УДК 51-77, 330.4

НЕПАРАМЕТРИЧЕСКАЯ ОЦЕНКА ВОЛАТИЛЬНОСТИ И ЕЕ ПАРАМЕТРИЧЕСКИЕ АНАЛОГИ

А.В. Добровидов, В.Э. Тевосян

Предложен метод непараметрической оценки стохастической волатильности и его сравнение с другими широко используемыми в эконометрике алгоритмами. Основным преимуществом данного подхода является возможность получения оценки волатильности в случае, когда ее распределение вероятностей полностью неизвестно. Показано, что разработанный метод имеет лучшие характеристики по сравнению с известными параметрическими алгоритмами, построенными на основе модели GARCH и фильтра Калмана.

Ключевые слова: стохастическая волатильность, непараметрическая оценка сигналов, фильтр Калмана; GARCH, модель Тейлора.

ВВЕДЕНИЕ

Прежде чем перейти к описанию моделей стохастической волатильности, рассмотрим несколько подробнее само понятие. Под словом *волатильность* могут пониматься различные финансовые явления. В зависимости от контекста и точки зрения волатильность может иметь смысл явления изменчивости цен в виде стандартного отклонения доходностей, параметра финансового риска или ценовой модели либо случайного процесса определенного вида. Появившаяся в 1973 г. модель оценки опционов Блэка — Шоулза — Мертона [1] изменила представление о волатильности. Блэк и Шоулз в своей работе применили непрерывное геометрическое броуновское движение для построения математической модели цен. Успех данной работы привел к широкому распространению стохастических моделей в эконометрике. В данной работе под термином волатильность будем понимать случайный процесс, описывающий изменчивость цен.

Анализ доходности американских инвестиционных фондов показал, что основной характеристикой, отличающей успешные фонды, является умение управлять волатильностью портфелей таким образом, чтобы в периоды высокой изменчивости цен акций снижать долю наиболее волатильных активов в портфеле [2]. Также стоит заметить, что основанные на динамическом измерении волатильности активные инвестиционные стратегии приносят значимую выгоду [3, 4]. Все эти выводы говорят о том, что знание волатильности имеет принципиальное значение для эффективного контроля портфелем акций, построения инвестиционных стратегий и управления рисками компаний наряду с другими вопросами.

В последнее время начали появляться работы по непараметрической оценке волатильности. Примерами могут служить прогноз волатильности методом локального экспонентного оценивания для модели Яо и Тонга [5], предложенный Зигельманном [6], или оценка параметров стохастических моделей Дюффе и др. [7] и Жакье и др. [8] через инфинитезимальные моменты в статье Банди и Рено [9]. В отличие от этих работ настоящая статья опирается на теорию принятия решений Вальда, а именно на наиболее распространенном в статистике критерии минимизации функции риска. Авторами рассматривается модель волатильности в виде ненаблюдаемого стохастического процесса с неизвестными распределением и уравнением состояния, который (при определенных условиях) можно оценивать на основе эмпирического байесовского подхода, развитого применительно к задачам теории непараметрической фильтрации, интерполяции и прогноза частично наблюдаемых марковских процессов [10]. В этой теории стохастические модели состояния полезных сигналов предполагаются неизвестными, а модели наблюдения, описывающие приборы исследователя, предполагаются полностью известными. Если полезный сигнал в чистом виде не наблюдаем, то в общем случае восстановить его распределение невозможно. Следовательно, напрямую воспользоваться оптимальной байесовской процедурой для восстановления полезного сигнала по наблюдениям не удается. Основное преимущество данного подхода по сравнению с вышеуказанными параметрическими оценками заключается в возможности рассмотрения процесса волатильности любой частоты без предположений о модели ненаблюдаемой составляющей.

1. МОДЕЛИ ОПИСАНИЯ СТОХАСТИЧЕСКОЙ ВОЛАТИЛЬНОСТИ

Цена актива на финансовом рынке S_t в момент времени t, согласно определению, положительна и обычно рассматривается в представлении $S_t = S_0 e^{R_t}$, где $R_t = \sum_{i=1}^{t} r_i$, $t = 0, 1, 2, ...$ Величина r_t называется логарифмическим доходом в момент времени t и определяется как

$$r_i = \log\frac{S_t}{S_{t-1}}.$$

Простейшее предположение о вероятностном характере величины r_t состоит в том, что ее модель описывается стохастическим уравнением $r_t = \sigma\varepsilon_t$, где $\varepsilon = (\varepsilon_t)$ — последовательность независимых одинаково распределенных $\mathcal{N}(0, 1)$ случайных величин, а σ — параметр, который в финансовой литературе принято называть волатильностью.

Дальнейшие исследования показали, что волатильность сама по себе «волатильна» и ее поведение хорошо описывается случайной функцией времени. Волатильность, как характеристика изменчивости, сама по себе не наблюдается, но может быть представлена случайным процессом (σ_t) в модели наблюдаемой доходности

$$r_t = \mu_t + \sigma_t\varepsilon_t, \tag{1}$$

где μ_t и σ_t — заданные случайные процессы. В настоящей работе рассматриваются две параметрические условно-гауссовские модели волатильности в дискретном времени, определенные в рамках концепции эффективного рынка: Generalized AutoRegressive Conditional Heteroscedasticity (GARCH) [11], предложенная Боллерслевом, и созданная Тейлором Stochastic Volatility (SV) [12]. Первая из них наиболее популярна на практике, так как не только описывает саму волатильность, но и предсказывает ее значение на шаг вперед. Однако в этой модели волатильность σ_t является заданной функцией от наблюдаемой доходности r_t, что несколько ограничивает возможности ее изменения. Особенность второй модели состоит в том, что волатильность σ_t в ней представлена в виде отдельного независимого от дохода r_t стохастического процесса [13]. Такое предположение позволяет существенно расширить разнообразие поведения волатильности и применить для ее описания богатые и хорошо развитые методы теории случайных процессов и теории мартингалов [14]. Далее в п. 1.1 и 1.2 эти две модели рассматриваются подробнее.

1.1. Модель GARCH

Предположим, что последовательность $\varepsilon = (\varepsilon_t)_{t\geq 1}$ в модели (1) является единственным источником случайности на рынке, а условное среднее и условная дисперсия равны соответственно

$$\mu_t = \mathsf{E}(r_t|r_{t-1}, ..., r_1) = 0,$$

$$\sigma_t^2 = \mathsf{E}(r_t^2|r_{t-1}, ..., r_1) = \alpha_0 + \sum_{i=1}^{p}\alpha_i r_{t-i}^2, \tag{2}$$

где $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, ..., p$ и $r_{1-p}, ..., r_0$ — заданные начальные константы. Условное среднее μ_t приравнивается нулю, потому что среднесуточная доходность и внутридневные доходности акций и валют в среднем равны нулю [15].

Такая модель была введена в 1982 г. Робертом Энглем и называется ARCH(p) (AutoRegressive Conditional Heteroskedastic Model — авторегрессионная модель условной неоднородности). Она оказалась весьма удачной при объяснении ряда нетривиальных свойств временных финансовых рядов таких, как эффект кластеризации значений величин r_t.

Определение 1. Модель ARCH(p) — это последовательность $r = (r_t)$ вида $r_t = \sigma_t\varepsilon_t$, где $\varepsilon = (\varepsilon_t)$ — последовательность независимых нормально распределенных случайных величин $\varepsilon_t \sim \mathcal{N}(0, 1)$, а σ_t^2 определяется из рекуррентного уравнения (2). ♦

Успех модели ARCH(p) привел к появлению различных ее обобщений и модификаций. Одной из таких разновидностей является введенная в 1986 г. Тимом Боллерслевом модель GARCH (Generalized AutoRegressive Conditional Heteroskedastic Model — обобщенная авторегрессионная модель условной неоднородности). Преимущество модели GARCH (p, q) по сравнению с предыдущей моделью состоит в том, что для подготовки ARCH к реальным данным требуется брать большие значения p, а обобщенная модель позволяет ограничиться небольшими значениями p и q.

Пусть по-прежнему считается, что $\mu_t = 0$, но в отличие от выражения (2) σ_t описывается уравнением

$$\sigma_t^2 = (r_t^2|r_{t-1}, ..., r_1) = \\ = \alpha_0 + \sum_{i=1}^{p}\alpha_i r_{t-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2, \tag{3}$$

с $\alpha_0 > 0$, α_i, $\beta_i \geq 0$ и начальными условиями (r_{1-p}, ..., r_0), (σ^2_{1-q}, ..., σ^2_0), которые могут быть константами либо сгенерированы заранее.

Определение 2. Модель GARCH(p, q) — это последовательность $r = (r_t)$, $r_t = \sigma_t\varepsilon_t$, где $\varepsilon = (\varepsilon_t)$ — последовательность независимых одинаково распределенных $\mathcal{N}(0, 1)$ случайных величин, а σ_t подчиняется соотношению (3).

1.2. Модель Тейлора

В моделях типа ARCH существует только один источник шума, который задается гауссовской последовательностью независимых случайных величин $\varepsilon = (\varepsilon_t)_{t \geq 1}$. При этом дисперсия процесса доходности предполагается в той или иной форме зависимой от прошлых его реализаций. Альтернативный способ моделирования состоит в том, чтобы задать динамику цен простой моделью, например, дифференциальным уравнением, как это сделано в модели Блэка — Шоулза, но волатильность в нем считать не параметром, а отдельным стохастическим процессом. Таким образом, появляются два независимых источника случайности. Модели стохастической волатильности включают в себя еще один источник шума $\delta = (\delta_t)_{t \geq 1}$, который в простейшем случае также состоит из независимых $\mathcal{N}(0, 1)$-распределенных случайных величин.

Пусть, как и в предыдущих моделях, выполняется первое из условий (2). Тогда

$$r_t = \sigma_t\varepsilon_t, \tag{4}$$

где ε_t являются $\mathcal{N}(0, 1)$-распределенными случайными величинами.

Положим,

$$\sigma_t = e^{\frac{1}{2}\Delta_t}, \tag{5}$$

отсюда следует, что σ_t — положительная случайная величина. Наибольшей популярностью пользуются модели зависимых случайных величин, в которых последовательность $(\Delta_t)_{t \geq 1}$ представлена в виде авторегрессионной модели

$$\Delta_t = a_0 + a_1\Delta_{t-1} + ... + a_p\Delta_{t-p} + c\delta_t, \tag{6}$$

где c — дисперсия шумовой составляющей. Из выражений (4)—(6) следует, что условное распределение r_t при фиксированных предыдущих значениях r_{t-1}, ..., r_1 является гауссовским с параметрами 0 и σ^2_t:

$$f(r_t|r_{t-1}, ..., r_1) = \mathcal{N}(0, \sigma^2_t).$$

Определение 3. Модель стохастической волатильности — это последовательность $r = (r_t)$ вида $r_t = \sigma_t\varepsilon_t$, где $\varepsilon = (\varepsilon_t)$ — последовательность независимых одинаково распределенных $N(0, 1)$ случайных величин, а σ_t определяется уравнением авторегрессии Δ_t, возбуждаемой «белым» $\mathcal{N}(0, 1)$ шумом.

2. ФИЛЬТР КАЛМАНА

Одним из широко применяемых в современной эконометрике способов решения задачи оценки волатильности является фильтр Калмана [16, 17]. Рассмотрим его применительно к задаче оценки стохастической волатильности. Данный метод оценивания применим только к линейным гауссовским моделям, поэтому необходимо привести уравнение (4) к линейному виду. Воспользовавшись формулой (5), получим

$$r_t = \sigma_t\varepsilon_t = e^{\frac{1}{2}\Delta_t}\varepsilon_t.$$

Логарифмируя квадрат этого выражения, получим

$$\log(r^2_t) = \log(e^{\Delta_t}\varepsilon^2_t) = \log(e^{\Delta_t}) + \log(\varepsilon^2_t) =$$
$$= \Delta_t + \log(\varepsilon^2_t).$$

Обозначим $b_t = \log(r^2_t)$, а $\xi_t = \log(\varepsilon^2_t)$. Тогда уравнение (4) представляется в виде

$$b_t = \Delta_t + \xi_t, \tag{7}$$

Для простоты уравнение (6) будет представлено в виде авторегрессии первого порядка с нулевым математическим ожиданием:

$$\Delta_t = a_1\Delta_{t-1} + c\delta_t. \tag{8}$$

Одно из обязательных условий применения фильтра Калмана состоит в гауссовости шумовых составляющих, но случайная величина $\xi_t = \log(\varepsilon^2_t)$ в модели состояния не является гауссовской, поэтому ее заменяют на гауссовскую ξ^g_t со средним $E\xi_t = E\xi^g_t = -1{,}27$ и дисперсией $D\xi_t = D\xi^g_t = 4{,}93 = \pi^2/2$ [18] (вывод этих значений приведен в Приложении). Тогда система (7), (8) заменяется на систему

$$\begin{cases} b_t = \Delta_t + \xi^g_t, \\ \Delta_t = a_1\Delta_{t-1} + c\delta_t. \end{cases}$$

В итоге фильтр Калмана для модели стохастической волатильности будет описываться системой уравнений:

$$\hat{\Delta}_t = a_1\hat{\Delta}_{t-1} + \\ + \frac{2c^2 + 2a_1^2\gamma_{t-1}}{\pi^2 + 2c^2 + 2a_1^2\gamma_{t-1}}(b_n + 1{,}27 - a_1\hat{\Delta}_{t-1}), \quad (9)$$

$$\gamma_t = \frac{\pi^2(c^2 + a_1^2\gamma_{t-1})}{\pi^2 + 2c^2 + 2a_1^2\gamma_{t-1}}$$

с начальными условиями

$$\hat{\Delta}_1 = \frac{2c^2}{\pi^2(1 - a_1^2) + 2c^2}b_1,$$

$$\gamma_1 = \frac{\pi^2c^2}{\pi^2(1 - a_1^2) + 2c^2}.$$

В данной системе оценка $\hat{\Delta}_t$ является рекуррентной, что удобно для вычислений, но замена на гауссовский шум может приводить к дополнительным ошибкам.

Полученные из системы уравнений (9) значения подставляются в формулу (5) для получения оценки самой волатильности:

$$\hat{\sigma}_t = \exp\left(a_1\hat{\Delta}_{t-1} + \frac{2c^2 + 2a_1^2\gamma_{t-1}}{\pi^2 + 2c^2 + 2a_1^2\gamma_{t-1}} \times \right. \\ \left. \times (b_n + 1{,}27 - a_1\hat{\Delta}_{t-1})\right).$$

3. УРАВНЕНИЕ ОПТИМАЛЬНОЙ ФИЛЬТРАЦИИ ПРИ НЕИЗВЕСТНОМ РАСПРЕДЕЛЕНИИ ПОЛЕЗНОГО СИГНАЛА

В 1960—1970-х гг. предпринималось достаточно много усилий для разработки строгих методов обучения и самообучения автоматических систем. Одно из направлений работы было нацелено на решение задач самообучения при статистической обработке сигналов с неизвестными моделями состояния. Оно привело к созданию теории непараметрического оценивания сигналов, одна из первых работ которой опубликована в 1983 г. [19]. В этой теории стохастические модели состояния полезных сигналов (S_t) предполагаются неизвестными, а модели наблюдения, описывающие измерительные приборы, предполагаются полностью известными. Если полезный сигнал в чистом виде не наблюдаем, то в общем случае восстановить его распределение невозможно. Следовательно, напрямую воспользоваться оптимальной байесовской процедурой для восстановления полезного сигнала по наблюдениям не удается. В параметрических моделях делается предложение о математических моделях ненаблюдаемого сигнала. К примеру, в моделях GARCH и стохастической волатильности вводятся уравнения состояния (3) и (6) соответственно. Основное преимущество непараметрического метода оценивания заключается в возможности получения оценки ненаблюдаемого стохастического процесса в случае, когда его распределение полностью неизвестно, т. е. решается задача фильтрации ненаблюдаемого сигнала без введения моделей описания процесса и предположений о его природе.

Пусть $(X_t, S_t)_{t \geq 1}$ — частично наблюдаемый двухкомпонентный процесс, где $(X_t \in R)_{t\geq 1}$ и $(S_t \in R)_{t\geq 1}$ — наблюдаемые и ненаблюдаемые компоненты, статистически связанные условной плотностью распределения $f(x_t|s_t) = f(x_t|S_t = s_t)$. Задать условную плотность, значит задать два объекта: распределение шумов η_t; модель наблюдения, т. е. уравнение связи наблюдения X_t с полезным сигналом S_t и шумом η_t, $X_t = \phi(S_t, \eta_t)$. В зависимости от формы модели наблюдения и распределения шумов получаются различные классы задач фильтрации. Рассмотрим класс задач, в которых статистическое уравнение связи и распределение шумов подобраны так, что условная плотность $f(x_t|s_t)$ принадлежит экспонентному семейству:

$$f(x_t|s_t) = \hat{C}(s_t)g(x_t)\exp\{T(x_t)Q(s_t)\} \quad (10)$$

при фиксированном полезном сигнале $S_t = s_t$. Здесь T, Q и g — заданные борелевские функции, а $\hat{C}(s_t)$ — нормирующий сомножитель. Параметрическое семейство (10) содержит, прежде всего, гауссовскую плотность распределения, единственную, для которой известны оптимальные среднеквадратические оценки в явном виде. Кроме того, туда входят χ^2-распределение, бета-распределение, часть системы распределений Пирсона и др. Если случайная величина S_t имеет вероятностный смысл, то такую параметризацию семейства называют *естественной*. Семейство распределений (10) допускает и более простую, так называемую *каноническую* параметризацию, когда в качестве параметра берется $\theta_t = Q(S_t)$, линейно входящее в выражение под экспонентой. В этом случае

$$f(x_t|\theta_t) = C(\theta_t)g(x_t)\exp\{T(x_t)\theta_t\} \quad (11)$$

и нормирующий сомножитель определяется выражением

$$C^{-1}(\theta_t) = \int_R g(x_t)\exp\{T(x_t)\theta_t\}dx_t. \qquad (12)$$

При этом множество значений θ_t, для которых существует экспонентная плотность $f(x_t|\theta_t)$, ограничивается множеством тех θ_t, для которых интеграл (12) конечен.

Основная идея построения алгоритма оценивания в этом случае состоит в том, чтобы вместо процесса S_t оценить процесс $(\theta_t = Q(S_t))_{t \geq 1}$, порождающий более простую каноническую параметризацию в семействе распределений (11), а затем получить оценку $\hat{S}_t$ в виде $Q^{-1}(\hat{\theta}_t)$ с помощью обратной функции Q^{-1}, существование которой обязательно предполагается. Оценка $\hat{S}_t = Q^{-1}(\hat{\theta}_t)$ также будет оптимальной, но с функцией потерь $(Q(\theta_t) - Q(\hat{\theta}_t))^2$.

Проблема оценивания процесса θ_t будет решена в рамках предположения, что ненаблюдаемая последовательность $(S_t)_{t \geq 1}$ является марковской. Отсюда следует, что двухкомпонентный процесс $(Z_t)_{t \geq 1} = (S_t, X_t)_{t \geq 1}$ также является марковским.

При полностью известной статистической информации для оценки θ_t используется оптимальная байесовская оценка в виде условного среднего

$$\hat{\theta}_t = \int_{\Theta_t} \theta_t w_t(\theta_t|x_1^t)d\theta_t, \qquad (13)$$

где Θ_t — множество значений процесса θ_t, а w_t — апостериорная плотность, удовлетворяющая рекуррентному уравнению

$$w_t(\theta_t|x_1^t) = \frac{f(x_t|\theta_t)}{f(x_t|x_1^{t-1})}w_t(\theta_t|x_1^{t-1}), \quad t \geq 2, \qquad (14)$$

где $w_t(\theta_t|x_1^{t-1}) = \int_{\Theta_{t-1}} \tilde{p}(\theta_t|\theta_{t-1})w_{t-1}(\theta_{t-1}|x_1^{t-1})d\theta_{t-1}$ — прогнозная апостериорная плотность, а — $\tilde{p}(\theta_t|\theta_{t-1})$ переходная плотность вероятности марковской последовательности $(\theta_t)_{t \geq 1}$.

Используя байесовскую формулу, найдем начальное условие для уравнения (14):

$$w_1(\theta_1|x_1) = \frac{f(x_1|\theta_1)\tilde{p}(\theta_1)}{\int_{\Theta_t} f(x_1|\theta_1)\tilde{p}(\theta_1)d\theta_1}. \qquad (15)$$

В случае полной статистической информации для получения оптимальной апостериорной оценки интегралы в уравнениях (14) и (15) заменяются на суммы. Таким образом получается фильтр Кушнера — Стратоновича [20, 21].

Теперь проинтегрируем по θ_t обе части уравнения (14) и перенесем влево сомножитель, который зависит только от наблюдений:

$$f(x_t|x_1^{t-1}) = \int_{\Theta_t} f(x_t|\theta_t)w_t(\theta_t|x_1^{t-1})d\theta_t. \qquad (16)$$

Продифференцируем полученное выражение по x_t:

$$f'_{x_t}(x_t|x_1^{t-1}) = \int_{\Theta_t} f'_{x_t}(x_t|\theta_t)w_t(\theta_t|x_1^{t-1})d\theta_t. \qquad (17)$$

Как упоминалось ранее, мы предполагаем, что условная плотность распределения $f(x_t|\theta_t)$ принадлежит экспонентному семейству. Тогда

$$\left(\frac{f(x_t|x_1^{t-1})}{g(x_t)}\right)'_{x_t} = \frac{\partial}{\partial x_t}\int_{\Theta_t} C(\theta_t)\exp(T(x_t)\theta_t)w_t(\theta_t|x_1^{t-1})d\theta_t =$$
$$= \int_{\Theta_t} T'_{x_t}(x_t)\theta_t C(\theta_t)\exp(T(x_t)\theta_t)w_t(\theta_t|x_1^{t-1})d\theta_t, \qquad (18)$$

где штрих означает дифференцирование по x_t.

Умножим левую и правую части уравнения (18) на $g(x_t)/f(x_t|x_1^{t-1})$:

$$\frac{\partial}{\partial x_t}\left(\ln\frac{f(x_t|x_1^{t-1})}{g(x_t)}\right) =$$
$$= \frac{T'_{x_t}(x_t)}{f(x_t|x_1^{t-1})}\int_{\Theta_t} \theta_t C(\theta_t)g(x_t)\exp(T(x_t)\theta_t)w_t(\theta_t|x_1^{t-1})d\theta_t.$$

Воспользовавшись определением байесовской оценки, получим уравнение оптимальной фильтрации относительно оценки $\hat{\theta}_t$:

$$T'_{x_t}(x_t)\hat{\theta}_t = \frac{\partial}{\partial x_t}\left(\ln\frac{f(x_t|x_1^{t-1})}{g(x_t)}\right). \qquad (19)$$

Это уравнение не зависит в явном виде от априорной и переходной плотностей распределения ненаблюдаемой последовательности S_t, а значит, с его помощью можно получить оптимальную оценку $\hat{\theta}_t$ по значениям наблюдаемой компоненты X_t без явного задания переходной плотности.

Единственной неизвестной характеристикой в уравнении оптимальной фильтрации, полученном для класса моделей (11), является условная плотность $f(x_t|x_1^{t-1})$ наблюдения x_n при фиксированных предыдущих наблюдениях $x_1^{t-1} = (x_1, x_2, ..., x_{t-1})$. Поскольку (X_t) — наблюдаемый процесс, то по его реализации эту характеристику можно восстановить непараметрическими ядерными методами. При этом знания функциональной формы условной плотности $f(x_t|x_1^{t-1})$ не требуется. Для применения техники ядерного оценивания необходимо лишь задать два параметра: коэффициент размытости и параметр регуляризации. В данной работе для этих целей применяется метод сглаженной кросс-валидации [22] и регуляризация Тихонова [23] соответственно. Имея в распоряжении методы оценивания коэффициентов размытости и параметров регуляризации вместе с непараметрическими алгоритмами фильтрации, мы получаем рабочий инструмент для решения задач выделения неизвестных полезных сигналов на фоне помех, который зависит только от наблюдаемой выборки. Именно поэтому непараметрические методы выделения сигналов можно назвать автоматическими или самообучающимися.

При неизвестном уравнении состояния условная плотность наблюдений $f(x_t|x_1^{t-1})$ не может быть вычислена точно. Однако ее можно приближенно вычислить при помощи непараметрической ядерной процедуры с заданной степенью точности по зависимым наблюдениям x_1^t. Согласно этой процедуре мы должны заменить неизвестную плотность $f(x_t|x_1^{t-1})$ на «урезанную» плотность $f(x_t|x_{t-\tau}^{t-1})$, где τ — степень зависимости наблюдаемого процесса (X_t). По существу, τ представляет собой порядок связности марковского процесса, аппроксимирующего немарковский процесс (X_t) (Dobrovidov, et al., 2012). По определению $f(x_t|x_{t-\tau}^{t-1}) = f(x_{t-\tau}^t)/f(x_{t-\tau}^{t-1})$. Тогда

$$(\ln(f(x_t|x_{t-\tau}^{t-1}))'_{x_n} = \frac{f'_{x_t}(x_{t-\tau}^t)}{f(x_{t-\tau}^t)}. \qquad (20)$$

Знаменатель последней формулы представляет собой $(\tau + 1)$-мерную плотность.

Непараметрическая ядерная оценка для такой плотности имеет вид:

$$\hat{f}(x_{t-\tau}^t) = \\ = \frac{1}{(t-\tau-1)h^{\tau+1}} \sum_{i=1}^{t-\tau-1} \prod_{j=1}^{\tau+1} K\Big(\frac{x_{t-j+1} - x_{t-j-i+1}}{h_t}\Big). \qquad (21)$$

Непараметрическая оценка числителя (21) записывается как

$$\hat{f}'_{x_t t}(x_{t-\tau}^t) = \frac{1}{(t-\tau-1)h_{1t}^{\tau+3}} \sum_{i=1}^{t-\tau-1} K'\Big(\frac{x_t - x_{t-i}}{h_{1t}}\Big) \times \\ \times \prod_{j=2}^{\tau+1} K\Big(\frac{x_{t-j+1} - x_{t-j-i+1}}{h_{1t}}\Big), \qquad (22)$$

где $\hat{f}'_{x_t}$ и K' означает частную производную по переменной x_t.

Таким образом, для непараметрической оценки логарифмической производной плотности получаем выражение:

$$(\ln(\hat{f}(x_t|x_{t-\tau}^{t-1}))'_{x_n} = \frac{\hat{f}'(x_{t-\tau}^t)}{\hat{f}(x_{t-\tau}^t)}.$$

Такая оценка называется оценкой подстановки (plug-in estimate). Для вычисления отношения(20) остается только выбрать коэффициенты размытости h_t для плотности (21) и h_{1t} — для ее частной производной (22) [24].

4. НЕПАРАМЕТРИЧЕСКАЯ ОЦЕНКА ВОЛАТИЛЬНОСТИ

Для линейного варианта модели Тейлора (7), (8) условная плотность распределения b_t при фиксированном Δ_t будет иметь вид (вывод приводится в Приложении):

$$p(b_t|\Delta_t) = \frac{\exp(b_t/2)}{\sqrt{2\pi}\exp(\Delta_t/2)} \exp\Big(-\frac{\exp(b_t)}{2\exp(\Delta_t)}\Big). \qquad (23)$$

Формула (23) важна для построения непараметрической оценки волатильности, так как одним из условий является принадлежность условного распределения к экспонентному семейству.

Нетрудно проверить, что плотность (23) относится к экспонентному семейству (11), где

$$g(b_t) = \frac{\exp(b_t/2)}{\sqrt{2\pi}}; \quad g'_{b_t}(b_t) = \frac{1}{2}\frac{\exp(b_t/2)}{\sqrt{2\pi}} = \frac{1}{2}g(b_t);$$

$$C(\Delta_t) = \frac{1}{\exp(\Delta_t/2)};$$

$$T(b_t) = -\frac{\exp(b_t)}{2}; \quad T'_{b_t}(b_t) = -\frac{\exp(b_t)}{2};$$

$$\theta(\Delta_t) = \frac{1}{\exp(\Delta_t)}.$$

Уравнение оптимальной фильтрации (19) принимает вид:

$$\theta(\Delta_t)\times T'_{b_t}(b_t) = \frac{\partial}{\partial b_t}\left(\ln\left(\frac{f(b^t_{t-\tau})}{g(b_t)}\right)\right) =$$

$$= \frac{f'(b^t_{t-\tau})}{f(b^t_{t-\tau})} - \frac{g'(b_t)}{g(b_t)}. \qquad (24)$$

Подставим в уравнение (24) полученные выше выражения:

$$\frac{\exp(b_t)}{2\exp(\Delta_t)} = \frac{1}{2} - \frac{f'(b^t_{t-\tau})}{f(b^t_{t-\tau})}.$$

В итоге непараметрическая оценка

$$\Delta_t = \log\left(\exp(b_t)\Big/\left(1-\frac{2f'(b^t_{t-\tau})}{f(b^t_{t-\tau})}\right)\right),$$

а оценка волатильности принимает вид:

$$\sigma_t = \exp\left(\frac{\Delta_t}{2}\right) = \sqrt{\frac{\exp(b_t)}{2}\Big/\left(1-\frac{2f'(b^t_{t-\tau})}{f(b^t_{t-\tau})}\right)}. \qquad (25)$$

Покажем, что знаменатель под корнем принимает только положительные значения, т. е.

$$1 - \frac{2f'(b^t_{t-\tau})}{f(b^t_{t-\tau})} > 0.$$

Для этого умножим обе части неравенства на $f(b^t_{t-\tau})$:

$$f(b^t_{t-\tau}) - 2f'(b^t_{t-\tau}) > 0.$$

Затем поделим на $f(b^t_{t-\tau})$:

$$\frac{f(b^t_{t-\tau})}{f(b^{t-1}_{t-\tau})} - \frac{2f'(b^t_{t-\tau})}{f(b^{t-1}_{t-\tau})} = f(b_t|b^{t-1}_{t-\tau}) - 2f'(b_t|b^{t-1}_{t-\tau}) > 0.$$

С учетом выражений (16) и (17) последнее неравенство можно представить в виде:

$$\int_{\Delta_t} p(b_t|\Delta_t)w_t(\Delta_t|b_1^{t-1})d\Delta_t - 2\int_{\Delta_t} p'(b_t|\Delta_t)w_t(\Delta_t|b_1^{t-1})d\Delta_t =$$

$$= \int_{\Delta_t}(p(b_t|\Delta_t) - 2p'(b_t|\Delta_t))w_t(\Delta_t|b_1^{t-1})d\Delta_t > 0,$$

где $p'(b_t|\Delta_t)$ производная по b_t. Поскольку $w_t(\cdot) \ge 0$, остается доказать, что

$$p(b_t|\Delta_t) - 2p'(b_t|\Delta_t) > 0.$$

Для этого воспользуемся формулой (23) и получим:

$$\frac{\exp(b_t/2)}{\sqrt{2\pi}\exp(\Delta_t/2)}\exp\left(-\frac{\exp(b_t)}{2\exp(\Delta_t)}\right) -$$

$$- \frac{\exp(b_t/2)}{\sqrt{2\pi}\exp(\Delta_t/2)}\exp\left(-\frac{\exp(b_t)}{2\exp(\Delta_t)}\right) +$$

$$+ \frac{\exp(3b_t/2)}{\sqrt{2\pi}\exp(3\Delta_t/2)}\exp\left(-\frac{\exp(b_t)}{2\exp(\Delta_t)}\right) =$$

$$= \frac{\exp(3b_t/2)}{\sqrt{2\pi}\exp(3\Delta_t/2)}\exp\left(-\frac{\exp(b_t)}{2\exp(\Delta_t)}\right) > 0.$$

Ввиду положительности всех входящих в это выражение функций неравенство справедливо при любых значения b_t и Δ_t.

Для непараметрического варианта точного уравнения (25) следует заменить $f(b^t_{t-\tau})$ и $f'(b^t_{t-\tau})$ на их непараметрические оценки $\hat{f}(b^t_{t-\tau})$ и $\hat{f}'\, b^t_{t-\tau})$ согласно формулам (21) и (22). При такой замене знаменатель под корнем в формуле (25) может оказаться неположительным. Чтобы избежать этого на него налагается дополнительное условие отделимости от неположительных значений, которое формулируется следующим образом: $(\upsilon)^+ = ((\upsilon$ при $\upsilon > 0) \vee (\varepsilon_0^+$ при $\upsilon \le 0))$, где ε_0^+ — некоторое малое положительное число. Тогда формула для непараметрической оценки волатильности представляется в виде:

$$\hat{\sigma}_t = \sqrt{\frac{\exp(b_t)}{2}\Big/\left(1-\frac{2\hat{f}'(b^t_{t-\tau})}{\hat{f}(b^t_{t-\tau})}\right)}. \qquad (26)$$

5. ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ

С целью численной реализации разработанного алгоритма оценивания в соответствии с моделью Тейлора (8) были сгенерированы $t = 1500$ значений истинной волатильности, на основании которых впоследствии были получены данные для доходов (4). Первые 1000 значений использовались для обучения моделей, на оставшихся 500 проводилось тестирование результатов. Для проверки качества оценки волатильности $\hat{\sigma}_i$ для каждого метода было проведено 100 повторных испытаний, по которым вычислялся корень квадратный от среднеквадратической ошибки (КСКО) отклонения от истинной волатильности σ_i:

$$\text{КСКО}(\hat{\sigma}) = \sqrt{\frac{\sum\limits_{i=1000}^{1500}(\sigma_i - \hat{\sigma}_i)^2}{t}}.$$

По этому критерию ниже строится таблица сравнения различных методов оценивания.

5.1. Непараметрическая оценка

Непараметрическая оценка волатильности (26) (светлая линия) в сравнении с истинной волатильностью (темная линия) представлена на рис. 1.

Заметим, что непараметрическая оценка уловила все участки кластеризации (инерции) волатильности, которые характерны для экономических кризисов и которые сигнализируют о сильных длительных колебаниях доходности в эти периоды. Для 100 различных тестов среднее значение КСКО составило 2,8471. При численной реализации алгоритма число ε_0^+ выбиралось равным 0,005.

5.2. Фильтр Кушнера — Стратоновича

На рис. 2 представлена оценка фильтром Кушнера — Стратоновича (13), (14) (светлая линия) и истинная сгенерированная волатильность (темная линия). В данном случае КСКО = 2,4674, что лучше непараметрической оценки. Такой результат был очевиден, поскольку оценка фильтром Кушнера — Стратоновича при использовании полной статистической информации является оптимальной.

5.3. Фильтр Калмана

Результат оценки фильтром Калмана представлен на рис. 3 (светлая линия). Как и в предыдущих двух случаях удалось уловить все участки кластеризации, однако КСКО = 3,1614 почти в два с половиной раза больше, чем у непараметрического

Рис. 1. Непараметрическая оценка волатильности

Рис. 2. Оценка волатильности фильтром Кушнера — Стратоновича

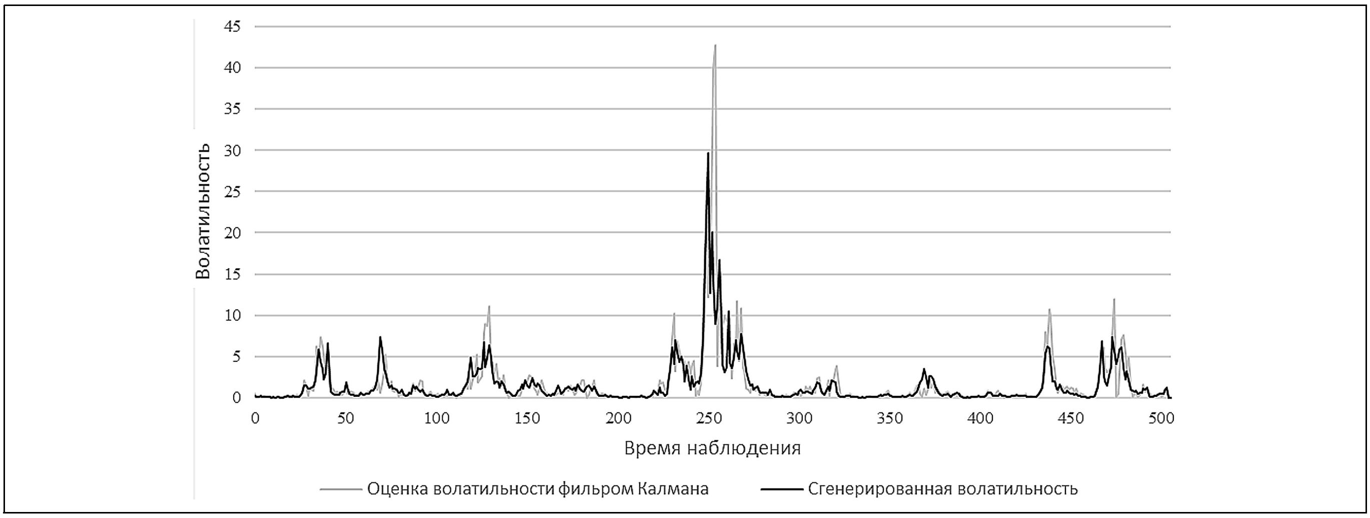

Рис. 3. Оценка волатильности фильтром Калмана

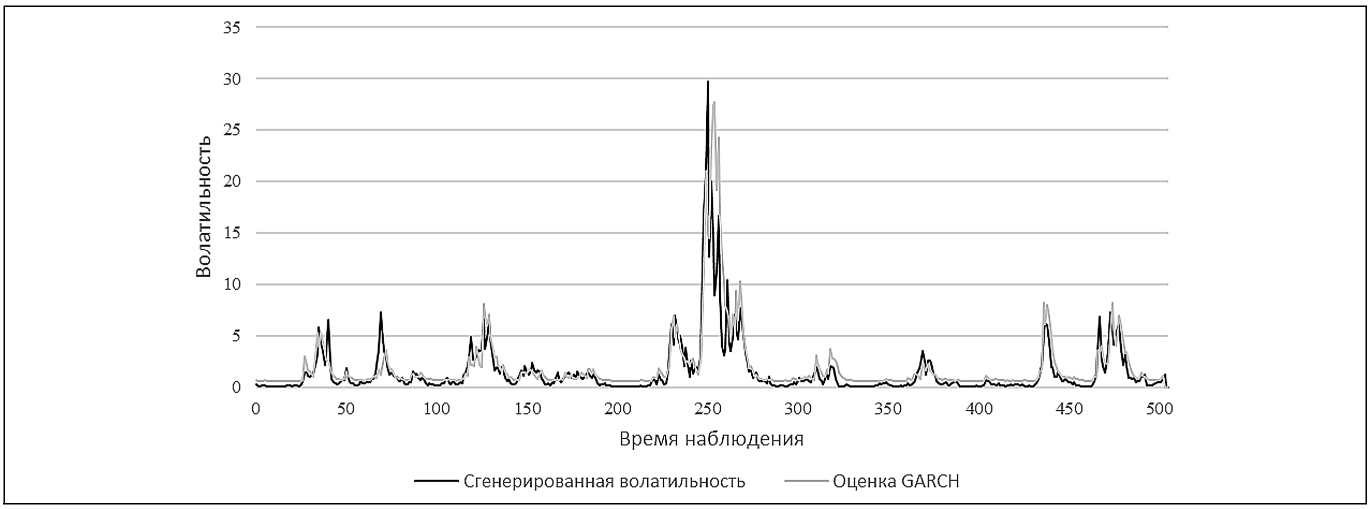

Рис. 4. Оценка волатильности с помощью GARCH

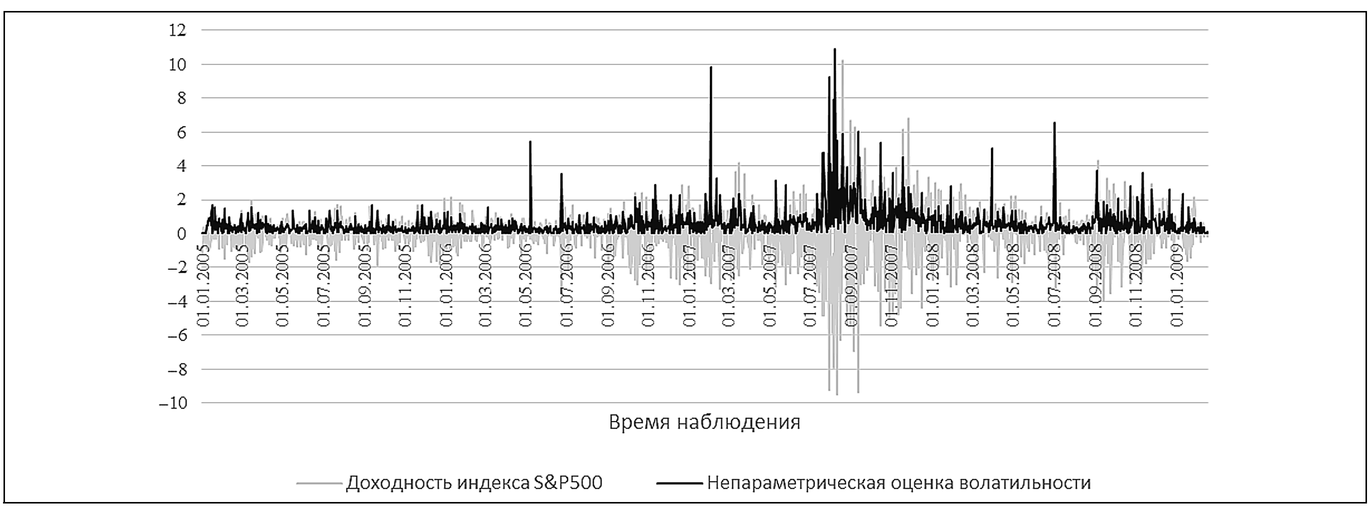

Рис. 5. Непараметрическая оценка волатильности индекса S & P500

аналога. Предположительно это связано с двумя причинам:

— для возможности применения фильтра Калмана в работе была сделана замена истинного распределения на нормальное распределение с теми же первыми двумя моментами;

— оценивалась не волатильность напрямую, а ее логарифм.

5.4. GARCH

Наконец, рассмотрим оценку волатильности, описанную в п. 2.1 моделью GARCH. Численная реализация выполнена с помощью стандартных функций garchset (настраивает модель по переданным на вход значениям p) q и garchfit (делает оценку волатильности согласно настроенной модели и переданному на вход вектору значений дохода) пакета прикладных программ Matlab. На рис. 4 представлен результат оценки волатильности моделью GARCH (светлая линия) в сравнении с истинной волатильностью (темная линия). Значение КСКО составило 4,0727, а значит GARCH, как и фильтр Калмана, уступает непараметрическому методу. Численный алгоритм выбирал значения p, q и при которых КСКО был минимальным.

5.5. Применение непараметрической оценки на реальных данных

Для того, чтобы быть уверенными, что предложенные методы непараметрического оценивания действительно работают, проверим их на реальных данных. Рассмотрим значение ежедневного индекса S&P 500 за период с 01.01.2005 г. до 31.12.2010 г. В 2008 г. начался мировой экономический кризис, во время которого наблюдались сильные колебания на рынке акций. На рис. 5 видно, что, на участке с 900 по 1100 день, который соответствует глобальной рецессии 2009 года, непараметрическая оценка имеет участок кластеризации, а значит, поставленная цель достигнута и для реальных данных.

ЗАКЛЮЧЕНИЕ

Ниже представлена таблица сравнения трех различных методов с КСКО для фильтра Кушнера — Стратоновича (К — С):

$$Res_i = \frac{\text{КСКО}_i - \text{КСКО}_{\text{К–С}}}{\text{КСКО}_{\text{К–С}}},$$

где $i = 1, 2, 3$ — это непараметрическая оценка, фильтр Калмана и GARCH соответственно. Таким образом, наилучшая оценка волатильности была получена непараметрическим методом. И это естественно, так как только в данном подходе можно работать даже при неизвестном распределении волатильности, т. е. нет необходимости строить относительно неточные модели для ее описания. В параметрических моделях для построения оценки вводятся уравнения состояния полезного сигнала. В случае с фильтром Калмана также приходится заменять негауссовские шумы гауссовскими с соответствующими параметрами. В то же время непараметрический метод показывает хорошие результаты при неизвестном распределении этого сигнала. Кроме того, данный результат можно объяснить тем, что в случае непараметрического метода оценка может строиться и в нелинейном варианте.

Сравнение различных методов оценки волатильности с фильтром Кушнера—Стратоновича

Метод оценки	Непараметрическая оценка	Фильтр Калмана	GARCH
Результат Res_i	0,15	0,28	0,65

Хотя и был получен удовлетворительный результат, в настоящей работе рассматривалась только диффузионная составляющая доходности и не учитывались разрывная компонента и память эффектов предыдущих периодов. В дальнейшем будет сделана попытка проанализировать и эти аспекты, основываясь на развиваемых в данной работе подходах теории непараметрического оценивания сигналов.

ПРИЛОЖЕНИЕ

П.1. Распределение $\xi_t = \log(\varepsilon_t^2)$

В уравнении (7) случайная величина $\xi_t = \log(\varepsilon_t^2)$, где $\varepsilon_t \sim \mathcal{N}(0, 1)$. Введем случайную величину $\zeta_t = \varepsilon_t^2$, которая имеет χ^2-распределение с одной степенью свободы. Для получения распределения $\xi_t = \log(\zeta_t)$ воспользуемся формулой преобразования плотностей распределения $p(\xi_t)d\xi_t = \phi(\zeta_t)d\zeta_t$, где $\phi(\zeta_t)$ — плотность χ^2-распределения. Эта плотность записывается в виде:

$$\phi(\zeta_t(\xi_t)) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\xi_t}{2}\right)\exp\left(-\frac{\exp(\xi_t)}{2}\right).$$

Поскольку дифференциал

$$\frac{d\zeta_t}{d\xi_t} = \frac{d(\exp(\xi_t))}{d\xi_t} = \exp(\xi_t),$$

$$\text{то } p(\xi_t) = \frac{\exp(\xi_t/2)}{\sqrt{2\pi}}\exp\left(-\frac{\exp(\xi_t)}{2}\right).$$

Из уравнения (7) следует, что $b_t - \Delta_t = \xi_t$. Так как в данном случае $\frac{d\xi_t}{db_t} = 1$, то условная (при фиксированном Δ_t) плотность наблюдений b_t

$$p(b_t|\Delta_t) = \frac{\exp((b_t - \Delta_t)/2)}{\sqrt{2\pi}}\exp\left(-\frac{\exp(b_t - \Delta_t)}{2}\right) =$$

$$= \frac{\exp(b_t/2)}{\sqrt{2\pi}\exp(\Delta_t/2)}\exp\left(-\frac{\exp(b_t)}{2\exp(\Delta_t)}\right).$$

П.2. Параметры полученного распределения

Математическое ожидание и дисперсию для ξ_t удобно вычислить через производящую функцию моментов $M_\xi(t) = E[\exp(t\xi)]$, для которой справедлива формула:

$$\mathrm{E}[\xi_t^k] = \left.\frac{d^k M_\xi(t)}{dt^k}\right|_{t=0}.$$

Производящая функция для $\xi_t = \log(\zeta_t)$ имеет вид:

$$M_\xi(t) = \mathrm{E}[\exp(t\log(\zeta))] = \mathrm{E}[\zeta^t].$$

В свою очередь, производящая функция для χ^2 со степенью свободы m

$$M_\zeta(t) = (1 - 2t)^{-m/2},$$

а момент порядка t

$$\mathrm{E}[\zeta^t] = 2^t \frac{\Gamma(m/2 + t)}{\Gamma(m/2)}.$$

Тогда

$$\mathrm{E}(\xi_t) = \left.\frac{dM_\xi(t)}{dt}\right|_{t=0} = \left.\frac{d\mathrm{E}(\zeta^t)}{dt}\right|_{t=0} = \left.\frac{d\left(2^t\frac{\Gamma(1/2+t)}{\Gamma(1/2)}\right)}{dt}\right|_{t=0} =$$

$$= \left.\left(2^t\log 2\frac{\Gamma(1/2+t)}{\Gamma(1/2)} + 2^t\frac{\Gamma(1/2+t)\psi(1/2+t)}{\Gamma(1/2)}\right)\right|_{t=0} =$$

$$= \log 2 + \psi(1/2) \approx -1{,}2704,$$

где $\psi(1/2)$ — дигамма-функция.

Для подсчета дисперсии сначала выведем $\mathrm{E}(\xi_n^2)$:

$$\mathrm{E}(\xi_n^2) = \left.\frac{d^2 M_\xi(t)}{dt^2}\right|_{t=0} = \left(\log 2\left(2^t\log 2\frac{\Gamma(1/2+t)}{\Gamma(1/2)} + \right.\right.$$

$$+ 2^t\frac{\Gamma(1/2+t)\psi(1/2+t)}{\Gamma(1/2)}\Big) + \psi(1/2+t)\Big(2^t\log 2\frac{\Gamma(1/2+t)}{\Gamma(1/2)} +$$

$$\left.\left.+ 2^t\frac{\Gamma(1/2+t)\psi(1/2+t)}{\Gamma(1/2)}\right) + 2^t\frac{\Gamma(1/2+t)}{\Gamma(1/2)}\psi'(1/2+t)\right)\Bigg|_{t=0} =$$

$$= \log 2(\log 2 + \psi(1/2)) + \psi(1/2)(\log 2 + \psi(1/2)) + \psi'(1/2) =$$

$$= (\log^2 + \psi'(1/2))^2 + \psi'(1/2).$$

В итоге

$$D(\xi_t) = \mathrm{E}(\xi_t^2) - (E(\xi_t))^2 = \psi'(1/2) = \pi^2/2 \approx 4{,}9348.$$

ЛИТЕРАТУРА

1. *Merton E.* Theory of rational optiopricing // Bell Journal of Economics and Management Science. — 1973. — Vol. 4. — P. 141—183.
2. *Busse J.* Volatility timing in mutual funds: Evidence from daily returns // Review of Financial Studies. — 1999. — Vol. 12, N 5. — P. 1009—1041.
3. *Fleming J., Kirby C., Ostdiek B.* The economic value of volatility timing // Journal of Finance. — 2001. — Vol. 56, N 1. — P. 329—352.
4. *Fleming J., Kirby C., Ostdiek B.* The economic value of volatility timing using «realized» volatility // Journal of Finance. — 2003. — Vol. 67, N 3. — P. 473—509.
5. *Yao Q., Tong H.* Quantifying the influence of initial values on nonlinear prediction // Journal of the Royal Statistical Society. — 1994. — Ser. B 56. — P. 701—725.
6. *Ziegelmann F.* Nonparametric Estimation of Volatility Functions: The Local Exponential Estimator // Econometric Theory. — 2002. — Vol. 18. — P. 985—91.
7. *Duffie D., Pan J., Singleton K.* Transform analysis and asset pricing for affine ne jump-diffusions // Econometrica. — 2000. — Vol. 68, N 6. — P. 1343—1376.
8. *Jacquier E., Polson N., Rossi P.* Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics. — 1994. — Vol. 12. — P. 371—389.
9. *Bandi F., Reno R.* Nonparametric Stochastic Volatility. — 2016. — Working paper. — URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1158438 (дата обращения: 4.07.2017).
10. *Dobrovidov A.V., Koshkin G.M., Vasiliev V.A.* Non-parametric models and statistical inference from dependent observations. — USA, Kendrick Press, 2012.
11. *Bollerslev T.* Generalized autoregressive conditional heteroskedasticy. — Journal of Econometrics. — 1986. — Vol. 31. — P. 307—327.
12. *Taylor S.* Modelling Financial Times Series. — N.-Y.: Wiley, 1986.
13. *Ширяев А.Н.* Основы стохастической финансовой математики. — М.: Фазис, 1998.
14. *Ширяев А.Н.* Вероятность: 2-е изд., перераб. и доп. — М.: Наука, 1989.
15. *Mills T.C.* The Econometric Modelling of Financial Time Series: 2nd ed. — Cambridge (UK): Cambridge University Press, 1999.
16. *Harvey A.C.* Forecasting structural Time Series Models and the Kalman Filter. — Cambridge (UK): Cambridge University Press, 1989.
17. *Kalman R.* A new approach to linear filtering and prediction problems // Journal of Basic Engineering. — 1960. — Vol. 82. — P. 34—45.
18. *Theoret E., Eacicot F.* Forecasting stochastic Volatility using the Kalman filter: an application to Canadian Interest Rates and Price-Earnings Ratio. — Munich Personal EePEc Archive, 2010. — URL: https://mpra.ub.uni-muenchen.de/35911/1/MPRA_paper_35911.pdf (дата обращения: 4.07.2017).
19. *Добровидов А.В.* Непараметрические методы нелинейной фильтрации стационарных случайных последовательностей // Автоматика и телемеханика. — 1983. — № 6. — С. 85—98.
20. *Kushner H.J.* On the dynamical equations of conditional probability density functions with applications to optimal stochastic control theory // Journal of Mathematical Analysis and Applications. — 1964. — Vol. 8, N 2. — P. 332—344.
21. *Stratonovich E.L.* Conditional Markovian processes and their application to the optimal control theory. — M.: Moscow University Press, 1966.
22. *Тихонов А.Н.* О некорректных задачах линейной алгебры и устойчивом методе их решения // Доклады академии наук СССР. — 1965. — Т. 163, № 3. — С. 591—594.
23. *Hall P., Marron J.S., Park B.U.* Smoothed Cross-Validation // Probability theory and related fields. — 1992. — Vol. 90. — P. 1—20.
24. *Добровидов А.В., Рудько И.М.* Выбор ширины окна ядерной функции в непараметрической оценке производной плотности методом сглаженной кросс-валидации // Автоматика и телемеханика. — 2010. — № 2. — С. 45—57.

Статья представлена к публикации членом редколлегии В.В. Клочковым.

Добровидов Александр Викторович — д-р физ. мат. наук, зав. лабораторией, ✉ dobrovidov@gmail.com,

Тевосян Вачик Эдвардович — математик, аспирант, ✉ vachik.tevosian@gmail.com,

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва.