Econophysica

ПРОГНОЗ ВОЛАТИЛЬНОСТИ ПРИ ПОМОЩИ HAR-RV МОДЕЛИ НА ОСНОВЕ РЕЖИМОВ

Автор работы

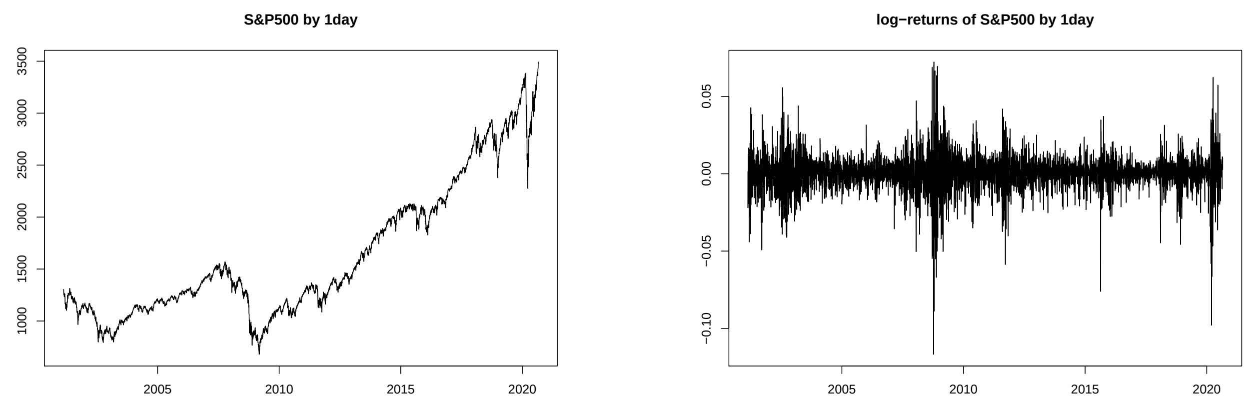
_____________ А. В. Пупков
Подпись

Томск – 2020

ОГЛАВЛЕНИЕ

1 Рассматриваемый процесс

Рассматривается последовательность значений индекса $S\&P500$ за 19-летний период с 16.02.2001 до 27.08.2020 (см. рис. 1). Данные: *https://www.finam.ru/*.



a) b)

Рисунок 1 — a) Последовательность значений индекса $S\&P500$ (с 16.02.2001 до 27.08.2020) b) Последовательность логарифмических доходностей индекса $S\&P500$ (с 19.02.2001 до 27.08.2020)

Для оценивания реализованной волатильности (см. ниже) используются данные индекса по 5-минутым промежуткам. На рис. 2 представлена последовательность логарифмических доходностей по 5-минутным интервалам. Серые линии, разделяющие временную шкалу, обозначают конец торгового дня. Более широкие промежутки между линиями, связаны с отсутствием торгов на выходных или с праздничными днями.

2 HAR-RV модель

В данной работе для прогноза волатильности используется модель, предложенная Corsi в 2009г. в статье [1]. Автором был рассмотрен диффузионный процесс вида

$$dp(t) = \mu(t)dt + \sigma(t)dW(t),$$

в котором $\sigma(t)$ некий случайный процесс, описывающий изменение дисперсии броуновского движения. Интегрированная дисперсия за день t имеет вид

$$IV_t^{(d)} = \int_{t-1d}^{t} \sigma^2(x)dx, \tag{1}$$

где $1d$ – полный день торгов.

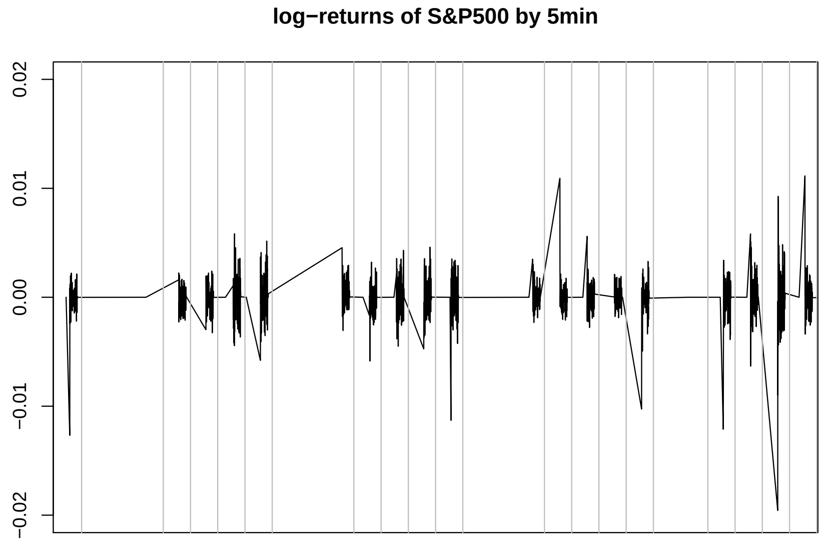


Рисунок 2 — Последовательность логарифмических доходностей $S\&P500$ по 5-минутным интервалам с 17.02.2001 до 15.03.2001 (серые линии обозначают конец торгового дня)

В работах, упомянутых в литературе статьи [1], приводится следующий результат. Состоятельная оценка величины (1) имеет вид

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

где t – день, M – количество внутридневных доходностей, $\Delta = 1d/M$, P_t – значение индеса в момент времени t. Реализованная волатильность за неделю и за месяц вычисляются соответственно как

$$RVW_t = \frac{1}{5}\sum_{j=0}^{4} RV_{t-j}, \qquad RVM_t = \frac{1}{22}\sum_{j=0}^{21} RV_{t-j}.$$

Последовательность реальзованных волатильностей $\{RV_t\}$ представлена на рис. 3.

HAR-RV модель имеет следующий вид

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RVW_t + \beta_3 RVM_t + \varepsilon_{t+1}.$$

где ε_t –шум модели в день t.

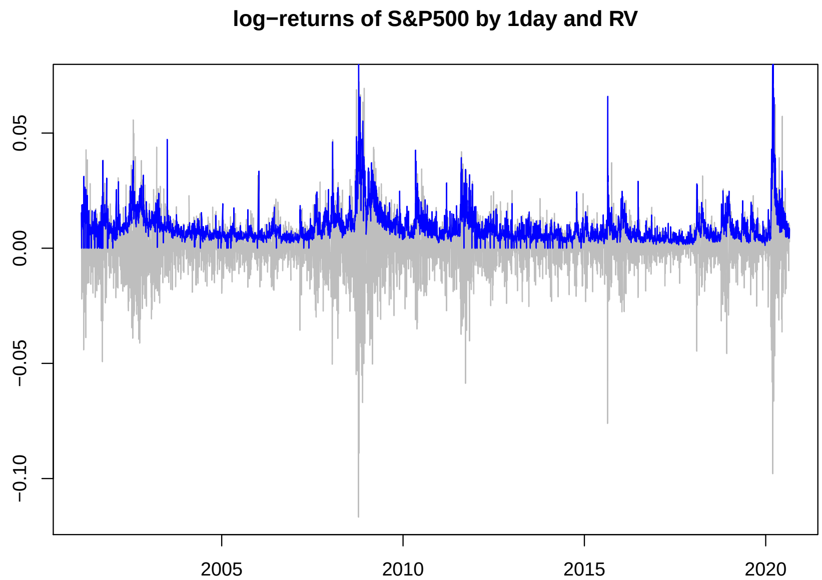


Рисунок 3 — Последовательность логарифмических доходностей $S\&P500$ по дневным интервалам с 19.02.2001 до 27.08.2020 (серый график) и последовательность реализованных волатильностей $\{RV_t\}$ (синий график)

3 Деление на режимы

Для деления исследуемого процесса на режимы используем скрытую марковскую цепь (HMM).

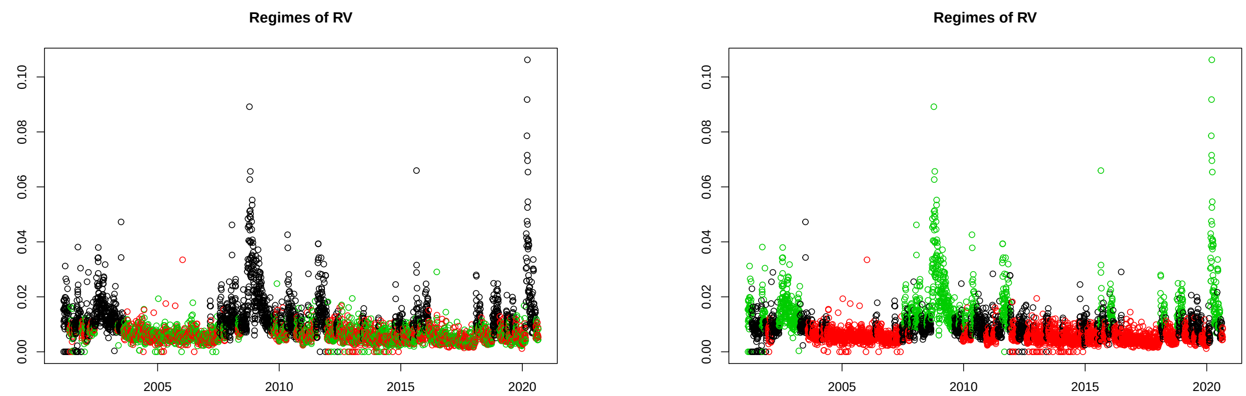
Сначала представим данные в дискретном виде для использования их в HMM-алгоритмах. Введем коэффицинт α и преобразуем последовательность $\{RV_t\}$ следующим образом

$$V_t = \begin{cases} HV, & \text{если } RV_t > \alpha, \\ LV, & \text{если } RV_t \leq \alpha, \end{cases} \tag{2}$$

Затем при помощи алгоритма Баума–Велша найдем параметры скрытой марковской цепи. Определим через $n = 3$ – количество скрытых состояний, а через $k = 2$ – количество значений, в которые преобразуем наблюдения $\{RV_t\}$ по формуле (2).

Далее, используя алгоритм Витерби, найдем наиболее вероятное скрытое состояние, в котором находится процесс в каждый момент времени.

Стоит отметить, что реализация процедур очень сильно зависит от начальных значений матрицы переходов и эмиссионной матрицы скрытой марковской цепи, которые используются в алгоритме Баума–Велша. Например, на рис. 4 представлены два варианта деления процесса реализованных волатильностей на режимы.



a) b)

Рисунок 4 — Деление последовательности $\{RV_t\}$ на режимы в зависимости от различных начальных значений матрицы переходов и эмиссионной матрицы (с 19.02.2001 до 27.08.2020)

4 Оценивание модели HAR-RV при помощи процедуры деления на режимы

Оценим параметры HAR-RV модели при помощи метода наименьших квадратов (МНК) по данным из разных режимов. Обозначим через $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)'$ вектор неизвестных параметров. Тогда МНК-оценка вектора параметров имеет вид

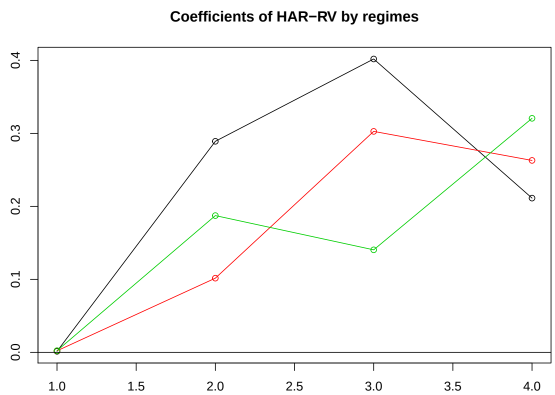
$$\hat{\beta}_t = (A_t' A_t) A_t' y_{t+1},$$

где

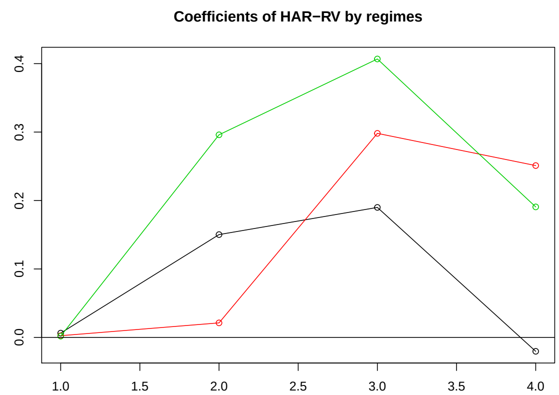
$$A_t = \begin{bmatrix} 1 & y_t & \frac{1}{5}\sum_{i=0}^{4} y_{t-i} & \frac{1}{22}\sum_{i=0}^{21} y_{t-i} \\ 1 & y_{t-1} & \frac{1}{5}\sum_{i=0}^{4} y_{t-1-i} & \frac{1}{22}\sum_{i=0}^{21} y_{t-1-i} \\ \dots & \dots & \dots\dots\dots & \dots\dots\dots \\ 1 & y_0 & \frac{1}{5}\sum_{i=0}^{4} y_{-i} & \frac{1}{22}\sum_{i=0}^{21} y_{-i} \end{bmatrix} \qquad y_{t+1} = \begin{bmatrix} y_{t+1} \\ y_t \\ \dots \\ y_1 \end{bmatrix}.$$

Оценив параметры HAR-RV модели на каждом из режимов, получим n векторов оценок параметров. Каждый из векторов соответствует одному из n

скрытых состояний. На рис. 5 представлены графики векторов параметров. На оси абсцисс в точке 1 находятся значения оценок $\{\hat{\beta}_0^i\}_{1\leq i\leq n}$ для каждого из n режимов (верхний индекс описывает принадлежность оценки параметра к режиму i), в точке 2 – оценки $\{\hat{\beta}_1^i\}_{1\leq i\leq n}$, в 3 – оценки $\{\hat{\beta}_2^i\}_{1\leq i\leq n}$ и в 4 – $\{\hat{\beta}_3^i\}_{1\leq i\leq n}$. Параметры, вычисленные в одном режиме, выделены одним цветом. Графики 4.a) и 5.a), а также 4.b) и 5.b) рассчитаны на одинаковых реализациях и соответствуют друг другу по цвету режимов.



a)



b)

Рисунок 5 — Оценки параметров HAR-RV модели в зависимости от режима a) соответствует реализации алгоритма на рис. 4.a), b) соответствует реализации алгоритма на рис. 4.b)

В следующем разделе рассмотрим задачу прогнозирования реализованных волатильностей на 1 день.

5 Прогнозирование волатильности при помощи HAR-RV модели

Прогноз релизованной волатильности на 1 день описывается выражением

$$\overline{RV}_{t+1} = \beta_0^* + \beta_1^* RV_t + \beta_2^* RVW_t + \beta_3^* RVM_t,$$

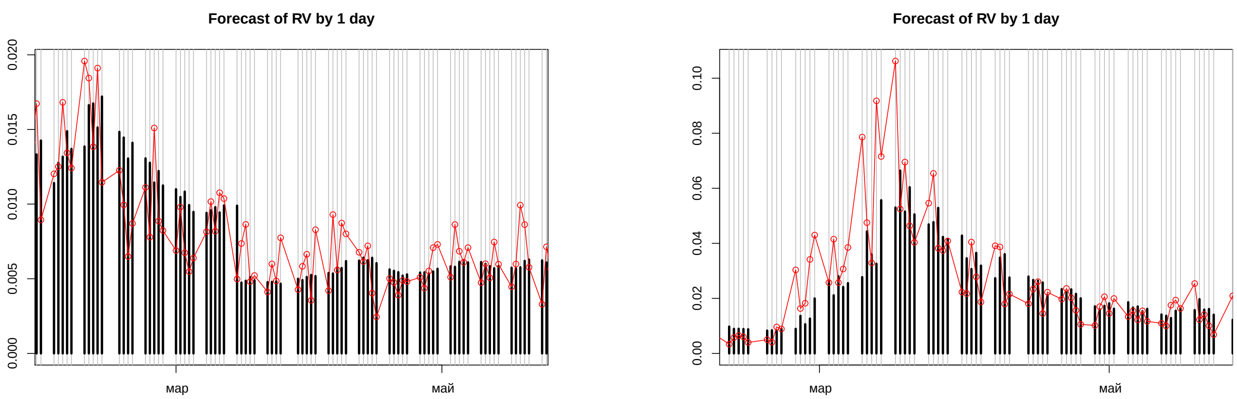
где

$$\beta_0^* = \beta_0^{\mathrm{Reg}(RV_t)}, \quad \beta_1^* = \beta_1^{\mathrm{Reg}(RV_t)}, \quad \beta_2^* = \frac{1}{5}\sum_{i=0}^{4}\beta_2^{\mathrm{Reg}(RV_{t-i})}, \quad \beta_3^* = \frac{1}{22}\sum_{i=0}^{21}\beta_3^{\mathrm{Reg}(RV_{t-i})},$$

$\mathrm{Reg}(x)$ – операция, определяющая в каком режиме находится наблюдение x. Т.е. имеем модель с динамически меняющимися коэффициентами в зависимости от режима, в котором находится наблюдение в момент t. Более того, коэффициенты β_2 и β_3 усредняются по наблюдениям, если алгоритм находится в точке перехода из режима в режим.

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 HMM и за ее пределами. Прогноз строился за пределами выборки оценивания параметров HAR-RV.

На рис. 6 представлены прогнозы на 1 день в разные временные интервалы. На рис. 6.a) рассмотрена ситуация перехода с высоковолатильного режима на низковолатильный. На рис. 6.b) 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.



a) b)

Рисунок 6 — Прогноз реализованной волатильности на 1 день (красная линия – реализованная волатильность, черные столбцы – прогноз на 1 день): a) с 01.02.2016 до 20.05.2016, b) с 12.02.2020 до 20.05.2020

ЛИТЕРАТУРА

1. Corsi F. A Simple Approximate Long-Memory Model of Realized Volatility // Journal of Financial Econometrics. – 2009. – Vol. 7, № 2. – P. 174 – 196.

2. Feng Ma. Volatility forecasting: long memory, regime switching and heteroscedasticity / Feng Ma, Xinjie Lu, Ke Yang, Yaojie Zhang // Applied Economics, DOI:10.1080/00036846.2019.1589645.