МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

МАТЕРИАЛЫ
VIII Международной молодежной научной
конференции
«МАТЕМАТИЧЕСКОЕ
И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ
ИНФОРМАЦИОННЫХ,
ТЕХНИЧЕСКИХ
И ЭКОНОМИЧЕСКИХ СИСТЕМ»

Томск, 26–30 мая 2021 г.

Под общей редакцией И.С. Шмырина

Томск
Издательство Томского государственного университета
2021

Действительно, обеспечивается точность оценивания $1/H$. В сравнении привели данные оценки [1], и видим, что гарантированная точность наблюдается хуже при росте θ . Также наблюдаем, что среднее время оценивания параметра при улучшенной оценке меньше, чем при другой.

Зачем данное улучшение? Ответ очевиден. Чем больше данных наблюдений мы имеем, тем точнее оценку мы получим. Но не всегда может быть большое количество наблюдений, относительно которых мы оцениваем параметр авторегрессии. Отметим, что намного чаще используются задачи с малым количеством данных. Данная проблема была решена в [2], где была построена оценка параметра авторегрессии, которая гарантирует заданную точность оценивания параметра независимо от количества наблюдений. Также данная оценка обеспечивает минимальное время оценивания параметра, что тоже немало важно.

Заключение

В заключении хочется отметить следующее: оценка параметра авторегрессии [2] лучше оценки [1] в смысле гарантированного среднеквадратического отклонения и времени оценивания, время оценивания параметра улучшенной оценки при неизвестной дисперсии шумов не радикально превышает время оценивания при известной дисперсией шумов, оценки сохраняют свои свойства при увеличении h и при гауссовском и двойном экспоненциальном шумах.

ЛИТЕРАТУРА

1. Дмитриенко А.А., Конев В.В. О гарантированном оценивании параметров авторегрессии при неизвестной дисперсии помех // Автоматика и телемеханика. – 1994. – № 2. – С. 87–99.
2. Vorobeychikov S.E., Burkatovskaya Y.B. Non-asymptotic Confidence Estimation of the Autoregressive Parameter in AR(1) Process with an Unknown Noise Variance // Austrian Journal of Statistics. – 2020. – 49(4). – P. 19–26.
3. Борисов В.З., Конев В.В. О последовательном оценивании параметров дискретных процессов // Автоматика и телемеханика. – 1977. – № 10. – С. 58–64.

DOI: 10.17223/978-5-907442-42-9-2021-11

НAR-RV МОДЕЛЬ И ПРОГНОЗ ВОЛАТИЛЬНОСТИ

Пупков А.В.

Томский государственный университет
andrewpupkov@gmail.com

Введение

Многие динамические процессы описываются стохастическими дифференциальными и стохастическими разностными уравнениями. Как правило, моделирование производится с помощью аппроксимации случайного процесса некоторой моделью определенной структуры. Наиболее известные модели дискретного времени, используемые для таких целей – это модель скользящего среднего (MA(q)), модель авторегрессии (AR(p)), семейство GARCH(p,q) моделей и всевозможные их сочетания. Также существует семейство непрерывных моделей. Наиболее значимым на сегодняшний день среди них является модель броуновского движения или так называемый винеровский процесс, на основе которого строится множество моделей, например, классическое геометрическое броуновское движение, используемое для моделирования финансовых рынков. Во многих случаях процесс моделирования требует достаточно много априорной информации о процессе, например, для аппроксимации броуновским движением некоторого процесса необходимо, чтобы математическое ожидание и приращения этого процесса удовлетворяли гауссовости. В случаях, когда ситуация не позволяет знать распределение процесса, на помощь приходят непараметрические модели ядерного типа. Аппроксимация случайных процессов упомянутыми моделями производится во

многих областях науки и техники, например, в экономике [1–3], физике [4], теории информации и кодирования [5], страховании [6] и т.д.

Для моделирования финансовых рынков, как правило, используются модели для описания поведения основного тренда процесса, а также дополнительная составляющая, описывающая динамику дисперсии процесса или так называемую волатильность. Для описания волатильности используются модели GARCH(p,q) семейства, а также HAR-RV модель реализованной волатильности, которая отличается простотой и интерпретируемостью.

Цель данной работы – представить результаты моделирования последовательности реализованных волатильностей на основе данных динамики индекса S&P500 за 19-летний период с 16.02.2001 по 27.08.2020, используя деление рассматриваемого процесса на режимы.

1. Понятие реализованной волатильности

В данной работе для прогноза волатильности используется модель, предложенная Corsi в 2009 г. в [7]. Автором был рассмотрен диффузионный процесс вида $dp(t) = \mu(t)dt + \sigma(t)dW(t)$, где $p(t)$ – цены актива в момент t , $\mu(t)$ – процесс, описывающий тренд цены актива, $W(t)$ – стандартное броуновское движение, $\sigma(t)$ – процесс, не зависящий от $W(t)$. В качестве описательной характеристики выступает следующая величина

$$IV_t^{(d)} = \int_{t-d}^t \sigma^2(\omega) d\omega, \quad (1)$$

называемая интегрированной дисперсией, где d – полный день торгов. Состоятельная оценка (1) имеет вид

$$RV_t = \sqrt{\sum_{j=0}^M r_{t-j\Delta}^2}, \quad r_{t-j\Delta} = \ln \left(\frac{P_{t-j\Delta}}{P_{t-(j+1)\Delta}} \right), \quad (2)$$

где t – день, M – количество внутридневных доходностей, $\Delta = 1/M$, P_t – значение индекса в момент времени t . Реализованная волатильность за неделю и за месяц вычисляются соответственно как $RVW_t = \frac{1}{5} \sum_{k=0}^4 RV_k$, $RVW_t = \frac{1}{22} \sum_{k=0}^{21} RV_k$, где среднее рассчитывается по количеству торговых дней за неделю и за месяц соответственно.

2. HAR-RV модель

HAR-RV модель имеет вид $RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RVW_t + \beta_3 RVM_t + \varepsilon_{t+1}$, где $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)^T$ – вектор параметров модели. HAR-RV модель показывает гораздо более качественные результаты при прогнозировании реализованной волатильности, по сравнению со всем семейством GARCH моделей [8].

3. Рассматриваемые данные и вычисление реализованных волатильностей

Рассматриваются значения индекса S&P500 за 19-летний период с 16.02.2001 по 27.08.2020. На рис. 1 представлена последовательность значений индекса S&P500 и последовательность ежедневных логарифмических доходностей индекса.

Рис. 1. Последовательность значений индекса S&P500 (слева)
и последовательность ежедневных логарифмических доходностей (справа)

Для вычисления последовательности реализованных волатильностей необходимо иметь последовательность внутридневных логарифмических доходностей, часть из которых представлены на рис. 2.

Рис. 2. Последовательность внутридневных доходностей r_t
(вертикальные серые линии означают конец торгового дня)

На рис. 3 представлена последовательность логарифмических доходностей (светло-серый график) и соответствующая последовательность реализованных волатильностей, вычисленная по (2) (черный график).

Рис. 3. Последовательность логарифмических доходностей (светло-серый график)
и последовательность $\{RV_t\}$ (черный график)

Из рис. 3 видно, что последовательность $\{RV_t\}$ достаточно хорошо описывает последовательность ежедневных логарифмических доходностей индекса.

4. Деление последовательности реализованных волатильностей на режимы

Для деления процесса на режимы использовалась скрытая марковская цепь. Для ее применения необходимо последовательность $\{RV_t\}$ преобразовать к последовательности, которая будет принимать конечное дискретное число значений. В данной работе предлагается подобное преобразование произвести по незамысловатой схеме вида

$$V_t = \begin{cases} HV, & \text{если } RV \geq \alpha, \\ LV, & \text{если } RV < \alpha, \end{cases}$$

где α – некоторое положительное число.

Далее при помощи алгоритма Баума – Велша находится значение эмиссионной матрицы скрытой марковской цепи. На основе полученной эмиссионной матрицы, при помощи алгоритма Витерби, вычисляется наиболее вероятное скрытое состояние, в которой находится скрытая марковская цепь в момент t . Стоит отметить, что значение полученной эмиссионной матрицы, довольно сильно зависит от ее начального значения. На рис. 4 представлены два варианта деления последовательности реализованных волатильностей на режимы в зависимости от различных начальных значений эмиссионной матрицы.

Рис. 4. Последовательность $\{RV_t\}$, разделенная на 3 режима. Приводятся два варианта деления на режимы

5. Оценивание HAR-RV модели на основе данных из разных режимов

Оценивание параметра β HAR-RV модели будем производить для данных из каждого режима, т.е. сначала разделим данные на три кластера, а затем на каждом из этих кластеров оценим параметры HAR-RV модели. Оценивание будем производить при помощи метода наименьших квадратов. Оценка имеет вид $\hat{\beta}_t^i = (A_{t,i}^T A_{t,i})^{-1} A_{t,i}^T y_{t+1,i}$, $i = \overline{1,3}$, где

$$A_{t,i} = \begin{bmatrix} 1 & RV_{t,i} & \frac{1}{5} \sum_{j=0}^4 RV_{t-j,i} & \frac{1}{22} \sum_{j=0}^{21} RV_{t-j,i} \\ 1 & RV_{t-1,i} & \frac{1}{5} \sum_{j=0}^4 RV_{t-j-1,i} & \frac{1}{22} \sum_{j=0}^{21} RV_{t-j-1,i} \\ \dots & \dots & \dots & \dots \\ 1 & RV_{0,i} & \frac{1}{5} \sum_{j=0}^4 RV_{-j,i} & \frac{1}{22} \sum_{j=0}^{21} RV_{-j,i} \end{bmatrix}, \quad y_{t+1,i} = \begin{bmatrix} RV_{t+1,i} \\ RV_{t,i} \\ \dots \\ RV_{1,i} \end{bmatrix},$$

здесь $RV_{t,i}$ – значение реализованной волатильности в день t в i -м кластере.

В следующем разделе рассмотрим прогнозирование последовательности реализованных волатильностей на 1 день вперед.

6. Прогнозирование волатильности при помощи HAR-RV модели

Прогноз реализованной волатильности на 1 день описывается выражением $\overline{RV}_{t+1} = \beta_0^* + \beta_1^* RV_t + \beta_2^* RVW_t + \beta_3^* RVM_t$, где $\beta_0^* = \beta_0^{R(RV_t)}$, $\beta_1^* = \beta_1^{R(RV_t)}$, $\beta_2^* = \frac{1}{5} \sum_{i=0}^4 \beta_2^{R(RV_{t-i})}$,

$\beta_3^* = \frac{1}{22} \sum_{i=0}^{21} \beta_3^{R(RV_{t-i})}$, $R(x)$ – режим, в котором находится наблюдение x .

В рассматриваемой ситуации параметры скрытой марковской цепи оценивались по 3000 первых наблюдений реализованной волатильности. Значения реализованной волатильности разделялись на режимы и внутри выборки оценивания параметров НММ и за ее пределами. Прогноз строился за пределами выборки оценивания параметров HAR-RV.

На рис. 5 представлены прогнозы на 1 день в разные временные интервалы. На рис. 5 (слева) рассмотрена ситуация перехода с высоковолатильного режима на низковолатильный. На рис. 5 (справа) рассмотрена ситуация перехода с низковолатильного на высоковолатильный режим с последующим снижением до низковолатильного. Визуально видно, что прогнозная сила алгоритма ухудшается в ситуации, когда происходит резкое повышение волатильности.

Рис. 5. Прогноз реализованной волатильности на 1 день (серая ломанная линия – реализованная волатильность, черные столбцы – прогноз на 1 день): график слева – с 01.02.2016 до 20.05.2016, график справа – с 12.02.2020 до 20.05.2020

Заключение

В настоящей работе представлен алгоритм прогнозирования реализованной волатильности процесса, описывающего динамику значения индекса S&P500, на основе деления процесса на режимы. Стоит отметить, что рассматриваемая модель ведет себя гораздо лучше в прогнозе реализованных волатильностей, чем все семейство GARCH моделей.

ЛИТЕРАТУРА

1. Ширяев А.Н. Основы стохастической финансовой математики: в 2 т. – Москва: ФАЗИС, 1998.
2. Кошкин Г.М., Пивен И.Г. Непараметрическая идентификация стохастических объектов. – Хабаровск: РАН, Дальневосточное отделение, 2009. – 336 с.
3. Tsay R.S. An Introduction to Analysis of Financial Data with R. – John Wiley & Sons, 2013. – 390 p.
4. Френкель Я.И. Статистическая физика. – Москва Ленинград: АН СССР, 1948. – 760 с.
5. Чечётка С.И. Введение в дискретную теорию информации и кодирования. – М: МЦНМО, 2011. – 224 с.
6. Asmussen S. Albrecher H. Ruin Probabilities. – World Scientific Publishing Co. Pte. Ltd., 2010. – 602 p.
7. Corsi F. A Simple Approximate Long-Memory Model of Realized Volatility // Journal of Financial Econometrics. – 2009. – V. 7. – N. 2. – P. 174 – 196.
8. Аганин А.Д. Сравнение GARCH и HAR-RV моделей для прогноза реализованной волатильности на российском рынке // Прикладная эконометрика. – 2017. – Т. 48. – С. 63 – 84.