Эконометрический ликбез: моделирование временных рядов

Объекты неструктурного моделирования временных рядов*

Станислав Анатольев†

Российская экономическая школа, Москва, Россия

При моделировании динамики временных рядов решение о выборе класса и типа модели сильно зависит от моделируемого объекта. Данное эссе рассказывает о специфике моделирования различных объектов, таких как условное на предыстории среднее, условная дисперсия, условный квантиль, условная вероятность и условная плотность. Перечисляются наиболее популярные классы моделей для каждого объекта, обсуждаются особенности их анализа. При этом внимание уделяется как одномерному, так и многомерному случаям, а также даются ссылки на узкотематические, но более подробные обзоры.

1 Введение

На начальном шаге неструктурного[1] моделирования динамики временных рядов приходится решать, какой объект распределения рассматриваемой переменной[2], условного на предыстории этой переменной, интересен исследователю. Под фразами «построить модель для ВВП» и «построить модель для доходности» чаще всего понимаются не только совершенно разные модели, но и разные моделируемые объекты. Если обозначить ВВП как y_t, доходность финансового актива за r_t, а предысторию рассматриваемого ряда как I_{t-1}, то «модель для ВВП» скорее всего означает модель для условного среднего $\mathbb{E}\left[\Delta y_t | I_{t-1}\right]$, являющегося наилучшим прогнозом (при среднеквадратических потерях) для роста ВВП, а «модель для доходности» – модель для условной дисперсии $\mathbb{V}\left[r_t | I_{t-1}\right]$, являющейся одной из мер волатильности доходностей. То есть, перед моделированием временного ряда необходимо определиться, какой именно объект, какая именно регрессия на предыстории, представляет интерес для исследователя.[3] Упомянутые выше объекты не являются единственными интересными при моделировании ВВП и доходности. Несколько реже мы можем видеть модели для условной (на предыстории) вероятности $\mathbb{P}\left\{\Delta y_t > \delta | I_{t-1}\right\}$, прогнозирующие подъемы и спады, или для $f_{t-1}\left(r_t\right)$, условной плотности доходностей.

---

*Работа основана на лекциях, читаемых автором в РЭШ. Цитировать как: Анатольев, Станислав (2013). «Объекты неструктурного моделирования временных рядов», Квантиль, №11, стр. 1–11. Citation: Anatolyev, Stanislav (2013). “Objects of nonstructural time series modeling,” Quantile, No.11, pp. 1–11.

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

[1]Неструктурным называтся моделирование динамики, не основанное на экономической теории. Интересен в таком случае анализ какой-то характеристики условного распределения, где в качестве информационного множества выступает собственная предыстория, чаще всего начиная с предыдущего периода.

[2]Переменная может быть скалярной или векторной. В первом случае анализ называется однопеременным или одномерным, во втором — многопеременным или многомерным.

[3]Иногда приходится читать работы, в которых автор сначала моделирует условное среднее (обычно с помощью линейной авторегрессии), а затем приступает к моделированию условной дисперсии (обычно с помощью ARCH-моделей), обосновывая свои шаги тем, что «вот что-то нашли в среднем, теперь логично смоделировать дисперсию». Подобный исследователь не имеет понятия, что он(а) исследует и, главное, зачем.

В настоящем эссе мы рассказываем, какие объекты обычно представляют интерес для исследователя, анализирующего конкретные типы данных. Для каждого объекта мы перечисляем наиболее актуальные и популярные модели и указываем на некоторые особенности их анализа. Большинство моделей представлены в своей простейшей конфигурации: например, авторегрессия имеет первый порядок, модель с режимами — два режима и т.д. Мы также приводим ссылки на соответствующую литературу, преимущественно обзоры, содержащие более подробный материал по каждому классу моделей.

Упомянем также, что является критерием предсказуемости временного ряда. Временной ряд x_t с предысторией I_{t-1} считается предсказуемым в смысле среднего, если $\mathbb{E}\left[x_t|I_{t-1}\right]$ зависит от I_{t-1}, предсказуемым в смысле дисперсии, если $\mathbb{V}\left[x_t|I_{t-1}\right]$ зависит от I_{t-1}, предсказуемым в смысле знака, если $\mathbb{P}\left\{x_t > 0|I_{t-1}\right\}$ зависит от I_{t-1}, предсказуемым в смысле α-квантиля, если $\mathbb{Q}_\alpha\left(x_t|I_{t-1}\right)$ зависит от I_{t-1} (где $\mathbb{Q}_\alpha$ обозначает α-квантильную функцию), и т.д. Под «просто» предсказуемостью можно понимать зависимость условной плотности $f_{t-1}\left(r_t\right)$ от I_{t-1}. Заметим, что предсказуемость временного ряда в одном смысле необязательно влечет за собой предсказуемость в другом смысле (см. примеры в Анатольев 2006). Видим, что критерием предсказуемости или непредсказуемости выступает факт зависимости или независимости объекта от информации в предыстории, но ни в коем случае не численный размер объекта (так что, например, большое, но постоянное значение $\mathbb{E}\left[x_t|I_{t-1}\right]$ *не означает* предсказуемости среднего, а $\mathbb{P}\left\{x_t > 0|I_{t-1}\right\} = 1$ *не означает* предсказуемости знака). Чтобы протестировать на предсказуемость, необходимо прогнать регрессию (возможно, нелинейную) соответствующего объекта на некоторых переменных из информационного множества I_{t-1} и протестировать на их невключение в нее. Например, для тестирования предсказуемости среднего надо прогнать регрессию среднего для x_t, для тестирования предсказуемости знака – прогнать бинарную регрессию для индикатора $\mathbb{I}\{x_t > 0\}$, и т.д. Некоторые подробности см. в эссе Анатольев (2006).

На протяжении данного эссе мы предполагаем, что моделируются (возможно, после взятия приращений логарифмов или применения других преобразований) строго стационарные эргодические ряды.

2 Моделирование условного среднего

Чаще всего целью моделирования макроэкономического ряда y_t является прогнозирование его уровня. При обычно используемых среднеквадратических потерях[4] оптимальным прогнозом является условное на предыстории среднее $\mathbb{E}\left[y_t|I_{t-1}\right]$. Поскольку речь чаще всего идет про спокойные, низкочастотные макроэкономические данные типа ВВП или безработицы, обычно, хоть и не всегда, накладывают условие условной гомоскедастичности $\mathbb{V}\left[y_t|I_{t-1}\right] = \sigma^2$. Конечный интерес чаще всего представляют функции импульсных откликов (IRF, от англ. impulse response function), показывающие распространение влияния шоков для различных горизонтов h. Малые значения h соответствуют краткосрочным эффектам, умеренные h — среднесрочным, а $h = \infty$ — долгосрочному. Долгосрочный эффект равен нулю, раз анализируется стационарный эргодичный ряд.

Самым простым классом моделей для условного среднего является класс линейных авторегрессий (AR, от англ. autoregression)

$$\mathbb{E}\left[y_t|I_{t-1}\right] = \mu + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \cdots + \rho_p y_{t-p},$$

или

$$y_t = \mu + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \cdots + \rho_p y_{t-p} + \varepsilon_t,$$

[4]То есть когда функция потерь равна e^2, где e — ошибка прогноза, а прогноз является оптимальным, если минимизируется $\mathbb{E}[e^2]$. Другие примеры функций потерь можно найти в Маккракен (2006).

где (здесь и далее) ε_t — шок со свойством непрогнозируемости среднего $\mathbb{E}\left[\varepsilon_t|I_{t-1}\right]=0$. Динамика переменной y_t линейная, а импульсные отклики пропорциональны размеру шока и не зависят от предыстории:

$$\mathrm{IRF}(h)=\frac{\partial y_{t+h}}{\partial \varepsilon_t}.$$

Например, в случае линейной авторегрессии первого порядка ($p=1$) функция импульсного отклика равна $\mathrm{IRF}(h)=\rho_1^h$.

Линейная авторегрессия в многомерном контексте называется векторной авторегрессией (VAR, от англ. vector autoregression). Векторные авторегрессии позволяют не только отслеживать динамику исследуемых переменных, но и анализировать их влияние друг на друга в том же периоде. См., например, главу 11 в учебнике Hamilton (1994).

Линейные модели слишком просты. В частности, импульсные отклики пропорциональны размеру шока, при перемене знака шока меняется только их знак, они неспособны учесть разную реакцию на шоки в разных состояниях системы. Нелинейную динамическую модель создать очень просто — достаточно в линейную модель ввести компоненты, нелинейные по шоку. Примерно таким образом можно сформировать класс билинейных моделей (англ. bilinear models). Несмотря на их простоту, в эконометрике они не получили особой популярности ввиду их плохой интерпретируемости. Неструктурная модель хороша, когда она способна учесть, или «объяснить», какие-то экономические стилизованные факты, присущие рассматриваемому типу данных.

Для нелинейных моделей эффекты откликов характеризуются обобщённой функцией импульсных откликов (GIRF, от англ. generalized IRF)

$$\mathrm{GIRF}(h,\varepsilon,I_{t-1})=\mathbb{E}\left[y_{t+h}|\varepsilon_t=\varepsilon,I_{t-1}\right]-\mathbb{E}\left[y_{t+h}|\varepsilon_t=0,I_{t-1}\right].$$

В определении фигурирует так называемый пересмотр прогноза, то есть как меняется прогноз при реализации ненулевого шока по сравнению с ситуацией, когда шок отсутствует. Эту функцию даже для простых нелинейных моделей аналитически вычислить очень сложно, поэтому обычно она рассчитывается численно с помощью симуляционных методов.

Наиболее популярными классами нелинейных моделей для условного среднего макроэкономических переменных, подверженных цикличности (ВВП, безработица), являются те, которые предусматривают наличие разных режимов, соответствующих разным стадиям деловых циклов, ведь это как раз и является ключевым стилизованным фактом для макроэкономических данных. В каждом из режимов динамика линейна, но она разная в разных режимах, а принадлежность к тому или иному режиму определяется эндогенно в зависимости от значений переменной.

Ярким представителем моделей с режимами является пороговая авторегрессия (TAR, от англ. threshold autoregression). Ее двухрежимная версия в простой конфигурации выглядит следующим образом:

$$y_t=\left\{\begin{array}{ll}\mu_1+\rho_1 y_{t-1}, & \text{если } q_{t-1}\le\gamma\\ \mu_2+\rho_2 y_{t-1}, & \text{если } q_{t-1}>\gamma\end{array}\right.+\varepsilon_t,$$

где q_{t-1} — пороговая переменная, наблюдаемая в момент $t-1$, а γ — значение порога. Модель TAR легко обобщается на случай большего числа лагов, большего количества режимов и нескольких пороговых переменных. Доступное изложение материала о пороговых авторегрессиях содержится в Hansen (1997). Возможно идею пороговых переключений режимов рассматривать и в контексте векторных авторегрессий; см., например, Balke (2000).

Похожа на пороговую модель с гладкими переходами (STAR, от англ. smooth transition autoregression). Простейшая двухрежимная версия выглядит как

$$y_t=(\mu_1+\rho_1 y_{t-1})(1-G(q_{t-1},\gamma,\delta))+(\mu_2+\rho_2 y_{t-1})G(q_{t-1},\gamma,\delta)+\varepsilon_t,$$

где $G(\cdot, \gamma, \delta)$ — функция перехода, q_{t-1} — переменная перехода, γ — точка локализации перехода, а δ — дополнительный параметр его гладкости. Модель STAR можно обобщить на случаи большего количества переходов и нескольких переменных перехода. Подробный обзор моделей с гладкими переходами содержится в работе Franses, Teräsvirta & van Dijk (2002).

В приведенных двух моделях переходы между режимами управляются наблюдаемой (возможно, с точностью до параметров) переменной q_{t-1}. Это непривлекательное предположение, и более гармоничной выглядит формулировка модели с Марковскими переключениями режимов (MSW, от англ. Markov switching). Аналогичная двухрежимная версия выглядит следующим образом:

$$y_t = \begin{cases} \mu_1 + \rho_1 y_{t-1}, & \text{если } s_t = 0 \\ \mu_2 + \rho_2 y_{t-1}, & \text{если } s_t = 1 \end{cases} + \varepsilon_t,$$

где s_t — ненаблюдаемая бинарная переменная состояния, о которой лишь известно, что её динамика следует Марковской цепи первого порядка с матрицей переходов

$$P = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix}$$

с неизвестными переходными вероятностями p и q. Платой за концептуальную привлекательность такой модели является относительная трудность ее оценивания: в то время как в моделях TAR и STAR функцию (квази-)правдоподобия выписать легко, в случае модели MSW функция (квази-)правдоподобия рассчитывается рекурсивно с помощью так называемого фильтра Гамильтона (по имени создателя модели, см. Hamilton, 1989). Причиной является как раз наличие ненаблюдаемого процесса s_t. Модель MSW можно обобщить на случаи большего количества режимов, более сложной динамики переменной состояния и зависимых от времени переходных вероятностей. Введение в модели с Марковскими переключениями режимов содержится в работе Куан (2013) в этом выпуске «Квантиля».

Модель с Марковскими переключениями является членом широкого класса моделей, называемых моделями с ненаблюдаемыми компонентами. В каждой такой модели присутствует латентный процесс, динамика и распределение которого специфицированы (в случае модели MSW это s_t), а дополнительное уравнение связывает наблюдаемые переменные (в случае модели MSW это y_t) с этими латентными переменными. Например, исследователь может наблюдать рост ВВП, но осознавать, что вклад в наблюдаемые данные по росту вносят несколько факторов с принципиально разной динамикой. Здесь наличие латентных факторов осложняет оценивание, к которому существует несколько подходов, основным из которых является фильтр Кальмана. См. подробный обзор в работе Цыплаков (2011).

Интересным классом являются модели с мультипликативной ошибкой (MEM, от англ. multiplicative error model). Модели MEM строятся для временных рядов, которые положительны по построению. Положительных экономических и особенно финансовых рядов довольно много: процентные ставки, дюрации между последовательными трансакциями, различные меры волатильности (например, реализованная волатильность), спрэды и т.п. Структура модели MEM напоминают структуру моделей GARCH (см. раздел 3):

$$y_t = \psi_t \eta_t,$$

где $\eta_t \sim \mathcal{D}$ для некоего положительного распределения $\mathcal{D}$ с единичным средним, а динамика условного среднего ψ_t задается авторегрессионным образом, например GARCH-подобным уравнением

$$\psi_t = \omega + \alpha y_{t-1} + \beta \psi_{t-1}.$$

См. обзор моделей MEM в свежей работе Brownlees, Cipollini & Gallo (2012).

Привычные стационарные (в частности, линейные) авторегрессии обладают тем свойством, что автокорреляционная функция падает по показательному закону, то есть асимптотически как a^j для некоторого $-1 < a < 1$. В ряде случаев такое падение автокорреляций слишком быстрое по сравнению с реальным поведением выборочных автокорреляций. Иногда по этой причине динамику (речь идёт в первую очередь о долгосрочной динамике) моделируют с длинной памятью (англ. long memory), в результате чего автокорреляционная функция падает уже по степенному закону, то есть асимптотически как j^{d-1} для некоторого $0 < d < 1$. Преобразование исходной переменной принимает форму дробной интегрированности (англ. fractional integration), т.е. вместо самой переменной y_t или её приращения Δy_t моделируется дробное приращение $\Delta^d y_t$, причём параметр d также оценивается наряду с остальными параметрами модели. См. обзор Baillie (1996).

3 Моделирование условной дисперсии

Условная дисперсия интересна при динамическом моделировании поведения финансовых данных вроде доходностей (которые будем обозначать через r_t), поскольку интуитивно ассоциируется с волатильностью. Для того чтобы адекватно смоделировать условную дисперсию, необходимо сначала (или одновременно) верно специфицировать условное среднее, иначе ошибки спецификации условного среднего вольются в дисперсионное уравнение и исказят инференцию о волатильности. Поскольку обычно рассматриваются эффективные рынки, достаточно в уравнении для среднего учесть (возможно присутствующую) небольшую серийную корреляцию с помощью авторегрессионного слагаемого и определенные сезонные компоненты с помощью фиктивных переменных.

Самый известный и популярный класс моделей условной дисперсии – это модели из класса ARCH (от англ. autoregressive conditional heteroskedasticity), введенного Робертом Энгелем в начале 80-х гг.; см. недавние обзоры, например, в Teräsvirta (2009) и Росси (2010). Они имеют следующую структуру:

$$r_t = \mu_t + \varepsilon_t,$$

где μ_t — условное среднее, $\varepsilon_t = \sigma_t \eta_t$ — непрогнозируемый (в среднем) шок, σ_t — условное стандартное отклонение, η_t — стандартизованный шок, имеющий распределение $\mathcal{D}$ с нулевым средним, динамика условной дисперсии σ_t^2 задается авторегрессионно, причём движущей силой являются шоки прошлых периодов (например, ε_{t-1}) или (реже) стандартизованные шоки прошлых периодов (например, η_{t-1}). В случае модели GARCH (от англ. generalized ARCH)

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2.$$

Существует великое множество спецификаций дисперсионного уравнения (см. Bollerslev 2009). Например, оно может формулироваться не для условной дисперсии, а для её логарифма, для квадратного корня из неё (т.е. для условного стандартного отклонения) или другой её функции. Инициировать процесс условной дисперсии может квадрат предыдущих шоков, как в базовой ARCH-постановке, и тогда так называемая кривая влияния новостей (NIC, от англ. news impact curve) представляет из себя параболу. Более правдоподобна кривая влияния новостей, несимметричная относительно знака предыдущего шока, что иногда интерпретируется как эффект рычага (от англ. leverage effect). Например, одна из популярных спецификаций, учитывающих несимметричность, выглядит следующим образом:

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 \mathbb{I}\{\varepsilon_{t-1} > 0\} + \beta \sigma_{t-1}^2.$$

Это так называемая спецификация GJR-GARCH (три первые буквы — аббревиатуры фамилий авторов). Кроме асимметричного влияния шоков, в дисперсионное уравнение модели

GARCH иногда закладывают переключения режимов, плавные переходы или структурные сдвиги параметров. Также можно ввести в дисперсионное уравнение свойство длинной памяти, в результате чего возникает модель FIGARCH (от англ. fractionally integrated GARCH). Спецификация условного распределения $\mathcal{D}$ несущественна для оценивания и прогнозирования волатильности, но имеет значение для моделирования динамики хвостов и условной плотности (см. разделы 4 и 6).

Базовую модель ARCH можно обогатить возможностью так называемых скачков, способных объяснить большое количество резких изменений цен на финансовом рынке. Шоки делятся на два типа: «гладкие» как ранее и «скачки»:

$$r_t = \mu_t + \varepsilon_t + \sum_{k=1}^{K_t} \zeta_{t,k},$$

где скачками являются шоки $\zeta_{t,k}$, которые реализуются в периоде t и распределены независимо друг от друга и от «гладких» шоков. Их количество K_t подчиняется распределению (например) Пуассона с параметром интенсивности, скажем, λ_t, который в свою очередь следует авторегрессионному процессу. См., например, приложение в работе Белоусов (2006).

Наконец, параллельно классу моделей ARCH существуют более привлекательные (концептуально, но не с точки зрения оценивания) модели стохастической волатильности (SV, от англ. stochastic volatility). Отличие модели SV от модели ARCH в том, что в дисперсионном уравнениии для первой наличествует еще один источник неопределённости, т.е. ещё один шок, независимый от обычного стандартизованного шока, в то время как дисперсионное уравнениие модели ARCH является детерминистической функцией от прошлых (почти) наблюдаемых шоков:

$$\log \sigma_t^2 = \gamma + \rho \log \sigma_{t-1}^2 + \nu_t,$$

где $\nu_t \sim \mathcal{N}(0, \sigma_\eta^2)$. Подробно о моделях стохастической волатильности см. Цыплаков (2010).

При многомерном моделировании волатильности, то есть матрицы условных дисперсий и ковариаций Σ_t, возникает проблема перепараметризации, то есть быстрого роста количества параметров с ростом количества анализируемых переменных. Существует несколько представлений для динамики дисперсионной матрицы, но все они в той или иной степени подвержены этой проблеме. Другой аспект — матрица Σ_t должна быть по построению положительно определена во все моменты времени. Например, популярная модель BEKK (от аббревиатур фамилий создателей) для многомерного аналога GARCH(1,1) имеет вид

$$\Sigma_t = \Omega + A\varepsilon_{t-1}\varepsilon_{t-1}' A' + B\Sigma_{t-1}B',$$

где Ω — симметричная положительно определенная матрица констант, а A и B — матрицы констант. Популярной альтернативой в настоящее время является так называемая модель динамических условных корреляций (DCC, от англ. dynamic conditional correlations), см. обзор Bauwens, Laurent & Rombouts (2006).

4 Моделирование условных квантилей

Необходимость динамического моделирования условных квантилей может возникнуть в двух случаях. В первом случае исследователь использует несимметричную кусочно-линейную функцию потерь

$$L(u) = \mathbb{E}\left[(\alpha \mathbb{I}\{u \geq 0\} - (1-\alpha)\mathbb{I}\{u < 0\})|u|\right],$$

по отношению к которой α-квантиль $\mathbb{Q}_\alpha\left(x_t|I_{t-1}\right)$ является оптимальным прогнозом. Во втором случае исследователю интересны квантили сами по себе, как, например, при анализе стоимостной меры риска (VaR, от англ. value at risk).

Основной способ расчёта условный квантилей — косвенный, через моделирование условной плотности (см. раздел 6) или её хвостов. В простейших случаях дело сводится к моделированию ещё более простых объектов. Например, если доходности подчиняются простому соотношению $r_t/\sigma_t \sim iid\,\mathcal{D}$, то $\mathbb{Q}_\alpha\left(r_t|I_{t-1}\right) = \phi_0 + \phi_\sigma\sigma_t$, где постоянные ϕ_0 и ϕ_σ зависят только от распределения $\mathcal{D}$. Значит, условные квантили можно смоделировать, смоделировав динамику условного стандартного отклонения и безусловную плотность стандартизованных доходностей.

Не так давно Engle & Manganelli (2004) предложили моделировать условные квантили непосредственно с помощью авторегрессии:

$$\mathbb{Q}_\alpha\left(r_t|I_{t-1}\right) = \xi + \alpha g(r_{t-1}, r_{t-2}, ...) + \beta\mathbb{Q}_\alpha\left(r_{t-1}|I_{t-2}\right),$$

где $g(r_{t-1}, r_{t-2}, ...)$ — движущая сила процесса, определённая функция от предыстории. Такая модель, будучи заточена на анализ стоимостной меры риска, имеет сокращение CAViAR (от англ. conditional autoregressive value at risk). В качестве движущей силы процесса авторы предлагают четыре варианта её спецификации, все, впрочем, не очень чётко мотивированные. Одна из простейших спецификаций подразумевает включение регрессоров r_{t-1}^- и r_{t-1}^+, где $a^- = \min\{a, 0\}$ и $a^+ = \max\{a, 0\}$.

Обобщение CAViAR-модели на многомерный случай содержится в White, Kim & Manganelli (2010). Kuester, Mittnik & Paolella (2006) сравнивают различные подходы в точности прогнозирования стоимостной меры риска.

5 Моделирование условных вероятностей

Иногда вместо моделирования поведения непрерывно распределённой переменной (такой как рост ВВП или доходность финансового актива) интересно моделировать факт превышения этой переменной какого-то порога (соответственно, превышение ростом определённого «естественного уровня» при анализе бизнес-циклов или положительность доходности при анализе направления движения рынка). В таком случае исследуемой переменной является индикатор, скажем, d_t, принимающий значения 1 или 0 (например, $d_t = \mathbb{I}\left\{\Delta y_t > \delta\right\}$ при известном пороге δ). Естественным кандидатом при моделировании этого объекта является аналог модели бинарного выбора, популярной в кросс-секционном анализе, пробита или логита. В данном контексте такие модели называют автопробитом и автологитом, соответственно, причём при анализе бизнес-циклов принято использовать пробит-версию, а при анализе направления движения финансовых рынков — логит-версию. Например, автологит выглядит следующим образом:

$$\mathbb{P}\left\{d_t = 1|I_{t-1}\right\} = \frac{1}{1 + \exp(-\theta_t)},$$

где переменная θ_t параметризуется как (обычно линейная) функция от наблюдаемых в предшествующий период переменных. Стильно использовать предыдущие индикаторы типа d_{t-1}, поскольку моделируются именно индикаторы, но включение переменных, лежащих в основе построения индикатора (Δy_t в примере выше), тоже возможно. См., например, модель Rydberg & Shephard (2003) для направления движения рынка от трансакции к трансакции. В работе Anatolyev (2009) предлагается обобщение автологит-модели на случай одновременного анализа направления движения нескольких рынков.

Аналогично можно строить авторегрессии для дискретных, но не бинарных переменных, например целочисленных. См., например, работу Rydberg & Shephard (2003) применительно к моделированию количества скачков цены от трансакции к трансакции.

6 Моделирование условных плотностей

Вся информация о динамике интересующей исследователя переменной содержится в эволюции её условного распределения. Мы будем считать последнее непрерывным и говорить о моделировании динамики условной плотности. Поскольку плотность — заметно более сложный объект, чем моменты, квантили или вероятности, моделировать её динамику значительно сложнее. С другой стороны, из оцененной адекватной модели условной плотности можно вывести динамику всех остальных, более простых объектов, например, условных квантилей (более того, это и есть ведущий способ их косвенного моделирования)[5].

Сначала моделируется условное среднее μ_t (например, как линейная авторегрессия), затем условное стандартное отклонение σ_t (например, в рамках GARCH-класса). Предполагается, что стандартизованные шоки $\eta = \sigma_t^{-1}(r_t - \mu_t)$ распределены, условно на предыстории, согласно (негауссовому) распределению $\mathcal{D}(\theta)$ со нулевым средним, единичной дисперсией и вектором дополнительных параметров θ, отвечающих за форму распределения (например, количество степеней свободы, показатель скошенности и т.д.). Чаще всего используется распределения Стьюдента или GED (от англ. generalized error distribution, введено в обиход в GARCH-контексте в работе Nelson 1991). Эти распределения способны подстроиться под типично тяжелохвостовое условное поведение финансовых доходностей, за что отвечает дополнительный параметр тяжести хвостов (для распределения Стьюдента этот параметр — количество степеней свободы). В то же время указанные распределения симметричны и поэтому не учитывают скошенность доходностей. Примерами распределений, учитывающими и асимметрию, являются скошенные распределения Стьюдента и GED. Первое используется довольно интенсивно, эту моду ввёл Брюс Хансен в Hansen (1994). Мы же здесь воспроизведём форму плотности скошенного GED-распределения (использованного, например, в работе Anatolyev & Shakin 2007):

$$f\left(\eta;\kappa,\varphi\right) = \frac{\kappa}{2\Lambda\Gamma(1/\kappa)} \exp\left(-\left|\frac{\eta-\Delta}{\Lambda\left(1+\operatorname{sgn}\left(\eta-\Delta\right)\varphi\right)}\right|^{\kappa}\right),$$

где Λ и Δ — определённые функции κ и φ, $\kappa > 0$ — параметр тяжести хвостов, а $-1 < \varphi < 1$ — дополнительный параметр скошенности.

Иногда в качестве параметрических плотностей используются такие, которые произошли из различных разложений. Пример подобной плотности — так называемое распределение Грама—Шарлье:

$$f\left(\eta;s,k\right) = \phi(\eta)\left(1 + \frac{s}{3!}(\eta^3 - 3\eta) + \frac{k-3}{4!}(\eta^4 - 6\eta^2 + 3)\right),$$

где $\phi(\eta)$ — плотность стандартного нормального распределения, а s и k — параметры скошенности и тяжести хвостов. Видно, что аналитическая форма напоминает некоторое усечённое разложение (произвольной плотности вокруг гауссовой). Данная плотность интегрируется в единицу, матожидание равно нулю, дисперсия — единице, а третий и четвёртый моменты — ровно s и k. Не всё, однако, здесь гладко. Строго говоря, такая плотность — не плотность, так как она может принимать отрицательные значения. Чтобы избавиться от этой неприятности, получила распространение другая форма, где выражение в скобках возводится в квадрат, и получившийся результат нормируется. Однако в результате такого преобразования утрачивается первоначальная идея разложения, и, в частности, портится интерпретация параметров s и k.

Описанные выше подходы обладают существенным недостатком: хотя форма распределения отличается от формы нормального распределения, динамика условных моментов высоких порядков (например, третьих и четвёртых) жёстко привязана к динамике условного

[5] Речь не идёт об условном среднем и условной дисперсии, поскольку обычно при моделировании плотности в первую очередь моделируются эти два объекта явным образом, см. ниже.

среднего и условной дисперсии. Чтобы разорвать эту связь, можно вектор параметров θ в спецификации распределения $\mathcal{D}(\theta)$ не сохранять постоянной величиной, а разрешить ему эволюционировать, параметризовав его как функцию от времени, а точнее, от наблюдаемой предыстории. Такой подход предложен в работе Hansen (1994), а соответствующая модель называется ARCD (от англ. autoregressive conditional density). Hansen (1994) показал на двух эмпирических примерах, как можно создать такую параметризацию. Для «плотности» Грам–Шарлье, приведённой выше, можно придумать уравнение динамики параметров s и k. Вполне естественно, что подобное занятие гораздо более затруднительно, чем придумать логичную спецификацию динамики первых двух условных моментов. Условное среднее и условная дисперсия обладают интуитивной интерпретацией в терминах предсказуемости и волатильности, а также обилием стилизованных фактов относительно связанных с ними явлений. Параметры же условного распределения чаще всего лишены такой интерпретации и стилизованных фактов. Кроме того, во многих случаях попытки эмпирически оценить такие спецификации приводят к статистически незначимым коэффициентам, сигнализируя о постоянстве во времени этих параметров или по крайней мере несущественной их динамике. Несмотря на это, ожидается, что финансовые эконометристы в ближайшее время уделят данному подходу больше внимания и обнаружат интересные и нетривиальные закономерности.

Описанный выше подход можно применить и к классу MEM-моделей (см. раздел 2), да и к другим классам, где фигурирует параметры, отвечающие за форму распределения (см., например, цитированную выше работу Anatolyev 2009).

При моделировании поведения динамики многомерных величин используются те же принципы; см. применение ряда многомерных плотностей в работе Балаев (2011). Здесь присутствуют, однако, дополнительные трудности. Одна из них — отсутствие в литературе обобщений некоторых сложных одномерных распределений на многомерный случай; существующие же обобщения нередко имеют очень сложное аналитическое представление. Вторая трудность типична для многомерного моделирования — это резкое увеличение количества параметров с ростом числа моделируемых переменных.

Обе проблемы помогает решить популярный инструментарий, получивший название копульного моделирования; см., например, обзор Patton (2012). Копула (англ. copula) представляет собой многомерное распределение $C(u_1, u_2, ..., u_p)$ с областью определения $[0,1]^p$, которое описывает зависимость компонент, а в качестве аргументов используются маржинальные распределения $F(y_j)$ компонент. В результате имеем многомерное распределение $C(F(y_1), F(y_2), ..., F(y_p))$. Таким образом, задача многомерного моделирования и оценивания распадается на p задач одномерного моделирования и оценивания распределений компонент и задачу моделирования и оценивания копулы. Копула обычно выбирается из многочисленных классов копул, обладающих наиболее подходящими свойствами, описывающими зависимость компонент. Параметров в спецификации копулы обычно мало, чем избегается перепараметризация модели. Более того, моделирование маржинальных компонент может быть непараметрическим, а моделирование копулы в то же самое время — параметрическим или полупараметрическим. Определённые трудности вызывает моделирование с помощью копул, когда некоторые (особенно если только некоторые) из компонент распределены дискретно; в качестве примера см. Anatolyev & Gospodinov (2010).

При моделировании условных плотностей принято тестировать модель на правильную спецификацию используемого семейства плотностей. С современным состоянием этой сферы можно ознакомиться в обзоре Хейфец (2011).

Список литературы

Анатольев, С. (2006). Тестирование на предсказуемость. *Квантиль* 1, 39–42.

Балаев, А. (2011). Моделирование многомерных параметрических плотностей финансовых доходностей. *Квантиль* 9, 39–60.

Белоусов, С. (2006). Моделирование волатильности со скачками: применение к российскому и американскому фондовым рынкам. *Квантиль* 1, 101–110.

Куан, Ч.-М. (2013). Модель с марковскими переключениями. *Квантиль* 11, 13–39.

Маккракен, М. (2006). Парные тесты на одинаковую точность прогнозов. *Квантиль* 1, 53–62.

Росси, Э. (2010). Одномерные GARCH-модели: обзор. *Квантиль* 8, 1–67.

Хейфец, И. (2011). Тестирование распределений. *Квантиль* 9, 25–34.

Цыплаков, А. (2010). Сделать тайное явным: искусство моделирования с помощью стохастической волатильности. *Квантиль* 8, 69–122.

Цыплаков, А. (2011). Введение в моделирование в пространстве состояний. *Квантиль* 9, 1–24.

Anatolyev, S. (2009). Multi-market direction-of-change modeling using dependence ratios. *Studies in Nonlinear Dynamics & Econometrics* 13, статья 5.

Anatolyev, S. & N. Gospodinov (2010). Modeling financial return dynamics via decomposition. *Journal of Business & Economic Statistics* 28, 232—245.

Anatolyev, S. & D. Shakin (2007). Trade intensity in the Russian stock market: dynamics, distribution and determinants. *Applied Financial Economics* 17, 87–104.

Baillie, R.T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5–59.

Balke, N.S. (2000). Credit and economic activity: credit regimes and nonlinear propagation of shocks. *Review of Economics and Statistics* 82, 344-–349.

Bauwens, L., S. Laurent & J.V.K. Rombouts (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21, 79–109.

Bollerslev, T. (2009). Glossary to ARCH (GARCH). Глава в *Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle* (под редакцией T. Bollerslev, J. R. Russell & M. Watson). Oxford University Press.

Brownlees C.T., F. Cipollini & G.M. Gallo (2012). Multiplicative error models. Глава в *Volatility Models and Their Applications* под редакцией L. Bauwens, C. Hafner & S. Laurent, Wiley.

Engle, R.F. & S. Manganelli (2004). CAViaR: conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics* 22, 367–381.

Franses, P.H., T. Teräsvirta & D. van Dijk. (2002). Smooth transition autoregressive models – a survey of recent developments. *Econometric Reviews* 21, 1–47.

Franses, P. & D. van Dijk (2000). *Nonlinear Time Series Models in Empirical Finance.* New York: Cambridge University Press.

Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357–384.

Hamilton, J.D. (1994). *Time Series Analysis.* Princeton University Press.

Hansen, B.E. (1997). Inference in TAR models. *Studies in Nonlinear Dynamics & Econometrics* 2, 1—14

Kuester, K., S. Mittnik & M.S. Paolella (2006). Value-at-risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics* 4, 53–89.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Journal of Econometrics* 43, 227–251.

Patton, A. (2012). A review of copula models for economic time series. *Journal of Multivariate Analysis* 110, 4–18.

Rydberg, T.H. & N. Shephard (2003). Dynamics of trade-by-trade price movements: decomposition and models. *Journal of Financial Econometrics* 1, 2–25.

Teräsvirta, T. (2009). An introduction to univariate GARCH models. Глава в *Handbook of Financial Time Series* (под редакцией T. Andersen, R. Davis, J.-P. Kreiss & T. Mikosch). Springer.

White, H., T.-H. Kim & S. Manganelli (2010). Modeling autoregressive conditional skewness and kurtosis with multi-quantile CAViaR. Глава в *Volatility and Time Series Econometrics: A Festschrift in Honor of Robert F. Engle* под редакцией J. Russell & M. Watson, Oxford University Press.

Objects of nonstructural time series modeling

Stanislav Anatolyev

New Economic School, Moscow, Russia

When modeling time series dynamics one has to decide on the class and type of models to use which depends much on the object to be modeled. This essay briefly overviews the specifics of time series modeling of various objects like conditional mean, conditional variance, conditional quantiles, conditional probabilities and conditional densities. We pay attention to both univariate and multivariate cases. References to narrower but more detailed surveys are given.