УДК 519.2

А. А. БОСОВ, О. В. ГОРБОВА (ДІІТ)

ПОБУДОВА МАТЕМАТИЧНОЇ МОДЕЛІ НА ОСНОВІ ПАСИВНОГО ЕКСПЕРИМЕНТУ

Запропоновано алгоритм побудови математичних моделей на основі пасивного експерименту.
Ключові слова: алгоритм побудови, математичні моделі, пасивний експеримент

Предложен алгоритм построения математических моделей на основе пассивного эксперимента.
Ключевые слова: алгоритм построения, математические модели, пассивный эксперимент

An algorithm for constructing the mathematical models based on a passive experiment is proposed.
Keywords: algorithm for constructing, mathematical models, passive experiment

Математичне моделювання стає одним із головних інструментів дослідження структур складних систем, визначення засобів активного впливу на них, вивчення впровадження систем майбутнього, особливо тоді, коли система є представником великої кількості взаємозалежних елементів. Інформація про ці елементи складається з результатів пасивного експерименту у вигляді матриці спостережень. У даному випадку проблема моделювання полягає у структурній невизначеності системи, коли взаємозв'язки між елементами невідомі, що викликає нові ускладнення у виборі незалежних змінних, котрі можуть бути предикторами для системи. Вибір кращих предикторів, за певними критеріями, реалізується у процесі розв'язання задачі параметричної ідентифікації, що приводить до надзвичайного перебору варіантів математичних моделей, навіть невеликій розмірності вхідних матриць [1].

У роботі А. Лебєга «Інтегрування та пошук примітивних функцій» наводиться поняття «співіснуючих величин», яке було введено Коші.

Визначення: Величини називаються співіснуючими, якщо вони визначені одними й тими ж умовами, геометричними або фізичними, або соціально-економічними.

В математичному плані співіснуючі величини являють собою деякі функції множини. Було помічено, що якщо величини піддаються безпосередньому вимірюванню, то завжди виявляється, що вони є функціями множин.

Нехай x_i, $i = \overline{1,n}$, набір показників, що характеризують деякий об'єкт (станцію), тоді згідно ідеї Коші, даному показнику x_i, відповідає деяка множина $A(x_i)$ властивостей об'єкт, що розглядається.

Тоді якщо

$$A(x_i) \cap A(x_j) \neq \varnothing,$$

то можна стверджувати, що між x_i та x_j існує деякий зв'язок.

Найпростішим варіантом кількісної оцінки зв'язку є коефіцієнт кореляції між цими показниками.

Побудова класів толерантності дозволяє визначати властивості об'єкту, а пошук класів еквівалентності, дозволяє визначати незалежні процеси, що протікають у об'єкті.

Розглянемо алгоритм визначення кращих предикторів.

Вхідними даними для алгоритму побудови є матриця даних, що була отримана методом пасивного експерименту X, що складається з m – кількість експериментів та n – кількість показників за якими буде виконуватись математичне моделювання.

$$X = \begin{pmatrix} x_{1,1} & x_{2,1} & \dots & x_{n,1} \\ x_{1,2} & x_{2,2} & \dots & x_{n,2} \\ \dots & \dots & \dots & \dots \\ x_{1,m} & x_{2,m} & \dots & x_{n,m} \end{pmatrix}.$$

Алгоритм побудови пошуку наборів предикторних змінних:

п. 1: Розрахувати матрицю кореляції ($R = (r_{i,j})$):

$$r_{i,j} = \frac{\overline{x_i \cdot x_j} - \overline{x}_i \cdot \overline{x}_j}{S_i \cdot S_j},$$

де
$$\overline{x}_i = \frac{1}{(m-1)} \sum_{k=1}^{m} x_{i,k} S_i,$$

$$\overline{x_i x_j} = \frac{1}{(m-1)} \sum_{k=1}^{m} x_{i,k} x_{j,k},$$

$$S_i = \sqrt{\frac{1}{(m-1)}\sum (x_{i,k} - \overline{x}_i)^2}\ .$$

п. 2: Виконати аналіз матриці R з позиції лінійно залежних елементів;

п. 3: Ввести значення довірчої ймовірності та розрахувати матрицю толерантності ($T = (\tau_{i,j})$):

$$\tau_{i,j} = \begin{cases} 1, \left|r_{i,j}\right| > r_{кр}(\alpha) \\ 0, \left|r_{i,j}\right| \le r_{кр}(\alpha) \end{cases}.$$

п. 4: Розрахувати класи толерантності за допомогою алгоритму М. Е. Пиша [2].

п. 5: Визначити класи еквівалентності за допомогою матриці толерантності

п. 6: На основі класів еквівалентності знайти набори предикторних змінних

п. 7: Спираючись на розраховану матрицю толерантності, звести дані до розв'язання задачі векторної оптимізації та за допомогою методу найменших квадратів знайти коефіцієнти при не нульових значеннях у матриці толерантності.

п. 8: Розрахувати нові значення матриці X за допомогою значень, отриманих у п.7 та підрахуємо середню абсолютну-відносну похибку кожного із показників.

п. 9: Розрахувати нові значення матриці X за допомогою значень, отриманих у п.5 та підрахуємо середню абсолютну-відносну похибку кожного із показників.

п. 10: На основі даних, отриманих у п. 8 обрати той клас предикторних змінних, де значення похибки є найменшою величиною.

Описаний алгоритм дає можливість автоматизації пошуку класів толерантності, еквівалентності та наборів предикторних змінних. Що дає можливість у автоматизації побудови математичної моделі відповідного процесу.

Згідно описаного вище алгоритму була розроблена програма, що виконує розрахунок відповідних матриць та пошук відповідних даних, спираючись на дані, отримані шляхом пасивного експерименту.

Розглянемо роботу програми на наборі статистичних даних станції Джанкой (Придніпровської залізниці) за період 1900-2008 р.р.. Нехай є вісім показників роботи цієї станції вказаний період[3]:

1. переробка вагонів на горки в середньому за добу;
2. оправлено вагонів в середньому за добу;
3. оправлення поїздів в середньому за добу;
4. великовагових поїздів;
5. довжина складених поїздів;
6. транзит з роботою (фактичний), год.;
7. транзит без роботи (фактичний), год.;
8. простой під однією вантажною операцією (фактичний), год.

Тоді маємо матрицю вхідних даних X (19x8):

$$X = \begin{pmatrix}
2910 & 3418 & 66 & 11441 & 3295 & 6{,}5 & 1 & 19{,}4 \\
2843 & 3203 & 61 & 10924 & 3088 & 6{,}4 & 1 & 21{,}7 \\
2549 & 3194 & 53 & 9625 & 2142 & 6{,}4 & 1{,}1 & 21{,}7 \\
1999 & 2253 & 44 & 8039 & 1278 & 6{,}8 & 1{,}4 & 26{,}2 \\
1179 & 1449 & 29 & 5317 & 480 & 8{,}9 & 1{,}6 & 23{,}6 \\
1040 & 1200 & 23 & 4633 & 546 & 9{,}1 & 1{,}8 & 17{,}1 \\
828 & 936 & 19 & 3949 & 473 & 10{,}8 & 2{,}2 & 16{,}8 \\
712 & 767 & 14 & 3330 & 535 & 12{,}1 & 2{,}3 & 15{,}4 \\
803 & 904 & 15 & 3270 & 791 & 11 & 2{,}2 & 13{,}9 \\
919 & 1098 & 17 & 3331 & 992 & 13{,}1 & 3{,}3 & 16{,}7 \\
1007 & 1206 & 18 & 3556 & 1028 & 11{,}1 & 3{,}9 & 22 \\
954 & 1174 & 18 & 3377 & 1231 & 5{,}4 & 3 & 25{,}5 \\
1214 & 1569 & 25 & 3623 & 2341 & 10{,}1 & 4{,}9 & 28{,}2 \\
1509 & 2136 & 34 & 4800 & 3190 & 8{,}8 & 3{,}8 & 33{,}7 \\
1437 & 1695 & 27 & 3387 & 3140 & 8{,}18 & 3{,}59 & 24{,}3 \\
1149 & 1329 & 21 & 2865 & 2368 & 7{,}1 & 1{,}76 & 19{,}8 \\
1274 & 1376 & 22 & 3292 & 2229 & 1{,}43 & 1{,}43 & 19{,}45 \\
1501 & 1566 & 24 & 4392 & 1868 & 9{,}62 & 0{,}87 & 25{,}11 \\
1650 & 1720 & 26 & 3892 & 3019 & 10{,}57 & 0{,}87 & 31{,}58
\end{pmatrix}$$

Згідно алгоритму далі, за допомогою програми, знайдемо матрицю кореляції R:

$$R = \begin{pmatrix} 1,00 & 0,99 & 0,98 & 0,93 & 0,66 & -0,46 & -0,49 & 0,25 \\ 0,99 & 1,00 & 0,98 & 0,93 & 0,67 & -0,45 & -0,39 & 0,29 \\ 0,98 & 0,98 & 1,00 & 0,97 & 0,57 & -0,44 & -0,44 & 0,19 \\ 0,93 & 0,93 & 0,97 & 1,00 & 0,38 & -0,36 & -0,50 & 0,04 \\ 0,66 & 0,67 & 0,57 & 0,38 & 1,00 & -0,41 & -0,07 & 0,55 \\ -0,46 & -0,45 & -0,44 & -0,36 & -0,41 & 1,00 & 0,36 & -0,14 \\ -0,49 & -0,39 & -0,44 & -0,50 & -0,07 & 0,36 & 1,00 & 0,19 \\ 0,25 & 0,29 & 0,19 & 0,04 & 0,55 & -0,14 & 0,19 & 1,00 \end{pmatrix}$$

Після виконання аналізу з позиції лінійної залежності отримуємо матрицю, що дорівнює матриці кореляції, тому переходимо до розрахунку матриці толерантності, при довірчій ймовірності 0,95 та $r_{кр}$=0,5.

$$\tau = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Спираючись на значення, отримані з матриці толерантності, та використовуючи алгоритм М. Е. Пиша отримуємо класи толерантності.

{x6}

{x5 x8}

{x4 x7}

{x1 x2 x3 x5}

{x1 x2 x3 x4}

Отримуємо п'ять класів толерантності та робимо висновок, що в даному об'єкті (станція) можна відзначити не менше п'яти властивостей, по числу класів толерантності. Інженерно - економічне тлумачення даних властивостей виходить за рамки даної статті та представляє задачу предметного розгляду відповідних спеціалістів.

Далі, аналізуючи дані матриці толерантності, визначаємо рефлективність, транзитивність та симетричність даних у матриці. У заданому прикладі всі властивості значень матриці підтвердилися і за допомогою програми було знайдено класи еквівалентності:

{x1, x2, x3, x4, x5, x7, x8}

{x6}

Як видно для заданих даних лише два класи еквівалентності. Спираючись на отримані класи еквівалентності, було виділено відповідні підматриці із матриці толерантності та було знайдено відповідні значення предикторних змінних:

{x5 x7 x8}

{x4 x5 x7}

{x3 x7 x8}

{x2 x7 x8}

{x1 x7 x8}

{x6}

На основі отриманих даних можна побудувати шість математичних моделей для вказаного об'єкту.

Далі використаємо метод найменших квадратів та зведемо задачу до векторної оптимізації та за допомогою методу Гауса отримаємо коефіцієнти для пошуку значень вхідної матриці X. Матриця коефіцієнтів, що отримані за допомогою методу найменших квадратів має вигляд наведений у табл. 1.

Таблиця 1

Матриця коефіцієнтів, що отримані за допомогою методу найменших квадратів

	a0	a1	a2	a3	a4	a5	a6	a7	a8
x1	5.80e-15	0	0.725	6,508	0	0	0	0	0
x2	4.56e-16	0.0297	0	51.1	0	0	0	0	0
x3	-5.99e-16	-4.96e-13	0.016	0	4.77e-13	1.403	0	0	0
x4	-7.10e-14	0	0	182.7	0	0	0	0	0
x5	60	0	0	0	0	0	0	0	0
x6	1.137	0	0	0	0	0	0	0	0
x7	0.275	0	0	0	0	0	0	0	0
x8	1.925	0	0	0	0	0	0	0	0

За допомогою програми виконаємо розрахунок нових значень матриці пасивного експерименту, для визначення абсолютно-відносної похибки розрахунку коефіцієнтів матриці А. Абсолютна – відносна похибка, отримана за допомогою програми для різних *x*, має вигляд:

- для x1, похибка ε_1=0.021;
- для x2, похибка ε_2=0.026;
- для x3, похибка ε_3=0.086;
- для x4, похибка ε_4=0.04;
- для x5, похибка ε_5=2.960;
- для x6, похибка ε_6=0.795;
- для x7, похибка ε_7=0.580;
- для x8, похибка ε_8=1.19

За обчисленою абсолютно-відносною похибкою та знайденими наборами предикторних змінних маємо, що параметри *x1*, *x2, x3* та *x4* приймають найменші значення. Тоді обчислимо суму значень абсолютно-відносних похибок всіх наборів предикторних змінних:

– для { x5 x7 x8} $\rightarrow$ {3.73};
– для { x4 x5 x7} $\rightarrow$ {3.58};
– для { x3 x7 x8} $\rightarrow$ {1.85};
– для { x2 x7 x8} $\rightarrow$ {1.796};
– для { x1 x7 x8} $\rightarrow$ {1.791};
– для { x6} $\rightarrow$ {0.795}.

Тоді з наборів предикторних змінних обираємо набір – {x5 x7 x8}. До елементів цього набору не входять відповідні значення похибок елементів *x1*, *x2, x3* та *x4,*а також сума відносних похибок елементів є найгіршою – {3,73}.

Таким чином пропонується наступні коефіцієнти математичної моделі:

– для x1: a0=5.80e-15, a2 = 0.725, a3=6,508;
– для x2: a0=4.56e-16, a1 = 0.0297, a3=51.1;
– для x3: a0=5.99e-16, a1=-4.96e-13, a2=0.016, a4=4.77e-13, a5=1.403;
– для x4: a0=-7.10e-14, a3=182.7.

БІБЛІОГРАФІЧНИЙ СПИСОК

1. Льюнг, Л. Идентификация систем. Теория для пользователя [Текст] / Л. Льюнг. – М.: Наука, 1991 – 433 с.
2. Лекции по теории графов [Текст] / В. А. Емеличев [и др.]. – М.: Наука, 1990. – 383 с.
3. Босов, А. А. Математичне моделювання раціонального використання ресурсів залізничної станції [Текст] / А. А. Босов, К. В. Єлісєєнко, О. І. Харченко // Вісник Дніпропетр. нац. ун-ту залізн. трансп. ім. акад. В. Лазаряна. – 2009. – Вип. 27. – Д.: Вид-во ДНУЗТ, 2009. – С. 205-209.

Надійшла до редколегії 09.11.2010.
Прийнята до друку 12.11.2010.