ВЫСШАЯ ШКОЛА ЭКОНОМИКИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

Лекция 3. Модели волатильности

Анализ данных с временной структурой

Почему линейных не хватает?

На практике в данных могут возникать самые разные феномены, которые линейная модель описать не может.
Например, если мы смотрим на цены, то могут возникать: кластеризация, катастрофическое изменение, тяжёлые хвосты распределений, наличие «долгой памяти» у цен и её свойств и так далее. Для того, чтобы как-то описать их, обращаются к нелинейным моделям.

Волатильность

Волатильность представляет собой меру риска используемого финансового инструмента на заданном промежутке времени.

Данная величина, как правило, рассчитывается как абсолютное или относительное значение от начальной стоимости в виде статистического показателя выборочного стандартного отклонения.

Волатильность

Волатильность акций — это изменение дисперсии доходности акций в течение определенного периода времени.

Обычно сравнивают волатильность одной акции с другой, чтобы получить представление о том, какая может иметь больший риск, или с рыночным индексом, чтобы понять волатильность акций относительно рынка.

Как правило, чем выше волатильность, тем рискованнее инвестиции в эту акцию. Необходимо отметить, что она не является постоянной и изменяется с течением времени.

Логарифмическая доходность

Если $S(t)$ является ценой базового актива (акции, индекса), то доходность на интервале Δt:

$$r_{\Delta t}(t) = \frac{S(t+\Delta t) - S(t)}{S(t)};$$

Так как

$$\frac{S(t+1) - S(t)}{S(t)} \approx \ln S(t+1) - \ln S(t),$$

то на практике часто работают с логарифмической доходностью:

$$\ln S(t+1) - \ln S(t).$$

Логарифмическая доходность

- Плюс использования доходности по сравнению с ценой самого актива - нормализация. Сравнивать абсолютные цены двух активов менее корректно, чем сравнивать их доходности.
- Расчет корреляции разных активов корректнее проводить на доходностях, а не на ценах актива.

Логарифмическая доходность

Мы используем логарифмические доходности:

1. Логнормальность: если мы предполагаем, тчо цены распределены логнормально (стандарная модель для рынка), то $\log(1+r_i)$ распределены нормально (r_i – доходность):

$$1 + r_t = \frac{S(t+1)}{S(t)} = \exp^{\log(\frac{S(t+1)}{S(t)})}.$$

2. В случае доходностей, близких к нулю, мы можем безболезненно использовать чистые приращения, потому что:

$$\log(1+r) \approx r, r \ll 1$$

Логарифмическая доходность

Мы используем логарифмические доходности:

3. Аддитивность по времени:
   рассмотрим n последовательных сделок. Накопленная доходность:

$$(1+r_1)(1+r_2)\cdots(1+r_n) = \prod_i (1+r_i)$$

Данной формулой пользоваться неудобно, так как произведение нормальных величин не является нормальной случайно величиной в общем случае. Однако сумма таких нормально распределенных величин будет нормальной (они некоррелированы, например,по условию):

$$\log(1+r_t) = \log(\frac{S(t+1)}{S(t)}) = \log(S(t+1)) - log(S(t))$$

Логарифмическая доходность

Поэтому накопленные лог доходности будут распределены нормально и

$$\sum_i \log(1+r_i) = \log(1+r_1) + \log(1+r_2) + \cdots + \log(1+r_n) =$$

$$= \log(S(n)) - \log(S(0))$$

Стандартное отклонение ежедневной логарифмической доходности актива принято называть дневной волатильностью актива.

Например, для индекса DJIA, **историческая** волатильность примерно равна $\sigma = 0.011$.

Оценка волатильности по разным данным

Мы можем оценить волатильность по разным данным:

- дневная доходность акций за каждый торговый день
- внутридневная динамика торгов
- стоимость опционов на акции Роснефти (если они есть)

Оценка волатильности по разным данным

Упомянутые на прошлом слайде источники дают три возможных оценки волатильности:

1. Волатильность как условное стандартное отклонение ежедневных доходностей. Это стандартное определение волатильности, которое мы будем использовать сегодня.
2. Подразумевая волатильность (implied volatility). Используя рыночные цены опционов и одну из моделей ценообразования опционов (к примеру, Black-Scholes), мы можем оценить волатильность. Однако это оценка зависит от модели, которую вы используете для определения цены опциона.

Оценка волатильности по разным данным

3. Реализованная волатильность. Если у вас есть high frequency данные, вы можете оценивать внутридневную волатильность доходностей. К примеру, вы можете использовать, 5 минутные промежутки времени, для того, чтобы оценить дневную волатильность.

Обычно волатильность сообщается на годовом уровне (% в год). Если вы оценили дневную волатильность, то вы можете оценить годовую волатильность, умножив ее на $\sqrt{252}$, то есть квадратный корень из количества торговых дней в году (примерно равно 16).

Как моделировать доходность актива? Плюсы, минусы, подводные камни

Анализ ACF/PACF зачастую указывает на отсутствие автокорреляции остатков в временных рядах.

Однако такой анализ не учитывает, что после больших скачков, как правило, следуют большие скачки.

Этот факт называют **volatility clustering**. Для его выявления нужно построить коррелограмму квадратов или модулей остатков.

Как моделировать доходность актива? Плюсы, минусы, подводные камни

И другие эмпирические факты:

1. Цены активов нестационарны. Доходности стационарны.
2. Автокорреляция в рядах доходностей слабая.
3. Гипотеза о независимости между квадратами значений ряда доходностей обычно отвергается в пользу альтернативы о наличия нелинейной зависимости.
4. Волатильность доходностей кластеризована.
5. Нормальность рядов доходностей отвергается в пользую альтернативы о тяжелых хвостах распределения.
6. Эффект «рычага» - изменения цен акций отрицательно коррелируют с изменениями волатильности.

Условная гетероскедастичность

Набор элементов (к примеру, части временного ряда) является **гетероскедастичным**, если определенные подгруппы этих элементов имеют разную дисперсию.

Структура финансового рынка и поведение участников приводит к дополнительным причинам, почему увеличение дисперсии приводит к еще большему увеличению дисперсии в реальной жизни.

Если гетероскедастичность имеет автокорреляцию, то есть условна в зависимости от периода роста волатильности, тогда наблюдается **условная гетероскедастичность**.

Условная гетероскедастичность

Как и раньше, пусть $(\Omega, \mathcal{F}, P)$ — исходное вероятностное пространство, а $\varepsilon = (\varepsilon_n)$ — последовательность iid случайных величин с стандартным нормальным распределением (они будут моделировать «случайность» в рассматриваемых далее моделях). Далее, введём фильтрацию $(\mathcal{F}_n)_{n>0}$ по правилу:
$\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_n = \sigma(\varepsilon_1, \dots, \varepsilon_n)$.

Доходности

Пусть есть стационарная серия доходности r_t, тогда

Безусловное среднее (uncondtional mean) – $E(r_t)$ – ожидаемое значение доходности.
Условное среднее (conditioanl mean) – это ожидаемое значение доходности в момент времени t, учитывая всю более раннюю информацию $E(r_t|\mathcal{F}_{t-1})$. Условное среднее зависит от времени.

Безусловная дисперсия (unconditional variance) $\sigma^2 = Var(r_t)$ – рассчитывается по стандартной формуле.
Условная дисперсия (conditional variance) $\sigma_t^2 = Var_t(r_t|\mathcal{F}_{t-1})$ – оценки дисперсии в момент времени t с учетом всей предыдущей информации.

ARCH и GARCH модели

Эффект условной гетероскедастичности сложно оценить по корелограммам. Можно использовать модели волатильности, такие как ARCH или GARCH

Одномерные ARCH модели появились в литературе с работой Engle (1982), и были вскоре обобщены в модель GARCH в работе Bollerslev (1986). Изначально их использовали для предсказания волатильности ряда инфляции, однако модели оказались релевантны и для предсказания волатильности финансовой доходности, наблюдаемой на месячных и более частых данных. Это позволяет изучать межвременное соотношение между риском и ожидаемой доходность.

Когда можно использовать ARCH?

Мы использовали корелограмму доходностей для идентификации AR(1)-процесса. Точно также можно использовать корелограмму квадратов доходностей для идентификации ARCH

- ARCH необходимо использовать, если вы уже подобрали адекватную модель, которая “оставляет” после себя остатки, более или менее похожие на белый шум. То есть, сначала необходимо определить модель для среднего доходностей, а потом моделировать дисперсию.
- ARCH не имеет смысла использовать для серий, которые имеют сезонные и/или трендовые эффекты. Сначала “уберите” сезонность/тренд с помощью ARIMA (SARIMA), экзогенных регрессоров, а после этого оценивайте ARCH.

Достоинства ARCH

+ Модель может “создавать” кластеры волатильности
+ Модель может “создавать” heavy tails

Недостатки ARCH

- Модель трактует положительные и отрицательные шоки одинаковым образом с точки зрения влияния на волатильность.
- Модель ARCH накладывает достаточно сильные ограничения на значения коэфф-ов, что ограничивает возможности моделировать более высокие моменты.
- Модель ARCH не дает возможность оценки источников шоков, это просто механистический способ оценки поведения условной дисперсии.
- Модель ARCH недооценивает волатильность, так как достаточно медленно реагирует на большие по значению шоки.
- Модель на фактических требует достаточно большого количество параметров для оценки, что увеличивает вероятность overfitting.

ARCH(p)

Модель **ARCH** порядка p имеет следующий вид:

Будем называть последовательность случайных величин $h = (h_n)$ ARCH(p)-моделью, если $h_n = \sigma_n \varepsilon_n$, где

$$\sigma_n^2 = \alpha_0 + \sum_{k=1}^{p} \alpha_i h_{n-k}^2,$$

$\varepsilon = (\varepsilon_n)$ — последовательность iid случайных величин с стандартным нормальным распределением,

$\alpha_0 > 0, \alpha_i > 0, h_0 = h_0(\omega)$ — случайная величина, не зависящая от $\varepsilon = (\varepsilon_n)_{n \geq 1}$.

Обычно h_0 полагают либо константой, либо случайной величиной, для которой второй момент выбирается из соображений «стационарности» значений $E[h_n^2]$.

GARCH (Generalized ARCH)

GARCH(p,q):

$$h_n = \sigma_n \varepsilon_n,$$

$$\sigma_n^2 = a_0 + \sum_{i=1}^{q} a_i h_{n-i}^2 + \sum_{j=1}^{p} b_j \sigma_{n-j}^2.$$

Основное преимущество GARCH(p, q)-моделей перед ARCH(p)-моделью, состоит в подборе параметров модели. На практике периодически может оказаться так, что при подгонке статистических данных моделями ARCH(p) параметр p становится слишком большим (что усложняет анализ модели), в то время как при подгонке GARCH(p, q)-моделями можно ограничиваться небольшими значениями p и q.

GARCH – это ARMA-процесс для дисперсии серии.

Использованные источники

1. Субботин А.В. Моделирование волатильности: от условной гетероскедастичности к каскадам на множественных горизонтах. – Журнал Прикладная Эконометрика. – 2009.

2. Ширяев А.Н. Основы стохастической финансовой математики. Т. 1. Факты. Модели. – М.: «ФАЗИС». – 1998.

3. Bulcsú Sándor. Time-scale effects on the gain-loss asymmetry in stock indices.: Physical Review. – 2016.

4. Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. – 2001. – pp. 223-236.