

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 (маржинальными) 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

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

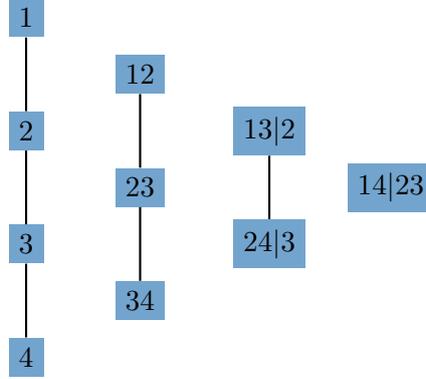
можно построить следующее разложение:

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

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.

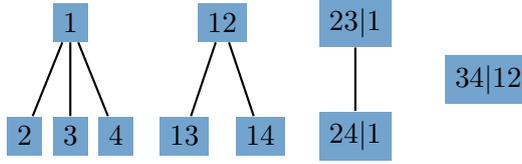
Пример 3 Следующее разложение соответствует изображенному объединению деревьев (лесу). Когда каждое дерево в лесу является линейным графом, разложение называется *D-вьюн (D-vine)*.

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



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

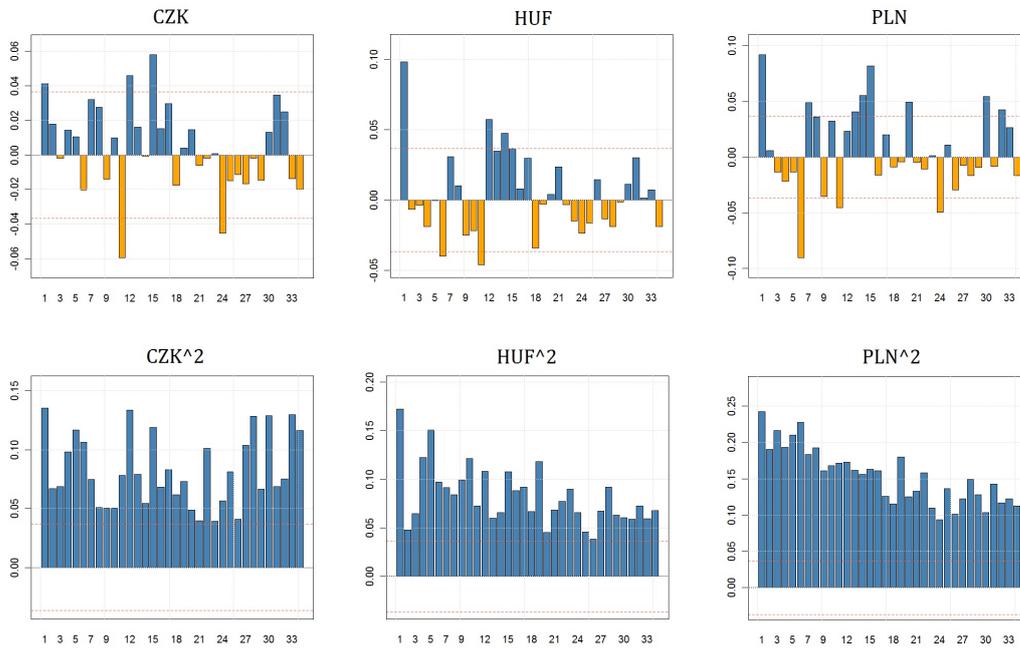
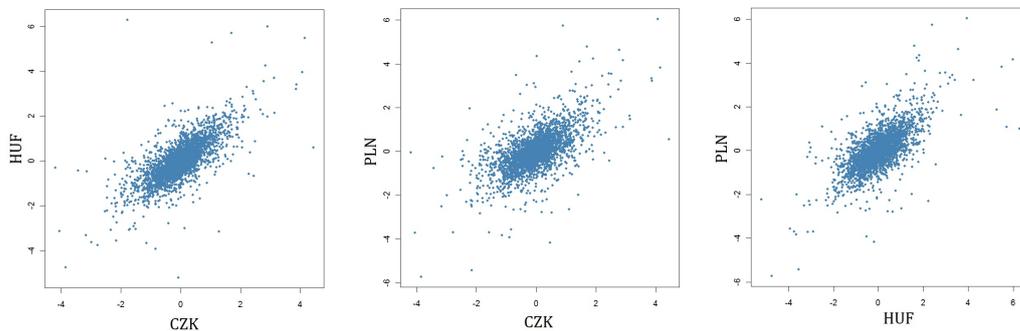


График 2: Диаграммы рассеяния доходностей пар USD/CCY



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:	<i>AIC</i>	2,128	2,129	2,129	2,129
	<i>BIC</i>	2,139	2,141	2,141	2,144
HUF:	<i>AIC</i>	2,450	2,443	2,367	2,364
	<i>BIC</i>	2,460	2,456	2,382	2,380
PLN:	<i>AIC</i>	2,249	2,247	2,247	2,244
	<i>BIC</i>	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.

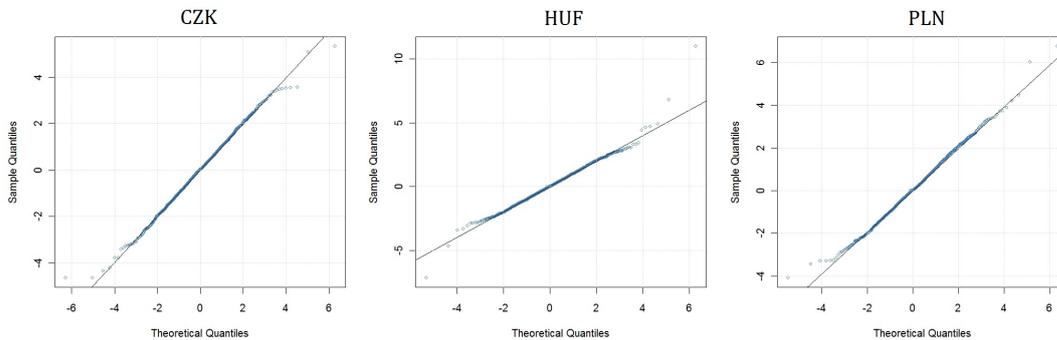
На графике 3 изображены QQ-диаграммы, которые демонстрируют разумное качество подгонки данных.

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



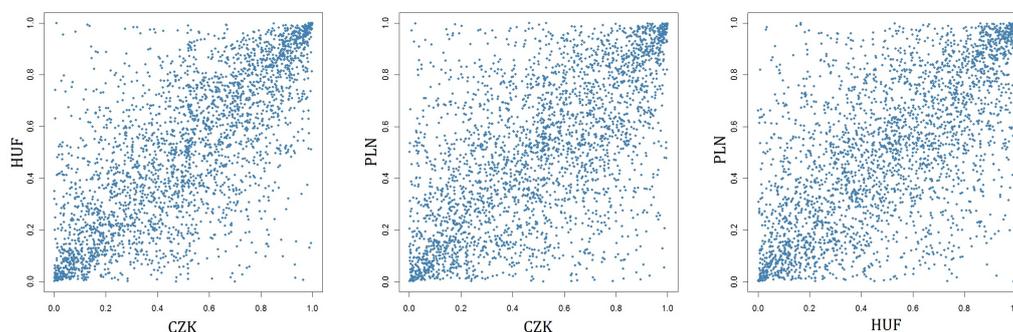
7 Моделирование копул

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	p-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.

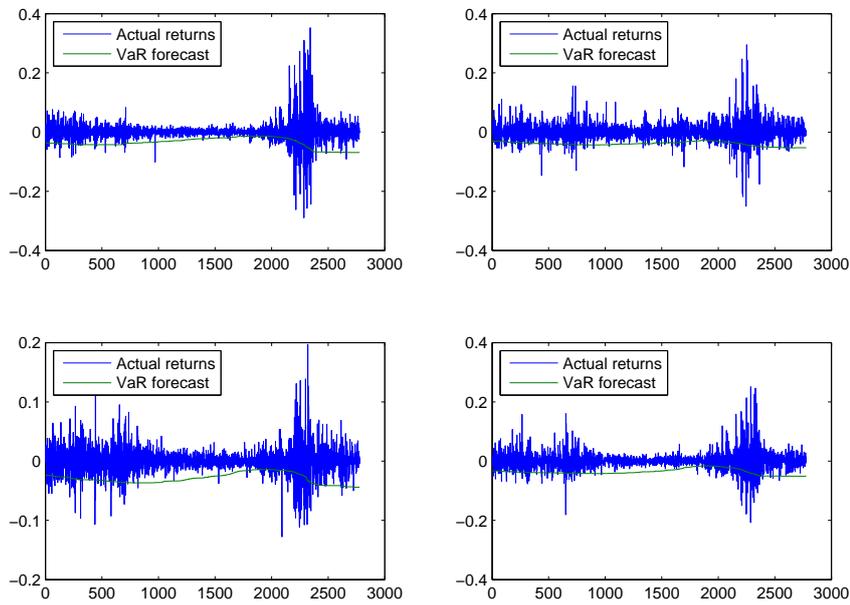


Рис. 1: HS-VaR для акций BAC, F, GE, JPM

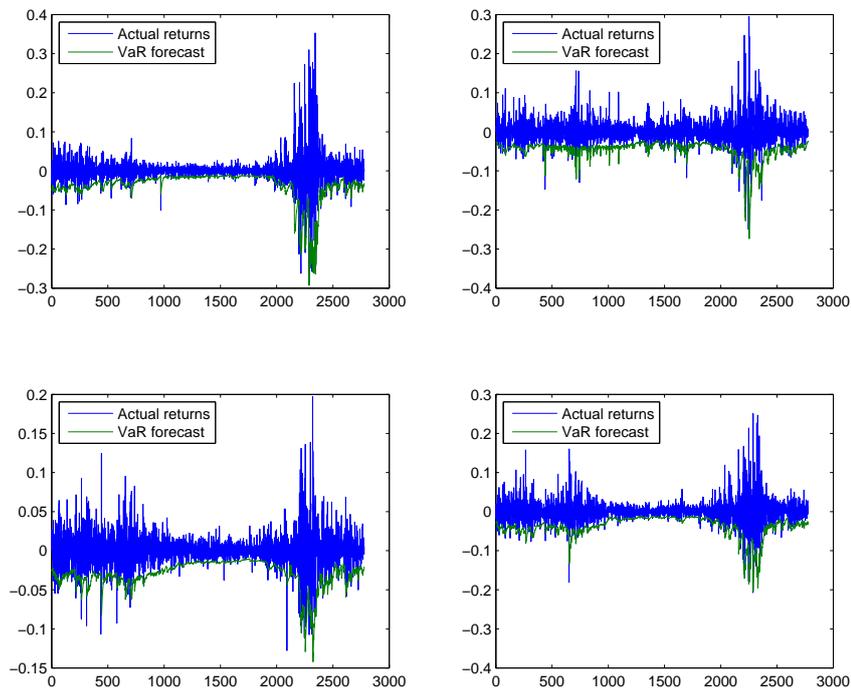


Рис. 2: SGE-ACD-VaR для акций BAC, F, GE, JPM

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 к 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 скопленным 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 скопленным 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