

АЛГОРИТМИЧЕСКАЯ И ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ПОСТРОЕНИЯ ОПТИМАЛЬНЫХ МОДЕЛЕЙ ПРОГНОЗА

Рассматривается методика отыскания моделей прогноза с минимальными ошибками прогноза. Основу методики составляют полиномиальные модели параметрического прогноза и алгоритмическая схема поиска оптимальных моделей прогноза.

Пусть мы имеем матрицу исследования размера $N \times M$,
где N – количество строчек-наблюдений,

M – количество параметров-столбиков исследования.

Будем строить параметрические модели прогноза [1-3] и осуществлять прогноз по ним по следующей алгоритмической схеме:

1. На наблюдениях с 1-го по 3-е строим параметрическую полиномиальную модель прогноза.

2. Осуществляем по построенной модели прогноз исследуемого параметра на одно или два и т.д. до пяти наблюдений вперед.

Поскольку модель для прогноза построена на наблюдениях с 1-го по 3-е, то осуществим прогноз, например, на два наблюдения вперед; посчитаем прогнозируемые значения исследуемого параметра в наблюдениях четвертом и пятом.

3. Вычислим абсолютные ошибки прогноза в наблюдениях четвертом и пятом (истинные значения исследуемого параметра в этих наблюдениях известны).

4. Вычислим среднюю абсолютную ошибку прогноза по наблюдениям четыре и пять.

5. Строим модель прогноза на наблюдениях с 1-го по 4-е.

6. Осуществляем прогноз в наблюдениях 5 и 6, вычисляем среднюю абсолютную ошибку прогноза.

7. По аналогичной схеме строим модель прогноза на наблюдениях с 1-го по 5-е, делаем прогноз для наблюдений 6 и 7 и т.д.

Допустим, мы осуществили прогноз по указанной схеме для десяти интервалов прогноза. То есть модель прогноза для 10-го интервала прогноза строилась с 1-го наблюдения по 12-е наблюдение, а прогноз осуществлялся для наблюдений 13-го и 14-го.

8. Строим модели прогноза по указанной схеме со второго наблюдения:

– модель со 2-го по 4-е наблюдение, прогноз 5-е и 6-е наблюдение;

– модель со 2-го по 5-е наблюдение, прогноз 6-е и 7-е наблюдение и т.д.

– модель со 2-го по 13-е наблюдение, прогноз 14-е и 15-е наблюдение.

9. Строим модели прогноза по указанной схеме с 3-го, 4-го и т.д. наблюдений матрицы исследования.

Очевидно, построение моделей прогноза будет заканчиваться на наблюдении N_2 .

Последний прогноз будет осуществляться для наблюдений N_1, N .

При программной реализации представленного алгоритма был создан специальный файл, в который записывались для каждого прогноза:

– границы отрезка построения модели прогноза,

– количество наблюдений для построения модели прогноза,

– средняя абсолютная ошибка при построении модели прогноза,

– средняя абсолютная ошибка при прогнозе.

Данные файла с ошибками прогноза упорядочивались по возрастанию этих ошибок прогноза.

После этого из файла с упорядоченными по возрастанию ошибками прогноза выбирались ошибки прогноза с первым появлением номера наблюдения правого конца отрезка построения модели.

Очевидно, если номер наблюдения правого конца отрезка построения модели прогноза встречается в файле с упорядоченными по возрастанию ошибками прогноза несколько раз, то первое появление этого номера будет соответствовать минимальной

ошибке прогноза из всех отрезков построения модели прогноза, которые в качестве правого конца имеют этот номер.

Например, пусть мы имеем фрагмент файла с упорядоченными ошибками прогноза:

2 8 7 0,5 0,70 – 1-я строчка

1 8 8 0,9 0,95 – 2-я строчка

3 8 6 1 1,50 – 3-я строчка.

В первой строчке фрагмента:

[2, 8] – концы отрезка построения модели прогноза;

7 – количество наблюдений для построения модели прогноза;

0,5 – средняя абсолютная ошибка при построении модели прогноза;

0,7 – средняя абсолютная ошибка прогноза на наблюдениях 9, 10 по модели, построенной на отрезке [2, 8].

Во второй строчке фрагмента:

[1, 8] – концы отрезка построения модели прогноза;

8 – количество наблюдений для построения модели прогноза;

0,9 – средняя абсолютная ошибка при построении модели прогноза;

0,95 – средняя абсолютная ошибка прогноза на наблюдениях 9, 10 по модели, построенной на отрезке [1, 8].

Аналогично рассматриваются данные третьей строчки фрагмента.

Поскольку ошибки прогноза в этом фрагменте упорядочены по возрастанию, то первое появление отрезка с правым концом 8 (номер наблюдения) для построения модели прогноза будет соответствовать минимальной ошибке прогноза из всех отрезков построения модели прогноза с правым концом 8.

После формирования файла с первыми появлениями номера правого конца отрезка построения модели прогноза записи этого файла упорядочивались по возрастанию номера наблюдения правого конца отрезка построения модели прогноза.

Получившийся после ранжирования файл представляет собой координаты концов отрезков построения моделей прогноза, для которых средняя абсолютная ошибка прогноза на последующих наблюдениях будет минимальной.

Разработанная методика была использована для построения оптимальных моделей прогноза количества вызовов скорой помощи по приливным изменениям силы тяжести.

Оптимальные модели прогноза количества вызовов скорой помощи по астме

Для построения моделей прогноза была взята матрица исследования со следующими параметрами-столбиками:

– количество нулей первой производной приливных изменений силы тяжести,

– количество нулей второй производной приливных изменений силы тяжести,

– количество нулей третьей производной приливных изменений силы тяжести,

– количество нулей четвертой производной приливных изменений силы тяжести,

– общее количество вызовов по астме в день.

Строчками-наблюдениями были значения параметров исследования по всем дням 2006 года (данные скорой помощи города Оренбурга).

Параметрами-аргументами в полиномиальных моделях были количества нулей 1, 2, 3, 4-й производных приливных изменений силы тяжести.

В качестве зависимой переменной рассматривалось общее количество вызовов по астме в день.

Отрезки для построения оптимальных моделей прогноза общего количества вызовов по астме в день (фрагмент):

1,000000	8,000000	4,400000
5,000000	9,000000	4,424946
.....
37,000000	67,000000	2,718140E-01
.....
293,000000	353,000000	2,384252E-01
288,000000	354,000000	2,919523E-01
288,000000	355,000000	9,278207E-01
264,000000	356,000000	1,657394
265,000000	357,000000	1,731236
326,000000	358,000000	1,990170
317,000000	359,000000	1,989979
317,000000	360,000000	2,350253

Первые два столбика во фрагменте – это номера наблюдений концов отрезков пост-

роения оптимальных моделей прогноза для соответствующего правого конца отрезка.

Третий столбик – это средняя абсолютная ошибка прогноза на последующих пяти наблюдениях после правого конца отрезка построения модели прогноза.

Модели прогноза и результаты прогноза по астме

Вариант прогноза-1

Начальный номер наблюдения для построения модели прогноза – 1.

Количество наблюдений для построения модели прогноза – 8.

Коэффициенты при аргументах в оптимальной модели прогноза общего количество вызовов по астме в день для правого конца отрезка построения модели прогноза – 8:

аргумент – 1
 количество нулей производной – 1
 степень – 4 коэффициент = 0,1951071
 степень – 3 коэффициент = -0,2536301
 степень – 2 коэффициент = -4,5416839
 степень – 1 коэффициент = 10,9388985

аргумент – 2
 количество нулей производной – 2
 степень – 4 коэффициент = 0,1039694
 степень – 3 коэффициент = -0,9065053
 степень – 2 коэффициент = 1,6537975
 степень – 1 коэффициент = 1,2694698

аргумент – 3
 количество нулей производной – 3
 степень – 4 коэффициент = -0,1201782
 степень – 3 коэффициент = -0,0653564
 степень – 2 коэффициент = 2,5656792
 степень – 1 коэффициент = 7,0830884

аргумент – 4
 количество нулей производной – 4
 степень – 4 коэффициент = 0,0418836
 степень – 3 коэффициент = 0,1605077
 степень – 2 коэффициент = -4,4820457
 степень – 1 коэффициент = 9,5923769

свободный член модели = -13,5284657

средняя абсолютная ошибка при построении модели = 0,62500

Таблица 1. Вклады параметров-аргументов

Номер:	Название параметра	Вклад
1:	122. количество нулей производной	-1: 0,29461:
2:	123. количество нулей производной	- 2: 0,19056:
3:	124. количество нулей производной	- 3: 0,18721:
4:	125. количество нулей производной	- 4: 0,32762:

Результаты прогноза:

интервал прогноза – 9 13
 истина = 10,0000 прогноз = 13,5000 ошибка = 3,5000
 истина = 21,0000 прогноз = 13,5000 ошибка = 7,5000
 истина = 19,0000 прогноз = 13,5000 ошибка = 5,5000
 истина = 10,0000 прогноз = 9,5935 ошибка = 0,4065
 истина = 17,0000 прогноз = 11,9065 ошибка = 5,0935

средняя абсолютная ошибка прогноза = 4,40000

Вариант прогноза-2

Начальный номер наблюдения для построения модели прогноза – 5

Количество наблюдений для построения модели прогноза – 5

Коэффициенты при аргументах в оптимальной модели прогноза общего количество вызовов по астме в день для правого конца отрезка построения модели прогноза – 9:

аргумент – 1
 количество нулей производной – 1
 степень – 2 коэффициент = -0,0623659
 степень – 1 коэффициент = 2,3741951

аргумент – 2
 количество нулей производной – 2
 степень – 2 коэффициент = -1,0200570
 степень – 1 коэффициент = 1,1403989

аргумент – 3
 количество нулей производной – 3
 степень – 2 коэффициент = 0,2027514
 степень – 1 коэффициент = 2,6431058

аргумент – 4
количество нулей производной – 4
степень – 2 коэффициент = -0,8587239
степень – 1 коэффициент = 3,7285149

свободный член модели = 2,2694656

средняя абсолютная ошибка при построении модели = 0,00000

Таблица 2. Вклады параметров-аргументов

Номер:	Название параметра	Вклад
1:	122. количество нулей производной – 1:	0,10907:
2:	123. количество нулей производной – 2:	0,27114:
3:	124. количество нулей производной – 3:	0,18339:
4:	125. количество нулей производной – 4:	0,43640:

Результаты прогноза:

интервал прогноза – 10 14

истина = 21,0000, прогноз = 14,0000,
ошибка = 7,0000

истина = 19,0000, прогноз = 14,0000,
ошибка = 5,0000

истина = 10,0000, прогноз = 12,0624,
ошибка = 2,0624

истина = 17,0000, прогноз = 9,9376, ошибка = 7,0624

истина = 13,0000, прогноз = 14,0000,
ошибка = 1,0000

средняя абсолютная ошибка прогноза = 4,42495

Используя данные файла с координатами концов для построения оптимальных моделей прогноза, мы можем осуществить прогноз общего количества вызовов скорой помощи по астме с минимальной ошибкой на любые пять дней (идут подряд) соответствующего года рассмотрения.

Разработанная методика прогноза может использоваться для любых многопараметрических объектов исследования.

Список использованной литературы:

1. Бендат Д. Ж., Пирсол А. Измерение и анализ случайных процессов. – М.: Мир, 1974.
2. Драйпер Н., Смит Г. Прикладной регрессионный анализ. – М.: Статистика, 1973.
3. Brandon D. B. Developing Mathematical Models for Computer Control, USA Journal, 1959, V.S,N7.

Статья рекомендована к публикации 28.12.07