

Санкт-Петербургский политехнический университет Петра Великого
Институт компьютерных наук и технологий
Высшая школа программной инженерии

Отчет по лабораторной работе 6

Вариант 5

Выполнил студент гр. в3530904/00321	<подпись>	В.Я. Копылов
Руководитель	<подпись>	В.С. Тутыгин

Санкт-Петербург
2022г.

Оглавление

Цель работы	3
Программа работы	4
Результат работы	6
Выводы	8

Цель работы

Целью работы является освоение методики получения адекватного аналитического описания зависимости $Y=f(X)$, по данным, содержащим случайные погрешности.

Наиболее совершенным программным средством для решения данной задачи является пакет программ MATLAB. MATLAB является средой разработки программ обработки данных и, одновременно, содержит большое количество готовых программ, в частности, программу подгонки кривых CurveFitting.

Рисунок 1 - цель работы

Массив зашумленных функциональных зависимостей: **Censu5**

Программа работы

- а) Запустите MATLAB2019b ;
- б) загрузите mat-файл данных censuN.mat :
Home/ImportData/ <путь>censuN.mat;
- в)откройте панель Curve Fitting Tool для чего выберите **Apps/ CurveFitting**;
- г) на панели CurveFittingTool в окнах Xdata, Ydata станут доступными составляющие cdate (X) и pop (Y) из файла censuN.mat;
- д) произведите выбор cdate и pop, выберите вид аппроксимирующей функции и активизируйте кнопку **Fit**;

Для отображения графика погрешностей аппроксимации:
View/Residuals

Рисунок 2. Задание

Этапы оценки качества подгонки и выбор лучшей модели. Этап 1.2.

При визуальной оценке качества подгонки (+ или -) следует принимать во внимание также характер графиков погрешностей аппроксимации. Подгонку можно считать удовлетворительной, если график погрешности шумоподобный с нулевым средним значением. Для отображения графика погрешности:
View/ResidualsPlot

Рисунок 3. Задание

Этап 2 оценки качества подгонки и выбор лучшей модели

2. Выбор лучшей кривой в каждом классе по Adjusted R-square

Лучшей кривой в классе является та, которой соответствует максимальное значение Adjusted R-square. Если максимальное значение соответствует двум или более кривым – выбирается кривая с более простым аналитическим описанием.

Рисунок 4. Задание

Этап 3 оценки качества подгонки и выбор лучшей модели

3. Окончательный выбор аппроксимирующей зависимости

Выбрать самую простую модель из лучших в каждом классе с условием примерно тех же значений диапазона погрешностей, R-square и RMSE (Root Mean Squared Error).

Если подгонка кривых выполняется для обработки результатов физического эксперимента – окончательный выбор модели производит конечный пользователь (физик).

Рисунок 5. Задание

Результаты работы

Для текущего датасета была построена таблица, в которой оценивается качество “подгонки” кривой, при разных аппроксимирующих функций.

Вид модели	Порядок модели	Качество подгонки (+ / -)	Диапазон погрешностей	R-square	Adjusted R-square	RMSE
Полиномиальная	1	-	-10..20	0.8496	0.8417	10.33
	2	-	-3..3	0.9962	0.9958	1.677
	3	+	-0.03..0.03	1	1	0.01424
	4	+	-0.03..0.03	1	1	0.01411
	5	+	-0.03..0.03	1	1	0.01456
Фурье	1	+	-3..3	0.9962	0.9956	1.726
	2	+	-0.03..0.03	1	1	0.01456
	3	+	-0.03..0.03	1	1	0.01545
	4	+	-0.03..0.03	1	1	0.01661
	5	+	-0.02 ..0.02	1	1	0.01167
Экспоненциальная	1	+	-5..3	0.9918	0.9914	2.406
	2	-	-15..10	0.949	0.9399	6.365
Гауссиан	1	+	-0.5..0.5	0.9998	0.9998	0.3695
	2	+	-0.5..0.5	0.9998	0.9998	0.4047
Показательная	1	+	-0.2..0.6	0.9999	0.9999	0.3027
	2	+	-0.03..0.03	1	1	0.01404
Рациональная	1/1	-	-15..25	0.8496	0.8329	10.62
	1/5	-	0..80	-0.8505	-1.644	42.23
	2/1	+	-4..4	0.9962	0.9956	1.726
	2/5	+	-0.2..0.3	1	1	0.136
	3/1	-	-4..3	0.9967	0.9959	1.664
	4/1	+	-0.03..0.03	1	1	0.01336
	5/5	+	-0.06..0.04	1	1	0.02745

Из данной таблицы, для каждой из видов моделей, была выбрана “лучшая кривая” исходя из критерия “Adjusted R-square”.

Вид модели	Порядок модели	Качество подгонки (+ / -)	Диапазон погрешностей	R-square	Adjusted R-square	RMSE
Полиномиальная	3	+	-0.03..0.03	1	1	0.01424
Фурье	2	+	-0.03..0.03	1	1	0.01456
Экспоненциальная	1	+	-5..3	0.9918	0.9914	2.406
Гауссиан	1	+	-0.5..0.5	0.9998	0.9998	0.3695
Показательная	2	+	-0.03..0.03	1	1	0.01404
Рациональная	2/5	+	-0.2..0.3	1	1	0.136

Из данной таблицы, была выбрана, как окончательный вариант аппроксимативной модели для датасета - **модель Фурье**, из-за легкости ее построения, низкого **RMSE** и малого диапазона погрешностей.

Выводы

После проведенных испытаний **была выбрана модель Фурье второй степени**, из-за легкости ее построения, низкого RMSE и малого диапазона погрешностей.