ВІЙСЬКОВІ НАУКИ

УДК 623.4.016

ЗАСТОСУВАННЯ МЕТОДУ ПРОГРЕСУЮЧОГО ЕТАЛОНУ ДО АНАЛІЗУ ДАНИХ РІЗНОЇ РОЗМІРНОСТІ

Кузьменко Р.В.
Національна академія сухопутних військ
імені гетьмана Петра Сагайдачного

У статті проведено статистичний аналіз результатів застосування методу прогресуючого еталону. Вивчено розподіл коефіцієнтів цільової функції у залежності від вимірності вихідних даних за їх рівномірного розподілу. Встановлено, що результат застосування методу прогресуючого еталону може бути формалізовано у вигляді вектора параметрів, розмірність якого на одиницю менша від кількості показників якості, також встановлено, що за умови рівномірно-розподіленого випадкового набору даних загальна кількість випадків, коли метод не дозволяє отримати резу-льтат, не перевищує 0,05% за умови наявності трьох еталонів, у випадку збільшення кількості цей показник буде зменшуватись, що в свою чергу вказує на те, що він може бути ефективно застосований для побудови лінійної функції корисності для широкого діапазону вихідних даних.

Ключові слова: рівень технічної досконалості, метод прогресуючого еталона, метод Монте-Карло.

Постановка проблеми та аналіз останніх досліджень і публікацій. Метод прогресуючого еталону (МПЕ) використовується для порівняння якості об'єктів на основі заданого вектору їх технічних характеристик [1, 2]. Вихідними даними виступає послідовність еталонів, характеристики (далі – показники якості – ПЯ) яких відомі, а самі еталони впорядковані таким чином, що кожен наступний еталон вважається кращим за попередній. Результатом застосування методу є побудова функції, що зростає на послідовності еталонів; ця функція у подальшому може бути використана для визначення рівня технічної досконалості однотипних об'єктів: на векторі характеристик більш досконалих об'єктів вона набуватиме більші значення. Розширення МПЕ на випадок, коли інформація про технічні дані окремих об'єктів відсутня або носить інтервальний характер, розглянуто у роботі [3].

Метод прогресуючого еталону дозволяє побудувати лінійну функцію корисності для оцінки і порівняння однотипних об'єктів на основі відомостей про їх технічні показники. Алгоритм методу оцінює тенденцію еволюції показників, враховуючи їхню взаємну кореляцію. Застосування методу неявно передбачає, що у послідовності еталонів існує певна, статистично значуща закономірність, котру можна виявити і описати математично.

Прогрес технічної досконалості однотипних зразків пов'язаний із відповідними змінами кожної із їхніх характеристик. Результат застосування методу до послідовності еталонів, характеристики яких утворюють монотонні послідовності є передбачуваний і очевидний. Чисельні результати дають змогу лише формалізувати виявлені тенденції до зменшення одних показників та збільшення інших. У ідеалізованому випадку міг би існувати єдиний вектор розвитку у просторі технічних характеристик об'єктів, а усі еталони описувались би точками у згаданому просторі, розташованими уздовж прямої, напрям якої задає вектор розвитку. У такому випадку кількість еталонів та їх характеристик не змінювала б принципово картину.

Геометрична інтерпретація методу полягає у знаходженні напрямного вектору прямої у просторі ПЯ еталонів, що проходить максимально близько (в сенсі методу найменших квадратів) від точок-еталонів [1, 2].

Виділення невирішених раніше частин загальної проблеми. З теоретичної точки зору значний інтерес представляє протилежний випадок, коли точки характеристик еталонів не можна вважати такими, що лежать уздовж певної прямої. Розгляд деяких обмежень, коли алгоритм методу не дозволяє однозначно виявити вектор розвитку, наведено у роботі [3]. Питання того, який результат матиме застосування МПЕ для випадкових вихідних даних заданої розмірності, досі детально не вивчалось.

Метою статті є аналізі можливостей застосування методу прогресуючого еталону до даних, що носять «некоректний» характер, не властивий логіці технічного розвитку, котра приводить до росту одних характеристик і зменшення інших.

Основний матеріал. Розглядається задача: наскільки формально застосовний алгоритм МПЕ на послідовності випадкових даних та як залежить результат від розмірності вхідних даних. З цією метою використаємо метод Монте-Карло до аналізу результатів використання МПЕ за різного набору вхідних даних, а саме попередньо провівши класифікацію можливих розв'язків задачі, дослідимо, як впливає кількість еталонів, на основі яких проводимо розрахунок, на результат застосування методу.

Методика дослідження. Чисельні дослідження проведено у програмному пакеті системи комп'ютерної алгебри *Maple*. Випадкову вихідну матрицю отримуємо за допомогою процедури *randmatrix (n,m)*, котра по замовчуванню дозволяє отримати матрицю розміру $n \times m$, елементи

© Кузьменко Р.В., 2016

якої є цілі числа, рівномірно розподілені на проміжку [−99; 99], тобто ймовірність pk того, що елементом матриці є число k $(-99 \le k \le 99)$ рівна 1/199. Для нормалізації даних використовуємо стандартні процедури пошуку екстремальних та середніх значень [4]. Пошук власних значень та власних векторів виконуємо використовуючи процедури *eigenvalues* та *eigenvectors*. Усі обчислення виконуємо із точністю у 10 значущих цифр (змінна *Digits* у програмному пакеті *Maple* рівна 10).

Вихідними даними застосування МПЕ можна вважати матрицю характеристик, кожен стовбець якої відповідає певному ПЯ, а кожен рядок якої являє собою вектор характеристик одного з еталонів, причому перший рядок відповідає найменш, а останній рядок − найбільш технічно досконалому еталону. Результатом застосування методу є вектор у просторі, спряженому до простору ПЯ об'єктів, що задає лінійний функціонал, значення якого є чисельним ПЯ об'єкта. Геометрично цей вектор можна трактувати як напрям «прогресу». У випадку двох ПЯ результат може бути описаний одним числом, що має сенс кута, утвореного вектором, перпендикулярним до ліній рівня функції корисності, із додатнім напрямом осі, що відповідає першому ПЯ. Будемо позначати вихідну матрицю характеристик $X = (x_{ij})$, а вектор коефіцієнтів функції результату $w = (w_j)$. Проведемо чисельне моделювання того, як випадковий набір вихідних даних X впливає на результат w.

На рис. 1. схематично відображено основні кроки та операції виконання МПЕ.

Очевидні можливі обмеження вихідних даних, коли МПЕ незастосовний, зводяться до випадку, якщо відсутні зміни по окремому ПЯ ($m_j = M_j$ для певного j) − тоді неможливо провести нормалізацію. Не очевидні обмеження методу полягають у невизначеності результату обчислень − вектору w, що може трапитись у випадку, якщо найбільше власне значення матриці V є кратним.

Зазначимо, що у варіанті реалізації МПЕ, розглянутого у роботі [1], перший етап нормалізації проводиться по більш широкій множині, котра окрім еталонів містить також інші об'єкти для дослідження. Їх врахування не впливає принципово на висновки про застосовність методу і цілком не змінює подальших міркувань та позначень у випадку якщо екстремальні значення показників досягаються саме на еталонах.

Характер очікуваного розв'язку. Розглянемо особливості застосування МПЕ до аналізу даних різної розмірності.

Тривіальним є випадок двох еталонів E_1 і E_2, що характеризуються одним ПЯ: $E_1 = (x_{11})$, $E_2 = (x_{21})$. У цьому випадку більш досконалим приймається об'єкт із більшим ПЯ, якщо $x_{11} < x_{21}$, і із меншим ПЯ, якщо $x_{11} > x21$. Якщо кількість ПЯ більша від одного, то МПЕ дозволяє враховувати їх кореляцію, а якщо кількість еталонів більша двох, то МПЕ дозволяє враховувати можливу немонотонність значень окремого показника.

Розглянемо класифікацію очікуваного результату. Із геометричної інтерпретації методу відомо [1], що результат використання МПЕ можна трактувати, як напрям зміни рівня якості виробу у просторі характеристик об'єктів. Оскільки у n-вимірному евклідовому просторі довільний напрям можна задати за допомогою $n - 1$ незалежних параметрів, котрі можна трактувати як напрямні косинуси апроксимуючої лінії, у випадку двох ПЯ результат можна описати єдиним параметром, що лежить у діапазоні $[0, 2\pi]$, у випадку трьох ПЯ − двома параметрами, один із яких лежить у діапазоні $[0, 2\pi]$ (в силу специфіки алгоритму далі зводиться до діапазону $[-\pi/2, \pi/2]$), а інший − у діапазоні $[0, \pi]$ або $[-\pi/2, \pi/2]$ і т.д. Обмежимося аналізом впливу кількості еталонів на результат за мінімальної кількості ПЯ у нетривіальному випадку. Оскільки результат застосування МПЕ по суті є власним вектором деякої матриці, а власний вектор визначено із точністю до відмінного від нуля множника, при вивченні закономірностей розподілу результату не зменшуючи загальності можна прийняти, що перший елемент вектору w дорівнює одиниці (масштабувавши вектор результату). Випадок, коли одна із компонент вектору w рівна нулю, зводиться фактично до того, що за відповідним ПЯ немає прогресу і такий показник повинен апріорі виключатись із розгляду, приводячи до тривіального випадку, розглянутого вище. Разом із тим, такий випадок є одним із можливих результатів застосування методу, тому його виключати із розгляду не будемо, а перейдемо до параметру, що дозволяє його описати.

Випадок двох еталонів. У цьому випадку (для двох еталонів із двома ПЯ) після масштабування власного вектору завжди отримуємо один із результатів:

$$w = (1, 1) \text{ або } w = (1, -1). \qquad (1)$$

Випадок трьох еталонів. У цьому випадку після масштабування власного вектору отримуємо один із випадків:

$$w = (1, \operatorname{tg} \alpha) \text{ або } w = (1, -\operatorname{tg} \alpha). \qquad (2)$$

Тут параметр α має зміст кута, який утворює вектор w із додатнім напрямом осі, уздовж якої

Рис. 1. Алгоритм МПЕ

відкладаємо значення першого ПЯ. У даному випадку питання відсутності прогресу по одному із показників не є тривіальним, оскільки не зводиться до випадку рівності усіх компонент. Тому рівність одної із компонент вектора w нулеві є допустимою, хоча й малоймовірною подією, що в подальшому було підтверджено експериментом. У сенсі подання (2) це відповідає одному із випадків$\{\alpha = 0, \alpha = \pi/2, \alpha = -\pi/2\}$.

Зазначимо, що випадок двох еталонів (1) можна трактувати як частковий випадок подання (2), де параметр α може приймати два значення: $\pi/4$ та $-\pi/4$.

Було проведено 8 серій випробувань для випадку трьох еталонів. Кожна серія випробувань полягала у генерації 10000 випадкових матриць, до яких застосовували МПЕ. Результатом застосування могла бути реалізація одного із наступних випадків «прогресу»:

А). Некоректна матриця (усі дані одного стовпця однакові, теоретична ймовірність цієї події 198^{-2}).

Б). Відсутність прогресу по одному з ПЯ ($\alpha = \pi/2, \alpha = 0, \alpha = -\pi/2$); отримуємо тривіальний випадок моделі із меншою кількістю показників.

В). Прогрес спостерігається при зростанні обох ПЯ ($0 < \alpha < \pi/2$).

Г). Прогрес спостерігається при зростанні одного і зменшенні другого ПЯ ($0 < -\alpha < \pi/2$).

Статистика випадання цих подій, а також деякі статистичні показники отриманих значень α наведено у табл. 1. Аналіз отриманих значень параметра α дозволяє виявити його симетрію відносно значення 0, крім того, розподіли додатніх значень α практично співпадають із розподілами від'ємних значень, зміщених на π. Таким чином, аналіз розподілу фактично можна проводити на прикладі одного із випадків В) або Г). У таблиці також наведено значення середнього (AVE), дисперсії (VAR), асиметрії (SKEW) та ексцесу (KURT) по вибірці окремо додатніх та від'ємних значень із кожної серії.

Рис. 2. Розподіл значень α усіх серій випадку трьох еталонів і двох ПЯ

Розподіл значень ілюструє гістограма на рис. 2. По горизонтальній осі відкладено значення кута α у діапазоні $(-\pi/2; \pi/2)$, вертикальна вісь відображає відносну частоту результату. Приймаючи значення α, отримане у окремому чисельному

Таблиця 1

Результати випробувань для двох ПЯ і трьох еталонів ($n = 10^4 \times 8$)

Випадок	Серія 1	Серія 2	Серія 3	Серія 4	Серія 5	Серія 6	Серія 7	Серія 8
А	2	0	2	1	0	1	0	1
Б	0	4	3	2	0	0	1	2
В	4985	4897	4954	4937	4960	5055	4926	4973
Г	5013	5099	5041	5060	5040	4944	5073	5024
AVE (<0)	-0.7841	-0.7858	-0.7836	-0.7876	-0.7815	-0.7866	-0.7851	-0.7831
VAR (<0)	0.0222	0.0224	0.0205	0.0211	0.0206	0.0203	0.0183	0.0213
SKEW (<0)	0.1197	0.0387	0.0234	-0.2419	0.1618	0.0180	0.0324	0.3157
KURT (<0)	10.0303	9.9302	10.5403	9.7883	10.5234	10.4719	11.2810	10.4451
AVE (>0)	0.7857	0.7864	0.7868	0.7866	0.7866	0.7862	0.7877	0.7885
VAR (>0)	0.0215	0.0215	0.0213	0.0212	0.0211	0.0211	0.0213	0.0214
SKEW (>0)	-0.0019	0.0580	0.1179	0.0818	0.0793	0.0481	0.1834	0.2450
KURT (>0)	10.4863	10.3989	10.4328	10.5057	10.6792	10.5578	10.6934	10.7315

Таблиця 2

Результати випробувань для двох ПЯ і чотирьох еталонів ($n = 10^4$)

Випадок	Серія 1	Серія 2	Серія 3	Серія 4	Серія 5	Серія 6	Серія 7	Серія 8
А	0	0	0	0	0	0	0	0
Б	0	2	0	0	1	0	0	1
В	4956	4887	4979	4962	4977	4963	4937	5119
Г	5044	5111	5021	5038	5022	5037	5063	4880
AVE (<0)	-0.7846	-0.7854	-0.7849	-0.7810	-0.7860	-0.7872	-0.7847	-0.7876
VAR (<0)	0.0387	0.0390	0.0389	0.0393	0.0377	0.0400	0.0379	0.0384
SKEW (<0)	0.0607	-0.0667	-0.1092	0.0799	-0.0108	0.1090	0.0614	0.1305
KURT (<0)	3.9547	3.8747	4.0375	4.0678	4.1020	4.0405	3.9406	3.7227
AVE (>0)	0.7876	0.7821	0.7788	0.7851	0.7890	0.7824	0.7840	0.7865
VAR (>0)	0.0374	0.0361	0.0393	0.0378	0.0382	0.0401	0.0372	0.0387
SKEW (>0)	0.0648	-0.0100	-0.1533	-0.2064	0.0221	0.0604	0.0189	0.0867
KURT (>0)	3.9530	4.1038	4.0285	4.2505	4.0326	3.6938	4.2444	4.0332

експерименті, за реалізацію деякої випадкової величини, отриманий розподіл можна трактувати як емпіричну функцію густини деякої випадкової величини; проінтегрувавши фрагмент цієї функції між фіксованими значеннями α_1 та α_2, знайдемо ймовірність p того, що $\alpha_1 < \alpha < \alpha_2$.

Аналізуючи, наведену на рис. 2., діаграму можна зробити висновок, розподіл значень кутів α (α > 0) зосереджений біля π/4, розподіл α (α < 0) аналогічний з точністю до знаку. Цей розподіл характеризується високим, порівняно із нормальним розподілом, ексцесом. Близкість середнього значення кожної групи (α > 0 і α < 0) до ±π/4 дозволяє трактувати випадок, що відповідає формулі (2) як збурення випадку (1) для моделі із двома еталонами.

Зазначимо, що загальна кількість випадків, коли МПЕ не дозволяє отримувати результат у вигляді власного вектору (випадки А і Б у табл. 1), не перевищує п'яти (серія 3), що складає 0,05%.

Випадок чотирьох еталонів. У цьому випадку отримуємо ситуацію, аналогічну до випадку трьох еталонів, коли результат описується співвідношенням (2) із деяким іншим розподілом параметру *α*. Результат чисельного дослідження наведено у таблиці 2.

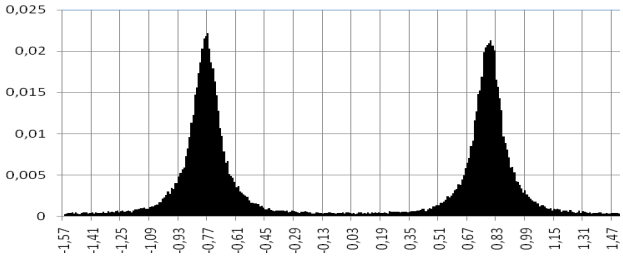

Рис. 3. Розподіл значень α усіх серій випадку чотирьох еталонів і двох ПЯ

Гістограма усіх даних зображена на рис. 3. Порівнюючи табл. 1 та рис. 1 відповідно із табл. 2 та рис. 2 можна відмітити аналогічну симетрію розподілу по від'ємних і додатніх значеннях, можливість розбиття на множини додатніх та від'ємних значень, медіана кожного з яких близька до ±π/4 відповідно. Із відмінностей відзначимо збільшення дисперсії та зменшення ексцесу.

Зазначимо, що загальна кількість випадків, коли МПЕ не дозволяє отримувати результат у вигляді власного вектору (випадки А і Б у табл. 2), не перевищує двох (серія 2), що становить 0,02%.

Вплив збільшення числа еталонів. У випадку п'яти і більшої кількості еталонів отримуємо ситуацію, аналогічну до попередніх випадків. Для випадків від 6 до 15 еталонів було проведено по одній серії випробувань із 104 експериментів. Результат чисельного дослідження для п'яти еталонів наведено на рис. 4а. Подальші дослідження підтвердили тенденцію до «розмивання» даних відносно векторів у формулі (1). У випадку, коли кількість еталонів становила 16 було проведено N = 65535 випробувань, результати яких демонструє діаграма на рис. 4б.

У таблиці 3 наведено результати обчислення описової статистики для додатніх значень результату (α > 0) для однієї серії випробувань у залежності від кількості еталонів.

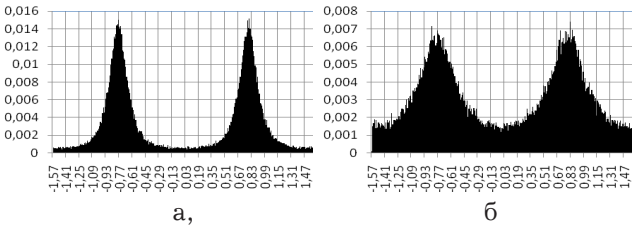

Рис. 4. Розподіл значень α усіх серій випадків для 5-ти (N = 8 · 10⁴) та 16-ти (N = 65535) еталонів

Динаміку зміни середнього значення, стандартного відхилення та ексцесу подано на рис. 5. Як бачимо, найбільш помітні зміни відбуваються при збільшенні числа еталонів із 3-х до 4-х: ексцес зменшується майже утричі, а стандартне відхилення зростає на 32%. Коли число еталонів сягне 10, ексцес стає від'ємний,

Таблиця 3

Основні статистичні показники розподілу параметра α (α > 0) для різного числа еталонів

Статистичний параметр	Кількість еталонів													
	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Середнє	0.781	0.788	0.792	0.787	0.783	0.777	0.781	0.786	0.787	0.791	0.785	0.783	0.792	0.791
Стандартна похибка	0.002	0.003	0.003	0.004	0.004	0.004	0.005	0.005	0.005	0.004	0.005	0.005	0.005	0.005
Медіана	0.785	0.786	0.790	0.785	0.786	0.781	0.785	0.788	0.783	0.788	0.786	0.785	0.789	0.793
Мода	0.785	0.928	1.063	1.413	1.015	1.548	1.039	0.959	1.230	1.111	1.064	0.980	1.121	1.108
Стандартне відхилення	0.146	0.193	0.240	0.267	0.293	0.302	0.322	0.323	0.334	0.335	0.343	0.348	0.350	0.347
Дисперсія	0.021	0.037	0.058	0.071	0.086	0.091	0.103	0.104	0.111	0.112	0.118	0.121	0.123	0.121
Ексцес	11.074	3.953	1.906	1.003	0.479	0.317	0.058	-0.038	-0.180	-0.151	-0.271	-0.322	-0.318	-0.350
Асиметрія	-0.421	0.065	0.006	-0.004	-0.037	-0.046	0.015	0.011	-0.001	0.035	0.001	-0.003	-0.020	-0.021
Інтервал	1.559	1.560	1.570	1.563	1.569	1.567	1.569	1.568	1.567	1.570	1.570	1.570	1.570	1.570
Кількість	4985	4956	4912	4961	4970	4900	4942	4988	4982	4958	5047	5031	5088	5049

а стандартне відхилення практично не змінюється – його коливання становлять 0-3%. Разом із тим, кут-результат α залишається зосередженим навколо вихідних значень ±π/4, а не стає рівномірно розподіленим на усьому проміжку, як можна було б очікувати, враховуючи рівномірний розподіл точок-еталонів всередині куба характеристик.

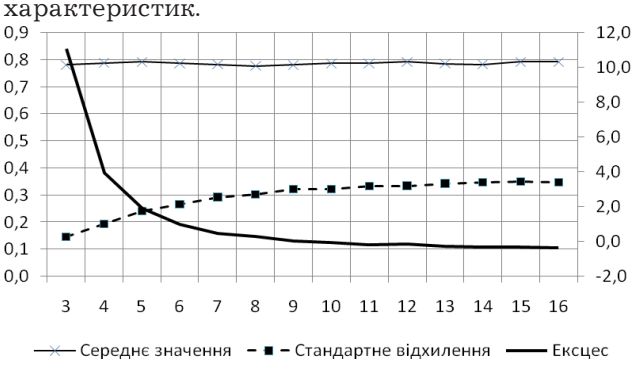

Рис. 5. Середнє, статистичне відхилення (ліва вісь) та ексцес (права вісь) для додатних значень параметру α за різного числа еталонів (3-16)

Висновки. МПЕ може бути ефективно застосований для побудови лінійної функції корисності для широкого діапазону вихідних даних. За рівномірно розподіленого випадкового набору даних загальна кількість випадків, коли метод не дозволяє отримати результат, не перевищує 0,05% у випадку трьох еталонів та 0,02% у випадку чотирьох еталонів;

Результат застосування МПЕ може бути формалізовано у вигляді вектора параметрів, розмірність якого на одиницю менша від кількості ПЯ, а значення компонент лежать у діапазоні (−π/2; π/2). У випадку двох ПЯ результат описується одним числом із згаданого діапазону – кутом-результатом α, значення якого розташовані навколо незбурених значень ±π/4;

Збільшення кількості еталонів приводить до «розмивання» значень результату α по відрізку (−π/2; π/2), причому таке «розмивання» відбувається повільно: у випадку п'яти еталонів стандартне відхилення становить 0,24 при інтервалі у 1.57, або 64% значень кута-результату потрапляють у 20%-й інтервал.

Список літератури:

1. Черноскутов А. І. Комплексная оценка качества изделий с помощью метода прогрессирующего эталона / А. І. Черноскутов, В. Н. Мурзин, В. Е. Ширенков // Надежность и контроль качества. – 1985. – № 5. – С. 12-16.
2. Лук'янов П. О. Порівняльна оцінка технічної ефективності рухомих засобів технічного обслуговування з використанням методу прогресуючого еталона // Збірник наукових праць ЦНДІ ОВТ ЗСУ. – Вип. 3 – К: ЦНДІ ОВТ 1999 – С. 89-94.
3. Кузьменко Р. В. Застосування методу прогресуючого еталону до систем з неповною інформацією / Р. В. Кузьменко, О. М. Зеленюх // Військово-технічний збірник. – Вип. 13 – Львів: НАСВ, 2015. – С. 46-51.
4. Карташов М. В. Теорія ймовірностей та математична статистика – Київ, ТВіМС, 2004. – 307 с.

Кузьменко Р.В.
Национальная академия сухопутных войск
имени гетмана Петра Сагайдачного

ПРИМЕНЕНИЕ МЕТОДА ПРОГРЕССИРУЮЩЕЙ ЭТАЛОНА К АНАЛИЗУ ДАННЫХ РАЗНОЙ РАЗМЕРНОСТИ

Аннотация
В статье проведен статистический анализ результатов применения метода прогрессирующего эталона. Изучено распределение коэффициентов целевой функции в зависимости от размерности исходных данных по их равномерного распределению. Установлено, что результат применения метода прогрессирующего эталона может быть формализована в виде вектора параметров, размерность которого на единицу меньше количества показателей качества, также установлено, что при равномерно-распределенном, случайном наборе данных общее количество случаев, когда метод не позволяет получить результат, не превышает 0,05% при условии наличия трех эталонов, в случае увеличения количества этот показатель будет уменьшаться, что в свою очередь указывает на то, что он может быть эффективно применен для построения линейной целевой функции для широкого диапазона исходных данных.
Ключевые слова: уровень технического совершенства, метод прогрессирующего эталона, метод Монте-Карло.

Kuzmenko R.V.
National Petro Sahaydachnyi Army Academy

APPLICATION OF PROGRESSIVE ETALON METHOD TO ANALYSIS OF DATA OF VARIOUS DIMENSIONS

Summary
The article presents statistical analysis of the results of applying the progressive etalon method. The distribution of coefficients of the target function depending on the dimension of the source data is investigated for uniform distribution of data. It was found that the result of applying the method of progressive etalon can be formalized as a vector whose dimension is one less than the number of quality indicators. There found also that for uniformly distributed random data the total number of cases for which the method does not produce a result, is no more than 0.05% provided that there are three etalons, and less in the case of greater amount of etalons confirming the method efficiency for linear target function constructing for a wide range of source data.
Keywords: technical improving level, progressive etalon method, Monte Carlo method.