**Лабораторная работа 3**

**Восстановление дискретно заданной функций**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | *x* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| *f* | 0.6 | 2.3 | 2.1 | 3.5 | 4.2 | 6.0 | 7.8 | 9.2 |
|  | *x* | Этот лист заданий заполняется студентами и остаётся у преподавателя | | | | | | | |
| *f* |
| 2 | *x* |  |  |  |  |  |  |  |  |
| *f* |  |  |  |  |  |  |  |  |
| 3 | *x* |  |  |  |  |  |  |  |  |
| *f* |  |  |  |  |  |  |  |  |
| 4 | *x* |  |  |  |  |  |  |  |  |
| *f* |  |  |  |  |  |  |  |  |
| 5 | *x* |  |  |  |  |  |  |  |  |
| *f* |  |  |  |  |  |  |  |  |
| 6 | *x* |  |  |  |  |  |  |  |  |
| *f* |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| i | x | y | Линейная регрессия | Квадратичная регрессия | Кубическая регрессия | Степенная регрессия | Показательная регрессия | Логарифмическая регрессия | Гиперболическая регрессия |
| 0.3 |  | -0.5425 | 0.6301 | 0.3365 | 0.1533 | 0.8185 | -5.2146 | -18.8248 | 0.8185 |
| 1 | 1 | 0.6 | 0.2917 | 0.9542 | 0.8667 | 0.6712 | 1.0383 | -0.6087 | -0.6738 |
| 2 | 2 | 2.3 | 1.4833 | 1.5780 | 1.6405 | 1.5705 | 1.4586 | 2.0431 | 3.2158 |
| 3 | 3 | 2.1 | 2.6750 | 2.3911 | 2.4786 | 2.5825 | 2.0490 | 3.5942 | 4.5123 |
| 4 | 4 | 3.5 | 3.8667 | 3.3935 | 3.4310 | 3.6751 | 2.8783 | 4.6948 | 5.1605 |
| 5 | 5 | 4.2 | 5.0583 | 4.5851 | 4.5476 | 4.8321 | 4.0432 | 5.5485 | 5.5495 |
| 6 | 6 | 6 | 6.2500 | 5.9661 | 5.8786 | 6.0431 | 5.6797 | 6.2459 | 5.8088 |

Линейная регрессия

Уравнение регрессии:  
\widehat{y}=ax+b

Коэффициент a:  
a&=\frac{\sum x_i \sum y_i- n\sum x_iy_i}{\left(\sum x_i\right)^2-n\sum x_i^2}

Коэффициент b:  
b&=\frac{\sum x_i \sum x_iy_i-\sum x_i^2\sum y_i}{\left(\sum x_i\right)^2-n\sum x_i^2}

Коэффициент линейной парной корреляции:

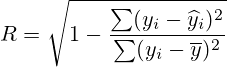
Коэффициент детерминации:  
R^2=r_{xy}^2

Средняя ошибка аппроксимации:  
\overline{A}=\dfrac{1}{n}\sum\left|\dfrac{y_i-\widehat{y}_i}{y_i}\right|\cdot100\%

Квадратичная регрессия

Уравнение регрессии:  
\widehat{y}=ax^2+bx+c

Система уравнений для нахождения коэффициентов a, b и c:

Коэффициент корреляции:  
,  
где  
\overline{y}= \dfrac{1}{n}\sum y_i

Коэффициент детерминации:  
R^2

Средняя ошибка аппроксимации:  
\overline{A}=\dfrac{1}{n}\sum\left|\dfrac{y_i-\widehat{y}_i}{y_i}\right|\cdot100\%

Кубическая регрессия

Уравнение регрессии:  
\widehat{y}=ax^3+bx^2+cx+d

Система уравнений для нахождения коэффициентов a, b, c и d:

Коэффициент корреляции, коэффициент детерминации, средняя ошибка аппроксимации - используются те же формулы, что и для квадратичной регрессии.

Степенная регрессия

Уравнение регрессии:  
\widehat{y}=a\cdot x^b

Коэффициент b:  
b=\dfrac{n\sum(\ln x_i\cdot\ln y_i)-\sum\ln x_i\cdot\sum\ln y_i }{n\sum\ln^2x_i-\left(\sum\ln x_i\right)^2 }

Коэффициент a:  
a=\exp\!\left(\dfrac{1}{n}\sum\ln y_i-\dfrac{b}{n}\sum\ln x_i\right)

Коэффициент корреляции, коэффициент детерминации, средняя ошибка аппроксимации — используются те же формулы, что и для квадратичной регрессии.

Показательная регрессия

Уравнение регрессии:  
\widehat{y}=a\cdot b^x

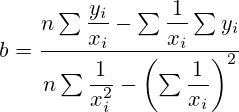
Коэффициент b:  
b=\exp\dfrac{n\sum x_i\ln y_i-\sum x_i\cdot\sum\ln y_i }{n\sum x_i^2-\left(\sum x_i\right)^2 }

Коэффициент a:  
a=\exp\!\left(\dfrac{1}{n}\sum\ln y_i-\dfrac{\ln b}{n}\sum x_i\right)

Коэффициент корреляции, коэффициент детерминации, средняя ошибка аппроксимации — используются те же формулы, что и для квадратичной регрессии.

Гиперболическая регрессия

Уравнение регрессии:  
\widehat{y}=a + \frac{b}{x}

Коэффициент b:  


Коэффициент a:  
a=\dfrac{1}{n}\sum y_i-\dfrac{b}{n}\sum\dfrac{1}{x_i}

Коэффициент корреляции, коэффициент детерминации, средняя ошибка аппроксимации - используются те же формулы, что и для квадратичной регрессии.

Логарифмическая регрессия

Уравнение регрессии:  
\widehat{y}=a + b\ln x

Коэффициент b:  
b=\dfrac{n\sum(y_i\ln x_i)-\sum\ln x_i\cdot \sum y_i }{n\sum\ln^2x_i-\left(\sum\ln x_i\right)^2 }

Коэффициент a:  
a=\dfrac{1}{n}\sum y_i-\dfrac{b}{n}\sum\ln x_i

Коэффициент корреляции, коэффициент детерминации, средняя ошибка аппроксимации - используются те же формулы, что и для квадратичной регрессии.

Экспоненциальная регрессия

Уравнение регрессии:  
\widehat{y}=e^{a+bx}

Коэффициент b:  
b=\dfrac{n\sum x_i\ln y_i-\sum x_i\cdot\sum\ln y_i }{n\sum x_i^2-\left(\sum x_i\right)^2 }

Коэффициент a:  
a=\dfrac{1}{n}\sum\ln y_i-\dfrac{b}{n}\sum x_i

Коэффициент корреляции, коэффициент детерминации, средняя ошибка аппроксимации - используются те же формулы, что и для квадратичной регрессии.

Вывод формул

Сначала сформулируем задачу:  
Пусть у нас есть неизвестная функция y=f(x), заданная табличными значениями (например, полученными в результате опытных измерений).  
Нам необходимо найти функцию заданного вида (линейную, квадратичную и т. п.) y=F(x), которая в соответствующих точках принимает значения, как можно более близкие к табличным.  
На практике вид функции чаще всего определяют путем сравнения расположения точек с графиками известных функций.

Полученная формула y=F(x), которую называют эмпирической формулой, или уравнением регрессии y на x, или приближающей (аппроксимирующей) функцией, позволяет находить значения f(x) для нетабличных значений x, сглаживая результаты измерений величины y.

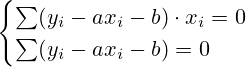
Для того, чтобы получить параметры функции F, используется метод наименьших квадратов. В этом методе в качестве критерия близости приближающей функции к совокупности точек используется суммы квадратов разностей значений табличных значений y и теоретических, рассчитанных по уравнению регрессии.

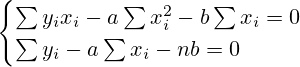
Таким образом, нам требуется найти функцию F, такую, чтобы сумма квадратов S была наименьшей:  
S=\sum\limits_i(y_i-F(x_i))^2\rightarrow min

Рассмотрим решение этой задачи на примере получения линейной регрессии F=ax+b.  
S является функцией двух переменных, a и b. Чтобы найти ее минимум, используем условие экстремума, а именно, равенства нулю частных производных.

Используя формулу производной сложной функции, получим следующую систему уравнений:

Для функции вида F(x,a,b)=ax+b частные производные равны:  
F^\prime_a=x,  
F^\prime_b=1

Подставив производные, получим:  


Далее:  


Откуда, выразив a и b, можно получить формулы для коэффициентов линейной регрессии, приведенные выше.  
Аналогичным образом выводятся формулы для остальных видов регрессий.