УДК 621.396.664

М. В. Коваленко, В. В. Воротніков

ФОРМУВАННЯ ВЕКТОРА ПЕРВИННИХ ЗНАЧУЩИХ ПАРАМЕТРІВ ДЛЯ ОЦІНКИ ТЕХНІЧНОГО СТАНУ СКЛАДНИХ ІНФОРМАЦІЙНИХ СИСТЕМ

У статті запропоновано розв'язок задачі вибору узагальнених параметрів для перевірки працездатності складного технічного об'єкта на етапі проектування його системи контролю. Групування простору ознак проводиться за допомогою методу головних компонент (МГК). Наведено переваги застосування МГК для задач встановлення технічного стану з урахуванням визначених кореляційних зв'язків між первинними параметрами.

Постановка проблеми. У даний час працездатність складних технічних об'єктів, як правило, оцінюється за результатами аналізу статистичної інформації про стан їх підсистем, одержуваної у ході процесу контролю первинних параметрів.

За результатами аналізу приймається рішення про технічний стан об'єкта контролю. Для складних інформаційних систем (СІС), що працюють у реальному масштабі часу, детальний аналіз вимірюваної інформації здійснюється, зазвичай, після закінчення процесу контролю. У результаті ідентифікація аварійних ситуацій технічного об'єкта проходить із запізненням.

Огляд останніх досліджень з даної проблеми [1–3] свідчить про те, що для розв'язання задачі контролю працездатності технічного стану в СІС формуються вимоги до параметрів призначення складних систем у технічних завданнях (ТЗ) і документації як показники якості та допуски на них:

$$K_{jH} \le K_j \le K_{jB},\ j = 1,2,...,l, \tag{1}$$

де K_{jH}, K_{jB} – відповідно нижня і верхня межа зміни j-го показника якості;

l – кількість показників якості, визначених у ТЗ.

Виконання умови (1) досягається, з одного боку, за рахунок використання високонадійних (і, отже, дорогих) комплектуючих елементів та їх резервування, а з іншого, – за рахунок організації контролю (керування показниками якості).

Перший шлях далеко не завжди прийнятний, тому що призводить до різкого зростання ціни апаратури.

Більш актуальним є другий шлях. Для виконання умови (1) за рахунок організації контролю формується вектор контрольованих параметрів (ознак технічних станів об'єкта):

$$X = (X_1, X_2, ..., X_n),$$

де n – кількість контрольованих параметрів.

Для взаємозв'язку між показниками якості (вектор $K = K_1, K_2, ..., K_j, ..., K_l)$ і первинними параметрами необхідна наявність моделі об'єкта функціонування підсистем СІС [1–3]:

$$K = \varphi(X, B), \tag{2}$$

де $B = (B_1, B_2, ..., B_n)$ – вектор, що характеризує чинники зовнішніх умов у момент експлуатації і проведення контролю працездатності СІС. Якщо безпосередньо параметри вектора X не піддаються вимірам, то здійснюється перехід до контрольованих характеристик (2).

Остаточно співвідношення (1–2) утворюють модель системи контролю СІС за первинними параметрами.

У ході аналізу великої кількості статистичних даних, особливо в умовах роботи в реальному масштабі часу, виникає потреба у зменшенні оброблюваної інформації без втрати або при незначній втраті самої інформації. Враховуючи те, що СІС містить велику кількість контрольованих параметрів, доцільно, щоб інформація про технічний стан СІС зосереджувалася в деяких узагальнених параметрах (УП). Актуальною є задача зменшення розмірності початкових даних, що характеризують технічний стан об'єкта контролю. У статистиці вона відома як задача зменшення ознакового простору об'єкта або як задача стиснення інформації [4, 5]. Її розв'язують за допомогою методів дисперсійного або факторного аналізу.

Факторний аналіз як і метод дисперсійного аналізу – метод головних компонент (МГК), що дозволяє зменшити вимірність ознакового простору СІС. Ці методи є ефективними способами дослідження взаємозв'язків між первинними параметрами. Основна розбіжність між ними полягає у тому, що головні компоненти є лінійними функціями від первинних параметрів, а факторний аналіз подає кореляційну структуру матриці вхідних змінних у термінах латентної (гіпотетичної) моделі.

Крім того, факторний аналіз, на відміну від МГК, не дає однозначного розв'язку задачі, тому що подання кореляційної матриці факторами можна здійснити нескінченною кількістю способів – ортогональне перетворення матриці вхідних даних призводить до нової факторизації [5]. Тому задачу зменшення вимірності ознакового простору СІС пропонується розв'язати за допомогою МГК.

Метою даної роботи є зменшення кількості контрольованих параметрів на етапі проектування системи контролю складного технічного об'єкта без втрати або з мінімальною втратою інформації про його технічний стан.

Виклад основного матеріалу. Для розв'язання задачі задамося початковими даними:

вхідною матрицею первинних параметрів $X(n \times m)$ як вибіркою випадкових величин з багатовимірним сумісним розподілом:

$$X = \begin{pmatrix} x_{11} & \dots & x_{1m} \\ \vdots & \vdots & \vdots \\ x_{n1} & \dots & x_{nm} \end{pmatrix},$$

де m – кількість вимірів первинних параметрів за період контролю t_k.

Знайдемо структурну залежність між первинними параметрами у вигляді лінійної комбінації вхідних даних, що включені до узагальнених параметрів U_j:

$$U_j = \sum_{l=1}^{p} v_l \cdot X_l + \xi_l, \; p<n, \tag{3}$$

де v_l – коефіцієнти лінійної моделі;

ξ_l – некорельовані похибки засобів вимірювань за кожним первинним параметром з нульовим математичним сподіванням $M(\xi_l) = 0$ та однаковою дисперсією $D(\xi_l) = \sigma_\xi^2$;

p – кількість первинних параметрів, включених до узагальненого параметра U_j.

Розв'язання задачі. Для цього з матриці вхідних даних $X(n \times m)$ знаходяться вектори середніх значень $\overline{X}$ та середньоквадратичних відхилень σ_x первинних параметрів складної інформаційної системи:

$$\overline{X} = \left(\overline{X}_1, \overline{X}_2, ..., \overline{X}_n\right), \tag{4}$$

де $\overline{X}_i = \frac{1}{m}\sum_{k=1}^{m} x_{ik}$ – середнє значення i-ого первинного параметра, i=1, m;

$$\sigma_x = \left(\sigma_{x1}, \sigma_{x2}, ..., \sigma_{xn}\right), \tag{5}$$

де $\sigma_{Xi} = \sqrt{D_i}$ – середньоквадратичне відхилення вимірювань i-ого первинного параметра;

$D_i = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}\left(X_i - \overline{X}_i\right)^2}$ – дисперсія вимірювань i-ого первинного параметра.

Враховуючи те, що МГК ґрунтується на стандартизації змінних [5], введемо припущення, що сумісний розподіл первинних параметрів – багатовимірний нормальний. Доцільність такого припущення обумовлена насамперед тим, що лінійні комбінації нормально розподілених первинних параметрів мають, у свою чергу, нормальний розподіл та визначаються вектором середніх значень та кореляційною матрицею [6].

З метою стандартизації змінних введемо змінні Z_i [4]:

$$Z_i = \left(X_{ij} - \overline{X}_i\right) \cdot D_i. \tag{6}$$

Тоді вираз (3) подамо як

$$U_j = \sum_{l=1}^{p} v_l \cdot Z_l + \xi_l, \; p<n, \tag{7}$$

де v_l можна подати як власний вектор, який відповідає системі рівнянь.

$$\left(Z^T \cdot Z - K_Z \cdot I\right) \cdot V = 0, \tag{8}$$

де $Z^T \cdot Z$ – кореляційна матриця;

I – одинична діагональна матриця;

K_Z – характеристичні корені рівняння (8).

Знаходження характеристичних коренів рівняння (8) є розв'язанням задачі побудови лінійної моделі головних компонент.

Для прикладу використання методу головних компонент з метою зменшення розмірності матриці даних розглянемо випадковий вектор $X[X_1, X_2]$ з однаковими дисперсіями $D_1 = D_2 = D_X$ і коваріаційною матрицею R_X:

$$R_X = M\left[\left(X - \overline{X}\right)\left(X - \overline{X}\right)^T\right] = D_X \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \tag{9}$$

де ρ – коефіцієнт кореляції величин X_1 та X_2.

Знайдемо відносну похибку відновлення початкового вектора X за значеннями узагальненого параметра U.

Для побудови лінійної моделі узагальненого параметра U з виразу (8) знайдемо характеристичні корені рівняння.

Враховуючи те, що вектор коефіцієнтів моделі V не нульовий [4], розв'язком (8) буде

$$Z^T \cdot Z - K_Z \cdot I = 0\,. \tag{10}$$

Крім того, для матричної форми запису добуток $K_Z \cdot I$ – ортогональна матриця Φ з елементами φ_i, для скалярного добутку яких правильно [7]:

$$\left(\varphi_i \times \varphi_j\right) = \begin{cases} 1 \;\; при \;\; i \neq j \\ 0 \;\; при \; i = j \end{cases} \tag{11}$$

Тоді задача зменшення розмірності ознакового простору є задачею знаходження найкращого лінійного перетворення первинних параметрів за допомогою матриці ортогональних перетворень Φ за умови їх ортогональності $\left(\varphi_1, \varphi_2\right) = \sum_{j=1}^{n} \varphi_{1j}^2 = 1$, що використовується як обмеження у вигляді рівностей та знаходження максимуму функції χ^2 значущості кореляційної матриці $Z^T \cdot Z$.

Для розв'язання такої задачі оптимізації введемо функцію Лагранжа [8, 9]:

$$L(\varphi) = \varphi_1' \cdot R_X \cdot \varphi_1 - \lambda_1 \cdot \left(\varphi_1' \cdot \varphi_1 - 1\right),$$

де λ_1 – множник Лагранжа.

Необхідну умову екстремуму отримаємо, якщо прирівняємо до нуля похідну

$$\partial L/\partial\varphi = 2 \cdot \left(R_X \cdot \varphi - \lambda \cdot \varphi\right) = 2 \cdot \left(R_X - \lambda \cdot I\right) \cdot \varphi = 0$$

та визначимо вектор матриці R_X :

$$| R_X - \lambda \cdot I | = \left| D_X \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = \lambda^2 - 2 \cdot D_X \cdot \lambda + D_X^2 \cdot \left(1 - \rho^2\right) = 0$$

$$\lambda_1 = D_X (1 + \rho)\;, \lambda_2 = D_X (1 - \rho)\,.$$

Згідно з (8) розв'яжемо систему рівнянь:

$$\left[R_X - \lambda \cdot I\right] \varphi_1 = 0 \quad \text{або} \quad \left[R_X - \lambda \cdot I\right] \begin{bmatrix} \varphi_{11} \\ \varphi_{21} \end{bmatrix} = 0\,,$$

за умови ортогональності компонент $\left(\varphi_1{}', \varphi_1\right) = \varphi_{11}^2 + \varphi_{21}^2 = 1$ отримаємо:

$$-\varphi_{11} + \varphi_{21} = 0\,,\; {\varphi_{11}}^2 + {\varphi_{21}}^2 = 1\,,$$

звідки $\varphi_{11} = \dfrac{1}{\sqrt{2}}$, $\varphi_{12} = \dfrac{1}{\sqrt{2}}$.

Власний вектор визначається як $\varphi_1 = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\varphi_2 = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Матриця ортогонального перетворення у такому випадку має вигляд:

$$\Phi = \left[\varphi_1, \varphi_2\right]' = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Матриця Φ переводить вектор X у вектор $U = \Phi' X = [U_1, U_2]$ з незалежними нормальними компонентами U_1 та U_2 :

$$U = \Phi' X = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} X_1 + X_2 \\ X_1 - X_2 \end{bmatrix}.$$

У вигляді узагальненого параметра обирається перша головна компонента, для якої дисперсія σ_{Ui}^2 найбільша. Якщо $\sigma_{U1}^2 > \sigma_{U2}^2$, стиснення даних здійснюється шляхом відкидання компоненти з меншою дисперсією:

$$\hat{U} = \frac{1}{\sqrt{2}}\begin{bmatrix} X_1 + X_2 \\ 0 \end{bmatrix}. \qquad (12)$$

Оцінку відновленого вектора X можна подати у такому вигляді:

$$\hat{X} = \Phi' \hat{U} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} . \frac{1}{\sqrt{2}}\begin{bmatrix} X_1 + X_2 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} X_1 + X_2 \\ X_1 + X_2 \end{bmatrix}. \qquad (13)$$

Відносна похибка відновлення вектора X по U визначається виразом:

$$D = \frac{\| \Delta X \|^2}{\| X \|^2} = \frac{\sigma_{U2}^2}{\sigma_{U1}^2 + \sigma_{U2}^2} = \frac{D_X (1 - \rho)}{D_X (1 + \rho + 1 - \rho)} = \frac{1 - \rho}{2}. \qquad (14)$$

Приклад. Для оцінки запропонованих рішень як початкові дані розглядалась вибірка з 15 значень трьох первинних параметрів СІС, розподілених за нормальним законом, що мають однакову дисперсію D та коефіцієнт кореляції ρ=0.8 (табл. 1).

Таблиця 1

*X*1	6,49	6,49	6,49	6,49	6,49	6,49	6,49	6,49	6,49	6,49	6,49	5,88	6,49	5,58	5,58
*X*2	-2,63	-2,35	-2,35	-2,35	-2,35	-2,35	-2,35	-2,35	-2,35	-2,35	-2,35	-2,55	-2,35	-2,36	-2,35
*X*3	-6,57	-6,57	-6,57	-6,83	-6,83	-6,83	-6,83	-6,83	-6,83	-6,83	-7,08	-7,08	-7,08	-7,08	-7,08

Відповідно до виразів (12–13) знаходимо відновлені значення змінних *X*1, *X*2 та *X*3, що наведені у табл. 2.

Таблиця 2

*X*1	6,494	6,494	6,494	6,490	6,490	6,490	6,490	6,490	6,490	6,490	6,486	6,045	6,406	6,576	6,576
*X*2	-2,34	-2,21	-2,21	-2,34	-2,34	-2,34	-2,34	-2,34	-2,34	-2,34	-2,36	-2,36	-2,36	-2,36	-2,36
*X*3	-6,84	-6,70	-6,70	-6,83	-6,83	-6,83	-6,83	-6,83	-6,83	-6,83	-6,95	-6,95	-6,95	-6,95	-6,95

Відносна похибка для теоретичних розрахунків для розглянутого прикладу становить

$$D = \frac{1-\rho}{2} = \frac{1-0{,}8}{2} = 0{,}1 \ ; \xi = \sqrt{D} = 0{,}316 .$$

Значення відносної похибки, що було отримано за результатами моделювання, дорівнює ξ_{X1}=0,135, ξ_{X2}=0,272, ξ_{X3}=0,209 та підтверджується теоретичними розрахунками ($\xi_t = 0{,}316$).

За результатами моделювання було побудовано залежності похибок відновлення векторів X по U за теоретичними (рис. 1) та експериментальними розрахунками (рис. 2).

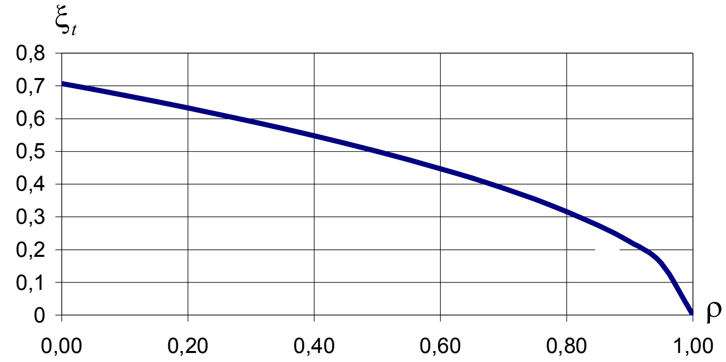

Рис. 1

З аналізу рис. 1 видно, що при кореляції між первинними параметрами відносна похибка відновлення значень первинних параметрів за результатами оцінки узагальненого параметра не перевищує 0,2. В області середньої кореляції при $0.1 < \rho < 0.9$ похибка ξ змінюється в межах від 0.2 до 0.65.

На рис. 2 подано залежність відносної похибки відновлення векторів X за значеннями головної компоненти U за експериментальними розрахунками. У зв'язку з тим, що кореляція між первинними параметрами – величина ймовірна, то відносна похибка відновлення досягає свого найбільшого (теоретичного) значення від значення коефіцієнта кореляції лише у виняткових випадках.

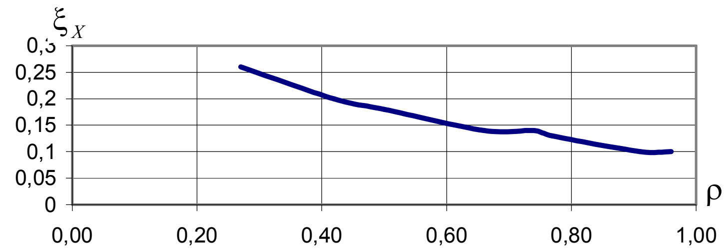

Рис. 2

Висновки

1. Запропоноване рішення дозволяє розв'язувати задачі зменшення вектора первинних параметрів системи контролю СІС, особливо при тісному зв'язку між початковими даними при $\rho > 0.9$ та однакових похибках однотипних засобів вимірювань. Це дає можливість зменшити інформаційну надмірність при проектуванні систем контролю СІС.

2. Недоліком є те, що використовуваний МГК не інваріантний відносно масштабу шкал вимірювань первинних змінних. Тому в таких випадках перспективним є перехід до стандартизованих за середньоквадратичним відхиленням параметрів.

СПИСОК ЛІТЕРАТУРИ

1. Кудрицкий В. Д. Автоматизация контроля РЭА / В. Д. Кудрицкий, М. А. Синица, П. И. Чинаев. – М. : Советское радио, 1977. – С. 65–83.

2. Надежность и эффективность в технике : справочник в 10 т. Т.9: Техническая диагностика; под ред. В. В. Клюева, П. П. Пархоменко. – М. : Машиностроение, 1987. – 350 с.
3. Коваленко М. В. Застосування МГВА для вибору узагальнених параметрів при синтезі систем контролю складних інформаційних систем / М. В. Коваленко, В. В. Воротніков // Вісник ЖІТІ. – Житомир, 1998. – № 9. – С. 110–116.
4. Пат. Україна, МПК G06F17/00,G06F19/00. Спосіб діагностування технічних об'єктів за допомогою узагальнених параметрів / Воротніков В. В., Бобунов А. І., Пашковський В. В. – № 56874А ; 15.05.03, Бюл. № 5.
5. Афифи А. Статистический анализ. Подход с использованием ЭВМ / А. Афифи, С. Эйзен. – М. : Мир, 1982. – 488 с.
6. Воротніков В. В. Алгоритм оптимізації допускової області первинних параметрів складної інформаційної системи / В. В. Воротніков // Збірник наукових праць КВІТІ. – К., 2003. – № 5. – С. 36–45.
7. Ахмед Н. Ортогональные преобразования при обработке цифровых сигналов / Н. Ахмед, К. Р. Рао ; пер. с англ. – М. : Связь, 1980. – 248 с.
8. Рейклетис Г. Оптимизация в технике : В 2-ух кн. Кн. 1 / Г. Рейклетис, А. Рейвиндран, К. Регсдел ; пер. с англ. – М. : Мир, 1986. – 352 с.
9. Базара М. Нелинейное программирование: Теория и алгоритмы / М. Базара, К. Шести. – М. : Мир, 1982. – 354 с.

Подано 25.08.08

Н. В. Коваленко, В. В. Воротников

ФОРМИРОВАНИЕ ВЕКТОРА ПЕРВИЧНЫХ ЗНАЧИМЫХ ПАРАМЕТРОВДЛЯ ОЦЕНКИ ТЕХНИЧЕСКОГО СОСТОЯНИЯ СЛОЖНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

В статье предложено решение задачи выбора обобщенных параметров для проверки работоспособности сложного технического объекта на этапе проектирования его системы контроля. Группирование пространства признаков проводится с помощью метода главных компонент (МГК). Приведены преимущества применения МГК для задач установления технического состояния с учетом определенных корреляционных связей между первичными параметрами.

M. V. Kovalenko, V. V. Vorotnikov

THE FORMING OF VECTOR PRIMARY MEASURES FOR VALUATION OF TECHNICAL CONDITION OF COMPOUND INFORMATION SYSTEM

The article gives problem solving of choosing general measures for control of the work of compound technical object of the level of projecting its control system. The features are conducting with the help of Method Main Components (M.M.C.). The advantages of using M.M.C. for tasks of determination of technical condition due to the determined connections between primary measures.