https://doi.org/10.31891/2219-9365-2023-75-26
УДК 519.2+004.8

ПИСАРЧУК Олексій
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
БАРАН Данило
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
ТУГАНСЬКИХ Олександр
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
https://orcid.org/0009-0007-9024-2570

МЕТОДИКА СТАТИСТИЧНОГО НАВЧАННЯ ПАРАМЕТРІВ ЕКСПОНЕНЦІЙНОЇ МАТЕМАТИЧНОЇ МОДЕЛІ

Запропоновано методика аналітичного визначення параметрів експоненційної моделей за статистичною вибіркою вимірів для задач Data Science. Методика базується на використанні диференціально-нетейлорівських перетворень. Наведено приклад застосування запропонованого підходу з навчання експоненційної моделі та доведено його ефективність, порівняно із класичними методами чисельних ітераційних розрахунків.

Ключові слова: модель, статистичне навчання, Data Science, диференціальні перетворення.

PYSARCHUK Oleksii, BARAN Danylo, TUHASKYKH Oleksandr
National Technical University of Ukraine "Ihor Sikorskyi Kyiv Polytechnic Institute"

METHOD OF STATISTICAL LEARNING OF THE PARAMETERS OF THE EXPONENTIAL MATHEMATICAL MODEL

In the practice of data processing in many applied fields: control of production processes; automation of dynamic object management; applied economic analysis - strict requirements are put forward for the accuracy, efficiency and reliability of the assessment and forecasting of the development of the researched processes. This requires the complication of mathematical models used in statistical learning technologies in the direction of increasing their nonlinearity. This is possible in the development of methodological support for Data Science technologies with the introduction of analytical approaches to determining the parameters of nonlinear models.

Analytical method of determining parameters for nonlinear models using statistical samples in Data Science is suggested here. It is based on statistical learning methodology in scheme of differential non-taylor transformations. As an example of suggested method usage, the exponential model was built. The effectiveness of its application was confirmed using classical numerical iterative calculations as comparison.

Key words: model, statistical learning, Data Science, differential transformations.

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

В практиці обробки даних в багатьох прикладних галузях: контроль виробничих процесів; автоматизація управління динамічними об'єктами; прикладний економічний аналіз – висуваються жорсткі вимоги до точності, оперативності та достовірності оцінювання і прогнозування розвитку досліджуваних процесів. Це потребує ускладнення математичних моделей, що використовуються в технологіях статистичного навчання в напрямку підвищення їх нелінійності. Зазначене можливо у розвитку методологічного забезпечення технологій Data Science із впровадженням впровадженням аналітичних підходів до визначення параметрів нелінійних моделей.

Аналіз досліджень та публікацій

Традиційно, для вирішення задач побудови нелінійних моделей використовуються чисельні методи апроксимації [1]. Суть їх полягає в ітераційному наближення оцінок до значення, що відповідає заданим вимогам до точності. Такі підхлди потребують апріорних знань про початкові значення оцінок параметрів моделі мають значну обчислювальну складність. Це унеможливлює їх використання для задач сучасних задач Data Science. Тому ***актуальною*** є задача удосконалення методів статистичного навчання у напрямку подолання недоліків традиційних підходів.

Формулювання цілей статті

Метою досліджень є розробка методики статистичного навчання параметрів експоненційної математичної моделі.

Виклад основного матеріалу

Суть методики аналітичного статистичного навчання параметрів експоненційної математичної моделі полягає в імплементація теоретико-емпіричного підходу до аналізу нелінійних процесів. Він базується на перенесенні якостей спрощеної моделі досліджуваного процесу на нелінійну, яка є більш адекватною до теоретичного опису процесу. Для отримання параметрів спрощеної моделі можуть пропонується використати будь-який метод статистичного навчання, наприклад метод найменших квадратів (МНК) Теоретичний опис математичної моделі може бути отриманий з аналітичних спостережень за природою вибірки даних. Значна, складність нелінійного теоретичного опису моделі унеможливлює застосування класичних статистичних алгоритмів згладжування. Запропонований підхід полягає в багатоетапному отримання параметрів нелінійної моделі з використанням диференційних перетворень [3] наближення властивостей поліноміальної та нелінійної моделей. Такий підхід дозволяє поєднати відносну простоту поліноміальної моделі з більш прогностично стійкою нелінійною моделлю.

Вихідними даними для аналізу нелінійного процесу є дискретна вибірка параметрів:

$$y = \{y_1, y_2, y_3, \dots, y_n\}, \tag{1}$$

Аналітично визначена нелінійна модель, позначено $f(t, c)$, та задана функціональною залежністю або диференціальним рівнянням. З теоретичної точки зору, це найбільш точно описує природу досліджуваного процесу. Необхідно побудувати нелінійну модель досліджуваного процесу $f(t, c)$ за вибіркою експериментальних даних (1).

Далі, використовуючи обраний статистичний алгоритм згладжування МНК, формуються параметри експериментальної моделі (2).

$$\hat{z}(t) = \Phi(y), \tag{2}$$

де Φ позначає операції алгоритму визначення параметрів моделі $\hat{z}(t)$ за вибіркою експериментальних даних (1) відповідно до обраного методу статистичного аналізу. Форма моделі $\hat{z}(t)$ є, як правило, поліноміальною в обраному базисі, що має властивості і недоліки традиційних підходів статистичного навчання:

$$y(x) = \sum_{i=1}^{m} c_i \varphi_i(x_i), \tag{3}$$

де c_i – коефіцієнти полінома; $\varphi_i(x_i)$ – базисні функції полінома.

Таким чином отримано дві різні моделі. $f(t, c)$ є теоретичною формою, яка більш адекватно відображає властивості досліджуваного процесу, але для неї не відомі ключові параметри $c = \{c_i\}, i = 1 \dots m$. В той час $\hat{z}(t)$ є апроксимуючою експериментальною моделлю, форма і параметри якої відомі з використання класичних статистичних методів, але вона менш адекватна процесу та містить більший вплив випадкових помилок.

Операція перенесення властивостей з однієї моделі на іншу можна реалізувати з використанням методу диференційних перетворень (ДП) [3]. Це реалізується за балансом диференціальних спектрів (БДС) з ДП тейлорівського базису.

Диференціальні перетворення в загальному випадку – це функціональні перетворення вигляду:

$$\boldsymbol{Z(k) = P\{z(t)\}_{t^*} = \frac{H^k}{k!}\left[\frac{d^k z(t)}{dt^k}\right]_{t^*}} \tag{4}$$

$$\boldsymbol{z(t) = f(t, c)} \tag{5}$$

де t^* – значення аргументу, при якому проводиться перетворення; $Z(k)$ – дискретна функція аргументу $k = 0,1,2, \dots$; H – відрізок аргументу, на якому розглядається функція $z(t)$; $f(t, c)$ – апроксимуюча функція; c– сукупність вільних коефіцієнтів c_i.

Вираз (4) забезпечує отримання за оригіналом $z(t)$ його зображення $Z(k)$ (пряме перетворення). Обернене перетворення (5) надає можливість відновлення оригіналу $z(t)$ у вигляді апроксимуючої функції. Диференціальне зображення $Z(k)$ називається диференціальним спектром (ДС) або Р-спектром, а значення $Z(k)$ при конкретних значеннях аргументу – дискретами ДС (Р-дискретами). У найпростішому випадку функція $f(t, c)$ має вигляд багаточлена, а відновлення оригіналу зводиться до підсумовування дискрет Р-спектра у вигляді відрізка ряду Тейлора - диференціально–тейлорівські (ДТ) перетворення).

Відповідно до БДС в схемі ДТ перетворень – визначення параметрів нелінійної моделі реалізується шляхом формування і розв’язку системи рівнянь

$$\boldsymbol{P\{z(t)\}_{t^*} \Rightarrow Z(k) = F(k,c) \Leftarrow P\{f(t,c)\}_{t^*}} \quad (6)$$

Покладемо, що процес, який має експоненційну природу заданий у формі набору вимірів (1) має загальну аналітичну форму:

$$f(t,a) = a_0 + a_1 t + a_2 e^{a_3 t}, \quad (7)$$

де t – це залежний параметр моделі; a – це вектор незалежних параметрів моделі.

Надалі проводиться згладжування вибірки (1) статистичним методом - МНК.

Визначення невідомих параметрів апроксимуючої функції за МНК реалізується за матричним виразом [2]:

$$C = (A^T A)^{-1} A^T Y, \quad (8)$$

де Y – вектор вибірки виміряних даних; A – матриця базисних функцій; C – вектор шуканих параметрів моделі.

Результатом застосування (8) є набір незалежних параметрів експериментальної моделі, статистично «навчені» за вибіркою вимірів. Надалі визначаються параметри нелінійної моделі за параметрами поліноміальної за схемою (6):

$$V(t,c) = c_0 + c_1 t + c_2 t^2 + c_3 t^3, \quad (9)$$

$$F(t,a) = (a_0 + a_2) + (a_1 + a_2 + a_3)t + \frac{a_2 {a_3}^2}{2} t^2 + \frac{a_2 {a_3}^3}{6} t^3, \quad (10)$$

де $V(t,c)$, $F(t,a)$ поліноміалні моделі, утворені з диференціальних Р-спектрів початкових моделей.

Наближення двох моделей реалізується за методом БДС, що дає можливість утворити систему рівнянь:

$$\begin{cases} c_0 - (a_0 + a_2) = 0, \\ c_0 - (a_1 + a_2 + a_3) = 0, \\ c_0 - \frac{a_2 {a_3}^2}{2} = 0, \\ c_0 - \frac{a_2 {a_3}^3}{6} = 0. \end{cases} \quad (11)$$

Розв’язком системи рівнянь (11) є параметри нелінійної моделі (7), пов’язані із параметрами поліноміальної моделі (9)

$$\begin{cases} a_0 = c_0 - a_2, \\ a_1 = c_1 - (a_2 * a_3), \\ a_2 = \frac{2c_2}{{a_3}^2}, \\ a_3 = \frac{3c_3}{c_2}. \end{cases} \quad (12)$$

Для перевірки та порівняння якості отриманих результатів з іншими методами, було виконано моделювання з наступними параметрами:

Ідеальний тренд: $f(x) = 0.0000005x^2, x \in 0{,}1 \ldots 10000$;

Стохастичні відхилення: нормальний закон, $\mu = 0, \sigma = 5$;

Аномальні відхилення: нормальний закон, $\mu = 0, \sigma = 15$;

На рис. 1 візуалізовано ідеальний тренд $f(x)$ з різними відхиленнями позначеними різними кольорами.

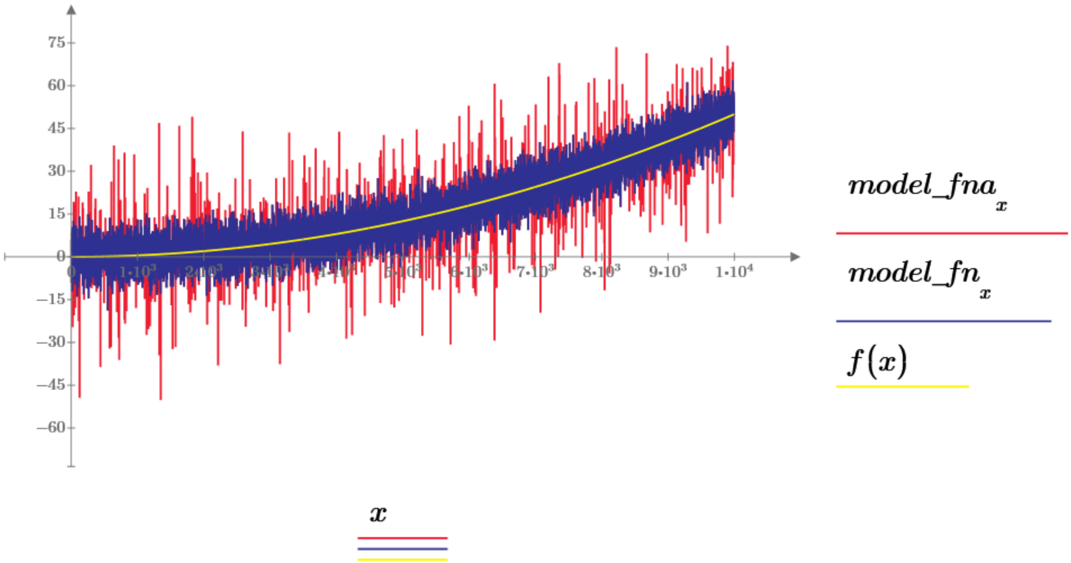

Рис. 1. Ідеальний тренд $f(x)$ з різними відхиленнями

До обробки моделі вхідних даних – визначення лінії тренду було застосовано три підходи:

- ✓ МНК – 1;
- ✓ запропонований метод – 2;
- ✓ чисельний метод з використанням алгоритму Левенберга-Маркардта – 3;

Отримані результати кожного з методів були оцінені за показниками: середньоквадратичне відхилення (СКВ), лінійного відхилення (ЛВ) та часу обрахунку T. Значення показники занесено в табл. 1.

Таблиця 1

Методи	СКВ	ЛВ	T (мкс)
1	37.842	0.102	436
2	44.489	0.098	568
3	37.808	0.102	4647

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

Отримані дані розрахунків та моделювання показують, що застосування розробленого методу для визначення параметрів нелінійної моделі є ефективним за показниками часу та адекватності моделі, порівняно із чисельними підходами. Запропонований підхід базується на використанні методу диференціальних перетворень та може бути масштабований для інших типів нелінійних моделей. Цей перехід дозволяє покращити також прогностичні характеристики моделі.

Література

1. Chapra, Steven C., and Raymond P. Canale. Numerical methods for engineers. Vol. 1221. New York: Mcgraw-hill, 2011. – 987p.

2. Пухов Г.Е. Диференциальные преобразования и математическое моделирование физических процессов. – Киев: Наук. Думка, 1986. – 159 с.

3. Hastie, Trevor, et al. The elements of statistical learning: data mining, inference, and prediction. Vol. 2. New York: springer, 2009. – 764p.

References

1. Chapra, Steven C., and Raymond P. Canale. Numerical methods for engineers. Vol. 1221. New York: Mcgraw-hill, 2011. – 987p.

2. Pukhov H.E. Dyferentsyalnыe preobrazovanyia y matematycheskoe modelyrovanye fyzycheskykh protsessov. – Kyev: Nauk. Dumka, 1986. – 159 s.

3. Hastie, Trevor, et al. The elements of statistical learning: data mining, inference, and prediction. Vol. 2. New York: springer, 2009. – 764p.