

РАЗНОЕ

ВОССТАНОВЛЕНИЕ ОДНОМЕРНЫХ ФУНКЦИОНАЛЬНЫХ ЗАВИСИМОСТЕЙ ПО ВЫБОРКАМ ОГРАНИЧЕННОГО ОБЪЕМА

Д-р техн. наук, проф. В.Н. Сызранцев, ст.препод. Т.Р. Змызгова, д-р техн. наук, проф. С.Л. Голофаст

Рассмотрен алгоритм построения аппроксимирующих функций по выборкам ограниченного объема, основанный на теории минимизации функционала эмпирического риска. Алгоритм направлен на решение таких задач, когда точный вид аппроксимирующей функции заранее не известен, могут иметься лишь общие представления о ее характере, диапазоне изменения аргументов, существенности их вклада в общую зависимость.

Многие задачи восстановления функциональных зависимостей сводятся к одной и той же схеме - непосредственной аппроксимации экспериментальных точек аналитической зависимостью. Успех применения данных методов в существенной мере определяется корректным выбором класса приближающих функций и наиболее естественным упорядочением их сложности.

Традиционно при аппроксимации экспериментальных точек аналитической зависимостью искомое приближение строится на основе его выбора из заранее фиксированного пространства гладких функций, который, чаще всего, определяется квалификацией исследователя, его опытом и возможностями наиболее распространенного при решении подобных задач метода наименьших квадратов в части решения получаемых при этом систем нормальных уравнений. В качестве альтернативы может быть использован метод сплайн-аппроксимации на неравномерной сетке узлов [1]. Результаты обработки экспериментальных данных свидетельствуют, что нередко искомая зависимость может быть настолько сложной, что при построении аппроксимирующей функции практически не используются строгие математические критерии оптимальности. Их заменяют качественные представления о целесообразности той или иной формы

связи, опирающиеся на субъективные оценки результатов. Другая проблема возникает из-за ограниченности объема эмпирических данных.

Предлагаемый алгоритм построения аппроксимирующей функции основан на теории минимизации функционала эмпирического риска [2]. Алгоритм направлен на решение таких задач, когда точный вид аппроксимирующих функций заранее не известен, могут иметься лишь более или менее грубые представления об ее характере, диапазоне изменения аргументов, существенности их вклада в общую зависимость. Оптимальная зависимость строится автоматически, наилучшим образом соотнося качество приближения экспериментальных данных и сложность выбранной функции. Реализация предложенного алгоритма может быть осуществлена только в процессе компьютерного моделирования.

Сущность метода структурной минимизации эмпирического риска сводится к тому, что на первом этапе в качестве начальной приближающей функции выбирается предельно грубая модель. Далее эта модель постепенно усложняется до достижения оптимального соотношения между точностью аппроксимации эмпирического материала и надежностью результата в условиях ограниченного объема данных. Алгоритмическое построение приближения функции регрессии заключается в поиске функции, минимизирующей ее среднее квадратичное отклонение от экспериментально наблюдаемых значений исследуемой зависимости, т.е. эмпирический риск

$$I_{\varepsilon}(y) = \frac{1}{n} \sum_{i=1}^n \frac{(y_i - y(x_i))^2}{\sigma_i^2}, \quad (1)$$

где (x_i, y_i) , $i = 1, \dots, n$, - независимые пары значений функции y и ее аргумента x , σ_i^2 - дисперсии замеров y_i .

Теоретическое обоснование данного критерия позволяет утверждать [2], что задача восстановления искомой зависимости может быть сведена к математической процедуре выбора размерности пространства сглаживающих функций. Оптимальной размерностью будет считаться та, при которой достигается минимум функционала эмпирического риска по экспериментальным данным, который соответствует минимуму среднеквадратической ошибки сглаживания.

В целях сокращения объема вычислений при определении параметров искомой зависимости в качестве базисных функций удобно использовать ортонормированные функции [3]. При использовании многочленов высокой степени предпочтительным является метод последовательного повышения степени многочлена, предложенный П.Л. Чебышевым [2, 4].

В качестве искомой регрессионной функции берется алгебраический полином степени k , который имеет вид

$$y(x) = \sum_{j=0}^k \alpha_j Q_j(x), \quad (2)$$

где $Q_j(N) = \cos(j \cdot \arccos(N))$ - полином Чебышева степени j .

Чтобы отказаться от использования тригонометрических соотношений для нахождения полиномов Чебышева любого порядка, легко установить для них рекуррентное соотношение вида

$$Q_{j+1}(x) = 2xQ_j(x) - Q_{j-1}(x). \quad (3)$$

Суть метода состоит в построении многочленов различных степеней, каждый из которых для своей степени минимизирует функционал эмпирического риска $I_3(y)$, и в выборе из них полинома, для которого данный функционал принимает наименьшее значение. В этом случае функционал эмпирического риска можно записать следующим образом:

$$I_3(y) = \frac{1}{n} \sum_{i=1}^n \frac{\left(y_i - \sum_{j=0}^k \alpha_j Q_j(x_i) \right)^2}{\sigma_i^2}. \quad (4)$$

Очевидно, что этот функционал соответствует минимуму среднеквадратической ошибки построения приближающей зависимости и фактически представляет собой меру адекватности построенного приближения значениям $(x_1, y_1, \dots, x_n, y_n)$.

Для решения задачи восстановления полиномиальной зависимости сначала определяется степень искомого полинома k , а затем в классе полиномов этой степени восстанавливается регрессия. При фиксированной степени k минимум функционала находится путем решения нормальной системы линейных уравнений с симметрической матрицей коэффициентов, полученных методом наименьших квадратов, относительно параметра α , при котором достигается минимум функционала эмпирического риска

$$B^T \cdot B\alpha = B \cdot B^T \delta. \quad (5)$$

Здесь $\alpha = (\alpha_0, \dots, \alpha_k)^T$ - искомый вектор коэффициентов разложения функции регрессии по полиномам Чебышева, $y = (y_1, y_2, \dots, y_n)^T$ - вектор экспериментальных значений исследуемой зависимости, $B = \{Q_j(x_i)/\sigma_i^2\}$ - матрица размера $n \times (k+1)$ значений полиномов Чебышева в экспериментальных точках x_i , $i = 1, \dots, n$, σ_i^2 - дисперсии замеров y_i .

Разрешая эту систему уравнений в матричном виде относительно параметра α , находим, что

$$\alpha^* = (B^T B)^{-1} B^T \delta. \quad (6)$$

Тогда достигнутая величина функционала эмпирического риска будет равна

$$I_3(\alpha^*) = \frac{1}{n} \sum_{i=1}^n \frac{\left(y_i - \sum_{j=0}^k \alpha_j^* Q_j(x_i) \right)^2}{\sigma_i^2}. \quad (7)$$

Именно это значение характеризует качество построенного приближения искомой зависимости. Оценка качества осуществляется по той же выборке, для которой это приближение строилось. На основании теорем, приведенных в [2], можно утверждать, что качество построенного соотношения можно определить выражением

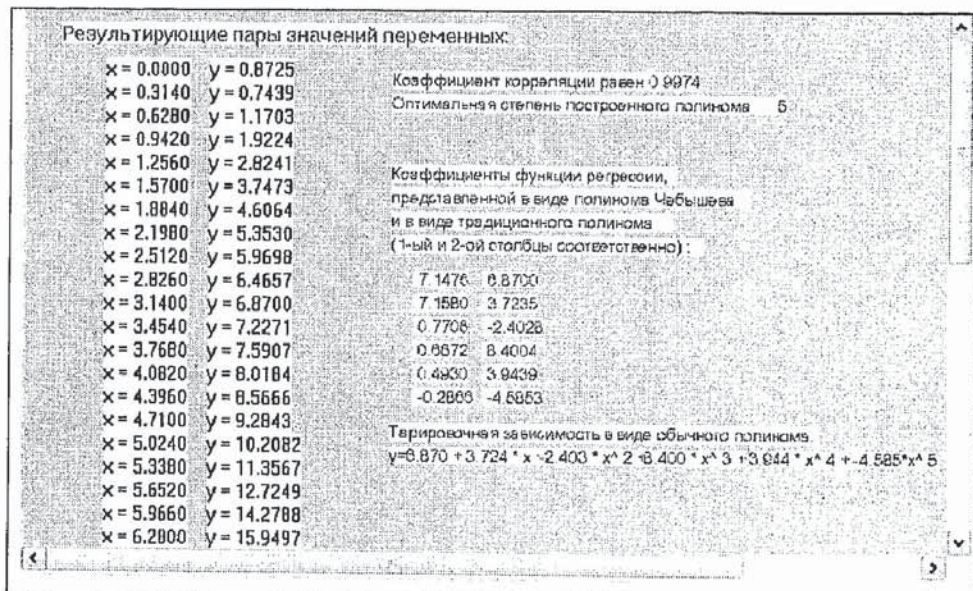
$$J(k) = \frac{I_3(\alpha^*)}{1 - \sqrt{C}}, \text{ где } C = \frac{(k+1) \left(\ln \left(\frac{n}{k+1} \right) + 1 \right) - \ln \eta}{n}. \quad (8)$$

Эта оценка имеет место для любой случайной выборки. Здесь $1 - \eta$ - вероятность, с которой она справедлива. Следует отметить, что величина критерия $J(k)$ зависит от степени полинома k . Степень, при которой значение критерия будет наименьшим, и является оптимальной степенью построенной полиномиальной зависимости. При этом сама функция регрессии аппроксимируется полиномом такого же порядка.

Входными параметрами программы, реализующей описанный выше алгоритм, служат число экспериментальных данных, массив значений независимой переменной, упорядоченный в порядке неубывания, и соответствующий массив значений зависимой переменной. Результатом работы программы являются оптимальная степень полученного полинома, массив значений его коэффициентов в виде разложения по полиномам Чебышева и в виде обычного многочлена.

Для компьютерного тестирования рассмотренного алгоритма была использована задача построения полиномиального приближения к регрессии, значения которой заданы со случайной помехой следующим образом:

- общее число экспериментальных пар точек $L = 21$;
- массив значений независимой переменной: 0.000; 0.314; 0.628; 0.942; 1.256; 1.570; 1.884; 2.198; 2.512; 2.826; 3.140; 3.454; 3.768; 4.082; 4.396; 4.710; 5.024; 5.338; 5.652; 5.966;



6.280;

- массив значений зависимой переменной: 0.670; 1.000; 1.492; 1.915; 2.167; 3.500; 4.816; 5.736; 6.242; 6.500; 6.725; 7.000; 7.325; 8.000; 8.750; 9.500; 10.444; 11.000; 12.500; 14.612; 15.857.

Рис. 1. Коэффициенты аппроксимирующей зависимости

Результаты компьютерной реализации алгоритма и вид полученной аппроксимирующей зависимости приведены соответственно на рис. 1 и рис. 2.

Согласно полученным данным искомое приближение можно представить в виде разложения по полиномам Чебышева:

$$y = 7.1476 + 7.1580 \cdot Q_1(x) + 0.7706 \cdot Q_2(x) + 0.6672 \cdot Q_3(x) + 0.4930 \cdot Q_4(x) - 0.2856 Q_5(x) \quad (9)$$

где $Q_j(x) = \cos(j \cdot \arccos(x))$, причем $Q_{j+1}(x) = 2xQ_j(x) - Q_{j-1}(x)$.

Очевидно, что предпочтительнее будет использование итоговой зависимости в виде обычного полинома:

$$y = 6.8700 + 3.7235 \cdot x - 2.4028 \cdot x^2 + 8.4004 \cdot x^3 + 3.9439 \cdot x^4 - 4.5853x^5. \quad (10)$$

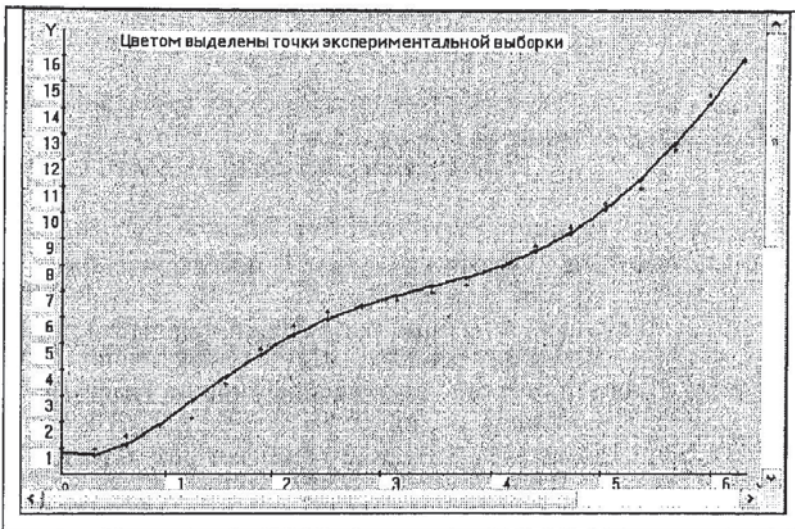


Рис. 2. Вид аппроксимирующей кривой

Анализ рис.2 показывает, что данная методика демонстрирует хорошее приближение экспериментальных точек построенной кривой. Таким образом, задача определения одномерных функциональных зависимостей может быть разложена на этапы, на каждом из которых с успехом могут быть реализованы формализованные методы обработки

данных. Разработанный алгоритм и реализующая его программа применялись в Курганском государственном университете и Тюменском государственном нефтегазовом университете при обработке данных, полученных в процессе научно-исследовательских работ по оценке нагруженности и ресурса деталей и металлоконструкций машин с помощью датчиков деформаций интегрального типа.

СПИСОК ЛИТЕРАТУРЫ

1. Методы и алгоритмы аппроксимации решений на неравномерной сетке: Методические рекомендации. — Минск: Институт технической кибернетики АН БССР, 1991. —142с.
2. Алгоритмы и программы восстановления зависимостей/ Под ред. Вапника В. Н. — М.: Наука, Главная редакция физико-математической литературы, 1984. —816с.

3. Хардле В. Прикладная непараметрическая регрессия: Пер. с англ. М., Мир, 1993.- 349 с.
4. Виттих В.А., Сергеев В.В., Соيفер В.А. Обработка изображений в автоматизированных системах научных исследований. М.: Наука, 1982. –216с.

К ВОПРОСУ ОПТИМИЗАЦИИ ОРГАНИЗАЦИОННЫХ СТРУКТУР НА ОСНОВЕ ПОЛОЖЕНИЙ СТАНДАРТА ГОСТ Р ИСО 9001

Д-р техн. наук, проф. С.Л. Голофаст, эксперт В.А.Булахов

Полноценная реализация требований систем менеджмента качества [1] невозможна без проведения трансформации организационных структур, основанных на функциональном подходе, т.е. приведения их к процессному виду. С одной стороны процессный подход является базовым принципом менеджмента качества, нарушение которого резко снижает результативность внедряемых систем менеджмента качества, с другой стороны реализация этого принципа является прямым требованием стандарта.

Практически, функциональный и процессный подходы представляют собой диалектически связанные формы организационной структуры и, как правило, дополняют друг друга, а система управления именуется по своей доминирующей составляющей части.

В отечественной практике основное распространение получили функциональные организационные структуры. Для такой системы характерно разделение функциональных подразделений по виду ресурсов, которые они предоставляют в общее пользование. Организационные структуры таких систем имеют иерархическую (пирамидальную) структуру. Данная структура позволяет сглаживать межфункциональные противоречия через общие органы управления, находящиеся на более высоких уровнях иерархии. В крупных организациях пирамиды управления могут содержать несколько уровней иерархии. Главной проблемой таких структур является их низкая оперативная управляемость. Если учесть, что по теории информации пропускная способность каналов связи в иерархических системах "снизу-вверх" хуже в N^2 раз, чем канала "сверху-вниз" (где N – кратность подчинённости), становится понятным то, что высшие уровни