

В.В. Воротніков, М.В. Коваленко

### ЗАСТОСУВАННЯ МЕТОДУ ГРУПОВОГО ВРАХУВАННЯ АРГУМЕНТІВ ДЛЯ ВИБОРУ УЗАГАЛЬНЕНИХ ПАРАМЕТРІВ

*Запропоновано використання методу групового врахування аргументів для вибору узагальнених параметрів при синтезі систем контролю складних інформаційних систем. Наведені переваги застосування методів теорії об'єктивного системного аналізу для моделювання складних систем в умовах обмеженості інформації.*

При синтезі системи контролю [1, 2, 3] характеристики визначаються насамперед вимогами точності і достовірності контролю  $D$ . Визначимося, з яких міркувань пред'являються вимоги до точності і достовірності контролю.

Однією з найважливіших задач планування або синтезу системи контролю є вибір раціональної або оптимальної точності вимірювань. Збільшення допустимих похибок вимірювань, зменшуючи вартість засобів вимірювання системи контролю, збільшує вірогідності помилок при контролі, що в кінцевому випадку збільшує собівартість об'єкта контролю. Очевидно, що існує оптимальна точність, яка відповідає мінімуму суми вартості збитків від браку і вартості контролю.

Наступною важливою задачею планування контролю є вибір правила прийняття рішення за результатами вимірювань.

Виходячи з вимог, які пред'являються до системи контролю, сформулюємо задачу синтезу системи контролю. Постановка задачі синтезу системи контролю зводиться до пошуку оптимальної кількості контрольованих параметрів, контроль яких дасть нам інформацію про технічний стан об'єкта контролю з достовірністю  $D \geq D_{\text{в}}$  — з одного боку, і вартості системи контролю, яка побудована за цією кількістю контрольованих параметрів, буде обмежена  $C_{\text{max}}$  — з другого. Але із збільшенням кількості контрольованих параметрів водночас необмежено збільшується вартість і час контролю. В результаті може виникнути, що незначне збільшення достовірності пов'язано з різким погіршенням його вартісних і часових показників. Іншими словами, поставлена задача побудови системи контролю передбачає рішення оптимізаційної задачі за критерієм "складність-вартість". Ефективність такої системи контролю залежить від вибору оптимальної кількості контрольованих параметрів.

В межах загальної постановки задачі, коли основний показник якості контролю —  $\beta$  (помилка II роду), глибина контролю буде оптимальною, коли безпомилковий контроль вибраної сукупності  $m \leq n$  параметрів (де  $n$  — загальна кількість, а  $m$  — вибрана кількість параметрів) забезпечує

$$\beta_k(m) \leq \beta_k.\text{доп} \text{ при } C \rightarrow \min, \quad (1)$$

де  $\beta_{\text{доп}}$  — максимально допустима помилка II роду процедури контролю,  $C$  — вартість операцій контролю.

Для рішення задачі синтезу системи контролю або удосконалення реально існуючої системи інформаційних систем задамося вихідними даними.

Нехай задана вихідна повна сукупність параметрів (для удосконалення системи контролю інформаційної системи — це сукупність контрольованих параметрів):

$$X[n] = \{X_1, X_2, \dots, X_n\} \quad (2)$$

з загальною щільністю розподілу  $F_n(x)$  і допусковою областю  $S_p(n)$ , яка визначає працездатність об'єкта контролю в просторі станів вектора технічних станів об'єкта контролю [2].

Помилка II роду з'являється тільки за умови, що всі  $m$  контрольованих параметрів знаходяться в межах допускової області, а хоча б один з  $n - m$  неконтрольованих (а у випадку удосконалення системи контролю інформаційних систем — неврахованих) — за її межами, тобто простору з  $m$  контрольованих параметрів відповідає допускова область  $S_p(m)$  і функція щільності розподілу контрольованих параметрів  $f_m(x)$ , а простору з  $(n - m)$  неконтрольованих — допускова область  $S_p(n - m)$  і функція щільності розподілу  $f_{n-m}(x)$ .

Тоді вірогідність того, що хоча б один з неконтрольованих параметрів вийде за межі допуску, якщо  $X[m] \in S_m$ , запишемо як:

$$\beta_K(m) = 1 - \iiint_{S_p(m)} f_{n-m}(x(n-m) / x(m) \in S_p dx(n-m) \quad (3)$$

Задача синтезу системи контролю у загальному вигляді передбачає її рішення з урахуванням реально існуючих кореляційних зв'язків між контрольованими параметрами.

Практика показує, що більшість складностей, які з'являються при організації контролю працездатності багатовимірного об'єкта контролю, є наслідком стохастичних зв'язків об'єктивно існуючих між його фізичними параметрами. Нехтування цими об'єктивно існуючими зв'язками між параметрами неминуче призводить до помилок.

Одним з шляхів вирішення задачі вибору необхідної кількості контрольованих параметрів для багатовимірного об'єкта із збереженням допустимої точності процесу контролю є перехід від описуючого стану об'єкта із залежними складовими вектора  $X = \{X_1, X_2, \dots, X_n\}$  до деякого іншого вектора  $U = \{U_1, \dots, U_n\}$ , адекватно описуючого стан об'єкта контролю, відносно складових якого припущення про незалежність контрольованих параметрів буде в певній мірі вправдано.

Суть запропонованого способу вибору узагальнених параметрів зводиться до наступного.

Вводиться вимога некорельованості складових вектора  $U_n$ . Оскільки кореляційний зв'язок між параметрами звично буває найбільш сильним (в порівнянні із зв'язками більш високих порядків), гарантія про її відсутність робить припущення про незалежність випадкових величин  $U_j, j = 1, 2, 3, \dots, n$  значно більш суворим. У випадку нормального розподілу вектора  $X_n$  незалежність безпосередньо витікає з некорельованості.

Для початку здійснимо перетворення вибірки значень вихідних, у загальному випадку, залежних перемінних в масив узагальнених незалежних перемінних  $U_j, j = 1, \dots, n$ , які мають властивості

$$\begin{aligned} M[U_j] &= 0; \\ M[U_i * U_j] &= 0, i \neq j; \\ M[U_j * U_j] &= D_j, j = 1, 2, 3, \dots, n \end{aligned} \quad (4)$$

де  $M[U_j]$  – математичне очікування значень  $j$ -ої перемінної, а  $D_j$  – її дисперсія,  $M[U_i * U_j]$  – математичне очікування значень добутку двох будь-яких узагальнених параметрів.

Треба підкреслити, що описування за допомогою вектора первинних параметрів  $X_i, i = 1, 2, 3, \dots, m$ , і вектора узагальнених параметрів  $U_j, j = 1, 2, 3, \dots, n$  цілком адекватні.

Приведення вектора  $X_i$  до вектора  $U_j$  з некорельованими складовими може здійснюватися будь-яким із відомих методів ортогоналізації [1].

Враховуючи, що масив вимірюваних первинних параметрів – сукупність досить різних перемінних (це обумовлено різною топологією первинних параметрів, а також різницею в розмірності), переведемо перемінні  $X_i$  з натуральних первинних перемінних в кодовані  $X'_i$  з обмеженням:  $X'_i \leq 1$ :

$$\begin{aligned} X_i &= (X_i - X_{i0}) / \Delta X_i, \\ X_{i0} &= (X_{i \max} + X_{i \min}) / 2, \\ i &= 1, 2, 3, \dots, n \end{aligned} \quad (5)$$

де  $X_{i0}$  – вихідна точка  $i$ -ого первинного параметра;  $\Delta X_i$  – інтервал варіювання  $i$ -ого первинного параметра, обумовлений дискретністю АЦП.

Для вибору узагальнених параметрів пропонується застосувати метод теорії об'єктивного системного аналізу, відомого під назвою методу групового врахування аргументів (МГВА) [4].

МГВА – це метод математичного моделювання, заснований на так званому принципі самоорганізації моделей на ЕОМ. Згідно з цим принципом здійснюється цілеспрямований перебір багатьох моделей-претендентів різної складності за рядом критеріїв. В результаті знаходиться модель оптимальної структури у вигляді одного або системи рівнянь.

Принцип самоорганізації математичних моделей складних систем дозволяє визначити протирічний характер рівнянь моделі при їх перевірці на різних вибірках даних. Непротиричною є модель, яка, будучи отриманою щонайменш на двох вибірках, дає на всіх даних (на сумарній

виборці) майже один і той же результат ідентифікації. Якщо модель, отримана на одній частині експериментальних даних, дає один результат, а отримана на другій – цілком інший, то така модель є протирічною і вірити їй не можна. Можна пропонувати багато різних зовнішніх критеріїв самоорганізації моделей, виходячи з різних вимог.

В МГВА основним зовнішнім критерієм є критерій непротиричності моделей. Цей критерій вимагає, щоб моделі задовольняли вказаній раніше властивості непротиричності. Вимога непротиричності може бути сформульована за допомогою різних математичних виразів. Найбільш розповсюджена наступна формула, названа критерієм мінімуму зміщення:

$$N_3 M^2 = \frac{\sum_{i=1}^n (Y_a - Y_b)^2}{\sum_{i=1}^n Y_{\phi}^2} \rightarrow \min, \quad (6)$$

де  $Y_a$  – вихідна величина моделі, оцінки коефіцієнтів якої отримані на підвибірці А;  $Y_b$  – те саме, але на підвибірці В;  $Y_{\phi}$  – фактичні дані вибірки,  $n$  – кількість вимірювань первинних параметрів.

Повернувшись до рішення задачі, розглянемо застосування самоорганізованих моделей для вибору узагальнених параметрів. Для цього визначимося, що між кодованими первинними елементами  $X_i$  і кінцевою характеристикою  $Y$  об'єкта контролю існує об'єктивний взаємозв'язок:

$$Y = \varphi(x_1, x_2, \dots, x_n), \quad (7)$$

$i = 1, 2, 3, \dots, n$

отже, буде існувати і об'єктивний взаємозв'язок між первинними параметрами  $X_i$  і узагальненими параметрами  $U_j$ :

$$U_j = \eta(x_1, x_2, \dots, x_n), \quad (8)$$

$j = 1, 2, 3, \dots, m, m \leq n$

Можна припустити, що функція відгуку  $\eta$  належить до класу алгебраїчних поліномів (так званий повний степінний поліном Колмогорова-Габора) степеня  $q$  від  $n$  перемінних:

$$\eta = \beta_0 + \sum_{i=1}^n \beta_i^* x_i + \sum_{i=1}^n \sum_{j=1}^n \beta_{ij}^* \beta_{j^*}^* x_i^* x_j + \dots \quad (9)$$

Доцільність такого припущення обумовлена двома основними передумовами: по-перше – загальністю моделі, за допомогою якої можна описати достатньо широкий клас залежностей;

по-друге – простотою моделі, що дає можливість її широкого використання.

При використанні ортогональності  $X_i^*$  вираз (9) приймає вигляд:

$$\eta = \beta'_0 + \sum_{i=1}^n \beta'_i x_i + \sum_{i=1}^n \sum_{j=1}^n \beta'_i \beta'_j x_i^* x_j + \dots \quad (10)$$

В матричному позначенні вираз (10) представляється у вигляді:

$$\eta = f^T(X) \beta', \quad (11)$$

де  $f^T(x)$  – відомий вектор перемінних,  $\beta'$  – вектор невідомих коефіцієнтів.

Об'єднання групи первинних перемінних  $X_i$  в один з узагальнених  $U_j$  засновано на наступному. Вибірка даних  $X_i$  розбивається на дві підвибірки: навчальну А і перевіірочну В.

Згідно з [5] припустимо:

$$N_a = N_b = N / 2, \quad (12)$$

де  $N_a, N_b$  – кількість вимірювань первинних параметрів, які потрапили до навчальної і перевіірочної підвибірки відповідно, а  $N$  – загальна кількість вимірювань (рядок в масиві).

Пошук моделей узагальнених параметрів проводимо за критерієм незміщення (6). Для цього коефіцієнти кожної моделі, розраховані на навчальній підвибірці А, зрівнюються з коефіцієнтами такої ж за структурою моделі розрахованими на перевіірочній підвибірці В. У випадку,

якщо величина різниці коефіцієнтів, а значить і різниця значень узагальнених параметрів, які характеризують поведінку групи первинних перемінних, менше критерію незміщення  $N_{зм}^2$  – задача вибору першого узагальненого параметра вирішена. Для того, щоб перейти до пошуку наступного узагальненого параметра, з вибірки даних виключаються стовпці значень первинних параметрів, які потрапили в попередній узагальнений. Після чого процес відбору повторюється. Параметри, які не потрапили в узагальнені, виключаються як надмірні. Таким чином, синтез системи контролю на цьому етапі закінчується.

Для оцінки коефіцієнтів узагальнених параметрів використовується метод найменших квадратів (МНК), тому що МНК-оцінка є незміщеною і має мінімальну дисперсію.

МНК-оцінки  $\hat{\beta}$  коефіцієнтів  $\beta$  знаходяться за виразом:

$$\hat{\beta} = (X^T * X)^{-1} * X^T * Y, \tag{13}$$

а коваріаційна матриця дисперсій  $D(\hat{\beta})$  вектора оцінок  $\beta$  дорівнює:

$$\begin{aligned} D(\hat{\beta}) &= (X^T * X)^{-1} * D\{Y\} * X * (X^T * X)^{-1} = \\ &= \sigma^2 (X^T * X)^{-1} * X^T * In * X * (X^T * X)^{-1} = \sigma^2 (X^T * X)^{-1} \end{aligned} \tag{14}$$

Функція відгуку (2.16), яка задовольняє лінійній моделі, є параметричною і допускає МНК- оцінку, яка у відповідності до теореми Гаусса-Маркова дорівнює:

$$\hat{\eta} = f^T(X) * \hat{\beta} = \sum_{i=1}^n \hat{\beta}_i * X_i. \tag{15}$$

Знаходженням матриці коефіцієнтів моделі узагальнених параметрів і коваріаційної матриці дисперсій вектора оцінок  $\hat{\beta}$  закінчується етап синтезу системи контролю.

На підставі викладених теоретичних положень відносно вибору узагальнених параметрів при синтезі систем контролю інформаційних систем розроблений алгоритм вибору узагальнених параметрів і програмно реалізований на ПЭОМ ІВМ АТ 486DX.

Використання МГВА розглядалось на наступному прикладі.

Нехай є вибірка значень первинних параметрів  $X_1(t)$ ,  $X_2(t)$ ,  $X_3(t)$ ,  $X_4(t)$  (табл.), на які накладаються помилки вимірювань, розподілені за нормальним законом. Стовпці  $X_5(t)$ ,  $X_6(t)$ ...,  $X_{10}(t)$  – це значення додатків полінома першого степеня, за допомогою яких будуть визначатися моделі узагальнених параметрів.

Таблиця

	X1(t)	X2(t)	X3(t)	X4(t)	X5(t)	X6(t)	X7(t)	X8(t)	X9(t)	X10(t)
t1	2.46	2.90	15.62	42.08	36.13	194.63	524.32	45.30	122.03	657.29
t2	12.80	3.08	19.74	38.95	39.42	252.67	498.56	60.80	119.97	768.87
t3	12.01	2.65	17.03	39.03	31.83	204.53	468.75	45.13	103.43	664.68
t4	10.90	2.81	17.59	40.07	30.63	191.73	436.76	49.43	112.60	704.83
t5	11.46	3.90	15.62	42.08	44.69	179.01	482.24	60.92	164.1	657.29
t6	12.40	3.58	19.74	38.95	44.39	244.78	482.98	70.67	139.44	768.87
t7	12.91	2.95	17.03	39.03	38.08	219.86	503.88	50.24	115.1	664.68
t8	10.06	3.01	17.59	40.07	30.28	176.96	403.10	52.95	120.61	704.83
t9	12.56	2.99	15.62	42.08	37.55	196.19	528.52	46.70	125.82	657.29
t10	11.80	3.48	19.74	38.95	41.06	232.93	459.61	68.70	135.55	768.87
t11	10.01	2.25	17.03	39.03	22.52	170.47	390.69	38.32	97.82	664.68
t12	11.90	3.10	17.59	40.07	36.89	209.32	476.83	54.53	124.22	704.83

Кінцева величина  $Y(t)$ , яка раніш невідома, розраховується за формулою:

$$Y(t) = 0.65 * X_1(t) + 0.92 * X_2(t) - 0.95 * X_1(t) * X_3(t). \tag{16}$$

Необхідно вибрати узагальнений параметр, що адекватно описує поведінку об'єкта контролю.

Помітимо, що у визначенні кінцевої величини  $Y(t)$  бере участь доданок  $X_1(t) * X_3(t)$ , який у явному вигляді не вимірюється. Мовою теорії об'єктивного системного аналізу це означає, що існуючий взаємозв'язок між первинними параметрами  $X_1(t)$  і  $X_3(t)$  здійснює прямий вплив на кінцеву функцію. Згідно із застосуванням методу МГВА для вибору узагальнених параметрів,

першим кроком необхідно розширити масив вимірюваних первинних перемінних у відповідності до вибраної опорної функції.

Потім для того, щоб спростити розрахунок вірогідності справного стану об'єкта контролю, позбудемося від взаємозв'язку між первинними перемінними. Для цього використаємо будь-який з відомих методів ортогоналізації. Представлення первинних параметрів у такому вигляді дає можливість розглядати їх як незалежні.

Проведений відбір моделей узагальнених параметрів дав наступні результати.

По розширеній вибірці первинних параметрів відібрані за критерієм мінімуму зміщення 3 узагальнені параметри, в які потрапили наступні первинні перемінні:

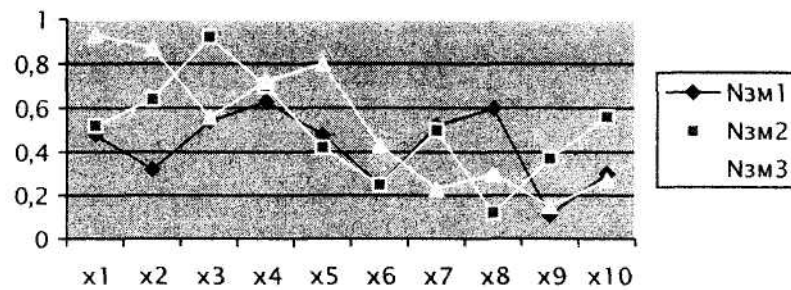
1-ий узагальнений параметр складається з 1, 2, 6, 9 первинних перемінних;

2-ий узагальнений параметр складається з 5, 6, 8 первинних перемінних;

3-ій узагальнений параметр складається з 6, 7, 9 первинних перемінних.

На рисунку зображені зміни критерію зміщення відібраних моделей трьох узагальнених параметрів. По осі абсцис відкладені первинні перемінні, по осі ординат – значення коефіцієнтів зміщення.

Малюнок



Необхідно помітити, що функція залежності коефіцієнта зміщення моделі узагальненого параметра від відбираємої первинної перемінної, яка представлена на рисунку функціями Nзм<sub>1</sub>, Nзм<sub>2</sub>, Nзм<sub>3</sub>, проходить через геометричний мінімум. Місце геометричного мінімуму відповідає відібраній структурі моделі узагальненого параметра.

Застосування МГВА для відбору первинних перемінних в один або декілька узагальнених параметрів дає ряд переваг в порівнянні з відовими методами, викладеними в [1, 2].

У випадку синтезу системи контролю складної інформаційної системи, коли приходится контролювати велику кількість перемінних, при побудові системи контролю стикаються із значними складностями.

По-перше, суб'єктивний характер, який вноситься обмеженими знаннями інженерів у процес синтезу системи контролю, приводить до побудови алгоритмів, які застосовують виключно точнісні та інтегральні критерії. У результаті знаходяться моделі узагальнених параметрів, найточніші по заданій вибірці даних. Характерним є довіра до однієї вибірки. Проте, апріорно невідомо, що нові дані, які надходять незабаром після заданої вибірки, приведуть до іншої моделі. Звідси і з'являються задачі самоорганізації (у нашому випадку вибору) деякої усередненої моделі, яка ураховує таку апріорну інформацію.

Якщо вибірка даних мала (10–20 вимірювань), кращі результати забезпечує "робастний" підхід. Поняття "робастний" звично застосовується до методів, нечутливих до розподілу перешкод. В МГВА цьому поняттю надається більш широкий сенс: робастний розуміється як малочутливий до зміни підвибірки даних. Моделі, які отримані за критерієм непротирічності, більш прості і стабільні на наступних підвибірках даних. Справа в тому, що непротирічність мовби зменшує точність на заданій малій вибірці, але у визначених умовах забезпечує більш вищу точність, розраховану на деяких майбутніх вибірках даних, які, звичайно, відсутні до моменту моделювання.

По-друге, нехтування об'єктивно існуючими стохастичними зв'язками між фізичними параметрами об'єкта контролю неминуче призводить до помилок. Відомі методики вибору узагальнених параметрів (1, 6) обмежуються вимогами некорельованості складових вектора узагальне-

них параметрів. Однак, апарат парних кореляційних функцій, які використовуються для вибору найбільш ефективних аргументів, має наступні недоліки:

- застосовується тільки для лінійних залежностей;
- дає тільки окремі значення кореляційного аналізу.

Більш удосконаленим є апарат множинної кореляції або лінійного регресійного аналізу, вирішуемого МГВА. Застосування цього апарату дозволяє визначити вплив стохастичних зв'язків більш вищих порядків, існуючих між первинними параметрами на узагальнений параметр, в який вони включені, або на об'єкт контролю цілком.

В результаті застосування цього методу ми отримали групу з  $n$  перемінних, адекватно описуючих поведінку об'єкта контролю, за допомогою формування узагальненого параметра, який дозволяє визначити технічний стан об'єкта контролю з достовірністю  $D$ .

Побудову системи контролю можна вважати закінченою, якщо контроль  $m < n$  числа узагальнених параметрів дозволяє характеризувати технічний стан об'єкта контролю з  $D \geq D_{\text{в}}$ . Якщо ця вимога не виконується, необхідно продовжити визначення наступного  $(m + 1)$  узагальненого параметра за таким же принципом (однак параметри, які потрапили в попередній узагальнений параметр, у визначенні наступного не беруть участь).

#### ЛІТЕРАТУРА:

1. Кудрицкий В.Д., Сеница М.А., Чинаев П.И. Автоматизация контроля РЭА. -- 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. Пашковский Г.С. Задачи оптимального обнаружения и поиска отказов в РЭА. -- М.: Радио и связь, 1981. -- 288 с.
4. Юрачковский Ю.П. Восстановление полиномиальных зависимостей на основе самоорганизации // Автоматика. -- 1985. -- № 5. -- С. 39–41.
5. Ивахненко А.Г., Юрачковский Ю.П. Моделирование сложных систем по экспериментальным данным. -- М.: Радио и связь, 1987. -- С. 76–113.
6. Годлевский В.С., Заварин А.Н. Многоэтапные методы контроля промышленной продукции // Метрология. -- 1978. -- № 1. -- С. 31–37.

ВОРОТНИКОВ Володимир Володимирович – помічник начальника науково-дослідного та редакційно-видавничого відділу Житомирського військового інституту радіоелектроніки.

Наукові інтереси:

- моделювання і оцінка технічного стану складних інформаційних систем.

КОВАЛЕНКО Микола Вікторович – доктор технічних наук, професор Житомирського інженерно-технологічного інституту.

Наукові інтереси:

- радіотехніка.