

УДК 681.5:629.3

DOI <https://doi.org/10.32782/3041-2080/2025-3-10>

МЕТОДИКА КОРЕКТУВАННЯ ПАРАМЕТРІВ БАЗИ ДАНИХ ТЕХНОЛОГІЧНОГО ПРОЦЕСУ

Разживін Олексій Валерійович,

кандидат технічних наук, доцент,
доцент кафедри автоматизації, електро- та робототехнічних систем
ТОВ «Технічний університет «Метінвест Політехніка»
ORCID ID: 0000-0002-1371-2651
Scopus ID: 57672166200

Койфман Олексій Олександрович,

кандидат технічних наук, доцент,
завідувач кафедри автоматизації, електро- та робототехнічних систем
ТОВ «Технічний університет «Метінвест Політехніка»
ORCID ID: 0000-0003-2075-7417
Scopus ID: 55808117000

Мірошниченко Вікторія Ігорівна,

кандидат технічних наук, доцент,
доцент кафедри автоматизації, електро- та робототехнічних систем
ТОВ «Технічний університет «Метінвест Політехніка»
ORCID ID: 0000-0002-5956-7867
Scopus ID: 57205494697

Новіков Дмитро Сергійович,

аспірант
Донбаської державної машинобудівної академії
ORCID ID: 0009-0003-3945-0754

У статті розкрито науково-технічно проблему управління складним та унікальним обладнанням. Під час автоматизації технологічного процесу використовують таблично задані параметри, за допомогою яких регламентують дії операторів та забезпечують необхідні результати процесу. Математичне моделювання, а також організація роботи АСУ ТП можливі за допомогою цифрових двійників об'єктів автоматизації. Сучасна тенденція синтезу цифрових двійників виходить з бази даних параметрів «вхід – вихід об'єкта».

Складання бази даних параметрів «вхід – вихід об'єкта» у вигляді таблиці, яка може містити: розроблені інструкції або вказівки щодо регулювання параметрів технологічного процесу, що охоплюють усі ймовірні ситуації; параметричну зміну керованої величини залежно від зміни вхідних (керівних) параметрів.

Тому процес створення табличної інформації потребує певних обмежень щодо достатності. Під достатністю варто розуміти достатність обсягу інформації, використання якого гарантує сприятливий результат процесу чи високу точність побудови цифрового двійника об'єкта автоматизації. До того ж таблиця, що містить значення параметрів процесу, повинна охоплювати всі вірогідні ситуації. Однак практична реалізація такого завдання вимагає проведення великої кількості експериментальних і модельних досліджень. Як правило, на практиці не вдається заповнити всю таблицю через обмежену кількість вимірів.

З огляду на викладені особливості, завдання визначеності ступеня інформативності бази даних опису стану об'єктів автоматизації є актуальним.

Ключові слова: база даних, коректування, таблиця, параметр, пропуск, метод найменших квадратів.

Razhyvin Oleksii, Koyfman Oleksiy, Miroshnychenko Viktoriia, Novikov Dmytro. Methodology for adjusting database parameters of the technological process

The article reveals the scientific and technical problem of managing complex and unique equipment. When automating a technological process, tabulated parameters are used, with the help of which the actions of operators are regulated and the necessary results of the process are ensured. Mathematical modeling, as well as the organization of ACS TP work, is possible with the help of digital duplicates of automation objects. The modern trend of synthesis of digital doubles comes from the database of input-output parameters of the object.

Compilation of a database of input-output parameters of the object in the form of a table, which may contain: developed instructions or instructions for regulating the parameters of the technological process, covering all situations that are likely to occur; parametric change of the controlled quantity depending on the change of the input (control) parameters.

Therefore, the process of creating tabular information requires certain limitations regarding sufficiency. In this case, sufficiency should be understood as the sufficiency of the amount of information, the use of which guarantees a favorable outcome of the process or high accuracy of building a digital double of the automation object. At the same time, the table containing the values of the process parameters should cover all situations that are likely to occur. However, the practical implementation of such a task requires a large number of experimental and model studies. As a rule, in practice, it is not possible to fill the entire table due to the limited number of measurements

In connection with the stated features, the task of determining the degree of informativeness of the database describing the state of automation objects is an urgent scientific and technical task

Key words: database, adjustment, table, parameter, omission, method of least squares.

Вступ. Таблично задані параметри, за допомогою регламенту яких діють і забезпечують необхідні результати процесу відновлення періоду плавки в технологічних комплексах, досить широко застосовують у практиці управління. До того ж таблиця, що містить значення параметрів процесу, повинна охоплювати всі вірогідні ситуації. Однак практична реалізація такого завдання вимагає проведення великої кількості експериментальних і модельних досліджень. Як правило, на практиці не вдається заповнити всю таблицю через обмежену кількість вимірів, тому розв'язання цієї проблеми є актуальним науково-технічним завданням.

Методи та методики дослідження. Виконуючи поставлене завдання управління тепловим режимом плавки в технологічних комплексах «Піч-Ківш», використовують накопичену в базі даних інформацію. Ця інформація необхідна для коригувальних дій. Вона представлена у вигляді численних даних, сформованих двома шляхами [1]:

- за допомогою пасивного експерименту із записом таких взаємозв'язаних параметрів: зміна швидкості нагріву розплаву, величина підвідної потужності (номер ступеня напруги пічного трансформатора) й маса розплаву;
- за допомогою моделювання потужності теплових втрат.

Інформаційна база параметрів технологічного процесу плавки містить відомості про такі параметри:

- швидкість нагріву розплаву залежно від кількості ступенів пічного трансформатора й маси розплаву;
- потужності теплових втрат залежно від швидкості нагрівання й маси розплавлення.

Для реалізації директивного графіка плавки потрібно встановити певну швидкість нагріву шляхом вибору номерів ступеня пічного трансформатора. З огляду на це, необхідна база даних, у якій містяться відомості про швидкість (інтенсивність) нагрівання металу за різних

ступенів напруги пічного трансформатора й різної маси рідкого металу в ківші.

В експериментальних дослідженнях швидкість нагріву визначилася за виразом [1]:

$$\Delta T_m = \frac{T_i - T_{i-1}}{\Delta t}, \quad (1)$$

де T_i , T_{i-1} – температура рідкого металу в i -проміжок часу, °C; Δt – проміжок часу між вимірами температури, хв.

Заміри температури металу відбувалися в такі періоди плавки:

- до початку плавки (вимірювалася початкова температура металу);
- під час дугового підігріву металу (вимірювалося значення величини швидкості нагріву рідкого металу, фіксувалися номер ступеня пічного трансформатора та маса металу);
- під час витримки металу (з інтервалом 15–20 хв);
- під час охолодження металу (самоохолодження) з піднятим та закритим склепінням (визначалося значення швидкості самоохолодження рідкого металу);
- під час охолодження металу, без дугового підігріву та продування аргоном;
- під час охолодження металу з дуговим підігрівом та продувці аргоном.

Результати. Під час створення бази даних теплових втрат за допомогою пасивного експерименту не вдається заповнити всю таблицю через обмежену кількість вимірювань. З огляду на це, для заповнення прогалів у таблиці передбачено отримати відсутні дані шляхом моделювання, а також у процесі подальшої експлуатації. Під час заповнення прогалів у таблицях даних під час експлуатації можлива ситуація, коли отримані дані, що описують залежності зміни теплових втрат від швидкості нагрівання та маси розплаву, стануть неадекватними. Варто також урахувати, що за наявності великої кількості прогалів даних регресійний аналіз майже не застосовується через велику різноманітність можливого розташування пропущених даних.

З огляду на це пропонується така методика заповнення прогалін у таблицях бази даних теплових втрат [2; 3].

На першому етапі заповнення прогалін у базі даних необхідно оцінити інформативність таблиці через ентропію та функцію бажаності Харрінгтона [2; 3]. Ця оцінка дає змогу формалізувати процедуру ухвалення рішення про достатність інформації та оцінити точність процесу заповнення бази ітераційним методом.

На другому етапі реалізується алгоритм заповнення прогалін у таблицях даних, заснований на моделюванні безлічі даних різноманітностями M малої розмірності.

Вектор даних x (потужність теплових втрат $P_{\text{тп}}$) з k -пробілами представляється як тривимірне лінійне різноманіття L_x , паралельне трьом координатним осям (ΔT_M , $M_{\text{металу}}$, $P_{\text{тепл. втрат}}$), які відповідають пропущеним даним. За наявності апріорних обмежень на пропущені значення місце L_x посідає прямокутний паралелепіпед $P_x \subset L_x$ (див. рис. 1).

Особливу роль у подальшому розгляді відіграють дві точки: $Pr_M(x)$ – найближча до x (відповідно до L_x або P_x) точка M , та для некомплектних даних – $Im_M(x)$ – найближча до M точка L_x (або, відповідно, P_x).

З використанням різноманіття M виконуються такі операції:

- заповнення прогалін – заміна x на $Im_M(x)$,
- ремонт даних (заміщення даних моделлю) – заміна x на $Pr_M(x)$.

Регресійна залежність однієї частини даних від іншої визначається через заповнення

прогалін, відповідних залежним змінним, за відомими значеннями незалежних.

Ітераційний процес моделювання даних у тому, що з вихідних даних будується найкраща (у певному сенсі) модель – різноманіття M малої розмірності. Далі з даних x віднімаються проєкції $Pr_M(x)$. Водночас отримуємо відхилення від першої моделі. Для цього безлічі ухилень знову будується проста модель і т.д., поки всі ухилення не стануть досить близькими до нуля.

Таблиця бази даних має прямокутну форму, клітини або заповнені дійсними числами, або знаком @, що означає відсутність даних. Потрібно правдоподібно відновити відсутні дані. За більш детального розгляду виникають три завдання:

- 1) заповнити перепустки в таблиці;
- 2) відредагувати таблицю – змінити значення відомих даних так, щоб моделі працювали якнайкраще;

- 3) побудувати за таблицею обчислювач, що заповнює прогалини в рядку даних для аналізу.

Для виконання цих завдань пропонується використовувати метод послідовного наближення множини векторів даних (рядок таблиці) прямими. Основна процедура зводиться до виконання завдання пошуку найкращого наближення таблиці з перепустками $A = (a_{ij})$ матрицею виду $x_i y_j + b_j$ методом найменших квадратів

$$\Phi = \sum_{\substack{i,j \\ a_{ij} \neq @}} (a_{ij} - x_i y_j - b_j)^2 \rightarrow \min. \quad (2)$$

Якщо фіксовано два з трьох векторів x_i , y_j та b_j , то третій легко розміщується за явними формулами.

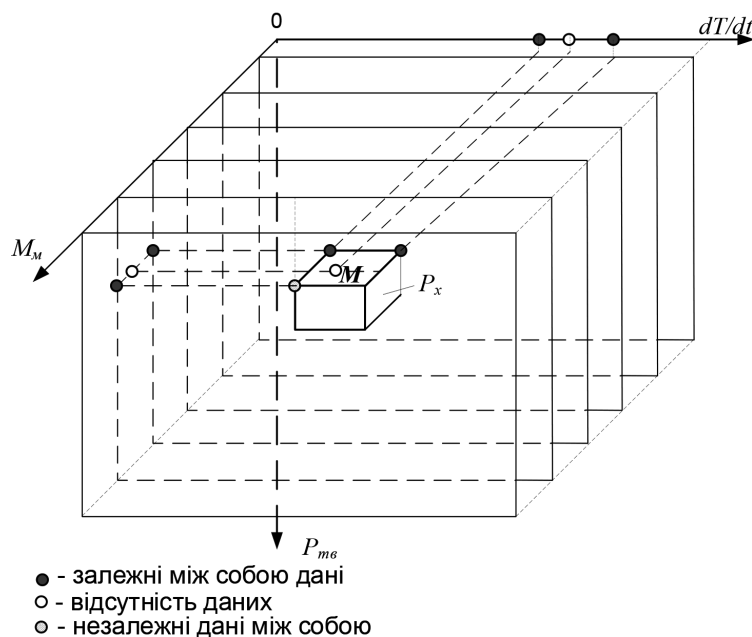


Рис. 1. Тривимірне різноманіття вектора даних

Задаючись майже довільними початковими наближеннями для двох, шукаємо значення третього, далі оголошуємо невідомим інший вектор з трьох, знаходимо його значення, нарешті знаходимо третій і т. д. (по колу) – ці прості ітерації, очевидно, сходяться. Ба більше, за фіксованим x_i можна відразу за явними формулами порахувати значення y_j та b_j . Так, розщеплення проводиться не на три, а на два складники.

За фіксованих векторів y_j та b_j значення x_i , що доставляють мінімум формі (2), визначаються з рівностей $\partial\Phi/\partial x_i = 0$ так:

$$x_i = \left(\sum_{a_j \neq @} (a_j - b_j) y_j \right) / \left(\sum_{a_j \neq @} (y_j)^2 \right). \quad (3)$$

За фіксованого вектора x_i значення y_j та b_j , що доставляють мінімум формі (2), визначаються з двох рівностей $\partial\Phi/\partial y_j = 0$ та $\partial\Phi/\partial b_j = 0$ так:

Для кожного j маємо систему з двох рівнянь щодо y_j та b_j :

$$\begin{cases} y_j A_{01}^j + b_j A_{00}^j = B_0^j \\ y_j A_{11}^j + b_j A_{10}^j = B_1^j \end{cases}, \quad (4)$$

де $A_{kl}^j = \sum_{a_j \neq @} x_i^{k+l} a_j x_i^k$, $B_k^j = \sum_{a_j \neq @} a_j x_i^k$, $k = 0 \dots 1$, $l = 0 \dots 1$.

Виділяючи з першого рівняння b_j і підставляючи отримане значення в друге, отримаємо:

$$y_j = \frac{B_1^j - B_0^j \frac{A_{10}^j}{A_{00}^j}}{A_{11}^j - A_{01}^j \frac{A_{10}^j}{A_{00}^j}}, \quad b_j = \frac{B_0^j - y_j A_{01}^j}{A_{00}^j}. \quad (5)$$

Початкові значення:

y – випадковий, нормований на 1 (тобто $\sum_j y_j^2 = 1$).

$b_j = \frac{1}{n_j} \sum_{a_j \neq @} a_j$, де $n_j = \sum_{a_j \neq @} 1$ (кількість відомих

даних у j -му стовпці), тобто b_j визначається як середнє значення в стовпці.

Критерієм зупинки є трохи відносного поліпшення $\Delta\Phi/\Phi$, де $\Delta\Phi$ – отримане за цикл зменшення значення Φ , а Φ – саме поточне значення. Другий критерій – трохи самого значення Φ . Остаточо: процедура зупиняється, якщо $\Delta\Phi/\Phi < \varepsilon$ або $\Phi < \delta$ для деяких ε при $\delta \rightarrow 0$.

Послідовне вичерпання матриці A . Для цієї матриці A знаходиться найкраще наближення матрицею P_1 виду $x_i y_j + b_j$. Далі для $A - P_1$ шукаємо найкраще наближення цього ж виду P_2 і т. д. Контроль ведеться, наприклад, щодо

залишкової дисперсії стовпців.

Q -факторне заповнення перепусток є їх визначенням із суми Q отриманих матриць виду $x_i y_j + b_j$.

Q -факторний «ремонт» таблиці – заміна її на суму Q отриманих матриць виду $x_i y_j + b_j$.

У результаті описаного процесу будується послідовність матриць P_q виду $x_i y_j + b_j$. ($P_q = x_i^q y_j^q + b_j^q$), вичерпна вихідна матриця A із заданою точністю.

Опишемо операцію відновлення даних у рядку, що надходить на обробку a_j з пробілами (деякі $a_j = @$). Для кожного q по заданому рядку визначається число $x^q(a)$ та вектор a_j^q :

$$\begin{aligned} a_j^0 &= a_j \quad (a_j \neq @); \\ x^1(a) &= \left(\sum_{a_j \neq @} (a_j^0 - b_j^1) y_j^1 \right) / \left(\sum_{a_j \neq @} (y_j^1)^2 \right); \\ a_j^1 &= a_j^0 - b_j^1 - x^1(a) y_j^1 \quad (a_j \neq @); \\ x^q(a) &= \left(\sum_{a_j \neq @} (a_j^{q-1} - b_j^q) y_j^q \right) / \left(\sum_{a_j \neq @} (y_j^q)^2 \right); \\ a_j^q &= a_j^{q-1} - b_j^q - x^q(a) y_j^q \quad (a_j \neq @); \end{aligned} \quad (6)$$

Тут різноманіття M – пряма, координати точок якої на M задаються параметричним рівнянням $z_j = t y_j + b_j$, а проєкція $\text{Pr}_M(a)$ визначається згідно з (6):

$$\begin{aligned} \text{Pr}(a) &= t(a) y_j + b_j; \\ t(a) &= \left(\sum_{a_j \neq @} (a_j - b_j) y_j \right) / \left(\sum_{a_j \neq @} (y_j)^2 \right). \end{aligned} \quad (7)$$

Для Q -факторного відновлення даних рахуємо:

$$\bar{a}_j = \sum_{q=1}^Q x^q(a) y_j^q + b_j^q, \quad (a_j \neq @). \quad (8)$$

Якщо прогалини відсутні, то описаний метод призводить до звичайних основних компонентів – сингулярного розкладання вихідної таблиці даних. У цьому разі, починаючи з $q = 2$, $P_q = x_i^q y_j^q$ ($b = 0$). Загалом це не так, і центрування до даних з пробілами не застосовується.

Висновок. Варто врахувати, що за відсутності прогалин, отримані прямі будуть ортогональні, тобто отримаємо ортогональну систему факторів (прямих). З огляду на це, при неповних даних можливий процес ортогоналізації отриманої системи факторів, який полягає в тому, що вихідна таблиця відновлюється шляхом перерахунку наново, але вже на повних даних.

ЛІТЕРАТУРА:

1. Разживін О. В., Храмов С. О. Аналіз розподілу енергетичних параметрів при дугового плавлі металу. *Вісник ДДМА*. Краматорськ : ДДМА, 2018. № 1(43). С. 65–69.
2. Сікора Л. С., Ткачук Р. Л., Довгун О. С. Моделі нечітких і ймовірнісних мір на розбитті простору станів складних систем в умовах невизначеності. *Зб. наук. пр., Вип. 30*. Київ, ІПМЕ. 2005. С. 132–140.
3. Borda M. *Fundamentals in Information Theory and Coding*. Springer, Berlin. 2011. 486 p. DOI 10.1007/978-3-642-20347-3
4. Суботін О. В., Разживін О. В. Підвищення достовірності первинної інформації в системах комплексної автоматизації прокатних станів. *Обробка матеріалів тиском: Зб. наук. пр. ДДМА, Краматорськ, 2010. № 3(24)*. С. 205–211.
5. Кучеренко Є. І., Кучеренко В. Є., Глушенкова І. С., Творошенко І. С. *Методи, моделі та інформаційні технології оцінювання станів складних об'єктів: монографія*; ХНАМГ, ХНУРЕ. Харків : ХНАМГ; ХНУРЕ, 2012. 278 с.
6. Кватер Т. *Нейромережні інформаційні технології контролю і діагностики динамічних об'єктів в умовах невизначеності*. Львів : ДНДТ, 2005. 267 с.

REFERENCES:

1. Razzhyvin, O.V., & Khramov, S.O. (2018). Analiz rozpodilu enerhetychnykh parametriv pry duhovoho plavtsi metalu [Analysis of the distribution of energy parameters during arc melting of metal]. *Visnyk DDMA*. – Kramators'k: DDMA, 2018. № 1(43). pp. 65–69. [in Ukrainian].
2. Sikora, L.S., Tkachuk, R.L., & Dovhun, O.S. (2005). Modeli nechitkykh i ymovirnisnykh mir na rozbytti prostoru staniv skladnykh system v umovakh nevyznachenosti [Models of fuzzy and probabilistic measures on the partitioning of the state space of complex systems under conditions of uncertainty]. *Zb. nauk. pr., Vyp. 30*. Kyiv, IPME. pp. 132–140. [in Ukrainian].
3. Borda, Monica (2011). *Fundamentals in Information Theory and Coding*. Springer. ISBN 978-3-642-20346-6. URL: <http://surl.li/aqhxii>
4. Subotin, O.V., & Razzhyvin, O.V. (2010). Pidvyshchennya dostovirnosti pervynnoyi informatsiyi v systemakh kompleksnoyi avtomatyzatsiyi prokatnykh staniv [Increasing the reliability of primary information in systems of complex automation of rolling mills]. *Obrobka materialiv tyskom: Zb. nauk. pr. DDMA Kramators'k, 2010. № 3(24)*. S. 205–211. [in Ukrainian].
5. Kucherenko, YE.I., Kucherenko, V.YE, Hlushenkova, I.S. & Tvoroshenko, I. S. (2012). *Metody, modeli ta informatsiyini tekhnolohiyi otsinyuvannya staniv skladnykh ob"yektiv: monohrafiya* [Methods, models and information technologies for assessing the state of complex objects: monograph]. /KHNAMH, KHNURE. KH. : KHNAMH; KHNURE. 278 p. [in Ukrainian].
6. Kvater, T. (2005). *Neyromerezhni informatsiyini tekhnolohiyi kontrolyu i diahnostyky dynamichnykh ob"yektiv v umovakh nevyznachenosti* [Neural network information technologies of control and diagnostics of dynamic objects under conditions of uncertainty]. Lviv : DNDT. 267 p. [in Ukrainian].