Статистический анализ быстроменяющихся нестационарных случайных процессов на основе многосвязных цепей Маркова

Жук Е.Е.; Башкевич А.Ю.;
Кафедра математического моделирования и анализа данных
Белорусский государственный университет
Минск, Республика Беларусь
e-mail: zhukee@mail.ru, bashkevichay@gmail.com

***Аннотация*—В данной работе исследуется поведение временных рядов на основе многосвязных цепей Маркова, а также их применение при анализе и статистическом описании математических моделей различных реальных объектов.**

Ключевые слова: многосвязные цепи Маркова; начальное распределение вероятностей; матрица вероятностей одношаговых переходов.

I. Введение

Для анализа экономических процессов удобно использовать математические модели случайных процессов. При наличии адекватной математической модели, описывающей экономический процесс, по результатам анализа данной модели можно производить оценку параметров модели, ее уточнение и использовать данные о поведении модели для прогнозирования характеристик наблюдаемого процесса.

В данной работе исследуется поведение быстроменяющихся случайных процессов на основе многосвязных цепей Маркова. Статистические оценки параметров этих цепей применяются для прогнозирования динамики цен на финансовых и валютных рынках.

II. Многосвязные цепи Маркова

Определение. Временной ряд $x_1,...,x_T, x_T,... \in S$, где $S=\{1,...,L\}$, называется многосвязной цепью Маркова [1] с глубиной k и пространством состояний S, образованным из $L \geq 2$ состояний, если

$$\begin{aligned} &P\{x_{t+1}=d_{t+1} \mid x_t=d_t, x_{t-1}= \\ &=d_{t-1},...,x_1=d_1\}= \\ &=P\{x_{t+1}=d_{t+1} \mid x_t=d_t, x_{t-1}= \\ &=d_{t-1},...,x_{t-k+1}=d_{t-k+1}\}= \\ &=p_{(d_{t-k+1},...,d_{t-1},d_t),d_{t+1}}, \end{aligned} \quad (1)$$

$$\forall d_{t-k+1}, d_{t-k+2},... \in S, \forall t=1,...T,...$$

Для того, чтобы однозначно определить k-связную цепь Маркова, необходимо задать две вероятностные характеристики [1]:

1) начальное распределение вероятностей:

$$\begin{aligned} &\pi^{(1)}_{(i_1,...,i_k)}=P\{x_1=i_1,...,x_k=i_k\}, \\ &\sum_{i_j \in S} \pi^{(1)}_{i_j}=1; i_j \in S, j=\overline{1,k}; \end{aligned} \quad (2)$$

2) матрицу вероятностей одношаговых переходов:

$$\begin{aligned} &P=(p_{(i_1,...,i_k),j})i_r, j \in S; r=\overline{1,k}, \\ &p_{(i_1,...,i_k),j}=P\{x_{t+1}=j \mid x_t= \\ &=i_k,...,x_{t-k+1}=i_1\}, \\ &\sum_{j \in S} p_{(i_1,...,i_k),j}=1, i_r \in S, r=\overline{1,k}. \end{aligned} \quad (3)$$

Распределение вероятностей реализации $X=\{x_1,...,x_T\}$ длительности $T \geq 2$ k-связной цепи Маркова (1) – (3) однозначно определяется через начальное распределение вероятностей $\pi^{(1)}=(\pi^{(1)}_{(i_1,...i_k)})_{i_j \in S, j=\overline{1,k}}$ и матрицу вероятностей одношаговых переходов $P=(p_{(i_1,...i_k),j})i_r, j \in S; r=\overline{1,k}$:

$$\begin{aligned} &P_T(d_1,...,d_T)= \\ &=P\{x_1=d_1,...,x_T=d_T\}= \\ &=P\{x_1=d_1, x_2=d_2,...,x_k=d_k\} \times \\ &\times \prod_{t=k}^{T-1} P\{x_{t+1}=d_{t+1} \mid x_t=d_t,...,x_1=d_1\} \\ &=\pi^{(1)}_{(d_1,...,d_k)} \prod_{t=k}^{T-1} p_{(d_{t-k+1},...,i_t),d_{t+1}}, \\ &d_1,...,d_T \in S. \end{aligned} \quad (4)$$

III. Оценка матрицы вероятностей одношаговых переходов методом максимального правдоподобия

Пусть имеется реализация $X=\{x_1,...,x_T\}$ длительности $T \geq 2$ k-связной цепи Маркова (1) – (3)

Ставится задача оценить матрицу вероятностей одношаговых переходов. Распределение вероятностей $\pi^{(1)}$ начального состояния оценить невозможно, поскольку оно встречается один раз. Для построения оценок $\hat{P}$ по реализации X воспользуемся методом максимального правдоподобия, основываясь на распределении вероятностей для X из (4).

Введем следующие обозначения:

$$n_{(i_1,...,i_k),j} = \\ = \sum_{t=k}^{T-1} \delta_{(x_{t-k+1},...,x_t),(i_1,...,i_k)} \delta_{x_{t+1},j} \quad (5)$$

– число одношаговых переходов из состояний $(i_1,...,i_k)$ в состояние j $(i_r, j \in S)$, встретившихся в реализации $X = \{x_1,...,x_T\}$ длительности $T \geq 2$;

$$n_{(i_1,...,i_k)} = \sum_{j \in S} n_{(i_1,...,i_k),j} \quad (6)$$

– число последовательных состояний $(i_1,...,i_k)$, встретившихся в реализации $X = \{x_1,...,x_T\}$.

Теорема. Пусть для реализации $X = \{x_1,...,x_T\}$ длительности $T \geq 2$ k-связной цепи Маркова (1) – (3) выполняется условие: $n_{(i_1,...,i_k)} > 0, i_r \in S, r = \overline{1,k}$ тогда оценка максимального правдоподобия (МП-оценка) матрицы вероятностей одношаговых переходов задается соотношением

$$\hat{P} = (\hat{p}_{(i_1,...,i_k),j})_{i_r, j \in S; r=\overline{1,k}},$$
$$\hat{p}_{(i_1,...,i_k),j} = \frac{n_{(i_1,...,i_k),j}}{n_{(i_1,...,i_k)}}. \quad (7)$$

IV. Построение модели на основе многосвязных цепей Маркова для анализа динамики цен

Необходимо проверить, возможно ли описание процесса изменения цены на финансовом рынке с помощью многосвязных цепей Маркова.

Пусть имеется реализация временного ряда $X = \{x_1,...,x_T\}$ длительности $T \geq 2$, где x_i – цена определенного продукта в момент времени i.

На основании ряда X построим новый ряд $Y = \{y_1,...,y_{T-1}\}$, длительности $T-1 \geq 2$, где каждое y_i принимает одно из значений множества $Q = \{0,1\}$ по следующему правилу: если $x_{i-1} \leq x_i$, то $y_i = 1$, если $x_{i-1} > x_i$, то $y_i = 0$.

Таким образом, построена цепь Маркова $Y = \{y_1,...,y_{T-1}\}$, размерности $T-1$ с пространством состояний $S = \{0,1\}$.

С помощью оценок максимального правдоподобия (7) оценим матрицу вероятностей одношаговых переходов, где $n_{(i_1,...,i_k),j}$ – количество переходов цепи Маркова из последовательных состояний $(i_1,...,i_k)$ в состояние j, а $n_{(i_1,...,i_k)}$ - количество переходов из состояний $(i_1,...,i_k)$ в любое другое, где все допустимые состояния имеют вид:

$$\begin{aligned} l_0 &= (0,0,...,0,0) \\ l_1 &= (0,0,...,0,1) \\ &... \\ l_{k-2} &= (0,1,...,1,1) \\ l_{k-1} &= (1,1,...,1,1) \end{aligned} \quad (8)$$

Данная матрица будет иметь нулевые элементы, которые соответствуют вероятностям недостижимых состояний, т.к. если цепь Маркова последовательно находилась в состояниях $(i_1,...,i_k)$, то в момент времени k+1 она может находиться в одном из состояний: $(i_2,...,i_{k-1},0)$ или $(i_2,...,i_{k-1},1)$.

На основе данной матрицы и осуществляется дальнейшее прогнозирование поведения цепи Маркова по алгоритму: если предыдущие состояния цепи $(i_{t-k+1},...,i_t)$, то следующее будет j:

$$p_{(i_1,...,i_k),j} = \\ = \max\{p_{(i_1,...,i_k),1}; p_{(i_1,...,i_k),0}\}. \quad (9)$$

И если $j = 1$, то произойдет рост цены либо она не измениться, а если $j = 0$, то прогнозируется падение цены.

[1] Ю.С. Харин, Е.Е. Жук. Математическая и прикладная статистика. – Мн: БГУ, 2005, С. 279.

[2] В.А. Таран. Играть на бирже просто?! 3-е изд. – СПб: Питер, 2007. С. 272.

[3] Ю.С. Харин. Теория вероятностей, математическая статистика. Задачи, упражнения, тестовые задания : учеб. пособие. – Мн: БГУ, 2009. С. 302.