

УДК 517.534

Б.В. Шамша, Т.Б. Шатовская, Л.А. Христова

Харьковский национальный университет радиоэлектроники, Харьков

ПРОГНОЗИРОВАНИЕ ВОЛОТИЛЬНОСТИ ДЛЯ ОЦЕНКИ СТЕПЕНИ РИСКА В ДИЛИНГОВЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

В статье рассматривается проблема оценки риска в стохастических временных рядах в условиях гетероскедастичности. Предлагается модель прогнозирования волатильности ряда с учетом погрешности оценки модели на предыдущих лагах. Предложен принцип построения GARCH моделей для различных временных рядов.

волатильность, риск, GARCH модель, гетероскедастичность, временной ряд

Введение

Постановка проблемы в общем виде. Управление финансовыми ресурсами в различных предметных областях связано с наличием риска. В связи с этим управление предприятиями в условиях риска требует качественно новых подходов принятия решений. В настоящее время существует множество понятий «риск». Часто «риск» интерпретируют как «опасность потерь». Возникает ситуация, когда наступление некоторых событий приводит к наличию материальных, физических, социальных, моральных и т.п. потерь. В таких условиях требуется предсказать появление неблагоприятных событий и измерить величину потерь. Другими словами необходимо построить модель риска и по ней измерить величину ущерба.

Временные ряды финансовых и экономических показателей характеризуются наличием стохастической волатильности, «тяжелых хвостов» в распределениях, прошлой зависимостью, кластеризацией экстремумов, фрактальностью, хаотичностью.

Состояние объекта будем называть рискованным, если действие различных возмущений на объект приводит к неопределенности. Принятие решений в таких условиях может привести объект к неблагоприятным последствиям, т.е. эффективность функционирования объекта становится низкой и приводит к потерям, убыткам. Измерение риска сводится к измерению потерь.

Анализ последних исследований и публикаций. Концепция Value-At-Risk [1] предлагает строить вероятностную оценку потерь. Риск связан с принятием решений в условиях неопределенности, которая не поддается математическому описанию и не измеряется ни количественно, ни качественно. Оценки риска можно провести исходя из теории рационального выбора и оценки предпочтения состояния. Неопределенность можно описать как конечное множество взаимоисключающих состояний $S = \{S_1, S_2, \dots, S_n\}$, при этом каждому состоянию S_i , $i = \overline{1, n}$ определяется его вероятностная оценка $P(S_i)$. Реализация конкретного состояния S_i полно-

стью определяет значение всех экзогенных переменных (т.е. состояние внешней среды). Субъект совершает выбор, способен ранжировать свои предпочтения в зависимости от вероятностных оценок.

Для вычисления меры риска используются различные методы. Наиболее распространенными из них являются метод постоянных ковариаций, метод экспоненциально-взвешенных ковариаций, методы, в основе которых лежит построение GARCH моделей, полупараметрических моделей, историческое моделирование, модели, использующие теорию экстремальных значений.

Цель статьи. Экономические временные ряды характеризуются изменчивостью статистических характеристик, в частности дисперсии, которые меняются в зависимости от типа рынка, времени, возмущений и других факторов.

В некоторые периоды наблюдается изменчивость в широком и в узком смысле стационарности характеристик. Так в случае стационарности имеет место низкая дисперсия, а в другие периоды наблюдается увеличение дисперсии, которая образует своего рода «пучки», «clustering volatility» (кластеризация волатильности)

Интерес представляет исследование и оценка применимости методов и моделей VAR для различных временных рядов рынков.

Для прогнозирования риска используем свойство условной дисперсии процесса. Пусть условная дисперсия стохастического скалярного процесса ε_t описывается линейной функцией квадратов предшествующих реализаций ряда вида

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2. \quad (1)$$

Для положительности σ_t^2 необходимо, чтобы

$$\omega > 0, \quad \alpha_1, \dots, \alpha_p \geq 0.$$

Исходя из (1), высокие по абсолютному значению реализации процесса в недалеком прошлом влекут увеличение условной дисперсии в момент t , и следовательно, условной вероятности появления вновь высокой по модулю реализации процесса ε .

Небольшие значения $\varepsilon_{t-1}, \dots, \varepsilon_{t-p}$ приводят к снижению этой вероятности. За большими (по абсолютному значению) наблюдениями вновь последуют большие наблюдения, за малыми – малые. Таким образом, по изменчивости условной дисперсии финансовых временных рядов можно судить об уровне риска, неопределенности, которая связана с прогнозированием основных индикаторов рынка.

Основной материал исследования

Рассмотрим обобщенный ARCH процесс [2] в виде GARCH модели

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \cdot \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \cdot \sigma_{t-j}^2.$$

Такая модель имеет бесконечную память и обладает более экономичной параметризацией.

GARCH модель представим в виде авторегрессии скользящего среднего в виде

$$\varepsilon_t^2 = \omega + (\alpha(L) + \beta(L))\varepsilon_t^2 + (1 - \beta(L))v_t, \quad (2)$$

$$v_t = \varepsilon_t^2 - \sigma_t^2.$$

Квадраты ошибок подчиняются ARMA модели.

Следует заметить, что согласно леввередж эффекту (leverage effect) в ситуациях, которые приводят к снижению доходов, увеличивается волатильность и повышается рискованность вложений в фирму. Согласно модели (2) существует корреляция между ε_t и σ_{t+1}^2 . В GARCH моделях имеет место, что распределение ε симметрично, и будущее значение условной дисперсии не коррелирует с текущей ошибкой прогноза. Это на практике не всегда имеет место.

Значения параметров модели (2) имеют ограничения, что приводит к трудностям в их оценке.

Модель ARCH-M предполагает функциональную зависимость условного среднего случайной величины y_t от собственной условной дисперсии. В таких моделях при нарастании неопределенности доходность увеличивается. Для решения задачи прогнозирования волатильности интерес представляет разработка модели с учетом корректировки ошибок модели.

Рассмотрим построение GARCH моделей прогнозирования с помощью методов экспоненциального сглаживания, авторегрессии и проинтегрированного скользящего среднего (АРПСС) с коррекцией прогнозируемого значения ошибки предсказания, так как эти методы основаны на знании информации о предыстории процесса.

Сложность прогнозирования динамических рядов, характеризующих сущность изменения валютных курсов, биржевых цен, ценных бумаг и других экономических показателей состоит во временном изменении данных. Действие различных возмуще-

ний на такие ряды непостоянны и могут приводить ряд из однородного в неоднородный, и если имеет место коррелированность между наблюдениями в момент времени t_1 , то в момент времени t_2 она может отсутствовать.

Таким образом, с течением времени такой ряд может менять как качественные характеристики, так и количественные. В соответствии с вышеупомянутым, выбор метода прогнозирования будет зависеть от статистических характеристик исходного ряда. Пользователь, на основании информации о эффективности алгоритмов и методов прогнозирования, для каждой конкретной ситуации должен автоматически или автоматизировано выбрать наиболее оптимальный метод. В этой связи возникают несколько глобальных проблем.

Первая проблема – текущая оценка состояния данных и их классификация.

Вторая проблема – оценка эффективности методов для каждого состояния данных.

Третья проблема – выбор метода прогноза для конкретного состояния.

Четвертая проблема – оценка адекватности модели прогноза и определение целесообразности использования ее принятия решений.

Таким образом, динамический ряд в общем виде представлен в виде последовательных числовых значений, зафиксированных в определенные промежутки времени, которые изменяют свои статистические характеристики. Другими словами динамический ряд может быть представлен сезонными составляющими, которые со временем могут меняться и даже вообще исчезать, с течением времени возможно изменение направленности тренда и далее его отсутствие.

Для описания динамического объекта или явления часто используется управление распределенного лага

$$Y_t + \sum_{i=1}^k P_i Y_{t-i} = \sum_{i=1}^k P_i X_{t-r-i},$$

где Y_t – входной сигнал, в момент времени t зависит от выходных сигналов на k шагов назад и от входного сигнала X на $r-i$ шагов назад во времени; r – чистое запаздывание объекта.

В общем случае динамическую модель можно представить в виде обобщенной регрессионной модели в виде:

$$Y = F_{y,x} \beta + \varepsilon \quad (3)$$

где $Y^T (Y_t, Y_{t+1}, \dots, Y_{t+N-1})$, $X^T (X_t, X_{t+1}, \dots, X_{t+N-1})$, $\varepsilon^T (\varepsilon_t, \varepsilon_{t+1}, \dots, \varepsilon_{t+N-1})$ – $(N \times 1)$ мерные вектора наблюдения выходного сигнала, входного сигнала и возмущающего воздействия соответственно;

$\hat{\beta}^T = (-P_1, -P_2, \dots, -P_k, q_1, q_2, \dots, q_k)$ – вектор оцінюваних параметрів, розмірністю $(R \times 1)$.

Уравнение (3) в отличие от регрессионной модели статистики содержит в матрице плана различные значения входных и выходных сигналов в различные моменты времени.

Для такой модели характерна корреляция компонентов эквивалентного возмущающего воздействия ε друг с другом.

Другая причина появления неоднородности в динамической модели – это корреляция между частью элементов матрицы плана и элементами эквивалентного возмущения ε . Корреляция возникает из-за зашумленности входного сигнала Y в некоторых столбцах матрицы $F_{y,x}$, в силу чего эти элементы, как и элементы ε , оказываются зависимы от компонент вектора шума.

Природа рынка валют на мировых биржах подвержена действию или зависимости дисперсии $\delta^2(y)$ от его математического ожидания $E\{y\}$. Более того, дисперсия данных зависит от временного фактора.

Модель АРПСС прогноза величины σ_{t+l}^2 в момент t с упреждением l получим из разностного уравнения с учетом условного математического ожидания в виде:

$$\sigma_{t+l}^2 = p_1[\sigma_{t+l-1}^2] + \dots + p_{p+d}[\sigma_{t+l-p-d}^2] - q_1[a_{t+l-1}] - \dots - q_q[a_{t+l-q}] + [a_{t+l}] + \frac{1}{k} \sum_{z=0}^k \hat{\varepsilon}_{t-z}(1+z). \quad (4)$$

Для модели АР прогноз получим из уравнения

$$[\sigma_{t+l}^2] = p_1[\sigma_{t+l-1}^2] + \dots + p_{p+d}[\sigma_{t+l-p-d}^2] + [a_{t+l}] + \frac{1}{k} \sum_{z=0}^k \hat{\varepsilon}_{t-z}(1+z). \quad (5)$$

где k – количество данных, участвующих в построении модели прогноза.

Величина $\frac{1}{k} \sum_{z=0}^k \hat{\varepsilon}_{t-z}(1+z)$ – это средневзвешенная величина ошибки прогноза, которую необходимо идентифицировать.

Допустим мы прогнозируем величину $\sigma_t^2(1)$ в момент t с упреждением 1. В момент t мы знаем реальное значение σ_t^2 . Ошибка прогноза составляет $\sigma_t^2 - \hat{\sigma}_{t-1}^2(1) = e_t$. С упреждением $l = 2$ из точки t получить значение ошибки прогноза в точке $t+1$ невозможно, так как неизвестно реальное значение ряда в точке $t+1$. Но прогнозное значение получить реально, если данные ряда до точки t представить в

виде прогнозных значений ошибок из точек $t-1, t-2, \dots, t-n$ с упреждением $l = 1, l = 2, l = 3, \dots, l = k$. Заметим, что прогнозное

значение $\hat{\sigma}_t^2(l)$ есть случайная величина и мало вероятно, что в будущем истинное значение ряда σ_{t+1}^2 совпадает с прогнозным $\hat{\sigma}_t^2(l)$. Поэтому при прогнозировании ошибки e_t устанавливаются доверительный интервал значений, куда попадает истинное значение с заданной вероятностью. Этот интервал выразим через прогнозное значение $\hat{e}_t(l)$ и оценку дисперсии δ_ζ^2 величины $e_{t+1} - e_{t+1} - \hat{e}_t(1) = \zeta_{t+1}$.

Трудность в определении δ_ζ^2 состоит в оценке величины упреждения для получения динамического ряда. Полученные прогнозное значение будем использовать для корректировки прогнозов.

Подправление прогноза на l упреждений можно реализовать двумя способами. В первом способе используем информацию для подправления прогнозов об ошибке на предыдущих шагах.

Из (5) представим математические модели прогноза с момента t для упреждений $1, 2, \dots, l$.

Для $l = 1$ с учетом условного математического ожидания получим:

$$\hat{\sigma}_t^2(1) = p_1\sigma_t^2 + p_2\sigma_{t-1}^2 + \dots + p_{p+d}\sigma_{t-p-d}^2 - q_1a_t - q_2a_{t-1} - \dots - q_qa_{t+1-q};$$

для $l = 2$

$$\hat{\sigma}_t^2(2) = p_1\hat{\sigma}_t^2(1) + p_2\sigma_t^2 + \dots + p_{p+d}\sigma_{t+2-p-d}^2 - q_2a_t - \dots - q_qa_{t+2-q};$$

для $l = 3$

$$\hat{\sigma}_t^2(3) = p_1\hat{\sigma}_t^2(2) + p_2\hat{\sigma}_t^2(1) + \dots + p_{p+d}\sigma_{t+3-p-d}^2 - q_3a_t - \dots - q_qa_{t+3-q};$$

для $l = 4$

$$\hat{\sigma}_t^2(4) = p_1\hat{\sigma}_t^2(3) + p_2\hat{\sigma}_t^2(2) + p_3\hat{\sigma}_t^2(1) + \dots + p_{p+d}\sigma_{t+4-p-d}^2 - q_4a_t - \dots - q_qa_{t+4-q}$$

и так далее. С увеличением величины упреждения l корректирующие значения величины ошибки сдвигаются к t . При $l > k$ значения a_{t+l-q} будут равны нулю.

Если порядок полинома $Q(B)$ будет больше величины упреждения l , то количество корректирующих ошибок определяется значением $t+l-q$. При $l = q$ коррекция будет определяться значениями a_t, a_{t-1} .

Итак прогнозы с большими упреждениями зависят от a_t, a_{t-1}, a_{t-2}, l и q .

При получении действительного значения

σ_{t+1}^2 , вычисляется новая совокупность прогнозов с учетом реальной ошибки $a_{t+1} = \sigma_{t+1}^2 - \hat{\sigma}_t^2(1)$. Появление нового реального значения, допустим, σ_{t+2}^2 , позволит скорректировать новую последовательность прогнозов.

Второй способ корректировки прогнозов основывается на знании весов ψ_ℓ , которые вычисляются в соответствии с выражениями (1).

Прогноз с точки $t+1$ с упреждением ℓ будет равен прогнозу с точки t с упреждением $\ell+1$ с добавлением ошибки $a_{t+\ell} = \sigma_{t+\ell}^2 - \hat{\sigma}_t^2(\ell)$ с весом ψ_1 , т.е

$$\sigma_{t+1}^2(\ell) = \sigma_t^2(\ell+1) + \psi_\ell + a_{t+1}.$$

Итак, новый прогноз $\hat{\sigma}_{t+1}^2(\ell)$ для упреждения ℓ последним способом найти нельзя, но он легко вычисляется по прогнозам с меньшими упреждениями из разностного уравнения.

Зная величину ошибки a_{t+1} и используя ψ_ℓ для соответствующего упреждения ℓ можно вычислить прогноз с упреждением ℓ .

И в первом и во втором случае величина ошибки прогноза зависит от величины упреждения ℓ , оценки ошибки a_t , величины $\hat{\sigma}_{t+1}^2$, коэффициентов ψ_ℓ .

Исследуем влияние величины упреждения на коррелированность ошибок прогноза для метода авторегрессии и проинтегрированного скользящего среднего (АРПСС). Заметим, что ошибки прогноза с утверждением, равным единице, некоррелированные импульсы.

Зная коэффициенты p и q , вычисляем веса ψ по формулам:

$$\psi_0 = 1; \psi_1 = p_1 - q_1; \psi_2 = p_1\psi_1 + p_2 - q_2;$$

$$\psi_j = p_1\psi_{j-1} + \dots + p_p + \psi_j - p - d - q_j;$$

$$q_j = 0 \text{ при } j > q;$$

веса ψ_j необходимы для вычисления доверительных интервалов и корректировки прогноза.

Пусть известно реальное значение величины σ_t^2 в точке $t+1$, σ_{t+1}^2 известно, тогда прогноз на момент $t+1$ может быть скорректирован с учетом новой ошибки $a_{t+1} = \sigma_{t+1}^2 - \hat{\sigma}_t^2(1)$ и предсказанные значения с точки $t+1$ вычисляются следующим образом

$$\hat{\sigma}_{t+1}^2(\ell) = p_1[\sigma_{t+1+\ell-1}^2] + \dots + p_{p+q}[\sigma_{t+1+\ell-p-q}^2] + a_{t+1} + q_1[a_{t+1+\ell-1}] - \dots - q_q[a_{t+1+\ell-q}]$$

Учитывая веса ψ_j и используя прогнозы $\hat{\sigma}_t^2(1), \hat{\sigma}_t^2(2), \dots, \hat{\sigma}_t^2(L)$ с момента t , получим прогнозы с момента $t+1$ из формулы

$$\hat{\sigma}_{t+1}^2(\ell) = \sigma_t^2(\ell-1) + \psi_\ell a_{t+1} = \sigma_t^2(\ell+1) + \psi_\ell[\sigma_{t+1}^2 - \hat{\sigma}_t^2(1)].$$

Итак, ошибка прогноза в предыдущей точке t используется для корректировки прогноза в точке $t+1$. Такой подход улучшает прогноз, но имеет смысл использовать для прогноза в точке ℓ еще предсказанные значения ошибки. Для вывода такого выражения определим вначале величину корреляции между ошибками, сделанные в один и тот же момент времени с различными упреждениями.

Предположим, что имеет место прогнозы с момента t для различных упреждений. Тогда ошибки таких прогнозов представим в виде

$$e_t(\ell) = \sigma_{t+1}^2 - \hat{\sigma}_t^2(\ell) = a_{t+\ell} + \psi a_{t+\ell-1} + \dots + \psi_{\ell-1} a_{t+1},$$

$$e_t(\ell+j) = \sigma_{t+\ell+j}^2 - \hat{\sigma}_t^2(\ell+j) = a_{t+\ell+j} + \psi a_{t+\ell+j-1} + \dots + \psi_j a_{t+1} + \psi_{j+1} a_{t+\ell-1} + \dots + \psi_{\ell+j-1} a_{t+1}$$

Определим ковариацию между прогнозами на момент t с упреждениями ℓ и $\ell+j$ в виде

$$\text{Cov}[e_t(\ell), e_t(\ell+j)] = \delta_a^2 \sum_{i=1}^{\ell-1} \Psi_i \Psi_{i+j}.$$

Коэффициент корреляции с момента t и с упреждениями

$$r_{\ell, \ell+j} = \frac{\sum_{i=0}^{\ell-1} \Psi_i \Psi_{i+j}}{\left(\sum_{h=0}^{\ell-1} \Psi_h^2 \sum_{g=0}^{i+j-1} \Psi_g^2 \right)^{1/2}}$$

Рассмотрим прогнозы для упреждения ℓ , сделанные в моменты t , и $t-j$, где j – положительное целое число. При $j = \ell, \ell+1, \ell+2, \dots$, ошибки прогноза не будут содержать общих компонент, однако для $j = 1, 2, \dots, \ell-1$ некоторые из a будут входить в ошибки обоих прогнозов.

Итак:

$$e_t(\ell) = \sigma_{t+\ell}^2 - \hat{\sigma}_t^2(\ell) = a_{t+\ell} + \psi_1 a_{t+\ell-1} + \dots + \psi_{\ell-1} a_{t+1};$$

$$e_{t-j}(\ell) = \sigma_{t+\ell-j}^2 - \hat{\sigma}_{t-j}^2(\ell) = a_{t-j+\ell} +$$

$$\psi_1 a_{t-j+\ell-1} + \dots + \psi_{\ell-1} a_{t-j+1}$$

и для автоковариации ошибок прогноза с упреждением ℓ для задержки j равна

$$\text{Cov}[e_t(\ell), e_{t-j}(\ell)] = \delta_a^2 \sum_{i=j}^{\ell-1} \Psi_i \Psi_{i-j},$$

где $\Psi_0 = 1$. Соответствующие корреляции равны

$$r[e_t(\ell), e_{t-j}(\ell)] = \frac{\sum_{i=j}^{\ell-1} \Psi_i \Psi_{i-j}}{\sum_{i=0}^{\ell-1} \Psi_i^2};$$

$$0 \leq j < \ell; j \geq 1$$

Генерация числовых последовательностей временных рядов без сезонных составляющих показала, что ошибки прогноза $e_t(\ell)$ и $e_t(\ell+j)$, сделанные для разных упреждений с одного и того же момента времени t , коррелированы. Функция прогноза в

этом случае лежит целиком выше, либо целиком ниже фактических значений ряда.

Ошибки прогноза $e_t(\ell)$ и $e_t(t-j)$, сделанные при том же упреждении в различные моменты времени ℓ и $t-j$ также коррелированы. Оптимальные ошибки прогноза с упреждением на шаг в перед ($\ell=1$) некоррелированы, а ошибки прогноза с большим упреждением будут коррелированы.

Выводы

На основании модели волатильности возможно определить степень риска. Заметим, что построить модель предсказания по известным ошибкам прогноза во времена $a_t, a_{t-1}, a_{t-2}, \dots$ возможно только тогда, когда ошибки прогноза коррелированы. В общем случае для построения модели прогноза волатильности можно использовать методы авторег-

рессии, скользящего среднего и экспоненциального сглаживания.

Список литературы

1. Engle R.F. Estimates of the variance of US inflation based on the ARCH model // *Journal of money, credit and banking*. – 1983. – №15. – Pp. 286-301.

2. Bollerslev, Tim, Ray Y. Chou and Kenneth F. Kroner. ARCH Modeling in Finance: A review of the theory and empirical evidence // *Journal of econometrics*. – 1992. – № 52. – Pp. 5-59.

Поступила в редколлегию 22.10.2007

Рецензент: д-р техн. наук, проф. Е.Г. Петров, Харьковский национальный университет радиоэлектроники, Харьков.