

УДК 621.3

Левченко А. О., канд. техн. наук

ГЕНЕРАТОР ЕТАЛОННИХ ВИБІРОК ЧИСЛОВИХ ДАНИХ З ЗАДАНИМ СТУПЕНЕМ ЗАБРУДНЕННЯ ДОВІЛЬНИМИ ЗНАЧЕННЯМИ

Анотація. В статті наведено порядок формування вибірок числових даних з заданим ступенем забруднення довільними значеннями. Генератор вибірок необхідний для перевірки працездатності програмних компонентів, які використовують “IEEE Standart for Binary Floating-Point Arithmetic@ (ANSI / IEEEStd 7C4 – 1985)” та інші стандарти представлення даних на його основі. Використання вказаного порядку генерування талонних вибірок дозволяє підтвердити той факт, що загальна похибка обчислень перевищує межі відносної похибки представлення найбільшого числа.

Ключові слова: числові дані, двійкова арифметика з плаваючою комою, генератор, недостовірні дані, довга арифметика

Левченко А. О., канд. техн. наук

ГЕНЕРАТОР ЭТАЛОННЫХ ВЫБОРОК ЧИСЛОВЫХ ДАННЫХ С ЗАДАННОЙ СТЕПЕНЬЮ ЗАГРЯЗНЕНИЯ ПРОИЗВОЛЬНЫМИ ЗНАЧЕНИЯМИ

Аннотация. В статье приведен порядок формирования выборки числовых данных (значений) с наперед заданной степенью зашумления недостоверными (неточными значениями). Генератор выборок необходим для проверки работоспособности программных компонентов, которые используют “IEEE Standart for Binary Floating-Point Arithmetic@ (ANSI / IEEEStd 7C4 – 1985)” и его разновидности. Использование изложенного порядка получения эталонных выборок позволяет объективно подтвердить факт того что общая ошибка превышает пределы относительной погрешности представления наибольшего числа

Ключевые слова: числовые данные, двоичная арифметика с плавающей запятой, генератор, недостоверные данные, длинная арифметика

Levchenko A. A., PhD

THE GENERATOR OF MAKING THE NUMERICAL DATA WITH THE PRESCRIBED DEGREE OF NOISE WITH INACCURATE VALUES

Abstract. This article gives the order of making the numerical data (value) with the prescribed degree of noise with inaccurate (wrong) values. Generator of Samples is required for testing operability of software components applying “IEEE Standard for Binary Floating-Point Arithmetic@ (ANSI/IEEE 7C4 - 1985)” and its upgrades (variations). Using the described order of obtaining sample reference enables to confirm objectively the fact that the overall error exceeds limits of relative error of the greatest number represented.

Keywords: numerical data, binary floating-point arithmetic, generator, inaccurate values, long arithmetic.

Одним із шляхів розв'язку задачі отримання аналітичних залежностей, що характеризуються масивами числових даних, кожне значення з яких включає похибку її отримання, є використання методів апроксимації. Ці шляхи ґрунтуються на тому, що масив даних замінюють простою функцією, яка описує тенденції зміни цих даних та забезпечує мінімум назначеному критерію відбору функції [1].

Задача знаходження структури та числових значень коефіцієнтів апроксимуючого поліному, з появою обчислювальної техніки, частіше вирішується по комбінаторному алгоритму. Тобто за допомогою перебору всіх можливих

варіантів структур моделі з повного базису та вибору найкращої моделі по заданому зовнішньому критерію (набору критеріїв) при використанні всіх можливих варіантів перебору даних вхідної вибірки.

Згідно з теоремою Вейєрштрасса (Стоуна-Вейєрштрасса), для будь якої безперервної функції можна підібрати послідовність многочленів, рівномірно збіжних до цієї функції на відрізьку. Для розв'язання задачі достатньо знати загальну кількість всіх можливих варіантів структур, кількість всіх можливих варіантів перебору вибірок вхідних даних з загальної вхідної вибірки та критерій відбору структури і коефіцієнтів поліному.

Для розрахунків з плаваючою точкою, що використовуються сучасними процесорами та програмними засобами асоціацією IEEE (Institute of Electrical and Electronics Engineers) розробтан стандарт [2,3,4].

В [5] доведено, що твердження про межу похибки результату при розрахунках з плаваючою точкою яка не повинна перевищувати межу відносної похибки представлення найбільшого числа є не правдивим. Під час розрахунків з плаваючою точкою похибки накопичуються, їх значення сумуються. Такі похибки як «брудний нуль» та «небезпечна редукція» роблять похибки розрахунків занадто великими [5].

Проблема виявляється більш складною коли при реалізації різних алгоритмів побудови апроксимуючих залежностей за допомогою ЕОМ внаслідок неминучих округлень результату обчислень, які пов'язані з обмеженою довжиною розрядної сітки ЕОМ (що зазвичай дорівнює 32 або 64 біта) вже при складанні системи рівнянь та обчисленні поліному 4 степені починають накопичуватися помилки які приводять до виродження матриці і унеможливають розв'язок системи рівнянь [5,6]. Аналогічне явище з'являється й при спробі розрахунків числових значень коефіцієнтів сігмоїди під час побудови нейронних мереж.

Для проведення розрахункових експериментів [6,7] з визначення межі втрати працездатності відповідних алгоритмів необхідно генерування еталонних вибірок числових значень з наперед заданою ступінню забруднення помилковими значеннями.

Метою статті визначено опис програмного генератора вибірок нормального закону розподілу з наперед заданим ступенем забруднення помилковими значеннями.

Функціональна блок-схема генератора зображена на рисунку 1.

Програмний продукт передбачає три режими роботи отримання вибірки даних:

- ручний режим обчислення;

- автоматичний режим обчислення для заданого параметра шуму;

- автоматичний режим обчислення для заданого діапазону значень шуму.



Рис. 1. Функціональна блок-схема генератора вибірок нормального закону розподілу.

При ручному режимі обчислення користувач самостійно задає діапазон значень аргументу, математичне очікування та значення середньоквадратичного очікування закону розподілу.

Результатом роботи програмного продукту при ручному режимі обчислення є генерування вибірки шуму та сумарної вибірки з заданим значенням математичного очікування та середньо квадратичного відхилення нормального закону розподілу.

При автоматичному режимі обчислення для заданого параметра шуму користувач самостійно задає діапазон значень аргументу, математичне очікування нормального закону розподілу, початкове значення шуму та точність підбору розрахункового абсолютного середнього значення функції нормального закону розподілу.

Результатом роботи програмного продукту при автоматичному режимі обчислення для заданого параметра шуму є генерування вибірки шуму та сумарної вибірки з заданим коефіцієнтом забруднення даних.

При автоматичному режимі обчислення для заданого діапазону значень шуму користувач задає діапазон значень шуму та точність підбору розрахункових абсолютних

середніх значень функції нормального закону розподілу.

Результатом роботи програмного продукту при автоматичному режимі обчислення для заданого діапазону значень шуму є генерування вибраного діапазону вибірок шуму та сумарних вибірок з заданим коефіцієнтом забруднення даних з автоматичним збереженням вибраних вибірок.

Додатково в даному режимі користувачеві надається можливість генерування вибірки забруднення даних при різних варіантах розміщення значень вибірки шуму.

Завантажувати вхідні дані в даний програмний продукт можна безпосередньо з файлів *.xls Microsoft Excel.

В зразку вхідного файлу перший рядок – значення аргументу, другий рядок – значення функції.

При завантаженні вхідних даних програмний продукт автоматично вираховує кількість значень кривої.

Програмний модуль виводу результативної інформації забезпечує вивід проміжної інформації, відображення графічних залежностей значень функції від аргументів. Передачу розрахованих даних та графічних залежностей в програмний продукт «Microsoft Excell».

Програмний модуль управління забезпечує вибір шляху до вхідних даних, налаштування автоматичного розрахунку та налаштування виводу.

Генератор вибірок нормального закону розподілу розроблений в системі візуального об'єктно-орієнтованого програмування C++Builder та реалізований на мові C++.

В програмному модулі генерування вибірки нормального закону розподілу в ручному режимі обчислень двохмірний масив вихідної вибірки представлений як масив з елементами типу **long double** довжиною по 8 байт розмірністю [KilkRow][KilkCol]. Змінні Xz, Yz, TmpModChysla, ModChysla, Krok описані як числа з плаваючою крапкою типу **float**.

При цьому назва файлу вибірки формується наступним чином: до назви

файлу додається шум для якого проводиться розрахунок та порядковий номер вибірки для даного шуму. Наприклад: Shum5%_3.

Програмний модуль читання з архіву вхідних даних забезпечує перетворення та завантаження вхідних даних з встановлених форм програмного продукту довжиною 4 байти.

Змінні KilkRow, KilkCol, KilZnach, описані як числа цілого типу **int** довжиною 4 байти.

В якості вхідних даних в даний алгоритм передається мінімальне (**Min**) та максимальне (**Max**) значення аргументу функції нормального закону розподілу для якої буде розраховуватись абсолютне середнє значення функції (**ModChysla**), розрахункові значення математичного очікування (**MatOch**) та середньо квадратичного відхилення (**Sigma**) та кількість розрахункових значень функції (**KilZnach**).

Суть роботи даного алгоритму полягає в наступному: спочатку розраховується приріст аргументу (**Krok**) функції нормального закону розподілу, а потім в циклі розраховується значення функції нормального закону розподілу (**Yz**) в заданій точці (**Xz**). По завершенню розрахунку розраховується абсолютне середнє значення функції нормального закону розподілу.

Приклад виводу графічних залежностей зображений на рисунку 2:



Рис. 2. Приклад виводу графічних залежностей

Блок виводу табличних даних передбачає вивід вибірки шуму для вказаного діапазону значень з заданим математичним очікуванням та середнім квадратичним відхиленням (рис.3).

Вибірка шуму									
	1	2	3	4	5	6	7	8	9
1		1,28571426E	1,57142853E	1,85714280E	2,14285707E	2,42857146E	2,71428585E	3,00000023E	3,28571462E
	0,12076462E	0,13638474E	0,15116967E	0,16445100E	0,17558252E	0,18399202E	0,18922984E	0,19100871E	0,18922984E

Рис. 3. Вибірка шуму

В програмному модулі генерування вибірки нормального закону розподілу в автоматичному режимі обчислень проводиться генерування одномірного масиву вихідної вибірки представленого також як масив з елементами типу **long double** довжиною по 8 байт розмірністю [KilkRow]. Змінні Xz, TmpModChysla, ModChysla, Krok, Pryrist, TmpPryrist, tmp описані як числа з плаваючою крапкою типу **float** довжиною 4 байти.

Змінні KilkRow, KilZnach, i, j, kounter описані як числа цілого типу **int** також довжиною 4 байти.

В якості вхідних даних в даний алгоритм передається мінімальне (**Min**) та максимальне (**Max**) значення аргументу функції нормального закону розподілу для якої буде розраховуватись абсолютне середнє значення функції (**ModChysla**), розрахункові значення математичного очікування (**MatOch**), кількість розрахункових значень функції (**KilZnach**), початкове значення шуму (**Pochatkove**), точність підбору розрахункового абсолютного середнього значення шуму (**Tochnist**) та значення приросту аргументу функції нормального закону розподілу (**Pryrist**).

Для спрощення опису програмного модулю генерування вибірки нормального закону розподілу в автоматичному режимі обчислень блок виводу графічних залежностей та табличних даних розглядатися не буде.

Суть роботи даного алгоритму полягає в наступному: спочатку розраховується приріст аргументу (**Krok**) функції нормального закону розподілу, та потім в циклі розраховується значення функції нормального закону розподілу (**c[i]**) в заданій точці (**Xz**). По завершенню розрахунку розраховується абсолютне середнє значення функції нормального

закону розподілу. Далі проводиться перевірка на справедливість рівності **|Pochatkove-ModChysla|<Tochnist** чи задовольняє розрахункове абсолютне середнє значення функції при заданому початковому значенні шуму заданій точності обчислень. І у випадку коли рівність справджується – проводиться вивід графічних залежностей, табличних даних та вихід з алгоритму. А коли рівність не справджується - проводиться перерахунок діапазону обчислень значень аргументу функції та рекурентний виклик даного алгоритму з новими розрахунковими вхідними даними.

Попередньо при розробці генератору виявлено, що сумування вхідних даних з отриманою вибіркою нормального закону розподілу проходить з накопиченням помилки в середині вибірки. Для виключення виявленого явища запропоновано технологію при якій запропоновано технологію при якій суміщення двох вибірок проводиться за допомогою генератору випадкових чисел, що в свою чергу забезпечує випадкове положення значень шуму в підсумковій еталонній вибірці.

Результатом роботи даного алгоритму є масив A[] елементи якого є неповторювальні випадкові числа в діапазоні від 1 до N+1, де N – кількість значень кривої (**KilZnach**).

Отже, підсумкове отримання значень забруднених вхідних даних з заданим ступенем забруднення отримується наступним чином: до початкових вхідних даних в порядку визначеному в масиві A[] додається вибірка шуму згенерована по нормальному закону розподілу з вказаним початковим значенням шуму.

Для роботи програмного модуля генерування вибірки нормального закону розподілу в автоматичному режимі обчислень для заданого діапазону значень

шуму використовуються такі ж алгоритми як і для роботи програмного модуля генерування вибірки нормального закону розподілу в автоматичному режимі обчислень для одного значення шуму, за виключенням того, що генерується задана кількість вибірок для вибраного значення

шуму та згенеровані вибірки записуються за адресом визначеним в меню збереження даних.

Приклад вихідного файлу згенерованої вибірки для рівня шуму 5% наведено в таблицях 1...3:

Таблиця 1. Параметри вибірки шуму

Вибірка 5% шуму	X min:	X max:	Математичне очікування:	Середнє квадратичне відхилення	Розрахункове абсолютне середнє значення
	-3	1	-1	2,74156	0,13333

Таблиця 2. 5% вибірка шуму

-3	-2,5556	-2,1111	-1,6667	-1,2222	-0,7778	-0,3333	0,11111	0,55556
0,11152	0,12388	0,13404	0,14128	0,14504	0,14504	0,14128	0,13404	0,12388

Таблиця 3. Вибірка сигналу + 5% вибірка шуму

-3	-2,5	-2	-1,5	-1	-0,5	0	0,5	1
9,13404	6,36152	4,14128	2,39504	1,12388	0,39504	0,14128	0,38404	1,12388

Підсумок. Таким чином, з завершенням розроблення генератору еталонних вибірок числових даних з заданим ступенем забруднення довільними значеннями, створено умови проведення розрахункових експериментів з апробацією використання довгої арифметики [6], під якою розуміється математичні дії над числами, які за своїм розміром перевищують обмеження стандартних типів даних та відповідно усувають похибки округлень, що в свою чергу дозволяє вирішувати за допомогою методів класичної математики на ЕОМ системи рівнянь великої розмірності в завданнях проектування систем [8, 9].

Для розрахункового експерименту пропонується метод який буде

використовувати прямі методи розв'язання систем лінійних рівнянь з виконанням обчислень за допомогою довгої арифметики з нескінченною кількістю розрядів (мається на увазі, при точних значеннях коефіцієнтів системи). На практиці при використанні ЕОМ обчислення проводяться з обмеженою кількістю знаків, що визначається розрядністю ЕОМ.

Список використаної літератури

1. Бусленко, Н. П. Моделирование сложных систем: Изд-во «Наука» / Н. П. Бусленко // 1968. – С. 356.
2. IEEE. Стандарт для двоичной арифметики с плавающей точкой

//ANSI/IEEE Std 754/ Вид. IEEE. – New York, NY 10017, USA, East 47th Street: – 1985.

3. IEEE. Стандарт для двоичной арифметики с плавающей точкой //ANSI/IEEE Std 754/ Вид. IEEE. – New York, NY 10017, USA, East 47th Street: – 2008.

4. IEC 60559. Стандарт: двоичная арифметика с плавающей точкой для микропроцессорных систем / Вид. International Electronics Comittii, Inc 345. – New York, NY 10017, USA, East 47th Street: – 1989.

5. <http://www.softelectro.ru/teoriy.html>.

6. Левченко А.О., Войтенков Р.М. Витоки втрати працездатності систем діагностики ОБТ другого роду з представленням чисел з плаваючою комою ISSN International Center / А.О. Левченко, Р.М. Войтенков // Сборник научных трудов SWORLD. – Одеса: – 2014. – № 4(37). – С. 27-35.

7. Максименко, Ю. А. Аналіз залежності ефективності складної системи радіоуправління технічними засобами від тимчасових характеристик управління / Ю. А. Максименко // Збірник наукових праць Військової академії (м. Одеса). – Одеса: – 2015. – Вип. 1(3). – С. 75-80.

8. Levchenko A.O. Determination of the dependence between the reliability values of control means and objects necessary for latter's attainment of definite tasks ISSN International Center / A.O. Levchenko, V.V. Khakhula // Сборник научных трудов SWORLD. – Одеса: – 2015. – № 1(38). – С. 4-12.

9. Гуткин, Л. С. Проектирование радиосистем и радиоустройств / Л. С. Гуткин //М.: Радио и связь. – 1986. – С. 288.

3. IEEE. Standard for Binary Floating-Point Arithmetic // Copyright by The Institute of Electrical and Electronics Engineers Std 754. – New York, NY 10017, USA, East 47th Street: – 2008.

4. IEC 60559. Binary floating-point arithmetic for microprocessor systems // Copyright by The International Electronics Comittii, Inc 345. New York, NY 10017, USA, East 47th Street: –1989.

5. <http://www.softelectro.ru/teoriy.html>.

6. Levchenko A.O., Voitenkov R.M. Souse of working capacity losses of diagnostic system of armament and military equipment of the second grade including the numbers with Floating-Point ISSN International Center / A.O. Levchenko, R.M. Voitenkov. // Sbornik nauchnykh trudov SWORLD. – Odessa: – 2014. – № 4(37). – С. 27-35.

7. Yu. A. Maksymenko. Analiz zalezhnosti efektyvnosti skladnoi systemy radioupravlinnia tekhnichnymy zasobamy vad tymchasovykh kharakterystyk upravlinnia. [Analysis of effectiveness depending on a complex system of radio technical devices of temporary management characteristics.] Zbirnyk naukovykh prats Viiskovoi akademii (m. Odesa), Odesa, 2015, 1 (13). (In Ukrainian).

8. Levchenko A.O., Khakhula V.V. Determination of the dependence between the reliability values of control means and objects necessary for latter's attainment of definite tasks ISSN International Center / A.O. Levchenko, V.V. Khakhula // Sbornik nauchnykh trudov SWORLD. – Odessa: – 2015. – № 1(38). – С. 4-12.

9. L. S. Hutkin. Proektirovanie radiosistem i radioustroistv. [Designing of radio systems and radio.] *Radio i sviaz, Moskow, 1986.* (In Russian)

Отримано 29.04.2016

References

1. N. P. Buslenko Modelirovanie slozhykh sistem. [Modeling of complex systems]. *Nauka, 1968.* (In Russian).
2. IEEE. Standard for Binary Floating-Point Arithmetic // Copyright by The Institute of Electrical and Electronics Engineers Std 754. – New York, NY 10017, USA, East 47th Street: – 1985.



Левченко Андрей
Олександрович, к.т.н.,
доцент, начальник
кафедры, Военная
академия (г.Одеса)
Т.:067-776-0753,
E-mail:
AndreyLevchenko2007@Yandex.ua