

ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ВЫСШАЯ ШКОЛА ЭКОНОМИКИ

В.А. Лапшин

**АЛГОРИТМ ОЦЕНКИ ПАРАМЕТРОВ
СТОХАСТИЧЕСКОЙ ДИНАМИКИ
СРОЧНОЙ СТРУКТУРЫ
ПРОЦЕНТНЫХ СТАВОК**

Препринт WP16/2009/02

Серия WP16

Финансовая инженерия,
риск-менеджмент и актуарная наука

Москва

Государственный университет – Высшая школа экономики

2009

УДК 336.78
ББК 65.262.2
Л24

Редакторы серии WP16
«Финансовая инженерия,
риск-менеджмент и актуарная наука»
С.Н. Смирнов, А.Г. Шоломицкий

Л24 **Лапшин В.А. Алгоритм оценки параметров стохастической динамики срочной структуры процентных ставок:** Препринт WP16/2009/02. — М.: Изд. дом Государственного университета – Высшей школы экономики, 2009. — 64 с.

В настоящее время модели стохастической динамики срочной структуры процентных ставок становятся всё более сложными. К сожалению, их практическое применение затруднено трудоёмкостью оценки параметров таких моделей по реальным рыночным данным. В работе предлагается алгоритм оценки параметров разработанной автором сложной динамической модели, основанной на моделировании динамики кривых доходностей при помощи стохастических дифференциальных уравнений в бесконечномерном функциональном пространстве, а также статистической обработки поступающей информации о сделках и котировках с использованием логической (нечёткой) интерпретации вероятности, отражающей их неполную достоверность. Описывается возможность применения технологий параллельных вычислений для ускорения расчётов.

УДК 336.78
ББК 65.262.2

Препринты Государственного университета – Высшей школы экономики размещаются по адресу: <http://new.hse.ru/C3/C18/preprintsID/default.aspx>

© Лапшин В.А., 2009
© Оформление. Издательский дом
Государственного университета –
Высшей школы экономики, 2009

Содержание

1	Введение	4
1.1	Ценообразование облигаций	6
2	Постановка задачи	9
2.1	Бесконечномерный случай	9
2.2	Конечномерный случай	13
3	Обзор методов оценки параметров	17
3.1	Краткий обзор методов оценки	17
3.2	Оценка параметров модели	22
4	Описание метода Монте-Карло для марковских цепей	23
5	Описание шагов алгоритма	26
5.1	Начальное приближение	26
5.2	Выборка $\{c_i\}$	28
5.3	Выборка $\{\sigma^s\}$	28
5.4	Выборка ν	30
5.5	Выражения для производных	30
5.6	Фиктивные наблюдения	35
6	Применение предварительных и параллельных вычислений	36
6.1	Предварительные вычисления	36
6.2	Параллельные вычисления	37
7	Иллюстрация	37
8	Заключение	54

1 Введение

В финансовом анализе при рассмотрении будущих потоков платежей часто используют понятие текущего значения (present value). Рассмотрим альтернативу: гарантированно получить 1 рубль через 1 год или d рублей сейчас. Значение d , при котором альтернативы будут равнозначны, называют текущим значением 1 рубля, полученного через 1 год. Это та сумма, которую среднестатистический инвестор согласится получить или отдать сейчас в обмен на платёж в будущем, текущая рыночная стоимость 1 рубля, полученного через 1 год. Зависимость текущей стоимости 1 рубля от срока до платежа x называют функцией дисконтирования и обозначают $d(x)$, а её значения в отдельных точках x_i называют коэффициентами дисконтирования и обозначают d_i . Исходя из экономического смысла, можно постулировать следующие свойства функции дисконтирования:

1. $d(0) = 1$: 1 рубль немедленно стоит ровно один рубль.
2. $d(x_1) < d(x_2)$, если $x_1 < x_2$: мы предпочитаем получить деньги раньше, чем позже.
3. $d(x) > 0$: деньги стоят сколько-то через любой промежуток времени.
4. $\lim_{x \rightarrow +\infty} d(x) = 0$: деньги бесконечно далеко нас не интересуют.

Вместо функции дисконтирования часто используют различные процентные ставки, так как они более наглядны. Так, процентная ставка на срок x , подразумеваемая коэффициентом дисконтирования $d(x)$ — это размер ставки по вкладу на срок x , который должен начислять банк, чтобы вклад размером $d(x)$ через время x вырос до размера 1. Конкретное числовое значение процентной ставки $r(x)$ зависит от способа начисления процентов. При непрерывном начислении процентов

выполнено соотношение

$$d(x) = \exp(-r(x)x),$$

а при дискретном раз в δ :

$$d(x) = \frac{1}{(1 + r(x)\delta)^{\frac{x}{\delta}}}.$$

Выбор того или иного способа начисления процентов — вопрос конвенции. Мы будем использовать непрерывное начисление процентов, так как это удобнее с математической точки зрения. Отметим, что рыночные котировки чаще даются с использованием дискретного начисления процентов раз в 3, 6 или 12 месяцев.

Мгновенная форвардная процентная ставка на срок x , $f(x)$, связана с процентной ставкой $r(\cdot)$ следующим соотношением:

$$r(x) = \frac{1}{x} \int_0^x f(\tau) d\tau.$$

Это мгновенная процентная ставка ($r(0)$), которая должна реализоваться через время x , чтобы обеспечить эквивалентность двух стратегий: положить деньги на срок x под процент $r(x)$ и непрерывно вкладывать на бесконечно малый срок под постоянно меняющийся процент $r(0)$ в течение всего срока x . Мы будем записывать текущее (астрономическое) время нижним индексом, обозначая текущий момент через t : $r_t(0)$ — текущее значение мгновенной процентной ставки, а $r_{t+x}(0)$ — её значение через промежуток времени x . Таким образом, в мире, лишённом случайности, выполнялось бы равенство

$$r_{t+x}(0) = f_t(x).$$

Говоря о срочной структуре процентных ставок, обычно подразумевают кривую процентных ставок $r(x)$ (или бескупонную кривую доходностей¹; это название будет объяснено ниже), тем не менее, для ряда целей (например, для оценки производных финансовых инструментов) гораздо важнее кривая мгновенных форвардных процентных ставок $f(x)$.

1.1 Ценообразование облигаций

Чтобы использовать котировки облигаций для определения срочной структуры процентных ставок, необходимо определить связь между этими объектами. Рассмотрим бескупонную облигацию номиналом 1 со сроком до погашения x . Бескупонная облигация — это та, по которой будет проведена всего одна выплата. Пусть эта выплата имеет размер 1 и произойдёт через время x . И пусть облигация безрисковая, т.е. отсутствуют кредитные риски (риски неуплаты), риски ликвидности (невозможности продать бумагу) и другие. Тогда её текущая стоимость будет равна $d(x)$. Кривая доходностей, соответствующая этой функции дисконтирования ($r(x) = -\frac{1}{x} \ln d(x)$), называется безрисковой кривой доходностей.

Если у нас теперь есть реальная облигация, её цена будет ниже, чтобы компенсировать возможные риски, а кривая доходностей, построенная по таким облигациям, будет лежать выше безрисковой кривой доходностей. Тем не менее, для оценки финансовых инструментов, схожих с данной облигацией по кредитному качеству и ликвидности, скорее подойдёт именно эта кривая, а не безрисковая.

Если облигация купонная, т.е. по ней имеются периодические выплаты процентов (купонов), или амортизационная, т.е. основной долг выплачивается по частям, то потоков платежей будет больше. Пусть

¹ Часто используют термин “кривая доходности”, однако мы будем придерживаться указанного написания, чтобы подчеркнуть тот факт, что доходности на разные сроки разные.

имеется n_s потоков платежей во времена $\tau_1, \dots, \tau_{n_s}$ размерами F_1, \dots, F_{n_s} соответственно. Логично предположить, что стоимость такой облигации будет равна стоимости портфеля (набора) из n_s бескупонных облигаций со сроками до погашения $\tau_1, \dots, \tau_{n_s}$ и номиналами F_1, \dots, F_{n_s} соответственно. Владение портфелем бескупонных облигаций даёт больше возможностей, нежели владение одной купонной облигацией: так, портфель можно продать частями, а купонную облигацию только целиком, но это стандартное допущение. В таком случае, стоимость нашей купонной облигации P будет определяться соотношением

$$P = \sum_{s=1}^{n_s} F_s d(\tau_s).$$

В реальности на ценообразование очень сильное влияние оказывают вопросы кредитного качества, ликвидности, налогообложения. Первый вопрос мы рассмотрим позже, а второй и третий не будем рассматривать в рамках данной работы ввиду отсутствия на сегодняшний день адекватного их описания (как математического, так и экономического).

Обычно модели кривой доходностей используются для трёх целей:

- Ценообразование и хеджирование. Это то, зачем модели используют трейдеры. Им важна точная оценка текущей стоимости инструментов для выявления среди них недо- и переоценённых. Также они используют модели для хеджирования — уменьшения (в идеале — сведения к нулю) риска, связанного с занятием той или иной позиции на рынке. Это самая многочисленная группа пользователей моделей кривой доходностей.
- Риск-менеджмент. Риск-менеджеры используют модели для проработки различных вариантов развития событий, чтобы дать оценку максимально неблагоприятного варианта развития событий.

- Объяснение изменений кривой доходностей. Если бы мы могли объяснить, почему (в рамках модели) процентные ставки изменились так или иначе, это дало бы нам возможность глубже понять экономические взаимосвязи между процентными ставками и различными экономическими переменными.

Какой должна быть модель? В связи со скудностью и плохим качеством доступных данных, хорошее их отражение не может быть критерием качества модели: любую мало-мальски сложную модель можно откалибровать под имеющиеся на настоящий момент данные. Калибровка модели по историческим данным (которых, разумеется, гораздо больше), вполне вероятно, будет давать очень разные результаты в разные периоды. И хотя определить лучшую модель весьма сложно, можно выделить общие качества, которыми должна обладать хорошая модель.

- Точная оценка простых финансовых инструментов. Если этого не будет, модель будет бесполезна и даже опасна. Рынок очень быстро заметит банк (или другой финансовый институт), который выставляет неправильные цены и весьма скоро банк понесёт серьёзные убытки от неправильной оценки инструментов.
- Лёгкость калибровки к рынку. Если параметры модели невозможно оценить, она бесполезна. Если этот процесс занимает много времени, полезность модели сильно ограничена.
- Устойчивость модели. Модель должна работать на максимально широком диапазоне рынков и рыночных ситуаций. Например, модель с нормально распределёнными изменениями доходностей становится неустойчивой при приближении процентных ставок к нулю. Использовать такую модель для анализа рынка Японии в 1998 году — плохая идея.

- Возможность оценки других финансовых инструментов, в том числе и новых. Такая возможность — это серьёзное преимущество на рынке. Некоторые экзотические инструменты не могут быть оценены простыми моделями.
- Устойчивость значений параметров во времени. Это полезно, если данных недостаточно, когда некоторые инструменты могут не торговаться в течение продолжительного времени.

2 Постановка задачи

Ниже будет очень кратко изложена суть рассматриваемой модели динамики кривой доходностей. Краткое её описание можно найти в работах [1, 2]; более подробное будет доступно в скором времени.

2.1 Бесконечномерный случай

Фазовой переменной модели является кривая мгновенных форвардных процентных ставок $f_t(x)$, $x \in [0, T]$, которая рассматривается как элемент гильбертова пространства $\mathcal{H} = W_2^1[0, T]$, где $T = const$ — горизонт моделирования (очевидно, что данные заданы на конечном отрезке времени; вкуче с известным результатом [3, 4], говорящим, что в большинстве разумных случаев мгновенные форвардные ставки стремятся к пределу: $\lim_{x \rightarrow 0} f_t(x) = L_t$).

Предполагается, что стохастическая динамика бесконечномерной переменной — кривой мгновенных форвардных процентных ставок — описывается стохастическим дифференциальным уравнением в пространстве \mathcal{H} :

$$\begin{cases} df_t &= (Df_t + \sum_{s=1}^{\infty} \mathcal{S}(\sigma^s f_t) - \sum_{s=1}^{\infty} \gamma^s \sigma^s f_t) dt + \sum_{s=1}^{\infty} \sigma^s f_t d\beta_t^s \\ f_0 &= f_0, \end{cases}$$

где β_t^s — независимые одномерные винеровские процессы, D — инфинитезимальный генератор полугруппы сдвигов $S(t) : (S(t)f)(x) = f(\min(x + t, T))$, \mathcal{S} — оператор $(\mathcal{S}f)(x) = f(x) \int_0^x f(\tau) d\tau$, $\sigma^s \in \mathcal{H}$, $s = 1, \dots, \infty$ и $\gamma^s \in \mathbb{R}$, $s = 1, \dots, \infty$ — параметры модели: σ^s суть коэффициенты волатильности, а γ^s — постоянные рыночные цены риска, связанные с s -ым случайным фактором.

После замены переменной $\xi_t = \ln f_t$ и линеаризации $e^\xi = e^{\xi_0} + e^{\xi_0}(\xi - \xi_0) + \bar{o}(\xi - \xi_0)$ мы имеем

$$d\xi_t = ((A + D)\xi_t + b) dt + \sum_{s=1}^{\infty} \sigma^s d\beta_t^s, \quad (1)$$

где

$$b(x) = \sum_{s=1}^{\infty} \sigma^s(x) \int_0^x f_0(\tau)(1 - \xi_0(\tau))\sigma^s(\tau) d\tau - \sum_{s=1}^{\infty} \gamma^s(x)\sigma^s(x) - \frac{1}{2} \sum_{s=1}^{\infty} \sigma^s(x)^2,$$

D — оператор дифференцирования, а оператор A действует на функцию ξ по закону

$$(A\xi)(x) = \sum_{s=1}^{\infty} \sigma^s(x) \int_0^x f_0(\tau)\sigma^s(\tau)\xi(\tau) d\tau.$$

Предполагается также, что наблюдения (сделки) происходят в известные (неслучайные) моменты времени t_i . Информация, заключённая в наблюдении, состоит из (верхний индекс i означает, что данные относятся к i -му наблюдению):

- цен облигаций P_k^i , $k = 1, \dots, n_k^i$;
- котировок спроса и предложения на них b_k^i , a_k^i , $k = 1, \dots, n_k^i$;
- статической информации об облигациях, т.е. о расписании τ_s^i , $s = 0, \dots, n_s^i$ и объёмах $F_{s,k}^i$, $s = 0, \dots, n_s^i$, $k = 1, \dots, n_k^i$ платежей.

Сроки выплат τ_s^i , отсчитывающиеся от времени наблюдения t^i , едины для всех облигаций: через срок τ_s^i , т. е. в момент времени $t_i + \tau_s^i$ по k -ой облигации из наблюдаемых будет выплачена сумма $F_{s,k}^i$. Если времена не едины, введём фиктивные нулевые платежи, где это необходимо.

В этих предположениях уравнение ценообразования облигаций будет выглядеть так:

$$P_k^i = \sum_{s=0}^n F_{s,k}^i \exp \left[- \int_0^{\tau_s^i} f_{t_i}(x) dx \right].$$

Чтобы не загромождать запись, начиная с этого времени, мы будем опускать верхний индекс i , показывающий, что информация относится к i -ому наблюдению).

Далее предполагается, что достоверность информации, несомой наблюдением, ставится под сомнение. Степень достоверности (credibility) этой информации может зависеть от различных факторов:

- от разницы котировок спроса и предложения (т.н. bid-ask спреда) — обратно пропорционально;
- от объёма сделки/котировки — нелинейная зависимость: тем достовернее, чем ближе к среднему объёму, характерному для рынка;
- от любых других параметров.

Чтобы учесть это в модели, мы будем считать, что величины P_k наблюдаются с нормально распределённым шумом $\epsilon_k \sim N(0, \delta_k^2)$. Такой подход соответствует логической интерпретации вероятности: вероятность — степень достоверности утверждения.

Для простоты мы остановимся лишь на bid-ask спреде: выберем среднеквадратичное отклонение шума равным ему, чтобы связать точность с ликвидностью рынка.

Это предположение является довольно сильным упрощением ситуации. На самом деле, распределение цены сделки по отношению к котировкам bid-ask сильно зависит как от типа рынка (наличия на нём market-maker'ов), так и от доступных данных: при полных данных мы будем наблюдать сделки исключительно по котировкам, но агрегированная (скажем, за день) цена с большой вероятностью будет лежать внутри агрегированных котировок, однако этот вопрос будет исследован в другой работе.

И наконец, последнее предположение относительно наблюдений заключается в том, что кривые доходностей, используемые участниками рынка для расчёта цены сделки, являются достаточно гладкими. Подобно статистической механике, будем считать, что правдоподобность того, что восприятие рынком сделки приведёт к кривой h , будет пропорциональна $e^{-\alpha E(h)}$, где $E(h)$ - некоторая мера негладкости кривой h . Если мы предположим, что у участников рынка есть среднее мнение относительно того, насколько гладкой должна быть форвардная кривая, то наше предположение соответствует распределению негладкости с максимальной энтропией при фиксированном среднем значении α^{-1} .

Функционал $E(h)$ может быть выбран произвольным образом, чтобы отразить наше представление о желаемой кривой доходностей. Нам будет удобно выбрать в качестве меры негладкости величину $E(h) = \|(\sqrt{h})'\|_{L_2}^2$, что соответствует [5] и [6]. Она кажется нелогичной на первый взгляд, тем не менее, именно такое выражение позволит нам получить ключевое совпадение в дальнейшем. Разумеется, возможны и другие формализации негладкости.

Это находится в согласии с принципом регуляризации (см. [7]): задача оценки даже самой кривой доходности, не говоря уже о параметрах её стохастической динамики, как подмечено во введении, является некорректно поставленной, кроме того, гладкость кривой мгновенных форвардных ставок — разумное предположение с экономической точки

зрения.

Условное распределение наблюдаемой цены P_k при известной кривой мгновенных форвардных ставок $f(\cdot)$ будет равно

$$P_k \sim N \left(\sum_{s=0}^n F_{s,k} \exp \left[- \int_0^{\tau_s} f(x) dx \right], a_k - b_k \right),$$

где a_k, b_k — соответственно котироваемые цены продавца и покупателя k -ой облигации.

Таким образом, наблюдение в формализуется в терминах функции правдоподобия следующим образом:

$$-\ln \frac{dP_{f_{t_i}}}{dP_{f_{t_i-0}}} \propto \sum_{k=1}^N w_k \left(\sum_{s=0}^n F_{s,k} \exp \left[- \int_0^{\tau_s} f_{t_i}(x) dx \right] - P_k \right)^2 + \alpha \int_0^T (\sqrt{f_{t_i}(x)})'^2 dx,$$

где $w_k = (a_k - b_k)^{-1}$. Единственный параметр, подлежащий заданию, — α , мера желаемой гладкости кривой, некоторый аналог температуры. Мы будем считать его заданным извне, например, пользователем системы, но можно оценить его и на основании статистических данных.

2.2 Конечномерный случай

Аналитическое решение уравнения (1) не выписывается, так как оператор $e^{t(A+D)}$ представляет собой оператор решения уравнения в частных производных

$$\frac{\partial \xi(x, t)}{\partial t} = \sum_{s=1}^{\infty} \sigma^s(x) \int_0^x f_0(\tau) \sigma^s(\tau) \xi(\tau, t) d\tau + \frac{\partial \xi(x, t)}{\partial x},$$

аналитическое решение для которого пока не найдено.

Поэтому для практической реализации необходимо ввести в пространстве \mathcal{H} базис и все дальнейшие вычисления производить с разложением по этому базису. Для этого мы используем аппарат кратномасштабного анализа (см., например, [8, 9] или любой учебник по вейвлет-анализу).

В пространстве L_2 любую функцию можно разложить как

$$f(x) = \sum_j c_j \varphi_j(x) + \sum_{i=-\infty}^0 \sum_k d_{i,k} \psi_{i,k}(x),$$

где

$$\varphi_j(x) = \varphi(x - j), \quad \psi_{i,k} = 2^{k/2} \psi(2^i x - k),$$

а $\psi(\cdot)$ — собственно вейвлет. Если он принадлежит к классу так называемых ортогональных вейвлетов, то все функции $\{\varphi_j\}, \{\psi_{i,k}\}$ попарно ортогональны.

С учётом требования ограниченности $\|f\|_{\mathcal{H}}$, а также в связи с тем фактом, что все функции, с которыми нам предстоит работать, будут достаточно гладкими, мы отбросим всю вторую сумму, что равносильно работе в пространстве аппроксимации. Так как нулевой уровень можно выбирать произвольным образом, это не ограничивает общность рассуждений.

Более подробно применение аппарата вейвлет-анализа в статистическом моделировании описано в [10].

Так как мы работаем на ограниченном интервале $[0, T]$, суммы по индексу k будут конечными. Осложняет задачу то, что наше пространство \mathcal{H} — не $L_2[0, T]$, а $W_2^1[0, T]$, поэтому функции φ_k не будут в нём ортогональными. Так как в нашу задачу входит исключительно получение приемлемого для практического применения разложения, и мы не собираемся в полной мере использовать аппарат вейвлет-анализа, для

решения этой проблемы подойдёт ортогонализация по Грамму-Шмидту. Она, разумеется, испортит как пространственную, так и частотную локализацию базисных функций, но в рамках данной задачи это не принципиально.

Таким образом, мы имеем разложение

$$\xi = \sum_{j=1}^N c_j e_j,$$

где $\{e_j\}_{j=1}^N$ — ортонормированный базис в пространстве V_0 “достаточно гладких” функций. В частности, норма в \mathcal{H} этих функций ограничена, так как по построению модели ограничена их норма в L_∞ , а в конечномерном пространстве, которым является V_0 , все нормы эквивалентны.

Итак, мы положим $f_t = \sum_{j=1}^N c_j e_j$ и перепишем уравнение динамики покоординатно (c — вектор координат проекции f_t на V_0).

$$dc = ((A + D)c + b) dt + \sum_{s=1}^N \sigma^s d\beta^s = ((A + D)c + b) dt + \Sigma dW.$$

Оно имеет явное решение

$$c \sim N \left(e^{t(A+D)} c^0 + \int_0^t e^{(t-u)(A+D)} b \, du, \int_0^t e^{(t-u)(A+D)} BB^T \left(e^{(t-u)(A+D)} \right)^T \, du \right).$$

$$dc = ((A_\sigma + D)c + b) ds + \Sigma dW,$$

где

$$A_\sigma = D + \sum_{s=1}^N A_4(\sigma^s, \sigma^s, \cdot, \cdot),$$

D — матрица оператора дифференцирования, а A_4 — 4-линейная форма:

$$A_4(i, j, l, k) = \langle e_k, e_i I(f_0 e_j e_l) \rangle_{\mathcal{H}}, \quad (I f)(x) = \int_0^x f(\tau) d\tau,$$

$$b = \sum_{s=1}^N A_4(\sigma^s, \sigma^s, 1 - \xi_0, \cdot) - \frac{1}{2} \sum_{s=1}^N A_3(\sigma^s, \sigma^s, \cdot) - \Sigma \gamma,$$

$$A_3(i, j, k) = \langle e_k, e_i e_j \rangle.$$

При этом наблюдения происходят в произвольные (но неслучайные) моменты времени и описываются случайными величинами P_k :

$$P_k \sim N \left(\sum_{s=0}^n F_{s,k} \exp \left[- \int_0^{\tau_s} \exp \{ \xi_t \} (x) dx \right], w_k \right).$$

Кроме того, при каждом наблюдении правдоподобность того, что для определения справедливой цены при сделке была использована кривая, соответствующая набору параметров c , составляет $\alpha \langle c, DDc \rangle$, где $DD(i, j) = \int_0^T \exp(\xi_0(x)) e'_i(x) e'_j(x) dx$.

Наша модель очень проста с одной стороны, так как уравнение динамики фазовой переменной ξ линейное, но с другой стороны слишком сложна, так как наблюдается не непосредственно значение процесса, а весьма сложная нелинейная функция от него. Если по формуле Ито выписать динамику наблюдаемых величин, чтобы свести задачу к обычно рассматриваемой задаче оценки параметров диффузии по частично наблюдаемым значениям, то задача сразу станет достаточно сложной.

Ближайшая по порядку сложности подобная модель, параметры которой оценивались в литературе, описана в [11], но даже она гораздо проще: во первых, она дискретная, а во-вторых, там наблюдаются значения непосредственно процесса (пусть даже его части).

3 Обзор методов оценки параметров

Оценка параметров любой модели — нетривиальная задача. Причём её сложность непропорционально возрастает со сложностью модели. Ниже дан краткий обзор различных методов, применяющихся для оценки параметров стохастических дифференциальных уравнений. Этот обзор приводится исключительно для законченности изложения и не претендует на полноту. Подробный обзор можно найти в работах [12], [13] и [14]. Все утверждения, делаемые здесь без явного указания источника, взяты из одной из вышеперечисленных работ.

3.1 Краткий обзор методов оценки

На вершине иерархии методов находится оценка максимального правдоподобия, традиционно используемая. Она заключается в минимизации минус логарифма функции правдоподобия (здесь и далее в этом пункте θ — вектор неизвестных параметров):

$$-\ln \mathcal{L}(\theta) = -\ln p(\xi^0, \xi^1, \dots, \xi^n | \theta),$$

где $p(\{\xi^i\} | \theta)$ — плотность вероятности наблюдать данные значения ξ^i при фиксированном параметре θ . Если процесс является марковским (как это обычно бывает), то формула упрощается:

$$-\ln \mathcal{L}(\theta) = -\ln p(\xi^0 | \theta) - \sum_{i=0}^{n-1} p(\xi^{i+1} | \xi^i, \theta),$$

где $p(\xi^{i+1} | \xi^i, \theta)$ — плотность переходной вероятности при известном состоянии ξ^i и известных параметрах θ . К сожалению, аналитическое выражение для неё доступно крайне редко, что привело к разработке в течение 90-х годов XX века широкого спектра методов, не требующих

аналитического выражения для переходной плотности.

Численное решение уравнения Фоккера-Планка

Известно, что плотность переходной вероятности удовлетворяет уравнению Фоккера-Планка:

$$\frac{\partial f}{\partial t} = \left[- \sum_{i=1}^N \frac{\partial}{\partial x_i} D_i^1(x_1, \dots, x_N) + \sum_{i=1}^N \sum_{j=1}^N \frac{\partial^2}{\partial x_i \partial x_j} D_{ij}^2(x_1, \dots, x_N) \right] f,$$

где

$$D_i^1(x, t) = \mu_i(x, t), \quad D_{ij}^2(x, t) = \frac{1}{2} \sum_k \sigma_{ik}(x, t) \sigma_{jk}(x, t).$$

с подходящими начальными и граничными условиями. В простейших случаях оно решается аналитически, однако в большинстве приложений необходимо численное его решение. С подробностями можно ознакомиться в работах [15], [16] и [17].

Дискретная оценка максимального правдоподобия

Главная идея этого метода — приближение переходной плотности выражением, полученным из дискретизации стохастического дифференциального уравнения, описывающего динамику. Простейшим примером является разложение Эйлера-Маруямы:

$$X_{i+1} = X_k + \mu(X_i; \theta)\Delta + \sigma(X_i; \theta)\epsilon_i,$$

где $\epsilon_i \sim N(0, \Delta)$ — независимые нормально распределённые случайные величины (шум).

Несмотря на то, что оценки, полученные по этому методу, оказываются смещёнными (см. обсуждение этого смещения в [18]). Тем не менее, в силу простоты, а также того, что зачастую этот подход даёт

оценки в аналитическом виде, без необходимости численных процедур минимизации, дискретные оценки максимального правдоподобия широко используются на практике (см., например, [19] или хороший обзор в [20]). В работе [21] предложено разложение коэффициента сноса в степенной ряд первого члена, что соответствует нашему выбранному способу линеаризации.

Разложение по многочленам Эрмита

В работах [22, 23, 13] разработан метод аналитического приближения переходной плотности путём её разложения в ряд по многочленам Эрмита в окрестности стандартной нормальной плотности. Более того, показано, как аналитически перегруппировать слагаемые по степеням Δt . Автор утверждает, что разложения до $(\Delta t)^2$ достаточно в абсолютном большинстве приложений. В применении к многомерным диффузиям этот метод требует возможности привести её к единичной матрице диффузии. В противном случае решение даётся в виде двойного ряда (см. [13]).

Выборочная оценка максимального правдоподобия

Этот метод был разработан в [24] и [25], с вариацией в [26]. Суть метода — в приближении переходной плотности её выборочным аналогом. Для этого численно моделируется большое количество траекторий системы, после чего переходная плотность восстанавливается, например, путём ядерных оценок. Часто этот метод комбинируют с выборкой по значимости (importance sampling) — см., например, [27]. Естественным обобщением этого метода выступает

Метод Монте-Карло для марковских цепей (Markov Chain Monte Carlo, МСМС)

Основная идея этого метода — в доразбитии интервала наблюдения данных $[t_i, t_{i+1}]$ на M равномерных меньших отрезков путём введения $(M - 1)$ ненаблюдаемых величин в промежуточные моменты времени. Затем алгоритм функционирует следующим образом:

1. Выбираются ненаблюдаемые значения X_u^{i+1} из распределения $P(X_u|X_o, \theta^i)$ (X_o — наблюдаемые значения процесса).
2. Выбираются значения параметров θ^{i+1} из $P(\theta|X_o, X_u^i)$.

При довольно слабых условиях регулярности марковская цепь, описанная выше, имеет предельное стационарное распределение $P(X_u, \theta|X_o)$, которое можно легко преобразовать в интересующее нас $P(\theta|X_o)$. Литература на эту тему весьма обширна: обзор методов Монте-Карло можно найти в книгах [28, 29, 30]. Применение этих методов к задачам оценки параметров диффузии — в [31, 32, 33, 34, 35, 36, 37, 38]. Основным недостатком методов Монте-Карло — чрезвычайная ресурсоёмкость по построению. Тем не менее, эти методы широко используются из-за своей универсальности, хорошей масштабируемости, способности учитывать ненаблюдаемые переменные, ошибки наблюдения и т.п., а также возможности применения параллельных вычислений для ускорения процесса.

Обобщённый метод моментов

Впервые этот метод был предложен в [39]. Его ядро — в указании некоторого количества условий $\psi_k(X; \theta)$, таких, что $\mathbb{E}[\psi_k(X_i; \theta^*)] = 0$, где θ^* — истинное значение вектора параметров. Пусть $\Psi_{k,i}(X; \theta) = \psi_k(X_i|\theta)$. Тогда оценка вектора θ получается путём минимизации функционала $J(\theta) = \mathbb{E}[\Psi(X, \theta)]^T \Sigma \mathbb{E}[\Psi(X; \theta)]$, где Σ — положительно определённая

матрица взвешивания. В [40] предложены различные способы построения этой матрицы. Способы получения подходящих функций ψ_k описаны в [41, 42]. Достоинствами этого метода являются простота реализации, скорость и общность (применимость к широкому классу диффузий), однако он не учитывает всю доступную информацию, а также страдает от погрешности дискретизации (используемой для получения функций ψ_k).

Непрямая оценка (Indirect Estimation)

В этих методах параметры оцениваются не напрямую через (обычно недоступную) плотность переходной вероятности, а косвенно — при помощи вспомогательной модели, которая с одной стороны легко оценивается методом максимального правдоподобия, а с другой даёт разумное приближение к истинной функции правдоподобия. Однако мы учитываем, что вспомогательная модель указана неверно, и поэтому оценки $\hat{\theta}$ связаны с истинными значениями θ^* посредством функции $\hat{\theta} = \varphi(\theta^*)$, которая определяется путём численного моделирования. К этому классу относятся методы, предложенные в [43] и [44] (этот метод известен под названием “эффективный метод моментов”, Efficient method of moments, ЕММ). Производительность методов и качество получаемых оценок сильно зависят от вспомогательной модели. Если модель выбрана хорошо, то качественные результаты получаются даже при небольших размерах выборки. В противном случае производительность существенно падает [26].

Метод характеристических функций

В работах [45, 46, 47] была предпринята попытка построения оценок на основе решения уравнения Фоккера-Планка в терминах не плотности

переходной вероятности, а её характеристической функции:

$$\tilde{f}(p, t) = \int e^{-px} f(x, t) dx.$$

Тем не менее, известность этому методу придали не эти исследования (т.к. остаётся необходимость обратного преобразования в плотность вероятности), а возможность получения большого количества условий для метода моментов. К этому же классу стоит отнести условия, полученные из инфинитезимального оператора в работе [48].

Метод маргинальной плотности

В [49] предложено оценивать θ путём сравнения маргинальной плотности стохастического дифференциального уравнения и её непараметрической оценкой по наблюдаемым данным. Недостатками этого метода является то, что реальные данные, скорее всего, чрезвычайно сильно коррелированы, тогда как для оценки методом ядер предполагается, что наблюдения независимы и одинаково распределены. В [50] подмечено, что в таком случае качество оценки неудовлетворительно.

3.2 Оценка параметров модели

В силу специфики задачи мы выберем подход [21] для построения функции правдоподобия и метод Монте-Карло для марковских цепей для собственно оценки. Эта линеаризация, уже проведённая раньше для оценки значения кривой доходности, пригодится и для оценки параметров, так как позволит использовать вместо сложных и вычислительно трудоёмких методов, описанных выше, оценку максимального правдоподобия. Отметим, что с точки зрения точности самым рациональным методом представляется оценка, основанная на разложении плотности переходной вероятности по многочленам Эрмита [13]: так как дина-

мика почти линейная, отклонения от нормальной плотности должны быть небольшими и легко захватываемыми даже небольшим количеством членов ряда.

4 Описание метода Монте-Карло для марковских цепей

Опишем кратко алгоритм Метрополиса-Гастингса, используемый для выборки случайных величин со сложным распределением. Этот алгоритм был предложен практически одновременно с методом Монте-Карло вычисления определённого интеграла в работах [51, 52], затем развит в [53]. Подробное описание можно найти в работах [54, 55].

Предположим, что нам нужно выбрать случайный элемент из некоторого абсолютно непрерывного распределения $p(\theta)$, где $p(\theta) = f(\theta)/K$, а нормализующая константа K неизвестна или очень трудна для вычисления. Алгоритм Метрополиса, предложенный в 1953 году [51, 52], генерирует последовательность элементов, выбранных из этого распределения, следующим образом:

1. В качестве начального значения θ_0 выбирается любое значение с $f(\theta_0) > 0$.
2. Используя текущее значение θ , выбрать вспомогательную точку θ^* из некоторого вспомогательного распределения $q(\theta_1, \theta_2)$, которое есть вероятность получить следующую точку θ_2 при условии, что текущая точка — θ_1 . Это распределение называют “proposal”, “jumping” или “candidate-generating”, русскоязычная терминология не сложилась. Единственное ограничение на плотность q — её симметричность: $q(\theta_1, \theta_2) = q(\theta_2, \theta_1)$.

3. Для данного значения θ^* вычисляется отношение плотностей распределения p во вспомогательной (θ^*) и текущей (θ_{t-1}) точках.

$$\alpha = \frac{p(\theta^*)}{p(\theta_{t-1})} = \frac{f(\theta^*)}{f(\theta_{t-1})}.$$

Так как мы рассматриваем отношение значений одной плотности в двух точках, нормализующие константы сокращаются.

4. Если предложенное значение более вероятно ($\alpha > 1$), то принимаем его как новое значение ($\theta_t = \theta^*$) и возвращаемся к шагу 2. Иначе, если оно уменьшает плотность вероятности ($\alpha < 1$), то мы принимаем его с вероятностью α , а иначе отклоняем его ($\theta_t = \theta_{t-1}$). После этого возвращаемся к шагу 2.

Эта процедура генерирует марковскую цепь $\theta_0, \theta_1, \dots, \theta_k, \dots$, так как вероятности перехода от θ_t к θ_{t+1} зависят только от θ_t , а не от $\theta_0, \dots, \theta_{t-1}$. В 1970 году Гастингс [53] обобщил этот метод на произвольную вспомогательную функцию $q(\theta_1, \theta_2) = P(\theta_1 \rightarrow \theta_2)$ путём модификации вероятности принятия нового значения:

$$\alpha = \min \left(\frac{f(\theta^*)q(\theta^*, \theta_{t-1})}{f(\theta_{t-1})q(\theta_{t-1}, \theta^*)}, 1 \right).$$

Этот алгоритм называют алгоритмом Метрополиса-Гастингса. При достаточно мягких технических условиях получаемая марковская цепь имеет стационарное распределение, совпадающее с распределением p . Таким образом, для получения выборки из p необходимо, запустив цепочку из любого начального значения, откинуть некоторое количество начальных значений, чтобы цепочка успела сойтись к стационарному распределению, а после этого значения θ_t можно рассматривать как независимые реализации случайной величины с распределением p .

Неудачный выбор начального значения или вспомогательного рас-

пределаения может очень сильно ухудшить производительность алгоритма.

Вариант, предложенный в [56], известный как Gibbs sampler, заключается в том, что выборка производится отдельно по каждому измерению: так, если требуется распределение $p(x, y)$, то последовательно выбираются значения

$$x_t \sim p(x|y = y_{t-1}),$$

$$y_t \sim p(y|x = x_t).$$

Повторение этого процесса много раз даёт марковскую цепь, сходящуюся к совместному распределению $p(x, y)$. Мощь этого метода - в том, что выборку на каждом этапе можно проводить при помощи алгоритма Метрополиса-Гастингса. Полученную схему обычно называют Metropolis-within-Gibbs.

Мы будем использовать именно этот вариант. Для фиксированного A мы будем генерировать марковскую цепь, выбирая случайные элементы сначала с распределением $P(c|\sigma, \nu)$, затем - с $P(\nu|\sigma, c)$ и потом с $P(\sigma|c, \nu)$. С более подробным описанием теоретических основ метода и практических аспектов его реализации можно ознакомиться в литературе, перечисленной выше при кратком обзоре методов оценки параметров.

Так как ни один из упомянутых шагов не является тривиальным, ниже мы рассмотрим каждый из них в отдельности. Для практической реализации необходимы выражения для производных функции правдоподобия по параметрам c, σ, ν . Эти выражения получаются простым дифференцированием, но весьма громоздки и поэтому вынесены в отдельный пункт, так как лишь загромождают текст, затрудняя понимание алгоритма.

5 Описание шагов алгоритма

5.1 Начальное приближение

Несложно получить вид функции правдоподобия (суммирование идёт по наблюдениям, для упрощения записи мы пишем индекс i вместо t_i для обозначения того, что данные относятся к i -му наблюдению, в момент t_i):

$$\begin{aligned}
 -\ln p(\{c_i\}, \{\sigma^s\}, \gamma) = & \frac{1}{2} \sum_i \{ \langle c_i - m_i, B_i^{-1}(c_i - m_i) \rangle + \\
 & + \sum_{k=1}^{n_k} w_k \left(\sum_{s=0}^n F_{s,k} \exp \left[- \int_0^{\tau_s} f_c(x) dx \right] - P_i^k \right)^2 + \\
 & + \alpha \langle c_i, DDc_i \rangle \}, \quad (2)
 \end{aligned}$$

где

$$\begin{aligned}
 m_i &= e^{(t_i - t_{i-1})A} c_{i-1} + \int_0^{t_i - t_{i-1}} e^{(t_i - t_{i-1} - u)A} b(\sigma^s, \gamma) du \\
 B_i &= \int_0^{t_i - t_{i-1}} e^{(t_i - t_{i-1} - u)A} \Sigma \Sigma^T \left(e^{(t_i - t_{i-1} - u)A} \right)^T du \\
 b(\{\sigma^s\}, \gamma) &= \sum_{s=1}^N A_4(\sigma^s, \sigma^s, 1 - \xi_0, :) - \frac{1}{2} \sum_{s=1}^N A_k(\sigma^s, \sigma^s, :) - \sum_{s=1}^N \gamma_s \sigma^s.
 \end{aligned}$$

(Здесь знаком $:$ обозначено измерение, остающееся после свёртки). Зависимость очень сложная, особенно сложна зависимость первого слагаемого от $\{\sigma^s\}$ через A . В то же время, вычисления показывают, что оно вносит наименьший вклад в правдоподобность. Поэтому мы построим алгоритм итеративно: оценка $\{\sigma^s\}$ будет проводиться для фиксирован-

ного значения A , затем пересчёт A и новая оценка $\{\sigma^s\}$. Однако на практике оказывается, что при разумном начальном приближении достаточно всего двух итераций.

При этом, вместо γ мы будем оценивать величину $\nu = \Sigma\gamma$. В [57] подробно рассматривается, почему такая параметризация будет улучшать качество оценки.

Сперва мы построим итерационный алгоритм нахождения начального приближения:

1. На первом шаге мы добавим слагаемое εI к выражению для B_i :

$$B_i = \varepsilon I + \int_0^{t_i - t_{i-1}} e^{(t_i - t_{i-1} - u)A} \Sigma \Sigma^T \left(e^{(t_i - t_{i-1} - u)A} \right)^T du, \quad (3)$$

где ε имеет порядок 10^{-4} , и проведём оценку $\{c_i\}$ методом максимального правдоподобия при фиксированных σ^s .

2. Затем, убрав это слагаемое, оценим $\{\sigma^s\}$ и γ при фиксированных $\{c_i\}$.
3. Процесс повторяется, но на каждом шаге значение ε уменьшается. Обычно достаточно 3–4 итераций, за время которых ε уменьшается от 10^{-4} до 10^{-6} .

Теперь с найденными значениями $\{\sigma^s\}$ оценим методом максимального правдоподобия значения $\{c_i\}$, а затем, при фиксированных $\{\sigma^s\}, \{c_i\}$, — ν . Начальное приближение получено.

Замечание. В случае, если матрица Σ не квадратная, а конкретно, если размерность пространства больше размерности шума, матрица B_i будет иметь неполный ранг. Для устойчивости вычислений слагаемое εI следует добавлять в качестве регуляризации и в остальных расчётах. Численный эксперимент показывает разумными значения 10^{-6} – 10^{-7} .

5.2 Выборка $\{c_i\}$

При фиксированных $\{\sigma^s\}$, ν плотность условного распределения $\{c_i\}$ имеет вид

$$-\ln p(\{c_i\}|\{\sigma^s\}, \nu) = \frac{1}{2} \sum_i \langle c_i - m_i, B_i^{-1}(c_i - m_i) \rangle + J_2^i(c_i) + \alpha \langle c_i, DDc_i \rangle,$$

$$J_2^i(c) = \frac{1}{2} \sum_{k=1}^{n_k} w_k \left(\sum_{s=0}^n F_{s,k} \exp \left[- \int_0^{\tau_s} r_c(\tau) d\tau \right] - P_i^k \right)^2,$$

причём первое слагаемое для $i = 0$ (начального значения) отсутствует.

Чтобы уменьшить размерность $\{c_i\}$, мы разобьём c_i на группы по i : $c_{i_0} \dots c_{i_1}$, $c_{i_1} \dots c_{i_2}$, ..., $c_{i_{K-1}} \dots c_{i_K}$. Для уменьшения автокорреляции и ускорения перемешивания границы групп на каждой итерации определяются случайно со средним размером группы в 6 измерений (при 20-30 членах в разложении по базису это отвечает размерности порядка 100-200). Для генерации случайной величины с нужным распределением мы используем алгоритм Метрополиса-Гастингса со вспомогательной нормальной плотностью, центрированной в максимуме плотности и с матрицей ковариаций равной минус обратной матрице вторых производных логарифма плотности (т.к. два слагаемых из трёх – квадратичные). Таким образом, выборка $\{c_i\}$ сводится к многократной выборке из многомерного (но не слишком) нормального распределения с параметрами, определяемыми процедурой минимизации.

5.3 Выборка $\{\sigma^s\}$

Оценка $\{\sigma^s\}$ а priori трудна, т.к. параметры ковариации плохо оцениваются. К тому же, их много. Для идентифицируемости мы наложим ограничение $\sigma_j^s = 0$ при $j < s$. Переход к единичной матрице ковариаций, как описано в [13] в данном случае нецелесообразен, хотя и возмо-

жен. Выборку также нельзя осуществлять по алгоритму Метрополиса-Гастингса, так как найти удовлетворительную вспомогательную плотность не представляется возможным: подход, описанный в [11], не работает из-за гораздо более сложной структуры модели.

Вместо этого мы используем алгоритм, описанный в [58]. Суть его в том, что мы производим выборку случайной точки вдоль некоторого направления так, чтобы результирующее распределение совпадало с целевым. Подробно эти вопросы разобраны в [58, 28, 59, 60], один современный метод описан в [61] здесь мы остановимся на более практических аспектах. Для выбора направления мы используем два алгоритма, случайно выбирая один из них.

Первый:

1. Стартуем с текущей точки.
2. Получаем градиент логарифма плотности и строим матрицу ковариаций: дисперсия вдоль градиента в 15 раз больше, чем вдоль остальных направлений.
3. Генерируем нормально распределённую точку с центром в текущей точке и только что посчитанной матрицей ковариаций.
4. Направление - из текущей точки в новую.

Второй:

1. Стартуем со случайной точки.
2. Из неё делаем несколько шагов градиентной максимизации логарифма плотности, получаем точку x_1 .
3. Направление - из точки x_1 - в текущую.

Для выборки точки с нужным распределением вдоль заданной линии мы используем алгоритм ARS ([62]), а точнее — его обобщение ARMS ([63]).

5.4 Выборка ν

Заметим, что в силу линейности уравнения динамики можно выделить всю зависимость от ν и построить плотность следующим образом:

$$-\ln p(\nu|\{\sigma^s\}, \{c_i\}) = -\ln p(\{c_i - \Delta c_i(\nu)\}|\{\sigma^s\}, \nu = 0),$$

где $\Delta c_i(\nu)_i = e^{t_i(A+D)}\nu$. Теперь выборка ν производится тем же способом, что и $\{c_i\}$, но без разбиения на блоки, так как размерность задачи и так достаточно мала (ν не зависит от времени t).

5.5 Выражения для производных

Логарифм функции правдоподобия имеет следующий вид:

$$\begin{aligned} -\ln p(\{c_i\}, \{\sigma^s\}, \nu) = & \frac{1}{2} \sum_i \{ \langle c_i - m_i, B_i^{-1}(c_i - m_i) \rangle + \\ & + \sum_{k=1}^{N_i} w_k \left(\sum_{s=0}^n F_{s,k} \exp \left[- \int_0^{\tau_s} r_{c_i}(\tau) d\tau \right] - P_i^k \right)^2 + \\ & + \alpha \langle c_i, DDc_i \rangle \}, \end{aligned}$$

где

$$\begin{aligned}
m_i &= e^{(t_i - t_{i-1})A_\sigma} c_{i-1} + \int_0^{t_i - t_{i-1}} e^{(t_i - t_{i-1} - s)A_\sigma} b \, ds; \\
B_i &= \int_0^{t_i - t_{i-1}} e^{(t_i - t_{i-1} - s)A_\sigma} \Sigma \Sigma^T \left(e^{(t_i - t_{i-1} - s)A_\sigma} \right)^T ds; \\
b &= \sum_{s=1}^N A_4(\sigma^s, \sigma^s, 1 - \xi_0, \cdot) - \frac{1}{2} \sum_{s=1}^N A_3(\sigma^s, \sigma^s, \cdot) - \Sigma \nu; \\
A_\sigma &= D + \sum_{s=1}^N A_4(\sigma^s, \sigma^s, \cdot, \cdot); \\
A_4(i, j, l, k) &= \langle e_k, e_i I(r_0 e_j e_l) \rangle_{\mathcal{H}}, \quad (If)(x) = \int_0^x f(\tau) \, d\tau; \\
A_3(i, j, k) &= \langle e_k, e_i e_j \rangle; \\
DD_{i,j} &= \int_0^T r_0(s) e'_i(s) e'_j(s) \, ds; \\
D_{i,j} &= \langle e_i, e'_j \rangle_{\mathcal{H}}; \\
r_c(x) &= \exp(\langle c, e(x) \rangle),
\end{aligned}$$

где $e(x)$ — вектор из значений базисных функций в точке x .

Сперва рассмотрим производные по c_i .

$$\begin{aligned}
\frac{\partial(-\ln p(c, \Sigma, \nu))}{\partial c_i} &= (c_i - m_i)^T (B_i^{-1})^T - (c_{i+1} - m_{i+1})^T (B_{i+1}^{-1})^T e^{(t_{i+1} - t_i)A_\sigma} + \\
&\quad \sum_{k=1}^N w_k \left(\sum_{s=0}^n F_{s,k} \exp \left[- \int_0^{\tau_s} r_{c_i}(\tau) \, d\tau \right] - P_i^k \right) \cdot \\
&\quad \sum_{s=0}^n F_{s,k} \exp \left[- \int_0^{\tau_s} r_{c_i}(\tau) \, d\tau \right] \left(- \int_0^{\tau_s} \exp(\langle c_i, e(\tau) \rangle) e(\tau)^T \, d\tau \right) + \\
&\quad \alpha c_i^T (DD)^T.
\end{aligned}$$

Для вторых производных имеем

$$\frac{\partial^2(-\ln p(c, \sigma, \nu))}{\partial c_i \partial c_i} = (B_i^{-1}) + \sum_{k=1}^N w_k [(q_k(c_i) - P_i^k) (q_k(c_i) E(c_i) E(c_i)^T + q_k(c_i)(-E_2(c_i))) + q_k(c_i)^2 E(c_i) E(c_i)^T] + \alpha DD;$$

$$q_k(c) = \sum_{s=0}^n F_{s,k} \exp \left[- \int_0^{\tau_s} r_c(\tau) d\tau \right];$$

$$E(c) = \int_0^{\tau_s} \exp(\langle c, e(\tau) \rangle) e(\tau) d\tau;$$

$$E_2(c) = \int_0^{\tau_s} \exp(\langle c, e(\tau) \rangle) e(\tau) e^T(\tau) d\tau; \quad \frac{\partial^2(-\ln p(c, \sigma, \nu))}{\partial c_i \partial c_{i+1}} = e^{(t_i - t_{i-1})A_\sigma} B_{i+1}^{-1},$$

все остальные вторые производные равны 0. Очевидно, что при i , равном количеству наблюдений, слагаемое с c_{i+1} просто будет отсутствовать.

Теперь рассмотрим производные по ν .

$$\frac{\partial(-\ln p(c, \Sigma, \nu))}{\partial \nu} = - \sum_i (c_{i+1} - m_{i+1})^T (B_{i+1}^{-1})^T \left(\int_0^{t_{i+1} - t_i} e^{(t_{i+1} - t_i - s)A_\sigma} ds \right)^T \Sigma^T.$$

$$\frac{\partial^2(-\ln p(c, \Sigma, \nu))}{\partial \nu^2} = - \sum_i \Sigma \left(\int_0^{t_{i+1} - t_i} e^{(t_{i+1} - t_i - s)A_\sigma} ds \right) (B_{i+1}^{-1})^T \left(\int_0^{t_{i+1} - t_i} e^{(t_{i+1} - t_i - s)A_\sigma} ds \right)^T \Sigma^T.$$

Следующий этап — производные по Σ . Первая производная

$\frac{\partial(-\ln p(c, \Sigma, \nu))}{\partial \Sigma}$ — матрица того же размера, что и Σ , однако вторая производная — 4-линейная форма. Для компьютерной реализации мы представим матрицу Σ в виде одного вектора. Тогда вторая производная будет матрицей размера $N'N \times N'N$. Для того, чтобы сделать запись воспринимаемой, мы будем использовать тензорную нотацию (с верхними и нижними индексами). Так, матрица будет иметь индексы Σ_i^j , вектор-столбец из тех же элементов (в любой заранее зафиксированной последовательности) — $\Sigma_{i,j}$, а вектор-строка — $\Sigma^{i,j}$. Одинаковый индекс сверху и снизу означает перемножение и суммирование. Операции опускания и поднятия индексов соответствуют физическим изменениям в структуре матрицы. Так, опускание индекса j в выражении A_i^j означает выписывание всех элементов матрицы A в один столбец. В выражении $\frac{\partial x}{\partial y}$ индексы x становятся нижними, а индексы y — верхними вне зависимости от того, какими они были изначально. Так, $\frac{\partial A_{i,j}^k}{\partial X_l^m} = \left(\frac{\partial A}{\partial X}\right)_{i,j,k}{}^{l,m}$.

Приведём несколько вспомогательных соотношений, которые легко устанавливаются вычислением частной производной по одному конкретному элементу матрицы.

$$\begin{aligned} \frac{\partial \ln \det X}{\partial X_{i,j}} &= (X^{-1})^{j,i}; \\ \frac{\partial (AXB)_{a,b}}{\partial X_{i,j}} &= \frac{\partial A_a^k X_k^l B_l^b}{\partial X_{i,j}} = A_a^i B_j^b = A_a^i (B^T)_b^j; \\ \frac{\partial (XX^T)_{a,b}}{\partial X_{i,j}} &= \delta_a^i X_b^j + X_a^j \delta_b^i; \\ \frac{\partial (X^{-1})_{a,b}}{\partial X_{i,j}} &= -(X^{-1})_a^i ((X^{-1})^T)_b^j; \\ \frac{\partial^2 X^{-1}_{a,b}}{\partial X_{i,j} \partial X_{k,l}} &= (X^{-1})_a^k ((X^{-1})^T)_b^j (X^{-1})^{l,i} + (X^{-1})_a^i ((X^{-1})^T)_b^l (X^{-1})^{j,k}; \end{aligned}$$

Теперь вычислим производные величины

$$J_{(s)} = \frac{1}{2} \langle c_{(s)} - m_{(s)}, B^{-1}(c_{(s)} - m_{(s)}) \rangle.$$

Производные всей функции правдоподобия получатся суммированием по s . Напомним, что величину A_σ мы считаем не зависящей от σ на этом этапе.

$$\begin{aligned} \frac{\partial J_{(s)}}{\partial \Sigma_{i,j}} &= \frac{\partial J}{\partial (B_{(s)}^{-1})_{a,b}} \frac{\partial (B_{(s)}^{-1})_{a,b}}{\partial B_{(s),k,l}} \frac{\partial B_{(s),k,l}}{\partial \Sigma_{i,j}} + \frac{\partial J_{(s)}}{\partial m_{(s),f}} \frac{\partial m_{(s),f}}{\partial \Sigma_{i,j}}; \\ \frac{\partial J_{(s)}}{\partial (B^{-1})_{a,b}} &= \frac{1}{2} (c - m)^a (c - m)^b; \\ \frac{\partial (B_{(s)}^{-1})_{a,b}}{\partial B_{(s),k,l}} &= -(B_{(s)}^{-1})_a^k (B_{(s)}^{-1})_{(s)l}^b. \end{aligned}$$

Заметим, что зависимость $B_{(s)}$ от Σ — квадратичная. Коэффициенты этой зависимости могут быть явно вычислены:

$$\Xi_{k,l}^{i,j,u,v} = \left[\int_0^{t_i - t_{i-1}} e^{(t_i - t_{i-1} - s)A_\sigma} \Theta(i, j) \Theta(u, v)^T \left(e^{(t_i - t_{i-1} - s)A_\sigma} \right)^T ds \right]_{k,l},$$

где $\Theta(i, j)$ — матрица с единицей в позиции (i, j) и нулями в остальных позициях.

$$\begin{aligned} B_{k,l}^{(s)} &= \Xi_{k,l}^{i,j,u,v} \Sigma_{i,j} \Sigma_{u,v}; \\ \frac{\partial B_{(s),k,l}}{\partial \Sigma_{i,j}} &= 2 \Xi_{k,l}^{i,j,u,v} \Sigma_{u,v}; \\ \frac{\partial J_{(s)}}{\partial m_f} &= -(B_{(s)}^{-1})_f^g (c - m)_g; \\ \frac{\partial m_f}{\partial \Sigma_{i,j}} &= \delta_i^f \nu_j. \end{aligned}$$

Вторые производные вычисляются аналогично:

$$\begin{aligned}
\frac{\partial^2 J_{(s)}}{\partial \Sigma_{i,j} \partial \Sigma_{u,v}} &= \frac{\partial J}{\partial (B_{(s)}^{-1})_{a,b}} \frac{\partial^2 (B_{(s)}^{-1})_{a,b}}{\partial B_{(s),k,l} \partial B_{(s),k',l'}} \frac{\partial B_{(s),k,l}}{\partial \Sigma_{i,j}} \frac{\partial B_{(s),k',l'}}{\partial \Sigma_{u,v}} + \\
&\quad \frac{\partial J}{\partial (B_{(s)}^{-1})_{a,b}} \frac{\partial (B_{(s)}^{-1})_{a,b}}{\partial B_{(s),k,l}} \frac{\partial^2 B_{(s),k,l}}{\partial \Sigma_{i,j} \partial \Sigma_{i,j}} + \\
&\quad \frac{\partial^2 J_{(s)}}{\partial m_{(s),f} \partial m_{(s),f'}} \frac{\partial m_{(s),f}}{\partial \Sigma_{i,j}} \frac{\partial m_{(s),f'}}{\partial \Sigma_{u,v}} + \\
&\quad \frac{\partial^2 J_{(s)}}{\partial m_{(s),f} \partial (B_{(s)}^{-1})_{g,h}} \frac{\partial (B_{(s)}^{-1})_{g,h}}{\partial B_{x,y}} \frac{\partial B_{x,y}}{\partial \Sigma_{u,v}}; \\
\frac{\partial^2 J_{(s)}}{\partial m_{(s),f} \partial (B_{(s)}^{-1})_{g,h}} &= \delta^{f,g} (c - m)^h; \\
\frac{\partial^2 J_{(s)}}{\partial m_{(s),f} \partial m_{(s),f'}} &= (B_{(s)}^{-1})^{f,f'}.
\end{aligned}$$

Путём надлежащей перегруппировки сомножителей все эти выражения могут быть сведены к операциям умножения матриц $A_i^j B_j^k$ и изменения структуры данных (опускания/поднимания индексов), что позволяет эффективно реализовать вычисления на компьютере.

5.6 Фиктивные наблюдения

В случае, если наблюдения случаются слишком редко, для улучшения качества линейных приближений, которое зависит от величины интервалов между наблюдениями, можно вводить фиктивные наблюдения в промежуточные моменты времени. При использовании методов Монте-Карло для марковских цепей ненаблюдаемые значения в итоге будут проинтегрированы и не будут влиять на результат. Эта методика известна под названием “Data augmentation”; впервые появившись в ра-

боте [64], она приобрела современные очертания в [65]. Можно также отметить работы [66] и [67]. Более практически ориентированное описание можно найти в работах [68, 32, 27, 31, 55, 69].

6 Применение предварительных и параллельных вычислений

Алгоритм, описанный выше, требует большого количества вычислений. Для того, чтобы максимально уменьшить время работы, можно использовать предварительные и параллельные вычисления.

6.1 Предварительные вычисления

В процессе построения марковской цепи приходится проводить много однотипных вычислений. При надлежащей перегруппировке некоторую часть их часть можно провести заранее, один раз. Так, заранее можно вычислить A_4 и A_3 , записав их в таблицу: эти величины меняться не будут. Кроме того, согласно нашему алгоритму, в рамках одной итерации верхнего уровня остаются неизменными величины A и $\exp(tA)$, которые тоже можно вычислить заранее ($\exp(tA)$ – на сетке по t , которая должна позволять проводить численное интегрирование на отрезках от 0 до Δt_i). Помимо этого, очевидно, что

$$B_i(\Sigma) = \int_0^{t_i - t_{i-1}} e^{(t_i - t_{i-1} - u)A} \Sigma \Sigma^T \left(e^{(t_i - t_{i-1} - u)A} \right)^T du$$

является квадратичной формой от Σ ; матрицу этой квадратичной формы тоже можно вычислить один раз и далее использовать готовую. Список этим не исчерпывается, но остальное полностью аналогично.

6.2 Параллельные вычисления

Для того, чтобы эффективно использовать несколько процессоров при вычислениях, достаточно каждый из них загрузить построением независимой марковской цепи. После отбрасывания начального отрезка на разных узлах мы будем иметь выборки из одного и того же распределения, которые можно просто объединить. Это означает практически неограниченный потенциал распараллеливания алгоритма при условии, что у каждой цепи будет достаточно большой начальный отрезок, чтобы та пришла в стационарное состояние. Однако этим положительный эффект от параллелизации не исчерпывается: во втором алгоритме выборки вспомогательной точки для $\{\sigma^s\}$ вместо случайной точки можно взять текущую точку одной из параллельных марковских цепей (случайно выбранной). Таким образом мы используем идеи работы [59], что позволяет ещё улучшить перемешивание.

7 Иллюстрация

В качестве иллюстрации рассмотрим модельный пример.

Предположим, что на рынке торгуется 20 бескупонных облигаций со, равномерно распределённых по срочности на отрезке $[0, T]$. Предположим также, что их цена наблюдается с ошибкой порядка 2% от цены (это соответствует ошибке в доходности к погашению порядка 100 базисных пунктов, что есть нормальная ситуация). Модельные данные генерировались в предположении, что истинная кривая доходности имеет динамику, описываемую моделью с функциями σ^s , $s = 1, 2, 3$, приведёнными на рис. 1, и рыночными ценами риска, соответствующими после репараметризации функции ν , показанной на рис. 2. Рассматривался период в 200 измерений с промежутком в 200 минут.

Было сгенерировано 5 цепочек по 1500 элементов в каждой, первые

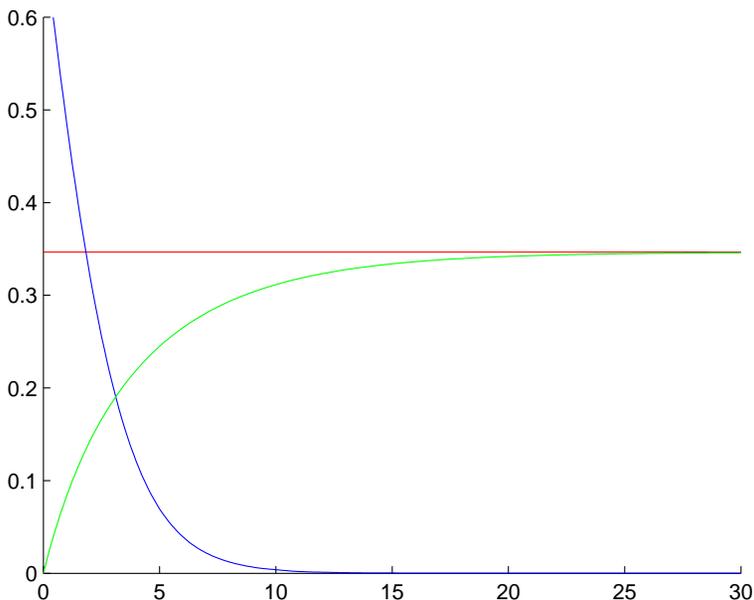


Рис. 1: Истинные значения $\sigma^s(\cdot)$.

500 элементов каждой были откинута. Результаты оценки σ приведены на рис. 3. Оценка ν приведена на рис. 4. На рис. 5 показана оценка собственно кривой доходности вместе с истинным её значением на 100-е измерение. На рис. 6 показана оценка на тот же самый момент при использовании приближённого алгоритма (в течение всех 100 измерений), а на рис. 7 и 8 — оценки мгновенной форвардной процентной ставки этими двумя алгоритмами.

Видно, что значение ν определилось достаточно хорошо, значения σ определились с точностью до линейной комбинации, что тоже достаточ-

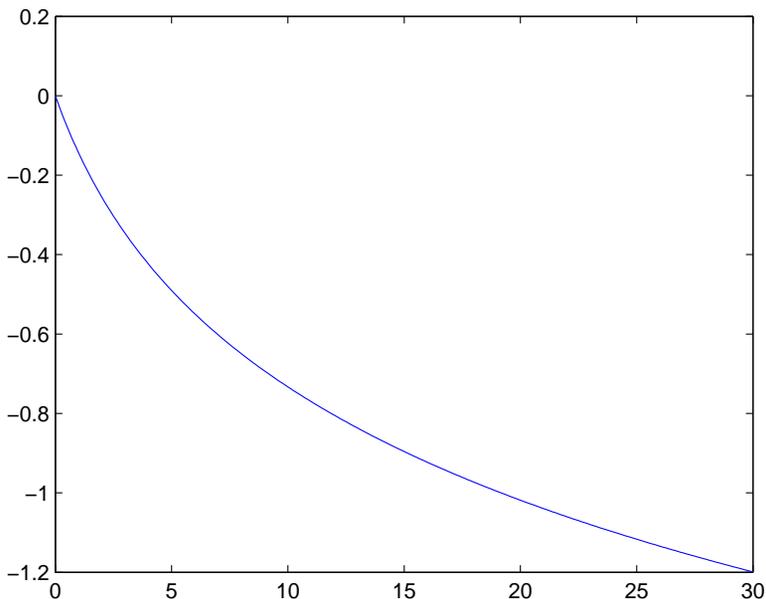


Рис. 2: Истинное значение $\nu(\cdot)$.

но хороший результат. Значение же ξ определилось с превосходной точностью (небольшая ошибка на коротком конце связана с большой ошибкой в наблюдении доходности одной из бумаг). Также видно, что столь же качественно определяются и мгновенные форвардные процентные ставки.

Для иллюстрации того, что используемый метод может быть использован для определения размерности броуновского движения, приведём результат оценки параметров для тех же исходных данных, но для 4 и 2 функций σ^s .

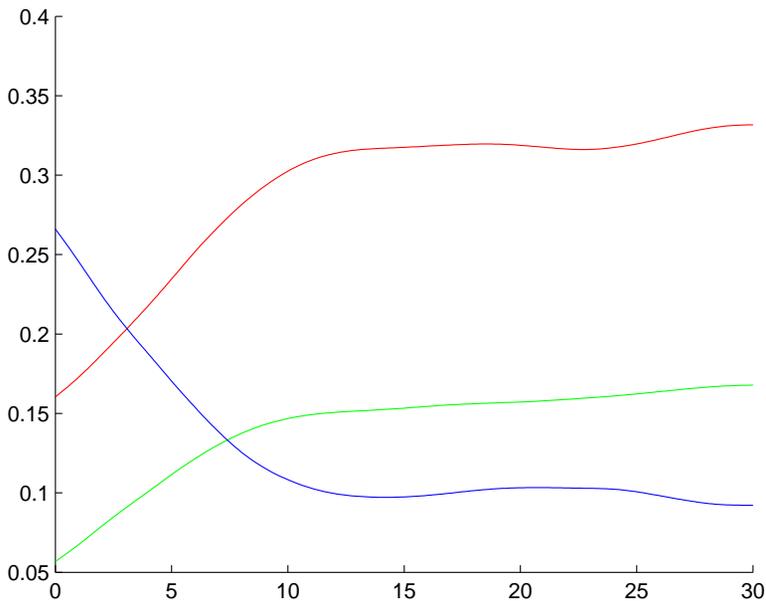


Рис. 3: Оценённые значения $\sigma^s(\cdot)$.

Из иллюстраций 9—12 видно, что занижение размерности шума, влияющего на динамику процесса, влечёт лишь увеличение абсолютного значения оставшихся компонент в то время, как качество оценок остаётся высоким.

На рисунках 13—16 видно, что четвёртая компонента шума явно была определена как лишняя, а качество оценки опять на высоте.

Замечание. Переменную размерность можно органично внести в алгоритм оценки параметров, добавив на каждом шаге некоторую вероятность увеличения или уменьшения размерности. Для простоты мы не

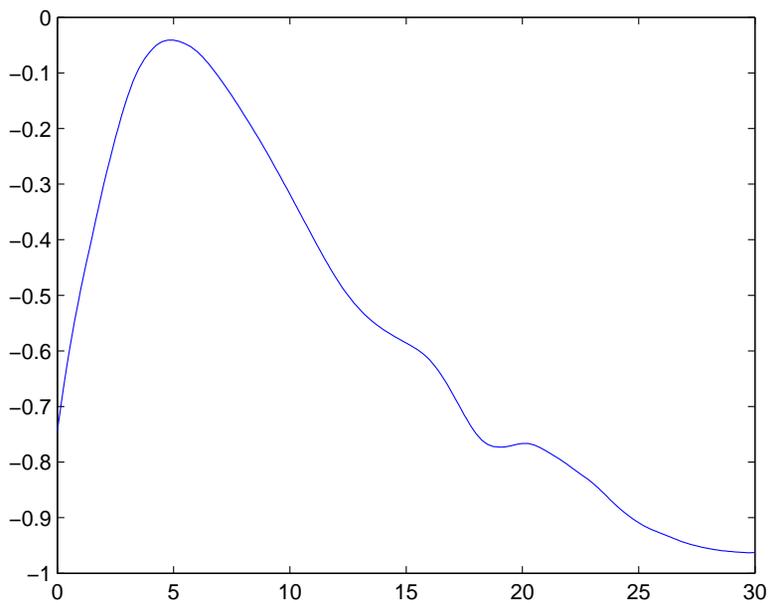


Рис. 4: Оценённое значение $\nu(\cdot)$.

стали этого делать.

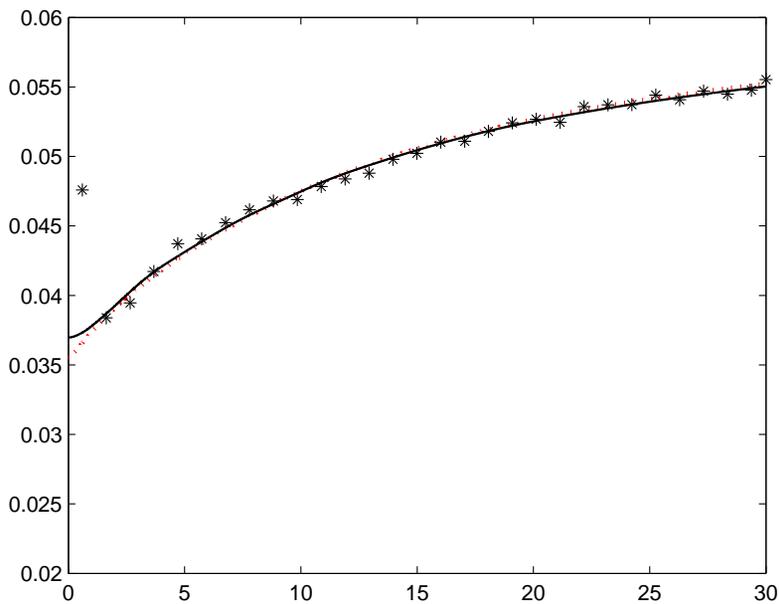


Рис. 5: Оценённое значение кривой доходности в момент t_{100} , $r_{100}(\cdot)$ и истинное значение.

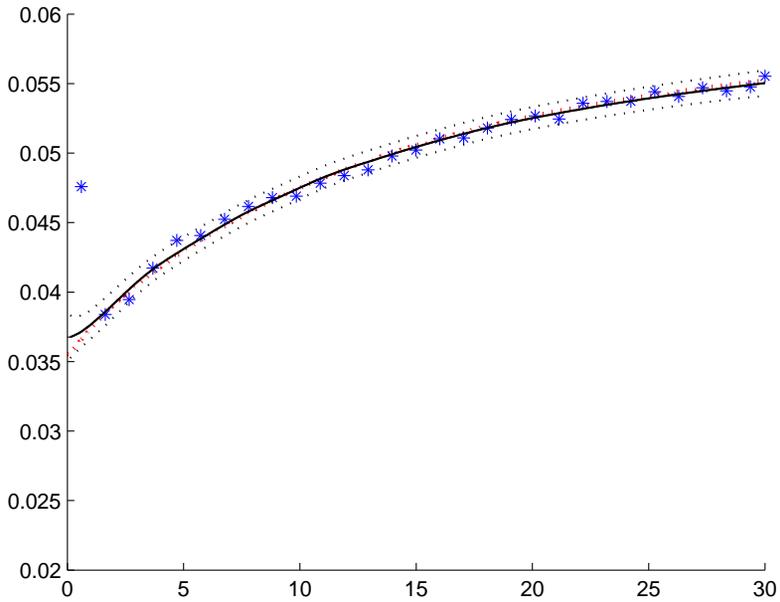


Рис. 6: Оценённое приближённым алгоритмом значение кривой доходности в момент t_{100} , $r_{100}(\cdot)$ и истинное значение.

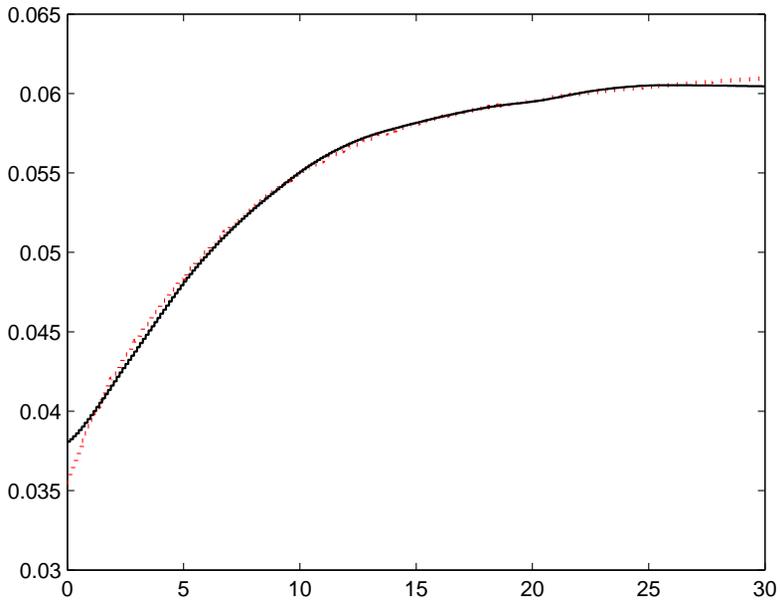


Рис. 7: Оценённое значение мгновенной форвардной процентной ставки $\xi_{100}(\cdot)$ и истинное значение.

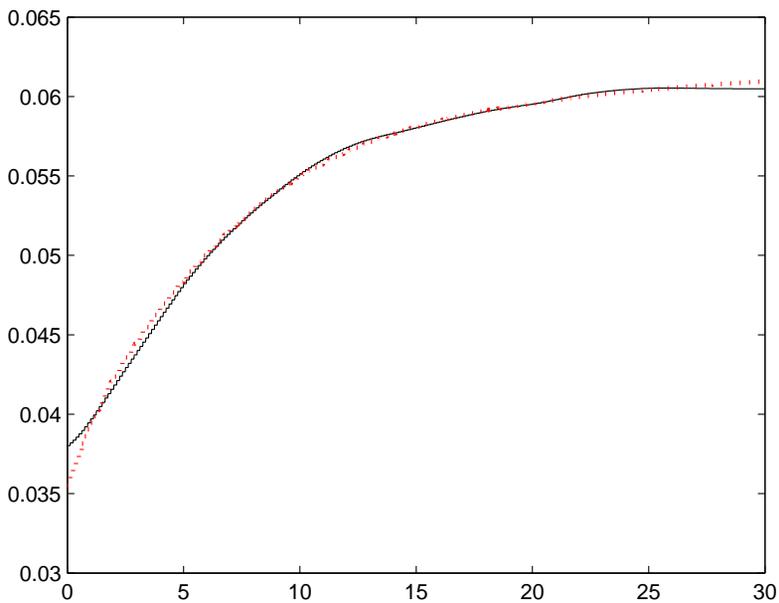


Рис. 8: Оценённое приближённым алгоритмом значение мгновенной форвардной процентной ставки $\xi_{100}(\cdot)$ и истинное значение.

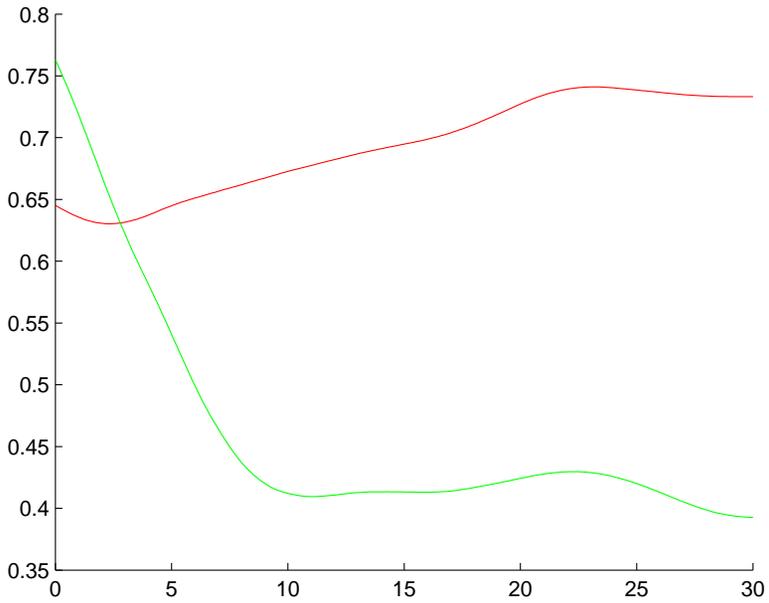


Рис. 9: Оценённые значения $\sigma^s(\cdot)$ ($N = 2$).

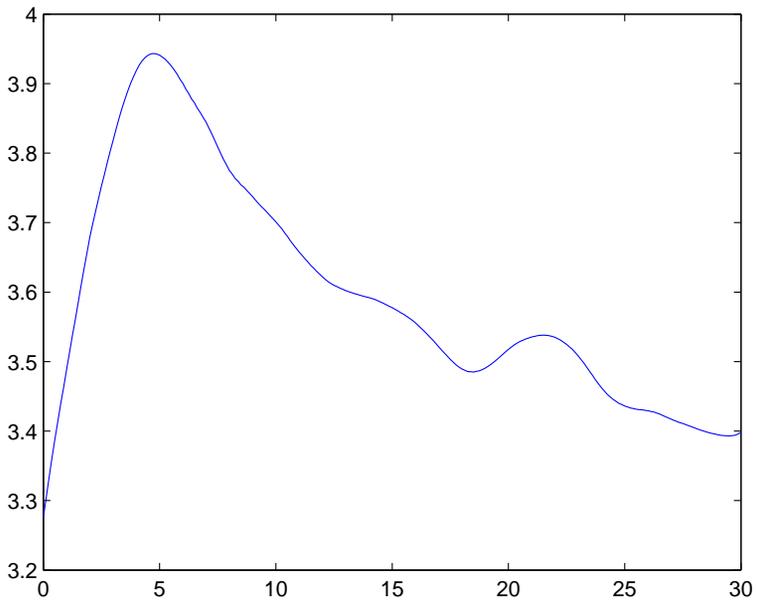


Рис. 10: Оценённое значение $\nu(\cdot)$ ($N = 2$).

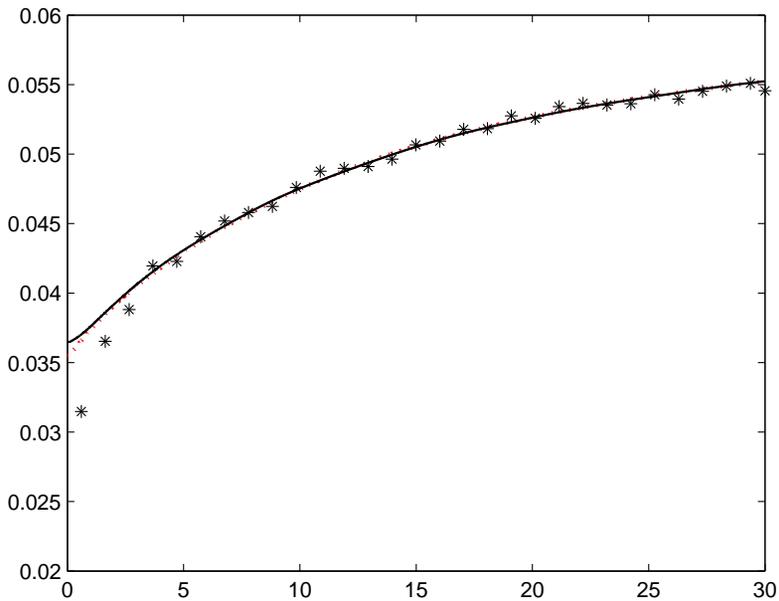


Рис. 11: Оценённое значение кривой доходности в момент t_{100} , $r_{100}(\cdot)$ и истинное значение ($N = 2$).

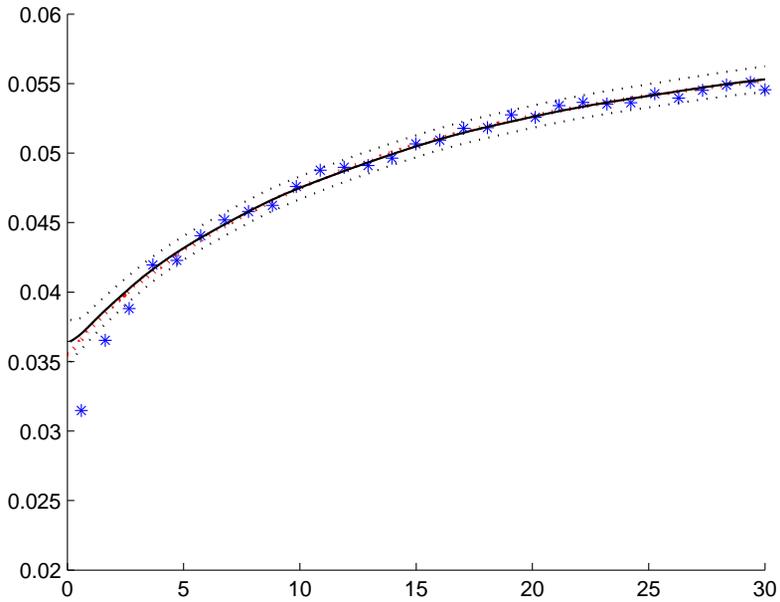


Рис. 12: Оценённое приближённым алгоритмом значение кривой доходности в момент t_{100} , $r_{100}(\cdot)$ и истинное значение ($N = 2$).

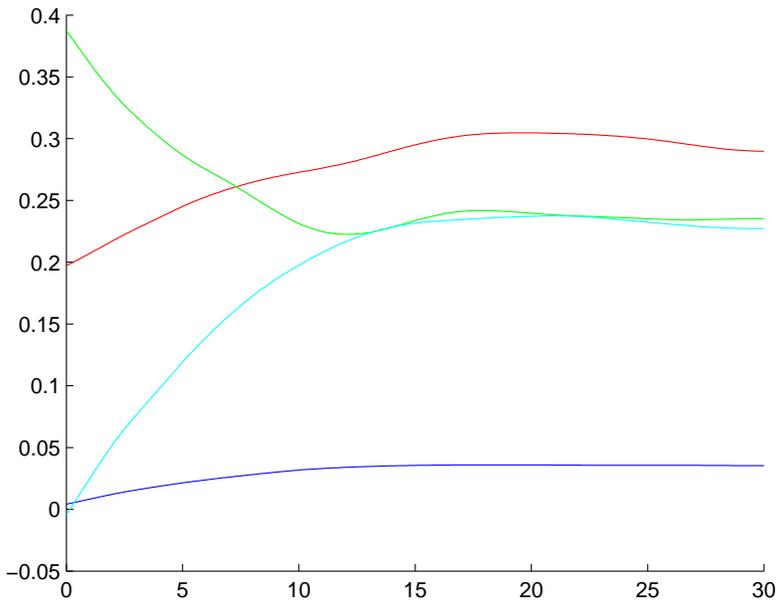


Рис. 13: Оценённые значения $\sigma^s(\cdot)$ ($N = 4$).

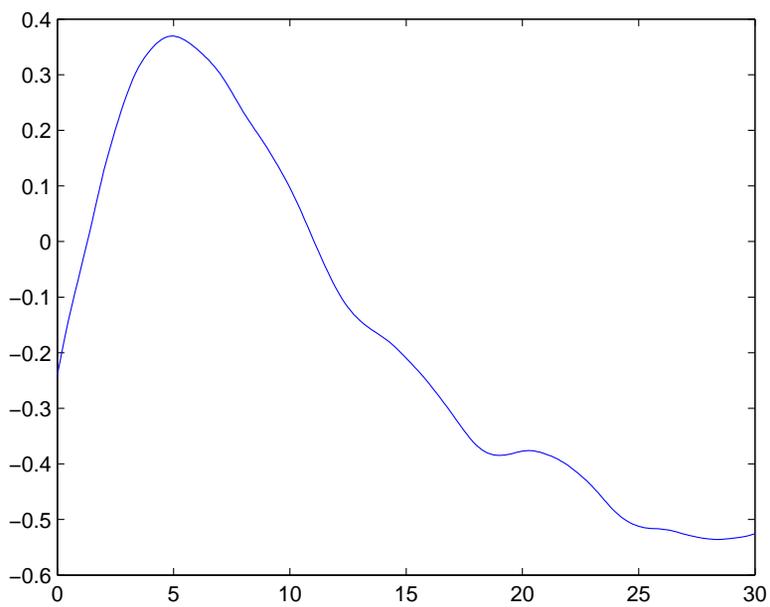


Рис. 14: Оценённое значение $\nu(\cdot)$ ($N = 4$).

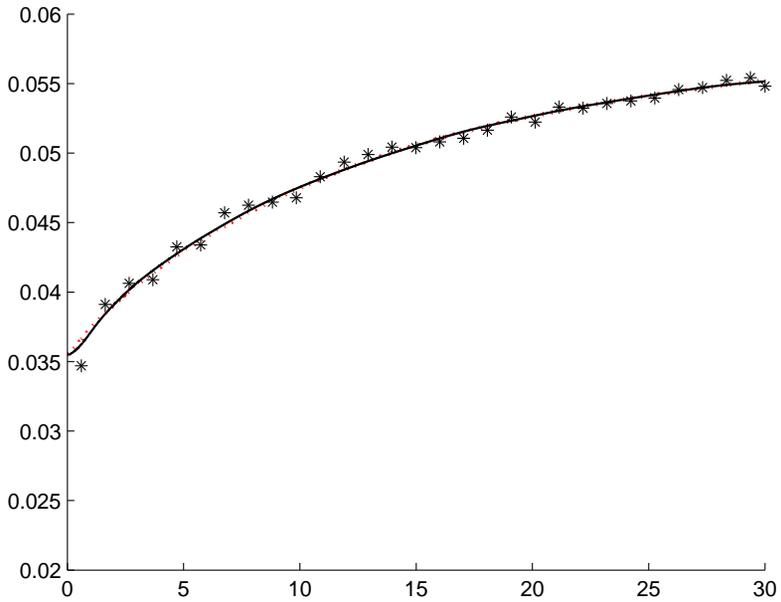


Рис. 15: Оценённое значение кривой доходности в момент t_{100} , $r_{100}(\cdot)$ и истинное значение ($N = 4$).

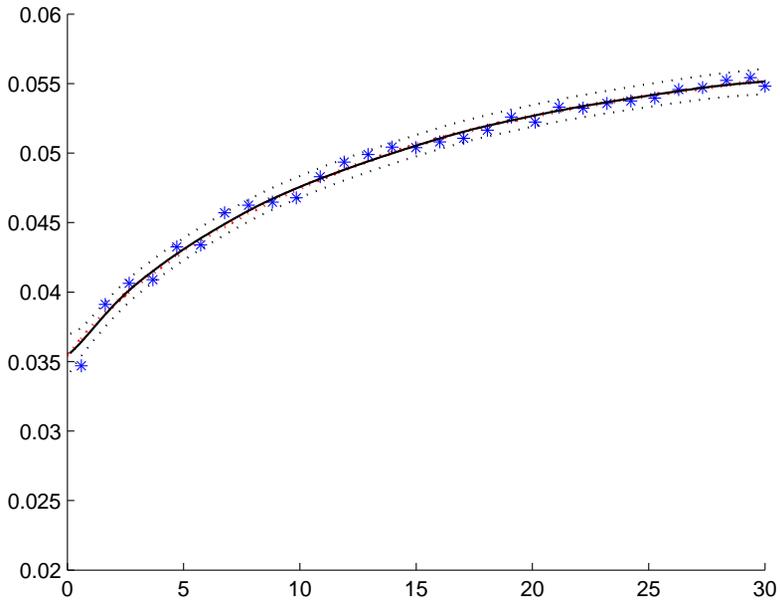


Рис. 16: Оценённое приближённым алгоритмом значение кривой доходности в момент t_{100} , $r_{100}(\cdot)$ и истинное значение ($N = 4$).

8 Заключение

Описанный метод непараметрической оценки параметров динамики срочной структуры процентных ставок даёт возможность использовать на практике сложные модели, наподобие описанной выше. К сожалению, качество идентификации, как показывает иллюстрация, оставляет пространство для улучшений: необходим более тонкий способ идентификации компонент $\{\sigma^s\}$, т.к. текущий вариант обеспечивает идентификацию лишь с точностью до линейной комбинации, хотя даже сейчас точность оценки вполне удовлетворительна. Высокая ресурсоёмкость алгоритма влечёт необходимость использования технологий параллельных вычислений.

Список литературы

- [1] *Лапшин В.* Определение срочной структуры процентных ставок // *Вестн. моск. ун-та . Сер. 15, Вычислительная математика и кибернетика.* — 2009. — № 4. — С. 37–43.
- [2] *Лапшин В.* Непараметрическая модель стохастической динамики процентных ставок // *Вестн. РУДН. Серия Математика. Информатика. Физика.* — 2009. — № 4. — С. 25–37.
- [3] *Dybvig P., Ingersoll Jr J., Ross S.* Long Forward and Zero-Coupon Rates Can Never Fall // *The Journal of Business.* — 1996. — Vol. 69, no. 1. — Pp. 1–25.
- [4] *Livingston M., Jain S.* Flattening of Bond Yield Curves for Long Maturities // *Journal of Finance.* — 1982. — Vol. 37, no. 1. — Pp. 157–167.
- [5] *Smirnov S., Zakharov A.* A Liquidity-Based Robust Spline Fitting of

Spot Yield Curve Providing Positive Forward Rates: Tech. rep.: European Bond Commission Working Paper, 2003.

- [6] *Ланшин В.* О задачах, связанных с определением срочной структуры процентных ставок // Вестник молодых ученых “Ломоносов”. — М.: Макс-пресс, 2006. — Т. 3. — С. 66–71.
- [7] *Тихонов А., Арсенин В.* Методы решения некорректных задач. — М.: Наука, 1979.
- [8] *Mallat S.* Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(R)$ // *Transactions of the American Mathematical Society*. — 1989. — Vol. 315, no. 1. — Pp. 69–87.
- [9] *Новиков И., Стечкин С.* Основные конструкции всплесков // *Фундаментальная и прикладная математика*. — 1997. — Т. 3, № 4. — С. 999–1028.
- [10] *Vidakovic B.* Statistical modeling by wavelets. — Wiley Series in Probability and Statistics, 1999.
- [11] *Chib S., Nardari F., Shephard N.* Analysis of high dimensional multivariate stochastic volatility models // *Journal of Econometrics*. — 2006. — Vol. 134, no. 2. — Pp. 341–371.
- [12] *Nielsen J., Madsen H., Young P.* Parameter estimation in stochastic differential equations: an overview // *Annual Reviews in Control*. — 2000. — Vol. 24. — Pp. 83–94.
- [13] *Ait-Sahalia Y.* Closed-Form Likelihood Expansions for Multivariate Diffusions // *The Annals of Statistics*. — 2008. — Vol. 36, no. 2. — Pp. 906–937.

- [14] *Hurn A., Jeisman J., Lindsay K.* Seeing the wood for the trees: A critical evaluation of methods to estimate the parameters of stochastic differential equations // *Journal of Financial Econometrics*. — 2007. — Vol. 5, no. 3. — Pp. 390–455.
- [15] *Lo A.* Maximum Likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data // *Econometric Theory*. — 1988. — Vol. 4, no. 2. — Pp. 231–247.
- [16] *Hurn A., Lindsay K.* Estimating the parameters of stochastic differential equations // *Mathematics and computers in simulation*. — 1999. — Vol. 48, no. 4. — Pp. 373–384.
- [17] *Jensen B., Poulsen R.* Transition densities of diffusion processes: Numerical comparison of approximation techniques // *The Journal of Derivatives*. — 2002. — Vol. 9. — Pp. 18–32.
- [18] *Broze L., Scaillet O., Zakoïan J.* Quasi-indirect inference for diffusion processes // *Econometric Theory*. — 1998. — Vol. 14, no. 02. — Pp. 161–186.
- [19] *Elerian O.* A note on the existence of a closed form conditional transition density for the milstein scheme: Tech. rep.: Nuffield College, Oxford University, 1998.
- [20] *Durham G., Gallant A.* Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes // *Journal of Business and Economic Statistics*. — 2002. — Vol. 20, no. 3. — Pp. 297–338.
- [21] *Shoji I., Ozaki T.* Estimation for nonlinear stochastic differential equations by a local linearization method // *Stochastic Analysis and Applications*. — 1998. — Vol. 16, no. 4. — Pp. 733–752.

- [22] *Ait-Sahalia Y.* Transition Densities for Interest Rate and Other Non-linear Diffusions // *The Journal of Finance*. — 1999. — Vol. 54, no. 4. — Pp. 1361–1395.
- [23] *Ait-Sahalia Y.* Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach // *Econometrica*. — 2002. — Vol. 70, no. 1. — Pp. 223–262.
- [24] *Pedersen A.* A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations // *Scandinavian journal of statistics*. — 1995. — Vol. 22, no. 1. — Pp. 55–71.
- [25] *Brandt M., Santa-Clara P.* Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets // *Journal of Financial Economics*. — 2002. — Vol. 63, no. 2. — Pp. 161–210.
- [26] *Hurn A., Lindsay K., Martin V.* On the efficacy of simulated maximum likelihood for estimating the parameters of stochastic differential equations // *Journal of Time Series Analysis*. — 2003. — Vol. 24, no. 1. — Pp. 45–63.
- [27] *Elerian O., Chib S., Shephard N.* Likelihood Inference for Discretely Observed Nonlinear Diffusions // *Econometrica*. — 2001. — Vol. 69, no. 4. — Pp. 959–993.
- [28] *Chen M., Shao Q., Ibrahim J.* Monte Carlo methods in Bayesian computation. — Springer, 2000.
- [29] *Congdon P.* Applied Bayesian Modelling. — John Wiley & Sons Inc, 2003.
- [30] *Häggström O.* Finite Markov chains and algorithmic applications. — Cambridge University Press, 2002.

- [31] *Eraker B.* MCMC analysis of diffusion processes with application to finance // *Journal of Business & Economic Statistics.* — 2001. — Vol. 19. — Pp. 177–191.
- [32] *Roberts G., Stramer O.* On inference for partially observed nonlinear diffusion models using the Metropolis-Hastings algorithm // *Biometrika.* — 2001. — Vol. 88, no. 3. — Pp. 603–621.
- [33] *Chib S., Nardari F., Shephard N.* Markov chain Monte Carlo methods for stochastic volatility models // *Journal of Econometrics.* — 2002. — Vol. 108, no. 2. — Pp. 281–316.
- [34] *Jacquier E., Polson N., Rossi P.* Bayesian analysis of stochastic volatility models with fat-tails and correlated errors // *Journal of Econometrics.* — 2004. — Vol. 122, no. 1. — Pp. 185–212.
- [35] Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion) / A. Beskos, O. Papaspiliopoulos, G. Roberts, P. Fearnhead // *Journal of the Royal Statistical Society: Series B (Statistical Methodology).* — 2006. — Vol. 68, no. 3. — Pp. 333–382.
- [36] *Kalogeropoulos K.* Likelihood-based inference for a class of multivariate diffusions with unobserved paths // *Journal of Statistical Planning and Inference.* — 2007. — Vol. 137, no. 10. — Pp. 3092–3102.
- [37] *Golightly A., Wilkinson D.* Bayesian sequential inference for nonlinear multivariate diffusions // *Statistics and Computing.* — 2006. — Vol. 16, no. 4. — Pp. 323–338.
- [38] *Altissimo F., Mele A.* Simulated Non-Parametric Estimation of Dynamic Models // *Review of Economic Studies.* — 2009. — Vol. 76, no. 2. — Pp. 413–450.

- [39] *Hansen L.* Large Sample Properties of Generalized Method of Moments Estimators // *Econometrica*. — 1982. — Vol. 50, no. 4. — Pp. 1029–1054.
- [40] *Hansen L., Heaton J., Yaron A.* Finite-Sample Properties of Some Alternative GMM Estimators // *Journal of Business & Economic Statistics*. — 1996. — Vol. 14, no. 3. — Pp. 262–280.
- [41] *Hansen L., Scheinkman J.* Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes // *Econometrica*. — 1995. — Vol. 63, no. 4. — Pp. 767–804.
- [42] An empirical comparison of alternative models of the short-term interest rate / K. Chan, G. Karolyi, F. Longstaff, A. Sanders // *Journal of Finance*. — 1992. — Vol. 47, no. 3. — Pp. 1209–1227.
- [43] *Gourieroux C., Monfort A., Renault E.* Indirect Inference // *Journal of Applied Econometrics*. — 1993. — Vol. 8. — Pp. S85–S118.
- [44] *Gallant A., Tauchen G.* Which Moments to Match? // *Econometric Theory*. — 1996. — Vol. 12, no. 4. — Pp. 657–681.
- [45] *Singleton K.* Estimation of affine asset pricing models using the empirical characteristic function // *Journal of Econometrics*. — 2001. — Vol. 102, no. 1. — Pp. 111–141.
- [46] *Jiang G., Knight J.* Estimation of continuous-time processes via the empirical characteristic function // *Journal of Business and Economic Statistics*. — 2002. — Vol. 20, no. 2. — Pp. 198–212.
- [47] *Chacko G., Viceira L.* Spectral GMM estimation of continuous-time processes // *Journal of Econometrics*. — 2003. — Vol. 116, no. 1-2. — Pp. 259–292.

- [48] *Kessler M., Sørensen M.* Estimating equations based on eigenfunctions for a discretely observed diffusion process // *Bernoulli*. — 1999. — Pp. 299–314.
- [49] *Ait-Sahalia Y.* Testing Continuous-Time Models of the Spot Interest Rate // *Review of Financial Studies*. — 1996. — Vol. 9, no. 2. — Pp. 385–426.
- [50] *Pritsker M.* Nonparametric density estimation and tests of continuous time interest rate models // *Review of Financial Studies*. — 1998. — Vol. 11, no. 3. — Pp. 449–487.
- [51] *Metropolis N., Ulam S.* The monte carlo method // *Journal of the American Statistical Association*. — 1949. — Vol. 44, no. 247. — Pp. 335–341.
- [52] Equation of state calculations by fast computing machines / N. Metropolis, A. Rosenbluth, M. Rosenbluth et al. // *The journal of chemical physics*. — 1953. — Vol. 21, no. 6. — P. 1087.
- [53] *Hastings W.* Monte Carlo sampling methods using Markov chains and their applications // *Biometrika*. — 1970. — Pp. 97–109.
- [54] *Chib S., Greenberg E.* Understanding the metropolis-hastings algorithm // *The American Statistician*. — 1995. — Vol. 49, no. 4. — Pp. 327–335.
- [55] *Chib S., Greenberg E.* Markov Chain Monte Carlo Simulation Methods in Econometrics // *Econometric Theory*. — 1996. — Vol. 12, no. 3. — Pp. 409–431.
- [56] *Geman S., Geman D.* Stochastic relaxation, gibbs distributions, and the bayesian restoration of images // *IEEE Transactions on Pattern*

- Analysis and Machine Intelligence*. — 1984. — Vol. 6, no. 6. — Pp. 721–741.
- [57] *Papaspiliopoulos O., Roberts G., Skold M.* A General Framework for the Parametrization of Hierarchical Models // *Statistical Science*. — 2007. — Vol. 22, no. 1. — Pp. 59–73.
- [58] *Liu J., Liang F., Wong W.* The Multiple-Try Method and Local Optimization in Metropolis Sampling // *Journal of the American Statistical Association*. — 2000. — Vol. 95, no. 449. — Pp. 121–134.
- [59] *Gilks W. R., Roberts G. O., George E. I.* Adaptive direction sampling // *Journal of the Royal Statistical Society. Series D (The Statistician)*. — 1994. — Vol. 43, no. 1. — Pp. 179–189.
- [60] *Gilks W., Richardson S., Spiegelhalter D.* Markov chain Monte Carlo in practice. — Chapman & Hall/CRC, 1996.
- [61] Adaptive radial-based direction sampling: some flexible and robust Monte Carlo integration methods / L. Bauwens, C. Bos, H. van Dijk, R. van Oest // *Journal of Econometrics*. — 2004. — Vol. 123, no. 2. — Pp. 201–225.
- [62] *Gilks W., Wild P.* Adaptive Rejection Sampling for Gibbs Sampling // *Applied Statistics*. — 1992. — Vol. 41, no. 2. — Pp. 337–348.
- [63] *Gilks W., Best N., Tan K.* Adaptive Rejection Metropolis Sampling within Gibbs Sampling // *Applied Statistics*. — 1995. — Vol. 44, no. 4. — Pp. 455–472.
- [64] *Dempster A., Laird N., Rubin D.* Maximum Likelihood from Incomplete Data via the EM Algorithm // *Journal of the Royal Statistical Society. Series B (Methodological)*. — 1977. — Vol. 39, no. 1. — Pp. 1–38.

- [65] *Tanner M., Wong W.* The Calculation of Posterior Distributions by Data Augmentation // *Journal of the American Statistical Association.* — 1987. — Vol. 82, no. 398. — Pp. 528–540.
- [66] *van Dyk D., Meng X.* The Art of Data Augmentation // *Journal of Computational and Graphical Statistics.* — 2001. — Vol. 10, no. 1. — Pp. 1–50.
- [67] *Jones C.* Bayesian estimation of continuous-time finance models: Tech. rep.: Working paper, Rochester University, 1998.
- [68] *Jones C.* A Simple Bayesian Method for the Analysis of Diffusion Processes // *SSRN eLibrary.* — 1998. <http://ssrn.com/paper=111488>.
- [69] *Ghosh-Dastidar B., Schafer J.* Multiple edit/multiple imputation for multivariate continuous data // *Journal of the American Statistical Association.* — 2003. — Vol. 98, no. 464. — Pp. 807–817.

Препринт WP16/2009/02
Серия WP16
Финансовая инженерия,
риск-менеджмент и актуарная наука

Лапшин В.А.

Алгоритм оценки параметров стохастической динамики срочной структуры процентных ставок

Зав. редакцией оперативного выпуска *А.В. Заиченко*
Технический редактор *Ю.Н. Петрина*

ЛР № 020832 от 15 октября 1993 г.
Отпечатано в типографии Государственного университета –
Высшей школы экономики с представленного оригинал-макета.
Формат 60×84 ¹/₁₆. Бумага офсетная. Тираж 150 экз. Уч.-изд. л. 3
Усл. печ. л. 3,72. Заказ № . Изд. № 1139.

Государственный университет – Высшая школа экономики
125319, Москва, Кочновский проезд, 3
Типография Государственного университета – Высшей школы экономики
125319, Москва, Кочновский проезд, 3
Тел.: (495) 772-95-71; 772-95-73

Для заметок
