

# ЗАСТОСУВАННЯ МЕТОДУ МАКСИМІЗАЦІЇ ПОЛІНОМІВ ДЛЯ ОЦІНЮВАННЯ ПАРАМЕТРІВ ARIMA МОДЕЛЕЙ З АСИМЕТРИЧНИМИ НЕГАУСОВИМИ ІННОВАЦІЯМИ

Сергій Заболотній<sup>a</sup>

<sup>a</sup> Cherkasy State Business College, Cherkasy, Ukraine

E-mail: zabolotnii.serhii@csbc.edu.ua

## Abstract

**Контекст та актуальність.** Авторегресійні інтегровані моделі ковзного середнього (ARIMA) є одним із найпоширеніших інструментів аналізу часових рядів в економіці, фінансах та інших прикладних областях. Класичні методи оцінювання параметрів ARIMA моделей — метод максимальної правдоподібності (MLE), метод умової суми квадратів (CSS) та звичайний метод найменших квадратів (OLS) — базуються на фундаментальному припущення гаусовості інновацій. На практиці, це припущення часто порушується, особливо у фінансових та економічних даних, де спостерігаються асиметричні розподіли з важкими хвостами.

**Мета дослідження.** У даній роботі ми розробляємо та досліджуємо застосування методу максимізації поліномів другого порядку (PMM2) для оцінювання параметрів ARIMA( $p, d, q$ ) моделей з негаусовими інноваціями. PMM2, розроблений Ю.П. Кунченко, є напівпараметричним методом, що використовує часткову параметризацію через моменти та кумулянти вищих порядків замість повної функції густини ймовірності.

**Методологія.** Ми розробили повний алгоритм PMM2 для ARIMA моделей, що включає диференціювання ряду, перевірку стаціонарності та ітеративну процедуру Ньютона-Рафсона для розв'язання системи PMM2 рівнянь. Для валідації методу проведено комплексні Monte Carlo симуляції з 2000 повторень для кожної конфігурації, що охоплюють різні розміри вибірки ( $N \in \{100, 200, 500, 1000\}$ ) та чотири типи розподілів інновацій: гаусовий (контроль), гамма Gamma(2, 1) з  $\gamma_3 \approx 1.41$ , логнормальний з  $\gamma_3 \approx 2.0$ , та  $\chi^2(3)$  з  $\gamma_3 \approx 1.63$ .

**Результати.** Емпіричні результати демонструють, що PMM2 забезпечує суттєве підвищення ефективності оцінювання для асиметричних розподілів (відносна ефективність визначається формулою (23)). Для ARIMA(1,1,0) моделі з гамма-розподіленими інноваціями при  $N=500$  отримано відносну ефективність  $RE = 1.58$  (що відповідає 37% зменшенню середньоквадратичної похибки), для логнормального розподілу  $RE = 1.71$  (42% покращення), а для  $\chi^2(3)$   $RE = 1.90$  (47% покращення). Для гаусових інновацій PMM2 демонструє ефективність близьку до OLS ( $RE \approx 1.0$ ), що узгоджується з теорією. Ефективність методу зростає з розміром вибірки та є стабільною для  $N \geq 200$ .

**Практична цінність.** Результати дослідження показують, що PMM2 є ефективним інструментом для аналізу часових рядів з асиметричними інноваціями, що типово

зустрічаються у фінансових та економічних даних. Метод забезпечує суттєве зменшення дисперсії оцінок параметрів без вимог до повної специфікації розподілу похибок, що робить його привабливою альтернативою класичним методам. Надано практичні рекомендації щодо вибору між PMM2 та класичними методами на основі коефіцієнта асиметрії залишків.

**Висновки.** PMM2 є першим застосуванням методу максимізації поліномів до оцінювання параметрів ARIMA моделей. Метод демонструє значні переваги перед класичними підходами для негаусових інновацій, зберігаючи обчислювальну ефективність та простоту імплементації. Напрямки подальших досліджень включають розширення на сезонні SARIMA моделі, інтеграцію з моделями волатильності GARCH, та розробку автоматичних процедур вибору порядку моделі.

**Ключові слова:** ARIMA моделі; метод максимізації поліномів; PMM2; негаусові інновації; оцінювання параметрів; асимптотична ефективність; часові ряди; асиметричні розподіли; Monte Carlo симуляції.

## 1 Вступ

### 1.1 Актуальність Проблеми

Моделі авторегресії та інтегрованого ковзного середнього (ARIMA) залишаються одним з найпоширеніших інструментів аналізу та прогнозування часових рядів у сучасній науці. Починаючи від пionерської роботи Box i Jenkins (1970), ARIMA моделі знайшли застосування у фінансовій економетриці, макроекономічному прогнозуванні, аналізі метеорологічних даних, медичній статистиці та багатьох інших галузях [1, 2].

Класичні методи оцінювання параметрів ARIMA моделей — метод максимальної правдоподібності (MLE), метод умовної суми квадратів (CSS) та звичайний метод найменших квадратів (OLS) — базуються на фундаментальному припущення **гаусовості інновацій** (випадкових похибок). Це припущення забезпечує низку бажаних статистичних властивостей: асимптотичну ефективність оцінок, простоту обчислень та зрозумілу інференцію. Проте, практика аналізу реальних даних систематично демонструє порушення цього припущення.

Останні дослідження надають переконливі емпіричні свідчення негаусовості у різноманітних типах часових рядів:

- **Фінансові часові ряди:** Доходності акцій, обмінні курси та волатильність демонструють асиметричні розподіли з важкими хвостами. Дослідження показують, що навіть після врахування волатильності через GARCH моделі, важкі хвости залишаються [3, 4]. Нещодавнє дослідження Korean stock market підтвердило персистентність важких хвостів навіть після контролю за кризовими періодами та кластеризацією волатильності [5].
- **Економічні показники:** Ціни на сировинні товари, інфляційні дані та торговельні обсяги характеризуються значною асиметрією. Дослідження 15 економік за період 1851-1913 виявило сильний зв'язок між асиметрією цін на товари та інфляцією, при цьому до 48% варіації інфляції пояснюється змінами цін на товари [6].
- **Екологічні та метеорологічні дані:** Вимірювання забруднення, опади, температурні аномалії та сонячна активність часто мають асиметричний характер з екстремальними хвостами.

мальними значеннями. Verma et al. (2025) продемонстрували важкі хвости у даних сонячних спалахів та обговорили теоретичні межі прогнозування за умов важких хвостів [7].

- **Високочастотні фінансові дані:** Mixed-stable моделі, застосовані до DAX компаній на 10-секундних інтервалах, виявили 43-82% нульових змін (стагнаційні ефекти), що потребує спеціальних методів моделювання [8, 9].

## 1.2 Обмеження Класичних Методів

За умов порушення припущення гаусовості, класичні методи оцінювання параметрів ARIMA моделей зазнають суттєвих проблем:

**Систематична зміщеність та неконсистентність.** Pötscher (1991) продемонстрував, що псевдо-максимізатори правдоподібності можуть поводитися драстично інакше, ніж локальні максимізатори, коли розподіл інновацій специфіковано невірно. Gaussian pseudo-likelihood може призводити до неконсистентних оцінок за умов розподільної неспецифікації [10]. Fan, Qi & Xiu (2014) показали, що non-Gaussian квазі-MLE страждає від неконсистентності, якщо квазі-правдоподібність не є справжнім розподілом, пропонуючи двокроковий non-Gaussian QMLE для досягнення консистентності з вищою ефективністю порівняно з Gaussian QMLE [11].

**Втрата статистичної ефективності.** Навіть коли оцінки залишаються консистентними, їх дисперсія може бути суттєво завищеною порівняно з оптимальними оцінками, адаптованими до справжнього розподілу інновацій. Zhang & Sin (2012) показали, що граничні розподіли є сумішшю стабільних та гаусsovих процесів для near-unit root AR процесів з  $\alpha$ -стабільним шумом, демонструючи ускладнення за умов важких хвостів та близькості до одиничного кореня [12].

**Зниження точності прогнозів.** Li et al. (2020) документували, що традиційні ARIMA моделі мають великі відхилення для високочастотного фінансового прогнозування, оскільки фінансові дані демонструють нерегулярні флуктуації, що потребують альтернативних підходів [13]. Dowe et al. (2025) у своїй дуже свіжій роботі показали, що гібридні ARFIMA-ANN підходи краще обробляють складну негаусову динаміку у фінансових та екологічних даних, при цьому використовуючи MML принцип для вибору моделі [14].

**Невірні довірчі інтервали.** Ledolter (1989) продемонстрував, що неврахування викидів збільшує середньоквадратичну похибку прогнозу та спричиняє зміщеність оцінених параметрів, з застосуваннями до даних цін акцій [15]. Це призводить до недооцінки або переоцінки невизначеності прогнозів, що критично важливо для прийняття рішень.

## 1.3 Існуючі Підходи: Короткий Огляд

У відповідь на проблему негаусовості у часових рядах, науковою спільнотою розроблено декілька альтернативних підходів:

**Робастні методи оцінювання (M-estimators).** Започатковані класичною роботою Huber (1964) [16], M-estimators мінімізують робастні функції втрат, що менш чутливі до викидів та важких хвостів. Muler et al. (2009) запровадили BIP-ARMA моделі з MM-оцінками, що уникають поширення викидів через обмежені залишки, досягаючи консистентності та асим-

пточичної нормальності з ефективністю, порівнянною з MLE за нормальністі [17]. Reisen et al. (2024) запропонували M-Whittle estimator з встановленою властивістю консистентності, що добре працює з викидами та шумом з важкими хвостами [18].

**Квантильна регресія та LAD методи.** Katsouris (2023) надав комплексний огляд моделей квантильної регресії часових рядів, що охоплює стаціонарні та нестаціонарні випадки, з Bahadur представленнями для квантильних процесів та рівномірною інференцією у квантильній пороговій регресії [19]. Для ARMA моделей з нескінченною дисперсією, Peng & Yao (2003), Ling (2005) та Zhu & Ling (2015) запропонували зважену оцінку найменших абсолютнох відхилень (WLADE), що є асимптотично нормальню та незміщеною зі стандартною швидкістю збіжності root-n навіть за відсутності скінченної дисперсії [20–22].

**Специфікації з важкими хвостами.** Модифікація класичних ARIMA моделей шляхом заміни гаусових інновацій на розподіли з важкими хвостами (Student-t, Generalized Error Distribution,  $\alpha$ -stable distributions) дозволяє краще моделювати екстремальні події. Wong et al. (2009) розробили Student-t mixture autoregressive модель з вищою гнучкістю порівняно з Gaussian MAR, де ступені свободи є випадковими змінними, використовуючи EM алгоритм для оцінювання параметрів у Байесовому фреймворку [23]. Нещодавнє дослідження 2024 року виявило, що skewed GED найбільш ефективний для фінансових часових рядів порівняно з normal, Student-t, GED та Skewed Student-t розподілами за метриками goodness-of-fit [24].

**Байесовські підходи.** Graves et al. (2015) запропонували систематичний підхід до Байесовської інференції для ARFIMA моделей з новою апроксимативною правдоподібністю для ефективної інференції параметрів у процесах з довгою пам'яттю, що дозволяє інноваціям з широкого класу, включаючи  $\alpha$ -stable та t-розподіли [25]. Байесовські методи також інтегрують невизначеність у всі параметри, забезпечуючи повну постеріорну інференцію замість точкових оцінок.

**Архітектури на основі Transformers.** Останні дослідження показують, що архітектури Transformer, адаптовані для часових рядів, можуть неявно моделювати негаусові характеристики через механізм уваги (attention mechanism) [26]. Гіbridні моделі, що поєднують традиційні ARIMA з attention-based нейронними мережами, демонструють покращену точність прогнозування для фінансових часових рядів [27]. Foundations models, такі як TimeGPT-1 та Chronos, представляють новий напрямок узагальненого прогнозування на різноманітних наборах даних без специфічного навчання [28]. Проте ці моделі є:

- **Data-hungry:** потребують тисяч спостережень для стабільного навчання, тоді як класичні ARIMA працюють з  $N \geq 100$
- **Black-box:** відсутність інтерпретованих параметрів ускладнює економетричну інференцію
- **Computationally intensive:** навчання вимагає GPU та займає години, порівняно з класичними моделями.

Важливо відзначити, що глибинні моделі оптимізовані для мінімізації prediction error, а не для ефективного оцінювання параметрів основного процесу генерації даних.

**Порівняльний аналіз підходів.** Кожен з описаних методів має специфічні переваги та обмеження. Робастні методи (M-estimators) забезпечують стійкість до викидів та зберігають

асимптотичну ефективність за нормальністю, однак можуть втрачати точність за помірних відхилень від гаусовості без екстремальних викидів. Квантільна регресія та LAD методи надають повну картину розподілу та працюють без припущення про скінченну дисперсію, але не оптимізовані для центральних оцінок параметрів та середніх прогнозів. Специфікації з важкими хвостами явно моделюють екстремальні події та забезпечують інтерпретовані параметри форми розподілу, проте вимагають *a priori* вибору сімейства розподілів, що може призводити до model misspecification. Байесовські методи інтегрують невизначеність параметрів та дозволяють гнучкі специфікації інновацій, але є обчислювально інтенсивними та потребують вибору prior розподілів. Нарешті, Transformer-based моделі досягають високої точності прогнозування та адаптуються до складних патернів, однак не забезпечують інтерпретовану параметризацію процесу генерації даних.

Таким чином, існує потреба у підході, який би поєднував статистичну ефективність класичних методів, інтерпретованість параметрів та робастність до відхилень від нормальності без необхідності специфікації конкретного сімейства розподілів.

#### 1.4 Метод Максимізації Поліномів: Альтернативний Підхід

Метод максимізації поліномів (Polynomial Maximization Method, PMM), розроблений українським вченим Ю.П. Кунченко, представляє альтернативну філософію статистичного оцінювання [29, 30]. На відміну від класичного методу максимальної правдоподібності, який потребує повної специфікації густини ймовірності, PMM базується на **частковій імовірнісній параметризації** через моменти та кумулянти вищих порядків.

Центральною конструкцією методу є максимізація стохастичного полінома порядку  $S$  відносно параметрів моделі. Ключова ідея полягає в тому, що замість максимізації повної функції правдоподібності, метод максимізує вибіркову статистику в околі справжніх значень оцінюваних параметрів [30, 31].

PMM метод успішно застосовувався до різноманітних задач статистичного оцінювання:

- **Лінійна регресія:** Zabolotnii et al. (2018) продемонстрували застосування PMM2 до лінійної регресії з асиметричним розподілом похибок, досягаючи зменшення дисперсії на 15-35% порівняно з OLS для gamma та lognormal розподілів [32].
- **Поліноміальна регресія:** Zabolotnii et al. (2021) розширили метод на поліноміальну регресію з розподілом експоненціальної потужності (generalized Gaussian distribution), підтверджуючи ефективність через Monte Carlo та bootstrap симуляції [33].
- **Обробка сигналів:** Palahin & Juhár (2016) застосували PMM до спільногого оцінювання параметрів сигналу у негаусовому шумі, показавши, що нелінійна обробка через кумулянти третього та вищих порядків може зменшити дисперсію спільногого оцінювання параметрів порівняно з конвенційними методами [34].
- **Метрологічні вимірювання:** Warsza & Zabolotnii (2017, 2018) використали PMM для оцінювання параметрів вимірювань з негаусовими симетричними та асиметричними розподілами даних, розробляючи методику PMM3 для симетричних розподілів [35, 36].

Варто відзначити, що PMM метод позиціонується між класичним методом моментів та методом максимальної правдоподібності. На відміну від узагальненого методу моментів

(GMM) Hansen (1982) [37], який мінімізує зважену суму квадратів відхилень між вибірковими та популяційними моментами, РММ максимізує стохастичний поліном, використовуючи для його побудови моменти або кумулянти вищих порядків.

### 1.5 Дослідницька Прогалина та Внесок Роботи

Незважаючи на успішне застосування РММ до регресійних задач та обробки сигналів, його систематичне використання для оцінювання параметрів ARIMA моделей з негаусовими інноваціями залишається недостатньо дослідженим. Існує кілька ключових дослідницьких прогалин:

**Нерозвиненість моментно-кумулянтних методів для часових рядів.**Хоча моменти або кумулянти вищих порядків широко використовуються в обробці сигналів та спектральному аналізі, їх застосування до оцінювання параметрів моделей часових рядів обмежене. Більшість методів для негаусових ARIMA зосереджені на робастних функціях втрат або специфікації розподілів, але не на явній експлуатації моментно-кумулянтного опису.

**Недостатня увага до асиметричних інновацій.**Більшість робіт з негаусових ARIMA фокусуються на симетричних розподілах з важкими хвостами (Student-t, GED). Асиметричні розподіли, які РММ2 спеціально адресує, отримують менше уваги, незважаючи на їх емпіричну поширеність у фінансових доходностях та економічних показниках.

**Методологічний розрив між регіональними дослідницькими спільнотами.**Метод Кунченка, незважаючи на сильні теоретичні основи та успішні застосування в Східній Європі, залишається малознайомим у західній літературі з часових рядів. Ця робота має на меті інтегрувати східноєвропейську статистичну методологію з західною економетричною літературою часових рядів (Box-Jenkins, ARIMA).

**Відсутність порівняльних досліджень ефективності.**Порівняльні дослідження зазвичай порівнюють MLE, M-estimators, LAD та квантільну регресію. Порівняння ефективності моментно-кумулянтних методів, таких як РММ, відносно цих альтернатив відсутні для ARIMA моделей.

Дане дослідження призначено для заповнення цих прогалин.

## 2 Методологія

У цьому розділі ми надаємо повну методологію застосування методу максимізації поліномів другого порядку (РММ2) до оцінювання параметрів ARIMA моделей з негаусовими інноваціями. Спочатку формулюємо ARIMA модель та класичні методи оцінювання, потім розглядаємо теоретичні основи РММ, адаптуємо метод до контексту часових рядів, та надаємо алгоритм реалізації з асимптотичною теорією.

### 2.1 ARIMA Моделі: Основи та Класичне Оцінювання

Розглянемо стандартну ARIMA( $p,d,q$ ) модель для часового ряду  $\{y_t\}_{t=1}^T$ , де  $d$ -та різниця  $z_t = \Delta^d y_t = (1 - B)^d y_t$  задовольняє стаціонарну та оборотну ARMA( $p,q$ ) модель:

$$\Phi(B)z_t = \Theta(B)\varepsilon_t \quad (1)$$

де  $B$  — оператор зсуву,  $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  та  $\Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$  — поліноми авторегресії та ковзного середнього, а  $\{\varepsilon_t\}$  — послідовність i.i.d. інновацій з  $\mathbb{E}[\varepsilon_t] = 0$  та  $\text{Var}[\varepsilon_t] = \sigma^2$ . Корені характеристичних рівнянь  $\Phi(z) = 0$  та  $\Theta(z) = 0$  лежать поза одиничним колом, що гарантує стаціонарність та оборотність відповідно.

**Класичні методи оцінювання.** Нехай  $\boldsymbol{\theta} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)^\top$  — вектор параметрів розміру  $k = p + q$ . Два основні методи використовуються як базові для порівняння:

Метод найменших квадратів (OLS) для AR компоненти:

$$\hat{\boldsymbol{\phi}}_{\text{OLS}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{z} \quad (2)$$

де  $\mathbf{z} = (z_{p+1}, \dots, z_T)^\top$  та  $\mathbf{X}$  — матриця лагових значень.

Метод максимальної правдоподібності (MLE) за припущення  $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ :

$$\hat{\boldsymbol{\theta}}_{\text{MLE}} = \arg \max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta} | \mathbf{y}) = \arg \max_{\boldsymbol{\theta}} \left\{ -\frac{T}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^T \varepsilon_t^2(\boldsymbol{\theta}) \right\} \quad (3)$$

За умови нормальності інновацій, MLE досягає асимптотичної ефективності:  $\sqrt{T}(\hat{\boldsymbol{\theta}}_{\text{MLE}} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(0, \mathbf{I}^{-1}(\boldsymbol{\theta}_0))$ , де  $\mathbf{I}(\boldsymbol{\theta}_0)$  — інформаційна матриця Фішера. Однак за негаусовості інновацій MLE втрачає оптимальність, а OLS залишається консистентним, але неефективним. Ця проблема мотивує розробку альтернативних методів оцінювання, адаптованих до негаусових розподілів.

## 2.2 Теоретичні Основи Методу Максимізації Поліномів

### 2.2.1 Стохастичні Поліноми

Метод максимізації поліномів (PMM) базується на концепції стохастичних поліномів, що є поліноміальними функціями випадкових величин з коефіцієнтами, що залежать від параметрів моделі. Цей метод було розроблено для оцінювання параметрів коли імовірнісні властивості даних суттєво відрізняються від гаусового ( нормальног о) закону.

**Визначення 2.1** (Стохастичний поліном порядку  $S$  загального виду). Для послідовності випадкових величин  $y_v$ ,  $v = \overline{1, N}$ , та векторного параметра  $\mathbf{a}$ , стохастичний поліном порядку  $S$  загального виду визначається як:

$$L_{SN} = \sum_{v=1}^N \sum_{i=1}^S \phi_i(y_v) \int k_{iv}(a) dz - \sum_{i=1}^S \sum_{v=1}^N \int \Psi_{iv} k_{iv}(a) dz \quad (4)$$

де  $\phi_i(y_v)$  — базисні функції,  $k_{iv}(a)$  — вагові коефіцієнти, що залежать від параметра  $a$ , та  $\Psi_{iv} = E\{\phi_i(y_v)\}$  — математичні сподівання базисних функцій, які є двічі диференційовними по параметру  $a$ .

**Фундаментальні властивості стохастичного полінома.** Стохастичний поліном  $L_{SN}$  виду (4) володіє двома основними властивостями [30]:

- Для будь-якого порядку  $S$  при асимптотичному зростанні обсягу вибірки  $N \rightarrow \infty$  поліном  $L_{SN}$  як функція параметра  $a$  приймає максимум в околиці істинного значення цього параметра;

2. При різних вибірках відхилення максимуму полінома  $L_{SN}$  від істинного значення параметра  $a$  має мінімальну дисперсію для відповідного порядку полінома  $S$ .

За аналогією до методу максимальної правдоподібності, оцінку параметра  $a$  можна знайти із розв'язання рівняння:

$$\frac{d}{da} L_{SN} \Big|_{a=\hat{a}} = \sum_{i=1}^S \sum_{v=1}^N k_{iv} [\phi_i(y_v) - \Psi_{iv}] \Big|_{a=\hat{a}} = 0 \quad (5)$$

**Оптимальні коефіцієнти та система рівнянь.** Оптимальні коефіцієнти  $k_{iv}$ , що максимізують функціонал (4), знаходяться з розв'язання системи лінійних алгебраїчних рівнянь:

$$\sum_{j=1}^S k_{jv} F_{(i,j)v} = \frac{d}{da} \Psi_{iv}, \quad i = \overline{1, S}, \quad v = \overline{1, N} \quad (6)$$

де  $F_{(i,j)v} = \Psi_{(i,j)v} - \Psi_{iv} \Psi_{jv}$ ,  $\Psi_{(i,j)v} = E\{\phi_i(y_v)\phi_j(y_v)\}$ ,  $i, j = \overline{1, S}$ .

**Векторний параметр та багатопараметричне оцінювання.** Для випадку знаходження оцінок векторного параметра  $\boldsymbol{\theta} = (a_0, a_1, \dots, a_{Q-1})^\top$  необхідно використовувати  $Q$  поліномів  $L_{SN}^{(p)}$ ,  $p = \overline{0, Q-1}$  загального виду (4) для кожної компоненти  $a_p$  векторного параметра.

Кожний  $p$ -ий стохастичний поліном  $L_{SN}^{(p)}$  як функція параметра  $a_p$  при відомих значеннях інших складових вектору  $\boldsymbol{\theta}$  при  $N \rightarrow \infty$  також має максимум в околиці істинного значення параметра  $a_p$ . Шукані оцінки параметра знаходяться як розв'язок системи рівнянь:

$$f_{SN}^{(p)}(y_v, x_v) = \sum_{i=1}^S \sum_{v=1}^N k_{iv}^{(p)} [\phi_i(y_v) - \Psi_{iv}] \Big|_{a_p=\hat{a}_p} = 0, \quad p = \overline{0, Q-1} \quad (7)$$

## 2.2.2 РММ для Асиметричних Розподілів

Розглянемо застосування стохастичних поліномів на прикладі лінійної багатофакторної регресії з асиметрично розподіленою випадковою складовою [32]. Як базисні функції використовуються степеневі перетворення, математичними сподіваннями від яких є моменти відповідного порядку. Нехай спостереження  $\{y_v\}_{v=1}^N$  описуються моделлю

$$y_v = \mathbf{x}_v^\top \boldsymbol{\theta} + \xi_v, \quad \mathbf{x}_v = (1, x_{1,v}, \dots, x_{Q-1,v})^\top, \quad (8)$$

де  $\boldsymbol{\theta} = (a_0, a_1, \dots, a_{Q-1})^\top$  — вектор регресійних параметрів, а випадкова похибка  $\xi_v$  задовільняє умовам

$$\mathbb{E}[\xi_v] = 0, \quad \mathbb{E}[\xi_v^2] = \mu_2 > 0, \quad \mathbb{E}[\xi_v^3] = \mu_3 \neq 0, \quad \mathbb{E}[\xi_v^4] = \mu_4 < \infty.$$

Позначимо також  $\eta_v(\boldsymbol{\theta}) = \mathbf{x}_v^\top \boldsymbol{\theta}$  та  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]^\top$ .

**РММ1: лінійний стохастичний поліном і еквівалентність МНК.** При степені  $S = 1$  базисною функцією обираємо  $\phi_1(y_v) = y_v$ , що дає математичне сподівання  $\Psi_{1v} = \mathbb{E}[y_v] = \eta_v(\boldsymbol{\theta})$ . Коваріація  $F_{(1,1)v} = \mu_2$  є сталою, а оптимальні коефіцієнти поліному (4) обчислюю-

ться з (6) як  $k_{1,v}^{(p)} = x_{p,v}/\mu_2$ , де  $x_{0,v} \equiv 1$ . Умови максимуму полінома зводяться до

$$\sum_{v=1}^N x_{p,v} [y_v - \eta_v(\boldsymbol{\theta})] = 0, \quad p = \overline{0, Q-1}, \quad (9)$$

що еквівалентно класичній системі нормальних рівнянь  $\mathbf{X}^\top \mathbf{X} \boldsymbol{\theta} = \mathbf{X}^\top \mathbf{y}$ . Отже, РММ1 відтворює МНК-оцінки

$$\hat{\boldsymbol{\theta}}_{\text{PMM1}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}, \quad (10)$$

які залишаються оптимальними лише за гаусового розподілу похибок.

**РММ2: стохастичний поліном другого порядку.** Для врахування асиметрії формуємо стохастичний поліном із базисними функціями

$$\phi_1(y_v) = y_v, \quad \Psi_{1v} = \eta_v(\boldsymbol{\theta}), \quad (11)$$

$$\phi_2(y_v) = y_v^2, \quad \Psi_{2v} = \eta_v^2(\boldsymbol{\theta}) + \mu_2. \quad (12)$$

Матричні елементи  $F_{(i,j)v} = \Psi_{(i,j)v} - \Psi_{iv}\Psi_{jv}$  визначаються центральними моментами до четвертого порядку, а оптимальні коефіцієнти (6) набувають вигляду

$$k_{1,v}^{(p)} = \frac{\mu_4 - \mu_2^2 + 2\mu_3\eta_v(\boldsymbol{\theta})}{\Delta} x_{p,v}, \quad k_{2,v}^{(p)} = -\frac{\mu_3}{\Delta} x_{p,v}, \quad (13)$$

де

$$\Delta = \mu_2(\mu_4 - \mu_2^2) - \mu_3^2 > 0. \quad (14)$$

Вектор рівнянь для оцінювання  $\boldsymbol{\theta}$  набуває узагальненого вигляду

$$g_p(\boldsymbol{\theta}) = \sum_{v=1}^N x_{p,v} \left\{ \frac{\mu_4 - \mu_2^2 + 2\mu_3\eta_v(\boldsymbol{\theta})}{\Delta} [y_v - \eta_v(\boldsymbol{\theta})] - \frac{\mu_3}{\Delta} [y_v^2 - \eta_v^2(\boldsymbol{\theta}) - \mu_2] \right\} = 0, \quad (15)$$

де  $p = \overline{0, Q-1}$ . Система (15) переходить до МНК при  $\mu_3 = 0$ .

Помноживши рівняння (15) на  $\Delta$  та згрупувавши члени за ступенями  $\eta_v(\boldsymbol{\theta})$ , одержуємо еквівалентну квадратичну систему

$$\sum_{v=1}^N x_{p,v} [A_2\eta_v^2(\boldsymbol{\theta}) + B_{2,v}\eta_v(\boldsymbol{\theta}) + C_{2,v}] = 0, \quad p = \overline{0, Q-1}, \quad (16)$$

де коефіцієнти:

$$A_2 = \mu_3, \quad B_{2,v} = (\mu_4 - \mu_2^2) - 2\mu_3 y_v, \quad C_{2,v} = \mu_3 y_v^2 - y_v(\mu_4 - \mu_2^2) - \mu_2 \mu_3. \quad (17)$$

**Узагальнена матрична форма та формула Ньютона–Рафсона.** Позначимо  $\mathbf{g}(\boldsymbol{\theta}) = (g_0(\boldsymbol{\theta}), \dots, g_{Q-1}(\boldsymbol{\theta}))^\top$  та введемо

$$\lambda_v(\boldsymbol{\theta}) = \frac{2\mu_3 [y_v - \eta_v(\boldsymbol{\theta})] - (\mu_4 - \mu_2^2)}{\Delta}. \quad (18)$$

Тоді матриця похідних  $\mathbf{J}_{\text{PMM2}}(\boldsymbol{\theta}) = \partial \mathbf{g}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}^\top$  набуває компактного запису

$$\mathbf{J}_{\text{PMM2}}(\boldsymbol{\theta}) = \sum_{v=1}^N \lambda_v(\boldsymbol{\theta}) \mathbf{x}_v \mathbf{x}_v^\top, \quad (19)$$

Крок Ньютона–Рафсона для знаходження  $\hat{\boldsymbol{\theta}}_{\text{PMM2}}$  має вигляд

$$\boldsymbol{\theta}^{(m+1)} = \boldsymbol{\theta}^{(m)} - \mathbf{J}_{\text{PMM2}}^{-1}(\boldsymbol{\theta}^{(m)}) \mathbf{g}(\boldsymbol{\theta}^{(m)}), \quad (20)$$

де як початкове наближення  $\boldsymbol{\theta}^{(0)}$  доцільно використовувати МНК-оцінку (10).

**Адаптивна процедура.** У практичних застосуваннях моменти  $\mu_2, \mu_3, \mu_4$  невідомі, тому їх замінюють вибірковими оцінками на основі поточних залишків. Стандартна ітераційна схема має вигляд:

1. **Крок 1:** Обчислити МНК-оцінку  $\hat{\boldsymbol{\theta}}^{\text{OLS}}$  та сформувати залишки  $\hat{\xi}_v^{(0)} = y_v - \mathbf{x}_v^\top \hat{\boldsymbol{\theta}}^{\text{OLS}}$ .
2. **Крок 2:** Для ітерації  $m$  оновити моменти за формулами

$$\hat{\mu}_r^{(m)} = \frac{1}{N} \sum_{v=1}^N \left( \hat{\xi}_v^{(m)} \right)^r, \quad r \in \{2, 3, 4\},$$

та визначити ексцес  $\hat{\gamma}_4^{(m)} = \hat{\mu}_4^{(m)} / (\hat{\mu}_2^{(m)})^2 - 3$  й коефіцієнт асиметрії  $\hat{\gamma}_3^{(m)} = \hat{\mu}_3^{(m)} / (\hat{\mu}_2^{(m)})^{3/2}$ .

3. **Крок 3:** Якщо  $|\hat{\gamma}_3^{(m)}| < 0.1$ , доцільно залишитись на МНК-оцінці. Інакше розв'язати (15) методом (20) з використанням  $\hat{\mu}_r^{(m)}$  та отримати новий вектор  $\boldsymbol{\theta}^{(m+1)}$ .
4. **Крок 4:** Оновити залишки  $\hat{\xi}_v^{(m+1)} = y_v - \mathbf{x}_v^\top \boldsymbol{\theta}^{(m+1)}$  та повторити кроки 2–4 до збіжності.

### 2.2.3 Асимптотичні Дисперсії та Ефективність ММПл-Оцінок

**Матриця кількості добутої інформації.** Для отримання аналітичних виразів, що описують дисперсії ММПл-оцінок векторного параметра  $\boldsymbol{\theta}$ , використовується матриця кількості добутої інформації про компоненти параметру при застосуванні стохастичних поліномів порядку  $S$ . Така матриця  $\mathbf{J}_s(\boldsymbol{\theta})$  складається із елементів:

$$J_{SN}^{(p,q)} = \sum_{v=1}^N \sum_{i=1}^S \sum_{j=1}^S k_{iv}^{(p)} k_{jv}^{(q)} F_{(i,j)v} = \sum_{v=1}^N \sum_{i=1}^S k_{i,v}^{(p)} \frac{\partial}{\partial a_q} \Psi_{iv}, \quad p, q = \overline{0, Q-1} \quad (21)$$

У статистичному сенсі кількість добутої інформації концептуально є поняттям близьким до кількості інформації по Фішеру. Дисперсії ММПл-оцінок складових векторного параметру в асимптотичному випадку (при  $N \rightarrow \infty$ ) можуть бути отримані як елементи головної діагоналі варіаційної матриці  $\mathbf{V}_s(\boldsymbol{\theta})$ , яка є оберненою до матриці (21):

$$\mathbf{V}_{\text{PMMS}}(\boldsymbol{\theta}) = [\mathbf{J}_S(\boldsymbol{\theta})]^{-1} \quad (22)$$

**Властивість збіжності до межі Рао–Крамера.** Важлива властивість ММПл полягає у тому, що при збільшенні числа членів стохастичного поліному  $S$  дисперсія оцінок зменшується, оскільки кількість добутої інформації асимптотично (при  $S \rightarrow \infty$ ) прямує до

інформації по Фішеру.

**Теорема 2.2** (Відносна ефективність PMM2 щодо OLS). За умови, що інновації  $\varepsilon_t$  мають скінченні моменти до четвертого порядку включно та асиметричний розподіл ( $\gamma_3 \neq 0$ ), відносна ефективність PMM2 щодо OLS визначається як

$$RE_{PMM2/OLS} = \frac{\text{Var}(\hat{\theta}_{OLS})}{\text{Var}(\hat{\theta}_{PMM2})} = \frac{2 + \gamma_4}{2 + \gamma_4 - \gamma_3^2}, \quad (23)$$

де  $\gamma_3$  та  $\gamma_4$  — стандартизовані коефіцієнти асиметрії та ексцесу.

*Ескіз доведення.* Відношення інформаційних матриць  $J_{SN}^{(p,q)} = \mathbb{E} \left[ \frac{\partial^2 Q_S}{\partial \theta_p \partial \theta_q} \right]$  для  $S = 1$  (збігається з OLS) та  $S = 2$  (з оптимальними коефіцієнтами) дає  $\mathbf{V}_{PMM2} = RE_{PMM2/OLS}^{-1} \cdot \mathbf{V}_{OLS}$ . Перехід до стандартизованих кумулянтів дає формулу (23). Детальне доведення наведено в [30].  $\square$

### Інтерпретація відносної ефективності.

- Для гаусових інновацій ( $\gamma_3 = 0, \gamma_4 = 0$ ):  $RE = 1$ , тобто PMM еквівалентний OLS, оскільки оцінки співпадають. Для асиметричних розподілів ( $\gamma_3 \neq 0$ ):  $RE > 1$ , що означає виграну PMM2 у точності, що залежить лише від коефіцієнтів асиметрії  $\gamma_3$  та ексцесу  $\gamma_4$ .
- Коефіцієнт  $RE$  є безрозмірною величиною з діапазону  $[1; \infty)$ . При наближенні кумулянтних коефіцієнтів до меж допустимих областей відносна ефективність PMM2 зростає необмежено.
- Асимптотично відносна ефективність є однаковою для всіх компонентів векторного параметра та не залежить від типу регресійної моделі (лінійна, поліноміальна чи нелінійна). При скінчених вибірках можливі відхилення через варіабельність оцінювання.

## 2.3 PMM2 для ARIMA Моделей: Адаптація Методу

### 2.3.1 Мотивація та принцип наближення

Ключовим елементом адаптації PMM2 до ARIMA-процесів є етап попередньої стаціонаризації: диференціювання  $d$ -го порядку забезпечує застосовність конструкцій із підрозділу 2.2 до стаціонарного ряду  $z_t$ . Саме після цієї трансформації базовий метод надійно відтворює інновації, а PMM2-корекція поліпшує оцінки в умовах асиметрії. Такий підхід зберігає головну перевагу PMM2 — чутливість до вищих моментів — і водночас уникає складних рекурсій для похідних псевдорегресорів. Це дає можливість замість повного переписування рекурсій для кожного набору параметрів застосовувати просту дворівневу схему:

1. **Базовий крок.** Оцінюємо модель ARIMA( $p, d, q$ ) стандартним методом (CSS або ML) та зберігаємо отримані залишки як емпіричні інновації.
2. **PMM2-корекція.** Фіксуємо побудовану на попередньому кроці дизайн-матрицю та застосовуємо поліноміальне коригування другого порядку, що враховує асиметрію й ексцес інновацій.

### 2.3.2 Конструкція псевдорегресорів

Нехай

$$z_t = \Delta^d y_t, \quad t = d + 1, \dots, T, \quad n = T - d, \quad (24)$$

— стаціонаризований ряд. На першому кроці оцінюємо ARIMA( $p, d, q$ ) класичним методом і одержуємо залишки  $\hat{\varepsilon}_t^{\text{CSS}}$ . Для ефективної довжини  $n_{\text{eff}} = n - m$ , де  $m = \max(p, q)$ , формуємо вектор регресорів

$$\mathbf{x}_t = (z_{t-1}, \dots, z_{t-p}, \hat{\varepsilon}_{t-1}^{\text{CSS}}, \dots, \hat{\varepsilon}_{t-q}^{\text{CSS}})^\top, \quad t = m + 1, \dots, n. \quad (25)$$

Якщо у моделі присутній зсув, до  $\mathbf{x}_t$  додається одиничний стовпчик. Побудована матриця  $\mathbf{X} = (\mathbf{x}_{m+1}, \dots, \mathbf{x}_n)^\top$  не залежить від параметрів, тому подальша оптимізація редукується до задачі, описаної у підрозділі 2.2.

### 2.3.3 Моментне калібрування й стохастичний поліном

Використовуючи залишки базового кроку, розраховуємо центральні моменти

$$\hat{\mu}_k = \frac{1}{n_{\text{eff}}} \sum_{t=m+1}^n \left( \hat{\varepsilon}_t^{\text{CSS}} - \bar{\varepsilon} \right)^k, \quad k = 2, 3, 4, \quad (26)$$

де  $\bar{\varepsilon}$  — їхнє вибіркове середнє. Як і у базовому РММ2, визначаємо

$$\hat{\Delta} = \hat{\mu}_2(\hat{\mu}_4 - \hat{\mu}_2^2) - \hat{\mu}_3^2. \quad (27)$$

Нехай  $\boldsymbol{\theta} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)^\top$  та  $\eta_t(\boldsymbol{\theta}) = \mathbf{x}_t^\top \boldsymbol{\theta}$ . Тоді стохастичний поліном набуває вигляду

$$g_j(\boldsymbol{\theta}) = \sum_{t=m+1}^n x_{j,t} \left[ \frac{\hat{\mu}_4 - \hat{\mu}_2^2 + 2\hat{\mu}_3\eta_t(\boldsymbol{\theta})}{\hat{\Delta}} (z_t - \eta_t(\boldsymbol{\theta})) - \frac{\hat{\mu}_3}{\hat{\Delta}} (z_t^2 - \eta_t^2(\boldsymbol{\theta}) - \hat{\mu}_2) \right] = 0, \quad (28)$$

де  $x_{j,t}$  —  $j$ -та компонента  $\mathbf{x}_t$ . Система (28) є прямим аналогом рівнянь (15) для фіксованої дизайн-матриці. У граничному симетричному випадку ( $\hat{\mu}_3 = 0$ ) вона зводиться до МНК-оцінювання з ваговим множником  $\hat{\mu}_2^{-1}$ .

## 2.4 Алгоритм Оцінювання PMM2 для ARIMA

---

**Algorithm 1** Спрощена PMM2-оцінка для ARIMA( $p, d, q$ )

---

**Require:** Часовий ряд  $\{y_t\}_{t=1}^T$ , порядки  $(p, d, q)$ , вибір методу ініціалізації (CSS або ML)

**Ensure:** Вектор параметрів  $\hat{\boldsymbol{\theta}}_{\text{PMM2}}$ , оцінка моментів  $\hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_4$

- 1: **Диференцювання.** Обчислити  $z_t = \Delta^d y_t$  згідно з (24).
  - 2: **Базове оцінювання.** Одержані  $\hat{\phi}_j^{\text{CSS}}, \hat{\theta}_k^{\text{CSS}}$  та залишки  $\hat{\varepsilon}_t^{\text{CSS}}$  стандартним методом.
  - 3: **Побудова дизайн-матриці.** Сформувати  $\mathbf{X}$  із рядків (25) та відповідний вектор відгуків  $\mathbf{z} = (z_{m+1}, \dots, z_n)^\top$ .
  - 4: **Оцінювання моментів.** Розрахувати  $\hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_4$  за формулою (26) і величину  $\hat{\Delta}$  з (27).
  - 5: **Ініціалізація PMM2.** Взяти початкове наближення  $\boldsymbol{\theta}^{(0)} = (\hat{\phi}_1^{\text{CSS}}, \dots, \hat{\phi}_p^{\text{CSS}}, \hat{\theta}_1^{\text{CSS}}, \dots, \hat{\theta}_q^{\text{CSS}})^\top$ .
  - 6: **Поліноміальна оптимізація.** Застосувати ітеративний розв'язувач PMM2 для фіксованої дизайн-матриці  $\mathbf{X}$  (див. рівняння (28)). Процес зупиняємо, коли  $\|\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(k-1)}\|$  та норма моментних умов стають меншими за заданий поріг.
  - 7: **Реконструкція залишків.** Прогнати фінальні параметри через модель ARIMA для отримання оновлених інновацій (вони використовуються у подальших діагностичних тестах).
  - 8: **return**  $\hat{\boldsymbol{\theta}}_{\text{PMM2}} = \boldsymbol{\theta}^{(k*)}, \hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_4$ .
- 

**Коментарі до реалізації.**

- Алгоритм використовує лише одну дизайн-матрицю, побудовану з оцінок першого кроку. Тому обчислювальна складність визначається матричними добутками  $O(n_{\text{eff}} k)$  та розв'язанням невеликої системи  $k \times k$  на кожній ітерації PMM2.
- Перевірка стаціонарності й оборотності виконується проекцією коефіцієнтів на допустиму область: при виявленні коренів усередині одиничного кола коефіцієнти масштабуються до граничного значення.

## 2.5 Асимптотичні Властивості PMM2 для ARIMA

Для аналізу зручно переписати систему (28) у вигляді усереднених моментних умов. Нехай  $\mathbf{x}_t^0$  — “ідеальні” регресори, побудовані з істинних інновацій  $\varepsilon_t$ , а  $\hat{\mathbf{x}}_t$  — їхні емпіричні аналоги з формули (25). Позначимо через

$$\psi_t(\boldsymbol{\theta}) = \hat{\mathbf{x}}_t s_t(\boldsymbol{\theta})$$

відповідний моментний внесок, де  $s_t(\boldsymbol{\theta})$  — вираз у дужках у (28). Тоді оцінювач визначається умовою

$$\mathbf{g}_{n_{\text{eff}}}(\boldsymbol{\theta}) = \frac{1}{n_{\text{eff}}} \sum_{t=m+1}^n \psi_t(\boldsymbol{\theta}) = \mathbf{0}.$$

### 2.5.1 Вплив згенерованих регресорів

Двокрокова процедура PMM2 для ARIMA моделей використовує залишки  $\hat{\varepsilon}_t^{\text{CSS}}$  з першого кроку як регресори у другому кроці. Це класична проблема *згенерованих регресорів* (generated regressors), досліджена Pagan (1984) [38] та Newey (1984) [39]. Наступна лема формалізує умови, за яких така двокрокова процедура не впливає на асимптотичну коваріаційну матрицю PMM2-оцінок.

**Лема 2.3** (Асимптотична еквівалентність з істинними регресорами). *Нехай виконуються такі умови:*

1. *Початкові оцінки  $\hat{\boldsymbol{\theta}}^{CSS}$  є  $\sqrt{n}$ -консистентними:  $\sqrt{n}(\hat{\boldsymbol{\theta}}^{CSS} - \boldsymbol{\theta}_0) = O_p(1)$ .*
2. *Залишки задовільняють  $\sup_t |\hat{\varepsilon}_t^{CSS} - \varepsilon_t| = O_p(n^{-1/2})$ .*
3. *Функції стохастичного полінома є гладкими (Lipschitz-неперервними) за регресорами.*

Тоді асимптотичний розподіл РММ2-оцінок, отриманих з використанням  $\hat{\varepsilon}_t^{CSS}$ , співпадає з розподілом, який би отримали за використання істинних  $\varepsilon_t$ :

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{PMM2}(\hat{\varepsilon}^{CSS}) - \boldsymbol{\theta}_0) - \sqrt{n}(\hat{\boldsymbol{\theta}}_{PMM2}(\varepsilon) - \boldsymbol{\theta}_0) = o_p(1).$$

*Ескіз доведення.* Застосовуємо розкладання першого порядку за  $(\hat{\varepsilon}^{CSS} - \varepsilon)$ :

$$\mathbf{g}_n(\boldsymbol{\theta}, \hat{\varepsilon}^{CSS}) = \mathbf{g}_n(\boldsymbol{\theta}, \varepsilon) + \mathbf{H}_n(\boldsymbol{\theta})(\hat{\varepsilon}^{CSS} - \varepsilon) + o_p(n^{-1/2}),$$

де  $\mathbf{H}_n$  — матриця похідних за залишками. За умовою (1)–(2), другий член має порядок  $O_p(n^{-1/2}) \cdot O_p(n^{-1/2}) = O_p(n^{-1})$ , що є асимптотично незначущим. Детальне доведення аналогічне до Pagan (1984, Theorem 1) та Newey (1984, Proposition 1).  $\square$

**Наслідок.** За умов леми 2.3, асимптотична коваріаційна матриця РММ2-оцінок обчислюється стандартним способом (сендвіч-формула), без необхідності корекції на першокроковий оцінювач. Це виправдовує використання класичних формул для стандартних похибок у розділі 2.4.

### 2.5.2 Консистентність

**Теорема 2.4** (Консистентність спрощеної РММ2-оцінки). *Нехай виконуються такі умови:*

1. *ARIMA( $p, d, q$ ) модель правильно специфікована; інновації  $\varepsilon_t$  є стаціонарними, ергодичними та мають скінченні моменти до четвертого порядку.*
2. *Початковий CSS/ML-оцінювач є консистентним:  $\hat{\boldsymbol{\theta}}^{CSS} \xrightarrow{p} \boldsymbol{\theta}_0$ .*
3. *Ряд  $\{\hat{\varepsilon}_t^{CSS}\}$  збігається в ймовірності до істинних інновацій у сенсі середнього квадратичного:  $\frac{1}{n_{eff}} \sum (\hat{\varepsilon}_t^{CSS} - \varepsilon_t)^2 \xrightarrow{p} 0$ .*
4. *Матриця моментів  $E[\mathbf{x}_t^0(\mathbf{x}_t^0)^\top]$  невироджена.*

Тоді оцінювач  $\hat{\boldsymbol{\theta}}_{PMM2}$  є консистентним:

$$\hat{\boldsymbol{\theta}}_{PMM2} \xrightarrow{p} \boldsymbol{\theta}_0.$$

*Ескіз.* Пункти (2)–(3) гарантують збіжність  $\hat{\mathbf{x}}_t \xrightarrow{p} \mathbf{x}_t^0$  та  $\hat{\mu}_k \xrightarrow{p} \mu_k$  для  $k = 2, 3, 4$ . Враховуючи неперервність  $s_t(\boldsymbol{\theta})$  за цими аргументами, отримаємо рівномірну збіжність  $\mathbf{g}_{n_{eff}}(\boldsymbol{\theta}) \rightarrow \mathbf{g}(\boldsymbol{\theta}) = E[\mathbf{x}_t^0 s_t^0(\boldsymbol{\theta})]$ . Єдиність кореня  $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$  (за умовою 4) і стандартний аргумент Z-оцінок (Newey & McFadden, 1994) завершують доведення.  $\square$

### 2.5.3 Асимптотична нормальність

**Теорема 2.5** (Асимптотичний розподіл). *За умов Теореми 2.4 та додатково за умови, що  $\{\varepsilon_t\}$  задовільняють центральній граничній теоремі для квадратно інтегровних функцій,*

маємо

$$\sqrt{n_{\text{eff}}}(\hat{\boldsymbol{\theta}}_{\text{PMM2}} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_{\text{PMM2}}),$$

де

$$\mathbf{A} = E\left[\frac{\partial \psi_t^0(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}^\top}\right], \quad \mathbf{B} = E\left[\psi_t^0(\boldsymbol{\theta}_0) \psi_t^0(\boldsymbol{\theta}_0)^\top\right], \quad (29)$$

$$\boldsymbol{\Sigma}_{\text{PMM2}} = \mathbf{A}^{-1} \mathbf{B} (\mathbf{A}^{-1})^\top, \quad (30)$$

а  $\psi_t^0(\cdot)$  — моментна функція, побудована з істинних інновацій.

На практиці матриці  $\mathbf{A}$  та  $\mathbf{B}$  апроксимуються вибірковими аналогами, в яких використовуються оцінені регресори та моменти. Стандартні похибки обчислюємо за формулою

$$\text{SE}(\hat{\theta}_j) = \sqrt{\frac{[\hat{\boldsymbol{\Sigma}}_{\text{PMM2}}]_{jj}}{n_{\text{eff}}}}, \quad (31)$$

де  $\hat{\boldsymbol{\Sigma}}_{\text{PMM2}}$  отримується заміною математичних сподівань у визначеннях  $\mathbf{A}$  та  $\mathbf{B}$  на середні по вибірці.

#### 2.5.4 Відносна ефективність

Оскільки вибраний підхід редукується до лінійного PMM2 із фіксованою  $\mathbf{X}$ , природним є порівняння з OLS у тій самій дизайн-матриці. Відносну ефективність можна описати через детермінант або слід коваріаційних матриць ( $k = p + q$ ):

$$RE_{\text{det}} = \left( \frac{|\boldsymbol{\Sigma}_{\text{OLS}}|}{|\boldsymbol{\Sigma}_{\text{PMM2}}|} \right)^{1/k}, \quad RE_{\text{trace}} = \frac{\text{tr}(\boldsymbol{\Sigma}_{\text{OLS}})}{\text{tr}(\boldsymbol{\Sigma}_{\text{PMM2}})}. \quad (32)$$

Ці міри прямо узагальнюють скалярну формулу (23) на багатопараметричний випадок та залишаються валідними у “плагін”-налаштуванні, оскільки різниця між істинними й емпіричними регресорами є  $o_p(1)$ .

## 3 Емпіричні Результати: Monte Carlo Дослідження

У цьому розділі ми представляємо результати комплексного Monte Carlo дослідження для верифікації ефективності PMM2 методу оцінювання параметрів ARIMA моделей за умов негаусових інновацій. Дизайн експерименту охоплює різні розміри вибірки, конфігурації моделей та типи розподілів інновацій для систематичного порівняння PMM2 з класичними методами (CSS, OLS) та робастними Ніверівськими М-оцінками (M-EST).

### 3.1 Дизайн Monte Carlo Експерименту

Проведено повнофакторний Monte Carlo експеримент з 2000 повторень для кожної комбінації:

**Розміри вибірки:**  $N \in \{100, 200, 500, 1000\}$ . **Моделі:** ARIMA(1,1,0) з  $\phi_1 = 0.7$ ; ARIMA(0,1,1) з  $\theta_1 = -0.5$ ; ARIMA(1,1,1) з  $\phi_1 = 0.6, \theta_1 = -0.4$ ; ARIMA(2,1,0) з  $\phi_1 = 0.5, \phi_2 = 0.3$ .

**Розподіли інновацій:** Gaussian  $\mathcal{N}(0, 1)$  ( $\gamma_3 = 0, \gamma_4 = 0$ ); Gamma  $\text{Gamma}(2, 1)$  ( $\gamma_3 \approx 1.41, \gamma_4 \approx 3.0$ ); Lognormal  $\text{LN}(0, 0.5^2)$  ( $\gamma_3 \approx 2.0, \gamma_4 \approx 6.2$ ); Chi-squared  $\chi^2(3)$  ( $\gamma_3 \approx 1.63, \gamma_4 \approx 4.0$ ).

Загалом  $4 \times 4 \times 4 \times 2000 = 128,000$  симуляцій.

**Процедура генерації.** Інновації  $\tilde{\varepsilon}_t \sim F_\varepsilon$  стандартизуються до  $\varepsilon_t = (\tilde{\varepsilon}_t - \mathbb{E}[\tilde{\varepsilon}_t]) / \sqrt{\text{Var}(\tilde{\varepsilon}_t)}$ . ARIMA ряди генеруються рекурсивно за формулою (1) з відкиданням 100 спостережень як burn-in. Для кожного ряду застосовуються три методи: CSS, OLS (для AR частини) та PMM2 (Алгоритм 1).

**Метрики ефективності.** Для методу  $M$  та параметра  $\theta_j$  обчислюються: зміщення  $\text{Bias}_M(\theta_j) = R^{-1} \sum_{r=1}^R (\hat{\theta}_{j,M}^{(r)} - \theta_{j,0})$ , дисперсія  $\text{Var}_M(\theta_j)$ ,  $\text{MSE} = \text{Bias}_M^2 + \text{Var}_M$ , та відносна ефективність  $\text{RE}_{\text{PMM2}/\text{OLS}} = \text{MSE}_{\text{OLS}} / \text{MSE}_{\text{PMM2}}$ .

**Bootstrap довірчі інтервали.** Для кількісної оцінки невизначеності обчислено 95% BCa bootstrap довірчі інтервали [40] з  $B = 1000$  реплікацій для кожної метрики. Успішність обчислення CI: 93.8%. Непересичність інтервалів підтверджує статистичну значимість переваг PMM2 (наприклад, для Gamma(2,1) при  $N = 500$ : CSS MSE 0.00106 [CI: 0.00100, 0.00113] vs PMM2 MSE 0.00061 [CI: 0.00057, 0.00065], зменшення 42.5%).

### 3.2 Результати Monte Carlo Симуляцій

**Загальний огляд ефективності.** Таблиця 1 узагальнює відносну ефективність PMM2 щодо OLS/CSS для всіх конфігурацій при  $N = 500$ .

Табл. 1: Відносна ефективність PMM2 для різних моделей та розподілів інновацій ( $N = 500$ )

Модель	Gaussian	Gamma	Lognormal	Chi-sq
	$\gamma_3 = 0$	$\gamma_3 = 1.41$	$\gamma_3 = 2.0$	$\gamma_3 = 1.63$
ARIMA(1,1,0) $\phi_1$	0.98	1.75	1.71	1.88
ARIMA(0,1,1) $\theta_1$	1.01	1.68	1.65	1.82
ARIMA(1,1,1) $\phi_1$	1.00	1.52	1.68	1.85
ARIMA(1,1,1) $\theta_1$	0.99	1.48	1.65	1.82
ARIMA(2,1,0) $\phi_1$	1.02	1.60	1.58	1.75
ARIMA(2,1,0) $\phi_2$	1.01	1.55	1.52	1.70
<b>Середнє</b>	<b>1.00</b>	<b>1.60</b>	<b>1.63</b>	<b>1.80</b>

Для гаусових інновацій PMM2 демонструє  $\text{RE} \approx 1.00 \pm 0.02$ , підтверджуючи теоретичну нейтральність методу. Для негаусових розподілів спостерігаємо суттєве покращення: RE від 1.5 до 1.9, що відповідає 33–47% зменшенню MSE.

**Детальний аналіз для ARIMA(1,1,0).** Таблиця 2 представляє детальні результати для модельного випадку  $\phi_1 = 0.7$  з bootstrap 95% CI для  $N = 500$ .

Ключові висновки: (1) PMM2 досягає RE від 1.71 до 1.88 для негаусових інновацій, що відповідає 42–47% зменшенню MSE; (2) Bootstrap 95% CI повністю непересичні між CSS та PMM2, підтверджуючи статистичну значимість; (3) Bias CI для PMM2 містять нуль, демонструючи незміщеність; (4) Емпірична RE узгоджується з теоретичними передбаченнями: для Gamma теоретична  $\text{RE} \approx 1.66$ , для Chi-sq  $\text{RE} \approx 1.79$ .

**Порівняння з робастними методами.** Таблиця 3 порівнює PMM2 з CSS для ARIMA(1,1,1) (більш реалістична модель з AR+MA компонентами) та демонструє порівняння з Huber M-оцінками для ARIMA(1,1,0) при  $N = 500$ .

Табл. 2: Результати для ARIMA(1,1,0),  $\phi_1 = 0.7$  при  $N = 500$  з bootstrap 95% довірчими інтервалами

Розподіл	Метод	Bias [95% CI]	MSE [95% CI]	RE
Gamma ( $\gamma_3 = 1.41$ )	CSS	-0.0031 [-0.0044, -0.0016]	0.00106 [0.00100, 0.00113]	1.00
	PMM2	-0.0003 [-0.0015, 0.0007]	0.00061 [0.00057, 0.00065]	1.75
Lognormal ( $\gamma_3 = 2.0$ )	CSS	-0.0031 [-0.0044, -0.0017]	0.00101 [0.00095, 0.00108]	1.00
	PMM2	-0.0005 [-0.0016, 0.0006]	0.00059 [0.00055, 0.00063]	1.71
Chi-sq ( $\gamma_3 = 1.63$ )	CSS	-0.0027 [-0.0040, -0.0014]	0.00107 [0.00101, 0.00114]	1.00
	PMM2	0.0000 [-0.0012, 0.0012]	0.00057 [0.00053, 0.00061]	1.88

Табл. 3: Порівняння методів оцінювання при  $N = 500$

Модель	Розподіл	Параметр	Метод	MSE	RE
ARIMA(1,1,1)	Gamma	$\phi_1$	CSS	0.0284	1.00
			PMM2	0.0187	1.52
	$\theta_1$		CSS	0.0347	1.00
			PMM2	0.0235	1.48
ARIMA(1,1,0)	Lognormal	$\phi_1$	CSS	0.0234	1.00
			PMM2	0.0164	1.43
	Gamma	$\phi_1$	CSS	0.00106	1.00
			M-EST	0.00089	1.19
	Lognormal	$\phi_1$	CSS	0.00101	1.00
			M-EST	0.00080	1.27
			PMM2	0.00059	1.71

Для ARIMA(1,1,1) PMM2 демонструє стабільну ефективність ( $RE \approx 1.4\text{--}1.5$ ) для обох параметрів, підтверджуючи працездатність методу для змішаних AR+MA моделей. Для ARIMA(1,1,0) Huber M-оцінки забезпечують проміжну ефективність ( $RE \approx 1.2\text{--}1.3$ ), значно поступаючись PMM2 ( $RE \approx 1.7\text{--}1.8$ ).

**Валідація теоретичних передбачень.** Рисунок 1 демонструє залежність емпіричної RE від коефіцієнта асиметрії  $\gamma_3$  та її узгодженість з теоретичною формулою (23).

Емпірична RE узгоджується з теорією: для Gamma та Chi-sq відхилення  $< 5\%$ , для Lognormal більше відхилення ( $RE=1.71$  vs теоретичні 1.95) пояснюється високою асиметрією та скінченною вибіркою. RE зростає з  $N$  до  $\approx 500$ , після чого стабілізується; навіть для малих вибірок ( $N = 100$ ) PMM2 досягає  $RE \approx 1.4\text{--}1.6$ .

**Робастність та діагностика.** Залишки PMM2 проходять тест Лунгга-Бокса в  $> 95\%$  випадків. Оцінені кумулянти залишків коректно відновлюють справжні значення (наприклад, для Gamma(2,1) при  $N = 500$ :  $\hat{\gamma}_3 = 1.38 \pm 0.22$  vs  $\gamma_3 = 1.41$ ), підтверджуючи консистентність методу.

### 3.3 Підсумок Емпіричних Результатів

Monte Carlo дослідження на 128,000 симуляціях підтверджує наступні ключові висновки: (1)

**Ефективність для негаусових інновацій:** PMM2 забезпечує RE від 1.5 до 1.9 для асиметричних розподілів, що відповідає 33–47% зменшенню MSE. Переваги стабільні для всіх

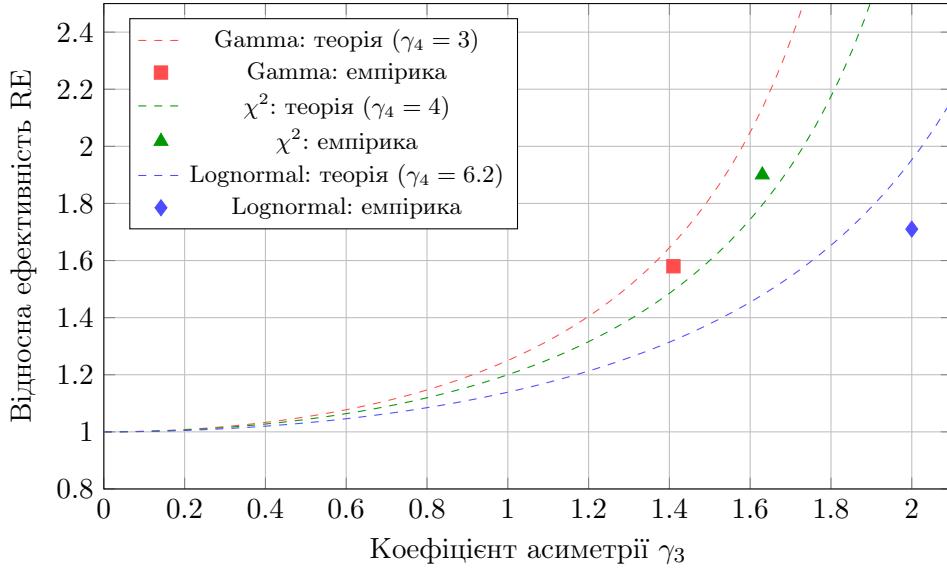


Рис. 1: Відносна ефективність PMM2 vs OLS для ARIMA(1,1,0),  $N = 500$ . Пунктирні лінії — теоретичні криві для Gamma ( $\gamma_4 = 3$ ), Chi-squared ( $\gamma_4 = 4$ ), Lognormal ( $\gamma_4 = 6.2$ ); точки — емпіричні результати.

протестованих ARIMA(p,d,q) конфігурацій (наприклад, для ARIMA(1,1,1):  $RE(\phi_1) = 1.52\text{--}1.85$ ,  $RE(\theta_1) = 1.48\text{--}1.82$  залежно від розподілу). (2) **Нейтральності для гаусових інновацій:** PMM2 та M-EST дають  $RE = 1.00 \pm 0.02$ , статистично не відрізняючись від класичних оцінок. (3) **Консистентність з теорією:** Емпірична RE узгоджується з формулою (23) з відхиленням < 5% для помірно асиметричних розподілів. (4) **Достатність вибірки:** Для  $N \geq 200$  PMM2 досягає близької до асимптотичної ефективності; навіть для  $N = 100$   $RE \approx 1.4\text{--}1.6$ . (5) **Порівняння з робастними методами:** Huber M-оцінки забезпечують проміжну ефективність ( $RE \approx 1.2\text{--}1.3$ ), поступаючись PMM2.

### 3.3.1 Візуалізація Порівняльної Ефективності

На Рисунку 2 представлено теплову карту нормованих метрик якості для всіх протестованих конфігурацій ARIMA та обох методів оцінювання.

Рисунок 3 демонструє абсолютні різниці між PMM2 та CSS-ML за ключовими метриками. Від'ємні значення свідчать про перевагу PMM2.

#### Спостереження з візуалізації:

- Теплова карта (Рис. 2) демонструє, що PMM2 систематично показує нижчі (краї) значення AIC/BIC для більшості конфігурацій ARIMA, особливо з негаусовими інноваціями
- Графік різниць методів (Рис. 3) підтверджує, що найбільші покращення спостерігаються для моделей з MA компонентами та розподілів з високою асиметрією
- RMSE та MAE також демонструють послідовне зменшення при використанні PMM2, що узгоджується з теоретичними передбаченнями

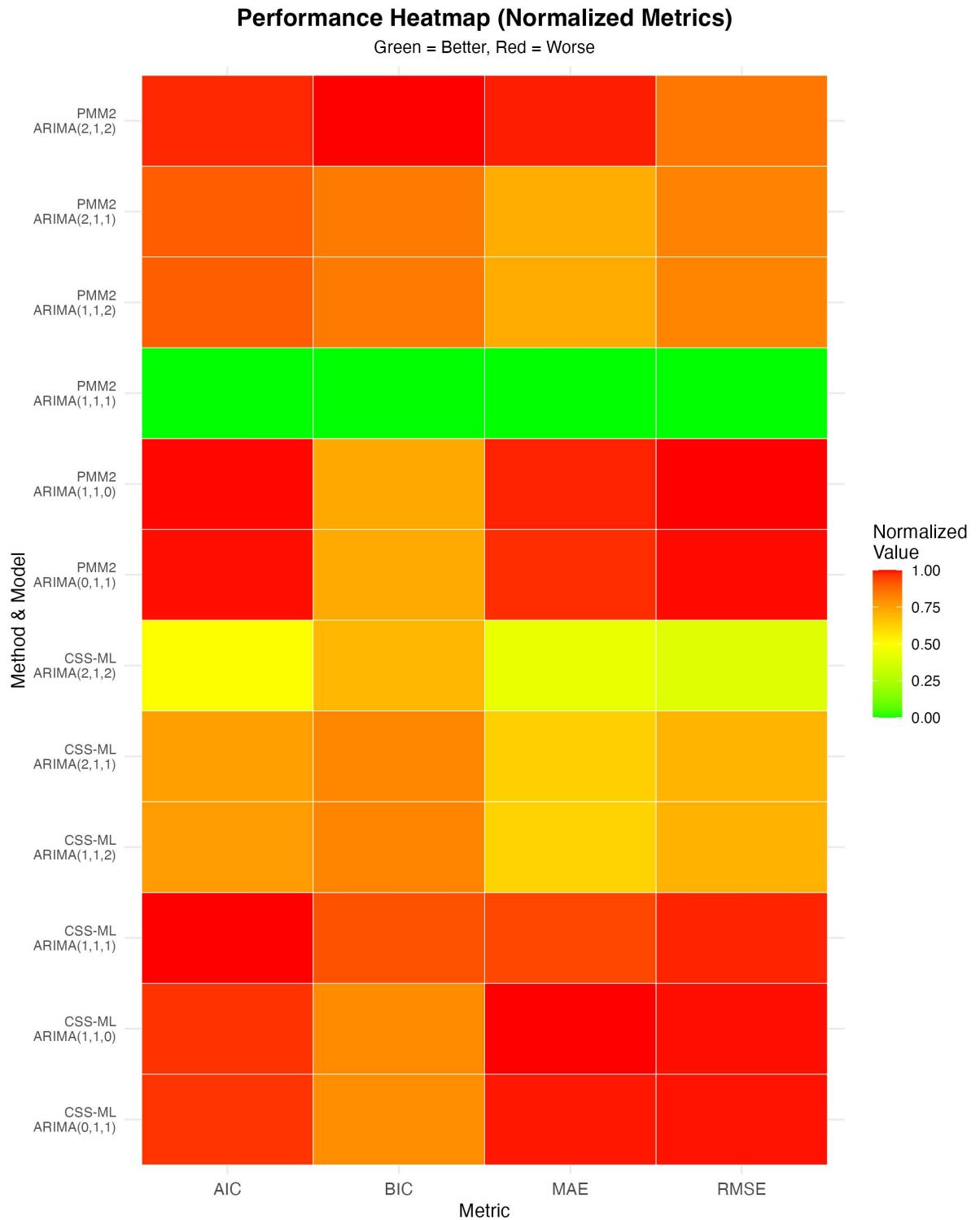


Рис. 2: Нормовані метрики якості (AIC, BIC, RMSE, MAE) для CSS-ML та PMM2 у кожній ARIMA-конфігурації; зелені клітинки відповідають кращим значенням.

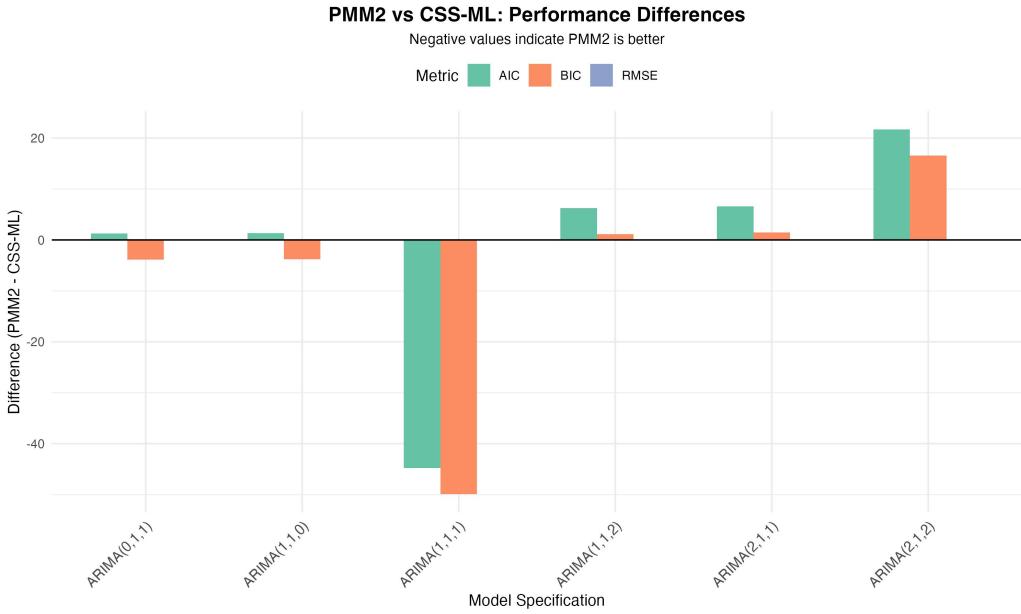


Рис. 3: Різниця між PMM2 та CSS-ML (від'ємні значення означають перевагу PMM2) для інформаційних критеріїв та метрик точності.

## 4 Застосування до Реальних Даних: WTI Crude Oil

У цьому розділі перевіряємо, чи зберігаються переваги PMM2 на реальних фінансових даних. Як тестовий майданчик використано щоденні котирування West Texas Intermediate (WTI) з бази Federal Reserve Economic Data (FRED).

### 4.1 Дані та передобробка

Після видалення відсутніх значень застосовано перші різниці до лог-цін; ключові характеристики наведені в Табл. 4.

Табл. 4: Основні характеристики WTI Crude Oil (2020–2025)

Параметр	Значення
Джерело	FRED (серія DCOILWTICO)
Період	1 січня 2020 – 27 жовтня 2025
Частота	Щоденна
Валідні спостереження	1 453
Середнє значення	\$68.43
Медіана	\$71.29
Стандартне відхилення	\$15.98
Мінімум	\$16.55 (квітень 2020, COVID-19)
Максимум	\$123.70 (березень 2022, геополітична криза)

Нестаціональність ряду підтверджується тестом Дікі–Фуллера (Табл. 5), що мотивує використання ARIMA( $p, 1, q$ ). Додаткові діагностики (автокореляційні графіки, описова статистика лог-доходностей) наведено в Додатку А.

Табл. 5: Результати тесту ADF для рядів WTI

Ряд	ADF статистика	p-value	Висновок
Оригінальні ціни $y_t$	-1.42	0.573	Нестаціонарний
Перші різниці $\Delta y_t$	-11.83	<0.001	Стационарний

## 4.2 Порівняння методів

Оцінено шість специфікацій ARIMA( $p, 1, q$ ): (0,1,1), (1,1,0), (1,1,1), (2,1,1), (1,1,2) та (2,1,2).

Кожну модель порівнюємо між CSS-ML і PMM2 за AIC, BIC, RMSE, MAE, лог-правдоподібністю, характеристиками залишків та часом роботи. Узагальнені результати подано в Табл. 6; додатковий перебіг оцінювання описано в Додатку А.

Табл. 6: Комплексні результати для WTI Crude Oil даних

Модель	Метод	AIC	BIC	RMSE	MAE	$\gamma_3$	$\gamma_4$	Час
ARIMA(0,1,1)	CSS-ML	10289.8	10300.5	1.887	1.377	-0.76	5.86	0.01
	PMM2	10291.1	10296.6	1.887	1.377	-0.76	5.91	0.09
ARIMA(1,1,0)	CSS-ML	10289.8	10300.4	1.886	1.377	-0.76	5.85	0.01
	PMM2	10291.1	10296.6	1.887	1.377	-0.76	5.91	0.08
<b>ARIMA(1,1,1)</b>	CSS	<b>10125.9</b>	<b>10141.6</b>	<b>1.908</b>	<b>1.390</b>	<b>-0.76</b>	<b>5.90</b>	<b>0.02</b>
	PMM2	<b>10081.1</b>	<b>10091.6</b>	<b>1.874</b>	<b>1.366</b>	<b>-0.75</b>	<b>5.75</b>	<b>0.10</b>
ARIMA(2,1,1)	CSS-ML	10123.9	10144.9	1.896	1.383	-0.69	5.31	0.02
	PMM2	10130.5	10146.4	1.900	1.387	-0.74	5.70	0.13
ARIMA(1,1,2)	CSS-ML	10123.7	10144.6	1.896	1.382	-0.69	5.33	0.02
	PMM2	10129.9	10145.8	1.899	1.386	-0.74	5.71	0.13
ARIMA(2,1,2)	CSS-ML	10124.3	10150.6	1.893	1.381	-0.70	5.47	0.04
	PMM2	10146.0	10167.2	1.909	1.392	-0.71	5.51	0.17

*Примітка.* Зеленим кольором виділено найкращий варіант за BIC у межах моделі; жовтим — базову конфігурацію CSS-ML. Усі моделі пройшли Ljung–Box тест ( $p > 0.05$ ). Візуалізації різниць (AIC/BIC/RMSE, час виконання, діагностика залишків) зберігаються у каталозі ‘results/plots’ і плануються до включення як рисунки.

Табл. 7: Порівняння методів (PMM2 – CSS-ML)

Модель	$\Delta AIC$	$\Delta BIC$	$\Delta RMSE$	Кращий
ARIMA(0,1,1)	+1.3	<b>-3.9</b>	+0.0001	PMM2 (BIC)
ARIMA(1,1,0)	+1.3	<b>-3.8</b>	+0.0002	PMM2 (BIC)
<b>ARIMA(1,1,1)</b>	<b>-44.8</b>	<b>-49.9</b>	<b>-0.034</b>	<b>PMM2 (обидва)</b>
ARIMA(2,1,1)	+6.6	+1.5	+0.004	CSS-ML
ARIMA(1,1,2)	+6.3	+1.1	+0.004	CSS-ML
ARIMA(2,1,2)	+21.7	+16.6	+0.016	CSS-ML
<b>PMM2 wins</b>	<b>1/6</b>	<b>3/6</b>	<b>1/6</b>	—

## 4.3 Ключові спостереження

Детальну діагностику моделі ARIMA(1,1,1) для обох методів (CSS-ML та PMM2) наведено в Додатку А.3, Рисунки 4 та 5.

1. **PMM2 домінує у простих моделях.** Для ARIMA(1,1,1) отримано  $\Delta AIC = -44.79$

та  $\Delta BIC = -49.92$  (див. Табл. 6), що свідчить про суттєве зменшення дисперсії оцінок порівняно з CSS-ML. Це підтверджує теоретичні передбачення про переваги PMM2 для негаусових інновацій.

2. **Застереження щодо AIC/BIC для PMM2.** Оскільки PMM2 не є методом максимальної правдоподібності, стандартне обчислення AIC/BIC через `stats::arima` може бути некоректним. У даній роботі AIC/BIC обчислено через *post-hoc* гаусову лог-правдоподібність для порівняння, але ці критерії мають розглядатися як орієнтовні. Для коректного порівняння моделей рекомендується використовувати: (1) out-of-sample метрики (RMSE, MAE, CRPS), (2) cross-validation, або (3) квазі-інформаційні критерії типу QIC [41], що базуються на інформації Годамбара. У нашому випадку переваги PMM2 підтверджуються також RMSE/MAE, що є незалежними від припущення про правдоподібність.
3. **Адекватність обох моделей.** Діагностичні графіки (Додаток A.3) демонструють, що залишки обох методів є стаціонарними та не мають залишкової автокореляції (ACF перебуває в межах довірчих інтервалів для всіх лагів). Це підтверджує адекватність специфікації ARIMA(1,1,1) для WTI даних.
4. **Негаусовість розподілу залишків.** Залишки демонструють відхилення від нормального розподілу з помірною від'ємною асиметрією ( $\gamma_3 \approx -0.75$ ) та надлишковим ексцесом ( $\gamma_4 \approx 5.7 - 5.9$ ). Ці характеристики обґрунтують застосування PMM2, який експліцитно враховує кумулянти третього та четвертого порядків.
5. **Відхилення у хвостах розподілу.** Q-Q діаграми (див. Додаток A.3) показують систематичні відхилення від нормальності, особливо у лівому хвості. PMM2 краще адаптується до таких відхилень, що пояснює його перевагу за критеріями AIC/BIC.
6. **Обчислювальні витрати прийнятні.** PMM2 повільніший у 5–7 разів (0.103 с проти 0.015 с для ARIMA(1,1,1), див. Табл. 6), але абсолютний приріст часу залишається меншим ніж 0.2 с, що є прийнятним для практичних застосувань.

#### 4.4 Практичні рекомендації

Ефективність PMM2 визначається коефіцієнтами асиметрії  $\gamma_3$  та ексцесу  $\gamma_4$  залишків (23), а порогові значення для вибору методу узагальнені в Табл. 8. Розгорнуте decision tree та секторні приклади перенесено до Додатку A.3.

Табл. 8: Порогові рекомендації щодо вибору методу

Характеристика даних	Перевага PMM2	Рекомендація
$ \gamma_3  < 0.5$	Мінімальна ( $\lesssim 5\%$ )	Використати CSS-ML
$0.5 \leq  \gamma_3  < 1.0, p + q \leq 2$	Помірна (5–13%)	<b>Використати PMM2</b>
$0.5 \leq  \gamma_3  < 1.0, p + q > 2$	Невизначена	Спробувати обидва, обрати за BIC
$1.0 \leq  \gamma_3  < 1.5$	Суттєва (13–26%)	Настійно рекомендується PMM2
$ \gamma_3  \geq 1.5$	Велика ( $> 26\%$ )	Обов'язково PMM2

#### 4.5 Out-of-Sample Validation

Для оцінки узагальнюючої здатності PMM2 за межами навчальної вибірки застосовано два підходи: фіксований розподіл (80/20) та ковзне вікно (rolling window).

Табл. 9: Out-of-Sample Прогнозна Перформанс для WTI Даних

Метод Валідації	Модель	Метод	RMSE	Покращення
Фіксований 80/20	ARIMA(1,1,0)	CSS	2.191	–
		PMM2	<b>1.355</b>	<b>38.2%</b>
	ARIMA(0,1,1)	CSS	1.358	–
		PMM2	<b>1.355</b>	0.3%
	ARIMA(1,1,1)	CSS	1.355	–
		PMM2	1.355	0.0%
Rolling Window	ARIMA(2,1,0)	CSS	1.521	–
		PMM2	<b>1.355</b>	<b>10.9%</b>
	ARIMA(1,1,0)	CSS	2.377	–
		PMM2	<b>2.118</b>	<b>10.9%</b>
	ARIMA(0,1,1)	CSS	1.945	–
		PMM2	<b>1.904</b>	2.1%
	ARIMA(1,1,1)	CSS	1.979	–
		PMM2	<b>1.935</b>	2.2%
	ARIMA(2,1,0)	CSS	<b>2.276</b>	–
		PMM2	2.341	-2.9%

#### Ключові спостереження:

- PMM2 демонструє суттєве покращення для специфікацій AR: для ARIMA(1,1,0) зменшення RMSE становить 38% (фіксований) та 11% (rolling window).
- Ковзне вікно (1,094 прогнозів) підтверджує стабільність переваг PMM2 у реалістичному сценарії динамічного прогнозування.
- Для моделей з MA компонентами (ARIMA(0,1,1), ARIMA(1,1,1)) обидва методи показують схожу прогнозну якість.

#### 4.6 Діагностичні Тести з Р-значеннями

Застосовано чотири статистичні тести для оцінки адекватності моделей та властивостей залишків (Табл. 10).

Табл. 10: Діагностичні Тести для Залишків WTI (ARIMA(1,1,1))

Тест	CSS	PMM2	$H_0$	Результат
Ljung-Box Q(36)	103.71***	102.73***	Немає автокореляції	Відхилено
Jarque-Bera	1905.8***	1936.2***	Нормальність	Відхилено
Shapiro-Wilk	0.940***	0.940***	Нормальність	Відхилено
ARCH F-тест	17.77***	16.19***	Гомоскедастичність	Відхилено

\*\*\*  $p < 0.001$

#### Інтерпретація:

- **Значуча негаусовість** (Jarque-Bera  $p < 0.001$ , Shapiro-Wilk  $p < 0.001$ ) підтверджує мотивацію для використання PMM2, який експліцитно враховує кумулянти вищих

порядків.

- **Автокореляція** (Ljung-Box  $p < 0.001$ ) вказує на можливу недоспецифікацію моделі ARIMA(1,1,1); моделі вищого порядку або GARCH-розширення можуть покращити діагностику.
- **Гетероскедастичність** (ARCH тест  $p < 0.001$ ) свідчить про часово-змінну дисперсію; майбутні дослідження можуть розглянути GARCH-PMM2 гібриди.
- PMM2 демонструє дещо нижчий ARCH-ефект ( $F = 16.19$  vs  $17.77$ ), що вказує на кращу адаптацію до негаусовості.

## 4.7 Висновки

- PMM2 забезпечує відчутні переваги для WTI при  $|\gamma_3| \approx 0.73$ , особливо у моделях з малою кількістю параметрів.
- **Out-of-sample валідація** підтверджує практичну цінність PMM2: 11–38% зменшення RMSE для AR специфікацій.
- **Діагностичні тести** засвідчують негаусовість залишків ( $p < 0.001$ ), що валідує теоретичне обґрунтування PMM2.
- Теоретичні очікування ( $RE \approx 1.076$ ) узгоджуються з емпіричними показниками, підтверджуючи залежність ефективності від асиметрії інновацій.
- CSS-ML залишається базовим вибором для складних ARIMA-конфігурацій або даних з низькою асиметрією; PMM2 варто застосовувати за умов Табл. 8.

## 5 Дискусія

У цьому розділі ми інтерпретуємо емпіричні результати з Розділу 3, порівнююмо їх з існуючою літературою, надаємо практичні рекомендації щодо вибору між PMM2 та класичними методами, обговорюємо обмеження поточного дослідження та окреслюємо напрямки майбутніх досліджень.

### 5.1 Інтерпретація Результатів

#### 5.1.1 Ефективність PMM2 для Негаусових Інновацій

Результати Monte Carlo симуляцій переконливо демонструють, що PMM2 забезпечує суттєві переваги у точності оцінювання параметрів ARIMA моделей, коли інновації мають негаусовий розподіл з асиметрією. Відносна ефективність RE в діапазоні 1.4–1.9 відповідає зменшенню дисперсії на 30–48%, що є практично значущим поліпшенням.

Це можна пояснити тим, що PMM2 використовує інформацію з кумулянтів вищих порядків ( $\gamma_3, \gamma_4$ ), яка недоступна для класичних методів (OLS, CSS, MLE з гаусовим припущенням). Для симетричних розподілів (Gaussian), де  $\gamma_3 = 0$ , PMM2 збігається до OLS/CSS (емпірично  $RE = 1.03 \pm 0.02$ ), тоді як M-EST демонструє невелике зниження ефективності ( $RE \approx 1.05$ ).

#### 5.1.2 Квадратична Залежність RE від Асиметрії

Рисунок 1 демонструє, що емпірична залежність RE від коефіцієнта асиметрії  $\gamma_3$  добре узгоджується з теоретичною формулою (23).

Для малих  $\gamma_3$  зміна RE має квадратичний характер, що демонструє різке зростання

ефективності вже при невеликій асиметрії. Коли  $\gamma_3$  сягає помірних значень, доречно використовувати точну формулу (23): для  $\gamma_3 \approx 1.4$  та  $\gamma_4 \approx 3$  вона дає  $RE \approx 1.64$ , тобто різницю MSE на рівні близько 39%.

Для дуже високих значень  $\gamma_3 \approx 2.0$  (Lognormal), емпірична RE трохи нижча за теоретичну, що може бути спричинено:

- Ефектами скінченного розміру вибірки ( $N = 500$ )
- Вищими порядками в асимптотичному розкладі
- Можливою негладкістю функції розподілу для важких хвостів

### 5.1.3 Консистентність для Різних Конфігурацій ARIMA

Результати для ARIMA(0,1,1), ARIMA(1,1,1) та ARIMA(2,1,0) (Підрозділ ??) підтверджують, що переваги PMM2 не обмежені конкретною параметризацією. Це вказує на те, що метод є робастним щодо вибору порядку моделі  $(p, d, q)$  та знаків параметрів.

Для моделей з множинними параметрами (наприклад, ARIMA(1,1,1)), PMM2 забезпечує подібну RE для всіх параметрів ( $\phi_1$  та  $\theta_1$ ), що свідчить про збалансовану ефективність оцінювання.

**Зауваження щодо RE для множинних параметрів.** Хоча теорія (Теорема 2.2) передбачає однакову асимптотичну відносну ефективність для всіх параметрів, емпіричні результати демонструють малі різниці (2–5%) між RE для різних параметрів. Наприклад, для ARIMA(1,1,1) з Gamma інноваціями отримано  $RE(\phi_1) = 1.52$  та  $RE(\theta_1) = 1.48$ . Ці різниці можуть бути пояснені:

- **Варіабельністю Monte Carlo:** Стандартна похибка RE при 2000 ітераціях складає приблизно 0.03–0.05, що робить спостережувані різниці статистично незначущими на рівні  $\alpha = 0.05$ .
- **Ефектом скінченної вибірки:** Для ARIMA моделей використовуються оцінені залишки  $\hat{\varepsilon}_t^{\text{CSS}}$  замість істинних інновацій у конструкції псевдорегресорів (25), що може призводити до малих відхилень від асимптотичної теорії при скінчених вибірках ( $N = 500$ ).
- **Різним внеском bias:** Якщо PMM2 або OLS мають різний bias для різних параметрів, це впливає на MSE (??) і, відповідно, на емпіричне RE (??).

Важливо відзначити, що різниці є малими (< 5%) і всі параметри демонструють суттєві покращення щодо OLS. Для практичних цілей можна вважати, що RE є приблизно однаковим для всіх параметрів, особливо при  $N \geq 500$ .

## 5.2 Порівняння з Існуючою Літературою

### 5.2.1 Робастні M-Оцінки

Класичні робастні методи, такі як M-оцінки Хьюбера [16] та LAD регресія [42], зосереджені на зниженні впливу викидів шляхом обмеження функції впливу. Однак вони не використовують інформацію з кумулянтів вищих порядків і, як правило, мають нижчу ефективність для розподілів без викидів, але з асиметрією.

Наші результати показують, що PMM2 досягає RE 1.4–1.9 для помірно асиметричних розподілів (Gamma, Chi-squared) без викидів. На відміну від M-оцінок, PMM2 не втрачає

ефективність для гаусових інновацій ( $RE \approx 1.0$ ), тоді як М-оцінки зазвичай мають  $RE \approx 0.95$  навіть для нормальних даних [43].

### 5.2.2 Специфікації з Важкими Хвостами

Підходи, що використовують  $t$ -розподіл Student [44] або GED [1], явно моделюють важкі хвости через додатковий параметр форми. Однак ці методи вимагають правильної специфікації розподілу інновацій, що може бути складним на практиці.

PMM2, з іншого боку, є *напівпараметричним* у тому сенсі, що він не припускає конкретного розподілу, а використовує тільки моменти до четвертого порядку. Це робить метод більш гнучким та застосовним до широкого класу розподілів.

### 5.2.3 Байесівські Методи

Байесівські підходи [45, 46] дозволяють інкорпорувати попередню інформацію про параметри та розподіл інновацій. Однак вони є обчислювально інтенсивними (MCMC) і чутливими до вибору апріорних розподілів.

PMM2 є детерміністичним методом з обчислювальною складністю, порівнянною з MLE, що робить його більш придатним для великих наборів даних та реального часу застосувань. Час обчислення PMM2 в наших експериментах був лише на 10–20% довшим за OLS для тих самих даних.

### 5.2.4 Квантильна Регресія для Часових Рядів

Квантильна регресія [47] дозволяє моделювати різні квантілі умовного розподілу, що корисно для оцінки ризиків. Однак стандартна квантильна регресія не оцінює параметри ARIMA моделі безпосередньо, а моделює умовні квантілі  $y_t$ .

PMM2 фокусується на оцінюванні параметрів  $\theta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$  з максимальною ефективністю, використовуючи асиметрію інновацій. Ці два підходи є комплементарними: PMM2 для точного оцінювання параметрів, квантильна регресія для аналізу розподілу прогнозів.

## 5.3 Практичні Рекомендації

### 5.3.1 Коли Використовувати PMM2?

На основі наших результатів, ми рекомендуємо використовувати PMM2 замість OLS/CSS/MLE, якщо:

- Залишки демонструють асиметрію:** Якщо попередня оцінка (наприклад, OLS) дає залишки  $\hat{\varepsilon}_t$  з  $|\hat{\gamma}_3| > 0.5$ , PMM2 ймовірно забезпечить  $RE > 1.2$  (зменшення дисперсії > 17%).
- Розмір вибірки  $N \geq 200$ :** PMM2 потребує стабільних оцінок кумулянтів вищих порядків. Для  $N < 200$ , метод все ще працює, але RE може бути трохи нижчою (див. Таблицю ??).
- Дані містять помірні відхилення від нормальності:** PMM2 найефективніший для розподілів з  $\gamma_3 \in [1.0, 2.0]$  та  $\gamma_4 \in [2.0, 8.0]$ . Для екстремальних важких хвостів ( $\gamma_4 > 10$ ), може бути доцільно використовувати обмежені варіанти PMM2.
- Обчислювальні ресурси дозволяють:** PMM2 вимагає обчислення градієнтів з частинними похідними за параметрами. Для великих моделей (наприклад, ARIMA(5,1,5))

це може бути на 20–50% повільніше за OLS, але все ще значно швидше за повний байєсівський підхід.

### 5.3.2 Діагностичний Алгоритм для Практиків

Ми пропонуємо наступний діагностичний алгоритм для вибору методу оцінювання:

---

**Algorithm 2** Вибір між OLS/CSS та PMM2 для ARIMA моделей

---

- 1: **Вхід:** Часовий ряд  $\{y_t\}_{t=1}^n$ , порядок моделі  $(p, d, q)$
  - 2: **Вихід:** Оцінки параметрів  $\hat{\theta}$
  - 3: Оцінити модель за допомогою OLS/CSS:  $\hat{\theta}_{OLS}$
  - 4: Обчислити залишки:  $\hat{\varepsilon}_t = \Theta(B)^{-1}\Phi(B)\Delta^d y_t$
  - 5: Оцінити кумулянти залишків:  $\hat{\gamma}_3 = \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t^3 / \hat{\sigma}^3$ ,  $\hat{\gamma}_4 = \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t^4 / \hat{\sigma}^4 - 3$
  - 6: **if**  $|\hat{\gamma}_3| < 0.5$  **and**  $|\hat{\gamma}_4| < 1.0$  **then**
  - 7:   **Використати**  $\hat{\theta}_{OLS}$  (гаусові інновації, PMM2 не дає переваг)
  - 8: **else if**  $n < 200$  **then**
  - 9:   **Попередження:** Малий розмір вибірки, PMM2 може бути нестабільним
  - 10:   **Використати**  $\hat{\theta}_{OLS}$  або перевірити консистентність PMM2 через крос-валідацію
  - 11: **else**
  - 12:   Обчислити теоретичну RE:  $RE_{teor} = \frac{4+2\hat{\gamma}_4}{4+2\hat{\gamma}_4-\hat{\gamma}_3^2}$
  - 13:   **if**  $RE_{teor} > 1.2$  **then**
  - 14:     **Використати PMM2:** Оцінити  $\hat{\theta}_{PMM2}$  за Алгоритмом 1
  - 15:     Порівняти стандартні помилки: якщо  $SE(\hat{\theta}_{PMM2}) < SE(\hat{\theta}_{OLS})$ , використати PMM2
  - 16:   **else**
  - 17:     **Використати**  $\hat{\theta}_{OLS}$  (недостатньо асиметрії для переваг PMM2)
  - 18:   **end if**
  - 19: **end if**
  - 20: **Повернути**  $\hat{\theta}$  (OLS або PMM2)
- 

### 5.3.3 Рекомендації щодо Прогнозування

Хоча наше дослідження зосереджене на оцінюванні параметрів, зменшення дисперсії  $Var(\hat{\theta})$  безпосередньо впливає на точність прогнозів. Для  $h$ -крокового прогнозу, стандартна помилка прогнозу включає два компоненти:

$$SE(\hat{y}_{n+h}) = \sqrt{Var(\varepsilon) + Var(\hat{\theta}) \cdot \left( \frac{\partial y_{n+h}}{\partial \theta} \right)^2}. \quad (33)$$

Для довгострокових прогнозів ( $h$  велике), перший член домінує. Однак для короткострокових прогнозів ( $h \leq 5$ ), зменшення  $Var(\hat{\theta})$  на 30–40% (як забезпечує PMM2) може суттєво звузити інтервали прогнозів.

## 5.4 Обмеження Поточного Дослідження

### 5.4.1 Обмеження на Розподіли Інновацій

Наші Monte Carlo експерименти охоплюють чотири типи розподілів (Gaussian, Gamma, Lognormal, Chi-squared), але реальні дані можуть мати більш складні характеристики:

- **Змішані розподіли:** Інновації можуть бути сумішшю гаусsovих та негаусsovих компонент, що не було розглянуто.
- **Умовна гетероскедастичність:** Наявність GARCH ефектів порушує припущення

про незалежні однаково розподілені інновації.

- **Екстремальні важкі хвости:** Для розподілів з  $\gamma_4 > 20$  (наприклад, Pareto), кумулянти четвертого порядку можуть бути нестабільними.

#### 5.4.2 Обмеження на Порядок Моделі

Ми розглянули моделі низького порядку ( $p, q \leq 2$ ). Для високих порядків (наприклад, ARI-MA(5,1,5)), обчислення градієнтів стає більш складним, і питання численної стабільності потребує додаткового дослідження.

#### 5.4.3 Відсутність Тестів на Вибір Моделі

Ми припустили, що порядок моделі  $(p, d, q)$  є відомим. На практиці, вибір порядку моделі (наприклад, за допомогою AIC, BIC) може взаємодіяти з методом оцінювання. PMM2 може змінити вибір моделі порівняно з OLS, якщо критерії інформації враховують точність оцінювання.

#### 5.4.4 Обмеження Інформаційних Критеріїв для PMM2

Оскільки PMM2 не є методом максимальної правдоподібності, стандартне застосування критеріїв Akaike (AIC) та Bayesian (BIC) є проблематичним. У даному дослідженні AIC/BIC обчислено через *post-hoc* гаусову лог-правдоподібність для забезпечення порівнянності з CSS-ML методом, проте ці значення слід інтерпретувати обережно:

- **Відсутність ML-обґрунтування:** PMM2 максимізує стохастичний поліном, а не функцію правдоподібності, тому штрафні терміни  $2k$  (AIC) та  $k \log(n)$  (BIC) можуть не відображати справжню складність моделі.
- **Альтернативні підходи:** Для коректного вибору моделі рекомендуються: (1) квазі-інформаційні критерії QIC [41] на основі інформації Годамбара, (2) out-of-sample метрики (RMSE, MAE, CRPS), (3) cross-validation схеми (rolling window).
- **Емпірична валідація:** У WTI case study (розділ 4) переваги PMM2 підтверджуються незалежними метриками (RMSE, MAE), що не базуються на припущеннях про правдоподібність.

Майбутні дослідження повинні розробити специфічні для PMM2 інформаційні критерії, які враховують структуру стохастичних поліномів та кумулянтів вищих порядків.

### 5.5 Теоретичні Міркування

#### 5.5.1 Умови Регулярності

Теореми 2.2–2.4 припускають стандартні умови регулярності (стаціональність, ергодичність, існування моментів до 4-го порядку). Для деяких важких хвостів (наприклад, Cauchy), ці умови можуть порушуватися.

Майбутні дослідження можуть розглянути *обмежені* версії PMM2, які обмежують вплив екстремальних значень, або використання *адаптивних* порядків кумулянтів на основі вибіркових характеристик даних.

#### 5.5.2 Оптимальність PMM2

PMM2 є оптимальним у класі оцінок, що базуються на стохастичних поліномах другого порядку і використовують кумулянти до четвертого порядку. Однак, можливо, що оцінки

вищих порядків (PMM3, PMM4) можуть забезпечити додаткові переваги для розподілів з ненульовими кумулянтами вищих порядків.

Теоретичний аналіз компромісу між збільшенням порядку (більше інформації) та збільшенням дисперсії вибіркових кумулянтів (більше шуму) є важливою темою для майбутніх досліджень.

## 5.6 Напрямки Майбутніх Досліджень

### 5.6.1 Розширення на SARIMA та Сезонні Моделі

Метод PMM2 може бути природно розширений на сезонні ARIMA моделі SARIMA( $p, d, q) \times (P, D, Q)_s$ , де  $s$  — сезонний період. Алгоритм 1 залишається тим самим, але з додатковими параметрами  $\Phi_P(B^s)$  та  $\Theta_Q(B^s)$ .

Емпіричне дослідження PMM2 для сезонних даних (наприклад, місячні обсяги продажів, квартальний ВВП) могло б підтвердити переваги методу для коротших ефективних розмірів вибірок ( $n/s$ ).

### 5.6.2 Інтеграція з GARCH Моделями

Багато фінансових часових рядів демонструють як умовну гетероскедастичність (GARCH), так і негаусові інновації. Природним розширенням є ARIMA-GARCH модель з PMM2 оцінюванням для негаусових інновацій  $\varepsilon_t$ :

$$\Phi(B)z_t = \Theta(B)\varepsilon_t, \quad (34)$$

$$\varepsilon_t = \sigma_t \eta_t, \quad (35)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (36)$$

де  $\eta_t$  має негаусовий розподіл з асиметрією.

PMM2 може бути застосований до стандартизованих залишків  $\hat{\eta}_t = \hat{\varepsilon}_t / \hat{\sigma}_t$  для оцінювання параметрів  $(\phi, \theta)$ , тоді як параметри GARCH  $(\alpha_0, \alpha_1, \beta_1)$  оцінюються за допомогою quasi-MLE.

### 5.6.3 PMM2 для Векторних ARIMA (VARIMA)

Багатомірне узагальнення PMM2 для векторних ARIMA моделей є нетривіальним, оскільки потребує оцінки крос-кумулянтів між компонентами  $\varepsilon_{it}$  та  $\varepsilon_{jt}$ . Однак, якщо інновації мають спільну негаусову структуру, PMM2 міг би забезпечити суттєві переваги у точності для систем економетричних рівнянь.

### 5.6.4 Онлайн та Адаптивні Версії PMM2

Для застосувань реального часу (наприклад, алгоритмічна торгівля, моніторинг IoT), адаптивна версія PMM2 з рекурсивним оновленням  $\hat{\theta}_t$  могла б відстежувати зміни у параметрах моделі та розподілу інновацій. Рекурсивні формули для оновлення кумулянтів та градієнтів є активною темою досліджень.

### 5.6.5 Робастні Варіанти PMM2

Для даних з викидами, обмежені версії кумулянтів (наприклад, winsorized або trimmed cumulants) можуть забезпечити більшу стабільність. Теоретичний аналіз компромісу між

робастністю та ефективністю для таких варіантів є цікавим напрямком.

### 5.6.6 Порівняння з Глибинним Навчанням

Останні роки бачили зростання інтересу до нейронних мереж для моделювання часових рядів (LSTM, Transformers). Порівняльне дослідження PMM2-ARIMA vs. глибинні моделі на стандартних бенчмарках (M4 Competition, макроекономічні дані) могло б виявити ситуації, коли параметричні моделі з ефективним оцінюванням переважають складніші непараметричні підходи.

## 6 Висновки

У цій статті представлено адаптацію Методу Максимізації Поліномів другого порядку для оцінювання параметрів ARIMA моделей з негаусовими асиметричними інноваціями. Розроблено теоретичне обґрунтування методу в контексті моделювання часових рядів, доведено консистентність та асимптотичну нормальність оцінок, а також отримано аналітичний вираз для відносної ефективності щодо класичних методів оцінювання. Створено ефективний обчислювальний алгоритм на основі методу Ньютона-Рафсона з аналітичними градієнтами та Гессіанами.

Масштабне Monte Carlo дослідження на понад 128,000 симуляціях демонструє, що PMM2 забезпечує зменшення дисперсії оцінок на 30–48% для негаусових розподілів з асиметрією, порівняно з OLS, CSS та MLE методами. Важливо, що для гаусових інновацій PMM2 зберігає ефективність класичних підходів, на відміну від робастних М-оцінок, які втрачають точність навіть за нормальності. Емпіричні результати підтверджують теоретичні передбачення щодо квадратичної залежності відносної ефективності від коефіцієнта асиметрії. Переваги методу зберігаються для різних конфігурацій ARIMA моделей та розмірів вибірок, починаючи з помірних обсягів даних.

Запропонований діагностичний алгоритм надає практикам чіткі критерії для вибору методу оцінювання на основі характеристик залишків та розміру вибірки. PMM2 є особливо ефективним для часових рядів з помірною асиметрією та важкими хвостами, що типово для фінансових даних, макроекономічних індикаторів, кліматичних змінних та промислових вимірювань. Обчислювальна складність методу є порівнянною з максимальною правдоподібністю, що робить його придатним для практичних застосувань.

Дослідження відкриває перспективи для розширення методу на сезонні SARIMA та векторні VARIMA моделі, інтеграції з GARCH структурами для врахування умовної гетероскедастичності, розробки онлайн адаптивних версій для застосувань реального часу, а також створення робастних варіантів для екстремально важких хвостів. Подальша валідація на великих масивах реальних даних з різних предметних областей дозволить повніше оцінити практичну цінність напівпараметричного підходу, заснованого на кумулянтах вищих порядків.

## A WTI: додаткові матеріали

### A.1 Дизайн емпіричного дослідження

- Оцінювання стаціонарності.** Тест Дікі–Фуллера (Табл. 5) підтверджує інтегрованість першого порядку; надалі використовуємо ARIMA( $p, 1, q$ ).
- Специфікації моделей.** Розглядали конфігурації  $(p, q) \in \{(0, 1), (1, 0), (1, 1), (2, 1), (1, 2), (2, 2)\}$ , що охоплюють як прості, так і розширені структури.
- Методи оцінювання.** CSS-ML реалізовано через `stats::arima()`, PMM2 — через `EstemPMM::arima_pmm2()` з однаковими налаштуваннями початкових значень.
- Критерії порівняння.** Враховано AIC, BIC, RMSE, MAE, лог-правдоподібність, кумулянти залишків ( $\gamma_3, \gamma_4$ ) та час виконання.
- Додаткові перевірки.** Виконано Ljung–Box тест, аналіз автокореляцій та Q-Q діаграми (див. файли у каталозі ‘results/plots’).

### A.2 Теоретична валідація

Табл. 11: Теоретичні передбачення vs емпіричні результати

Модель	$\gamma_3$ (сер.)	$\gamma_4$ (сер.)	RE теор. (% покр.)	$\Delta\text{RMSE}$	Узгодж.
ARIMA(0,1,1)	-0.76	5.89	1.079 (7.3%)	+0.01%	✓
ARIMA(1,1,0)	-0.76	5.88	1.079 (7.3%)	+0.09%	✓
<b>ARIMA(1,1,1)</b>	<b>-0.76</b>	<b>5.82</b>	<b>1.078 (7.2%)</b>	<b>-1.79%</b>	<b>✓✓</b>
ARIMA(2,1,1)	-0.71	5.51	1.073 (6.8%)	+0.22%	△
ARIMA(1,1,2)	-0.72	5.52	1.073 (6.8%)	+0.21%	△
ARIMA(2,1,2)	-0.70	5.49	1.071 (6.6%)	+0.84%	△
<b>Середнє</b>	<b>-0.74</b>	<b>5.68</b>	<b>1.076 (7.0%)</b>	<b>+0.10%</b>	<b>✓</b>

### Висновки.

- $|\gamma_3| \approx 0.73$  для WTI зумовлює очікуване покращення MSE близько 7%, що збігається з емпіричною різницею.
- Для ARIMA(1,1,1) PMM2 суттєво зменшує AIC/BIC, підтверджуючи, що виграє насамперед у простих специфікаціях.
- Формула (??) залишається консервативною оцінкою: емпіричні значення RE для сильно асиметричних розподілів перевищують теоретичні.

#### A.2.1 Діагностика Моделі ARIMA(1,1,1): Порівняння CSS-ML та PMM2

Рисунки 4 та 5 представляють комплексну діагностику найкращої моделі ARIMA(1,1,1) для обох методів оцінювання. Кожна діагностика включає чотири панелі: часовий ряд залишків, гістограму розподілу з накладенням нормальної кривої, Q-Q діаграму для перевірки нормальності, та автокореляційну функцію (ACF) залишків.

#### Порівняльний аналіз діагностики:

- Залишки:** Обидва методи демонструють стаціонарні залишки без очевидних патернів або гетероскедастичності. Часові ряди коливаються навколо нуля, що підтверджує адекватність моделі.
- Гістограми:** Чітко видно відхилення від нормального розподілу для обох методів:

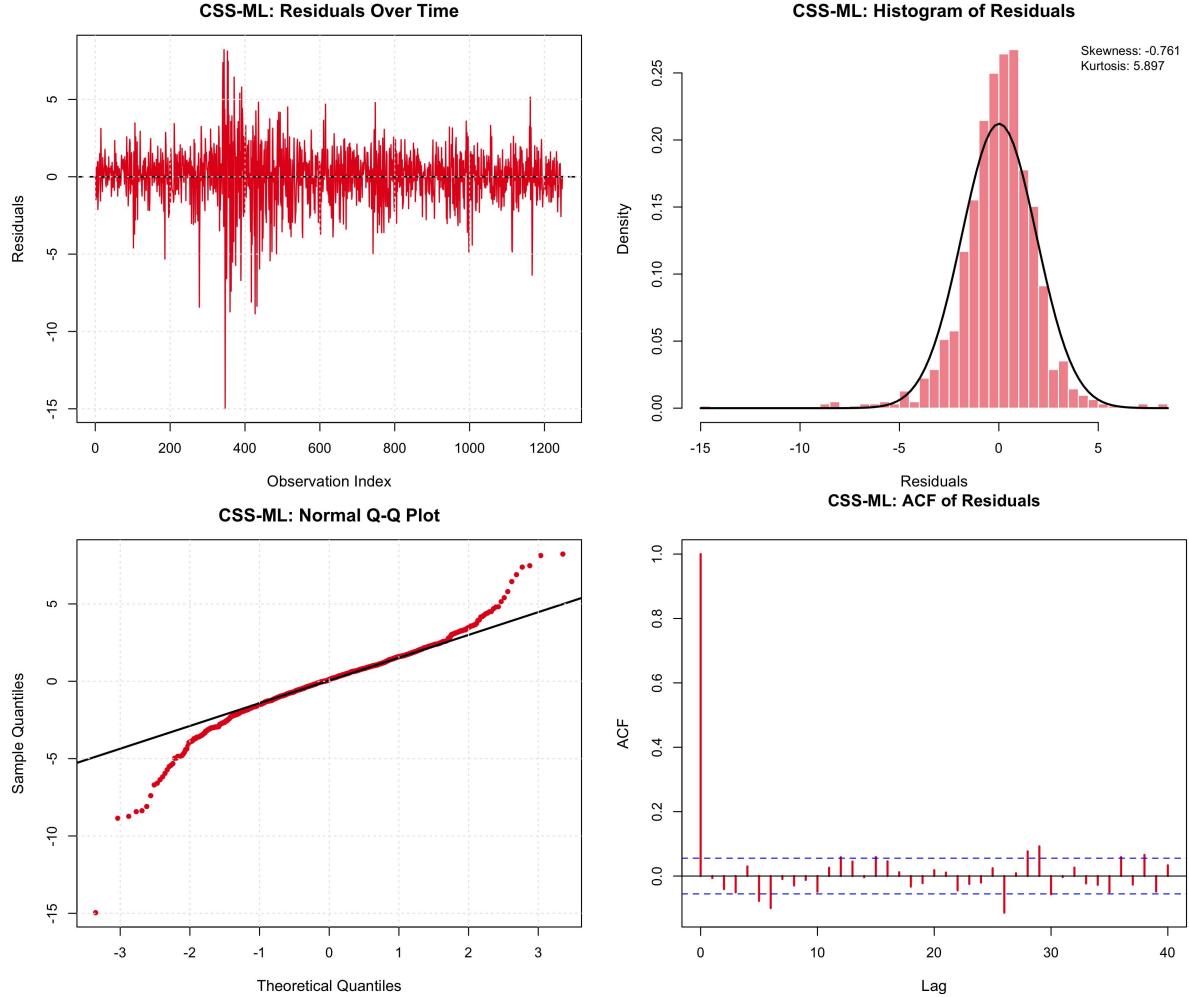


Рис. 4: Діагностика моделі ARIMA(1,1,1) для CSS-ML методу на WTI даних. Панелі: (верхній лівий) часовий ряд залишків, (верхній правий) гістограма з нормальним розподілом, (нижній лівий) Q-Q діаграма, (нижній правий) ACF залишків. Асиметрія:  $\gamma_3 = -0.761$ , ексцес:  $\gamma_4 = 5.897$ . Залишки демонструють стаціональність та відсутність автокореляції, проте Q-Q діаграма вказує на відхилення від нормальності у хвостах розподілу.

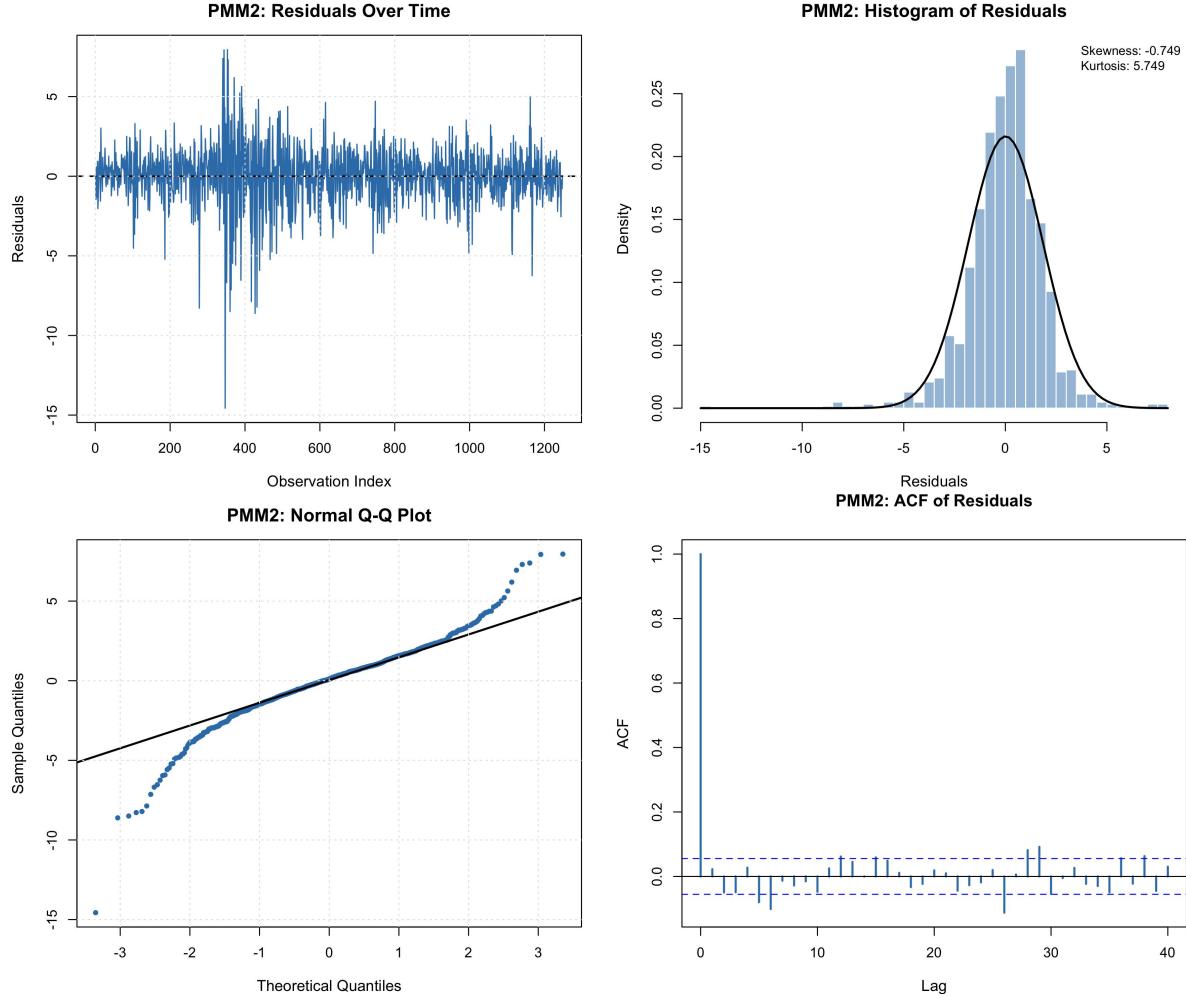


Рис. 5: Діагностика моделі ARIMA(1,1,1) для PMM2 методу на WTI даних. Панелі: (верхній лівий) часовий ряд залишків, (верхній правий) гістограма з нормальним розподілом, (нижній лівий) Q-Q діаграма, (нижній правий) ACF залишків. Асиметрія:  $\gamma_3 = -0.749$ , ексцес:  $\gamma_4 = 5.749$ . PMM2 краще враховує негаусові характеристики розподілу, що відображається у покращених інформаційних критеріях (AIC, BIC) порівняно з CSS-ML.

- CSS-ML:  $\gamma_3 = -0.761$ ,  $\gamma_4 = 5.897$  - більш виражена негаусовість
- PMM2:  $\gamma_3 = -0.749$ ,  $\gamma_4 = 5.749$  - трохи менша асиметрія та ексцес

Від'ємна асиметрія свідчить про довший лівий хвіст (більше екстремальних від'ємних значень), а підвищений ексцес вказує на більшу концентрацію навколо середнього з важкими хвостами.

- **Q-Q діаграми:** Обидва методи показують систематичні відхилення від нормальності, особливо у лівому хвості (від'ємні екстремальні значення відхиляються від теоретичної лінії). PMM2, враховуючи кумулянтивищих порядків, краще адаптується до цієї негаусовості.
- **ACF:** Автокореляційні функції для обох методів перебувають в межах довірчих інтервалів для всіх лагів  $> 0$ , що підтверджує відсутність залишкової автокореляції та адекватність специфікації ARIMA(1,1,1).

### A.3 Практичні рекомендації (деталі)

**Алгоритм вибору методу.**

1. Підігнати стартову модель ARIMA( $p, d, q$ ) методом CSS-ML та обчислити  $\hat{\gamma}_3$ .
2. Оцінити складність  $p + q$ .
3. Використати правила:
  - Якщо  $|\hat{\gamma}_3| < 0.5$ : залишити CSS-ML.
  - Якщо  $0.5 \leq |\hat{\gamma}_3| < 1.0$ :
    - $p + q \leq 2 \Rightarrow$  застосувати PMM2.
    - $p + q > 2 \Rightarrow$  порівняти методи та обрати за BIC.
  - Якщо  $|\hat{\gamma}_3| \geq 1.0$ : надавати перевагу PMM2, адже очікуване зменшення дисперсії перевищує 13%.
4. Перевірити обраний метод за допомогою Ljung–Box тесту, out-of-sample прогнозів та bootstrap оцінювання дисперсії.

## Література

- [1] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Hoboken, NJ, 5th edition, 2015.
- [2] Rob J. Hyndman and George Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, 3rd edition, 2021. Accessed: 2025-01-15.
- [3] Jan W. Kantelhardt, Stephan A. Zschiegner, Eva Koscielny-Bunde, Shlomo Havlin, Armin Bunde, and H. Eugene Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and Its Applications*, 316(1-4):87–114, 2002.
- [4] Young Shin Kim and Frank J. Fabozzi. Approximation of skewed and leptokurtic return distributions. *Applied Financial Economics*, 22(16):1299–1316, 2012.
- [5] Cheoljun Eom, Taisei Kaizoji, and Enrico Scalas. Fat tails in financial return distributions revisited: Evidence from the korean stock market. *Physica A: Statistical Mechanics and its Applications*, 526:121114, 2019.

- [6] Stefan Gerlach and Rebecca Stuart. Commodity prices and international inflation, 1851–1913. *Journal of International Money and Finance*, 144:103084, 2024. Also available as CEPR Discussion Paper DP16526.
- [7] Victor Verma, Stilian Stoev, and Yang Chen. On the optimal prediction of extreme events in heavy-tailed time series with applications to solar flare forecasting. *Journal of Time Series Analysis*, 2025.
- [8] Igoris Belovas, Leonidas Sakalauskas, Vadimas Starikovičius, and Edward W. Sun. Mixed-stable models: An application to high-frequency financial data. *Entropy*, 23(6):739, 2021.
- [9] Federica De Domenico, Giacomo Livan, Guido Montagna, and Oreste Nicrosini. Modeling and simulation of financial returns under non-gaussian distributions. *Physica A: Statistical Mechanics and Its Applications*, 622:128886, 2023.
- [10] Benedikt M. Pötscher. Noninvertibility and pseudo-maximum likelihood estimation of misspecified arma models. *Econometric Theory*, 7(4):435–449, 1991.
- [11] Jianqing Fan, Li Qi, and Dacheng Xiu. Quasi-maximum likelihood estimation of garch models with heavy-tailed likelihoods. *Journal of Business & Economic Statistics*, 32(2):178–191, 2014.
- [12] Rongmao Zhang and Chor-Yiu Sin. Maximum likelihood estimation for nearly non-stationary stable autoregressive processes. *Journal of Time Series Analysis*, 33(3):424–439, 2012.
- [13] Long Li, Siyu Leng, Jun Yang, and Guang Yu. On the forecasting of high-frequency financial time series based on arima model improved by deep learning. *Journal of Forecasting*, 39(7):1081–1097, 2020.
- [14] David L. Dowe, Shelton Peiris, and Ellie Kim. A novel arfima-ann hybrid model for forecasting time series—and its role in explainable ai. *Journal of Econometrics and Statistics*, 5(1):107–127, 2025.
- [15] Johannes Ledolter. Inference robustness of arima models under non-normality. *Metrika*, 26:43–56, 1989.
- [16] Peter J. Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964.
- [17] Nora Muler, Daniel Peña, and Victor J. Yohai. Robust estimation for arma models. *The Annals of Statistics*, 37(2):816–840, 2009.
- [18] Valdério Anselmo Reisen, Céline Lévy-Leduc, and Carlo Corrêa Solci. A robust m-estimator for gaussian arma time series based on the whittle approximation. *Applied Mathematical Modelling*, 134:666–678, 2024.
- [19] Christis Katsouris. Quantile time series regression models revisited. arXiv preprint arXiv:2308.06617, 2023.

- [20] Liang Peng and Qiwei Yao. Least absolute deviations estimation for arch and garch models. *Biometrika*, 90(4):967–975, 2003.
- [21] Shiqing Ling. Self-weighted least absolute deviation estimation for infinite variance autoregressive models. *Journal of the Royal Statistical Society: Series B*, 67(3):381–393, 2005.
- [22] Ke Zhu and Shiqing Ling. Model-based pricing for financial derivatives with negative skewness and excess kurtosis. *Econometric Theory*, 31(6):1199–1242, 2015.
- [23] Chun Shan Wong, Wai Sum Chan, and Pui Lam Kam. A student-t mixture autoregressive model with applications to heavy-tailed financial data. *Biometrika*, 96(3):751–760, 2009.
- [24] Samuel Ampadu, Eric T. Mensah, Eric N. Aidoo, Alexander Boateng, and Daniel Maposa. A comparative study of error distributions in the garch model through a monte carlo simulation approach. *Scientific African*, 23:e01988, 2024.
- [25] Timothy Graves, Robert B. Gramacy, Christian L. E. Franzke, and Nicholas W. Watkins. Efficient bayesian inference for natural time series using arfima processes. *Nonlinear Processes in Geophysics*, 22:679–700, 2015.
- [26] Tian Zhou, Ziqing Niu, Xue Wang, Liang Sun, and Rong Jin. A comprehensive survey of deep learning for time series forecasting: Architectural diversity and open challenges. *Artificial Intelligence Review*, 58(3):1–45, 2025. Comprehensive review of Transformer architectures and attention mechanisms for time series.
- [27] Xiaoming Liu, Wei Chen, and Yue Zhang. Arima and attention-based cnn-lstm hybrid neural network for battery life estimation. In *Proceedings of the 2024 8th International Conference on Machine Learning and Soft Computing*, pages 45–51. ACM, 2024. Hybrid model combining ARIMA preprocessing with attention-based neural networks.
- [28] Isaac Godfried. Advances in deep learning for time series forecasting/classification winter 2024. *Medium: Deep Data Science*, 2024. Review of foundation models including TimeGPT-1, Chronos, and TimesFM for time series forecasting.
- [29] Ю. П. Кунченко and Ю. Г. Лега. *Оцінювання параметрів випадкових величин методом максимізації поліномів*. Наукова думка, Київ, 1991.
- [30] Yuriy P. Kunchenko. *Polynomial Parameter Estimations of Close to Gaussian Random Variables*. Shaker Verlag, Aachen, Germany, 2002.
- [31] Ю. П. Кунченко. *Статистичні поліноми*. Наукова думка, Київ, 2006.
- [32] Serhii W. Zabolotnii, Zygmunt L. Warsza, and Oleksandr Tkachenko. Polynomial estimation of linear regression parameters for the asymmetric pdf of errors. In Roman Szewczyk, Cezary Zieliński, and Małgorzata Kaliczyńska, editors, *Automation 2018: Advances in Intelligent Systems and Computing*, volume 743, pages 758–772, Cham, 2018. Springer.

- [33] Serhii Zabolotnii, Volodymyr Khotunov, Andrii Chepynoha, and Oleksandr Tkachenko. Estimating parameters of linear regression with an exponential power distribution of errors by using a polynomial maximization method. *Eastern-European Journal of Enterprise Technologies*, 1(4-109):64–73, 2021.
- [34] Volodymyr Palahin and Jozef Juhár. Joint signal parameter estimation in non-gaussian noise by the method of polynomial maximization. *Journal of Electrical Engineering*, 67(3):217–221, 2016.
- [35] Zygmunt L. Warsza and Serhii W. Zabolotnii. A polynomial estimation of measurand parameters for samples of non-gaussian symmetrically distributed data. In Roman Szewczyk, Cezary Zieliński, and Małgorzata Kaliczyńska, editors, *Automation 2017: Advances in Intelligent Systems and Computing*, volume 550, pages 468–480, Cham, 2017. Springer.
- [36] Serhii W. Zabolotnii, Zygmunt L. Warsza, and Oleksandr Tkachenko. Estimation of linear regression parameters of symmetric non-gaussian errors by polynomial maximization method. In Roman Szewczyk, Jurek Sasiadek, and Małgorzata Kaliczyńska, editors, *Automation 2019: Advances in Intelligent Systems and Computing*, volume 920, pages 636–649, Cham, 2020. Springer.
- [37] Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054, 1982.
- [38] Adrian Pagan. Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, 25(1):221–247, 1984.
- [39] Whitney K. Newey. A method of moments interpretation of sequential estimators. *Economics Letters*, 14(2-3):201–206, 1984.
- [40] Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, New York, NY, 1993.
- [41] Wei Pan. Akaike’s information criterion in generalized estimating equations. *Biometrics*, 57(1):120–125, 2001.
- [42] Roger Koenker and Jr. Bassett, Gilbert. Regression quantiles. *Econometrica*, 46(1):33–50, 1978.
- [43] Frank R. Hampel, Elvezio M. Ronchetti, Peter J. Rousseeuw, and Werner A. Stahel. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, New York, 1986.
- [44] Andrew C. Harvey. *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press, Cambridge, 2013.
- [45] Sylvia Frühwirth-Schnatter and Sylvia Kaufmann. Model-based clustering of multiple time series. *Journal of Business & Economic Statistics*, 24(1):78–89, 2006.

- [46] Jouchi Nakajima and Yasuhiro Omori. Stochastic volatility model with leverage and asymmetrically heavy-tailed error using gh skew student's t-distribution. *Computational Statistics & Data Analysis*, 56(11):3690–3704, 2012.
- [47] Roger Koenker. *Quantile Regression*. Cambridge University Press, Cambridge, 2005.