

МАТЕМАТИЧЕСКИЙ АНАЛИЗ ЭКОНОМИЧЕСКИХ МОДЕЛЕЙ

Парная кусочно-линейная регрессионная модель с фиксированными узлами

© 2024 г. К.М. Зубрилин

К.М. Зубрилин,

Филиал ФГБОУ ВО «Керченский государственный морской технологический университет», Феодосия;
e-mail: kzubrilin@yandex.ru

Поступила в редакцию 25.05.2023

Аннотация. В работе представлен метод построения парной регрессионной модели в классе кусочно-линейных функций с фиксированными узлами. Вводится понятие линейного разбиения корреляционного поля, его частичных участков и его узлов. С помощью данного разбиения корреляционного поля определяется класс кусочно-линейных функций с фиксированными узлами. Линейное разбиение выявляется в ходе визуального анализа корреляционного поля. С помощью метода наименьших квадратов составляется система линейных уравнений для нахождения точечных оценок параметров аппроксимирующей функции. Исключением двух неизвестных угловых коэффициентов эта система приводится к трехдиагональной системе уравнений, неизвестными которой являются узловое значения аппроксимирующей функции. Трехдиагональная система решается методом прогонки. Сконструирован алгоритм для демонстрации работы метода. Его исходными данными являются массивы значений объясняющей и зависимой переменных, а также массив номеров правых концов интервалов, определяющих узлы. Из массива значений объясняющей переменной и массива номеров правых концов интервалов строится массив фиксированных узлов. Далее строится массив точечных оценок параметров аппроксимирующей функции. Данный алгоритм реализован на языке Python в виде пользовательских функций.

Ключевые слова: корреляционное поле, линейное разбиение корреляционного поля, кусочно-линейная функция с фиксированными узлами, кусочно-линейная аппроксимация, парная регрессионная модель, метод наименьших квадратов, трехдиагональная система, метод прогонки.

Классификация JEL: C20.

УДК: 330.43.

Для цитирования: Зубрилин К.М. (2024). Парная кусочно-линейная регрессионная модель с фиксированными узлами // *Экономика и математические методы*. Т. 60. № 4. С. 113–124. DOI: 10.31857/S0424738824040103

ВВЕДЕНИЕ

Регрессионное моделирование представляет собой математический аппарат исследования экономических процессов, целями которого являются анализ изучаемого процесса, прогноз экономических показателей, моделирование поведения процесса и выработка управленческих решений. Например, регрессионное моделирование используется для прогнозирования финансовых временных рядов, моделирования показателей перевозочного процесса на транспорте, исследования пропускной способности аэропортов и т. п.

Исходными данными парной регрессионной модели являются наборы значений объясняющей и зависимой переменных, которые геометрически изображаются точками координатной плоскости, образуя корреляционное поле.

Моделирование начинается с определения вида аппроксимирующей функции. В результате предварительного анализа корреляционного поля выявляется класс функций, среди представителей которого находится аппроксимирующая функция. Представитель класса — математическое выражение, содержащее объясняющую переменную и некоторые параметры. Когда параметры независимо друг от друга пробегут по множествам допустимых значений, представитель пробежит весь класс.

Следующий этап — применение метода наименьших квадратов для нахождения оценок параметров аппроксимирующей функции. Эти оценки образуют решение оптимизационной задачи о минимуме суммы квадратов отклонений наблюдаемых значений объясняемой переменной от теоретических значений, рассчитанных на основании аппроксимирующей функции.

На заключительном этапе моделирования проверяется адекватность построенной модели, т.е. соответствует ли построенная модель априорным данным.

Постмодельный этап использует основные предпосылки регрессионного анализа для проверки значимости параметров аппроксимирующей функции — достаточно ли объясняющей переменной для описания зависимой переменной.

Парная линейная регрессионная модель является самой распространенной ввиду простоты получения ее оценок как точечных, так и интервальных. Тем не менее реальные физические и социально-экономические процессы, как правило, имеют более сложную, нелинейную природу. Существуют априорные приемы выявления класса нелинейных функций (степенных, показательных, логарифмических и т.д.), к которому принадлежит аппроксимирующая функция. Парные нелинейные регрессионные модели, перечисленные выше, относятся к числу линеаризующихся моделей. Они надлежащей заменой переменных приводятся к линейной вспомогательной модели. Нахождение оценок параметров нелинейной модели, а также проверка значимости этих параметров осуществляется на основании вспомогательной линейной модели.

В большинстве случаев изучаемые нелинейные модели не линеаризуются. Здесь можно применить полиномиальную модель и попытаться улучшить ее адекватность повышая степень аппроксимирующего многочлена. Но это ведет к усложнению точечных оценок. Вопрос же о получении интервальных оценок параметров и проверки их значимости остается открытым.

Более эффективный путь — использование кусочно-непрерывной аппроксимации с известными кусками. Если в качестве кусков брать линейные функции, то получатся кусочно-линейные модели. Они различаются методами построения аппроксимирующих кусочно-линейных функций.

В работе (Боровской, Костелей, 2017, с. 73) применяется алгоритм кусочно-линейной аппроксимации при прогнозировании финансовых временных рядов. Здесь строится набор трендовых линий, разделяющих область интервала и не соединенных между собой. Затем для получения непрерывной трендовой модели используется механизм сглаживания. В результате получается трендовая модель, состоящая из последовательно соединенных трендов.

В работах (Носков, 2013, с. 135; Иванова, Лебедева, Носков, 2016, с. 107; Носков, Лоншаков, 2008, с. 63; Носков, Хоняков, 2019, с. 47) рассмотрена кусочно-линейная регрессионная модель, называемая также производственной функцией Леонтьева, или функцией с нулевой эластичностью замены ресурсов. Идентификация параметров этой модели осуществлена при помощи метода наименьших модулей, реализация которого сводится к задаче линейно-булевого программирования.

В работе (Носков, 2023, с. 218) осуществляется формализация вложенной кусочно-линейной регрессии двух типов, являющихся комбинациями традиционной кусочно-линейной регрессии и функции риска. Нахождение оценок параметров вложенной кусочно-линейной регрессии сводится к решению задачи линейно-булевого программирования.

В настоящей работе разрабатывается метод построения аппроксимирующей функции в классе кусочно-линейных функций с фиксированными узлами.

1. КЛАСС КУСОЧНО-ЛИНЕЙНЫХ ФУНКЦИЙ С ФИКСИРОВАННЫМИ УЗЛАМИ

Пусть имеются наборы $(x_i)_{i=0, \dots, n-1}$ и $(y_i)_{i=0, \dots, n-1}$ значений объясняющей X и зависимой Y переменных. Точки (x_i, y_i) , $i = 0, \dots, n-1$ координатной плоскости $x \times y$ образуют корреляционное поле $\mathbf{K}$.

Будем предполагать, что последовательность $(x_i)_{i=0, \dots, n-1}$ строго возрастает, т.е. $x_0 < x_1 < \dots < x_{n-1}$. При необходимости последовательности $(x_i)_{i=0, \dots, n-1}$ и $(y_i)_{i=0, \dots, n-1}$ можно перенумеровать так, чтобы условие возрастания выполнялось.

Для точек u' и u'' таких, что $\{u', u''\} \cap \{x_i : i = 0, \dots, n-1\} = \emptyset$ и $u' < u''$, множество $K_{u'; u''}$ точек (x_i, y_i) , $i = 0, \dots, n-1$ корреляционного поля $\mathbf{K}$, удовлетворяющих условию $u' < x_i < u''$, будем называть *полосовым участком*, а точки u' и u'' его узлами, левым и правым соответственно. Случай бесконечных узлов не исключается.

Узлы полосового участка $K_{u'; u''}$ определяются неоднозначно. Если $u' < x_0$, то в качестве левого узла можно взять любое число интервала $(-\infty; x_0)$ или $-\infty$. Если $x_0 < u' < x_{n-1}$, то найдется такое натуральное число k , что $0 < k < n-1$ и $x_{k-1} < u' < x_k$. В этом случае каждое число интервала $(x_{k-1}; x_k)$ может быть

взято в качестве левого узла. Если $x_{n-1} < u''$, то правым узлом может быть любое число интервала $(x_{n-1}; +\infty)$ или $+\infty$. Если $x_0 < u'' < x_{n-1}$, то найдется такое натуральное число l , что $0 < l \leq n-1$ и $x_{l-1} < u'' < x_l$. Тогда каждое число интервала $(x_{l-1}; x_l)$ может быть взято в качестве правого узла. Вопрос выбора хороших узлов, т.е. узлов, удовлетворяющих нужным условиям (например, оптимальности), остается открытым. В общем случае, когда на узлы не накладываются условия, левым узлом будем брать $-\infty$ или середину интервала $(x_{k-1}; x_k)$, а правым узлом $+\infty$ или середину интервала $(x_{l-1}; x_l)$.

С геометрической точки зрения для конечных узлов u' и u'' прямые $x = u'$ и $x = u''$ вырезают из корреляционного поля полосовой участок. Эти прямые будем называть *узловыми*.

Полосовым разбиением корреляционного поля будем называть конечную последовательность непустых попарно непересекающихся полосовых участков, объединение которых совпадает с корреляционным полем. При этом для каждого полосового участка этой последовательности, кроме последнего, его правый узел совпадает с левым узлом следующего участка данной последовательности. Элементы полосового разбиения будем называть *частичными участками разбиения*. Таким образом, последовательность $\mathbf{R} = (K_{u'_j; u''_j}^{(j)})_{j=1, \dots, m}$ полосовых участков является полосовым разбиением, если:

- 1) $K_{u'_j; u''_j}^{(j)} \neq \emptyset \quad \forall j = 1, \dots, m$;
- 2) $K_{u'_j; u''_j}^{(j)} \cap K_{u'_l; u''_l}^{(l)} = \emptyset \quad \forall j = 1, \dots, m; \quad \forall l = 1, \dots, m; \quad j \neq l$;
- 3) $u''_j = u'_{j+1} \quad \forall j = 1, \dots, m-1$.

Если число частичных участков полосового разбиения равно m , то их попарно различные узлы, расположенные в порядке возрастания, обозначим как

$$a, u^{(0)}, u^{(1)}, \dots, u^{(m-2)}, b. \quad (1)$$

Причем в общем случае, $a = -\infty$ и $b = +\infty$. Тогда $K_{a; u^{(0)}}$ — первый частичный участок разбиения; $K_{u^{(0)}; u^{(1)}}$ — второй, ..., $K_{u^{(m-2)}; b}$ — участок m . Найдется такая возрастающая последовательность $(s_j)_{j=1, \dots, m-1}$ первых $n-1$ натуральных чисел, что $x_{s_{j+1}-1} < u^{(j)} < x_{s_{j+1}}$ для каждого $j = 0, \dots, m-1$. Для удобства введем обозначения: $s_0 = 0$, $s_m = n$. Тогда:

$$\begin{aligned} K_{a; u^{(0)}} &= \{(x_i, y_i) : i = s_0, \dots, s_1 - 1\}; \\ K_{u^{(j-1)}; u^{(j)}} &= \{(x_i, y_i) : i = s_j, \dots, s_{j+1} - 1\}, \quad j = 1, \dots, m-2; \\ K_{u^{(m-2)}; b} &= \{(x_i, y_i) : i = s_{m-1}, \dots, s_m - 1\}. \end{aligned}$$

Геометрически прямые $x = u^{(0)}$, $x = u^{(1)}$, ..., $x = u^{(m-2)}$ разрезают корреляционное поле на полосовые участки, составляющие его полосовое разбиение.

Полосовой участок $K_{u'; u''}$ будем называть *линейным*, если, рассматривая его как отдельное корреляционное поле, можно эффективно применять линейную аппроксимацию. Полосовое разбиение корреляционного поля будем называть *линейным*, если каждый его частичный участок является линейным.

На данный момент нет четкого алгоритма построения линейного разбиения корреляционного поля. Можно предложить следующие правила: во-первых, линейность участка определяется визуально, а во-вторых, число таких линейных участков должно быть наименьшим. Поэтому первичной обработкой исходных данных является визуальный анализ корреляционного поля, результатом которого будет последовательность $(l_j)_{j=0, \dots, m-2}$ номеров правых концов интервалов $(x_{l_{j-1}}; x_{l_j})$, $j = 0, \dots, m-2$ для определения узлов $u^{(j)} \in (x_{l_{j-1}}; x_{l_j})$, $j = 0, \dots, m-2$ линейного разбиения $\mathbf{R}$.

Для каждого частичного участка линейного разбиения $\mathbf{R}$ на отрезке с концами в узлах построим линейную функцию таким образом, чтобы для любых двух соседних частичных участков данного разбиения в их общем узле значения линейных функций совпадали. В результате на интервале $(a; b)$ будет построена кусочно-линейная функция φ с фиксированными узлами (1):

$$\varphi(x) = \begin{cases} k^{(0)}x + l^{(0)}, & x < u^{(0)}, \\ \dots & \\ k^{(j)}x + l^{(j)}, & u^{(j-1)} < x < u^{(j)}, \quad j = 1, \dots, m-2, \\ \dots & \\ k^{(m-1)}x + l^{(m-1)}, & u^{(m-2)} < x, \end{cases} \quad (2)$$

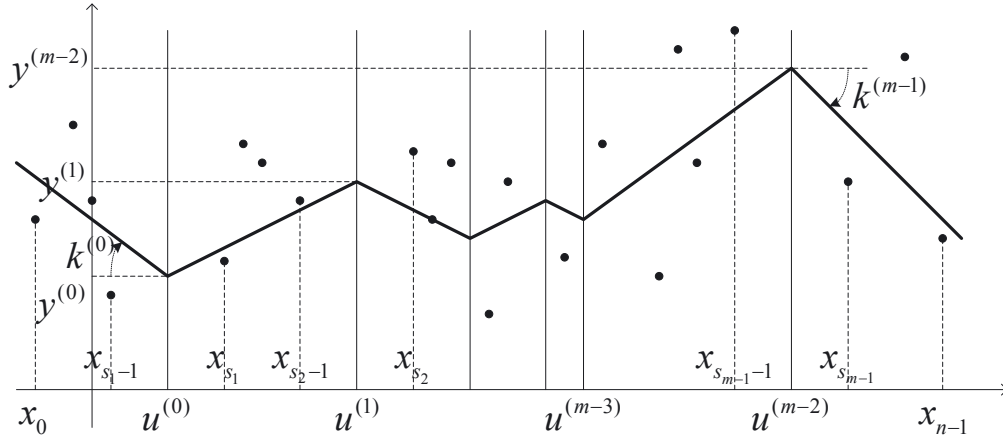


Рис. 1. Кусочно-линейная функция, определенная на линейном разбиении корреляционного поля

где

$$k^{(j)}u^{(j)} + l^{(j)} = k^{(j+1)}u^{(j)} + l^{(j+1)}, \quad j = 0, \dots, m-2. \quad (3)$$

Значения кусочно-линейной функции φ в узлах $y^{(j)} = \varphi(u^{(j)}) = k^{(j)}u^{(j)} + l^{(j)}$ называют узловыми значениями.

Используя элементарные преобразования, с учетом условий (3), легко получить выражения для коэффициентов $l^{(0)}$, $k^{(j)}$, $l^{(j)}$, $j = 1, \dots, m-2$, $l^{(m-1)}$ через узлы и узловые значения. В таком случае кусочно-линейная функция φ примет вид

$$\varphi(x) = \begin{cases} y^{(0)} + k^{(0)}(x - u^{(0)}), & x \leq u^{(0)}, \\ \dots \\ \frac{y^{(i)}}{u^{(i)} - u^{(i-1)}}(x - u^{(i-1)}) - \frac{y^{(i-1)}}{u^{(i)} - u^{(i-1)}}(x - u^{(i)}), & u^{(i-1)} \leq x \leq u^{(i)}, \quad i = 1, \dots, m-2 \\ \dots \\ y^{(m-2)} + k^{(m-1)}(x - u^{(m-2)}), & u^{(m-2)} \leq x. \end{cases} \quad (4)$$

Геометрически (рис. 1) левый кусок изображается лучом с началом в точке $(u^{(0)}; y^{(0)})$, правый кусок — лучом с началом в точке $(u^{(m-2)}; y^{(m-2)})$, промежуточный кусок i — отрезком с концами в точках $(u^{(i-1)}; y^{(i-1)})$ и $(u^{(i)}; y^{(i)})$ для каждого $i = 1, \dots, m-2$. Эти отрезки являются звеньями ломаной линии.

Из (4) видно, что кусочно-линейная функция φ зависит от $m+1$ параметров. Из них $m-1$ узловых значений $y^{(0)}, \dots, y^{(m-2)}$, перемещающих вершины ломаной вдоль узловых прямых $x = u^{(0)}, \dots, x = u^{(m-2)}$; $k^{(0)}$ — угловой коэффициент левого луча, поворачивающий его вокруг начала в точке $(u^{(0)}, y^{(0)})$; $k^{(m-1)}$ — угловой коэффициент правого луча, позволяющий поворачивать его вокруг начала $(u^{(m-2)}, y^{(m-2)})$.

Таким образом, мы построили класс кусочно-линейных функций с фиксированными узлами. Именно в этом классе, методом наименьших квадратов, мы будем искать аппроксимирующую функцию.

2. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ В КЛАССЕ КУСОЧНО-ЛИНЕЙНЫХ НЕПРЕРЫВНЫХ ФУНКЦИЙ С ФИКСИРОВАННЫМИ УЗЛАМИ

Сумма квадратов отклонений точек корреляционного поля K от кусочно-линейной непрерывной функции (4) имеет вид:

$$\begin{aligned} S &= \sum_{i=0}^{n-1} (y_i - \varphi(x_i))^2 = \sum_{j=0}^{m-1} \sum_{i=s_j}^{s_{j+1}-1} (y_i - \varphi(x_i))^2 = \sum_{j=0}^{m-2} (y_i - \varphi(x_i))^2 + \sum_{j=1}^{m-2} \sum_{i=s_j}^{s_{j+1}-1} (y_i - \varphi(x_i))^2 + \sum_{i=s_{m-1}}^{s_m-1} (y_i - \varphi(x_i))^2 = \\ &= \sum_{i=s_0}^{s_1-1} (y_i - y^{(0)} - k^{(0)}(x_i - u^{(0)}))^2 + \sum_{j=1}^{m-2} \sum_{i=s_j}^{s_{j+1}-1} \left(y_i - \frac{y^{(j)}}{u^{(j)} - u^{(j-1)}}(x_i - u^{(j-1)}) + \frac{y^{(j-1)}}{u^{(j)} - u^{(j-1)}}(x_i - u^{(j)}) \right)^2 + \\ &\quad + \sum_{i=s_{m-1}}^{s_m-1} (y_i - y^{(m-2)} - k^{(m-1)}(x_i - u^{(m-2)}))^2. \end{aligned}$$

Обозначим $x_i^{(j)} = x_i - u^{(j)}$, $j = 0, \dots, m-2$, тогда

$$S = \sum_{i=s_0}^{s_1-1} (y_i - y^{(0)} - k^{(0)} x_i^{(0)})^2 + \sum_{j=1}^{m-2} \sum_{i=s_j}^{s_{j+1}-1} \left(y_i - \frac{y^{(j)}}{u^{(j)} - u^{(j-1)}} x_i^{(j-1)} + \frac{y^{(j-1)}}{u^{(j)} - u^{(j-1)}} x_i^{(j)} \right)^2 + \sum_{i=s_{m-1}}^{s_m-1} (y_i - y^{(m-2)} - k^{(m-1)} x_i^{(m-2)})^2. \quad (5)$$

Согласно методу наименьших квадратов оценки параметров функции регрессии, представленной кусочно-линейной непрерывной функцией с фиксированными узлами, находятся как решение оптимизационной модели $S = S(k^{(0)}, k^{(m-1)}, y^{(0)}, y^{(1)}, \dots, y^{(m-2)}) \rightarrow \min$.

Необходимые условия локального экстремума функции от нескольких переменных имеют вид:

$$\frac{\partial S}{\partial k^{(0)}} = 0, \quad \frac{\partial S}{\partial y^{(0)}} = 0, \quad \dots, \quad \frac{\partial S}{\partial k^{(m-1)}} = 0. \quad (6)$$

Введем в рассмотрение величины групповых сумм:

$$\begin{aligned} S_p(x^{(p)2}) &= \sum_{i=s_p}^{s_{p+1}-1} x_i^{(p)2}, \quad S_p(x^{(p)}y) = \sum_{i=s_p}^{s_{p+1}-1} x_i^{(p)} y_i, \quad S_{p+1}(x^{(p)}y) = \sum_{i=s_{p+1}}^{s_{p+2}-1} x_i^{(p)} y_i, \quad S_{p+1}(x^{(p)2}) = \sum_{i=s_{p+1}}^{s_{p+2}-1} x_i^{(p)2}, \\ S_p(x^{(p-1)}x^{(p)}) &= \sum_{i=s_p}^{s_{p+1}-1} x_i^{(p-1)} x_i^{(p)}, \quad p = 0, \dots, m-2; \\ S_0(x^{(0)}) &= \sum_{i=s_0}^{s_1-1} x_i^{(0)}, \quad S_0(y) = \sum_{i=s_0}^{s_1-1} y_i, \quad S_{m-1}(x^{(m-2)}) = \sum_{i=s_{m-1}}^{s_m-1} x_i^{(m-2)}, \quad S_{m-1}(y) = \sum_{i=s_{m-1}}^{s_m-1} y_i. \end{aligned}$$

Тогда частные производные от S , с учетом (5), примут следующий вид.

Частная производная по $k^{(0)}$:

$$\frac{\partial S}{\partial k^{(0)}} = -2 \sum_{i=s_0}^{s_1-1} (y_i - y^{(0)} - k^{(0)} x_i^{(0)}) x_i^{(0)}.$$

Отсюда

$$0,5 \frac{\partial S}{\partial k^{(0)}} = -S_0(x^{(0)}y) + y^{(0)} S_0(x^{(0)}) + k^{(0)} S_0(x^{(0)2}). \quad (7)$$

Частная производная по $y^{(0)}$:

$$\frac{\partial S}{\partial y^{(0)}} = -2 \sum_{i=s_0}^{s_1-1} (y_i - y^{(0)} - k^{(0)} x_i^{(0)}) + 2 \sum_{i=s_1}^{s_2-1} \left(y_i - \frac{y^{(1)}}{u^{(1)} - u^{(0)}} x_i^{(0)} + \frac{y^{(0)}}{u^{(1)} - u^{(0)}} x_i^{(1)} \right) \frac{1}{u^{(1)} - u^{(0)}} x_i^{(1)}.$$

Тогда

$$0,5 \frac{\partial S}{\partial y^{(0)}} = k^{(0)} S_0(x^{(0)}) + y^{(0)} \left(n^{(0)} + \frac{S_1(x^{(1)2})}{(u^{(1)} - u^{(0)})^2} \right) - y^{(1)} \frac{S_1(x^{(0)}x^{(1)})}{(u^{(1)} - u^{(0)})^2} - S_0(y) + \frac{S_1(x^{(1)}y)}{u^{(1)} - u^{(0)}}. \quad (8)$$

Для произвольного $p = 1, \dots, m-3$ находим частную производную по $y^{(p)}$:

$$\begin{aligned} \frac{\partial S}{\partial y^{(p)}} &= -2 \sum_{i=s_p}^{s_{p+1}-1} \left(y_i - \frac{y^{(p)}}{u^{(p)} - u^{(p-1)}} x_i^{(p-1)} + \frac{y^{(p-1)}}{u^{(p)} - u^{(p-1)}} x_i^{(p)} \right) \frac{1}{u^{(p)} - u^{(p-1)}} x_i^{(p-1)} + \\ &+ 2 \sum_{i=s_{p+1}}^{s_{p+2}-1} \left(y_i - \frac{y^{(p+1)}}{u^{(p+1)} - u^{(p)}} x_i^{(p)} + \frac{y^{(p)}}{u^{(p+1)} - u^{(p)}} x_i^{(p+1)} \right) \frac{1}{u^{(p+1)} - u^{(p)}} x_i^{(p+1)}. \end{aligned}$$

Тогда

$$\begin{aligned} 0,5 \frac{\partial S}{\partial y^{(p)}} &= -y^{(p-1)} \frac{S_p(x^{(p-1)}x^{(p)})}{(u^{(p)} - u^{(p-1)})^2} + y^{(p)} \left(\frac{S_p(x^{(p-1)2})}{(u^{(p)} - u^{(p-1)})^2} + \frac{S_{p+1}(x^{(p+1)2})}{(u^{(p+1)} - u^{(p)})^2} \right) - \\ &- y^{(p+1)} \frac{S_{p+1}(x^{(p)}x^{(p+1)})}{(u^{(p+1)} - u^{(p)})^2} + \frac{S_{p+1}(x^{(p+1)}y)}{u^{(p+1)} - u^{(p)}} - \frac{S_p(x^{(p-1)}y)}{u^{(p)} - u^{(p-1)}}. \end{aligned} \quad (9)$$

Частная производная по $y^{(m-1)}$:

$$\frac{\partial S}{\partial y^{(m-2)}} = -2 \sum_{i=s_{m-2}}^{s_{m-1}-1} \left(y_i - \frac{y^{(m-2)} x_i^{(m-3)}}{u^{(m-2)} - u^{(m-3)}} + \frac{y^{(m-3)} x_i^{(m-2)}}{u^{(m-2)} - u^{(m-3)}} \right) \frac{x_i^{(m-3)}}{u^{(m-2)} - u^{(m-3)}} - 2 \sum_{i=s_{m-1}}^{s_m-1} (y_i - y^{(m-2)} - k^{(m-1)} x_i^{(m-2)}).$$

Тогда

$$0,5 \frac{\partial S}{\partial y^{(m-2)}} = -y^{(m-3)} \frac{S_{m-2}(x^{(m-3)}x^{(m-2)})}{(u^{(m-2)} - u^{(m-3)})^2} + y^{(m-2)} \left(\frac{S_{m-2}(x^{(m-3)2})}{(u^{(m-2)} - u^{(m-3)})^2} + n^{(m-1)} \right) +$$

$$+ k^{(m-1)} S_{m-1}(x^{(m-2)}) - \frac{S_{m-2}(x^{(m-3)}y)}{u^{(m-2)} - u^{(m-3)}} - S_{m-1}(y). \quad (10)$$

Частная производная по $k^{(m-1)}$:

$$\frac{\partial S}{\partial k^{(m-1)}} = -2 \sum_{i=s_{m-1}}^{s_{m-1}-1} (x_i^{(m-2)} y_i - y^{(m-2)} x_i^{(m-2)} - k^{(m-1)} x_i^{(m-2)2}).$$

Тогда

$$0,5 \frac{\partial S}{\partial k^{(m-1)}} = y^{(m-2)} S_{m-1}(x^{(m-2)}) + k^{(m-1)} S_{m-1}(x^{(m-2)2}) - S_{m-1}(x^{(m-2)}y). \quad (11)$$

Подставляя правые части равенств (7)–(11) в систему (6), получим систему уравнений:

$$k^{(0)} S_0(x^{(0)2}) + y^{(0)} S_0(x^{(0)}) = S_0(x^{(0)}y),$$

$$k^{(0)} S_0(x^{(0)}) + y^{(0)} \left(n^{(0)} + \frac{S_1(x^{(1)2})}{(u^{(1)} - u^{(0)})^2} \right) - y^{(1)} \frac{S_1(x^{(0)}x^{(1)})}{(u^{(1)} - u^{(0)})^2} = S_0(y) - \frac{S_1(x^{(1)}y)}{u^{(1)} - u^{(0)}}, \quad p=1, \dots, m-3;$$

$$-y^{(p-1)} \frac{S_p(x^{(p-1)}x^{(p)})}{(u^{(p)} - u^{(p-1)})^2} + y^{(p)} \left(\frac{S_p(x^{(p-1)2})}{(u^{(p)} - u^{(p-1)})^2} + \frac{S_{p+1}(x^{(p+1)2})}{(u^{(p+1)} - u^{(p)})^2} \right) - y^{(p+1)} \frac{S_{p+1}(x^{(p)}x^{(p+1)})}{(u^{(p+1)} - u^{(p)})^2} =$$

$$= \frac{S_p(x^{(p-1)}y)}{u^{(p)} - u^{(p-1)}} - \frac{S_{p+1}(x^{(p+1)}y)}{u^{(p+1)} - u^{(p)}}, \quad (12)$$

$$-y^{(m-3)} \frac{S_{m-2}(x^{(m-3)}x^{(m-2)})}{(u^{(m-2)} - u^{(m-3)})^2} + y^{(m-2)} \left(\frac{S_{m-2}(x^{(m-3)2})}{(u^{(m-2)} - u^{(m-3)})^2} + n^{(m-1)} \right) + k^{(m-1)} S_{m-1}(x^{(m-2)}) =$$

$$= \frac{S_{m-2}(x^{(m-3)}y)}{u^{(m-2)} - u^{(m-3)}} + S_{m-1}(y),$$

$$y^{(m-2)} S_{m-1}(x^{(m-2)}) + k^{(m-1)} S_{m-1}(x^{(m-2)2}) = S_{m-1}(x^{(m-2)}y).$$

Из первого уравнения системы (12) находим

$$k^{(0)} = \frac{S_0(x^{(0)}y)}{S_0(x^{(0)2})} - y^{(0)} \frac{S_0(x^{(0)})}{S_0(x^{(0)2})}. \quad (13)$$

Правую часть равенства (13) подставим во второе уравнение системы (12):

$$y^{(0)} \left(n^{(0)} + \frac{S_1(x^{(1)2})}{(u^{(1)} - u^{(0)})^2} - \frac{S_0(x^{(0)})^2}{S_0(x^{(0)2})} \right) - y^{(1)} \frac{S_1(x^{(0)}x^{(1)})}{(u^{(1)} - u^{(0)})^2} = S_0(y) - \frac{S_1(x^{(1)}y)}{u^{(1)} - u^{(0)}} - \frac{S_0(x^{(0)}y)S_0(x^{(0)})}{S_0(x^{(0)2})}.$$

Фактически мы из второго уравнения системы (12) исключили $k^{(0)}$. Подобным образом из последнего равенства системы (12) будем иметь

$$k^{(m-1)} = \frac{S_{m-1}(x^{(m-2)}y)}{S_{m-1}(x^{(m-2)2})} - y^{(m-2)} \frac{S_{m-1}(x^{(m-2)})}{S_{m-1}(x^{(m-2)2})}. \quad (14)$$

С помощью равенства (14) из предпоследнего уравнения системы (12) исключаем $k^{(m-1)}$ и приходим к уравнению:

$$-y^{(m-3)} \frac{S_{m-2}(x^{(m-3)}x^{(m-2)})}{(u^{(m-2)} - u^{(m-3)})^2} + y^{(m-2)} \left(\frac{S_{m-2}(x^{(m-3)2})}{(u^{(m-2)} - u^{(m-3)})^2} + n^{(m-1)} - \frac{S_{m-1}(x^{(m-2)})^2}{S_{m-1}(x^{(m-2)2})} \right) =$$

$$= \frac{S_{m-2}(x^{(m-3)}y)}{u^{(m-2)} - u^{(m-3)}} + S_{m-1}(y) - \frac{S_{m-1}(x^{(m-2)}y)S_{m-1}(x^{(m-2)})}{S_{m-1}(x^{(m-2)2})}.$$

К двум уравнениям, полученным из второго и предпоследнего уравнения системы (12) исключением $k^{(0)}$ и $k^{(m-1)}$, добавляем неиспользованные уравнения системы (12). Приходим к системе уравнений:

$$\begin{cases} a_{0,0}y^{(0)} + a_{0,1}y^{(1)} = b_0, \\ a_{p,p-1}y^{(p-1)} + a_{p,p}y^{(p)} + a_{p,p+1}y^{(p+1)} = b_p, \quad p = 1, \dots, m-3, \\ a_{m-2,m-3}y^{(m-3)} + a_{m-2,m-2}y^{(m-2)} = b_{m-2}, \end{cases} \quad (15)$$

коэффициенты которой определяются равенствами:

$$a_{0,0} = n^{(0)} + \frac{S_1(x^{(1)2})}{(u^{(1)} - u^{(0)})^2} - \frac{S_0(x^{(0)2})}{S_0(x^{(0)2})}, \quad a_{0,1} = -\frac{S_1(x^{(0)}x^{(1)})}{(u^{(1)} - u^{(0)})^2}, \quad b_0 = S_0(y) - \frac{S_1(x^{(1)}y)}{u^{(1)} - u^{(0)}} - \frac{S_0(x^{(0)}y)S_0(x^{(0)})}{S_0(x^{(0)2})}; \quad (16)$$

$$a_{p,p-1} = -\frac{S_p(x^{(p-1)}x^{(p)})}{(u^{(p)} - u^{(p-1)})^2}, \quad a_{p,p} = \frac{S_p(x^{(p-1)2})}{(u^{(p)} - u^{(p-1)})^2} + \frac{S_{p+1}(x^{(p+1)2})}{(u^{(p+1)} - u^{(p)})^2}, \quad a_{p,p+1} = -\frac{S_{p+1}(x^{(p)}x^{(p+1)})}{(u^{(p+1)} - u^{(p)})^2}, \quad (17)$$

$$b_p = \frac{S_p(x^{(p-1)}y)}{u^{(p)} - u^{(p-1)}} - \frac{S_{p+1}(x^{(p+1)}y)}{u^{(p+1)} - u^{(p)}}, \quad p = 1, \dots, m-3;$$

$$a_{m-2,m-3} = -\frac{S_{m-2}(x^{(m-3)}x^{(m-2)})}{(u^{(m-2)} - u^{(m-3)})^2}, \quad a_{m-2,m-2} = \frac{S_{m-2}(x^{(m-3)2})}{(u^{(m-2)} - u^{(m-3)})^2} + n^{(m-1)} - \frac{S_{m-1}(x^{(m-2)2})}{S_{m-1}(x^{(m-2)2})}, \quad (18)$$

$$b_{m-2} = \frac{S_{m-2}(x^{(m-3)}y)}{u^{(m-2)} - u^{(m-3)}} + S_{m-1}(y) - \frac{S_{m-1}(x^{(m-2)}y)S_{m-1}(x^{(m-2)})}{S_{m-1}(x^{(m-2)2})}, \quad p = m-2.$$

Именно точечные оценки параметров $y^{(0)}, \dots, y^{(m-2)}$ находятся как решения системы уравнений (15).

Система уравнений (15) называется *трехдиагональной*. Известен *метод прогонки* решения такой системы (см., например, (Ильин, Кузнецов, 1985, с. 50)). Для полноты изложения приведем его в разд. 3.

3. МЕТОД ПРОГОНКИ РЕШЕНИЯ ТРЕХДИАГОНАЛЬНОЙ СИСТЕМЫ УРАВНЕНИЙ

Главные миноры матрицы системы (15) обладают свойствами:

$$M_0 = a_{0,0}, \quad M_1 = \begin{vmatrix} a_{0,0} & a_{0,1} \\ a_{1,0} & a_{1,1} \end{vmatrix} = a_{0,0}a_{1,1} - a_{1,0}a_{0,1} = a_{1,1}M_0 - a_{1,0}a_{0,1}. \quad (19)$$

Раскладывая по последней строке, будем иметь

$$M_p = \begin{vmatrix} a_{0,0} & a_{0,1} & 0 & \dots & 0 & 0 & 0 \\ a_{1,0} & a_{1,1} & a_{1,2} & \dots & 0 & 0 & 0 \\ 0 & a_{2,1} & a_{2,2} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{p-2,p-2} & a_{p-2,p-1} & 0 \\ 0 & 0 & 0 & \dots & a_{p-1,p-2} & a_{p-1,p-1} & a_{p-1,p} \\ 0 & 0 & 0 & \dots & 0 & a_{p,p-1} & a_{p,p} \end{vmatrix} = a_{p,p}M_{p-1} - a_{p,p-1}a_{p-1,p}M_{p-2}.$$

Вспомогательные определители получаются из главных миноров заменой последнего столбца столбцом свободных членов:

$$M_{0;b} = b_0, \quad M_{1;b} = \begin{vmatrix} a_{0,0} & b_0 \\ a_{1,0} & b_1 \end{vmatrix} = a_{0,0}b_1 - a_{1,0}b_0 = b_1M_0 - a_{1,0}M_{0;b}. \quad (20)$$

Раскладываем по последней строке

$$M_{p;b} = \begin{vmatrix} a_{0,0} & a_{0,1} & 0 & \dots & 0 & 0 & b_0 \\ a_{1,0} & a_{1,1} & a_{1,2} & \dots & 0 & 0 & b_1 \\ 0 & a_{2,1} & a_{2,2} & \dots & 0 & 0 & b_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{p-2,p-2} & a_{p-2,p-1} & b_{p-2} \\ 0 & 0 & 0 & \dots & a_{p-1,p-2} & a_{p-1,p-1} & b_{p-1} \\ 0 & 0 & 0 & \dots & 0 & a_{p,p-1} & b_p \end{vmatrix} = b_pM_{p-1} - a_{p,p-1}M_{p-1;b}.$$

Из первого уравнения системы (15) получим

$$y^{(0)} = \frac{b_0}{a_{0,0}} - \frac{a_{0,1}}{a_{0,0}} y^{(1)}. \quad (21)$$

Методом математической индукции доказывается, что выполняются равенства

$$y^{(p)} = \frac{M_{p;b}}{M^p} - \frac{M_{p-1} a_{p,p+1}}{M^p} y^{(p+1)}, \quad p = 1, \dots, m-3; \quad (22)$$

$$y^{(m-3)} = \frac{M_{m-3;b}}{M_{m-3}^p} - \frac{M_{m-4} a_{m-3,m-2}}{M_{m-3}^p} y^{(m-2)}, \quad p = m-3. \quad (23)$$

Уравнение (23) и последнее уравнение системы (15) приводят к системе двух линейных уравнений с переменными $y^{(m-3)}$ и $y^{(m-2)}$.

Методом Крамера решения систем линейных уравнений приходим к равенству

$$y^{(m-2)} = M_{m-2;b} / M_{m-2}. \quad (24)$$

Подставляя правую часть (24) в (23), находим $y^{(m-3)}$. Применяя к равенству (22) индукцию спуска, получим все значения $y^{(p)}$ от $p = m-4$ до $p = 1$ включительно. Подставляя $y^{(1)}$ в (21), находим $y^{(0)}$, что завершает решение системы (15).

Для решения системы (12) необходимо выполнить еще два шага. Значение $y^{(m-2)}$ подставить в равенство (14) и определить $k^{(m-1)}$, а $y^{(0)}$ подставить в равенство (13) и найти $k^{(0)}$. Решив систему (12), мы найдем точечные оценки параметров аппроксимирующей функции (4).

Равенства (21)–(24) совместно с (13) и (14) составляют каркас алгоритма определения точечных оценок параметров аппроксимирующей функции. Остановимся на этом подробнее.

4. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АППРОКСИМИРУЮЩЕЙ ФУНКЦИИ В КЛАССЕ КУСОЧНО-ЛИНЕЙНЫХ ФУНКЦИЙ С ФИКСИРОВАННЫМИ УЗЛАМИ

Входными данными являются массивы значений $(x_i)_{i=0,\dots,n-1}$, $(y_i)_{i=0,\dots,n-1}$ и $(u^{(j)})_{j=0,\dots,m-2}$. По ним вычисляются групповые суммы, которые сохраняются в специальных переменных и массивах. Затем находим коэффициенты и свободные члены системы (15) и сохраняем их значения в отдельном массиве. После этого рассчитываем значения главных миноров и вспомогательных определителей и сохраняем их в двух массивах. Применяя равенства (21)–(24) к элементам двух последних массивов, мы найдем $y^{(p)}$ для $p = 0, \dots, m-2$. В заключение равенствами (13) и (14) получаем $k^{(0)}$, $k^{(m-1)}$. Алгоритм возвращает массив значений параметров $y^{(p)}$, $p = 0, \dots, m-2$, $k^{(0)}$, $k^{(m-1)}$.

Данный алгоритм реализован на языке Python в виде функции `Gluing_m_Linear_Pieces()`.

Вычисление значения аппроксимирующей функции в точке реализуется с помощью функции `mLinearGluing()`, а на векторе — с помощью функции `mLinearGluingV()`.

Напомним, что последовательность узлов $(u^{(j)})_{j=0,\dots,m-2}$ построена в разд. 1 по номерам l_j , $j = 0, \dots, m-2$ правых концов интервалов $(x_{l_{j-1}}; x_{l_j})$. Функция `IndexKnot()` принимает массивы $(l_j)_{j=0,\dots,m-2}$ и $(x_i)_{i=0,\dots,n-1}$ и возвращает массив узлов $(u^{(j)})_{j=0,\dots,m-2}$ как середины $u^{(j)} = 0,5(x_{l_{j-1}} + x_{l_j})$, $j = 0, \dots, m-2$ соответствующих интервалов.

Также организована функция `mLinearGluingDeter()` для вычисления коэффициента детерминации R^2 .

5. ПРИМЕР

Работу алгоритма из разд. 4 проиллюстрируем на следующем примере. В табл. 1 приведены значения $(x_i)_{i=0,\dots,45}$, $(y_i)_{i=0,\dots,45}$ соответствующих значений объясняющей X и зависимой Y переменных. Согласно алгоритму, составляем корреляционное поле (рис. 2).

Проводим визуальный анализ на обнаружение линейного разбиения данного корреляционного поля. Выбираем интервалы: $(6;6,5)$, $(9,5;10)$, $(15;15,5)$, $(17;17,5)$, $(20,5;21)$. Находим их середины

$$u^{(0)} = 0,5(6 + 6,5) = 6,25; \quad u^{(1)} = 0,5(9,5 + 10) = 9,75; \quad u^{(2)} = 0,5(15 + 15,5) = 15,25; \\ u^{(3)} = 0,5(17 + 17,5) = 17,25; \quad u^{(4)} = 0,5(20,5 + 21) = 20,75.$$

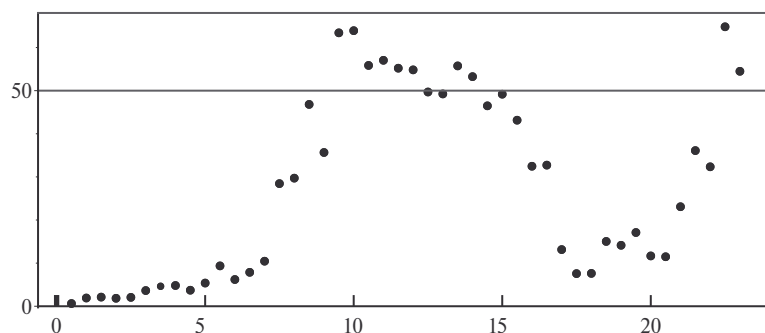


Рис. 2. Корреляционное поле

Примечание. Рис. 2–4 выполнены в Python с помощью пакета расширений matplotlib.

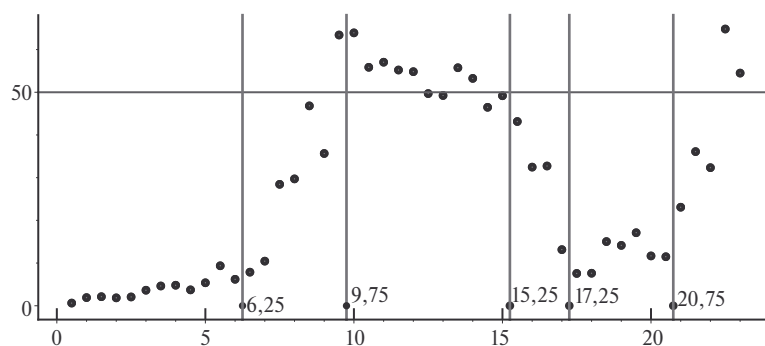


Рис. 3. Визуальный анализ корреляционного поля с построением линейного разбиения

Таблица 1. Значения соответствующих значений объясняющей X и зависимой Y переменных

i	x	y	i	x	y	i	x	y
1	0,5	0,63	17	8,5	46,81	33	16,5	32,73
2	1	1,91	18	9	35,66	34	17	13,14
3	1,5	2,12	19	9,5	63,41	35	17,5	7,57
4	2	1,85	20	10	63,91	36	18	7,62
5	2,5	2,06	21	10,5	55,85	37	18,5	15,03
6	3	3,64	22	11	57,04	38	19	14,14
7	3,5	4,63	23	11,5	55,21	39	19,5	17,1
8	4	4,81	24	12	54,82	40	20	11,67
9	4,5	3,72	25	12,5	49,69	41	20,5	11,48
10	5	5,38	26	13	49,24	42	21	23,08
11	5,5	9,36	27	13,5	55,74	43	21,5	36,1
12	6	6,2	28	14	53,25	44	22	32,34
13	6,5	7,87	29	14,5	46,48	45	22,5	64,81
14	7	10,45	30	15	49,18	46	23	54,49
15	7,5	28,44	31	15,5	43,15			
16	8	29,71	32	16	32,47			

Вертикальные прямые $x = 6,25$, $x = 9,75$, $x = 15,25$, $x = 17,25$, $x = 20,75$ разбивают корреляционное поле на шесть полосовых участков. В пределах каждого участка можно эффективно проводить линейную аппроксимацию. Следовательно, все полосовые участки являются линейными, образуя линейное разбиение данного корреляционного поля, а точки 6,25; 9,75; 15,25; 17,25 и 20,75 — его конечными узлами (рис. 3).

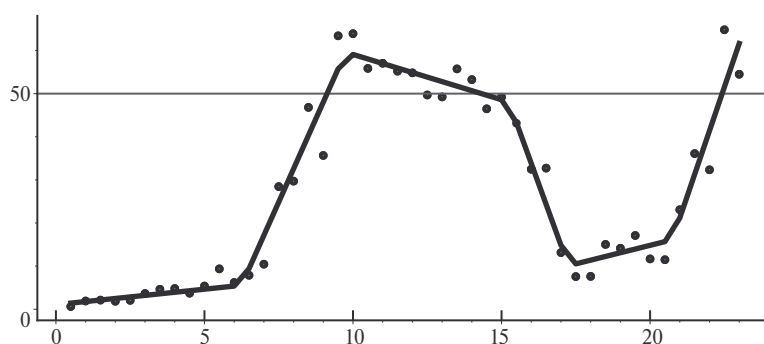


Рис. 4. Корреляционное поле и кусочно-линейная аппроксимирующая функция

Таблица 2. Значения кусочно-линейной аппроксимирующей функции

i	$\varphi(x_i)$	i	$\varphi(x_i)$	i	$\varphi(x_i)$	i	$\varphi(x_i)$	i	$\varphi(x_i)$
1	1,42	11	5,008	21	58,042	31	43,314	41	15,676
2	1,779	12	5,366	22	56,991	32	33,822	42	21,163
3	2,137	13	9,408	23	55,94	33	24,331	43	31,279
4	2,496	14	17,133	24	54,89	34	14,839	44	41,394
5	2,855	15	24,857	25	53,839	35	10,523	45	51,509
6	3,214	16	32,582	26	52,788	36	11,381	46	61,625
7	3,572	17	40,307	27	51,737	37	12,24		
8	3,931	18	48,031	28	50,687	38	13,099		
9	4,29	19	55,756	29	49,636	39	13,958		
10	4,649	20	59,093	30	48,585	40	14,817		

К массивам $(x_i)_{i=0,\dots,45}$, $(y_i)_{i=0,\dots,45}$, $(u^{(j)})_{j=0,\dots,4}$ применяем функцию `Gluing_m_LinearPieces()`. Находим массив параметров (коэффициентов) аппроксимирующей функции $y^{(0)} = 5,545729$; $y^{(1)} = 59,61817$; $y^{(2)} = 48,05973$; $y^{(3)} = 10,0931$; $y^{(4)} = 16,10552$; $k^{(0)} = 0,717556$; $k^{(4)} = 20,23084$.

Далее вычисляем значения кусочно-линейной функции в точках x массива $(x_i)_{i=0,\dots,45}$ с помощью функции `mLinearGluing()` и получаем массив значений (табл. 2). С помощью найденных значений строим график аппроксимирующей функции (рис. 4). С помощью функции `mLinearGluingDeter()` вычисляем коэффициент детерминации $R^2 = 0,9571$.

6. ВЫВОДЫ И ПЕРСПЕКТИВЫ ДАЛЬНЕЙШИХ ИССЛЕДОВАНИЙ

Разработан метод построения парной регрессионной модели в классе кусочно-линейных функций с фиксированными узлами. В результате визуального анализа корреляционного поля определяются фиксированные узлы кусочно-линейной функции и строятся групповые суммы. По формулам (16)–(18) вычисляются коэффициенты системы (15), которая решается методом прогонки. Из равенств (13) и (14) находим угловые коэффициенты.

Представлен алгоритм вычисления точечных оценок параметров аппроксимирующей функции в классе кусочно-линейных функций с фиксированными узлами. Этот алгоритм реализован на языке Python в виде пользовательской функции.

Улучшить полученные результаты можно решив следующие задачи.

1. Нахождение оптимальных фиксированных узлов (1). Оптимальность понимается в смысле обеспечения минимума суммы квадратов отклонений (5). При этом областью изменения узла $u^{(j)}$ является интервал $(x_{s_{j+1}-1}; x_{s_{j+1}})$, $j = 0, \dots, m-2$. Если рассматривать сумму квадратов отклонений S как функцию $2m$ переменных $S(k^{(0)}, k^{(m-1)}, y^{(0)}, \dots, y^{(m-2)}, u^{(0)}, \dots, u^{(m-2)})$, то частные производные $\partial S / \partial u^{(\alpha)}$, $\alpha = 0, \dots, m-2$ выражаются через переменные $k^{(0)}, k^{(m-1)}, y^{(0)}, y^{(1)}, \dots, y^{(m-2)}, u^{(0)}, u^{(1)}, \dots, u^{(m-2)}$.

нелинейно. В таком случае уравнения $\partial S / \partial u^{(\alpha)} = 0$, $\alpha = 0, \dots, m-2$ относительно указанных переменных являются нелинейными. Так что при решении системы для отыскания неизвестных оценок можно столкнуться с серьезными трудностями.

2. Используя основные предпосылки регрессионного анализа определить интервальные оценки для параметров $k^{(0)}$, $k^{(m-1)}$, $y^{(0)}$, ..., $y^{(m-2)}$, а также для математического ожидания объясняемой переменной y , и прогнозного значения этой переменной.

3. Нахождение оптимального числа фиксированных узлов кусочно-линейной аппроксимирующей функции. Под оптимальностью понимается наименьшее значение m , при котором достигается наибольшее значение коэффициента детерминации R^2 . Можно попробовать следующий алгоритм. Определяем коэффициент детерминации R_0^2 линейной регрессионной модели. Далее каждый новый узел $u^{(j)}$ добавляем так, чтобы коэффициент детерминации R^2 увеличивался.

4. Разработка универсального алгоритма по определению оптимального числа оптимальных фиксированных узлов кусочно-линейной функции регрессии и нахождение точечных и интервальных оценок ее параметров.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Боровской И.Г., Костелей Я.В.** (2017). Прогнозная модель финансовых рядов на основе кусочно-линейной аппроксимации // *Доклады ТУСУРа*. Т. 20. № 2. С. 73–75. [**Borovskoy I.G., Kosteley Y.V.** (2017). Use of piecewise-linear approximation to identify trends in the financial market. *Proceedings of TUSUR University*, 20, 2, 73–75 (in Russian).]
- Иванова Н.К., Лебедева С.А., Носков С.И.** (2016). Идентификация параметров некоторых негладких регрессий // *Информационные технологии и проблемы математического моделирования в управлении сложными системами*. № 17. С. 107–110. [**Ivanova N.K., Lebedeva S.A., Noskov S.I.** (2016). Identification of parameters of some nonsmooth regressions. *Information Technologies and Problems of Mathematical Modeling in the Management of Complex Systems in the Management of Complex Systems*, 7, 107–110 (in Russian).]
- Ильин В.П., Кузнецов Ю.И.** (1985). *Трёхдиагональные матрицы и их приложения*. Москва: Наука. [**Ilyin V.P., Kuznetsov Yu.I.** (1985). *Tridiagonal matrices and their applications*. Moscow: Nauka (in Russian).]
- Носков С.И.** (2013). Оценивание параметров аппроксимирующей функции с постоянными пропорциями // *Современные технологии. Системный анализ. Моделирование*. № 2. С. 135–136. [**Noskov S.I.** (2013). Estimation of parameters of the approximating function with constant proportions. *Modern Technologies. System Analysis. Modeling*, 2, 135–136 (in Russian).]
- Носков С.И.** (2023). Подход к формализации вложенной кусочно-линейной регрессии // *Международный журнал гуманитарных и естественных наук*. № 1–2 (76). С. 218–220. [**Noskov S.I.** (2023). Approach to formalizing nested piece-linear regression. *International Journal of Humanities and Natural Sciences*, 1–2 (76), 218–220 (in Russian).]
- Носков С.И., Лоншаков Р.В.** (2008). Идентификация параметров кусочно-линейной регрессии // *Информационные технологии и проблемы математического моделирования в управлении сложными системами*. № 6. С. 63–64. [**Noskov S.I., Lonshakov R.V.** (2008). Identification of parameters of piecewise linear regression. *Information Technologies and Problems of Mathematical Modeling in the Management of Complex Systems*, 6, 63–64 (in Russian).]
- Носков С.И., Хоняков А.А.** (2019). Программный комплекс построения некоторых типов кусочно-линейных регрессий // *Информационные технологии и математическое моделирование в управлении сложными системами*. № 3. С. 47–55. [**Noskov S.I., Khonyakov A.A.** (2019). A software package for constructing some types of piecewise linear regressions. *Information Technologies and Mathematical Modeling in the Management of Complex Systems*, 3, 47–55 (in Russian).]

Paired piecewise linear regression model with fixed nodes

© 2024 K.M. Zubrilin

K.M. Zubrilin,

Branch of Federal State Budgetary Educational Institution of Higher Education «Kerch State Maritime Technological University» in Feodosiya, Feodosiya, Russia; e-mail: kzubrilin@yandex.ru

Received 25.05.2023

Abstract. In this paper we develop a method for constructing a paired regression model in the class of piecewise linear continuous functions with fixed nodes. The concept of linear division of the correlation field, its partial sections and its nodes is introduced into consideration. Using the linear division of the correlation field, a class of piecewise linear functions with fixed nodes is determined. The linear division is revealed during the visual analysis of the correlation field. Using the least squares method, a system of linear equations is compiled to find point estimates of the parameters of the approximating function. With the exception of two unknown angular coefficients (unknown) this system is reduced to a tridiagonal system of equations, the unknowns of which are the nodal values of the approximating function. The tridiagonal system is solved by the run-through method. An algorithm was constructed to demonstrate the operation of the developed method. Its initial data is arrays of the corresponding values of the explanatory and dependent variables, as well as an array of numbers of the right ends of the intervals defining the nodes. An array of fixed nodes is constructed from an array of values of the explanatory variable and an array of numbers of the right ends of the intervals. Next, an array of point estimates of the parameters of the approximating function is constructed. This algorithm is implemented in Python in the form of user-defined functions.

Keywords: correlation field, linear division of the correlation field, piecewise linear function with fixed nodes, piecewise linear approximation, paired regression model, least squares method, tridiagonal system, run-through method.

JEL Classification: C20.

UDC: 330.43.

For reference: **Zubrilin K.M.** (2024). Paired piecewise linear regression model with fixed nodes. *Economics and Mathematical Methods*, 60, 4, 113–124. DOI: 10.31857/S0424738824040103 (in Russian).