Министерство образования Республики Беларусь

Белорусский национальный технический университет

Факультет информационных технологий и робототехники

Кафедра «Высшая математика»

«МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ТЕХНИЧЕСКИХ ОБЪЕКТОВ И ПРОЦЕССОВ С ИСПОЛЬЗОВАНИЕМ КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ»

Электронное учебно-методическое пособие для магистрантов специальности 1-36 80 02 «Инновационные технологии в машиностроении»

Минск, 2020

Составители:

Г.И.Лебедева, О.Л.Зубко

Рецензент:

З.Н.Примичева – кандидат физ.-матем.наук, доцент БГУИР

В учебном пособии рассматриваются вопросы парного корреляционно-регрессионного анализа (линейная и нелинейная зависимость), многофакторного корреляционно-регрессионного анализа (построение моделей, анализ полученных моделей, оценка коэффициентов уравнений регрессии), теории игр.

Учебное пособие содержит основные математические методы, знание которых необходимо для магистрантов технических и экономических специальностей.

Вместе с теоретическими материалами в пособии рассматривается достаточное количество примеров решения задач с реальным содержанием. В конце каждого раздела приведены упражнения для самостоятельной работы.

Предназначено для магистрантов высших технических учебных заведений различных типов, преподавателей, аспирантов, бакалавров.

Белорусский национальный технический университет
Пр-т Независимости 65, г. Минск, Республика Беларусь
Тел.(017)292-80-75
E-mail: matematics@bntu.by
Регистрационный № БНТУ/ФИТР48-02.2020

Содержание

ПРЕДИСЛОВИЕ

Изменившиеся экономические условия в республике требуют совершенствования подготовки высококвалифицированных специалистов для народного хозяйства. Задачи, выдвигаемые ныне производством и практикой, предполагают наличие исследовательских навыков и творческого подхода к их решениям. В связи с этим важное место занимает владение будущими специалистами различными математическими методами для решения инженерных задач. В настоящее время ощущается недостаток в литературе, в которой компактно приводились бы математические методы, наиболее часто используемые при решении прикладных математических задач. В данном учебном пособии рассматриваются и излагаются: корреляционно-регрессионный анализ и теория игр.

Наряду с изложением теоретического материала приводится достаточное количество примеров решения задач с реальным содержанием. Большое внимание уделено практической направленности рассматриваемых математических методов. Пособие написано для читателей, которые уже изучили общий курс математики в техническом высшем учебном заведении и владеют необходимым математическим аппаратом для понимания излагаемых методов и их применения для решения прикладных задач. Изложение материала ведется на уровне, доступном широкому кругу читателей. В конце каждого раздела даны упражнения для самостоятельного решения. Упражнения нумеруются отдельно в каждой главе. Нумерация формул, рисунков и примеров приведена также по главам.

ГЛАВА 1. КОРРЕЛЯЦИОННО-РЕГРЕССИОННЫЙ АНАЛИЗ

1.1. Общие сведения

Корреляционно-регрессионный анализ широко используется при исследовании различных зависимостей между статистическими рядами. На автомобильном транспорте с помощью корреляционно-регрессионного анализа, например, можно исследовать зависимость:

а) подвижности населения от влияющих на нее факторов (численности населения, плотности маршрутной сети, среднего расстояния передвижения, типа подвижного состава и т.д.);

б) эксплуатационной скорости движения автобусов на маршруте от длины маршрута, средней длины перегона, количества остановок, состояния дорожной сети и т.д.;

в) себестоимости перевозок от эксплуатационных затрат, коэффициента использования пробега, коэффициента использования грузоподъемности, уровня механизации погрузочно-разгрузочных работ и т.д.;

г) уровня дорожно-транспортных происшествий от интенсивности движения транспортных средств, состояния проезжей части дороги, наличия пересечений и перекрестков, скорости движения транспортного потока, квалификации водителей и т.д.

И таких примеров из области автомобильного транспорта можно привести множество.

В отличие от функциональной, корреляционная зависимость не является строго определенной. Например, известно, что производительность подвижного состава возрастает с увеличением его грузоподъемности. Однако при одной и той же грузоподъемности мы можем получить различные значения производительности, так как на нее еще влияют дорожные условия, опыт водителя и т.д.

Тем не менее, общая закономерность увеличения производительности от грузоподъемности проявляется отчетливо, хотя и не строго.

В зависимости от количества рассматриваемых факторов корреляционно-регрессионный анализ подразделяется на парный и многофакторный.

Парный корреляционно-регрессионный анализ устанавливает связь между двумя (парой) факторами: $y_i = f(x_i)$, *многофакторный* – между n факторами, один из которых – зависимый, а остальные – независимые: $y_i = f(x_{i1}, x_{i2},\ldots,x_{in})$.

Корреляционный анализ используется для установления тесноты связей, а регрессионный – для установления формы связей (зависимостей). Вообще корреляционно-регрессионный анализ применяется не ко всем факторам, а только к коррелированным, т.е., к факторам, между которыми связь существует. Например, в процессе обследования режимов движения автомобилей получили два статистических ряда (значений эксплуатационной скорости $v_э$ и плотности потока автомобилей P):

$y = v_э$:	= 35,4,	36,2,	37,2,	38,9,	40,0;
$x=P$:	=2,8	2,82	2,75	2,69	2,5.

Заданные величины являются корреляционными, так как известно, что эксплуатационная скорость зависит от плотности потока автомобилей.

Нам дано пять пар:

35,4	36,2	37,2	38,9	40,0
↓	↓	↓	↓	↓
2,8	2,82	2,75	2,69	2,5

Если нанести полученные результаты на координатную плоскость, получим поле корреляции.

Эмпирическая линия регрессии на рис. 1.1 соответствует графику 1. Линия, к которой стремится эмпирическая линия регрессии при неограниченном числе наблюдений, называется *теоретической линией регрессии*. На рисунке она соответствует прямой II.

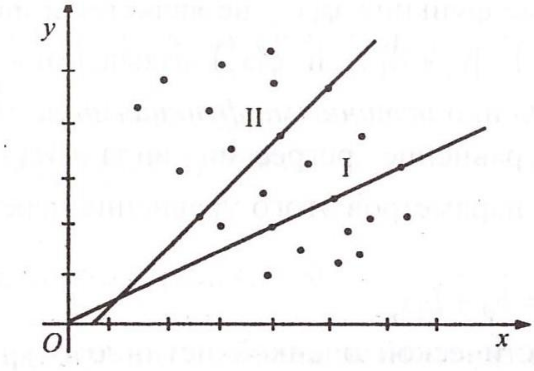

Рис. 1.1

1.2. Парный корреляционно-регрессионный анализ

Пусть проведен эксперимент, в результате которого зафиксировано n пар значений (x_i, y_i), $i = \overline{1,n}$.

Результаты измерений представляют собой выборку некоторой двумерной генеральной совокупности. Требуется установить зависимость между переменными x и y генеральной совокупности по результатам рассматриваемой выборки. При установлении зависимости следует исходить из двух моделей.

М о д е л ь 1. Пусть генеральная совокупность, образующая систему случайных величин (x, y) является нормальной. В этом случае условный дифференциальный закон составляющей y (при условии, что составляющая x приняла некоторое фиксированное значение x_i) является нормальным с математическим ожиданием $M(y/x_i) = \beta_0 + \beta_1 x_i$ и дисперсией σ^2.

Условное математическое ожидание зависимой переменной y

$$M(y/x_i) = \beta_0 + \beta_1 x_i \tag{1.1}$$

называется *функцией регрессии*. В целом, рассматривая двумерную нормальную совокупность, можно говорить о двух функциях регрессии:

- функции регрессии y на x: $M(y/x_i) = \beta_0 + \beta_1 x_i$;
- функции регрессии x на y: $M(x/y_i) = \alpha_0 + \alpha_1 x_i$.

М о д е л ь 2. Пусть закон распределения системы случайных величин x, y не является нормальным. Тогда математическое ожидание составляющей y при условии, что составляющая x приняла некоторое фиксированное значение x_i, является функцией $\varphi(x_i)$:

$$M(y/x_i) = \varphi(x_i).$$

В общем случае функция $\varphi(x_i)$ не является линейной. Функции регрессии $M(y/x_i) = \beta_0 + \beta_1 x_i$ и $\varphi(x_i)$ называются *функциями регрессии первого рода* или *истинными функциями регрессии*.

Рассмотрим уравнение регрессии вида $M(y/x_i) = \beta_0 + \beta_1 x_i$. Обозначив оценки параметров этого уравнения через b_0 и b_1, получим уравнение

$$y' = M(y/x_i) = b_0 + b_1 x_i, \tag{1.2}$$

являющееся статистической оценкой истинного уравнения регрессии или *уравнением регрессии*.

Линейная функция регрессии довольно часто встречается на практике. Уравнение (1.2) называется *эмпирическим*.

b_0 называется свободным членом уравнения. Он показывает, в какой точке линия регрессии пересекает ось Oy.

b_1 – коэффициент уравнения регрессии. Он показывает, на какую величину изменится переменная y при изменении x на единицу.

Коэффициенты b_0 и b_1 могут быть определены по формулам:

$$b_1 = \frac{\sum_{i=1}^{n} x_1 y_1 - n\overline{x}\,\overline{y}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2};$$

$$b_0 = \overline{y} - b_1\overline{x},$$

где $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$; $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

В общем случае коэффициенты b_0 и b_1 определяются либо методом наименьших квадратов, либо методом максимального правдоподобия.

Сущность ***метода наименьших квадратов*** заключается в том, что сумма квадратов отклонений расчетных значений от фактических есть величина минимальная:

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

Для регрессии $y' = b_0 + b_1 x_i$ имеем:

$$S = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2 \to \min .$$

Значит, имеется функция двух переменных (b_0 и b_1), стремящаяся к минимуму.

Действуя по стандартным правилам математического анализа, получим систему

$$\frac{\partial S}{\partial b_0} = -2\sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) = 0 ;$$

$$\frac{\partial S}{\partial b_1} = -2\sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)(x_i) = 0 .$$

Раскрывая скобки и производя элементарные преобразования, получаем систему двух линейных уравнений с двумя неизвестными

$$n b_0 + b_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i ;$$

$$b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} (x_i)^2 = \sum_{i=1}^{n} x_i y_i .$$

Решая данную систему, находим искомые коэффициенты уравнения регрессии.

Достоинством метода наименьших квадратов является то, что полученные оценки (коэффициенты b_i) являются эффективными, несмещенными и состоятельными.

Оценка называется *несмещенной*, если математическое ожидание оцениваемого параметра равно истинному значению

$M(b_i) = \beta_i$,

где β_i – истинное значение.

Оценка является *эффективной*, если

$M(b_i - \beta_i)^2 \to 0$.

Оценка b_i называется *состоятельной оценкой* параметра β_i, если при $n \to \infty$ $b_i \to \beta_i$ по вероятности:

$$\underset{n\to\infty}{b_i \xrightarrow{P} \beta_i}.$$

Метод максимального правдоподобия (для линейного уравнения регрессии) основан на использовании функции правдоподобия

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\delta_{y.x}} \exp\left[-\frac{1}{2}\frac{(y_i - b_0 - b_1 x_i)^2}{\delta_{y.x}^2}\right],$$

где $\delta_{y.x}$ – среднее квадратическое отклонение;

$$\delta_{y.x}^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - y')^2 \; ; \; y' = b_0 + b_1 x \, .$$

Оценки коэффициентов b_i и $\delta_{y.x}$ выбираются так, чтобы функция правдоподобия достигала своего максимального значения. Для определения коэффициентов берутся производные от L по b_i и $\delta_{y.x}$ и приравниваются к нулю. С целью упрощения вычислений удобнее брать производные от предварительно прологарифмированной функции правдоподобия:

$$\begin{cases} \dfrac{\partial \ln L}{\partial b_0} = -\dfrac{1}{\delta_{y.x}^2}\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2 = 0; \\ \dfrac{\partial \ln L}{\partial b_1} = -\dfrac{1}{\delta_{y.x}^2}\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)x_i = 0; \\ \dfrac{\partial \ln L}{\partial b_0} = -\dfrac{n}{\delta_{y.x}} + \dfrac{1}{\delta_{y.x}^3}\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2 = 0. \end{cases}$$

Решая полученную систему, находим:

$$b_0 = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2};$$

$$b_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2};$$

$$\delta_{y,x}^2 = \frac{\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2}{n}.$$

Оценки b_i, полученные по методу максимального правдоподобия и по методу наименьших квадратов, совпадают. Оценка $\delta_{y.x}^2$ имеет различие. В методе максимального правдоподобия оценка $\delta_{y.x}^2$ является смещенной.

Если закон распределения системы случайных величин не является нормальным, то функция регрессии будет нелинейной. К нелинейным функциям относятся:

$y = b_0 + b_1 \dfrac{1}{x}$ – гиперболическая;

$y = b_0 x^{b_1}$ – степенная;

$y = b_0 + b_1 \ln x$ – логарифмическая;

$y = b_0 + b_1 x + \dots + b_n x^n$ – параболическая и др.

1.3. Линейные уравнения регрессии

Линейными уравнениями регрессии называются уравнения вида

$$y = M(y/x) = b_0 + b_1 x . \qquad (1.3)$$

Это уравнения регрессии y на x. Для уравнений (1.3) расчет коэффициентов b_i производится либо с помощью метода наименьших квадратов, либо с помощью метода максимального правдоподобия.

По методу наименьших квадратов получаем

$$b_0 = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2};$$

$$b_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}.$$

После определения b_0 и b_1 оценивается теснота связи между рассматриваемыми y и x. Для этого определяется выборочный коэффициент корреляции

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left(n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right)\left(n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right)}}.$$

Коэффициент корреляции не зависит от выбора единиц измерения случайных величин. Значения данного коэффициента корреляции колеблются от –1 до +1. Если $|r| \le 1$, то между случайными величинами x и y существует линейная связь. Если $r = 0$, то между случайными величинами x и y линейная связь не существует. Если вычисленное значение коэффициента корреляции меньше 0,4–0,5, то целесообразно перейти к рассмотрению других зависимостей.

Полученное уравнение регрессии должно согласоваться с фактическими данными, полученными экспериментальным путем. Соответствие уравнения экспериментальным данным определяется с помощью *F-критерия* (*критерия Фишера*):

$$F=\frac{b_1^2(n-2)\sum\limits_{i=1}^{n}(x_i-\bar{x})^2}{\sum\limits_{i=1}^{n}(y_i-b_0-b_1x_i)^2}.$$

Расчетное значение F сравнивается с табличным F_{α},v_1,v_2, где F_{α}, v_1, v_2 – табличное значение (прил. 3); α – уровень значимости. Обычно принимается $\alpha=0{,}01; 0{,}10; 0{,}15; 0{,}20$, $v_1=P$, $v_2=n-P-1$; P – порядок рассматриваемой функции. Для линейной модели $P=1$. Если $F>F_{\alpha},v_1,v_2$, то полученное уравнение регрессии согласуется с данными эксперимента. В противном случае полученное уравнение непригодно для практического применения.

После построения уравнения регрессии, имеющего согласование с экспериментальными данными, строятся доверительные интервалы для коэффициентов уравнения и линии регрессии. Построение осуществляется исходя из заданной доверительной вероятности $(1-\alpha)$.

Доверительным интервалом называется интервал, который покрывает неизвестный параметр с заданной надежностью $(1-\alpha)$. Чем меньше длина доверительного интервала, тем точнее оценка. Если длина доверительного интервала велика, то оценка малопригодна для практики.

Доверительный интервал для коэффициентов b_i определяется по формуле

$$b_i \pm t_{\frac{\alpha}{2},n-2}\cdot S_{b_i},$$

где $t_{\frac{\alpha}{2},n-2}$ – статистика Стьюдента, выбираемая по таблице (прил. 4); S_{b_i} – среднеквадратическое отклонение оценки коэффициента b_i:

$$S_{b_0}=S_{y.x}\sqrt{\frac{\sum\limits_{i=1}^{n}x_i^2}{n\sum\limits_{i=1}^{n}x_i^2-\left(\sum\limits_{i=1}^{n}x_i\right)^2}};$$

$$S_{b_1}=S_{y.x}\sqrt{\frac{1}{\sum_{i=1}^{n}(x_i-\bar{x})^2}};$$

$$S_{y.x}=\sqrt{\left\{\sum_{i=1}^{n}(y_i-\bar{y})^2-\frac{\left[\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\right]^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right\}/(n-2)}\,.$$

Доверительная зона, накрывающая истинную регрессию в интервале [m, k], где m – нижняя граница значений x; k – верхняя граница значений x (независимой переменной), определяется по формуле:

$$Y_{\min}=b_0+b_1x-u_{n-2}(P,\lambda)S_{y.x}\sqrt{\frac{1}{n}+\frac{(x-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}};$$

$$Y_{\max}=b_0+b_1x+u_{n-2}(P,\lambda)S_{y.x}\sqrt{\frac{1}{n}+\frac{(x-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}},$$

где $P=(1-\alpha)$ – доверительная вероятность; λ – параметр, определяемый по формуле

$$\lambda=\sqrt{\frac{1}{2}\left[1-\frac{1+nCD^2}{\sqrt{\left(1+nC^2\right)\left(1+nD^2\right)}}\right]};$$

$$C=\frac{m-\bar{x}}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}};$$

$$D=\frac{k-\bar{x}}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}};$$

$u_{n-2}(P,\lambda)$ – критическое значение, определяемое по таблице (прил. 3).

В настоящее время достаточно хорошо разработано математическое обеспечение корреляционно-регрессионного анализа, и все вычисления производятся на компьютере. При этом в рассмотрение включается не одна модель, а группа моделей. По полученным расчетным характеристикам выбирается та зависимость, которая точнее отражает связь между исследуемыми величинами. Основными оценочными критериями являются: коэффициент корреляции, критерий Фишера, дисперсия и среднеквадратическое отклонение. Широкое применение корреляционно-регрессионного анализа в исследованиях связано с тем, что он имеет:

1) сравнительную простоту использования;

2) наличие хорошо разработанного математического аппарата и программ для компьютера;

3) малые затраты времени на получение частных решений при известных значениях независимой переменной;

4) возможность как отдельного, так и совместного исследования влияния различных факторов на исследуемую величину;

5) возможность использования полученных зависимостей (моделей) для прогнозирования.

Под *прогнозированием* понимается предсказание значения исследуемой зависимой переменной для последующего периода при известном значении влияющих на нее факторов.

Модели, полученные с помощью корреляционно-регрессионного анализа, имеют исследовательский характер и представляются в наглядном виде.

Пример 1.1. Исследовать зависимость прочности материала в зависимости от содержания легирующих компонентов (табл. 1.1):

Таблица 1.1

№ п/п	Прочность H_b,	Количество легирующих элементов, %
1	35,1	35,5
2	37,1	37,0
3	27,3	28,5
4	24,5	25,0
5	31,0	31,0
6	39,0	40,0
7	19,8	20,0
8	22,8	22,5
9	35,0	33,5
10	38,5	38,0

1. Определить параметры линейного уравнения регрессии.

2. Проверить тесноту связи между исследуемыми параметрами.

3. Проверить согласованность построенного уравнения регрессии с данными эксперимента.

4. Построить 90%-й доверительный интервал для коэффициентов линейного уравнения регрессии.

5. Построить 95%-ю доверительную зону для истинного уравнения регрессии.

Р е ш е н и е . 1. Определим параметры b_0 и b_1:

$$b_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\,\bar{y}}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2} = \frac{100748 - 10\cdot 31{,}1\cdot 31{,}01}{10100 - 10\cdot 31{,}1^2} = 1{,}00652;$$

$$\sum_{i=1}^{n} x_i y_i = 1007{,}48\ ; \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 31{,}1\ ; \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = 31{,}01\ ;$$

$$\sum_{i=1}^{n} x_i^2 = 10100;\quad b_0 = \bar{y} - b_1\bar{x} = 31{,}01 - 1{,}00652\cdot 31{,}1 = -0{,}292778.$$

Следовательно, эмпирическое уравнение регрессии имеет вид

$$y' = b_0 + b_1 x = -0{,}292778 + 1{,}00652x\,.$$

2. Коэффициент корреляции:

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left(n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right)\left(n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right)}} =$$

$$= \frac{10\cdot 1007{,}48 - 311\cdot 310{,}1}{\sqrt{(10\cdot 10100 - 311^2)(10\cdot 10055{,}09 - 310{,}1^2)}} = 0{,}99\ ;$$

$$\sum_{i=1}^{n} x_i = 311; \qquad \sum_{i=1}^{n} y_i^2 = 10055{,}09; \qquad \sum_{i=1}^{n} y_i = 310{,}1\,.$$

Полученное значение коэффициента корреляции свидетельствует о тесной линейной связи между исследуемыми параметрами

3. F-критерий:

$$F=\frac{b_1^2(n-2)\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(y_i-b_0-b_1x_i)^2}=\frac{1{,}00652^2\cdot 8\cdot 427{,}9}{\sum_{i=1}^{n}(y_i+0{,}292778-1{,}00652x_i)^2}=78{,}5\ .$$

$$\sum_{i=1}^{n}(x_i-\bar{x})^2=427{,}9\ .$$

Примем α = 0,10. По таблице значений *F*-критерия находим: $F_{0,1;1;8}$ = 3,46. Видим, что $F>F_{0,1;1;8}$. Следовательно, построенная модель хорошо согласуется с экспериментальными данными.

4. Для построения 90%-го доверительного интервала для коэффициентов истинного уравнения регрессии вычислим

$$S_{y.x}^2=\frac{1}{n-2}\left[\sum_{i=1}^{n}(y_i-\bar{y})^2-\frac{\left[\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\right]^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right]=$$

$$=\frac{1}{8}\left[438{,}889-\frac{430{,}69^2}{427{,}9}\right]=0{,}67385$$

$$\sum_{i=1}^{n}(y_i-\bar{y})^2=438{,}889;\ \sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=430{,}69;\ S_{y.x}=0{,}820884.$$

$$S_{b_0}=S_{y.x}\sqrt{\frac{\sum_{i=1}^{n}x_i^2}{n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2}}=0{,}820884\sqrt{\frac{10100}{10\cdot 10100-311^2}}=1{,}2116;$$

$$S_{b_1}=S_{y.x}\sqrt{\frac{1}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}=0{,}820884\sqrt{\frac{1}{427{,}9}}=0{,}0396836$$

По таблице Стьюдента для данной вероятности (1–α) = 0,90 и числу степеней свободы (*n*–2) находим:

$$t_{\frac{\alpha}{2},n-2}=t_{0,05;8}=1{,}860.$$

Следовательно, 90%-е доверительные интервалы равны:

$$b_0 \pm t_{\frac{\alpha}{2},n-2} \cdot S_{b_0} = -0{,}292778 \pm 2{,}345134 = [-2{,}63791;\ 2{,}05235];$$

$$b_1 \pm t_{\frac{\alpha}{2},n-2} \cdot S_{b_1} = 1{,}006522 \pm 1{,}860 \cdot 0{,}0396836 =$$

$$= 1{,}006522 \pm 0{,}073792 = [0{,}93273;\ 1{,}08031].$$

5. Определим 95%-ю доверительную зону, накрывающую истинную линию регрессии:

$$C = \frac{m - \bar{x}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{20 - 31{,}1}{\sqrt{427{,}9}} = -0{,}539;$$

$$D = \frac{k - \bar{x}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{40 - 31{,}1}{\sqrt{427{,}9}} = 0{,}432;$$

$$\lambda = \sqrt{\frac{1}{2}\left[1 - \frac{1 + nCD^2}{\sqrt{\left(1 + nC^2\right)\left(1 + nD^2\right)}}\right]} = \sqrt{\frac{1}{2}\left[1 - \frac{1 - 2{,}328}{\sqrt{(1 + 2{,}91)\,(1 + 1{,}87)}}\right]} = 0{,}87\,.$$

По таблице (прил. 3) находим:

$$u_{n-2}(P, \lambda) = u_8(0{,}95;\ 0{,}87) = 2{,}946.$$

Тогда

$$y_{\min} = 0{,}293 + 1{,}0065x - 0{,}104\sqrt{427{,}9 + 10(x - 31{,}1)^2};$$

$$y_{\max} = -0{,}293 + 1{,}0065x + 0{,}104\sqrt{427{,}9 + 10(x - 31{,}1)^2}.$$

1.4. Нелинейные уравнения регрессии

Уравнения регрессии, имеющие нелинейный вид, называются *нелинейными*. В настоящее время существует два основных метода подбора криволинейных эмпирических функций регрессии:

1) аппроксимацией эмпирических зависимостей функциями, допускающими линеаризацию (показательная, степенная, логарифмическая и др.);

2) аппроксимацией исследуемых зависимостей многочленами (параболами). В первом случае с помощью специальных преобразований мы получаем линейную модель.

Например, для гиперболической зависимости $y = b_0 + b_1 \frac{1}{x}$, заменив $\frac{1}{x}$ на t, получим $y = b_0 + b_1 t$ – линейная модель. Для показательной модели $y = b_0\, b_1^{\,x}$, прологарифмировав обе части, получим $\ln y = \ln b_0 + x \ln b_1$. Обозначив, $y^* = \ln y$, $b_0^* = \ln b_0$, $b_1^* = \ln b_1$ получим: $y^* = b_0^* + b_1^* x$ – линейная модель и т.д. Методика расчета линейных моделей рассмотрена в предыдущем параграфе.

Во втором случае рассматриваются параболы различного порядка:

$$y = b_0 + b_1 x + b_2 x^2 + \dots + b_p x^p.$$

Коэффициенты b_i находятся по методу наименьших квадратов, т.е. из системы нормальных уравнений

$$\frac{\partial S}{\partial b_0} = -2\sum_{i=1}^{n}\left(y_i - b_0 - b_1 x_i - b_2 x_i^2 - \dots - b_p x_i^p\right) = 0;$$

$$\frac{\partial S}{\partial b_1} = -2\sum_{i=1}^{n}\left(y_i - b_0 - b_1 x_i - b_2 x_i^2 - \dots - b_p x_i^p\right) x_i = 0;$$

$$\frac{\partial S}{\partial b_2} = -2\sum_{i=1}^{n}\left(y_i - b_0 - b_1 x_i - b_2 x_i^2 - \dots - b_p x_i^p\right) x_i^2 = 0;$$

$$\frac{\partial S}{\partial b_p} = -2\sum_{i=1}^{n}\left(y_i - b_0 - b_1 x_i - b_2 x_i^2 - \dots - b_p x_i^p\right) x_i^p = 0,$$

где $S = \sum_{i=1}^{n} \ell_i^2 = \sum_{i=1}^{n}\left(y_i - b_0 - b_1 x_i - b_2 x_i^2 - \dots - b_p x_i^p\right)^2 \to \min$.

Сделав элементарные преобразования, получим:

$$b_0 n + b_1 \sum_{i=1}^{n} x_i + b_2 \sum_{i=1}^{n} x_i^2 + \dots + b_2 \sum_{i=1}^{n} x_i^p = \sum_{i=1}^{n} y_i\ ;$$

$$b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 + b_2 \sum_{i=1}^{n} x_i^3 + \dots + b_p \sum_{i=1}^{n} x_i^{p+1} = \sum_{i=1}^{n} x_i y_i\ ;$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots$$

$$b_0 \sum_{i=1}^{n} x_i^p + b_1 \sum_{i=1}^{n} x_i^{p+1} + b_2 \sum_{i=1}^{n} x_i^{p+2} + \dots + b_p \sum_{i=1}^{n} x_i^{2p} = \sum_{i=1}^{n} x_i^p y_i\ .$$

Из указанной системы уравнений методом последовательного исключения находятся коэффициенты b_i, $i = \overline{0,P}$. Порядок параболы *p* устанавливается путем последовательного рассмотрения парабол, начиная с нулевого порядка. Процесс увеличения порядка параболы идет до тех пор, пока остаточная сумма квадратов не станет меньше 1 или среднеквадратическое отклонение $S_{y.x}$ не станет наименьшим:

$$S_{y.x} = \frac{\sum_{i=1}^{n}(y_i - b_0 x_i - b_1 x_i^2 - \cdots - b_p x_i^p)}{\sqrt{n-p-1}}.$$

При достаточно большом порядке параболы вычисления становятся громоздкими. В этом случае лучше воспользоваться матричной формой записи системы уравнений

$$(X^* X)B = X^* Y \ ,$$

где

$$(X^* X) = C = \begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \cdots & \sum_{i=1}^{n} x_i^p \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \cdots & \sum_{i=1}^{n} x_i^{p+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^{n} x_i^p & \sum_{i=1}^{n} x_i^{p+1} & \sum_{i=1}^{n} x_i^{p+2} & \cdots & \sum_{i=1}^{n} x_i^{2p} \end{bmatrix}$$

является симметричной квадратной матрицей размерности $(p+1)\times(p+1)$;

$$B = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{bmatrix}; \qquad (X^* Y) = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \vdots \\ \sum_{i=1}^{n} x_i^p y_i \end{bmatrix}.$$

Тогда

$$B = (X^* X)^{-1}(X^* Y) = C^{-1}(X^* Y).$$

Матрица C^{-1}, обратная матрице *C*, называется *дисперсионно-*

ковариационной. После определения коэффициентов b_i проверяется теснота криволинейной связи между y и x. Теснота криволинейной связи определяется по корреляционному отношению

$$\eta = \sqrt{1 - \frac{S_{y.x}^2}{S_y^2}},$$

где

$$S_{y.x} = \frac{1}{n-p-1}\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i - \cdots - b_p x_i^p)^2.$$

$$S_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2.$$

Чем ближе η к единице, тем теснее криволинейная связь между исследуемыми случайными величинами. Если $\eta = 0$, то между y и x корреляционная связь отсутствует.

Для проверки согласованности полученных зависимостей с данными эксперимента используется *статистика Стьюдента t*. Для этого вычисляем значение t:

$$t = \frac{\eta\sqrt{n-2}}{1-\eta^2}$$

и сравниваем его с табличным значением $t_{\alpha,\, n-p-1}$.

Если вычисленное значение

$$t < t_{\alpha,\, n-p-1},$$

где $t_{\alpha,\, n-p-1}$ – табличное значение статистики Стьюдента, то корреляционная связь между рассматриваемыми y и x отсутствует. В противном случае полученная модель является согласованной с данными эксперимента и может быть рекомендована для практического применения.

Согласованность построенной зависимости с данными эксперимента можно осуществлять и по *критерию Фишера*:

$$F = \frac{\sum_{i=1}^{n}(y' - \bar{y})^2 (n-p-1)}{\sum_{i=1}^{n}(y' - \bar{y})^2 \cdot p},$$

где p – порядок параболы.

С помощью данного критерия проверяется гипотеза

H_0: $\beta_0 = \cdots = \beta_p = 0$.

Если

$$F \geq F_{\alpha,\, p,\, n-p-1},$$

где $F_{\alpha,\, p,\, n-p-1}$ – критическое значение распределения Фишера, соответствующее уровню значимости α, порядку p и числу степеней свободы $n-p-1$, то нулевая гипотеза H_0: $\beta_0 = \cdots = \beta_p = 0$ отвергается, т.е. считается, что построенная парабола порядка P согласуется с данными эксперимента.

В противном случае считается, что парабола не согласуется с данными эксперимента.

Если построенное уравнение хорошо согласуется с данными эксперимента, переходим к следующему этапу – проверке значимости коэффициентов b_i. Значимость коэффициентов b_i проверяется с помощью статистики t':

$$t' = \frac{|b_i|}{S_{b_i}},$$

где S_{b_i} – среднеквадратическое отклонение для коэффициента b_i;

$$S_{b_i} = S_{y.x}\sqrt{c_{ii}^{-1}},$$

где c_{ii}^{-1} – элементы матрицы c^{-1}, стоящие на пересечении i-й строки и i-го столбца (диагональные элементы матрицы C^{-1}).

Если вычисленное значение $t' \geq t_{\alpha,\, n-p-1}$, взятое по таблице (прил. 4), то коэффициент b_i – значимый.

Если

$$t' < t_{\alpha,\, n-p-1},$$

то коэффициент b_i – несущественный и может быть исключен из уравнения регрессии.

Доверительные интервалы для коэффициентов b_i определяются следующим образом:

$$b_i \pm t_{\frac{\alpha}{2},\, n-p-1} \cdot S_{b_i},$$

где p – порядок параболы.

Для условного математического ожидания доверительный интервал равен

$$y' \pm t_{\frac{\alpha}{2},\, n-p-1} \cdot S_{y.x}\sqrt{1+\sum C_{ik}^{-1} x^i x^k}\ ,$$

где $y' = b_0 + b_1 x + b_2 x^2 + \cdots + b_p x^p$.

Пример 1.2. В результате проведенного исследования зависимости шероховатости поверхности от скорости резания, получили следующие данные (табл. 1.2).

Таблица 1.2

$v_э$	28,4	28,1	28,0	27,7	29,1	30	27,5	27,2	27,0
n_k	2	3	4	5	1	0	6	7	8
$v_э$	26,8	26,5	26,3	26,1	25,7	25,3	24,8	24,0	
n_k	9	10	11	12	13	14	15	16	

Требуется построить зависимость между исследуемыми параметрами $v_э$ от n_k.

Р е ш е н и е . Не имея никаких сведений о характере распределения случайных величин $(v_{э_i}, n_{k_i})$, аппроксимацию будем делать с помощью уравнений параболического типа.

1. Предположим, что связь между $y = v_э$ и $x = n_k$ отсутствует. В этом случае эмпирическое уравнение регрессии имеет вид параболы нулевого порядка.

А именно, $y' = b_0$; $b_0 = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{458}{17} = 26{,}97.$

Несмещенная оценка дисперсии переменной y будет равна:

$$S_x^2 = \frac{\sum\limits_i (y_i - y')^2}{n-1} = \frac{\sum\limits_i (y_i - 26{,}97)^2}{17-1} = 1{,}548.$$

Среднеквадратическое отклонение $S_{y.x} = \sqrt{1{,}548} = 1{,}244 > 1$.

2. Увеличим порядок параболы на единицу и рассмотрим параболу первого порядка:

$y' = b_0 + b_1 x$.

Система нормальных уравнений в этом случае имеет вид:

$$n b_0 + \sum_{i=1}^{n} x_i b_1 = \sum_{i=1}^{n} y_i;$$

$$b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i;$$

где $\sum_{i=1}^{n} x_i = 136$; $\sum_{i=1}^{n} y_i = 458{,}5$; $\sum_{i=1}^{n} x_i y_i = 3545{,}1$; $\sum_{i=1}^{n} x_i^2 = 1496$.

Подставляя значения сумм в рассматриваемую систему уравнений, получим:

$$\begin{cases} 17b_0 + 136b_1 = 458{,}5; \\ 136b_0 + 1496b_1 = 3545{,}1. \end{cases}$$

Решая данную систему, получим:

$b_0 = 29{,}38$; $b_1 = -0{,}3012$; $y' = 29{,}38 - 0{,}3012x$.

Остаточная сумма квадратов:

$$S = \sum_{i=1}^{n} (y_i - 29{,}38 + 0{,}3012x_i)^2 = 1{,}3349 > 1 .$$

Среднеквадратическое отклонение

$$S_{y.x} = \sqrt{\frac{\sum (y_i - y')^2}{n - p - 1}} = \sqrt{\frac{1{,}3349}{17 - 2}} = 0{,}299.$$

Посмотрим, можем ли мы уменьшить среднеквадратическое отклонение на единицу.

3. Рассмотрим для этого параболу второго порядка:

$y' = b_0 + b_1 x + b_2 x^2$.

Коэффициенты b_i находим из системы уравнений:

$$\begin{cases} n b_0 + b_1 \sum_{i=1}^{n} x_i + b_2 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i \; ; \\ b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 + b_2 \sum_{i=1}^{n} x_i^3 = \sum_{i=1}^{n} x_i y_i \; ; \\ b_0 \sum_{i=1}^{n} x_i^2 + b_1 \sum_{i=1}^{n} x_i^3 + b_2 \sum_{i=1}^{n} x_i^4 = \sum_{i=1}^{n} x_i^2 y_i \; . \end{cases}$$

где $\sum_{i=1}^{n} x_i^3 = 18496$; $\sum_{i=1}^{n} x_i^4 = 243848$; $\sum_{i=1}^{n} x_i^2 y_i = 38374{,}5$.

Подставляя значения сумм в рассматриваемую систему уравнений, получим:

$$\begin{cases} 17b_0 + 136b_1 + 1496b_2 = 458{,}5; \\ 136b_0 + 1496b_1 + 18496b_2 = 3545{,}1; \\ 1496b_0 + 18496b_1 + 243848b_2 = 38374{,}5. \end{cases}$$

Решая систему, находим:

$b_0 = 29{,}34;\ b_1 = -0{,}2866;\ b_2 = -0{,}000916;$

$y' = 29{,}34 - 0{,}2866x - 0{,}000916x^2.$

Остаточная сумма квадратов

$$S = \sum_{i=1}^{n}(y_i - b_0 - b_1x_i - b_2x_i^2)^2 = 1{,}2868 > 1\,.$$

Среднеквадратическое отклонение $S_{y.x} = \sqrt{\dfrac{1{,}2868}{17-3}} = 0{,}318$ – оно также не близко к нулю.

4. Повысим порядок параболы до третьей степени:

$y' = b_0 + b_1x + b_2x^2 + b_3x^3;$

$$\sum_{i=1}^{n}x_i^5 = 334776; \qquad \sum_{i=1}^{n}x_i^6 = 47261936; \qquad \sum_{i=1}^{n}x_i^3y_i = 469396{,}5.$$

Система нормальных уравнений для этой зависимости будет иметь вид:

$$\begin{cases} 17b_0 + 136b_1 + 1496b_2 + 18496b_3 = 458{,}5; \\ 136b_0 + 1496b_1 + 18496b_2 + 243848b_3 = 3545{,}1; \\ 1496b_0 + 18496b_1 + 243848b_2 + 334776b_3 = 38374{,}5; \\ 18496b_0 + 243848b_1 + 334776b_2 + 47261936b_3 = 469396{,}5. \end{cases}$$

Решая систему, находим, что

$b_0 = 29{,}34;\ b_1 = -0{,}7133;\ b_2 = 0{,}0678;\ b_3 = 0{,}0029;$

$y' = 29{,}34 - 0{,}7133x + 0{,}0678x^2 + 0{,}0029x^3.$

Остаточная сумма квадратов S = 0,4249 < 1.

Среднеквадратическое отклонение

$$S_{y.x} = \sqrt{\frac{0{,}4249}{17-4}} = 0{,}182.$$

Оно уже близкое к нулю.

На этом процесс увеличения порядка параболы прекратим и проведем дальнейшее исследование параболы третьего порядка.

5. Корреляционное отношение

$$\eta = \sqrt{1 - \frac{S_{y.x}^2}{S_y^2}} = \sqrt{1 - \frac{0{,}0331}{5{,}769}} = 0{,}997,$$

где $S_y = \dfrac{1}{16} \cdot 38{,}4423 = 2{,}402.$

Это свидетельствует о том, что у нас определена тесная связь между x и y.

6. Проверим согласованность полученной модели с экспериментальными данными. Для этого вычислим статистику t:

$$t = \frac{\eta\sqrt{n-2}}{1-\eta^2} = \frac{0{,}997\sqrt{15}}{1-0{,}997^2} = 6{,}53.$$

По таблице (прил. 2) находим $t_{табл}$=2,110.

Наше расчетное t > $t_{табл}$=2,11. Следовательно, полученное уравнение согласуется с данными эксперимента.

7. Проверим значимость коэффициентов b_i построенной параболы третьего порядка по формуле:

$$t' = \frac{|b_i|}{S_{b_i}}.$$

Для этого сначала определим среднеквадратические отклонения

$$S_{b_i} = S_{y.x}\sqrt{C_{ii}^{-1}}.$$

Матрица C для нашего примера имеет вид

$$C = \begin{bmatrix} 17 & 136 & 1496 & 18496 \\ 136 & 1496 & 18496 & 243848 \\ 1496 & 18496 & 243848 & 3347776 \\ 18496 & 243848 & 3347776 & 47261936 \end{bmatrix}.$$

Матрица C^{-1} будет равна:

$$C^{-1} = \begin{bmatrix} 0{,}624 & -0{,}282 & +0{,}034 & -0{,}0010 \\ -0{,}282 & +0{,}195 & -0{,}027 & -0{,}0010 \\ +0{,}034 & -0{,}027 & +0{,}004 & -0{,}0002 \\ -0{,}0010 & +0{,}001 & -0{,}0002 & +0{,}000007 \end{bmatrix}.$$

Следовательно,

$$S_{b_0} = 0{,}182\sqrt{0{,}624} = 0{,}144;$$

$$S_{b_1} = 0{,}182\sqrt{0{,}195} = 0{,}08;$$

$$S_{b_2} = 0{,}182\sqrt{0{,}004} = 0{,}0012;$$

$$S_{b_3} = 0{,}182\sqrt{0{,}000007} = 0{,}0005.$$

Теперь найдем t':

$$t'_0 = \frac{|29{,}821|}{0{,}144} = 207{,}08;$$

$$t'_1 = \frac{|-0{,}7133|}{0{,}08} = 8{,}916;$$

$$t'_2 = \frac{|0{,}0678|}{0{,}0012} = 56{,}5;$$

$$t'_3 = \frac{|0{,}0029|}{0{,}0005} = 5{,}8.$$

Примем доверительную вероятность (1 – α) = 0,95. По таблице распределения Стьюдента (прил. 4) находим:

$$t^{(0)}_{0,05;16} = 1{,}746;$$

$$t^{(1)}_{0,05;15} = 1{,}753;$$

$$t^{(2)}_{0,05;14} = 1{,}761;$$

$$t^{(3)}_{0,05;13} = 1{,}771.$$

Так как вычисленные значения t' больше табличных:

$$t'_0 = 207{,}08 > t^{(0)}_{0,05;16} = 1{,}746;$$

$$t'_1 = 8{,}916 > t^{(1)}_{0,05;15} = 1{,}753;$$

$$t'_2 = 56{,}5 > t^{(2)}_{0,05;14} = 1{,}761;$$

$$t'_3 = 5{,}8 > t^{(3)}_{0,05;13} = 1{,}771,$$

то все полученные коэффициенты b_i являются значимыми.

8. Найдем доверительные интервалы для коэффициентов b_i: $b_i \pm t_{\frac{\alpha}{2}, n-p-1} \cdot S_{b_i}$.

Табличное значение $t_{\frac{\alpha}{2}, n-p-1} = t_{0,025;\,17-3-1} = 2{,}16$.

Доверительные интервалы для коэффициентов будут равны:

$$\left[b_0 \pm t_{\frac{\alpha}{2},\, n-p-1} \cdot S_{b_0} \right] = [29{,}82 \pm 2{,}16 \cdot 0{,}144] = [29{,}502; 30{,}131];$$

$$\left[b_1 \pm t_{\frac{\alpha}{2},\, n-p-1} \cdot S_{b_1} \right] = [-0{,}7133 \pm 2{,}16 \cdot 0{,}08] = [-0{,}886; -0{,}540];$$

$$\left[b_2 \pm t_{\frac{\alpha}{2},\, n-p-1} \cdot S_{b_2} \right] = [0{,}0678 \pm 2{,}16 \cdot 0{,}0012] = [0{,}065; 0{,}070];$$

$$\left[b_3 \pm t_{\frac{\alpha}{2},\, n-p-1} \cdot S_{b_3} \right] = [0{,}0029 \pm 2{,}16 \cdot 0{,}0005] = [-0{,}0038; 0{,}0018].$$

На этом исследование модели закончено.

1.5. Многофакторный корреляционно-регрессионный анализ

Многофакторный корреляционно-регрессионный анализ используется для установления одновременной зависимости между y и n факторами $x_1, \ldots, x_n$, влияющими на исследуемую случайную величину y:

$$y = f(x_1, x_2, \ldots, x_n),$$

где

$$y = (y_1, y_2, \ldots, y_m);$$

$$x_1 = (x_{11}, x_{12}, \ldots, x_{1m});$$

$$x_2 = (x_{21}, x_{22}, \ldots, x_{2m});$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$x_n = (x_{n1}, x_{n2}, \ldots, x_{nm}).$$

Например, метод применяют для установления зависимости подвижности населения (y) от влияющих на нее факторов: численности населения (x_1), плотности маршрутной сети (x_2), плотности транспортных средств общественного пассажирского транспорта (x_3) и т.д.

Для данного случая зависимость в общем виде будет иметь вид

$$y = f(x_1, x_2, x_3).$$

Многофакторные модели широко используются при анализе и планировании работы автомобильного транспорта.

В настоящее время многофакторные корреляционные модели разработаны для:

1) планирования технико-эксплуатационных показателей работы подвижного состава (по транспортным управлениям Министерства автомобильного транспорта РФ);

2) планирования фондоотдачи по транспортным управлениям (по Министерству автомобильного транспорта РФ);

3) оценки экономических показателей деятельности грузовых автопредприятий (по ряду областей страны);

4) определения подвижности сельского населения на внутрирайонных, внутрихозяйственных и школьных связях (Республика Беларусь) и т.д.

Разработанные корреляционные модели подлежат к использованию только на тех объектах, для которых они разработаны, так как разработка моделей осуществлялась по конкретным данным исследуемых объектов. Как и в случае парных зависимостей, следует различать: линейные многофакторные модели

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_p x_p;$$

нелинейные: гиперболическую

$$y = b_0 + b_1 \frac{1}{x_1} + b_2 \frac{1}{x_2} + \cdots + b_p \frac{1}{x_p};$$

показательную

$$y = b_0 \cdot b_1^{x_1} \cdot b_2^{x_2} \cdots b_p^{x_p};$$

степенную

$$y = b_0 x_1^{b_1} \cdot x_2^{b_2} \cdots x_p^{b_p}$$ и др.

При рассмотрении нелинейных моделей чаще всего их сводят к линейному виду (как и в случае парных зависимостей).

Создание многофакторных моделей требует конкретизации включаемых в модель факторов. Поэтому корреляционно-регрессионному анализу предшествует всесторонний теоретический анализ возможности существования связи между исследуемыми явлениями – факторный анализ.

Следует выделить две стадии факторного анализа: качественную и количественную. На стадии качественного факторного анализа отбираются наиболее важные, на взгляд исследователя, факторы, качественно связанные с исследуемой проблемой, численные значения которых можно определить. При этом никаких ограничений в выборе факторов не существует. Допускается даже их дублирование на этом этапе.

На втором этапе (стадии количественного анализа) отбираются факторы, влияние которых на исследуемую проблему существенно. В уравнении множественной регрессии существенными обычно оказываются те факторы, которые имеют существенную корреляционную связь с результативным признаком (хороший коэффициент корреляции), а между собой – несущественную (значение коэффициента корреляции малое). На данном этапе рассчитывается корреляционная матрица (табл. 1.3).

Таблица 1.3

	y	x_1	x_2	$\dots$	x_n
y	1	r_{yx_1}	r_{yx_2}	$\dots$	r_{yx_n}
x_1		1	$r_{x_1x_2}$	$\dots$	$r_{x_1x_n}$
x_2			1	$\dots$	$r_{x_2x_n}$
x_3				$\dots$	$r_{x_3x_n}$
$\vdots$					$\vdots$
x_n					1

В табл. 1.3 $r_{x_ix_j}$ – коэффициент корреляции между факторами x_i и x_j:

$$r_{x_ix_j}=\frac{\sum\limits_{k=1}^{n}\left[(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j)\right]}{\sqrt{\sum\limits_{k=1}^{n}(x_{ik}-\bar{x}_i)^2(x_{jk}-\bar{x}_j)^2}}.$$

Матрица является симметрической. Поэтому в процессе исследования требуется заполнить или верхнюю, или нижнюю часть таблицы.

Покажем на примере, как по корреляционной матрице отбираются существенные факторы.

Предположим, что в процессе исследования получили следующую корреляционную матрицу (табл. 1.4):

Таблица 1.4

	y	x_1	x_2	x_3	x_4
y	1	0,5	0,3	0,7	0,8
x_1		1	0,4	0,6	0,5

x_2			1	0,4	0,2
x_3				1	0,3
x_4					1

У нас один y и четыре независимых фактора x_i. Из таблицы видим, что коэффициент корреляции между y и x_1 равен 0,5; между y и x_2 – 0,3; между y и x_3 – 0,7; между y и x_4 – 0,8. Для фактора x_2 (рассматриваем по столбцам) связь с фактором x_1 сильнее, чем с y ($r_{x_1x_2} = 0{,}4 > r_{yx_2} = 0{,}3$). Следовательно, фактор x_2 является несущественным и его нужно исключить из дальнейшего рассмотрения. Для фактора x_3 связь с x_1 меньше, чем с y ($r_{x_1x_3} = 0{,}60 < r_{yx_3} = 0{,}7$). Следовательно, фактор x_3 – существенный. Аналогично для фактора x_4: связь x_4 с y теснее, чем с x_1 и x_3 ($r_{yx_4} = 0{,}8 > r_{x_1x_4}$ и $r_{yx_4} = 0{,}8 > r_{x_3x_4} = 0{,}3$). Таким образом, окончательно получаем, что существенными факторами для дальнейшего исследования являются: x_1, x_3 и x_4.

После отбора факторов осуществляется построение многофакторных моделей. Рациональным является рассмотрение ряда моделей, из которых затем с помощью специальных критериев оценки выбирается наиболее точная.

Коэффициенты b_i эмпирического уравнения регрессии определяются с помощью метода наименьших квадратов или максимального правдоподобия.

Наиболее удобным является метод наименьших квадратов. Согласно ему коэффициенты b_i, $i = \overline{0,P}$, в эмпирическом уравнении регрессии находятся при условии, что

$$S = \sum_{j=1}^{n} \ell_j = \sum_{i=1}^{n} (y_j - b_0 - b_1x_1 - b_2x_2 \cdots - b_px_p)^2 \to \min \cdot$$

Взяв $\dfrac{\partial S}{\partial b_i}$ и приравняв их нулю, после преобразования получим систему уравнений:

$$b_0 n + b_1\sum x_1 + b_2\sum x_2 + \dots + b_p\sum x_p = \sum y;$$
$$b_0\sum x_1 + b_1\sum x_1^2 + b_2\sum x_1 x_2 + \dots + b_p\sum x_1 x_p = \sum x_1 y;$$
$$b_0\sum x_2 + b_1\sum x_2 x_1 + b_2\sum x_2^2 + \dots + b_p\sum x_2 x_p = \sum x_2 y;$$
$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots$$
$$b_0\sum x_p + b_1\sum x_p x_1 + b_2\sum x_p x_2 + \dots + b_p\sum x_p^2 = \sum x_p y.$$

Решая данную систему, находим значения неизвестных параметров $b_0, b_1, \dots, b_p$.

Теснота связей между *y* и исследуемыми факторами x_i, $i = \overline{1,P}$, оценивается с помощью коэффициента множественной корреляции:

$$R = \sqrt{1 - \frac{\sum_{j=1}^{n} \ell_j^2}{\sum_{j=1}^{n} (y_j - \bar{y})^2}},$$

где $\ell_j = y_j - y'_x$; $y' = b_0 + b_1 x_1 + \dots + b_p x_p$; y_j – фактические данные; $\bar{y} = \frac{1}{n}\sum_{j=1}^{n} y_j$.

Значения коэффициента множественной корреляции колеблются от 0 до +1. Чем ближе значения *R* к единице, тем теснее связь между *y* и влияющими на него факторами.

Согласованность полученных моделей с данными эксперимента проверяется по *F*-критерию:

$$F = \frac{R^2(n - p - 1)}{(1 - R^2)\cdot p},$$

где *p* – число независимых переменных x_i.

Вычисленное значение *F* сравнивается с табличным (прил. 3) F_{α, v_1, v_2}, где $v_1 = p$ и $v_2 = n - p - 1$; α – принятый уровень значимости. Если вычисленное значение $F \ge F_{\alpha, v_1, v_2}$ – модель согласуется с фактическими данными.

Если вычисленное значение $F < F_{\alpha, v_1, v_2}$ – модель не согласуется с фактическими данными.

Значимость коэффициентов эмпирического уравнения регрессии проверяется по статистике:

$$t = \frac{|b_i|}{S_{b_i}};$$

$$S_{b_i} = S_\ell \sqrt{C_{ii}^{-1}};$$

$$S_\ell = \frac{\sqrt{\sum_{j=1}^{n} \ell_j^2}}{n - p - 1},$$

где C_{ii}^{-1} – элементы матрицы C^{-1}, стоящие на пересечении i-й строки и i-го столбца (диагональные).

Если

$$t \ge t_{\alpha, n-p-1},$$

где $t_{\alpha, n-p-1}$ – табличное значение, то коэффициент b_i является значимым.

Если

$$t < t_{\alpha, n-p-1},$$

то коэффициент b_i с переменной x_i могут быть исключены из уравнения регрессии.

Интервальные оценки коэффициентов уравнения регрессии определяются следующим образом:

$$b_i \pm t_{\frac{\alpha}{2}, n-p-1} \cdot S_{b_i}.$$

Упражнения

1. Составить парные уравнения регрессии (табл. 1.5)и оценить их. В рассмотрение включить зависимости:

$$y = b_0 + b_1 x, \quad y = b_0 + b_1 \frac{1}{x}, \quad y = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots$$

Таблица 1.5

	Прочность материала	Значения факторов по вариантам							
		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	365	3,5	1805	18,05	50	27,7	0,61	4,0	3,9
2	400	2,9	537	44,7	36,3	31,1	1,01	3,0	3,8
3	730	4,0	1230	87,8	29,1	45,5	0,87	4,0	3,8
4	702	3,0	804	80,4	50	30	0,71	3,0	3,5
5	356	3,2	345	31,3	60	29	0,6	3,0	3,5
6	980	4,3	910	53,5	54,9	39	0,7	4,0	3,6
7	780	4,0	456	57	26,1	50,4	0,91	4,0	3,6
8	240	6,0	1470	91,8	22,8	22,4	0,83	5,6	4,5
9	250	5,3	1800	71,0	28,8	35,3	0,65	5,0	4,5
10	740	3,9	3083	82,0	12	55	0,6	3,6	2,8
11	200	4,5	1620	75,0	24	36	1,01	4,0	3,6
12	315	3,2	1600	160	50	37,5	0,67	3,3	3,5
13	262	4,0	1200	157	21,6	75,0	0,5	4,0	4,0
14	780	4,0	1690	89	29	37,7	0,46	4,0	4,0
15	300	4,2	800	70	24	25,2	0,6	4,1	4,1
16	648	2,1	3970	209	66,5	23,6	0,75	2,0	2,0
17	816	3,5	1360	97	39,3	44,1	0,8	3,0	2,5
18	484	3,9	1000	100	20	40	0,53	3,5	3,0
19	390	5,3	985	164	12,5	40,6	0,87	5,0	4,5
20	995	4,5	1500	190	20,5	32	0,7	4,0	4,0
21	900	4,6	1740	250	18,4	40	0,6	4,0	4,0

ГЛАВА 2. ТЕОРИЯ ИГР

2.1. Общие сведения о теории игр

Теория игр представляет собой математическую теорию конфликтных ситуаций. Ситуация называется *конфликтной*, если в ней сталкиваются интересы двух и более сторон. Приведем примеры конфликтных ситуаций.

1. Боевые действия. Есть нападающая и есть обороняющаяся стороны. И каждая желает выиграть.

2. Требования к количеству и вместимости подвижного состава со стороны пассажиров и автотранспортных предприятий. С точки зрения пассажиров необходимо, чтобы на маршрутах было как можно больше автобусов и как можно большей вместимости. АТП заинтересованы в том, чтобы перевозки осуществлять как можно меньшим количеством автобусов и меньшей вместимости с тем, чтобы затраты были наименьшими.

3. Требования к скорости движения автомобилей со стороны водителя и работника ГАИ. Водитель заинтересован в повышении скорости движения, а работник ГАИ – в снижении с целью избежания дорожно-транспортных происшествий.

Цель теории игр – выработка рекомендаций по разумному поведению участников конфликтов. При этом делается предположение о полной разумности противников. Одна реализация игры называется *партией*. По количеству участников игры делятся на *парные* (когда 2 участника) и множественные (число участников равно 3 и более). Во множественных играх участники могут образовывать *коалиции*.

Развитие игры во времени представляется рядом последовательных ходов. *Ход* – это выбор участником действия и его осуществление. Следует различать личные и случайные ходы. *Личный ход* – когда участник игры сам выбирает и осуществляет какой-либо вариант действия. *Случайный ход* – когда действие осуществляется с помощью какого-либо механизма случайного выбора (бросание монеты, вынимание карты из колоды и т.д.).

Теория игр распространяется только на игры, в которых есть личные ходы. Игры с личными ходами называются *стратегическими*. *Стратегией* игрока называется совокупность правил, определяющих выбор варианта действия при каждом личном ходе в зависимости от сложившейся ситуации.

В зависимости от числа стратегий игры делятся на конечные и бесконечные. *Конечная игра* – это игра, имеющая конечное число стратегий. Стратегия называется *оптимальной*, если она обеспечивает игроку наилучшее положение в данной игре.

По количеству информации игры делятся на игры с полной и неполной информацией. Игра называется игрой *с нулевой суммой*, если сумма выигрышей всех игроков равна нулю (каждый игрок выигрывает только за счет других). Существуют различные способы задания игры: матричные; аналитические и др. Каждый из способов имеет свои особенности.

2.2. Матричные игры

Матричной игрой называется игра, разыгрываемая с помощью специальной матрицы. Рассмотрим игру двух игроков (A и B) с нулевой суммой. Интересы игроков являются прямо противоположными: выигрыш игрока A равен проигрышу игрока B и наоборот. Если игрок A желает достичь максимума выигрыша в игре, то игрок B – минимума проигрыша. Число стратегий у игроков является конечным. Пусть у игрока A имеется m возможных стратегий: $A_1, A_2, \dots, A_m$, а у игрока B – n стратегий: $B_1, B_2, \dots, B_n$. Обозначим a_{ij} – выигрыш игрока A при применении им стратегии A_i, а игроком B – стратегии B_j. Если выигрыш a_{ij} известен для всех возможных пар стратегий A_i, B_j, то игру можно свести к матричной. Для этого составляется табл.2.1. Эта таблица еще называется *платежной матрицей*.

Таблица 2.1

A_i \ B_j	B_1	B_2	…	B_n
A_1	a_{11}	a_{12}	…	a_{1n}
A_2	a_{21}	a_{22}	…	a_{2n}
A_3	a_{31}	a_{32}	…	a_{3n}
…	…	…	…	…
A_m	a_{m1}	a_{m2}	…	a_{mn}

Тем самым игра сводится к одноходовой: от каждого игрока требуется сделать только один ход – выбрать стратегию.

Решение матричных игр, как правило, начинается с упрощения исходной матрицы. Для этого из таблицы исключаются все дублирующие и доминирующие стратегии – столбцы и строки. Строка A_k (столбец B_k) называется *дублирующей* строку A_l (столбец B_l), если все элементы строки A_k (столбца B_k) равны соответствующим элементам строки A_l (столбца B_l).

Например, в матрице (табл. 2.2) строка A_1 является дублирующей строку A_2, а столбец B_1 – дублирующим столбец B_2.

Таблица 2.2

A_i \ B_j	B_1	B_2	B_3	B_4	B_5
A_1	8	8	7	5	4
A_2	8	8	7	5	4
A_3	8	8	3	3	8

Поэтому из дальнейшего рассмотрения следует исключить строку A_1 и столбец B_1. Матрица при этом упростится (табл. 2.3).

Таблица 2.3

A_i \ B_j	B_2	B_3	B_4	B_5
A_2	8	7	5	4
A_3	8	3	3	8

Стратегия A_i игрока *А* называется *доминирующей* над стратегией A_k, если выигрыши в строке A_i больше либо равны соответствующим выигрышам строки A_k. Обязательным условием является наличие хотя бы одного элемента в строке A_i, который больше соответствующего элемента в строке A_k. Стратегия B_j называется доминирующей над стратегией B_m, если в столбце B_j стоят выигрыши, не превосходящие соответствующие выигрыши столбца B_m, и один из них обязательно меньше, чем соответствующий выигрыш в столбце B_m.

Например, пусть нам дана парная игра со следующей платежной матрицей (табл. 2.4) Требуется упростить матрицу.

Таблица 2.4

A_i \ B_j	B_1	B_2	B_3	B_4	B_5	B_6
A_1	3	5	6	7	9	8
A_2	2	4	6	7	9	8
A_3	3	5	6	7	8	8
A_4	3	6	1	2	4	3
A_5	3	5	6	7	8	8

В рассматриваемой матрице стратегия A_5 дублирует стратегию A_3, поэтому ее можно исключить из дальнейшего рассмотрения. В строке A_1 все элементы больше или равны соответствующим элементам строки A_2. Следовательно, стратегия A_1 является доминирующей над стратегией A_2. Поэтому ее исключаем из таблицы. Больше доминирующих строк не осталось. Рассмотрим теперь столбцы. Столбец B_1 является доминирующим над столбцами B_2, B_5 и B_6. Следовательно, мы можем исключить его из таблицы. Столбец B_3 является доминирующим над столбцом B_4, а столбец B_4 является доминирующим над столбцами B_5 и B_6. Исключая столбцы B_3 и B_4, получим табл. 2.5.

Таблица 2.5

A_i \ B_j	B_2	B_5	B_6
A_2	4	9	8
A_3	5	8	8
A_4	6	4	3

В полученной матрице стратегия B_6 является доминирующей над стратегией B_5 . Поэтому исключаем B_6 из рассмотрения. Наша таблица примет следующий вид (табл. 2.6).

Таблица 2.6

A_i \ B_j	B_2	B_5
A_2	4	9
A_3	5	8
A_4	6	4

Из полученной табл. 2.6 видим, что в ней нет больше ни дублирующих, ни доминирующих строк и столбцов. Следовательно, на этом процесс упрощения платежной матрицы прекращается.

После упрощения табл. 2.6 начинается решение игровой задачи. Основным принципом, которым руководствуется теория игр, является принцип «минимакса». Для игрока *A* этот принцип означает, что надо выбрать ту стратегию, при которой его минимальный выигрыш максимален. Очевидно, противник (*B*) должен выбирать ту стратегию, при которой его максимальный проигрыш будет минимальным.

Пусть теперь нам задана матричная игра двух лиц (табл. 2.7).

Таблица 2.7

A_i \ B_j	B_1	B_2	B_3	B_4	B_5	α_i
A_1	3	4	6	2	3	2
A_2	1	8	4	3	4	1
A_3	10	3	1	7	6	1
A_4	5	5	4	6	8	4
β_i	10	8	6	7	8	

Поставим в правом добавочном столбце (α_i) минимальные значения выигрышей игрока *A*. В нижней добавочной строке (β_i) поставим максимальные значения проигрышей игрока *B*. Среди элементов α_i выбирается максимальный (в данном случае $\alpha_4 = 4$).

Этот элемент определяет гарантированный выигрыш игрока A при применении им стратегии A_4. Меньше игрок A получить не может. Этот выигрыш называется *нижней ценой игры*. Затем среди элементов строки β_i выбирается минимальный. В данном случае $\beta_i = 6$ соответствует стратегии B_3. Этот элемент определяет проигрыш, больше которого игрок B не отдаст. Этот проигрыш является гарантированным и называется *верхней ценой игры*. Следовательно, в приведенной игре игрок A должен сначала выбрать стратегию A_4, игрок B – B_3. Затем игрок A должен применить стратегию A_1, игрок B – B_4 и т.д.

Если в матричной игре элемент $a_{ij} = \alpha_i = \beta_j$, то точка (A_i, B_j) называется *седловой*. Любая игра с полной информацией имеет седловую точку.

Седловая точка определяет оптимальные стратегии. Поэтому прежде чем решать задачу теории игр, надо проверить, имеется ли седловая точка. Если седловой точки нет, то руководствуются принципом «минимакса».

Следует отметить, что в платежной матрице может быть не одна, а несколько седловых точек. Например, в матрице (табл.2.8) имеется 6 седловых точек, для которых $\alpha_i = \beta_j = 2$.

Таблица 2.8

A_i \ B_j	B_1	B_2	B_3	B_4	α_i
A_1	2	2	3	2	2
A_2	2	2	3	3	2
A_3	2	2	2	2	2
β_i	2	2	3	3	

Им соответствуют следующие пары оптимальных стратегий: (A_1, B_1), (A_1, B_2), (A_2, B_1), (A_2, B_2), (A_3, B_1), (A_3, B_2).

Все седловые точки дают одно и то же значение выигрыша.

2.3. Геометрическое решение задач теории игр

Часто для решения задач теории игр используется графический метод решения. По сравнению с численными методами он дает наглядное представление решения задачи, позволяет изобразить как верхнюю, так и нижнюю цену игры.

Если число стратегий одного из игроков равно 2, то графическое решение не вызывает затруднений. При большем числе стратегий у обоих игроков графическое решение становится затруднительным. В этом случае переходят к численным методам решения.

Рассмотрим примеры графического решения задач теории игр.

Пример 2.1. Пусть нам дана парная игра игроков *A* и *B*, заданная матрицей (2 ×2) (табл. 2.9):

Таблица 2.9

	B_1	B_2
A_1	a_{11}	a_{12}
A_2	a_{21}	a_{22}

Каждый игрок имеет две стратегии. Игрок *A* имеет выигрыши a_{ij}, которые для игрока *B* являются проигрышами.

Р е ш е н и е . Геометрическое решение заданной задачи осуществляется следующим образом (рассматриваем для игрока *A*).

1. На оси *Ox* откладываем единичный отрезок *OC* (рис. 2.1). Через точки *O* и *C* проводим прямые I-I и II-II, перпендикулярные оси *Ox*. На прямой I-I будем откладывать выигрыши при стратегии A_1, а на прямой II-II – при стратегии A_2.

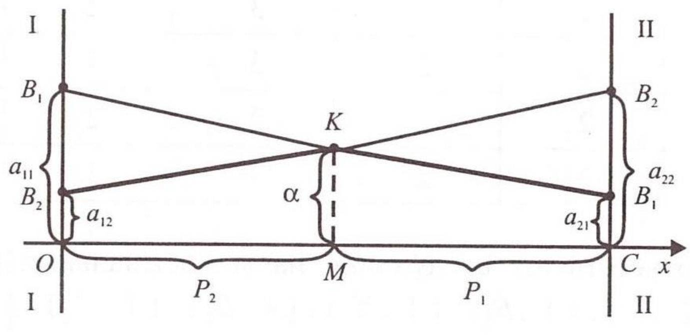

Рис. 2.1

2. Предположим, что противник *B* применил стратегию B_1. Тогда выигрыши игрока *A* при стратегиях A_1 и A_2 будут равны a_{11} и a_{21} соответственно. Откладываем их на соответствующих прямых:

выигрыш α_{11} на прямой I-I, выигрыш α_{21} – на прямой II-II.

3. Полученные точки обозначим в соответствии с принятой игроком *B* стратегией B_1 и соединим их отрезком.

4. Аналогичным образом поступаем со стратегией B_2 игрока *B*: на прямой I-I откладываем значение a_{12}, а на прямой II-II – значение a_{22}. Полученные точки обозначаем B_2 и соединяем между собой отрезком.

Наши прямые пересеклись. Точку пересечения обозначим буквой *K*. В соответствии с принципом «минимакса» оптимальная стратегия игрока *A* будет соответствовать точке, в которой минимальный выигрыш игрока *A* является максимальным. Минимальный выигрыш игрока *A* определяется нижней ценой игры. Для нашей задачи нижней ценой игры является ломаная B_2KB_1. Самой верхней точкой на ней является точка *K*. Следовательно, эта точка и определяет решение и цену игры. Ценой игры является ордината найденной точки *K*. Решение определяют вероятности P_i, на которые проекция точки *K* на ось *Ox* делит отрезок *OC*. Причем вероятность P_1 соответствует правому отрезку (*MC*), а вероятность P_2 – левому (*OM*). Оптимальная стратегия игрока *A* равна $P=(P_1,P_2)$. Мы рассмотрели решение задачи для игрока *A*.

При решении задачи для игрока *B* исходную матрицу следует рассматривать по строкам. Построения осуществляются аналогично рассмотренному методу. Оптимальная стратегия в этом случае будет определяться по минимуму верхней границы (рис. 2.2).

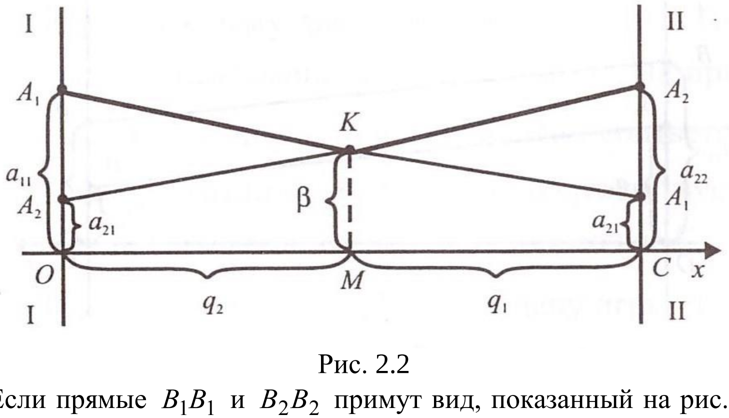

Рис. 2.2

Если прямые B_1B_1 и B_2B_2 примут вид, показанный на рис. 2.3, то максимальной точкой, лежащей на ломаной B_1KB_2, соответствующей нижней цене игры, является точка B_2. В этом случае оптимальная стратегия игрока A определяется только стратегией A_2: $P = (P_2)$.

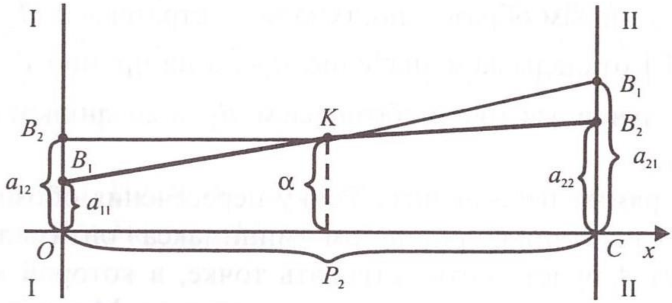

Рис. 2.3

Если при геометрическом решении задачи получится, что прямые B_1B_1 и B_2B_2 не пересекаются (рис. 2.4), то из этого следует, что один из игроков имеет явно выгодную стратегию. Для примера, приведенного на рис. 2.4, явно выгодную стратегию имеет игрок B.

Если один из игроков имеет больше или 2 стратегии, а второй – 2, то решение осуществляется аналогичным образом. А именно, берется единичный отрезок, к нему проводятся два перпендикуляра и строятся отрезки для всех возможных стратегий. Затем определяется нижняя граница выигрыша для игрока A или верхняя граница выигрыша для игрока B. На построенной границе выбирается точка, соответствующая оптимальной стратегии игрока (для игрока A – это наибольшая, для игрока B – наименьшая).

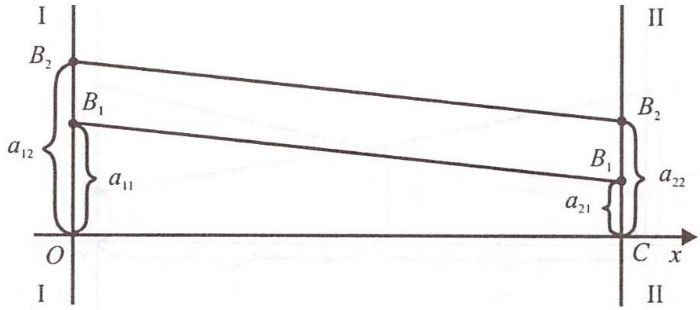

Рис. 2.4

Пример 2.2. Дана игра двух игроков, заданная матрицей (2 ×6) (табл. 2.10).

Таблица 2.10

A_i \ B_j	B_1	B_2	B_3	B_4	B_5	B_5
A_1	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
A_2	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}

Требуется найти решение графическим способом.

Р е ш е н и е. Строим для игрока A числовую ось Ox (рис. 2.5). Откладываем на ней единичный отрезок OC ($|OC| = |I| = I$).

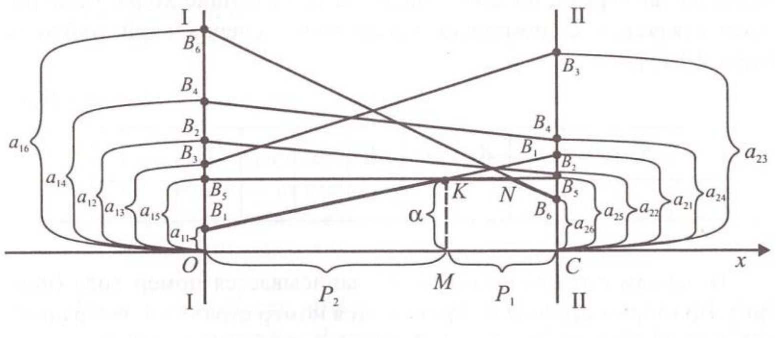

Рис. 2.5

Восстанавливаем к нему два перпендикуляра I-I и II-II. На прямой I-I откладываем выигрыши для стратегии A_1, на прямой II-II – для стратегии A_2. Полученные точки соединяем соответствующими отрезками. Нижней границе выигрыша соответствует ломаная B_1KNB_6. Наибольшей точкой на ней является точка K. Она и определяет оптимальное решение $P = (P_1, P_2)$ и цену игры $\alpha = |KM|$.

2.4. Численные методы решения задач теории игр

Численные методы широко применяются для решения игровых задач. Наиболее распространенным из них является *метод итераций Брауна-Робинсона*. Рассмотрим этот метод.

Пусть у нас имеется конечная парная игра двух игроков A и B.

Игра начинается с выбора игроком A некоторой стратегии A_i. Противник (в данном случае – B) должен ответить той из своих стратегий B_j, которая хуже всего для игрока A. Затем игрок A в ответ на стратегию B_j отвечает некоторой стратегией A_i. Игрок B снова выбирает такую стратегию B_j, которая хуже всего для игрока A и т.д.

Сущность метода Брауна-Робинсона заключается в том, что вместо того, чтобы каждый раз вычислять средний выигрыш, используют выигрыш, «накопленный» за предыдущие ходы. Решение осуществляется с помощью составления специальной таблицы (табл. 2.11).

Таблица 2.11

K	i	B_1	B_2	…	B_n	j	A_1	A_2	…	A_m	$C_н$	$C_в$	$C_{ср}$

В первом столбце таблицы (K) записывается номер хода (партии). Во втором столбце (i) указывается номер стратегии, выбранной игроком. В последующих столбцах указывается «накопленный выигрыш» игрока A за предыдущие ходы при соответствующих стратегиях игрока B (B_i). «Накопленный выигрыш» получается прибавлением элементов строки, соответствующей выбранной стратегии игрока A, к тому, что было строкой выше в таблице. Из элементов накопленного выигрыша выбирается наименьший. Если таких минимальных элементов несколько, выбирается любой из них.

Столбец, соответствующий выбранному минимальному элементу, определяет ответную стратегию игрока B. Следующим за стратегиями (B_i) является столбец (j), в котором записывается номер стратегии, выбранный игроком B.

Дальше следуют столбцы, в которых записывается «накопленный проигрыш» игрока B при соответствующих стратегиях (A_i) игрока A. «Накопленный проигрыш» получается прибавлением элементов столбца исходной таблицы, соответствующего выбранной стратегии B_j к тому, что было строкой выше в заполняемой таблице. Затем из элементов «накопленного проигрыша» выбирается наибольший (если их несколько, выбирается любой). Столбец, соответствующий выбранному элементу, определяет номер стратегии, которую должен применить игрок A в ответ на только что сделанную игроком B стратегию. И так далее.

Нижняя оценка цены игры $C_{\text{н}}$ для каждого хода определяется как частное от деления минимального накопленного выигрыша на число сделанных партий. Верхняя оценка цены игры $C_{\text{в}}$ равна максимально накопленному проигрышу, деленному на число сделанных партий. Среднее арифметическое $C_{\text{ср}} = \dfrac{C_{\text{в}} + C_{\text{н}}}{2}$ может служить приближенным значением цены игры. Доказано, что такой метод сходится: при увеличении числа партий средний выигрыш на одну партию стремится к цене игры, а при увеличении частоты применения стратегий – к их вероятностям в оптимальных смешанных стратегиях игроков.

Особенностью рассмотренного итерационного метода является то, что с увеличением размерности игры трудоемкость вычислений возрастает сравнительно медленно, в то время как для метода линейного программирования она растет очень быстро.

Пример 2.3. Пусть имеется конечная игра игроков A и B, имеющих по 3 стратегии. В качестве игроков A приняты автотранспортные предприятия, а игроков B – обслуживаемые грузоотправители. Выигрыши игрока A для всех пар стратегий A_i и B_j известны и представлены в табл. 2.12.

Таблица 2.12

A_i \ B_j	B_1	B_2	B_3
A_1	7	2	8
A_2	2	8	0
A_3	9	0	10

Элементами матрицы является доход АТП от выполнения определенной транспортной работы (в условных единицах). Требуется разыграть 15 партий, начиная со стратегий A_3 игрока A.

Р е ш е н и е . Составим расчетную таблицу (табл. 1.13).

1. В первом столбце (K) ставим номер первой партии – 1.

2. Во втором столбце ставим номер стратегии, которую должен применить игрок A. По условию задачи игра начинается со стратегии A_3. Следовательно, во втором столбце ставим число 3. В 3–5

столбцах записываем строки, соответствующие стратегии A_3 (из исходной таблицы).

3. Минимальный элемент из написанных выигрышей в столбцах (3–5) таблицы равен нулю. Он соответствует стратегии B_2 игрока B.

Таблица 2.13

K	i	B_1	B_2	B_n	j	A_1	A_2	A_m	$C_{\text{н}}$	$C_{\text{в}}$	$C_{\text{ср}}$
1	3	9	0	10	2	2	8	0	0	8	4
2	2	11	8	10	2	4	16	0	4	8	6
3	2	13	16	10	3	12	16	10	3,3	5,3	4,3
4	2	15	24	10	3	18	16	20	2,5	5	3,75
5	3	24	24	20	3	26	16	30	4	6	5
6	3	33	24	30	2	28	24	30	4	5	4,5
7	1	40	26	38	2	30	32	30	3,7	4,57	4,13
8	2	42	34	38	2	32	40	30	4,25	5	4,62
9	2	44	42	38	3	41	40	40	4,2	4,55	4,37
10	1	51	44	46	2	43	48	40	4,4	4,8	4,6
11	2	53	52	46	3	51	48	50	4,18	4,6	4,39
12	1	60	54	54	2	53	56	50	4,5	4,67	4,58
13	2	62	62	54	3	62	56	60	4,15	4,8	4,47
14	1	69	64	62	3	70	56	70	4,42	5	4,71
15	3	78	64	72	2	72	64	70	4,26	4,8	4,53

4. В столбце 6 ставим $j = 2$.

5. В столбцах (7–9) записываем элементы столбца B_2 из исходной таблицы.

6. Максимальный элемент среди проигрышей, записанных в столбцах (7–9), равен 8. Он соответствует стратегии A_2 игрока A. Следовательно, при следующем ходе игрок A делает стратегию A_2 .

7. Записываем в первом столбце номер партии, равный 2, во втором – номер стратегии, равный 2 (для A_2).

8. В столбцах (3–5) записываем суммарный выигрыш, получаемый прибавлением к элементам верхней строки элементов строки, соответствующих стратегии A_2 из исходной таблицы:

$B_1 = 9 + 2 = 11;$

$B_2 = 0 + 8 = 8;$

$B_3 = 10 + 0 = 10.$

9. Минимальный элемент из полученных выигрышей равен 8. Он соответствует стратегии B_2.

10. В столбце для *j* записываем 2.

11. В столбцах (7–9) записываем суммы, полученные прибавлением элементов столбца B_2 к элементам верхней строки:

$A_1 = 2 + 2 = 4;$

$A_2 = 8 + 8 = 16;$

$A_3 = 0 + 0 = 0.$

12. Максимальный элемент среди полученных проигрышей равен 16. Он соответствует стратегии A_2 игрока *А*. Следовательно, в следующей партии игрок *А* применит стратегию A_2 и так до конца таблицы.

13. Нижняя цена игры для первой партии $C_{\text{н}} = \frac{0}{1} = 0$, верхняя – $C_{\text{в}} = \frac{8}{1} = 8$, средняя цена игры $C_{\text{ср}} = \frac{(C_{\text{н}} + C_{\text{в}})}{2} = \frac{(0+8)}{2} = 4$.

Для второй партии $C_{\text{н}} = \frac{8}{2} = 4{,}0$, $C_{\text{в}} = \frac{16}{2} = 8$, $C_{\text{ср}} = \frac{(4{,}0 + 8)}{2} = 6{,}0$ и так далее.

14. Частоты применения стратегий игроком *А* равны: для A_1: $P_1 = \frac{4}{15} = 0{,}26$, для A_2: $P_2 = \frac{7}{15} = 0{,}468$, для A_3: $P_3 = \frac{4}{15} = 0{,}26$.

15. Частоты применения стратегий игроком *В*: для B_1: $q_1 = \frac{0}{15} = 0$, для B_2: $q_2 = \frac{8}{15} = 0{,}534$, для B_3: $q_3 = \frac{7}{15} = 0{,}466$, что не намного отличается от соответствующих вероятностей.

Если игрок *В* будет пользоваться стратегией *P* = (0; 0,534; 0,466), то выигрыш *А* будет равен 4,53 (последнее значение $C_{\text{ср}}$).

2.5. Связь теории игр с линейным программированием

Теория игр для двух лиц с нулевой суммой имеет непосредственную связь с задачей линейного программирования. Покажем это.

Пусть у нас имеется парная игра с платежной матрицей

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Обозначим стратегию игрока *A*

$$P = (P_1, P_2, \ldots, P_m);$$

стратегию игрока *B*

$$q = (q_1, q_2, \ldots, q_n).$$

Предположим, что игрок *A* применил стратегию *P*. Она должна обеспечить максимальный выигрыш (α). Тогда для j-й стратегии второго игрока *B* имеем:

$$P_1 a_{1j} + P_2 a_{2j} + \cdots + P_m a_{mj} \ge \alpha \ . \qquad (2.5)$$

В матрице *A* только положительные элементы. Если будут отрицательные числа, то путем прибавления некоторого положительного числа ко всем элементам матрицы, делаем их положительными. Прибавляемое положительное число должно быть больше либо равно наименьшему отрицательному элементу. Выигрыш $\alpha > 0$. Разделим полученное уравнение на α:

$$a_{1j} \frac{P_1}{\alpha} + a_{2j} \frac{P_2}{\alpha} + \cdots + a_{mj} \frac{P_m}{\alpha} \ge 1, \quad j = \overline{1, n} \ .$$

Обозначим

$$\frac{P_i}{\alpha} = x_i \ge 0 .$$

Тогда получим:

$$a_{1j} x_1 + \cdots + a_{mj} x_m \ge 1, \quad j = \overline{1, n}; \qquad (2.6)$$

$$x_1 + x_2 + \cdots + x_m = F \to \min, \quad x_i \ge 0, \quad i = \overline{1, m} \ .$$

А это есть задача линейного программирования, решая которую, мы получим и решение задачи теории игр, а именно:

$$\alpha = \frac{1}{\sum_{i=1}^{m} x_i}, \quad P_i = x_i \cdot \alpha \ .$$

Для двойственной задачи в уравнениях знак меняется на противоположный, т.е. получим

$$a_{1j} y_1 + a_{2j} y_2 + \cdots + a_{nj} y_n \le 1 ,$$

где $y_1, y_2, \ldots, y_n$ – неотрицательные переменные, равные $y_j = \dfrac{q_j}{\alpha}$.

Целевая функция

$$Z' = y_1 + y_2 + \cdots + y_n \to \max.$$

Решая полученные задачи линейного программирования, находим решение и исходной задачи теории игр.

Верхняя цена игры $\beta = \dfrac{1}{\sum_{j=1}^{n} y_j}$; вероятности $q_j = y_j \cdot \beta$.

Пример 2.4. Используя метод линейного программирования, найти решение игры, заданной платежной матрицей (табл. 2.14).

Таблица 2.14

A_i \ B_j	B_1	B_2	B_3
A_1	2	–4	6
A_2	–2	8	–4
A_3	6	–2	8

Р е ш е н и е . Так как в матрице есть отрицательные элементы, путем прибавления некоторого неотрицательного числа, например $c = 4$, ко всем элементам матрицы, сделаем их положительными (табл. 2.15).

Таблица 2.15

A_i \ B_j	B_1	B_2	B_3
A_1	6	0	10
A_2	2	12	0
A_3	10	2	12

Решение от этого не изменяется, а цена игры изменится – увеличится на 4 единицы (на величину прибавляемого числа): $\alpha' = \alpha + 4$.

Условия (2.1) примут вид

$6x_1 + 10x_3 \geq 1;$

$2x_1 + 12x_2 \geq 1;$

$10x_1 + 2x_2 + 12x_3 \geq 1.$

Целевая функция запишется в виде $F = x_1 + x_2 + x_3 \to \min$.

Для решения полученной задачи линейного программирования применим табличный симплекс-метод. В качестве базисных введем неотрицательные переменные x_4, x_5 и x_6. Тогда система ограничений примет вид:

$6x_1 + 10x_3 - x_4 = 1;$

$2x_1 + 12x_2 - x_5 = 1;$

$10x_1 + 2x_2 + 12x_3 - x_6 = 1.$

Из последней системы уравнений выразим базисные переменные через свободные x_1, x_2 и x_3:

$x_4 = -1 + 6x_1 + 10x_3;$

$x_5 = -1 + 2x_1 + 12x_2;$

$x_6 = -1 + 10x_1 + 2x_2 + 12x_3.$

Составим симплексную таблицу (табл. 2.16).

Таблица 2.16

БП	1	СП		
		$-x_1$	$-x_2$	$-x_3$
x_4	–1	–6	0	–10
x_5	–1	–2	–12	0
x_6	–1	–10	–2	–12
F	0	– 1	– 1	– 1

Так как в столбце свободных членов у нас есть отрицательные, то для определения опорного решения фиксируем строку (x_6) и по минимальному отношению, равному $\min\left(\frac{1}{6}, \frac{1}{2}, \frac{1}{10}\right) = \frac{1}{10}$, выбираем разрешающим столбец (x_1). Элемент –10, стоящий на пересечении разрешающей строки и разрешающего столбца, является разрешающим. С ним делаем шаг обычных симплексных преобразований (табл. 2.17).

Таблица 2.17

БП	1	СП		
		$-x_6$	$-x_2$	$-x_3$
x_4	$-4/10$	$-6/10$	$+12/10$	–2,8
x_5	$-8/10$	$-2/10$	–9,6	2,4
x_1	$1/10$	$-1/10$	$2/10$	$12/10$
F	$1/10$	$-1/10$	–0,8	0,2

Так как в столбце свободных членов есть отрицательные элементы, полученный план не является опорным. Фиксируем строку (x_5) и в ней отрицательный элемент –9,6 в столбце ($-x_2$). Разрешающим элементом будет элемент (–9,6). С полученным разрешающим элементом снова делаем шаг симплексных преобразований (табл. 2.18).

Таблица 2.18

БП	1	СП		
		$-x_6$	$-x_5$	$-x_3$
x_4	–0,18	–0,6	0,12	–2,5
x_2	0,08	0,02	–0,01	–0,25
x_1	0,14	0,14	0,021	+1,15
F	0,167	0,067	–0,08	1,8

Полученное решение также не является опорным. Оно содержит отрицательный элемент в столбце свободных членов. Поэтому делаем дальнейшие преобразования симплекс-метода. Фиксируем строку (x_4), содержащую отрицательный элемент в столбце свободных членов. В этой строке фиксируем элемент (– 2,5), содержащийся в столбце ($-x_3$). Он же будет и разрешающим элементом. Делаем с ним шаг симплексных преобразований (табл. 2.19).

Таблица 2.19

БП	1	СП		
		$-X_6$	$-X_5$	$-X_4$
X_3	0,07			
X_2	0,06			
X_1	0,03			
F	0,03	– 0,368	– 0,004	– 0,720

Так как в F-строке нет положительных элементов, полученное решение $x^* = (x_1 = 0{,}03;\ x_2 = 0{,}06;\ x_3 = 0{,}07;\ x_4 = 0;\ x_5 = 0;\ x_6 = 0)$ является оптимальным.

Цена игры с матрицей (табл. 14.15) равна

$$\alpha' = \frac{1}{\sum x_i} = \frac{1}{0{,}16} = 6{,}25\ .$$

Цена исходной игры $\alpha = \alpha' - 4 = 6{,}25 - 4 = 2{,}25$.

Тогда вероятности стратегий игрока A будут равны:

$P_1 = x_1 \cdot \alpha' = 0{,}03 \cdot 6{,}25 = 0{,}19$; $P_2 = x_2 \cdot \alpha' = 0{,}06 \cdot 6{,}25 = 0{,}37$; $P_3 = x_3 \cdot \alpha' = 0{,}07 \cdot 6{,}25 = 0{,}44$.

Получили, что оптимальная стратегия игрока A равна:

$P = (P_1, P_2, P_3) = (0{,}19;\ 0{,}37;\ 0{,}44)$.

Упражнения

1. С помощью метода Брауна-Робинсона разыграть 15 ходов парной игры, представленной в матричной форме. Розыгрыш начинать с первой стратегии A_1.

1.1.

A_i \ B_j	B_1	B_2	B_3	B_4	B_5
A_1	9	7	6	6	9
A_2	0	6	5	7	6
A_3	3	3	3	8	7
A_4	4	2	2	9	4

1.2.

A_i \ B_j	B_1	B_2	B_3	B_4
A_1	2	3	5	6
A_2	7	2	0	4
A_3	8	9	10	2
A_4	6	0	1	3

1.3.

B_j \ A_i	B_1	B_2	B_3	B_4
A_1	8	9	6	13
A_2	0	10	8	12
A_3	4	12	0	15
A_4	5	10	0	9
A_5	8	14	12	0

1.4.

B_j \ A_i	B_1	B_2	B_3	B_4
A_1	0	2	6	8
A_2	10	12	10	8
A_3	6	0	4	6
A_4	10	4	0	3
A_5	2	2	4	0

1.5.

B_j \ A_i	B_1	B_2	B_3	B_4	B_5
A_1	9	3	10	0	9
A_2	0	6	8	9	6
A_3	4	0	6	3	5
A_4	6	8	9	4	2

1.6.

B_j \ A_i	B_1	B_2	B_3	B_4
A_1	6	8	0	8
A_2	0	10	14	6
A_3	3	0	16	10
A_4	6	12	10	8

1.7.

$A_i \backslash B_j$	B_1	B_2	B_3	B_4	B_5	B_6
A_1	6	5	7	10	16	10
A_2	0	4	8	0	14	8
A_3	7	3	0	12	13	6
A_4	2	6	9	13	12	4

1.8.

$A_i \backslash B_j$	B_1	B_2	B_3	B_4	B_5	B_6
A_1	8	6	0	12	9	1
A_2	0	10	8	10	0	2
A_3	4	2	18	6	3	4
A_4	6	8	4	8	1	5

1.9.

$A_i \backslash B_j$	B_1	B_2	B_3	B_4	B_5
A_1	9	0	8	4	10
A_2	10	16	0	5	12
A_3	12	8	12	0	13
A_4	0	4	6	10	10

1.10.

$A_i \backslash B_j$	B_1	B_2	B_3	B_4	B_5
A_1	8	8	4	1	3
A_2	0	6	5	3	4
A_3	9	8	0	2	5
A_4	4	5	2	6	8

1.11.

$A_i \backslash B_j$	B_1	B_2	B_3	B_4	B_5
A_1	9	0	8	4	8
A_2	0	8	6	10	4
A_3	6	12	10	0	3
A_4	4	8	6	12	1
A_5	8	10	8	8	2

1.12.

$A_i \backslash B_j$	B_1	B_2	B_3	B_4	B_5
A_1	9	8	4	10	1
A_2	0	16	0	2	3
A_3	12	10	3	1	5
A_4	10	8	6	6	6
A_5	6	6	5	4	4

1.13.

$A_i \backslash B_j$	B_1	B_2	B_3	B_4	B_5
A_1	4	8	0	8	10
A_2	6	3	9	0	2
A_3	8	2	6	3	0
A_4	10	10	4	1	9

1.14.

$A_i \backslash B_j$	B_1	B_2	B_3	B_4	B_5
A_1	1	9	0	0	2
A_2	2	3	6	6	3
A_3	0	2	4	4	4
A_4	8	1	3	3	6

2. Задачи 1.1–1.14 решить графическим способом.

3. Найти решение задач 1.1–1.14 с помощью методов линейного программирования.

ЛИТЕРАТУРА

1. Горлаг Б.А., Шахов В.Г. Математическое моделирование. Построение моделей и численная реализация. М.: Высшая школа. 2018. – 280с.
2. Королев А.В Экономика – математические методы и моделирование. Учебник и практикум для бакалавриата и магистратуры. М.: Высшая школа. 2018 – 280 с.
3. Лебедева Г.И. Прикладная математика. Математические модели в транспортных системах /Г.И. Лебедева, Н.А. Микулик, Минск: Асар, 2009.

Приложение 1

Значения *F* – критерия

(значения *F* даны для $\alpha = 0{,}05$, $v_1 = P$, $v_2 = n - p - 1$)

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	12	14	16	20
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,41	19,42	19,43	19,44
3	10,19	9,55	9,28	9,12	9,01	8,94	8,86	8,84	8,81	8,78	8,74	8,71	8,69	8,66
4	7,71	6,94	6,59	6,36	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,87	5,84	5,80
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,68	4,64	4,60	4,56
6	5,90	5,14	4,76	4,58	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,96	3,92	3,87
7	5,60	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,57	3,52	3,40	3,44
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,28	3,26	3,20	3,15
9	5,12	4,26	3,86	3,63	3,48	3,37	3,20	3,23	3,10	3,10	3,07	3,02	2,98	2,98
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,91	2,86	2,82	2,77
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,79	2,74	2,70	2,65
12	4,76	3,88	3,40	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,60	2,60	2,60	2,54
13	4,67	3,80	3,41	3,16	3,02	2,92	2,84	2,77	2,72	2,67	2,60	2,55	2,51	2,46
14	4,60	3,74	3,34	3,11	2,96	2,86	2,77	2,70	2,65	2,60	2,53	2,48	2,44	2,39
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,48	2,43	2,39	2,33
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,48	2,42	2,37	2,33	2,28
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,38	2,33	2,29	2,23
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,48	2,41	2,34	2,29	2,25	2,19
19	4,36	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,31	2,26	2,21	2,15

Окончание приложения 1

v_1 / v_2	1	2	3	4	5	6	7	8	9	10	12	14	16	20
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,28	2,23	2,18	2,12
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,20	2,15	2,09
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,23	3,18	2,13	2,07
23	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28	2,20	2,14	2,10	2,04
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,18	2,13	2,09	2,02
25	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24	2,16	2,11	2,06	2,00
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,10	2,05	1,99
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25	2,20	2,13	2,08	2,04	1,97
28	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,12	2,06	2,02	1,96
29	4,18	3,33	2,93	2,70	2,54	2,40	2,35	2,28	2,22	2,18	2,10	2,05	2,00	1,94
30	4,17	3,32	2,92	2,68	2,53	2,40	2,34	2,27	2,21	2,16	2,09	2,04	1,99	1,93
32	4,15	3,30	2,90	2,67	2,51	2,40	2,32	2,25	2,19	2,14	2,07	2,02	1,97	1,91
34	4,13	3,28	2,88	2,65	2,49	2,38	2,30	2,23	2,17	2,12	2,05	2,00	1,95	1,89
36	4,11	3,26	2,86	2,63	2,48	2,36	2,28	2,21	2,15	2,10	2,02	1,98	1,93	1,87
38	4,10	3,25	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	2,02	1,96	1,92	1,85
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,00	1,95	1,90	1,84
42	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	1,99	1,94	1,89	1,82
44	4,06	3,21	2,82	2,58	2,43	2,31	2,23	2,16	2,10	2,05	1,98	1,92	1,88	1,81
46	4,05	3,20	2,81	2,57	2,42	2,30	2,22	2,14	2,09	2,04	1,97	1,91	1,87	1,80
48	4,04	3,19	2,80	2,56	2,41	2,30	2,21	2,14	2,08	2,03	1,96	1,90	1,86	1,79

50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,95	1,90	1,85	1,78
55	4,02	3,17	2,78	1,54	2,38	2,27	2,18	2,11	2,05	2,00	1,93	1,88	1,83	1,76

Приложении 2

t – распределение Стьюдента.

(Значения t даны в зависимости от числа степеней свободы v и вероятности α)

β / v	0,40	0,20	0,10	0,05	0,025	0,01	0,005	0,0005
1	0,325	1,379	3,078	6,314	12,706	31,821	63,667	636,619
2	0,289	1,061	1,886	2,920	4,303	6,965	9,925	31,598
3	0,277	0,978	1,638	2,353	3,182	4,541	5,841	12,941
4	0,271	0,941	1,533	2,132	2,776	3,747	4,604	8,010
5	0,267	0,320	1,476	2,015	2,571	3,365	4,032	6,859
6	0,265	0,906	1,440	1,943	2,447	3,148	3,707	5,959
7	0,263	0,879	1,414	1,895	2,365	2,998	3,499	5,405
8	0,262	0,889	1,397	1,860	2,306	2,896	3,355	5,041
9	0,261	0,883	1,383	1,833	2,262	2,821	3,250	4,784
10	0,260	0,879	1,372	1,812	2,228	2,764	3,169	4,587
11	0,260	0,876	1,363	1,796	2,202	2,718	3,106	4,437
12	0,259	0,873	1,356	1,782	2,179	2,681	3,055	4,318
13	0,259	0,870	1,350	1,771	2,160	2,650	3,012	4,221
14	0,258	0,768	1,345	1,761	2,145	2,624	2,977	4,140
15	0,258	0,860	1,341	1,753	2,131	2,602	2,947	4,073
16	0,258	0,865	1,337	1,746	2,120	2,583	2,921	4,015
17	0,257	0,863	1,333	1,740	2,110	2,567	2,798	3,965
18	0,257	0,862	1,330	1,734	2,101	2,552	2,878	3,922
19	0,257	0,861	1,328	1,729	2,093	2,539	2,861	3,888
20	0,257	0,860	1,325	1,725	2,036	2,528	2,845	3,850
21	0,257	0,859	1,323	1,721	2,030	2,518	2,831	3,819
22	0,256	0,858	1,321	1,717	2,074	2,508	2,810	3,792
23	0,256	0,858	1,319	1,714	2,069	2,500	2,807	3,767
24	0,256	0,857	1,318	1,711	2,064	2,492	2,797	3,745
25	0,256	0,856	1,316	1,708	2,060	2,485	2,787	3,725
26	0,256	0,856	1,315	1,706	2,056	2,479	2,779	3,707
27	0,256	0,855	1,314	1,703	2,052	2,478	2,771	3,690
28	0,256	0,855	1,313	1,701	2,048	2,467	2,763	3,674
29	0,256	0,854	1,311	1,699	2,045	2,462	2,756	3,659
30	0,256	0,854	1,310	1,697	2,042	2,457	2,750	3,646
40	0,255	0,851	1,304	1,684	2,021	2,423	2,704	3,551
50	0,254	0,848	1,296	1,671	2,000	2,403	2,660	3,485

Приложение 3

Критические значения U_v $P = 0{,}95$, λ

$v = n - 2$	λ										
	0	0,10	0,20	0,30	0,40	0,50	0,60	0,70	0,80	0,90	1,00
1	12,706	13,919	15,014	16,001	16,885	17,669	18,355	18,941	19,422	19,782	19,975
2	4,303	4,582	4,844	5,087	5,312	5,517	5,701	5,864	6,00	6,107	6,164
3	3,182	3,356	3,518	3,668	3,809	3,939	4,059	4,166	4,258	4,330	4,371
4	2,176	2,915	3,040	3,159	3,217	3,373	3,469	3,556	3,692	3,692	3,727
5	2,571	2,692	2,802	2,904	3,00	3,089	3,173	3,249	3,316	3,371	3,402
6	2,447	2,558	2,659	2,751	2,838	2,919	2,995	3,065	3,127	3,178	3,207
7	2,365	2,469	2,563	2,650	2,730	2,806	2,877	2,943	3,002	3,051	3,078
8	2,306	2,406	2,496	2,578	2,654	2,726	2,793	2,856	2,913	2,959	2,986
9	2,262	2,358	2,445	2,524	2,597	2,666	2,731	2,792	2,846	2,891	2,918
10	2,228	2,322	2,406	2,482	2,552	2,618	2,681	2,741	2,794	2,838	2,865
12	2,179	2,269	2,350	2,422	2,489	2,552	2,612	2,669	2,720	2,763	2,788
14	2,145	2,233	2,311	2,381	2,446	2,506	2,564	2,616	2,669	2,710	2,735
16	2,120	2,206	2,282	2,351	2,414	2,473	2,529	2,583	2,631	2,672	2,696
18	2,101	2,186	2,260	2,328	2,389	2,447	2,502	2,555	2,603	2,643	2,666
20	2,086	2,170	2,243	2,209	2,370	2,427	2,481	2,533	2,580	2,620	2,643
25	2,060	2,141	2,213	2,277	2,336	2,392	2,444	2,494	2,540	2,579	2,602
50	2,009	2,086	2,154	2,215	2,271	2,323	2,372	2,419	2,463	2,501	2,523
100	1,984	2,060	2,126	2,186	2,240	2,290	2,338	2,384	2,426	2,462	2,485
более 100	1,960	2,064	2,099	2,156	2,209	2,258	2,304	2,349	2,390	2,426	2,448