

МАЛИК І.В., ІВАСЮК Р.В.

ПММ ТА НПМ У ЧАСОВИХ РЯДАХ

Основна увага в роботі сфокусована на розгляд, так званих, прихованих ланцюгів Маркова та їх аналогів і узагальнень. Зокрема розглянуто вплив прихованих ланцюгів Маркова та напівмарковських прихованих моделей на моделі часових рядів, які описують вартості акцій топових компаній станом на 2024 рік. Під час дослідження вдалося з'ясувати, що розгляд узагальненіших моделей дозволяє більш точно описувати динаміку вартості акцій, а отже, більш адекватно визначати основні характеристики реального процесу.

Ключові слова і фрази: приховані ланцюги Маркова, напівмарковські приховані моделі.

Чернівецький національний університет імені Юрія Федьковича, Чернівці, Україна
e-mail: *i.malyk@chnu.edu.ua* (Малик І.В.), *ivasiuk.roman@chnu.edu.ua* (Івасюк Р.В.)

Вступ

Теорія прихованих ланцюгів Маркова інтенсивно розвивається із середини минулого століття і розвиток даної теорії насамперед пов'язано із наявністю великої кількості реальних процесів, динаміка яких залежить від інших (прихованих) випадкових процесів. Дана концепція знайшла своє впровадження в багатьох прикладних задачах, включаючи прогнозування фінансових часових рядів [1, 2, 3, 4], прогнозування погоди [5], прогнозування кліматичних змін на основі часових рядів [6, 7, 8, 9] тощо. Таким чином, випадкові процеси із врахуванням прихованих складових являються ще одним покращенням опису реальних процесів та врахування невідомих факторів в різних прикладних задачах.

Слід відмітити, що класичні моделі процесів, що містять приховані складові в основному ґрунтуються на припущенні про марковість прихованого процесу, тобто відсутність післядії. Дане припущення з одного боку дозволяє значно спростити оцінку параметрів моделі, з іншого боку ж звужує коло реальних процесів, для яких може бути використана теорія процесів із прихованими складовими. Враховуючи це, в роботі розглянуто напівмарковські приховані моделі, в яких час перебування в стані описується

УДК 517.956

2010 *Mathematics Subject Classification:* 35k35, 35k20.

Information on some grant ...

не показниковим розподілом (як в класичній моделі), а довільним розподілом, що визначений на R_+ , тобто

$$P(\theta_n \geq 0) = 1,$$

де $\theta_n(u)$ - час перебування в стані u .

1 ПРИХОВАНІ МАРКОВСЬКІ МОДЕЛІ У ЧАСОВИХ РЯДАХ

Приховані марковські моделі (ПММ) використовуються для моделювання часових рядів, де ймовірнісні розподіли спостережень визначаються невидимими станами скінченного стану Марковського ланцюга. У цих моделях робиться припущення, що спостереження X_t у момент часу t залежать від деякого неспостережуваного стану S_t , який є елементом скінченного набору станів $\{1, 2, \dots, N\}$. Послідовність станів $\{S_t\}$ утворює ланцюг Маркова, що означає, що наступний стан залежить лише від поточного стану, а не від попередніх станів [5, 1, 2, 3, 4].

Ймовірність переходу між станами описується матрицею перехідних ймовірностей P , яка має розмір $N \times N$ і елементи p_{ij} , що представляють ймовірність переходу з стану i до стану j

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1N} \\ p_{21} & p_{22} & \dots & p_{2N} \\ \dots & \dots & \dots & \dots \\ p_{N1} & p_{N2} & \dots & p_{NN} \end{pmatrix}$$

а самі ймовірності визначаються співвідношенням

$$p_{ij} = P(S_{t+1} = j | S_t = i).$$

Для кожного стану $S_t = j$ існує відповідний розподіл спостережень прихованого ланцюга X_t . Наприклад, якщо спостереження є нормально розподіленими, то $X_t \sim N(\mu_j, \sigma_j^2)$, де μ_j і σ_j^2 є середнім і дисперсією спостережень у стані j .

Загальна формула для ймовірності спостереження X_t за умови стану S_t може бути записана як

$$P(X_t = x_t | S_t = j) = f_j(x_t),$$

де $f_j(x_t)$ – функція щільності ймовірності для стану j .

Параметри оцінюються за допомогою методу максимальної правдоподібності [5], часто реалізованого за допомогою алгоритму Баум-Велча (Baum-Welch algorithm), який є особливим випадком ЕМ-алгоритму. Ці моделі виявилися корисними для моделювання різких змін у фінансових часових рядах, але мали обмеження у відображенні тривалості перебування в станах через показниковий розподіл.

2 НАПІВМАРКОВСЬКІ ПРИХОВАНІ МОДЕЛІ У ЧАСОВИХ РЯДАХ

Напівмарковські приховані моделі (НПМ) узагальнюють ПММ, дозволяючи більш гнучкі розподіли тривалості перебування в стані (часу перебування в даному стані). Ця гнучкість вирішує ключове обмеження ПММ, яке полягає в показниковому розподілі тривалості перебування.

Як і в ПММ, НПМ [5, 1] має скінченну кількість прихованих станів $S = \{s_1, s_2, \dots, s_M\}$. Кожен стан відповідає певному режиму в часовому ряді. Перехід між станами описується матрицею перехідних ймовірностей $P = (p_{ij})$, $\{i, j\} \subset \{1, 2, \dots, N\}$ але, на відміну від ПММ, НПМ дозволяє задавати тривалість перебування в кожному стані за допомогою розподілу, відмінного від показникового. У НПМ перехідна ймовірність визначається як ймовірність переходу в стан j , якщо відбувся вихід з поточного стану i . Ключова відмінність між НПМ і ПММ полягає у розподілі часу перебування в кожному стані. У ПММ тривалість перебування в стані геометрично розподілена, тоді як у НПМ може бути довільний розподіл тривалості перебування в стані j , який визначається функцією

$$d_j(u) = P(\text{перебування в стані } j \text{ триває } u \text{ періодів}),$$

де $d_j(u)$ – це ймовірність того, що процес залишатиметься в стані j протягом u часових періодів. Кумулятивна функція розподілу або функція виживання визначається як

$$D_j(u) = \sum_{v \geq u} d_j(v).$$

Для кожного стану S_i визначено розподіл спостережень. Наприклад, якщо спостереження мають нормальний розподіл в кожному стані, то

$$X_t \sim N(\mu_j, \sigma_j^2), S_t = j,$$

де X_t – спостереження в момент часу t , а μ_j і σ_j^2 – середнє та дисперсія в стані $S_t = j$.

Подібно до ПММ, ймовірність спостережень залежить від послідовності станів. Ймовірність послідовності спостережень $O = \{O_1, O_2, \dots, O_T\}$ задається як

$$P(O|\lambda) = \sum_S P(O|S, \lambda)P(S|\lambda),$$

де $\lambda = (P, d_j(u), B)$ – набір параметрів НПМ, включаючи матрицю перехідних ймовірностей P , розподіл тривалості перебування $d_j(u)$, та розподіл спостережень B .

НПМ краще моделюють властивість фінансових часових рядів - повільне спадання автокореляції у квадратах приростів - оскільки їх гнучкість дозволяє враховувати довгострокові залежності. ПММ, у свою чергу, обмежені своїм фіксованим набором станів та короткостроковими залежностями, тому не можуть точно змоделювати цю характеристику.

Квадрати приростів для часових рядів - це квадрат значення приросту (r_t) у кожний момент часу. Цей показник використовується для оцінки волатильності (мінливості) фінансових часових рядів, що є важливим аспектом аналізу ризику.

У статті розглянуто дві конкретні НПМ та зроблено порівняння їх відповідності до щоденних рядів прибутків з 18 європейських секторних індексів [10], демонструючи, що НПМ забезпечують значно кращу відповідність порівняно з ПММ. Також є аналіз прилягання моделей до часових рядів даних, що включає перевірку автокореляції у квадратах приростів.

Ці моделі показали кращу відповідність даним, зокрема у випадках повільного спадання автокореляції в квадратах прибутків. Вони забезпечують більш гнучке моделювання тривалості перебування в станах, що робить їх більш придатними для фінансових часових рядів.

3 ЧАСТИННІ ВИПАДКИ НПМ МОДЕЛЕЙ

НПМ з негативно-біноміальним розподілом часу перебування складається з двох процесів: прихованого марківського процесу (стани) і видимого процесу (спостереження). Видимий процес залежить від невидимого стану, який змінюється відповідно до марківського процесу. Негативно-біноміальний розподіл використовується для опису часу перебування в кожному стані.

У НПМ з основним нормальним розподілом (SMN розподіл) спостереження в кожному стані мають нормальний розподіл. Якщо модель знаходиться в стані $S_t = j$, то спостереження X_t буде нормально розподіленим

$$X_t | S_t = j \sim N(\mu_j, \sigma_j^2),$$

де μ_j – середнє значення, а σ_j^2 – дисперсія в стані j .

У НПМ з основним розподілом Стьюдента (SMt розподіл) спостереження мають t -розподіл. Якщо модель знаходиться в стані $S_t = j$, то спостереження X_t буде розподіленим за t -розподілом з параметрами

$$X_t | S_t = j \sim t(v_j, \mu_j, \sigma_j^2),$$

де v_j – параметр ступенів свободи t -розподілу, μ_j – середнє значення, а σ_j^2 – масштабний параметр (дисперсія) в стані j .

t -розподіл забезпечує важчі "хвости" у розподілі, що робить його більш придатним для моделювання фінансових даних із високою волатильністю і частими екстремальними значеннями. Правдоподібність для моделі SMN чи SMt розраховується шляхом підсумовування повної правдоподібності даних за всіма можливими шляхами, які може приймати процес переходу станів.

На відміну від попередніх моделей, SMN та SMt мають більшу гнучкість у моделюванні часу перебування в станах, що дозволяє їм точніше відображати залежності в даних, особливо для фінансових часових рядів, які характеризуються тривалими періодами як низької, так і високої волатильності. Завдяки цьому ці моделі здатні краще моделювати автокореляційні функції.

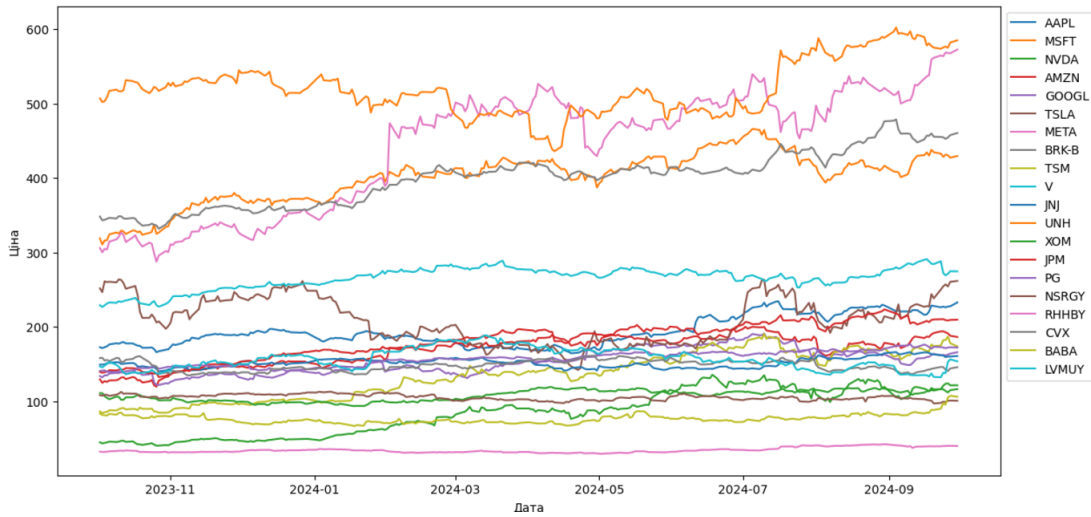


Рис. 1. Зміни вартості акцій різних компаній протягом року

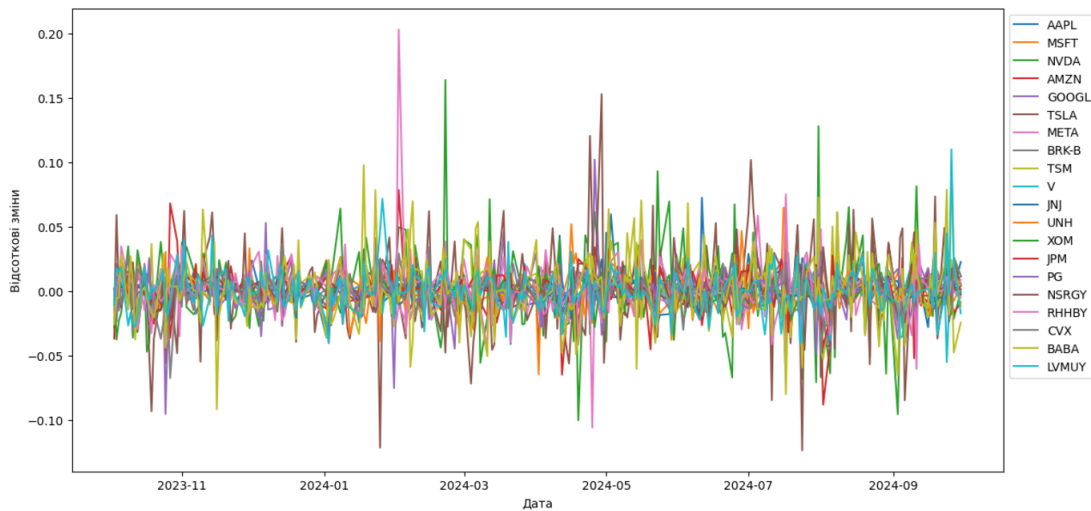


Рис. 2. Відсоткові зміни вартості акцій під час того самого періоду

4 ТЕСТИ ДЛЯ ОЦІНКИ СТАЦІОНАРНОСТІ

Розглянемо основні статистичні тести для перевірки стаціонарності часового ряду, які ґрунтуються на припущенні скінченної післядії. Розширений тест Діккі - Фуллера (ADF-тест) - це статистичний тест, який використовується для перевірки стаціонарності часового ряду. Стаціонарність є ключовою властивістю, яка означає, що статистичні властивості часового ряду, такі як середнє значення та дисперсія, залишаються постійними в часі. Нульова гіпотеза (H_0) стверджує, що часовий ряд має одиничний корінь і є нестаціонарним, тоді як альтернативна гіпотеза (H_1) припускає, що часовий ряд є стаціонарним. ADF-тест є розширенням класичного тесту Діккі-Фуллера. Він включає додаткові лаги залежної змінної, щоб враховувати можливу автокореляцію залишків. Це робить тест більш надійним для практичного використання.

ADF-тест часто використовується в економіці та фінансах для аналізу стаціонарності фінансових часових рядів, таких як ціни акцій, валютні курси та інші економічні

індикатори. Якщо ряд не є стаціонарним, для подальшого аналізу його часто перетворюють, наприклад, шляхом взяття різниць або логарифмування.

KPSS-тест - це статистичний тест, що використовується для перевірки стаціонарності часового ряду, але на відміну від ADF-тесту, його нульова гіпотеза припускає, що ряд є стаціонарним. Цей тест допомагає виявити, чи має часовий ряд одиничний корінь, тобто чи є він нестаціонарним. KPSS-тест оцінює дисперсію помилок у регресії, щоб визначити, чи є часовий ряд стаціонарним. Він перевіряє, чи є дисперсія постійною (стабільною) в часі. Але вимагає вибір кількості лагів для регресії залишків, що впливає на результати тесту. Зазвичай використовується автоматичний вибір оптимального лагу. KPSS-тест часто використовується разом з ADF-тестом, оскільки вони мають протилежні нульові гіпотези. Це дозволяє провести більш всебічний аналіз стаціонарності часового ряду. Якщо ADF-тест відхиляє нульову гіпотезу (нестаціонарність), а KPSS не відхиляє (стаціонарність), це означає, що часовий ряд, ймовірно, є стаціонарним. KPSS-тест доповнює ADF-тест, оскільки вони перевіряють різні гіпотези також він є надійним для рядів з детермінованими трендами. Однак, вибір кількості лагів може впливати на результати, а також існує чутливість до вибору тренду в даних. KPSS-тест є потужним інструментом для перевірки стаціонарності часових рядів. Його використання разом з іншими тестами, такими як ADF, допомагає отримати більш точні та надійні результати щодо природи часового ряду. Як ми можемо бачити з наступної таблиці, майже всі розглянуті відсоткові ставки являються стаціонарними, що забезпечує більш точний аналіз моделей.

Акція	ADF $p - value$	KPSS тест $p - value$	Стаціонарність Result
AAPL	5.74e-27	0.15	Yes
MSFT	6.17e-30	0.25	Yes
NVDA	6.10e-30	0.10	Yes
AMZN	1.44e-24	0.14	Yes
GOOGL	7.45e-29	0.08	Yes
TSLA	1.12e-28	0.25	Yes
META	2.71e-24	0.08	Yes
BRK-B	2.57e-24	0.04	No
TSM	1.05e-29	0.07	Yes
V	1.75e-22	0.25	Yes
JNJ	7.44e-27	0.10	Yes
UNH	1.46e-26	0.14	Yes
XOM	8.13e-15	0.16	Yes
JPM	1.89e-28	0.12	Yes
PG	2.13e-11	0.04	No
NSRGY	6.30e-05	0.02	No
RHHBY	6.18e-17	0.16	Yes
CVX	1.42e-26	0.08	Yes
BABA	4.03e-29	0.43	Yes
LVMUY	3.18e-26	0.10	Yes

5 ІНФОРМАЦІЙНІ КРИТЕРІЇ

Поряд із статистичними тестами, розглянемо дві метрики, на основі яких будемо визначати точність моделі – критерій Акаїке та критерій Шварца.

Критерій Акаїке (AIC) використовується для порівняння різних статистичних моделей, що застосовуються для опису одного і того ж набору даних. Його головна мета - знайти модель, яка найкраще описує дані з найменшими втратами інформації. AIC враховує як якість підгонки моделі, так і її складність. Формула для розрахунку AIC виглядає наступним чином:

$$AIC = -2\ln(L) + 2k$$

де L – функція правдоподібності моделі, а k – кількість параметрів у моделі. Чим менше значення AIC, тим краща модель. Важливо відзначити, що AIC не дає абсолютної міри якості моделі, а лише дозволяє порівнювати моделі між собою.

Критерій Шварца або Баєсівський інформаційний критерій (BIC) є подібним до AIC, але враховує розмір вибірки і більш жорстко штрафує моделі з великою кількістю параметрів, що допомагає уникати перенавчання. Формула для розрахунку BIC:

$$BIC = -2\ln(L) + 2k\ln(n)$$

де L – максимальна ймовірність моделі, k – кількість параметрів у моделі, і n – кількість спостережень у вибірці. Як і у випадку з AIC, модель з меншим значенням BIC вважається кращою. Оскільки в BIC враховується розмір вибірки, він частіше обирає менш складні моделі, особливо при великому значенні n .

Також AIC використовує коефіцієнт $2k$, тоді як BIC використовує $k\ln(n)$. Це означає, що BIC більше штрафує моделі за складність. AIC зазвичай краще для ситуацій, коли важливо максимально точно передбачити дані, а BIC - коли потрібно визначити найбільш вірогідну структуру даних.

Обидва критерії корисні при виборі моделі, але конкретний вибір залежить від цілей аналізу - AIC краще для передбачень, а BIC - для структурної інтерпретації.

Ticker	AIC (ПММ)	BIC (ПММ)	AIC (НПМ)	BIC (НПМ)
AAPL	-1864.5	-1819.17	-1870.43	-1825.10
MSFT	-1781.06	-1735.72	-1786.52	-1741.19
NVDA	-1400.8	-1355.47	-1398.98	-1353.64
AMZN	-1621.33	-1576.0	-1620.79	-1575.46
GOOGL	-1667.12	-1621.79	-1670.87	-1625.54
TSLA	-1283.05	-1237.71	-1285.33	-1240.0
META	-1607.59	-1562.26	-1604.56	-1559.23
BRK-B	-2091.91	-2046.58	-2124.63	-2079.29
TSM	-1610.71	-1565.37	-1618.68	-1573.35
V	-2032.99	-1987.66	-2054.40	-2009.07

Ticker	AIC (ПММ)	BIC (ПММ)	AIC (НПМ)	BIC (НПМ)
JNJ	-2033.42	-1988.09	-2035.43	-1990.1
UNH	-1858.96	-1813.63	-1875.58	-1830.25
XOM	-1775.22	-1729.89	-1786.14	-1740.8
JPM	-1922.17	-1876.84	-1914.54	-1869.21
PG	-2048.42	-2003.09	-2074.86	-2029.53
NSRGY	-1975.75	-1930.42	-1997.34	-1952.0
RHHBY	-1901.05	-1855.71	-1909.57	-1864.24
CVX	-1828.25	-1782.91	-1839.76	-1794.43
BABA	-1496.71	-1451.38	-1500.2	-1454.86
LVMUY	-1690.33	-1644.1	-1695.03	-1649.7

6 ВИСНОВОК

У роботі розглянуто математичні моделі марковських та напівмарковських прихованих моделей з використанням до реальних даних. Як було виявлено під час дослідження, напівмарковські приховані моделі є більш точними для реальних процесів, про що свідчать розглянуті статистичні тести та відповідні метрики. У майбутніх роботах в даному напрямку планується розглянути різні розподіли прихованих моделей, що узагальнюють нормальний розподіл та розподіл Стюдента, також розробити алгоритм оцінки параметрів даних розподілів.

СПИСОК ЛІТЕРАТУРИ

- [1] Sansom J., Thomson P. *Fitting hidden semi-Markov models to breakpoint rainfall data*. J. Appl. Probab. 2001, **38A** (Supplement), 142–157.
- [2] Hamilton J.D. *Analysis of time series subject to changes in regime*. J. Econometrics 1990, **45** (1-2), 39–70.
- [3] Linne T. *A Markov switching model of stock returns: an application to the emerging markets in central and eastern Europe*. In: East European Transition and EU Enlargement. Physica-Verlag, Heidelberg, 2002, pp. 371–384.
- [4] Turner C.M., Startz R., Nelson C.R. *A Markov model of heteroskedasticity, risk, and learning in the stock market*. J. Financial Econom. 1989, **25** (1), 3–22.
- [5] Baum L.E., Petrie T., Soules G., Weiss N. *A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains*. Ann. Math. Statist. 1970, **41** (1), 164–171.
- [6] Prudhomme C., Wilby R.L., Crooks S., Kay A.L., Reynard N.S. *Scenario-neutral approach to climate change impact studies: Application to flood risk*. J. Hydrology 2010, **390** (1-2), 198–209.
- [7] Yu C., Yu L. *Hidden Markov Models for Environmental Prediction*. J. Hydrological Engineering 2010, **15** (3), 140–151.
- [8] Carpenter S.R., Cole J.J., Kitchell J.F. *A Bayesian Hidden Semi-Markov Model with Covariate-Dependent State Duration Parameters for High-Frequency Environmental Data*. Environ. Model. Software 2020, **125**, 104–113.
- [9] Gomez-Navarro L., Salgado D., Cañedo R.S. *Markov Chain Ensemble for Climate Model Prediction*. Geophys. Meteorol. 2020, **157**, 265–271.

[10] Yahoo Finance. Retrieved October 1, 2024, <https://finance.yahoo.com>

Надійшло 28.10.2024

Malik I.V., Ivasiuk R.V. *HMM and HSMM in time series*, Bukovinian Math. Journal. **12**, 2 (2024), 119–127.

The main focus of the work is on the study of so-called hidden Markov chains (hidden Markov models, HMM) and their analogs and generalizations. In particular, the research examines the impact of HMM and semi-Markov hidden models (HSMM) on time series models describing the stock prices of top companies as of 2024. The study revealed that considering more generalized models allows for a more accurate description of stock price dynamics and, consequently, a more accurate determination of the key characteristics of the actual process.

The research employs both HMM and HSMM frameworks to analyze financial data, demonstrating their capacity to capture key features of stock price volatility, including sharp transitions between periods of high and low market variability. A series of tests and metrics were conducted to evaluate the performance of these models, including the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), which indicate superior fit for HSMMs. Additionally, methods such as the Augmented Dickey-Fuller (ADF) test and KPSS tests were used to validate the stationarity properties of the time series.

The study's results emphasize that semi-Markov extensions provide a significant improvement over classical HMMs when analyzing financial market data, allowing for better detection of long-term dependencies and accurate modeling of asset price trends. The findings open avenues for further applications in financial risk analysis and forecasting tasks, showcasing the potential of HSMMs to deliver more robust insights into market behavior.