

квантиль

*международный эконометрический журнал
на русском языке*

**№12
февраль 2014 г.**

СОДЕРЖАНИЕ ВЫПУСКА

Приглашение на конференцию

Международная конференция «Современный эконометрический
инструментарий и приложения»

Эконометрический ликбез: динамические стохастические модели общего равновесия

Микушева Анна. Оценивание динамических стохастических моделей общего равновесия	1
Джонс Коллум, Кулиш Мариано. DSGE-моделирование в пакете Dynare: практическое введение	23

В помощь изучающим эконометрику

Цыплаков Александр. Мини-словарь англоязычных эконометрических терминов, часть 3	45
---	----

Статьи: финансовая эконометрика

Грошев Олег. Непостоянные во времени выходящие копулы в многомерном анализе доходностей	53
Франгуриди Григорий. Динамика условных моментов высоких порядков и прогнозирование стоимостной меры риска	69

Некролог

Александр Иванович Саичев

Квантиль

№12, февраль 2014 г.

ISSN 2309-3013

Сайт в Интернете: <http://quantile.ru>

Адрес электронной почты: quantile@quantile.ru

Доступ к журналу бесплатный и неограниченный

РЕДАКТОР

Станислав Анатольев

Российская Экономическая Школа (Москва, Россия)

РЕДАКЦИОННЫЙ СОВЕТ

Виктория Зинде-Уолш

Университет МакГилл (Монреаль, Канада)

Рустам Ибрагимов

Имперский колледж (Лондон, Великобритания)

Анна Микушева

Массачусетский технологический институт (Кэмбридж, США)

Алексей Онацкий

Кэмбриджский университет (Кэмбридж, Великобритания)

Александр Цыплаков

Новосибирский государственный университет (Новосибирск, Россия)

Виктор Черножуков

Массачусетский технологический институт (Кэмбридж, США)

К СВЕДЕНИЮ АВТОРОВ

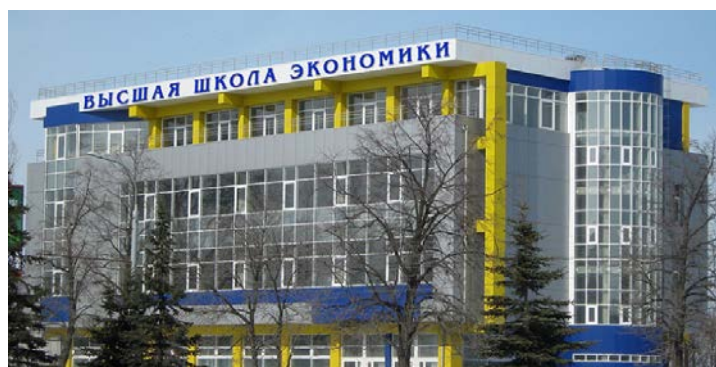
Рукописи для публикации в разделе «Статьи» принимаются в электронном виде по адресу submit@quantile.ru. Работы могут принадлежать любой прикладной сфере экономической науки. Главным требованием является интенсивное использование адекватных эконометрических методов. Рукопись должна быть написана на русском (для русскоязычных авторов) или на английском (для остальных авторов) языке в формате *Microsoft Word* или (предпочтительнее) *LaTeX*, и по объему не превышать 30 страниц формата А4 с двойным междустрочным интервалом. Работы подвергаются контролю качества членами редакционного совета и независимыми референтами. Перспективная работа может быть при необходимости возвращена автору на доработку. Редакция также приглашает к сотрудничеству экспертов по эконометрике, готовых внести вклад в методологические рубрики журнала.

При публикации статьи или методологического эссе в журнале «Квантиль» передача авторских прав не происходит ни полностью, ни частично.

**Международная конференция
«Современный эконометрический инструментарий и его приложения»**



Высшая школа экономики, Российская экономическая школа и журнал «Квантиль» приглашают всех любителей эконометрики принять участие в международной конференции «Современный эконометрический инструментарий и приложения», которая состоится в Нижнем Новгороде 18–20 сентября 2014 года на базе нижегородского филиала ВШЭ.



В конференции примут участие учёные-эконометристы из разных стран мира. Предусматриваются регулярные сессии по самым разным эконометрическим темам, как теоретическим, так и прикладным, а также гостевые мастер-классы по некоторым важным вопросам инструментария современной эконометрики. Рабочий язык конференции — английский.

Статьи следует присылать по электронному адресу ec2014@hse.ru до 31 марта 2014 г. Решения по включению в программу будут приняты до 1 мая 2014 г. При желании посетить конференцию без доклада следует зарегистрировать своё присутствие до 1 июня 2014 г.

Подробная информация о конференции находится на сайте ec2014.hse.ru.

Эконометрический ликбез: динамические стохастические модели общего равновесия

Оценивание динамических стохастических моделей общего равновесия*

Анна Микушева[†]

Массачусеттский Технологический Институт, Кембридж, США

Данное эссе является обзором основных эконометрических методов оценивания динамических стохастических моделей общего равновесия, широко используемых современными центральными банками и федеральными резервными системами. Статья посвящена обсуждению основных эконометрических проблем, возникающих при статистическом оценивании и формировании статистических выводов о параметрах линеаризованных моделей общего равновесия. В статье мы рассматриваем три основных метода оценивания динамических моделей — метод наименьшего расстояния, метод максимального правдоподобия и Байесовский метод. Особое внимание уделено проблемам, которые возникают в связи со скудностью макроэкономических данных. Вопросы макроэкономического моделирования и вопросы решения стохастических динамических систем, хотя и являются исключительно интересными и сложными, находятся за рамками данного эссе.

1 Динамические стохастические модели общего равновесия

Последнее десятилетие принесло большие изменения в эмпирической практике макроэкономических исследований. Произошел качественный сдвиг от использования векторных авторегрессионных моделей (VAR) к массовому использованию динамических стохастических моделей общего равновесия (DSGE). Модели DSGE являются современной попыткой использовать разумные структурные ограничения, накладываемые экономической теорией, при оценивании взаимоотношений между различными ведущими макроиндикаторами. Большим прорывом последнего десятилетия стало развитие численных методов решения динамических систем, которые позволили решать сложные теоретические модели численными методами и тем самым сократили разрыв между теоретическими моделями, которые можно решить, и моделями, которые можно использовать при оценивании.

Основная разница между VAR моделями и DSGE моделями в том, что первые являются достаточно общими линейными моделями, накладывающими минимальные структурные ограничения на взаимодействия между входящими макроэкономическими переменными. В противовес им DSGE модели являются полностью параметризованными моделями с микроэкономическим фундаментом, то есть основанными на полной спецификации действий и поступков всех агентов в экономике (с описанием их функций полезности) а также условий равновесия на всех рынках. Условия максимизации полезности всеми агентами и условия равновесия всех рынков накладывают очень сильные ограничения на возможные взаимодействия между макропеременными и устанавливают функциональные зависимости между

*Цитировать как: Микушева, Анна (2014). «Оценивание динамических стохастических моделей общего равновесия», Квантиль, №12, стр. 1–21. Citation: Mikusheva, Anna (2014). “Estimation of dynamic stochastic general equilibrium models,” *Quantile*, No.12, pp. 1–21.

[†]Адрес: 50 Memorial Drive, Cambridge, MA-02138, USA. Электронная почта: amikushe@mit.edu

коэффициентами приведенной формы. Так например, многие линеаризованные DSGE модели могут быть записаны в форме VAR, однако, коэффициенты соответствующего VAR представления будут не свободными, а строго функционально связанными между собой параметрами. Вопрос о том, когда DSGE модель может быть записана в VAR форме с ограничениями, обсуждается в статье Fernandez-Villaverde, Rubio-Ramirez, Sargent & Watson (2007).

Более жесткая параметризация DSGE моделей имеет свои преимущества. Во-первых, дополнительные ограничения позволяют надеяться на более эффективные оценки всех структурных взаимоотношений. Во-вторых, спецификация функций полезности агентов позволяет изучать эффекты различных типов макроэкономических вмешательств и макроэкономической политики на благосостояние агентов, что невозможно сделать в VAR моделях. И в третьих, дополнительные ограничения на взаимодействия переменных дают надежду на более качественные прогнозы. Основными недостатками DSGE моделей является повышенная вероятность неверной спецификации и сложности с нелинейным оцениванием. DSGE модели легче фальсифицируемы, так как накладывают больше ограничений на данные. Некоторые статьи делающие эмпирические сравнения DSGE и VAR моделей — Chari, Kehoe & McGrattan (2005), Christiano, Eichenbaum & Vigfusson (2006) и Del Negro, Schorfheide, Smets & Wouters (2007).

Статистическое оценивание DSGE модели является многоступенчатым процессом, который включает в себя процесс моделирования (написание динамической макромоделли общего равновесия), решение модели, приведения решения к виду, подходящему для статистического оценивания, и, наконец, эконометрическое оценивание. В этом эссе мы обсудим только последний шаг, а именно, только эконометрическую составляющую этого процесса. Следующая глава дает краткое описание доэконометрических стадий.

На данный момент существует три основных подхода к оцениванию DSGE моделей, два из них классических и один Байесовский. Классические подходы используют либо обобщенный метод моментов и метод наименьшего расстояния, включая случаи, когда для вычисления целевой функции приходится использовать симуляции, либо метод максимального правдоподобия. Байесовское оценивание полагается главным образом на MCMC алгоритм, который является численным методом для получения постериорного распределения. Ниже мы обсудим преимущества различных методов, особенности их применения, а также сложности, с которыми сталкиваются макроэконометристы.

В последнее время было выпущено немало книг и обзорных статей, которые позволяют ознакомиться с теорией и практикой оценивания DSGE моделей. Я бы хотела обратить внимание читателей на книги Canova (2007) и DeJong & Dave (2007), а также лекции Ларри Кристиано,¹ которые являются достаточно полными и интересными источниками по данной теме. Проблеме оценивания DSGE моделей также посвящена обзорная лекция Джима Стока на летней школе NBER в 2008.² Другим интересным источником является лекция Франка Шорфалда (Schorfheide, 2010) на мировом конгрессе, посвященная нерешенным эконометрическим проблемам DSGE моделей.

Для практического использования DSGE моделей я рекомендую ознакомиться с бесплатной софтверной программой Dynare³, которая является надстройкой в пакете MatLab. Программа Dynare позволяет автоматически линеаризовать и решать стандартные DSGE модели, а также стандартно реализовать оценивание методом максимального правдоподобия и Байесовским методом. Dynare поддерживается большим числом академических ученых, а также исследователями, работающими в центральных банках и резервных системах различных стран. Сообщество Dynare поддерживает активный форум, где можно получить ответы на практические вопросы, связанные с оцениванием DSGE моделей.

¹Доступны на веб-сайте faculty.wcas.northwestern.edu/~lchrist/

²www.nber.org/minicourse/_2008.html

³www.dynare.org

2 Стадии, предшествующие эконометрическому анализу

2.1 Моделирование и линеаризация модели

Типичная задача макроэкономического моделирования начинается с описания всех агентов в экономике и их функций полезности или оптимизационных задач. Такими агентами могут быть домохозяйства, которые потребляют конечный продукт, предлагают свои услуги на рынке труда, инвестируют свои сбережения, и, например, фирмы, которые нанимают рабочих, инвестируют в капитал и продают конечную продукцию. Необходимо также описать все бюджетные ограничения, накладываемые на агентов. Делаются также предположения о различных рынках и их несовершенствах. Модель может содержать (и обычно содержит) стохастическую составляющую, которая заключается в описании всех структурных шоков в экономике и каким образом они воздействуют на различные переменные. Построение теоретической макроэкономической модели является важным шагом и полностью находится в ведении макроэкономики. В данной статье мы откажемся от обсуждения различных моделей и их предположений.

На этом этапе наша задача заключается в получении системы уравнений, состоящей из условий первого порядка для оптимизационных задач агентов, уравнений балансов, условий равновесия на всех рынках и уравнений, описывающих динамику всех шоков в экономике. Данная система уравнений описывает динамические зависимости между всеми макропеременными. К сожалению, в подавляющем большинстве случаев мы не можем разрешить данную систему уравнений из-за следующих сложностей: во-первых, многие из уравнений нелинейны, во-вторых, многие уравнения включают условные математические ожидания будущих значений переменных.

Для разрешения первой сложности большинство статей решает не изначальную модель, а ее приближение, полученное в результате лог-линеаризации уравнений вокруг стационарного состояния (steady state). Впрочем, в настоящее время появляются исследования использующие также нелинейные модели; примером статьи, решающей нелинейную DSGE модель, может служить Fernandez-Villaverde & Rubio-Ramirez (2005).

В данном эссе в качестве примера мы будем использовать лог-линеаризованную DSGE модель, предложенную в статье Galí, Lopez-Salido & Valles (2003). Это одна из простейших моделей, на которой построены более сложные современные DSGE модели, включающие в себя большее число макро переменных и позволяющие оценивать большее число структурных параметров.

Пример 1. Предположим, что переменная x_t обозначает логарифм совокупной продукции (и совокупного потребления) в экономике в момент t , переменная π_t обозначает логарифм инфляции, а переменная r_t — реальный процент. Эволюция этих трех макропеременных описывается следующей лог-линеаризованной системой уравнений:

$$\begin{cases} -(r_t - \mathbb{E}_t \pi_{t+1} - \rho \Delta a_t) + \mathbb{E}_t x_{t+1} - x_t = 0, \\ b \mathbb{E}_t \pi_{t+1} + \kappa x_t - \pi_t + \varepsilon_t = 0, \\ \lambda r_{t-1} + (1 - \lambda) \varphi_\pi \pi_t + (1 - \lambda) \varphi_x x_t + u_t = r_t. \end{cases} \quad (1)$$

Первое уравнение в этой системе является линеаризованным уравнением Эйлера, или условием первого порядка в задачи оптимизации полезности потребителем. Второе уравнение в системе (1) — это так называемая кривая Филлипса (Phillips curve). Это уравнение получается как решение модели Калво (Calvo model), в которой фирмы не могут свободно устанавливать цены на свою продукцию каждый период, а меняют цены только в экзогенно заданные случайные моменты времени. Третье уравнение имеет название «правило Тейлора» (Taylor Rule) и является характеристикой монетарной политики, проводимой Федеральной резервной системой США. Символ \mathbb{E}_t обозначает условное математическое ожидание,

использующее всю информацию, известную в момент t ; часто такую парадигму называют «рациональными ожиданиями» (rational expectations).

Помимо эндогенно меняющихся макропеременных, чья эволюция описана в системе (1), в систему также входят экзогенные переменные, являющиеся шоками. Таковыми в данной системе являются технологический шок Δa_t , монетарный шок u_t и шок потребления ε_t . Дополнительно к системе (1) мы должны описать динамическую эволюцию шоков. Предположим, что шоки удовлетворяют следующим допущениям:

$$\Delta a_t = \rho \Delta a_{t-1} + \varepsilon_{a,t}; \quad u_t = \delta u_{t-1} + \varepsilon_{u,t}; \quad (\varepsilon_t, \varepsilon_{a,t}, \varepsilon_{u,t})' \sim i.i.d. N(0, \Sigma); \quad \Sigma = \text{diag}(\sigma^2, \sigma_a^2, \sigma_u^2).$$

В рассматриваемой модели имеется 10 параметров, обычно называемых структурными параметрами. Часть из них имеет экономическое значение, такие например, как фактор дисконтирования (discount rate) b , параметер Калво κ , и параметры правила Тейлора φ_x, φ_π и λ . Остальные параметры, $\delta, \rho, \sigma, \sigma_a, \sigma_u$, статистически описывают эволюцию экзогенных переменных. Обозначим набор всех неизвестных структурных параметров β . Задача эконометрического оценивания и тестирования состоит в том, чтобы, имея данные $\mathcal{Y}_T = \{x_t, \pi_t, r_t, t = 1, \dots, T\}$, оценить, построить доверительные интервалы и протестировать интересующие нас гипотезы о структурных параметрах модели β .

2.2 Решение системы рациональных ожиданий

Модель (1) является линейной, но при этом содержит условные математические ожидания будущих значений переменных. Упомянутые условные ожидания ненаблюдаемы, и с этой точки зрения модель (1) неудобна для оценивания. Следующим шагом является решение линейной системы рациональных ожиданий. В общем виде модель, стандартно получаемую в результате лог-линеаризации DSGE модели, можно записать как

$$\Gamma_0(\beta)Z_t = \Gamma_1(\beta)\mathbb{E}_t Z_{t+1} + \Gamma_2(\beta)Z_{t-1} + \Gamma_3(\beta)\eta_t,$$

где Z_t содержит все эндогенные переменные (как наблюдаемые, так и не наблюдаемые), а η_t содержит все экзогенные шоки. Матрицы $\Gamma_i(\beta)$ являются известными функциями структурных параметров β .

Решить линейную модель рациональных ожиданий означает записать ее в виде стандартной линейной модели пространства состояний (state-space representation), состоящей из двух уравнений

$$\begin{cases} Y_t = \Psi_0(\beta) + \Psi_1(\beta)t + \Psi_2(\beta)S_t, \\ S_t = \Phi_1(\beta)S_{t-1} + \Phi_2(\beta)\eta_t. \end{cases} \quad (2)$$

Здесь S_t — это вектор переменных, описывающих состояние динамической системы (state variable), а Y_t — набор переменных, наблюдаемых эконометристом в момент t . Первое из уравнений системы (2) называется уравнением измерения (measurement equation), а второе — уравнением перехода (transition equation). В случае Примера 1, $S_t = (\pi_t, x_t, r_t, \Delta a_t, u_t, \varepsilon_t)$.

Существует целый ряд различных алгоритмов, позволяющих численно решать линейные системы с рациональными ожиданиями; наиболее известны из них методы, предложенные в статьях Blanchard & Kahn (1980), Anderson & Moore (1985), Klein (2000) и Sims (2001). Большинство из вышеупомянутых методов запрограммированы и являются частью стандартного пакета Dynare.⁴ Следует сказать, что решается система численно, а не аналитически. Это значит, что для каждого значения параметра β можно численно посчитать соответствующие матрицы $\Gamma_i(\beta)$ и решить систему рациональных ожиданий, т.е. численно найти значение матриц $\Psi_i(\beta)$ и $\Phi_i(\beta)$.

⁴www.dynare.org

Проблема решения системы рациональных ожиданий очень интересна и сложна. Для некоторых значений параметров система не имеет решения, а в некоторых случаях решений несколько, что соответствует нескольким потенциальным равновесиям. Стандартная на данный момент методология состоит в ограничении области определения параметра только теми значениями, при которых существует только единственное устойчивое решение линейной системы ожиданий. В данной статье мы не будем обсуждать как решать системы рациональных ожиданий, а предположим, что DSGE модель уже линеаризована и решена в виде (2).

3 Методы оценивания: метод наименьшего расстояния

Наиболее ранним является классический метод оценивания DSGE моделей, основанный на методе наименьшего расстояния и впервые в данном контексте примененный в статье Rotemberg & Woodford (1997); примеры более современного использования метода можно найти в статьях Voivin & Giannoni (2006), Christiano, Eichenbaum & Evans (2005) и Schorfheide (2000). Альтернативными названиями в эконометрической литературе являются английские термины *matching moments* и *indirect inference*.

Метод состоит в минимизации расстояния между некоторыми эмпирически значимыми характеристиками процесса оцененными на данных и полученными из теоретической модели. Чаще всего такими характеристиками являются функция импульсного отклика (*impulse response function*), автоковариационная и спектральная функции процесса.

Пусть β_0 является истинным (неизвестным) значением параметра β , которое мы пытаемся оценить. Предположим, что характеристика процесса, которую мы будем приближать, скажем $\delta = \delta(\beta)$, зависит от структурных параметров модели β . Пусть $\delta_0 = \delta(\beta_0)$ является истинным значением характеристики для наблюдаемого процесса, а $\hat{\delta} = \hat{\delta}(\mathcal{Y}_T)$ является состоятельной оценкой этой характеристики, основанной на доступных данных \mathcal{Y}_T . Предположим, что данная оценка удовлетворяет условию

$$\sqrt{T}(\hat{\delta} - \delta_0) \Rightarrow N(0, \Omega).$$

Оценка $\hat{\beta}$ по методу наименьшего расстояния определяется как

$$\hat{\beta} = \arg \min_{\beta} (\hat{\delta} - \delta(\beta))' \Xi (\hat{\delta} - \delta(\beta)),$$

где Ξ — положительно определенная симметричная матрица. В классических предположениях, что структурный параметр β идентифицируем, значение β_0 является внутренней точкой пространства параметров, что матрица первых производных $R = \partial\delta(\beta_0)/\partial\beta$ имеет полный ранг, равный размерности параметра β , и что матрица $R'\Xi R$ имеет обратную, классический результат гарантирует, что оценка $\hat{\beta}$ состоятельна и

$$\sqrt{T}(\hat{\beta} - \beta_0) \Rightarrow N(0, (R'\Xi R)^{-1} R'\Xi \Omega \Xi R (R'\Xi R)^{-1}).$$

Оптимальным выбором матрицы взвешивания разных характеристик является $\Xi = \Omega^{-1}$, при котором асимптотическая матрица ковариаций равна $(R'\Omega^{-1}R)^{-1}$.

Целевая функция, используемая для оценивания, может также использоваться для тестирования спецификации в случае, когда размерность характеристики δ выше размерности структурного параметра β . В частности в предположении правильной спецификации мы можем показать, что статистика $\min_{\beta} (\hat{\delta} - \delta(\beta))' \Omega^{-1} (\hat{\delta} - \delta(\beta))$ асимптотически сходится к χ^2 -распределению с числом степеней свободы, равным разности между размерностями параметров δ и β .

Метод наименьшего расстояния хорошо изучен в эконометрике. Здесь мы обсудим только особенности его применения для оценивания DSGE моделей. Метод наименьшего расстояния использует для оценивания только часть возможных статистических связей между переменными, вытекающих из модели (только те связи, которые описаны характеристиками δ , используемыми для оценивания). В этом смысле метод наименьшего расстояния является процедурой с ограниченной информацией (limited information procedure). Из этого факта проистекают как преимущества, так и недостатки. Основным недостатком является потеря эффективности по сравнению с методом максимального правдоподобия, который использует всю совокупность статистических связей наиболее эффективным образом. Принимая во внимание, что даже более эффективный метод максимального правдоподобия страдает от проблемы слабой идентификации, которую мы обсудим в деталях в следующей секции, метод наименьшего расстояния еще более уязвим к этой проблеме. С другой стороны, для задания функции правдоподобия необходимы предположения о том, как распределены шоки в системе, такие как предположения о гауссовости в Примере 1. Для метода наименьшего расстояния предположения о распределениях не нужны, и таким образом, такой метод оценивания более робастен к неверной спецификации модели.

Идея оценивания путем минимизации расстояния между важными характеристиками наблюдаемыми в данных и предсказанными моделью была предложена в статье Kydland & Prescott (1982); в частности и за это авторы были награждены Нобелевской премией по экономике. Kydland & Prescott (1982) были против методов, основанных на правдоподобии из соображений что в случае максимального правдоподобия наше понимание того, хорошо ли описывает модель данные, основывается на статистических критериях, которые не всегда совпадают с практическими и эмпирически значимыми целевыми функциями. Например, если некая модель построена для объяснения эффектов монетарной политики на макроиндикаторы, то разумно использовать в качестве теста спецификации целевую функцию, измеряющую расстояние между функциями импульсного отклика оцененными в данных и предсказанными моделью. Этот тест может противоречить тесту, основанному на функции правдоподобия, так как на последнюю могут влиять много факторов, плохо объясняемых моделью, большая часть из которых нам может быть неинтересна.

3.1 Оценивание характеристики δ на данных: метод

Заметим, что если DSGE модель решена в виде системы уравнений (2), то вычисление для любого значения параметра β теоретических (предписанных моделью) автоковариационной и спектральной функций для наблюдаемых переменных, а также функций импульсного отклика является относительно простой задачей. Действительно, например, чтобы получить теоретическое значение импульсного отклика Y_{t+h} на единичный шок η_t при значении структурного параметра β , достаточно положить $S_{t-1} = 0$, $\eta_{t+s} = 0$ для всех $s > 0$ и сравнить значения Y_{t+h} , получаемые из системы (2) при $\eta_t = 1$ и $\eta_t = 0$.

Основная сложность метода состоит в оценивании соответствующих характеристик δ на данных \mathcal{U}_T . Остановимся на наиболее частом случае, когда характеристика δ , на которой основано оценивание, есть функция импульсного отклика на структурные шоки в экономике. В таком случае оценка $\hat{\delta}$ обычно получается как результат оценивания структурной VAR модели. Оценивание структурной VAR модели состоит из двух шагов. На первом шаге оценивается приведенная VAR модель, т.е. производится регрессия по методу наименьших квадратов переменных $Y_t = (x_t, \pi_t, r_t)$ на их лаги:

$$Y_t = A_0 + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + e_t. \quad (3)$$

Используя данную регрессию, мы можем получить функцию импульсного отклика для при-

веденной модели:

$$\theta_{i,h} = \frac{\partial Y_{t+h}}{\partial e_{t,i}} = A_1 \theta_{i,h-1} + \dots + A_p \theta_{i,h-p}.$$

В данном случае мы используем предположение что $\theta_{i,h} = 0$ для всех $h < 0$ и $\theta_{i,h} = 1$ для $h = 0$. Пусть Θ_h обозначает матрицу, содержащую все импульсные отклики всех наблюдаемых переменных Y_{t+h} на все шоки e_t приведенной модели (на горизонте h). Проблема заключается в том, что шоки e_t не являются структурными шоками. Вторым шагом в оценивании структурной VAR модели является оценивание матрицы C , связывающей шоки приведенной модели со структурными шоками $\eta_t = Ce_t$. Функция импульсного отклика структурной модели δ_h связана с функцией импульсного отклика приведенной модели θ_h линейно: $\delta_h = \theta_h C^{-1}$.

Для оценивания импульсного отклика на структурные шоки требуется наложить дополнительные ограничения, позволяющие идентифицировать матрицу C . Данные ограничения должны быть согласованы со структурной DSGE моделью. Как оказывается, существует достаточная свобода в выборе условий идентификации, и конечный результат, а именно оценка функции импульсного отклика, очень чувствителен к этому выбору. Для ознакомления с теорией структурных VAR моделей я советую читателю ознакомиться со статьей Нобелевского лауреата Sims (1980). Пример наложения краткосрочных идентификационных условий можно найти в King, Plosser, Stock & Watson (1991), а пример долгосрочных идентификационных условий — в Blanchard & Quah (1989).

3.2 Оценивание характеристики δ на данных: проблемы

Существует несколько теоретических и практических сложностей при использовании метода наименьшего расстояния путем мэтчинга функции импульсного отклика, что делает этот метод проблематичным на практике.

Первая сложность заключается в том, что стандартная DSGE модель, записанная в виде системы (2), вообще говоря не является VAR моделью конечного порядка, она является моделью типа VARMA, т.е. у нее есть компонента типа скользящего среднего. Таким образом, стандартная теоретическая DSGE модель может быть записана как VAR модель бесконечного порядка (включающая в себя бесконечное число лагов). В этом смысле оценивание структурной VAR модели, описанное выше, является оцениванием неверно специфицированной модели. Как показали Chari, Kehoe & McGrattan (2005), погрешность вносимая неверной спецификацией динамики процесса, может быть огромной, и оценка импульсного отклика, полученная в результате оценивания структурной VAR модели, будет до неузнаваемости отличаться от истинной оцениваемой функции импульсного отклика.

Вторая сложность состоит в том что оценки функции импульсного отклика являются смещенными в конечной выборке, даже если абстрагироваться от факта неверной спецификации VAR модели. Смещенность оценки имеет несколько истоков, один из них — смещенность в конечной выборке оценки VAR коэффициентов редуцированной VAR модели. Оценивание методом наименьших квадратов системы временных рядов типа (3) обычно дает смещенную оценку, впрочем смещение для оценок VAR коэффициентов не очень велико. Однако, если хотя бы один из характеристических корней близок к единичному корню, то это небольшое смещение оценок коэффициентов транслируется в существенное смещение оценки функции импульсного отклика (см. Mikusheva, 2012). Другой источник смещения оценки функции импульсного отклика заключается в том, что импульсный отклик является очень сильно нелинейной функцией VAR коэффициентов, а любое нелинейное преобразование несмещенных оценок неминуемо влечет появление смещения в конечных выборках. Попыткой уменьшить влияние смещения оценки импульсного отклика на структурную оценку β , получаемую ме-

тодом наименьшего расстояния, является коррекция Когли–Нейсона, предложенная в статье Cogley & Nason (1995).

Наконец, дополнительной практической сложностью является выбор характеристик процесса δ , на основе которых строится оценка наименьшего расстояния. Предположим, что мы решили минимизировать расстояние между теоретической функцией импульсного отклика и ее оценкой в данных. Возникает вопрос, какие именно импульсные отклики использовать: отклики каких переменных на какие шоки и на каком горизонте. Использование каждого дополнительного импульсного отклика может нести дополнительную информацию о структурном коэффициенте β , но в то же время это вносит дополнительную неопределенность так как оценка этого отклика в данных имеет ошибку. Оптимальный выбор характеристик для оценки по методу наименьшего расстояния должен балансировать эти два фактора. В статье Hall, Inoue, Nason & Rossi (2012) предложен информационный критерий, который помогает сделать данный выбор.

4 Методы оценивания: Метод максимального правдоподобия

Второй часто используемый классический метод оценивания DSGE моделей — метод максимального правдоподобия. Примеры использования метода максимального правдоподобия для оценивания DSGE моделей можно найти в статьях Ingram, Kocherlakota & Savin (1994), Ireland (2004) и Lindé (2005).

Метод состоит в нахождении функции правдоподобия $\mathcal{L}(\mathcal{Y}_T|\beta)$ и последующей ее максимизации по отношению к параметру β . Значение $\hat{\beta}_{MLE}$ параметра β , в котором достигается максимум, называется оценкой максимального правдоподобия. При некоторых условиях регулярности модели оценка максимального правдоподобия является состоятельной и асимптотически нормальной:

$$\sqrt{T}(\hat{\beta}_{MLE} - \beta_0) \Rightarrow N(0, \mathcal{J}^{-1}), \quad (4)$$

здесь \mathcal{J} обозначает так называемую информацию Фишера. В регулярных случаях состоятельной оценкой информации Фишера является наблюдаемая информация, равная отрицательной второй производной логарифма функции правдоподобия:

$$\hat{J} = -\frac{1}{T} \frac{\partial^2}{\partial \beta \partial \beta'} \log \mathcal{L}(\mathcal{Y}_T|\hat{\beta}_{MLE}). \quad (5)$$

Теория максимального правдоподобия также позволяет тестировать различные гипотезы, используя один из трех классических тестов — тест Вальда, тест основанный на мультипликаторе Лагранжа (Lagrange multiplier test или score test) и тест отношения правдоподобия (likelihood ratio test).

Условия регулярности, при которых верны указанные выше асимптотические утверждения в случае временных рядов, менее тривиальны, чем классическая теория максимального правдоподобия, верная для независимых одинаково распределенных наблюдений. Хорошим учебником по данной теме является Hall & Heyde (1980).

Если DSGE модель линеаризована и решена в виде модели пространства состояний (2), то функция правдоподобия может быть получена при помощи применения фильтра Калмана. Хорошим учебником, освещающим теорию фильтрации является Harvey (1989). Фильтр Калмана в применении к линеаризованной DSGE модели реализован в программе Dynare. В случае, когда модель нелинейна, но решена в виде модели в пространстве состояний, проблема получения функции правдоподобия гораздо более сложна, обзор возможных методов может быть найден в статье Цыплакова (2011). В частности, один из часто употребляемых методов для получения функции правдоподобия — это использование так называемых корпускулярных фильтров (particle filters).

4.1 Проблемы метода максимального правдоподобия

Теперь мы обсудим особенности и проблемы применения метода максимального правдоподобия к оцениванию DSGE моделей.

Первой проблемой, с которой чаще всего сталкивается прикладной исследователь, является проблема численного поиска максимума функции правдоподобия. Большинство известных и решенных на данный момент DSGE моделей обладают крайне нерегулярной функцией правдоподобия. Типичная функция правдоподобия для DSGE модели имеет множество локальных максимумов, участки типа «плато», где функция почти не меняется, а также области резкого изменения производной («обрывы»). Все это, а также тот факт, что типичная DSGE модель обычно имеет большое число структурных параметров (наша простейшая модель в Примере 1 имеет 10 неизвестных параметров), осложняет численную задачу поиска оценки максимального правдоподобия.

В большинстве случаев использование численных оптимизаторов, основанных на градиентных методах, не позволяют найти максимум функции правдоподобия. Практические рекомендации включают в себя предложение использовать неградиентные методы (например, симплекс-метод) и использовать оптимизационный алгоритм несколько раз, стартуя из разных начальных точек, в надежде найти глобальный (а не локальный) максимум.

Другой частой проблемой является то, что исследователь редко догматично верит в модель, и скорее рассматривает ее как аппроксимацию к действительности. В этом смысле оцениваемая модель зачастую неверна буквально, т.е. неверно специфицирована. Таким образом, метод максимального правдоподобия более чувствителен к проблемам неверной спецификации по сравнению с обсуждаемым выше методом наименьшего расстояния. Метод максимального правдоподобия использует абсолютно все связи, налагаемые моделью на наблюдаемые переменные, а также требует спецификацию распределения всех шоков (гауссовость в Примере 1). В этом смысле правильнее рассматривать функцию правдоподобия не как истинную, а как функцию квази-правдоподобия. Эта идея была введена в статье White (1982). При некоторых условиях регулярности оценка максимального правдоподобия является состоятельной оценкой псевдоистинного параметра β^* . К сожалению, в сложных моделях связь между псевдоистинным параметром β^* и изначальным параметром β неочевидна, что усложняет интерпретацию результатов такого оценивания.

White (1982) также указывает, что если рассматривать оценку максимального правдоподобия с точки зрения потенциально неверной спецификации, то следует использовать стандартные ошибки, робастные к неверной спецификации. Робастная формула для оценки асимптотической матрицы вариаций \mathcal{J}^{-1} в утверждении типа (4) дана формулой $\hat{J}^{-1}\hat{I}\hat{J}^{-1}$, где

$$\hat{I} = -\frac{1}{T} \sum_{t=1}^T \left(\frac{\partial}{\partial \beta} \log f(Y_t | \mathcal{Y}_{t-1}, \hat{\beta}_{MLE}) \right) \left(\frac{\partial}{\partial \beta} \log f(Y_t | \mathcal{Y}_{t-1}, \hat{\beta}_{MLE}) \right)', \quad (6)$$

а \hat{J} задана формулой (5).

Дополнительной проблемой оценивания DSGE моделей методом максимального правдоподобия является проблема стохастической сингулярности. Если представить себе DSGE модель, в которой число случайных ненаблюдаемых шоков меньше числа наблюдаемых переменных, то часто одним выводом модели может быть наличие некой детерминистической (то есть выполняемой во всех реализациях данных) связи между наблюдаемыми переменными. Статья Ingram, Kocherlakota & Savin (1994) дает очень простой пример такой ситуации. Однако, подобная детерминистическая связь между наблюдаемыми переменными практически никогда не верна в реальных данных, и как результат, модель является неверно специфицированной и будет отвергнута в большой выборке. Из данного рассуждения следует частая практическая рекомендация (которая сама по себе все же не гарантирует отсутствие

стохастической сингулярности), что число шоков в DSGE модели должно быть не меньше числа наблюдаемых переменных.

Существует два подхода к разрешению проблемы сингулярности. Один из них состоит в введении в систему дополнительных системных шоков; этот подход предложен в статье Ingram, Kocherlakota & Savin (1994). Другим подходом является введение ошибок измерения. Например, предположим, что некая модель устанавливает детерминистическую зависимость вида $c_t = \alpha y_t + \beta c_{t-1}$, где c_t и y_t — наблюдаемые переменные потребления и производства, а α и β — параметры. В реализации данных, очевидно, такая зависимость не может быть удовлетворена при всех t ни для каких фиксированных значений α и β . Введение ошибок измерения состоит в том, чтобы не менять модель (и ее выводы), а предположить что на самом деле настоящие данные о потреблении c_t недоступны, вместо этого эконометрист наблюдает переменную c_t^* , которая является измерением c_t с ошибкой $c_t^* = c_t + e_t$.

5 Эконометрические проблемы: слабая идентификация

В данном разделе мы поговорим о вопросах идентификации и слабой идентификации, которые очень часто возникают при оценивании DSGE моделей, и создают проблему для классического оценивания (так называемого подхода фреквентиста). Вопросы слабой идентификации являются на данный момент активной областью эконометрических исследований, и решения для многих из них пока неизвестны.

5.1 Идентификация

Параметр β считается (глобально) идентифицируемым в точке $\beta = \beta_0$, если для любого другого значения параметра $\beta^* \neq \beta_0$ распределение данных \mathcal{Y}_T , сгенерированных из модели с параметром β_0 , отличается от распределения данных, сгенерированных из модели, где параметр равен β^* . Эконометрическая литература, посвященная вопросам идентификации, очень быстро растет в последние годы. Однако, большая ее часть посвящена кросс-секционным данным, в то время как идентификация в моделях временных рядов имеет свои особенности.

Примером сложности идентификации в моделях временных рядов может служить, например, идентификация простейшего процесса скользящего среднего.

Пример 2. Рассмотрим процесс $x_t = e_t + \theta e_{t-1}$, где e_t являются независимыми одинаково распределенными гауссовскими случайными величинами $e_t \sim N(0, \sigma^2)$. Предположим, что эконометрист наблюдает данные $\mathcal{X}_T = \{x_1, \dots, x_T\}$. Несложно показать, что вектор \mathcal{X}_T является гауссовским и параметры $\theta = \delta, \sigma^2 = \sigma_1^2$ и $\theta = 1/\delta, \sigma^2 = \delta^2 \sigma_1^2$ дают одинаковое распределение вектора \mathcal{X}_T . Таким образом, без дополнительных ограничений параметры процесса скользящего среднего являются неидентифицируемыми.

Другим примером исчезновения идентификации может служить пример сокращения корней авторегрессионной части и скользящего среднего.

Пример 3. Предположим, что нас интересует процесс ARMA(1,1): $x_t = \alpha x_{t-1} + e_t - \beta e_{t-1}$, где $e_t \sim N(0, 1)$ — независимые и одинаково распределенные шоки, α и β — неизвестные параметры. Мы наблюдаем вектор $\mathcal{X}_T = \{x_1, \dots, x_T\}$. Легко показать, что если $\alpha = \beta$, то $x_t = e_t$, и параметры неидентифицируемы.

Оба примера выявляют сложности с идентификацией динамики процесса в моделях временных рядов. Если DSGE модель линейна и разрешена в виде (2), то наблюдаемый процесс является процессом типа VARMA с ограничениями на коэффициенты. Статья Komunjer &

Ng (2011) устанавливает критерии идентификации таких моделей. Альтернативный подход представлен в статье Qu & Tkachenko (2012), которые формулируют условия идентификации линейных DSGE моделей в терминах спектральной плотности. Iskrev (2008, 2010) связывает условия идентификации с невырожденностью матрицы информации Фишера и указывает на простой способ вычисления ее для логлинеаризованных DSGE моделей.

Следует отметить, что все эти подходы устанавливают условия для локальной (а не глобальной) идентификации. Под локальной идентификацией понимается следующее свойство: модель локально идентифицируема при значении параметра β_0 , если существует открытая окрестность параметра β_0 такая, что модель, ограниченная на данную окрестность, является идентифицируемой. Так, модель в Примере 2 является локально идентифицируемой, так как после ограничения пространства параметров до $|\theta| < 1$ она становится идентифицируемой. В то же время модель в Примере 3 является ни глобально, ни локально неидентифицируемой. Условия для глобальной идентификации DSGE моделей на данный момент неизвестны.

Как говорилось ранее, большинство DSGE моделей являются нелинейными по входящим макропеременным, что осложняет их решение и оценивание. Большая часть DSGE литературы посвящена решению и оцениванию не изначальных моделей, а их линеаризованных версий. Однако вполне возможно, что в процессе линеаризации идентифицируемая модель может потерять идентификацию. Это мотивирует поиск и развитие методов решения и оценивания нелинейных моделей.

5.2 Слабая идентификация

Одна из важных и нерешенных на данный момент проблем оценивания DSGE моделей является проблема слабой идентификации. Термин «слабая идентификация» хотя и применяется часто в эконометрической литературе, не имеет до сей поры строгого и четкого определения. Обычно предполагается, что модель является идентифицированной в смысле, описанном в предыдущем подразделе, однако это само по себе не гарантирует, что оценка неизвестных параметров будет хороша, или что классические тесты будут иметь правильный уровень значимости. Термин «слабая идентификация» чаще всего описывает ситуацию, когда распределения в конечных выборках классических оценок и тестовых статистик далеки от распределений, предписываемых классической асимптотической теорией, часто вследствие того, что данные не информативны в достаточной степени о неизвестном параметре. Иными словами, термин «слабая идентификация» зачастую используется как замена утверждению «асимптотика не работает».

Хорошо изученным примером слабой идентификации является случай так называемых слабых инструментов, когда рассматривается регрессия, оцениваемая методом инструментальных переменных, но инструменты очень слабо коррелированы с эндогенным регрессором. В таком случае стандартная классическая асимптотика дает очень плохое приближение распределениям в конечных выборках. Так, например, распределение двухшаговой оценки наименьших квадратов в модели со слабыми инструментами достаточно сильно отличается от гауссовского распределения, оно сильно смещено в сторону несостоятельной оценки наименьших квадратов и имеет более тяжелые хвосты. Распределения классических статистик (t - и F -статистик) также существенно отличаются от своих асимптотических пределов. Использование классических методов инструментальной регрессии в этом случае ведет к неверным статистическим выводам. Вместо этого необходимо использовать методы, робастные к слабым инструментам.

Существует несколько обобщений феномена слабых инструментов на нелинейные модели. Один подход выражен в статье Stock & Wright (2000), которая вводит понятие слабо идентифицированного обобщенного метода моментов (ОММ). В упомянутой статье функция, задающая моментное условие в ОММ, почти не меняется при изменении параметра.

Асимптотически это моделируется так, что функция зависит от размера выборки, в предположении, что в пределе производная моментной функции равна нулю. При этом неизвестный параметр предполагается идентифицируемым в конечных выборках, но информация о нем не накапливается асимптотически. Как результат, ОММ оценка этого параметра не является состоятельной, ее распределение нестандартно, и классические тесты имеют распределения, отличающиеся от классических асимптотических пределов.

Другим обобщением понятия слабой идентификации является подход, предложенный в статье Andrews & Cheng (2012). Этот подход предполагает, что модель идентифицируема за исключением одного специального значения некоторого параметра, при котором другое значение параметра теряет свою идентификацию. Термином слабая идентификация Andrews & Cheng (2012) называют ситуацию, когда настоящее значение параметра приближается к точке потери идентификации. Как пример такой ситуации рассмотрим процесс, описанный в Примере 3. Мы видели что если $\alpha = \beta$, то ни α ни β не идентифицируемы. Предположим вместо этого, что $\beta = \alpha + \delta$, и назовем новой параметризацией пару (α, δ) . Тогда если $\delta = 0$, то параметр α не идентифицируем. Ситуация, когда $\delta \neq 0$, но при этом параметр δ очень близок к нулю, описывает ситуацию, где данные обладают очень малой информацией о параметре α . Асимптотикой слабой идентификации считается набор моделей, где параметр $\delta_n = \text{const}/\sqrt{n}$ меняется с величиной выборки n и асимптотически сходится к нулю. В данной асимптотике параметр α не может быть оценен состоятельно, стандартные оценки (например оценка максимального правдоподобия) не являются асимптотически нормальными, и классические статистики имеют нестандартные предельные распределения.

Общее для этих двух подходов то, что асимптотическая теория дает очень плохое приближение для распределений оценок и статистик в конечных выборках из-за недостаточности информации в данных относительно какого-то параметра или группы параметров. Похожий феномен присутствует в DSGE моделях. Впервые это было замечено в статье Canova & Sala (2009), которые обратили внимание, что типичная функция правдоподобия для типичной DSGE модели имеет области постоянства, где она очень слабо меняется при существенном изменении параметров. В такой области постоянства может находиться несколько локальных (а также и глобальный) максимумов. Эта особенность DSGE моделей очень хорошо известна прикладным макроэкономистам. Так например Fernández-Villaverde (2010) характеризует проблему следующим образом: (1) поверхность функции правдоподобия нерегулярна, обладает многочисленными локальными экстремумами и имеет области постоянства, что затрудняет поиск численного значения оценки; (2) даже когда оценка найдена, возникают проблемы по поиску стандартной ошибки, характеризующей точность данной оценки, так как стандартные формулы дают странные значения; (3) получаемые доверительные интервалы имеют уровень доверия, сильно отличающийся от декларируемого.

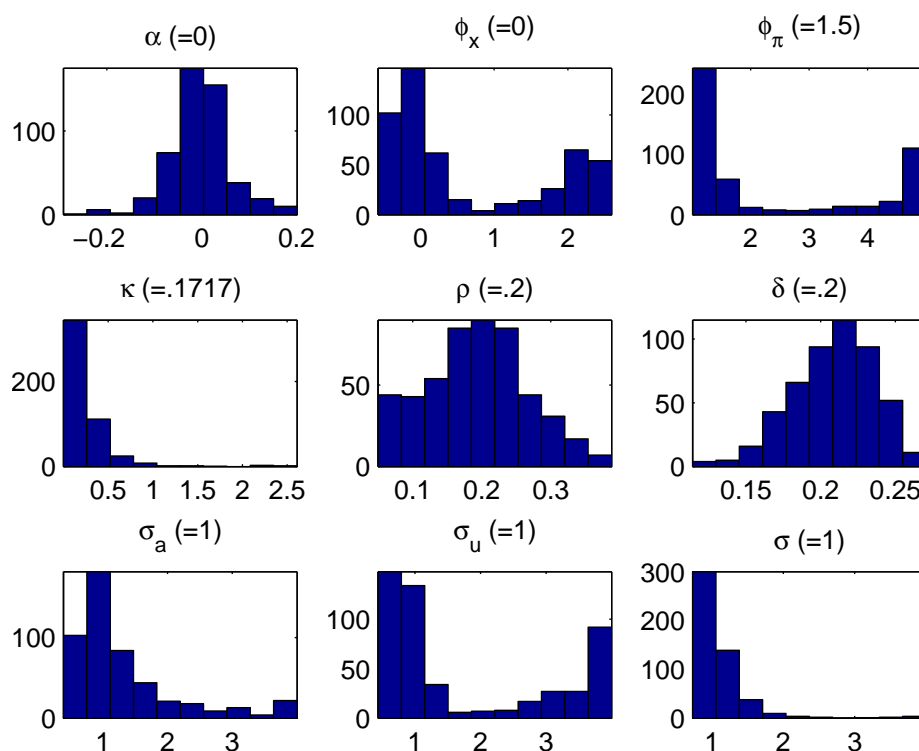
Чтобы продемонстрировать возникающие проблемы, воспроизведем эксперимент Монте-Карло, описанный в статье Andrews & Mikusheva (2011).

Пример 1 (продолжение). Рассмотрим простейшую логлинеаризованную DSGE модель, описанную в Примере 1. Предположим, что настоящие значения экономически значимых параметров равны следующим величинам (широко согласующимся с макроэкономическим консенсусом): коэффициент дисконтирования $b = 0,99$, Калво параметр $\kappa = 0,1717$, параметры правила Тейлора $\varphi_x = 0$, $\varphi_\pi = 1,5$ и $\lambda = 0$. Положим значения параметров статистического описания равным следующим значениям: $\rho = 0,2$ и $\delta = 0,2$ (чтобы внести автокорреляцию шоков, при этом сохранив стационарность) и $\sigma = \sigma_a = \sigma_u = 1$.

Из описанной модели мы генерируем выборки размера 200 периодов, что является типичным размером для макроэкономических данных. На полученных выборках мы будем производить оценивание по методу максимального правдоподобия и тестировать различные гипотезы, которые выполняются в популяции. В данном Монте-Карло эксперименте мы бу-

дем считать значение коэффициента дисконтирования b известным, так как в практике он очень часто оценивается вне модели и калибруется. Оставшиеся 9 неизвестных параметров мы назовем вектором β . Оценивание по методу максимального правдоподобия происходит через численное решение модели, нахождение функции правдоподобия фильтром Кальмана и максимизацией функции правдоподобия. Эксперименты производятся с использованием пакета Dynare.

Рис. 1: Гистограмма распределения оценок по методу максимального правдоподобия. Истинные значения параметров приведены в скобках над каждым графиком. Размер выборки $T = 200$. Число симуляций 500.

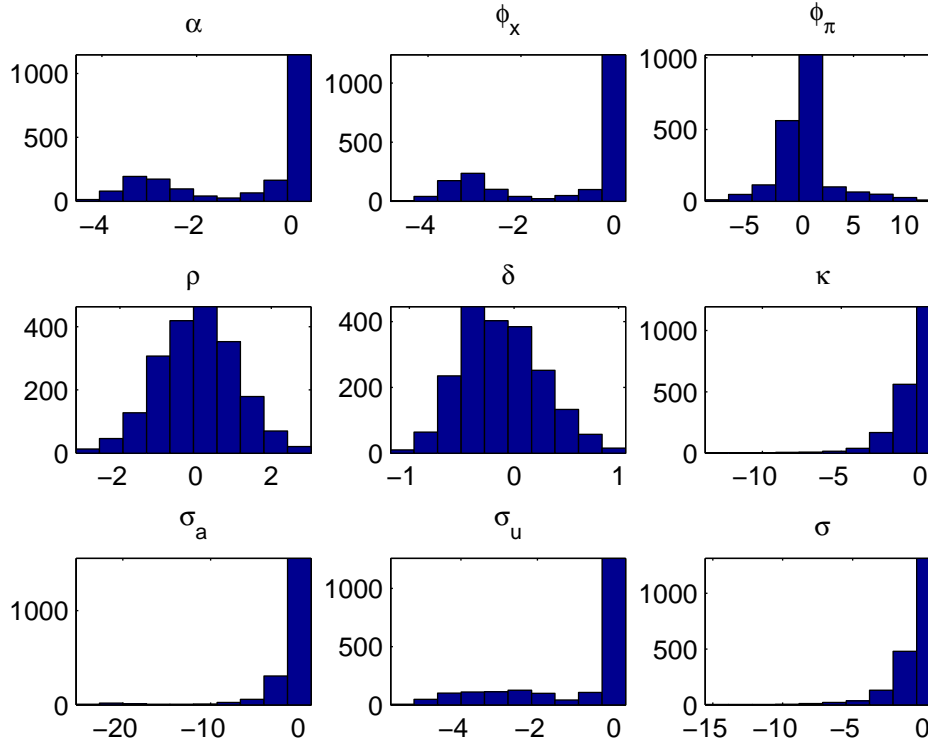


Для начала рассмотрим распределение в конечной выборке (размера 200 периодов) оценки максимального правдоподобия, получающуюся в результате оптимизации функции максимального правдоподобия по 9-мерному параметру β . Гистограммы соответствующих распределений компонент β приведены на Рис. 1. Согласно асимптотической теории, в больших выборках распределение оценки должно быть близко к нормальному. Однако, как мы видим на Рис. 1, для большинства параметров распределение достаточно далеко от нормального. В этом смысле асимптотическая теория дает неверное представление о статистической неопределенности ассоциированной с оценкой наибольшего правдоподобия для данной модели.

Любая оценка должна иметь характеристику статистической погрешности. Часто эту роль выполняет матрица ковариаций оценки. В случае метода максимального правдоподобия асимптотическая теория предписывает, что данная матрица асимптотически равна обратной матрице к матрице информации Фишера. В случае корректной спецификации модели, а в случае нашего эксперимента Монте-Карло модель корректно специфицирована, информация Фишера может быть оценена на практике двумя способами: через квадратную вариацию первой производной (см. формулу (6)) и через вторую производную функции правдоподобия (см. формулу (5)). Однако на практике эти два способа дают существенно разные ответы, и оба способа дают очень странные значения для матрицы Фишера.

На Рис. 2 изображены гистограммы распределения в конечной выборке t -статистик для

Рис. 2: Гистограмма распределения t -статистик для тестирования отдельных параметров. Размер выборки $T = 200$. Число симуляций 500. Информация Фишера подсчитана как вариация первой производной функции правдоподобия.



тестирования верного значения каждого параметра в отдельности. Здесь мы использовали формулу (6) для вычисления матрицы Фишера. Если асимптотическая теория дает хорошее приближение, то данные распределения должны быть близки к стандартному нормальному. Однако, как можно видеть, распределения еще более далеки от нормальности, чем распределения на Рис. 1.

Таблица 1: Фактический уровень значимости классических тестов метода максимального правдоподобия для простой гипотезы $H_0 : \beta = \beta_0$.

	LR	Wald ($\hat{J}(\theta_0)$)	Wald ($\hat{J}(\hat{\theta})$)	Wald ($\hat{I}(\theta_0)$)	Wald ($\hat{I}(\hat{\theta})$)
Размер 5%-го теста	3,20%	65,45%	63,05%	68,05%	68,15%
Размер 10%-го теста	7,05%	67,20%	64,30%	70,80%	71,00%

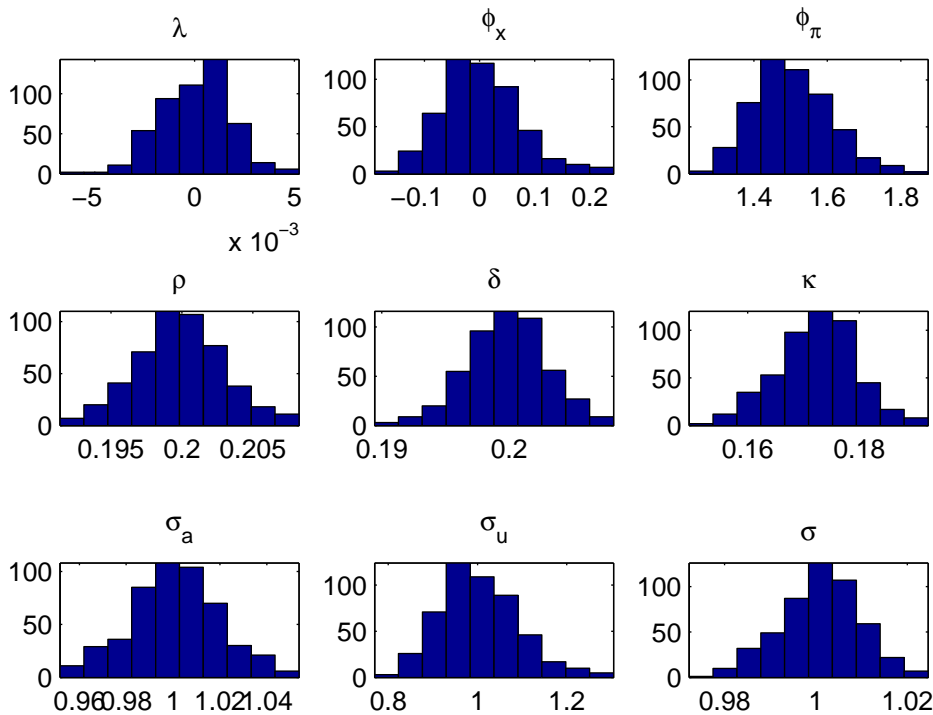
Замечания: Декларируемый уровень значимости 5% и 10%. Число симуляций 2000.

Наконец, обратимся к классическим тестам. Рассмотрим здесь только простейший случай, а именно проблему тестирования простой гипотезы $H_0 : \beta = \beta_0$, где все 9 неизвестных параметров тестируются одновременно. Таблица 1 дает фактический уровень значимости классических теста отношения правдоподобия (LR) и теста Вальда. Последний рассчитан в нескольких видах в зависимости от того, как считается информация Фишера. В данном случае тест, основанный на статистике отношения правдоподобия, является слегка консервативным, а все варианты теста Вальда слишком часто отвергают верную гипотезу и ведут к совершенно неверным статистическим выводам.

Таким образом видим, что результаты классической асимптотической теории, применен-

ные к нашей простой DSGE модели, дают статистически неверные выводы и не могут быть использованы для оценивания и статистических выводов. Может ли наблюдаемый эффект быть назван слабой идентификацией? Скорее да, чем нет. Проблема заключается в том, что на данный момент не существует тестов на слабую идентификацию в общем случае. Но есть несколько соображений, указывающих на то, что ситуация полностью аналогична известным моделям со слабой идентификацией.

Рис. 3: Гистограмма распределения оценок по методу максимального правдоподобия при размере выборки $T = 100000$.



Во-первых, данная модель локально идентифицируема, что следует из применения к ней критериев идентификации, предложенных Komunjer & Ng (2011) и Iskrev (2008, 2010). Во-вторых, наблюдаемая неспособность асимптотики аппроксимировать настоящие распределения относится только к малым выборкам, так например, Рис. 3 изображает гистограммы распределения оценки максимального правдоподобия в случае, когда размер выборки $T = 10^5$. Как видим, в этом случае распределения близки к нормальным. Наконец, Andrews & Mikusheva (2011) показали, что все известные на данный момент модели со слабой идентификацией обладают одним общим интересным свойством, а именно для них две оценки матрицы информации Фишера (данные в формулах (5) и (6)) сильно отличаются друг от друга. Более конкретно, White (1982) показал, что если функция правдоподобия правильно специфицирована, то $\mathbb{E}\hat{I} = \mathbb{E}\hat{J}$. Однако для моделей со слабой идентификацией оценка \hat{I} может быть нормализована так, что она сходится асимптотически к детерминистической невырожденной матрице, в то время как \hat{J} при той же нормализации сходится к случайной матрице, которая не обязана быть положительно определена. Последнее особенно парадоксально, учитывая что \hat{J} является оценкой положительно определенной матрицы информации Фишера. В нашем Монте-Карло эксперименте мы обнаружили, что при размере выборки $T = 200$ в 95% выборок оценка \hat{J} не является положительно определенной матрицей.

5.3 Методы, робастные к слабой идентификации

Метод считается робастным к слабой идентификации, когда он дает тесты с фактическим уровнем значимости, не превышающем (с небольшой погрешностью) декларируемый уровень значимости даже в ситуациях, когда классические методы не работают. Большинство методов так или иначе опираются на асимптотические приближения, однако качество подобных аппроксимаций может сильно варьироваться. Например, в случае слабых инструментов робастные методы опираются на предположение, что центральная предельная теорема, примененная к моментному условию, дает хорошее асимптотическое приближение для распределения нормализованной частичной суммы. То есть считается, что выборка достаточно велика, чтобы простейшие асимптотические законы работали, в то время как выборка не достаточна для того, чтобы были верны асимптотические приближения более сложных нелинейных функций от частичных сумм.

Так, согласно Hall & Heyde (1980), качество асимптотической аппроксимации в классической теории максимального правдоподобия основывается на следующих предположениях: (1) функция правдоподобия хорошо аппроксимируется квадратичной функцией, (2) центральная предельная теорема дает хорошее приближение для некоторых частичных сумм (часто речь идет о первой производной функции правдоподобия), и (3) нормированная вторая производная функции правдоподобия близка к неслучайной константе (информации Фишера). Из предыдущих дискуссий видно, что условия (1) и (3) вряд ли выполняются для DSGE моделей в небольших выборках. Существует несколько статей, пытающихся найти методы для построения статистических выводов, не использующих данные предположения, которые тем не менее опираются на предположения типа (2).

На данный момент существует несколько методов тестирования простых гипотез (гипотез вида $H_0 : \beta = \beta_0$), робастных к слабой идентификации. Andrews & Mikusheva (2011) показали, что один из классических тестов, а именно тест, основанный на мультипликаторе Лагранжа (LM), является робастным к слабой идентификации, если в качестве оценки матрицы Фишера используется \hat{I} . Qu (2011) предложил использовать приближенный тест Лагранжа, основанный на приближении спектральной плотности процесса. Guerгон-Quintana, Inoue & Kilian (2012) предложили использовать тест отношения правдоподобия для приведенной модели, которая пренебрегает многими структурными связями, накладываемыми DSGE моделью ради робастности к слабой идентификации. Наконец, Dufour, Khalaf & Kichian (2009) предложили робастный тест, который отходит от принципа максимального правдоподобия и использует метод наименьшего расстояния. Таблица 2 приводит фактический уровень значимости для всех упомянутых тестов в модели из Примера 1 при размере выборки 200 периодов. Как можно видеть, в то время как метод максимального правдоподобия дает неверные результаты, все четыре робастных теста имеют уровень значимости, очень близкий к декларируемому.

Таблица 2: Фактический уровень значимости робастных тестов для тестирования простой гипотезы в Примере 1.

Декларируемый уровень значимости	AM (2011)	Qu (2011)	GQIK (2012)	DKK (2009)
5%	4,8%	5,1%	6,6%	6,9%
10%	8,9%	10,4%	13,1%	12,4%

Замечания: Число симуляций 1000. Размер выборки 200 периодов.

К сожалению, прогресс в создании робастных тестов относится только к тестированию простых гипотез, т.е. гипотез, которые одновременно тестируют значения всех неизвестных параметров (девятимерный параметр в нашем примере). На практике мы чаще заинтересованы в тестировании гипотез о каждом параметре в отдельности, что является составной ги-

потезой (composite hypothesis). Возможность тестирования каждого одномерного параметра отдельно необходимо также для построения отдельных доверительных интервалов для каждого параметра. К сожалению, задача робастного тестирования сложных гипотез на данный момент неразрешена. Предположим, что в Примере 1 мы хотим протестировать гипотезу $H_0 : \kappa = 0,1717$ (которая в нашем примере верна). В таком случае все остальные восемь параметров задачи являются неизвестными и не оговариваются гипотезой. Они называются шумовыми параметрами. Если предположить что все эти шумовые параметры являются сильно идентифицируемыми, то существуют модификации всех упомянутых выше робастных тестов для тестирования сложной гипотезы. К сожалению, предположение о сильной идентификации шумовых параметров, скорее всего, неверно для типичной DSGE модели, и к тому же не существует на данный момент тестов, которые могли бы протестировать данное предположение. Таким образом, вопрос о том, как производить статистические выводы в случае слабой идентификации модели, остается открытой эконометрической проблемой.

6 Методы оценивания: Байесовский подход

Частично из-за провала классических методов (слабая идентификация), частично из-за появления новых численных алгоритмов, самым популярным методом оценивания DSGE моделей на данный момент является Байесовский метод. Хорошим изложением методологии является статья An & Schorfheide (2007). Наиболее цитируемый пример оценивания DSGE модели Байесовским способом приведен в статье Smets & Wouters (2007).

Входными компонентами Байесовского анализа являются корректно специфицированная функция правдоподобия $\mathcal{L}(\mathcal{Y}_T|\beta)$ и априорное распределение (или вера) $p(\beta)$. На основе их получается постериорное распределение параметра

$$p(\beta|\mathcal{Y}_T) = \frac{\mathcal{L}(\mathcal{Y}_T|\beta)p(\beta)}{\int \mathcal{L}(\mathcal{Y}_T|\tilde{\beta})p(\tilde{\beta})d\tilde{\beta}}. \quad (7)$$

В качестве Байесовской оценки обычно берется либо среднее либо медиана постериорного распределения, а в качестве доверительного интервала- множество, чья постериорная вероятность равна уровню доверия.

Долгое время основным препятствием к использованию Байесовского метода была техническая сложность в вычислении многомерных интегралов, стоящих в знаменателе формулы (7). Постериорное распределение было известно только в ограниченном числе случаев при очень специализированных априорных распределениях. В последнее десятилетие был сделан огромный прорыв, позволяющий применять Байесовский подход в широком наборе случаев. Этот прорыв заключается в разработке алгоритма MCMC (Markov Chain Monte-Carlo). Алгоритм состоит в построении Марковской цепи β_1, β_2, \dots таким образом чтобы для некоего большого числа N (в пределе) случайная величина β_N имела распределение совпадающее с постериорным распределением. Таким образом, вместо того чтобы считать нормализующую константу в уравнении (7), MCMC алгоритм ставит своей целью получить случайную выборку (любого необходимого размера) из постериорного распределения. Имея случайную выборку из постериорного распределения легко получить Байесовские оценки и сделать Байесовские статистические выводы. Например, среднее такой выборки приблизительно равно среднему постериорного распределения, при этом последнее может быть подсчитано с любой точностью.

Существует несколько версий MCMC алгоритма, наиболее часто употребляющиеся для оценки DSGE моделей это алгоритм Метрополис-Хастинга и сэмплер Гиббса. Первый реализован в стандартном пакете Dynare. Для ознакомления с MCMC алгоритмами я советую прочитать статьи Chib (2001) и Chib & Greenberg (1995, 1996), которые дают очень доступное изложение материала.

6.1 Откуда берутся априорные распределения?

Основной сложностью на данный момент в применении Байесовского метода для оценивания DSGE моделей является ответ на вопрос, откуда берутся априорные распределения. С философской точки зрения априорное распределение или вера является отображением текущего уровня знаний или ожиданий ученого относительно вероятностей, что параметр принимает то или иное значение, до того как ученый увидел данные. Постериорное распределение является коррекцией априорной веры с учетом выборки, имеющейся у исследователя. К сожалению, большинство прикладных макроэкономистов признают, что у них чаще всего нет содержательных кандидатов для априорной веры.

На практике многие статьи, оценивающие DSGE модели Байесовским методом, используют априорные распределения, предложенные в статье Smets & Wouters (2007), или аналогичные им. Принцип построения таких априорных распределений следующий: во-первых, априорно все одномерные параметры модели считаются независимо распределены (т.е. многомерная априорная плотность равна произведению априорных плотностей всех параметров). Во-вторых, одномерные распределения выбраны из стандартных семейств распределений, по возможности сопряженных (conjugate) с моделью, т.е. например коэффициенты линейной регрессии считаются априорно нормально распределенными, дисперсии берутся из обратного Гамма-распределения, авторегрессионные коэффициенты считаются априори распределенными по Бета-закону. В-третьих, если есть какие-то оценки параметров, по которым достигнут консенсус среди макроэкономистов, то априорное распределение выбирается центрированным на данной оценке. И наконец, по возможности выбирается распределение, позволяющее широкий разброс значений (или большую дисперсию).

Классическим результатом Байесовской теории считается теорема Бернштейна–фон Мизеса, которая показывает, что асимптотически выбор априорного распределения неважен, так как с точки зрения фреквентистского подхода при широких условиях Байесовское постериорное распределение сходится (когда размер выборки стремится к бесконечности) к гауссовскому распределению вокруг оценки максимального правдоподобия. Эта теорема часто служит утверждением, что Байесовский метод может применяться и сторонником фреквентистского (классического) подхода как удобный численный механизм, коль скоро результаты не сильно отличаются от результатов метода максимального правдоподобия.

К сожалению, эта логика не вполне применима к моделям со слабой идентификацией, а значит и к DSGE моделям в частности. Теорема Бернштейна–фон Мизеса существенным образом использует асимптотическую теорию максимального правдоподобия, которая дает очень неверное представление фактических распределений в конечных выборках для типичных DSGE моделей. В Байесовском методе информация агрегируется из двух источников: выборки и априорного распределения. Теорема Бернштейна–фон Мизеса говорит, что когда выборка растет, информация, приходящая из априорного распределения, асимптотически доминируется информацией из выборки и становится несущественной. В моделях со слабой идентификацией выборка является малоинформативной, и значение априорного распределения велико.

Тот факт, что в DSGE моделях априорное распределение влияет на результат оценивания, заостряет проблему выбора априорного распределения и делает ее центральной. Распространенным заблуждением является утверждение, что если при построении априорного распределения взять одномерные распределения параметров, имеющие большую дисперсию, то оно будет неинформативным, и все результаты будут зависеть главным образом от данных, а не от априорного распределения. Müller (2012) показал, что это не вполне верно. Так например, само предположение, что априори различные одномерные параметры независимы друг от друга, является очень информативным.

Интересным вопросом с этой точки зрения является вопрос о том, сколько информации приходит в результат из данных, а сколько из априорного распределения, и насколько резуль-

тат чувствителен к выбору априорного распределения. Попыткой ответить на этот вопрос является статья Müller (2012). Müller (2012) применил эти методы к модели из Примера 1 и типичным данным и показал, что вся информация о Калво параметре приходит из априорного распределения, а данные главным образом не информативны о нем. Как результат, Байесовская оценка параметра Калво очень чувствительна к выбору априорной веры.

7 Заключение

Большие структурные модели становятся все более и более привлекательными для макроэкономистов. Однако, макроэкономические данные не очень информативны и фактически не позволяют сколь либо хорошо оценивать сколь нибудь сложные модели. Это порождает очень сложные и интересные эконометрические проблемы, которые на данный момент не решены и открывают большую область для деятельности эконометристов.

На мой взгляд, интересными проблемами являются задача поиска статистических методов, робастных к слабой идентификации, вопрос измерения информативности данных и априорной веры в Байесовском подходе, изучение фреквентистских свойств Байесовских процедур а также задача построения неинформативных априорных распределений. Прорыв в решении любой из этих задач способен изменить практику оценивания DSGE моделей и оказать существенное влияние на развитие прикладной макроэконометрики.

Список литературы

- Цыплаков, А. (2011). Введение в моделирование в пространстве состояний. *Квантиль* 9, 1–24.
- An, S. & F. Schorfheide (2007). Bayesian analysis of DSGE models. *Econometric Reviews* 26, 113–172.
- Anderson, G. & G. Moore (1985). A linear algebraic procedure for solving linear perfect foresight models. *Economics Letters* 17, 247–252.
- Andrews, D.W.K. & X. Cheng (2012). Estimation and inference with weak, semi-strong and strong identification. *Econometrica* 80, 2153–2211.
- Andrews, I. & A. Mikusheva (2011). Maximum likelihood inference in weakly identified DSGE models. Unpublished manuscript, MIT.
- Blanchard, O.J. & C.M. Kahn (1980). The solution of linear difference models under rational expectations. *Econometrica* 48, 305–311.
- Blanchard, O.J. & D. Quah (1989). Dynamic effects of aggregate demand and supply disturbances. *American Economic Review* 79, 655–673.
- Boivin, J. & M. Giannoni (2006). Has monetary policy become more effective? *The Review of Economics and Statistics* 88, 445–462.
- Canova, F. (2007). *Methods for Applied Macroeconomic Research*. Princeton University Press.
- Canova, F. & L. Sala (2009). Back to square one: identification issues in DSGE models. *Journal of Monetary Economics* 56, 431–449.
- Chib, S. (2001). Markov chain Monte Carlo methods: computation and inference. Глава 5 в *Handbook of Econometrics* том 5 под редакцией J.J. Heckman & E. Leamer, 3564–3634. Amsterdam: North-Holland.
- Chib, S. & E. Greenberg (1995). Understanding the Metropolis–Hastings algorithm. *American Statistician* 49, 327–335.
- Chib, S. & E. Greenberg (1996). Markov chain Monte Carlo simulation methods in econometrics. *Econometric Theory* 12, 409–431.
- Chari, V., P. Kehoe & E. McGrattan (2005). A critique of structural VARs using business cycle theory. Research Department Staff Report 364, Federal Reserve Bank of Minneapolis.
- Christiano, L.J. (2007). A short course on estimation, solution and policy analysis using equilibrium monetary models. Lecture Notes, Northwestern University.

- Christiano, L., Eichenbaum, M. & C. Evans (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113, 1–45.
- Christiano, L., M. Eichenbaum & R. Vigfusson (2006). Assessing structural VARs. *NBER Macroeconomics Annual* 21, 1–72.
- Cogley, T. & J. Nason (1995). Output dynamics in real-business-cycle models. *American Economic Review* 85, 492–511.
- DeJong, D. N. & C. Dave (2007). *Structural Macroeconometrics*, Princeton University Press.
- Del Negro, M. & F. Schorfheide (2004). Priors from general equilibrium models for VARs. *International Economic Review* 45, 643–673.
- Del Negro, M., F. Schorfheide, F. Smets, & R. Wouters (2007). On the fit of new Keynesian models. *Journal of Business & Economic Statistics* 25, 123–143.
- Dufour, J.M., L. Khalaf & M. Kichian (2009). Structural multi-equation macroeconomic models: identification-robust estimation and fit. Working Paper, Bank of Canada.
- Fernández-Villaverde, J. (2010). The econometrics of DSGE models. *SERIES: Journal of the Spanish Economic Association* 1, 3–49.
- Fernandez-Villaverde, J. & J. Rubio-Ramirez (2005). Estimating dynamic equilibrium economies: linear versus nonlinear likelihood. *Journal of Applied Econometrics* 20, 891–910.
- Fernandez-Villaverde, J., J. Rubio-Ramirez, T. Sargent & M. Watson (2007). The ABC and (D's) to understand VARs. *American Economic Review* 97, 1021–1026.
- Gali, J., J.D. Lopez-Salido & J. Valles (2003). Technology shocks and monetary policy: assessing the Fed's performance. *Journal of Monetary Economics* 50, 723–743.
- Guerron-Quintana, P., A. Inoue & L. Kilian (2012). Frequentist inference in weakly identified DSGE models. *Quantitative Economics*, в печати.
- Hall, A., A. Inoue, J. Nason & B. Rossi (2012). Information criteria for impulse response function matching estimation of DSGE models. *Journal of Econometrics* 170, 499–518.
- Hall, P. & C.C. Heyde (1980). *Martingale Limit Theory and its Application*. Academic Press.
- Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and Kalman Filter*. Cambridge University Press.
- Ingram, B.F., N. Kocherlakota & N.E. Savin (1994). Explaining business cycles: a multi-shock approach. *Journal of Monetary Economics* 34, 415–428.
- Ireland, P. (2004). A method for taking models to data. *Journal of Economic Dynamics and Control* 28, 1205–26.
- Iskrev, N. (2008). Evaluating the information matrix in linearized DSGE models. *Economics Letters* 99, 607–610.
- Iskrev, N. (2010). Evaluating the strength of identification in DSGE models: an a priori approach. Working paper, Bank of Portugal.
- King, R.G., C.I. Plosser, J.H. Stock & M.W. Watson (1991). Stochastic trends and economic fluctuations. *American Economic Review* 81, 819–840.
- Klein, P. (2000). Using the generalized Schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control* 24, 1405–1423.
- Komunjer, I. & S. Ng (2011). Dynamic identification of dynamic stochastic general equilibrium models. *Econometrica* 79, 1995–2032.
- Kydland, F. & E. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica* 50, 1345–1370.
- Lindé, J. (2005). Estimating new-Keynesian Phillips curves: a full information maximum likelihood approach. *Journal of Monetary Economics* 52, 1135–1149.
- Mikusheva, A. (2012). One-dimensional inferences in autoregressive models in potential presence of unit root. *Econometrica* 80, 173–212.
- Müller, U. (2012). Measuring prior sensitivity and prior informativeness in large Bayesian models. *Journal of Monetary Economics* 59, 581–597.
- Qu, Z. (2011). Inference and specification testing in DSGE models with possible weak identification. Unpublished manuscript, Boston University.

- Qu, Z. & D. Tkachenko (2012). Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models. *Quantitative Economics* 3, 95–132.
- Rabanal, P. & J. Rubio-Ramirez (2005). Comparing new Keynesian models of the business cycle: a Bayesian approach. *Journal of Monetary Economics* 52, 1151–1166.
- Rotemberg, J. & M. Woodford (1997). An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual* 12, 297.
- Schorfheide, F. (2000). Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics* 15, 645–670.
- Schorfheide, F. (2010). Estimation and evaluation of DSGE models: progress and challenges. Invited Lecture at 2010 Econometric Society World Congress in Shanghai.
- Sims, C.A. (1980). Macroeconomics and reality. *Econometrica* 48, 1–48.
- Sims, C.A. (2001). Solving linear rational expectations models. *Computational Economics* 20, 1–20.
- Smets, F. & R. Wouters (2007). Shocks and frictions in US business cycles: a Bayesian DSGE approach. *American Economic Review* 97, 586–606.
- Stock, J.H. & J.H. Wright (2000). GMM with weak identification. *Econometrica* 68, 1055–1096.
- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* 50, 1–25.

Estimation of dynamic stochastic general equilibrium models

Anna Mikusheva

Massachusetts Institute of Technology, Cambridge, USA

This essay is a survey of the main econometric approaches to estimation of the dynamic stochastic general equilibrium (DSGE) models widely used by central banks and federal reserves. The paper discusses in detail the main econometric problems arising in inferences about the parameters of the log-linearized DSGE models. We examine three main estimation approaches: the minimum distance method, the maximum likelihood method and the Bayesian approach. We focus on the problems of weak identification that are due to scarcity of macro data. The issues of economic modelling and methods of solving dynamic models are beyond the scope of the current essay.

DSGE-моделирование в пакете Dynare: практическое введение*

Коллум Джонс[†]

Нью-Йоркский Университет, Нью-Йорк, США

Мариано Кулиш[‡]

Университет Нового Южного Уэльса, Сидней, Австралия

Настоящий документ представляет собой практическое введение в пакет Dynare. В нём рассказывается об установке Dynare, записи динамической стохастической модели общего равновесия в обозначениях Dynare, а также обсуждаются выдача результатов оценивания модели в рабочем пространстве MATLAB и частые ошибки, возникающие при использовании Dynare. Работа Dynare демонстрируется на конкретном примере с кратким обсуждением оценивания и прогнозирования. Документ завершает эмпирическое приложение.

1 Установка Dynare

Dynare представляет собой набор функций MATLAB. Пакет считывает уравнения модели DSGE, решает или оценивает модель с заданным набором параметров и выдаёт результаты.

Для скачивания последней стабильной версии Dynare необходимо пройти по ссылке www.dynare.org/download: установочный файл должен называться `dynare-*-win.exe`, где символ `*` обозначает номер версии. Установка начнётся при запуске `.exe`-файла. Папкой установки по умолчанию является, как правило, `C:\dynare*`, где символ `*` обозначает номер версии. Все файлы MATLAB, необходимые для запуска Dynare, находятся в `C:\dynare*\matlab`.

После установки пакета Dynare необходимо указать MATLAB путь на его файлы. Для этого в MATLAB кликните `File` → `Set Path`, затем `Add Folder...` и выберите папку `C:\dynare*\matlab`. Все файлы Dynare должны теперь появиться в пути поиска MATLAB (`MATLAB search path`), как на Рис. 1. Нажмите `Save` и `Close`.

Для тестирования корректности установки Dynare наберите `dynare` в командной строке MATLAB. Если Dynare установлен правильно, вы получите сообщение об ошибке `??? Input argument 'fname' is undefined`. Если Dynare установлен неправильно, вы получите сообщение об ошибке `??? Undefined function or variable 'dynare'`.

2 Модель в Dynare

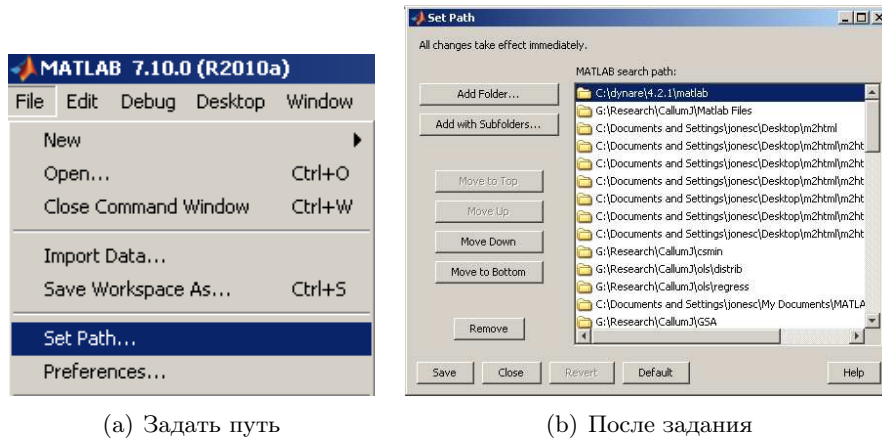
Dynare читает текстовые файлы формата `.mod` и преобразует их в файлы MATLAB, которые затем вызывают функции Dynare для анализа модели. В настоящем разделе мы продемонстрируем это на примере линеаризованной системы уравнений модели и покажем, как записать модель в `.mod`-файл, читаемый Dynare.

*Цитировать как: Джонс, Коллум & Мариано Кулиш (2014). «DSGE-моделирование в пакете Dynare: практическое введение», Квантиль, №12, стр. 23–44. Citation: Jones, Callum & Mariano Kulish (2014). “A practical introduction to DSGE modeling with Dynare,” *Quantile*, No.12, pp. 23–44.

[†]Адрес: Department of Economics, 19 West 4th Street, 6th floor, New York, USA. Электронная почта: jonescallum@gmail.com

[‡]Адрес: West Lobby Level 4, Australian School of Business Building, University of New South Wales, Sydney NSW 2052, Australia. Электронная почта: m.kulish@unsw.edu.au

Рис. 1: Задание пути поиска MATLAB



(a) Задать путь

(b) После задания

Рассматриваемая модель предложена в работе Ireland (2004). Используя обозначения автора, можно записать линейаризованные уравнения модели в следующем виде:

$$\hat{x}_t = E_t \hat{x}_{t+1} - (\hat{r}_t - E_t \hat{\pi}_{t+1}) + (1 - \omega)(1 - \rho_a) \hat{a}_t, \quad (1)$$

$$\hat{\pi}_t = \beta E_t \hat{\pi}_{t+1} + \psi \hat{x}_t - \hat{e}_t, \quad (2)$$

$$\hat{r}_t = \hat{r}_{t-1} + \rho_\pi \hat{\pi}_t + \rho_g \hat{g}_t + \rho_x \hat{x}_t + \varepsilon_{r,t}, \quad (3)$$

$$\hat{x}_t = \hat{y}_t - \omega \hat{a}_t, \quad (4)$$

$$\hat{g}_t = \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t, \quad (5)$$

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{a,t}, \quad (6)$$

$$\hat{e}_t = \rho_e \hat{e}_{t-1} + \varepsilon_{e,t}, \quad (7)$$

$$\hat{z}_t = \varepsilon_{z,t}. \quad (8)$$

Уравнение (1) — это кривая IS, уравнение (2) — кривая Филлипса, уравнение (3) — правило Тэйлора, уравнение (4) определяет разрыв ВВП, уравнение (5) определяет рост ВВП, а уравнения (6), (7) и (8) определяют спрос, издержки и технологические шоки соответственно. Детали вывода уравнения см. в работе Ireland (2004) или в подробных лекциях этого автора.

Запишем теперь модель в форме, которую Dynare сможет интерпретировать. В первой части кода Dynare нужно указать эндогенные переменные модели. Всего имеем восемь уравнений, а значит и восемь переменных: \hat{y}_t , \hat{x}_t , \hat{g}_t , \hat{r}_t , $\hat{\pi}_t$, \hat{a}_t , \hat{e}_t , \hat{z}_t . Синтаксис Dynare: `var [variable1, variable2, ...]`; строка должна оканчиваться точкой с запятой. В случае модели Ireland необходимо написать:

```
var y, x, g, r, pi, a, e, z ;
```

Далее определим экзогенные переменные модели. В модели Ireland четыре шока: $\varepsilon_{r,t}$, $\varepsilon_{a,t}$, $\varepsilon_{e,t}$ и $\varepsilon_{z,t}$. Синтаксис Dynare: `varexo [variable1, variable2, ...]`; строка должна оканчиваться точкой с запятой. В случае модели Ireland необходимо написать:

```
varexo eps_r, eps_a, eps_e, eps_z ;
```

Теперь зададим параметры модели. В модели Ireland восемь структурных параметров и четыре стандартных отклонения шоков, т.е. всего двенадцать параметров. Синтаксис Dynare: `parameters [parameter1, parameter2, ...]`; строка должна оканчиваться точкой с запятой. В случае модели Ireland необходимо написать:

```
parameters beta, omega, psi, rho_pi, rho_g, rho_x, ...
           rho_a, rho_e, sig_r, sig_z, sig_a, sig_e ;
```

Наконец, зададим значения параметров. Ireland оценивает большую часть параметров методом максимального правдоподобия на квартальных данных США с 1947 года. Синтаксис задания параметров `[parameter] = [value]`; совпадает с синтаксисом MATLAB. В случае модели Ireland напишем:

```
beta    = 0.99 ;
psi     = 0.1  ;
omega   = 0.0617 ;
rho_pi  = 0.3597 ;
rho_g   = 0.0000 ;
rho_x   = 0.0347 ;
rho_a   = 0.9470 ;
rho_e   = 0.9625 ;
sig_a   = 0.0405 ;
sig_e   = 0.0012 ;
sig_z   = 0.0109 ;
sig_r   = 0.0031 ;
```

Запишем теперь линеаризованные уравнения модели. Уравнения записываются практически в таком же виде, в каком они появляются в статье. Раздел задания уравнений начинается с `model (linear)` ; и заканчивается с `end ;`. Некоторые особенности синтаксиса:

1. Переменные, которые умножаются (делятся) на параметр, отделяются от него с помощью символа `*` (`/`).
2. Переменные текущего момента времени не имеют нижнего временного индекса в коде Dynare. Лаги переменных x_{t-i} записываются как `x(-i)`, а условные математические ожидания $E_t x_{t+i}$ записываются как `x(+i)`.
3. Уравнения заканчиваются двоеточием ;.

Уравнения модели Ireland в форме Dynare имеют вид:

```
model (linear) ;
  x = x(+1) - ( r - pi(+1) ) + ( 1 - omega ) * ( 1 - rho_a ) * a ;
  pi = beta * pi(+1) + psi * x - e ;
  r = r(-1) + rho_pi * pi + rho_g * g + rho_x * x + eps_r ;
  x = y - omega * a ;
  g = y - y(-1) + z ;
  a = rho_a * a(-1) + eps_a ;
  e = rho_e * e(-1) + eps_e ;
  z = eps_z ;
end ;
```

После спецификации уравнений необходимо сообщить Dynare, откуда начинать симуляции и считать функции импульсного отклика, т.е. начальные значения переменных. Аналогично заданию стартовых значений параметров, начнём секцию со строки `initval ;` и закончим её строкой `end ;`, задав значения переменных. Обычно начальные значения соответствуют состоянию равновесия линеаризованной модели. В модели Ireland все переменные заданы в Dynare в отклонениях от своих равновесных значений, и поэтому равновесию модели соответствуют нулевые начальные значения. В данном случае, вместо указания начального

значения для каждой переменной отдельно, можно написать `steady ;`, после чего Dynare начнёт симуляцию из состояния равновесия.

Далее мы должны определить шоки и задать их дисперсии либо произвольным образом, либо соответствующей оценкой. Начинаем секцию кода со строки `shocks ;` и заканчиваем её строкой `end ;`. Синтаксис для задания шока: `var [shock] = [shock variance] ;`. Для модели Ireland соответствующий блок кода будет таким:

```
shocks ;
    var eps_a = sig_a^2 ;
    var eps_e = sig_e^2 ;
    var eps_z = sig_z^2 ;
    var eps_r = sig_r^2 ;
end ;
```

Наконец, мы решаем уравнения и получаем результаты Dynare с настройками по умолчанию. Используем команду `stoch_simul` без дополнительных опций, справку по которым можно получить в справочнике Dynare¹. Таким образом, заключительная строка `.mod`-файла имеет вид

```
stoch_simul ;
```

Этого достаточно для запуска Dynare, решения модели и получения результатов. Содержимое `.mod`-файла приведено в приложении.

3 Запуск Dynare и вывод результатов

Чтобы запустить созданный `.mod`-файл, сначала сохраните текстовый файл с расширением `.mod` в текущую папку MATLAB, например, как `ireland.mod`. Чтобы запустить Dynare, в командной строке MATLAB наберите

```
>> dynare ireland.mod
```

Вы увидите, что Dynare начал работу. Пакет читает содержимое `.mod`-файла с описанием модели и решает линейную модель рациональных ожиданий. Dynare выводит массу аналитической информации в командное окно MATLAB, сохраняя большую часть в рабочее пространство, и генерирует графики функций импульсных откликов на шоки к каждой переменной. Пакет также сохраняет весь сгенерированный контент в рабочую папку.

Первое, что печатает Dynare, — это информация обработки. Строки `[mex]` означают, что Dynare нашёл `.mex`-файлы — исполняемые файлы MATLAB, которые производят некоторые вычисления вне оболочки MATLAB. Следует проверить, чтобы количество уравнений, которые Dynare находит в `.mod`-файле, совпадало с количеством эндогенных переменных модели. В модели Ireland таких переменных восемь.

```
Configuring Dynare ...
[mex] Generalized QZ.
[mex] Sylvester equation solution.
[mex] Kronecker products.
[mex] Sparse kronecker products.
[mex] Bytecode evaluation.
[mex] k-order perturbation solver.
[mex] k-order solution simulation.
```

¹www.dynare.org/documentation-and-support/manual/manual.pdf

```

Starting Dynare (version 4.2.0).
Starting preprocessing of the model file ...
Found 8 equation(s).
Evaluating expressions...done
Computing static model derivatives:
- order 1
Computing dynamic model derivatives:
- order 1
- order 2
Processing outputs ...done
Preprocessing completed.
Starting MATLAB/Octave computing.

```

Затем Dynare напечатает равновесные значения всех переменных и таблицу с количеством переменных.

STEADY-STATE RESULTS:

```

y    0
x    0
g    0
r    0
pi   0
a    0
e    0
z    0

```

MODEL SUMMARY

```

Number of variables:      8
Number of stochastic shocks: 4
Number of state variables: 4
Number of jumpers:       2
Number of static variables: 2

```

Далее, Dynare сгенерирует ковариационную матрицу шоков. Значения на диагонали соответствуют специфицированным `eps`-значениям в `.mod`-файле.

MATRIX OF COVARIANCE OF EXOGENOUS SHOCKS

Variables	eps_r	eps_a	eps_e	eps_z
eps_r	0.000010	0.000000	0.000000	0.000000
eps_a	0.000000	0.001640	0.000000	0.000000
eps_e	0.000000	0.000000	0.000001	0.000000
eps_z	0.000000	0.000000	0.000000	0.000119

Dynare вычисляет решение модели как

$$\hat{y}_t = A\hat{y}_{t-1} + Bu_t,$$

где \hat{y}_t — вектор эндогенных переменных в отклонениях от равновесных значений, \hat{y}_{t-1} — отклонение y_{t-1} от равновесного значения, а u_t — вектор экзогенных шоков. Первые четыре строки выдачи соответствуют матрице A^T , а последние четыре строки соответствуют матрице B^T .

POLICY AND TRANSITION FUNCTIONS

	y	x	g	r	pi	a	e	z
y(-1)	0	0	-1.000	0	0	0	0	0
r(-1)	-2.923	-2.923	-2.923	0.623	-0.764	0	0	0
a(-1)	0.182	0.124	0.182	0.014	0.027	0.947	0	0
e(-1)	5.939	5.939	5.939	-0.385	-1.643	0	0.962	0
eps_r	-2.923	-2.923	-2.923	0.624	-0.764	0	0	0
eps_a	0.193	0.131	0.193	0.015	0.029	1.000	0	0
eps_e	6.171	6.171	6.171	-0.400	-1.708	0	1.000	0
eps_z	0	0	1.000	0	0	0	0	1.000

Следующая секция выдачи Dynare содержит моменты модели, вычисленные с использованием решения модели и специфицированных стандартных ошибок экзогенных шоков.

THEORETICAL MOMENTS

VARIABLE	MEAN	STD. DEV.	VARIANCE
y	0.0000	0.0436	0.0019
x	0.0000	0.0422	0.0018
g	0.0000	0.0186	0.0003
r	0.0000	0.0070	0.0000
pi	0.0000	0.0056	0.0000
a	0.0000	0.1261	0.0159
e	0.0000	0.0044	0.0000
z	0.0000	0.0109	0.0001

Далее Dynare выдаёт разложение дисперсии эндогенных переменных, т.е. долю дисперсии эндогенных переменных, которая может быть приписана экзогенным шокам. Таким образом, для каждой переменной сумма долей по всем четырём шокам равна 100%. Например, для модели Ireland с откалиброванными/оцененными значениями параметров, 47% роста ВВП можно отнести на счёт технологических шоков $\varepsilon_{z,t}$.

VARIANCE DECOMPOSITION (in percent)

	eps_r	eps_a	eps_e	eps_z
y	7.08	8.67	84.25	0.00
x	7.55	2.53	89.91	0.00
g	29.12	20.00	16.69	34.19
r	12.42	45.99	41.60	0.00
pi	29.43	7.03	63.54	0.00
a	0.00	100.00	0.00	0.00
e	0.00	0.00	100.00	0.00
z	0.00	0.00	0.00	100.00

Затем Dynare выдаёт матрицу теоретических корреляций между эндогенными переменными. Эти значения вычисляются на основе ковариационной матрицы эндогенных переменных, которую Dynare не возвращает в командное окно.

MATRIX OF CORRELATIONS

Var	y	x	g	r	pi	a	e	z
y	1.0000	0.9841	0.1407	-0.5297	-0.4875	0.2656	0.9121	0.0000
x	0.9841	1.0000	0.1393	-0.6683	-0.5140	0.0900	0.9422	0.0000
g	0.1407	0.1393	1.0000	-0.1070	0.0608	0.0329	0.0800	0.5847
r	-0.5297	-0.6683	-0.1070	1.0000	0.2581	0.6569	-0.6303	0.0000
pi	-0.4875	-0.5140	0.0608	0.2581	1.0000	0.0561	-0.7689	0.0000
a	0.2656	0.0900	0.0329	0.6569	0.0561	1.0000	0.0000	0.0000
e	0.9121	0.9422	0.0800	-0.6303	-0.7689	0.0000	1.0000	0.0000
z	0.0000	0.0000	0.5847	0.0000	0.0000	0.0000	0.0000	1.0000

Далее Dynare показывает автокорреляции до пятого порядка, т.е. корреляции эндогенных переменных с собственными лагами p -го порядка, где $p = \{1, \dots, 5\}$.

COEFFICIENTS OF AUTOCORRELATION

Order	1	2	3	4	5
y	0.9398	0.8902	0.8476	0.8099	0.7756
x	0.9454	0.8993	0.8589	0.8226	0.7891
g	-0.0578	-0.0386	-0.0265	-0.0189	-0.0140
r	0.9438	0.8942	0.8493	0.8081	0.7697
pi	0.7871	0.6483	0.5558	0.4925	0.4476
a	0.9470	0.8968	0.8493	0.8043	0.7616
e	0.9625	0.9264	0.8917	0.8582	0.8260
z	0.0000	0.0000	0.0000	0.0000	0.0000

Total computing time : 0h00m13s

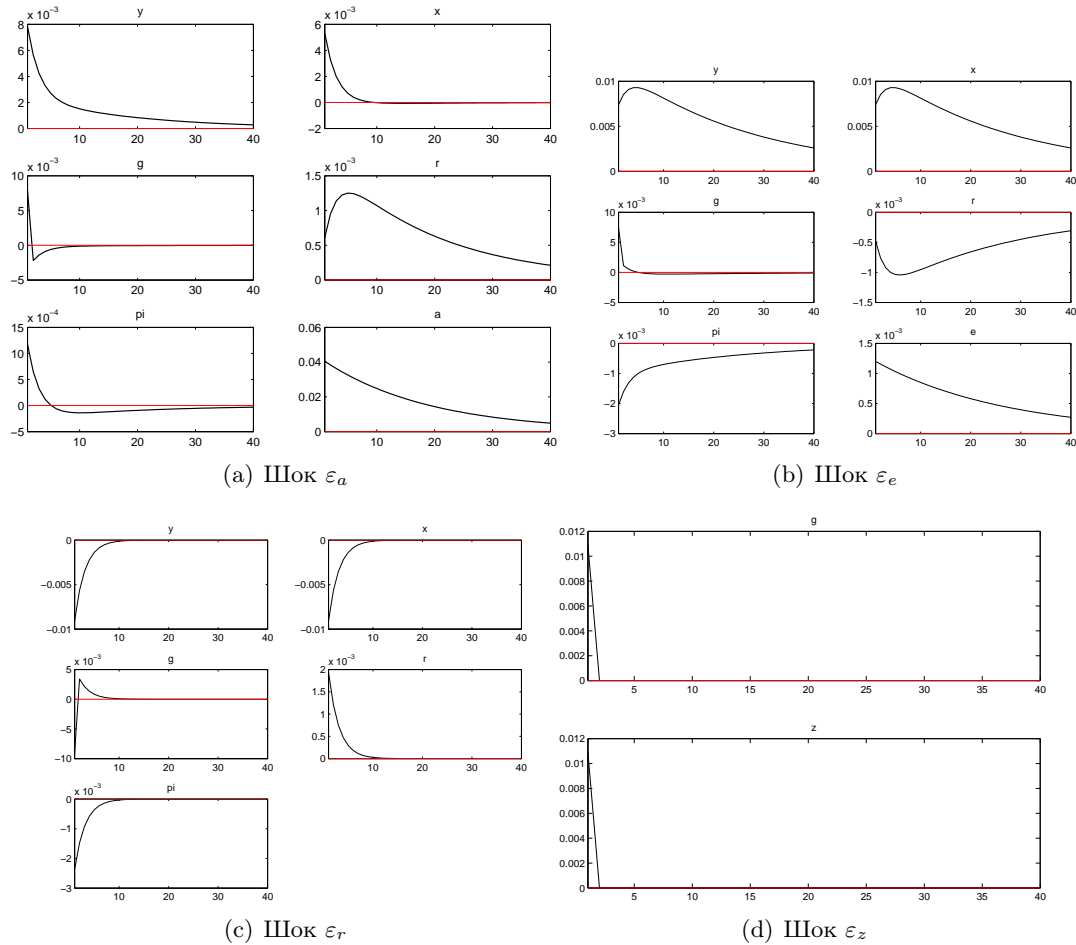
Наконец, Dynare генерирует графики функций импульсного отклика эндогенных переменных на экзогенные шоки. Dynare отфильтровывает функции отклика, незначимо отличающиеся от нуля на области вычисления. В каждом окне рисунка MATLAB будет нарисовано не более 9 графиков импульсных функций; если в модели более 9 эндогенных переменных, Dynare нарисует импульсные отклики сверх 10-го в последующих окнах рисунка MATLAB. Функции импульсного отклика для модели Ireland приведены на Рис. 2. Видно, что количество импульсных откликов соответствует количеству нулевых элементов матрицы B из policy-функции и переходной функции (см. выше).

4 Рабочее пространство MATLAB

В таблице 1 представлена информация о некоторых из переменных, которые Dynare сохраняет в рабочее пространство MATLAB. Вообще говоря, Dynare сохраняет в рабочее пространство все значения параметров и импульсные отклики всех эндогенных переменных на шоки.²

²MATLAB может хранить данные в объектах класса `struct`, состоящих из полей, в которые можно поместить любые объекты MATLAB, например матрицы, строки, структуры и т.д. Чтобы обратиться к полю `field` структуры `struct` из командной строки, наберите `struct.field`. Большая часть выходной информации Dynare хранится в структурах.

Рис. 2: Графики функций импульсных откликов



5 Выдача результатов решения

Необходимо понимать, как Dynare упорядочивает переменные модели в вычисленных матрицах решения (назовём такой порядок DR-порядком, от слов *decision rule*). Dynare отличает эндогенные переменные по их типу, т.е. по отнесению в одну из следующих категорий:

- **Предопределённые переменные:** появляющиеся в модели только в виде текущего значения или лагов. В модели Ireland четыре предопределённых переменных: y , r , a и e . Их количество записывается в `oo_.dr.npred`.
- **Переменные будущих периодов:** появляющиеся в модели только в виде текущего или будущих значений. В модели Ireland две переменные будущих периодов: π и x . Их количество записывается в `oo_.dr.nfwr`.
- **Смешанные переменные:** появляющиеся как в виде текущего значения и лагов, так и в виде будущих значений. Таких переменных в модели Ireland нет. Их количество записывается в `oo_.dr.nboth`.
- **Статические переменные:** появляющиеся в модели только в виде текущего значения. В модели Ireland таких переменных две: g и z . Их количество записывается в `oo_.dr.nstatic`.

Таблица 1: Рабочее пространство MATLAB

Переменная	Описание
<code>oo_.exo_simul</code>	Симулированные серии экзогенных переменных
<code>oo_.endo_simul</code>	Симулированные серии эндогенных переменных
<code>oo_.dr</code>	Информация о переменных и policy-функции
<code>oo_.dr.order_var</code>	Отображение из порядка декларирования в DR-порядок
<code>oo_.dr.inv_order_var</code>	Отображение из DR-порядка в порядок декларирования
<code>oo_.dr.ghx</code>	Матрица A переходной функции (см. раздел 5)
<code>oo_.dr.ghu</code>	Матрица B переходной функции (см. раздел 5)
<code>oo_.steady_state</code>	Равновесные значения эндогенных переменных
<code>oo_.gamma_y{7,1}</code>	Разложение дисперсии из командного окна
<code>oo_.mean</code>	Средние эндогенных переменных
<code>oo_.var</code>	Ковариационная матрица
<code>oo_.autocorr</code>	Автокорреляционные матрицы эндогенных переменных (до 5 лагов)
<code>oo_.irfs</code>	Импульсные отклики эндогенных переменных на экзогенные шоки
<code>oo_.irfs.x_y</code>	Импульсный отклик x на шок в y .
<code>oo_.forecast</code>	Прогнозные значения эндогенных переменных
<code>M_</code>	Структура, содержащая информацию о модели
<code>M_.endo_names</code>	Имена эндогенных переменных в порядке декларирования
<code>M_.exo_names</code>	Имена экзогенных переменных в порядке декларирования
<code>M_.exo_nbr</code>	Количество экзогенных переменных
<code>M_.endo_nbr</code>	Количество эндогенных переменных
<code>options_</code>	Структура, содержащая все опции, распознаваемые Dynare

DR-порядок таков: статические переменные, предопределённые переменные, смешанные переменные, переменные будущих периодов. Отображение из порядка декларирования в секции `var` в DR-порядок определяется в `oo_.dr.order_var`. Обратное отображение определяется в `oo_.dr.inv_order_var`. Обозначим вектор эндогенных переменных за y_t , в котором переменные располагаются в порядке декларирования. Тогда для того, чтобы DR-упорядочить этот вектор, необходимо написать $y_{dr} = y(oo_.dr.order_var, :)$. Чтобы вернуться к порядку декларирования, надо написать $y_{dec} = y_{dr}(oo_.dr.inv_order_var, :)$. Тогда можно убедиться, что y_{dr} совпадает с y_{dec} .

Чтобы понять, для чего нужно упорядочивание, рассмотрим следующее решение:

$$\hat{y}_t = A\hat{y}_{t-1} + Bu_t.$$

Матрица A появляется в `oo_.dr.ghx`, а матрица B — в `oo_.dr.ghu`. Dynare располагает переменные по строкам в порядке декларирования. Столбцы A соответствуют переменным состояния. Отображение этих переменных состояния из порядка декларирования записывается в вектор

$$oo_.dr.order_var(oo.dr.nstatic+1 : oo_.dr.nstatic+oo_.dr.npred)$$

Эти переменные совпадают с лагированными переменными из policy-функции и переходной функции.

Итак, пусть мы хотим вручную вычислить функции импульсных откликов каждой эндогенной переменной на шок в размере одного стандартного отклонения для 20 кварталов,

используя матрицы решения, а затем получить результаты в порядке декларирования. Тогда мы должны написать следующий код. График для ненулевых функций представлен на Рис. 3. Он должен совпадать с графиком на Рис. 2.

```

horizon = 20 ;
shocks  = zeros(M_.exo_nbr,horizon) ;
irf     = zeros(M_.endo_nbr,horizon) ;

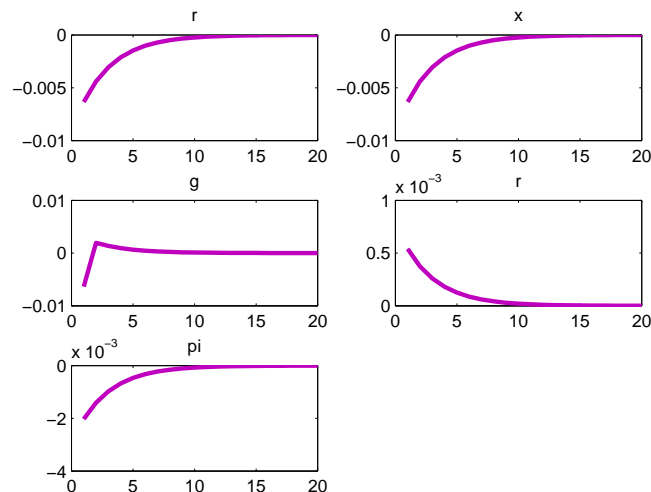
shocks(1,1) = sig_r; % Policy shock in period 1

irf(:,1) = oo_.dr.ghu(oo_.dr.inv_order_var,:)*shocks(:,1) ;

for t=2:horizon
    irf(:,t) = oo_.dr.ghx(oo_.dr.inv_order_var,:) * ...
        irf(oo_.dr.order_var(oo_.dr.nstatic+1:oo_.dr.nstatic+oo_.dr.npred),t-1)+...
end

```

Рис. 3: Графики функций импульсных откликов на шоки политики, вычисленные вручную



6 Некоторые распространённые сообщения об ошибке Dynare

К сожалению, Dynare известен своими неоднозначными сообщениями об ошибках. Вот некоторые распространённые ошибки в написании `.mod`-файла и соответствующие сообщения об ошибке Dynare. Заметим, что Dynare не улавливает логические ошибки, такие как ошибки при записи системы линейных уравнений в `.mod`-файл.

6.1 Ошибки парсинга

Как правило, если Dynare сталкивается с ошибкой парсинга, он выдаёт следующее сообщение в командное окно MATLAB:

```

ERROR: ireland.mod:10.1

??? Error using ==> dynare at 126
DYNARE: preprocessing failed

```

В строке `ERROR: ireland.mod:10.1` число 10.1 означает, что ошибка парсинга идентифицирована в строке 10, столбце 1 `.mod`-файла.

Следующая ошибка может появиться, если переменная или параметр не задекларирована в `var [.]`, но появляется в уравнениях модели. Например, здесь пропущена декларация `z`:

```
ERROR: ireland.mod:25.26: Unknown symbol: z
```

Такая ошибка может появиться также когда отсутствует строка `end` ; после описания модели. Если строка `end` ; пропущена после секции `shocks`, то Dynare пытается читать строку `stoch_simul`, которая не согласуется с синтаксисом `shocks` ;, что приводит к следующему сообщению об ошибке:

```
ERROR: ireland.mod:40.1-11: syntax error, unexpected NAME,
expecting CORR or END or VAR
```

Ошибка типа `unexpected NAME` может также возникнуть, если Dynare сталкивается с неизвестной ему опцией, например, опцией `nograp` вместо `nograph`. Такая ошибка может быть также следствием пропуска математического оператора (например, `*`, `-`, `+` или `/`) в декларации уравнений модели или если строки не разделяются символом `;`.

Если в модели больше уравнений, чем задекларированных эндогенных переменных, Dynare выбросит соответствующее сообщение об ошибке:

```
ERROR: There are 7 equations but 8 endogenous variables!
```

6.2 Ошибки декларирования

Если параметру не присвоено значение в `.mod`-файле, появится сообщение об ошибке, которое информирует об успехе парсинга и неудаче в решении модели из-за невозможности построить матрицы, входящие в модель. Это может также произойти из-за опечатки в названии параметра в блоке установки параметров. Например, здесь не присвоено значение переменной `beta`:

```
Warning: Some of the parameters have no value (beta) when using stoch_simul.
If these parameters are not initialized in a steadystate file,
Dynare may not be able to solve the model...
```

```
??? Error using ==> print_info at 36
MJDGGES returns the following error code: 6
```

Если параметрам присваиваются значения, которые не дают единственного равновесия, Dynare сообщает, что не выполнены условия Бланшара–Кана³. Например, присвоим параметрам значения `rho_pi = 0`, `rho_g = 0` и `rho_x = 0`. Тогда политика не соответствует правилу Тэйлора, что означает существование более чем одного равновесного пути. Dynare сообщает:

```
??? Error using ==> print_info at 42
Blanchard Kahn conditions are not satisfied: indeterminacy
```

Условия Бланшара–Кана могут быть не выполнены, если спецификация параметров приводит к отсутствию равновесия. Например, задекларируем авторегрессионный параметр при временном технологическом шоке `rho_a = 1.1`. Тогда любое решение модели будет взрывным, и Dynare сообщит:

```
??? Error using ==> print_info at 39
Blanchard Kahn conditions are not satisfied: no stable equilibrium
```

³Более подробно про равновесия можно прочитать в лекциях Симса.

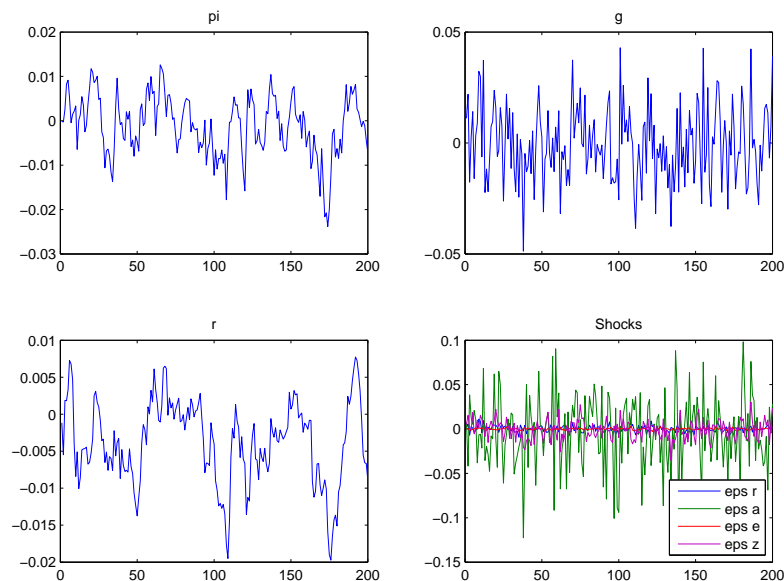
7 Использование Dynare для симуляции данных

Dynare может использовать модель для симуляции рядов эндогенных переменных. Чтобы сгенерировать 200 наблюдений, заменим `stoch_simul` на `stoch_simul(periods=200)`. Dynare сгенерирует 200 шоков экзогенных переменных и использует решение модели для симуляции эндогенных переменных. На Рис. 4 представлены симулированные шоки и симулированные переменные π_t , g_t и r_t . Сопоставьте размеры шоков и заданные значения параметров.

Симулированные эндогенные переменные хранятся в рабочем пространстве под своим именем. Например, в модели Ireland симулированный ряд для инфляции будет храниться под именем `pi`. Симулированные эндогенные переменные также хранятся под именем `endo_simul` в порядке декларирования из `.mod`-файла (т.е. в порядке, заданном `M_ .endo_names`). Симулированные шоки хранятся в `exo_simul` в порядке декларирования (т.е. в порядке, заданном `M_ .exo_names`).

На основе симулированных данных можно вычислить моменты, корреляции и автокорреляции. По умолчанию Dynare игнорирует первые 100 наблюдений, т.е. количество периодов должно быть больше 100. Важно, что новые версии Dynare устанавливают фиксированные начальные значения генератора случайных чисел, поэтому симуляции повторяются, что можно исправить командой `set_dynare_seed(x)`, где `x` — целое число, заданное выше команды `stoch_simul` в `.mod`-файле.

Рис. 4: Симуляции



8 Анализ

Как правило исследователю интересно симулировать множество рядов из одной модели или посмотреть, что происходит с моделью при изменении параметров. Здесь естественно прибегнуть к использованию циклов. В Dynare можно использовать операторы циклов MATLAB вместе с функциями Dynare непосредственно в `.mod`-файле после задания модели.

Пусть, например, необходимо сгенерировать 200 рядов с 200 наблюдениями роста ВВП \hat{g}_t и инфляции $\hat{\pi}_t$ из модели Ireland. Таким образом, нужно запустить `stoch_simul` 200 раз и

сохранить каждый симулированный ряд в столбец матрицы. Это можно сделать следующим образом:

```
for i=1:200
    set_dynare_seed(i);
    stoch_simul( periods=200, noprint, nograph );
    g_sim(:,i) = g;
    pi_sim(:,i) = pi;
end
```

Переменная `i` используется для задания различных начальных условий генератора случайных чисел, а также для индексации матриц `g_sim` и `pi_sim`. На первом шаге переменная цикла принимает значение `i=1`, программа решает модель, симулирует эндогенные переменные на 200 периодов с помощью `stoch_simul` и сохраняет `g` в первом столбце матрицы `g_sim` (аналогично для `pi`). После этого переменная `i` увеличивается на единицу и повторяет те же действия до тех пор, пока не достигнет значения `i=200`. Из соображений скорости исполнения используются опции `noprint` и `nograph` функции `stoch_simul`, подавляющие прорисовку графиков функций импульсных откликов и печать информации о модели в командное окно.

Чтобы пробежать по диапазону возможных значений параметра (при которых решение модели единственно), используется та же техника. Пусть требуется отследить, как функция импульсного отклика инфляции на шок монетарной политики изменяется с издержками на корректировку цен. Иными словами, необходимо проанализировать поведение `pi_eps_a` при изменении значений `psi`, где $\psi = (\theta - 1)/\varphi$, а параметр φ отвечает за издержки на корректировку цен. Определим вектор возможных значений параметра `psi` и используем технику, описанную выше:

```
psi_params = 0:0.05:1;

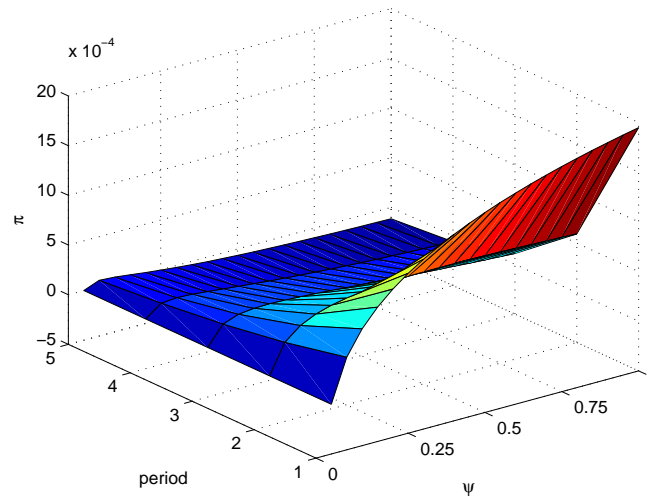
for i=1:length(psi_params)
    psi = psi_params(i);
    stoch_simul(noprint, nograph);
    pi_eps_a_mat(:,i) = pi_eps_a;
end
```

Возможные значения параметра `psi` хранятся в переменной `psi_params`: это множество чисел от 0 до 1 с шагом 0,05, т.е. вектор размерности 1×21 , где `length(psi_params) = 21`. Переменная цикла используется для индексации возможных значений параметра `psi`, а также для обращения к соответствующему столбцу в матрице, хранящей импульсные отклики. Параметр `psi` выбирается из вектора, вызывается функция `stoch_simul`, а затем сохраняется импульсный отклик `pi_eps_a`.

Для выработки интуитивного представления об импульсных откликах, полезно представить эффект изменения параметров с помощью трёхмерного графика. Используем следующую команду MATLAB:

```
>> surf(pi_eps_a_mat(1:5,:)); xlabel('\psi'); ylabel('period'); zlabel('\pi');
```

Исполнение этой команды нарисует каждое значение матрицы, хранящей импульсные отклики по 5 периодам для каждого значения параметра в `psi_params`. Соответствующий график показан на Рис. 5. В терминах модели график показывает, что фирмы имеют большие стимулы к повышению цен в ответ на шок спроса, когда издержки на корректировку малы (т.е. когда `psi` велико).

Рис. 5: Зависимость функции импульсного отклика от ψ 

9 Оценивание модели в Dynare

Dynare можно использовать для оценивания параметров и шоков по данным. В настоящем разделе мы покажем только как оценивать модель на практике. Заметим, что для того, чтобы модель можно было оценить, необходимо, чтобы размер выборки не превышал количество шоков в модели. Подробную информацию по оцениванию содержит файл *resources/Dynare-UserGuide-WebBeta.pdf*.

9.1 Метод максимального правдоподобия

Чтобы оценить модель методом максимального правдоподобия, сначала нужно сообщить, какие переменные модели наблюдаемы. Это делается командой `varobs VAR ;`. Модель Ireland разумно оценивать на основе роста ВВП, инфляции и процентной ставки:

```
varobs g pi r;
```

Далее, специфицируем параметры, которые хотим оценить, в блоке `estimated_params`. Всем неочищаемым переменным необходимо присвоить значение в секции установки параметров `.mod`-файла. В блоке `estimated_params` каждому оцениваемому параметру нужно задать начальное значение и, возможно, границы: соответствующий синтаксис `parameter, initial value [, lower bound, upper bound] ;`, где в квадратные скобки взяты необязательные части кода. Спецификация стандартных ошибок шоков аналогична: необходимо только добавить команду `stderr` перед экзогенной переменной.

Полностью блок `estimated_params` выглядит так:

```
estimated_params ;
  rho_a, 0.9, 0, 1 ;
  rho_e, 0.9, 0, 1 ;
  omega, 0.1, 0, 1 ;
  rho_pi, 0.4 ;
  rho_g, 0.3 ;
  rho_x, 0.05 ;
  stderr eps_a, 0.1, 0, 1 ;
  stderr eps_e, 0.1, 0, 1 ;
```



```

    stderr eps_z, 0.1, 0, 1 ;
    stderr eps_r, 0.1, 0, 1 ;
end ;

```

Наконец, нужно вызвать функцию оценивания, указав Dynare путь к файлу с данными:

```
estimation(datafile=data) ;
```

9.2 Байесовское оценивание

В байесовском подходе для каждого оцениваемого параметра специфицируется априорное распределение. Как и в случае метода максимального правдоподобия, укажем сначала наблюдаемые переменные:

```
varobs g pi r;
```

В блоке `estimated_params` для каждого оцениваемого параметра декларируем распределение и специфицируем параметры распределения. Синтаксис: `parameter, prior_shape, prior_μ, prior_σ [, p_3] [, p_4]` ;. В таблице 2 приведён список распределений и их параметров для декларирования в Dynare:

Таблица 2: Байесовские априорные распределения

Спецификация <code>prior_shape</code>	Распределение
<code>normal_pdf</code>	$N(\mu, \sigma)$
<code>gamma_pdf</code>	$G(\mu, \sigma, p_3)$
<code>beta_pdf</code>	$B(\mu, \sigma, p_3, p_4)$
<code>inv_gamma_pdf</code>	$IG(\mu, \sigma)$
<code>uniform_pdf</code>	$U(p_3, p_4)$

Эти распределения имеют очень разные формы и могут сильно зависеть от параметров. В качестве иллюстрации приведём Рис. 6 с пятью распределениями и различными параметрами. Чтобы посмотреть на зависимость формы распределения от параметров, откройте файл `Mathematica_resources/distributions_mathematica.nb` (при наличии у вас лицензии Mathematica). При оценивании важно тщательно подобрать форму априорного распределения параметров.

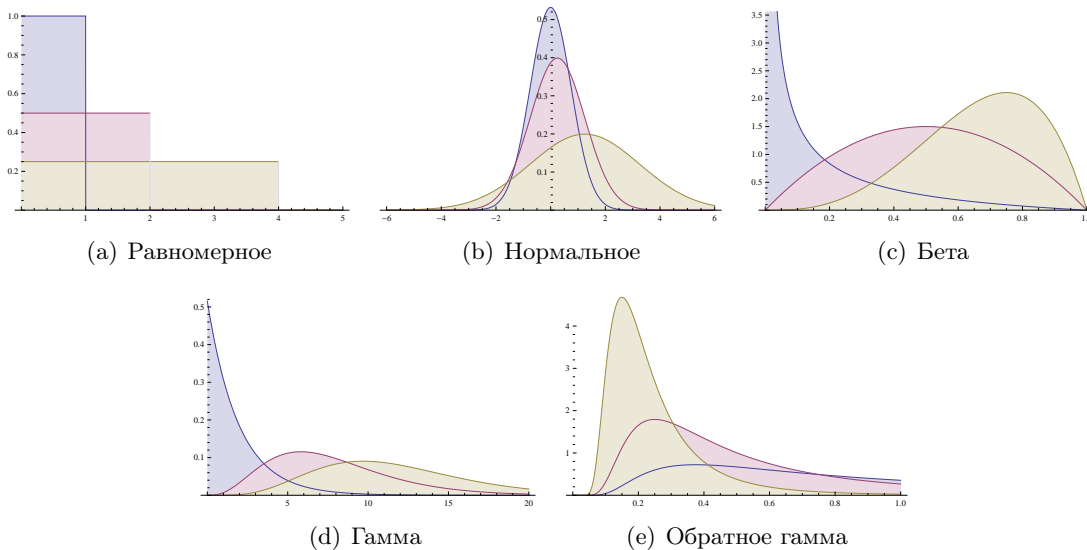
Проиллюстрируем декларирование распределений на примере модели Ireland. Заметим, что дополнительные запятые нужны команде `uniform_pdf` для спецификации третьего и четвёртого моментов.

```

estimated_params ;
    rho_a, beta_pdf, 0.7, 0.1 ;
    rho_e, beta_pdf, 0.7, 0.1 ;
    omega, normal_pdf, 0.1, 0.025 ;
    rho_pi, normal_pdf, 0.4, 0.1 ;
    rho_g, normal_pdf, 0.4, 0.1 ;
    rho_x, normal_pdf, 0.1, 0.05 ;
    stderr eps_a, uniform_pdf, , , 0, 0.1 ;
    stderr eps_e, uniform_pdf, , , 0, 0.1 ;
    stderr eps_z, uniform_pdf, , , 0, 0.1 ;
    stderr eps_r, uniform_pdf, , , 0, 0.1 ;
end ;

```

Рис. 6: Распределения



Наконец, вызовем функцию оценивания так же, как в случае метода максимального правдоподобия: `estimation(datafile=data);`. Для байесовского оценивания доступно множество опций; обзор опций см., например, в папке *resources/Dynare_Manual.pdf*.

9.3 Выдача результатов в рабочем пространстве

Результаты оценивания сохраняются в структуре `oo_` в рабочем пространстве. Описание некоторых полей этой структуры приведено в Таблице 3.

Таблица 3: Рабочее пространство MATLAB: результаты оценивания

Переменная	Описание
<code>oo_.posterior_mode</code>	Апостериорные стандартные ошибки параметров и шоков
<code>oo_.posterior_std</code>	Апостериорные стандартные ошибки оценок
<code>oo_.posterior_density</code>	Апостериорная плотность (значения x, y) для каждого параметра
<code>oo_.prior_density</code>	Априорная плотность (значения x, y) для каждого параметра
<code>oo_.SmoothedVariables</code>	Оценки эндогенных переменных всех периодов
<code>oo_.SmoothedShocks</code>	Оценки шоков всех периодов

9.4 Декомпозиция шока

Оценивание даёт набор структурных шоков, которые обычно представляют особый интерес. Команда `shock_decomposition` позволяет лучше понять, что представляют собой эти шоки. Она раскладывает динамику каждой эндогенной переменной модели на шоки. Иными словами, динамика каждой эндогенной переменной представляется в виде комбинации шоков модели, а команда `shock_decomposition` вычисляет вклад каждого шока в динамику каждой переменной. Команда вызывается строкой

```
shock_decomposition ;
```

10 Прогнозирование в Dynare

На основании решения модели Dynare может строить прогнозы. После команды `stoch_simul` наберите команду

```
forecast ;
```

Некоторые опции: `periods = x` вычисляет прогнозы на `x` периодов, `nograph` подавляет рисовку графиков. Эти опции совпадают с опциями команды `stoch_simul`, поэтому если они не специфицированы, то используются опции команды `stoch_simul`.

Результаты прогноза хранятся в `oo_.forecast.Mean`. Если переменные не симулированы, команда `forecast` просто генерирует прогноз, равный равновесным значениям переменных. Если же переменные симулированы, как в разделе 8, за начальный период прогнозирования берётся первый период симуляций.

11 Другие полезные команды

11.1 Сохранение данных в .m-файл

Если необходимо сохранить вектор или матрицу в скрипт MATLAB (.m-файл), можно использовать команду Dynare `datatomfile()`. Она может быть полезна, если, скажем, для оценивания модели мы хотим использовать симулированные данные. Например, исполняя следующую строку кода, Dynare сохранит три вектора симулированных данных `g`, `pi`, и `r` в файл `data.m` в рабочей папке:

```
datatomfile('data',{'g';'pi';'r'}) ;
```

11.2 Язык макросов Dynare: включение дополнительного кода в .mod-файл

Dynare умеет интерпретировать команды, написанные на языке макросов Dynare, преобразовывая .mod-файл с макросами в .mod-файл без макросов, который затем читается стандартным образом. Это полезно для автоматизации повторяющегося кода в .mod-файле или автоматической подстановке текста в .mod-файл.

В любом месте кода Dynare можно использовать команду `@#include x.mod`. Эта команда добавляет содержимое файла `x.mod` в файл Dynare. Файл `x.mod` должен соответствовать обычным требованиям к языку Dynare. Одно из применений данной команды — изменение структуры зарубежного сектора в модели маленькой открытой экономики. Мы написали функцию (файл) `MATLAB`, которая записывает уравнения из .mod-файла, описывающие векторную авторегрессию (VAR), оценённую для разного количества лагов и на разных выборках. Этот .mod-файл добавляется к первичному .mod-файлу с помощью команды `include`; остальное содержимое первичного .mod-файла при этом не затрагивается.

Чтобы сконструировать .mod-файл с авторегрессией, запишем сначала уравнения в виде строки. Затем зададим имя файла и откроем его с уровнем доступа `wt`, что означает возможность записи текста в файл. Далее, напечатаем строку в файл с опцией `%s`, что означает запись текста как строки. В конце закроем файл.

```
VAR = ( 'y = rho11*y(-1) + rho12*x(-1) + e_y ; ', ...
        'x = rho21*y(-1) + rho22*x(-1) + e_x ; ' ) ;

filename = 'WORLD.mod' ;           % Choose filename
fid1     = fopen(filename, 'wt') ; % Open the file
```

```
fprintf(fid1, '%s', VAR) ;           % Print to the new file
fclose(fid1) ;                       % Close the file
```

Теперь необходимо включить содержимое файла `WORLD.mod` в набор уравнений первичного `.mod`-файла, что осуществляется следующими командами:

```
model (linear) ;
#include "WORLD.mod"
```

Переменные и параметры, упоминаемые во включаемом файле, должны быть задекларированы в верхних строках `.mod`-файла стандартным образом, что также можно осуществить с помощью команды `@#include`.

Одно из важных применений команды `@#include` — это записывание оцененных значений параметров в `.mod`-файл и присваивание этих значений параметрам в `.mod`-файле. Таким образом, не приходится вписывать оценки параметров в `.mod`-файл вручную. Пусть, например, написан `.m`-файл, обращающийся к файлу `dynare ireland_estimation.mod`. Тогда можно написать скрипт, который записывает параметры как строки в `.mod`-файл `parameters.mod`, а затем использовать команду `@#include parameters.mod` в блоке задания значений параметров файла `ireland.mod`. Такой скрипт позволяет автоматически анализировать несколько моделей.

Наконец, опция `savemasro` позволяет сохранить `.mod`-файл с макросами: сохранённый таким образом `.mod`-файл `filename.mod` получит имя `filename-macroexp.mod`.

11.3 Генерирование L^AT_EX из Dynare

В Dynare встроены функции для записывания уравнений модели и оцененных параметров в файл L^AT_EX. Для того, чтобы сообщить Dynare, какие переменные L^AT_EX использовать, необходимо задать L^AT_EX-представление для каждой переменной и каждого параметра в блоке декларирования с помощью двойного символа доллара, например:

```
parameters beta $\beta$, omega $\omega$, psi $\psi$, ...
```

После декларирования L^AT_EX-эквивалентов для переменных и параметров, в конце кода модели Dynare, необходимо использовать команду

```
write_latex_dynamic_model ;
```

для создания `.tex`-файла с именем `MODELNAME_latex_dynamic.tex`, где `MODELNAME` — имя `.mod`-файла. Этот `.tex`-файл можно скомпилировать или использовать его в другом файле L^AT_EX.

12 Эмпирическое приложение: графические модели DSGE

Используем теперь изложенную технику на практике. В настоящем приложении мы используем линейризованные уравнения модели Ireland (2004) для получения кривых совокупного спроса и предложения, связывающих инфляцию и рост ВВП. Представление модели в таком виде позволяет изучить механизмы DSGE в знакомой постановке.⁴

⁴ Данное эмпирическое исследование базируется на работе Kulish & Jones (2011), которая доступна на странице ideas.repec.org/p/cpm/dynare/003.html наряду со всеми файлами Dynare и MATLAB, данными, результатами оценивания и документацией.

12.1 Кривые совокупного спроса и предложения

Линеаризованные уравнения, в которых инфляция и рост ВВП выражены в уровнях (а не в отклонениях от равновесных значений), имеют вид:

$$\hat{x}_t = E_t \hat{x}_{t+1} - (\hat{r}_t - E_t \hat{\pi}_{t+1}) + (1 - \omega)(1 - \rho_a) \hat{a}_t, \quad (9)$$

$$\pi_t = \pi + \beta E_t \hat{\pi}_{t+1} + \psi \hat{x}_t - \hat{e}_t, \quad (10)$$

$$\hat{r}_t = \hat{r}_{t-1} + \rho_\pi \hat{\pi}_t + \rho_g \hat{g}_t + \rho_x \hat{x}_t + \varepsilon_{r,t}, \quad (11)$$

$$\hat{x}_t = \hat{y}_t - \omega \hat{a}_t, \quad (12)$$

$$g_t = g + \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t, \quad (13)$$

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{a,t}, \quad (14)$$

$$\hat{e}_t = \rho_e \hat{e}_{t-1} + \varepsilon_{e,t}, \quad (15)$$

$$\hat{z}_t = \varepsilon_{z,t}, \quad (16)$$

где, в отличие от формы записи линеаризованных уравнений выше по тексту, π есть логарифм равновесной инфляции, а g — логарифм равновесного роста ВВП.

Преобразуем уравнения для получения функций совокупного спроса и предложения, связывающих инфляцию и рост ВВП. Чтобы найти функцию совокупного предложения, подставим уравнения (12) и (13) в соотношение (10) и получим

$$\pi_t = \psi g_t + \hat{s}_t + (\pi - \psi g), \quad (17)$$

где $\hat{s}_t = \beta E_t \hat{\pi}_{t+1} + \psi \hat{y}_{t-1} - \psi \hat{z}_t - \omega \psi \hat{a}_t - \hat{e}_t$. В пространстве роста ВВП и инфляции (g_t, π_t) уравнение (17) представляет инфляцию как линейную функцию роста ВВП с наклоном ψ и свободным членом $\hat{s}_t + (\pi - \psi g)$. Заметим, что изменяющийся во времени свободный член \hat{s}_t равен нулю, когда экономика находится на пути сбалансированного роста. Также отметим, что наклон кривой зависит от номинальной жёсткости цен: при $\psi \rightarrow \infty$ цены становятся абсолютно гибкими, что даёт вертикальную кривую совокупного предложения. Обратно, при росте издержек корректировки цен, $\psi \rightarrow 0$, кривая совокупного предложения становится всё более полой.

Чтобы получить кривую совокупного спроса, подставим уравнения (11), (12) и (13) в соотношение (9). Тогда

$$\pi_t = -\frac{1+\rho_g+\rho_x}{\rho_\pi} g_t + \hat{d}_t + \pi + \frac{1+\rho_g+\rho_x}{\rho_\pi} g, \quad (18)$$

где $\hat{d}_t = -\frac{1}{\rho_\pi} \hat{r}_{t-1} + \frac{1}{\rho_\pi} E_t \hat{x}_{t+1} + \frac{1}{\rho_\pi} E_t \hat{\pi}_{t+1} - \frac{1+\rho_x}{\rho_\pi} \hat{y}_{t-1} + \frac{1+\rho_x}{\rho_\pi} \hat{z}_t + \frac{\omega(1+\rho_x)+(1-\omega)(1-\rho_a)}{\rho_\pi} \hat{a}_t - \frac{1}{\rho_\pi} \varepsilon_{r,t}$. Заметим, что так же как и в случае изменяющегося во времени свободного члена в функции совокупного предложения, \hat{d}_t равен нулю, когда экономика находится на пути сбалансированного роста. Наклон кривой (18) зависит от параметров политики. Более сильная реакция на отклонение инфляции от целевого уровня ρ_π делает кривую более полой. Более сильная реакция на рост ВВП ρ_g и разрыв ВВП ρ_x , напротив, делает кривую более крутой.

12.2 Добавление кривых в Dynare

Уравнения модели записываются в `.mod`-файл. Полезно добавить в них переменные, сдвигающие кривые совокупного спроса \hat{d}_t и совокупного предложения \hat{s}_t . Тогда, чтобы нарисовать кривые, можно будет использовать результаты выдачи Dynare по этим переменным, а также параметрам и равновесным значениям. Итак, добавим в созданный ранее `ireland2004.mod`-файл две эндогенных переменных `adt` и `ast`:

$$\begin{aligned}
 \text{ast} &= \beta * \pi(+1) + \psi * y(-1) - \psi * z - \omega * \psi * a - e; \\
 \text{adt} &= 1/\rho_{\pi} * x(+1) + 1/\rho_{\pi} * \pi(+1) - (1+\rho_x)/\rho_{\pi} * y(-1) - \dots \\
 & 1/\rho_{\pi} * r(-1) + (1+\rho_x)/\rho_{\pi} * z + \dots \\
 & ((\omega * (1+\rho_x) + (1-\omega) * (1-\rho_a))/\rho_{\pi}) * a - (1/\rho_{\pi}) * \text{eps}_r;
 \end{aligned}$$

12.3 Оценивание на данных по США

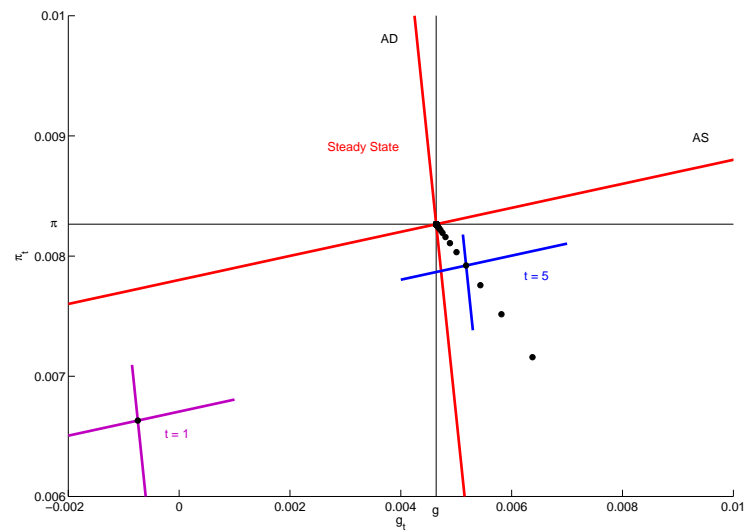
Для построения оцененных кривых совокупного спроса и предложения нужны оценки параметров модели. Мы следуем подходу Ireland (2004) с новыми данными. Калибровка даёт значения $\beta = 0,99$, $\psi = 0,1$ и $\omega = 0,06$. Оценка остальных параметров и соответствующие стандартные ошибки приведены в таблице 4. Из-за сильной оцененной реакции короткой процентной ставки на изменения роста ВВП и разрыва ВВП по сравнению с реакцией инфляции оценка наклона кривой совокупного спроса невелика: $-\frac{1+\rho_g+\rho_x}{\rho_{\pi}} = -4,4$. Кривая совокупного предложения получается довольно пологой из-за малого значения $\psi = 0,1$, что объясняет, почему в модели велик эффект шоков совокупного спроса.

Таблица 4: Оценки метода максимального правдоподобия

Параметр	Оценка	Стандартная ошибка
ρ_a	0,9489	0,0224
ρ_e	0,9470	0,0393
ρ_{π}	0,2885	0,0404
ρ_g	0,2090	0,0451
ρ_x	0,0679	0,0177
σ_a	0,0365	0,0122
σ_e	0,0008	0,0002
σ_z	0,0120	0,0015
σ_r	0,0026	0,0003

Как было замечено, в пространстве роста ВВП и инфляции динамику экономики можно описать пересечением кривых совокупного спроса и предложения в каждый момент времени. Шок сдвигает меняющиеся во времени свободные члены \hat{s}_t и \hat{d}_t . С течением времени положение кривых возвращается к равновесному. Таким образом, например, шок к $\varepsilon_{r,t}$ в размере одного стандартного отклонения приводит к значениям $\hat{s}_t = -0,001$ и $\hat{d}_t = -0,026$. Обе кривые движутся одновременно, что служит иллюстрацией важного факта: в анализе общего равновесия наличие ожиданий приводит к тому, что шоки одновременно сдвигают кривые совокупного спроса и предложения \hat{d}_t и \hat{s}_t . В период шока монетарной политики увеличение \hat{r}_t приводит к снижению совокупного спроса в силу коррекции уровня потребления агентами. Помимо этого агенты ожидают отрицательного разрыва ВВП в следующем периоде и инфляцию ниже равновесной, что, в силу уравнения Эйлера, ещё сильнее снижает совокупный спрос. Что касается предложения, то оно увеличивается из-за падения инфляционных ожиданий: так как фирмы несут издержки на корректировку цен, то при фиксированном объёме выпуска у них есть стимулы снижать цены сегодня.

Проиллюстрируем кривые и их динамику после шока монетарной политики в пространстве роста ВВП и инфляции. На Рис. 7 представлены кривые совокупного спроса и предложения в равновесном состоянии, в период шока монетарной политики ($t = 1$) и четыре периода спустя ($t = 5$). Видно, что после шока совокупный спрос снижается, а совокупное предложение увеличивается по сравнению с равновесными уровнями. Также показано, как сдвигаются кривые при возвращении экономики в равновесное состояние.

Рис. 7: Кривые совокупного спроса и предложения: шок ε_r 

Приложение: файл ireland.mod

```

var y, x, g, r, pi, a, e, z ;
varexo eps_r, eps_a, eps_e, eps_z ;
parameters beta, omega, psi, rho_pi, rho_g, rho_x, rho_a, ...
           rho_e, sig_r, sig_z, sig_a, sig_e ;

beta   = 0.99 ;   rho_pi = 0.3597 ;   rho_a  = 0.9470 ;
sig_e  = 0.0012 ; psi     = 0.1       ; rho_x  = 0.0347 ;
rho_e  = 0.9625 ; sig_z  = 0.0109 ;   omega = 0.0617 ;
rho_g  = 0.0000 ; sig_a  = 0.0405 ;   sig_r  = 0.0031 ;

model (linear) ;
    x = x(+1) - ( r - pi(+1) ) + ( 1 - omega ) * ( 1 - rho_a ) * a ;
    pi = beta * pi(+1) + psi * x - e ;
    r = r(-1) + rho_pi * pi + rho_g * g + rho_x * x + eps_r ;
    x = y - omega * a ;
    g = y - y(-1) + z ;
    a = rho_a * a(-1) + eps_a ;
    e = rho_e * e(-1) + eps_e ;
    z = eps_z ;
end ;
steady ;
shocks ;
    var eps_a = sig_a^2 ;
    var eps_e = sig_e^2 ;
    var eps_z = sig_z^2 ;
    var eps_r = sig_r^2 ;
end ;
stoch_simul ;

```

Список литературы

Ireland, P. (2004). Technology shocks in the New Keynesian model. *Review of Economics and Statistics* 86, 923–936.

Kulish, M. & C. Jones (2011). A Graphical representation of an estimated DSGE model. *Dynare Working Paper Series*, номер 3.

A practical introduction to DSGE modeling with Dynare

Callum Jones

New York University, New York, USA

Mariano Kulish

University of New South Wales, Sydney, Australia

This document is a practical introduction to Dynare. It shows how to install Dynare and write a DSGE model in Dynare notation, and goes through the output from running the model, where output is stored in the Matlab workspace, as well as common Dynare errors. We use Dynare to do some useful analysis. We briefly discuss estimation and forecasting using Dynare. The document closes with a research application.

В помощь изучающим эконометрику

Мини-словарь англоязычных эконометрических терминов, часть 3*

Александр Цыплаков[†]

Новосибирский государственный университет, Новосибирск, Россия

В этой части словаря комментируются англоязычные эконометрические термины *estimator*, *quantile*, *nested*, *marginal* и др. Как и в предыдущих частях, акцент делается на уточнении значения терминов с целью избежать возможной путаницы и некорректностей при интерпретации.

Estimator

Основная проблема с переводом терминов, связанных с получением статистических оценок и изучением их свойств, состоит в том, что русское слово «оценка» может означать по крайней мере три разные вещи. Во-первых, это собственно оценка, т. е. значение, полученное в конкретной выборке — по-английски *estimate*. Во-вторых, это метод (или функция) в соответствии с которым получают оценки в первом смысле — по-английски *estimator*. В-третьих, это процесс получения оценок в первом смысле — по-английски *estimation*. Во избежание путаницы *estimator* лучше переводить как «метод оценивания» или «оцениватель», а *estimation* — оценивание. К сожалению, использование термина «оценка» во втором смысле прочно укоренилось в русскоязычной статистической литературе. Если говорят о свойствах оценки, таких как несмещенность или эффективность, то на самом деле имеют в виду свойства соответствующего метода, поскольку бессмысленно, например, говорить об эффективности конкретного числа. Особенно следует следить за адекватностью перевода слова «оценка» с русского на английский, потому что если значение слова «оценка» в русском тексте обычно можно понять из контекста, то «consistent estimate», употребленное в неудачном месте, может вызвать у читателя недоумение.

Estimation: оценивание, оценка (как процесс).

Estimator: метод оценивания, способ оценивания, оценка, оцениватель.

Estimate (гл.): оценивать.

Estimate (сущ.): оценка.

Estimated: оцененный.

Estimating: оценивающий, оценочный.

Estimating equations: оценивающие уравнения, уравнения, задающие оценку.

Estimand: оцениваемая величина.

*Цитировать как: Цыплаков, Александр (2014) «Мини-словарь англоязычных эконометрических терминов, часть 3», Квантиль, №12, стр. 45–51. Citation: Tsyplakov, Alexander (2014) “A mini-dictionary of English econometric terminology III,” *Quantile*, No.12, pp. 45–51.

[†] Адрес: 630090, г. Новосибирск, Весенний проезд, 6–44. Электронная почта: tsy@academ.org

Quantile

Пусть случайная величина с носителем $[a, b]$ имеет функцию распределения $F(x)$, которая непрерывна и возрастает на $[a, b]$ (здесь a и b могут быть бесконечными). Тогда квантиль (quantile) случайной величины для вероятности p определяется как $F^{-1}(p)$. В частности, при $p = 1/2$ получаем медиану распределения.

Как аббревиатура слова **quantile** может использоваться буква **Q**. Например, **Q-Q plot** — это сокращение от **quantile-quantile plot**.

Если отрезок $[0, 1]$ разделить на n равных промежутков, то для соответствующих вероятностей $1/n, \dots, (n-1)/n$ имеем семейство «равномерных» квантилей. Второе значение слова **quantile** в английском языке — квантиль из такого семейства. В этом смысле i -я квантиль равна $F^{-1}(i/n)$, $i = 1, \dots, n-1$.

Для некоторых n у квантилей, соответствующих разбиению на n частей, есть специальные названия, происходящие от соответствующих латинских числительных. В частности, децили соответствуют разбиению на 10 частей, квартили — на 4 части, (более редко встречающиеся) квинтили — на 5 частей. Слова **quantile**, **quartile** и **quintile** легко спутать, но надо помнить, что первый из этих терминов общий, в отличие от остальных двух. (А еще есть английское слово **quantal** — «квантовый».)

Из квантилей во втором смысле особо выделяют нижнюю (**lower**), то есть первую, и верхнюю (**upper**), то есть $(n-1)$ -ю. Так нижняя квартиль соответствует вероятности $1/4$, а верхняя — вероятности $3/4$ (еще одна квартиль — вторая — это, конечно, просто медиана). Разность между верхней и нижней квартилями называют интерквартильным размахом (**interquartile range**). Соответственно, разность между девятой и первой децилями называют интердецильным размахом.

Процентиль (**percentile**, иногда просто **centile**) соответствует разбиению на 100 промежутков. Термин **percentile** может использоваться почти как синоним слова **quantile** в первом смысле. По-видимому, это связано с тем, что 100 промежутков — это уже «много» и более мелкое деление с практической точки зрения не всегда нужно. Квантили в привычном нам первом смысле по-английски называют также **fractile**. То есть “0.35-fractile” — это “35th percentile”. (**Fractile** желательно не путать со словом **fractal** — «фрактал».)

Support: носитель (распределения).

Quantile: квантиль.

Median: медиана.

Q-Q plot (quantile-quantile plot): диаграмма квантиль-квантиль, Q-Q диаграмма.

Quartile/quintile/decile: квартиль/квинтиль/дециль.

Upper/lower quartile: нижняя/верхняя квартиль.

Interquartile/interdecile: интерквартильный/интердецильный, межквартильный/междецильный.

Interquartile range: интерквартильный размах, межквартильный размах, интерквартильная широта.

Percentile (centile): процентиль, перцентиль, квантиль.

Fractile: фрактиль, квантиль.

Quantile: журнал *Квантиль*.

Nested, encompassing

Когда модель A является частным случаем модели B , т. е. A можно получить из B , наложив на параметры некоторые ограничения (и, возможно, сделав замену переменных), то говорят, что модель A вложена (*nested*) в модель B . Если же A и B не сводятся друг к другу с помощью наложения ограничений на параметры, то говорят, что A и B являются невложенными (*non-nested*); примером могут служить модели логит и пробит. Если правильность спецификации одной модели проверяют на основе сопоставления с невложенной альтернативной моделью, то говорят о невложенных гипотезах. Термин связан не с прямым значением глагола *to nest* — «гнездиться», а с переносным: «помещать один объект внутрь другого» (как одну матрешку внутрь другой) или «содержать другой объект». Соответственно *to be nested* (редко также *to nest*) «находиться внутри другого объекта».

В том же контексте выбора модели есть похожий термин *encompassing* (охват). *Принцип охвата* заключается в том, что если модель правильно специфицирована, то она должна быть способна объяснять результаты использования альтернативной модели. О выборе модели, невложенных и охватывающих моделях см. Gougieloux & Monfort (1994), Pesaran & Weeks (2001), Greene (2012).

Кроме вложенных моделей и гипотез в эконометрике известна также модель, называемая *nested logit*, в основе которой лежит иерархия вложенных друг в друга структур — множеств альтернатив. См. Greene (2012).

Nested/non-nested: вложенный/невложенный.

Nested model/hypothesis: вложенная модель/гипотеза.

Nested logit: вложенный логит.

Nest (гл.): вкладывать, включать (как частный случай).

Nest (сущ.): «гнездо», подмножество альтернатив (во вложенном логите).

Be nested: вкладываться, быть частным случаем.

Nesting: вложение (одной гипотезы/модели/структуры в другую).

Encompass: охватывать, заключать в себе.

Encompassing: охват; охватывающий, заключающий в себе.

Encompassing principle: принцип охвата.

Eigenvalue

Приставка *eigen-*, используемая в английских словах *eigenvalue* и *eigenvector* имеет немецкое происхождение и означает «собственный». Эти термины — частичный перевод с немецкого, где они имеют вид *Eigenwert* и *Eigenvektor*.

Eigenvalue: собственное значение, собственное (характеристическое) число.

Eigenvector: собственный вектор.

Marginal

В эконометрику термин **marginal** попал разными путями. С одной стороны, это общий статистический термин, относящийся к распределениям нескольких случайных величин. Например, если есть две случайные величины X и Y , то можно рассматривать их совместное распределение, условное распределение одной переменной относительно другой, а также безусловное распределение одной переменной. Последнее распределение называется в таком контексте маргинальным¹ или частным. Кроме распределений, маргинальными могут быть плотности, вероятности и т. п. С этим значением связаны производные термины **marginalize** и **marginalization**.

С другой стороны, в экономической теории есть понятие предельного влияния одной переменной на другую (предельная полезность, предельная производительность). С этим связано понятие предельного влияния (эффекта) объясняющей переменной на ожидаемое значение или другую характеристику объясняемой переменной в эконометрических моделях. Для модели линейной регрессии линейной по регрессорам предельный эффект регрессора — это его коэффициент. Но в более общем случае коэффициенты эконометрической модели могут не иметь такой очевидной интерпретации и тогда для анализа результатов оценивания имеет смысл использовать подходящим образом определенное предельное значение (маргинальное значение). Например в модели логит это может быть производная расчетной вероятности по регрессору.

Есть также использование термина **marginal** в общем значении «граничный», например **marginal significance level**.

Margin: край, граница, грань, предел, разница, преимущество.

Marginal: маргинальный, маржинальный, частный (в статистике), предельный (в экономической теории), граничный.

Marginal distribution: маргинальное (маржинальное) распределение, частное распределение, безусловное распределение (в отличие от условного).

Marginal (сущ.): маргинальное (маржинальное) распределение.

Joint distribution: совместное распределение.

Conditional distribution: условное распределение.

Unconditional distribution: безусловное распределение.

Marginalize: маргинализировать, получать маргинальное распределение из совместного распределения.

Marginalization: маргинализация, получение маргинального распределения из совместного.

Marginal effect: предельное (маргинальное) влияние, предельный (маргинальный) эффект.

Marginal value: предельное (маргинальное) значение.

Marginal significance level: граничный уровень значимости.

On the margin: на грани, с небольшой разницей.

¹Относительно термина «маргинальный» (буквально «лежащий на краю/границе») в русском языке есть неоднозначность. Вариант «маргинальный» отражает заимствование из средневековой латыни (наряду с «маргиналиями» — заметками на полях). Несколько менее распространен вариант «маржинальный», соответствующий терминам «маржа», «маржинализм», заимствованным из французского языка и отражающим французское произношение.

Smoothing, bandwidth, kernel

Непараметрические методы оценивания, дающие сглаженную оценку некоторой функции (плотности, регрессии или спектральной плотности), обычно имеют параметр, от которого зависит степень гладкости оценки. Для так называемых ядерных оценок естественным параметром сглаживания является **bandwidth** — ширина полосы или ширина окна, отвечающая за масштабирование ядерной функции. Ядерная оценка в данной точке — это взвешенное среднее несглаженных данных, причем вес, задаваемый масштабированной ядерной функцией, убывает до нуля при удалении от точки. «Далекие» данные входят с нулевым или незначительным весом. (Встречаются и волнообразные ядерные функции, и функции с областями отрицательных значений, но и они в пределе стремятся к нулю.) Когда ширина полосы большая, в оценке используется широкая окрестность данной точки и оценка получается гладкой, т. е. ее дисперсия низкой.

Название по-видимому связано с обработкой сигналов и спектральным анализом. **Band** в этом контексте — это полоса (т. е. интервал) частот, а **bandwidth** — это ширина такой полосы. При сглаживании в частотной области — получении ядерной оценки спектральной плотности на основе периодограммы — ядерную функцию принято называть (спектральным) окном. Отсюда термин «ширина окна» в русском языке. В английском языке **window width** тоже встречается, но не так часто. Хотя ядерные оценки плотности и ядерные регрессии не имеют отношения к полосе частот, но для соответствующего параметра сглаживания используют тот же термин **bandwidth**.

О непараметрической регрессии можно прочесть в Härdle (1990) (русский перевод Хардле 1993) и Расин (2008), о сглаженных оценках спектральной плотности — в Jenkins & Watts (1998) (русский перевод Дженкинс & Ваттс 1971) и Hamilton (1994).

Smoothing: сглаживание.

Undersmoothing/oversmoothing: недостаточное/избыточное сглаживание.

Smoothing parameter: параметр сглаживания.

Kernel: ядро.

Kernel estimation: ядерное оценивание.

Kernel function: ядерная функция.

Spectral window: спектральное окно.

Bandwidth: ширина полосы, ширина окна.

Window width: ширина окна.

Back-testing

Термин **back-testing** пришел из финансов. Если, например, трейдер хочет использовать новую торговую стратегию, то прежде, чем работать с реальными денежными средствами, ему имеет смысл опробовать ее на исторических данных. В эконометрике этот термин применяют к прогнозным моделям, когда проверяется, насколько качественные прогнозы некоторого (обычно финансового) показателя данная модель давала бы, если бы была использована в прошлом. В терминологии Дж. Стока и М. Уотсона (Stock & Watson 2011) такая имитация работы прогнозной модели в реальном времени — это псевдо-вневыборочное прогнозирование (**pseudo out-of-sample forecasting**).

Back-testing: обратное тестирование, историческое тестирование, бэктестинг.

Out-of-sample forecasting: вневыборочное прогнозирование.

Time domain, frequency domain

Само слово **domain** общенаучное и может переводиться как «область» или «сфера». В частности, в математике это область определения функции.

Разделение **time domain** и **frequency domain** (временная область и частотная область) пришло из инженерных приложений, в которых сигнал (например, звуковой) можно представить как функцию времени, т. е. временной ряд, а можно как функцию от частоты. Такие понятия как «лаг», «автоковариация», «автокорреляция» и т. п. относятся к временной области, а «периодограмма», «спектр» — к частотной. Слабо стационарный случайный процесс можно представить как во временной, так и в частотной области и описание процесса в одной области имеет аналоги в другой области. Связь между двумя областями определяется преобразованием Фурье. Во многом выбор области — это вопрос простоты представления; см. Hamilton (1994, p. 152). Например, определение причинности по Грейнджеру (Granger 1969) было изначально дано в частотной области, но сейчас практически всегда с ней работают во временной области.

Time domain: временная область.

Frequency domain: частотная область.

Domain: область, сфера.

Domain (of function): область определения функции.

Spectrum: спектр.

Spectral: спектральный.

Spectral density: спектральная плотность.

Power spectrum: спектр мощности.

Periodogram: периодограмма.

Fourier transform (transformation): преобразование Фурье.

Cross-spectrum: кросс-спектр.

Bispectrum: биспектр.

Granger causality: причинность/обусловленность по Грейнджеру.

Список литературы

Дженкинс, Г. & Д. Ваттс (1971). *Спектральный анализ и его приложения*. Москва: Мир.

Расин, Дж. (2008). Непараметрическая эконометрика: вводный курс. *Квантиль* №4, 7–56.

Хардле, В. (1993). *Прикладная непараметрическая регрессия*. Москва: Мир.

Gourieroux, C. & A. Monfort (1994). Testing non-nested hypotheses. Глава 44 в *Handbook of Econometrics* под редакцией R. F. Engle & D. L. McFadden, Elsevier, том 4, стр. 2583–2637.

- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438.
- Greene, W.H. (2012). *Econometric Analysis* (7th edition). Boston: Prentice Hall.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.
- Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.
- Jenkins, G. M. & D. Watts (1998). *Spectral Analysis and Its Applications*. Boca Raton: Emerson Adams.
- Pesaran, M. H. & M. Weeks (2001). Nonnested hypothesis testing: An overview. Глава в *A Companion to Theoretical Econometrics* под редакцией В. Н. Baltagi, Blackwell Publishing, 279–309.
- Stock, J. H. & M. W. Watson (2011). *Introduction to Econometrics* (3rd edition). Amsterdam: Addison-Wesley Longman.

A mini-dictionary of English econometric terminology III

Alexander Tsyplakov

Novosibirsk State University, Novosibirsk, Russia

This part of the dictionary comments on English econometric terms estimator, quantile, nested, marginal, and some others. Emphasis is again placed on accurate definitions of their meaning to avoid possible confusion and incorrect interpretation.

Статьи: финансовая эконометрика

Непостоянные во времени вьющиеся копулы в многомерном анализе доходностей*

Олег Грошев[†]

Barclays Capital, Москва, Россия

В данной работе анализируется многомерное распределение финансовых доходностей с помощью непостоянных во времени вьющихся копул. В ней представлена оценка d -шагового метода максимального правдоподобия (dSML), и показано, что эта оценка является не только состоятельной и асимптотически нормальной, но еще и более вычислительно привлекательной, чем оценка метода максимального правдоподобия или оценка Паттона. Используя оценку dSML, мы подгоняем вьющиеся копулы под доходности валют развивающихся стран, производим диагностику и выбор модели.

Ключевые слова: многомерные временные ряды, вьющиеся копулы, d -шаговый метод максимального правдоподобия, вычислительная эффективность

Классификация JEL: C13, C58, G15

1 Введение

Моделирование многомерных данных при помощи копул ни в коей мере не является новым. Начиная с оригинальной работы Sklar (1959), копулы находят все большее применение в статистике, эконометрике и финансах. Основная идея использования копул заключается в декомпозиции многомерной функции распределения на две компоненты. Первая соответствует маргинальным функциям распределения данного многомерного распределения. Вторая компонента называется функцией-копулой и отражает зависимость компонент многомерных данных. Таким образом, многомерная плотность можно оценить раздельно — сначала маргинальные плотности, потом копулу. Процесс оценивания особенно прост для случайной выборки (независимых и одинаково распределенных наблюдений).

В контексте временных рядов моделирование с помощью копул сложнее и является более недавним открытием. Паттон в своей диссертации (Patton, 2002) доказал аналог теоремы Склара для условных распределений. Эта теорема позволяет разложить логарифм функции правдоподобия на сумму логарифмов маргинальных функций правдоподобия и лог-плотности копулы. Разложение абсолютно аналогично случаю независимых и одинаково распределенных величин. Паттон предложил двухшаговый метод максимального правдоподобия (2SML) для оценки параметров многомерного распределения. Сначала, используя стандартный метод МП, требуется оценить параметры маргинальных распределений. Затем следует подставить полученные оценки параметров в логарифм функции правдоподобия копулы и, используя МП еще раз, найти параметры копулы. Паттон показал, что 2SML дает состоятельные оценки параметров; кроме того, он вывел формулу для асимптотической ковариационной матрицы. Замечателен подход Паттона тем, что он позволяет зависимости между временными рядами меняться со временем. Несколько спецификаций для условной зависимости были предложены в Jondeau & Rockinger (2006), а также в Patton (2006).

*Цитировать как: Грошев, Олег (2014). «Непостоянные во времени вьющиеся копулы в многомерном анализе доходностей», Квантиль, №12, стр. 53–67. Citation: Groshev, Oleg (2014). “Time varying vine copulas for multivariate returns,” *Quantile*, No.12, pp. 53–67.

[†]Адрес: 123056, г. Москва, Большая Грузинская ул., 71. Электронная почта: oleg.groshev@gmail.com

Несмотря на простоту идеи, теорию копул сложнее применять для многомерных данных в случае размерности большей двух. Существует множество двумерных копул; их обзор приведен в Trivedi & Zimmer (2005). В то же время моделирование в больших размерностях до недавних пор было ограничено эллиптическими копулами (например, Гауссовой или копулой Стьюдента). Хотя такие модели превосходят на финансовых данных модели, основанные на многомерном нормальном или Стьюдентовом распределении, они все равно являются «слишком линейными» в том смысле, что не могут уловить асимметричную зависимость в хвостах. Хорошее введение в эти вопросы содержится в Patton (2004).

Joe (1996) и Bedford & Cooke (2002) предложили использовать вьющиеся копулы для моделирования многомерных данных. Этот подход требует оценки $d(d-1)/2$ двумерных копул; многомерная копула затем конструируется из этих двумерных. Данный подход обладает большой свободой в том, как моделируется зависимость между компонентами многомерных данных. Недавно Aas, Czado, Frigressi & Bakken (2009) оценили несколько вьющихся копул с постоянными параметрами зависимости. Они использовали стандартный метод ML, который потребовал значительных вычислительных мощностей. В их статье была предложена итеративная процедура для нахождения хорошего начального значения для решения оптимизационной задачи ML. Оказалось возможным довести эту процедуру (по аналогии с методом 2SML Паттона) до dSML оценки. В данном случае d соответствует количеству шагов (равному размерности многомерных данных).

Основной вклад данного исследования – это строгое построение dSML оценки и доказательство ее состоятельности и асимптотической нормальности. Преимущество dSML составляет снижение размерности задачи и вычислительных издержек. Поскольку вьющиеся копулы имеют разложение по попарным копулам, одна многомерная оптимизационная задача разбивается на примерно $d(d+1)/2$ небольших задач. Кроме того, данный подход позволяет зависимости между рядами меняться со временем.

В эмпирической части работы моделируются валютные курсы трех европейских развивающихся стран (CZK Чехии, PLN Польши и HUF Венгрии) и зависимость между ними. Начав с моделей GARCH/GJR для маргинальных распределений, я двигаюсь к попарным копулам, соединяющим маргинальные распределения, и, наконец, к копуле, соединяющей парные копулы. В качестве парных копул используются копулы Стьюдента с эволюцией параметра зависимости из Tse & Tsui (2002).

Выбор этих конкретных финансовых рядов продиктован несколькими причинами. Во-первых, довольно мало исследований посвящено эконометрическому моделированию валютных курсов развивающихся стран. Во-вторых, моделирование зависимости валютных курсов развивающихся стран может пролить свет на существование финансового заражения среди этих стран. Например, Rodriguez (2007) использовал копулы с Марковским переключением режима для моделирования динамики цен акций в развивающихся странах во время кризиса. Наконец, я выбрал валюты трех стран из Центральной Европы по той причине, что, несмотря на принадлежность к рискованным активам, они демонстрируют более стабильное поведение. Большинство других валют развивающихся стран имеют жесткую привязку к твердой валюте (USD или EUR), которая при разрушении вызывает девальвацию, а также большую волатильность и куртозис. В действительности, в течение рассматриваемого периода HUF и PLN плавали относительно свободно, в то время как CZK был крепче привязан к евро.

Обнаружилась асимметрия в поведении данных обменных курсов, что делает их скорее похожими на цены акций или других рискованных активов, чем на обменные курсы развитых стран. А именно, модель GJR из Glosten, Jagannathan & Runkle (1993) подгоняется под HUF и PLN значительно лучше, чем стандартный GARCH. GJR предполагает, что существует асимметричная обратная связь для волатильности; это означает, что девальвация валюты приводит к более сильной волатильности на следующий день. Tse (1998) показал,

что этот эффект не наблюдается в обменных курсах развитых стран. Кроме того, существует асимметрия распределения дневных доходностей, выражающаяся в ненулевой скошенности распределения. Оказывается, что модели со скошенным распределением Стьюдента в качестве распределения ошибки подгоняются под HUF и PLN лучше, чем стандартный Стьюдент. Однако, асимметричные модели не обеспечивают лучшую подгонку под CZK, сильнее привязанному к евро. Этот факт косвенно подтверждает гипотезу о рискованном поведении валют развивающихся стран по сравнению с симметричным поведением валют развитых стран. Предыдущие исследования асимметричности валют развивающихся стран получали неоднозначные результаты. Так Longmore & Robinson (2004) находят асимметричность в поведении доллара Ямайки, в то время как Sandoval (2006) оспаривает прогнозные превосходство моделей с асимметричной волатильностью для валют Латинской Америки. Griebeler (2010) утверждает, что нет связи между качеством подгонки моделей с асимметричной волатильностью и принадлежностью валюты к развивающимся/развитым странам.

2 Условные копулы

Краеугольным камнем применения теории копул для временных рядов является теорема Паттона, аналог теоремы Склера для условных копул.

Теорема 1 (Patton 2002) Пусть F будет функцией распределения (y_1, \dots, y_d) условно на w , F_i будут (маргинальными) функциями распределения y_i условно на w . Предположим, что F непрерывна. Тогда существует $C(\cdot | w) : [0, 1]^d \rightarrow [0, 1]$, называемая (условной) копулой, такая, что

$$F(y_1, \dots, y_d | w) = C(F_1(y_1 | w), \dots, F_d(y_d | w) | w).$$

Если у F есть плотность f и, следовательно, у F_i есть плотности f_i , тогда следующая формула верна для плотностей:

$$f(y_1, \dots, y_d | w) = c(F_1(y_1 | w), \dots, F_d(y_d | w) | w) \prod_{i=1}^d f_i(y_i | w), \quad (1)$$

где $c = \partial_1 \dots \partial_d C$ и ∂_i обозначает частную производную по i -му аргументу.

Пусть $y^t = (y_1^t, \dots, y_d^t)$, $t = 1, \dots, T$ описывает многомерный временной ряд. Обозначим многомерную и маргинальную историю как

$$\mathcal{Y}^{t-1} = \{y^{t-1}, \dots, y^1\} \quad \text{и} \quad \mathcal{Y}_i^{t-1} = \{y_i^{t-1}, \dots, y_i^1\}.$$

Логарифм функции правдоподобия, характеризующей систему, есть

$$\mathcal{L} = \sum_{t=1}^T \ell^t = \sum_{t=1}^T \ln f(y^t | \mathcal{Y}^{t-1}; \theta).$$

Используя (1), можно разложить

$$\mathcal{L} = \sum_{i=1}^d \mathcal{L}_i + \mathcal{L}_{\text{Cоп}},$$

где

$$\mathcal{L}_i = \sum_{t=1}^T \ell_i^t, \quad \mathcal{L}_{\text{Cоп}} = \sum_{t=1}^T \ell_{\text{Cоп}}^t.$$

Здесь

$$\ell_i^t = \ln f_i(y_i^t | \mathcal{Y}_i^{t-1}; \theta_i),$$

в то время как лог-функция правдоподобия копулы вычисляется как

$$\ell_{\text{Cоп}}^t = \ln c(F_1(y_1^t | \mathcal{Y}_1^{t-1}; \theta_1), \dots, F_d(y_d^t | \mathcal{Y}_d^{t-1}; \theta_d) | \mathcal{Y}^{t-1}; \theta_{\text{Cоп}}).$$

Теперь мы хотим оценить $\theta = (\theta_1^\top, \dots, \theta_d^\top, \theta_{\text{Cоп}}^\top)^\top$. Patton (2002) предлагает следующую 2SML процедуру. Сначала оцениваем

$$\hat{\theta}_i = \operatorname{argmax}_{\theta_i} \mathcal{L}_i(\theta_i).$$

Затем подставляем $\hat{\theta}_i$ в $\ell_{\text{Cоп}}^t$ и максимизируем выбирая $\theta_{\text{Cоп}}$.

$$\hat{\theta}_{\text{Cоп}} = \operatorname{argmax}_{\theta_{\text{Cоп}}} \mathcal{L}_{\text{Cоп}}(\hat{\theta}_1, \dots, \hat{\theta}_d, \theta_{\text{Cоп}}).$$

Замечание 2 Patton (2002) доказал, что в довольно скромных предположениях 2SML производит состоятельную и асимптотически нормальную оценку. Хотя она и не является эффективной, она может быть приведена к ней с помощью состоятельно оцененной асимптотической ковариационной матрицы. Вычислительно 2SML является гораздо более эффективным, чем ML, примененный к многомерной плотности.

3 Вьющиеся копулы

Вьющиеся копулы были впервые предложены Джо (1996) для моделирования многомерных данных через конструкции с попарными копулами. Слегка изменив обозначения, будем обозначать d -мерную функцию распределения и плотность (y_1, \dots, y_d) как $F_{1\dots d}$ и $f_{1\dots d}$. Также обозначим за $F_{ij|\mathbf{k}}$ двумерную функцию распределения (y_i, y_j) условную на $(y_{k_1}, \dots, y_{k_p})$, где (k_1, \dots, k_p) обозначается мульти-индексом \mathbf{k} . Используя теорему Паттона, получаем

$$F_{ij|\mathbf{k}} = C_{ij|\mathbf{k}}(F_{i|\mathbf{k}}, F_{j|\mathbf{k}}),$$

или

$$f_{ij|\mathbf{k}} = c_{ij|\mathbf{k}}(F_{i|\mathbf{k}}, F_{j|\mathbf{k}})f_{i|\mathbf{k}}f_{j|\mathbf{k}}$$

для плотностей. Используя разложение плотностей типа

$$f_{1\dots d} = f_d f_{d-1|d} \dots f_{1|2\dots d} \tag{2}$$

и формулы Джо

$$F_{i|\mathbf{k}} = \partial_2 C_{ij|\mathbf{k}}(F_{i|\mathbf{k}}, F_{j|\mathbf{k}}) \tag{3}$$

и

$$f_{i|\mathbf{k}} = c_{ij|\mathbf{k}}(F_{i|\mathbf{k}}, F_{j|\mathbf{k}})f_{i|\mathbf{k}},$$

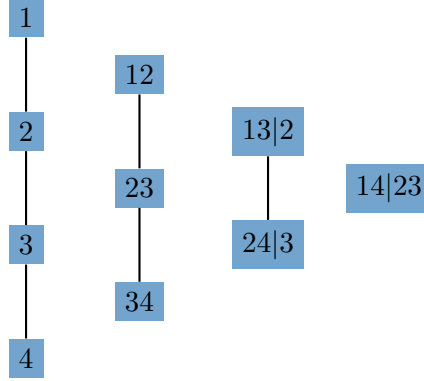
можно построить следующее разложение:

$$f_{1\dots d} = \prod_{k=1}^d f_k \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j|i+1,\dots,i+j-1}. \tag{4}$$

Это не единственное разложение такого типа. Используя другие факторизации (2), можно получить другие разложения. Bedford & Cooke (2002) предложили представлять формулы вида (4) с помощью диаграмм.

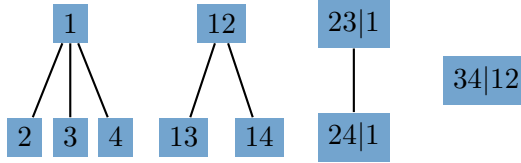
Пример 3 Следующее разложение соответствует изображенному объединению деревьев (лесу). Когда каждое дерево в лесу является линейным графом, разложение называется *D-вьюн (D-vine)*.

$$f_1 f_2 f_3 f_4 \cdot c_{12} c_{23} c_{34} \cdot c_{13|2} c_{23|4} \cdot c_{14|23}$$



Пример 4 Другой лес будет соответствовать другому разложению. Когда каждое дерево в лесу имеет один корень, а все остальные вершины являются листьями, разложение называется *канонический вьюн (C-vine)*.

$$f_1 f_2 f_3 f_4 \cdot c_{12} c_{13} c_{14} \cdot c_{23|1} c_{24|1} \cdot c_{34|12}$$



Замечание 5 Не каждый лес производит вьюн-разложение. Леса, которые могут производить вьюны, будем называть *плодородными*.

Пусть \mathcal{F} будет плодородным лесом. Обозначим $\text{Vert } \mathcal{F}$ вершины леса, $\nu \in \text{Vert } \mathcal{F}$ будет иметь вид $ij|\mathbf{k}$ для некоторых i, j, \mathbf{k} (некоторые из них могут быть пустыми как в вершине i соответствующей маргинальным f_i). Для удобства обозначим $c_i = f_i$. Тогда вьюн-разложение может быть записано как

$$f_{1\dots d} = \prod_{\nu \in \text{Vert } \mathcal{F}} c_\nu. \tag{5}$$

Каждому $\nu = ij|\mathbf{k}$ мы можем сопоставить степень $\deg \nu = \dim i + \dim j + \dim \mathbf{k}$. Здесь $\dim i$ равна 0, если i пусто и 1 иначе; $\dim \mathbf{k} = p$ для $\mathbf{k} = (k_1, \dots, k_p)$. Обозначим $\text{Vert}_s \mathcal{F} = \{\nu \in \mathcal{F} \mid \deg \nu = s\}$. Следовательно, мы имеем $\text{Vert } \mathcal{F} = \coprod_{s=1}^d \text{Vert}_s \mathcal{F}$. Таким образом, мы можем перегруппировать (5) в

$$f_{1\dots d} = \prod_{s=1}^d \prod_{\nu \in \text{Vert}_s \mathcal{F}} c_\nu. \tag{6}$$

4 Оценка dSML

В этом разделе опишем процедуру dSML. Вектор параметров, который нужно оценить, обозначим $\theta = (\theta_\nu^\top \mid \nu \in \text{Vert } \mathcal{F})^\top$. Обозначим $\theta_{[s]} = (\theta_\nu^\top \mid \nu \in \text{Vert } \mathcal{F}, \deg \nu \leq s)^\top$,

$\theta_{(s)} = (\theta_\nu^\top \mid \nu \in \text{Vert}_s \mathcal{F})^\top$. Также введем обозначение $y_{\mathbf{k}}^t = (y_{k_1}^t, \dots, y_{k_p}^t)$ для временных подрядов, а $\mathcal{Y}_{\mathbf{k}}^{t-1} = \{y_{\mathbf{k}}^{t-1}, \dots, y_{\mathbf{k}}^1\}$ для подыстории. Логарифм функции правдоподобия запишется как

$$\mathcal{L} = \sum_{t=1}^T \ell^t = \sum_{t=1}^T \ln f_{1\dots d}(y^t \mid \mathcal{Y}^{t-1}; \theta).$$

Теперь используя (6), разложим

$$\mathcal{L} = \sum_{s=1}^d \mathcal{L}_{(s)} = \sum_{s=1}^d \sum_{\nu \in \text{Vert}_s \mathcal{F}} \mathcal{L}_\nu, \quad \mathcal{L}_\nu = \sum_{t=1}^T \ell_\nu^t.$$

Здесь

$$\ell_{ij|\mathbf{k}}^t = \ln c_{ij|\mathbf{k}}(F_{i|\mathbf{k}}(y_i^t \mid \mathcal{Y}_{\mathbf{k}}^{t-1}; \theta_{[s-1]}), F_{j|\mathbf{k}}(y_j^t \mid \mathcal{Y}_{\mathbf{k}}^{t-1}; \theta_{[s-1]} \mid \mathcal{Y}_{\mathbf{k}}^{t-1}; \theta_{ij|\mathbf{k}}), \quad (7)$$

где $\deg ij|\mathbf{k} = s$. Причина, по которой $F_{i|\mathbf{k}}$ зависит только от $\theta_{[s-1]}$, — это формула Джо (3). А именно, пусть $\mathbf{k} = (k, \mathbf{l})$, затем

$$F_{i|\mathbf{k}\mathbf{l}} = \partial_2 C_{ik|\mathbf{l}}(F_{i|\mathbf{l}}(y_i^t \mid \mathcal{Y}_{\mathbf{l}}^{t-1}; \theta_{[s-2]}), F_{k|\mathbf{l}}(y_k^t \mid \mathcal{Y}_{\mathbf{l}}^{t-1}; \theta_{[s-2]} \mid \mathcal{Y}_{\mathbf{l}}^{t-1}; \theta_{ik|\mathbf{l}}).$$

Поскольку $\deg ik|\mathbf{l} = s - 1$, предложение доказано методом математической индукции.

Метод математической индукция дает идею, как оценить параметры. Сначала требуется оценить параметры маргинальных распределений:

$$\hat{\theta}_i = \operatorname{argmax}_{\theta_i} \mathcal{L}_i(\theta_i).$$

$\hat{\theta}_{[1]}$ теперь равняется $(\hat{\theta}_i^\top \mid i = 1, \dots, d)^\top$. Предполагая, что мы имеем оценку $\hat{\theta}_{[s-1]}$, можно подставить ее в (7), чтобы найти

$$\hat{\theta}_{ij|\mathbf{k}} = \operatorname{argmax}_{\theta_{ij|\mathbf{k}}} \mathcal{L}_{ij|\mathbf{k}}(\hat{\theta}_{[s-1]}, \theta_{ij|\mathbf{k}}).$$

Затем $\hat{\theta}_{[s]} = (\hat{\theta}_\nu^\top \mid \nu \in \text{Vert} \mathcal{F}, \deg \nu \leq s)^\top$, и процедура повторяется до шага $s = d$. Наконец, $\hat{\theta} = \hat{\theta}_{[d]}$.

Замечание 6 В приложении доказано, что в присутствии стандартных предположений $dSML$ приводит к состоятельной и асимптотически нормальной оценке (в точности как $2SML$). Также вычислена асимптотическая ковариационная матрица оценки, которую можно использовать для вывода скорректированной оценки Ньютона–Рафсона. Эта оценка уже будет эффективна.

5 Описание данных

Мы используем курсы цен покупки на закрытие дня валютных пар восточноевропейских стран USD/CZK, USD/HUF и USD/PLN, представленные на сайте FX-брокера OANDA Corp¹. Котировка USD/PLN 4.5 означает, что 1 доллар США обменивается на 4,5 польских злотых. Рассматриваемый временной промежуток длится с декабря 1999 г. до декабря 2010 г., образуя 2869 наблюдений. Обменные курсы S_t преобразуются в дневные логарифмические доходности в процентах $r_t = 100 \ln S_t / S_{t-1}$. Сводные статистики представлены в таблице 1. Положительная скошенность означает, что в среднем валюты испытывают более сильные девальвации, чем ревальвации, а положительный (добавочный) куртозис подтверждает обычное для финансовых доходностей распределение с толстыми хвостами.

Таблица 1: Сводные статистики доходностей USD/CCY

Статистика	CZK	HUF	PLN
Среднее	-0,022	-0,006	-0,012
Медиана	-0,008	-0,035	-0,033
Стандартное отклонение	0,771	0,883	0,872
Скошенность	0,138	0,456	0,448
Куртозис	3,14	5,07	5,56
Минимум	-4,20	-5,21	-5,73
Максимум	4,45	6,29	6,05

Замечания: Выборка с декабря 1999 г. до декабря 2010 г. (2869 наблюдений). Данные представлены для $r_t = 100 \ln S_t/S_{t-1}$.

Таблица 2: Тесты Льюнга–Бокса

Статистика Льюнга–Бокса	CZK		HUF		PLN	
$Q_r(50)$	76,57	[0,0091]	102,22	[0,0000]	176,98	[0,0000]
$Q_{rr}(50)$	983,32	[0,0000]	946,33	[0,0000]	2511,75	[0,0000]

Замечание: p-Значения представлены в квадратных скобках.

В таблице 2 представлен результаты тестов Льюнга–Бокса с 50 лагами, которых достаточно для примерно 3000 наблюдений (Hsieh 1989). Результаты тестов показывают автокорреляцию и гетероскедстичность доходностей.

Однако, если судить по графику 1, можно увидеть, что нет очевидной сезонности, соответствующей дням недели. Поэтому мы будем опускать переменные-индикаторы, соответствующие дням недели, в спецификации условного среднего. На графике 2 изображены диаграммы рассеяния доходностей валютных пар. Валюты стран центральной Европы демонстрируют сильную корреляцию обменных курсов.

6 Моделирование маргинальных распределений

Следуя теоретическому подходу, описанному в разделах 2 и 3, мы сначала моделируем одномерные ряды доходностей. Для этой цели мы используем модели GARCH и GJR. Условное среднее моделируется процессом AR(1); однако, в случае USD/CZK авторегрессионная компонента не является значимой, не улучшает подгонку, и поэтому отбрасывается. Условная дисперсия моделируется как GARCH(1,1) и GJR(1,1), модель, впервые предложенная в Glosten, Jagannathan & Runkle (1993). Нормализованные ошибки предполагаются условно распределенными по Стьюденту или скошенному Стьюденту. Скошенность вводится в распределение Стьюдента по Fernandez & Steel (1998):

$$p(z; \xi) = \frac{2}{\xi + 1/\xi} (f(z/\xi)\mathbf{I}_{\{z \geq 0\}} + f(z\xi)\mathbf{I}_{\{z \leq 0\}}),$$

где f — стандартное распределение Стьюдента с κ степеней свободы. Скошенность равняется

$$\mathbf{E}z^3 = \frac{\xi^4 - 1/\xi^4}{\xi + 1/\xi} \mathbf{M}_3,$$

¹www.oanda.com/currency/historical-rates

График 1: Функция автокорреляции для доходностей и квадратов доходностей

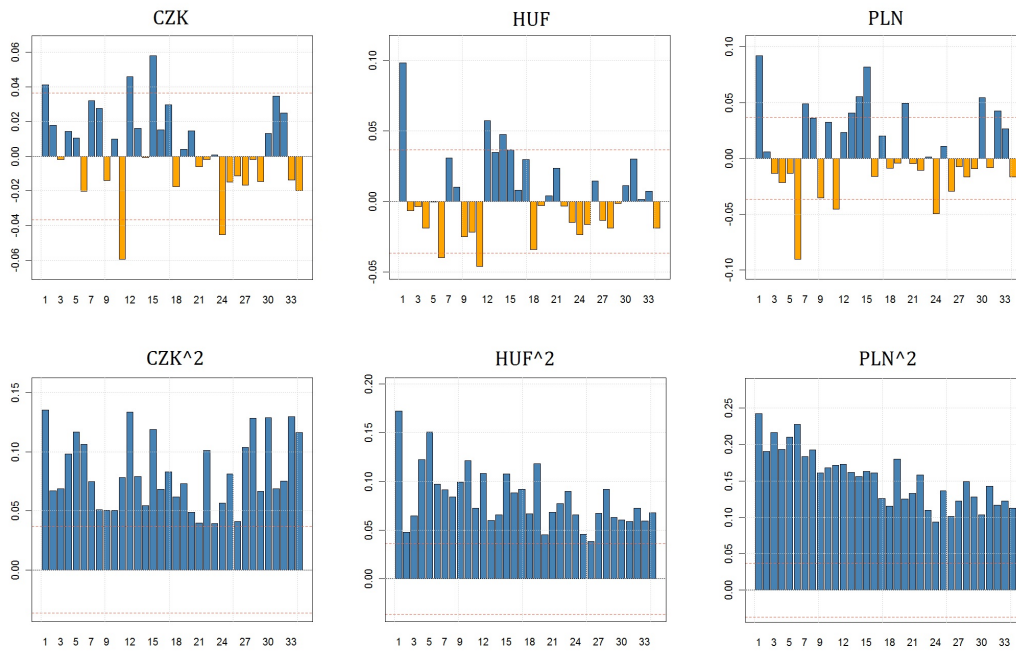
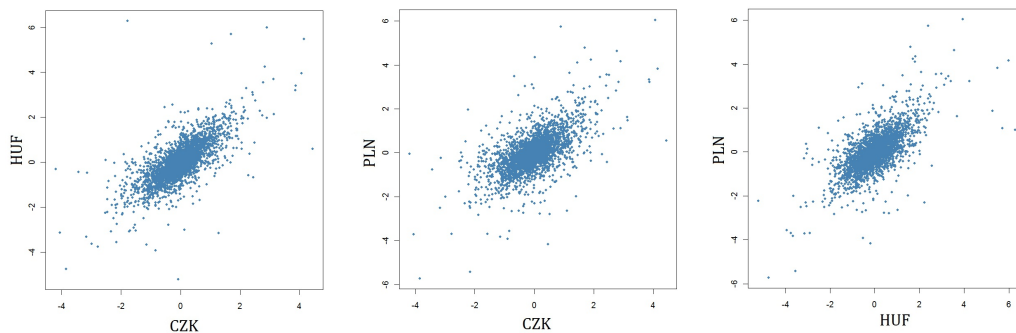


График 2: Диаграммы рассеяния доходностей пар USD/CCY



где

$$M_3 = 2 \int_0^{\infty} s^3 f(s) ds.$$

Таким образом, необходимо оценить параметры в спецификации для условного среднего

$$r_t = \mu + \varphi r_{t-1} + \epsilon_t, \quad \text{где } \epsilon_t = \sigma_t z_t \text{ и } z_t \propto St(\xi, \kappa).$$

и условной дисперсии GARCH(1,1)

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

или GJR(1,1)

$$\sigma_t^2 = \omega + \alpha(|\epsilon_{t-1}| - \gamma \epsilon_{t-1})^2 + \beta \sigma_{t-1}^2.$$

Коэффициенты в уравнении для дисперсии имеют простую экономическую интерпретацию. Постоянство волатильности соответствует β , дисперсия GARCH на длинных периодах равняется $\omega/1 - \alpha - \beta$, а γ измеряет величину эффекта рычага. То есть, отрицательная γ с большим абсолютным значением означает, что девальвация валюты вызывает больший всплеск волатильности, чем ревальвация.

Чтобы оценить уравнения, я использую метод максимального правдоподобия (ML). Вычисления выполнены в R с помощью пакета fGarch. В таблице 3 представлены значения информационных критериев Акаике и Шварца для оцененных моделей. Можно заметить, что скошенный GJR превосходит другие модели для USD/HUF и USD/PLN, в то время как стандартный GARCH является лучшей моделью для USD/CZK.

Таблица 3: Качество подгонки

		GARCH	Sk-GARCH	GJR	Sk-GJR
CZK:	AIC	2,128	2,129	2,129	2,129
	BIC	2,139	2,141	2,141	2,144
HUF:	AIC	2,450	2,443	2,367	2,364
	BIC	2,460	2,456	2,382	2,380
PLN:	AIC	2,249	2,247	2,247	2,244
	BIC	2,262	2,261	2,261	2,260

Замечания: Показаны значения информационных критериев Акаике и Шварца. Для USD/CZK условное среднее не включает в себя AR(1) компоненту. Жирным выделены минимальные значения критериев.

Статистическая значимость скошенности и эффекта рычага проверяется с помощью теста LR, что возможно из-за вложенности моделей. Тестовая статистика равняется $LR = 2(\max \mathcal{L} - \max \mathcal{L}_0)$, которая асимптотически распределена как хи-квадрат с числом степеней свободы равным количеству ограничений в гипотезе. Проверяются две гипотезы $\gamma = 0$ и $\xi = 1$ для USD/HUF и USD/PLN, и обе отвергаются на 1% уровне. Поскольку USD/CZK демонстрирует большую симметрию, проверяется совместная гипотеза $\gamma = 0$ & $\xi = 1$ для USD/CZK. Это необходимо сделать, ведь из того, что мы не можем отвергнуть обе гипотезы $\gamma = 0$ и $\xi = 1$, не следует, что мы не сможем отвергнуть более сильную гипотезу $\gamma = 0$ & $\xi = 1$. Таблица 4 показывает, что нет оснований отвергать нулевую скошенность и эффект рычага для CZK.

Таблица 4: Тест отношения правдоподобия

	CZK	HUF		PLN	
H_0	$\gamma = 0$ & $\xi = 1$	$\gamma = 0$	$\xi = 1$	$\gamma = 0$	$\xi = 1$
LR статистика	1,318	9,776	11,362	10,756	10,387
p-значение	0,5173	0,0018	0,0007	0,0010	0,0013

Замечания: Для гипотез $\gamma = 0$ и $\xi = 1$ мы сравниваем LR статистику с квантилями $\chi^2(1)$ распределения. Для $\gamma = 0$ & $\xi = 1$ мы сравниваем LR с квантилями $\chi^2(2)$.

В таблице 5 представлены оцененные параметры для предложенных моделей, а также некоторые диагностические значения. Нетрудно видеть, что скошенная модель GJR обеспечивает очень хорошую подгонку для USD/HUF и довольно хорошую для USD/PLN. Кроме того, модель GJR смогла объяснить почти всю гетероскедастичность в рядах. Малая значимость ω , а также непрохождение теста Льюнга-Бокса говорят, что GARCH плохо подгоняет USD/CZK. Похоже, что волатильность сохраняется еще дольше, поскольку $\hat{\alpha} + \hat{\beta} \approx 1$, однако модель IGARCH не дает улучшения качества подгонки (результаты не приводятся). В дополнение, я диагностирую качество подгонки тестом Колмогорова-Смирнова. С помощью оцененной функции плотности я перевожу стандартизированные ошибки в $u_i \in [0, 1]$ и вычисляю $D = \sup |\hat{G}(u) - u|$, которая имеет асимптотическое распределение Колмогорова.

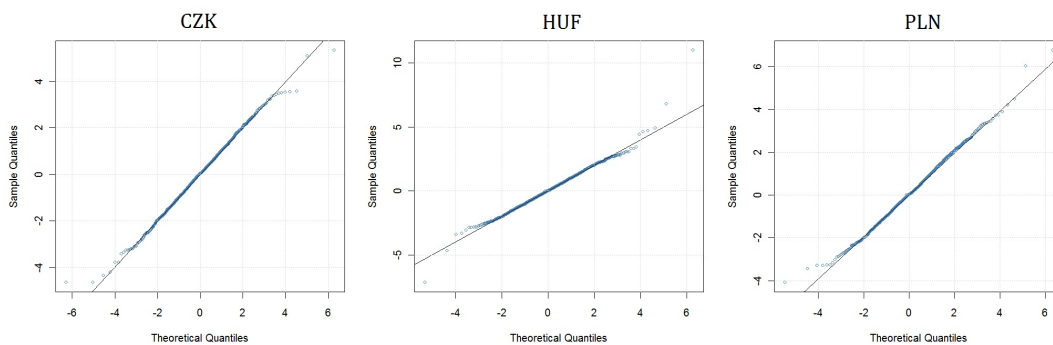
На графике 3 изображены QQ-диаграммы, которые демонстрируют разумное качество подгонки данных.

Таблица 5: Оценки параметров и диагностические значения

Parameter	CZK		HUF		PLN	
μ	-0,029*	(0,011)	-0,011	(0,014)	-0,028*	(0,013)
φ			0,075***	(0,018)	0,060**	(0,019)
ω	0,0017†	(0,0010)	0,0042*	(0,0017)	0,0152***	(0,0042)
α	0,038***	(0,007)	0,019**	(0,006)	0,072***	(0,014)
γ			-0,405*	(0,193)	-0,214**	(0,076)
β	0,962***	(0,006)	0,971***	(0,006)	0,902***	(0,016)
ξ			1,093***	(0,029)	1,083***	(0,027)
κ	5,47***	(0,57)	6,22***	(0,70)	6,02***	(0,70)
LogLik	-3047,8		-3382,9		-3210,8	
$Q_z(50)$	57,17	[0,1053]	52,13	[0,1359]	73,77**	[0,0018]
$Q_{zz}(50)$	86,23***	[0,0002]	14,69	[1]	34,92	[0,7723]
D	0,0167	[0,3983]	0,0089	[,09775]	0,0186	[0,2747]

Замечания: Стандартные ошибки приведены в круглых скобках, а р-значения в квадратных. Уровни значимости: 0,1% = ***, 1% = **, 5% = *, 10% = †. LogLik соответствует логарифму функции правдоподобия.

График 3: QQ-диаграммы стандартизированных остатков



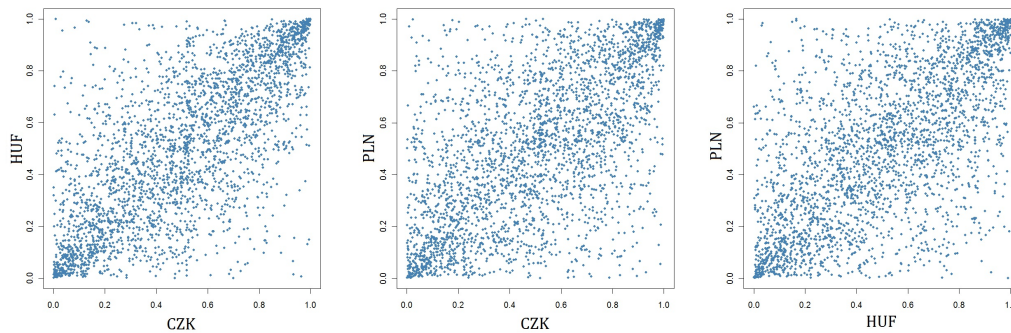
7 Моделирование копул

В данном разделе мы будем моделировать зависимость между валютами. Модели предполагаются одинаковыми на каждом шаге, а именно, копулы Стьюдента с эволюцией параметра зависимости как в Tse & Tsui (2002). Сначала оцениваются копулы для пары валют. Затем, используя вышеизложенную технику, я оцениваю копулу, которая соединит две попарные. Как было объяснено в разделе 3, существует несколько различных выходящих моделей для одних и тех же данных. В нашем случае ($d = 3$) существуют 3 модели: 1–2–3 12–23 13|2, 2–3–1 23–13 12|3, и 3–1–2 13–12 23|1. Здесь 1 соответствует CZK, 2 – HUF и 3 – PLN. Мы выберем лучшую модель из этих трех и протестируем качество подгонки. Для этого мы будем использовать методологию Diebold, Gunther & Tay (1998).

На графике 4 изображены остатки, преобразованные с помощью оцененных маргинальных распределений (маргинальное интегральное вероятностное преобразование). На графике не видно значительной асимметрии.

Я буду использовать копулу Стьюдента для моделирования, поскольку в нее легко зашить непостоянный по времени параметр зависимости, который можно интерпретировать

График 4: диаграммы трансформированных остатков



как корреляцию. Двумерная копула Стьюдента задается формулой

$$C(u, v | \theta, \nu) = T_{\theta, \nu}^{-1}(T_{\nu}^{-1}(u), T_{\nu}^{-1}(v)),$$

где $T_{\theta, \nu}$ является двумерной функцией распределения Стьюдента, T_{ν} является одномерной функцией распределения Стьюдента, $-1 < \theta < 1$, ν – число степеней свободы.

Параметр зависимости θ подчиняется следующему уравнению:

$$\theta_t = (1 - \varphi - \psi)\bar{\theta} + \varphi\tau_{t-1} + \psi\theta_{t-1}.$$

Здесь $\varphi, \psi \geq 0$, $\varphi + \psi \leq 1$, скользящая корреляция τ_{t-1} задается формулой

$$\tau_{t-1} = \frac{\sum_{s=1}^m u_{t-s}v_{t-s}}{(\sum_{s=1}^m u_{t-s}^2 \sum_{s=1}^m v_{t-s}^2)^{1/2}}.$$

Копула Стьюдента с описанной структурой зависимости подгоняется под данные методом из раздела 4. Результаты оценивания dSML представлены в таблице 6.

Таблица 6: Оценка попарных копул

	θ	φ	ψ	ν	Log Lik
12	0,78	0,60	0,01	5,47	1282,4
13	0,68	0,35	0,39	6,15	846,4
23	0,71	0,41	0,31	5,85	920,9
12 3	0,59	0,56	0,09	6,77	570,9
13 2	0,31	0,31	0,03	9,76	152,1
23 1	0,71	0,71	0,04	6,16	984,1

Используя аддитивность лог-правдоподобия, мы можем выбрать в качестве лучшей модель с максимальным лог-правдоподобием. Это модель 3–1–2 13–12 23|1, как видно из таблицы 7.

Таблица 7: Выбор выющейся модели

	LogLik
12 23 13 2	2355,4
23 13 12 3	2338,2
13 12 23 1	3112,9

Я провожу тесты Колмогорова–Смирнова для $u_{2|1}, u_{3|1}, u_{2|13}, u_{3|12}$ для проверки качества подгонки данной модели. Результаты являются положительными, как видно из таблицы 8.

Таблица 8: Тесты Колмогорова–Смирнова

	D	р-значение
$u_{2 1}$	0,0158	0,4713
$u_{3 1}$	0,0209	0,1614
$u_{2 13}$	0,0197	0,2157
$u_{3 12}$	0,0178	0,3214

8 Выводы

В данной работе предложен d -шаговый метод максимального правдоподобия (dSML) для оценивания условных выходящих копул. Доказано, что dSML дает состоятельную и асимптотически нормальную оценку. Кроме того, он предоставляет состоятельную оценку матрицы ковариации. Этот метод вычислительно эффективнее, чем стандартное оценивание методом максимального правдоподобия, и, таким образом, позволяет рассматривать более сложные модели (например, с непостоянными по времени параметрами зависимости).

Из исследования видно, что разные выходящие модели приводят к подгонкам различного качества. Было бы интересно придумать алгоритм, который бы находил оптимальную выходящую модель за разумное время.

Благодарности

Я признателен моему научному руководителю Станиславу Анатольеву за введение в круг вопросов и конструктивные советы. Я благодарен Сергею Воронцову, Ренату Хабибуллину, Игорю Хейфецу и участникам семинара про прогнозированию временных рядов за полезные комментарии и за вдохновляющую атмосферу.

Приложение: асимптотические свойства dSML

В этом приложении я опираюсь на теорию 2SML оценки из White (1994) для того, чтобы доказать асимптотическую нормальность оценки dSML. Вначале наложим ограничения на лог-функции правдоподобия ℓ_ν^t , \mathcal{L}_ν , функции вклада $s_\nu^t = \nabla_\nu \ell_\nu^t$, $\mathcal{S}_\nu = \sum_{t=1}^T s_\nu^t$, и гессианы $\nabla_\mu^T s_\nu^t$; градиент рассматривается как вектор-столбец. Пусть θ будет истинным параметром, \hat{f} обозначает $f(\hat{\theta})$, а \dot{f} обозначает $f(\dot{\theta})$ для любой функции $f(\theta)$. В последующем я буду опускать индекс T из $\mathcal{L}^T, \mathcal{S}^T, A^T, B^T$ для краткости. Данные предположения должны выполняться для всех $\nu, \mu \in \text{Vert } \mathcal{F}$, $\deg \nu = s$, $t = 1, \dots, T$, и произвольного натурального T .

- A1.
 - Θ_ν компактное подмножество Евклидова пространства
 - c_ν^t дважды непрерывно дифференцируема на Θ_ν почти наверное
- A2.
 - $\mathbb{E} \mathcal{L}_\nu(\dot{\theta}_{[s-1]}, \theta_\nu)$ максимизируется в единственной внутренней точке из Θ_ν
- A3.
 - $\mathbb{E} \mathcal{L}_\nu$ конечно и непрерывно на Θ_ν равномерно по T
 - $\{\ell_\nu^t\}$ подчиняется слабому равномерному закону больших чисел (РЗБЧ) на Θ_ν
 - $\mathbb{E} \mathcal{L}_\nu / T = O(1)$ при $T \rightarrow \infty$ равномерно по Θ_ν
- A4.
 - $\mathbb{E} \mathcal{S}_\nu$ конечно и непрерывно на Θ_ν равномерно по T
 - $\{s_\nu^t\}$ подчиняется слабому РЗБЧ на Θ_ν

- A5.**
- $\mathbb{E}\nabla_{\mu}^{\top}\mathcal{S}_{\nu}$ конечно и непрерывно на Θ_{ν} равномерно по T
 - $\{\nabla_{\mu}^{\top}s_{\nu}^t\}$ подчиняется слабому РЗБЧ на Θ_{ν}
 - $\mathbb{E}\nabla_{\nu}^{\top}\mathring{\mathcal{S}}_{\nu}/T = O(1)$ при $T \rightarrow \infty$ и является неотрицательно определенной

Теорема 7 (Состоятельность dSML) *Предположим, A1–A5 верны. Тогда $\hat{\theta} \xrightarrow{P} \mathring{\theta}$.*

Доказательство. Воспользуемся методом математической индукции, показав что $\hat{\theta}_{\nu} \xrightarrow{P} \mathring{\theta}_{\nu}$ для всех ν . Первый шаг – это состоятельность оценок максимального правдоподобия параметров маргинальных распределений $\hat{\theta}_i$. Оно следует из теоремы 3.13 в White (1994). Затем, предполагая, что $\hat{\theta}_{[s-1]}$ состоятельно оценивает $\mathring{\theta}_{[s-1]}$, мы применяем теорему 3.10 из White (1994) к 2SML оценке $\hat{\theta}_{\nu} = \operatorname{argmax}_{\theta_{\nu}} \mathcal{L}_{\nu}(\hat{\theta}_{[s-1]}, \theta_{\nu})$. ■

Следующее предположение необходимо для асимптотической нормальности $\hat{\theta}$.

- A6.**
- $\{s^t/\sqrt{T}\}$ подчиняется треугольной ЦПТ с матрицей ковариации \mathring{B}
 - $\mathring{B} = O(1)$ при $T \rightarrow \infty$

Матрица \mathring{B} положительно определена и имеет блочную структуру $(\mathring{B}_{\mu\nu} \mid \mu, \nu \in \operatorname{Vert} \mathcal{F})$, где $B_{\mu\nu} = \mathbb{E} \left[\sum_{t=1}^T s_{\mu}^t (s_{\nu}^t)^{\top} / T \right]$.

Теорема 8 (Асимптотическая нормальность dSML) *Пусть выполняются A1–A6. Тогда*

$$\mathring{B}^{-1/2} \mathring{A} \sqrt{T} (\hat{\theta} - \mathring{\theta}) \xrightarrow{d} \mathcal{N}(0, I),$$

где \mathring{A} следующая сложная блочная структура:

$$A = \begin{pmatrix} A_{(11)} & 0 & \dots & 0 \\ A_{(21)} & A_{(22)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{(s1)} & A_{(s2)} & \dots & A_{(ss)} \end{pmatrix}$$

и $A_{(qp)} = (\mathbb{E}\nabla_{\mu}^{\top}\mathcal{S}_{\nu}/T \mid \mu \in \operatorname{Vert}_p \mathcal{F}, \nu \in \operatorname{Vert}_q \mathcal{F})$.

Доказательство. Запишем разложение Тейлора функции вклада $\hat{\mathcal{S}}/T$ в точке $\mathring{\theta}$. Вспомним, что по построению $\hat{\theta}$ мы имеем $\hat{\mathcal{S}} = 0$. Таким образом,

$$\hat{\mathcal{S}}/T = \mathring{\mathcal{S}}/T + \bar{A}(\hat{\theta} - \mathring{\theta}) = 0,$$

где \bar{A} – это матрица A в точке $\bar{\theta}$, лежащей покомпонентно между $\hat{\theta}$ и $\mathring{\theta}$. Из этого уравнения следует, что

$$\bar{A} \sqrt{T} (\hat{\theta} - \mathring{\theta}) = -\mathring{\mathcal{S}}/\sqrt{T}.$$

Поскольку A непрерывно и $\bar{\theta} \xrightarrow{P} \mathring{\theta}$ при $T \rightarrow \infty$, получаем

$$\mathring{A} \sqrt{T} (\hat{\theta} - \mathring{\theta}) = -\mathring{\mathcal{S}}/\sqrt{T} + o_p(1).$$

Умножая на $\mathring{B}^{-1/2}$, мы получаем

$$\mathring{B}^{-1/2} \mathring{A} \sqrt{T} (\hat{\theta} - \mathring{\theta}) = -\mathring{B}^{-1/2} \mathring{\mathcal{S}}/\sqrt{T} + o_p(1) \xrightarrow{d} \mathcal{N}(0, I).$$

Предельное распределение следует из предположения A6. ■

Последнее предположение необходимо для получения состоятельных оценок асимптотической ковариационной матрицы.

- A7.**
- B конечна и непрерывна на Θ равномерно по T
 - $\{s^t(s^t)^\top\}$ удовлетворяет слабому РЗБЧ на Θ

Теорема 9 (Состоятельные оценки ковариационной матрицы) *Предположим, что A1–A7 выполнены. Тогда*

$$\hat{A}^{-1}\hat{B}\hat{A}^{-1} - \mathring{A}^{-1}\mathring{B}\mathring{A}^{-1} \xrightarrow{P} 0,$$

где $\hat{A}_{\mu\nu} = \nabla_{\mu}^\top \hat{S}_{\nu}$ и $\hat{B}_{\mu\nu} = \sum_{t=1}^T \hat{s}_{\mu}^t (\hat{s}_{\nu}^t)^\top / T$.

Доказательство. Состоятельность \hat{A} как оценки \mathring{A} следует из доказательства теоремы 8.24 в White (1994), в то время как состоятельность \hat{B} для \mathring{B} следует из доказательства теоремы 8.26. Результат следует из теоремы Манна–Вальда. ■

Список литературы

- Aas, K., C. Czado, A. Frigressi & H. Bakken (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* 44, 182–198.
- Bedford, T. & R. M. Cooke (2002). Vines – a new graphical model for dependent random variables. *Annals of Statistics* 30, 1031–1068.
- Diebold, F.X., T.A. Gunther, & A.S. Tay (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review* 39, 863–883.
- Fernandez, C. & M.F.J. Steel (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association* 93, 359–371.
- Glosten, L.R., R. Jagannathan & D.E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779–1801.
- Griebeler, M.C. (2010). Models for forecasting exchange rate volatility: a comparison between developed and emerging countries. Working Paper, IMPA.
- Hsieh, D.A. (1989). Modeling heteroskedasticity in daily foreign-exchange rates. *Journal of Business & Economic Statistics* 7, 307–317.
- Joe, H. (1996). Families of m -variate distributions given margins and $m(m-1)/2$ bivariate dependence parameters. Глава в *Distributions with Fixed Margins and Related Topics* под редакцией L. Rüschendorf, B. Schweizer & M.D. Taylor.
- Jondeau, E. & M. Rockinger (2006). The copula-GARCH model of conditional dependencies: an international stock market application. *Journal of International Money and Finance* 25, 827–853.
- Longmore, R. & W. Robinson (2004). Modelling and forecasting exchange rate dynamics in Jamaica: an application of asymmetric volatility models. Working Paper WP2004/03, Bank of Jamaica.
- Patton, A.J. (2002). *Applications of Copula Theory in Financial Econometrics*. PhD Thesis, University of California–San Diego.
- Patton, A.J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics* 2, 130–168.
- Patton, A.J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review* 47, 527–526.
- Rodriguez, J.C. (2007). Measuring financial contagion: A copula approach. *Journal of Empirical Finance* 14, 401–423.
- Sandoval, J. (2006). Do asymmetric GARCH models fit better exchange rate volatilities on emerging markets? Working Paper, Universidad Externado de Colombia. *Odeon* 3, 99–116.
- Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* 8, 229–231.
- Trivedi, P.K. & D.M. Zimmer (2005). Copula modeling: an introduction for practitioners. *Foundations and Trends in Econometrics* 1, 1–111.

- Tse, Y.K. (1998). The conditional heteroskedasticity of the yen-dollar exchange rate. *Journal of Applied Econometrics* 13, 49–55.
- Tse, Y.K. & A.K.C. Tsui (1998). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics* 20, 351–362.
- White, H. (1994). *Estimation, Inference and Specification Analysis*. Cambridge University Press.

Time varying vine copulas for multivariate returns

Oleg Groshev

Barclays Capital, Moscow, Russia

We analyze the multivariate distribution of financial returns using time varying conditional vine copulas. We present the d-Stage Maximum Likelihood (dSML) estimator which is shown to be not only consistent and asymptotically normal, but also more computationally attractive than the standard ML or Patton's 2SML. Using dSML, we fit vine copulas to returns of a portfolio on emerging market currencies.

Keywords: multidimensional time series, vine copulas, dSML estimator, computational efficiency

JEL Classification: C13, C58, G15

Динамика условных моментов высоких порядков и прогнозирование стоимостной меры риска*

Григорий Франгуриди†

Российская экономическая школа, Москва, Россия

В настоящей работе исследуется возможность улучшения качества прогнозов стоимостной меры риска VaR с помощью моделирования динамики скошенности и эксцесса условных распределений финансовых доходностей. Для однодневных прогнозов VaR, вычисленных для пяти ликвидных акций индекса S&P500 из различных секторов, мы сравниваем стандартную модель GARCH со скошенным студентовским распределением ошибок с набором моделей с динамическими моментами высоких порядков, таких как модели типа ARCD с нормальным обратным гауссовским и обобщённым скошенным студентовским распределением ошибок. В качестве тестов на предсказательную способность используются как скоринговые тесты для левых хвостов предсказанных плотностей, так и LR-тесты корректности прогнозов VaR. Мы предлагаем модель со скошенным обобщённым распределением ошибок с динамическими скошенностью и куртозисом, которая даёт по крайней мере не менее качественные прогнозы VaR, чем аналогичные модели, и в то же время является вычислительно менее сложной.

Ключевые слова: стоимостная мера риска, условное распределение, скошенность, куртозис, финансовые доходности

Классификация JEL: C22, C53, C58

1 Введение

Стоимостная мера риска value-at-risk (VaR) является в настоящее время одним из самых распространённых показателей риска в финансовой индустрии. По определению, VaR доверительной вероятности p (близкой к 1) портфеля ценных бумаг Π_{t+1} , оцениваемый на день $t+1$, есть квантиль уровня $1-p$ распределения логарифмической доходности $r_{t+1}^{\Pi} = \log(\Pi_{t+1}/\Pi_t)$. Таким образом, VaR можно понимать как значение, меньше которого завтрашняя доходность не упадёт с вероятностью p . VaR имеет несколько важных преимуществ над другими мерами риска. Во-первых, совокупные риски портфеля представляются всего одним числом с прозрачной интерпретацией. Во-вторых, VaR можно применять к портфелям, состоящим из ценных бумаг любого типа, будь то обыкновенные акции или экзотические опционы, что позволяет сравнивать риски портфелей разной структуры. В-третьих, расчёт VaR реализован почти во всех пакетах для финансовых вычислений, что позволяет использовать его финансистам, не владеющим программированием. Наконец, расчёт VaR является обязательным требованием Базельского комитета по банковскому надзору в рамках вычислений показателей достаточности капитала коммерческих банков. Так как во многих развитых странах банковские регуляторы требуют выполнения Базельских соглашений, подотчётные банки не могут обойтись без расчёта VaR. Более того, многие банки ежедневно рассчитывают VaR портфелей кредитов и ценных бумаг для внутренней отчётности по контролю за рисками.

*Цитировать как: Франгуриди, Григорий (2014) «Динамика условных моментов высоких порядков и прогнозирование стоимостной меры риска», Квантиль, №12, стр. 69–82. Citation: Franguridi, Grigory (2014) “Higher order conditional moment dynamics and forecasting value-at-risk,” Quantile, No.12, pp. 69–82.

†Адрес: 117418, г. Москва, Нахимовский проспект, 47, офис 1721(5). Электронная почта: gfranguridi@nes.ru

Однако нельзя не отметить и некоторые недостатки VaR. Очевидно, что сведение рисков портфеля к одному числу сильно *упрощает* истинную структуру рисков. VaR может также сильно *искажать* представление о «настоящих» рисках: например, два портфеля могут иметь одинаковый VaR, даже если при условии пробы VaR по одному из этих портфелей потери будут заведомо больше, чем по другому. Иными словами, VaR не может отразить разницу в левых хвостах распределений доходностей, находящихся за $(1 - p)$ -квантилем. Более того, VaR не всегда поощряет диверсификацию портфеля, т.е. не является *субаддитивным*: не для всех случайных величин (доходностей) r_1 и r_2 верно $\text{VaR}(r_1 + r_2) \leq \text{VaR}(r_1) + \text{VaR}(r_2)$ (см. контрпример в книге Föllmer & Schied, 2004).¹

В банковской индустрии обычно интересуются *условным* дневным VaR, т.е. квантилем завтрашнего распределения финансовой доходности с учётом наблюдаемой истории доходностей. Как правило, для прогнозирования условного VaR специфицируется и оценивается параметрическая модель доходностей, после чего прогноз условного VaR вычисляется как квантиль предсказанной условной плотности. Одними из наиболее распространённых моделей здесь являются разновидности GARCH-моделей с распределениями ошибок, характеризующимися тяжёлыми хвостами и скошенностью. В настоящей работе мы показываем, что учёт *динамики* условных моментов высоких порядков позволяет улучшить качество прогнозирования VaR (вне обучающей выборки) по сравнению с моделями с фиксированными скошенностью и куртозисом как в смысле поведения процесса пробоев VaR, так и в смысле корректности предсказания левых хвостов условных распределений доходностей. Мы также строим новую модель SGE-ACD со скошенным обобщённым распределением ошибок Theodossiou (2000) и авторегрессионной динамикой на (параметры, отвечающие за) условные моменты второго, третьего и четвёртого порядков, которая по качеству прогнозов VaR и левых хвостов условных распределений неотличима от более вычислительно сложных моделей с динамикой высоких моментов, таких как SGT-GARCH с обобщённым скошенным распределением Стюдента из работы Bali, Mo & Tang (2008) и NIG-ACD с нормальным обратным гауссовским распределением из работы Wilhelmsson (2009).

2 Моделирование стоимостной меры риска

2.1 Модели с постоянными условными моментами высоких порядков

Пусть дана выборка доходностей $\{r_t\}_{t=1}^T$, а задача состоит в оценивании однопериодного $\text{VaR}_{T+1|T}^p$ — истинного $(1 - p)$ -квантиля распределения доходности r_{T+1} , условно на предыстории r_T, r_{T-1}, \dots, r_1 .

Наиболее примитивный способ оценивания (*historically simulated, HS-VaR*) состоит в вычислении эмпирического $(1 - p)$ -квантиля по подвыборке $\{r_t\}_{t=T-L+1}^T$ длины $L \leq T$. Как отмечают Manganelli & Engle (2001), такой подход является логически противоречивым. Если мы предполагаем, что процесс $\{r_t\}$ эргодичен и стационарен, то $\text{HS-VaR}_{t+1|t}^p$ сходится к $(1 - p)$ -квантилю *безусловного* распределения доходностей при неограниченном увеличении длины окна L , и тогда разумно положить $L = T$. Однако, на практике обычно требуется оценка *условного* квантиля, и значит длина окна не может быть слишком большой: в противном случае оценка будет испорчена слишком давними наблюдениями. Даже если требуется оценить безусловный квантиль, качество оценки $\text{HS-VaR}_{t+1|t}^p$ в конечных выборках оставляет желать лучшего, особенно если для процесса $\{r_t\}$ характерна сильная кластеризация волатильности, что для финансовых доходностей верно почти всегда.

Для оценивания условного квантиля существует множество подходов, которые можно грубо разделить на две категории: моделирование условного распределения доходностей (в том

¹В работе Danielsson, de Haan, Peng & Vries (2001) показано, впрочем, что в большинстве практически релевантных случаев VaR является субаддитивной мерой риска.

числе полу- и непараметрические методы) и непосредственное моделирование условного квантиля. В настоящей работе мы сконцентрируем внимание только на моделях первого типа.

Положим $\mathcal{F}_{t-1} = \{r_{t-1}, r_{t-2}, \dots\}$, $\mu_t = \mathbb{E}[r_t | \mathcal{F}_{t-1}]$, $\sigma_t^2 = \mathbb{V}[r_t | \mathcal{F}_{t-1}]$ и представим динамику r_t в стандартном виде

$$\begin{aligned} r_t &= \mu_t + \eta_t, \\ \eta_t &= \sigma_t \varepsilon_t, \quad \{\varepsilon_t\} \sim iid(0, 1), \\ \varepsilon_t | \mathcal{F}_{t-1} &\sim F(\cdot | \theta). \end{aligned}$$

Описываемые ниже модели отличаются только в динамике условной дисперсии σ_t^2 и предположениях об условном распределении F стандартизированной ошибки ε_t . Условное среднее μ_t как правило незначимо отличается от нуля, поэтому моделируется редко, так что будем считать, что $\mu_t \equiv 0$. После получения оценок условной дисперсии $\hat{\sigma}_{t+1}^2$ и условной функции распределения $F(\cdot | \hat{\theta})$ оценка VaR вычисляется просто:

$$\widehat{\text{VaR}}_{t+1|t}^p = \hat{\sigma}_{t+1} F^{-1}(1 - p | \hat{\theta}).$$

Простейшая модель условной дисперсии — так называемая методология RiskMetrics, разработанная инвестиционным банком JPMorgan, см. Zumbach (2007). Условная дисперсия имеет динамику

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2.$$

Легко видеть, что условная дисперсия представляется в виде «экспоненциально сглаженной» суммы квадратов доходностей прошедших периодов:

$$\sigma_t^2 = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^i r_{t-i}^2,$$

а значит не обладает желательным свойством возвращения к среднему. Более того, структура волатильности в модели RiskMetrics линейная: условная дисперсия $\sigma_{t+k|t}^2$ доходности за k дней равна $k \sigma_{t+1|t}^2$.

Модель RiskMetrics является ограничением более общей модели обобщённой авторегрессионной условной гетероскедастичности GARCH(p, q), предложенной в работе Bollerslev (1986):

$$\sigma_t^2 = \omega + \alpha_1 \eta_{t-1}^2 + \dots + \alpha_q \eta_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2,$$

где p и q — натуральные числа. В отличие от RiskMetrics, модель GARCH обладает свойством возвращения к среднему:

$$\sigma_t^2 - \sigma^2 = \alpha(r_{t-1}^2 - \sigma^2) + \beta(\sigma_{t-1}^2 - \sigma^2),$$

где величина $\sigma = \omega / (1 - \alpha - \beta)$ интерпретируется как *долгосрочная волатильность*.

При моделировании распределения стандартизированной ошибки ε_t необходимо найти баланс между простотой вида распределения (а значит, и соответствующей функции правдоподобия) и гибкостью спецификации. Например, для нормального распределения $\mathcal{N}(0, 1)$ известен явный вид функции правдоподобия; более того, теория квазиравдоподобия для GARCH-моделей, развитая в работе Bollerslev & Wooldridge (1992), показывает, что даже в случае неверной спецификации нормального распределения оценивание может давать состоятельные оценки. С другой стороны, очевидным недостатком (стандартного) гауссовского

распределения является отсутствие параметров и неспособность уловить такой важный стилизованный факты о распределении доходностей, как скошенность².

В следующем параграфе будут описаны модели с динамическими моментами высоких порядков; параметрические распределения стандартизованных ошибок в этих моделях находят применение и в более простом случае постоянных моментов.

2.2 Модели с динамическими условными моментами высоких порядков

Многие работы, например, Harvey & Siddique (1999), Jondeau & Rockinger (2003) и Bali, Mo & Tang (2008) показывают, что условные скошенности и эксцессы финансовых доходностей не являются постоянными во времени. В настоящем параграфе делается обзор некоторых моделей динамики моментов высоких порядков.

Основополагающей работой, посвящённой моделированию динамики моментов высоких порядков, является статья Hansen (1994). В ней предложена общая методология моделирования условных плотностей, имеющих авторегрессионную структуру (*autoregressive conditional density, A/R/CD*), и получено обобщение распределение Стьюдента на случай нетривиальных скошенности и куртозиса. Плотность SGT-распределения (*skewed generalized t*) имеет вид

$$f(z|\lambda, \eta, \kappa) = C \left(1 + \frac{|z + \delta|^\kappa}{(\eta + 1)/\kappa (1 + \operatorname{sgn}(z + \delta)\lambda)^\kappa \theta^\kappa} \right)^{\frac{\eta+1}{\kappa}},$$

где

$$\begin{aligned} C &= 0.5\kappa \left(\frac{\eta + 1}{\kappa} \right)^{-\frac{1}{\kappa}} B \left(\frac{\eta}{\kappa}, \frac{1}{\kappa} \right)^{-1} \theta^{-1}, \\ \theta &= \frac{1}{\sqrt{g - \rho^2}}, \\ \rho &= 2\lambda B \left(\frac{\eta}{\kappa}, \frac{1}{\kappa} \right)^{-1} \left(\frac{\eta + 1}{\kappa} \right)^{\frac{1}{\kappa}} B \left(\frac{\eta - 1}{\kappa}, \frac{2}{\kappa} \right), \\ g &= (1 + 3\lambda^2) B \left(\frac{\eta}{\kappa}, \frac{1}{\kappa} \right)^{-1} \left(\frac{\eta + 1}{\kappa} \right)^{\frac{2}{\kappa}} B \left(\frac{\eta - 2}{\kappa}, \frac{3}{\kappa} \right), \\ \delta &= \rho\theta. \end{aligned}$$

Здесь параметр $\lambda \in (-1, 1)$ отвечает за скошенность, $\eta > 2$ — за толщину хвостов, $\kappa > 0$ — за куртозис, а за $B(\cdot)$ обозначена бета-функция. Чтобы избежать необходимости решать задачу максимального правдоподобия с ограничениями, Хансен предлагает использовать подходящие логистические преобразования параметров. Автор накладывает простую ARCH-динамику на параметры и показывает, что полученная модель показывает отличные результаты для дневных данных по обменному курсу USD/CHF. Jondeau & Rockinger (2003) приводят список возможных спецификаций динамики параметров в таких моделях, в том числе спецификации типа асимметричной GARCH-модели и спецификации для динамики третьего и четвёртого начальных моментов непосредственно.

Аналогичная методология применяется в работе Bali, Mo & Tang (2008). Авторы используют SGT-распределение и результаты Theodossiou (1998) для моделирования стандартизованных доходностей с динамикой скошенности, куртозиса и толщины хвостов. Условное среднее моделируется простой авторегрессией первого порядка AR(1), а условная дисперсия — с помощью различных GARCH-моделей. Параметры $\lambda_t, \eta_t, \kappa_t$ записываются в виде

²Здесь необходимо отметить, что речь идёт про *безусловную* скошенность, которая вряд ли возможна без *условной* скошенности. Что касается куртозиса, то можно показать, что даже в гауссовской модели GARCH безусловное распределение доходностей имеет тяжёлые хвосты; в этом отношении предположение о нормальности не является столь ограничительным.

$\lambda_t = -1 + 2/(1 + \exp(\tilde{\lambda}_t))$, $\eta_t = 2 + \exp(\tilde{\eta}_t)$ и $\kappa_t = \exp(\tilde{\kappa}_t)$, после чего на неограниченные вспомогательные параметры $\tilde{\lambda}_t$, $\tilde{\eta}_t$ и $\tilde{\kappa}_t$ накладывается обычная ARCH-динамика

$$\tilde{\lambda}_t = \lambda^0 + \lambda^1 z_{t-1} + \lambda^2 \tilde{\lambda}_{t-1}, \quad (1)$$

$$\tilde{\eta}_t = \eta^0 + \eta^1 z_{t-1} + \eta^2 \tilde{\eta}_{t-1}, \quad (2)$$

$$\tilde{\kappa}_t = \kappa^0 + \kappa^1 z_{t-1} + \kappa^2 \tilde{\kappa}_{t-1}, \quad (3)$$

Показано, что такая SGT-GARCH модель генерирует очень точные и робастные оценки VaR, причём гипотеза о постоянстве параметров во времени уверенно отвергается.

В работе Wilhelmsson (2009) предлагается использовать нормальное обратное гауссовское (*Normal Inverse Gaussian, NIG*) распределение для моделирования стандартизованных доходностей. NIG-распределение обладает двумя основными преимуществами. Во-первых, как показано в работе Jondeau & Rockinger (2003), по сравнению с SGT-распределением Hansen (1994) NIG-распределение позволяет моделировать более широкий класс комбинаций скошенности и куртозиса. Во-вторых, использование NIG-распределения имеет теоретическое обоснование: в модели Barndorff-Nielsen (1997) со скрытым процессом появления информации доходности имеют условное NIG-распределение. Плотность NIG-распределения имеет вид

$$f(x; \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} - \beta\gamma\right) q\left(\frac{x - \mu}{\delta}\right)^{-1} K_1\left(\delta\alpha q\left(\frac{x - \mu}{\delta}\right)\right) \exp(\beta x),$$

где $q(z) = \sqrt{1 + z^2}$, а $K_1(\cdot)$ — модифицированная функция Бесселя порядка 3 и индекса 1. Параметр α отвечает за куртозис, β — за скошенность, μ — за центр плотности на действительной оси, а σ — масштабирующий параметр. NIG-распределение замкнуто относительно взятия свёртки, а значит, если однодневные доходности NIG-распределены, то многодневные доходности также получаются NIG-распределёнными. Однако дисперсия NIG-распределения зависит от нескольких параметров, что неудобно для моделирования динамики волатильности. Чтобы избавиться от этой зависимости, Jensen & Lunde (2001) предлагают перепараметризацию $\bar{\alpha} = \alpha\delta$, $\bar{\beta} = \beta\delta$.

Приведём далее полную спецификацию модели NIG-ACD. Положим $\bar{\gamma}_t = \sqrt{\bar{\alpha}_t^2 + \bar{\beta}_t^2}$ и $\rho_t = \beta_t/\alpha_t \in [0, 1)$. Тогда доходности $\{r_t\}$ имеют вид

$$r_t = \mu + \bar{\gamma}_t^{1/2} \delta_t \rho_t + \varepsilon_t, \quad (4)$$

$$\varepsilon_t = \delta_t \eta_t, \quad \eta_t \sim NIG\left(\bar{\alpha}_t, \bar{\beta}_t, -\bar{\gamma}_t^{1/2} \rho_t, \bar{\gamma}_t^{3/2} / \bar{\alpha}_t\right), \quad (5)$$

$$\delta_t^v = c + b\delta_{t-1}^v + a(|\varepsilon_{t-1}| - \tau\varepsilon_{t-1})^v, \quad (6)$$

$$\bar{\gamma}_t = \exp\left(\lambda_0 + \lambda_1 \varepsilon_{t-1}^2 + \lambda_2 \log(\bar{\gamma}_{t-1})\right), \quad (7)$$

$$\tilde{\rho}_t = \theta_0 + \theta_1 \varepsilon_{t-1}. \quad (8)$$

где $\tilde{\rho}_t = \log\left(\frac{1+\rho_t}{1-\rho_t}\right) \in \mathbb{R}$ — удобная перепараметризация параметра скошенности ρ . Условное стандартное отклонение δ_t имеет динамику типа асимметричной степенной ARCH-модели (APARCH), условная «крутизна» $\bar{\gamma}_t$ эволюционирует согласно экспоненциальной ARCH-модели, а модифицированный параметр скошенности $\tilde{\rho}_t$ имеет простую MA(1)-динамику. Сложный вид спецификации обусловлен тем, что необходимо гарантировать нулевое среднее и единичную дисперсию ошибок η_t .

Условное распределение r_{t+1} в такой модели имеет вид

$$r_{t+1} | \mathcal{F}_t \sim NIG\left(\alpha_t, \beta_t, \mu, \bar{\gamma}_t^{3/2} / \alpha_t\right).$$

Таким образом, модель легко оценивается условным методом максимального правдоподобия.

Wilhelmsson (2009) показывает, что для дневных доходностей индекса S&P500 за период с 1962 по 2005 годы модель NIG-ACD с динамическими моментами высоких порядков показывает лучшие результаты, чем стандартные GARCH-модели с нормальным и NIG распределением ошибок, модели ARCD из работы Hansen (1994), а также NIG-ACD модели с постоянными моментами высоких порядков.

Мы предлагаем использовать спецификацию для динамики параметров λ_t и κ_t из модели SGT-GARCH и спецификацию условной волатильности из модели NIG-ACD для скошенного обобщённого распределения ошибок (*skewed generalized error, SGE*), предложенного в Theodossiou (2000), и показываем далее, что новая модель по крайней мере не менее успешна в смысле качества прогнозирования VaR, одновременно являясь вычислительно более простой. Плотность SGE-распределения имеет вид:

$$f(x|\mu, \sigma, \kappa, \lambda) = \frac{C}{\sigma} \exp\left(-\frac{|x - \mu + \delta\sigma|^\kappa}{[1 - \operatorname{sgn}(x - \mu - \delta\sigma)\lambda]^\kappa \theta^\kappa \sigma^\kappa}\right),$$

где

$$\begin{aligned} C &= \frac{\kappa}{2\theta} \Gamma(1/\kappa)^{-1}, \\ \theta &= \Gamma(1/\kappa)^{1/2} \Gamma(3/\kappa)^{-1/2} S(\lambda)^{-1}, \\ \delta &= 2\lambda A S(\lambda)^{-1}, \\ S(\lambda) &= \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2}, \\ A &= \Gamma(2/\kappa) \Gamma(1/\kappa)^{-1/2} \Gamma(3/\kappa)^{-1/2}. \end{aligned}$$

Здесь параметр $\lambda \in (-1, 1)$ отвечает за скошенность и контролирует скорость убывания плотности в окрестности моды распределения, параметр $\kappa > 0$ контролирует тяжесть хвостов и высоту пика распределения, а параметры $\mu = \mathbb{E}[x]$ и $\sigma = \sqrt{\mathbb{V}[x]}$. Частными случаями SGE-распределения являются распределение Субботина ($\lambda = 0$), распределение Лапласа ($\lambda = 0$, $\kappa = 1$), нормальное распределение ($\lambda = 0$, $\kappa = 2$); при $\kappa \rightarrow \infty$ и фиксированном $\lambda = 0$ распределение становится всё более «равномерным»³.

Таким образом, в новой SGE-ACD модели параметры распределения ошибок $\mu = 0$, $\sigma = 1$, параметры λ_t и κ_t имеют динамику (1) и (3) соответственно, а в качестве модели волатильности $\delta_t = \sqrt{\mathbb{V}[r_t|\mathcal{F}_{t-1}]}$ используется APARCH (6).

3 Тестирование предсказательной способности моделей VaR

После оценивания набора моделей и вычисления прогнозов VaR по каждой модели необходимо ответить на два вопроса: (1) можно ли для конкретной модели считать, что она «успешна» в прогнозировании VaR? (2) какая из моделей «лучше» с точки зрения прогнозирования VaR?

Ответ на первый вопрос зависит от определения «успешности» модели, под которой наиболее естественно понимать корректное поведение процесса пробоев прогнозов VaR. В этом случае можно воспользоваться классическими результатами Christoffersen (1998), который предлагает трёхшаговую процедуру тестирования качества построенных доверительных интервалов⁴. Сначала тестируется корректность *безусловного накрытия*, т.е. гипотеза о том, что вычисленные прогнозы VaR действительно пробиваются с вероятностью $1 - p$. Затем тестируется отсутствие кластеризации пробоев VaR, т.е. корректность *условного накрытия*. На последнем шаге тестируется *совокупная* корректность накрытия.

³Однако предел не определён, так как равномерного распределения на \mathbb{R} не существует.

⁴Заметим, что VaR можно рассматривать как односторонний интервал доверительной вероятности p вида $(\text{VaR}, +\infty)$.

Формально, обозначим за $I_t = \mathbb{I}\left(r_t > \widehat{\text{VaR}}_{t|t-1}^p\right)$, $t = 1, \dots, T$ индикатор отсутствия пробоя прогноза VaR в момент t . Тогда при нулевой гипотезе о корректном *безусловном* накрытии $I_t \sim \mathcal{B}(p)$, а значит статистика отношения правдоподобия

$$\text{LR}_{UC} = 2 \log \left[\frac{\hat{\pi}_1^{T_1} (1 - \hat{\pi}_1)^{T-T_1}}{(1 - \alpha)^{T_1} \alpha^{T-T_1}} \right] \xrightarrow{d} \chi^2(1),$$

где T_1 — количество единиц в последовательности индикаторов и $\hat{\pi}_1 = T_1/T$.

Для тестирования отсутствия кластеризации Christoffersen (1998) предлагает тестировать гипотезу о независимости индикаторов против альтернативной гипотезы о том, что последовательность индикаторов образует марковскую цепь первого порядка. В такой постановке статистика отношения правдоподобия

$$\text{LR}_{IND} = 2 \log \left[\frac{(1 - \hat{\pi}_{01})^{T_0 - T_{01}} \hat{\pi}_{01}^{T_{01}} (1 - \hat{\pi}_{11})^{T_1 - T_{11}} \hat{\pi}_{11}^{T_{11}}}{\hat{\pi}_1^{T_1} (1 - \hat{\pi}_1)^{T - T_1}} \right] \xrightarrow{d} \chi^2(1),$$

где $T_{ij} = \#\{I_t = i, I_{t+1} = j\}$ и $\hat{\pi}_{ij} = T_{ij}/T_i$.

Наконец, в статье показывается, что при корректном *совокупном* накрытии (эквивалентно, верности обеих нулевых гипотез выше или $I_t \sim iid \mathcal{B}(p)$) статистика

$$\text{LR}_{CC} = \text{LR}_{UC} + \text{LR}_{IND} \xrightarrow{d} \chi^2(2).$$

В работе Christoffersen & Pelletier (2004) отмечается, что во многих случаях асимптотическое приближение работает плохо, и предлагается бутстрапирование статистики LR_{CC} .

Что касается ответа на второй вопрос, то в качестве критерия для сравнения (двух) моделей VaR можно использовать некоторую меру качества предсказания левых хвостов условных распределений доходностей. Ясно, что хороший прогноз левого хвоста будет давать качественную оценку VaR, однако обратное неверно, а значит, тестирование на одинаковое качество прогноза левых хвостов — несколько избыточная задача с точки зрения меры VaR. Тем не менее, для параметрических моделей такой подход является естественным и позволяет косвенно судить о том, насколько одна из моделей более подходит для прогнозирования мер риска, более чувствительных к распределению потерь, например, Expected Shortfall.

Одну из методологий сравнения качества out-of-sample прогнозов распределений предлагают Diks, Panchenko & van Dijk (2011). Авторы модифицируют существующие скоринговые методы, основанные на взвешенном правдоподобии, чтобы обойти проблему смещения стандартных тестов в сторону прогнозов, которые приписывают интересующему исследователя хвосту большую совокупную вероятность. А именно, они определяют *цензурированный скор* $S^{csl}(\hat{f}_t(\cdot); r_{t+1})$ для прогноза $\hat{f}_t(\cdot)$ истинной условной плотности $f_{t+1|t}(\cdot)$ доходности r_{t+1} :

$$S^{csl}(\hat{f}_t(\cdot); r_{t+1}) = \mathbb{I}_{A_t}(r_{t+1}) \log \hat{f}_t(r_{t+1}) + \mathbb{I}_{\mathbb{R} \setminus A_t}(r_{t+1}) \log \int_{\mathbb{R} \setminus A_t} \hat{f}_t(s) ds,$$

где $\mathbb{I}_A(\cdot)$ — индикатор множества A . Нас интересуют левые хвосты, поэтому положим $A_t := (-\infty, q)$ для некоторого малого квантиля q распределения r . Авторы показывают, что при некоторых технических предположениях цензурированный скор S^{csl} является *корректно построенным* в том смысле, что для него выполняется аналог информационного неравенства: средний цензурированный скор любого прогноза $\hat{f}_t(\cdot)$ не может быть больше среднего цензурированного скор истинной плотности $f_t(\cdot)$, т.е.

$$\mathbb{E}_t S^{csl}(\hat{f}_t(\cdot); r_{t+1}) \leq \mathbb{E}_t S^{csl}(f_t(\cdot); r_{t+1}).$$

Таблица 1: Описательные статистики для дневных доходностей с 01.01.1996 до 31.12.2010

	ВАС	F	GE	JPM	PFE
Mean	0,000613	0,000651	0,000428	0,000673	0,000421
Std	0,0317	0,0291	0,0208	0,0275	0,0191
Min	-0,2897	-0,2500	-0,1279	-0,2072	-0,1115
Max	0,3527	0,2951	0,1970	0,2610	0,1017
Skew	0,910	0,655	0,310	0,828	-0,035
Kurt	27,18	16,24	10,65	15,23	6,13
JB	92584	27848	9274	23972	1547
ACF(1)	0,00208	0,00165	-0,00930	-0,07739	0,00276
ACF(2)	0,014	0,049	-0,016	-0,030	-0,0929
ACF(3)	-0,037	0,037	-0,008	-0,020	-0,022
Q(5)	28,70	31,91	11,65	31,14	36,02

Mean, Std, Min, Max, Skew и Kurt означают среднее, стандартное отклонение, минимум, максимум, эмпирическую скошенность и куртозис соответственно. JB означает статистику Харке–Бера (5%-ое критическое значение 5,9773). ACF(1), ACF(2), ACF(3) — первые три лага эмпирической автокорреляционной функции. Q(5) — статистика Льюнга–Бокса (5%-ое критическое значение 11,0705).

Для конкурирующих прогнозов $\hat{f}_t(\cdot)$ и $\hat{g}_t(\cdot)$ необходимо протестировать гипотезу о равенстве средних цензурированных скоров

$$H_0 : \mathbb{E}_t S^{csl}(\hat{f}_t(\cdot); r_{t+1}) = \mathbb{E}_t S^{csl}(\hat{g}_t(\cdot); r_{t+1}) \quad \forall t = 0, \dots, n-1.$$

Тестовая статистика имеет вид стандартизованной средней разности эмпирических цензурированных скоров двух прогнозов:

$$DM_n = \frac{\bar{d}_n^*}{\sqrt{\hat{\sigma}_n^2/n}},$$

где

$$\bar{d}_n^* = \frac{1}{n} \sum_{t=0}^{n-1} \left(S^{csl}(\hat{f}_t(\cdot); r_{t+1}) - S^{csl}(\hat{g}_t(\cdot); r_{t+1}) \right),$$

а $\hat{\sigma}_n^2$ — состоятельная оценка асимптотической дисперсии величины \bar{d}_n^* . Применяя классический результат работы Diebold & Mariano (1995), авторы показывают, что при нулевой гипотезе статистика DM_n имеет стандартное нормальное асимптотическое распределение.

Таблица 2: P -значения тестов на предсказательную способность VaR для различных моделей

	BAC	F	GE	JPM	PFE
HS-VaR					
Безусловный	0,01	0,07*	0,04	0,00	0,22*
Условный	0,00	0,03	0,00	0,00	0,01
Совместный	0,00	0,01	0,00	0,00	0,02
N-GARCH(1,1)-VaR					
Безусловный	0,56*	0,49*	0,35*	0,44*	0,02
Условный	0,34*	0,35*	0,29*	0,06*	0,39*
Совместный	0,59*	0,61*	0,15*	0,13*	0,07*
SGT-GARCH(1,1)-VaR с постоянными моментами высоких порядков					
Безусловный	0,93*	0,89*	0,13*	0,64*	0,06*
Условный	0,95*	0,39*	0,42*	0,04	0,84*
Совместный	0,99*	0,68*	0,24*	0,13*	0,16*
NIG-GARCH(1,1)-VaR с постоянными моментами высоких порядков					
Безусловный	0,88*	0,81*	0,43*	0,53*	0,19*
Условный	0,94*	0,28*	0,43*	0,08*	0,80*
Совместный	0,99*	0,65*	0,36*	0,18*	0,20*
SGT-GARCH(1,1)-VaR с динамическими моментами высоких порядков					
Безусловный	0,93*	0,89*	0,13*	0,64*	0,06*
Условный	0,95*	0,39*	0,42*	0,04	0,84*
Совместный	0,99*	0,68*	0,24*	0,13*	0,16*
NIG-ACD-VaR с динамическими моментами высоких порядков					
Безусловный	0,88*	0,81*	0,43*	0,53*	0,19*
Условный	0,94*	0,28*	0,43*	0,08*	0,80*
Совместный	0,99*	0,65*	0,36*	0,18*	0,20*
SGE-ACD-VaR с динамическими моментами высоких порядков					
Безусловный	0,83*	0,71*	0,67*	0,62*	0,49*
Условный	0,88*	0,31*	0,29*	0,14*	0,23*
Совместный	0,89*	0,59*	0,43*	0,46*	0,25*

Нулевые гипотезы тестов на корректное безусловное, условное и совокупное покрытие, соответственно: $\mathbb{E}(I_t) = 1-p$, $I_t \sim$ независимы и $I_t \sim \text{iid } \mathcal{B}(1-p)$, где I_t — индикатор пробоя VaR. Звёздочками обозначены случаи неотвержения нулевой гипотезы на 5%-ом уровне значимости.

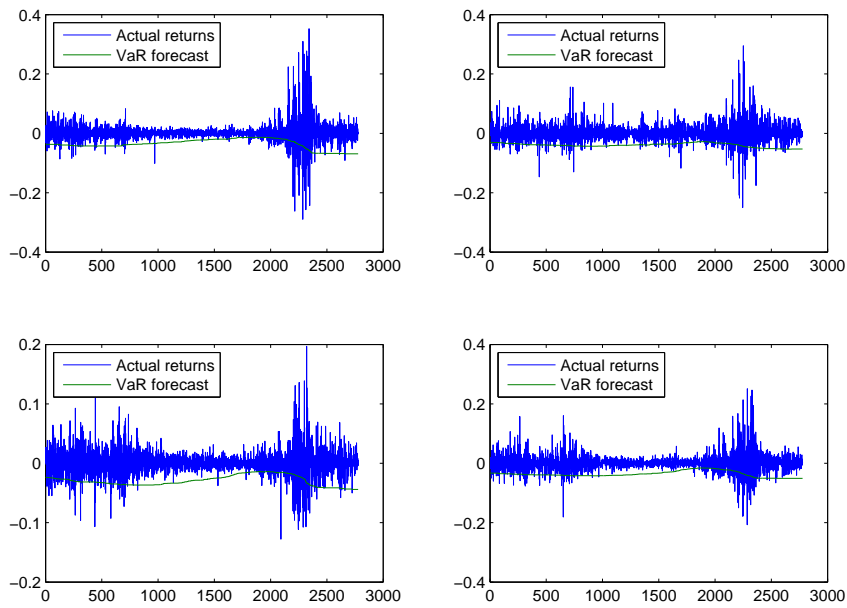


Рис. 1: HS-VaR для акций BAC, F, GE, JPM

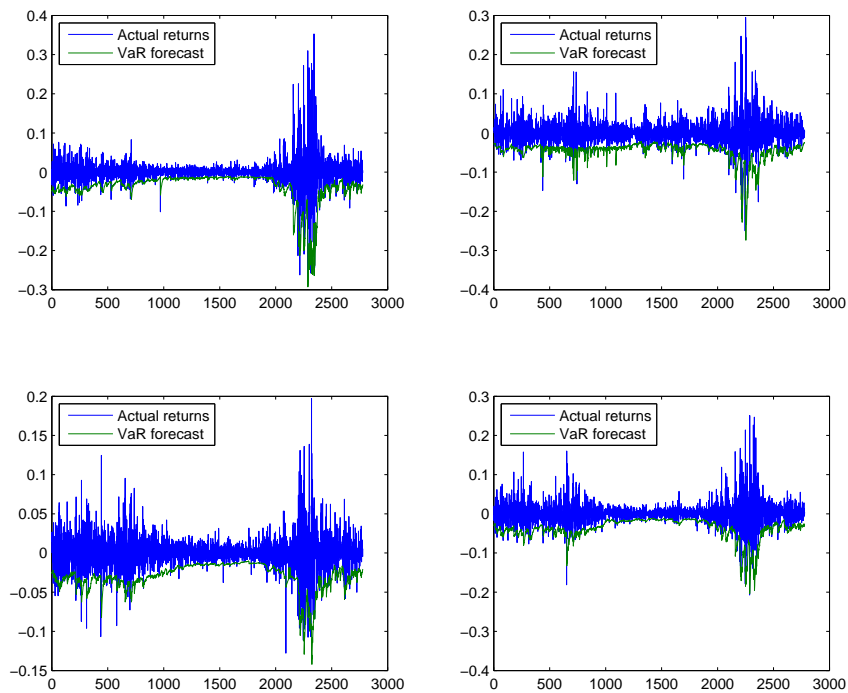


Рис. 2: SGE-ACD-VaR для акций BAC, F, GE, JPM

Таблица 3: P -значения тестов на одинаковую предсказательную способность для левых хвостов распределений для пар моделей с динамическими моментами высоких порядков

	VAC	F	GE	JPM	PFE
NIG-ACD и SGT-GARCH(1,1)	0,12*	0,04	0,07*	0,08*	0,12*
SGE-ACD и SGT-GARCH(1,1)	0,25*	0,63*	0,12*	0,32*	0,19*
SGE-ACD и NIG-ACD	0,86*	0,78*	0,56*	0,42*	0,53*

Нулевая гипотеза $H_0 : \mathbb{E}_t S(\hat{f}_t; r_{t+1}) = \mathbb{E}_t S(\hat{g}_t; r_{t+1})$ для всех t (одинаковая предсказательная способность). Звёздочками обозначены случаи неотвержения H_0 на 5%-ом уровне значимости.

4 Эмпирическое сравнение моделей VaR с динамическими моментами высоких порядков

4.1 Данные

Мы используем дневные доходности по ценам закрытия нескольких ликвидных⁵ акций, входящих в индекс широкого рынка S&P500, — Bank of America, Ford Motor Co., General Electric, J.P. Morgan и Pfizer, — с 1996 по 2010 годы включительно, что даёт 3778 наблюдений каждого ряда доходностей. Рассматриваемый период включает турбулентность финансового кризиса, что позволяет отследить реакцию различных способов оценивания VaR на отрицательные рыночные шоки. Описательные статистики представлены в таблице 1. Видно, что безусловное распределение доходностей далеко от гауссовского. Тест Льюнга–Бокса отвергает гипотезу об отсутствии автокорреляции на 5%-м уровне значимости, что может служить первым признаком наличия динамических моментов. Эмпирическая скошенность положительна для всех бумаг, кроме одной, что необычно, т.к. считается, что скошенность безусловного распределения доходностей должна быть отрицательной, см., например, Manganelli & Engle (2001). Эмпирический куртозис значительно больше 3 (значения, соответствующего гауссовскому распределению), что может свидетельствовать о наличии тяжёлых хвостов у распределения доходностей.

4.2 Оценивание и тестирование моделей

Для каждой модели прогноз 5%-го однодневного VaR вычисляется в скользящем окне длины⁶ $L = 1000$ с 10.03.1997 до конца выборки 31.12.2010.

Оценим для начала VaR простым выборочным квантилем (HS-VaR), в целях иллюстрации неадекватности метода. В первой секции таблицы 2 приведены результаты тестов на корректность накрытия. Видно, что HS-VaR плох не только как оценка условного VaR, но и для оценивания безусловного VaR. Причиной этого может быть недостаточность длины окна или нестационарность доходностей. Неудовлетворительное качество оценивания условного VaR видно на Рис. 1: HS-VaR не способен быстро отразить недавние шоки, и поэтому характеризуется большими запаздываниями.

Посмотрим теперь на результаты прогнозирования VaR с помощью моделей GARCH(1,1) с нормальным и скошенным студентовским распределением ошибок. Результаты тестирования моделей представлены во второй и третьей секциях таблицы 2. Видно, что GARCH(1,1) даёт значительно более точные прогнозы VaR, чем выборочный квантиль. Как условно, так

⁵См. веб-страницу www.marketvolume.com/stocks/mosttraded.asp?s=nyse

⁶Такая длина окна стандартна для вычисления 5%-го однодневного VaR, см., например, Kuuster, Mittnik & Paoletta (2006).

и безусловно модель показывает себя на удивление точной: только для акций PFE N-GARCH имеет неточное безусловное покрытие (однако для всех рядов нельзя отвергнуть гипотезу о точном совместном покрытии). Ответы VaR на отрицательные шоки для GARCH моделей по сравнению с наивным историческим квантилем характеризуются значительно меньшими задержками.

Теперь перейдём к моделям с динамическими условными моментами высоких порядков. Прогнозы VaR в модели SGT-GARCH хороши как в смысле условного, так и безусловного покрытия, и доминируют p -значения модели NIG-ACD, см. таблицу 2. Последняя показывает равномерно лучшие результаты, чем простая модель NIG-GARCH(1,1) с постоянными моментами, что согласуется с выводами работы Wilhelmsson (2009). То же верно и для предложенной нами модели SGE-ACD (см. последнюю секцию таблицы 2). Рис. 2 показывает, что модель SGE-ACD позволяет получить прогнозы VaR с меньшей кластеризацией и более быстрой реакцией на неожиданные скачки волатильности.

Так как p -значения тестирования на корректность покрытия для модели SGE-ACD близки к p -значениям для моделей SGT-GARCH и NIG-ACD, встаёт естественный вопрос: имеет ли модель SGE-ACD значимо лучшую (худшую) предсказательную способность для VaR, чем модели SGT-GARCH и NIG-ACD? Скоринговый метод Diks, Panchenko & van Dijk (2011), описанный в предыдущем разделе, идеально подходит для разрешения этого вопроса. Результаты в таблице 3 показывают, что доминирование модели SGT-GARCH над моделью NIG-ACD может быть ложным. В самом деле, только для акций Ford нулевая гипотеза об одинаковой предсказательной способности в левом хвосте распределения может быть отвергнута на 5% уровне значимости, причём соответствующее p -значение всего лишь немногим меньше 5%. Сравнивая аналогичным образом модель SGE-ACD попарно с моделями SGT-GARCH(1,1) и NIG-ACD, из таблицы 3 видим, что тест не может различить предсказательную способность трёх моделей в хвостах распределения, и значит, скорее всего, не сможет различить качество предсказаний условных квантилей VaR.

5 Заключение

Разрушительные последствия глобальной рецессии не оставляют сомнений в том, что эффективный риск-менеджмент должен быть важнейшим приоритетом мировых финансовых институтов. Сложность динамики финансовых рынков делает необходимым использование продвинутых моделей для расчёта различных мер риска, таких как VaR, который получил в последнее время широкое распространение в бизнесе из-за концептуальной простоты и регуляторных требований к его вычислению.

В настоящей работе производится сравнение некоторых моделей финансовых доходностей с динамическими условными моментами высоких порядков, таких как скошенность и куртозис. Так как считается, что непостоянство условных моментов критически сказывается на предсказании условных распределений доходностей, в качестве меры качества моделей выступают прогнозы VaR. Мы действительно находим, что по качеству условных прогнозов VaR модели с динамическими моментами, такие как GARCH(1,1) со скошенным студентовским (SGT) распределением ошибок и модель авторегрессионной условной нормальной обратной гауссовской плотности NIG-ACD превосходят свои более традиционные аналоги с постоянными условными моментами. Предложенная нами комбинация данных моделей (SGE-ACD) со скошенным обобщённым распределением ошибки позволяет получить столь же качественные прогнозы VaR, в то же время являясь вычислительно менее сложной, особенно в сравнении с моделью NIG-ACD.

В заключение необходимо отметить, что в последнее время набирают популярность полупараметрические методы оценивания VaR, а также использование теории экстремальных значений для более точного оценивания хвостов условных распределений, см., например,

Kuester, Mittnik & Paoletta (2006). Насколько нам известно, в настоящее время не существует достаточно полного сравнения таких методов с параметрическим подходом, а значит, разработка методологии для такого сравнения может являться перспективным направлением дальнейших исследований.

Благодарности

Автор искренне благодарит своих научных руководителей в Российской экономической школе Станислава Анатольева и Станислава Храпова, а также участников исследовательского проекта РЭШ по анализу временных рядов 2012–2013 гг. за ценные замечания и плодотворные обсуждения.

Список литературы

- Bali, T. G., H. Mo & Y. Tang (2008). The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. *Journal of Banking & Finance* 32, 269–282.
- Barndorff-Nielsen, O. E. (1997). Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* 24, 1–13.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.
- Bollerslev, T. & J. M. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11, 143–172.
- Christoffersen P. & D. Pelletier (2004). Backtesting value-at-risk: a duration based approach. *Journal of Financial Econometrics* 2, 84–108.
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review* 39, 841–862.
- Danielsson, J., D. Jorgensen, S. Mandira, G. Samorodnitsky & C. Vries (2005). Subadditivity re-examined: the case for value-at-risk. Working paper. Financial Markets Group, London School of Economics and Political Science.
- Diebold, F. & R. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13, 253–263.
- Diks, C., V. Panchenko & D. van Dijk (2011). Likelihood-based scoring rules for comparing density forecasts in tails. *Journal of Econometrics* 163, 215–230.
- Föllmer, H. & A. Schied (2004). *Stochastic finance. An introduction in discrete time*. Walter de Gruyter, Berlin.
- Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review* 35, 705–730.
- Harvey, C. & A. Siddique (1999). Autoregressive conditional skewness. *Journal of Financial & Quantitative Analysis* 34, 465–487.
- Jensen, M. B. & A. Lunde (2001). The NIG-S&ARCH model: a fat-tailed, stochastic and autoregressive conditional heteroskedastic volatility model. *Econometrics Journal* 4, 319–342.
- Jondeau, E. & M. Rockinger (2003). Conditional volatility, skewness and kurtosis: existence, persistence and comovements. *Journal of Economic Dynamics and Control* 27, 1699–1737.
- Kuester, K., S. Mittnik & M. Paoletta (2006). Value-at-risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics* 4, 53–89.
- Manganelli, S. & R. Engle (2001). Value-at-risk models in finance. Working paper, European Central Bank.
- Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science* 44, 1650–1661.
- Theodossiou, P. (2000). Skewed generalized error distribution of financial assets and option pricing. SSRN working paper 219679.
- Wilhelmsson, A. (2009). Value-at-risk with time-varying variance, skewness and kurtosis – the NIG-ACD model. *Econometrics Journal* 12, 82–104.
- Zumbach, G. (2007). A gentle introduction to the Risk Metrics 2006 methodology. New York, RiskMetrics group.

Higher order conditional moment dynamics and forecasting value-at-risk

Grigory Franguridi

New Economic School, Moscow, Russia

We empirically investigate the possibilities for enhancing value-at-risk predictions by explicit modelling conditional higher order moment dynamics of financial returns. Using one-day-ahead VaR forecasts for 5 highly liquid constituents of the S&P500 index from different industrial sectors, we compare performances of the benchmark GARCH model with skewed generalized Student's innovations with a set of models allowing for time-varying asymmetry and kurtosis such as ARCD-type models with normal inverse gaussian and skewed generalized Student's errors. As predictive accuracy tests we exploit both the scoring rules for left tail forecasts and likelihood-ratio tests for correct (un)conditional quantile forecasts. We also propose a parsimonious ARCD model with the skewed generalized error distribution for innovations, asymmetric power ARCH for volatility and autoregressive dynamics for skewness and kurtosis related parameters which is shown to perform not worse than the aforementioned models in terms of VaR prediction accuracy, while being computationally less demanding.

Keywords: value-at-risk, conditional distribution, skewness, kurtosis, financial returns

JEL Classification: C22, C53, C58

Александр Иванович САИЧЕВ (1946–2013)



8 июня 2013 года на 67-м году жизни от нас ушёл Александр Иванович Саичев — ученый с мировым именем, лауреат Государственной премии РФ 1997 г., автор более двухсот статей и монографий в области теории случайных процессов.

Александр Иванович родился в Нижнем Новгороде 12 декабря 1946 г. В 1969-м он окончил радиофизический факультет Нижегородского государственного университета им. Н.И. Лобачевского, где впоследствии защитил кандидатскую, а затем и докторскую диссертации. Александр Иванович работал в должности заведующего кафедрой, а потом профессора кафедры математики, был руководителем научной школы, почётным профессором ННГУ.

Область научных интересов Александра Ивановича была обширна и включала в себя теорию диффузии, процессы, связанные с распространением волн в нелинейных случайных средах, квантовый хаос, геофизические процессы и многое другое. В течение своей научной деятельности он активно сотрудничал с учёными из многих стран, в том числе из Швеции, США, Израиля, Франции и Швейцарии.

С 2007 г. Александр Иванович занимался проблемами моделирования стохастических процессов финансовых рынков на кафедре предпринимательских рисков в Швейцарской высшей технической школе Цюриха. Его работы опубликованы во многих научных изданиях, в том числе и в журнале «Квантиль». Он был прекрасным светлым человеком с добрым сердцем и щедрой душой, настоящим другом и учителем. Мы всегда будем помнить о нём.

Quantile

No. 12, February 2014

ISSN 2309-3013

English page in the world wide web: <http://quantile.ru/eng>

Electronic mail address: quantile@quantile.ru

Access to the journal is free and unlimited

EDITOR

Stanislav Anatolyev
New Economic School (Moscow, Russia)

EDITORIAL COUNSEL

Victoria Zinde-Walsh
McGill University (Montréal, Canada)

Rustam Ibragimov
Imperial College London (London, UK)

Anna Mikusheva
Massachusetts Institute of Technology (Cambridge, USA)

Alexey Onatsky
Cambridge University (Cambridge, UK)

Alexander Tsyplakov
Novosibirsk State University (Novosibirsk, Russia)

Victor Chernozhukov
Massachusetts Institute of Technology (Cambridge, USA)

GUIDE TO AUTHORS

Manuscripts for publication in the “Articles” section should be submitted by electronic mail to the address submit@quantile.ru. Submitted work may be drawn from any applied field within the economics profession. The main requirement is correct usage of adequate econometric methodology. The manuscript should be written in Russian (for Russian-speaking persons) or in English (for all others) in the *Microsoft Word* or (preferably) *LaTeX* formats, and not exceed 30 double-spaced A4 pages. All submissions are subject to quality control by the editorial counsel and independent referees. A promising manuscript may be returned to the author(s) for polishing or rewriting. The editor also invites econometrics experts worldwide to contribute to the methodological sections of the journal.

Articles and methodological material published in “Quantile” do not transfer original copyright, neither in full, nor in part.

Quantile

*international econometric journal
in Russian language*

**No. 12
February 2014**

IN THIS ISSUE

Conference invitation

International conference “Modern econometric tools and applications”

Econometric literacy: dynamic stochastic general equilibrium models

- | | |
|---|----|
| Mikusheva, Anna. Estimation of dynamic stochastic general equilibrium models | 1 |
| Jones, Callum; Kulish, Mariano. A practical introduction to DSGE modeling with Dynare | 23 |

Advice to econometrics students

- | | |
|--|----|
| Tsyplakov, Alexander. A mini-dictionary of English econometric terminology III | 45 |
|--|----|

Articles: financial econometrics

- | | |
|---|----|
| Groshev, Oleg. Time varying vine copulas for multivariate returns | 53 |
| Franguridi, Grigory. Higher order conditional moment dynamics and forecasting value-at-risk | 69 |

Obituary

Alexander Saichev