УДК 519.688

А. О. Долгіх, О. Г. Байбуз

Дніпропетровський національний університет імені Олеся Гончара

ОГЛЯД СУЧАСНИХ РОЗРОБОК ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ З ВИКОРИСТАННЯМ ПРИХОВАНИХ МАРКІВСЬКИХ МОДЕЛЕЙ

Виконано огляд сучасних розробок у сфері використання марківських процесів до вирішення задачі прогнозування часових рядів. Розглянуто основні проблеми, які виникають у процесі використання прихованих марківських моделей (ПММ) у прогнозуванні, та алгоритми їх розв'язку. Переставлено базовий алгоритм прогнозування часових рядів з використанням ПММ. Зроблено огляд модифікацій базового алгоритму. Розглянуто коло відкритих питань у цій сфері досліджень.

Ключові слова: *прогнозування часових рядів; приховані марківські моделі; алгоритм прямого – зворотного ходу; алгоритм Вітербі; алгоритм Баума – Велша; фінансові часові ряди.*

Выполнен обзор современных разработок в области использования марковских процессов к решению задачи прогнозирования временных рядов. Рассмотрены основные проблемы, возникающие в процессе использования марковских моделей (СММ) в прогнозировании, и алгоритмы их решения. Представлен базовый алгоритм прогнозирования временных рядов с использованием СММ. Сделан обзор модификаций базового алгоритма. Рассмотрен круг открытых вопросов в этой области исследований.

Ключевые слова: *прогнозирование временных рядов; скрытые марковские модели; алгоритм прямого – обратного хода; алгоритм Витерби; алгоритм Баума – Велша; финансовые временные ряды.*

The review of modern developments in the area of time series forecasting using Markov processes has been performed. The main problems arising in the process of using Markov models (HMM) in forecasting and algorithms of their solution are considered. The basic algorithm of time series forecasting using HMM has been presented. The review of basic algorithm modifications has been performed. The range of questions to investigate in this area of research has been discussed.

Keywords: *time series forecasting; hidden Markov models; forward – backward algorithm; Viterbi algorithm; Baum – Welch algorithm; financial time series.*

Вступ. Ідея застосування марківських моделей до прогнозування часових рядів заснована на припущенні, що в основі коливань часового ряду лежить деякий марківський випадковий процес.

Марківським називається випадковий процес, який має таку властивість: уся передісторія процесу повністю зосереджена у теперішньому стані, лише теперішній стан впливає на майбутнє, і не важливо, як процес розвивався у минулому [1].

Прогнозування у випадку марківських моделей зводиться до того, щоб, проаналізувавши функціонування системи протягом деякого періоду часу, знайти значення параметрів випадкового процесу та, враховуючи поточний стан системи, з певною ймовірністю, передбачити, як система буде поводитися у майбутньому.

У звичайних марківських моделях кожному фізичному явищу відповідає певний стан системи. Спостерігач знає поточний стан системи, тому потрібно лише оцінити ймовірності переходів системи. Ця модель проста у реалізації, але, на жаль, занадто обмежена. Використовуючи її, неможливо розв'язати багато актуальних задач.

На практиці для вирішення багатьох реальних завдань використовують приховані марківські моделі. В даній статті наведено опис прихованих марківських моделей та алгоритм їх використання для розв’язку задачі прогнозування часових рядів.

Постановка задачі. У випадку прихованих марківських моделей (ПММ) розглядається система, яка представляє собою результат функціонування двох випадкових процесів. Перший – прихований процес, який спостерігач не може бачити, але цей процес безпосередньо впливає на другий процес – послідовність подій, яку бачить спостерігач.

На рис. 1 зображено схему прихованої марківської моделі. На представленій схемі ми бачимо, що система переходить з одного прихованого стану в інший $x(t-1)$, $x(t)$, $x(t+1)$ ... у той час, як насправді ми спостерігаємо події $y(t-1)$, $y(t)$, $y(t+1)$ ...

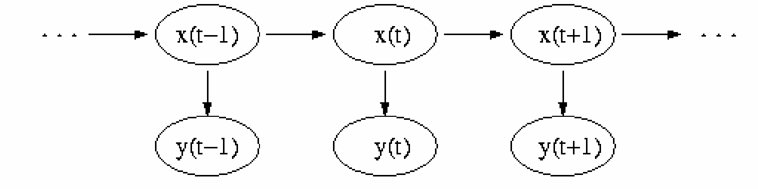

Рисунок 1 – Прихована марківська модель

Приховану марківську модель (ПММ) подають у вигляді $\lambda = \{S, \Omega, \pi, A, B\}$, де $S = \{s_1, ..., s_N\}$ – множина станів системи, $\Omega = \{w_1, ..., w_K\}$ – можливі події, $\pi = \{\pi_1, ..., \pi_N\}$ – початкові ймовірності, $A = \{a_{ij}\}$ – матриця переходів, а $B = \{b_{iwl}\}$ – ймовірність спостереження події w_l після переходу до стану s_i [10].

У процесі використання марківських моделей до вирішення задач аналізу даних та прогнозування постають три такі основні задачі [3]:

1) Нехай задана спостережувана послідовність $O = \{o_1, o_2, ..., o_T\}$ і прихована марківська модель $\lambda = \{A, B, \pi\}$. Необхідно обчислити $P(O | \lambda)$ – імовірність послідовності подій $O = \{o_1, o_2, ..., o_T\}$ для даної моделі λ. Для вирішення цієї проблеми використовують алгоритм прямого – зворотного ходу.

2) Нехай задана спостережувана послідовність $O = \{o_1, o_2, ..., o_T\}$ і модель $\lambda = \{A, B, \pi\}$. Необхідно підібрати послідовність станів системи $S = \{s_1, ..., s_T\}$, яка найкраще відповідає послідовності, що спостерігається, тобто «пояснює» спостережувану послідовність. Абсолютно точно визначити цю послідовність неможливо. Тут можна говорити лише про припущення з відповідним ступенем достовірності. Для розв’язку цієї задачі використовують алгоритм Вітербі.

3) Остання третя задача полягає в тому, щоб підібрати параметри моделі $\lambda = \{A, B, \pi\}$ таким чином, щоб максимізувати $P(O | \lambda)$. Це задача є основною для більшості проектованих ПММ. Вона полягає в оптимізації параметрів моделі на основі навчальної спостережуваної послідовності. Для навчання ПММ використовують алгоритм Баума – Велша.

Аналіз літературних даних та основний матеріал. Теорія прихованих марківських моделей не нова. Її основи опублікував Баум і його колеги в кінці 60-х – на початку 70-х років минулого століття. Тоді ж, на початку 70-х, Бейкер і Джелінек з колегами з IBM застосували ПММ у розпізнаванні мови.

Проте широке розповсюдження ПММ отримали нещодавно. За цей час вони показали досить якісні результати у таких сферах прикладних досліджень, як розпізнавання мови, біоінформатика, обробка аудіо- та відеосигналів. В останній час зростає інтерес до використання цих моделей для вирішення задачі прогнозування часових рядів.

У порівнянні з іншими моделями марківські моделі мають ряд переваг: їх легко будувати на експериментальних даних; вони не потребують розуміння внутрішніх механізмів динаміки змін системи; їх застосування не потребує великих витрат машинного часу.

У переважній більшості досліджень у сфері прогнозування за допомогою марківських моделей було розглянуто проблему прогнозування фінансових часових рядів. Це насамперед пов’язано з тим, що марківські моделі базуються на ймовірності подій, що дозволяє враховувати невизначеність та нелінійність, притаманну фондовим ринкам.

Розглянемо більш детально алгоритми, які використовуються для вирішення основних проблем ПММ.

Алгоритм прямого – зворотного ходу. Алгоритм прямого – зворотного ходу використовують для обчислення ймовірності $P(O\,|\,\lambda)$, тобто ймовірність побачити послідовності $O=\{o_1,o_2,...,o_T\}$ подій для побудованої моделі λ. Для того щоб обчислити значення $P(O\,|\,\lambda)$, можна використовувати пряму або зворотну процедуру. Нижче подано їхній більш детальний опис.

Розрахунок прямих ймовірностей. Для обчислення прямих ймовірностей вводять змінну $a_i(t)$, яка характеризує ймовірність того, що в момент часу t система знаходилася у стані S_i, і під час переходів системи сформувалася послідовність подій $\{o_1,o_2,...,o_{T-1}\}$, тобто $\alpha_i(t)=P(O_{1,t-1}\,|\,q_t=s_i)$ [2, 6].

Розрахунок ймовірності $P(O\,|\,\lambda)$ з використанням прямої процедури складається з трьох кроків:

1) Ініціалізація змінних α_i, тобто обчислення значень $\alpha_i(1)$.
2) Індукція, тобто обчислення значень $\alpha_i(t)$ для усіх можливих станів $1\le i\le N$.
3) Завершення роботи, тобто обчислення значення $P(O\,|\,\lambda)$, використовуючи значення $\alpha_i(t)$.

У момент часу $t=1$ $\alpha_i(1)$ беруться з початкового розподілу ймовірностей станів, тобто $\alpha_i(1)=\pi_i$. Для наступних моментів часу t $\alpha_i(t)$ можна отримати за рекурентою формулою:

$$\alpha_i(t) = P(O_{1,t} \mid X_t = s_i) = \sum_{j \in S} P(O_{1,t} \mid X_t = i \cap X_{t-1} = j) =$$

$$\sum_{j \in S} P(O_{1,t-1} \mid X_{t-1} = j) \cdot P(X_{t-1} = j \mid X_t = i) \cdot P(O_t = o_t \mid X_t = i) =$$

$$\sum_{j \in S} \alpha_j(t-1) \cdot a_{ji} \cdot b_{iot} = b_{iot} \sum_{j \in S} \alpha_j(t-1) \cdot a_{ji}$$

Наведемо схему обчислень прямої процедури.

Крок 1. Ініціалізація

$$\alpha_i(1) = \pi_i,\ 1 \le i \le N,$$
$$t = 1.$$

Крок 2. Індукція

$$While\ (\ t < T\ ),$$
$$\alpha_j(t+1) = b_{jot} \sum_{i=1}^{N} \alpha_i(t) \cdot a_{ij},\ 1 \le i \le N,$$
$$t = t + 1.$$

Крок 3. Завершення

$$P(O \mid \lambda) = \sum_{j=1}^{N} \alpha_j(T)$$

Розрахунок зворотних ймовірностей. Для обчислення зворотних ймовірностей вводять змінну $\beta_i(t)$, яка характеризує ймовірність того, що в момент часу t система знаходилася у стані S_i, і після цього ми будемо спостерігати послідовність подій $\{O_{t+1},...,O_T\}$, тобто $\beta_i(t) = P(O_{t+1,T} \mid X_t = s_i)$ [2,6].

Розрахунок ймовірності $P(O \mid \lambda)$ з використанням зворотної процедури складається з трьох кроків:

1) Ініціалізація змінних β_i, тобто обчислення значень $\beta_i(T)$.
2) Індукція, тобто обчислення значень $\beta_i(t)$ для усіх можливих станів $1 \le i \le N$.
3) Завершення роботи, тобто обчислення значення $P(O \mid \lambda)$, використовуючи значення $\beta_i(t)$.

У момент часу $t = T + 1$ $\beta_i(t) = 1$, як ймовірність отримати довільну послідовність подій. Для попередніх моментів часу $\beta_i(t)$ можна отримати за рекурентною формулою:

$$\beta_i(t) = P(O_{t+1,T} \mid X_t = s_i) =$$

$$\sum_{j \in S} P(O_{t+2,T} \mid X_{t+1} = j) \cdot P(X_{t+1} = j \mid X_t = i) \cdot P(o_{t+1} \mid X_t = j) =$$

$$\sum_{j \in S} \beta_j(t+1) \cdot a_{ij} \cdot b_{jot+1}$$

Розглянемо схему обчислень зворотної процедури.

Крок 1. Ініціалізація

$$\beta_i(T+1) = 1,\ 1 \le i \le N,$$

$$t = T$$

Крок 2. Індукція

$$While\ (\ t > 0\ ),$$

$$\beta_i(t) = \sum_{i=1}^{N} \beta_j(t+1) \cdot a_{ij} \cdot b_{jot+1},\ 1 \le i \le N,$$

$$t = t - 1.$$

Крок 3. Завершення

$$P(O \mid \lambda) = \sum_{j=1}^{N} \beta_j(1) \cdot \pi_j$$

Алгоритм Вітербі. Для вирішення другої проблеми ПММ – проблеми знаходження найбільш вірогідної послідовності прихованих станів – існують декілька підходів. Перший, найбільш очевидний, це сформувати послідовність зі станів, кожен з яких є найбільш вірогідним у момент часу t, $\forall 1 \le t \le N$. Проте цей підхід має суттєвий недолік: він не враховує значення ймовірностей переходів між станами, тому згенерована таким чином послідовність прихованих станів може бути зовсім не правдоподібною. Наприклад, якщо в певній точці $a_{ij} = 0$, то послідовність станів, отримана за допомогою цього методу, буде неправильною. Тому для знаходження оптимальної послідовності прихованих станів використовують алгоритм Вітербі [2, 7].

До розгляду вводять дві змінні: $\delta_i(t)$ та $\varphi_i(t)$. Змінна $\delta_i(t)$ зберігає ймовірність того, що на t-му кроці система знаходиться в i-му стані. Змінна $\varphi_i(t)$ зберігає номер найбільш ймовірного стану на $t-1$-му кроці. Алгоритм складається з трьох кроків:

1) Обчислюємо значення $\delta_i(t)$ та $\varphi_i(t)$ у момент часу $t=1$. $\delta_i(1)=\pi_i \cdot b_{io1}$, як ймовірність отримати першу подію з початкового розподілу ймовірностей. $\varphi_i(1)=0$, тут ми використовуємо 0 деяке фіктивне значення.
2) Обчислюємо значення $\delta_i(t)$ та $\varphi_i(t)$ для усіх значень *t*, використовуючи матрицю переходів *A* та матрицю емісій *B*.
3) Розглядаючи максимальні значення $\varphi_i(t)$ на кожному кроці, починаючи з останнього моменту часу, видаємо відповідь.

Наведемо схему обчислень більш детально.

Крок 1. Ініціалізація

$$\delta_i(1)=\pi_i \cdot b_{io1}, 1 \le i \le N,$$

$$\varphi_i(1)=0,\ 1 \le i \le N,$$

$$t=2.$$

Крок 2. Індукція

While (t < T),

$$\delta_j(t)=b_{ijot} \cdot \max_{1\le i\le N}(\delta_i(t-1)\cdot a_{ij}),\ 1 \le j \le N,$$

$$\varphi_j(t)=\arg\max_{1\le i\le N}(\delta_i(t-1)\cdot a_{ij}),\ 1 \le j \le N,$$

$$t=t+1.$$

Крок 3. Зчитування відповіді

$$s_T^*=\arg\max_{1\le i\le N}(\delta_i(T)),$$

$$t=T-1,$$

While (t > 2),

$$s_t^*=\varphi_{t+1}(s_{t+1}^*),$$

$$t=t-1.$$

Алгоритм Баума – Велша. Для навчання марківської моделі використовують алгоритм Баума – Велша. Цей алгоритм є окремим випадком алгоритму EM (expectation -maximization). На кожному кроці алгоритму значення функції $P(O \mid \lambda)$ поступово збільшується, і алгоритм сходиться до локального екстремуму. В результаті роботи алгоритму отримуємо вектор параметрів $\lambda = \{A, B, \pi\}$, які найкраще описують вихідну послідовність спостережуваних подій [7, 8].

Алгоритм Баума – Велша складається з таких кроків:

1) Обчислення значень $\alpha_i(t)$ та $\beta_i(t)$, $1 \le i \le N$, $1 \le t \le T$.
2) Обчислення значень тимчасових змінних $\gamma_i(t)$ та $\varepsilon_{ij}(t)$.
3) Використовуючи значення $\gamma_i(t)$ та $\varepsilon_{ij}(t)$, обчислюємо нові значення π_i, a_{ij}, $b_i(o_k)$.

Кроки 1–3 повторюють до сходження алгоритму до єдиного розв'язку. Розглянемо схему обчислень алгоритму Баума – Велша більш детально.

Крок 1. Використовуючи поточні значення *A*, *B* та π, розраховуємо значення $\alpha_i(t)$ та $\beta_i(t)$:

$$\alpha_i(1) = \pi_i,\ 1 \le i \le N,$$

$$\alpha_i(\mathrm{t}+1) = b_{jot} \cdot \sum_{i=1}^{N} \alpha_i(\mathrm{t}) \cdot a_{ij},\ 1 \le i \le N,\ 2 \le t \le T,$$

$$\beta_i(\mathrm{T}+1) = 1,\ 1 \le i \le N,$$

$$\beta_i(\mathrm{t}) = \sum_{i=1}^{N} \beta_j(\mathrm{t}+1) \cdot a_{ij} \cdot b_{jot+1},\ 1 \le i \le N,\ T \ge t \ge 1.$$

Крок 2. Використовуючи отримані значення $\alpha_i(t)$ та $\beta_i(t)$, знаходимо значення $\gamma_i(t)$ та $\varepsilon_{ij}(t)$:

$$\gamma_i(t) = P(q_t = s_i \mid O, \lambda) = \frac{\alpha_i(t) \cdot \beta_i(t)}{\sum_{j=1}^{N} \alpha_j(t) \cdot \beta_j(t)},\ 1 \le i \le N,$$

$$\varepsilon_{ij}(t) = P(q_t = s_i, q_{t+1} = s_j \mid O, \lambda) =$$

$$\frac{\alpha_i(t) \cdot a_{ij} \cdot \beta_j(t+1) \cdot b_j(o_{t+1})}{\sum_{s=1}^{N} \sum_{l=1}^{N} \alpha_s(t) \cdot a_{sl} \cdot \beta_l(t+1) \cdot b_l(o_{t+1})},\ 1 \le i, j \le N.$$

Крок 3. За допомогою поточних значень $\gamma_i(t)$ та $\varepsilon_{ij}(t)$ обчислюємо нові значення *A*, *B* та π.

$$\pi_i = \gamma_i(1),\ 1 \le i \le N,$$

$$a_{ij} = \frac{\sum_{t=1}^{T} \varepsilon_{ij}(t)}{\sum_{t=1}^{T} \gamma_i(t)}, \ 1 \le i, j \le N,$$

$$b_i(o_k) = \frac{\sum_{t=1}^{T} b_i(o_k) \cdot \gamma_i(t)}{\sum_{t=1}^{T} \gamma_i(t)}, \ 1 \le i \le N, \ 1 \le k \le K.$$

Обчислення проводять поки значення *A*, *B* та π не перестануть істотно змінюватися.

Дискретні та неперервні ймовірності емісій. Значення b_{iwl}, тобто ймовірність спостереження події w_l після переходу до стану s_i, можуть бути дискретними лише у випадку, коли множина можливих подій є дискретною, наприклад, коли ми бажаємо спрогнозувати тенденцію часового ряду в майбутньому: зростання значень, спадання, ніяких істотних змін. Якщо ж множина можливих значень є неперервною, наприклад, коли ми бажаємо спрогнозувати не тенденцію, а конкретні значення показника у майбутньому, доцільно використовувати неперервні функції щільностей розподілів (*pdf*) для представлення ймовірності спостереження подій. В цьому випадку для представлення b_{iwl} використовують суміш нормальних розподілів [3, 9]:

$$b_{iwl} = \sum_{l=1}^{K} w_{il} \cdot N(O, \mu_{il}, \Sigma_{il}),$$

$$w_{il} \ge 0, \ 1 \le i \le N, \ 1 \le l \le K.$$

$$\sum_{l=1}^{K} w_{il} = 1, \ 1 \le i \le N.$$

де $N(O, \mu_{il}, \Sigma_{il})$ – багатовимірний нормальний розподіл з параметрами μ – вектор середніх значень та Σ – матриця коваріацій. w_{il} називається вагою l-го гаусіану для i-го стану. В цьому випадку алгоритм навчання ПММ змінюється: окрім обчислення значень π_i та a_{ij}, на кожному кроці переховують значення μ, Σ та w_{il}.

Алгоритм прогнозування часових рядів з використанням ПММ.

1. Сформувати навчальну послідовність.
2. Задати початкові значення для матриць A, B та вектора π випадковим чином або з урахуванням певної апостеріорної інформації.
3. Виконати тренування ПММ за допомогою алгоритму Баума – Велша.
4. Знайти найбільш вірогідний стан системи з використанням алгоритму Вітербі.
5. У якості прогнозного значення видати найбільш ймовірну подію для отриманого на попередньому кроці стану.

Огляд останніх наукових досліджень в області прогнозування часових рядів за допомогою ПММ. Сьогодні застосування прихованих марківських моделей до прогнозування часових рядів вважається перспективним напрямом наукових досліджень. Протягом останніх років було опубліковано цілу низку наукових праць з даного питання. Базовий алгоритм застосування марківських моделей до прогнозування часових рядів був представлений деякий час назад. Сьогодні багато дослідників з різних країн працюють над покращенням класичного методу.

Китайський дослідник Зенг у роботі [3] запропонував використовувати експоненційно зважений алгоритм ЕМ для навчання марківської моделі. Ідея цього алгоритму заснована на введенні нового параметра η, який залежить від часу та враховує старіння інформації. Вчений пропонує алгоритм обчислення нового параметра та доводить збіжність зваженого алгоритму. Окрім цього, дослідник показує, що з використанням модифікованого алгоритму з'являється нова проблема. Деякі часові ряди змінюють знак свого напрямку дуже швидко протягом короткого періоду часу. Новий алгоритм намагається підлаштуватися під значення ряду якнайкраще, і це призводить до появи проблеми дуже високої чутливості алгоритму, приклад якої показано на рис. 2 (жирною лінією позначено вихідний ряд, пунктиром – прогнозні значення).

Для розв'язку цієї проблеми автор пропонує використовувати подвійно зважений алгоритм ЕМ, ідея якого полягає у введенні другого параметра ς, який використовується як рівень впевненості для параметра η.

Рисунок 2 – Побудова некоректного прогнозу з використанням зваженого алгоритму ЕМ

Також у роботі [3] автори пропонують використовувати динамічний пул для навчання марківської моделі. В кінці кожного торгового дня у навчальний пул додається нове значення, наприклад, значення ціни деякого цінного паперу на момент закриття біржі. Це забезпечує можливість навчання моделі на найбільш сучасних даних. Слід зауважити, що останнє значення, яке додається в пул, використовується дослідником як для навчання моделі, так і в якості цільового значення, яке необхідно спрогнозувати.

Запропонований автором метод навчання ПММ був застосований для прогнозування як синтетичних даних, так і реальних значень рівня індексу NASDAQ в період за 1998–2000 роки. Результати його роботи були порівняні з результатами, отриманими при прогнозуванні за допомогою нейронних мереж та звичайного алгоритму Баума – Велша. В обох випадках запропонований автором експоненційно зважений алгоритм ЕМ показав у цілому більшу потужність, ніж нейронні мережі та традиційний метод навчання ПММ. Слід зауважити, що на досліджуваних автором даних не було помічено високої різниці між зваженим та подвійно зваженим ЕМ алгоритмом. У випадку застосування подвійно зваженого алгоритму результат був лише на 2 % кращим за результат, отриманий при використанні алгоритму, зваженого один раз.

Одними з перших дослідження в області прогнозування часових рядів за допомогою прихованих моделей Маркова почали австралійські науковці Рафуіл Хассан та Баікунс Нес [4]. Вони досліджували використання ПММ для прогнозування цін на акції міжнародних авіакомпаній. Для цього вчені використовували приховану марківську модель з 4 станами та неперервною матрицею емісій: для кожного стану значення $b_i(\mathrm{w}_k)$ представляють собою суміш трьох нормальних розподілів. У якості вхідних даних були використані 4 показники: ціна на момент відкриття біржі, найвища ціна за день, найнижча ціна за день, ціна на момент закриття біржі. Метою було побудувати прогноз щодо ціни на момент закриття біржі на наступний день. Спроектована модель була застосована для прогнозування цін на акції авіакомпанії British Airlines, Delta Airlines, Southwest Airlines, Ryanair Holdings Ltd. В якості навчальної вибірки використовувалися ціни в період з 18 грудня 2002 року по 29 вересня 2004 року. Метою було побудувати прогнозне значення ціни на момент закриття біржі на 30 вересня 2004 року. В результаті проведення експериментів було встановлено, що розроблений метод майже не поступається штучним нейронним мережам, а в деяких випадках будує більш точні прогнози.

Італійські дослідники з університету Сасарі Мануель Біцего, Енріко Гроссо та Едуардо Ортландо [5] запропонували прогнозувати не конкретне значення показника, а знак майбутнього тренду: зростання, спадання показника або ніяких істотних змін. Для цього вчені використовували не одну неперервну ПММ, а дві дискретні ПММ. Одна за моделей, λ_+, була натренована на прогнозування позитивних тенденцій, тобто значення показника зростає, а друга, λ_- – на ідентифікацію негативних тенденцій, тобто спад значень ряду. Для цього вхідні дані були поділені на три послідовності: послідовність F_+, яка закінчується на τ позитивних сигналів, послідовність F_-, яка закінчується на τ негативних сигналів, та послідовність $F_?$ яка представляє собою нерегулярні коливання останніх τ знаків.

Обидві моделі λ_+ та λ_- були натреновані з використанням послідовностей F_+ та F_-. Кількість станів ПММ була оцінена за допомогою інформаційного критерію Акаіке (AIC). Автори також запропонували алгоритм введення порогового правила, яке здатне характеризувати зони «невизначеності».

Запропонований підхід був протестований за допомогою індексу Доу Джонса (DJ) та інших цінних паперів з різними ризиками: GE (General Electric), IBM (International Machines Corp.) (акції з низьким рівнем ризику), BGP (Borders Group) та CGI (Commerc Group) (акції високого ризику). Для роботи з цими моделями використовували щоденні ціни в період з 30 листопада 1995 року до 5 лютого 2001. Значення цін у період з грудня 1995 року по червень 1998 року використовувалися для тренування моделей, решта значень серії була використана для тестування.

Результати роботи двох дискретних ПММ були порівняні з результатами, отриманими при використанні однієї неперервної ПММ та методу локального поліноміального тренду, натренерованих на однакових вхідних даних. За допомогою розробленого авторами методу вдалося отримати помітно більшу точність побудованих прогнозів.

Висновки. Результати експериментів [3, 4, 5] показують, що приховані моделі Маркова мають високий потенціал у сфері аналізу та прогнозування часових рядів. Перевагою цих моделей є те, що результати їх роботи можуть бути зрозумілі та пояснені з точки зору людини, на відміну від інших методів, а також те, що вони мають сильне математичне обґрунтування. Незважаючи на це, сьогодні існує широке коло відкритих питань в цій галузі досліджень. По-перше, існує проблема вибору оптимального розміру тренувального набору. З однієї сторони, достатньо велика навчальна послідовність містить більше інформації про історію коливань; отже, є більш корисною. З іншої – для обробки великої вибірки потрібно більше обчислень. Друга проблема – це проблема вибору функцій розподілу для обчислення значень емісійної матриці *B* в неперервному випадку. У багатьох роботах дослідники роблять припущення, що має місце нормальний розподіл або суміш декількох нормальних розподілів. Насправді ж таке твердження є сумнівним, тому що існує велика кількість часових рядів, коливання яких не дотримуються нормального закону. По-третє, більшість розроблених методів передбачають прогнозування конкретної події. Окрім цього, корисно було б розрахувати ймовірність кожної події. По-четверте, за допомогою марківських моделей можна було б відстежити, як коливання на одному, більш потужному ринку цінних паперів, впливає на інші ринки, тобто як один випадковий процес впливає на інший.

Бібліографічні посилання

1. Grinstead Charles M., J. Laurie Snell. Chapter 11. Markov Chains. // Introduction to Probability (Second Revised Edition) Published by Orient BlackSwan / Universities Press (2009). URL: http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter11.pdf

2. Jurafsky Daniel, James H. Marti. Hidden Markov Models. Speech and Language Processing. Stanford University and University of Colorado at Boulder. URL: https://web.stanford.edu/~jurafsky/slp3/9.pdf

3. Zhang Y. Prediction of financial time series with Hidden Markov Models. Master's thesis, The School of Computing Science, Simon Frazer

University, Canada, 2004. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.136.5871&rep=rep1&type=pdf

4. Md. Rafiul Hassan, Baikunth Nath. Stock Market Forecasting Using Hidden Markov Model: A New Approach // Proceedings of the 2005 5th International Conference on Intelligent Systems Design and Applications (ISDA'05). Wroclaw, Poland. 2005. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.491.9959&rep=rep1&type=pdf

5. Bicego Manuele, Grosso Enrico, Otranto Edoardo. A Hidden Markov Model Approach to Classify and Predict the Sign of Financial Local Trends // Proceedings of the 2008 Joint IAPR International Workshop on Structural, Syntactic, and Statistical Pattern Recognition. Orlando, Florida. December 04–06, 2008. URL: https://pdfs.semanticscholar.org/6cc0/87f66de03c8a3ab8c29475c65522ef122148.pdf

6. Nishanth Ulhas Nair, Sreenivas T.V. Forward/Backward algorithms for joint multi pattern speech recognition. // Signal Processing Conference, 2008 16th European 25–29 Aug. 2008. URL: http://ieeexplore.ieee.org/abstract/document/7080380/

7. Churbanov Alexander, Winters-Hilt Stephen. Implementing EM and Viterbi algorithms for Hidden Markov Model in linear memory // BMC Bioinformatics, Published: 30 April 2008. URL:https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-224

8. Miklós István, Meyer Irmtraud M. A linear memory algorithm for Baum-Welch training // BMC Bioinformatics. 2005. URL : https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1262623/pdf/1471-2105-6-231.pdf

9. Pfundstein Georg. Hidden Markov Models with Generalised Emission Distribution for the Analysis of High-Dimensional, Non-Euclidean Data. Master Thesis, University of Munich, 19th December 2011. URL: https://epub.ub.uni-muenchen.de/12710/1/MA_Pfundstein.pdf

10. Rank Erhard, Pernkopf Franz. Hidden Markov Models. Lecture Notes Speech Communication 2, SS 2004, Signal Processing and Speech Communication Laboratory, Graz University of Technology. URL: https://www.spsc.tugraz.at/system/files/specomm04_3.pdf

Надійшла до редколегії 06.07.2017