ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ
Сертификат: 175940[illegible]
Владелец: Евсюкова Наталья Юрьевна
Действителен: [illegible]

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования
«МОСКОВСКИЙ АВТОМОБИЛЬНО-ДОРОЖНЫЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ (МАДИ)»
ВОЛЖСКИЙ ФИЛИАЛ

Кафедра математики и информатики

Лабораторный практикум по дисциплине

«Моделирование АСОИиУ»

для направления
09.03.01 «Информатика и вычислительная техника»
профиля
«Автоматизированные системы обработки информации и управления»

ЧАСТЬ 2
Математическая теория планирования эксперимента

Для студентов очной формы обучения

Чебоксары 2015

ББК
М 171

М 171 Максимова А.П. Лабораторный практикум по дисциплине «Моделирование АСОИиУ». Часть 2. – Чебоксары: Волжский филиал Московского автомобильно-дорожного государственного технического университета (МАДИ), 2015. – с.

Методическое пособие составлено в соответствии с действующей программой курса «Моделирование АСОИиУ» для студентов направления 09.03.01 «Информатика и вычислительная техника». Предназначено для студентов очной формы обучения. В нем кратко изложены необходимые теоретические сведения и методические указания для выполнения лабораторных работ по дисциплине «Моделирование АСОИиУ».

© Максимова А.П., 2015
© Волжский филиал МАДИ, 2015

Лабораторная работа 1

ПОСТРОЕНИЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ

Цель работы: Изучение основных понятий, определений, принципов теории планирования экспериментов, приобретение навыков проведения экспериментов по построению математических моделей, ознакомление с методикой построения регрессионных моделей.

Основные понятия

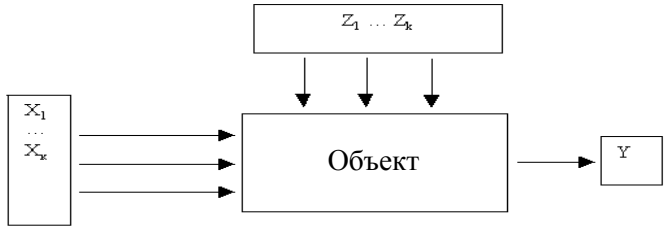
Эксперимент – метод научного исследования, когда исследователь активно и целенаправленно воздействует на объект исследования путем создания искусственных условий или использования естественных условий, необходимых для выявления конкретных свойств объекта.

Эксперименты делятся на пассивные и активные (управляемые). В пассивном эксперименте контролируемые (входные) параметры нельзя изменять, в активном – можно.

Планирование эксперимента – область знания, связанная с построением и оптимизацией математических моделей.

Объект исследования рассматривается как носитель некоторых неизвестных или подлежащих исследованию свойств и качеств – своеобразный «черный ящик». При этом вектор $X_1 \ldots X_k$ представляет собой группу контролируемых и управляемых величин, которые могут изменяться определенным образом в ходе эксперимента, а $Z_1 \ldots Z_k$ контролируемые характеристики. Характеристики ($X_1 \ldots X_k$) также называют факторами или управляемыми воздействиями. Функция Y – функция отклика (поверхность отклика), представляет собой реакцию системы на воздействие факторов. Также можно выделить и третью, не обозначенную на идеальной модели систему входных сигналов – это шумы или помехи, которые обусловлены многими факторами: ошибками обслуживающего персонала, влиянием внешней среды, погрешностью приборов и т.д. К этой же группе относятся воздействия, которые не могут контролироваться либо из-за их сложности, либо из-за незнания их природы и невозможности контроля.

Характеристики объектов имеют различную физическую природу, а, следовательно, и размерность, что затрудняет построения модели. Поэтому на практике значения факторов, которые имеют реальный физический смысл, нормируют (приводят к определенному ранее заданному набору значений). Для любого фактора X существует нижний X_{min} и верхний X_{max} уровни изменения значений.



Структурная схема объекта (процесса) при проведении активного эксперимента

Факторное пространство

Приведем алгоритм нормировки фактора:

- выбираем масштаб и положение осей координат таким образом, чтобы $X_{\text{i min}}$ соответствовало –1, а $X_{\text{i max}}$ +1;
- вычисляем значение $X_{\text{i}0}$ для данного фактора следующим образом

$$X_{\text{i}0} = \frac{X_{\text{i max}} + X_{\text{i min}}}{2};$$

- вычисляем интервал изменения фактора;

$$\text{d}x_{\text{i}} = X_{\text{i}0} - X_{\text{imin}} = X_{\text{imax}} - X_{\text{i}0};$$

- находим нормированное значение $X_{\text{н}}$ для каждого фактора:

$$X_{\text{iн}} = \frac{X_{\text{i}} - X_{\text{i}0}}{dx_{\text{i}}}.$$

Зависимость реакции объекта от точки факторного пространства называется функцией отклика Y, а ее геометрическое представление $Y(x_1, x_2,\ldots, x_{\text{i}})$ – поверхностью отклика. Векторов значений функции отклика может быть столько, сколько опытов.

Проведение эксперимента

Эксперимент состоит из опытов (воспроизведение исследуемого явления). Под планированием эксперимента понимают выбор плана эксперимента – совокупности данных, определяющих число, условия и порядок

реализации опытов. Каждый опыт эксперимента характеризуется определенным набором значений факторов.

$$X_{\mathrm{q}} = \left(X_{1\mathrm{q}}, X_{2\mathrm{q}}, X_{3\mathrm{q}}, ..., X_{\mathrm{nq}}\right)$$

Вектор, содержащий некоторый набор конкретных значений факторов X_{I}, определяет q-ю точку плана эксперимента. Совокупность векторов X_{q} (q = 1, 2, …, n) образует план эксперимента (матрица, содержащая k строк и n столбцов, каждая строка которой образует точку плана эксперимента, а столбец фактор эксперимента).

X_{11}	X_{21}	X_{31}	…	X_{N1}
X_{12}	X_{22}	X_{32}	…	X_{N2}
…	…	…	…	…
$X_{1\mathrm{K}}$	$X_{2\mathrm{K}}$	$X_{3\mathrm{K}}$	…	X_{NK}

Совокупность всех точек плана, отличающихся уровнем хотя бы одного фактора (различных строк матрицы планирования), называется спектром плана. Матрица, получаемая из всех различных строк плана – матрица спектра плана. Она отличается от приведенной выше матрицы только числом строк (из-за отсутствия повторяющихся точек плана). При количестве точек спектра плана G, ее размерность будет составлять: G строк на N столбцов. Применяется также матрица дублирования, размерность которой совпадает с размерностью матрицы спектра плана. Она имеет вид:

K_1	0	0	…	0
0	K_2	0	…	0
…	…	…	…	…
0	0	0	…	K_{g}

Здесь K_{j} – число параллельных опытов в точке спектра плана с номером j (j = 1, 2, …, N). Т.е. это число характеризует дублирование соответствующей строки в матрице спектра плана.

Построение регрессионных моделей

Для описания объектов управления часто используются полиномиальные модели. При этом в качестве базисного выражения используется ряд Тейлора, имеющий конечное число членов.

$$F(x) = F(a) + \frac{x - a}{1!} F'(a) + \frac{(x - a)^2}{2!} F''(a) + \ldots + \frac{(x - a)^n}{n!} F^{(n)}(a).$$

Но при использовании аппроксимирующего полинома Тейлора в приведенном выше виде возникает ряд проблем, связанных с нахождением производных, так как неизвестна функция, а известен только ряд ее значений. Поэтому заменим полином Тейлора на аналогичное ему уравнение регрессии

$$\hat{Y} = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i,j=1}^{k} b_{i,j} x_i x_j + \sum_{i=1}^{k} b_{i,i} x_i^2 + \sum_{i,j,n=1}^{k} b_{i,j,k} x_i x_j x_n + \ldots$$

где k – число столбцов в матрице планирования. Построим линейную регрессионную модель. Для ее экспериментального получения используем план первого порядка (факторный эксперимент первого порядка).

$$\hat{Y} = b_0 + \sum_{i=1}^{k} b_i x_i \ .$$

Для k-факторного эксперимента достаточно $k + 1$ опытов. При определении коэффициентов регрессии должны выполняться необходимые и достаточные условия:

1. Результаты измерений выходной величины Y в N точках факторного пространства – нормально распределенные величины.
2. Дисперсии реализации во всех точках факторного пространства одинаковы, то есть не зависят от абсолютного значения величины и от направления обхода факторного пространства.
3. Входные переменные (факторы) – это независимые величины, которые измеряются с бесконечно малой ошибкой по отношению к ошибке выходной величины.

Оценка выполняется по критерию Фишера.

Любой многофакторный эксперимент является результатом варьирования всех факторов.

Полный факторный эксперимент

Если в многофакторном эксперименте использованы все возможные комбинации уровней факторов, то такой эксперимент называется полным факторным экспериментом. Приведем таблицу (для линейного уравнения регрессии):

Количество факторов k	Количество неизвестных коэффициентов	Количество опытов в полном факторном эксперименте	Достаточное количество для определения коэффициентов
3	4	8	2^{3-1}
4	5	16	2^{4-1}
5	6	32	2^{5-2}
6	7	64	2^{6-3}

Полный факторный эксперимент (ПФЭ) включает в себя $2k$ опытов, которые при построении линейной модели могут полностью не использоваться. В общем случае ПФЭ позволяет найти $2k$ коэффициентов регрессии при $2k$ базисных функциях. Первые k + 1 базисные функции очевидны – они составляют линейную модель ($f0 = 1$ $f1 = X1$ $f2 = X2$ $f3 = X3$).

Приведем пример полного трехфакторного эксперимента (столбцы с первого по четвертый – первый столбец вводится искусственным путем и постоянен и равен 1). Эта матрица является матрицей базисных функций.

	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	$X_1X_2X_3$
1	+1	–1	–1	–1	+1	+1	+1	–1
2	+1	+1	–1	–1	–1	–1	+1	+1
3	+1	–1	+1	–1	–1	+1	–1	+1
4	+1	+1	+1	–1	+1	–1	–1	–1
5	+1	–1	–1	+1	+1	–1	–1	+1
6	+1	+1	–1	+1	–1	+1	–1	–1
7	+1	–1	+1	+1	–1	–1	+1	–1
8	+1	+1	+1	+1	+1	+1	+1	+1

Свойства полного факторного эксперимента

Матрица планирования ПФЭ обладает рядом свойств:

1) симметричность плана относительно центра эксперимента:

$$\sum_{i=1}^{n} x_{j,i} = 0,$$

то есть сумма значений уровней любого фактора (столбца) равна 0 ;

2) нормировка плана:

$$\sum_{i=1}^{n} x^2_{j,i} = N ,$$

сумма квадратов значений уровней любого фактора равна N (числу строк матрицы планирования ПФЭ);

3) ортогональность плана:

$$\sum_{i=1}^{n} x_{j,i} \cdot x_{u,i} = 0 ,$$

сумма по парных произведений значений уровней любых 2 факторов (кроме $j = u$) равна 0;

4) рототабельность плана – точность предсказания значений функции отклика одинакова на равном расстоянии от центра и не зависит от направления обхода.

Свойства ортогональности и рототабельности взаимоисключающие.

Дробный факторный эксперимент

В некоторых случаях нет необходимости использовать полный факторный эксперимент. В таких случаях усекают количество строк матрицы ПФЭ до количества коэффициентов регрессионной модели. Это производится в случаях линейной регрессионной модели. Дробный факторный эксперимент удовлетворяет всем свойствам полного факторного эксперимента.

Определение коэффициентов уравнения регрессии

После проведения опытов во всех точках факторного пространства необходимо найти коэффициенты уравнения регрессии. Для этого воспользуемся методом наименьших квадратов:

$$\Phi = \sum_{i=1}^{n} (\overset{\Lambda}{Y}_i - Y_i)^2 \to \min ;$$

$$\overset{\Lambda}{Y}_i = \varphi(X_1,...,X_k, b_0,...,b_k)$$

$$\Phi = \sum_{i=1}^{n} (\varphi(X_1,...,X_k, b_0,...,b_k) - Y_i)^2 , \text{ поскольку } \begin{cases} \dfrac{\partial \Phi}{\partial b_0} = 0 \\ \ldots \\ \dfrac{\partial \Phi}{\partial b_k} = 0 \end{cases} ,$$

то после дифференцирования получим

$$\begin{cases} \dfrac{\partial \Phi}{\partial b_0} = 2\sum\limits_{i=1}^{n} (\varphi(X_1,...,X_k,\ b_0,...,b_k) - Y_i)\dfrac{\partial \varphi}{\partial b_0} = 0, \\ ... \\ \dfrac{\partial \Phi}{\partial b_k} = 2\sum\limits_{i=1}^{n} (\varphi(X_1,...,X_k,\ b_0,...,b_k) - Y_i)\dfrac{\partial \varphi}{\partial b_k} = 0. \end{cases}$$

Для линейной регрессии при $k = 2$:

$$Y_i = \varphi(X_{1i}, X_{2i}, b_0, b_1, b_2), Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i}.;$$

продифференцировав по коэффициентам, получим:

$$\frac{\partial \varphi}{\partial b_0} = 1, \qquad \frac{\partial \varphi}{\partial b_1} = X_{1i}, \qquad \frac{\partial \varphi}{\partial b_2} = X_{2i}.$$

Запишем уравнения в полной форме:

$$\begin{cases} \sum\limits_{i=1}^{n} (b_0 + b_1 X_{1i} + b_2 X_{2i} - Y_i)\cdot 1 = 0, \\ \sum\limits_{i=1}^{n} (b_0 + b_1 X_{1i} + b_2 X_{2i} - Y_i)\cdot X_{1i} = 0, \quad \Leftrightarrow \\ \sum\limits_{i=1}^{n} (b_0 + b_1 X_{1i} + b_2 X_{2i} - Y_i)\cdot X_{2i} = 0. \end{cases}$$

$$\begin{cases} (\sum\limits_{i=1}^{n} 1) b_0 + (\sum\limits_{i=1}^{n} X_{1i})\, b_1 + (\sum\limits_{i=1}^{n} X_{2i}) b_2 = \sum\limits_{i=1}^{n} Y_i, \\ (\sum\limits_{i=1}^{n} X_{1i}) b_0 + (\sum\limits_{i=1}^{n} X_{1i}^2) b_1 + (\sum\limits_{i=1}^{n} X_{2i} X_{1i}) b_2 = \sum\limits_{i=1}^{n} X_{1i} Y_i, \\ (\sum\limits_{i=1}^{n} X_{2i}) b_0 + (\sum\limits_{i=1}^{n} X_{1i} X_{2i}) b_1 + (\sum\limits_{i=1}^{n} X_{2i}^2) b_2 = \sum\limits_{i=1}^{n} X_{2i} Y_i. \end{cases}$$

$\sum_{i=1}^{n} 1 = n$, разделим каждое уравнение на n

$$\begin{cases} b_0 + (\frac{1}{n}\sum_{i=1}^{n} X_{1i})b_1 + (\frac{1}{n}\sum_{i=1}^{n} X_{2i})b_2 = \frac{1}{n}\sum_{i=1}^{n} Y_i, \\ (\frac{1}{n}\sum_{i=1}^{n} X_{1i})b_0 + (\frac{1}{n}\sum_{i=1}^{n} X_{1i}^2)b_1 + (\frac{1}{n}\sum_{i=1}^{n} X_{2i}X_{1i})b_2 = \frac{1}{n}\sum_{i=1}^{n} X_{1i}Y_i, \\ (\frac{1}{n}\sum_{i=1}^{n} X_{2i})b_0 + (\frac{1}{n}\sum_{i=1}^{n} X_{1i}X_{2i})b_1 + (\frac{1}{n}\sum_{i=1}^{n} X_{2i}^2)b_2 = \frac{1}{n}\sum_{i=1}^{n} X_{2i}Y_i. \end{cases}$$

Отсюда, принимая в расчет свойства матрицы планирования, получим следующие формулы для вычисления коэффициентов

$$b_0 = \frac{1}{n}\sum_{i=1}^{n} Y_i,$$

$$b_1 = \frac{1}{n}\sum_{i=1}^{n} X_{1i}Y_i,$$

$$b_2 = \frac{1}{n}\sum_{i=1}^{n} X_{2i}Y_i.$$

ЗАДАНИЕ НА ЛАБОРАТОРНУЮ РАБОТУ

1. Используя программу генерации случайных чисел провести трехфакторный эксперимент в восьми точках (то есть сформировать три столбца и восемь строк в матрице планирования – заполнить ее случайным образом). Желательно взять ограничение до 20 при генерации случайных чисел, но учесть возможность его изменения по требованию преподавателя.

2. Определить значения нулевых уровней факторов, выполнить нормировку факторов.

3. Составить матрицу планирования для полного трехфакторного эксперимента с использованием дополнительного нулевого фактора ($X_0 = 1$).

4. Составить матрицу планирования для дробного трехфакторного эксперимента, пренебрегая взаимодействием факторов.

5. Провести эксперимент во всех точках ДФЭ, повторив 5 раз опыты в выбранных точках факторного пространства (найти значения функции отклика Y из таблицы 1 в соответствии с вариантом, выданным преподавателем).

6. Найти коэффициенты уравнения регрессии.

7. Проверить свойства полного факторного эксперимента: симметричность, нормировку, ортогональность и рототабельность.

8. Составить уравнение регрессии в кодированном виде, привести его к натуральному, используя значение интервалов варьирования.

ПОРЯДОК ПРОВЕДЕНИЯ РАБОТЫ

1. Используя генератор случайных чисел, найти значения факторов в точках, а также функцию отклика. Определить нулевой уровень фактора, провести нормировку.

2. Составить матрицу планирования для полного трехфакторного эксперимента с использованием дополнительного нулевого фактора ($X_0 = 1$), и заполнить таблицу кодированными значениями X_1, X_2 и X_3.

	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	$X_1X_2X_3$	$Y_{1,j}$	$Y_{2,j}$	$Y_{3,j}$	Y_j
1												
2												
3												
4												
5												
6												
7												
8												

3. Составить матрицу планирования для дробного трехфакторного эксперимента, пренебрегая взаимодействием факторов.

4. Провести эксперимент во всех точках ДФЭ (найти значения функции отклика Y). Для каждой точки плана провести по три экспери-

мента, значения функции отклика брать из таблицы 1 в соответствии с вариантом.

5. Получаем коэффициенты регрессии после упрощения системы уравнений b_0, b_1 , b_2 , b_3.

6. Уравнение регрессии будет иметь вид:

$$Y = b_0X_0 + b_1X1 + b_2X2 –+b_3X3 \text{ ; } (X_0 = 1).$$

7. Полученное в кодированном виде уравнение регрессии преобразовать в натуральный, используя значения интервалов варьирования.

Вариант 1

3,004	3,031	3,035	3,039	3,001
5,193	5,152	5,177	5,209	5,151
3,927	3,950	3,936	3,898	3,897
7,141	7,099	7,111	7,138	7,097

Вариант 2

3,651	3,605	3,653	3,592	3,627
6,547	6,514	6,535	6,562	6,581
4,761	4,793	4,816	4,792	4,801
9,515	9,566	9,534	9,552	9,528

Вариант 3

2,124	2,150	2,139	2,140	2,157
3,382	3,394	3,368	3,374	3,372
2,705	2,652	2,655	2,674	2,713
4,307	4,242	4,276	4,317	4,255

Вариант 4

2,588	2,597	2,542	2,537	2,539
4,191	4,165	4,152	4,129	4,138
3,201	3,231	3,202	3,199	3,248
5,509	5,453	5,448	5,511	5,445

Вариант 5

3,072	3,028	3,080	3,049	3,069
5,193	5,159	5,163	5,220	5,168
3,932	3,955	3,893	3,915	3,939
7,094	7,126	7,149	7,102	7,158

Вариант 6

4,292	4,285	4,333	4,304	4,277
8,385	8,390	8,404	8,421	8,390
5,881	5,886	5,847	5,900	5,909
13,349	13,332	13,357	13,342	13,356

Вариант 7

4,307	4,284	4,284	4,316	4,286
8,387	8,396	8,430	8,389	8,404
5,832	5,873	5,856	5,843	5,862
13,329	13,304	13,328	13,340	13,312

Вариант 8

3,583	3,605	3,623	3,623	3,587
6,555	6,564	6,523	6,559	6,511
4,795	4,790	4,776	4,798	4,744
9,504	9,530	9,524	9,557	9,530

Вариант 9

3,054	3,032	3,024	3,046	3,019
5,147	5,170	5,178	5,190	5,177
3,926	3,895	3,937	3,931	3,915
7,117	7,121	7,101	7,130	7,091

Вариант 10

2,549	2,537	2,563	2,564	2,569
4,118	4,164	4,155	4,126	4,151
3,236	3,220	3,202	3,212	3,207
5,445	5,485	5,449	5,472	5,455

Вариант 11

2,164	2,165	2,145	2,150	2,163
3,347	3,338	3,322	3,318	3,358
3,950	3,932	3,908	3,935	3,901
6,855	6,870	6,875	6,872	6,907

Вариант 12

2,788	2,823	2,815	2,777	2,773
4,491	4,467	4,492	4,473	4,460
3,485	3,510	3,515	3,524	3,475
5,883	5,879	5,863	5,870	5,877

Вариант 13

2,132	2,114	2,160	2,146	2,120
3,373	3,324	3,377	3,327	3,385
3,978	3,928	3,905	3,948	3,904
6,898	6,908	6,887	6,940	6,904

Вариант 14

2,567	2,587	2,585	2,527	2,583
4,148	4,183	4,155	4,144	4,169
4,998	4,949	4,950	4,947	4,968
9,758	9,689	9,701	9,711	9,686

Вариант 15

3,073	3,033	3,062	3,065	3,029
5,191	5,186	5,221	5,156	5,198
3,884	3,932	3,929	3,914	3,899
14,701	14,690	14,734	14,754	14,674

СОДЕРЖАНИЕ И ОФОРМЛЕНИЕ ОТЧЕТА

1. Титульный лист, содержащий информацию о студенте (группа, фамилия, номер варианта).
2. Результаты подготовки (выбранные по варианту значения экспериментальных данных).
3. Основные теоретические положения (используемые формулы).
4. Результаты подготовки (матрица планирования в виде таблицы).
5. Листинг программы (язык программирования не имеет значения).
6. Ответы на контрольные вопросы.
7. Результат выполнения работы.
8. Выводы по лабораторной работе.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Что называется экспериментом?
2. Какие бывают эксперименты?
3. Что называется планированием эксперимента?
4. Что образует план эксперимента?
5. Что называется спектром плана?
6. Чем характеризуется объект исследования? Дайте определение факторному пространству.
7. Что такое регрессионные полиномы и где они применяются?
8. Перечислите условия, необходимые для определения коэффициентов регрессии.
9. Что называется полным факторным экспериментом?

Лабораторная работа 2

ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

Цель работы: Изучение методики проведения экспериментов по ортогональному плану, овладение теорией проверки адекватности модели оригиналу.

Основные понятия

Нахождение построчной дисперсии

Предположим, что в каждой точке факторного пространства, которой соответствует одна из строк матрицы планирования, проводится серия из m опытов. Для любой i-й точки вычисляется среднее значение выходной величины:

$$y_{\mathrm{i}} = \sum_{u=1}^{m} \frac{y_{\mathrm{i\,u}}}{m} \qquad (1)$$

и построчную дисперсию выходной величины:

$$S^2\{y_{\mathrm{i}}\} = \sum_{u=1}^{m} \frac{\left(y_{\mathrm{i\,u}} - y_i\right)^2}{m-1}. \qquad (2)$$

Рассмотрим этапы обработки результатов эксперимента на примере.

N	X_0	X_1	X_2	X_1X_2	$Y_{1\mathrm{i}}$	$Y_{2\mathrm{i}}$	$Y_{3\mathrm{i}}$	Y_{i}	$S^2\{y_{\mathrm{i}}\}$
1	+1	–1	–1	+1	43	35	48	42	43
2	+1	+1	–1	–1	90	86	94	90	16
3	+1	–1	+1	–1	10	16	16	14	12
4	+1	+1	+1	+1	56	54	58	56	4

Среднее значение выходной величины Y_{i} в каждой точке определим по формуле (1) (m = 3):

$Y_1 = (43 + 35 + 48) / 3 = 42,$
$Y_2 = (90 + 86 + 94) / 3 = 90,$
$Y_3 = (10 + 16 + 16) / 3 = 14,$
$Y_4 = (56 + 54 + 58) / 3 = 56.$

Определим по формуле (2) построчную дисперсию:

$$S^2\{y_1\} = [(43 - 42)^2 + (35 - 42)^2 + (48 - 42)^2] / 2 = 43,$$
$$S^2\{y_2\} = [(90 - 90)^2 + (86 - 90)^2 + (94 - 90)^2] / 2 = 16,$$
$$S^2\{y_3\} = [(10 - 14)^2 + (16 - 14)^2 + (16 - 14)^2] / 2 = 12,$$
$$S^2\{y_4\} = [(56 - 56)^2 + (54 - 56)^2 + (58 - 56)^2] / 2 = 4.$$

Проверка однородности по критерию Кохрена

Среди всей совокупности рассчитанных построчных дисперсий выбирается максимальная $S^2\{y_i\}$max и берется отношение данной дисперсии к сумме всех построчных дисперсий $S^2\{y_i\}$, т.е. определяется расчетное значение коэффициента Кохрена:

$$G_p = \frac{S^2\{y_i\}\max}{\sum_{i=1}^{N} S^2\{y_i\}},$$

который показывает, какую долю в общей сумме построчных дисперсий занимает максимальная из них. В случае идеальной однородности построчных дисперсий коэффициент *Gp* стремился бы к значению $1 / N$, где N – число опытов (количество строк в матрице планирования).

Расчетное значение коэффициента Кохрена сравнивается с табличным значением G – критерия, которое выбирается из таблиц для принятого уровня значимости α и для чисел степени свободы соответственно числителя f_1 и знаменателя f_2:

$$f_1 = m - 1; \quad f_2 = N.$$

Для этого значение f_1 находится в горизонтальном заголовке таблицы (выбирается столбец), а f_2 выбирается слева в вертикальном заголовке таблицы (выбирается строка) и на пересечении получаем табличное значение $G_т$ коэффициента Кохрена. Если выполняется условие

$$G_p < G_т,$$

то с выбранным уровнем статистической значимости α (с достоверностью $1 - \alpha$) все построчные дисперсии признаются однородными. В противном случае гипотезу отвергают.

По данным из нашего примера определим расчетное значение коэффициента

$$G_{\text{р}} = 43 / (43 + 16 + 12 + 4) = 0{,}57.$$

В соответствии с таблицей коэффициентов для $\alpha = 0{,}05$; $f_1 = 3 - 1 = 2$; $f_2 = 4$, находим $G_{\text{т}} = 0{,}77$; $G_{\text{т}} > G_{\text{р}}$, т.е. условие выполняется.

Проверка нуль-гипотезы по критерию Стьюдента

После проверки однородности переходят к определению оценок коэффициентов по формуле:

$$a_{\text{k}} = \sum_{i=1}^{N} \frac{y_{\text{i k}} x_{\text{i k}}}{N},$$

где k – номер вектор-столбца.

В нашем примере имеем:

$a_0 = (42 + 90 + 14 + 56) / 4 = 50{,}5,$
$a_1 = (-42 + 90 - 14 + 56) / 4 = 22{,}5,$
$a_2 = (-42 - 90 + 14 + 56) / 4 = -15{,}5,$
$a_{12} = (42 - 90 - 14 + 56) / 4 = -1{,}5.$

Найденные таким образом коэффициенты уравнения регрессии необходимо оценить на статистическую значимость. Оценка производится по t-критерию Стьюдента. Для каждого коэффициента a_{k} вычисляется коэффициент (a_{k} – коэффициент уравнения регрессии):

$$t_{\text{k}} = \frac{|a_{\text{k}}|}{S\{a_{\text{k}}\}},$$

т. е. проверяется отклонение от нуля найденной оценки. $S\{a_{\text{k}}\}$ – оценка среднего квадратичного отклонения погрешности определения коэффициента.

Оценка дисперсии коэффициентов, найденных по экспериментальным данным:

$$S^2\{a_{\text{k}}\} = \frac{S^2_{\text{в}}}{N \cdot m},$$

$$S\{a_{\text{k}}\} = \sqrt{S^2\{a_{\text{k}}\}}\ .$$

Оценкой генеральной дисперсии воспроизводимости $S^2_{\text{в}}$, характеризующая точность одного измерения, является средняя из всех построчных дисперсий:

$$S_{\text{в}}{}^2 = \sum_{i=1}^{N} \frac{S^2\{y_{\text{i}}\}}{N}.$$

При выбранном уровне статистической значимости α по таблицам распределения Стьюдента при числе степеней свободы $f = N\,(m - 1)$ находят табличное значение коэффициента $t_{\text{табл}}$. Найденное табличное значение сравнивается с расчетным значением коэффициента. Если выполняется неравенство $t_{\text{табл}} > t_{\text{k}}$, то принимается нуль-гипотеза, т.е. считается, что найденный коэффициент a_{k} является статистически незначительным и его следует исключить из уравнения регрессии.

Для рассматриваемого примера:

$$S^2_{\text{в}} = (43 + 16 + 12 + 4) / 4 = 18{,}75;$$
$$S^2\{a_{\text{k}}\} = 18{,}75 / (4 \cdot 3) = 1{,}56; \qquad S\{a_{\text{k}}\} = 1{,}25.$$

Определим расчетные значения коэффициента Стьюдента:

$$t_0 = 50{,}5 / 1{,}25 = 40{,}4,$$
$$t_1 = 22{,}5 / 1{,}25 = 18,$$
$$t_2 = 15{,}5 / 1{,}25 = 12{,}4,$$
$$t_{12} = 1{,}5 / 1{,}25 = 1{,}2.$$

Из таблиц при уровне статистической значимости $\alpha = 0{,}05$ и числе степеней свободы $f = 4(3 - 1) = 8$, определим табличное значение коэффициента. Оно равно $t_{\text{т}} = 2{,}3$. Сопоставим расчетные значения t_{k} с табличным $t_{\text{т}}$. Неравенство выполняется для t_{12}. Следовательно, можно предположить, что a_{12} статистически незначим и его можно исключить из уравнения регрессии.

Уравнение регрессии, содержащее статистически значимые коэффициенты, будет (в кодированной системе)

$$Y' = 50{,}5 + 22{,}5x_1 - 15{,}5x_2.$$

Проверка адекватности по критерию Фишера

Полученное уравнение регрессии необходимо проверить на адекватность исследуемому объекту. Для этой цели необходимо оценить, насколько отличаются средние значения y_{i} выходной величины, полученной в точках факторного пространства, и значения y_{i} ,полученного из уравне-

ния регрессии в тех же точках факторного пространства. Для этого используют дисперсию адекватности

$$S^2_{\text{ад}} = \frac{m}{N-L}\sum_{i=1}^{N}(y_{\text{i}} - y'_{\text{i}})^2 ,$$

где l – число значимых коэффициентов.

Адекватность модели проверяют по F-критерию Фишера $F_{\text{р}} = S^2_{\text{ад}} / S^2_{\text{в}}$.

Найденное расчетным путем $F_{\text{р}}$ сравнивают с табличным значением $F_{\text{т}}$, которое определяется при уровне значимости α и числе степеней свободы $f_{\text{ад}} = N - l$ и $f_{\text{в}} = N(m - 1)$. Если $F_{\text{р}} < F_{\text{т}}$, то полученная математическая модель с принятым уровнем статистической значимости α адекватна экспериментальным данным.

Для рассматриваемого примера получаем:

$$Y'_1 = 50{,}5 + 22{,}5\ (-1) - 15{,}5\ (-1\ ) = 43{,}5,$$
$$Y'_2 = 50{,}5 + 22{,}5\ (+1) - 15{,}5\ (-1\ ) = 88{,}5;$$
$$Y'_3 = 50{,}5 + 22{,}5\ (-1) - 15{,}5\ (+1\ ) = 12{,}5,$$
$$Y'_4 = 50{,}5 + 22{,}5\ (+1) - 15{,}5\ (+1\ ) = 57{,}5.$$

Рассчитаем оценку дисперсии адекватности:

$$S^2_{\text{ад}} = 3[(42 - 43{,}5)^2 + (90 - 88{,}5)^2 + (14 - 12{,}5)^2 + (56 - 57{,}5)^2]/(4 - 3) = 27,$$
$$F_{\text{р}} = S^2_{\text{ад}} / S^2_{\text{в}} = 27 / 18{,}75 = 1{,}44.$$

Табличное значение коэффициента Фишера при уровне статистической значимости α =0,05 и числе степеней свободы $f_{\text{ад}} = (4 - 3) = 1$ и $f_{\text{в}} = 4(3 - 1) = 8$ будет $F_{\text{т}} = 5{,}32$. Следовательно, при выбранном уровне статистической значимости полученная в результате эксперимента:

$$y' = 50{,}5 + 22{,}5x_1 - 15{,}5x_2$$

адекватна исследуемому объекту.

ЗАДАНИЕ НА ЛАБОРАТОРНУЮ РАБОТУ

1. Составить матрицу планирования для полного трехфакторного эксперимента с использованием дополнительного нулевого фактора ($X_0 = 1$).

2. Провести эксперимент, во всех точках факторного пространства повторив 5 раз опыты во всех точках факторного пространства (найти значения функции отклика Y из таблицы 1 согласно варианту, выданному преподавателем).

3. Проверить однородность дисперсии по критерию Кохрена.

4. Найти коэффициенты уравнения регрессии.

5. С помощью критерия Стьюдента оценить значимость коэффициентов регрессии.

6. Составить уравнение регрессии в кодированном виде и проверить его адекватность с помощью критерия Фишера.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Опишите план нахождения построчной дисперсии выходной величины?

2. Для чего нужно расчетное значение коэффициента Кохрена и как он находится?

3. Что такое критерий Стьюдента и где он используется?

4. Для чего оценивают, насколько отличаются средние значения y_i выходной величины, полученной в точках факторного пространства, и значения y_i, полученного из уравнения регрессии в тех же точках факторного пространства?

5. Чем определяется F-критерий Фишера и как его применяют?

Вариант 1

3,004	3,031	3,035	3,039	3,001
5,193	5,152	5,177	5,209	5,151
3,927	3,950	3,936	3,898	3,897
7,141	7,099	7,111	7,138	7,097
4,684	4,697	4,688	4,730	4,729
9,135	9,123	9,166	9,134	9,117
6,371	6,403	6,343	6,339	6,337
14,672	14,680	14,695	14,668	14,672

Вариант 2

3,651	3,605	3,653	3,592	3,627
6,547	6,514	6,535	6,562	6,581
4,761	4,793	4,816	4,792	4,801
9,515	9,566	9,534	9,552	9,528
5,828	5,847	5,842	5,905	5,886
13,041	13,081	13,051	13,089	13,063
8,364	8,371	8,338	8,365	8,366
25,575	25,563	25,611	25,578	25,534

Вариант 3

2,124	2,150	2,139	2,140	2,157
3,382	3,394	3,368	3,374	3,372
2,705	2,652	2,655	2,674	2,713
4,307	4,242	4,276	4,317	4,255
3,107	3,089	3,096	3,119	3,137
5,081	5,148	5,123	5,092	5,073
3,948	3,901	3,914	3,951	3,919
6,873	6,920	6,932	6,858	6,869

Вариант 4

2,588	2,597	2,542	2,537	2,539
4,191	4,165	4,152	4,129	4,138
3,201	3,231	3,202	3,199	3,248
5,509	5,453	5,448	5,511	5,445
3,793	3,830	3,850	3,789	3,852
6,718	6,752	6,760	6,709	6,743
4,963	4,966	5,001	4,952	5,007
9,738	9,753	9,702	9,746	9,737

Вариант 5

3,072	3,028	3,080	3,049	3,069
5,193	5,159	5,163	5,220	5,168
3,932	3,955	3,893	3,915	3,939
7,094	7,126	7,149	7,102	7,158
4,740	4,704	4,668	4,698	4,724
9,163	9,167	9,160	9,133	9,191
6,336	6,396	6,369	6,405	6,357
14,676	14,668	14,725	14,722	14,741

Вариант 6

4,292	4,285	4,333	4,304	4,277
8,385	8,390	8,404	8,421	8,390
5,881	5,886	5,847	5,900	5,909
13,349	13,332	13,357	13,342	13,356
7,389	7,368	7,439	7,419	7,442
20,252	20,271	20,271	20,258	20,310
11,282	11,269	11,293	11,249	11,254
66,571	66,613	66,562	66,585	66,620

Вариант 7

4,307	4,284	4,284	4,316	4,286
8,387	8,396	8,430	8,389	8,404
5,832	5,873	5,856	5,843	5,862
13,329	13,304	13,328	13,340	13,312
7,379	7,415	7,415	7,368	7,368
20,255	20,278	20,304	20,279	20,261
11,226	11,238	11,271	11,234	11,273
66,599	66,605	66,588	66,595	66,562

Вариант 8

3,583	3,605	3,623	3,623	3,587
6,555	6,564	6,523	6,559	6,511
4,795	4,790	4,776	4,798	4,744
9,504	9,530	9,524	9,557	9,530
5,855	5,839	5,827	5,881	5,863
13,040	13,011	13,045	13,061	13,036
8,328	8,301	8,303	8,319	8,310
25,586	25,544	25,578	25,562	25,556

Лабораторная работа 3

ПОСТРОЕНИЕ ДВУХФАКТОРНОГО ЭКСПЕРИМЕНТА С ИСПОЛЬЗОВАНИЕМ КВАДРАТИЧНОЙ МОДЕЛИ

Цель работы: Изучение методики построения квадратичных моделей объектов на основе планов второго порядка, теории композиционного планирования.

Основные понятия

Планы второго порядка

Планы второго порядка предназначены для получения регрессионной модели в виде полного квадратичного полинома (полинома второй степени):

$$\varphi(X_1, X_2, ..., X_n) = B_0 + \sum_{i=1}^{n} B_i X_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} B_{ij} X_i X_j + \sum_{i=1}^{n} B_i^2 X_i^2 .$$

Для факторов x_i в стандартизированном масштабе эта модель будет иметь вид:

$$\varphi(x_1, x_2, ..., x_n) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \beta_{ij} x_i x_j + \sum_{i=1}^{n} \beta_i^2 x_i^2 .$$

Подобные планы применяют, как правило, либо в том случае, когда использование планирования первого порядка не позволило получить адекватную регрессионную модель, и выяснилась необходимость ее усложнения, либо если заранее известно, что объект исследования обладает существенными нелинейными свойствами.

По сравнению с планами первого порядка планы второго порядка являются более сложными по структуре, имеют большее число точек в спектре плана и уровней варьирования для каждого фактора, требуют при своей реализации увеличенного количества опытов. Действительно, квадратичная модель содержит $k = 1 + 2n + C_n^2 = \frac{[(n+1)(n+2)]}{2}$ членов, что в $(n + 2) / 2$ раз больше, чем в линейной модели. Значит, соответственно возрастает и минимально необходимое количество точек в спектре плана. Для получения квадратичной зависимости каждый из факторов должен изменяться, по крайней мере, на трех уровнях.

Центрально-композиционный план

Центрально-композиционные планы (ЦКП) любой модификации состоят из трех частей. Первая часть – основа или ядро плана – это ПФЭ 2^n или ДФЭ 2^{n-p}, где n – количество неизвестных коэффициентов регрессии, $p = 0,1,2$. При этом требуется, чтобы ядро плана обеспечивало раздельную оценку коэффициентов регрессии и всех парных взаимодействий. Данное условие накладывает весьма жесткое ограничение на возможную степень дробности используемого ДФЭ. В частности, при $n \le 4$, как показывают расчеты, может применяться лишь ПФЭ 2^n; если $5 \le n \le 7$, то кроме ПФЭ 2^nможно использовать и ДФЭ 2^{n-1}, а для $n > 7$ допустим также и ДФЭ 2^{n-2}. Вторая часть ЦКП – так называемые «звездные» точки, расположенные на координатных осях на расстоянии $\pm\alpha$ от центра эксперимента. Общее число таких точек равно $2n$. Третья часть ЦКП – опыты в центре плана; число таких опытов $N_0 \ge 1$. Произвольный симметричный ЦКП приведен в таблице:

Составная часть ЦКП	G	Фактор				Число точек
		x_1	x_2	…	x_n	
Ядро плана (ПФЭ $2n$ или ДФЭ $2n-p$)	1	–1	–1	…	–1	2^{n-p} $p = 0; 1; 2;$
	2	+1	–1	…	–1	
	3	–1	+1	…	–1	
	4	+1	+1	…	–1	
	…	…	…	…	…	
	$2n$-p	+1	+1	…	–1	
«Звездные» точки	$2n$-p + 1	$-\alpha$	0	…	0	$2n$
	$2n$-p + 2	$+\alpha$	0	…	0	
	$2n$-p + 3	0	$-\alpha$	…	0	
	$2n$-p + 4	0	$+\alpha$	…	0	
	…	…	…	…	…	
	$2n$-p + $2n$-1	0	0	…	$-\alpha$	
	$2n$-p + $2n$	0	0	…	$+\alpha$	
Центральные точки	$2n$-p + $2n$ + 1	0	0	…	0	N_0
	…	…	…	…	…	
	$2n$-p + $2n$ + N0	0	0	…	0	

Таким образом, ЦКП для двух факторов и значения $\alpha = 1$ ($\pm\alpha$ – расстояние от «звездных» точек до центра эксперимента), имеет вид:

g	X_0	x_1	x_2	x_1x_2	x_1^2	x_2^2	Y'
1	+1	–1	–1	+1	+1	+1	Y'_1
2	+1	+1	–1	–1	+1	+1	Y'_2
3	+1	–1	+1	–1	+1	+1	Y'_3
4	+1	+1	+1	+1	+1	+1	Y'_4
5	+1	–1	0	0	+1	0	Y'_5
6	+1	+1	0	0	+1	0	Y'_6
7	+1	0	–1	0	0	+1	Y'_7
8	+1	0	+1	0	0	+1	Y'_8
9	+1	0	0	0	0	0	Y'_9

При этом оценки коэффициентов регрессии мы будем находить, решая матричное уравнение $\Phi \cdot B = (F^{\mathrm{T}}\ Y)$, где:

$$(F^{\mathrm{T}}Y) = \left\| \begin{array}{l} \sum\limits_{i=1}^{N} x_{0\mathrm{i}} y'_{\mathrm{i}} \\ \sum\limits_{i=1}^{N} x_{1\mathrm{i}} y'_{\mathrm{i}} \\ \dots \\ \sum\limits_{i=1}^{N} x_{\mathrm{ni}}^2 y'_{\mathrm{i}} \end{array} \right\|.$$

Это уравнение можно переписать в виде

$$B = \Phi^{-1}\left(F^{\mathrm{T}}Y\right),$$

где Φ^{-1} – нормированная обратная информационная матрица Фишера центрального композиционного плана второго порядка.

Ортогональный центрально-композиционный план

Конкретные значения α и N_0 выбираются исходя из тех или иных критериев оптимальности регрессионных экспериментов (α – звездное плечо, N_0 – количество экспериментов в центре плана). В связи с этим принято выделять ортогональные (ОЦКП) и рототабельные (РЦКП) центрально-композиционные планы.

В ОЦКП, как правило, $N_0 = 1$, а план целиком строится с учетом критерия ортогональности (сумма попарных произведений значений уровней двух любых факторов (столбцов) равна нулю). Тогда информационная матрица Фишера должна быть диагональной, для чего необходимо, как это следует из вида информационной матрицы Фишера для произвольного центрально-композиционного плана, принять специальные меры для обеспечения по парной ортогональности столбцов, отвечающих свободному члену β_0 и квадратичным коэффициентам β_{i2} , i = 1, 2, …, *n*, а также столбцов, отвечающих квадратичным членам между собой.

С этой целью, прежде всего, несколько видоизменяют систему базисных функций, а именно – ищут регрессионную модель в виде:

$$\varphi(x) = \beta_0{}' + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n-1} \sum_{j=1}^{n} \beta_{ij}\, x_i x_j + \sum_{i=1}^{n} \beta_i^2 \left(x_i^2 - x'^2_i \right),$$

где

$$x'^2_i = \frac{1}{N} \sum_{g=1}^{N} x_{ig}^2 = \frac{2^{n-p} + 2\alpha^2}{N};$$

N – общее число точек плана: $N = 2^{n-p} + 2n + 1$; $\beta_0{}' = \beta_0 + \sum_{i=1}^{n} \beta_{ii}\, x'^2_i$ (где p – число, определяющее дробность эксперимента, а β_i – коэффициенты уравнения регрессии). Как видно, в этой модели при квадратичных коэффициентах используются центрированные переменные. Переход к таким переменным обеспечивает ортогональность столбца матрицы F численных значений базисных функций, соответствующего свободному члену уравнения регрессии, и любого из столбцов центрированных квадратов (базисная функция вида $\tilde{x}_i^2 = x_i^2 - x'^2_i$). Действительно, для указанных столбцов имеет место следующее равенство:

$$\sum_{g=1}^{N} f_0 \left(x_{ig}^2 - x'^2_i \right) = \sum_{g=1}^{N} x_{ig}^2 - N x'^2_i = \sum_{g=1}^{N} x_{ig}^2 - N \frac{1}{N} \sum_{g=1}^{N} x_{ig}^2 = 0.$$

Это равенство справедливо независимо от конкретного значения α. Однако, при произвольном α, остаются неортогональными столбцы матрицы F, отвечающие различным центрированным квадратичным пере-

менным. Поэтому, в ОЦКП числовое значение α и выбирается как раз из условия ортогональности именно этих столбцов, т.е. исходя из условия:

$$\sum_{\substack{g=1 \\ i \neq j}}^{N} \left(x_{ig}^2 - x_i'^2 \right) \left(x_{jg}^2 - x_j^2 \right) = 0,$$

или, в развернутом виде:

$$2^{n-p} \left(1 - \frac{2^{n-p} + 2\alpha^2}{N} \right)^2 + 4 \left[\alpha^2 - \frac{2^{n-p} + 2\alpha^2}{N} \right] \left[- \frac{2^{n-p} + 2\alpha^2}{N} \right] +$$

$$+ \left[- \frac{2^{n-p} + 2\alpha^2}{N} \right]^2 (2n - 3) = 0.$$

После несложных преобразований получаем уравнение для требуемого значения α:

$$\alpha = \sqrt{\sqrt{2^{n-p-2} \left(2^{n-p} + 2n + 1 \right)} - 2^{n-p-1}} = \sqrt{\left(\sqrt{N \cdot 2^{n-p}} - 2^{n-p} \right) / 2}$$

С помощью этой формулы найдены конкретные числовые значения α при $n = 2 \div 8$:

1	2	3	4	5		6		7		8		
Ядро ЦПК	ПФЭ 2^2	ПФЭ 2^3	ПФЭ 2^4	ПФЭ 2^5	ДФЭ 2^{5-1}	ПФЭ 2^6	ДФЭ 2^{6-1}	ПФЭ 2^7	ДФЭ 2^{7-1}	ПФЭ 2^8	ДФЭ 2^{8-1}	ДФЭ 2^{8-2}
N	9	15	25	43	27	77	45	143	79	273	145	81
α	1,000	1,215	1,414	1,596	1,547	1,761	1,724	1,909	1,885	2,045	2,029	2,000

Общее количество опытов N в ОЦКП равно $N = 2^{n-p} + 2n + N_0$. Таким образом, переходя к квадратичной модели с центрированными квадратичными переменными и используя указанные значения α, можно добиться полной ортогонализации столбцов матрицы F.

Оценки коэффициентов регрессии, полученные с помощью ОЦКП, некоррелированы между собой, что, впрочем, характерно для любого ортогонального плана.

Оценки коэффициентов регрессии для соответствующих групп равны:

– для свободного члена $\beta_0' = \dfrac{\sum_{i=1}^{N} x_{0i} y'_i}{N}$;

– для линейных слагаемых $\beta_k = \dfrac{\sum_{i=1}^{N} x_{ki} y'_i}{2^{n-p} + 2\alpha^2}$, $k = 1, 2, \ldots, n$;

– для попарных взаимодействий $\beta_{jk} = \dfrac{\sum_{i=1}^{N} x_{jki} y'_i}{2^{n-p}}$, $j, k = 1, 2 \ldots n; j \neq k$;

Для центрированных квадратичных переменных $\beta_k^2 = \dfrac{\sum_{i=1}^{N} x_{ki}^2 y'_i}{2\alpha^4}$.

Приведем теперь уравнение регрессии к более привычному для нас виду:

$$\varphi(x) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n-1} \sum_{j=1}^{n} \beta_{ij} x_i x_j + \sum_{i=1}^{n} \beta_i^2 x_i^2,$$

где

$$\beta_0 = \beta_0' - x_i'^2 \sum_{i=1}^{n} \beta_i^2 = \frac{\sum_{i=1}^{N} x_{0i} y'_i}{N} - \frac{2^{n-p} + 2\alpha^2}{N} \sum_{i=1}^{n} \beta_i^2 .$$

Условие нормировки в случае ортогонального ЦКП не соблюдается, т. к. $\sum_{i=1}^{N} x_{ik}^2 \neq N$ (k – номер любого столбца, кроме нулевого). Это значит, что точность оценки коэффициентов регрессии для разных групп неодинакова.

Оценки дисперсий для каждой из четырех однородных групп для m параллельных опытов подсчитываются по следующим формулам:

$$S^2\{\beta_i\} = \frac{S_в^2}{m\left(2^{n-p} + 2\alpha^2\right)},$$

$$S^2\{\beta_{ij}\} = \frac{S_в^2}{m \cdot 2^{n-p}},$$

$$S^2\{\beta_i^2\} = \frac{S_в^2}{m \cdot 2\,\alpha^4},$$

$$S^2\{\beta_0\} = \frac{S_в^2}{m\left(N + 2\alpha^4\right)},$$

где $S_в^2$ – дисперсия воспроизводимости.

Таким образом, дисперсия оценки Y' функции отклика в некоторой точке факторного пространства зависит не только от расстояния этой точки до центра плана ρ, но и от ее положения на гиперсфере. Значит, ОЦКП не удовлетворяет условию рототабельности. Поэтому, если не предъявляются особые требования к точности предсказания выходной величины по уравнению регрессии в любом направлении факторного пространства от базовой точки, предпочтительно применение ортогонального ЦКП ввиду его простоты.

ЗАДАНИЕ НА ЛАБОРАТОРНУЮ РАБОТУ

1. Составить матрицу планирования ортогонального центрально-композиционного плана для двух факторов с использованием дополнительного нулевого фактора ($X_0 = 1$).

2. Провести эксперимент, во всех точках факторного пространства, повторив 5 раз опыты во всех точках факторного пространства (найти значения функции отклика Y из таблицы 1 согласно варианту, выданному преподавателем).

3. Проверить однородность дисперсии по критерию Кохрена и, если необходимо, подобрать такое m (m – кратность проведения опытов, не больше 5), чтобы дисперсия была однородной.

4. Найти коэффициенты уравнения регрессии для нормализованной системы координат.

5. С помощью критерия Стьюдента оценить значимость коэффициентов регрессии.

6. Проверить адекватность модели оригиналу с помощью критерия Фишера.

7. Привести уравнение регрессии к натуральному виду.

СОДЕРЖАНИЕ И ОФОРМЛЕНИЕ ОТЧЕТА

1. Титульный лист, содержащий информацию о студенте (группа, фамилия, номер варианта).
2. Результаты подготовки (выбранные по варианту значения экспериментальных данных).
3. Основные теоретические положения (используемые формулы).
4. Результаты подготовки (матрица планирования в виде таблицы).
5. Проверка ортогональности столбцов матрицы.
6. Результат проверки однородности дисперсии по критерию Кохрена.
7. Коэффициенты регрессии β_i .
8. Результат проверки значимости коэффициентов регрессии.
9. Результат проверки адекватности модели оригиналу с помощью критерия Фишера.
10. Уравнение регрессии в натуральном виде.
11. Ответы на контрольные вопросы.
12. Выводы по лабораторной работе.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Почему в планах второго порядка возрастает минимально необходимое количество точек в спектре плана? Как определяется число членов квадратичной модели?
2. В каких случаях используют квадратичную модель объекта?
3. Дайте определение ЦКП.
4. Цель натурализации уравнения регрессии.
5. Чем обеспечивается ортогональность столбцов матрицы F?
6. Определение ОЦКП. Каким образом для ОЦКП выбирается числовое значение α (звездного плеча).
7. Объясните, почему точность оценки коэффициентов регрессии для разных групп неодинакова.

Вариант 1

3,004	3,031	3,035	3,039	3,001
5,193	5,152	5,177	5,209	5,151
3,927	3,950	3,936	3,898	3,897
7,141	7,099	7,111	7,138	7,097
4,684	4,697	4,688	4,730	4,729
9,135	9,123	9,166	9,134	9,117
6,371	6,403	6,343	6,339	6,337
14,672	14,680	14,695	14,668	14,672
5,828	5,847	5,842	5,905	5,886

Вариант 2

3,651	3,605	3,653	3,592	3,627
6,547	6,514	6,535	6,562	6,581
4,761	4,793	4,816	4,792	4,801
9,515	9,566	9,534	9,552	9,528
5,828	5,847	5,842	5,905	5,886
13,041	13,081	13,051	13,089	13,063
8,364	8,371	8,338	8,365	8,366
25,575	25,563	25,611	25,578	25,534
5,081	5,148	5,123	5,092	5,073

Вариант 3

2,124	2,150	2,139	2,140	2,157
3,382	3,394	3,368	3,374	3,372
2,705	2,652	2,655	2,674	2,713
4,307	4,242	4,276	4,317	4,255
3,107	3,089	3,096	3,119	3,137
5,081	5,148	5,123	5,092	5,073
3,948	3,901	3,914	3,951	3,919
6,873	6,920	6,932	6,858	6,869
6,718	6,752	6,760	6,709	6,743

Вариант 4

2,588	2,597	2,542	2,537	2,539
4,191	4,165	4,152	4,129	4,138
3,201	3,231	3,202	3,199	3,248
5,509	5,453	5,448	5,511	5,445
3,793	3,830	3,850	3,789	3,852
6,718	6,752	6,760	6,709	6,743
4,963	4,966	5,001	4,952	5,007
9,738	9,753	9,702	9,746	9,737
7,094	7,126	7,149	7,102	7,158

Вариант 5

3,072	3,028	3,080	3,049	3,069
5,193	5,159	5,163	5,220	5,168
3,932	3,955	3,893	3,915	3,939
7,094	7,126	7,149	7,102	7,158
4,740	4,704	4,668	4,698	4,724
9,163	9,167	9,160	9,133	9,191
6,336	6,396	6,369	6,405	6,357
14,676	14,668	14,725	14,722	14,741
8,385	8,390	8,404	8,421	8,390

Вариант 6

4,292	4,285	4,333	4,304	4,277
8,385	8,390	8,404	8,421	8,390
5,881	5,886	5,847	5,900	5,909
13,349	13,332	13,357	13,342	13,356
7,389	7,368	7,439	7,419	7,442
20,252	20,271	20,271	20,258	20,310
11,282	11,269	11,293	11,249	11,254
66,571	66,613	66,562	66,585	66,620
7,379	7,415	7,415	7,368	7,368

Вариант 7

4,307	4,284	4,284	4,316	4,286
8,387	8,396	8,430	8,389	8,404
5,832	5,873	5,856	5,843	5,862
13,329	13,304	13,328	13,340	13,312
7,379	7,415	7,415	7,368	7,368
20,255	20,278	20,304	20,279	20,261
11,226	11,238	11,271	11,234	11,273
66,599	66,605	66,588	66,595	66,562
13,040	13,011	13,045	13,061	13,036

Вариант 8

3,583	3,605	3,623	3,623	3,587
6,555	6,564	6,523	6,559	6,511
4,795	4,790	4,776	4,798	4,744
9,504	9,530	9,524	9,557	9,530
5,855	5,839	5,827	5,881	5,863
13,040	13,011	13,045	13,061	13,036
8,328	8,301	8,303	8,319	8,310
25,586	25,544	25,578	25,562	25,556
4,701	4,682	4,690	4,718	4,719

Вариант 9

3,054	3,032	3,024	3,046	3,019
5,147	5,170	5,178	5,190	5,177
3,926	3,895	3,937	3,931	3,915
7,117	7,121	7,101	7,130	7,091
4,701	4,682	4,690	4,718	4,719
9,150	9,159	9,115	9,162	9,156
6,390	6,383	6,384	6,378	6,378
14,677	14,670	14,718	14,690	14,693
6,721	6,714	6,741	6,704	6,722

Вариант 10

2,549	2,537	2,563	2,564	2,569
4,118	4,164	4,155	4,126	4,151
3,236	3,220	3,202	3,212	3,207
5,445	5,485	5,449	5,472	5,455
3,825	3,812	3,790	3,782	3,781
6,721	6,714	6,741	6,704	6,722
4,951	4,989	4,955	4,941	4,981
9,735	9,693	9,705	9,711	9,726
3,950	3,932	3,908	3,935	3,901

Лабораторная работа 4

ПОСТРОЕНИЕ ДВУХФАКТОРНОГО ЭКСПЕРИМЕНТА С ИСПОЛЬЗОВАНИЕМ РОТОТАБЕЛЬНОГО ЦЕНТРАЛЬНО-КОМПОЗИЦИОННОГО ПЛАНА

Цель работы: Изучение методики построения квадратичных моделей объектов на основе планов второго порядка, теории композиционного планирования.

Основные понятия

При исследовании экстремальной области часто интерес представляет оценка не коэффициентов полученной регрессионной модели, а самой функции отклика. Кроме того, на практике часто можно значительно упростить регрессионную модель путем поворота координатных осей, т.е. преобразованием координат. Рототабельное планирование, обеспечивающее погрешность предсказания выходной величины по уравнению регрессии, зависящую лишь от расстояния точки факторного пространства до центра эксперимента, позволяет предсказывать с одинаковой точностью значение функции отклика, а следовательно, преобразовывать систему координат с целью упрощения уравнения регрессии.

Основным условием рототабельности планов является инвариантность нормированной информационной матрицы Φ и корреляционной матрицы Φ^{-1} к вращению прямоугольных осей относительно начала координат, помещенного в базовую точку. Исходя из условия инвариантности матриц к вращению системы координат, точность оценок коэффициентов регрессии при вращении также не будет изменяться. Следует при этом отметить, что изменение момента масштаба входных переменных приводит к потере свойства рототабельности. Таким образом, необходимо поддерживать постоянство масштаба задания независимых переменных при проведении всего эксперимента.

Нормированная матрица Φ должна обладать некоторыми свойствами, чтобы быть инвариантной к ортогональному преобразованию (вращению).

Будем называть моментами плана элементы нормированной информационной матрицы:

$$\frac{1}{N}\sum_{i=1}^{N} x_{i1}^{p_1} x_{i2}^{p_2} \ldots x_{ij}^{p_j} \ldots x_{in}^{p_n},$$

а порядком момента величину:

$$P_{\text{ku}} = \sum_{\text{j}=1}^{n} p_{\text{j}}$$

Момент будет четным, если все степени p_j четные, и нечетные, если хотя бы одна степень p_j элемента нормированной информационной матрицы нечетна.

Для рототабельного плана второго порядка все нечетные моменты вплоть до четвертого порядка включительно должны быть равны нулю. К таковым относятся следующие моменты:

1. Первого порядка

$$\sum_{\text{i}=1}^{\text{N}} x_{\text{i j}} = 0,\ j = \overline{1,n}.$$

2. Второго порядка

$$\sum_{\text{i}=1}^{\text{N}} x_{\text{i j}} x_{\text{i u}} = 0,\ j \neq u,\ j = \overline{1,}\ .$$

3. Третьего порядка

$$\sum_{\text{i}=1}^{\text{N}} x_{\text{i j}} x_{\text{i u}} x_{\text{i k}} = \sum_{\text{i}=1}^{\text{N}} x_{\text{i j}}^2 x_{\text{i k}} = \sum_{\text{i}=1}^{\text{N}} x_{\text{i j}}^3 = 0,\ j \neq u \neq k;\ j,u,k = \overline{1,n}.$$

4. Четвертого порядка

$$\sum_{\text{i}=1}^{\text{N}} x_{\text{i j}}^3 x_{\text{i u}} = \sum_{\text{i}=1}^{\text{N}} x_{\text{i j}}^2 x_{\text{i u}} x_{\text{i k}} = \sum_{\text{i}=1}^{\text{N}} x_{\text{i j}} x_{\text{iu}} x_{\text{i k}} x_{\text{i g}} = 0;$$

$$j \neq u \neq k \neq g;\ j,u,k,g = \overline{1,n}.$$

Все четные моменты такого плана должны удовлетворять таким соотношениям:

1. Второго порядка

$$\sum_{\text{i}=1}^{\text{N}} x_{\text{i j}}^2 = N\Lambda_2,\ j = \overline{1,n}$$

2. Четвертого порядка

$$\sum_{\text{i}=1}^{\text{N}} x_{\text{i j}}^2 x_{\text{i u}}^2 = N\Lambda_{22},\ j \neq u;\ j = \overline{1,n}$$

$$\sum_{\text{i}=1}^{\text{N}} x_{\text{i j}}^4 = 3N\Lambda_4,\ j = \overline{1,n}$$

Константа Λ_2 выбирается из условия выбора масштаба плана, а константа Λ_4 выбирается из условия невырожденности информационной матрицы.

Если все точки плана расположить на поверхности сферы с радиусом ρ и с центром в начале координат нормированного факторного пространства, то дли любой *I*-ой точки данного плана можно записать:

$$\rho^2 = \sum_{j=1}^{n} x_{ij}^2, \; i = \overline{1,n}.$$

В соответствии с матрицей Фишера (подматрица с линейными членами):

$$\rho^2 = \sum_{j=1}^{n}\sum_{i=1}^{N} x_{ij}^2 = n\lambda_2.$$

Таким образом, точки рототабельного плана, в которых реализуются опыты, должны быть расположены на нескольких (как минимум двух) концентрических гиперсферах с общим центром.

Для построения рототабельного ЦКП второго порядка необходимо выбрать соответствующие значения «звездного» плеча α и количество центральных точек N_0.

Запишем общие выражения для всех четных моментов

1. Нулевого порядка

$$\sum_{i=1}^{N} x_{i0} = N.$$

2. Второго порядка

$$\sum_{i=1}^{N} x_{ij}^2 = 2n + 2\alpha^2, \; i = \overline{1,n}$$

3. Четвертого порядка

$$\sum_{i=1}^{N} x_{ij}^2 x_{iu}^2 = 2^n; \; j,u = \overline{1,n}; j \neq u;$$

$$\sum_{i=1}^{N} x_{ij}^4 = 2^n + 2\alpha^4, \; j = \overline{1,n}.$$

Таким образом, $N\Lambda_2 = 2^n + 2\alpha^2$, $N\Lambda_4 = 2^n$, $3N\Lambda_4 = 2^n + 2\alpha^4$.

Отсюда находим условие рототабельности:

$$\frac{3\cdot 2^{n}}{N}=\frac{2^{n}+2\alpha^{4}}{N},$$

или $\alpha^4 = 2^n$.

Следовательно, чтобы ЦКП второго порядка обладал свойствами рототабельности, значение «звездного» плеча должно составлять:

$$\alpha = 2^{n/4}.$$

Как и в случае ортогональных ЦКП, α зависит от числа n входных величин. Для определения числа опытов в центре плана («нулевой» точке) необходимо исходить из условия невырожденности информационной матрицы Фишера для рототабельных планов:

$$\frac{\lambda_4}{\lambda_2^2}=\frac{2^{n}\left(2^{n}+2n+N_0\right)}{\left(2^{n}+2\alpha^{2}\right)^{2}}=\frac{2^{n}+2n+N_0}{\left(2^{1/2}+2\right)^{2}}\neq\frac{n}{n+2}.$$

Из последнего соотношения видно, что для построения рототабельного ЦКП с невырожденной информационной матрицей Фишера достаточно в центре плана проводить один опыт. Увеличение числа N_0 опытов в центре плана приводит к увеличению числителя и позволяет усилить неравенство до требуемой степени.

Часто исследователя интересует информация о функции отклика в некоторой окрестности центра плана, т.е. требуется, чтобы информация о выходной величине, полученная на основании уравнения регрессии, была практически одинаковой (постоянной) внутри гипершара радиуса $\rho = 1$ для $\rho \in [0,1]$. Такое планирование называется униформ-рототабельным. Для его получения достаточно обеспечить равенство дисперсии в центре плана ($\rho = 0$) и на поверхности гиперсферы радиуса $\rho = 1$. Этого добиваются подбором числа наблюдений N_0 в центре плана.

В случае, когда число факторов велико, то в качестве «ядра» рототабельного ЦКП выбирается матрица ДФЭ. Оптимальное значение «звездного» плеча при этом определяется так:

$$\alpha = 2^{(n-p)/4}.$$

В таблице приведены значения α и N_0 для РЦКП, причем значения N_0 приведены для униформ-рототабельного плана.

1	2	3	4	5		6		7	
Ядро ЦПК	ПФЭ 2^2	ПФЭ 2^3	ПФЭ 2^4	ПФЭ 2^5	ДФЭ 2^{5-1}	ПФЭ 2^6	ДФЭ 2^{6-1}	ПФЭ 2^7	ДФЭ 2^{7-1}
N_0	5	6	7	10	6	15	9	21	14
N	13	20	31	52	32	91	53	163	92
α	1,414	1,682	2,000	2,378	2,000	2,828	2,378	3,364	2,828

Для нахождения методом наименьших квадратов оценки коэффициентов регрессии будем исходить из известного соотношения:

$$CB = F^T Y,$$

откуда

$$B = C^{-1} F^T Y.$$

Это равенство может быть переписано в виде

$$B = \frac{1}{N} \Phi^{-1} F^T Y,$$

где Φ^{-1} – матрица, обратная нормированной информационной матрице Φ любого рототабельного плана.

Ввиду того, что подматрицы с линейными факторами, а также с их линейными взаимодействиями диагональные, то они легко обращаются (аналогично ортогональным планам).

В таблице приведена матрица рототабельногоn-мерного плана второго порядка, в которой используется обозначение:

$$B = \left(2\lambda_4^2 \left[(n+2)\lambda_4 - n\lambda_2^2\right]\right)^{-1}.$$

Запишем формулы для определения оценок коэффициентов:

$$\hat{b}_0 = \frac{B}{N}\left\{\left[2\Lambda_4^2(n+2)\sum_{i=1}^{N} y_i\right] - 2\Lambda_2\Lambda_4 \sum_{j=1}^{n}\sum_{i=1}^{N} x_{ij}^2 y_i\right\};$$

$$\hat{b}_j = \frac{\sum_{i=1}^{N} x_{ij} y_i}{N\Lambda_2};$$

$$\hat{b}_{ju} = \frac{\sum_{i=1}^{N} x_{ij} x_{iu} y_i}{N\Lambda_4};$$

$$\hat{b}_{\text{j}}^2 = \frac{B}{N}\left\{\left[(n+2)\Lambda_4 - n\Lambda_2^2\right]\sum_{\text{i}=1}^{\text{N}} x_{\text{i}\,\text{j}}^2 y_{\text{i}} + \left[\left(\Lambda_2^2 - \Lambda_4\right)\sum_{\text{u}=1}^{\text{n}}\sum_{\text{i}=1}^{\text{N}} x_{\text{i}\,\text{u}}^2 y_{\text{i}}\right] - 2\Lambda_2\Lambda_4\sum_{\text{i}=1}^{\text{N}} y_{\text{i}}\right\};$$

$j,u = \overline{1,n};\ j < u.$

Точность вычисленных оценок определяется их дисперсиями. Для определения дисперсий оценок коэффициентов воспользуемся таблицей, в которой представлена матрица Φ^{-1}, рассмотрев ее диагональные и недиагональные элементы. Поскольку в матрице недиагональные элементы не нулевые, то оценки коэффициентов регрессии квадратичных членов и оценка свободного члена остаются взаимозависимыми.

Исходя из системы оценок коэффициентов с учетом кратности m проведения опытов, получим:

$$S^2\left\{\hat{b}_0\right\} = \frac{\left[2B\Lambda_4^2(n+2)\right]S_{\text{в}}^2}{Nm};$$

$$S^2\left\{\hat{b}_{\text{j}}\right\} = \frac{S_{\text{в}}^2}{Nm\Lambda_2};$$

$$S^2\left\{\hat{b}_{\text{ju}}\right\} = \frac{S_{\text{в}}^2}{Nm\Lambda_4};$$

$$S^2\left\{\hat{b}_{\text{j}}^2\right\} = \frac{\left[(n+1)\Lambda_4 - (n-1)\Lambda_2^2\right]BS_{\text{в}}^2}{Nm}.$$

Вторые слагаемые в выражениях оценок $\hat{b}_0$ и $\hat{b}_{\text{j}}^2$ обусловлены наличием корреляционной связи между ними. Найдем эту зависимость:

$$r\left\{\hat{b}_0, \hat{b}_{\text{j}}^2\right\} = \frac{\left(-2B\Lambda_2\Lambda_4\right)S_{\text{в}}^2}{Nm};$$

$$r\left\{\hat{b}_{\text{j}}^2, \hat{b}_{\text{u}}^2\right\} = \frac{B\left(\Lambda_2^2 - \Lambda_4\right)S_{\text{в}}^2}{Nm};\ j,u = \overline{1,n};\ j < u.$$

Поскольку остальные подматрицы нормированной матрицы Фишера Φ и обратной матрицы Φ^{-1} диагональные или нулевые, то все остальные корреляционные моменты равны нулю. Из этого следует, что оценки коэффициентов $\hat{b}_{\text{j}}$ и $\hat{b}_{\text{ju}}$ при факторах и их линейных взаимодействиях не-

зависимы. Поэтому их можно проверять независимо от полученных оценок других групп коэффициентов на статистическую значимость на основании t-критерия Стьюдента, как и при ортогональном планировании. Если какой-либо из коэффициентов $\hat{b}_{\mathrm{j}}$ или $\hat{b}_{\mathrm{ju}}$ окажется статистически незначимым, то его можно исключить из уравнения регрессии, не пересчитывая оценки других коэффициентов. Значимость оценки $\hat{b}_0$ свободного члена или оценки $\hat{b}_{\mathrm{j}}^2$ квадратичных коэффициентов регрессии должна проверяться при фиксированных значениях всех остальных коэффициентов из этой группы с помощью F-критерия. Если какая-либо из оценок коэффициентов $\hat{b}_0$ или $\hat{b}_{\mathrm{j}}^2$ окажется статистически незначимой, то для ее исключения из уравнения регрессии требуется пересчет остающихся оценок данной группы.

Рассмотрим задачу проверки гипотезы адекватности полученной регрессионной модели, содержащей значимые коэффициенты. Вначале будем предполагать, что параллельные опыты не ставятся, т. е. $m = 1$. Тогда остаточная сумма квадратов может быть записана:

$$S_{\mathrm{R}} = \sum_{\mathrm{i}=1}^{\mathrm{N}} \left(y_{\mathrm{i}} - y'_{\mathrm{i}}\right)^2 = \sum_{\mathrm{i}=1}^{\mathrm{N}-\mathrm{N}_0} \left(y_{\mathrm{i}} - y'_{\mathrm{i}}\right)^2 + \sum_{\mathrm{i}=\mathrm{N}-\mathrm{N}_0+1}^{\mathrm{N}} \left(y_{\mathrm{i}} - y'_{\mathrm{i}}\right)^2,$$

т.е. остаточная сумма разбита на сумму отклонений в точках ПФЭ (или ДФЭ), в «звездных» точках и на сумму отклонений опытных y_i и расчетных y'_i значений выходной величины в центре плана. Y'_{i}

Перепишем последнее выражение в виде:

$$S_{\mathrm{R}} = \sum_{\mathrm{i}=1}^{\mathrm{N}-\mathrm{N}_0} \left(y_{\mathrm{i}} - y'_{\mathrm{i}}\right)^2 + \sum_{\mathrm{i}=\mathrm{N}-\mathrm{N}_0+1}^{\mathrm{N}} \left(y_{\mathrm{i}} - y_0 + y_0 - y'_{\mathrm{i}}\right)^2 =$$

$$= \sum_{\mathrm{i}=1}^{\mathrm{N}-\mathrm{N}_0} \left(y_{\mathrm{i}} - y'_{\mathrm{i}}\right)^2 + \sum_{\mathrm{i}=\mathrm{N}-\mathrm{N}_0+1}^{\mathrm{N}} \left(y_0 - y'_{\mathrm{i}}\right)^2 + \sum_{\mathrm{i}=\mathrm{N}-\mathrm{N}_0+1}^{\mathrm{N}} \left(y_{\mathrm{i}} - y'_{\mathrm{o}}\right)^2.$$

В данном выражении первые два слагаемых связаны с общим рассеянием результатов наблюдений отклика относительно оценки регрессионной модели. Общее рассеяние связано со случайными погрешностями на-

блюдений, возникающими в результате влияния неконтролируемых факторов и систематическими погрешностями в случае неадекватности регрессионной модели и функции отклика. Третий член остаточной суммы связан с дисперсией, характеризующейся только случайной погрешностью опыта. Следовательно, с дисперсией адекватности (остаточной дисперсией) связана сумма:

$$S_{\text{ост}} = S_{\text{R}} - \sum_{\text{i=N}-\text{N}_0+1}^{\text{N}} (y_{\text{i}} - y'_0)^2 = \sum_{\text{i=1}}^{\text{N}} (y_{\text{i}} - y'_{\text{i}})^2 - \sum_{\text{i=N}-\text{N}_0+1}^{\text{N}} (y_{\text{i}} - y'_0)^2 .$$

Подставим уравнение регрессии в это выражение:

$$S_{\text{ост}} = \sum_{\text{i=1}}^{\text{N}} \left[y_{\text{i}} - \left(\hat{b}_0 + \sum_{\text{j=1}}^{\text{n}} \hat{b}_{\text{j}} x_{\text{j}} + \sum_{\substack{\text{j,u=1} \\ \text{j}\neq\text{u}}}^{\text{n}} \hat{b}_{\text{ju}} x_{\text{j}} x_{\text{u}} + \sum_{\text{j=1}}^{\text{n}} \hat{b}_{\text{j}}^2 x_{\text{j}}^2 \right) \right]^2 - \sum_{\text{i=N}-\text{N}_0+1}^{\text{N}} (y_{\text{i}} - y'_0)^2 .$$

и получим остаточную сумму квадратов, связанную с дисперсией адекватности и имеющую число степеней свободы:

$$f_{\text{ад}} = N - \frac{(n+2)(n+1)}{2} - (N_0 - 1).$$

Если в каждой точке рототабельного центрального композиционного плана проводилось m параллельных опытов, то, проделав аналогичные выкладки, можно получить выражение для остаточной суммы, связанной с дисперсией адекватности. Оно будет отличаться заменой результатов y'_{i} единичных наблюдений в точках плана на средние арифметические $y_{\text{i}} = \sum_{\text{k=1}}^{\text{m}} \frac{y_{\text{ik}}}{m}$ единичных наблюдений, а среднего арифметического y_0 из N_0 параллельных наблюдений в центре плана на общее среднее арифметическое $\overline{\overline{y}}_0$ из mN_0 таких наблюдений, т. е.:

$$S_{\text{ост}} = m\left\{\sum_{i=1}^{N}\left[y_i - \left(\hat{b}_0 + \sum_{j=1}^{n}\hat{b}_j x_j + \sum_{\substack{j,u=1\\ j\neq u}}^{n}\hat{b}_{ju}x_j x_u + \sum_{j=1}^{n}\hat{b}_j^2 x_j^2\right)\right]^2 - \sum_{i=N-N_0+1}^{N}\left(y_i - \bar{\bar{y}}_0\right)^2\right\}.$$

Данная остаточная сумма имеет то же число степеней свободы $f_{\text{ад}}$, что и предыдущая.

Далее для проверки адекватности модели необходимо для отношения дисперсии адекватности $S^2_{\text{ад}} = \dfrac{S_{\text{ост}}}{f_{\text{ад}}}$ и дисперсии воспроизводимости применить F-критерий Фишера, как и в общем случае регрессионного анализа. Полученная адекватная модель позволяет не только предсказать с равной точностью независимо от направления значение величины отклика, но и оценить ординаты точки экстремума. Ввиду свойства рототабельности плана эта задача облегчается – можно от полинома второго порядка, полученного в результате эксперимента, преобразованием системы координат (поворотом координатных осей) перейти к стандартному каноническому уравнению.

ЗАДАНИЕ НА ЛАБОРАТОРНУЮ РАБОТУ

1. Составить матрицу планирования ортогонального центрально-композиционного плана для двух факторов с использованием дополнительного нулевого фактора ($X_0 = 1$).

2. Провести эксперимент повторив 3 раза опыты во всех точках факторного пространства (найти значения функции отклика Y из таблицы 1 согласно варианту, выданному преподавателем).

3. Проверить однородность дисперсии по критерию Кохрена и, если необходимо, подобрать такое m (m – кратность проведения опытов, не больше 5), чтобы дисперсия была однородной.

4. Найти коэффициенты уравнения регрессии для нормализованной системы координат.

5. С помощью критерия Стьюдента оценить значимость коэффициентов регрессии.

6. Проверить адекватность модели оригиналу с помощью критерия Фишера.

7. Привести уравнение регрессии к натуральному виду.

СОДЕРЖАНИЕ И ОФОРМЛЕНИЕ ОТЧЕТА

1. Титульный лист, содержащий информацию о студенте (группа, фамилия, номер варианта).
2. Результаты подготовки (выбранные по варианту значения экспериментальных данных).
3. Основные теоретические положения (используемые формулы).
4. Результаты подготовки (матрица планирования в виде таблицы).
5. Проверка ортогональности столбцов матрицы.
6. Результат проверки однородности дисперсии по критерию Кохрена.
7. Коэффициенты регрессии β_i .
8. Результат проверки значимости коэффициентов регрессии.
9. Результат проверки адекватности модели оригиналу с помощью критерия Фишера.
10. Уравнение регрессии в натуральном виде.
11. Ответы на контрольные вопросы.
12. Выводы по лабораторной работе.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Почему в планах второго порядка возрастает минимально необходимое количество точек в спектре плана? Как определяется число членов квадратичной модели?
2. В каких случаях используют квадратичную модель объекта?
3. Дайте определение ЦКП.
4. Цель натурализации уравнения регрессии.
5. Чем обеспечивается ортогональность столбцов матрицы F?
6. Определение ОЦКП. Каким образом для ОЦКП выбирается числовое значение α (звездного плеча).

Лабораторная работа 5

ПРИМЕНЕНИЕ ПОЛНОГО ФАКТОРНОГО ЭКСПЕРИМЕНТА ПРИ ПРОВЕДЕНИИ ИССЛЕДОВАНИЙ

Цель работы: ознакомление студентов с использованием полного факторного эксперимента при исследовании технологических процессов.

Основные понятия

Студенты должны приобрести навыки постановки эксперимента, определения исследуемых факторов и области планирования эксперимента, составления матриц планирования полного факторного эксперимента, расчетов коэффициентов регрессии. При выполнении лабораторной работы студенты должны научиться работать с полученными математическими моделями. Перед студентами стоит задача изучения методов планирования полного факторного эксперимента применительно к технологическим задачам. Студенты должны освоить принципы составления матрицы планирования полного факторного эксперимента, проводить расчет коэффициентов регрессии, использовать статистические критерии для оценки однородности, нормальности экспериментальных данных, значимости коэффициентов и адекватности полученной математической модели. Студенты должны научиться работать с полученной моделью, строить изолинии.

Теоретическая часть

Планирование эксперимента – это процедура выбора числа и условий проведения опытов, необходимых и достаточных для получения математической модели процесса. При этом важно учитывать следующее: стремление к минимизации числа опытов; одновременное варьирование всех переменных, определяющих процесс; выбор четкой стратегии, позволяющей принимать обоснованные решения после каждой серии экспериментов. Перед проведением планирования активного эксперимента необходимо собрать дополнительную информацию об исследуемом объекте, для получения которой используются навыки и знания, которые получены ранее в предыдущих исследованиях или описаны в литературе.

При использовании метода активного планирования весь эксперимент обычно разбивается на несколько этапов. Информация, полученная после каждого этапа, используется для планирования исследований на следующем этапе. Планирование эксперимента позволяет варьировать

ряд факторов и получать одновременно количественные оценки всех проявляющихся эффектов. При этом в отличие от классического регрессионного анализа, избежать корреляции между коэффициентами уравнения регрессии. При статистическом подходе математическая модель объекта или процесса представляется в общем виде полиномом *n*-степени, т. е. отрезком ряда Тейлора, в который разлагается неизвестная функция.

Под планированием эксперимента понимают процесс определения числа и условий проведения опытов, необходимых и достаточных для решения поставленной задачи с требуемой точностью. Эффективность использования МПЭ при исследовании процессов в различных объектах объясняется тем, что их наиболее важные характеристики являются случайными величинами, распределение которых близко к нормальному.

Основная задача исследований – оптимизация, заключающаяся в нахождении совокупности варьируемых параметров, при которых выбранная целевая функция (параметр оптимизации) принимает экстремальное значение.

При решении задачи предполагается, что оптимизируется один параметр, каждый из факторов управляем, результаты опытов должны воспроизводиться. Для построения математической модели используется полный или дробный факторный план эксперимента, обладающий оптимальной матрицей планирования.

При планировании эксперимента цель исследования должна быть четко сформулирована и должна иметь количественную оценку. Характеристику цели, заданную количественно, называют параметром оптимизации.

Границы изменения факторов определяются так, чтобы обеспечить условия физической реализации процесса. Факторы должны быть управляемыми, независимыми (устанавливаться на любом уровне независимо друг от друга), однозначными (не являться функциями других). Кроме того, должны быть соблюдены некоторые ограничения:

1. Ограничения, связанные с условиями соблюдения техники безопасности при изучении данного процесса.

2. Ограничения, связанные с изменением экологической ситуации (использование веществ свыше предельно допустимой концентрации; проведением экспериментов, повлекших за собой ухудшение экологической ситуации).

3. Ограничения, связанные с технико-экономическими соображениями (дефицитность отдельных элементов, стоимость сырья, и т. д.).

4. На выбор интервала варьирования также накладываются ограничения: он не может быть меньше ошибки, с которой экспериментатор фиксирует уровень фактора, и не может быть настолько большим, что верх-

ний и нижний уровни оказались за пределами области определения. Необходимо отметить, что при выборе верхнего и нижнего уровней факторов необходимо учитывать ограничения, связанные с свойствами объекта исследования.

5. Принципиальные ограничения (например, если исследуемый фактор «температура», то ее нижний предел не может быть ниже абсолютного нуля).

6. Ограничения, связанные с конкретными условиями проведения процесса (например, верхний уровень температуры нельзя поднять выше температуры плавления материала, из которого сделан реактор).

7. Ограничения, связанные с условиями деградации процесса либо деструкцией изучаемого материала (параметры процесса после его полного завершения; свойства жидкости после ее испарения, свойства композиции после ее разрушения).

8. Ограничения, связанные с фазовыми переходами вещества, либо составляющих его компонентов.

Полный факторный эксперимент (2^k) целесообразно проводить в том случае, если он не продолжителен по времени и требует небольших затрат (число варьируемых факторов k мало). Для каждого включенного в эксперимент факторов $(x_1, x_2, .. x_k)$ устанавливают только два уровня: верхний x_{imax} и нижний x_{imin}. Основной уровень вычисляется:

$$x_{осн} = \frac{x_{i\,max} + x_{i\,min}}{2}, \qquad (1)$$

а интервал варьирования по выражению:

$$J_i = \frac{x_{i\,max} - x_{i\,min}}{2}. \qquad (2)$$

Полный факторный эксперимент

Для полного факторного эксперимента, в котором реализуются все возможные сочетания уровней факторов, число опытов определяется по следующей формуле:

$$N = 2^k, \qquad (3)$$

где N – число опытов; k – число факторов.

Эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется полным факторным экспериментом.

Вводят условное обозначение верхнего (+), нижнего (–) и основного уровня (0). Затем строится план эксперимента (например, для двух факторов, приведенный в таблице 1):

Таблица 1 – Матрица планирования

Номер точки плана	Фактор		Значение параметра оптимизации
	X_1	X_2	
1	–	–	Y_1
2	+	–	Y_2
3	–	+	Y_3
4	+	+	Y_4

В первом столбце знаки чередуются через один, во втором – через два(в третьем – через четыре и т.д. по степеням 2, если необходимо построить план для трех и более факторов).

Составленная матрица планирования должна соответствовать свойствам полного факторного эксперимента.

1. Симметричность относительно центра эксперимента:

$$\sum X_{ij} = 0, \tag{4}$$

где j – номер фактора; N – число опытов, $j = 1 \ldots k$.

2. Условие нормировки:

$$\sum X_{ij}^{\,2} = N. \tag{5}$$

Ортогональность матрицы:

$$\sum X_{ij} X_{iu} = 0. \tag{6}$$

$j = u$; $j, u = 0, 1, 2 \ldots k$.

Дробный факторный эксперимент

Количество опытов в полном факторном эксперименте значительно превосходит число определяемых коэффициентов линейной модели. Другими словами, полный факторный эксперимент обладает большой избыточностью опытов.

Чтобы сократить число опытов нужно новому фактору присвоить вектор-столбец матрицы, принадлежащий взаимодействию, которым можно пренебречь. Тогда значение нового фактора в условиях опытов определяется знаками этого столбца. Так, план дробного факторного эксперимента 2^{3-1} будет иметь вид, приведенный в таблице 2.

Таблица 2 – Матрица планирования дробного факторного эксперимента

Номер точки плана	Факторы			Значение параметра оптимизации
	X_1	X_2	$(X_3)\ X_1X_2$	
1	–	–	+	Y_1
2	+	–	–	Y_2
3	–	+	–	Y_3
4	+	+	+	Y_4

Запись 2^{k-m} обозначает: 2 – количество уровней факторов, k – количество факторов, m – количество факторов, введенных вместо эффектов взаимодействия. Эффект взаимодействия, подлежащий замене выбирается из условия максимального разрешения или априорной информации, имеющейся об эффектах взаимодействия.

Построение линейной модели процесса

Записываем уравнение процесса в виде:

$$Y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + \ldots + b_{1..k}x_1 \cdot \ldots \cdot x_k \qquad (7)$$

Для движения к точке оптимума необходимо вычислить коэффициенты линейной модели по следующей формуле:

$$b_j = \sum x_{ij}\ y_i / N. \qquad (8)$$

Следует учесть, что для расчетов используются кодированные значения факторов, которые определяются по формуле:

$$\tilde{x}_j = (x_j - x_{j0}) / J_j, \quad (9)$$

где $\tilde{x}_j$ – кодированное значение фактора, x_j – натуральное значение фактора, x_{j0} – натуральное значение основного уровня, J_j – интервал варьирования, j – номер фактора.

Коэффициент b_0 есть среднее арифметическое значений параметра оптимизации.

Каждый эксперимент содержит элемент неопределенности вследствие ограниченности экспериментального материала. Постановка параллельных опытов не дает полностью совпадающих результатов, так как всегда существует ошибка опыта. Ошибка опыта может быть определена по следующим показателям:

1. Среднее арифметическое результатов:

$$\bar{y} = (y_1 + y_2 + y_3 + \ldots + y_n) / n = \sum y_i / n, \quad (10)$$

где y_i – результаты экспериментов, n – количество опытов в серии.

2. Дисперсия – среднее значение квадрата отклонений величины от ее среднего значения:

$$S^2 = \sum (y_i - y)^2 / (n - 1), \quad (11)$$

где $(n - 1)$ – число степеней свободы, равное количеству опытов минус единица.

2. Квадратичная ошибка или стандарт:

$$S = \sqrt{\sum \frac{(y_i - y)^2}{n - 1}}. \quad (12)$$

Стандарт имеет размерность той величины, для которой он вычислен.

Для определения ошибок опыта используем критерий Стьюдента:

$$\frac{y - \bar{y}}{S} \geq t, \quad (13)$$

где t – табличное значение критерия Стьюдента.

Полученные дисперсии проверяем на однородность. Однородность дисперсий означает, что среди всех суммируемых дисперсий нет таких, которые бы значительно превышали все остальные. Проверка однородности дисперсий производится с помощью различных статистических критериев. Простейшим из них является критерий Фишера, предназначенный для сравнения двух дисперсий. При его использовании в случае, когда дисперсий более 2 критерий Фишера (Ф-критерий) представляет собой отношение большей дисперсии к меньшей. Полученная величина сравнивается с табличной.

Если критерий Фишера, полученный расчетным путем меньше табличного значения, то дисперсии однородны.

При n параллельных опытах и однородности дисперсии, расчет дисперсии параметра оптимизации проводится по следующей формуле:

$$S_y^2 = \frac{\sum_{i=1}^{N} S_i}{N}. \qquad (14)$$

После построения модели необходимо провести проверку ее адекватности.

С этой целью вычисляем дисперсию адекватности по формуле:

$$S_{ад}{}^2 = \Delta\sum y_i^2 / f. \qquad (15)$$

Остаточная дисперсия, или дисперсия адекватности $S_{ад}{}^2$ – это остаточная сумма квадратов, деленная на число степеней свободы.

Числом степеней свободы в статистике называется разность между числом опытов и числом коэффициентов, которые уже вычислены по результатам этих опытов независимо друг от друга. Число степеней свободы вычисляется по следующей формуле:

$$f = N - (k + 1), \qquad (16)$$

где N – число серий опытов, k – количество факторов.

Для проверки адекватности модели используется $Ф$-критерий Фишера, который определяется следующей формулой:

$$F = S_{ад}^2 / S_y^2. \qquad (17)$$

Если табличное значение критерия больше расчетного, модель адекватна. Столбцы таблицы связаны с определенным числом степеней свободы для числителя f_1 , строки – для знаменателя f_2.

Если полученная линейная модель неадекватна, то необходимо учесть эффекты взаимодействия или проверить все ли факторы учтены.

После проверки адекватности модели проводим проверку значимости каждого коэффициента. Для этого необходимо рассчитать дисперсии коэффициента регрессии.

Дисперсия коэффициента регрессии вычисляется по следующей формуле:

$$S^2_{\{bj\}} = S^2_y / N. \quad (18)$$

Из формулы видно, что дисперсии всех коэффициентов равны друг другу, так как они зависят только от ошибки опыта и числа опытов.

На основе полученной дисперсии коэффициентов регрессии строим доверительный интервал по формуле:

$$b\Delta_j S\{b \cdot t \pm =_j\}, \quad (19)$$

где $S\{b_j\}$ – квадратичная ошибка коэффициента регрессии; t – табличное значение критерия Стьюдента.

Коэффициент является значимым, если его абсолютная величина больше доверительного интервала.

Если построенная модель адекватна и большинство факторов значимы, то переходим к движению по градиенту, на основе которого исследуем поведение объекта в точках, отличных от заданных. Изменяя факторы пропорционально величинам коэффициентов регрессии с учетом их знака, мы будем двигаться в направлении градиента функции отклика по самому крутому пути. Поэтому процедура движения к точке стационарной области называется крутым восхождением. Крутое восхождение заканчивают достижением оптимума.

Порядок проведения работы

Определить область планирования эксперимента, число действующих факторов, функцию отклика.

1. Провести проверку экспериментальных данных на однородность и нормальность.

2. Провести расчет матрицы планирования полнофакторного эксперимента, занести матрицу в протокол. Необходимо обратить внимание, что матрица заполняется в строгом соответствии с планом эксперимента.

3. Получить уравнение регрессии. Занести результаты в протокол. Провести сравнение экспериментальных и расчетных значений. Занести в протокол полученную таблицу.

4. Провести оценку значимости коэффициентов регрессии и оценку адекватности полученного уравнения. Результаты занести в протокол.

5. Рассчитать значения выходного параметра в 4 точках. Занести результаты в протокол.

6. Рассчитать отклонение расчетного значения выходного параметра от экспериментальных данных в центре плана.

7. Написать выводы о проделанной работе.

8. Подготовиться к отчету лабораторной работы преподавателю.

Контрольные вопросы

1. Какие вы знаете методы активного эксперимента. Назовите основные отличия методов пассивного и активного эксперимента?

2. Какой порядок выборов действующих факторов, области определения эксперимента?

3. Какой имеют вид уравнения регрессии при полном факторном эксперименте?

4. Как производится построение матрицы планирования полного факторного эксперимента? Ее назначение?

5. Каков порядок расчет коэффициентов математической модели?

6. Как производится определение значимости коэффициентов регрессии?

Лабораторная работа 6

ИНТЕРПОЛЯЦИЯ

Цель работы: Научиться применять формулы для приближенного вычисления значения функции в заданной точке, проектировать и программно реализовывать рекуррентные формулы.

Формулы Маклорена и Тейлора

Общие сведения: Формулой Тейлора для произвольной функции $f(x)$ называется формула:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_n(x), \qquad (1)$$

Здесь x_0 некоторая произвольная (опорная) точка, в которой может быть вычислена функция $f(x)$ и все ее производные:

$$R_n(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k - \qquad (2)$$

остаточный член формулы Тейлора.

Формула Тейлора применяется для приближенного вычисления значения функции в заданной точке. Для этого выбирают такое количество слагаемых n суммы в формуле (1), чтобы модуль остаточного члена (2) не превышал некоторой заданной величины (точности) ε.

Так как остаточный член $R_n(x)$ не может быть точно вычислен, его принимают равным величине:

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}, \qquad (3)$$

где ξ – некоторая средняя точка интервала $[x_0, x]$. На практике погрешность вычисления значения функции по формуле (2) обычно оценивают

модулем первого слагаемого остаточного члена (что совпадает со случаем $\xi = x_0$):

$$\Delta_{f(x)} = \left| \frac{f^{(n+1)}(x_0)}{(n+1)!}(x - x_0)^{n+1} \right|. \quad (4)$$

Формула Маклорена является частным случаем формулы Тейлора при $x_0 = 0$. В этом случае обычно производные $f^{(k)}(x)$ принимают достаточно простой вид и могут быть легко вычислены, поэтому формулу Маклорена чаще применяют на практике. Формулы (1) и (4) для данного случая приобретают вид:

$$f(x) \approx \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k + R_n(x), \quad (5)$$

$$\Delta_{f(x)} \approx \left| \frac{f^{(n+1)}(0)}{(n+1)!} x^{n+1} \right|. \quad (6)$$

Ряд в формуле Маклорена для большинства элементарных функций (e^x, sin$x$, cos$x$, ln$x$ и т. д.) сходится при $x \in (-\infty, +\infty)$ (для ln$x$ – $x \in (0, +\infty)$), однако, чем меньше x , тем меньшее количество слагаемых требуется для достижения заданной точности.

Ниже приведены формулы Маклорена для некоторых элементарных функций:

В случае, если $x >> 1$, вычисление значения с помощью формулы Маклорена до заданной точности может потребовать значительного количества слагаемых. В этом случае удобно выразить аргумент следующим образом: $x = E(x) + q(x)$, где $E(x)$ – большое число, для которого можно легко вычислить значение $f(x)$, а $q(x)$ – оставшаяся часть числа x (обычно $|q(x)| < 1$). Тогда достаточно применить формулу только для $q(x)$. Например, для вычисления $e^{12,43}$, представим аргумент в виде $x = 12 + 0{,}43$. В этом случае $e^{12,43} = e^{12} \cdot e^{0,43}$. В этом выражении e^{12} вычисляем, просто перемножая 12 раз число e, а для $e^{0,43}$ применяем формулу Маклорена, при этом количество вычислений снижается на 1–2 порядка по сравнению с решением «в лоб».

Пример:

Вычислить $\sqrt{e}$ с точностью до 0,0001.

Решение

$f(x) = e^{\frac{1}{2}}$. Запишем ряд для этой функции:

$$f(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + K + \frac{x^n}{n!} + K$$

Обозначим через u_k k-й член этой суммы. Тогда $u_k = \frac{x^k}{k!}$. Выразим u_k через u_{k-1}, получим: $u_k = \frac{x}{k} u_{k-1}$. Таким образом, можно составить следующую рекуррентную схему:

$$u_0 = 1,$$

$$u_k = \frac{x}{k} u_{k-1},$$

пока

$$|u_k| > \varepsilon.$$

Результаты работы программы, производящей вычисления по заданной формуле будут иметь вид:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \qquad \ln x = -2\sum_{k=1}^{\infty} \frac{1}{2k-1}\left(\frac{1-x}{1+x}\right)^{2k-1}$$

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \qquad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

$$(1+x)^m = \sum_{k=0}^{\infty} \frac{m(m-1)K(m-k+1)}{k!} x^k \qquad x \in (-1,\ 1)$$

$$\operatorname{arctg} x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \qquad x \in (-1,\ 1)$$

k	u_k	$\sum u_k$
0	1,00000000	1,00000000
1	0,50000000	1,50000000
2	0,12500000	1,62500000
3	0,02083333	1,64583333
4	0,00260417	1,64843750
5	0,00026042	1,64869792
6	0,00002170	

Ответ округляем до одного сомнительного знака. В качестве погрешности берем модуль последнего полученного значения u_k, округляя до 2 знаков в большую сторону.

Ответ: $\sqrt{e}$ = 1,64870 с точностью Δ = 0,000022. Для проверки вычислим $\sqrt{e}$, используя стандартную функцию языка Паскаль exp(1/2).

Проверка: 1.6487212707.

Варианты заданий

№	Вычислить	С точностью	№	Вычислить	С точностью
1	$\frac{1}{\sqrt[5]{e}}$	0,00001	13*	ctg 0,25	0,0001
2	$\cos 18^0$	0,0001	14	$\sqrt{2}$	0,000001
3	$\sqrt[5]{1{,}1}$	0,0001	15	$\sqrt{3}$	0,000001
4	$\sqrt[3]{130}$	0,001	16	$\sqrt[3]{3}$	0,000001
5	ln 1,04	0,0001	17	$\sqrt[3]{10}$	0,000001
6	$\operatorname{arctg}\frac{1}{3}$	0,001	18	ln (3e)	0,0001
7	ln 2	0,0001	19	ln (11$\sqrt{e}$)	0,0001
8	ln 5	0,0001	20	5ln 5	0,0001
9	e	0,00001	21	sh0,3	0,00001
10	sin 9°	0,0001	22	sh0,4	0,0001
11	$\sqrt{27}$	0,001	23*	tg28°	0,001

Многочленные приближения (формулы Чебышева)

Общие сведения: Формулы Чебышева являются одним из методов интерполяции функции многочленом. Главным достоинством этих формул является достаточно фиксированная высокая точность вычислений при постоянных заранее известных коэффициентах и фиксированном числе слагаемых. Ниже приведены формулы, основанные на многочленах Чебышева для некоторых элементарных функций.

1) $e^x = \sum_{k=0}^{7} a_k x^k$

с точностью $2 \cdot 10^{-7}$ для $|x| \le 1$.

где $a_0 = 0{,}9999998$ $a_1 = 1{,}0000000$ $a_2 = 0{,}5000063$

$a_3 = 0{,}1666674$ $a_4 = 0{,}0416350$ $a_5 = 0{,}0083298$

$a_6 = 0{,}0014393$ $a_7 = 0{,}0002040$

2) $\ln(1+x) = \sum_{k=1}^{7} a_k x^k$

с точностью $2{,}2 \cdot 10^{-7}$ для $0 \le x \le 1$.

где $a_1 = 0{,}999981028$ $a_2 = -0{,}499470150$ $a_3 = 0{,}328233122$

$a_4 = 0{,}225873284$ $a_5 = 0{,}134639267$ $a_6 = -0{,}055119959$

$a_7 = 0{,}010757369$

3) $\sin x = \sum_{k=0}^{4} a_{2k+1} x^{2k+1}$

с точностью $6\cdot10^{-9}$ для $|x| \le \frac{\pi}{2}$.

где $a_1 = 1{,}00000002$ $a_3 = -0{,}166666589$ $a_5 = 0{,}008333075$

$a_7 = -0{,}000198107$ $a_9 = 0{,}000002608$

4) $\cos x = \sum_{k=0}^{5} a_{2k} x^{2k}$

с точностью $2\cdot10^{-9}$ для $|x| \le 1$.

где $a_0 = 1{,}000000000000$ $a_2 = -0{,}499999999942$

$a_4 = 0{,}041666665950$ $a_6 = -0{,}001388885683$

$a_8 = 0{,}000024795132$ $a_{10} = -0{,}000000269591$

5) $\operatorname{tg} x = \sum_{k=0}^{6} a_{2k+1} x^{2k+1}$

с точностью $2\cdot10^{-8}$ для $|x| \le \frac{\pi}{4}$.

где $a_1 = 1{,}00000002$ $a_3 = 0{,}33333082$ $a_5 = 0{,}13339762$

$a_7 = 0{,}05935836$ $a_9 = 0{,}02457096$ $a_{11} = 0{,}00294045$

$a_{13} = 0{,}00947324$

В качестве вариантов заданий могут быть использованы задания лабораторной работы 12.

Варианты заданий

№	Вычислить	С точностью	№	Вычислить	С точностью
1	$\frac{1}{\sqrt[5]{e}}$	0,00001	13*	ctg 0,25	0,0001
2	$\cos 18^0$	0,0001	14	$\sqrt{2}$	0,000001
3	$\sqrt[5]{1{,}1}$	0,0001	15	$\sqrt{3}$	0,000001
4	$\sqrt[3]{130}$	0,001	16	$\sqrt[3]{3}$	0,000001
5	ln 1,04	0,0001	17	$\sqrt[3]{10}$	0,000001
6	$\operatorname{arctg}\frac{1}{3}$	0,001	18	ln (3e)	0,0001
7	ln 2	0,0001	19	ln (11$\sqrt{e}$)	0,0001
8	ln 5	0,0001	20	5ln 5	0,0001
9	e	0,00001	21	sh0,3	0,00001
10	sin 9°	0,0001	22	sh0,4	0,0001
11	$\sqrt{27}$	0,001	23*	tg28°	0,001
12*	tg 0,14	0,0001	24*	ctg 74°	0,001

Интерполяционный многочлен Лагранжа

Общие сведения: Для приближенного вычисления значений неизвестной функции по заданной таблице ее значений $y_1 = f(x_1)$, K, $y_n = f(x_n)$. Интерполяционный ***многочлен Лагранжа*** предлагает наиболее общую формулу для решения данной задачи:

$$f(x) \approx L_n(x) = \sum_{i=1}^{n} y_i \prod_{i=1} \frac{x - x_j}{x_i - x_j}.$$

Остаточный член интерполяционной формулы с многочленом Лагранжа равен:

$$R_n = \frac{f^{(n)}(\xi)}{n!} \prod_{i=1} \frac{x - x_j}{x_i - x_j},$$

где $\xi \in [x_1, x_n]$. .

Варианты заданий

1–5. Вычислить конечные разности и значения функций в точках 0,2, 0,45, 0,55, 0,82, 1,15, 1,25, 1,5 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция	X1	X2	X3	X4	X5	X6	X7	X8
	x	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2
1	ex	1,64872	1,82212	2,01375	2,22554	2,45960	2,71828	3,004717	3,32012
2	$\sin x$	0,47943	0,56464	0,64422	0,71736	0,78333	0,84147	0,89121	0,93204
3	$\cos x$	0,87758	0,82534	0,76484	0,69671	0,62161	0,54030	0,45360	0,36236
4	$\mathrm{tg}\, x$	0,54630	0,68414	0,84229	1,02964	1,26016	1,55741	1,96476	2,57215
5	$\ln x$	–0,69315	–0,51083	–0,35667	–0,22314	–0,10536	0	0,09531	0,18232

Сделать сравнительный анализ методов

6–8. Вычислить конечные разности и значения функций в точках 0,7, 0,9, 1,1, 1,5, 2,3, 2,5, 2,7 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция	X1	X2	X3	X4	X5	X6	X7	X8
	x	1,0	1,2	1,4	1,6	1,8	2,0	2,2	2,4
6	ex	2,71828	3,32012	4,05519	4,95303	6,04964	7,38905	9,02501	11,02317
7	$\sin x$	0,84147	0,93204	0,98544	0,99957	0,97384	0,90929	0,80849	0,67546
8	$\cos x$	0,54030	0,36236	0,16996	–0,02919	–0,22720	–0,41614	–0,58850	–0,73739

Сделать сравнительный анализ методов

Схема Эйткена

Общие сведения: Для приближенного вычисления значений неизвестной функции по заданной таблице ее значений $y_1 = f(x_1)$, K, $y_n = f(x_n)$.

Схема Эйткена предлагает более удобную форму вычисления $f(x)$ по формуле Лагранжа. Основная идея данного метода заключается в следую-

щем. На первом этапе вычисляются многочлены $L_{1,2}(x)$, $L_{2,3}(x)$, K, $L_{2n-2,\ n-1}(x)$, построенные на каждой паре соседних узлов:

$$L_{i,\,i+1} = \frac{1}{x_{i+1} - x_i} \begin{vmatrix} y_i & x_i - x \\ y_{i+1} & x_{i+1} - x \end{vmatrix},\ i = \overline{1, n-1}.$$

Затем на их основе вычисляются многочлены, построенные на тройках соседних узлов:

$$L_{i,\,i+1,i+2}(x) = \frac{1}{x_{i+2} - x_i} \begin{vmatrix} L_{i,\,i+1}(x) & x_i - x \\ L_{i+1,\,i+2}(x) & x_{i+2} - x \end{vmatrix},\ i = \overline{1, n-2},$$

и т.д., пока не получится один многочлен, построенный на всех узлах интерполяции:

$$L_{1,\,2,\,K,\,n}(x) = \frac{1}{x_n - x_1} \begin{vmatrix} L_{1,\,2,\,K,\,n\text{-}1}(x) & x_1 - x \\ L_{\,2,\,3,\,K,\,n}(x) & x_n - x \end{vmatrix}.$$

Нетрудно убедиться, что $L_{\text{n}}(x) \equiv L_{1,\,2,\,K,\,n}(x)$.

Варианты заданий

1–5. Вычислить конечные разности и значения функций в точках 0,2, 0,45, 0,55, 0,82, 1,15, 1,25, 1,5 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция	X1	X2	X3	X4	X5	X6	X7	X8
	x	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2
1	ex	1,64872	1,82212	2,01375	2,22554	2,45960	2,71828	3,004717	3,32012
2	$\sin x$	0,47943	0,56464	0,64422	0,71736	0,78333	0,84147	0,89121	0,93204
3	$\cos x$	0,87758	0,82534	0,76484	0,69671	0,62161	0,54030	0,45360	0,36236
4	$\text{tg}\, x$	0,54630	0,68414	0,84229	1,02964	1,26016	1,55741	1,96476	2,57215
5	$\ln x$	–0,69315	–0,51083	–0,35667	–0,22314	–0,10536	0	0,09531	0,18232

Сделать сравнительный анализ методов

6–8. Вычислить конечные разности и значения функций в точках 0,7, 0,9, 1,1, 1,5, 2,3, 2,5, 2,7 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция	X1	X2	X3	X4	X5	X6	X7	X8
	x	1,0	1,2	1,4	1,6	1,8	2,0	2,2	2,4
6	ex	2,71828	3,32012	4,05519	4,95303	6,04964	7,38905	9,02501	11,02317
7	$\sin x$	0,84147	0,93204	0,98544	0,99957	0,97384	0,90929	0,80849	0,67546
8	$\cos x$	0,54030	0,36236	0,16996	–0,02919	–0,22720	–0,41614	–0,58850	–0,73739

Сделать сравнительный анализ методов.

Первая интерполяционная формула Ньютона

Общие сведения: Для приближенного вычисления значений неизвестной функции по заданной таблице ее значений $y_1 = f(x_1)$, K, $y_n = f(x_n)$.

Первая формула Ньютона используется для интерполяции и экстраполяции в начале таблицы:

$$f(x) \approx y_1 + q\Delta y_1 + \frac{q(q-1)}{2!}\Delta^2 y_1 + K + \frac{q(q-1)K(q-n+2)}{(n-1)!}\Delta^{n-1} y_1,$$

где

$$q = \frac{x - x_1}{h}.$$

Остаточный член первой формулы Ньютона равен

$$R_n(x) = h^n \frac{q(q-1)K(q-n+1)}{n!} f^{(n)}(\xi).$$

Варианты заданий

1–5. Вычислить конечные разности и значения функций в точках 0,2, 0,45, 0,55, 0,82, 1,15, 1,25, 1,5 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция	X1	X2	X3	X4	X5	X6	X7	X8
	x	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2
1	ex	1,64872	1,82212	2,01375	2,22554	2,45960	2,71828	3,004717	3,32012
2	$\sin x$	0,47943	0,56464	0,64422	0,71736	0,78333	0,84147	0,89121	0,93204
3	$\cos x$	0,87758	0,82534	0,76484	0,69671	0,62161	0,54030	0,45360	0,36236
4	$\mathrm{tg}\, x$	0,54630	0,68414	0,84229	1,02964	1,26016	1,55741	1,96476	2,57215
5	$\ln x$	–0,69315	–0,51083	–0,35667	–0,22314	–0,10536	0	0,09531	0,18232

Сделать сравнительный анализ методов

6-8. Вычислить конечные разности и значения функций в точках 0,7, 0,9, 1,1, 1,5, 2,3, 2,5, 2,7 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция	X1	X2	X3	X4	X5	X6	X7	X8
	x	1,0	1,2	1,4	1,6	1,8	2,0	2,2	2,4
6	ex	2,71828	3,32012	4,05519	4,95303	6,04964	7,38905	9,02501	11,02317
7	sin x	0,84147	0,93204	0,98544	0,99957	0,97384	0,90929	0,80849	0,67546
8	cos x	0,54030	0,36236	0,16996	–0,02919	–0,22720	–0,41614	–0,58850	–0,73739

Сделать сравнительный анализ методов.

Вторая интерполяционная формула Ньютона

Общие сведения: Для приближенного вычисления значений неизвестной функции по заданной таблице ее значений $y_1 = f(x_1)$, K, $y_n = f(x_n)$.

Вторая формула Ньютона применяется для интерполяции и экстраполяции в конце таблицы:

$$f(x) \approx y_n + q\Delta y_{n-1} + \frac{q(q+1)}{2!}\Delta^2 y_{n-2} + K + \frac{q(q+1)K(q+n-2)}{(n-1)!}\Delta^{n-1} y_1,$$

где

$$q = \frac{x - x_n}{h}.$$

Остаточный член второй формулы Ньютона равен

$$R_n(x) = h^n \frac{q(q+1)K(q+n-1)}{n!} f^{(n)}(\xi).$$

Варианты заданий

1–5. Вычислить конечные разности и значения функций в точках 0,2, 0,45, 0,55, 0,82, 1,15, 1,25, 1,5 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция								
	x	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2
1	ex	1,64872	1,82212	2,01375	2,22554	2,45960	2,71828	3,004717	3,32012
2	$\sin x$	0,47943	0,56464	0,64422	0,71736	0,78333	0,84147	0,89121	0,93204
3	$\cos x$	0,87758	0,82534	0,76484	0,69671	0,62161	0,54030	0,45360	0,36236
4	$\operatorname{tg} x$	0,54630	0,68414	0,84229	1,02964	1,26016	1,55741	1,96476	2,57215
5	$\ln x$	–0,69315	–0,51083	–0,35667	–0,22314	–0,10536	0	0,09531	0,18232

Сделать сравнительный анализ методов

6–8. Вычислить конечные разности и значения функций в точках 0,7, 0,9, 1,1, 1,5, 2,3, 2,5, 2,7 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

	Функция								
№	x	1,0	1,2	1,4	1,6	1,8	2,0	2,2	2,4
6	ex	2,71828	3,32012	4,05519	4,95303	6,04964	7,38905	9,02501	11,02317
7	sin x	0,84147	0,93204	0,98544	0,99957	0,97384	0,90929	0,80849	0,67546
8	cos x	0,54030	0,36236	0,16996	–0,02919	–0,22720	–0,41614	–0,58850	–0,73739

Сделать сравнительный анализ методов

Первая интерполяционная формула Гаусса

Общие сведения: Для приближенного вычисления значений неизвестной функции по заданной таблице ее значений $y_1 = f(x_1)$, K, $y_n = f(x_n)$.

Интерполяционные формулы Гаусса используются для интерполяции в середине таблицы. Для $x \in (x_k, x_{k+1})$ ***первая формула Гаусса*** имеет вид:

$$f(x) \approx G_{2i}(x) = y_k + \frac{\Delta y_k}{1!}t + \frac{\Delta^2 y_{k-1}}{2!}t(t-1) + \frac{\Delta^3 y_{k-1}}{3!}t(t-1)(t+1) + K +$$

$$+ \frac{\Delta^{2i} y_{k-i}}{(2i)!}t(t-1)(t+1)K(t-(i-1))(t+(i-1))(t-i),$$

где $t = \frac{x - x_n}{h}$, а i определяется требуемым количеством слагаемых либо количеством узлов (таблица конечных разностей имеет треугольную форму).

Варианты заданий

1–5. Вычислить конечные разности и значения функций в точках 0,2, 0,45, 0,55, 0,82, 1,15, 1,25, 1,5 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция								
	x	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2
1	ex	1,64872	1,82212	2,01375	2,22554	2,45960	2,71828	3,004717	3,32012
2	sin x	0,47943	0,56464	0,64422	0,71736	0,78333	0,84147	0,89121	0,93204
3	cos x	0,87758	0,82534	0,76484	0,69671	0,62161	0,54030	0,45360	0,36236
4	tg x	0,54630	0,68414	0,84229	1,02964	1,26016	1,55741	1,96476	2,57215
5	ln x	–0,69315	–0,51083	–0,35667	–0,22314	–0,10536	0	0,09531	0,18232

Сделать сравнительный анализ методов.

6–8. Вычислить конечные разности и значения функций в точках 0,7, 0,9, 1,1, 1,5, 2,3, 2,5, 2,7 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция								
	x	1,0	1,2	1,4	1,6	1,8	2,0	2,2	2,4
6	ex	2,71828	3,32012	4,05519	4,95303	6,04964	7,38905	9,02501	11,02317
7	sin x	0,84147	0,93204	0,98544	0,99957	0,97384	0,90929	0,80849	0,67546
8	cos x	0,54030	0,36236	0,16996	–0,02919	–0,22720	–0,41614	–0,58850	–0,73739

Лабораторная работа 7

ИНТЕРПОЛЯЦИЯ СПЛАЙНАМИ

Цель работы: Изучить интерполяцию сплайн-функциями, научиться строить сплайны для функции, заданной таблично.

Общие сведения: Сплайном называется непрерывная функция, принимающая в узлах интерполяции соответствующие значения y_1, y_2K, y_n и описываемая на отдельных отрезках $[x_i, x_{i+1}]$ ($i = \overline{1, n-1}$) некоторыми полиномами $P_i(x)$ невысокого порядка (обычно второго или третьего). Сплайны хороши тем, что дают возможность получить приближенный аналитический вид функции, при этом степень полинома остается невысокой. Главным недостатком сплайнов является то, что на каждом интервале $[x_i, x_{i+1}]$ функция приближается отдельным полиномом. Таким образом, для всего промежутка $[x_i, x_{i+1}]$ требуется построить $n - 1$ полином. При увеличении степени сплайна повышается точность интерполяции, однако, увеличивается количество вычислений и порядок полинома. Простейшим случаем сплайн-функции является линейный сплайн, получающийся соединением точек (x_1, y_1), …, (x_n, y_n) ломаной линией. В этом случае

$$P_i(x) = a_i x + b_i,$$

где коэффициенты a_i и b_i можно найти по формулам:

$$a_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \quad b_i = y_i - a_i x_i .$$

Параболический сплайн строится на полиномах второго порядка:

$$P_i(x) = a_i x^2 + b_i x + c_i.$$

При построении параболического сплайна обычно считается заданной производная функции в первом узле $f'(x_1)$. Коэффициенты полиномов такого сплайна могут быть вычислены по следующим рекуррентным формулам:

$$a_1 = \frac{1}{x_2 - x_1}\left(\frac{y_2 - y_1}{x_2 - x_1} - f'(x_1)\right), \quad b_1 = f'(x_i) - 2a_1 x_1 ,$$

$$c_1 = y_1 = f'(x_1)x_1 + a_1 x_1^2 ;$$

$$a_i = \frac{1}{x_{i+1} - x_i}\left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i} - \beta_i\right),\ b_1 = \beta_i - 2a_i x_i,\ c_i = y_i - \beta_i x_i + a_i x_i^2,$$

где $\beta_i = 2a_{i-1}x_i + b_{i-1}$.

Чаще всего на практике используют кубический сплайн, так как при достаточно низком порядке полиномов он дает удовлетворительную точность вычислений. Для полиномов $P_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i$ кубического сплайна коэффициенты могут быть получены по следующей рекуррентной схеме:

$$a_1 = \frac{1}{(x_2 - x_1)^2}\left(\frac{y_2 - y_1}{x_2 - x_1} - \frac{1}{2}f''x_i(x_2 - x_1) - f'(x_1)\right),\ b_1 = \frac{1}{2}f''(x_i) - 3a_i x_i,$$

$$c_1 = f'(x_1) - 3a_1 x_1^2 - 2b_1 x_1,$$

$$d_1 = y_1 - a_1 x_1^3 - b_1 x_1^2 - c_1 x_1,$$

где значения $f'(x_1)$ и $f''(x_1)$ считаются заданными (на практике часто полагают $f''(x_1) = 0$).

$$a_i = \frac{1}{(x_{i+1} - x_i)^2}\left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i} - y_i(x_{i+1} - x_i) - \beta_i\right),\ b_i = y_i - 3a_i x_i,$$

$$c_i = \beta_i - 3a_i x_i^2 - 2b_i x_i,$$

$$d_i = y_i - a_i x_i^3 - b_i x_i^2 - c_i x_i,$$

где $\beta_i = 3a_{i-1}x_i^2 + 2b_{i-1}x_i + c_{i-1},\ y_i = 3a_{i-1}x_i + b_{i-1}$.

Варианты заданий

1–10. По заданной таблице значений (из предыдущей лабораторной работы) составить линейный, параболический и кубический сплайны ($f'(x_1)$ и $f''(x_1)$ вычислить, используя проверочную функцию). В одной системе координат построить график функции и графики всех сплайнов. Сделать анализ представленных методов.

Варианты заданий

1–5. Вычислить конечные разности и значения функций в точках 0,2, 0,45, 0,55, 0,82, 1,15, 1,25, 1,5 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция								
	x	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2
1	ex	1,64872	1,82212	2,01375	2,22554	2,45960	2,71828	3,004717	3,32012
2	sin x	0,47943	0,56464	0,64422	0,71736	0,78333	0,84147	0,89121	0,93204
3	cos x	0,87758	0,82534	0,76484	0,69671	0,62161	0,54030	0,45360	0,36236
4	tg x	0,54630	0,68414	0,84229	1,02964	1,26016	1,55741	1,96476	2,57215
5	ln x	–0,69315	–0,51083	–0,35667	–0,22314	–0,10536	0	0,09531	0,18232

Сделать сравнительный анализ методов.

6–10. Вычислить конечные разности и значения функций в точках 0,7, 0,9, 1,1, 1,5, 2,3, 2,5, 2,7 с использованием каждого из описанных методов интерполяции по заданной таблице значений. Сделать проверку, используя стандартную функцию языка программирования:

№	Функция								
	x	1,0	1,2	1,4	1,6	1,8	2,0	2,2	2,4
6	ex	2,71828	3,32012	4,05519	4,95303	6,04964	7,38905	9,02501	11,02317
7	sin x	0,84147	0,93204	0,98544	0,99957	0,97384	0,90929	0,80849	0,67546
8	cos x	0,54030	0,36236	0,16996	–0,02919	–0,22720	–0,41614	–0,58850	–0,73739

Сделать сравнительный анализ методов

Лабораторная работа 8

АППРОКСИМАЦИЯ ФУНКЦИЙ

Цель работы: Изучить аппроксимацию функции, научиться применять МНК при решении задач.

Общие сведения

Пусть функциональная зависимость $y = f(x)$ задана в виде таблицы

x	X_1	X_2	…	X_n
y	Y_1	Y_2	…	Y_n

На практике часто бывает необходимо знать значения функции и в других точках, причём нет возможности для их непосредственного вычисления или измерения. Например, для этого нужно было бы решить очень сложную задачу, провести трудоёмкие или дорогостоящие эксперименты и т.д. Поэтому очень важным является построение такого алгоритма, который позволил бы по имеющимся табличным данным найти значение функции в любой другой точке заданного отрезка или вне его. Вот следующая задача: построить функцию $y' = F(x)$, которая с некоторой степенью точности может приближенно заменить (аппроксимировать) данную функцию $y = f(x)$. Эта аппроксимирующая функция должна быть такой, чтобы её отклонение в некотором смысле, от данной функции на заданном множестве $\{x_i\}$ было минимальным. На практике имеют место две основные задачи приближения функций. Одна из них состоит в том, что аппроксимирующая функция строится таким образом, чтобы её значения в точках $\{x_i\}$ в точности совпадали с табличными данными $\{y_i\}$. Это соответствует нахождению уравнения $y' = F(x)$ сплошной линии, которая проходит через заданные точки $\{x_i, y_i\}$. Такая задача называется интерполированием функции. Она применяется в основном в тех случаях, когда приведённые табличные данные имеют высокий порядок точности.

Далее рассматривается другая задача. Наиболее распространенным методом решения этой задачи является метод наименьших квадратов, в соответствии с которым ищем такую функциональную зависимость $y' = F(x)$, при которой:

$$S = \sum_{i=1}^{n} (y_i - y')^2 \to \min.$$

Из курса математики известно, что в минимуме функции её производная равна нулю. Рассматривая S как функцию двух переменных a и b, приравниваем нулю частные производные от S по a и по b. Т. е.

$$\frac{\partial S}{\partial a} = 0,$$

$$\frac{\partial S}{\partial b} = 0.$$

Пример 1.

В результате эксперимента получены величины (x_i, y_i) $(i = 1, \ldots, n)$, которые связаны отношением $y = a \cdot e^{b/x}$. Определить a и b таким образом, чтобы кривая $y = a \cdot e^{b/x}$ была самой близкой к точкам (x_i, y_i) в смысле наименьших квадратов.

Т. к. дана показательная регрессия, сделаем следующие преобразования:

$$y = a\, e^{b/x} \;\rightarrow\; \ln y = \ln a + b/x.$$

Сделаем замену: $c = \ln a$ и будем решать задачу относительно новых неизвестных c и b.

В соответствии с методом наименьших квадратов c и b выбирают так, чтобы

$$S = \sum_{i=1}^{n} (\ln y_i - c - b / x_i)^2 \;\rightarrow\; \min.$$

Из курса математики известно, что в минимуме функции её производная равна нулю. Рассматривая S как функцию двух переменных c и b, приравниваем нулю частные производные от S по c и по b:

$$\begin{cases} \dfrac{\partial S}{\partial c} = -2\sum\limits_{i=1}^{n} (\ln y_i - c - b / x_i) = 0, \\ \dfrac{\partial S}{\partial b} = -2\sum\limits_{i=1}^{n} (\ln y_i - c - b / x_i)\dfrac{1}{x_i} = 0 \end{cases}$$

Упростим выражения, и решим систему уравнений:

$$\begin{cases} \sum\limits_i \ln y_i = nc + b\sum\limits_i 1/x_i, \\ \sum\limits_i \ln y_i / x_i = c\sum\limits_i 1/x_i + b\sum\limits_i 1/x_i^2 \end{cases}$$

Решим эту систему уравнений методом Крамера:

$$\Delta = \begin{vmatrix} n & \sum_i 1/x_i \\ \sum_i 1/x_i & \sum_i 1/x_i^2 \end{vmatrix},\ \Delta b = \begin{vmatrix} n & \sum_i \ln y_i \\ \sum_i 1/x_i & \sum_i \ln y_i / x_i \end{vmatrix},$$

$$\Delta c = \begin{vmatrix} \sum_i \ln y_i & \sum_i 1/x_i \\ \sum_i \ln y_i / x_i & \sum_i 1/x_i^2 \end{vmatrix}$$

$$b = \frac{\Delta b}{\Delta}, c = \frac{\Delta c}{\Delta}$$

$$b = \frac{n\sum_i \ln y_i / x_i - \sum_i \ln y_i \sum_i 1/x_i}{n\sum_i 1/x_i^2 - \left(\sum_i 1/x_i\right)^2},$$

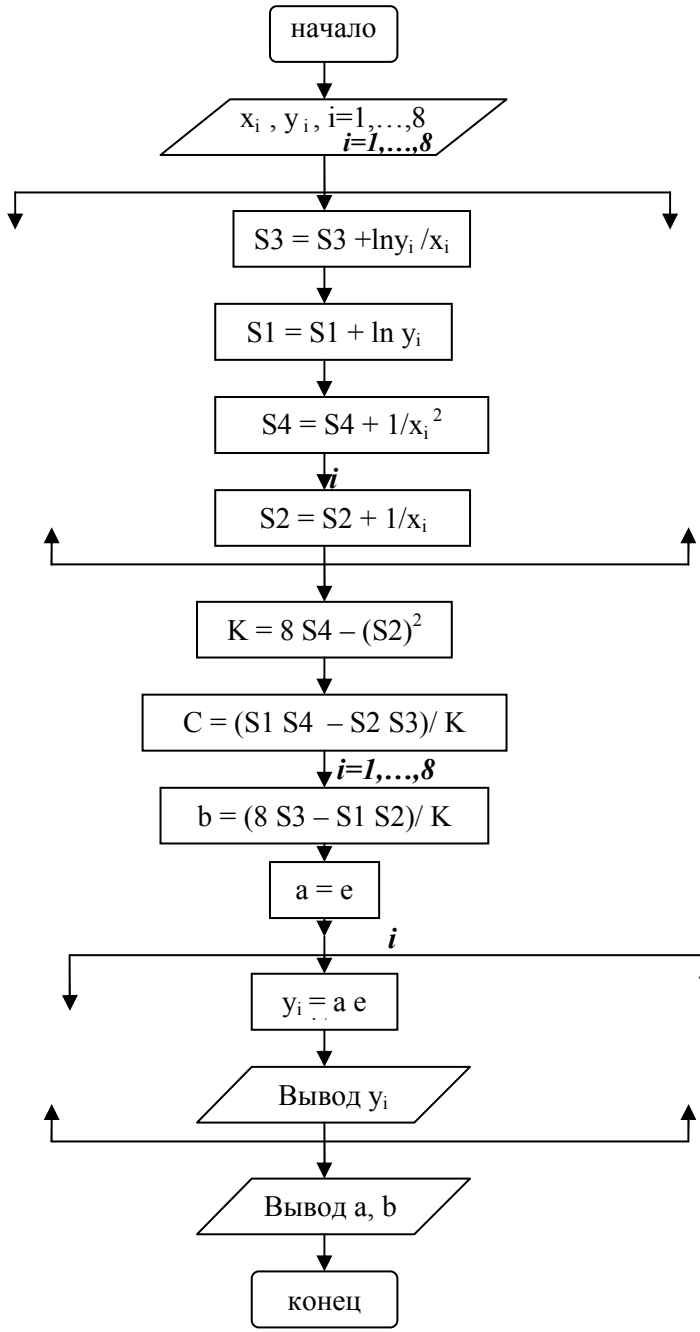
$$c = \frac{\sum_i \ln y_i \sum_{ш} 1/x_i^2 - \sum_i \ln y_i / x_i \sum_i 1/x_i}{n\sum_i 1/x_i^2 - \left(\sum_i 1/x_i\right)^2}$$

$$a = \exp(c).$$

Обозначения:

$$S2 = \sum_i 1/x_i, S1 = \sum_i \ln y_i, S4 = \sum_i 1/x_i^2, S3 = \sum_i \ln y_i / x_i$$

Блок-схема


начало
x_i , y_i , i=1,…,8
i=1,…,8
S3 = S3 +lny_i /x_i
S1 = S1 + ln y_i
S4 = S4 + 1/x_i^2
i
S2 = S2 + 1/x_i
K = 8 S4 – (S2)^2
C = (S1 S4 – S2 S3)/ K
i=1,…,8
b = (8 S3 – S1 S2)/ K
a = e
i
y_i = a e
Вывод y_i
Вывод a, b
конец

Исходные данные:

X	1	2	3	4	5	6	7	8
Y	14,2	11,4	2,9	2,5	2,1	2	1,9	1,8

Результаты вычислений.

X	1	2	3	4	5	6	7	8
Y	14,2	11,4	2,9	2,5	2,1	2	1,9	1,8
Y'	18,9	5,13	3,31	2,67	2,34	2,15	2,02	1,93

$$S = \sum_{i=1}^{8} (y_i - y'_i)^2 = 61.85$$

$A = 1.39,\ B = 2.61$

Пример 2.

В результате эксперимента получены величины (x_i, y_i) (i = 1, …, n), которые связаны отношением $y = ax^2+b\mathrm{tg}x$. Определить a и b таким образом, чтобы кривая $y = ax^2+b\mathrm{tg}x$ была самой близкой к точкам (x_i, y_i) в смысле наименьших квадратов.

В соответствии с методом наименьших квадратов a и b выбирают так, чтобы:

$$S = \sum_{i=1}^{n} (y_1 - ax_i^2 - b\,\mathrm{tg}\,x_i)^2 \rightarrow \min$$

Из курса математики известно, что в минимуме функции её производная равна нулю. Рассматривая S как функцию двух переменных a и b, приравниваем нулю частные производные от S по a и по b:

$$\begin{cases} \dfrac{\partial S}{\partial a} = -2\sum\limits_{i=1}^{n} (y_i - ax_i^2 - b\,\mathrm{tg}\,x_i)x_i^2 = 0, \\ \dfrac{\partial S}{\partial b} = -2\sum\limits_{i=1}^{n} (y_i - ax_i^2 - b\,\mathrm{tg}\,x_i)\,\mathrm{tg}\,x_i = 0 \end{cases}$$

Упростим выражения, и решим систему уравнений:

$$\begin{cases} \sum_i y_i x_i^2 = a\sum_i x_i^4 + b\sum_i x_i^{\,2}\,\mathrm{tg}\,x_i, \\ \sum_i y_i\,\mathrm{tg}\,x_i = a\sum_i x_i^2\,\mathrm{tg}\,x_i + b\sum_i \mathrm{tg}^2\,x_i \end{cases}$$

Решим эту систему уравнений методом Крамера:

$$\Delta = \begin{vmatrix} \sum_i x_i^4 & \sum_i x_i^{\,2}\,\mathrm{tg}\,x_i \\ \sum_i x_i^2\,\mathrm{tg}\,x_i & \sum_i \mathrm{tg}^2\,x_i \end{vmatrix}, \Delta b = \begin{vmatrix} \sum_i x_i^4 & \sum_i y_i x_i^2 \\ \sum_i x_i^2\,\mathrm{tg}\,x_i & \sum_i y_i\,\mathrm{tg}\,x_i \end{vmatrix},$$

$$\Delta a = \begin{vmatrix} \sum_i y_i x_i^2 & \sum_i x_i^{\,2}\,\mathrm{tg}\,x_i \\ \sum_i y_i\,\mathrm{tg}\,x_i & \sum_i \mathrm{tg}^2\,x_i \end{vmatrix}$$

$$a = \frac{\Delta a}{\Delta}, b = \frac{\Delta b}{\Delta}$$

$$a = \frac{\sum_i y_i x_i^2 \sum_i \mathrm{tg}^2\,x_i - \sum_i y_i\,\mathrm{tg}\,x_i \sum_i x_i^2\,\mathrm{tg}\,x_i}{\sum_i x_i^4 \sum_i \mathrm{tg}^2\,x_i - (\sum_i x_i^2 tgx_i)^2},$$

$$b = \frac{\sum_i x_i^4 \sum_i y_i\,\mathrm{tg}\,x_i - \sum_i x_i^2 y_i \sum_i x_i^2\,\mathrm{tg}\,x_i}{\sum_i x_i^4 \sum_i \mathrm{tg}^2\,x_i - (\sum_i x_i^2\,\mathrm{tg}\,x_i)^2}$$

Обозначения:

$$S1 = \sum_i x_i^{\,2} y_i,\ S2 = \sum_i x_i^{\,4},\ S3 = \sum_i x_i^2\,\mathrm{tg}\,x_i,\ S4 = \sum_i y_i\,\mathrm{tg}\,x_i,\ S5 = \sum_i \mathrm{tg}^2\,x_i$$

Имеем:

a = (S1*S5 – S4*S3) / (S2*S5 – S3^2)

B = (S2*S4 – S1*S3) / (S2*S5 – S3^2)

Задания

Получить систему линейных уравнений для определения параметров a и b, составить блок-схему и написать программу для определения a и b.

№	Исходные данные	Вид зависимости
1	X 1,5 2 2,5 3 3,5 4 4,5 У 4,3 5,2 6,5 7,9 9,7 11,9 14,6	$y' = a\,b^x$
2	X 0,15 0,25 0,3 0,4 0,5 0,6 У 5,02 5,00 4,96 4,91 4,81 4,7	$y' = a + b\cos x$
3	X 1 2 3 4 5 6 7 8 У 3,73 3,86 4,04 4,22 4,41 4,63 4,87 5,13	$y' = a + be^{0,1x}$
4	X 0,2 0,6 0,8 1,2 1,5 1,8 1,9 2,1 У 0,08 0,25 0,44 1,41 3,30 7,9 10,6 18,8	$y' = 10^{ax+b}$
5	X 1 2 3 4 5 6 7 У 3,6 7,5 14,0 23,0 34,7 49,0 66	$y' = a + bx^2$
6	X 1 2 3 4 5 6 7 8 9 У 9.5 6,8 6 5,5 5,3 5,2 4,9 4,8 4,7	$y' = a + \frac{b}{x}$
7	X 1 2 3 4 5 6 7 8 У 4,5 3,2 2,4 1,9 1,6 1,4 1,2 1,1	$y' = \frac{a}{b+x}$
8	X 1 2 3 4 5 6 7 У 1,2 1,8 2,1 2,4 2,6 2,7 2,9	$y' = a + b\cdot\ln x$
9	X 1 2 3 4 5 6 У 168,8 19,8 9,7 6,8 5,5 4,7	$y' = a\cdot 10^{b/x}$
10	X 0,1 0,15 0,2 0,25 0,3 0,35 0,4 0,45 У 2,5 2,6 2,73 2,81 2,9 3,0 3,1 3,2	$y' = a + b\sin x$
11	X 1,2 1,4 1,8 2,0 2,4 2,6 3,0 3,2 У 2,4 3,1 5,1 6,4 10,5 13,3 21,7 27,8	$y' = ae^{bx}$

Содержание

Лабораторная работа 1.
Построение регрессионных моделей 3

Лабораторная работа 2.
Обработка результатов эксперимента 15

Лабораторная работа 3.
Построение двухфакторного эксперимента с использованием модели 23

Лабораторная работа 4.
Построение двухфакторного эксперимента с использованием рототабельного центрально-композиционного плана 34

Лабораторная работа 5.
Применение полного факторного эксперимента при проведении исследований..... 44

Лабораторная работа 6.
Интерполяция 53

Лабораторная работа 7.
Интерполяция сплайнами 66

Лабораторная работа 8.
Аппроксимация функций..... 69

Максимова Алина Петровна

ЛАБОРАТОРНЫЙ ПРАКТИКУМ

по дисциплине

«Моделирование АСОИиУ»

для направления 09.03.01 «Информатика и вычислительная техника»
профиля
«Автоматизированные системы обработки информации и управления»

Часть 2
Математическая теория планирования эксперимента

Компьютерная верстка – Максимова А.П.

Подписано в печать . Формат 60x84/16.
Бумага писчая. Печать оперативная. Усл. печ.л.4 .
Тираж экз. Заказ № . Цена свободная

Федеральное государственное бюджетное образовательное учреждение
Высшего профессионального образования
«Московский автомобильно-дорожный
государственный технический университет (МАДИ)»
Волжский филиал