Министерство науки и высшего образования Российской Федерации

Новосибирский государственный технический университет

Кафедра вычислительной техники

Отчет по лабораторной работе № 3

по дисциплине «Теория систем и системный анализ»

на тему «Полный метод наименьших квадратов»

Студенты: Атласюк И.Р., Ириков Е.А.

Группа: АММ2-24

Преподаватель: Гошко Е. Ю.

Новосибирск, 2024

**Цель работы**

Произвести аппроксимацию линейной функции в случае, когда ошибка присутствует не только в зависимой переменной *y*, но и в независимой переменной *x*. Для аппроксимации использовать полный метод наименьших квадратов.

**Ход работы**

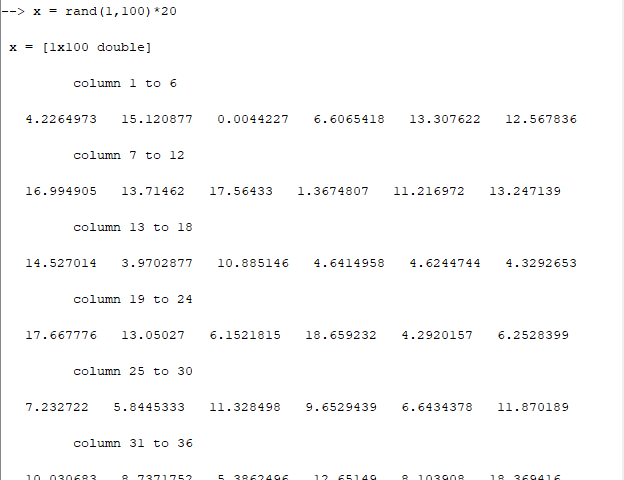
Выберем коэффициенты для линейной функции, которую будем аппроксимировать. Пусть это будет следующая функция:



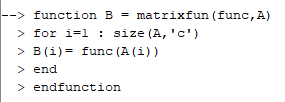
В среде Scilab данную функцию зададим следующим образом:



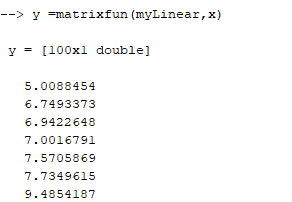
Далее зададим множество значений независимой переменной *x* размера 100 случайным образом. Значения в диапазоне от -20 до 20. Отсортируем значения по возрастанию:



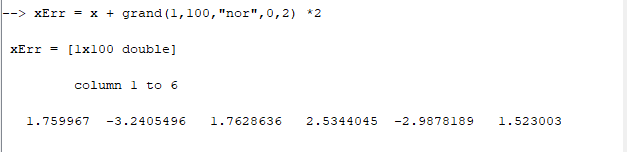
Далее идет функция, которая применяется к каждому значению вектора:

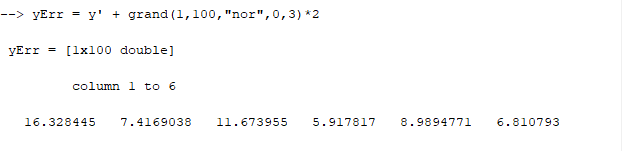


Затем применяем нашу функцию к вектору входных значение *x:*



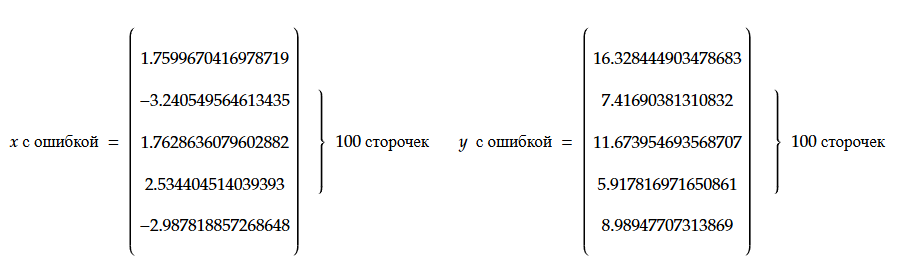
Сгенерируем вектор ошибки для обеих переменных *x* и *y* получаем, что ошибка присутствует во всех измеряемых нами параметрах. Ошибка имеет нормальное распределение, но зададим у функции ошибки для *x* и *y* разные дисперсии распределения, 2 и 3 соответственно. Математическое ожидание равно 0:



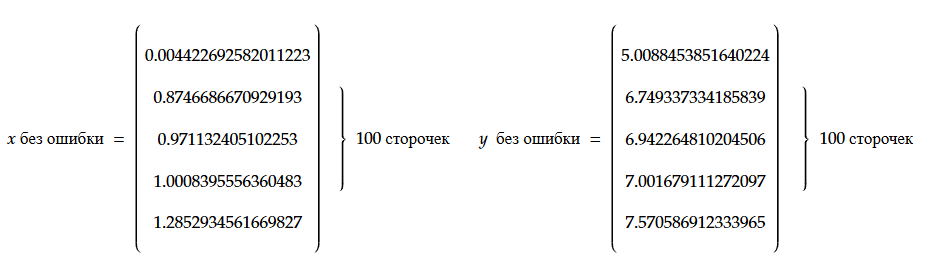


Получили следующие вектора для независимой и зависимой переменных:

Вектора с ошибкой:



Исходные вектора без ошибок:



## **Аппроксимация линейной функции полным методом наименьших квадратов**

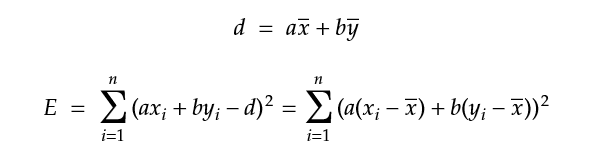
Этот метод применяется, в основном, для линейных функций. Идея заключается в том, что уравнение прямой можно представить в виде:



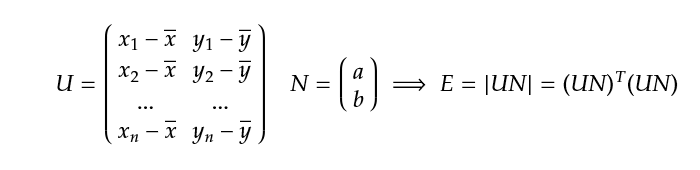
где 𝑎 и 𝑏 — компоненты вектора нормали к прямой, а 𝑑 — свободный член, который определяет её удалённость от начала координат. В таком случае расстояние от точки до прямой, т. е. длина перпендикуляра, опущенного из этой точки на прямую в двумерном пространстве, можно вычислить как:



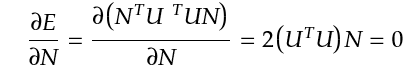
Цель метода состоит в минимизации суммы квадратов этих расстояний для всех точек. Это достигается аналогично стандартному методу наименьших квадратов — через обнуление частных производных функции ошибки. При этом дополнительно предполагается, что прямая должна проходить через центр масс заданного множества точек, из-за чего 𝑑 становится зависимой переменной от 𝑎 и 𝑏, а его значение определяется положением центра масс. Соответственно получаем:



Данная сумма представляется как произведение матрицы отклонений точек от среднего на вектор нормали прямой, возведенные в квадрат. Поэтому мы можем представить функцию, которую мы хотим минимизировать в виде скалярного произведения матрицы отклонений и вектора нормали, что и записано ниже:



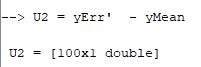
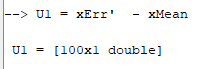
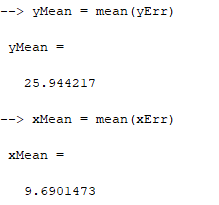
Далее просто минимизируем нашу функцию, находя ноль у частной производной по вектору нормали, который содержит коэффициенты *a* и *b:*

**

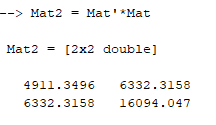
Для решения задачи строится квадратная матрица вторых моментов и находятся её собственные числа. Если бы существовало точное решение, то искомый собственный вектор соответствовал бы нулевому собственному значению. Однако, поскольку идеального решения нет, в качестве приближённого решения выбирается собственный вектор, соответствующий минимальному собственному значению. Этот вектор и определяет нормальные коэффициенты

𝑎 и 𝑏, которые задают наилучшую аппроксимацию прямой.

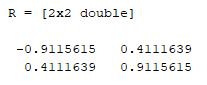
Делаем в Scilab матрицу *U:*

**

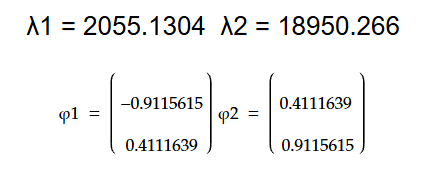
Находим матрицу вторых моментов U^T U:



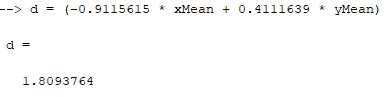
Находим собственные числа и вектора для неё:



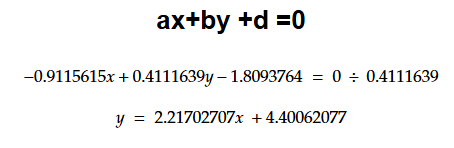
Получаем следующие собственные числа и вектора:



Берем первый собственный вектор в качестве решения, т. к. для него собственное число ближе к 0, т. е. к точному решению. Находим *d* как центр масс точек, с уже известными коэффициентами a и *b:*

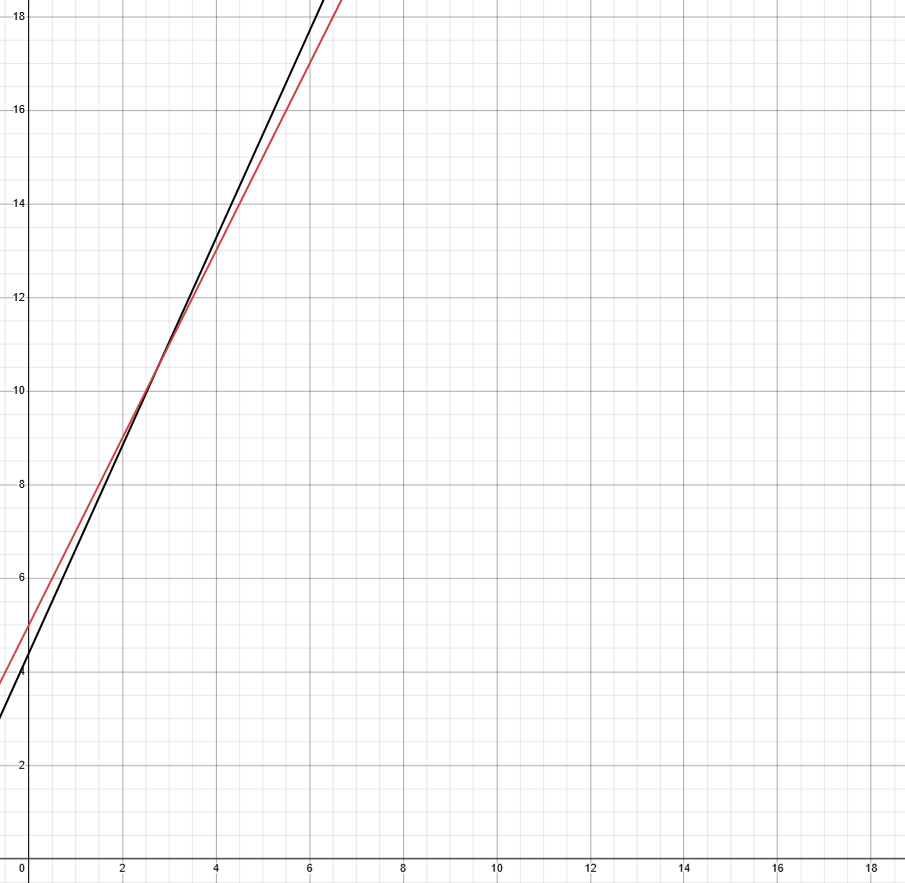


Получаем d = -1.8093764**.** Теперь восстанавливаем все коэффициенты, разделив на коэффициент перед *y:*

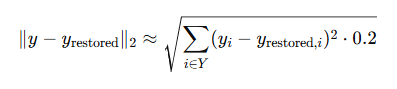
**

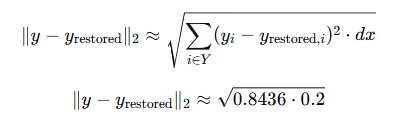
Приблизительно получили изначальное уравнение, достаточно точное.

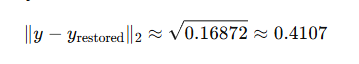
Графики прямых предоставлены ниже, черный – аппроксимированный, красный оригинальный.



Чтобы посчитать расстояние между функциями, нужна приблизительная ширина прямоугольника (расстояние между x). Но она здесь разная. Поэтому, для простоты, посчитает среднее расстояние между x и примем это расстояние за ширину прямоугольника, подставим формулу. Получаем следующую величину расстояния между приближенной и аппроксимированной формулой по Евклидовой норме:







**Заключение**

В рамках лабораторной была аппроксимирована линейная функция с использованием полного метода наименьших квадратов. Рассмотрен случай аппроксимации линейной функции, где ошибка присутствует как в зависимой, так и в независимой переменных.