УДК 004.67

АНАЛИЗ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ В ПРОЕКТАХ VB.NET С ИСПОЛЬЗОВАНИЕМ СПЕЦИАЛЬНЫХ ФУНКЦИЙ ТАБЛИЧНОГО ПРОЦЕССОРА MS EXCEL

Карчевская М.П., Рамбургер О.Л., Тархов С.В.

ФГБОУ ВО «Уфимский государственный авиационный технический университет», Уфа,
e-mail: karchevskaya.m@mail.ru, ramburger@mail.ru, tarkhov@inbox.ru

Рассмотрен метод, позволяющий значительно расширить возможности проектов, создаваемых в среде объектно-ориентированного программирования VB.NET на платформе Visual Studio, за счет использования процедур и функций других программных продуктов. Показаны возможности решения задач, связанных с аппроксимацией экспериментальных данных в табличном процессоре MS Excel на основе использования класса специальных проблемно-ориентированных функций. Описаны специальные проблемно-ориентированные функции табличного процессора MS Excel, предназначенные для нахождения аппроксимирующих функций, зависящих от нескольких переменных, вычисления значений в произвольных точках, принадлежащих заданному интервалу, а также функций для предсказания будущих значений на основе имеющихся массивов экспериментальных данных. Описан метод интеграции функций MS Excel в проекты VB.NET путем создания класса WorksheetFunction, который содержит все используемые в MS Excel функции. Описаны аналоги функций MS Excel, соответствующие методам класса WorksheetFunction в VB.NET. Приведен практический пример использования функций табличного процессора MS Excel, предназначенных для аппроксимации и прогнозирования экспериментальных данных для ***n*** независимых переменных в проекте VB.NET. Описанный метод анализа экспериментальных данных позволяет исключить непосредственное программирование алгоритмов их обработки.

Ключевые слова: эксперимент, данные, аппроксимирующая функция, прогнозирование, библиотека объектов, табличный процессор, MS Excel, класс Excel.Application, VB.NET

ANALYSIS OF EXPERIMENTAL DATA IN PROJECTS VB.NET WITH CONNECTION OF SPECIAL FUNCTIONS OF MS EXCEL

Karchevskaya M.P., Ramburger O.L., Tarkhov S.V.

Ufa State Aviation Technical University, Ufa,
e-mail: karchevskaya.m@mail.ru, ramburger@mail.ru, tarkhov@inbox.ru

We consider a method that allows you to significantly expand the capabilities of projects created in the environment of object-oriented programming of VB.NET on the Visual Studio platform by using the procedures and functions of other software products. The possibilities of solving problems related to the approximation of experimental data in a tabular MS Excel processor based on the use of a class of special problem-oriented functions are shown. Special problem-oriented functions of the MS Excel spreadsheet processor are described for finding approximating functions depending on several independent variables, calculating values at arbitrary points belonging to a given interval, as well as functions for predicting future values based on available experimental data arrays. The method of integrating the MS EXCEL table processor functions into VB.NET projects is described by creating the WorksheetFunction class in the project, which contains all the functions used in MS Excel. A practical example is given of using the functions of the MS Excel spreadsheet processor designed to approximate and predict experimental data for n independent variables in project VB.NET. The considered method of analysis of experimental data allows us to exclude the direct programming of algorithms for their processing.

Keywords: experiment, date, approximating function, prediction, library of objects, MS Excel, class Excel.Application, VB.NET

В научно-исследовательской деятельности и в инженерной практике часто возникают задачи, связанные с аналитической обработкой экспериментальных данных. Все задачи такого рода объединяет наличие массивов исходных данных, характеризующих состояние и динамику изменения параметров исследуемых объектов или процессов. Построение математических моделей по экспериментальным данным заключается в определении аналитической зависимости, т.е. подборе такой эмпирической формулы, вычисления значений по которой в максимально возможной степени соответствовали бы данным, полученным опытным путем [1]. При этом должны быть учтены случайные отклонения измеряемых величин. Для решения такого рода задач на первом этапе необходимо установить вид зависимости и, соответственно, определить вид аппроксимирующей функции, зависящей от одной или нескольких переменных в соответствии с имеющимися массивами исходных данных, полученных экспериментальным путем. В случае отсутствия каких-либо теоретических предпосылок для определения вида аппроксимирующей функции, обычно выбирают функциональную зависимость из числа наиболее простых, сравнивая их графики с графиком

результатов эксперимента [2]. При этом во многих случаях ограничиваются полиномом. В зависимости от количества параметров, определяемых размерностью массива экспериментальных данных, в качестве аппроксимирующей будет выбрана функция одной или от нескольких переменных.

В настоящее время одним из популярных инструментальных средств, позволяющих эффективно решать такого рода задачи, не прибегая к использованию языков программирования, является табличный процессор, в состав которого включена группа специальных проблемно-ориентированных функций для анализа данных [3].

Цель исследования: повышение эффективности процесса анализа экспериментальных данных в проектах VB.NET посредством интеграции функций табличного процессора MS Excel, что позволит исключить непосредственное программирование алгоритмов их обработки.

Материалы и методы исследования

Аппроксимация экспериментальных данных функцией одной переменной в MS Excel осуществляется путем построения точечного графика (x – имеет конкретные значения) с последующим подбором подходящей аппроксимирующей функции. Для осуществления аппроксимации на диаграмме экспериментальных данных необходимо построить линию тренда, выбрав в контекстном меню пункт «Добавить линию тренда» и в появившемся диалоговом окне указать вид аппроксимирующей функции. Здесь же задается параметр, который позволяет «увидеть» на диаграмме аналитический вид аппроксимирующей функции и достоверность аппроксимации.

В тех случаях, когда аппроксимирующая переменная y зависит от нескольких независимых переменных ($y = f(x_1, x_2,..., x_n)$), подход с построением линии тренда не дает решения. В таких случаях могут быть использованы специальные функции MS Excel ЛИНЕЙН и ЛГРФПРИБЛ (табл. 1). Функции ЛИНЕЙН и ЛГРФПРИБЛ имеют одинаковые аргументы (табл. 2).

Функция ЛГРФПРИБЛ работает подобно функции ЛИНЕЙН. Различия между ними состоят лишь в том, что ЛИНЕЙН определяет параметры прямой линии, наилучшим образом аппроксимирующей исходные данные, а функция ЛГРФПРИБЛ – экспоненциальной кривой. Чем больше график экспериментальных данных напоминает экспоненциальную кривую, тем лучше вычисленная кривая будет аппроксимировать данные. Функция ЛГРФПРИБЛ, так же как функция ЛИНЕЙН, возвращает массив, который описывает зависимость между значениями, но функция ЛИНЕЙН подгоняет прямую линию к имеющимся данным, а ЛГРФПРИБЛ подгоняет экспоненциальную кривую.

Таблица 1

Специальные функции MS Excel для нахождения аппроксимирующих функций, зависящих от нескольких переменных

Название функций	Назначение
ЛИНЕЙН	Аппроксимация совокупности заданных точек x_1, x_2,..., x_n линейной функцией вида $y = a_0 + a_1x_1 + a_2x_2 + ...+ a_nx_n$ с применением метода наименьших квадратов. Позволяет найти неизвестные коэффициенты a_0, a_1, a_2,..., a_n; коэффициент детерминации (R^2) и другие показатели
ЛГРФПРИБЛ	Аппроксимация совокупности заданных точек x_1, x_2,..., x_n показательными функциями вида $y = a_0a_1x^1a_2x^2...a_nx^n$. Позволяет найти неизвестные коэффициенты a_0, a_1, a_2,..., a_n; коэффициент детерминации (R^2) и другие показатели

Таблица 2

Аргументы функций ЛИНЕЙН и ЛГРФПРИБЛ

№	Параметр	Значение
1	<Известные_значения_y>	Множество наблюдаемых значений $y_1, y_2,..., y_n$
2	<Известные_значения_x>	Множество наблюдаемых значений $x_1, x_2,..., x_n$
3	<Конст>	Логическое значение, которое указывает, требуется ли вычислять a_0: – ЛОЖЬ, если a_0 принимается равным нулю для функции ЛИНЕЙН (прямая должна проходить через начало координат) и 1 для функции ЛГРФПРИБЛ; – ИСТИНА, если a_0 вычисляется обычным образом
4	<Статистика>	Логическая переменная, которая указывает, требуется или нет вернуть дополнительную статистику по регрессии. Если аргумент имеет значение ЛОЖЬ или опущен, то функция возвращает только найденные коэффициенты

Таблица 3

Специальные функции MS Excel для нахождения значений в произвольных точках на базе экспериментальных значений

Название функций	Назначение
ТЕНДЕНЦИЯ	Позволяет находить значения полученной линейной функции *y* в произвольных точках, принадлежащих заданному интервалу
РОСТ	Используется для предсказания будущих значений *y*. Позволяет находить значения полученной показательной функции *y* в произвольных точках, принадлежащих заданному интервалу. Рассчитывает прогнозируемый экспоненциальный рост на основании имеющихся данных. Возвращает значения *y* для последовательности новых значений *x*, задаваемых с помощью существующих *x*- и *y*-значений

Таблица 4

Аргументы функций ТЕНДЕНЦИЯ и РОСТ

№	Параметр	Значение
1	<Известные_значения_y>	Множество наблюдаемых значений $y_1, y_2, ..., y_n$
2	<Известные_значения_x>	Множество наблюдаемых значений $x_1, x_2, ..., x_n$
3	<Новые_значения_x>	Значения *x*, для которых необходимо определить соответствующие аппроксимирующие или предсказанные значения *y*
4	<Конст>	Логическая переменная, принимающая значение: – ЛОЖЬ, если a_0 принимается равным 0 при аппроксимации прямой и 1 – при аппроксимации показательной функцией; – ИСТИНА, если a_0 вычисляется обычным образом

Для прогнозирования исследуемых значений используются функции MS Excel ТЕНДЕНЦИЯ и РОСТ (табл. 3). Функция ТЕНДЕНЦИЯ рассчитывает прогнозные значения исследуемого показателя в соответствии с линейным трендом. Подобно функции ЛИНЕЙН она с помощью метода наименьших квадратов аппроксимирует прямой линией массивы известных значений *y* и известных значений *x*. Однако в отличие от функции ЛИНЕЙН она определяет точки, лежащие на этой линии, и прогнозирует значения *y* для вновь заданных значений *x* [4]. При этом функция ТЕНДЕНЦИЯ не приводит математического описания и статистических характеристик самой модели тренда. Функции ТЕНДЕНЦИЯ и РОСТ имеют одинаковые аргументы (табл. 4).

Если аргумент <Новые_значения_x> опущен, то предполагается, что его значения совпадают со значениями аргумента <Известные_значения_x>. Эти функции, если не задавать аргумент <Новые_значения_x>, возвращают массив вычисленных значений *y* для фактических значений *x в* соответствии с прямой или кривой. Прежде чем использовать функцию ТЕНДЕНЦИЯ можно сначала рассчитать значения *y* по линейному тренду, а для функции РОСТ – значения *y* по экспоненциальному тренду.

Использование специальных проблемно-ориентированных статистических функций табличного процессора MS Excel позволяет выполнять обработку данных, полученных в ходе проведения эксперимента. Однако если задачи аппроксимации экспериментальных данных необходимо реализовать в составе «большого проекта» на языке программирования VB.NET, то в этом случае возникает необходимость знания алгоритмов и непосредственного программирования соответствующих методов. Возможности проектов, разрабатываемых в интегрированной среде визуального событийного объектно-ориентированного программирования VB.NET, можно существенно расширить за счет использования функций других пакетов, в частности, функций табличного процессора MS Excel [5].

Метод интеграции функций MS Excel в проекты VB.NET предусматривает создание в проекте ссылки на библиотеку объектов MS Excel помощью команды «Add Reference» меню «Project» (рис. 1). В открывшемся диалоговом окне, во вкладке .Net, следует выбрать Microsoft.Office.Interop.Excel соответствующей версии. Далее необходимо создать экземпляр класса Excel.Application:

```
Dim EXL As New Microsoft.Office.Interop.Excel.Application()
```

После этого становится доступен класс WorksheetFunction, который и содержит все используемые в MS Excel функции [6].

Результаты исследования и их обсуждение

Рассмотрим пример практического использования функций табличного процессора MS Excel, предназначенных для анализа экспериментальных данных, полученных в результате следующего эксперимента [7]. Пусть в жидкость помещен источник радиоактивного излучения. На расстоянии 20, 50 и 100 см от источника расположены датчики. В качестве эксперимента проводились измерения интенсивности излучения через 1, 5 и 10 суток (табл. 5).

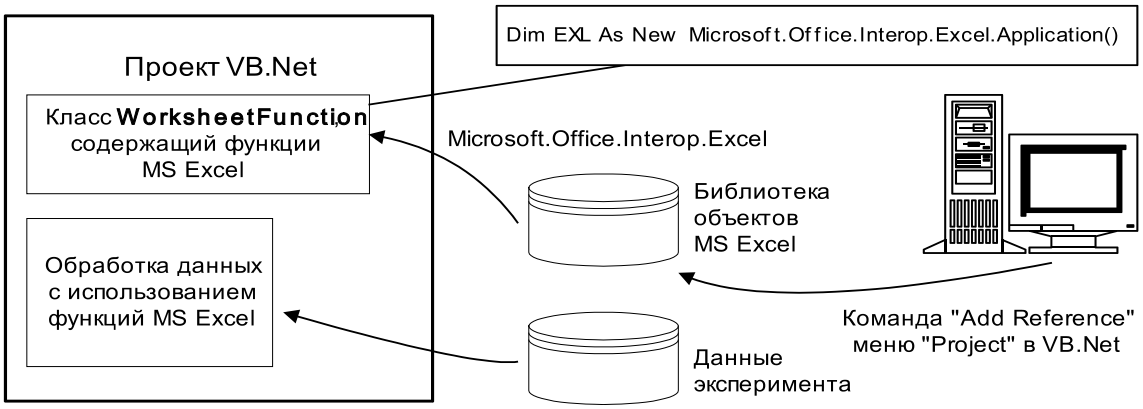


Рис. 1. Подключение и использование библиотеки объектов MS Excel в проекте VB.NET

Таблица 5

Результаты экспериментального исследования

Расстояние до источника *X*1, см	20	50	100	20	50	100	20	50	100
Периодичность измерения *X*2, сутки	1	1	1	5	5	5	10	10	10
Интенсивность излучения *Y*, МРн	61,2	33,6	12,3	43,6	24	8,8	28,3	15,6	5,7

Таблица 6

Задача обработки экспериментальных данных

Расстояние до источника *X*1, см	0	0	0	200	200	200
Периодичность измерения *X*2, сутки	0	1	3	0	1	3
Интенсивность излучения *Y*, МРн	?	?	?	?	?	?

Таблица 7

Параметры метода WorksheetFunction.LogEst

Параметр	Назначение параметра
*arg*1	множество наблюдаемых значений *Y*, полученных в результате эксперимента
*arg*2	множество значений X_1, X_2, для которых проводился эксперимент
*arg*3	логическое значение: 1 – коэффициент a_0 вычисляется обычным образом, 0 – устанавливается принудительно равным 1
*arg*4	логическая переменная: 1 – функция возвращает дополнительную информацию по регрессии, 0 – функция возвращает только коэффициенты регрессии.

Таблица 8

Параметры метода WorksheetFunction.Growth

Параметр	Назначение параметра
*arg*1	множество наблюдаемых значений *Y*, полученных в результате эксперимента
*arg*2	множество значений X_1, X_2, для которых проводился эксперимент
*arg*3	множество значений X_1, X_2, для которых необходимо определить соответствующие аппроксимирующие или предсказанные значения *Y*
*arg*4	логическая переменная: 1 – коэффициент a_0 вычисляется обычным образом, 0 – устанавливается принудительно равным 1

После окончания эксперимента возникла необходимость определить интенсивность излучения возле источника, на расстоянии 200 см, в момент установки источника, через одни сутки и через трое суток (табл. 6).

Для решения данной задачи аппроксимируем совокупность заданных точек показательной функцией, зависящей от двух переменных вида $y = a_0 a_1^{x1} a_2^{x2}$. Затем сделаем прогноз для тех значений x_1 и x_2, ко-

торые в эксперимент не вошли. Для решения такой задачи в MS Excel используются функция ЛГРФПРИБЛ, позволяющая найти неизвестные коэффициенты a_0, a_1, a_2 и коэффициент детерминации (R^2), и функция РОСТ, рассчитывающая прогнозируемые значения функции в произвольных точках на базе экспериментальных данных.

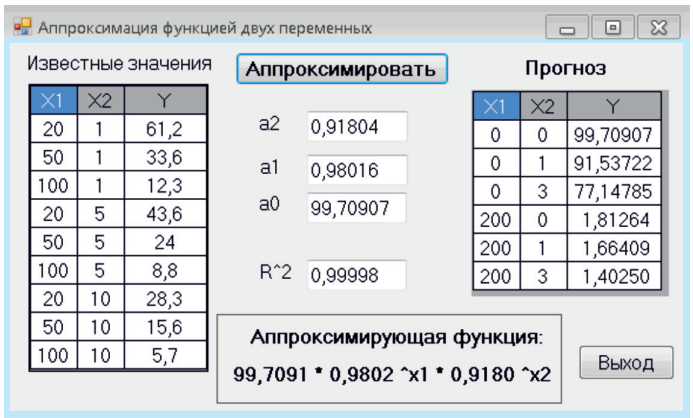
Аналогом функции MS Excel ЛГРФПРИБЛ в VB.NET является метод класса WorksheetFunction.LogEst (табл. 7), а аналогом функции РОСТ в VB.NET – метод WorksheetFunction.Growth (табл. 8).

Результаты работы методов помещаются в соответствующие объекты, к которым можно обращаться так же как к элементам двумерного массива. Для описанного эксперимента это матрицы размерностью 5×3. Фрагмент программного кода, написанный на VB.NET в Visual Studio 2008 представлен на рис. 2.

Объект LPR(,) содержит результаты работы функции LogEst: значения коэффициентов a_2, a_1 и a_0 помещаются в первую строку, стандартные ошибки для коэффициентов a_2, a_1, a_0 – во вторую строку. В первый столбец третьей строки помещается коэффициент детерминации R^2: если $R^2 >= 0,95$, то имеет место высокая точность корреляции с моделью, если коэффициент $R^2 < 0,6$, то уравнение регрессии неудачно и следует выбрать другой вид аппроксимирующей функции. Во второй столбец третьей строки помещается стандартная ошибка. В четвертую строку – значение критерия Фишера (1 столбец) и число степеней свободы (2 столбец). В пятую строку – сумма квадратов регрессии (1 столбец) и остаточная сумма квадратов (2 столбец). Объект Rost(,) содержит результаты работы функции Growth. Его первый столбец – это и есть рассчитанный прогноз. Результат работы программы приведен на рис. 3. Искомая аппроксимирующая функция имеет вид $y(x_1, x_2) = 99,7 \cdot 0,98^{x_1} \cdot 0,92^{x_2}$, точность аппроксимации очень высокая: $R^2 = 0,99998$.

```
'ЗАДАНИЕ ИСХОДНЫХ ДАННЫХ
'Двумерный массив, содержащий множество наблюдаемых значений X1, X2:
Dim x(,) As Double = {{20, 1}, {50, 1}, {100, 1}, {20, 5}, {50, 5}, {100, 5},
                      {20, 10}, {50, 10}, {100, 10}}
'Одномерный массив, содержащий множество наблюдаемых значений Y:
Dim y() As Double = {61.2, 33.6, 12.3, 43.6, 24, 8.8, 28.3, 15.6, 5.7}
'Двумерный массив, содержащий множество значений  X1, X2, для которых необходимо
'определить аппроксимирующие значения Y:
Dim x1(,) As Double = {{0, 0}, {0, 1}, {0, 3}, {200, 0}, {200, 1}, {200, 3}}
'СОЗДАНИЕ ЭКЗЕМПЛЯРА КЛАССА EXCEL.APPLICATION:
Dim EXL As New  Microsoft.Office.Interop.Excel.Application()
'Перед обращением к методам вектор-строку y() следует преобразовать в вектор-столбец y1():
Dim y1  As Object =  EXL.Application.WorksheetFunction.Transpose(y)
'Обращение к методу LogEst, результаты работы метода помещаются в объект LPR(,):
Dim LPR(,) As Object = EXL.Application.WorksheetFunction.LogEst(y1, x, 1, 1)
'Обращение к методу Growth, результаты работы метода помещаются в объект Rost(,)
Dim Rost(,) As Object = EXL.Application.WorksheetFunction.Growth(y1, x, x1, 1)
```

Рис. 2. Фрагмент программного кода на VB.NET



X1	X2	Y
20	1	61,2
50	1	33,6
100	1	12,3
20	5	43,6
50	5	24
100	5	8,8
20	10	28,3
50	10	15,6
100	10	5,7

X1	X2	Y
0	0	99,70907
0	1	91,53722
0	3	77,14785
200	0	1,81264
200	1	1,66409
200	3	1,40250

Рис. 3. Результат работы программного кода

Точно так же можно аппроксимировать экспериментальные данные для n независимых переменных показательной функцией вида $y = a_0 a_1^{x1} a_2^{x2} \ldots a_n^{xn}$. В этом случае объект LPR будет иметь размерность $5\times(n + 1)$, его первая строка будет содержать коэффициенты $a_n, a_{n-1}, \ldots, a_0$. Аналогичным образом можно подключить другие функции MS Excel. Необходимо только знать англоязычный аналог их русифицированного названия, способ передачи параметров функции и как обработать объект, в котором содержится результат ее работы.

Заключение

Решение задачи обработки экспериментальных данных в проектах VB.NET на основе парадигм императивного и аспектно-ориентированного программирования подразумевает разработку алгоритмов и написание исходного кода подпрограмм для реализации соответствующих методов. Описанный метод интеграции функций табличного процессора MS Excel в проекты VB.NET для реализации анализа экспериментальных данных позволяет исключить непосредственное программирование алгоритмов их обработки. Таким образом, метод позволяет снизить затраты на разработку проектов VB.NET и исключить возможные ошибки, которые могли быть допущены при разработке процедур обработки данных.

Список литературы

1. Демидович Б.П. Численные методы анализа. Приближение функций, дифференциальные и интегральные уравнения: учеб. пособие. СПб.: Лань, 2010. 400 с.

2. Борздова Т.В. Основы статистического анализа и обработка данных с применением Microsoft Excel: учеб. пособие. Минск: ГИУСТ БГУ, 2011. 75 с.

3. Просветов Г.И. Анализ данных с помощью Excel. Задачи и решения. М.: Альфа-Пресс, 2015. 160 с.

4. Скриба С.И., Скриба Н.Н. Экономико-статистическое моделирование и прогнозирование средствами MS Excel: учеб. пособие. Минск: БГЭУ, 2002. 171 с.

5. Библиотека документации для разработчиков под ОС MS Windows [Электронный ресурс]. URL: https://msdn.microsoft.com/ru-ru/library/wss56bz7.aspx (дата обращения 24.10.18).

6. Библиотека документации для разработчиков под ОС MS Windows [Электронный ресурс]. URL: https://msdn.microsoft.com/en-us/library/microsoft.office.interop.excel.worksheetfunction_methods.aspx (дата обращения 24.10.18).

7. Кондратьева Н.В., Карчевская М.П., Рамбургер О.Л. Основы планирования эксперимента и обработки экспериментальных данных: учеб. пособие. Уфа: УГАТУ, 2012. 252 с.