structural breaks in financial time series. Глава в *Handbook of Financial Time Series*, под редакцией T.G. Andersen, R.A. Davis, J.-P. Kreiss & T. Mikosch, Springer, Berlin.

Andrews, D.W.K. (1991). Heteroscedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

Boldeay, O. & A.R. Hall (2010). Estimation and inference in unstable nonlinear least squares models. Препринт, Univeristy of Manchester.

Clark, T.E. & M.W. McCracken (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of econometrics* 105, 85–110.

Corradi, V. & N.R. Swanson (2005). Predictive density evaluation. Глава в *Handbook of Economic Forecasting*, под редакцией C.W.J. Granger, G. Elliot & A. Timmerman, Elsevier.

Dedecker J. & C. Prieur (2004). Coupling for τ-dependent sequences and applications. *Journal of Theoretical Probability* 17, 861–885.

De Jong R.M. & J. Davidson (2000). The functional central limit theorem and weak convergence to stochastic integrals I. *Econometric Theory* 16, 621–642.

Giacomini, R. & B. Rossi (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics* 25, 595–620.

Giacomini, R. & H. White (2006). Tests for conditional predictive ability. *Econometrica* 74, 1545–1578.

Hirukawa M. (2010). A two-stage plug-in bandwidth selection and its implementation for covariance estimation. *Econometric Theory* 26, 710–743.

Perron, P. (2006). Dealing with structural breaks. Глава в *Palgrave Handbook of Econometrics 1. Econometric Theory*, под редакцией T.C. Mills & K. Patterson, Palgrave Macmillan.

Sunklodas, J. (2008). On normal approximation for strongly mixing random fields. *Theory of Probability and its Applications* 52, 125–132.

Sung, S.H., S. Lisawadi & A.Volodin (2008). Weak laws of large numbers under condition of uniform integrability. *Journal of Korean Mathematical Society* 45, 289–300.

West, K.D. (1996). Asymptotic inference about predictive ability. *Econometrica* 64, 1067–1084.

West, K.D. & M. McCracken (1998). Regression-based tests of predictive ability. *International Economic Review* 39, 817–40.

White, H. (2000). A reality check for data snooping, *Econometrica* 68, 1097–1126.

Приложение

Доказательство теоремы 1 основано на предложении 1, некоторых результатах о точности нормальной аппроксимации сумм слабозависимых величин, а также законе больших чисел для дискретных стохастических интегралов (см. ниже).

Предложение 1: Пусть выполнены условия (A1)–(A5). Тогда для любого $\varepsilon > 0$

$$\mathbb{P}\{|T_{mn}(\varepsilon)| = o\big(n^{1/2}\big)\} = 1 + o(1), \quad T_{mn}(\varepsilon) = \big\{t \in [m, n) : ||\hat{\theta}_t - \theta_{t+1}|| > \varepsilon\big\}.$$

На промежутке стационарности ряда $\{(x_t, y_t)\}_{t=1}^{\infty}$ имеем $\theta_t = (\theta'_{1t}, \theta'_{2t})' \equiv (\theta'_1, \theta'_2)' = \theta$ (величина θ своя для каждого промежутка стационарности и, как и выше, $\theta_i \in \mathbb{R}$). Обозначим

$$H_{it}(\theta) = \frac{1}{m} \sum_{s=t-m+1}^{t} \partial^i \phi_s(\theta).$$

Далее будем опускать зависимость от θ, т.е. писать g, подразумевая $g(\theta)$. Условия первого порядка экстремальной задачи, из которой определяются $\hat{\theta}_{it}$, влекут

$$\hat{\theta}_t - \theta = -\Big(H_{2t} + (\hat{\theta}_t - \theta)^* H_{3t}(\bar{\theta}_t)/2\Big)^{-1} H_{1t},$$

где координаты $\bar{\theta}_t$ лежат между соответствующими координатами θ и $\hat{\theta}_t$. Кроме того, имеет место разложение

$$\begin{aligned} f_{t+1}(\hat{\theta}_t) - \mathbb{E}f_{t+1} &= f_{t+1} - \mathbb{E}f_{t+1} + \partial f_{t+1}(\hat{\theta}_t - \theta) + \frac{1}{2}(\hat{\theta}_t - \theta)'\partial^2 f_{t+1}(\tilde{\theta}_t)(\hat{\theta}_t - \theta) \\ &= Z_t + I_t + R_t, \end{aligned}$$

где

$$Z_t = f_{t+1} - \mathbb{E}f_{t+1} - \mathbb{E}\partial f_{t+1}(\mathbb{E}\partial^2 \phi_t)^{-1} H_{1t},$$

$$\begin{aligned} I_t = {} & \mathbb{E}\partial f_{t+1}(\mathbb{E}\partial^2\phi_t)^{-1}(H_{2t} - \mathbb{E}\partial^2\phi_t)(\mathbb{E}\partial^2\phi_t)^{-1} H_{1t} - (\partial f_{t+1} - \mathbb{E}\partial f_{t+1})(\mathbb{E}\partial^2\phi_t)^{-1} H_{1t} \\ & + \frac{1}{2} H'_{1t}(\mathbb{E}\partial^2\phi_t)^{-1}\mathbb{E}\partial^2 f_{t+1}(\mathbb{E}\partial^2\phi_t)^{-1} H_{1t} \\ & - \frac{1}{2}\mathbb{E}\partial f_{t+1}(\mathbb{E}\partial^2\phi_t)^{-1}\big((\mathbb{E}\partial^2\phi_t)^{-1} H_{1t}\big)^* \,\mathbb{E}\partial^3\phi_t(\mathbb{E}\partial^2\phi_t)^{-1} H_{1t}, \end{aligned}$$

а остаток R_t состоит из всех слагаемых, не вошедших в Z_t и I_t.

Согласно теореме 4 статьи Sunklodas (2007), при указанных условиях

$$\sup_f |\mathbb{E}f\big(Z_{mn}/\Sigma_{mn}\big) - \mathbb{E}f(\xi)| = o(1), \tag{2}$$

где ξ распределено по закону $\mathcal{N}(0,1)$, верхняя грань берется по всем f с

$$|f(x)| \le 1, \quad \frac{|f(x) - f(y)|}{|x - y|} \le 1, \quad x, y \in \mathbb{R},$$

$$Z_{mn} = \frac{1}{\sqrt{n}} \sum_{t=m}^{n-1} Z_t, \quad \Sigma^2_{mn} = \mathbb{E}Z^2_{mn}.$$

С помощью техники сглаживания из (2) легко получается оценка вида

$$\sup_{x\in\mathbb{R}} |\mathbb{P}\big(Z_{mn} \le x\Sigma_{mn}\big) - \mathbb{P}(\xi \le x)| = o(1).$$

Теперь главную роль при доказательстве теоремы 1 играет приводимый ниже закон больших чисел (ЗБЧ) для дискретных стохастических интегралов. Данный ЗБЧ продолжает ряд предельных теорем о суммах вида $\sum_{i\le j} u_i v_j$, широко используемых в эконометрической литературе (см., например, De Jong & Davidson, 2000).

Для формулировки ЗБЧ нам понадобится определение τ-коэффициентов зависимости, обобщающих понятие сильного перемешивания. Следуя статье Dedecker & Prieur (2004), для интегрируемой случайной величины ξ и σ-алгебры $\mathcal{G}$ положим

$$\tau(\mathcal{G},\xi) = \sup_f \mathbb{E}\big|\mathbb{E}\big[f(\xi)|\mathcal{G}\big] - \mathbb{E}f(\xi)\big|,$$

где верхняя грань берется по всем функциям f с константой Липшица, не большей 1. Связь между $\tau(\mathcal{G},\xi)$ и коэффициентом сильного перемешивания $\alpha(\mathcal{G},\xi)$ иллюстрируется неравенством

$$\tau(\mathcal{G},\xi) \le \|\xi\|_p[\alpha(\mathcal{G},\xi)]^{1/q},$$

здесь $\|\xi\|_p = \big(\mathbb{E}|\xi|^p\big)^{1/p}$, $p, q > 1$ и $1/p + 1/q = 1$.

Определим дискретный стохастический интеграл как

$$I = \int_T U dV = \sum_{i=1}^{|T|-1} U(t_i)[V(t_{i+1}) - V(t_i)], \quad t_1 < t_2 < \ldots < t_{|T|}, \quad t_i \in T.$$

Предположим, что для некоторого $p > 1$ процессы U, V удовлетворяют неравенству

$$\|[U(z) - U(x)]\cdot[V(y) - V(z)]\|_p \le f(x,y), \quad z \in [x,y], \ \ x,y,z \in T,$$

где функция $f : \mathbb{R}^2 \to \mathbb{R}_+$ такова, что $f(x,x) = 0$, $f(x,y)$ симметрична, не убывает по y при $y \ge x$ и не возрастает по x при $x \le y$. Следуя работе Яськов (2010), введем интегральную характеристику вида

$$H_g(a,b) = \iint_{[a,b]^2} \frac{g(x,y)}{(x-y)^2}\, dxdy + \int_a^b \frac{g(a,y)}{y-a}\, dy + \int_a^b \frac{g(x,b)}{b-x}\, dx + g(a,b).$$

Положим $F(a,b) = H_{f\mathbf{1}_D}(a,b)$, где $\mathbf{1}_D$ - индикатор множества D, состоящего из всех (x,y) в $\mathbb{R}^2$, для которых отрезок с концами x, y содержит более одной точки T.

Пусть далее

$$\Delta_{ij} = \int_{T\cap[t_i,t_j]} U dV, \quad 1 \le i < j \le |T|, \quad \|\xi\|_p = \big(\mathbb{E}|\xi|^p\big)^{1/p}.$$

Фиксируем σ-алгебры $\mathcal{F}_k = \sigma\big(\Delta_{ij},\ i < j \le k\big)$ и для $n < |T|$ зададим

$$\tau_n = \sup\left\{\tau\left(\mathcal{F}_k, \frac{\Delta_{ij}}{\|\Delta_{ij}\|_p}\right) : \ j > i \ge k + n\right\},$$

$$\Gamma_n = \max\Big\{\|U(t_i)\cdot[V(t_j) - V(t_i)]\|_p : \ j - i \in (0,n]\Big\}.$$

Теорема 2: Для последовательности интегралов $I = I_N = \int_{T_N} U_N dV_N$ и всех $q \in [1,p)$ имеем

$$\mathbb{E}|I - \mathbb{E}I|^q = o(B), \quad n, N \to \infty,$$

где коэффициенты $B = B(n,N)$ таковы, что

$$(n^{q-1}\tau_n + 1)\sum_i F\big(t_{in+1}, t_{(i+1)n\wedge|T|}\big)^q + \frac{N\Gamma_n}{n} = O(B),$$

$$\max_i F\big(t_{in+1}, t_{i(n+1)\wedge|T|}\big)^q = o(B),$$

здесь суммирование и максимизация ведутся по всем i, для которых определены соответствующие точки t, а также для простоты записи опущен индекс N у F, T, τ_n, Γ_n и t.

Теорема 2 доказывается с использованием моментных неравенств статьи Яськов (2010), а также элементарного предложения, вытекающего из результатов работы Sung, Lisawadi & Volodin (2008).

Предложение 2: Допустим, что при всех натуральных n последовательность $\{X_{ni}\}$ есть мартингал-разность, здесь и далее i изменяется от 1 до k_n. Пусть также

$$\max_i \|X_{ni}\|_p^q = o\big(b_n\big), \quad \sum_i \|X_{ni}\|_p^q = O(b_n), \quad n \to \infty,$$

для некоторых констант $p > q \geq 1$ и $b_n \uparrow \infty$. Тогда

$$\mathbb{E}\Big|\sum_i X_{ni}\Big|^q = o(b_n).$$

В настоящей работе мы используем следующий частный случай теоремы 2.

Следствие 1: Пусть в условиях теоремы 2

$$f(x,y) = |x-y|^\lambda, \;\; \lambda > 1, \quad \tau_n = O(n^{-\gamma}), \;\; \gamma > 0, \quad t_i = i, \quad |T_N| = N, \quad \Gamma_n = O(n^{1/2}).$$

Тогда при $\delta = \min\{2+\gamma-q(1+\lambda), 1-q\lambda, -1/2\}$ имеем

$$\mathbb{E}|I - \mathbb{E}I|^q = o(Nn^{-\delta}), \quad n, N \to \infty.$$

Следствие 1 при подходящем выборе процессов U, V и множеств T (см. теорему 2) позволяет установить сходимость вида

$$I_{mn} = \frac{1}{\sqrt{n}} \sum_{t=m}^{n-1} (I_t - \mathbb{E}I_t + R_t) \xrightarrow{p} 0.$$

Это с учетом некоторых технических рассуждений завершает доказательство теоремы 1.

Testing for predictive ability in the presence of structural breaks

Pavel Yaskov

Moscow State University, Moscow, Russia

We propose a new approach to testing for predictive ability in the presence of structural breaks in data. Our approach extends the well-known results of West (1996) and West & McCracken (1998), and is alternative to methods developed in Giacomini & White (2006) and Giacomini & Rossi (2010).

Keywords: testing for predictive ability, forecast evaluation, structural change

JEL Classification: C22, C52, C53