Федеральное агентство по образованию РФ

Государственное образовательное учреждение высшего профессионального образования

Белгородский Государственный Технологический Университет им В. Г. Шухова

Кафедра программного обеспечения вычислительной техники и

автоматизированных систем

Расчётно-графическое задание по дисциплине

«Системный анализ»

**Регрессионный анализ**

Выполнение: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Защита: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Выполнил:

студент группы ПВ-31

Адаменко И. И.

Проверил:

к. т. н., профессор

Полунин А. И.

Белгород  
2015

Лабораторная работа №7

«Регрессионный анализ»

## Теоретические сведения

Основная задача регрессионного анализа — создание мат. модели объектов или явлений на основе экспериментов или наблюдений. Он является наиболее широко распространённым статистическим методом, позволяя строить мат. модели и проводить статистический анализ этой модели и результатов расчётов на предмет точности.

**Регрессионный анализ** — это получение требуемых зависимостей с использованием метода наименьших квадратов, плюс статистическая проверка полученных результатов.

Существует две схемы регрессионного анализа:

1. Уравнения регрессии определяется как функция безусловного мат. ожидания результатов измерения зависимостей переменной при детерминированных значениях независимой переменной.
2. Независимые переменные считаются случайными и уравнения регрессии ищутся как функция условного мат. ожидания зависимых переменных.

**Линейная регрессия** — оцениваемые коэффициенты являются множителями перед базовыми функциями.

Модель процесса или явления, которую надо найти по результатам экспериментов описывается алгебраическими зависимостями и зависит от большого числа аргументов. Одни аргументы сильно влияют на выходную величину, другие — менее, причём часто бывает ситуация, что какие-то факторы (аргументы) известны исследователю, а какие-то нет. В этом случае в регрессионном анализе принимается допущение, и регрессионная модель задаёт функцию мат. ожидания выходной величины. В этом случае модель имеет вид

а реальный процесс:

В регрессионном анализе возникает задача, правильно подобрать  — регрессоры. При определении модели возникает задача нахождения регрессоров, коэффициентов и .

При построении мат. модели (поиске регрессоров и коэффициентов) используют результаты наблюдений. Далее необходимо оценить коэффициенты модели. Будем считать, что известны

Выбор регрессоров зависит от характера исследования процесса. Если процесс монотонно-возрастающий, то чаще всего для его аппроксимации используют степенные функции (линейные, либо квадратичные, либо кубические). Если процесс периодический, то можно использовать тригонометрические функции. Допустим, что нам известны регрессоры модели, известен вектор измерений y (вектор наблюдений за процессом) и известны значения аргументов для каждого наблюдения. Тогда можем записать выражение для мат. модели вектора наблюдений.

Эксперименты бывают активные и пассивные. При активном эксперименте задают значения аргументов и получают выходную величину. При нем можно аргументы задавать так, чтобы обеспечить выполнение разных требований к эксперименту (мин. ошибку) или обеспечить диагональность корреляционной матрицы.

Для оценки коэффициентов a вводим функцию характеризующую качество аппроксимации — сумма квадратов отклонений результатов эксперимента минул эти результаты, вычисленные по модулю.

Необходимо, имея вектор измерений y и мат. модель этого вектора, подобрать такие значения коэффициентов a, чтобы функция J, задающая сумму квадратов разностей между измерением в эксперименте и соответствующим значением переменной, вычисленной по модели.

Если регрессионная модель выбрана правильно, то оценку шума вычисляю по формуле:

K — число элементов в векторе a^,

N — число измерений, при котором оценивали.

N – K — число степеней свободы.

## Проверка адекватности модели

Полученные формулы позволяют оценить коэффициенты модели в случае, когда регрессоры заданы, то есть задано количество регрессоров и их вид.

При поиске модели необходимо осуществить перебор различных видов модели. Вид модели часто берётся исходя из физический законов, описывающих поведение системы. Для выбора модели из всех возможных вариантов используется специальная процедура — проверка адекватности модели.

В дисперсионном анализе доказывается, что если для каждого уравнения факторов есть несколько наблюдений выходной переменной, то существует зависимость:

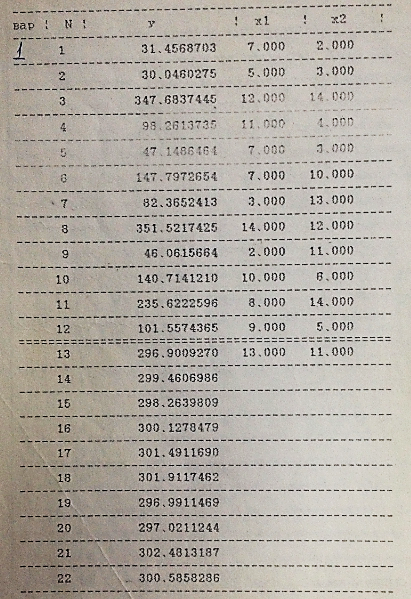
Величины Q, Qr, Qo характеризуют качество модели. Если модель хорошая, то Q примерно равно Q, а Qo = 0

## Работа по нахождению уравнения регрессии

1. На основе априорной информации выбирают предварительную структуру модели.
2. Используя экспериментальные данные и математическую модель получают методом наименьших квадратов оценки коэффициентов регрессии.
3. Вычисляют результаты дисперсионного анализа Qr, Qo. Вычисляют коэффициенты множественной корреляции R и проверяют его значимость. При незначимом коэффициенте R проводят анализ явления и меняют модель, включая новые факторы или усложняя структуру.
4. Перебирают разные модели и выписывают коэффициенты множественной корреляции R.
5. Выбирают больший или группу больших (личную модель). Проверяют значимость коэффициентов уравнения регрессии. Если есть незначимые, то их отбрасывают и снова оценивают Q, Qr, Qo, R, и проверяют значимость уравнения так до тех пор, пока все коэффициенты значимы.
6. Строят доверительные интервалы.

## Задание

По данным измерениям составить уравнение регрессии. Измерения 13-22 использовать для оценки сигма эпсилон.



## Результат выполнения программы

