*Государственное образовательное учреждение высшего профессионального образования*

*«Московский государственный технический университет имени Н. Э. Баумана»*

ОТЧЕТ

По лабораторной работе №3

По курсу «Математическая статистика»

Тема: «Метод наименьших квадратов»

Вариант 21

|  |  |
| --- | --- |
| Студент: | Янова Д.Ю. |
| Группа: | ИУ7-63 |
| Преподаватель: | Саркисян П. С. |

Москва, 2019

**1. Постановка задачи аппроксимации неизвестной зависимости по результатам наблюдений**

Пусть – случайная величина, – детерминированные величины. Если изменение значений влияет на значения случайной величины Y, то говорят, что стохастически зависит от . Задача регрессионного анализа – задача, связанная с установлением аналитических зависимостей между случайной величиной и детерминированными величинами , носящими количественный характер..

В регрессионном анализе используется модель черного ящика, как наиболее общая модель, ассоциируемая с понятием отображения. На вход поступает вектор , который посредством некоторого отображения и случайных возмущений преобразуется в вектор .

**2. Понятие МНК-оценки параметров линейной модели**

Рассмотрим частный случай: m=1, s=1, имеются результаты n наблюдений; ; требуется на основании этих данных подобрать функцию так, чтобы она «наилучшим» образом аппроксимировала неизвестную функцию Ф. В качестве часто выбирают функцию вида – некоторые известные (базисные) функции, а параметры необходимо подобрать. В этом случае регрессионная модель называется линейной по параметрам.

Будем предполагать, что систематические ошибки отсутствуют: . Оценка вектора параметров называется оценкой, полученной по методу наименьших квадратов (МНК-оценкой), если доставляет минимальное значение функции .

**3. Формулы вычисления МНК-оценки (в рассматриваемом случае)**

В рассматриваемом случае базисные функции имеют вид ; среднеквадратичное отклонение полученной модели: , где – экспериментально полученное значение Y при – значение Y, полученное с помощью МНК-оценки.

Для вычисления оценки вектора параметров используется формула , где .

**4. Листинг программы**

ti = csvread("data\_t.csv");

yi = csvread("data\_y.csv");

ti\_square = ti .\* ti;

Ksi = horzcat(ones(length(ti), 1), transpose(ti), transpose(ti\_square));

display(Ksi);

theta = inv(transpose(Ksi) \* Ksi) \* transpose(Ksi) \* transpose(yi);

y\_cap = theta(1) + theta(2) \* ti + theta(3) \* ti\_square;

delta = sqrt(sum((yi - y\_cap).^2));

fprintf("delta is %.2f\n thetas: %.2f, %.2f, %.2f\n", delta, theta(1), theta(2), theta(3));

plot(ti, yi, '.b');

hold on;

plot(ti, y\_cap, 'r');

hold off;

axis tight;

grid on;

**5. Результаты работы для индивидуальной выборки**

